SJÄLVSTÄNDIGA ARBETEN I MATEMATIK

MATEMATISKA INSTITUTIONEN, STOCKHOLMS UNIVERSITET

Three Families Indexed by Catalan Objects

av

Martin Willert

2020 - No K29

Three Families Indexed by Catalan Objects

Martin Willert

Självständigt arbete i matematik 15 högskolepoäng, grundnivå

Handledare: Per Alexandersson

2020

Within the field of algebraic combinatorics, there has been some recent interest in the bijective correspondence between families indexed by Catalan objects. In this thesis, we look at three such families: acyclic orientations, rook placements and perfect matchings. We search for bijections between these three families, and we use q-analogs to refine our results. We are particularly interested in bijections which behave well with respect to such q-analogs. We find that these families are in bijective correspondence, and we propose using this correspondence to assist future advances in some well-known open problems of the field, for example the Stanley–Stembridge Conjecture.

Contents

1 Introduction

The Catalan numbers, $C_0, C_1, C_2, C_3, \ldots$, form the infinite sequence 1, 1, 2, 5, 14, 42, 132, 429, 1430, ..., and within the field of combinatorics, this well-known sequence seem to appear just about anywhere and everywhere [Sta15]. Broadly speaking, Catalan numbers count the number of elements in the sets of various families, including Dyck paths, binary trees, ballot sequences, perfect matchings and many more, see Fig. 1. These Catalan families are commonly referred to as *Catalan objects*, for the sake of simplicity. Readers unfamiliar with these objects need not to worry, as they will be properly defined in Chapter 2.3.

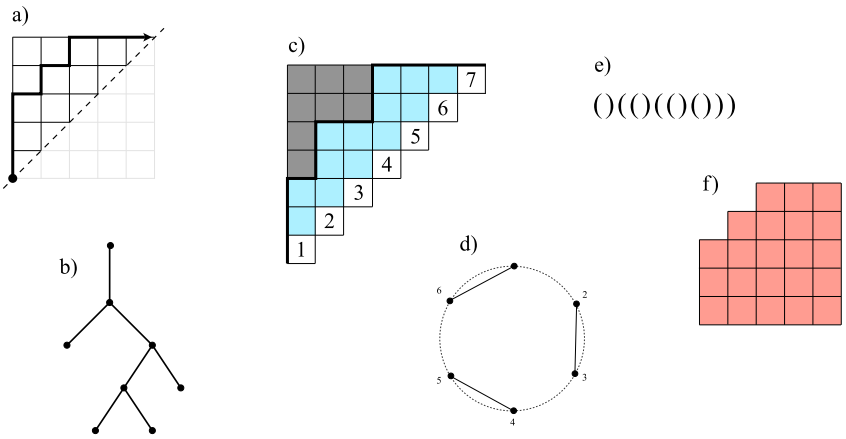


Figure 1: Examples of objects counted by the Catalan numbers. a) Dyck paths, b) binary trees, c) area sequences and unit interval graphs, d) perfect matchings (non-crossing), e) valid parenthesis words, f) Ferrers boards.

In mathematics, a *family* is generally understood as a "set of sets", and a *Catalan family*, $\mathbf{Cat} = (Cat_1, Cat_2, Cat_3, \ldots)$, is defined as a family, such that Cat_i has cardinality C_i, where C_i the i:th Catalan number. For example, let the Catalan family $\mathbf{Dyck} = (Dyck_1, Dyck_2, Dyck_3, \ldots)$ be the family of Dyck paths of size $(1, 2, 3, \ldots)$. It follows that each set $Dyck_i$ is restricted to Dyck paths of size i, and that the cardinality of $Dyck_i$ is C_i, see Fig. 2. Note that the *cardinality* of a set is just the number of elements in the set.

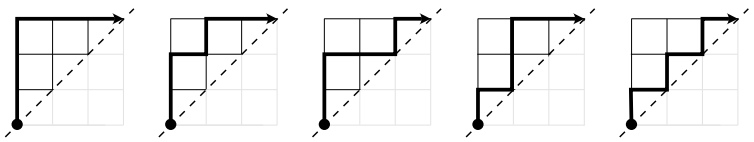


Figure 2: The $C_3 = 5$ Dyck paths of $Dyck_3$, i.e. the set of Dyck paths from $(0, 0)$ to $(3, 3)$.

Since the cardinality of *every* Catalan set Cat_i is C_i for all $i \geq 0$, all Catalan objects have a one-to-one correspondence on element-, set- and family level. In

other words, Catalan objects are in *bijective correspondence*, see Fig. 3. We use the terms *corresponding elements*, *corresponding sets* and *corresponding families* whenever we wish to describe these bijective relations.

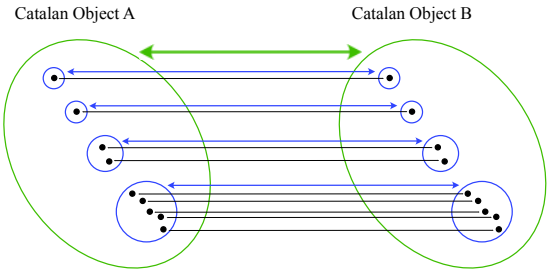


Figure 3: One-to-one correspondence on element level (black), set level (blue), and family level (green), i.e. a bijection.

The described bijective correspondence is sometimes easy to perceive, and at other times a bit harder to grasp using pure intuition, see Fig. 4. Such bijections have been thoroughly examined, for example in [Sta15].

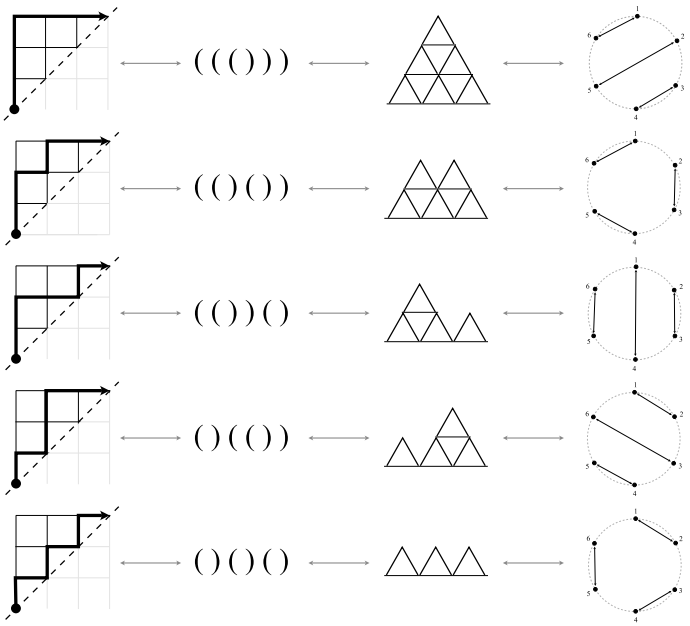


Figure 4: Bijections between the C_3 Dyck paths from $(0,0)$ to $(3,3)$, the C_3 valid parenthesis words of length 6, the C_3 ways to build a classic house of cards with 3 units in the bottom row, and the C_3 non-crossing perfect matchings of 6 points on a circle.

In this thesis, we are not primarily interested in Catalan objects per se, but rather certain families *indexed* by Catalan objects, that is $\{Fam_D : D \in Cat_i \text{ for some } i\}$, for which we introduce the name *C-Indexed families* or *CIF:s*. More specifically, we will look at the following C-Indexed families:

- **Acyclic Orientations (AO)**, indexed by the Catalan object *unit interval graphs* (UIG).
- **Rook Placements (RP)**, indexed by the Catalan object *Ferrers boards* (FB).
- **Perfect Matchings (PM)**, indexed by the Catalan object *perfect matching starting points* (PMSP).

Just as with Catalan objects, these families are in bijective correspondence, although it is not intuitively apparent. Some of this correspondence have been studied earlier, for example in [AP18]. However, to our knowledge, there is no publication which summarises these observations. In this thesis, our main interest is to illuminate certain bijections between the three families listed above.

To be more specific, in Chapter 4 we will look at the following: Consider the Catalan objects unit interval graphs, Ferrers boards and perfect matching starting points, and let $\text{UIG}_{\mathbf{a}}$, $\text{FB}_{\mathbf{a}}$, and $\text{PMSP}_{\mathbf{a}}$ be three *corresponding* elements in these Catalan families, see Fig. 5. For now, we simply state that these three elements are corresponding, since they all correspond to the same area sequence $\mathbf{a} = (a_1, a_2, a_3, \ldots, a_n)$, defined in Chapter 2.3. However, we will define this correspondence properly in Chapter 4.

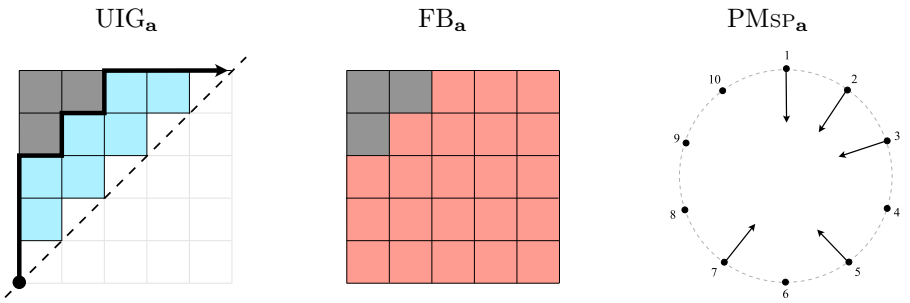


Figure 5: Three *corresponding* Catalan elements.

These three corresponding elements index the set of Acyclic Orientations $\text{AO}(\text{UIG}_{\mathbf{a}})$, rook placements $\text{RP}(\text{FB}_{\mathbf{a}})$, and perfect matchings $\text{PM}(\text{PMSP}_{\mathbf{a}})$. Since $\text{UIG}_{\mathbf{a}}$, $\text{FB}_{\mathbf{a}}$ and $\text{PMSP}_{\mathbf{a}}$ are *corresponding* Catalan elements, we prefer to simplify the notation, and so we write $\text{AO}(\mathbf{a})$, $\text{RP}(\mathbf{a})$ and $\text{PM}(\mathbf{a})$. Figure 6 shows one arbitrary element of each such set, i.e. one acyclic orientation $\vartheta \in \text{AO}(\mathbf{a})$, one rook placement $\Psi \in \text{RP}(\mathbf{a})$ and one perfect matching $\Omega \in \text{PM}(\mathbf{a})$.

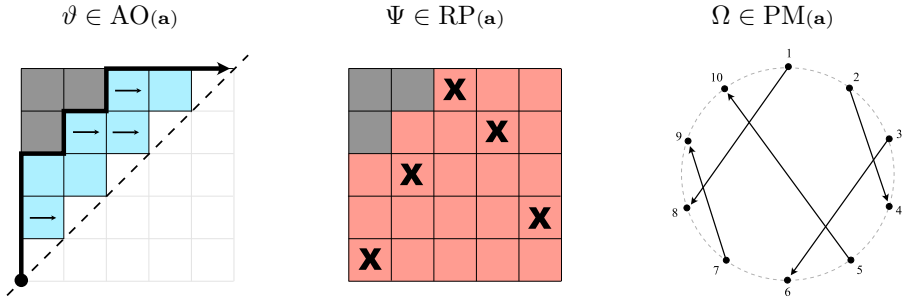


Figure 6: One acyclic orientation $\vartheta \in \mathrm{AO}(\mathbf{a})$, one rook placement $\Psi \in \mathrm{RP}(\mathbf{a})$ and one perfect matching $\Omega \in \mathrm{PM}(\mathbf{a})$.

Notably, we find that the cardinality of these sets coincide for every area sequence $\mathbf{a}$, that is

$$|\mathrm{AO}(\mathbf{a})| = |\mathrm{RP}(\mathbf{a})| = |\mathrm{PM}(\mathbf{a})| = \prod_{i=1}^{n}(a_i + 1)\,. \tag{1}$$

This implies that these families are in bijective correspondence. However, although (1) is a noteworthy result, we are interested in refining this result, by not only considering the cardinality of the sets, but also certain *statistics* within the sets.

Each single orientation $\vartheta \in \mathrm{AO}(\mathbf{a})$ has some number of ascending edges (asc). We can *refine* (or simply sort) our set $\mathrm{AO}(\mathbf{a})$ using the number of ascending edges in each orientation ϑ as statistics, so that

$$\mathrm{AO}(\mathbf{a}) = \{\vartheta : asc(\vartheta) = 0\} \cup \{\vartheta : asc(\vartheta) = 1\} \cup \{\vartheta : asc(\vartheta) = 2\} \dots\,.$$

Likewise, we can refine $\mathrm{RP}(\mathbf{a})$, using the number of inversions (inv) in each rook placement Ψ as statistics, so that

$$\mathrm{RP}(\mathbf{a}) = \{\Psi : inv(\Psi) = 0\} \cup \{\Psi : inv(\Psi) = 1\} \cup \{\Psi : inv(\Psi) = 2\} \dots\,,$$

and we can refine $\mathrm{PM}(\mathbf{a})$, using the number of inversions crossings (cr) in each perfect matching Ω, as statistics, so that

$$\mathrm{PM}(\mathbf{a}) = \{\Omega : cr(\Omega) = 0\} \cup \{\Omega : cr(\Omega) = 1\} \cup \{\Omega : cr(\Omega) = 2\} \dots\,.$$

As we will see in Chapter 3, q-analogs are highly suitable for describing such refinements. We define our q-analogs

$$|\mathrm{AO}(\mathbf{a})|_q = \sum_{\vartheta \in \mathrm{AO}(\mathbf{a})} q^{asc(\vartheta)}\,, \tag{2}$$

$$|\mathrm{RP}(\mathbf{a})|_q = \sum_{\Psi \in \mathrm{RP}(\mathbf{a})} q^{inv(\Psi)}\,, \tag{3}$$

$$|\mathrm{PM}(\mathbf{a})|_q = \sum_{\Omega \in \mathrm{PM}(\mathbf{a})} q^{cr(\Omega)}\,. \tag{4}$$

In (2), q^0 means acyclic orientations $\vartheta \in \mathrm{AO}(\mathbf{a})$ with 0 ascending edges, q^1 acyclic orientations $\vartheta \in \mathrm{AO}(\mathbf{a})$ with 1 ascending edge, and so on. The fundamentals of q-analogs and q-counting are explained in Chapter 3, for readers not familiar with these concepts.

