

ON THE SCALING OF RANDOM TAMARI INTERVALS AND SCHNYDER WOODS OF RANDOM TRIANGULATIONS (WITH AN ASYMPTOTIC D-FINITE TRICK)

Guillaume Chapuy*¹

¹*Université Paris Cité, CNRS, IRIF, F-75013, Paris, France.
guillaume.chapuy@irif.fr*

Submitted: May 6, 2024; Accepted: Dec 27, 2024; Published: Jul 15, 2025

© The author. Released under the CC BY license (International 4.0).

Abstract. We consider a Tamari interval of size n (i.e., a pair of Dyck paths which are comparable for the Tamari relation) chosen uniformly at random. We show that the height of a uniformly chosen vertex on the upper or lower path scales as $n^{3/4}$, and has an explicit limit law. By the Bernardi–Bonichon bijection, this result also describes the height of points in the canonical Schnyder trees of a uniform random plane triangulation of size n .

The exact solution of the model is based on polynomial equations with one and two catalytic variables. To prove the convergence from the exact solution, we use a version of moment pumping based on D-finiteness, which is essentially automatic and should apply to many other models. We are not sure to have seen this simple trick used before.

It would be interesting to study the universality of this convergence for decomposition trees associated to positive Bousquet–Mélou–Jehanne equations.

Keywords. Tamari lattice, interval, scaling limit, asymptotics, method of moments, algebraic functions, computer algebra, random triangulation, catalytic variables, discrete differential equation

Mathematics Subject Classifications. 05A15, 05A16, 60F99

1. Introduction and main results

For $n \geq 0$, a *Dyck path of size n* is a lattice path made of n up-steps and n down-steps, starting (and ending) at height 0, and that always stays nonnegative. See Figure 1.1. Dyck paths of

*Supported by the grants ANR-19-CE48-0011 “Combiné”, ANR-18-CE40-0033 “Dimers”, and ANR-23-CE48-0018 “CartesEtPlus”

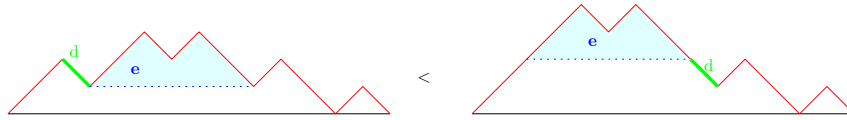


Figure 1.1: A Dyck path of size 7, with a marked down-step followed by an up-step. The path obtained by flipping this down-step with the shortest excursion following it is declared to be *larger*. The Tamari lattice is the partial order generated by all such relations.

size n are in immediate bijection with well-formed parentheses words of length $2n$, and they are counted by the Catalan number

$$Cat(n) := \frac{1}{n+1} \binom{2n}{n}.$$

The set \mathcal{D}_n of Dyck paths of size n can be endowed with several interesting (partial) orders. Maybe the most interesting and natural one is the *Tamari partial order*, whose covering relation is nothing but the edge-flip once interpreted through the classical bijection between Dyck paths and triangulations of a polygon. The Tamari partial order is a lattice, whose covering relation can be described as follows: Let P be a Dyck path, and let d be a down-step in P , followed by an up-step. Let e be the excursion following d in P (an excursion is a path staying higher than its starting point except for its last point). Then the Dyck path Q obtained from P by exchanging d and e is declared *larger* than P for the Tamari order. The reflexive transitive closure of this relation defines the Tamari lattice. See Figure 1.1.

The Tamari lattice plays an important role in many facets of algebraic combinatorics and discrete mathematics. To name a few: its Hasse-diagram is the 1-skeleton of a famous polytope called the associahedron, see e.g. [MHPS12, PSZ23]; determining its diameter has been a long-standing problem solved only recently [Pou14], despite mysterious connections to hyperbolic geometry [STT88]; and, quite excitingly, determining the mixing time on this graph is still an open problem despite recent partial results [MT99, EF23].

In this paper we are interested in yet another connection, related to *Tamari intervals*. We define an *interval* of size n in the Tamari lattice as a pair $(P, Q) \in (\mathcal{D}_n)^2$ such that $P \leq Q$ (for the Tamari partial order). Beware that in this definition an interval is a pair of elements (it is *not* defined as the set of all elements in the lattice that lie between them). Note also that the size of the interval is defined as the common size of the Dyck paths it contains.

We let \mathcal{I}_n be the set of Tamari intervals of size n . In a famous paper, Chapoton [Cha07] proved that the number of elements of \mathcal{I}_n is given by

$$|\mathcal{I}_n| = \frac{2}{n(n+1)} \binom{4n+1}{n-1}, \quad (1.1)$$

which is also the number of rooted planar triangulations¹ of size n [Tut62]. An elegant, and deep, direct bijective proof has been given by Bernardi and Bonichon in [BB09]. Since Chapo-

¹Rooted planar triangulations are embedded planar graphs with only faces of degree 3. These objects are much more complex than the triangulations of a single polygon Dyck paths are in bijection with.

ton’s discovery, the analogy between Tamari intervals and planar maps has been much developed. On the one hand, the enumeration of Tamari intervals can be refined (or weighted), giving rise to product formulas [BMCPR13] surprisingly similar to the ones counting planar constellations [BMS00], suggesting the existence of an underlying object maybe as rich as the GUE 1-matrix model, with a potentially integrable structure, which is yet to be discovered. On the other hand, many different variants of Tamari intervals have been introduced and enumerated. Their enumeration exhibits a universality phenomenon very similar to the one existing in the world of planar maps, and in fact many of these variants are equi-enumerated with some well-known families of planar maps. The understanding of this phenomenon is still very partial, see [Fan21] and references therein. More recently, non-lattice variants of the Tamari partial order have been introduced for which the similarity to maps is even more stunning [BMC24].

For integers a, b , we write $[a..b] := \{a, \dots, b\}$, $(a..b) := \{a + 1, \dots, b - 1\}$, and we use the shortcut notation $[n] := [1..n]$.

1.1. Main results

Although large random planar maps have been intensely studied in the last decades (see for example, among hundreds of references, [CS04, BDFG03, LG19]), it seems that the behaviour of random Tamari intervals has not been studied (and we are not sure to know why). It is however natural to ask this question:

What does a large uniformly random Tamari interval look like?

We are not aware of any results in this direction. In this paper we give a first answer to this question. If $P \in \mathcal{D}_n$ and $i \in [0..2n]$, we write $P(i)$ for the height of the point of P lying at abscissa i . We show:

Theorem 1.1. *Let (P, Q) be a Tamari interval of size n , chosen uniformly at random in \mathcal{I}_n . Let $I \in [0..2n]$ be an integer chosen uniformly at random, and let $Q_n(I)$ be the height of the point of the upper path Q lying at abscissa I . Then we have the convergence in law*

$$\frac{Q_n(I)}{n^{3/4}} \longrightarrow Z = (XY)^{1/4} \tag{1.2}$$

when n goes to infinity, where $X \sim \beta(\frac{1}{3}, \frac{1}{6})$ and $Y \sim \Gamma(\frac{2}{3}, \frac{1}{2})$ are independent random variables. In fact, we have the convergence of all moments: for integer $k \geq 0$,

$$\mathbf{E} \left[\left(\frac{Q_n(I)}{n^{3/4}} \right)^k \right] \longrightarrow \frac{\sqrt{3} \cdot 2^{-\frac{k}{4}-1} \Gamma(\frac{1}{4}k + \frac{1}{3}) \Gamma(\frac{1}{4}k + \frac{2}{3})}{\sqrt{\pi} \Gamma(\frac{1}{4}k + \frac{1}{2})}. \tag{1.3}$$

We recall that the random variables $\beta(a, b)$ and $\Gamma(\ell, \theta)$, defined for positive real numbers a, b, ℓ, θ , have respective densities $\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1} (1-x)^{b-1}$ on $(0, 1)$ and $\frac{x^{\ell-1} e^{-x/\theta}}{\Gamma(\ell)\theta^\ell}$ on \mathbb{R}_+ . Their respective k -th moments are $\frac{\Gamma(a+b)\Gamma(a+k)}{\Gamma(a)\Gamma(a+b+k)}$, and $\frac{\theta^k \Gamma(\ell+k)}{\Gamma(\ell)}$, see e.g. [JKB94, JKB95]. Therefore it is direct to check that (1.3) with k substituted by $4k$ is indeed equal to the k -th moment

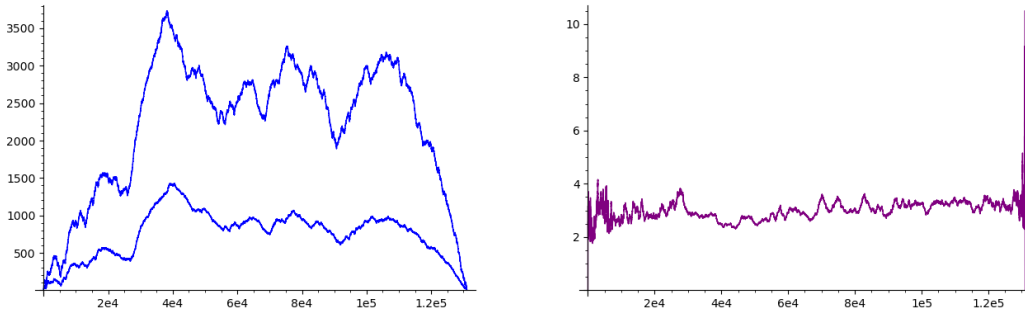


Figure 1.2: Left: A uniform random Tamari interval (P_n, Q_n) of size $n = 65536$ generated with a python code generously provided by Wenjie Fang. Right: plot of $Q_n(i)/P_n(i)$.

of XY . Note also that, because these moments do not grow too fast, the convergence of moments implies the convergence in law (see e.g. Theorems 30.1 and 30.2 in [Bil12], or [FS31]). Therefore in this paper we will only study the convergence of moments.

Interestingly, the random variable Z^4 already appears as a limit law in a (seemingly unrelated) physics context [PS02]. This reference also gives a closed form for the density (but does not identify it as a product of a Gamma and a Beta variable)².

One can also address the case of the lower path, at the price of more technical, and maybe less standard, methods. We obtain the following theorem.

Theorem 1.2. *Let (P, Q) be a Tamari interval of size n , chosen uniformly at random in \mathcal{I}_n . Let $I \in [0..2n]$ be an integer chosen uniformly at random, and let $P_n(I)$ be the height of the point of the lower path P lying at abscissa I . Then we have the convergence in law*

$$\frac{P_n(I)}{n^{3/4}} \longrightarrow \frac{1}{3}Z, \quad (1.4)$$

when n goes to infinity, where Z is as in Theorem 1.1. Again, we have the convergence of all moments.

In view of Theorems 1.1 and 1.2, it is natural to suspect that

$$P_n(I) = \frac{1}{3}Q_n(I) + o(Q_n(I)), \quad (1.5)$$

with a little- o in probability. If true, this means roughly speaking that the lower path is almost a third of the upper path, asymptotically almost surely and almost everywhere. See Figure 1.2 which shows a simulation for $n = 2^{16}$, which supports this conjecture.

