

A LATTICE ON DYCK PATHS CLOSE TO THE TAMARI LATTICE

JEAN-LUC BARIL¹, SERGEY KIRGIZOV¹, AND MEHDI NAIMA²

Dedicated to Jean Pallo on the occasion of his 75th birthday.

ABSTRACT. We introduce a new poset structure on Dyck paths where the covering relation is a particular case of the relation inducing the Tamari lattice. We prove that the transitive closure of this relation endows Dyck paths with a lattice structure. We provide a trivariate generating function counting the number of Dyck paths with respect to the semilength, the numbers of outgoing and incoming edges in the Hasse diagram. We deduce the numbers of coverings, meet and join irreducible elements. As a byproduct, we present a new involution on Dyck paths that transports the bivariate of the numbers of outgoing and incoming edges into its reverse. Finally, we give a generating function for the number of intervals, and we compare this number with the number of intervals in the Tamari lattice.

1. INTRODUCTION AND MOTIVATION

Many various classes of combinatorial objects are enumerated by the well-known Catalan numbers. For instance, it is the case of Dyck paths, planar trees, rooted binary trees, triangulations, Young tableaux, non-associative products, stack sortable permutations, permutations avoiding a pattern of length three, and so on. A list of over 60 types of such combinatorial classes of independent interest has been compiled by Stanley [27]. Generally, these classes have been studied in the context of the enumeration according to the length and given values of some parameters. Many other works investigate structural properties of these sets from order theoretical point of view. Indeed, there exist several partial ordering relations on Catalan sets which endow them with an interesting lattice structure [2, 3, 4, 21, 24, 29]. Of much interest is probably the so-called Tamari lattice [17, 29] which can be obtained equivalently in different ways. The coverings of the Tamari lattice could be different kinds of elementary transformations as reparenthesizations of letters products [16], rotations on binary trees [22, 25], diagonal flips in triangulations [25], and rotations on Dyck paths [5, 7, 11, 22]. The Tamari lattice appears in different domains. Its Hasse diagram is a graph of the polytope called associahedron; it is a Cambrian lattice underlying the combinatorial structure of Coxeter groups; and it has many enumerative properties with tight links with combinatorial objects such as planar maps [8, 9, 10, 12, 19].

In this paper, we introduce a new poset structure on Dyck paths where the covering relation is a particular case of the covering relation that generates the Tamari lattice. This is a new attempt to study a lattice structure on a Catalan set, and since it is close to the Tamari lattice, its study becomes natural. We prove that the transitive closure of this relation endows Dyck paths with a lattice structure. We provide a trivariate generating function

¹LIB, Université de Bourgogne, B.P. 47 870, 21078 Dijon Cedex France

²Sorbonne Université, CNRS, LIP6, F-75005 Paris, France

E-mail addresses: barjl@u-bourgogne.fr, sergey.kirgizov@u-bourgogne.fr, mehdi.naima@lip6.fr.

for the number of Dyck paths with respect to the semilength, and the statistics \mathbf{s} and \mathbf{t} giving the number of outgoing edges and the number of incoming edges, respectively. As a byproduct, we obtain generating functions and close forms for the numbers of meet and join irreducible elements, and for the number of coverings. An asymptotic is given for the ratio between the numbers of coverings in the Tamari lattice and those in our lattice. Also, we exhibit an involution on the set of Dyck paths that transports the bivariate (\mathbf{s}, \mathbf{t}) into (\mathbf{t}, \mathbf{s}) . Finally, we provide the generating function for the number of intervals and we offer some open problems.

2. NOTATION AND DEFINITIONS

In this section, we provide necessary notation and definitions in the context of Dyck paths, combinatorics and order theory.

Definition 1. *A Dyck path is a lattice path in \mathbb{N}^2 starting at the origin, ending on the x -axis and consisting of up steps $U = (1, 1)$ and down steps $D = (1, -1)$.*

Let \mathcal{D}_n be the set of Dyck paths of semilength n (i.e., with $2n$ steps), and $\mathcal{D} = \cup_{n \geq 0} \mathcal{D}_n$. The cardinality of \mathcal{D}_n is the n -th Catalan number $c_n = (2n)!/(n!(n+1)!)$ (see [A000108](#) in [26]). For instance, the set \mathcal{D}_3 consists of the five paths $UDUDUD$, $UUDDUD$, $UDUDD$, $UUDUDD$ and $UUUDDD$.

In this work, we will use the *first return decomposition* of a Dyck path, $P = URDS$, where $R, S \in \mathcal{D}$, and the *last return decomposition* of a Dyck path, $P = RUSD$, where $R, S \in \mathcal{D}$. Such a decomposition is unique and will be used to obtain a recursive description of the set \mathcal{D} . A Dyck path having a first return decomposition with S empty will be called *prime*, which means that the path touches the x -axis only at the origin and at the end.

Definition 2. *A peak in a Dyck path is an occurrence of the subpath UD . A pyramid is a maximal occurrence of $U^k D^\ell$, $k, \ell \geq 1$, in the sense that this occurrence cannot be extended in a occurrence of $U^{k+1} D^\ell$ or in a occurrence of $U^k D^{\ell+1}$.*

We say that a pyramid $U^k D^\ell$ is *symmetric* whenever $k = \ell$, and *asymmetric* otherwise. The *weight* of a symmetric pyramid $U^k D^k$ is k . For instance, the Dyck path in the south west of Figure 1 contains two symmetric pyramids and two asymmetric pyramids.

A *statistic* on the set \mathcal{D} of Dyck paths is a function \mathbf{f} from \mathcal{D} to \mathbb{N} , and a *multistatistic* is a tuple of statistics $(\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_t)$, $t \geq 2$. Given two statistics (or multistatistics) \mathbf{f} and \mathbf{g} , we say that they have the same *distribution* (or equivalently, are *equidistributed*) if, for any $k \geq 0$,

$$\text{card}\{P \in \mathcal{D}, \mathbf{f}(P) = k\} = \text{card}\{P \in \mathcal{D}, \mathbf{g}(P) = k\}.$$

Below, we define two important statistics for our study.

Definition 3. *Let \mathbf{s} be the statistic on \mathcal{D} where $\mathbf{s}(P)$ is the number of occurrences in P of $DU^k D^k$, $k \geq 1$. Let \mathbf{t} be the statistic on \mathcal{D} where $\mathbf{t}(P)$ is the number of occurrences in P of $U^k D^k D$, $k \geq 1$.*

For instance, for the path P in the south west of Figure 1, we have $\mathbf{s}(P) = 3$ and $\mathbf{t}(P) = 4$.

We end this section by defining the main concepts of order theory that we will use in this paper. We can find all these definitions in [16] for instance.

A poset \mathcal{L} is a set endowed with a partial order relation. Given two elements $P, Q \in \mathcal{L}$, a *meet* (or *greatest lower bound*) of P and Q , denoted $P \wedge Q$, is an element R such that $R \leq P$, $R \leq Q$, and for any S such that $S \leq P$, $S \leq Q$ then we have $S \leq R$. Dually, a *join* (or *least upper bound*) of P and Q , denoted $P \vee Q$, is an element R such that $P \leq R$, $Q \leq R$, and for any S such that $P \leq S$, $Q \leq S$ then we have $R \leq S$. Notice that join and meet elements do not necessarily exist in a poset. A *lattice* is a poset where any pair of elements admits a meet and a join.

