Counting Planar Tanglegrams

Dimbinaina Ralaivaosaona
Department of Mathematical Sciences, Stellenbosch University, Stellenbosch, South Africa
naina@sun.ac.za

Jean Bernoulli Ravelomanana
Department of Mathematical Sciences, Stellenbosch University, Stellenbosch, South Africa
jeanr@sun.ac.za

Stephan Wagner[1]
Department of Mathematical Sciences, Stellenbosch University, Stellenbosch, South Africa
swagner@sun.ac.za

Abstract

Tanglegrams are structures consisting of two binary rooted trees with the same number of leaves and a perfect matching between the leaves of the two trees. We say that a tanglegram is planar if it can be drawn in the plane without crossings. Using a blend of combinatorial and analytic techniques, we determine an asymptotic formula for the number of planar tanglegrams with n leaves on each side.

2012 ACM Subject Classification Mathematics of computing → Enumeration, Mathematics of computing → Generating functions

Keywords and phrases rooted binary trees, tanglegram, planar, generating functions, asymptotic enumeration, singularity analysis

Digital Object Identifier 10.4230/LIPIcs.AofA.2018.32

1 Introduction and statement of results

A *tanglegram* is a structure consisting of two (unordered, non-plane) binary trees with the same number of leaves and a perfect matching between the leaf sets. Tanglegrams occur naturally in the study of cospeciation and coevolution (see [10, 12]), where the two trees are phylogenetic trees, and also in computer science in the analysis of software projects and clustering problems [2]. Formally, we can define a tanglegram as a triplet (T, ϕ, S), where T and S are two rooted binary trees with the same number of leaves n, and ϕ is a bijection between the leaf sets. The size of a tanglegram is the number of leaves in each tree. We draw a tanglegram (T, ϕ, S) with one tree on top and the other on the bottom; the corresponding bijection ϕ is represented by inter-tree edges (see Figure 1 for an example of a tanglegram: edges between leaves are dashed). In such a representation, tree edges are not allowed to cross, while inter-tree edges may have crossings.

Note that a tanglegram can usually be drawn in many different ways. Figure 2 shows two representations of a tanglegram (the same as in Figure 1), where corresponding leaves are indicated by identical colours and labels. Formally, we consider two tanglegrams (T, ϕ, S) and (T', ϕ', S') isomorphic if there are (rooted tree) isomorphisms g from T to T' and h from S to S' such that $\phi' = g \circ \phi \circ h^{-1}$.

[1] The third author is supported by the National Research Foundation of South Africa under grant number 96236.

29th International Conference on Probabilistic, Combinatorial and Asymptotic Methods for the Analysis of Algorithms (AofA 2018).
Editors: James Allen Fill and Mark Daniel Ward; Article No. 32; pp. 32:1–32:18
Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

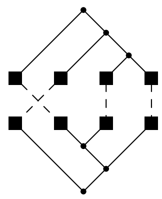

Figure 1 A tanglegram of size 4.

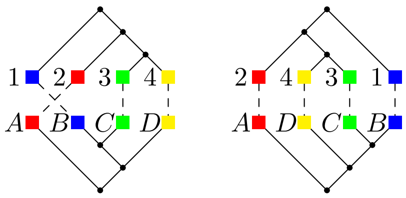

Figure 2 Two different representations of a tanglegram.

An equivalence class of tanglegrams under this definition of isomorphism formally corresponds to a double coset of the symmetric group, see [1] for details. We point out that a tanglegram isomorphism cannot interchange the top and the bottom tree.

It is desirable, both for aesthetic and practical purposes, to represent a tanglegram with a minimum number of crossings between inter-tree edges. For instance, the left representation in Figure 2 has one crossing, the right representation is crossing-free. The problem of determining the minimum number of crossings for a given tanglegram is known as the *Tanglegram Layout* (TL) problem [2]. This problem is, just like the crossing number problem for graphs, NP-hard in general [5].

In this paper, we consider a related enumerative question. In [1], Billey, Konvalinka and Matsen established a formula for the number t_n of tanglegrams of size n (up to isomorphism). The first few terms of the sequence t_n are given by $1, 1, 2, 13, 114, 1509, 25595, 535753, 13305590, 382728552, \ldots$, see also [11, A258620]. They also obtained an asymptotic formula for t_n: for $n \to \infty$, we have

$$t_n \sim \frac{2^{2n-\frac{3}{2}} \cdot n^{n-\frac{5}{2}}}{\sqrt{\pi} \cdot e^{n-\frac{1}{8}}}.$$

Based on the results of Billey, Konvalinka and Matsen, properties of random tanglegrams were investigated in [9].

Here, we ask a similar question: how many tanglegrams of size n (up to isomorphism) are there that can be drawn without crossing? In analogy to planar graphs, we will call them *planar tanglegrams*. For example, the tanglegram in Figure 1 has a crossing-free representation (as Figure 2 shows) and is thus planar. All tanglegrams of size $1, 2$, or 3 are easily seen to be planar. Among the thirteen tanglegrams of size 4, only two are not planar, see Figure 3. Indeed, it can be shown that a tanglegram is planar if and only if all of the subtanglegrams induced by four leaf pairs are planar, in analogy to Kuratowski's celebrated characterisation of planar graphs (see [3]). Here, an induced subtanglegram is a tanglegram obtained in the following way: pick some leaf pairs, then take the smallest subtree on each side that contains the respective leaves, and suppress internal vertices that are no longer needed.

Our approach to the enumeration of planar tanglegrams is based on generating functions. Our first main result characterises the generating function $T(x)$ of planar tanglegrams by means of a functional equation.

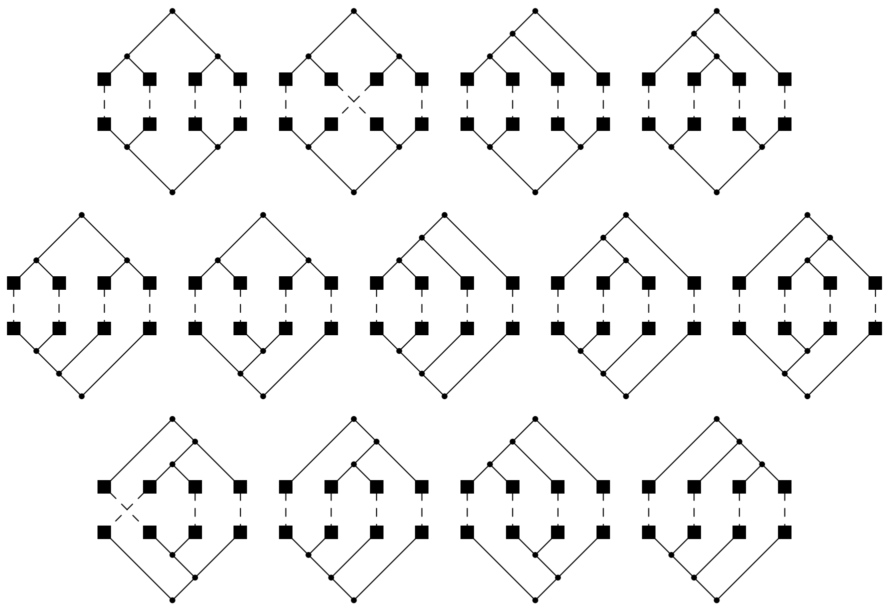

Figure 3 All tanglegrams of size 4. Only the second and tenth tanglegram cannot be drawn without crossing.