Remarkably, we find that (2), (3) and (4) coincide, so that

$$|\mathrm{AO}(\mathbf{a})|_q = |\mathrm{RP}(\mathbf{a})|_q = |\mathrm{PM}(\mathbf{a})|_q = \prod_{i=1}^{n} [a_i + 1]_q\,. \tag{5}$$

This result refines the statement in (1), and it provides a strong indication of a more general bijective correspondence between these three C-Indexed families, see Fig. 7.

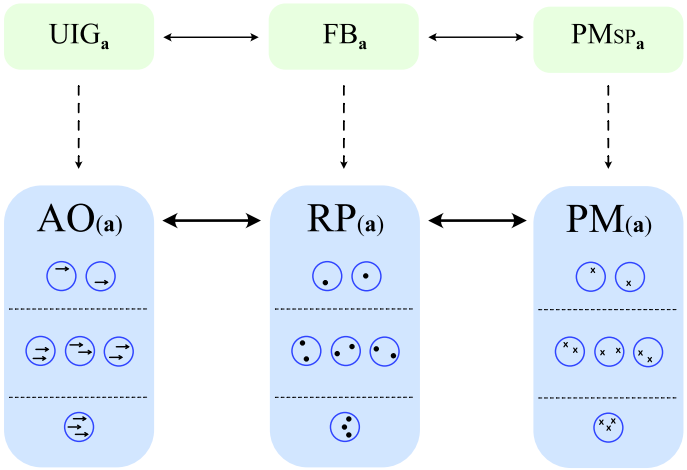


Figure 7: Schematic illustration of the bijections discussed in this thesis.

These findings are beautiful as they are, but they might also be useful for solving some well-known open problems within the field of algebraic combinatorics, for example the *Stanley–Stembridge Conjecture* [SS93], [Sta95]. The core of this problem relates to acyclic orientations. However, in the light of (5), it might be more fruitful to consider a corresponding family instead. This will be further expanded on in Chapter 5.

This thesis provides an overview of the bijections outlined above, with particular focus on the construction of transparent proofs and explanation models. A background on the properties of Catalan numbers is provided in Chapter 2, and q-analogs are explained in Chapter 3.

2 Introduction to Catalan Numbers

In order to understand the bijections between C-Indexed families, we first need to review the fundamental properties of Catalan numbers. Readers already familiar with these properties may browse this chapter.

2.1 Definition

The most historically important definition of Catalan numbers comes from counting triangulations of polygons [Sta15]. It is easy to show that a regular polygon with n edges, P_n, always has $(n-3)$ non-intersecting diagonals. The polygon P_n is said to be *triangulated* when a complete set of such non-intersecting diagonals are drawn within the polygon, thus dividing the polygon into $(n-2)$ smaller triangles, see Fig. 8.

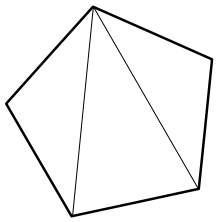


Figure 8: One possible triangulation of P_5 into 3 triangles, using 2 non-intersecting diagonals.

Consequently, the polygon P_{n+2} is triangulated into n triangles, using $(n-1)$ diagonals, see Fig. 9. In polygons with more than three sides, i.e. when $n > 1$, the triangulation could be achieved in more than one way. Define *the n:th Catalan number (C_n)* as *the number of ways to triangulate a regular polygon with* $(n+2)$ *edges.*

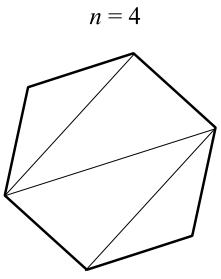


Figure 9: One possible triangulation of P_{n+2} into n triangles, using $n-1$ non-intersecting diagonals. The number of ways to triangulate P_{n+2} is given by the n:th Catalan number (C_n). Hence, P_6 can be triangulated in C_4 ways.

Even though there is no obvious interpretation of how to triangulate a 2-gon, C_0 is set to 1, see Fig. 10.

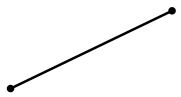


Figure 10: "Triangulation" of P_2. C_0 is set to 1.

2.2 Recurrence Relation

The Catalan numbers, $C_0, C_1, C_2, C_3, \ldots$, form the infinite sequence 1, 1, 2, 5, 14, 42, 132, 429, 1430, 4862, 16796, ..., which satisfies the recurrence explained below.

As stated above, C_0 is set to 1, and since a polygon with three sides (P_3) can be triangulated in exactly 1 way, $C_1 = 1$, see Fig. 11.

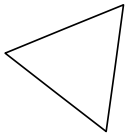


Figure 11: $C_1 = 1$.

As we add more sides to a polygon, it is helpful to think of the regular polygon P_{n+2} as a combination of two smaller polygons, Q_a and Q_b, by removal of some edge e, see Fig. 12. Note that $a, b \geq 2$ and that, since Q_a and Q_b have one mutual vertex, $a + b = n + 3$. This means that if Q_a has 2 edges, Q_b has $n + 1$ edges, and if Q_a has 3 edges, Q_b has n edges, and so on. Since Q_a is a polygon, it can be *internally* triangulated in C_{a-2} ways and, naturally, this is the case for Q_b as well. Hence, according to the multiplication principle, each *combination* of Q_a and Q_b can be triangulated in $C_{a-2} \cdot C_{b-2}$ ways, and the number of ways to triangulate the larger polygon P_{n+2} is equal to the sum of all of these triangulated combinations of Q_a and Q_b.

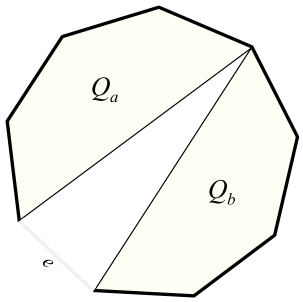


Figure 12: A regular polygon as a combination of two smaller polygons, Q_a and Q_b.

Although it seems like we have found the principle for the recurrence already, we choose to proceed slowly, building our case step by step. We thus continue with P_4. Starting at one of the vertices connected to e, P_4 is either a combination of first a 2-gon and then a 3-gon, or first a 3-gon and then a 2-gon, both alternatives shown in Fig. 13. The number of ways to triangulate the larger polygon P_4 is equal to the sum of all such triangulated combinations of Q_a and Q_b. Thus, $C_2 = C_0C_1 + C_1C_0 = 1 \cdot 1 + 1 \cdot 1 = 2$.

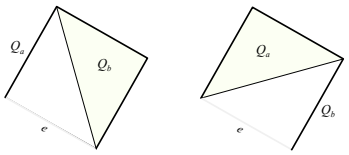


Figure 13: $C_2 = 2$.

Similarly, P_5 is either a combination of first a 2-gon and then a 4-gon, two 3-gons, or first a 4-gon and then a 2-gon, see Fig. 14. As we have seen, each 4-gon can be triangulated in $C_2 = 2$ ways. Thus, $C_3 = C_0C_2 + C_1C_1 + C_2C_0 = 1 \cdot 2 + 1 \cdot 1 + 2 \cdot 1 = 5$.

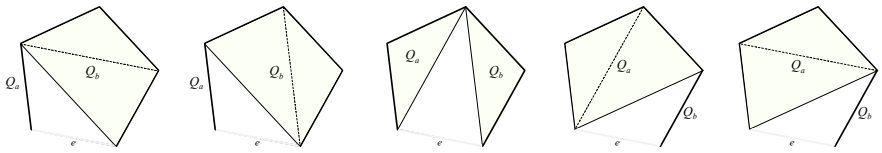


Figure 14: $C_3 = 5$.

The same argument can be used for any $n \geq 0$, and we see that C_{n+1} is always the sum of such mutually dependent combinations of C_0 through C_n. Hence,

$$C_{n+1} = C_0C_n + C_1C_{n-1} + \cdots + C_nC_0 = \sum_{k=0}^{n} C_k\, C_{n-k}\,. \tag{6}$$

2.3 Some Famous Catalan Objects

Although the family of triangulated polygons gives important historical background to Catalan numbers, it is not the most studied object. In fact, many other Catalan objects reveal the properties of Catalan numbers in a more intuitive and useful manner. Some of the most well-known objects are introduced in this section. For a more comprehensive list, see [Sta15].

2.3.1 Valid Parenthesis Words

A *valid parenthesis word* is a sequence of $2n$ parenthesis signs, "(" or ")", arranged in a valid order, according to the rules of parenthesis pairing. For example, (())() is one of the valid words of length 6, whereas ())()(is not.

Proposition 1. *The set of valid parenthesis words of length* $2n$ *has cardinality* C_n.

Proof. We note that any valid word must consist of an equal number of left and right signs, and thus we think of the word as consisting of n *pairs* of left/right signs. It is also evident that any such pair must start with a left parenthesis sign. Let A and B be sets of such valid pairs, and note that A and B can be empty. We realise that any valid word with $n \geq 1$ could be written on the form $(A)B$, i.e. where set A is locked within a fixed pair f. Since f is fixed, $|A|+|B| = n-1$, so if set A contains k pairs, set B contains $n-k-1$ pairs.

Let C'_n be the number of valid pairings of n parenthesis pairs. It is obvious that $C'_0 = 1$, since there is exactly 1 way to arrange an empty set. A word with 1 pair (C'_1) uses only the fixed pair f with two empty sets A and B. We have just learned that any empty set could be *internally arranged* in $C'_0 = 1$ ways, and so the number of valid pairings using two empty sets equals $C'_0 \cdot C'_0 = 1 \cdot 1 = 1$, according to the multiplication principle.

In a word with 2 pairs, we use f and one more pair, which could be placed either in set A *or* B. We know that if A contains k pairs, set B contains $n-k-1$ pairs, so if we choose B, we have that $k = 0$ and $n-k-1 = 1$. We have just learned that the number of ways to *internally arrange* a set with 1 pair is $C'_1 = 1$, and so the number of valid pairings when choosing B is $C'_0 \cdot C'_1 = 1 \cdot 1 = 1$, according to the multiplication principle. If we choose A instead, $k = 1$ and and $n-k-1 = 0$ and we have $C'_1 \cdot C'_0 = 1 \cdot 1 = 1$. Thus, the *total* number of ways to arrange 2 pairs of parenthesis is $C'_0 \cdot C'_1 + C'_1 \cdot C_0 = 2$.

In a word with 3 pairs (C'_3), we use f and two more pairs. We can choose to place both in A, both in B, or 1 in each set. Whenever 2 in a set, these can be *internally arranged* in C'_2 ways. Thus, $C'_3 = C'_0 \cdot C'_2 + C'_1 \cdot C'_1 + C'_2 \cdot C'_0 = 2 + 1 + 2 = 5$. The same argument can be used for words with any n number of pairs, for $n \geq 0$. We see that C'_{n+1} is always the sum of such mutually dependent combinations of C'_0 through C'_n. Consequently, the general recurrence is

$$C'_{n+1} = C'_0 C'_n + C'_1 C'_{n-1} + \cdots + C'_n C'_0 = \sum_{k=0}^{n} C'_k \, C'_{n-k} \,. \tag{7}$$

We see that C'_n satisfies the same initial condition and recurrence as C_n, for all $n \geq 0$, see (6). Hence, valid parenthesis words is a Catalan object. □

2.3.2 Dyck Paths

Consider a $n \times n$ grid, and define a *monotonic path* from $(0,0)$ to (n,n) as a path such that if one horizontal step has direction $(x,0)$, *all* horizontal steps have direction $(x,0)$, and if one vertical step has direction $(0,y)$, *all* vertical steps have direction $(0,y)$. In this thesis, we consistently use monotonic paths with the allowed steps $(1,0)$ and $(0,1)$, i.e. paths that "travel north-east". Define a *Dyck path* as a monotonic path from $(0,0)$ to (n,n), such that all steps of the path are above the main diagonal from $(0,0)$ to (n,n). A Dyck path may *touch* the main diagonal, but not surpass it at any time, see Fig. 15.

For practical reasons, monotonic paths are sometimes described as binary words, where up-steps are represented with 1, and right-steps with 0. Note that this is just a question of notation, and we do not need to distinguish between a Dyck path and its corresponding binary word.

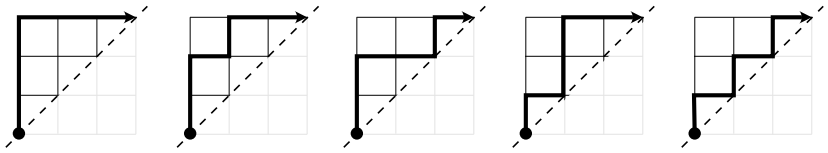


Figure 15: The C_3 Dyck paths from (0,0) to (3,3), analogue to the C_3 binary words $(1,1,1,0,0,0)$, $(1,1,0,1,0,0)$, $(1,1,0,0,1,0)$, $(1,0,1,1,0,0)$ and $(1,0,1,0,1,0)$.

Proposition 2. *The set of Dyck paths from* $(0,0)$ *to* (n,n) *has cardinality* C_n.

Proof. There is an obvious bijection between valid parenthesis words and Dyck paths. In fact, we can picture a Dyck path as just a graphical illustration of such a word, where a left sign is represented with an up-step, a right sign with a right-step. The condition that the pairs are *balanced*, and must start with a left sign translates to the restriction of all steps of the path being above the main diagonal. □

2.3.3 Area Sequences

Consider a Dyck path D and define the *area sequence*, $\mathbf{a}(D) = (a_1, a_2, a_3, \ldots, a_n)$ as the number of intact grid units between D and the main diagonal at row $1, 2, 3, \ldots, n$, see Fig. 16. We say that a_i is the *row area* at row i. It is obvious that each Dyck path defines exactly 1 area sequence, and it follows that the set of area sequences with n rows has cardinality C_n.