Unfortunately our methods based on functional equations make it hard to track the joint law of $(P_n(I), Q_n(I))$ so we cannot prove (1.5). However, we can prove a result of similar strength

²I thank Philippe Biane for help identifying the variable from its moments, and Thomas Budzinski for suggesting to look for them in the OEIS, which is how I found the reference [PS02], see the sequence OEIS:A064352. Note that using the duplication formula the R.H.S. of (1.3) can also be written $\frac{\Gamma(3k/4)}{\Gamma(k/2)} 2^{-1+k/4} 3^{1-3k/4}$, see [Cha], and see also (5.2) for yet another expression.

for a slight variant of the problem. For a Dyck path P of size n and an integer $j \in [n]$, we let $\tilde{P}(j)$ be the initial height of the j -th up-step of P . It turns out that we can write a functional equation for the generating function corresponding to the joint law of $(\tilde{P}_n(J), \tilde{Q}_n(J))$ where J is uniform on $[n]$. We can prove:

Theorem 1.3. *Let J be uniform on $[n]$, and (P, Q) be uniform on \mathcal{I}_n . Then one has, when n goes to infinity,*

$$\mathbf{E}[(\tilde{Q}_n(J) - 3\tilde{P}_n(J))^2] = O(n). \tag{1.6}$$

In particular, we have, in probability,

$$\tilde{P}_n(J) = \frac{1}{3}\tilde{Q}_n(J) + o(\tilde{Q}_n(J)). \tag{1.7}$$

Note that if one is interested only in the individual convergence, studying $P_n(I)$ and $\tilde{P}_n(J)$ does not make much of a difference, thanks to the following well-known lemma:

Lemma 1.4. *Let P_n be a fixed Dyck path of size n . Then it is possible to couple a uniform variable $J \in [n]$ and a uniform variable $I \in [0..2n - 1]$ such that*

$$|P_n(I) - \tilde{P}_n(J)| \leq 1.$$

Proof. First sample $J \in [n]$ uniformly. Let I_1 be the initial abscissa of the J -th up-step of P , and let I_2 be the initial abscissa of the down-step that matches this step, viewing P as a parentheses word. Then let I be equal to I_1 or I_2 with probability $\frac{1}{2}$. It is clear that I is uniform on $[0..2n - 1]$ – the converse operation consists in considering the step starting at I , which is part of a pair of up and down steps matched together in the underlying parentheses word, and letting J be the index of the up-step. □

From the lemma and the fact that the case $I = 2n$ has a negligible contribution to moments, the individual convergences (in distribution and of all moments)

$$\frac{\tilde{P}_n(J)}{n^{3/4}} \rightarrow \frac{Z}{3}, \quad \frac{\tilde{Q}_n(J)}{n^{3/4}} \rightarrow Z$$

are equivalent respectively to the one of $P_n(I)$ and $Q_n(I)$. Note however that this coupling argument does not apply to their joints laws, which is why (1.5) and (1.7) are a priori incompatible statements. However we could deduce one from the other if we had some control on the regularity of the paths in their natural scaling limit.

As we said, Bernardi and Bonichon [BB09] provided an explicit bijection between intervals in \mathcal{I}_n and rooted plane triangulations of size n . Such a triangulation can always be equipped in a canonical way with a *realizer* or *Schnyder wood*, which is a partition of its internal edges into three subsets (say red, blue, green), such that the edges of each subset form a tree, with certain conditions. See Figure 1.3. In the Bernardi–Bonichon bijection, if the interval (P, Q) is in correspondence with the triangulation T , then P encodes the blue tree in the canonical Schnyder wood of T . Therefore we have:

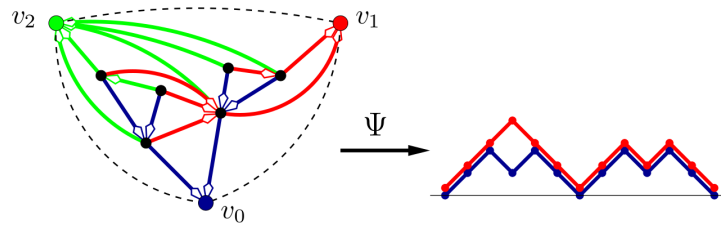


Figure 1.3: A figure taken from [BB09] (thanks to the authors). A rooted planar triangulation equipped with its minimal Schnyder-wood, and its image (a Tamari interval) by the Bernardi–Bonichon bijection. The lower path is nothing but the contour function of the blue tree.

Corollary 1.5. *Let T_n be a rooted plane triangulation of size n , chosen uniformly at random, and let $(T_n^{(1)}, T_n^{(2)}, T_n^{(3)})$ be its canonical Schnyder wood, that is to say, the one associated with its minimal orientation in the sense of [BB09]. Let V be a uniform random internal vertex of T_n and let $H_n^{(i)}$ be the height of the vertex V in the tree $T_n^{(i)}$. Then, for any $i \in \{1, 2, 3\}$ we have*

$$\frac{H_n^{(i)}}{n^{3/4}} \longrightarrow \frac{1}{3}Z,$$

where Z is as in Theorem (1.1).

Proof. The Bernardi–Bonichon bijection implies that $H_n^{(1)}$ has the same law as $1 + \tilde{P}_n(J)$ in previous notation. This implies the result for $i = 1$, and by symmetry for $i \in \{1, 2, 3\}$. \square

The proofs Theorems 1.1, 1.2, and 1.3 each have two parts: the first one consists in solving “exactly” the model, by obtaining an explicit algebraic equation for the generating function $f(t, s)$ of Tamari intervals having a marked abscissa, with control on the size n , and the random variable we want to control. This requires to solve an equation with catalytic variables, which is specific to each case. The second part is to deduce the asymptotic of moments from there, which is a problem of analytic combinatorics in 2 variables, for which we need to find good tools. We describe below a simple method that will do the trick, and that we believe is of independent interest.

1.2. A method to prove the convergence of moments from an algebraic equation

Assume that we have a combinatorial class \mathcal{A} equipped with a size function $|\cdot| : \mathcal{A} \rightarrow \mathbb{N}$ and an integer statistic $\mu : \mathcal{A} \rightarrow \mathbb{Z}$, and consider the random variable $X_n = \mu(A_n)$ where A_n is an object of size n in \mathcal{A} chosen uniformly at random. Consider the generating function

$$f(t, s) = \sum_{n \geq 0} \sum_{p \in \mathbb{Z}} a_{n,p} t^n s^p = \sum_{n \geq 0} \sum_{p \in \mathbb{Z}} a_n \mathbf{P}[X_n = p] t^n s^p, \quad (1.8)$$

with $a_{n,p} = |\{\alpha \in \mathcal{A} : |\alpha| = n, \mu(\alpha) = p\}|$, and $a_n = |\{\alpha \in \mathcal{A} : |\alpha| = n\}|$.

Let us assume that the generating function $f(t, s)$ is algebraic, i.e. there exists a nontrivial polynomial A such that

$$A(f(t, s), t, s) = 0.$$

For $k \geq 0$ we consider the generating function of factorial moments

$$f^{(k)} \equiv f^{(k)}(t) := \left(\left(\frac{d}{ds} \right)^k f(t, s) \right) \Big|_{s=1} = \sum_{n \geq 0} a_n \mathbf{E}[(X_n)_{(k)}] t^n,$$

with $(X_n)_{(k)} := X_n(X_n - 1) \cdots (X_n - k + 1)$. To study the convergence of the random variable X_n , a standard way called the method of moments consists in studying the asymptotics for fixed k of its k -th moment $\mathbf{E}[(X_n)^k]$ (or factorial moment $\mathbf{E}[(X_n)_{(k)}]$). See for example [Bil12, Sec. 30]. In our setting, this is equivalent to studying the asymptotics of the coefficient $[t^n]f^{(k)}(t)$, for fixed k . Now, since all the functions $f^{(k)}(t)$ are algebraic, they are amenable to singularity analysis in the sense of [FO90, FS09], as we will recall now.

In what follows, for a function g and $\alpha, c \in \mathbb{R}, \rho > 0$ we write

$$g(t) \sim c(1 - t/\rho)^\alpha, \quad t \rightarrow \rho,$$

if $g(t)$, as an analytic function, has no singularity on the closed circle of radius ρ except maybe at $t = \rho$, and if $g(t)$ has a Puiseux expansion of the form

$$g(t) = P(t) + c(1 - t/\rho)^\alpha + o((1 - t/\rho)^\alpha)$$

near $t = \rho$, where P is a polynomial. Let us now recall the classical transfer theorem, slightly reformulated³. Note that we allow $c = 0$ in the statement, and also $\alpha \in \mathbb{N}$ provided $c = 0$.

Proposition 1.6 (Transfer theorem for algebraic functions [FO90, FS09]). *Let $g(t) = \sum_{n \geq 0} g_n t^n$ be an algebraic function, with radius of convergence $\rho > 0$, and assume that $t = \rho$ is the unique singularity of g of minimal modulus. Assume that $g(t)$ has a singular expansion of the form*

$$g(t) \sim c(1 - t/\rho)^\alpha, \quad t \rightarrow \rho, \tag{1.9}$$

with $c \in \mathbb{R}$ and $\alpha \in \mathbb{R}$, and $c = 0$ if $\alpha \in \mathbb{N}$. Then the coefficients $g_n = [t^n]g(t)$ satisfy the asymptotics, when $n \rightarrow \infty$,

$$\frac{g_n}{n^{-1-\alpha}\rho^{-n}} \rightarrow \frac{c}{\Gamma(-\alpha)}.$$

Of course, when $\alpha \in \mathbb{N}$ and $c = 0$ we interpret the right-hand side as zero.

Thanks to the transfer theorem, to estimate moments we only need to determine the expansion (1.9) for each of the functions $g = f^{(k)}$. Here comes the main trick: since the function $f(t, s)$ is algebraic, it is also D-finite, i.e. the coefficients of its expansion in any variable satisfy a linear recurrence relation with polynomial coefficients (see e.g. [Com64] or [Sta80]). We will apply this to the coefficients of the expansion⁴ in the variable r such that $s = r + 1$. These coefficients are, up to a factorial factor, the functions $f^{(k)}$, i.e.

$$[r^k]f(t, 1 + r) = \frac{1}{k!} f^{(k)}(t).$$

³We will consider only the case of a unique singularity of minimal modulus. See Remark 1.9

⁴The function $f(t, 1 + r)$ is algebraic, therefore it is D-finite in the variable r . No notion of convergence is required to say this. Of course, one has to be careful about which branch of this function one considers when performing the actual calculations.

Therefore, we can write a recurrence formula of the form:

$$\sum_{d=0}^L P_d(t, k) \frac{f^{(k-d)}(t)}{(k-d)!} = 0, \quad k \geq L, \quad (1.10)$$

with $L \geq 1$ and P_0, \dots, P_L being bivariate polynomials over \mathbb{Q} . In practice, it will be useful to allow ourselves to work under some change of variable (of the variable t), and for this reason we will more generally assume that our algebraic function f satisfies (1.10) with P_0, \dots, P_L being polynomials in k whose coefficients are algebraic functions of t (this will be the case in our main applications of the method).

To sum up, one can compute the $f^{(k)}$ by induction, using

$$f^{(k)}(t) = \sum_{d=1}^L h_d(t, k) f^{(k-d)}(t), \quad k \geq L, \quad (1.11)$$

with $h_d(t, k) = -\frac{k!P_d(t, k)}{(k-d)!P_0(t, k)}$, which is a rational function in k whose coefficients are algebraic functions of t . This leads us to:

Method 1.7 (D-finite trick for moment pumping). *Given a bivariate algebraic function f , obtain a linear recurrence relation for its derivatives $f^{(k)}$ of the form (1.11) using standard computer algebra tools. Then, use it to determine the asymptotic of $f^{(k)}$ near $t = \rho$ by induction on k . Deduce the asymptotics of $[t^n]f^{(k)}$ using the transfer theorem (Proposition 1.6).*

The idea of this method is quite general, but of course one has to be careful to carry the analytical details in the induction (check that no unwanted singularity appears, for example; or be careful about multiple dominant singularities). Let us give a simple framework of application in the case of a unique dominant singularity in the next theorem, whose proof is essentially immediate.

Theorem 1.8 (D-finite trick for moment pumping, an instance). *Let $f(t, s)$ be a generating function of the form (1.8), and assume that $f(t, s)$ is an algebraic function. Then f is D-finite in the variable $s - 1$, therefore it satisfies an equation of the form (1.11), with $L \geq 1$ and where for $d \in [L]$, $h_d(t, k)$ is a rational function of k whose coefficients are algebraic functions of t . Assume that:*

(i) *There is $\beta > 0$ such that for each $d \in [L]$,*

$$h_d(t, k)(1 - t/\rho)^{\beta d} \rightarrow a_d(k), \quad t \rightarrow \rho,$$

with possibly $a_d(k) = 0$. Moreover, for $k \geq 1$, $h_d(t, k)$ has no other singularity than ρ on the closed disk of radius ρ .

(ii) *“Initial conditions”: there is $\alpha \in \mathbb{R} \setminus \mathbb{N}$ and numbers c_ℓ , with $c_0 > 0$, such that*

$$f^{(\ell)} \sim c_\ell (1 - t/\rho)^{\alpha - \beta \ell}$$

for all $0 \leq \ell \leq \ell_0$, with $\ell_0 := L + \max(\lfloor \alpha/\beta \rfloor, -1)$. Moreover, $c_\ell = 0$ if $\alpha - \beta \ell \in \mathbb{N}$.

Then one has $f^{(k)} \sim c_k(1 - t/\rho)^{\alpha - \beta k}$ for any $k \geq 0$, where c_k is given by the recurrence formula

$$c_k = \sum_{d=1}^L a_d(k)c_{k-d}, \quad k > \ell_0, \tag{1.12}$$

and the values of c_k for $k \leq \ell_0$ are given by the initial conditions. Moreover, for each $k \geq 0$,

$$\frac{\mathbf{E}[(X_n)^k]}{n^{\beta k}} \rightarrow \frac{c_k \Gamma(-\alpha)}{c_0 \Gamma(\beta k - \alpha)}. \tag{1.13}$$

Proof. We will show by induction that $f^{(\ell)}(t)$ has a unique singularity of minimal modulus at $t = \rho$, with $f^{(\ell)} = c_\ell(1 - t/\rho)^{\alpha - \beta \ell} + o((1 - t/\rho)^{\alpha - \beta \ell})$, for every $\ell \geq 0$ with $\ell > \frac{\alpha}{\beta}$.