▶ **Theorem 1.** *Let $T(x)$ be the (ordinary) generating function for the number of planar tanglegrams, counted up to isomorphism. The function $T(x)$ is uniquely determined by the following system of functional equations involving two auxiliary functions $A(x)$ and $H(x)$:*

$$A(x) = \sum_{r=1}^{\infty} \frac{1}{(r+1)^2} \binom{2r}{r}^2 x^r (1 - A(x))^{r+1}, \tag{1}$$

$$H(x) = \frac{xA(x)}{2}, \tag{2}$$

$$T(x) = H(T(x)) + x + \frac{T(x^2)}{2}. \tag{3}$$

It turns out that $A(x)$ in the theorem above is the generating function for the number of ordered pairs of triangulations of polygons without common diagonals. Moreover, the auxiliary function $H(x)$ that occurs in Theorem 1 has a natural combinatorial interpretation as well: it is the generating function for *irreducible* planar tanglegrams (to be defined in the following). We will see that unordered pairs of triangulations of polygons without common diagonals and irreducible tanglegrams are in one-to-one correspondence. In order to define irreducible planar tanglegrams, we first need the concept of proper subtanglegrams.

A binary subtree T' of a binary tree T is an induced binary tree consisting of a vertex and all its successors. We call the binary subtree T' a proper binary subtree if it is not a leaf and the root of T' is different from the root of T. A subtanglegram of a planar tanglegram consists of a binary subtree of the top tree and a binary subtree of the bottom tree with the same number of leaves, where each leaf of the top subtree is matched to a leaf of the bottom subtree. Moreover, a subtanglegram is called a proper subtanglegram if the two corresponding binary subtrees are proper. Figure 4 shows a proper subtanglegram of a planar tanglegram. An irreducible planar tanglegram is a planar tanglegram which does not contain any proper subtanglegrams and which has more than one leaf in each tree. For example, in Figure 3, the first, third, fifth, seventh, and ninth tanglegram contain proper subtanglegrams

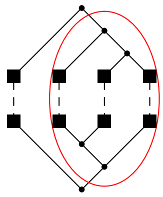

Figure 4 A proper subtanglegram of a planar tanglegram.

Table 1 The first 10 values of T_n and H_n.

n	1	2	3	4	5	6	7	8	9	10
T_n	1	1	2	11	76	649	6 173	63 429	688 898	7 808 246
H_n	0	1	1	5	34	273	2 436	23 391	237 090	2 505 228

of size 2; the seventh and eighth contain proper subtanglegrams of size 3. Only five of the eleven planar tanglegrams shown in the figure are irreducible.

Let T_n be the number of planar tanglegrams of size n, and let H_n be the number of irreducible planar tanglegrams of size n. It is easy (with the help of a computer algebra system) to determine the first few values of T_n and H_n from the functional equations in Theorem 1 – see Table 1. Figure 3 illustrates the values $T_4 = 11$ and $H_4 = 5$. The sequence H_n also occurs in a different context as the number of proper diagonals of the n-dimensional associahedron, see [8] and [11, A257887].

Several ingredients are needed in order to prove Theorem 1. In the next section, we show a bijection between pairs of triangulations of polygons without common diagonals and irreducible planar tanglegrams. Thereafter, we use this bijection to obtain functional equations for the generating function of irreducible planar tanglegrams and related generating functions. Finally, we derive a functional equation relating the generating function of planar tanglegrams with the generating function of irreducible planar tanglegrams. An important feature of irreducible planar tanglegrams is the fact that their embeddings in the plane are almost unique, see Proposition 5. Moreover, every planar tanglegram can be reduced to an irreducible planar tanglegram by contracting maximal proper subtanglegrams.

In order to determine the asymptotic behaviour of T_n, we study the analytic properties of its generating function and eventually apply singularity analysis. This is also done in several steps, starting from the function $A(x)$ that is closely related to an elliptic integral, from which the behaviour of $H(x)$ is derived. As a side result, we also obtain the asymptotic behaviour of the coefficients H_n (see Theorem 11). Our main analytic result regarding the generating function $T(x)$ reads as follows:

▶ **Theorem 2.** *The generating function T enumerating planar tanglegrams satisfies the following properties:*

(i) *Let ρ be the radius of convergence of T. There exist positive real numbers θ and ϵ such that T is analytic in*

$$\Delta = \{x : |x| < \rho + \epsilon \text{ and } |\operatorname{Arg}(x - \rho)| > \theta\},$$

and for $x \in \Delta$, we have:

$$T(x) = \alpha + C_1(\rho - x) + C_2(\rho - x)^2 + B(\rho - x)^2 \log(\rho - x) + O(|(\rho - x)^3 \log(\rho - x)|). \quad (4)$$

Here, C_1, C_2 and B are constants that can be computed numerically, and $\alpha = \frac{4-\pi}{4\pi}$.

(ii) *As $n \to \infty$, the n^{th} coefficient T_n of T satisfies the asymptotic formula*

$$T_n \sim C \cdot n^{-3} \cdot \rho^{-n},$$

where $\rho \approx 0.0633892927$, $\rho^{-1} \approx 15.7755349051$ and $C \approx 0.0078873668$.

▶ Remark. From the analytic behaviour of T given in the previous theorem, it follows that T cannot be algebraic (see [6, Theorem VII.7] and [6, Theorem VII.8]).

The first property allows us to use singularity analysis on the generating function $T(x)$ and obtain the asymptotic formula in the second statement. This analysis is outlined in Section 3. We remark that, once the enumeration problem has been solved, it will also be possible to study statistics of planar tanglegrams. This is left as a future project.

2 Deriving the functional equations

2.1 Irreducible tanglegrams and triangulations

First, we work with rooted plane binary trees, which are rooted binary trees with a plane embedding, so that left and right child of every vertex are distinguishable. It is well known that rooted plane binary trees are counted by the Catalan numbers. We denote by $\mathcal{C}_b$ the set of ordered pairs of rooted plane binary trees with the same number of leaves. If we label the leaves canonically (from left to right) and match leaves with the same label, we obtain a planar tanglegram. Every planar tanglegram can be obtained in this way, but of course several pairs of rooted plane binary trees may represent the same planar tanglegram. The next proposition relates pairs of triangulations of a polygon and elements of $\mathcal{C}_b$, based on the well-known bijection between rooted plane binary trees and triangulations.