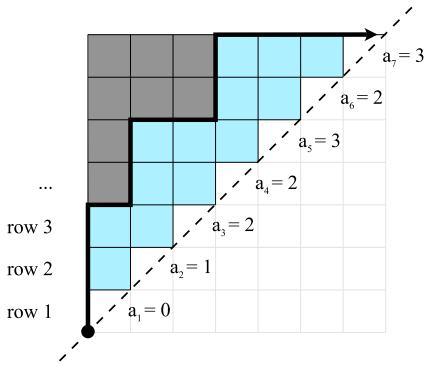


Figure 16: A Dyck path D from $(0,0)$ to $(7,7)$, and its associated area sequence $\mathbf{a}(D) = (0,1,2,2,3,2,3)$ in blue.

2.3.4 Non-crossing Perfect Matchings

A *perfect matching* is a pairing of of $2n$ elements into n pairs, visualised by the drawing of n connecting chords between $2n$ points on a circle. Define a *non-crossing perfect matching* as a perfect matching, such that there are no crossings of the chords. We can picture the set of non-crossing perfect matchings of $2n$ points as the possibilities for $2n$ persons, spread out around a table, to simultaneously shake hands in pairs of two, without any crossings of arms, see Fig. 17.

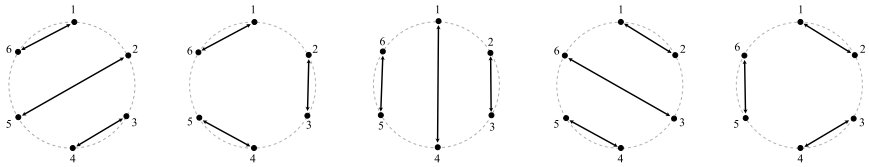


Figure 17: The C_3 non-crossing perfect matchings of 6 points on a circle.

Proposition 3. *The set of non-crossing perfect matchings of* $2n$ *points on a circle has cardinality* C_n*.*

Proof. Consider a circle with $2n$ points, and fix any of these points as our starting point p. We construct our first pair by connecting point p to some other point m, using a connecting chord. This divides the remaining points into two subsets, A and B, see Fig. 18. In order not to break the rules of non-crossing perfect matchings, m has to be chosen so that there is an *even* number of points in set A (and thus in set B), including the option of an empty set. The points of each of the subsets A and B can now be *internally* paired, such that there are no crossings of the connecting chords. It is clear that, for each acceptable point m, if A contains k pairs of points, B contains $n-k-1$ pairs of points. It is also clear that, for each acceptable point m, the number of perfect matchings of the original

$2n$ points is the product of the number of possibilities to *internally* pair the points of A and B. By moving point m to all acceptable positions, all possible combinations are covered, and we see that the *total* number of perfect matchings of the original $2n$ points is the sum of the products from each acceptable point m.

Let C'_n be the number of non-crossing perfect matchings of $2n$ points on a circle. There is exactly one way to connect the points of an empty set, so $C'_0 = 1$. When $n = 1$, A and B are both empty, and so $C'_1 = C'_0 \cdot C'_0 = 1$, see Fig. 18, top. As we add one more *pair* of points (i.e. when $n = 2$) and move m around the circle, this added pair becomes a part of either A *or* B, and thus $C'_2 = C'_0 \cdot C'_1 + C'_1 \cdot C'_0 = 2$, see Fig. 18, middle. Consequently, when $n = 3$ and m moves around the circle, A changes from 0 to 1 to 2 while B changes from 2 to 1 to 0. We have just learned that a set of 2 pairs has C'_2 matchings, so $C'_3 = C'_0 \cdot C'_2 + C'_1 \cdot C'_1 + C'_2 \cdot C'_0 = 5$, see Fig. 18, bottom. The same argument can be used for perfect matchings with any n number of matchings, for $n \geq 0$. We see that C'_{n+1} is always the sum of such mutually dependent combinations of C'_0 through C'_n. Consequently, the general recurrence

$$C'_{n+1} = C'_0 C'_n + C'_1 C'_{n-1} + \cdots + C'_n C'_0 = \sum_{k=0}^{n} C'_k \, C'_{n-k} \tag{8}$$

follows. We see that C'_n satisfies the same initial condition and recurrence as C_n, for all $n \geq 0$, see (6). Hence, non-crossing perfect matchings is a Catalan object.

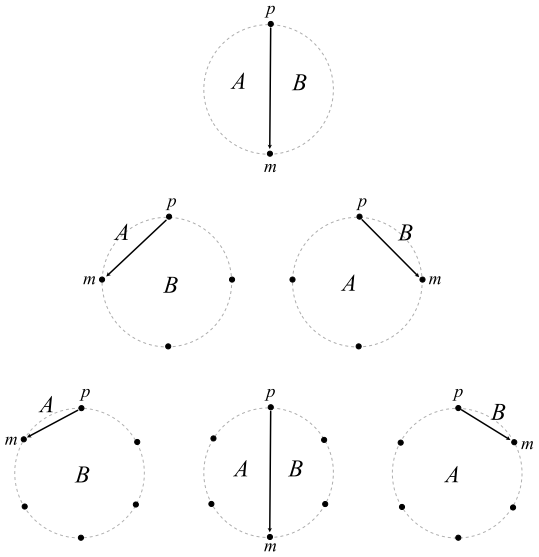


Figure 18: A step-by-step movement of point m generates all possible combinations of the subsets A and B, for C'_1 (top), C'_2 (middle) and C'_3 (bottom).

□

2.4 Explicit Formula

The *explicit formula for the n:th Catalan number* is

$$C_n = \frac{1}{n+1}\binom{2n}{n}, \tag{9}$$

where $\binom{2n}{n}$ is a binomial number.

This equals $\frac{1}{n+1}\,\frac{(2n)!}{(n!)^2}$, and since $(n+1)\,n! = (n+1)!$, the n:th Catalan number is sometimes defined as

$$C_n = \frac{(2n)!}{(n+1)!\,n!}, \tag{10}$$

or

$$C_n = \frac{1}{2n+1}\binom{2n+1}{n}, \tag{11}$$

via the equivalence $\frac{(2n)!}{(n+1)!\,n!} = \frac{1}{2n+1}\,\frac{(2n+1)!}{(n+1)!\,n!}$. All three forms (9), (10) and (11) are common in the literature.

Proof 1. This widespread proof, well explained in [Dav10] for example, uses Dyck paths and a somewhat developed version of *André's reflection method* [And87], which is a clever method for identifying *bad paths* from $(0,0)$ to (n,n), i.e. monotonic paths from $(0,0)$ to (n,n) that are *not* Dyck paths. Note that, by convention, this proof defines Dyck paths as monotonic paths from $(0,0)$ to (n,n), such that all steps of the path are *below* the main diagonal from $(0,0)$ to (n,n). It should be clear that this is just a mirroring of the same property.

The number of monotonic paths from $(0,0)$ to (n,n) equals the number of ways to position the n up-steps in the $2n$ string of steps, which is $\binom{2n}{n}$. Any path entering the forbidden area above the diagonal at any time is considered a bad path. Note that a path might illegally surpass the diagonal already at the first step. As soon as a bad path P has surpassed the diagonal for the first time, we *reflect* the remaining steps so that up-steps become right-steps and vice versa, see Fig. 19.

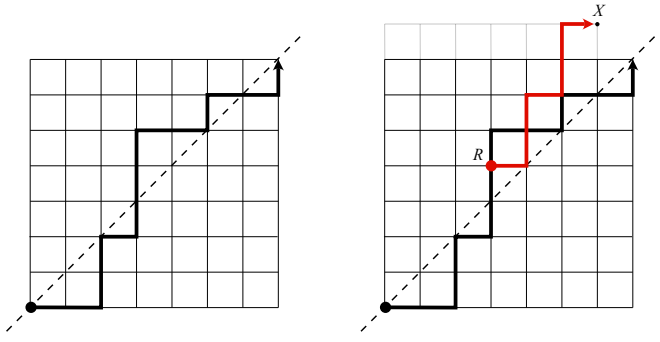


Figure 19: Left: The Original path P. Right: The new path P_r, reflected in point R.

When a path P has surpassed the diagonal it has used one more up-step than right-step by definition, and since P in the end has to use n steps of each, the remaining part of P (after R) contains one more right-step. When this part is reflected, it will instead contain one more up-step, and thus end in *point* X, defined as $(n-1,\ n+1)$. Note that *all* reflected paths end in point X, since every bad path must cross the diagonal by definition. Also note that each bad path is reflected into one *distinct* path ending in point X. Hence, the number of paths ending in point X must be equal to, or greater than, the number of bad paths (since there might be *other* paths ending in point X as well). Conversely, since the reflection process is reversible, *any* monotonic path from $(0,0)$ to point X could be reflected into a bad path from $(0,0)$ to (n,n). So, using the same argument, the number of bad paths must be equal to, or greater than, the number of paths ending in point X. However, we have already seen that $|\text{paths } (0,0) \to \text{point } X| \geq |\text{bad paths } (0,0) \to (n,n)|$. Hence $|\text{bad paths } (0,0) \to (n,n)| = |\text{paths } (0,0) \to \text{point } X|$.

The number of bad paths is thus equal to the total number of paths from $(0,0)$ to point X, which is $\binom{n-1+n+1}{n+1} = \binom{2n}{n+1} = \binom{2n}{n-1}$. The number of Dyck paths from $(0,0)$ to (n,n) could now be calculated by subtracting the number of bad paths from the total number of paths from $(0,0)$ to (n,n), i.e.

$$\begin{aligned}
\binom{2n}{n} - \binom{2n}{n+1} &= \binom{2n}{n} - \frac{1}{n+1} \cdot \frac{(2n)!}{n!(n-1)!} \\
&= \frac{(2n)!}{n!n!} - \frac{n}{n+1} \cdot \frac{(2n)!}{n!n!} \\
&= \frac{n+1}{n+1} \cdot \frac{(2n)!}{n!n!} - \frac{n}{n+1} \cdot \frac{(2n)!}{n!n!} \\
&= \frac{1}{n+1} \binom{2n}{n}, \qquad (12)
\end{aligned}$$

which is identical to the explicit formula for Catalan numbers, see (9). □

Proof 2 (Chung-Feller/Chen). Another, perhaps more interesting, proof provides a better intuitive understanding of the $\frac{1}{n+1}$ part of the formula. The proof was constructed by Young-Ming Chen in 2008 [Che08], with the purpose of making a more direct and bijective proof of the famous *Chung-Feller Theorem* from 1949 [CWF49], which was analytically proved.

As in proof 1, we consider the $\binom{2n}{n}$ monotonic paths from $(0,0)$ to (n,n), where Dyck paths are defined as monotonic paths such that all steps of the path are *below* the main diagonal from $(0,0)$ to (n,n). Each monotonic path can *exceed* the diagonal by $0, 1, 2, \ldots, n$ vertical steps *in total* (which obviously equals the number of exceeding horizontal steps as well). Thus, there are $n+1$ distinct *exceedance classes*, which we refer to as $\beta_0, \beta_1, \beta_2, \ldots, \beta_n$, see Fig. 20.

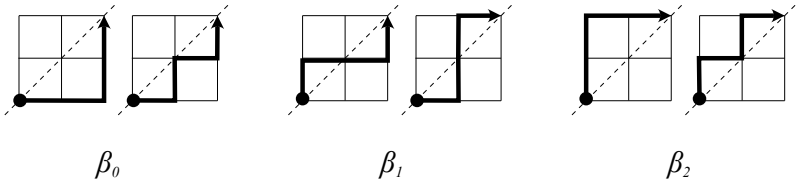


Figure 20: All monotonic paths from $(0,0)$ to $(2,2)$, sorted into the exceedance classes β_0, β_1 and β_2.

A Dyck path belongs to exceedance class β_0 by definition, and class $\beta_1, \beta_2 \ldots \beta_n$ contain all the bad paths. If we can prove that there is an equal number of paths in each class, i.e. that $|\beta_0| = |\beta_1| = |\beta_2| = \cdots = |\beta_n|$, formula (9) follows immediately, and we are done. Luckily, this statement is exactly the Chung-Feller Theorem, which Young-Ming Chen provides a bijective proof of.

The central thought of the Young-Ming Chen proof is this: Let us say we find an algorithm that transforms paths of exceedance class k (with k "flaws") to paths of exceedance class $k+1$, and we can show that this algorithm is *bijective*. Then all paths of exceedance class k are *potential* paths of exceedance class $k+1$ and vice versa. Consequently, all paths of exceedance class $k+1$ are potential paths of exceedance class $k+2$ and vice versa, and so on up to exceedance class n and down to exceedance class 0. This is only possible if there is an equal number of paths in each exceedance class.

Here is the algorithm from exceedance class k to $k-1$:

1. Start at $(0,0)$ and note when the path surpasses the diagonal for the first time.
2. As the path *touches* the diagonal *again*, denote the last horizontal step X. Also let A be the portion before X, and B the portion after X.
3. Construct the path $B - X - A$ (Swap A and B).

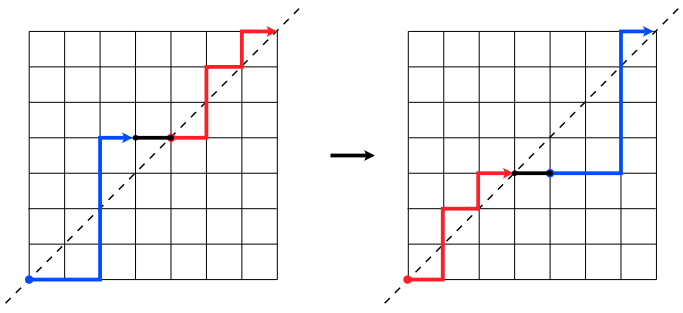


Figure 21: The clever algorithm by Young-Ming Chen.