Definition 4. An element $P \in \mathcal{L}$ is *join-irreducible* (resp., *meet-irreducible*) if $P = R \vee S$ (resp., $P = R \wedge S$) implies $P = R$ or $P = S$.

Definition 5. An interval I in a poset \mathcal{L} is a subset of \mathcal{L} such that for any $P, Q \in I$, and any $R \in \mathcal{L}$, if $P \leq R$ and $R \leq Q$, then R is also in I .

In 1962 [29], the Tamari lattice \mathcal{T}_n of order n is defined by endowing the set \mathcal{D}_n with the transitive closure \preceq of the covering relation

$$P \xrightarrow{\mathcal{T}} P'$$

that transforms an occurrence of $DUQD$ in P into an occurrence $UQDD$ in P' , where Q is a Dyck path (possibly empty). The top part of Figure 1 shows an example of such a covering, and Figure 5 illustrates the Hasse diagram of \mathcal{T}_n for $n = 4$ (the red edge must be considered). The number of meet (resp., join) irreducible elements is $n(n-1)/2$, and the number of coverings is $(n-1)c_n/2$ [15] where c_n is the n -th Catalan number. In 2006, Chapoton [10] proves that the number of intervals in \mathcal{T}_n is

$$\frac{2}{n(n+1)} \binom{4n+1}{n-1}.$$

Later, Bernardi and Bonichon [6] exhibit a bijection between intervals in \mathcal{T}_n and minimal realizers.

Now, we introduce a new partial order on \mathcal{D}_n for $n \geq 0$. We endow it with the ordered relation \leq defined by the transitive closure of the following covering relation

$$P \longrightarrow P'$$

that transforms an occurrence of $DU^k D^k$ in P into an occurrence $U^k D^k D$ in P' , where $k \geq 1$. For short, we will often use the notation

$$DU^k D^k \longrightarrow U^k D^k D, \quad k \geq 1$$

whenever we need to show where the transformation is applied.

Notice that the covering \longrightarrow is a particular case of the Tamari covering $\xrightarrow{\mathcal{T}}$ whenever we take $Q = U^{k-1} D^{k-1}$ in the transformation $DUQD \xrightarrow{\mathcal{T}} UQDD$. Let \mathcal{S}_n be the poset (\mathcal{D}_n, \leq) . The bottom part of Figure 1 shows an example of such a covering, and Figure 2 illustrates the Hasse diagram of \mathcal{S}_n for $n = 4$ (without the red edges which belongs to the Tamari lattice but not to this new poset). See also Figure 7 for an illustration of Hasse diagram of the Tamari lattice and this poset for the case $n = 6$.

In our study, the following facts will sometimes be used explicitly or implicitly.

Fact 1. If $P \longrightarrow P'$, then the path P' is above the path P , that is, for any points $(x, y) \in P'$ and $(x, z) \in P$ we have $y \geq z$.

Fact 2. If $P \longrightarrow P'$ then $\mathfrak{s}(P) - 1 \leq \mathfrak{s}(P') \leq \mathfrak{s}(P)$ and $\mathfrak{t}(P) \leq \mathfrak{t}(P') + 1$.

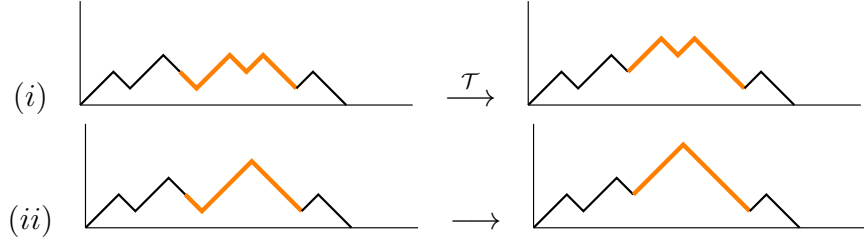


FIGURE 1. (i) corresponds to a covering relation for the Tamari Lattice, while (ii) corresponds to the covering relation for the new lattice of this study.

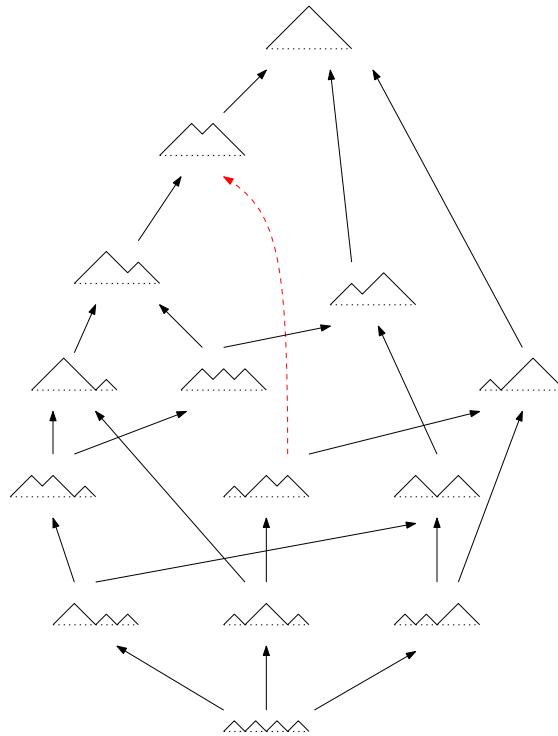


FIGURE 2. The Hasse diagram of $\mathcal{S}_4 = (\mathcal{D}_4, \leq)$. The Tamari lattice $\mathcal{T}_4 = (\mathcal{D}_4, \preceq)$ can be viewed by considering the dotted edge (in red).

3. LATTICE STRUCTURE OF $\mathcal{S}_n = (\mathcal{D}_n, \leq)$.

In this section, we prove that the poset $\mathcal{S}_n = (\mathcal{D}_n, \leq)$ is a lattice and provide some related results.

Lemma 1. *For $n \geq 2$, any Dyck path $P \in \mathcal{D}_n$, $P \neq U^n D^n$, contains at least one occurrence of $DU^k D^k$ for some $k \geq 1$.*

Proof. For $n \geq 2$, in any path from \mathcal{D}_n not equal to $U^n D^n$, there exists an occurrence of DU , and the rightmost occurrence of DU always starts an occurrence of $DUU^\ell D^\ell D$ for some $\ell \geq 0$. \square

Lemma 2. *For $n \geq 2$, any Dyck path $P \in \mathcal{D}_n$, $P \neq (UD)^n$, contains at least one occurrence of $U^k D^k D$ for some $k \geq 1$, and then P contains at least one occurrence of UDD .*

Proof. By contradiction, let us assume that P does not contain an occurrence UDD . This means that any peak UD is either at the end of P , or it precedes an up step U , which implies that a down step cannot be contiguous to another down step. Thus, $P = (UD)^n$ which contradicts the hypothesis $P \neq (UD)^n$. \square

Proposition 1. *For any Dyck path $P \in \mathcal{D}_n$, we have $P \leq U^n D^n$ and $(UD)^n \leq P$, which means that the poset (\mathcal{D}_n, \leq) admits a maximum element and a minimum element.*

Proof. It suffices to apply Lemma 1 and Lemma 2. Indeed, let P be a Dyck path in \mathcal{D}_n , $P \neq U^n D^n$. Using Lemma 1, P contains at least one occurrence of $DU^k D^k$, $k \geq 1$. Let P_1 be the Dyck path obtained from P after applying the covering $DU^k D^k \rightarrow U^k D^k D$ on this occurrence. Due to Fact 1, any point (x, y) in P is below the point (x, z) in P_1 (i.e., $y \leq z$). Iterating the process from P_1 , we construct a sequence of coverings $P \rightarrow P_1 \rightarrow \dots \rightarrow P_r$, $r \geq 1$, of Dyck paths that necessarily converges towards a Dyck path without occurrence of $DU^k D^k$ for some $k \geq 1$, i.e., towards $U^n D^n$, which implies $P \leq U^n D^n$.