▶ **Proposition 3.** *To every element (T_1, T_2) of $\mathcal{C}_b$ with n leaves corresponds a unique pair of triangulations of an $(n+1)$-gon. The tanglegram associated with (T_1, T_2) contains a proper subtanglegram if and only if the corresponding pair of triangulations has a common diagonal.*

Proof. The bijection between binary trees and triangulations is a classical application of the plane dual (see [6, Section I.5.3]): given a rooted plane binary tree with n leaves, draw non-intersecting lines from the root and all leaves to infinity. These lines and the edges of the tree divide the plane into regions. We place a vertex in each of these regions and connect two such vertices by an edge if the corresponding regions share part of their boundaries. The result is a triangulation of an $(n+1)$-gon, and the correspondence is bijective. A canonical way to label the vertices of the triangulation is to number them clockwise, starting from the root of the tree – see Figure 5.

This also yields a bijection between the elements of $\mathcal{C}_b$ and pairs of triangulations. It is not difficult to see that diagonals of triangulations correspond to proper subtrees, so that a common diagonal in a pair of triangulations corresponds to a pair of proper subtrees whose leaves are matched to each other, i.e. a proper subtanglegram. This is illustrated in Figure 6 for the tanglegram in Figure 4. ◀

We call an element of $\mathcal{C}_b$ that corresponds to an irreducible planar tanglegram a *representation* of that irreducible tanglegram. The next theorem relates irreducible planar tanglegrams and their representations to pairs of triangulations.

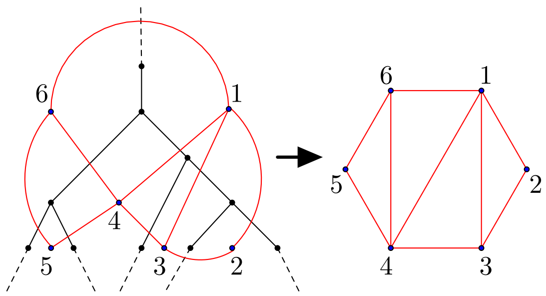

Figure 5 The correspondence between binary trees and triangulations

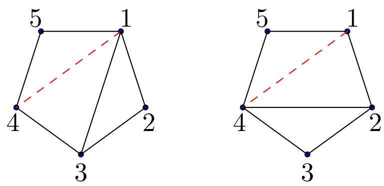

Figure 6 The triangulations corresponding to the two halves of the tanglegram in Figure 4 (the bottom tree is reflected by a horizontal axis before applying the bijection). The common diagonal is indicated in red and dashed.

▶ **Theorem 4.** *The following statements hold:*

(1) *To every representation of an irreducible planar tanglegram of size n corresponds a unique ordered pair of triangulations of an $(n+1)$-gon without common diagonals.*

(2) *There is a bijection between irreducible planar tanglegrams of size n and unordered pairs of triangulations of an $(n+1)$-gon that do not have a common diagonal.*

The first part of Theorem 4 is a consequence of Proposition 3. In order to prove the second part, we first have to show that an irreducible planar tanglegram has a unique representation up to homeomorphism.

▶ **Proposition 5.** *Every irreducible planar tanglegram with more than two leaves in each tree has precisely two possible representations, which are mirror images of each other.*

Proposition 5 is obtained by means of a famous theorem of Whitney:

▶ **Theorem 6** (Whitney [13])**.** *Every 3-connected planar graph has a unique plane embedding up to homeomorphism.*

The main idea is to apply Whitney's theorem to the graph obtained from a tanglegram by removing the leaves on each side (but leaving the connecting edges) and connecting the roots by an additional edge. Let us call this process *smoothing* – see Figure 7. We have the following proposition, whose proof is given in the appendix.

▶ **Proposition 7.** *The graph obtained by smoothing a tanglegram is* 3*-regular and* 3*-connected if the tanglegram is irreducible and has more than 2 leaves in each tree.*

We can now proceed to the proof of Proposition 5.

Proof of Proposition 5. Let I be an irreducible tanglegram with more than two leaves in each tree, and let T_1 and T_2 be the corresponding binary trees. Every representation yields a plane embedding of the graph that is obtained by smoothing. By Whitney's theorem and

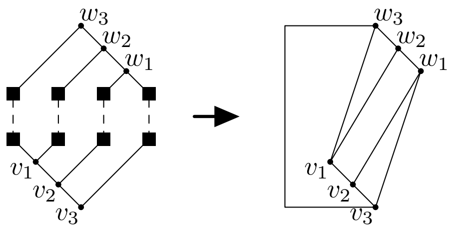

Figure 7 Smoothing the last tanglegram in Figure 3.

Proposition 7, there are only two possible representations, which are mirror images of each other. Suppose that the mirror images are identical. Then the mirror images of T_1 and T_2 are respectively the same as T_1 and T_2, which implies that the left and right branches of T_1 and T_2 are the same. So the branches of T_1 and T_2 induce proper subtanglegrams since they contain more than one leaf each. This contradicts the assumption that I is irreducible. Thus, we find that an irreducible tanglegram with more than two leaves on each side has precisely two distinct irreducible representations that are mirror images of each other. ◀

We conclude this section with the proof of part (2) of Theorem 4.

Proof of Theorem 4, part (2). For $n = 2$, the statement is clearly true since both sets contain exactly one element. Now suppose that $n > 2$. Let $\mathcal{P}_n$ be the set of pairs of triangulations of an $(n+1)$-gon without common diagonal, and let $\mathcal{I}_n$ be the set of representations of irreducible tanglegrams of size n. Moreover, denote by $\mathcal{P}'_n$ the set of unordered pairs of triangulations of an $(n+1)$-gon without common diagonal and $\mathcal{I}'_n$ the set of irreducible tanglegrams.

By Propositon 5, we know that to every element of $\mathcal{I}'_n$, there are two distinct corresponding elements of $\mathcal{I}_n$. Moreover, to every pair of triangulations in $\mathcal{P}'_n$, there are two distinct corresponding ordered pairs in P_n. This is because the two triangulations of an element of $\mathcal{P}'_n$ have to be distinct, as they would otherwise have a common diagonal. By the first part of the theorem, there is a bijection between $\mathcal{P}_n$ and $\mathcal{I}_n$. Since $\mathcal{I}'_n$ and $\mathcal{I}_n$ are in a 2–1 correspondence, as are $\mathcal{P}'_n$ and $\mathcal{P}_n$, it follows that there is a bijection between $\mathcal{P}'_n$ and $\mathcal{I}'_n$. ◀

▶ Remark. The only symmetric irreducible tanglegram (equal to its own mirror image) is the tanglegram with two leaves in each tree.