For $\frac{\alpha}{\beta} < \ell \leq \ell_0$, this is true by the initial conditions (note that $\alpha - \beta \ell < 0$ for these values of ℓ , so the polynomial part in the definition of the symbol \sim has a subdominating contribution).

Now let $k \geq \ell_0 + 1$ and assume the induction hypothesis for $\frac{\alpha}{\beta} < \ell < k$. Note that for all $d \in [L]$, we have

$$k - d \geq \ell_0 + 1 - L \geq \lfloor \frac{\alpha}{\beta} \rfloor + 1 > \frac{\alpha}{\beta}$$

we can thus use the induction hypothesis for all the functions $f^{(k-d)}$ appearing in the right-hand side of (1.11). Using (1.11), the hypothesis on the singularity of h_d , and induction, it then immediately follows that

$$f^{(k)} = c_k(1 - t/\rho)^{\alpha - \beta k} + o((1 - t/\rho)^{\alpha - \beta k})$$

with c_k given by (1.12), and moreover that $f^{(k)}$ has a unique dominant singularity at ρ . This concludes the induction, and shows that we have $f^{(\ell)} \sim c_\ell(1 - t/\rho)^{\alpha - \beta \ell}$ for all $\ell \geq 0$.

By the transfer theorem, this implies that $[t^n]f^{(\ell)} \sim \frac{c_\ell}{\Gamma(\beta \ell - \alpha)} n^{\beta \ell - \alpha - 1} \rho^{-n}$. Since $\mathbf{E}[(X_n)_{(k)}] = \frac{[t^n]f^{(k)}}{[t^n]f^{(0)}}$, we obtain from the transfer theorem that

$$\frac{\mathbf{E}[(X_n)_{(k)}]}{n^{\beta k}} \rightarrow \frac{c_k \Gamma(-\alpha)}{c_0 \Gamma(\beta k - \alpha)}.$$

Since $\beta > 0$, the asymptotics (1.13) follows from the triangular relation between moments and factorial moments. □

Remark 1.9. We do not try to provide minimal hypotheses in this theorem: our goal is to emphasize the method, that seems applicable to various situations, including some D-finite non algebraic functions. The only requirement for the method to work is that dominant singularities of $f^{(k)}$ are “not too hard” to track by induction from the linear recurrence relation (1.11). Note also that there is no direct objection to trying to apply the method for more than bi-variate examples, i.e. for the asymptotic of joint moments of variables of the form $\mathbf{E}X_{n,1}^{k_1} \cdots X_{n,j}^{k_j}$ for fixed k_1, \dots, k_j when $n \rightarrow \infty$, provided the corresponding $(j + 1)$ -variate generating function is algebraic.

Remark 1.10. We insist that Theorem 1.8 if applicable, is essentially automatic. Indeed, computer algebra software (e.g. the Maple package gfun [SZ94]) are able to provide a recurrence of the form (1.11) automatically from an algebraic equation for f . Also, since all the $f^{(k)}$ are algebraic with a computable algebraic equation, it is in principle automatic to check the initial conditions required in the theorem for $0 \leq \ell \leq \ell_0$. All this will be illustrated by the proofs of Theorems 1.1-1.2.

Example 1.11. As a simple (and somewhat artificial) application, let us rediscover the well-known limit law for the height H_n of a uniform random point on a uniform random Dyck path of size n . Using standard methods (e.g. last passage decompositions), the corresponding generating function is found to be

$$f(t, s) = \frac{E(t)^2}{1 - stE(t)^2}, \quad E(t) = 1 + tE(t)^2.$$

Using a computer algebra system, it is immediate to eliminate $E(t)$ to find an algebraic equation for $f(t, s)$. We then find automatically a linear recurrence relation for its derivatives at $s = 1$. The Maple package gfun [SZ94] gives us:

$$\frac{f^{(k)}(t)}{k!} = -\frac{f^{(k-1)}}{(k-1)!} + \frac{t}{1-4t} \frac{f^{(k-2)}}{(k-2)!}, \quad k \geq 2. \quad (1.14)$$

This recurrence (multiplied by $k!$) fits in the framework of the theorem with $L = 2$, $a_1 = 0$, $a_2 = \frac{1}{4}k(k-1)$, and $\beta = 1/2$. It is then easy to check explicitly the initial conditions, which hold with $\alpha = -\frac{1}{2}$, thus $\ell_0 = 1$, and with $c_0 = 2$ and $c_1 = 1$. The recurrence (1.12) becomes $c_k = \frac{1}{4}k(k-1)c_{k-2}$, so that $c_k = 2^{1-k}k!$, and (1.13) immediately gives:

$$\frac{\mathbf{E}[(H_n)^k]}{n^{k/2}} \rightarrow \frac{2^{-k}k!\Gamma(\frac{1}{2})}{\Gamma(\frac{k}{2} + \frac{1}{2})} = \Gamma\left(\frac{k}{2} + 1\right).$$

This implies that H_n/\sqrt{n} converges in distribution to a random variable, which we recognize from its moments as a Rayleigh law of parameter $\sigma = 1/\sqrt{2}$ (with density $\frac{x}{\sigma^2}e^{-x^2/(2\sigma^2)}$ on $(0, \infty)$). This fact is well known, see e.g. [Ald91, p. 49].

In passing, we remark that we can also apply the theorem with a value of β larger than $1/2$. In this case we get $a_d \equiv 0$, and $c_k = 0$ for $k > 0$. We conclude that all positive of moments of H_n/n^β go to zero – which is true but weaker than the statement we obtained with the optimal value $\beta = 1/2$.

To conclude this section, let us recall that using D-finiteness to compute the coefficients of algebraic functions in the univariate case is a well-known trick, that goes back at least to Comtet [Com64] and is extensively used in the context of computer algebra. Here we are only recycling this idea in the context of bivariate asymptotics.

We also insist that the key idea of “moment pumping” is a standard tool to analyze Pólya urns and related random processes. While the underlying moment generating functions are (in general) not algebraic nor D-finite, they often satisfy partial differential equations that are amenable

to a recursive description of the derivatives at 1, see the references [BMW20, KP16, HN02] or again [FS09]. As far as we understand, a difference with the present paper is that in the context of these papers the recurrence relations are often “naturally inherited” from the combinatorial model, while in our case we obtain them from the “black box” argument of D-finiteness of algebraic functions.

1.3. Comments, questions, and plan of the paper

When we brought Theorem 1.1 to the attention of Nicolas Curien, he mentioned a book in preparation (joint with Jean Bertoin and Armand Riera) containing (among other things) general convergence results for a wide class of decomposition trees including the peeling trees of random planar maps and critical parking trees. These results will apply as well to Tamari intervals and describe the scaling limit of the (whole) upper path Q in terms of a certain conditioned self similar Markov tree which they introduce. It is not clear however that these methods lead to the explicit form of the limit law for the typical height obtained in this paper. In the other direction, we suspect that, all calculations made, the methods of our paper should enable to obtain the limit law of the height of a uniform random point in other models in this class, for example critical random Cayley parking trees (but we have not tried to do so). Note that the limiting first moment of the height of a random point for Cayley parking trees was already computed from generating functions in [CC23].

A related question is to ask for the universality of the random variable $Z^{1/4}$ appearing in our asymptotic theorems. It is now well understood that bivariate generating functions $F(t, x)$ satisfying non-linear polynomial equations with one catalytic variable give rise to asymptotic polynomial terms of the form $n^{-5/2}$, under natural positivity hypotheses for the equation (see [DNY22]). To each such equation, one can naturally associate a decomposition tree, that roughly speaking tracks the recursive calculation of coefficients of the function $F(t, x)$. It is natural to expect that the random variable $Z^{1/4}$ and the scaling exponent $n^{3/4}$ are universal in this context, possibly with slightly more restrictive hypotheses. We leave this question open – studying the behaviour of factorial moments by induction from the equation as we do here, introducing an extra variable s appropriately, could be a way to try to prove this, but combining the estimates of [DNY22] with the aforementioned probabilistic methods could be more direct.

The paper is organized as follows. In Section 2 we solve the exact counting problem which underlies Theorem 1.1 by studying the classical equation with one catalytic variable for Tamari intervals, which we enrich by an extra variable marking the height of a marked point. See Theorem 2.3.

In Section 3 we solve the exact counting problem which underlies Theorem 1.2 which deals with the height of the lower path. See Theorem 3.2. Our proof is more technical, as we only manage to write an equation with *two* catalytic variables for this problem (enriched, once again, by an extra variable for the height). It turns out that this equation with two catalytic variables is not too hard to solve, as it can be viewed as two nested equations with one catalytic variable each, which can be solved by standard methods.

In Section 4 we solve the exact counting problem which underlies Theorem 1.3 in which we manage to compare the heights of steps on both paths. See Theorem 4.2. The method is similar

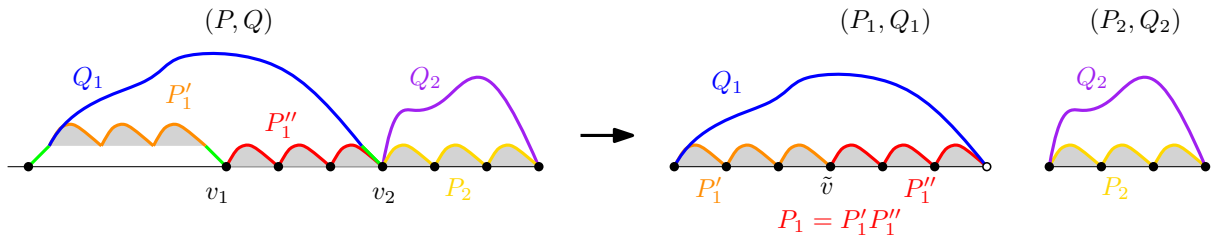


Figure 2.1: The classical decomposition of Tamari intervals. To the left, an interval of size $n + 1$, where v_1, v_2 are the first contacts of the lower and upper path, respectively. To the right, the decomposition gives rise to two Tamari intervals of total size n , the first of which has a marked contact, called here \tilde{v} . This construction is bijective.

to the one used in Section 3.

In Section 5, we prove Theorems 1.1 and 1.2 by performing the asymptotics of moments from the algebraic equations obtained in Theorems 2.3 and 3.2. The proofs are immediate verifications given Theorem 1.8 and computer algebra calculations. We also prove Theorem 1.3 which is a direct consequence of Theorem 4.2.

We will try in this paper to focus on the methods, and leave the actual calculations to the accompanying Maple worksheet [Cha]. Whether these calculations could be done or verified by hand will not be our concern.

2. Upper path: exact solution

2.1. The classical equation and its solution, after [Cha07, BMFPR11]

Let (P, Q) be a Tamari interval. Following [BMFPR11], we call *contact* of a path a vertex of that path lying on the x -axis. We now present a recursive decomposition of Tamari intervals based on contacts of the lower path, following [Cha07, BMFPR11]. See Figure 2.1 as a support for this discussion.

Let $n \geq 0$ and let $(P, Q) \in \mathcal{I}_{n+1}$ be a Tamari interval. Let v_1 and v_2 be respectively the leftmost contact of P and Q different from the origin. Then v_2 is also a contact of P and it is preceded by a down-step in both P and Q . We can thus split paths as $P = uP_0dP_2$, $Q = uQ_1dQ_2$, where we write paths as words with letters u/d representing up/down steps, respectively, and where P_2, Q_2 start at the contact v_2 . Moreover, v_1 is a contact of uP_0d (possibly equal to v_2), so we can write $uP_0d = uP'_1dP''_1$, with P''_1 possibly empty, and P'_1 starting at v_1 . We now define the path $P_1 := P'_1P''_1$, which is naturally equipped with a marked contact \tilde{v} . See Figure 2.1 again.