By applying a similar argument, we can prove easily the second inequality. \square

Proposition 2. *Given $P, Q \in \mathcal{D}_n$ satisfying $P \leq Q$, $P \neq Q$, and such that $P = RDS$ and $Q = RUS'$ (R is the maximal common prefix of P and Q). Let W the Dyck path obtained from P by applying the covering $P \rightarrow W$ on the leftmost occurrence of $DU^k D^k$, $k \geq 1$, in DS , then we necessarily have $W \leq Q$.*

Proof. Let $P_0 = P \rightarrow P_1 \rightarrow \dots \rightarrow P_k = Q$ be a sequence of coverings from P to Q . Let us suppose that the first covering is not applied on the leftmost occurrence of $DU^k D^k$, $k \geq 1$, in DS (we call D_0 the first step of this occurrence). Let $P_i \rightarrow P_{i+1}$ be the covering involving D_0 . Necessarily, the down step D_0 is followed by $U^\ell D^\ell$ in P_i where $\ell \geq k$. Then, all coverings between P and P_i occur on the right of $D_0 U^k$. We call them $\alpha_1, \dots, \alpha_a$, $a \geq 1$. So, there is another sequence of coverings from P to P_{i+1} by applying the following: we first apply the covering β involving $D_0 U^k D^k$ in P , then we obtain the path $RU^k D^k DS$; we apply the coverings $\alpha'_1, \dots, \alpha'_a$ (in the same order) where α'_i is the covering involving the same occurrence of $DU^b D^b$ moved by α_i , then we obtain $RU^k DU^{\ell-k} D^\ell S'$; finally, we apply an additional covering in order to obtain $P_{i+1} = RU^\ell D^\ell DS'$ (see below for an illustration of this process).

$$\begin{array}{ccccccc}
 P = RD_0 U^k D^k S & \xrightarrow{\alpha_1} & \cdots & \xrightarrow{\alpha_a} & P_i = RD_0 U^\ell D^\ell S' & \longrightarrow & P_{i+1} = RU^\ell D^\ell DS' \\
 \downarrow \beta & & & & & & \nearrow \\
 RU^k D^k DS & \xrightarrow{\alpha'_1} & \cdots & \xrightarrow{\alpha'_a} & RU^k DU^{\ell-k} D^\ell S' & &
 \end{array}$$

Therefore, we can reach Q from P by first applying the covering involving $D_0U^kD^k$, which completes the proof. \square

As a byproduct of Proposition 2 and by a straightforward induction, we have the following corollary.

Corollary 1. *The longest chain between $(UD)^n$ and U^nD^n is of length $n(n-1)/2$.*

Proof. Thanks to Proposition 2, the longest chain between $(UD)^n$ and U^nD^n is unique and it can be constructed from $(UD)^n$ by applying at each step the leftmost covering. So, the length of this chain is given by $1 + 2 + \dots + (n-1)$, which gives the expected result. \square

Theorem 1. *The poset (\mathcal{D}_n, \leq) is a lattice.*

Proof. For any $P, Q \in \mathcal{D}$, we need to prove that P and Q admit a join and a meet elements. Let us start to give the proof of the existence of a join element. We proceed by induction on the semilength of the Dyck paths. For $n \leq 3$, $\mathcal{S}_n = (\mathcal{D}_n, \leq)$ is isomorphic to the Tamari lattice.

Now, let us assume that $\mathcal{S}_n = (\mathcal{D}_n, \leq)$ is a lattice for $n \leq N$, and let us prove the result for $N+1$. Let P and Q two paths in \mathcal{D}_{N+1} . We distinguish two cases according to the form of the first return decompositions of P and Q .

- (1) If $P = URDS$ and $Q = UR'DS'$ where R and R' have the same length. Then we apply the recurrence hypothesis for R and R' (resp. S and S'), which means that $R \vee R'$ (resp., $S \vee S'$) exists. Therefore, the path $U(R \vee R')D(S \vee S')$ is necessarily the least upper bound of P and Q , which proves that $P \vee Q$ exists.
- (2) Now, let us suppose that $P = URDS$ and $Q = UR'DS'$ where the length r' of R' is strictly less than the length r of R . Let M be an upper bound of P and Q (there is at least one thanks to Proposition 1). Since $r' < r$ and due to Fact 1, M has necessarily a decomposition $M = UM_1DM_2$ where the length of M_1 is at least r . Therefore, in any sequence of coverings $Q \rightarrow \dots \rightarrow M$ from Q to M , there is necessarily a covering that involves and elevates the down-step just after R' . Due to the definition of the covering \longrightarrow , we can apply such a transformation only when this down-step is followed by an occurrence U^kD^k for some $k \geq 1$.

Assuming $S' = US_1DS_2$ and using Proposition 2, we deduce that the inequality $Q = UR'DUS_1DS_2 \leq M$ is equivalent to $UR'DU^kD^kS_2 \leq M$ where $k \geq 1$ is the semilength of US_1D . Moreover, this condition is equivalent to $Q_1 := UR'U^kD^kDS_2 \leq M$ (see Figure 3). It is worth noting that Q_1 does not depend on the upper bound M . Iterating this process with P and Q_1 , we can construct two Dyck paths P' and Q' such that the condition $P \leq M$ and $Q \leq M$ is equivalent to $P' \leq M$ and $Q' \leq M$ where P' and Q' (that do not depend on M) are two Dyck paths lying in the first case of the proof. Using the hypothesis recurrence, we conclude that $P' \vee Q' = P \vee Q$ exists.

Considering the two cases, the induction is completed.

The existence of greatest lower bound then follows automatically since the poset is finite with a least element and a greatest element. \square

4. COVERINGS, JOIN AND MEET IRREDUCIBLE ELEMENTS

For a given Dyck path P , the number of incoming edges (in the Hasse diagram) corresponds to the number $\mathfrak{t}(P)$ of occurrences U^aD^aD , $a \geq 1$, in P , and the number of outgoing edges

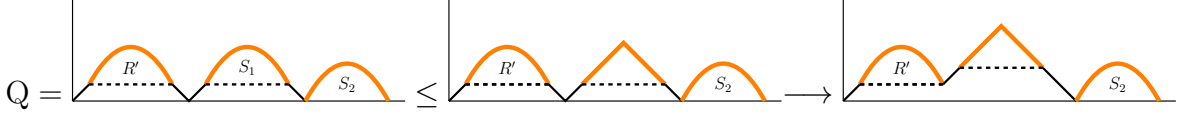


FIGURE 3. An illustration of the construction $Q = UR'DUS_1DS_2 \leq UR'DU^kD^kS_2 \rightarrow UR'U^kD^kDS_2$ of the case (ii) in the proof of Theorem 1.