Note that the algorithm is reversible. We leave it as a fun exercise for the reader to verify the algorithm, and to sort all the monotonic paths from, let us say, $(0,0)$ to $(3,3)$ into the relevant exceedance classes, using the algorithm. □

This proof makes it clear that Catalan numbers actually count the number of monotonic paths in *each* exceedance class β (and Dyck paths is just one of these classes).

3 Using q-analogs

Within the field of mathematics as a whole, q-analogs are used for a variety of purposes. In algebraic combinatorics however, we primarily use q-analogs as an efficient means of providing additional information to a set count with respect to some statistic we are interested in. In general terms, it is helpful to think of such a q-analog as describing a *refinement*, or a *set partition* with respect to this chosen statistic. We define a *q-analog* $|S|_q$ on the elements s in a set S, using $stat(s)$ as statistics, as

$$|S|_q = \sum_{s \in S} q^{stat(s)} \,. \tag{13}$$

Hence, if our statistic describes outcomes between 0 and n, the q-analog is a polynomial $A_1q^0 + A_2q^1 + A_3q^2 + \cdots + A_{n+1}q^n$, providing the refined information that the set S contains A_1 elements s such that $stat(s) = 0$, A_2 elements such that $stat(s) = 1$, and so on. It is easy to see that for $q = 1$, the sum of the polynomial equals the cardinality of the set.

Example 4. Let us say we have a set *Buck* of 7 *distinct* buckets b with some balls in them: 2 buckets have 0 balls, 3 buckets have 1 ball, 1 bucket has 2 balls and 1 bucket has 3 balls, see Fig. 22. Note that we already know that these buckets are distinct, and that the cardinality of *Buck* thus is 7. For some reason we are particularly interested in communicating how many buckets have 0 balls, how many have 1 ball etc. We do this by q-counting *Buck*, using the statistic $nb(b)$, i.e. "number of balls in bucket" . Using (13) we have that

$$|Buck|_q = \sum_{b \in Buck} q^{nb(b)} = 2q^0 + 3q^1 + 1q^2 + 1q^3 = 2 + 3q + q^2 + q^3 \,.$$

We see that the q-analog describes our desired additional information, and that if we set $q = 1$, the q-analog returns the total sum 7.

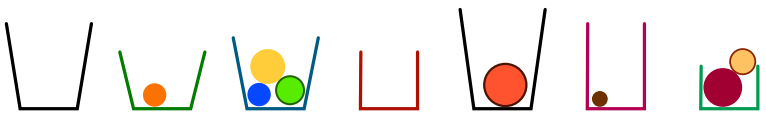


Figure 22: Seven distinct buckets.

The use of q-analogs becomes much more interesting when our statistic of interest is added as a series of independent choices. Note that this time we do *not* know the cardinality of the set (i.e. the number of distinct buckets) nor the number of balls in each bucket beforehand.

Example 5. Let us say we have one empty bucket, and we are given a number of specific choices to fill this bucket with balls. At choice 1 we are given three

alternatives: add $0, 1$ or 2 green balls. At choice 2 we are given two alternatives: add 0 or 1 blue ball, see Fig. 23. How many distinct buckets can we construct? and how is the statistic *number of balls in bucket* distributed within this set of distinct buckets? Well, it is trivial to see that the number of distinct buckets, $|Buck|$, is $3 \cdot 2 = 6$, according to the multiplication principle. However, this does not tell us anything about the number of balls in each bucket.

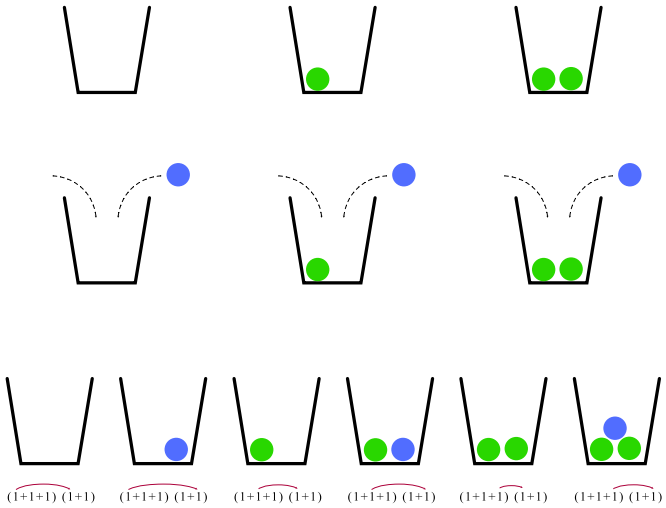


Figure 23: Top: Choice 1. Middle: Choice 2. Bottom: The set of 6 distinct buckets as a result of two independent choices.

The multiplication principle is really just parenthesis multiplication according to the distributive law, and so we could equally well write $|Buck| = (1{+}1{+}1)(1{+}1) = 1+1+1+1+1+1 = 6$. In this way, it is more apparent that, for example, the last "1" in the first parenthesis represents the alternative "2 green balls", and the last "1" in the second parenthesis represents the alternative "1 blue ball". The multiplication of these two alternatives represents the bucket with three balls, which is one of the 6 elements in *Buck*, i.e. one of the 6 distinct buckets.

However, in the parenthesis, "1" means simply "1 alternative", regardless of how many balls this alternative represents, and when we multiply "1" with "1", we get "1". Hence, we still do not know the number of balls in each bucket. This is because we lack the *refined* information held in $nb(b)$. We realise that we can refine our parenthesises by rewriting $(1+1+1+\dots)$ as $(q^0+q^1+q^2+\dots)$, and since exponentials are added under multiplication ($q^m \cdot q^n = q^{m+n}$), this will bring our statistic $nb(b)$ to the count. Hence,

$$\begin{aligned} |Buck|_q = (q^0+q^1+q^2)(q^0+q^1) &= (1+q+q^2)(1+q) \\ &= 1+q+q+q^2+q^2+q^3 \\ &= 1+2q+2q^2+q^3. \end{aligned}$$

We see that the q-analog describes our desired additional information, see Fig. 23, and that if we set $q = 1$, the q-analog returns the total sum 6. Note that the q-analog is a refinement of the statistic *number of balls in bucket*. The information of which ball came from which choice (i.e. the color in our example) is not held in the q-analog.

In general terms, for any integer $k > 0$, we define the q-analog $[k]_q$ of k as $[k]_q = q^0 + q^1 + q^2 + \cdots + q^{k-1}$. It follows that, just as in Example 5, when we have the exact $a + 1$ alternatives of adding $\{0, 1, 2, \ldots, a\}$ balls (or whatever is our chosen statistic), the q-analog is defined as $[a + 1]_q = q^0 + q^1 + q^2 + \cdots + q^a$. Hence, when the statistic of our interest is a result of a series of n independent choices, such that choice i has the exact $a_i + 1$ alternatives $\{0, 1, 2, \ldots, a_i\}$, we define the *q-analog* $|S|_q$ on the set S as

$$|S|_q = \prod_{i=1}^{n} [a_i + 1]_q \,. \tag{14}$$

4 C-Indexed Families

Let us now proceed to the main section of this paper. As stated in the introduction, the set of acyclic orientations $\mathrm{AO}(\mathbf{a})$, rook placements $\mathrm{RP}(\mathbf{a})$, and perfect matchings $\mathrm{PM}(\mathbf{a})$ correspond to the same area sequence $\mathbf{a} = (a_1, a_2, a_3, \ldots, a_n)$, defined in Chapter 2.3. In this chapter, we will look more closely at this correspondence. We begin by examining the exact relation between a Dyck path D and its associated area sequence $\mathbf{a}$.

Definition 6. Let D be a Dyck path from $(0,0)$ to (n,n). In Chapter 2.3, we defined $\mathbf{a}(D) = (a_1, a_2, a_3, \ldots, a_n)$ as the area sequence associated to D, where a_i is the row area at row i, for each $i \in \{1, 2, 3, \ldots, n\}$. In D, there are n up-steps and n right-steps. There is exactly 1 up-step per row. Define u_i as the up-step at row i, and define s_i as the *index* of u_i in the $2n$ Dyck path, thus forming the sequence $\mathbf{s}(D) = (s_1, s_2, s_3, \ldots, s_n)$. For example, the Dyck path $D = (1,1,0,1,0,1,0,0)$ has 4 up-steps, u_1, u_2, u_3 and u_4. Up-step u_1 has index 1, u_2 has index 2, u_3 has index 4 and u_4 has index 6 in D, so $s_1 = 1$, $s_2 = 2$, $s_3 = 4$ and $s_4 = 6$, thus forming the sequence $\mathbf{s}(D) = (1, 2, 4, 6)$, see Fig. 24, left.

Also, define a *lost square* as any grid unit above the Dyck path and define D_{max} as the unique Dyck path from $(0,0)$ to (n,n), such that $s_i = i$ for all $i \in \{1, 2, 3, \ldots, n\}$, or equivalently: the number of lost squares at row i equals 0, for all $i \in \{1, 2, 3, \ldots, n\}$, see Fig. 24.

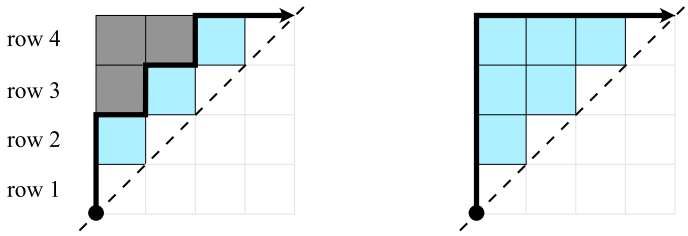


Figure 24: Left: The Dyck path $D = (1,1,0,1,0,1,0,0)$ has 1 lost square at row 3, and 2 lost squares at row 4. Right: D_{max}.

Lemma 7. *For a given Dyck path D from $(0,0)$ to (n,n), let a_i be the row area at row i, and let s_i be the the index of the i:th up-step u_i in D. Then*

$$a_i = 2i - 1 - s_i\,, \tag{15}$$

for each $i \in \{1, 2, 3, \ldots, n\}$.

Proof. By definition, a Dyck path D from $(0,0)$ to (n,n) has n up-steps $u_1, u_2, u_3, \ldots, u_n$, with exactly one up-step per row. We have defined s_i as

the *index* of u_i in D. Roughly speaking, the index s_i tells us "how early or late" the Dyck path moves from row i to row $i+1$. Hence, s_i has a direct impact on the row area at row i.

Let us now consider the unique Dyck path D_{max}. By definition, the number of lost squares at row i in D_{max} equals 0, for all $i \in \{1, 2, 3, \ldots, n\}$. Hence, the row area at row i in D_{max} is $i-1$ (i.e. all units above the diagonal at row i). We know that D_{max} occurs exactly when $s_i = i$ for all rows i, i.e. when all the up-steps comes first in D.

We now consider some other Dyck path D', so that $D' \neq D_{max}$. In D', $s_i \neq i$ for some i. To be more specific, since the i:th up-step cannot be positioned *before* we even reach row i, we must have that $s_i > i$ for some i. Whenever $s_i > i$, lost squares are induced at row i, so in D' there must be lost squares at row i. The number of lost squares at row i of D' equals the difference between the "earliest possible scenario" of D_{max} and the actual scenario of D', which is exactly $s_i - i$, see Fig. 24.

The same argument can be used independently for any row of any Dyck path. Hence, the row area a_i at row i for some Dyck path D equals the row area of D_{max} at row i minus the number of lost squares at row i, which equals $i - 1 - (s_i - i) = 2i - 1 - s_i$. □

Definition 8. Define the *column area sequence*, $\mathbf{b} = (b_1, b_2, b_3, \ldots, b_n)$ as the number of intact grid units between a Dyck path and the main diagonal from $(0, 0)$ to (n, n), at column $1, 2, 3, \ldots, n$.

Lemma 9. *Let* $\mathbf{a} = (a_1, a_2, a_3, \ldots, a_n)$ *be the area sequence corresponding to a Dyck path* D*, and let* $\mathbf{b} = (b_1, b_2, b_3, \ldots, b_n)$ *be the* column *area sequence of the same Dyck path* D*. Then* $\mathbf{b}$ *is a permutation of* $\mathbf{a}$*.*

Proof. This Lemma is easily understood through induction. Consider a Dyck path where there are no lost squares, as in Fig. 25, left. Due to the symmetric construction it is evident that sequences $\mathbf{a}$ and $\mathbf{b}$ are permutations of each other. They contain the same values, although in reverse order. Let us now introduce lost squares, one by one. The first lost square can only be introduced at one position, see Fig. 25, middle. When this lost square is introduced, the largest value of $\mathbf{a}$ and $\mathbf{b}$ diminishes by 1. Hence, $\mathbf{a}$ and $\mathbf{b}$ are still permutations.

We now introduce another lost square, as in Fig. 25, right. Let x the *horizontal* distance between the lost square and main diagonal, measured in area units. Now one element in $\mathbf{a}$ changes from value $x+1$ to x. However, in the figure, we see that the *vertical* distance between the lost square and main diagonal is also x, which means that one element in $\mathbf{b}$ changes from value $x+1$ to x as well. Hence, $\mathbf{a}$ and $\mathbf{b}$ are still permutations. This was not due to pure luck, but a direct consequence of the definition of the diagonal as the line where row position equals column position. So, from any square in the grid, the horizontal

distance to the diagonal equals the vertical distance to the diagonal. Hence, for every introduced lost square, sequence **a** and **b** will always be permutations of each other.

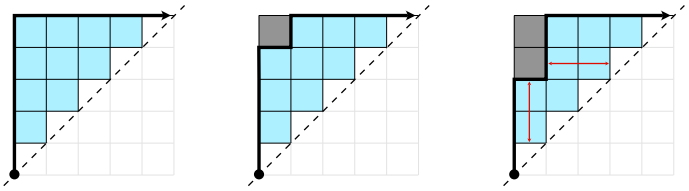


Figure 25: Left: $\mathbf{a} = (0, 1, 2, 3, 4)$ and $\mathbf{b} = (4, 3, 2, 1, 0)$.
Middle: $\mathbf{a} = (0, 1, 2, 3, 3)$ and $\mathbf{b} = (3, 3, 2, 1, 0)$.
Right: $\mathbf{a} = (0, 1, 2, 2, 3)$ and $\mathbf{b} = (2, 3, 2, 1, 0)$.