It is proved in [Cha07, BMFPR11] that (P_1, Q_1) and (P_2, Q_2) are two Tamari intervals. Moreover, this operation is a *bijection* (!) between intervals of size $n + 1$ and pairs of intervals of total size n , such that the first interval of the pair has a distinguished contact on its lower path. To translate this recursive construction into an equation for enumeration, one needs to introduce a two-parameter generating function of Tamari intervals:

$$F(x) \equiv F(t, x) := \sum_{n \geq 0} t^n \sum_{(P, Q) \in \mathcal{I}_n} x^{\text{contact}(P)}$$

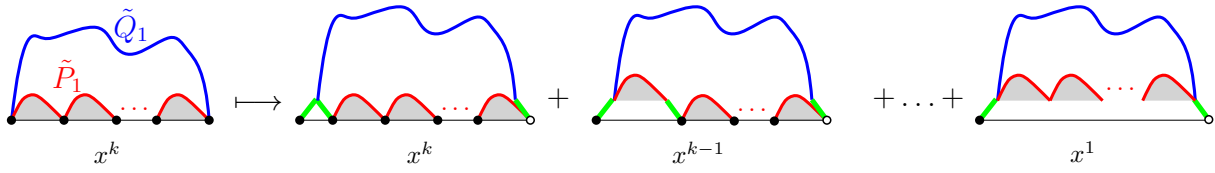


Figure 2.2: Tracking the number of contacts, and how the divided difference operator appears. On the left, the power of x marks all contacts, while on the right it only marks contacts which are not the last one.

where $\text{contact}(P)$ is the number of contacts of the lower path P .

Note that, if a path \tilde{P}_1 has k contacts, there are k possible ways to mark a contact in \tilde{P}_1 , thus decomposing it as $\tilde{P}_1 = \tilde{P}'_1 \tilde{P}''_1$. When going through these k choices, the number of *non-final* contacts of the path $\tilde{P}_0 := u\tilde{P}'_1 d\tilde{P}''_1$ goes through the values $1, 2, \dots, k$, see Figure 2.2. Therefore, at the level of generating functions, the operation of marking a contact in an interval $(\tilde{P}_1, \tilde{Q}_1)$ and transforming it into an interval of the form (\tilde{P}, \tilde{Q}) with this construction is taken into account by the operator:

$$\Delta_x : x^k \mapsto x^k + \dots + x^1 = x \frac{x^k - 1}{x - 1} \tag{2.1}$$

$$\Delta_x : f(x) \mapsto x \frac{f(x) - f(1)}{x - 1}. \tag{2.2}$$

This operator acts linearly on polynomials in x , and by extension on formal power series whose coefficients are polynomials in x . The reason why we choose to count only non-final contacts is to avoid counting the final contact twice when concatenating intervals (see Figure 2.1 again).

This observation being made, the recursive decomposition immediately translates into the following functional equation [Cha07, BMFPR11]

$$F(x) = x + xtF(x) \frac{F(x) - F(1)}{x - 1}. \tag{2.3}$$

Indeed, the first term accounts for the empty path (of size $n = 0$), and in the second term the factor $F(x)$ is the contribution of the interval (P_2, Q_2) , and the factor $x \frac{F(x) - F(1)}{x - 1} = \Delta_x F(x)$ is the contribution of the interval (P_1, Q_1) .

Equation (2.3) is a typical example of a *functional equation with one-catalytic variable*, which is the name given in the combinatorics literature⁵ to polynomial equations involving a bivariate series $F(t, x)$ as well as its specialisation at $x = 1$. Such equations are not easy to solve *a priori*, but fortunately a full theory has been developed in [BMJ06], which makes their resolution essentially automatic. In the present case, the solution can be written especially nicely with a rational parametrization as follows.

⁵We have been told that a part of the community is shifting towards the fancier terminology “discrete differential equation”.

Proposition 2.1 ([BMFPR11, Thm. 10]). *The generating functions $F(x)$ and $F(1)$ have the explicit rational parametrisations:*

$$F(x) = \frac{1+u}{(1+zu)(1-z)^3}(1-2z-z^2u) \quad , \quad F(1) = \frac{1-2z}{(1-z)^3}, \quad (2.4)$$

with

$$t = z(1-z)^3, \quad (2.5)$$

$$x = \frac{1+u}{(1+zu)^2}. \quad (2.6)$$

2.2. The enriched equation and its solution

One can apply the same decomposition as above to Tamari intervals in which the upper path Q carries a marked point: one only has to track recursively the position of the marked point into the different components of the decomposition. We introduce the generating function

$$H(x) \equiv H(t, x, s) := \sum_{n \geq 0} t^n \sum_{(P, Q) \in \mathcal{I}_n} x^{\text{contact}(P)} \sum_{i=0}^{2n} s^{Q(i)},$$

where we recall that $Q(i)$ is the height of the point of the path Q lying at abscissa i . It is clear that the height of the marked point can be tracked in the decomposition above. This leads to the functional equation:

Proposition 2.2. *The generating function $H(x) \equiv H(t, x, s)$ of Tamari intervals with a marked abscissa where s marks the upper height is solution of the equation:*

$$H(x) = F(x) + sxtF(x) \frac{H(x) - H(1)}{x-1} + xtH(x) \frac{F(x) - F(1)}{x-1}. \quad (2.7)$$

Proof. This follows directly from the combinatorial decomposition presented in the last section, applied to intervals with a marked abscissa, see Figure 2.3. The first term accounts for the case where the marked abscissa (and height) is zero. The second term accounts for the case where the marked abscissa appears before vertex v_2 . Through the decomposition the corresponding vertex of the upper path becomes a marked vertex of the path Q_1 , and its height is shifted by 1, hence a contribution of $s\Delta_x H(s, x)$ for the interval (P_1, Q_1) , while the contribution of the interval (P_2, Q_2) , which has no marked abscissa, is just $F(x)$. The third term accounts for the case where the marked abscissa appears at v_2 or after, in which case the corresponding vertex becomes a marked vertex of the path Q_2 , with no shift in the height, hence a contribution of $\Delta_x F(x)$ for (P_1, Q_1) and $H(x, s)$ for (P_2, Q_2) . \square

Our most skeptical readers might convince themselves (again) that we haven't forgotten any terms in the equation by noting that $H(t, x, 1) = (2\frac{xd}{dx} + 1)F(x)$, since there are $2n + 1$ possible abscissas to mark in an interval of size n , and that Equation (2.7) for $s = 1$ can be obtained by applying the operator $2\frac{xd}{dx} + 1$ to (2.3).

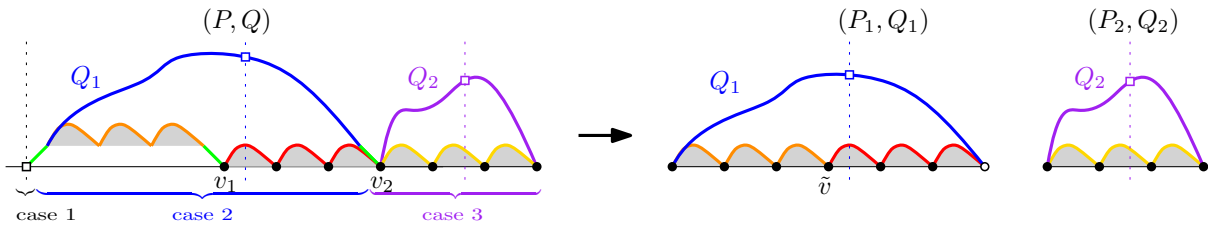


Figure 2.3: The decomposition of Figure 2.1, now with a marked abscissa. The three cases correspond to the three terms in Equation (2.7). In cases 2 and 3, the upper point at this abscissa becomes an upper point of the path Q_1 and Q_2 respectively, and in case 2 the height is shifted by one.

Note that, since $F(x)$ and $F(1)$ are known, this equation for $H(x)$ is again a polynomial equation with catalytic variable x , in which the variable s only plays the role of a parameter. In some sense, this equation is simpler than the previous one since it is *linear* in H . Linear equations with one catalytic variable are easy to solve, at least in principle, via the *kernel method*, which we now apply (see [FS09, p. 508] for an introduction to the kernel method).

We write (2.7) in kernel form as follows:

$$K(x)H(x) = F(x) - \frac{sxtF(x)H(1)}{x-1} \tag{2.8}$$

with

$$K(x) = \left(1 - sxt \frac{F(x)}{x-1} - xt \frac{F(x) - F(1)}{x-1} \right). \tag{2.9}$$

There is a unique formal power series $X_0(t) \in \mathbb{Q}[s][[t]]$ that is such that

$$K(X_0) = 0.$$

Indeed, multiplying by $x - 1$, this equation has the form $X_0(t) = 1 + t \text{Polynomial}(X_0(t), t, s)$ which enables one to compute the coefficients of $X_0(t)$ recursively. We will now apply the kernel method. The most obvious way to do it consists in substituting x for $X_0(t)$ and solving for $H(1)$, but we find more it convenient to work directly with the variables (u, z) rather than x, t . Expressing the kernel in terms of these variables we have (see [Cha] for the calculation)

$$\tilde{K} \equiv \tilde{K}(u) := K(x) \Big|_{\substack{t=z(1-z)^3 \\ x=\frac{1+u}{(1-zu)^2}}} = \frac{P(u, z, s)}{(uz+1)u(uz^2+2z-1)},$$

with

$$P(u, z, s) = s(1+u)^2z(uz^2+2z-1) + u(1-z)^3.$$

Since $P(u, z, s) = u - sz + O(z^2 + u^2 + zu)$, there is a unique series $U \equiv U(z, s) \in \mathbb{Q}[s][[z]]$ such that $P(U, z, s) = 0$. Its expansion starts as

$$U = sz + O(z^2).$$

Moreover, since P has degree 1 in s , we can also write s as a rational function of U and z ,

$$s = \frac{U(1-z)^3}{z(1+U)^2(1-Uz^2-2z)}. \quad (2.10)$$

We obtain

Theorem 2.3. *The series $H(1) \equiv H(t, 1, s)$ of Tamari intervals with a marked abscissa, where t marks the size, and s the upper height at this abscissa, has the following rational parametrisation:*

$$H(1) = \frac{(1-2z-Uz^2)^2(1+U)}{(1-z)^6} \quad (2.11)$$

with the change of variables $(t, s) \leftrightarrow (z, U)$ given by (2.5) and (2.10).

Proof. We substitute $t = z(1-z)^3$ and $x = \frac{1+u}{(1-zu)^2}$ in (2.8), and we further substitute $u = U(z, s)$, which cancels the left-hand side. From the right-hand side, and using the explicit expression (2.4) of $F(1)$ in terms of z we thus get an expression for $H(1)$ as a rational function of z , s , and $U = U(z, s)$, in which s can be substituted by (2.10), thus giving an expression for $H(1)$ as a rational function of U and z . The only thing to be done is to perform the explicit calculation. See [Cha]. \square

3. Lower path: exact solution

We will now apply the same decomposition as in the previous sections, but keep track of the height of points on the lower path P . In order to do this, we will have to treat differently the contacts of P which appear before or after the marked abscissa, which will force us to work with *two* catalytic variables. We write $\text{contact}_{<i}(P)$, $\text{contact}_{\geq i}(P)$ for the number of contacts of P strictly before, or weakly after, abscissa i , respectively.

3.1. The enriched equation

We introduce the generating function

$$G(x, y) \equiv G(t, x, y, w) := \sum_{n \geq 0} t^n \sum_{(P, Q) \in \mathcal{I}_n} \sum_{i=0}^{2n} w^{P(i)} x^{\text{contact}_{<i}(P)} y^{\text{contact}_{\geq i}(P)}.$$

We have

Proposition 3.1. *The generating function $G(x, y) \equiv G(t, x, y, w)$ of Tamari intervals with a marked abscissa where w marks the lower height is solution of the equation:*

$$\begin{aligned} G(x, y) = & F(y) + txw \frac{G(1, y) - G(1, 1)}{y-1} F(y) + tx \frac{F(y) - yF(1)}{y-1} F(y) \\ & + t \frac{x^2 G(x, y) - \frac{y}{x} F(x) - G(1, y) + yF(1)}{x-1} F(y) + tx \frac{F(x) - F(1)}{x-1} G(x, y). \end{aligned} \quad (3.1)$$

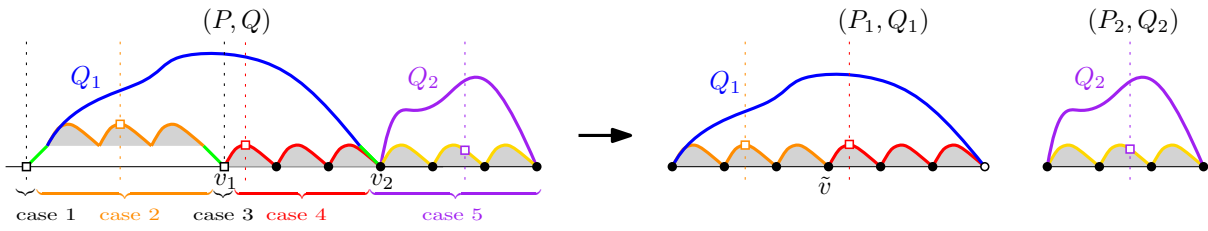


Figure 3.1: The decomposition of Figure 2.1, with a marked abscissa, but we now track the height of points of the lower path. There are now five cases, which correspond to the five terms in Equation (3.1).