(coverings) corresponds to the number $\mathfrak{s}(P)$ of occurrences DU^aD^a , $a \geq 1$, in P . Let $A(x, y, z)$ be the trivariate generating function where the coefficient of $x^n y^k z^\ell$ is the number of Dyck paths of semilength n having k possible coverings (or equivalently k outgoing edges), and ℓ incoming edges.

Theorem 2. *We have*

$$A(x, y, z) = \frac{R(x, y, z) - \sqrt{4x(xzy - xy - xz + 1)(xy + xz - x - 1) + R(x, y, z)^2}}{2x(xzy - xy - xz + 1)},$$

where $R(x, y, z) = x^2zy - x^2y - x^2z + x^2 - xy - xz + x + 1$.

Proof. For short, we set $A := A(x, y, z)$. We consider the last return decomposition of a nonempty Dyck path P , that is $P = RUSD$ where R and S are two Dyck paths. We distinguish six cases.

- (1) If R and S are empty, then $P = UD$ and the g.f. for this path is x .
- (2) If R is not empty and S is empty, the g.f. for these paths is $(A - 1)xy$.
- (3) If R is empty and $S = U^aD^a$ with $a \geq 1$, then the g.f. for these paths is $\frac{x^2z}{1-xz}$.
- (4) If R is not empty and $S = U^aD^a$ with $a \geq 1$, then the g.f. for these paths is $\frac{x^2z}{1-xz}(A - 1)y$.
- (5) If $S = S'U^aD^a$ with $a \geq 1$ and S' not empty (R is possibly empty), then the g.f. for these paths is $\frac{x^2z}{1-xz}(A - 1)yA$.
- (6) If S does not end with U^aD^a , $a \geq 1$, then the g.f. for these paths is AxB where B is the g.f. for nonempty Dyck paths that do not end with a pyramid U^aD^a , $a \geq 1$. Using the complement, we have $B = A - 1 - x - \frac{x^2z}{1-xz} - x(A - 1)y - \frac{x^2z}{1-xz}(A - 1)y$ where 1 is for the empty path; x is for UD ; $\frac{x^2z}{1-xz}$ is for the paths of the form U^aD^a , $a \geq 2$; $x(A - 1)y$ is for the paths RUD with R not empty; and $\frac{x^2z}{1-xz}(A - 1)y$ is for the paths RU^aD^a , $a \geq 2$.

Summarizing the three cases, we obtain the functional equation

$$A = 1 + x + (A - 1)xy + \frac{x^2z}{1 - xz} + \frac{x^2z}{1 - xz}(A - 1)y + \frac{x^2z}{1 - xz}(A - 1)yA + \\ + Ax \left(A - 1 - x - \frac{x^2z}{1 - xz} - x(A - 1)y - \frac{x^2z}{1 - xz}(A - 1)y \right),$$

and a simple calculation provides the result. \square

The first terms of the series expansion are

$$1 + x + (z + y)x^2 + (y^2 + 3yz + z^2)x^3 + (y^3 + 5y^2z + 5yz^2 + z^3 + 2yz)x^4 + \\ + (y^4 + 7y^3z + 13y^2z^2 + 7yz^3 + z^4 + 5y^2z + 5yz^2 + 3yz)x^5 + O(x^6).$$

Remark 1. Observe that $R(x, y, z) = R(x, z, y)$ and $A(x, y, z) = A(x, z, y)$, which means that (\mathbf{t}, \mathbf{s}) and (\mathbf{s}, \mathbf{t}) are equidistributed on \mathcal{D}_n , $n \geq 1$. In the following, we will show the existence of an involution on \mathcal{D}_n that transports the bivariate statistic (\mathbf{t}, \mathbf{s}) into (\mathbf{s}, \mathbf{t}) (see Definition 3).

Corollary 2. The generating function $E(x)$ where the coefficient of x^n is the total number of possible coverings over all Dyck paths of semilength n (or equivalently the number of edges in the Hasse diagram) is

$$E(x) = \frac{-1 + 4x + (1 - 2x)\sqrt{1 - 4x}}{2(1 - 4x)(1 - x)}.$$

The coefficient of x^n is given by

$$\sum_{k=0}^{n-2} \binom{2k+2}{k}.$$

The ratio between the numbers of coverings in \mathcal{T}_n and \mathcal{S}_n tends towards $3/2$.

Proof. We obtain $E(x)$ by calculating $\partial_y(A(x, y, 1))|_{y=1}$. Now we set $F(x) = \frac{E(x)(1-x)}{x}$. Noticing that $F(x) = x \cdot \partial_x(C(x))$ where $C(x)$ is the generating function for the Catalan numbers satisfying $C(x) = 1 + xC(x)^2$. So, we deduce directly that

$$f_n := [x^n]F(x) = \frac{n}{n+1} \binom{2n}{n}.$$

Then, we have

$$[x^n]E(x) = \sum_{k=1}^{n-1} f_k = \sum_{k=1}^{n-1} \frac{k}{k+1} \binom{2k}{k} = \sum_{k=0}^{n-2} \binom{2k+2}{k}.$$

Considering the asymptotics of $[x^n]E(x)$ and $\frac{2}{n(n+1)} \binom{4n+1}{n-1}$ (using classical method, see [13] for instance), the limit of the ratio between the number of coverings in \mathcal{T}_n and \mathcal{S}_n is $3/2$. \square

The first terms of the series expansion of $E(x)$ are

$$x^2 + 5x^3 + 20x^4 + 76x^5 + 286x^6 + 1078x^7 + 4081x^8 + 15521x^9 + O(x^{10}),$$

and the sequence of coefficients corresponds to [A057552](#) in [26].

Let $K(x)$ be the generating function where the coefficient of x^n is the number of meet irreducible elements (Dyck paths with only one outgoing edge). Using the symmetry $y \longleftrightarrow z$ in $A(x, y, z)$, $K(x)$ also is the generating function where the coefficient of x^n is the number of join irreducible elements (Dyck paths with only one incoming edge).

Corollary 3. We have

$$K(x) = \frac{x^2}{(x-1)(2x-1)},$$

and the coefficient of x^n is $2^{n-1} - 1$ for $n \geq 1$.