2.2 From bijections to generating functions

Since there is a bijection between irreducible tanglegrams and unordered pairs of triangulations of a polygon without common diagonals, the generating functions of the two combinatorial objects are the same. We derive a functional equation for the generating function of pairs of triangulations without common diagonals using the inclusion-exclusion method described in [6, Section III.7]. We consider ordered pairs of triangulations (of the same polygon) in which some of the common diagonals (not necessarily all and possibly none) are marked. Let $\mathcal{T}$ be the family of ordered pairs of triangulations (of the same polygon) without marked diagonals. For a pair (T_1, T_2) in $\mathcal{T}$, $M(T_1, T_2)$ is the set of all possible configurations of the pair (T_1, T_2) with marked diagonals. For every $m \in M(T_1, T_2)$, we denote by $N(m)$ the number of marked diagonals in m. Lastly, $n(T_1,T_2)$ is the number of triangles in each of the two triangulations T_1 and T_2. Now define a bivariate generating function $A(x,v)$ by

$$A(x,v) = \sum_{(T_1,T_2)\in\mathcal{T}} \Big(\sum_{m\in M(T_1,T_2)} v^{N(m)} \Big) x^{n(T_1,T_2)}. \tag{5}$$

We have the following key observation: if T_1 and T_2 are two triangulations of a polygon with $k(T_1,T_2)$ common diagonals, then

$$\sum_{m\in M(T_1,T_2)} v^{N(m)} = (1+v)^{k(T_1,T_2)}.$$

Indeed, we can choose to mark a common diagonal, which yields a factor v, or not to mark it, which yields a factor 1. Thus,

$$A(x,v) = \sum_{(T_1,T_2)\in\mathcal{T}} \Big(\sum_{m\in M(T_1,T_2)} v^{N(m)} \Big) x^{n(T_1,T_2)} = \sum_{(T_1,T_2)\in\mathcal{T}} (1+v)^{k(T_1,T_2)} x^{n(T_1,T_2)}. \tag{6}$$

If we plug in $v=-1$, all pairs with $k(T_1,T_2)\neq 0$ vanish, and we are left precisely with those ordered pairs of triangulations that have no common diagonals. Hence $A(x,-1)$ represents the generating function for ordered pairs of triangulations without common diagonals. Next we prove that $A(x)=A(x,-1)$ satisfies the functional equation (1).

Proof of (1). Let $\mathcal{A}$ be the set of all configurations consisting of two triangulations of a polygon with vertices labelled $1,2,\ldots,n$ with some of the common diagonals potentially marked. Then $A(x,v)$, as defined in (6), is the bivariate generating function corresponding to $\mathcal{A}$, where the exponents of x and v indicate the number of triangles in each triangulation and the number of common diagonals respectively.

We can decompose an element of $\mathcal{A}$ in the following way: the marked common diagonals divide the polygon into one or more subpolygons. One of them (let us call it P) contains the side from vertex 1 to vertex 2. This polygon P is bounded by edges of the larger polygon and marked diagonals that separate it from smaller elements of $\mathcal{A}$. Thus we have a decomposition into a pair of triangulations without marked diagonals (inside of polygon P) that is surrounded by sides of the larger polygon and elements of $\mathcal{A}$. Let r be the number of triangles in each of the triangulations of P; then P has $r+2$ sides. The number of possibilities for each of the triangulations is the Catalan number $C_r = \frac{1}{r+1}\binom{2r}{r}$, and each of the $r+1$ sides of P other than the side between vertices 1 and 2 is either a side of the whole polygon or a marked diagonal that separates off a smaller element of $\mathcal{A}$.

This decomposition can be translated to the functional equation

$$A(x,v) = \sum_{r=1}^{\infty} C_r^2 x^r (1+vA(x,v))^{r+1}. \tag{7}$$

From (6) we have

$$A(x,v) = \sum_{(T_1,T_2)\in\mathcal{T}} (1+v)^{k(T_1,T_2)} x^{n(T_1,T_2)},$$

where $n(T_1,T_2)$ is the number of triangles and $k(T_1,T_2)$ is the number of common diagonals in (T_1,T_2). Setting $v=-1$, all pairs of triangulations $(T_1,T_2)\in\mathcal{T}$ such that $k(T_1,T_2)\neq 0$ vanish, as mentioned before. This means that all pairs of triangulations (T_1,T_2) which have a common diagonal will not contribute to the sum for $A(x,-1)$. In other words, only the pairs of triangulations without common diagonal contribute to $A(x,-1) = A(x)$, i.e. $A(x)$ is the generating function for pairs of triangulations without common diagonals. The equation

$$A(x) = \sum_{r=1}^{\infty} \frac{1}{(r+1)^2} \binom{2r}{r}^2 x^r (1 - A(x))^{r+1} \tag{8}$$

follows immediately from (7). ◀

Recall that there is a 2–1 correspondence between ordered and unordered pairs of triangulations without common diagonals, except for the trivial case of triangulations of a triangle. This and Theorem 4 yield the following proposition.

▶ **Proposition 8.** *The generating function $H(x)$ of irreducible tanglegrams is given by*

$$H(x) = \frac{xA(x)}{2}.$$

Proof. The coefficient of x^r in $A(x)$ corresponds to pairs of triangulations without common diagonal and r triangles in each triangulation. When we transform a triangulation of an $(r+2)$-gon into a planted binary tree, we obtain a planted binary tree with $r+1$ leaves. So, by the first part of Theorem 4, the coefficient of x^r in $A(x)$ is the number of representations of irreducible tanglegrams which have $r+1$ leaves on each side. Multiplying $A(x)$ by x gives us the generating function of representations of irreducible tanglegrams. From Theorem 4, we know that to every irreducible tanglegram with more than two leaves on each side, there are two irreducible representations. The statement of the proposition follows. ◀

▶ Remark. The coefficient of x^2 in $H(x)$ is $\frac{1}{2}$. We maintain it as it will help us later to take symmetries into account when the irreducible tanglegram has two leaves.

Proposition 8 gives us equation (2). It only remains to prove (3) to complete the proof of Theorem 1.

Proof of (3)**.** We would like to prove the identity

$$T(x) = H(T(x)) + x + \frac{T(x^2)}{2}.$$

The term x accounts for the tanglegram with only one leaf in each tree. Now consider an arbitrary planar tanglegram T with more than one leaf in each tree. It has maximal proper subtanglegrams (with respect to inclusion) $T_1, T_2, \ldots, T_k$ for some nonnegative integer k (if all T_j's have size 1, then T is irreducible). For each of these subtanglegrams T_j we have two proper binary subtrees T_j' and T_j'' in the two trees that constitute T. Replace both of them by leaves, and include an inter-tree edge between these two leaves. Contracting each maximal proper subtanglegram to a single pair of leaves in this way, we obtain an irreducible tanglegram (see Figure 8). Conversely, if I is an irreducible planar tanglegram, we can replace each pair of matched leaves in I by some planar tanglegram (possibly of size 1, i.e. the leaves remain as they are) to obtain a new planar tanglegram.

Thus every planar tanglegram T can be decomposed uniquely into an irreducible planar tanglegram I and a collection of planar tanglegrams corresponding to the edges of I. We have two cases to consider:

- The irreducible tanglegram I has size greater than 2. Then I is not symmetric, as we have seen in the proof of Proposition 5. In the monomial x^r in $H(x)$, r represents the number of leaves, and replacing a pair of leaves by a planar tanglegram in the irreducible tanglegram translates to replacing x by $T(x)$ in $H(x) - \frac{x^2}{2}$.