Proof. Given an interval (P, Q) of size $n + 1$ with a marked abscissa, we apply again the decomposition of Section 2.1. As before we let v_1, v_2 , be the first non-initial contacts of the paths P and Q , respectively, and we also let i_1, i_2 be their abscissa. We let $i \in [0..2n]$ be the marked abscissa. We distinguish five cases (we will treat the third one last), see Figure 3.1,

Case 1: $i = 0$. Then we are only considering a Tamari interval, where all contacts are marked by the variable y . The corresponding contribution is $F(y)$.

Case 2: $i \in (0..i_1)$. In the decomposition, the corresponding vertex of P becomes a marked vertex of P_1 with a shift of 1 in the height, hence a contribution of w . Moreover, configurations in this case are obtained by applying the construction of Figure 2.2 but restricting it to contacts appearing after the marked abscissa, see Figure 3.2(top). Therefore, an interval $(\tilde{P}_1, \tilde{Q}_1)$ having a marked abscissa, with respectively k and ℓ contacts strictly before, and weakly after, this abscissa (thus having a contribution of $x^k y^\ell$ in $G(x, y)$) gives rise to ℓ intervals contributing to this case, with a contribution of

$$x^1 y^{\ell-1} + x^1 y^{\ell-2} + \dots + x^1 y^0 = x \frac{y^\ell - 1}{y - 1}.$$

In total, the contribution for the first interval is thus $x \frac{G(1,y) - G(1,1)}{y-1}$. The contribution of the second interval (P_2, Q_2) is just $F(y)$, since all corresponding contacts appear after the marked abscissa. In total, this case gives a contribution of $txw \frac{G(1,y) - G(1,1)}{y-1} F(y)$, which is the second term in (3.1).

Case 4: $i \in (i_1..i_2)$. In the decomposition, the vertex of P becomes a vertex of P_1 with no shift in height. Moreover, these configurations are obtained by applying the construction of Figure 2.2 but restricting it to contacts appearing before the marked abscissa, see Figure 3.2(bottom). More precisely, an interval $(\tilde{P}_1, \tilde{Q}_1)$ having a marked abscissa which is not the final one, with respectively k and ℓ contacts strictly before, and weakly after, this abscissa (thus having a contribution of $x^k y^\ell$ in $G(x, y)$) gives rise to k intervals contributing to this case, with a contribution of

$$x^{k+1} y^\ell + x^k y^{\ell-1} + \dots + x^2 y^{\ell-1} = \frac{x^2}{y} \frac{x^k - 1}{x - 1} y^\ell.$$

In total, the contribution for the first interval is thus

$$t \frac{x^2}{y} \frac{\tilde{G}(x, y) - \tilde{G}(1, y)}{x - 1},$$

where \tilde{G} is the same generating function as G , but in which the marked abscissa is not the last one. We have

$$\tilde{G}(x, y) = G(x, y) - \frac{y}{x} F(x),$$

since an interval with the last abscissa marked is the same as an unmarked interval (and the corresponding vertex has height zero), and since in such a case all contacts are counted with weight x except the last one which receives a weight y .

Moreover, the contribution of the second interval (P_2, Q_2) is just $F(y)$, since all corresponding contacts appear after the marked abscissa. In total, this case gives a contribution of $t \frac{x^2}{y} \frac{\tilde{G}(x, y) - \tilde{G}(1, y)}{x - 1} F(y)$, which is the fourth term in (3.1).

Case 5: $i \geq i_2$. In this case, the marked vertex becomes a marked vertex of Q_2 and there is not shift in height. The contribution of (P_2, Q_2) is simply $G(x, y)$, while the contribution of the first interval (P_1, Q_1) is $xt \frac{F(x) - F(1)}{x - 1}$ as in the classical case, since all its contacts appear before the marked abscissa. This gives the last term in (3.1).

Case 3: $i = i_1$ and $i_1 \notin \{0, i_2\}$. Such configurations are made of the concatenation of two intervals. The second one has no constraint, hence a contribution of $F(y)$. The first one has several constraints. First, its upper path has only two contacts. Arguing as in the classical decomposition, the generating function for intervals where the upper path has two contacts is

$$ty \frac{F(y) - F(1)}{y - 1}.$$

Second, in the first interval, the first nonzero contact of the lower path should be different from the one of the upper path (since we want $i_1 \neq i_2$ in the end). We thus have to subtract the corresponding contribution, and in total we find that the contribution of the first interval is

$$ty \frac{F(y) - F(1)}{y - 1} - tyF(1).$$

Putting things together, the generating function for case 3 is

$$\frac{x}{y} \left(ty \frac{F(y) - F(1)}{y - 1} - tyF(1) \right) F(y),$$

where the prefactor $\frac{x}{y}$ accounts for the initial contact. This is the third term in (3.1). \square

For skeptics, we have (successfully) tested the expansion of $G(x, y)$ given by Equation (3.1) up to order $n = 8$, using an exhaustive generation of Tamari lattices made by independent means.

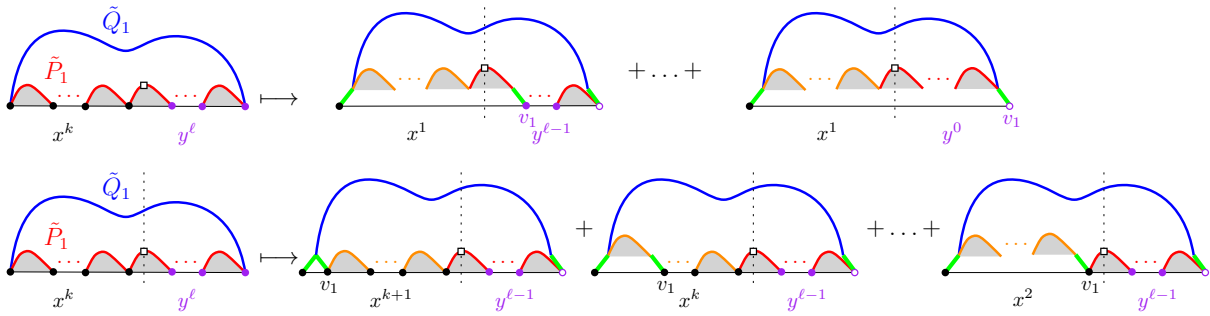


Figure 3.2: Refining the operator of Figure 2.2 to distinguish the case where the marked abscissa is before (top) or after (bottom) the contact v_1 .

3.2. Solution

Although equations with two catalytic variables are notoriously difficult, Equation (3.1) is of a very particular kind. First, it is again *linear* in the main unknown $G(x, y)$. Second, the equation involves the specializations $G(x, y)$, $G(1, y)$, and $G(1, 1)$, but not $G(x, 1)$. As we will see, this will enable us to treat this equation as two nested equations, each having only one catalytic variable⁶. The reader interested in more difficult cases of equations with two catalytic variables might consult, for example, the reference [BBMR21] for entry points into this fast growing literature.

3.2.1 First step: eliminating variable x (or u).

We start by writing Equation (3.1) in kernel form

$$K_2(x, y)G(x, y) = F(y) + txw \frac{G(1, y) - G(1, 1)}{y - 1} F(y) + tx \frac{F(y) - yF(1)}{y - 1} F(y) + t \frac{x^2 - \frac{y}{x} F(x) - G(1, y) + yF(1)}{x - 1} F(y), \quad (3.2)$$

with

$$K_2(x, y) = 1 - \frac{tx^2 F(y)}{y(x - 1)} - xt \frac{F(x) - F(1)}{x - 1}. \quad (3.3)$$

We will now work under the change of variables (2.5)–(2.6). Since we have two catalytic variables, we introduce a new variable v , which is to y what u is to x . To summarize, we write

$$t = z(1 - z)^3, \quad x = \frac{1 + u}{(1 + zu)^2}, \quad y = \frac{1 + v}{(1 + zv)^2}, \quad (3.4)$$

⁶After we prepared this work, Mireille Bousquet-Mélou has informed us of a paper of her in preparation with Hadrien Notarantonio, studying such equations with several catalytic variables where some variable specializations are missing. They also solve them iteratively as “nested” equations as we do here, and they prove that this works in a generic context with mild hypotheses. It is not clear that their theorems apply to our case since our equations also involve the known function $F(x)$ and are thus not directly polynomial, but we definitely recommend having a look at that paper when it is out!

We write respectively $\tilde{G}(u, v)$, $\tilde{G}_1(v)$, \tilde{G}_{11} , \tilde{K}_2 for the quantities $G(x, y)$, $G(1, y)$, $G(1, 1)$, K_2 expressed in the variables z, u, v after the substitutions (3.4).

By (2.4), the kernel \tilde{K}_2 is an explicit rational function of z, u, v . Its numerator is a polynomial $\tilde{k}_2(z, u, v, w)$ which, computations made [Cha], is given by

$$\tilde{k}_2(u, v) = ((vz^2 + z^2)u + vz^2 + 2z - 1)((vz^2 - z^2 + 3z - 1)u + vz^2 + z).$$

The equation $\tilde{k}_2 = 0$ has a unique root $U_0 \equiv U_0(z, v)$ which is a power series in z , given by the root of the second factor. Explicitly, it is given by $U_0 = \frac{z(1+uz)}{1-3z+z^2-vz^2} = z + O(z^2)$.

Using the substitutions (3.4) and using the known expressions of $F(x)$, $F(y)$, Equation (3.2) takes the shape

$$\tilde{K}_2(u, v)\tilde{G}(u, v) = \tilde{L}_2(z, u, v, \tilde{G}_1(v), \tilde{G}_{11})$$

for some rational function \tilde{L}_2 that can be written explicitly (its numerator has 228 terms, [Cha]). Substituting $u = U_0$ in this equation, we cancel the left-hand side, hence we also cancel the right-hand side. We are thus left with the following polynomial equation satisfied by $z, v, w, \tilde{G}_1(v)$ and \tilde{G}_{11} ,

$$\tilde{L}_2(z, U_0(z, v), v, \tilde{G}_1(v), \tilde{G}_{11}) = 0. \quad (3.5)$$

The numerator of this equation has 91 terms, but it has only degree one in $\tilde{G}_1(v)$. At this stage, we have eliminated the unknown $\tilde{G}(u, v)$ and the variable u .

3.2.2 Second step: eliminating variable y (or v).

At this stage, we have shown that the series $\tilde{G}_1(v)$ and \tilde{G}_{11} satisfy the explicit polynomial equation (3.5). But we recognize again this equation as an equation with one catalytic variable, which is now the variable v ! Even more, since the equation is linear in $\tilde{G}_1(v)$, we can just use the kernel method again. We write the numerator of (3.5) in the form

$$a(\tilde{G}_{11}, v, z, w)\tilde{G}_1(v) + b(\tilde{G}_{11}, v, z, w) = 0,$$

where in fact $a(\tilde{G}_{11}, v, z, w) = a(v, z, w)$ does not depend on \tilde{G}_{11} . A direct check shows that the equation $a(V_0, z, w) = 0$ as a unique solution $V_0(z)$ which is a formal power series in z , with $V_0 = zw + O(z^2)$. Substituting $v = V_0$ in the equation, we finally obtain the equation

$$b(\tilde{G}_{11}, V_0(z), z, w) = 0,$$

which shows that \tilde{G}_{11} is algebraic. Explicitly, we can eliminate the variable v from the equations $a(v, z, w)$ and $b(\tilde{G}_{11}, v, z, w)$, and we obtain a polynomial equation for the function \tilde{G}_{11} . It turns out that this equation is not even so big, and computations made [Cha] we finally obtain:

Theorem 3.2. *The generating function $G(x, y)$ is an algebraic function. Moreover, the function $G(1, 1)$ after the change of variables $t = z(1 - z)^3$ satisfies the polynomial equation $C(G(1, 1), z, w) = 0$ with*

$$\begin{aligned} C(h, z, w) = & wz(-1 + z)^9 h^3 + (-1 + z)^6 (2w^2 z^2 - w^2 z + 2z^2 + w - z) h^2 \\ & - (-1 + z)^3 (w^2 z^3 - 3w^2 z^2 - 2wz^3 + w^2 z - 2wz^2 + z^3 + 5wz - 3z^2 - 2w + z) h \\ & + 4wz^2 - 4wz + w. \end{aligned} \quad (3.6)$$

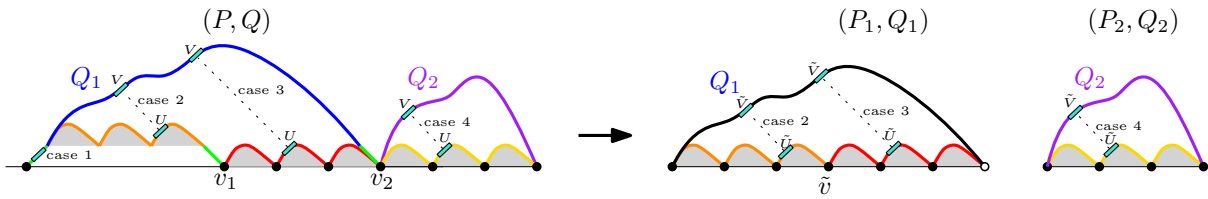


Figure 4.1: The decomposition of Figure 2.1, where the j -th up-step for some $j \in [n]$ is marked on both paths. There are four cases to consider, corresponding to the four terms in Equation (4.1).