Below, we define recursively an involution ϕ from \mathcal{D} into itself as follows:

$$\left\{ \begin{array}{lll} \phi(\epsilon) & = \epsilon, & (i) \\ \phi(U^k D^k) & = (UD)^k, & k \geq 1 \quad (ii) \\ \phi(RU^k D^k) & = \phi(R)^\sharp(UD)^{k-1}, & k \geq 2 \quad (iii) \\ \phi(R_0UR_1 \dots UR_kUD^{k+1}) & = \phi(R_0)U\phi(R_1) \dots U\phi(R_k)UD^{k+1}, & k \geq 1 \quad (iv) \\ \phi(R_0UR_1 \dots UR_kU^{r+1}D^{k+r+1}) & = \phi(R_0)U\phi(R_1) \dots U\phi(R_k)UD^{k+1}(UD)^r, & k, r \geq 1 \quad (v) \end{array} \right.$$

where R and R_k are a nonempty Dyck paths, and R_i , $0 \leq i \leq k-1$, are Dyck paths (possibly empty), and $\phi(R)^\sharp$ is obtained from $\phi(R)$ by inserting an occurrence UD on the last peak of $\phi(R)$, i.e., if $\phi(R) = R'UD^k$, then $\phi(R)^\sharp = R'UUDD^k$.

It is worth noticing that the definition of ϕ is given on a subset \mathcal{P}_1 of \mathcal{D} , consisting of Dyck paths of the forms ϵ , $U^k D^k$ with $k \geq 1$, $RU^k D^k$ with $k \geq 2$ and R nonempty, and Dyck paths in $\mathcal{D}^1 \cup \mathcal{D}^2$. Moreover, we have $\phi(\mathcal{D}^1) = \mathcal{D}^1$, and $\phi(\mathcal{P}_1 \setminus \mathcal{D}^1) = \mathcal{D} \setminus \mathcal{D}^1$. By considering the map ϕ on \mathcal{P}_1 , and its converse on $\mathcal{D} \setminus \mathcal{P}_1$, we obtain an involution ϕ on the entire set \mathcal{D} . For instance, we have $\phi(UUUDDD) = UDUDUD$, $\phi(UUDDUD) = UDUUDD$, $\phi(UDUUDD) = UUDDUD$, $\phi(UUDUDD) = UUDUDD$ and $\phi(UDUDUD) = UUUDDD$.

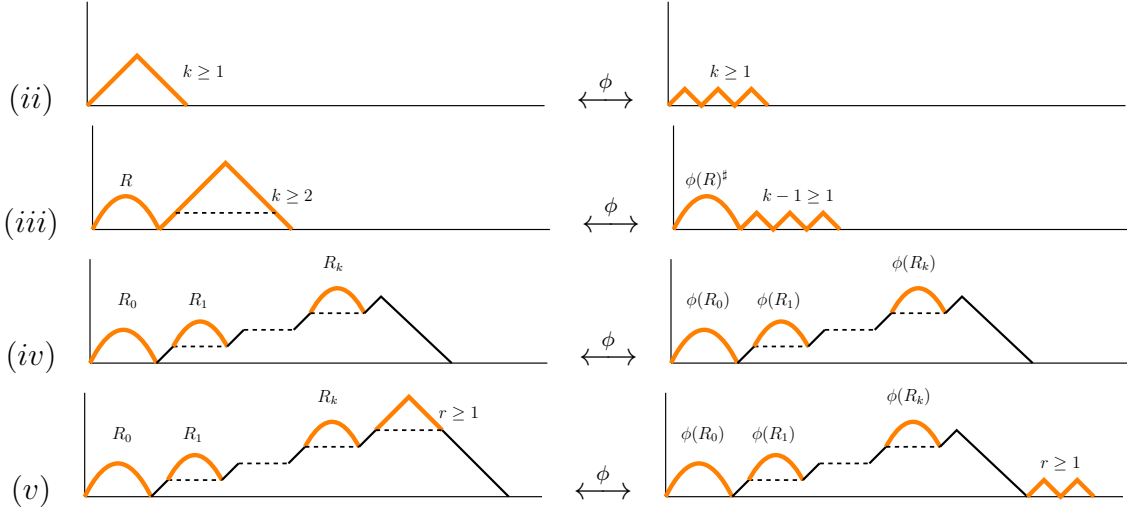


FIGURE 5. An illustration of the involution ϕ for the cases (ii)-(v).

Let P be a Dyck path, we recall (see Definition 3) that $\mathbf{s}(P)$ (resp. $\mathbf{t}(p)$) gives the number of occurrences $DU^k D^k$ (resp. $U^k D^k D$), $k \geq 1$, in P .

Theorem 3. *The map ϕ is an involution on \mathcal{D} that preserves the semilength, and that transports the bistatistic (\mathbf{s}, \mathbf{t}) into (\mathbf{t}, \mathbf{s}) . Moreover, ϕ preserves the number of asymmetric pyramids, and preserves the sum of the weights of symmetric pyramids.*

Proof. We proceed by induction on the semilength n of the paths. We assume the result for paths of semilength at most n , and we prove the result for the semilength $n+1$. We distinguish four cases according to the definition of ϕ (we omit the case (i) since $n+1 \geq 1$).

- (ii) If $P = U^{n+1}D^{n+1}$ then we have $\phi(P) = (UD)^{n+1}$, $\mathfrak{s}(P) = 0 = \mathfrak{t}(\phi(P))$, and $\mathfrak{t}(P) = n = \mathfrak{s}(\phi(P))$.
- (iii) If $P = RU^kD^k$, with R nonempty and $k \geq 2$, then we have $\phi(P) = \phi(R)^\sharp(UD)^{k-1}$, $\mathfrak{s}(P) = \mathfrak{t}(R) + 1$, and using the recurrence hypothesis, this equals to $\mathfrak{t}(\phi(R)) + 1 = \mathfrak{t}(\phi(P)^\sharp)$. Moreover, we have $\mathfrak{t}(P) = \mathfrak{t}(R) + k - 1$ and using the recurrence hypothesis, this equals to $\mathfrak{s}(\phi(R)) + k - 1 = \mathfrak{s}(\phi(P)^\sharp) + k - 1 = \mathfrak{s}(\phi(P)^\sharp(UD)^{k-1}) = \mathfrak{s}(\phi(P))$.
- (iv) If $P = R_0UR_1 \dots UR_kUD^{k+1}$, then we have $\phi(P) = \phi(R_0)U\phi(R_1) \dots U\phi(R_k)UD^{k+1}$, and $\mathfrak{s}(P) = \sum_{i=0}^k \mathfrak{s}(R_i) + 1$, and using the recurrence hypothesis, this equals to $\sum_{i=0}^k \mathfrak{t}(\phi(R_i)) + 1 = \mathfrak{t}(\phi(R_0)U\phi(R_1) \dots U\phi(R_k)UD^{k+1}) = \mathfrak{t}(\phi(P))$. A same argument allows us to prove $\mathfrak{t}(P) = \mathfrak{s}(\phi(P))$.
- (v) If $P = R_0UR_1 \dots UR_kUD^{k+1}(UD)^r$, then we have $\phi(P) = \phi(R_0)U\phi(R_1) \dots U\phi(R_k)UD^{k+1}(UD)^r$, and $\mathfrak{s}(P) = \sum_{i=0}^k \mathfrak{s}(R_i) + 1$, and using the recurrence hypothesis, this equals to $\sum_{i=0}^k \mathfrak{t}(\phi(R_i)) + 1 = \mathfrak{t}(\phi(R_0)U\phi(R_1) \dots U\phi(R_k)UD^{k+1}(UD)^r) = \mathfrak{t}(\phi(P))$. The equality $\mathfrak{t}(P) = \mathfrak{s}(\phi(P))$ is obtained in the same way.