□

4.1 Acyclic Orientations on Unit Interval Graphs

The area sequence $\mathbf{a} = (a_1, a_2, a_3, \ldots, a_n)$ of a Dyck path, uniquely defines a *unit interval graph*, $\text{UIG}_{\mathbf{a}}$, where the squares on the main diagonal represent vertices $1, 2, 3, \ldots, n$ (counted from the bottom left and upwards to the right). An area sequence unit (blue square) in the horizontal/vertical crossing of two vertices corresponds to an edge between these vertices. Thus, the unit interval graph to the left in Fig. 26 corresponds to the traditional graph drawing to the right. Outside the limitations of this thesis, it should be mentioned that unit interval graphs are traditionally defined through counting overlapping unit intervals on integers, hence the name. However, area sequences give an alternative definition, see [AP18].

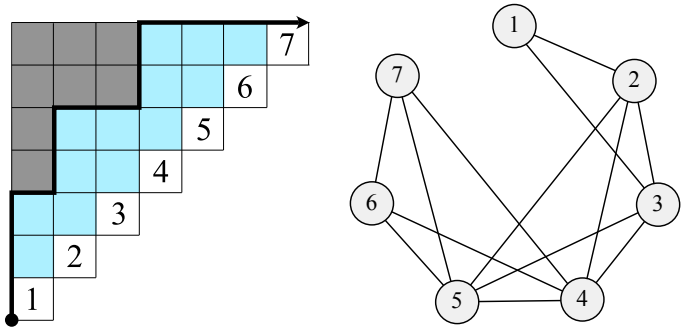


Figure 26: Left: The area sequence $\mathbf{a} = (0, 1, 2, 2, 3, 2, 3)$ defines a unit interval graph, $\text{UIG}_{\mathbf{a}}$. Right: Traditional drawing of the same graph $\text{UIG}_{\mathbf{a}}$.

By assigning an orientation to all edges of our graph $\text{UIG}_{\mathbf{a}}$, it turns into a *directed graph* with an *orientation* θ, see Fig. 27. The orientation of any *single edge* is either *ascending* (oriented from vertex $\alpha \to \beta$, where $\alpha < \beta$), or *descending*

($\alpha \to \beta$, where $\alpha > \beta$). An orientation θ on $\text{UIG}_{\mathbf{a}}$ may, or may not, contain cycles. Define a *cycle* as a subgraph of $\text{UIG}_{\mathbf{a}}$, where the orientation θ provides a directed path from some vertex u to some vertex v *and* a directed path from vertex v to vertex u, with $u \neq v$. Define an *acyclic orientation*, ϑ, as an orientation with no cycles.

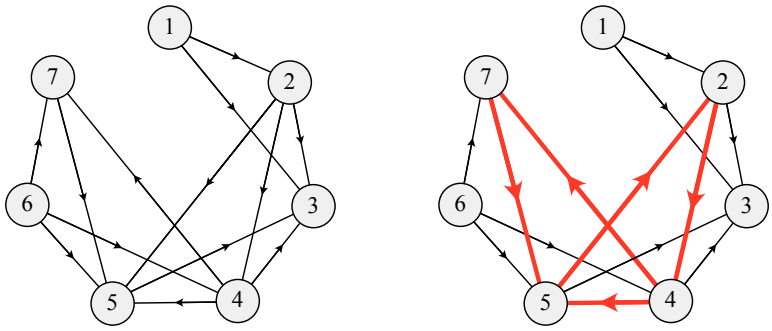


Figure 27: Left: One acyclic orientation on $\text{UIG}_{\mathbf{a}}$ for $\mathbf{a} = (0, 1, 2, 2, 3, 2, 3)$. Edge $1 \to 2$ is ascending and edge $5 \to 3$ is descending. Right: By reversing the orientation of edge $2 \to 5$, the two cycles $5 \to 2 \to 4 \to 5$ and $5 \to 2 \to 4 \to 7 \to 5$ are created on $\text{UIG}_{\mathbf{a}}$.

Let us now introduce the set of acyclic orientations on $\text{UIG}_{\mathbf{a}}$.

Definition 10. Define $\text{AO}(\mathbf{a})$ as the set of acyclic orientations on the unit interval graph $\text{UIG}_{\mathbf{a}}$, corresponding to the area sequence $\mathbf{a}$. For any given acyclic orientation $\vartheta \in \text{AO}(\mathbf{a})$, define $\mathbf{asc}(\vartheta) = (asc_1, asc_2, asc_3, \ldots, asc_n)$ as the sequence counting the number of ascending edges pointed at vertex $i \in \{1, 2, 3, \ldots, n\}$ in $\text{UIG}_{\mathbf{a}}$. Also, define $asc(\vartheta)$ as the total number of ascending edges in ϑ, i.e. the sum of the sequence $\mathbf{asc}(\vartheta)$. Note that $\mathbf{asc}(\vartheta)$ is a vector, and $asc(\vartheta)$ is an integer.

Remark 11. If not restricted to *acyclic* orientations, there might be more than one orientation corresponding to the same **asc**-sequence.

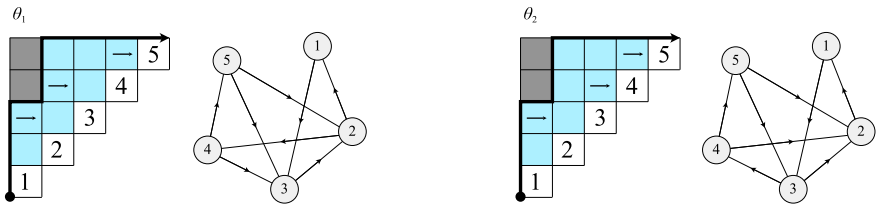


Figure 28: Two orientations, θ_1 and θ_2, that give the same sequence $\mathbf{asc} = (0, 0, 1, 1, 1)$. Ascending edges are marked with right-arrows in the diagram. Edges that are not ascending, are descending.

Definition 12. We prepare for Lemma 13 with a few definitions: In a unit interval graph $\mathrm{UIG}_{\mathbf{a}}$ with n vertices, define sub_i as the subgraph consisting of vertex i and its connected vertices v, such that $v < i$, i.e. all the vertices prior to vertex i such that they are connected to vertex i. Define $presub_i$ as the subgraph sub_i without vertex i. For example, in Fig. 28, sub_5 is the subgraph containing vertices 2, 3, 4 and 5, and $presub_5$ is the subgraph containing vertices 2, 3 and 4.

Also, define *complete graphs* as graphs where *every* distinct pair of vertices is connected by a unique edge. In graph theory, it is well-known that an acyclic orientation of a complete graph defines a *total ordering* on its vertices, so that the *in-degree* of every vertex in the graph is distinct. (The *in-degree* of a vertex v is just the number of edges pointed at v).

Lemma 13. *Consider the the set of acyclic orientations* $\mathrm{AO}_{(\mathbf{a})}$ *on the unit interval graph* $\mathrm{UIG}_{\mathbf{a}}$*, corresponding to the area sequence* $\mathbf{a}$*, and let* $\mathbf{asc}_{(\vartheta)} = (asc_1, asc_2, asc_3, \ldots, asc_n)$ *be the sequence counting the number of ascending edges pointed at vertex* i *in each acyclic orientation* $\vartheta \in \mathrm{AO}_{(\mathbf{a})}$*. Then each distinct sequence* $\mathbf{asc}_{(\vartheta)}$ uniquely *defines* one *acyclic orientation* ϑ*.*

Proof. Consider an orientation of a graph with n vertices. It is trivial to see that, for $n = 1$, the number of ascending edges is exactly 0. For $n = 2$, every orientation has to be acyclic since there cannot be cycles in graphs with less than 3 vertices. It is also evident that the unique edge between vertex 1 and vertex 2 is either ascending or descending, i.e. each acyclic orientation corresponds to exactly 1 distinct value on asc_2. Hence, the Lemma is true for $n = 1$ and $n = 2$.

Let us proceed by induction. Consider a unit interval graph $\mathrm{UIG}_{\mathbf{a}}$ with n vertices, and let us say we already have an acyclic orientation on the first $n - 1$ vertices. Now we want to include vertex n into this orientation, without introducing any cycles, see Fig. 29, left. We realise that since the $[n-1]$-graph is acyclic by definition, any cycles in $\mathrm{UIG}_{\mathbf{a}}$ can *only* be introduced in the subgraph sub_n, see Fig. 29, middle.

Due to the restrictions of legal steps in a monotonic path, $presub_n$ and sub_n have to be a *complete graphs*, and we know that an acyclic orientation of a complete graph defines a *total ordering* on its vertices. Hence, we can picture the acyclic orientation of $presub_n$ as an ordered line, into which vertex n now has to fit. The number of distinct total orderings of sub_n is equal to the number of positions we can choose for vertex n in this ordering, see Fig. 29, right.

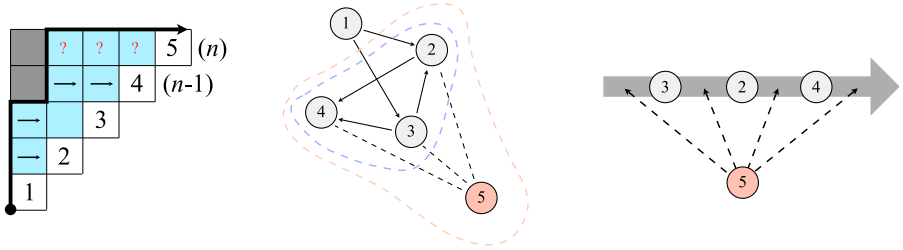


Figure 29: Left: The acyclic graph $\text{UIG}_{\mathbf{a}}$, before adding orientation to vertex n. Middle: $\text{UIG}_{\mathbf{a}}$, with the complete subgraphs sub_n (within the red dotted line), and $presub_n$ (within the blue dotted line). Right: The total ordering of $presub_n$, and the possible positions for vertex n.

The subgraph sub_n is acyclic if, and only if, there is a total ordering on its vertices, and each distinct total ordering uniquely defines one acyclic orientation. Hence, the number of ways to add vertex n to the acyclic orientation of the $[n-1]$-graph, equals the number of vertices in sub_n. Due to the restrictions of legal steps in a monotonic path, and since the horizontal distance to the main diagonal always equals the vertical distance to the main diagonal (see proof of Lemma 9), the number of vertices in sub_n is always equal to $a_n + 1$, see Fig. 29, left.

Now, let us say we choose position p in this ordering. Then all vertices prior to p, points at n, and since $n > n-1 > n-2 \ldots$, these edges must be ascending by definition. It is evident that, for each position p, there is a unique number of vertices prior to p in the ordering. We see that for each of the $a_n + 1$ legal ways to add vertex n to the acyclic graph, there is a *unique* number $\{0, 1, 2, \ldots, a_n\}$ of ascending edges pointed at vertex n in the unit interval graph, i.e. a distinct value $asc_n \in \{0, 1, 2, \ldots, a_n\}$ at index n in the sequence $\mathbf{asc}(\vartheta)$.

Naturally, the same argument can be used for any row $i \in \{1, 2, 3, \ldots, n\}$, and we thus see that each distinct sequence $\mathbf{asc}(\vartheta) = (asc_1, asc_2, asc_3, \ldots, asc_n)$ uniquely defines one acyclic orientation ϑ.

□

An alternative way of looking at Lemma 13, is this: Of all possible orientations corresponding to the same **asc**-sequence, exactly 1 is acyclic, see Remark 11.

Theorem 14. *Let* $\mathrm{AO}_{(\mathbf{a})}$ *be the set of acyclic orientations corresponding to the area sequence* $\mathbf{a}$ *with the* column *area sequence* $\mathbf{b}$*, and let* $|\mathrm{AO}_{(\mathbf{a})}|_q$ *be the* q*-analog of* $\mathrm{AO}_{(\mathbf{a})}$*, with respect to the number of ascending edges* $(asc(\vartheta))$ *in each acyclic orientation* $\vartheta \in \mathrm{AO}_{(\mathbf{a})}$*. Then*

$$|\mathrm{AO}_{(\mathbf{a})}|_q = \sum_{\vartheta \in \mathrm{AO}(\mathbf{a})} q^{asc(\vartheta)} = \prod_{i=1}^{n} [a_i + 1]_q = \prod_{i=1}^{n} [b_i + 1]_q \,, \tag{16}$$

for any area sequence $\mathbf{a}$*.*

Proof. We have defined $|\mathrm{AO}_{(\mathbf{a})}|_q$ as the q-analog on $\mathrm{AO}_{(\mathbf{a})}$, with respect to the number of ascending edges in each orientation $\vartheta \in \mathrm{AO}_{(\mathbf{a})}$. This is exactly

$$\sum_{\vartheta \in \mathrm{AO}(\mathbf{a})} q^{asc(\vartheta)} \,.$$

In the proof of Lemma 13, we saw that the number of ways to add vertex n to an acyclic orientation of a $[n-1]$-graph of $\mathrm{UIG}_{\mathbf{a}}$, equals $a_n + 1$. Let us say we already have r distinct acyclic orientations before we add vertex n. Each of these can be combined with the new possibilities at row n. We know that this is true for *any* row in the graph (Lemma 13), and we realise that, since the number of such possibilities at row i is $a_i + 1$ for any row $i \in \{1, 2, 3, \ldots, n\}$, the total number of acyclic orientations is the product of the possibilities at each row i, hence $\prod_{i=1}^{n}(a_i + 1)$. From Lemma 13 we have that each acyclic orientation ϑ *uniquely* defines one distinct sequence $\mathbf{asc}_{(\vartheta)}$, such that $asc_i \in \{0, 1, 2, \ldots, a_i\}$ for all $i \in \{1, 2, 3, \ldots, n\}$.