4. Tracking marked up-steps on both paths

We will now apply the same decomposition as in the previous sections, but consider paths with a marked up-step, rather than a marked abscissa. As mentioned in the introduction, the reason to do that is that in this setting we are able to study the joint law of the heights of the two paths. Our goal is to prove Theorem 1.3.

Recall that for a Dyck path P we let $\tilde{P}(j)$ be the height (of the initial point of) the j -th up-step of the path P . We write $\text{contact}'_{\leq j}(P)$, $\text{contact}'_{\geq j}(P)$ for the number of contacts of P weakly before, or after, the j -th up-step of P , respectively.

4.1. The enriched equation

We introduce the generating function

$$J(x, y) \equiv J(t, x, y, r, s) := \sum_{n \geq 0} t^n \sum_{(P, Q) \in \mathcal{I}_n} \sum_{j=1}^n r^{\tilde{P}(j)} s^{\tilde{Q}(j)} x^{\text{contact}'_{\leq j}(P)} y^{\text{contact}'_{\geq j}(P)}.$$

We have

Proposition 4.1. *The generating function $J(x, y) \equiv J(t, x, y, r, s)$ of Tamari intervals where the j -th up-step of each path is marked for some integer j , where r marks the lower height and s the upper height is solution of the equation:*

$$J(x, y) = x \frac{F(y) - y}{y} + rstx \frac{J(1, y) - J(1, 1)}{y - 1} F(y) + stx^2 \frac{J(x, y) - J(1, y)}{(x - 1)y} F(y) + xtJ(x, y) \frac{F(x) - F(1)}{x - 1}. \quad (4.1)$$

Proof. The proof is very similar to the proof of Proposition 3.1, except that we keep track of the height on both paths. It is actually less involved since there are less cases to consider. Given an interval (P, Q) of size $n + 1$ with a marked $j \in [n]$, we apply again the decomposition of Section 2.1. We let U, V be the j -th up-step of P and Q , respectively. As before we let v_1, v_2 , be the first non-initial contacts of the paths P and Q . We distinguish four cases, see Figure 4.1,

Case 1: $U = V$ is the first step of the path ($j = 1$). In this case we are only counting intervals by contacts as in the series $F(y)$. However we have to exclude the empty interval (hence a correcting term $-y$), and we should count all contacts with weight y except for the first one, hence a correcting factor x/y . This gives the first term in (4.1).

Case 2: U is not the initial step of P , but U appears before v_1 . In this case V belongs to Q_1 . In this case, after applying the decomposition, U, V give rise naturally to marked contacts \tilde{U}, \tilde{V} of P_1, Q_1 appearing before the vertex \tilde{v} . The height of both steps is shifted by 1, hence the factor rs . The remaining factors are understood in a similar way as in case 2 of the proof of Proposition 3.1. Namely, an interval $(\tilde{P}_1, \tilde{Q}_1)$ with marked up-steps \tilde{U}, \tilde{V} , with respectively k and ℓ contacts before and after \tilde{U} (thus having a contribution of $x^k y^\ell$ in $J(x, y)$) gives rise to ℓ intervals contributing to this case, with a contribution of

$$x^1 y^{\ell-1} + x^1 y^{\ell-2} + \dots + x^1 y^0 = x \frac{y^\ell - 1}{y - 1}.$$

In total, the contribution for the first interval is thus $x \frac{J(1, y) - J(1, 1)}{y - 1}$. The contribution of the second interval (P_2, Q_2) is just $F(y)$, since all corresponding contacts appear after the marked abscissa. In total, we obtain the second term in (4.1).

Case 3: U appears (weakly) after v_1 but before v_2 . In this case again V belongs to Q_1 . In this case, after applying the decomposition, U, V give rise naturally to marked contacts of P_1, Q_1 appearing (weakly) after the vertex \tilde{v} . Only the height of the steps of the upper path is shifted by 1, hence the factor s . The rest of factors are understood in a similar way as in case 4 of the proof of Proposition 3.1, but is simpler. Namely, an interval $(\tilde{P}_1, \tilde{Q}_1)$ with marked up-steps \tilde{U}, \tilde{V} , with respectively k and ℓ contacts before and after \tilde{U} (thus having a contribution of $x^k y^\ell$ in $G(x, y)$) gives rise to k intervals contributing to this case, with a contribution of

$$x^{k+1} y^\ell + x^k y^{\ell-1} + \dots + x^2 y^{\ell-1} = \frac{x^2}{y} \frac{x^k - 1}{x - 1} y^\ell.$$

In total, the contribution for the first interval is thus

$$t \frac{x^2}{y} \frac{J(x, y) - J(1, y)}{x - 1}.$$

(Note that the situation is simpler than in Proposition 3.1 since we do not have to deal with the subtle case where the vertex whose height is considered is the last on the path – since we mark only *up* steps). Moreover, the contribution of the second interval (P_2, Q_2) is just $F(y)$, and in total we obtain the third term in (4.1).

Case 4: U appears (weakly) after v_2 . In this case V belongs to Q_2 . This case is similar to the last case of the proof of Proposition 3.1 and directly gives the fourth term in (4.1) \square

4.2. The specialized equation

In view of proving Theorem 1.3, we will consider the following specialisation of the function $J(x, y)$:

$$M(x, y) \equiv M(t, x, y, w) := J(x, y) \Big|_{s=w, r=w^{-3}} \\ = \sum_{n \geq 0} t^n \sum_{(P, Q) \in \mathcal{I}_n} \sum_{j=1}^n w^{\tilde{Q}(j) - 3\tilde{P}(j)} x^{\text{contact}'_{\leq j}(P)} y^{\text{contact}'_{\geq j}(P)}.$$

From Proposition 4.1 it satisfies the equation

$$M(x, y) = x \frac{F(y) - y}{y} + w^{-2} t x \frac{M(1, y) - M(1, 1)}{y - 1} F(y) + w t x^2 F(y) \frac{M(x, y) - M(1, y)}{(x - 1)y} \\ + x t \frac{F(x) - F(1)}{x - 1} M(x, y). \quad (4.2)$$

We will now solve this equation.

4.3. Solving the equation

Our method to solve (4.2) is similar to the one used in Section 3.2, except that the equation to be solved in the second step will now be quadratic instead of linear.

4.3.1 First step: eliminating variable x (or u).

We start by writing Equation (4.2) in kernel form

$$\mathbf{K}_2(x, y) M(x, y) = x \frac{F(y) - y}{y} + w^{-2} t x \frac{M(1, y) - M(1, 1)}{y - 1} F(y) + w t x^2 F(y) \frac{-M(1, y)}{(x - 1)y} \quad (4.3)$$

with

$$\mathbf{K}_2(x, y) = 1 - w t x^2 F(y) \frac{1}{(x - 1)y} - x t \frac{F(x) - F(1)}{x - 1}. \quad (4.4)$$

We will again work under the change of variables (3.4). We write respectively $\tilde{M}(u, v)$, $\tilde{M}_1(v)$, \tilde{M}_{11} , $\tilde{\mathbf{K}}_2$ for the quantities $M(x, y)$, $M(1, y)$, $M(1, 1)$, \mathbf{K}_2 expressed in the variables z, u, v after the substitutions (3.4).

By (2.4), the kernel $\tilde{\mathbf{K}}_2$ is an explicit rational function of z, u, v . Its numerator is a polynomial $\tilde{\mathbf{k}}_2(u, v) \equiv \tilde{\mathbf{k}}_2(z, u, v, w)$ which, computations made [Cha], is given by

$$\tilde{\mathbf{k}}_2(u, v) = u^2 v^2 w z^4 + 2 u v^2 w z^4 + 3 u^2 v w z^3 + v^2 w z^4 - u^2 v w z^2 - u^2 z^4 \\ + 6 u v w z^3 + 2 u^2 w z^2 + 3 u^2 z^3 - 2 u v w z^2 + 3 v w z^3 - u^2 w z - 3 u^2 z^2 \\ + 4 u w z^2 - u z^3 - v w z^2 + u^2 z - 2 u w z + 3 u z^2 + 2 w z^2 - 3 u z - w z + u.$$

The equation $\tilde{\mathbf{k}}_2 = 0$ has a unique root $\mathbf{U}_0 \equiv \mathbf{U}_0(z, v)$ which is a power series in z . It satisfies $\mathbf{U}_0 = wz + O(z^2)$.

Making the substitutions (3.4) and using the known expressions of $F(x)$, $F(y)$, Equation (4.3) takes the form

$$\tilde{\mathbf{K}}_2(u, v)\tilde{M}(u, v) = \tilde{\mathbf{L}}_2(z, u, v, \tilde{M}_1(v), \tilde{M}_{11})$$

for some rational function $\tilde{\mathbf{L}}_2$ that can be written explicitly (its numerator has 99 terms, [Cha]). Substituting $u = \mathbf{U}_0$ in this equation, we cancel the left-hand side, hence we also cancel the right-hand side. We are thus left with the following polynomial equation satisfied by $z, v, w, \tilde{M}_1(v)$ and \tilde{M}_{11} ,

$$\tilde{\mathbf{L}}_2(z, \mathbf{U}_0(z, v), v, \tilde{M}_1(v), \tilde{M}_{11}) = 0. \quad (4.5)$$

We can eliminate \mathbf{U}_0 between this equation and $\tilde{\mathbf{k}}_2(\mathbf{U}_0(z, v), v)$, and we thus obtain [Cha] a polynomial equation of the form

$$B(z, v, w, \tilde{M}_1(v), \tilde{M}_{11}) = 0. \quad (4.6)$$

This polynomial has 854 terms, and it has degree two in $\tilde{M}_1(v)$. At this stage, we have eliminated the unknown $\tilde{M}(u, v)$ and the variable u .

4.3.2 Second step: eliminating variable y (or v).

Let us write polynomial equation (4.6) in the form.

$$\mathbf{a}(\tilde{M}_{11}, v, z, w)\tilde{M}_1(v)^2 + \mathbf{b}(\tilde{M}_{11}, v, z, w)\tilde{M}_1(v) + \mathbf{c}(\tilde{M}_{11}, v, z, w) = 0.$$

We will solve this equation using (a variant of) the quadratic method (see [BMJ06] and references therein). Introduce notation:

$$B(v) := \mathbf{a}(v)\tilde{M}_1(v)^2 + \mathbf{b}(v)\tilde{M}_1(v) + \mathbf{c}(v) \quad , \quad B_2(v) := 2\mathbf{a}(v)\tilde{M}_1(v) + \mathbf{b}(v), \quad (4.7)$$

where in the notation we only indicate dependencies in the variable v , for example $\mathbf{a}(v) = \mathbf{a}(\tilde{M}_{11}, v, z, w)$. We first claim that there exists a formal power series $V_0(z) = zw^{-1} + O(z^2)$ such that $B_2(V_0(z)) = 0$. To see this, we generate the first few terms of the series $\tilde{M}_1(v)$ to see [Cha] that

$$B_2(v) = a_{2,0}v^2 + a_{1,1}vz + a_{1,2}vz^2 + a_{2,1}v^2z + a_{3,0}v^3 + O_4(z, v)$$

where the $a_{i,j}$ are explicit polynomials in w and where $O_4(z, v)$ is a formal power series in z, v (and coefficients polynomial in w) with all terms of homogeneous degree at least 4. Therefore the equation $B_2(V_0) = 0$ can be written

$$V_0 = -(a_{2,0})^{-1}(a_{1,1}z + a_{1,2}z^2 + a_{2,1}V_0z + a_{3,0}V_0^2 + V_0^{-1}O_4(z, V_0)).$$

This is enough to see that the equation $B_2(V_0) = 0$ has a formal power series root $V_0(z)$ of the form $V_0(z) = \frac{-a_{1,1}}{a_{2,0}}z + O(z^2)$. Indeed all coefficients of V_0 can be computed inductively from this

germ. This proves the claim (note that this is essentially a Newton–Puiseux argument). Making constants explicit, we have $V_0(z) = zw^{-2} + O(z^2)$. There is also another power series root V_0 with constant coefficient -1 but we will not use it.