Considering all these cases, we obtain the result by induction. □

Notice that Theorem 3 allows us to retrieve the symmetry obtained by Theorem 2, but unfortunately, the involution ϕ does not induce a symmetry on the lattice (\mathcal{D}_n, \leq) .

5. ENUMERATION OF INTERVALS

In this section, we provide the generating function and a close form for the number of intervals in the lattice \mathcal{S}_n . The method is inspired by the work of Bousquet-Mélou and Chapoton [9]. We will use a catalytic variable that considers the size of last run of down-steps. We introduce the bivariate generating function

$$I(x, y) = \sum_{n, k \geq 1} a_{n, k} x^n y^k,$$

where $a_{n, k}$ corresponds to the number of intervals in \mathcal{S}_n such that the upper path ends with k down-steps exactly. We also define

$$J(x, y) = \sum_{n, k \geq 1} b_{n, k} x^n y^k,$$

where $b_{n, k}$ corresponds to the number of intervals in \mathcal{S}_n such that the upper path is prime and ends with k down-steps exactly (recall that a Dyck path is prime whenever it only touches the x -axis at its beginning and its end).

Lemma 3. *The following functional equation holds*

$$I(x, y) = J(x, y) + I(x, 1) \cdot J(x, y).$$

Proof. Let (P, Q) be an interval in \mathcal{S}_n where P is the lower bound and Q the upper bound. We distinguish two cases.

- (1) If Q is prime, then the contribution for these intervals is simply $J(x, y)$.

- (2) Otherwise, Q is not prime and it has a last return decomposition $Q = RUSD$ with $R, S \in \mathcal{D}$ and R not empty. This implies that P is of the form $P = P_1P_2$ where $P_1, P_2 \in \mathcal{D}$, and P_2 and USD have the same length. This induces a bijection from intervals in this case with pairs of intervals $I_1 := (P_1, R)$ and $I_2 := (P_2, USD)$, where I_1 is any interval of smaller length, and I_2 is any interval (of smaller length) lying into the case (1). Therefore, the contribution for the intervals in this case is $I(x, 1) \cdot J(x, y)$.

Considering the two cases, we obtain the expected result. □

Lemma 4. *The following functional equation holds*

$$J(x, y) = xy + xyI(x, y) + \frac{J(x, y) - J(x, 1)}{y - 1} \cdot C(xy)xy^2,$$

where $C(x)$ is the g.f. for Catalan numbers, i.e., $C(x) = 1 + xC(x)^2$.

Proof. Let (P, Q) be an interval in \mathcal{S}_n where P is the lower bound and Q the upper bound. We distinguish three cases.

- (1) If $P = UD$ and $Q = UD$, then the contribution is xy .
- (2) If P is prime, then we have $P = UP'D$ and we necessarily have $Q = UQ'D$ (i.e., Q is prime) where P' and Q' are nonempty Dyck paths. Thus, the contribution for these intervals is $xyI(x, y)$.
- (3) Otherwise, P is not prime which means that it has a last return decomposition $P = RUSD$ with $R, S \in \mathcal{D}$ and R not empty (see the bottom of Figure 6). As Q is prime, in any path of coverings $P \rightarrow P_1 \rightarrow \dots \rightarrow Q$ from P to Q , there is necessarily a covering that involves and elevates the up step just after R . We can apply such a transformation only when the suffix P delimited by this up step is of the form U^kD^k for some $k \geq 1$. This condition implies that Q is necessarily of the following form $Q = Q'U^kD^{k+\ell}$ where Q' is a prefix of Dyck path and $\ell \geq 1$ (see the top of Figure 6 for an illustration of the form of $Q = Q'U^kD^{k+\ell}$). Let $h \geq 1$ be the height of the right point of the last up step of Q' . Then, we necessarily have $h \geq \ell \geq 1$.

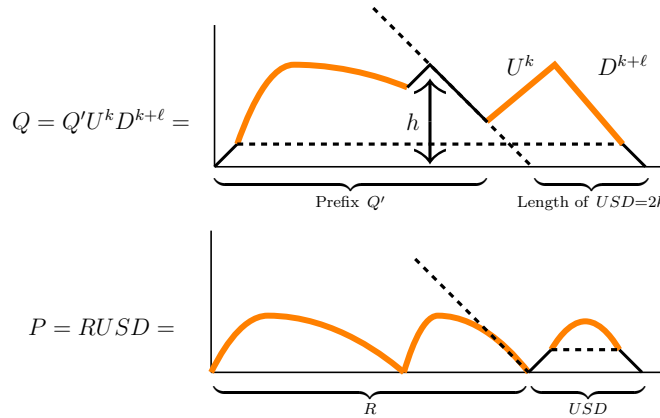


FIGURE 6. The form of the upper bound $Q = Q'U^kD^{k+\ell}$, $\ell \geq 1$, and the form of the lower bound $P = RUSD$.

On the other hand, any Q of this form (i.e., $Q = Q'U^kD^{k+\ell}$ with $h \geq \ell \geq 1$) is candidate to an upper bound of an interval $(P, Q) = (RUSD, Q'U^kD^{k+\ell})$ if and only if $(R, Q'D^\ell)$ and (USD, U^kD^k) are two intervals, which is equivalent to the condition $(R, Q'D^\ell)$ is an interval (indeed, (USD, U^kD^k) is always an interval).

So, for each interval of the form $(R, Q'D^\ell)$, $h \geq \ell \geq 1$, we can construct h intervals of the form $(RUSD, Q''U^kD^{k+\ell})$, $1 \leq \ell \leq h$, where Q'' is the greatest prefix of Q' ending by an up step, i.e., the intervals

$$\begin{aligned} & (RUSD, Q''D^{h-1}U^kD^{k+1}), \\ & (RUSD, Q''D^{h-2}U^kD^{k+2}), \\ & \dots\dots\dots \\ & (RUSD, Q''DU^kD^{k+h-1}), \\ & (RUSD, Q''U^kD^{k+h}). \end{aligned}$$

Considering this study, the contribution for the intervals in this case is given by

$$xyC(xy) \cdot \sum_{h \geq 1} \sum_{n \geq 1} b_{n,h} x^n (y + y^2 + \dots + y^h) = xy^2C(xy) \cdot \sum_{h \geq 1} \sum_{n \geq 1} b_{n,h} x^n \frac{y^h - 1}{y - 1},$$

that can be expressed as

$$xy^2C(xy) \cdot \left(\sum_{h \geq 1} \sum_{n \geq 1} b_{n,h} x^n \frac{y^h}{y - 1} - \sum_{h \geq 1} \sum_{n \geq 1} b_{n,h} x^n \frac{1}{y - 1} \right) = C(xy)xy^2 \cdot \frac{J(x, y) - J(x, 1)}{y - 1}.$$

Considering the three cases, we obtain the expected result. \square

From Lemma 3 and Lemma 4 and after a straightforward substitution, we obtain a system of equations.