Combinatorically speaking, the rows represent independent choices, and as we saw in (14), when the statistic *ascending edges* is the result of a series of n independent choices, such that the i:th choice has the exact $a_i + 1$ alternatives of adding $\{0, 1, 2, \ldots, a_i\}$ ascending edges, the q-analog is defined as $\prod_{i=1}^{n}[a_i + 1]_q$. For details on q-analog notation, see (13) and (14).

The q-analog $\prod_{i=1}^{n}[a_i + 1]_q$ is the product of a series of independent choices. Naturally, the *order* of these choices does not alter the refinement (Chapter 3). Hence, since the column area sequence $\mathbf{b}$ is a *permutation* of the area sequence $\mathbf{a}$ (Lemma 9), $\prod_{i=1}^{n}[a_i + 1]_q = \prod_{i=1}^{n}[b_i + 1]_q$.

□

Example 15. Consider the set of acyclic orientations $\mathrm{AO}_{(\mathbf{a})}$ on the unit interval graph $\mathrm{UIG}_{\mathbf{a}}$, corresponding to the area sequence $\mathbf{a} = (0, 1, 2, 2, 1)$, as in Fig. 30. Then the number of acyclic orientations $\vartheta \in \mathrm{AO}_{(\mathbf{a})}$ is

$$\prod_{i=1}^{n}(a_i + 1) = 1 \cdot 2 \cdot 3 \cdot 3 \cdot 2 = 36. \tag{17}$$

We refine the set cardinality by q-counting this set with respect to the number of ascending edges in each orientation. Using (16) we have that

$$\begin{aligned} |\mathrm{AO}_{(\mathbf{a})}|_q = \prod_{i=1}^{n}[a_i + 1]_q &= (q^0)(q^0 + q^1)(q^0 + q^1 + q^2)(q^0 + q^1 + q^2)(q^0 + q^1) \\ &= 1q^0 + 4q^1 + 8q^2 + 10q^3 + 8q^4 + 4q^5 + 1q^6. \end{aligned} \tag{18}$$

Hence, in this set we have 1 ϑ with no ascending edges, 4 ϑ with one ascending edge, 8 ϑ with two ascending edges, 10 ϑ with three ascending edges, 8 ϑ with four ascending edges, four ϑ with five ascending edges and 1 ϑ with six ascending edges. Any reader with plenty of free time is strongly encouraged to verify this result. For $q = 1$, the sum of (18) equals the cardinality of the set, which is 36.

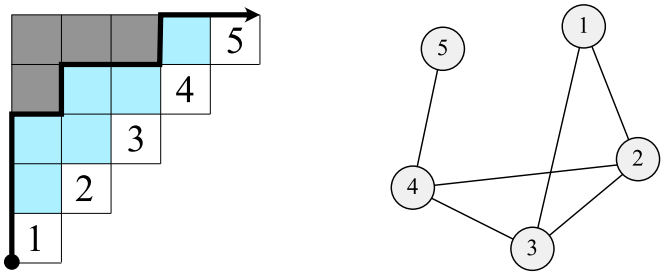


Figure 30: The unit interval graph $\mathrm{UIG}_{\mathbf{a}}$, corresponding to the area sequence $\mathbf{a} = (0, 1, 2, 2, 1)$.

4.2 Rook Placements on Ferrers Boards

Ferrers boards are traditionally used to visualise partitions of some integer n into m parts, where each part is represented with one row. Thus, a partition of the integer 7 into the parts $4+2+1$ would generate a Ferrers board with 3 rows of length 4, 2 and 1, see Fig. 31.

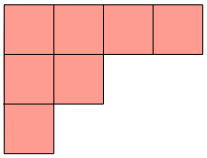


Figure 31: One partition of the integer 7, in the form of a Ferrers board. Note that the actual drawing of a Ferrers board can be rotated and mirrored.

More specifically, Ferrers boards are subsets of square grids, where each square has row position i and column position j, i.e. the squares of the board has coordinates i, j. Now, consider an area sequence $\mathbf{a} = (a_1, a_2, a_3, \ldots, a_n)$, as described in Chapter 4.1, see Fig. 32, left. We define the Ferrers board $\mathrm{FB}_{\mathbf{a}}$, corresponding to $\mathbf{a}$, as the unique Ferrers board containing *all* the squares with coordinates restricted by $1 \leq i \leq n$ and $n+1-(a_{(n+1-i)}+i) \leq j \leq n$, where the rows on the Ferrers board are counted *from the top*, and columns from the left. Consequently, in $\mathrm{FB}_{\mathbf{a}}$, define the *row length*, r_i, of the i:th row from the top, as $a_{(n+1-i)}+i$. To avoid confusion, note that r_1 is the top row of $\mathrm{FB}_{\mathbf{a}}$, whereas a_1 is the bottom row of $\mathbf{a}$. It is clear that each distinct area sequence $\mathbf{a}$ defines exactly one Ferrers board $\mathrm{FB}_{\mathbf{a}}$, see Fig. 32, middle. The traditional interpretation of $\mathrm{FB}_{\mathbf{a}}$ would be one possible partition of the integer $\sum_{i=1}^{n} r_i$ into n parts.

However, over the years, there has also been some interest in the properties of non-attacking rook placements on Ferrers boards [Sta11]. A *non-attacking rook placement*, Ψ, on a Ferrers board $\mathrm{FB}_{\mathbf{a}}$ with n rows, is a setup of n rooks, where no single rook can attack any other rook, according to the rules of chess (i.e. column position u and row position v are distinct in $X_{u,v}$, for all rooks $X \in \mathrm{FB}_{\mathbf{a}}$), see Fig. 32, right. In this chapter, we are interested in the set $\mathrm{RP}_{(\mathbf{a})}$ of such non-attacking rook placements on the Ferrers board $\mathrm{FB}_{\mathbf{a}}$.

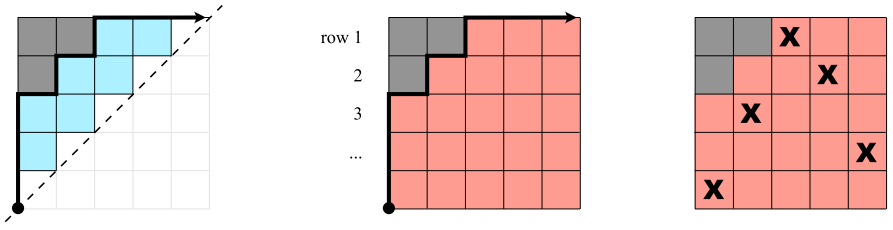


Figure 32: Left: The area sequence $\mathbf{a}$ in blue. Middle: The Ferrers board $\mathrm{FB}_{\mathbf{a}}$. Right: A rook placement Ψ on $\mathrm{FB}_{\mathbf{a}}$.

Definition 16. Within the context of non-attacking rook placements on Ferrers boards, define an *inversion* as one square $\boxed{\bullet}$ of the board, where these three conditions are met:

1. There is no rook on $\boxed{\bullet}$.
2. There are no rooks to the left of $\boxed{\bullet}$ in the same row.
3. There are no rooks above $\boxed{\bullet}$ in the same column.

Define $\mathbf{inv}(\Psi) = (inv_1, inv_2, inv_3, \ldots, inv_n)$ as the sequence counting the number of inversions at each row $i \in \{1, 2, 3 \ldots, n\}$ of a rook placement Ψ. Also, define $inv_{(\Psi)}$ as the total number of ascending edges in Ψ, i.e. the sum of the sequence $\mathbf{inv}_{(\Psi)}$. An example is provided in Fig. 33.

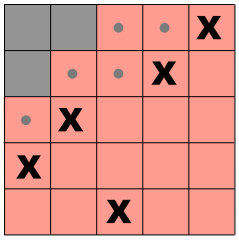


Figure 33: A rook placement Ψ, where $\mathbf{inv}(\Psi) = (2, 2, 1, 0, 0)$, and $inv_{(\Psi)} = 5$.

In Lemma 13, we saw that an acyclic orientation is *uniquely* defined by the number of ascending edges in each row. Analogously, a rook placement is uniquely determined by the number of inversions in each row.

Lemma 17. *Consider the set of rook placements using n rooks, $\mathrm{RP}_{(\mathbf{a})}$, on the Ferrers board $\mathrm{FB}_{\mathbf{a}}$, corresponding to the area sequence $\mathbf{a}$. For each rook placement $\Psi \in \mathrm{RP}_{(\mathbf{a})}$, let $\mathbf{inv}(\Psi) = (inv_1, inv_2, inv_3, \ldots, inv_n)$ be the number of inversions at each row of Ψ. Then Ψ is uniquely defined by $\mathbf{inv}(\Psi)$, i.e. for each distinct inversion sequence, there is one, and only one, rook placement Ψ.*

Proof. It is easy to see that there are $n!$ non-attacking rook placements on a "full" $n \times n$ Ferrers board. Since there are n rooks, any rook placement has exactly 1 rook per row and, placing one rook per row, the first can be placed in n positions, the second in $n-1$ positions etc., all the way down to the last rook, which can be placed in exactly 1 position. A similar approach can be used for Ferrers boards derived from *any* area sequence.

Consider a Ferrers board $\mathrm{FB}_{\mathbf{a}}$ with row lengths $r_1, r_2, r_3, \ldots, r_n$, counted from the top row, see Fig. 34, left, and note that the row lengths $(r_1, r_2, r_3, \ldots, r_n)$ are increasing by definition, so that $r_i \leq r_{i+1} \leq r_{i+2}$ and so on. Since the row lengths are $a_{(n+1-i)} + i$, the first rook can be placed in $a_n + 1$ positions, the

second in $(a_{n-1}+2)-1$ positions, the third in $(a_{n-2}+3)-2$ and so on all the way down to the last rook which can be placed in $(a_1+n)-(n-1)$ positions. Hence, for any row $i \in \{1,2,3,\ldots,n\}$, the number of rook placements is $a_{(n+1-i)}+1$.

We now want to prove that, for each given rook placement Ψ, there is a *unique* inversion sequence $\mathbf{inv}(\Psi)$, and we use the principle of induction. We construct a rook placement Ψ by adding one rook at a time, starting at row 1. This rook can be placed in any of the r_1 squares, and we choose square x. According to the definition of inversions, *all* squares to the left of x are inversions, see Fig. 34, middle. Thus, each position x uniquely defines the number of inversions at row 1.

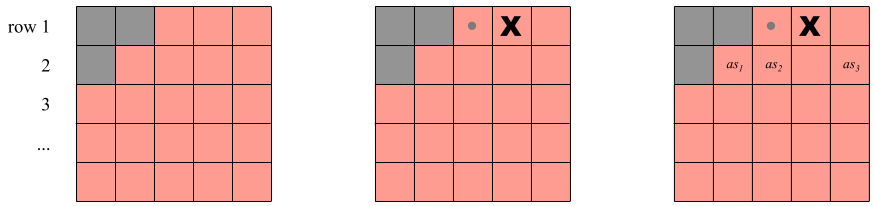


Figure 34: Left: The Ferrers board $\mathrm{FB}_{\mathbf{a}}$. Middle: Placement of the first rook at row 1. Squares $\boxed{\bullet}$ are inversions. Right: The three *allowed squares*, as_1, as_2 and as_3, at row 2.

We now proceed to row 2. On this row, there are r_2-1 squares for the rook to choose from (remembering that the square below x has been disqualified). Let $as_1, as_2, as_3, \ldots, as_{(r_2-1)}$ be such *allowed squares*, numbered from left to right, see Fig. 34, right. Since the definition of inversions is equivalent to *allowed squares to the left of the rook*, choosing as_1 is equivalent to having 0 inversions. If we abandon the alternative as_1, and choose as_2 instead, as_1 becomes an open square to the left of as_2, and thus an inversion. Choosing as_2 is therefore equivalent to having 1 inversion. Consequently, choosing as_3 means 2 inversions, and for the k:th *allowed square* from the left, there are exactly $k-1$ inversions. Hence, we see that for each rook position at row 2, there is a *unique* number $\{0,1,2,\ldots,r_2-2\}$ of inversions at row 2, i.e a distinct inv_2.

Naturally, this argument can be repeated for the remaining rows. So, for each of the $a_{(n+1-i)}+1$ rook positions at row $i \in \{1,2,3\ldots,n\}$, there is a *distinct* number $\{0,1,2,\ldots,r_i-i\}$ of inversions at row i. Remember that r_i is defined as $a_{(n+1-i)}+i$, and so we have that for each of the $a_{(n+1-i)}+1$ rook positions at row $i \in \{1,2,3\ldots,n\}$, there is a distinct number $\{0,1,2,\ldots,a_{(n+1-i)}\}$ of inversions at row i. Since inv_i is uniquely defined by the rook placement at row i, the full inversion sequence $\mathbf{inv}(\Psi) = (inv_1, inv_2, inv_3, \ldots, inv_n)$ is uniquely defined by the rook placement at row $1,2,3,\ldots,n$. □

Theorem 18. *Let* $\mathrm{RP}_{(\mathbf{a})}$ *be the set of rook placements corresponding to the area sequence* $\mathbf{a}$*, and let* $|\mathrm{RP}_{(\mathbf{a})}|_q$ *be the* q*-analog of* $\mathrm{RP}_{(\mathbf{a})}$*, with respect to the number of inversions* ($inv_{(\Psi)}$) *in each rook placement* $\Psi \in \mathrm{RP}_{(\mathbf{a})}$*. Then*

$$|\mathrm{RP}_{(\mathbf{a})}|_q = \sum_{\Psi \in \mathrm{RP}(\mathbf{a})} q^{inv(\Psi)} = \prod_{i=1}^{n} [a_i + 1]_q \,, \tag{19}$$

for any area sequence $\mathbf{a}$*.*

Proof. We have defined $|\mathrm{RP}_{(\mathbf{a})}|_q$ as the q-analog on $\mathrm{RP}_{(\mathbf{a})}$, with respect to the number of inversions in each rook placement $\Psi \in \mathrm{RP}_{(\mathbf{a})}$. This is exactly

$$\sum_{\Psi \in \mathrm{RP}(\mathbf{a})} q^{inv(\Psi)} \,.$$

According to the multiplication principle, the number of rook placements equals the product of the possibilities at each row, which is $\prod_{i=1}^{n}(a_{(n+1-i)} + 1)$, see Lemma 17. We know that each rook placement Ψ uniquely defines one inversion sequence $\mathbf{inv}(\Psi)$, and that for each of the $a_{(n+1-i)} + 1$ rook positions at row $i \in \{1, 2, 3 \ldots, n\}$, there is a distinct number $\{0, 1, 2, \ldots, a_{(n+1-i)}\}$ of inversions at row i (Lemma 17).