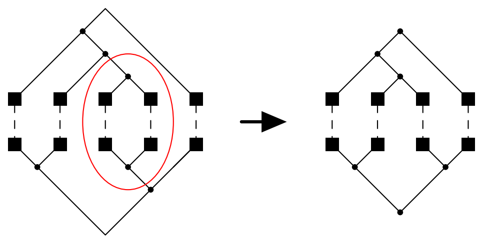

Figure 8 Contracting a proper subtanglegram to obtain an irreducible tanglegram.

- The irreducible tanglegram has size 2. We have to replace the two pairs of leaves by two planar tanglegrams; however, in view of the symmetry, the order is irrelevant, so this amounts to taking an unordered pair of tanglegrams. By Pólya's enumeration theorem (see [7, Section 2.4] or [6, Section I.6.1]) the generating function for these unordered pairs is given by $\frac{1}{2}(T(x)^2 + T(x^2))$.

Combining all cases, we get

$$T(x) = x + \left(H(T(x)) - \frac{T(x)^2}{2}\right) + \frac{1}{2}\big(T(x)^2 + T(x^2)\big) = H(T(x)) + x + \frac{T(x^2)}{2}.$$

This completes the proof of (3) and thus of Theorem 1. ◀

3 Asymptotic analysis

In this section, we consider analytic properties of the generating functions in Theorem 1. Since the proofs of the results in this section are all rather technical, they are deferred to the appendix. We will first work with H and deduce the properties of T in Theorem 2 from H. Setting $u(x) = x(1 - A(x))$, Equation (1) can be rewritten in the form

$$x = \sum_{r=0}^{\infty} C_r^2 \cdot u(x)^{r+1}.$$

This motivates the definition of a function ϕ by

$$\phi(u) = \sum_{r=0}^{\infty} C_r^2 \cdot u^{r+1} = \sum_{r=0}^{\infty} \frac{1}{(r+1)^2}\binom{2r}{r}^2 u^{r+1}.$$

The function u becomes the inverse of ϕ. We will obtain the analytic behaviour of the generating functions A and H by studying ϕ. Next, we note that the function ϕ is connected to the complete elliptic integral

$$k(x) = \int_0^{\pi/2} \frac{1}{\sqrt{1 - x\sin^2 t}}\, dt$$

by means of the identity (cf. [4, 19.5.1])

$$\sum_{r=0}^{\infty} \binom{2r}{r}^2 u^r = \frac{2}{\pi} k(16u),$$

which is valid for $|u| < \frac{1}{16}$. We can exploit this connection to obtain the following proposition:

▶ **Proposition 9.** *The function ϕ has an analytic continuation to the slit plane $\mathbb{C} \setminus [\frac{1}{16}, \infty)$. Moreover, when u tends to $\frac{1}{16}$, we have*

$$\phi(u) = \frac{4-\pi}{4\pi} - \frac{1}{4\pi}(1-16u) - \frac{1}{64\pi}\Big(5 - 8\log 2 + 2\log(1-16u)\Big)(1-16u)^2 + O\Big(\Big|(1-16u)^3 \log(1-16u)\Big|\Big). \quad (9)$$

Since we are interested in the inverse of ϕ, we also need the information given in the following lemma:

▶ **Lemma 10.** *The function ϕ is injective in $\mathbb{C} \setminus [\frac{1}{16}, \infty)$, and for all $u \in \mathbb{C} \setminus [\frac{1}{16}, \infty)$ we have $\phi'(u) \neq 0$.*

The branch cut singularity of ϕ at $\frac{1}{16}$ corresponds to a singularity of H at $\phi(\frac{1}{16}) = \frac{4-\pi}{4\pi}$. Inverting the asymptotic expansion of ϕ around the singularity by means of bootstrapping, we derive an asymptotic expansion for the generating function $H(x)$ at its dominant singularity, which also yields an asymptotic formula for the number of irreducible tanglegrams by a typical application of singularity analysis [6, Chapter VI].

▶ **Theorem 11.** *There exist constants $\theta' \in (0, \frac{\pi}{2})$ and $\epsilon' > 0$ such that H is analytic in*

$$\Delta' = \{x \mid |x| < \alpha + \epsilon' \text{ and } |\operatorname{Arg}(x-\alpha)| > \theta'\},$$

and for $x \in \Delta'$, we have

$$H(x) = C_0' + C_1'(\alpha - x) + C_2'(\alpha - x)^2 + B'(\alpha - x)^2 \log(\alpha - x) + O(|(\alpha - x)^3 \log(\alpha - x)|) \quad (10)$$

where $\alpha = \phi(\frac{1}{16}) = \frac{4-\pi}{4\pi}$, $C_0' = \frac{1}{2\pi} - \frac{5}{32}$, $C_1' = \frac{\pi}{8} - \frac{1}{2}$, $C_2' = -\frac{\pi^2}{32}(5 - 4\log 2 + 2\log \pi)$ and $B' = -\frac{\pi^2}{16}$. Thus, the number of irreducible planar tanglegrams is asymptotically given by

$$H_n = [x^n]H(x) \sim \frac{(\pi\alpha)^2}{8} \cdot n^{-3} \cdot \alpha^{-n}.$$

Finally, we move on to the analysis of the generating function $T(x)$ for planar tanglegrams, culminating in the proof of Theorem 2. Details of this proof can be found in the appendix, we focus on the main points. Recall that $T(x)$ satisfies the functional equation

$$T(x) = H(T(x)) + x + \frac{T(x^2)}{2}. \quad (11)$$

The radius of convergence ρ of $T(x)$ can be bounded above by the radius of convergence of $H(x)$ (since $T_n = [x^n]T(x) \geq [x^n]H(x) = H_n$ for all n in view of the combinatorial interpretation) and is thus less than 1. Pringsheim's Theorem guarantees that ρ is also a singularity. We note that $T(x^2)$ has radius of convergence $\sqrt{\rho} > \rho$, so it is an analytic function in a larger region than $T(x)$ itself.

One also finds that, importantly, $H'(x) \neq 1$ for all x inside the closed disk of convergence of H. This means that the implicit function theorem never fails and the dominant singularity of T is carried over from H: we reach a singularity when $T(x)$ equals the value of H's singularity, which is $\frac{4-\pi}{4\pi}$. This gives us the following characterisation of ρ:

$$T(\rho) = \frac{4-\pi}{4\pi}.$$

For practical computation of ρ, however, it is useful to plug this into (11). Since $H(\frac{4-\pi}{4\pi}) = \frac{1}{2\pi} - \frac{5}{32}$, we obtain

$$\rho + \frac{T(\rho^2)}{2} = \frac{1}{2\pi} - \frac{3}{32},$$

which can be solved numerically with high accuracy. The singular expansion of $T(x)$ around the singularity ρ can be obtained by means of the same bootstrapping process that is also used to prove Theorem 11. Finally, the asymptotic formula for T_n is another standard application of singularity analysis.