We can now apply (a variant of) the quadratic method. We have two equations involving the series V_0 : the one that defines it, and the one obtained by substituting $v = V_0$ in (4.6). Namely:

$$B(z, V_0, w, \tilde{M}_1(V_0), \tilde{M}_{11}) = 0 \quad , \quad B_2(z, V_0, w, \tilde{M}_1(V_0), \tilde{M}_{11}) = 0.$$

We can eliminate the series $\tilde{M}_1(V_0)$ between these two equations (which amounts to saying that the discriminant $b^2 - 4ac$ vanishes), and we obtain an equation of the form

$$P(z, V_0, w, \tilde{M}_{11}) = 0.$$

Computations made [Cha], this equation has three nontrivial irreducible factors, P_1, P_2, P_3 , where only P_3 depends (in fact, linearly) on \tilde{M}_{11} . It is a direct check that none of the v -roots of P_1 and P_2 is equal to V_0 (in fact none is a formal power series in z), from which we conclude that $P_3(z, V_0, w, \tilde{M}_{11}) = 0$. This equation can be written as a rational expression of \tilde{M}_{11} in the variable V_0 , namely [Cha] we have $\tilde{M}_{11} = \mathbf{R}(z, V_0)$ with

$$R(z, V) = \frac{(Vz^2 - z^2 + 3z - 1)Vw^2}{Vz^4 - 3Vz^3 + 3Vz^2 + z^3 - Vz - 3z^2 + 3z - 1}. \tag{4.8}$$

One can check that substituting $v = V_0$ and this expression of \tilde{M}_{11} cancels both \mathbf{b} and \mathbf{c} . Therefore it also cancels \mathbf{a} , which gives us an explicit equation satisfied by V_0 , namely [Cha] we have $S(V_0, z) = 0$ with:

$$\begin{aligned} S(V, z, w) = & V^5w^5z^5 + V^5w^4z^5 + V^5w^3z^5 + 5V^4w^5z^4 + V^5w^2z^5 - 2V^4w^5z^3 \\ & + 5V^4w^4z^4 - V^3w^4z^5 + V^5wz^5 - 2V^4w^4z^3 + 5V^4w^3z^4 + 2V^4w^2z^5 + 8V^3w^5z^3 \\ & + 3V^3w^4z^4 - V^3w^3z^5 - 2V^4w^3z^3 + 4V^4w^2z^4 + 2V^4wz^5 - 6V^3w^5z^2 + 5V^3w^4z^3 \\ & + 3V^3w^3z^4 + V^3w^2z^5 - 2V^2w^4z^4 - V^4w^2z^3 + 4V^4wz^4 + V^3w^5z - 5V^3w^4z^2 + 5V^3w^3z^3 \\ & + 8V^3w^2z^4 + V^3wz^5 + 4V^2w^5z^2 + 7V^2w^4z^3 - 2V^2w^3z^4 - V^4wz^3 + V^3w^4z - 5V^3w^3z^2 \\ & + V^3w^2z^3 + 8V^3wz^4 - 4V^2w^5z - 5V^2w^4z^2 + 7V^2w^3z^3 + 4V^2w^2z^4 + V^3w^3z - V^3w^2z^2 \\ & + V^3wz^3 + V^2w^5 + V^2w^4z - 5V^2w^3z^2 + 6V^2w^2z^3 + 4V^2wz^4 - V^3wz^2 - 2V^3z^3 + V^2w^3z \\ & - 5V^2w^2z^2 + 5V^2wz^3 + 4Vw^2z^3 + V^3z^2 + 3V^2w^2z - 2V^2wz^2 - 4V^2z^3 - 4Vw^2z^2 + 3Vwz^3 \\ & - V^2w^2 + 3Vw^2z - Vwz^2 - 2Vz^3 + V^2z - Vw^2 - 3Vz^2 + 2Vz - 2z^2 + z. \end{aligned} \tag{4.9}$$

We have thus proved:

Theorem 4.2. *The generating function $M(1, 1) \equiv M(1, 1; t, w)$ is an algebraic function given by*

$$M(1, 1) = R(z, V_0)$$

where the change of variables $(t, w) \leftrightarrow (z, V)$ is given by (2.5) and by the equation $S(V_0, z, w) = 0$.

By eliminating V_0 in (4.8)-(4.9) one obtains an explicit polynomial cancelling $M(1, 1)$. Written in the variables z and w , it has 601 terms and degree five in $M(1, 1)$. See [Cha]. Given the previous calculations it is not difficult to see that the full series $M(x, y)$ is algebraic but we will never use it.

We leave to our readers the task of proving algebraicity of the full function $J(1, 1)$ (with both variables r and s) rather than of its specialization $M(1, 1)$ (involving only w) as we do here. We suspect that this is possible with similar methods, but by lack of application in mind we have not tried to do so – therefore we do not know if the calculations are manageable nor if unexpected difficulties arise.

5. Asymptotics of moments

In this section we prove Theorems 1.1 and 1.2. In both cases this will be a direct application of the method of Section 1.2, in particular Theorem 1.8, up to computer algebra calculations done in [Cha]. We will also prove Theorem 1.3 which is a direct check.

5.1. Proof of Theorem 1.1

We start with Equation (2.7) in Theorem 2.3. This equation shows that $f(t) = H(t, 1, s) \equiv H(1)$ is algebraic, therefore $H(x, r + 1)$ is algebraic. Therefore (see e.g. [Com64] again), there exists a recurrence formula of the form (1.11) to compute its derivatives

$$f^{(k)} \equiv f^{(k)}(t) := \left(\left(\frac{d}{ds} \right)^k H(1) \right) \Big|_{s=1}.$$

Given the form of our parametrization, we will prefer to work under the variable z given by (2.5), and in fact it is even more convenient to introduce the variable $\delta = (1 - 4z)^{1/2}$. It is easily checked [Cha] that the main singularity of $H(1)$ is at $t = \rho := \frac{27}{256}$, which corresponds to $z = \frac{1}{4}$, i.e. $\delta = 0$. In what follows we will still write $f^{(k)}$ for the function $f^{(k)}(t)$ expressed in the variable δ .

Using the package gfun, we immediately find [Cha] that the $f^{(k)}$ satisfy an equation of the form (1.11) with $L = 6$, namely

$$f^{(k)}(t) = \sum_{d=1}^L h_d(t, k) f^{(k-d)}(t), \quad k \geq L,$$

where the $h_d(t, k)$ for $d = 1..6$ are Laurent polynomials in δ with coefficients which are rational functions of k . The degree of h_1, \dots, h_6 in δ^{-1} are respectively 0, 6, 6, 10, 10, 10. Since δ behaves as $(1 - t/\rho)^{1/4}$, this implies that hypothesis (i) of Theorem 1.8 holds with $\beta = \frac{3}{4}$. One can explicitly check the values of the corresponding constants $a_d(k)$, which are nothing but top coefficients of h_d in δ^{-1} up to a scaling factor. They are given [Cha] by

$$a_1(k), \dots, a_6(k) = 0, \frac{\sqrt{6}}{96}(3k - 4)(3k - 8), 0, 0, 0, 0. \quad (5.1)$$

Now it only remains to check the initial conditions (ii). From the explicit Equation (2.7), the dominant singularity of $f^{(0)}$, and indeed of each given $f^{(i)}$, can be computed automatically. We find that $f^{(0)} = \frac{64}{(\delta+3)^3}$, with singular expansion $f^{(0)} \sim c_0(1-t/\rho)^\alpha$ with $\alpha = \frac{1}{2}$ and $c_0 = -\frac{32\sqrt{6}}{27}$.

Given α, β we compute $\ell_0 = L - 0 = 6$, and to check the initial conditions we need to estimate the main singularity of $f^{(1)}, \dots, f^{(6)}$. This is done automatically and one gets $f^{(\ell)} \sim c_\ell(1-t/\rho)^{\alpha-\beta\ell}$ for $\ell \geq 6$, with

$$c_0, \dots, c_6 = -\frac{32}{27}\sqrt{6}, \frac{16}{27}3^{3/4}\sqrt[4]{2}, \frac{8}{27}, \frac{5}{54}\sqrt[4]{32}^{3/4}, \frac{8}{81}\sqrt{2}\sqrt{3}, \frac{385}{2592}3^{3/4}\sqrt[4]{2}, \frac{70}{81}.$$

We have now verified all hypotheses of Theorem 1.8. The main recurrence formula becomes

$$c_k = \frac{\sqrt{6}(3k-4)(3k-8)}{96}c_{k-2}, \quad k > 6,$$

with initial conditions above. It is a direct check that the solution of this recurrence is given by

$$c_k = \frac{16}{27} \frac{\Gamma\left(\frac{k}{2} + \frac{1}{3}\right) \Gamma\left(\frac{k}{2} - \frac{1}{3}\right) \sqrt{2} 4^{-k} 6^{3/4k}}{\pi}.$$

for all $k \geq 0$. Applying Theorem 1.8, we immediately get:

$$\mathbf{E} \left[\left(\frac{Q_n(I)}{n^{3/4}} \right)^k \right] \rightarrow \frac{\Gamma\left(\frac{k}{2} + \frac{1}{3}\right) \Gamma\left(\frac{k}{2} - \frac{1}{3}\right) 4^{-k} 6^{3/4k}}{\sqrt{3}\pi \Gamma\left(\frac{3k}{4} - \frac{1}{2}\right)} = \frac{\sqrt{3} \cdot 2^{-\frac{k}{4}-1} \Gamma\left(\frac{1}{4}k + \frac{1}{3}\right) \Gamma\left(\frac{1}{4}k + \frac{2}{3}\right)}{\sqrt{\pi} \Gamma\left(\frac{1}{4}k + \frac{1}{2}\right)}, \tag{5.2}$$

where the equality follows from the duplication formula for the Gamma function (see also [Cha]). This proves Theorem 1.1.

5.2. Proof of Theorem 1.2

The proof is similar, we will be more succinct (all calculations are in [Cha]). We now write $f(t) = G(1, 1)$. We start with Equation (3.6) in Theorem 3.2. We use this equation to compute a recurrence formula for the functions

$$f^{(k)} \equiv f^{(k)}(t) := \left(\left(\frac{d}{dw} \right)^k G(1, 1) \right) \Big|_{w=1}.$$

The main singularity of $f^{(0)}$ (which is the same as in the previous section) is at $t = \rho := \frac{27}{256}$, which corresponds to $z = \frac{1}{4}$, i.e. $\delta = 0$.

Using the package gfun, we immediately find [Cha] that the $f^{(k)}$ satisfy an equation of the form (1.11) with $L = 9$, where the $h_d(t, k)$ for $d \in [9]$ are rational functions in δ with coefficients which are rational functions of k . The poles of h_1, \dots, h_9 at $\delta = 0$ are respectively of order 0, 6, 6, 10, 10, 10, 10, 10, 10. This implies that hypothesis (i) of Theorem 1.8 holds with $\beta = \frac{3}{4}$. One can explicitly check the values of the corresponding constants $a_d(k)$, which

are nothing but the coefficients of the leading term of h_d in the expansion at $\delta^{-1} = \infty$, up to a scaling factor. They are given [Cha] by

$$a_1(k), \dots, a_9(k) = 0, \frac{\sqrt{6}(3k-4)(3k-8)}{864}, 0, 0, 0, 0, 0, 0, 0. \quad (5.3)$$

Now it only remains to check the initial conditions (ii). From the explicit Equation (2.7), the dominant singularity of $f^{(0)}$, and indeed of each given $f^{(i)}$, can be computed automatically. The function $f^{(0)} = \frac{64}{(\delta+3)^3}$ is the same as in the previous section, with singular expansion $f^{(0)} \sim c_0(1-t/\rho)^\alpha$ with $\alpha = \frac{1}{2}$ and $c_0 = -\frac{32\sqrt{6}}{27}$.