As we saw in (14), when our statistic *inversions* is a result of a series of n independent choices, such that the i:th choice has the exact $a_{(n+1-i)} + 1$ alternatives of adding $\{0, 1, 2, \ldots, a_{(n+1-i)}\}$ inversions, the q-analog is given by $\prod_{i=1}^{n}[a_{(n+1-i)} + 1]_q$. However, the *order* in which we multiply does not change the resulting q-analog, and for $i \in \{1, 2, 3 \ldots, n\}$, the rows $a_{(n+1-i)}$ is just the reverse order of the rows a_i. Hence, $\prod_{i=1}^{n}[a_{(n+1-i)} + 1]_q = \prod_{i=1}^{n}[a_i + 1]_q$. □

Corollary 19. *For any area sequence* $\mathbf{a}$*,*

$$|\mathrm{RP}_{(\mathbf{a})}|_q = |\mathrm{AO}_{(\mathbf{a})}|_q \,. \tag{20}$$

Proof. A given area sequence $\mathbf{a}$ *uniquely determines* the q-analog $\prod_{i=1}^{n}[a_i + 1]_q$. The corollary follows from Theorem 14 and Theorem 18. □

Example 20. Consider the set of non-attacking rook placements $\mathrm{RP}_{(\mathbf{a})}$ on the Ferrers board $\mathrm{FB}_{\mathbf{a}}$, corresponding to the area sequence $\mathbf{a} = (0, 1, 1, 2)$, as in Fig. 35, left. Then the number of rook placements $\Psi \in \mathrm{RP}_{(\mathbf{a})}$ is

$$\prod_{i=1}^{n}(a_i + 1) = 1 \cdot 2 \cdot 2 \cdot 3 = 12. \tag{21}$$

We refine the set cardinality by q-counting this set with respect to the number of inversions in each rook placement. Using (19) we have that

$$\begin{aligned} |\mathrm{RP}_{(\mathbf{a})}|_q = \prod_{i=1}^{n}[a_i + 1]_q &= (q^0)(q^0 + q^1)(q^0 + q^1)(q^0 + q^1 + q^2) \\ &= 1q^0 + 3q^1 + 4q^2 + 3q^3 + 1q^4. \end{aligned} \tag{22}$$

Hence, in this set we have 1 Ψ with no inversions, 3 Ψ with one inversion, 4 Ψ with two inversions, 3 Ψ with three inversions and 1 Ψ with four inversions. We encourage the reader to verify this result. For $q = 1$, the sum of (22) equals the cardinality of the set, which is 12. According to Corollary 19, this refinement mirrors the refinement on the set of acyclic orientations $\mathrm{AO}_{(\mathbf{a})}$ on the unit interval graph $\mathrm{UIG}_{\mathbf{a}}$, with respect to the number of ascending edges in each orientation $\vartheta \in \mathrm{AO}_{(\mathbf{a})}$, see Fig. 35, right.

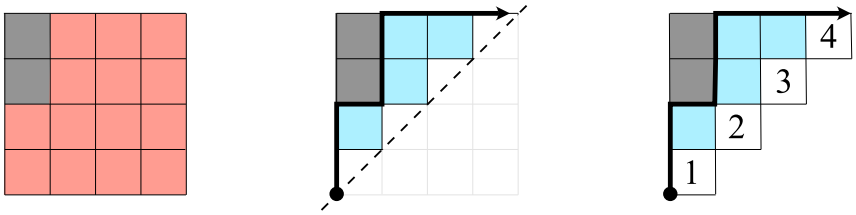


Figure 35: Left: The Ferrers board $\mathrm{FB}_{\mathbf{a}}$, corresponding to ... Middle: ... the area sequence $\mathbf{a} = (0, 1, 1, 2)$. Right: The unit interval graph $\mathrm{UIG}_{\mathbf{a}}$.

4.3 Perfect Matchings on PM-Starting Points

As stated earlier, a *perfect matching*, Ω, is a pairing of of $2n$ elements into n pairs, visualised by the drawing of n connecting chords between $2n$ points on a circle, see Fig. 36. Such perfect matchings can have one or more crossings, or be non-crossing. In Chapter 2.3 we saw that *non-crossing* perfect matchings and area sequences are Catalan objects, linked by a bijection. In this chapter we will see that, even though this is true, it is more helpful to think of non-crossing perfect matchings as just *one element* in a set of perfect matchings, and that each such set contains exactly 1 non-crossing perfect matching.

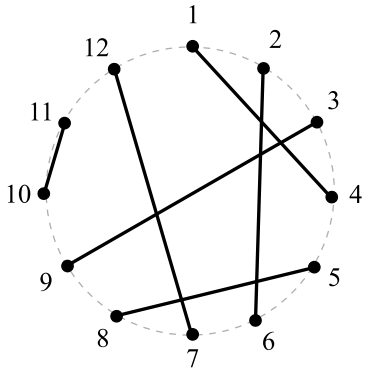
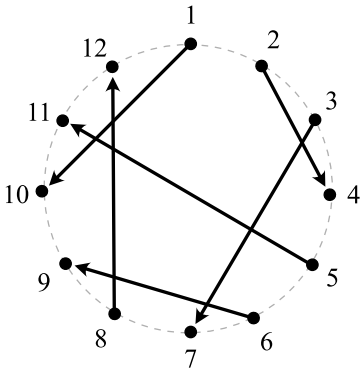


Figure 36: A perfect matching of 12 points into 6 pairs.

Definition 21. Consider a perfect matching Ω of $2n$ points, and define the *matching array of* Ω as an array of the n pairs $s_1 - e_1, s_2 - e_2, s_3 - e_3, \ldots, s_n - e_n$,

$$\Omega = \begin{bmatrix} e_1 & e_2 & e_3 & \ldots & e_n \\ s_1 & s_2 & s_3 & \ldots & s_n \end{bmatrix},$$

ordered so that the bottom row is strictly increasing, and the top row has the larger of the two values in each pair, i.e. such that $e_i > s_i$ for all $i \in [n]$ and $s_i < s_{i+1}$ for all $i \in [n-1]$, see Fig. 37.



$$\Omega = \begin{bmatrix} 10 & 4 & 7 & 11 & 9 & 12 \\ 1 & 2 & 3 & 5 & 6 & 8 \end{bmatrix}$$

Figure 37: Arbitrary example of a perfect matching Ω, and the matching array of Ω.

Arranged this way, the bottom row becomes a subsequence of *starting points* and the top row a subsequence of *ending points*. Define $\mathbf{s}(\Omega) = (s_1, s_2, s_3, \ldots, s_n)$ as the sequence of starting points, and $\mathbf{e}(\Omega) = (e_1, e_2, e_3, \ldots, e_n)$ as the sequence of ending points in the matching array of Ω. This ordering principle will prove to be useful, but note that any actual matching Ω is only dependent on its pairs, and it is thus independent of the ordering of these.

Definition 22. Although we have discussed *non-crossing* perfect matchings several times, we have not provided a definition yet. Given the set $\mathrm{PM}(\mathbf{a})$, we define the perfect matching $\Omega \in \mathrm{PM}(\mathbf{a})$ as *non-crossing* if, and only if, $e_{i-1} < e_i \Rightarrow e_{i-1} < s_i$, see Fig. 38. In other words, for every index i in the matching array of Ω: whenever there are elements $x \in \mathbf{e}(\Omega)$, such that $x < e_i$, *to the left* of index i, these elements also need to be such that $x < s_i$. In any set of perfect matchings, we let Ω^* denote the unique perfect matching of the set.

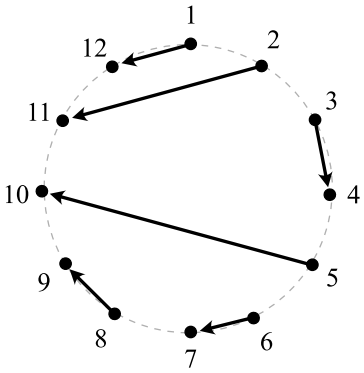


Figure 38: The unique non-crossing perfect matching with starting points $\mathbf{s}(\Omega^*) = (1, 2, 3, 5, 6, 8)$.

Definition 23. Let ψ be the function defined by the following algorithm: In a perfect matching Ω of $2n$ points, start at point 1 and visit each point $i \in \{1, 2, 3, \ldots, 2n\}$ in the numbered order. For each point i, let $x_i = 1$ if the chord starting at point i is encountered for the first time, and let $x_i = 0$ otherwise. Construct the binary word $\mathbf{x}(\Omega)$, see Fig. 39.

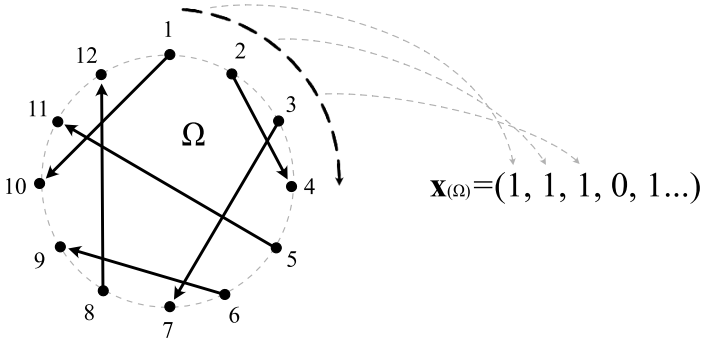


Figure 39: The function ψ at work.

Lemma 24. *The number of distinct binary words* $\mathbf{x}$ *of length* $2n$*, generated by the function* ψ*, equals the* n*:th Catalan number* C_n*.*

Proof. It is fairly easy to see that the number of distinct binary words generated by the function ψ is limited by the exact same rules as that of valid parenthesis words, described in Chapter 2.3. By definition, when we feed any perfect matching of $2n$ points into ψ, the output has to be *balanced* pairs of 1:s and 0:s and, according to the definition of ψ, any such pair must start with a "1".

To clarify, *balanced pairs* means that the total number of 1:s equals the total number of 0:s, and that, at any index $i \in \{1, 2, 3, \ldots, 2n\}$, the number of 1:s *so far* is equal to, or greater than, the number of 0:s *so far*. The rest follows from Chapter 2.3. We encourage the reader to compare to the proof of (7), for more details.

□

A direct consequence of Lemma 24 is that the set of distinct binary words $\mathbf{x}$ of length $2n$, generated by the function ψ, equals the set of Dyck path words of length $2n$. Let us consider one specific such binary word, corresponding to the area sequence $\mathbf{a}$ and the matching Ω. Remember that the sequence $\mathbf{s}_{(\mathbf{a})}$ indexes the up-steps in $\mathrm{DP}_{\mathbf{a}}$, i.e. $s_{i(\mathbf{a})}$ equals the index of the i-th "1" in the binary word, see Definition 6. Similarly, the subsequence $\mathbf{s}_{(\Omega)}$ indexes the starting points of Ω, i.e. $s_{i(\Omega)}$ *also* equals the index of the i-th "1" in the binary word. Hence, for one specific area sequence, the sequence $\mathbf{s}_{(\mathbf{a})}$, indexing the up-steps in $\mathrm{DP}_{\mathbf{a}}$, equals the subsequence $\mathbf{s}_{(\Omega)}$ of starting points in Ω, see Fig. 40. Naturally, this is why we let $\mathbf{s} = (s_1, s_2, s_3, \ldots, s_n)$ denote either of them.

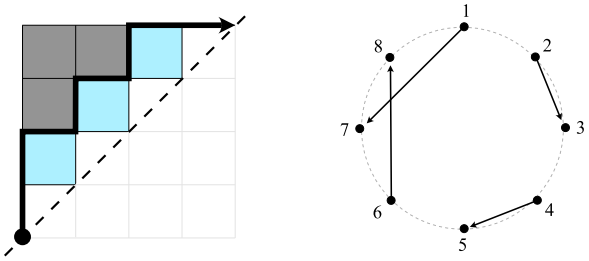


Figure 40: $\mathbf{s}_{(\mathbf{a})} = \mathbf{s}_{(\Omega)} = (1, 2, 4, 6)$.

However, each distinct binary word could correspond to more than one perfect matching, since we could just switch some of the chords ending points, without altering the binary word. For example, in Fig. 40, we could switch the ending points of the chord starting at 1 and the chord starting at 4. This would alter the perfect matching, but the binary word would not change. It follows that the number of perfect matchings Ω corresponding to the same binary word $\mathbf{x}$ equals the number of legal set-ups of ending points, i.e. sequences $\mathbf{e}_{(\Omega)}$ such that $e_i > s_i$ for all $i \in [n]$ and $s_i < s_{i+1}$ for all $i \in [n-1]$.

We are now ready to define the set of perfect matchings corresponding to $\mathbf{a}$.

Definition 25. For a given area sequence $\mathbf{a} = (a_1, a_2, a_3, \ldots, a_n)$, define the Catalan element $\mathrm{PMSP}_{(\mathbf{a})}$ as the specific set-up of starting points $\mathbf{s}_{(\mathbf{a})}$, corresponding to $\mathbf{a}$, and define $\mathrm{PM}_{(\mathbf{a})}$ as the set of perfect matchings on $\mathrm{PMSP}_{(\mathbf{a})}$. Note that, for any $\mathbf{a}$, *all* perfect matchings $\Omega \in \mathrm{PM}_{(\mathbf{a})}$ share the same starting points, see Fig. 41.