References

1 Sara C. Billey, Matjaž Konvalinka, and Frederick A. Matsen. On the enumeration of tanglegrams and tangled chains. *J. Combin. Theory Ser. A*, 146:239–263, 2017.

2 Kevin Buchin, Maike Buchin, Jaroslaw Byrka, Martin Nöllenburg, Yoshio Okamoto, Rodrigo I. Silveira, and Alexander Wolff. Drawing (complete) binary tanglegrams: hardness, approximation, fixed-parameter tractability. *Algorithmica*, 62(1-2):309–332, 2012.

3 Éva Czabarka, László Székely, and Stephan Wagner. A tanglegram Kuratowski theorem. Preprint, submitted; `arXiv:1708.00309`.

4 NIST Digital Library of Mathematical Functions. http://dlmf.nist.gov/, Release 1.0.15 of 2017-06-01. F. W. J. Olver, A. B. Olde Daalhuis, D. W. Lozier, B. I. Schneider, R. F. Boisvert, C. W. Clark, B. R. Miller and B. V. Saunders, eds. URL: `http://dlmf.nist.gov/`.

5 Henning Fernau, Michael Kaufmann, and Mathias Poths. Comparing trees via crossing minimization. In *Proceedings of the 25th International Conference on Foundations of Software Technology and Theoretical Computer Science*, FSTTCS '05, pages 457–469, Berlin, Heidelberg, 2005. Springer-Verlag.

6 Philippe Flajolet and Robert Sedgewick. *Analytic combinatorics.* Cambridge University press, 2009.

7 Frank Harary and Edgar M. Palmer. *Graphical Enumeration.* Academic Press, 1973.

8 Danièle Huguet and Dov Tamari. La structure polyédrale des complexes de parenthésages. *J. Combin. Inform. System Sci.*, 3(2):69–81, 1978.

9 Matjaž Konvalinka and Stephan Wagner. The shape of random tanglegrams. *Adv. in Appl. Math.*, 78:76–93, 2016.

10 Frederick A. Matsen IV, Sara C. Billey, Arnold Kas, and Matjaž Konvalinka. Tanglegrams: a reduction tool for mathematical phylogenetics. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 15(1):343–349, 2018.

11 N.J.A. Sloane, editor. The On-line Encyclopedia of Integer sequences. Published electronically at https://oeis.org, 2018.

12 Roderic D. M. Page. *Tangled trees: phylogeny, cospeciation, and coevolution.* University of Chicago Press, 2003.

13 Hassler Whitney. Non-separable and planar graphs. *Trans. Amer. Math. Soc.*, 34(2):339–362, 1932.

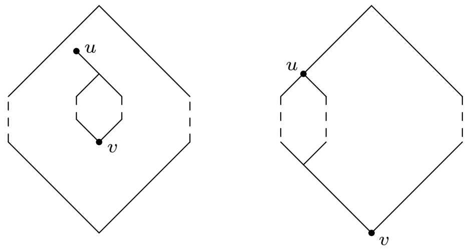

Figure 9 Components of $G \setminus \{u, v\}$ with only 2 edges ending in v

A Appendix: additional proofs

A.1 Proof of Proposition 7

First, notice that the process of smoothing an irreducible tanglegram does not create any parallel edges since the tanglegram would not be irreducible if that was the case (there would be a proper subtanglegram of size 2). After the process of smoothing, the remaining vertices (except the two roots) are all internal vertices, so they all have degree 3. The two roots are also of degree 3 because of the additional edge joining them. Thus, we have a 3-regular graph, which we will denote by G.

Next, let T_1, T_2 be the two halves of an irreducible tanglegram with more than two leaves on each side. We will prove that removing any pair of vertices u, v of the graph obtained from the smoothing process does not disconnect the graph.

- Suppose u, v are in the same tree, say T_1. Every vertex in T_2 is clearly still connected to T_2's root. Every vertex in T_1 has three connections to the root of T_2 that are disjoint within T_1: via the root of T_1 and via the two children. Removing u, v can only destroy at most two of them, so all vertices of T_1 are also still connected to the root of T_2. This means that $G - \{u, v\}$ is connected.
- Now, suppose that u, v are in different trees. Assume that u is a vertex of T_1, v is a vertex of T_2 and that removing disconnects the graph obtained from the tanglegram by smoothing. $T_1 \setminus u$ has up to three components: two corresponding to the children of u, and one containing the root. Some of these components might be empty. Every non-empty component has at least two edges going to the other half of the tanglegram. Suppose there are only two, and both of them have v as an end. Then we are in one of the following situations:

  Either way, there is a proper subtanglegram. So we can assume that every component of $T_1 \setminus u$ has an edge to $T_2 \setminus v$. The same applies to the components of $T_2 \setminus v$. Now consider the bipartite graph whose vertices are the components of $T_1 \setminus u$ and $T_2 \setminus v$, where we connect two components if there is an edge between them. If this graph is connected, then so is the graph $G \setminus \{u, v\}$. So call this graph G' and suppose it is disconnected. Note that the root components of $T_1 \setminus u$ and $T_2 \setminus v$ (if they exist) are connected in G' by definition (since there is an edge between the roots in G). So there must be a component of G' containing only child components of $T_1 \setminus u$ and $T_2 \setminus v$ respectively. This component must have one of the shapes in Figure 10, each corresponding to a proper subtanglegram, which is impossible. It follows that G' must actually be connected. Therefore, we can conclude that G is 3-connected.

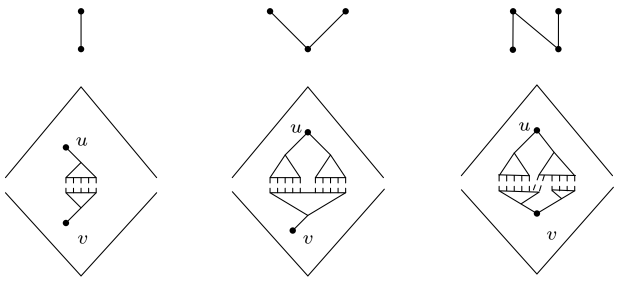

Figure 10 Component of G' containing only child components of $T_1 \setminus u$ and $T_2 \setminus v$.