Theorem 4. *The following system of functional equations holds:*

$$\begin{cases} I(x, y) &= \frac{J(x, y)}{1 - J(x, 1)}, \\ J(x, y) &= xy + xy \frac{J(x, y)}{1 - J(x, 1)} + \frac{J(x, y) - J(x, 1)}{y - 1} \cdot C(xy)xy^2. \end{cases}$$

In order to compute $J(x, 1)$, we use the kernel method [1, 23] on

$$J(x, y) \cdot \left(1 - \frac{xy}{1 - J(x, 1)} - \frac{C(xy)xy^2}{y - 1} \right) = xy - \frac{J(x, 1)}{y - 1} \cdot C(xy)xy^2.$$

This method consists in cancelling the factor of $J(x, y)$ by finding y as an algebraic function y_0 of $J(x, 1)$ and x . So, if we substitute y with y_0 in the right hand side of the equation, then it necessarily equals zero (in order to counterbalance the cancellation the left hand side). So, we deduce the following system of equations that allows us to determine $J(x, 1)$, and then $J(x, y)$:

$$\begin{cases} 1 - \frac{xy_0}{1 - J(x, 1)} - \frac{C(xy_0)xy_0^2}{y_0 - 1} &= 0, \\ xy_0 - \frac{J(x, 1)}{y_0 - 1} \cdot C(xy_0)xy_0^2 &= 0. \end{cases}$$

So, we deduce $y_0 = \frac{1+4x-\sqrt{1-8x}}{8x}$, and the following theorem.

Theorem 5. *The generating function $J(x, y)$ for the number of intervals (P, Q) where Q is prime, with respect to the semilength and the size of the last run of down steps is*

$$J(x, y) = \frac{xy(-1 + J(x, 1))(J(x, 1)C(xy)y - y + 1)}{J(x, 1)C(xy)xy^2 - C(xy)xy^2 - xy^2 - J(x, 1)y + xy + J(x, 1) + y - 1}$$

with

$$J(x, 1) = \frac{1 - \sqrt{1 - 8x}}{4},$$

and $C(x)$ is the g.f. for Catalan numbers, i.e. $C(x) = 1 + xC(x)^2$.

The series expansion of $J(x, 1)$ is

$$x + 2x^2 + 8x^3 + 40x^4 + 224x^5 + 1344x^6 + 8448x^7 + 54912x^8 + 366080x^9 + O(x^{10}),$$

where the sequence of coefficients corresponds to the sequence [A052701](#) in [26] that counts outerplanar maps with a given number of edges [14]. The n -th coefficient is given by the close form

$$2^{n-1}c_{n-1},$$

where $c_n = (2n)!/(n!(n+1)!)$ is the n -th Catalan number [A000108](#) in [26].

The series expansion of $J(x, y)$ is

$$yx + 2y^2x^2 + (5y + 3)y^2x^3 + (14y^2 + 15y + 11)y^2x^4 + (42y^3 + 61y^2 + 68y + 53)y^2x^5 + (132y^4 + 233y^3 + 325y^2 + 363y + 291)y^2x^6 + O(x^7),$$

Theorem 6. *The generating function $I(x, y)$ for the number of intervals (P, Q) with respect to the semilength and the size of the last run of down steps is*

$$I(x, y) = J(x, y) \cdot \frac{3 - \sqrt{1 - 8x}}{2(x + 1)},$$

and the generating function $I(x, 1)$ for the number of intervals (P, Q) with respect to the semilength is

$$I(x, 1) = \frac{1 - 2x - \sqrt{1 - 8x}}{2(x + 1)}.$$

The series expansion of $I(x, 1)$ is

$$x + 3x^2 + 13x^3 + 67x^4 + 381x^5 + 2307x^6 + 14589x^7 + 95235x^8 + 636925x^9 + O(x^{10}),$$

where the sequence of coefficients corresponds to the sequence [A064062](#) in [26] that counts simple outerplanar maps with a given number of vertices [14]. The n -th coefficient is given by the close form

$$\frac{1}{n} \sum_{m=0}^{n-1} (n-m) \binom{n+m-1}{m} 2^m.$$

An asymptotic approximation for the ratio of the numbers of intervals in \mathcal{T}_n and \mathcal{S}_n is

$$\frac{2^{5n+\frac{5}{2}}}{n \cdot 3^{3n+\frac{1}{2}}}.$$

The series expansion of $I(x, y)$ is

$$yx + (2y + 1)yx^2 + (5y^2 + 5y + 3)yx^3 + (14y^3 + 20y^2 + 20y + 13)yx^4 + (42y^4 + 75y^3 + 98y^2 + 99y + 67)yx^5 + O(x^6).$$

6. GOING FURTHER

We discussed here several open questions related to the lattice $\mathcal{S}_n = (\mathcal{D}_n, \leq)$.

Question 1. *Finding a combinatorial interpretation of the equality $J(x, 1) = x + 2J(x, 1)^2$.*

An *outerplanar map* [14] is a connected planar multigraph with a specific embedding in the 2-sphere, up to oriented homeomorphisms, where a root edge is selected and oriented, and such that all its vertices are in the outer face.

Question 2. *Find a nice bijection between intervals in \mathcal{S}_n and the simple outerplanar maps with a given number of vertices.*

The *distance* between two Dyck paths P and Q is the length of a shortest path between P in Q in the underlying undirected graph of the poset.

Question 3. *Is there a polynomial time algorithm to compute the distance between two Dyck paths in \mathcal{S}_n ?*

The *diameter* is the maximum distance between any two vertices.

Question 4. *For $n \geq 3$, we conjecture that the diameter of \mathcal{S}_n is $2n - 4$, and that this value corresponds to the distance between $(UD)^n$ and $UU(UD)^{n-2}DD$.*

The *Möbius function* [28] of \mathcal{S}_n , $\mu : \mathcal{S}_n \rightarrow \mathbb{Z}$ is defined recursively by

$$\mu(P) = - \sum_{Q < P} \mu(Q) \text{ if } P \neq (UD)^n, \text{ anchored with } \mu((UD)^n) = 1.$$

Question 5. *For $n \geq 2$, is there an efficient and non-recursive algorithm to compute the Möbius function of \mathcal{S}_n as in [20]?*

An m -Dyck path of size n is a path on \mathbb{N}^2 , starting at the origin and ending at $(2nm, 0)$, consisting of up steps (m, m) and down steps $(1, -1)$. In the literature, these paths have been endowed with a lattice structure that generalizes the Tamari lattice [7, 8].

Question 6. *Is it possible to generalize this study for m -Dyck paths?*

The Tamari covering transforms an occurrence of $DUPD$ into $UPDD$ without any constraints on the structure of a Dyck path P . The covering relation discussed in this paper can be viewed as the following restriction of the Tamari covering: we only allow $DUPD \rightarrow UPDD$ whenever the path P avoids consecutive pattern DU . This opens the way to the further generalization: *pattern-avoiding Tamari poset* is given by the transitive closure of the covering $DUPD \rightarrow UPDD$ where P avoids a given consecutive pattern μ . It is easy to prove that any Dyck path of semilength n can be obtained from $(UD)^n$ by a sequence of such pattern-aware transformations for any μ . For $k \geq 1, n \geq k + 2$, if $\mu = U^k$, then the poset will have at least two non comparable maximal elements $U^n D^n$ and $UDU^{k+1} D^{k+1}$. In general the situation is not that simple, for example, if $\mu = UDU$ the poset has a minimum and

maximum element, but some pairs of paths have not a meet. However, if $\mu = UU DU$ then it seems that we obtain a lattice structure. So, it becomes natural to ask the following.