We now have $\ell_0 = L + 0 = 9$, and to check the initial conditions we need to estimate the main singularity of $f^{(1)}, \dots, f^{(9)}$. This is done automatically and one gets $f^{(\ell)} \sim c_\ell(1-t/\rho)^{\alpha-\beta\ell}$ for $\ell \leq 9$, with c_ℓ given by

$$c_1, \dots, c_9 = \frac{16 \cdot 3^{\frac{3}{4}} 2^{\frac{1}{4}}}{81}, \frac{8}{243}, \frac{5 \cdot 2^{\frac{3}{4}} 3^{\frac{1}{4}}}{1458}, \frac{8\sqrt{6}}{6561}, \frac{385 \cdot 3^{\frac{3}{4}} 2^{\frac{1}{4}}}{629856}, \frac{70}{59049}, \frac{85085 \cdot 3^{\frac{1}{4}} 2^{\frac{3}{4}}}{181398528}, \frac{700\sqrt{6}}{1594323}, \frac{37182145 \cdot 3^{\frac{3}{4}} 2^{\frac{1}{4}}}{78364164096}.$$

We have now verified all hypotheses of Theorem 1.8. The main recurrence formula becomes

$$c_k = \frac{\sqrt{6}(3k-4)(3k-8)}{864} c_{k-2}, \quad k > 9,$$

with initial conditions above. It is direct to obtain the solution of this recurrence, and this concludes the proof of Theorem 1.2 in the same way as we did for Theorem 1.1.

5.3. Proof of Theorem 1.3

From Theorem 4.2 it is automatic to obtain an algebraic equation for $M(1, 1)$ at $w = 1$, or any of its derivatives at $w = 1$ of given order. In particular one checks [Cha] that the second derivative has a unique dominant singularity at $z = 1/4$ and an expansion of the form

$$\left(\frac{d}{dw}\right)^2 M(1, 1) \Big|_{w=1} = O((1-4z)^{-1})$$

at this point. Note that this is one order of magnitude more singular (in powers of $(1 - \frac{256}{27}t) \sim (1 - 4z)^2$) than the singularity of the main function

$$M(1, 1) \Big|_{w=1} \sim c(1-4z)$$

for some $c > 0$.

By transfer theorems, this immediately shows that the second moment of $\tilde{Q}_n(J) - 3\tilde{P}_n(J)$ is bounded by $O(n)$. This proves (1.6) in Theorem 1.3. By the Markov inequality⁷ this implies that $\frac{\tilde{Q}_n(J) - 3\tilde{P}_n(J)}{n^{3/4}}$ converges to zero in probability (i.e. is $o(1)$ in probability), therefore (1.7) in Theorem 1.3 follows from the previously proven convergence of $\frac{\tilde{P}_n(J)}{n^{3/4}}$ and $\frac{\tilde{Q}_n(J)}{n^{3/4}}$ to Z and $Z/3$ – note that Z is almost surely nonzero. This concludes all remaining proofs.

⁷more precisely the inequality $P(|X| \geq a) \leq E(X^2)/a^2$, for any fixed $a > 0$, applied to $X = \frac{\tilde{Q}_n(J) - 3\tilde{P}_n(J)}{n^{3/4}}$.

The order of magnitude $O(n)$ for the second moment, as well as the explicit computation of a few more higher moments, strongly suggests that $n^{-1/2}(\tilde{Q}_n(J) - 3\tilde{P}_n(J))$ has a Gaussian limit law, but we did not prove it. Theorem 1.8 does not immediately apply as there is an even/odd phenomenon in the form of asymptotic singularities of successive derivatives of $M(1, 1)$ (this is not surprising given the fact that odd moments of the standard Gaussian are null). We suspect that the general idea of Section 1.2 might apply with some adaptations. However, for Gaussian limit laws there are many available tools, in particular experts in classical ACSV or quasi-power theorems might be able to prove Gaussian convergence directly from Theorem 4.2, while we hope experts of the Bernardi–Bonichon bijection might be able to see it directly from the central limit theorem.

In any case, we believe it is a good point to end this paper.

Acknowledgements

Many thanks also to an anonymous referee for their help improving the quality of the paper, and for providing comments and references related to Pólya urns at the origin of the last paragraph of Section 1.2

References

- [Ald91] David Aldous. The continuum random tree. II. An overview. In *Stochastic analysis (Durham, 1990)*, volume 167 of *London Math. Soc. Lecture Note Ser.*, pages 23–70. Cambridge Univ. Press, Cambridge, 1991. doi:10.1017/CB09780511662980.003.
- [BB09] Olivier Bernardi and Nicolas Bonichon. Intervals in Catalan lattices and realizers of triangulations. *J. Combin. Theory Ser. A*, 116(1):55–75, 2009. doi:10.1016/j.jcta.2008.05.005.
- [BBMR21] Olivier Bernardi, Mireille Bousquet-Mélou, and Kilian Raschel. Counting quadrant walks via Tutte’s invariant method. *Comb. Theory*, 1:Paper No. 3, 77, 2021. doi:10.5070/C61055360.
- [BDFG03] J. Bouttier, P. Di Francesco, and E. Guitter. Geodesic distance in planar graphs. *Nuclear Phys. B*, 663(3):535–567, 2003. doi:10.1016/S0550-3213(03)00355-9.
- [Bil12] Patrick Billingsley. *Probability and measure*. Wiley Series in Probability and Statistics. John Wiley & Sons, Inc., Hoboken, NJ, anniversary edition, 2012. With a foreword by Steve Lalley and a brief biography of Billingsley by Steve Koppes.
- [BMC24] Mireille Bousquet-Mélou and Frédéric Chapoton. Intervals in the greedy Tamari posets. *Comb. Theory*, 4(1):Paper No. 5, 27, 2024. doi:10.5070/C64163842.
- [BMCP13] Mireille Bousquet-Mélou, Guillaume Chapuy, and Louis-François Prévaille-Ratelle. The representation of the symmetric group on m -Tamari intervals. *Adv. Math.*, 247:309–342, 2013. doi:10.1016/j.aim.2013.07.014.

- [BMFPR11] Mireille Bousquet-Mélou, Éric Fusy, and Louis-François Prévaille-Ratelle. The number of intervals in the m -Tamari lattices. *Electron. J. Combin.*, 18(2):Paper 31, 26, 2011. doi:10.37236/2027.
- [BMJ06] Mireille Bousquet-Mélou and Arnaud Jehanne. Polynomial equations with one catalytic variable, algebraic series and map enumeration. *J. Combin. Theory Ser. B*, 96(5):623–672, 2006. doi:10.1016/j.jctb.2005.12.003.
- [BMS00] Mireille Bousquet-Mélou and Gilles Schaeffer. Enumeration of planar constellations. *Adv. in Appl. Math.*, 24(4):337–368, 2000. doi:10.1006/aama.1999.0673.
- [BMW20] Cyril Banderier, Philippe Marchal, and Michael Wallner. Periodic Pólya urns, the density method and asymptotics of Young tableaux. *The Annals of Probability*, 48(4):1921–1965, 2020. doi:10.1214/19-AOP1411.
- [CC23] Alice Contat and Nicolas Curien. Parking on Cayley trees and frozen Erdős-Rényi. *Ann. Probab.*, 51(6):1993–2055, 2023. doi:10.1214/23-AOP1632.
- [Cha] Guillaume Chapuy. Maple worksheet accompanying this paper, available as here (.mw) or here (.pdf).
- [Cha07] F. Chapoton. Sur le nombre d’intervalles dans les treillis de Tamari. *Sém. Lothar. Combin.*, 55:Art. B55f, 18, 2005/07.
- [Com64] L. Comtet. Calcul pratique des coefficients de Taylor d’une fonction algébrique. *Enseign. Math. (2)*, 10:267–270, 1964.
- [CS04] Philippe Chassaing and Gilles Schaeffer. Random planar lattices and Integrated SuperBrownian Excursion. *Probab. Theory Related Fields*, 128(2):161–212, 2004. doi:10.1007/s00440-003-0297-8.
- [DNY22] Michael Drmota, Marc Noy, and Guan-Ru Yu. Universal singular exponents in catalytic variable equations. *J. Combin. Theory Ser. A*, 185:Paper No. 105522, 33, 2022. doi:10.1016/j.jcta.2021.105522.
- [EF23] David Eppstein and Daniel Frishberg. Improved mixing for the convex polygon triangulation flip walk. In *50th International Colloquium on Automata, Languages, and Programming*, volume 261 of *LIPICs. Leibniz Int. Proc. Inform.*, pages Art. No. 56, 17. Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, 2023. doi:10.4230/lipics.icalp.2023.56.
- [Fan21] Wenjie Fang. Bijective link between Chapoton’s new intervals and bipartite planar maps. *European J. Combin.*, 97:Paper No. 103382, 15, 2021. doi:10.1016/j.ejc.2021.103382.
- [FO90] Philippe Flajolet and Andrew Odlyzko. Singularity analysis of generating functions. *SIAM J. Discrete Math.*, 3(2):216–240, 1990. doi:10.1137/0403019.
- [FS31] M. Fréchet and J. Shohat. A proof of the generalized second-limit theorem in the theory of probability. *Trans. Amer. Math. Soc.*, 33(2):533–543, 1931. doi:10.2307/1989421.

- [FS09] Philippe Flajolet and Robert Sedgewick. *Analytic combinatorics*. Cambridge University Press, Cambridge, 2009. doi:10.1017/CB09780511801655.
- [HN02] Hsien-Kuei Hwang and Ralph Neininger. Phase change of limit laws in the quick-sort recurrence under varying toll functions. *SIAM J. Comput.*, 31(6):1687–1722, 2002. doi:10.1137/S009753970138390X.
- [JKB94] Norman L. Johnson, Samuel Kotz, and N. Balakrishnan. *Continuous univariate distributions. Vol. 1*. Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics. John Wiley & Sons, Inc., New York, second edition, 1994. A Wiley-Interscience Publication.
- [JKB95] Norman L. Johnson, Samuel Kotz, and N. Balakrishnan. *Continuous univariate distributions. Vol. 2*. Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics. John Wiley & Sons, Inc., New York, second edition, 1995. A Wiley-Interscience Publication.
- [KP16] Markus Kuba and Alois Panholzer. On moment sequences and mixed Poisson distributions. *Probab. Surv.*, 13:89–155, 2016. doi:10.1214/14-PS244.
- [LG19] Jean-François Le Gall. Brownian geometry. *Jpn. J. Math.*, 14(2):135–174, 2019. doi:10.1007/s11537-019-1821-7.
- [MHPS12] Folkert Müller-Hoissen, Jean Marcel Pallo, and Jim Stasheff, editors. *Associahedra, Tamari Lattices and Related Structures. Tamari Memorial Festschrift*, volume 299 of *Progress in Mathematics*. Springer, New York, 2012.
- [MT99] Lisa McShine and Prasad Tetali. On the mixing time of the triangulation walk and other Catalan structures. In *Randomization methods in algorithm design (Princeton, NJ, 1997)*, volume 43 of *DIMACS Ser. Discrete Math. Theoret. Comput. Sci.*, pages 147–160. Amer. Math. Soc., Providence, RI, 1999. doi:10.1090/dimacs/043/09.
- [Pou14] Lionel Pournin. The diameter of associahedra. *Adv. Math.*, 259:13–42, 2014. doi:10.1016/j.aim.2014.02.035.
- [PS02] K.A. Penson and A.I. Solomon. Coherent states from combinatorial sequences. In *Quantum Theory and Symmetries*, pages 527–530, 2002. doi:10.1142/9789812777850_0066.
- [PSZ23] Vincent Pilaud, Francisco Santos, and Günter M. Ziegler. Celebrating Loday’s associahedron. *Arch. Math. (Basel)*, 121(5-6):559–601, 2023. doi:10.1007/s00013-023-01895-6.
- [Sta80] R. P. Stanley. Differentiably finite power series. *European J. Combin.*, 1(2):175–188, 1980. doi:10.1016/S0195-6698(80)80051-5.
- [STT88] Daniel D. Sleator, Robert E. Tarjan, and William P. Thurston. Rotation distance, triangulations, and hyperbolic geometry. *J. Amer. Math. Soc.*, 1(3):647–681, 1988. doi:10.2307/1990951.

- [SZ94] Bruno Salvy and Paul Zimmermann. GFUN: a Maple package for the manipulation of generating and holonomic functions in one variable. *ACM Trans. Math. Softw.*, 20(2):163–177, 1994. doi:10.1145/178365.178368.
- [Tut62] W. T. Tutte. A census of planar triangulations. *Canadian J. Math.*, 14:21–38, 1962. doi:10.4153/CJM-1962-002-9.