A.2 Proofs for Section 3

Proof of Proposition 9. First of all, it is well known that

$$\sum_{r=0}^{\infty} \binom{2r}{r}^2 u^r = \frac{2}{\pi} k(16u) \tag{12}$$

for $|u| < \frac{1}{16}$, with the complete elliptic integral

$$k(x) = \int_0^{\pi/2} \frac{1}{\sqrt{1 - x \sin^2 t}}\, dt. \tag{13}$$

The integral $k(x)$ defines an analytic function on the slit plane $\mathbb{C} \setminus [1, \infty)$. Now, [4, 19.12.1] gives the series

$$k(1-x) = \sum_{m=0}^{\infty} \binom{m - 1/2}{m}^2 x^m \Big(-\frac{1}{2} \log x + d(m) \Big)$$

after some rewriting (note that [4] uses a different notation, where $K(u) = k(u^2)$ according to our notation), where $d(0) = 2 \log 2$, $d(m) = d(m-1) - \frac{1}{m(2m-1)}$, or equivalently $d(m) = \psi(1+m) - \psi(\frac{1}{2} + m)$ (here ψ is the Digamma function). Now since

$$\sum_{m=0}^{\infty} \binom{m - 1/2}{m}^2 x^m = \frac{2}{\pi} k(x),$$

which is in fact equivalent to (12), we can also write this as

$$k(1-x) = -\frac{1}{\pi} k(x) \log x + \sum_{m=0}^{\infty} \binom{m - 1/2}{m}^2 d(m) x^m,$$

which provides us with an analytic continuation around the branch cut for $|x| < 1$, $x \notin (-1, 0]$. In particular, we have the following asymptotic expansion around $u = \frac{1}{16}$ (by taking the first term in the series):

$$k(16u) = 2 \log 2 - \frac{1}{2} \log(1 - 16u) + O\Big(\Big|(1 - 16u) \log(1 - 16u)\Big|\Big).$$

Now

$$\phi(u) = \frac{2}{\pi} \int_0^u \frac{1}{v} \int_0^v k(16z)\, dz\, dv,$$

which provides an analytic continuation of ϕ to the slit plane $\mathbb{C} \setminus [\frac{1}{16}, \infty)$. The asymptotic expansion can be integrated termwise by writing

$$\int_0^v k(16z)\,dz = \int_0^{1/16} k(16z)\,dz - \int_v^{1/16} k(16z)\,dz,$$

cf. [6, Theorem VI.9]. We only need the values

$$\int_0^{1/16} k(16z)\,dz = \frac{1}{8} \quad \text{and} \quad \phi\Big(\frac{1}{16}\Big) = \frac{2}{\pi}\int_0^{1/16} \frac{1}{v}\int_0^v k(16z)\,dz\,dv = \frac{4-\pi}{4\pi},$$

which can be obtained by plugging in (13) and interchanging the order of integration. This gives us first

$$\int_0^v k(16z)\,dz = \frac{1}{8} - \frac{1}{32}\Big(1+4\log 2 - \log(1-16v)\Big)(1-16v) + O\Big(\Big|(1-16v)^2\log(1-16v)\Big|\Big).$$

Then, by multiplication with $\frac{1}{v} = 16 + 16(1-16v) + O\big(|1-16v|^2\big)$, we obtain

$$\frac{1}{v}\int_0^v k(16z)\,dz = 2 + \frac{1}{2}\Big(3-4\log 2 + \log(1-16v)\Big)(1-16v) + O\Big(\Big|(1-16v)^2\log(1-16v)\Big|\Big).$$

One more integration step yields (9). ◀

Proof of Lemma 10. We notice that

$$\phi'(u) = \frac{1}{4\pi u}\int_0^{\frac{\pi}{2}} \frac{1-\sqrt{1-16u\sin^2(t)}}{\sin^2(t)}\,dt = \frac{4}{\pi}\int_0^{\frac{\pi}{2}} \frac{1}{1+\sqrt{1-16u\sin^2(t)}}\,dt,$$

which follows from differentiating (13). Now we can make use of the fact that

$$\operatorname{Re}\Big(1+\sqrt{1-16u\sin^2(t)}\Big) > 0,$$

since $\operatorname{Re}(\sqrt{z}) > 0$ holds for every $z \in \mathbb{C} \setminus (-\infty, 0]$. Thus,

$$\operatorname{Re}\Big(\frac{1}{1+\sqrt{1-16u\sin^2(t)}}\Big) > 0,$$

which in turn means that $\operatorname{Re}(\phi'(u)) > 0$ for all $u \in \mathbb{C} \setminus [\infty, \frac{1}{16})$. In particular, $\phi'(u) \neq 0$ for all possible values of u. In the same way, we can show that $\operatorname{Im}(\phi'(u))$ has the same sign as $\operatorname{Im}(u)$ for all u, and the two combined imply that ϕ is injective on its domain of analyticity. Indeed, let $u, v \in \mathbb{C}$ such that $u \neq v$. Since ϕ is analytic in $\mathbb{C} \setminus [\frac{1}{16}, \infty)$, we have $\phi(u) = \phi(v) + \int_u^v \phi'(z)dz$, where we can integrate along any path joining u and v in the slit plane. We have several cases to consider, depending on the location of u and v. In each case, after integration one finds that either $\operatorname{Im}(\phi(u)) \neq \operatorname{Im}(\phi(v))$ or $\operatorname{Re}(\phi(u)) \neq \operatorname{Re}(\phi(v))$, which means $\phi(u) \neq \phi(v)$. ◀

Proof of Theorem 11. In order to simplify computations, we write $y = 1 - 16u$. We let u tend to $\frac{1}{16}$ so y tends to 0. Then, by Proposition 9, we have

$$\phi(u) = x = \alpha - \frac{1}{4\pi}y - \frac{1}{32\pi}y^2\log(y) + O(|y^2|). \tag{14}$$

By Lemma 10 and the implicit function theorem, ϕ is invertible and the inverse ϕ^{-1} is analytic in $\phi(\mathbb{C} \setminus [\frac{1}{16}, \infty))$. The function ϕ comes from the integration of k, which has a

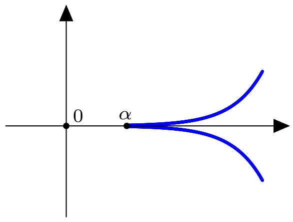

Figure 11 Two branches of ϕ.

branch cut of square root type. The cut $[\frac{1}{16}, \infty)$ is mapped to two branches (see Figure 11), corresponding to the two branches of the square root, and it is easily verified that $\operatorname{Re}\phi(u)$ and $\operatorname{Im}\phi(u)$ are monotone functions of u for both branches. In view of the expansion (14), it is possible to choose θ' in such a way that Δ' lies in the image $\phi(\mathbb{C} \setminus [\frac{1}{16}, \infty))$. Hence, $u = \phi^{-1}$ is well defined and analytic in Δ'.