Question 7. *Could we characterize patterns μ inducing a lattice structure in the pattern-avoiding version of the Tamari poset?*

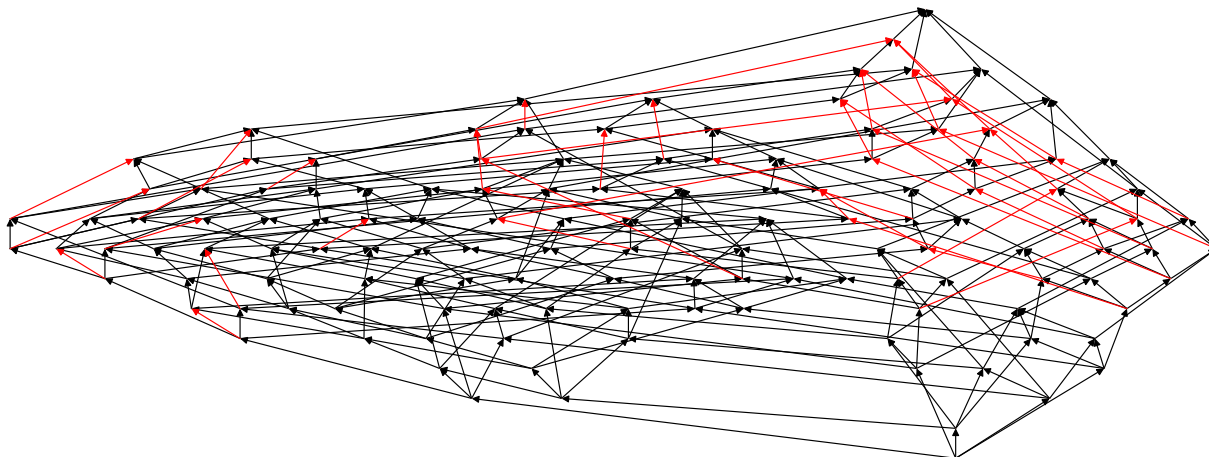


FIGURE 7. Hasse diagram of $\mathcal{S}_6 = (\mathcal{D}_6, \leq)$ (black edges). The Tamari lattice $\mathcal{T}_6 = (\mathcal{D}_6, \preceq)$ can be viewed by considering the red edges.

REFERENCES

- [1] C. Banderier, M. Bousquet-Mélou, A. Denise, P. Flajolet, and D. Gouyou-Beauchamps. Generating functions for generating trees. *Discrete Math.*, 246, 2002, 29–55.
- [2] E. Barucci, A. Bernini, L. Ferrari, and M. Poneti. A distributive lattice structure connecting Dyck paths, noncrossing partitions and 312-avoiding permutations. *Order*, 22(4) (2005), 311–328.
- [3] J.-L. Baril, and J.-M. Pallo. The Phagocyte Lattice of Dyck Words. *Order*, 23(2-3) (2006), 97–107.
- [4] J.-L. Baril, and J.-M. Pallo. The pruning-grafting lattice of binary trees. *Theoretical Computer Science*, 409(3) (2008), 382–293.
- [5] J.-L. Baril, and J.-M. Pallo. A Motzkin filter in the Tamari lattice. *Discrete mathematics*, 338 (2015), 1370–1378.
- [6] O. Bernardi, and N. Bonichon. Intervals in Catalan lattices and realizers of triangulations. *Journal of Combinatorial Theory, Series A*, 116 (2009), 55–75.
- [7] F. Bergeron, and L.-F. Préville-Ratelle. Higher trivariate diagonal harmonics via generalized Tamari posets. *J. comb.* 3(3) (2012), 317–341.
- [8] M. Bousquet-Mélou, E. Fusy, and L.-F. Préville-Ratelle. The number of intervals in the m -Tamari lattices. *Electron. J. Combin.*, 18(2) (2011), paper 31.
- [9] M. Bousquet-Mélou, and F. Chapoton. Intervals in the greedy Tamari posets. *ArXiv*, arXiv:2303.18077.
- [10] F. Chapoton. Sur le nombre d’intervalles dans les treillis de Tamari. *Sém. Lothar. Combin.*, 55 (2006), Art. B55f (electronic).
- [11] F. Chapoton. Some properties of a new partial order on Dyck paths. *Algebraic combinatorics*, 3(2) (2020), 433–463.
- [12] W. Fang, and L.-F. Préville-Ratelle. The enumeration of generalized Tamari intervals. *European J. Combin.*, 61 (2017), 69–84.
- [13] P. Flajolet, and R. Sedgewick. *Analytic Combinatorics*, Cambridge University Press, Cambridge, 2009.
- [14] I. Geffner, and M. Noy. Counting outerplanar maps. *The Electronic Journal of Combinatorics*, 24(2) (2017), P2.3.

- [15] W. Geyer. On Tamari lattices. *Disc. Math.*, 133 (1994), 99–122.
- [16] G. Grätzer. *General Lattice Theory*, Second edition, Birkhäuser, 1998.
- [17] S. Huang, and D. Tamari. Problems of associativity: A simple proof for the lattice property of systems ordered by a semi-associative law. *J. Combinatorial Theory Ser. A*, 13 (1972), 7–13.
- [18] Knuth, D. E. *The Art of Computer Programming, Volume 3: Sorting and Searching* 2nd ed., Addison-Wesley, 1998.
- [19] F. Müller-Hoissen, J.M. Pallo, and J. Stasheff (eds), Associahedra, Tamari Lattices and Related Structures, Progress in Mathematics vol. 299, Birkhäuser, 2012.
- [20] J.M. Pallo. An algorithm to compute the Möbius function of the rotation lattice of binary trees. *RAIRO Informatique Théorique et Applications*, 27(1993), 341–348.
- [21] J.M. Pallo, and R. Racca. A note on generating binary trees in A-order and B-order. *Intern. J. Computer Math.*, 18 (1985), 27–39.
- [22] J.M. Pallo. Right-arm rotation distance between binary trees. *Inf. Process. Lett.*, 87(4) (2003), 173–177.
- [23] H. Prodinger. The kernel method: a collection of examples. *Sém. Lothar. Combin*, 50 (2004), Paper B50f.
- [24] R. Simion, and D. Ullman. On the structure of lattice of noncrossing partitions. *Discrete Math.*, 98 (1991), 193–206.
- [25] D.D. Sleator, R.E. Tarjan, and W.P. Thurston. Rotation distance, triangulations and hyperbolic geometry. *J. Am. Math. Soc.*, 1 (1988), 647–681.
- [26] N. J. A. Sloane. The On-line Encyclopedia of Integer Sequences, available electronically at <http://oeis.org>.
- [27] R. P. Stanley. *Enumerative Combinatorics, Vol. 2*, Cambridge University Press, 1999.
- [28] R. P. Stanley. Supersolvable lattices. *Alg. Univ.*, 2 (1972), 197–217.
- [29] D. Tamari. The algebra of bracketings and their enumeration. *Nieuw Archief voor Wiskunde*, 10 (1962), 131–146.