By means of bootstrapping, we get

$$y = 4\pi(\alpha - x) - 2\pi^2(\alpha - x)^2 \log(\alpha - x) - \pi^2(5 - 4\log 2 + 2\log \pi)(\alpha - x)^2 + O\Big(\Big|(\alpha - x)^3 \log(\alpha - x)\Big|\Big).$$

Now, since $y = 1 - 16u$, we have

$$\begin{aligned} H(x) = \frac{x - u}{2} &= \frac{x}{2} - \frac{1}{32} + \frac{\pi}{8}(\alpha - x) - \frac{\pi^2}{16}(\alpha - x)^2 \log(\alpha - x) \\ &\quad - \frac{\pi^2}{32}(K + 2\log(4\pi))(\alpha - x)^2 + O\Big(\Big|(\alpha - x)^3 \log(\alpha - x)\Big|\Big) \\ &= C_0' + C_1'(\alpha - x) + C_2'(\alpha - x)^2 + B'(\alpha - x)^2 \log(\alpha - x) \\ &\quad + O\Big(\Big|(\alpha - x)^3 \log(\alpha - x)\Big|\Big). \end{aligned}$$

When $n > 2$, the coefficient of x^n in $(\alpha - x)^2$ vanishes, so asymptotically only the term $(\alpha - x)^2 \log(\alpha - x)$ contributes to $[x^n]H(x)$. The function H and the region Δ' satisfy the conditions of [6, Theorem VI.4], so we can apply singularity analysis to get

$$[x^n]H(x) \sim -2 \cdot \alpha^2 \cdot B' \cdot \alpha^{-n} \cdot n^{-3} = \frac{(\pi\alpha)^2}{8} \cdot n^{-3} \cdot \alpha^{-n}. \qquad ◀$$

Proof of Theorem 2. For the first property, we investigate each term in the functional equation for $T(x)$, which reads

$$T(x) = H(T(x)) + x + \frac{T(x^2)}{2}.$$

First, the term x on the right side represents an entire function. Next, $\rho \leq \alpha < 1$ since the coefficients of T are greater than or equal to the coefficients of H in view of their combinatorial interpretation. Since the radius of convergence of $T(x^2)$ is $\sqrt{\rho}$ it follows that $T(x^2)$ has a radius of convergence greater than $T(x)$. Thus, the dominant singularity of $T(x)$ is inherited from the dominant singularity of $H(x)$. $T(x)$ has non-negative coefficients, so by Pringsheim's theorem, the radius of convergence ρ of $T(x)$ is also a singularity. We know that $H(x)$ has its dominant singularity at $x = \alpha = \frac{4-\pi}{4\pi}$, so $H(T(x))$ has a singularity at any point x for which $T(x) = \alpha = \frac{4-\pi}{4\pi}$. Suppose that there exists a positive real number τ such that $\tau < \alpha$ and $H(T(x))$ is singular at $T(x_0) = \tau$ for some $x_0 > 0$. We define the bivariate function

$$F(t, x) = H(t) + x + \frac{T(x^2)}{2} - t,$$

and we have

$$\frac{\partial F(t, x)}{\partial t} = H'(t) - 1.$$

Since τ is a singularity of $T(x)$, the implicit function theorem has to fail at $(t, x) = (\tau, x_0)$ for $F(t, x)$. In other words, we must have

$$\frac{\partial F(t, x)}{\partial t}(\tau, x_0) = H'(\tau) - 1 = 0, \text{ i.e. } H'(\tau) = 1.$$

Next, we have

$$H'(x) = \frac{1 - u'(x)}{2}.$$

$H'(x)$ has non-negative coefficients, hence $H'(x)$ is an increasing function in $(0, \alpha]$. Moreover, we have $u'(\alpha) = \frac{\pi}{4}$. Hence, $H'(\alpha) = \frac{1}{2} - \frac{\pi}{8} < 1$. Thus, $H'(\tau) \leq H'(\alpha) < 1$ contradicting the assumption that $H'(\tau) = 1$. We conclude that the dominant singularity of T appears at ρ and $T(\rho) = \alpha$.

Now, we continue with the proof of the second property. Let $x \in B(0, \rho)$. Since T has positive coefficients, we have

$$|T(x)| = \Big|\sum_{n=0}^{\infty} T_n x^n\Big| \leq \sum_{n=0}^{\infty} T_n |x^n| < \sum_{n=0}^{\infty} T_n \rho^n = \alpha.$$

Hence $T(x) \in B(0, \alpha)$. Moreover, by the implicit function theorem, T can be continued analytically around each point x of the circle $C(0, \rho)$ of center 0 and radius ρ, except perhaps around ρ. However around ρ, T can be continued by Theorem 11. Thus, it is indeed possible to find ϵ and θ such that T is analytic in Δ as required. Let $G(x) = x + \frac{T(x^2)}{2}$. Using the same arguments as in the proof of the first property, $G(x)$ is still analytic around ρ. Hence, the Taylor expansion of $G(x)$ around ρ gives

$$G(x) = G(\rho) + G'(\rho)(x - \rho) + \frac{G''(\rho)}{2}(x - \rho)^2 + O(|x - \rho|^3).$$

For simplicity, we let $D_0 = G(\rho) = \rho + \frac{T(\rho^2)}{2}$, $D_1 = -G'(\rho) = -(1 + \rho T'(\rho^2))$ and $D_2 = \frac{G''(\rho)}{2} = \frac{T'(\rho)}{2} + \rho^2 T''(\rho^2)$. Since $\rho^2 < \rho$, $T'(\rho^2)$ and $T''(\rho^2)$ exist, and the power series for T converges exponentially at ρ^2, which allows for D_0, D_1 and D_2 to be determined with high numerical accuracy. By Theorem 11 and the functional equation for T given in Theorem 1, we have

$$\begin{aligned} T(x) &= C_0' + C_1'(\alpha - T(x)) + C_2'(\alpha - T(x))^2 \\ &\quad + B'(\alpha - T(x))^2 \log(\alpha - T(x)) + O(|(\alpha - T(x))^3 \log(\alpha - T(x))|) \\ &\quad + D_0 + D_1(\rho - x) + D_2(\rho - x)^2 + O(|(\rho - x)^3|). \end{aligned} \tag{15}$$

We note that when $x \to \rho$, we have $T(x) \to T(\rho) = \alpha$, hence $C_0' + D_0 = \alpha$. Again, by means of bootstrapping, we obtain

$$T(x) = \alpha + \frac{D_1}{1 + C_1'}(\rho - x) + \frac{B' \cdot D_1^2}{(1 + C_1')^3}(\rho - x)^2 \log(\rho - x) + C_2(\rho - x)^2 + O(|(\rho - x)^3 \log(\rho - x)|),$$

which gives us the asymptotic expansion (4) for T by setting $C_1 = \frac{D_1}{1+C_1}$ and $B = \frac{B' \cdot D_1^2}{(1+C_1')^3}$.

We remain with the proof of the asymptotic formula for T_n. As in the proof of Theorem 11, only $(\rho - x)^2 \log(\rho - x)$ contributes to the main term of $[x^n]H(x)$ when n is large. So by singularity analysis, we have

$$T_n = [x^n]T(x) \sim C \cdot \rho^{-n} \cdot n^{-3},$$

where $C = \frac{-2 \cdot B' \cdot D_1^2 \cdot \rho^2}{(1+C_1')^3}$. Here, $C > 0$ because $B' = -\frac{\pi^2}{16} < 0$, and $C_1' = \frac{\pi}{8} - \frac{1}{2} > -1$. ◀