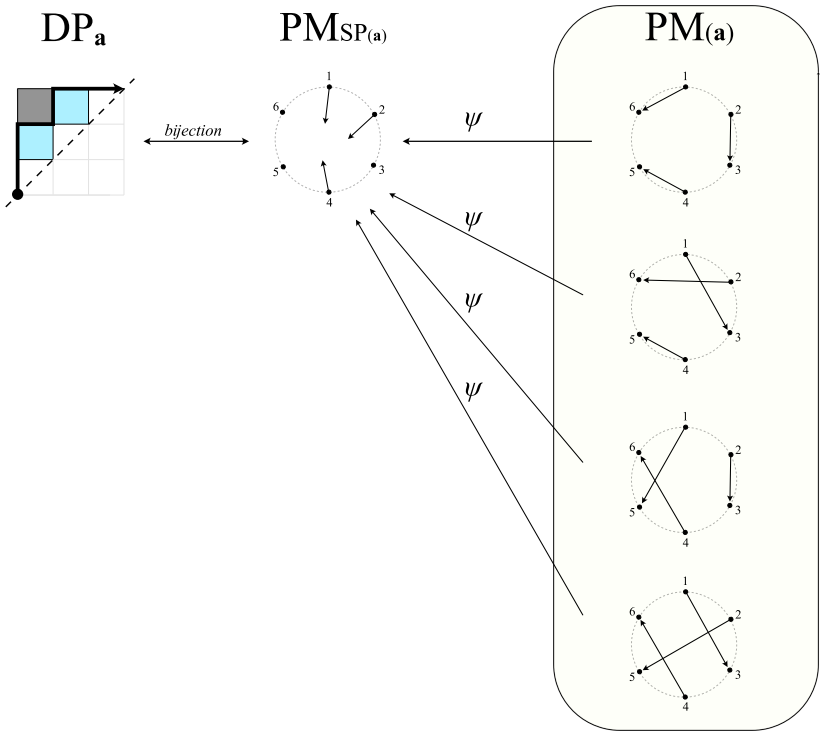


Figure 41: The perfect matchings of PM($\mathbf{a}$), corresponding to the area sequence $\mathbf{a} = (0, 1, 1)$. All the matchings have starting points 1, 2 and 4, i.e. $\mathbf{s}(\mathbf{a}) = (1, 2, 4)$. There are 4 matchings Ω, i.e. 4 possible permutations of $\mathbf{e}(\Omega)$.

Intuitively, we might think that a "crossing" of two chords involves both chords equally (and it does!). However, when we *count* crossings, we need to make sure we do not count every crossing twice.

Definition 26. In any intersection of two chords with starting points a and b, where $b > a$, we define the *crossing* as belonging to *only* the chord from b. For example, if we switch the ending points "4" and "11" in Fig. 38, the chord starting at "2" has 0 crossings, and the chord starting at "3" has 1 crossing.

Definition 27. In a perfect matching Ω, define the sequence $\mathbf{cr}(\Omega) = (cr_1, cr_2, cr_3, \ldots, cr_n)$ as the sequence counting the number of crossings belonging to the chords with starting points $s_1, s_2, s_3, \ldots, s_n$. Also, define $cr(\Omega)$ as the total number of ascending edges in Ω, i.e. the sum of the sequence $\mathbf{cr}(\Omega)$.

Lemma 28. *Consider the set* PM($\mathbf{a}$) *of perfect matchings, corresponding to the area sequence* $\mathbf{a}$*, and let the sequence* $\mathbf{cr}(\Omega) = (cr_1, cr_2, cr_3, \ldots, cr_n)$ *be the sequence counting the number of crossings belonging to the chords with starting points* $s_1, s_2, s_3, \ldots, s_n$*. Then each perfect matching* $\Omega \in$ PM($\mathbf{a}$) *uniquely defines one sequence* $\mathbf{cr}(\Omega)$*.*

Proof. Remember that all matchings in the set $\mathrm{PM}(\mathbf{a})$ share the same starting points. Now, starting with the unique non-crossing perfect matching $\Omega^* \in \mathrm{PM}(\mathbf{a})$, the chord from s_i to e_i could *potentially* cross exactly the chords with ending points larger than e_i and starting points smaller than s_i. All other crossings will belong to some other chord by definition, see Fig. 42.

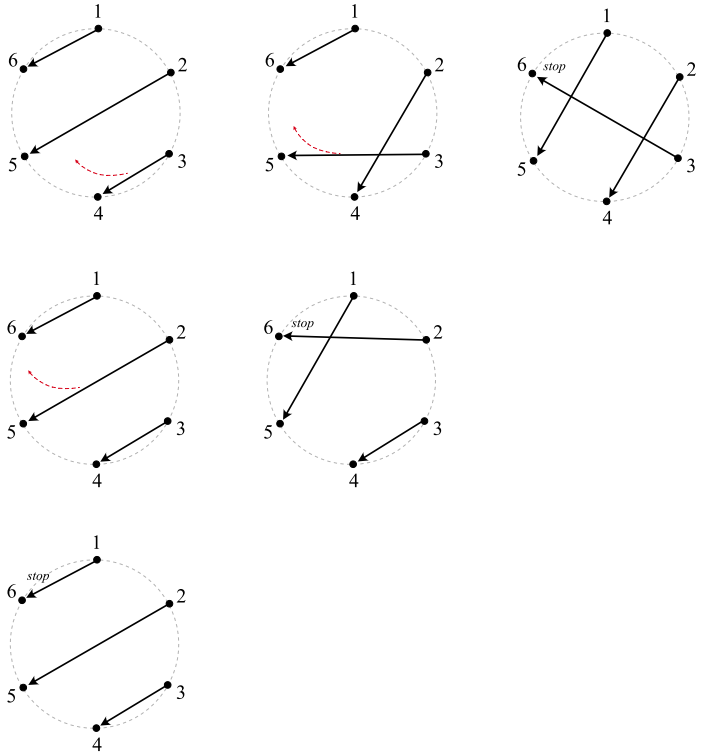


Figure 42: For this specific set-up of starting points: Top: Chord 3 can cross exactly 2 chords. Middle: Chord 2 can cross exactly 1 chord. Bottom: Chord 1 can cross exactly no chord.

In the matching array, any such legal crossing corresponds to switching e_i with some element $z \in \mathbf{e}(\Omega)$, such that z is to the left of position i and $z > e_i$. The number of such elements is $i - 1$, minus all elements x, such that $x < e_i$. We thus need to find the number of elements x, such that $x < e_i$, and we proceed by induction:

Let $\Omega^*(\mathbf{a})$ be the unique non-crossing perfect matching in $\mathrm{PM}(\mathbf{a})$, such that $s_i = i$ for all i and $e_1 > e_2 > e_3 \cdots > e_n$. For example,

$$\Omega^*(\mathbf{a}) = \begin{bmatrix} 6 & 5 & 4 \\ 1 & 2 & 3 \end{bmatrix}.$$

Since the sequence $\mathbf{e}(\Omega^*(\mathbf{a}))$ is strictly decreasing, it is obvious that to the left of position i in $\mathbf{e}(\mathbf{a})$, there are exactly 0 elements such that $x < e_i$.

Let us now alter the sequence of starting points and consider a non-crossing perfect matching $\Omega^*(\mathbf{b})$, corresponding to the area sequence $\mathbf{b}$. In this sequence, $s_i > i$ for some i, for example

$$\Omega^*(\mathbf{b}) = \begin{bmatrix} 6 & 3 & 5 \\ 1 & 2 & 4 \end{bmatrix}.$$

With this alteration, we have "forced" $s_i - i$ elements to switch from the row of starting points to the row of ending points, namely the element "3" in our example. All such forced elements are by definition elements x such that $x < s_i$, and they thus can only be positioned to the left of position i since $s_i < s_{i+1}$ for all $i \in [n-1]$ and $e_i > s_i$ for all $i \in [n]$. Hence, to the left of position i in $\mathbf{e}(\mathbf{b})$, there are $(s_i - i)$ elements x such that $x < e_i$.

Naturally, the same argument can be used for any sequence of starting points, and so we see that, for every given sequence $\mathbf{s}$ of starting points, the unique non-crossing perfect matching has exactly $s_i - i$ elements x such that $x < e_i$, for all $i \in \{1, 2, 3, \ldots, n\}$. Hence, the number of *potential* crossings belonging to chord i equals $i - 1 - (s_i - i) = 2i - s_i - 1$. Naturally, to *not* cross any other chord is also an option. Consequently, for cr_i, exactly $2i - s_i$ distinct values $cr_i \in \{0, 1, 2, \ldots, 2i - s_i - 1\}$ are possible.

We have already seen that the number of perfect matchings $\Omega \in \mathrm{PM}(\mathbf{a})$ is equal to the number of legal permutations of $\mathbf{e}(\Omega)$. We can use a similar approach as above to find the number of such permutations. When $s_i = i$ for all $i \in \{1, 2, 3, \ldots, n\}$, we have that the number of alternative ending points for s_i is i, see Fig. 43, left. Whenever $s_i > i$, the number of legal alternatives for s_i becomes $s_i - i$ less, see Fig. 43, middle and right, and the number of alternative ending points is thus $i - (s_i - i) = 2i - s_i$.

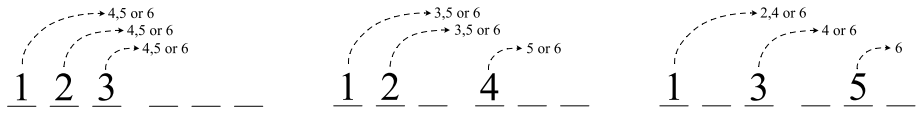


Figure 43: Left: $s_i = i$. Middle: $s_3 = 4$. Right: $s_2 = 3$ and $s_3 = 5$.

We see that the number of distinct choices to match chord i to some ending point equals the number of distinct values at cr_i. It is also clear that any such choice we make for chord i, generates a distinct number of crossings at chord i, since we must pass a distinct number of chords in each switch, see Fig. 42. According to the multiplication principle, the number of perfect matchings $\Omega \in \mathrm{PM}(\mathbf{a})$ is the product of the choices for all chords $i \in \{1, 2, 3, \ldots, n\}$. Each such combination corresponds to exactly one distinct sequence $\mathbf{cr}(\Omega)$, such that $cr_i \in \{0, 1, 2, \ldots, 2i - s_i - 1\}$ for all $i \in \{1, 2, 3, \ldots, n\}$. Hence, each sequence $\mathbf{cr}(\Omega)$ *uniquely* defines one perfect matching $\Omega \in \mathrm{PM}(\mathbf{a})$.

□

Theorem 29. *Let* $\mathrm{PM}_{(\mathbf{a})}$ *be the set of distinct perfect matchings corresponding to the area sequence* $\mathbf{a}$*, and let* $|\mathrm{PM}_{(\mathbf{a})}|_q$ *be the* q*-analog of* $\mathrm{PM}_{(\mathbf{a})}$*, with respect to the number of crossings* $(cr_{(\Omega)})$ *in each perfect matching* $\Omega \in \mathrm{PM}_{(\mathbf{a})}$*. Then*

$$|\mathrm{PM}_{(\mathbf{a})}|_q = \sum_{\Omega \in \mathrm{PM}(\mathbf{a})} q^{cr(\Omega)} = \prod_{i=1}^{n} [a_i + 1]_q \,, \tag{23}$$

for any area sequence $\mathbf{a}$*.*

Proof. We have defined $|\mathrm{PM}_{(\mathbf{a})}|_q$ as the q-analog on $\mathrm{PM}_{(\mathbf{a})}$, with respect to the number of crossings in each perfect matching $\Omega \in \mathrm{PM}_{(\mathbf{a})}$. This is exactly

$$\sum_{\Omega \in \mathrm{PM}(\mathbf{a})} q^{cr(\Omega)} \,.$$

In the proof of Lemma 28, we saw that the number of perfect matchings $\Omega \in \mathrm{PM}_{(\mathbf{a})}$ equals $\prod_{i=1}^{n}(2i - s_i)$, and that each perfect matching corresponds to exactly one distinct sequence $\mathbf{cr}_{(\Omega)}$, such that $cr_i \in \{0, 1, 2, \ldots, 2i - s_i - 1\}$, for all $i \in \{1, 2, 3, \ldots, n\}$. Since we know that, for any given area sequence $\mathbf{a}$, the sequence $\mathbf{s}_{(\Omega(\mathbf{a}))}$ of starting points equals the sequence $\mathbf{s}_{(\mathbf{a})}$ indexing the up-steps in $\mathrm{DP}_{\mathbf{a}}$ (see aftermath of Lemma 24), we can use the result from Lemma 7, which states that $2i - s_i = a_i + 1$.

As we saw in (14), when our statistic *crossings* is the result of a series of n independent choices, such that the i:th choice has the exact $a_i + 1$ alternatives of adding $\{0, 1, 2, \ldots, a_i\}$ crossings, the q-analog is given by $\prod_{i=1}^{n}[a_i + 1]_q$. □

Corollary 30. *For any area sequence* $\mathbf{a}$*,*

$$|\mathrm{PM}_{(\mathbf{a})}|_q = |\mathrm{RP}_{(\mathbf{a})}|_q = |\mathrm{AO}_{(\mathbf{a})}|_q \,. \tag{24}$$

Proof. A given area sequence $\mathbf{a}$ *uniquely determines* the q-analog $\prod_{i=1}^{n}[a_i + 1]_q$. The corollary follows from Theorem 14, Theorem 18 and Theorem 29. □

Example 31. Consider the set of perfect matchings $\mathrm{PM}_{(\mathbf{a})}$ on the unique set-up of starting points $\mathrm{PMSP}_{(\mathbf{a})}$, corresponding to the area sequence $\mathbf{a} = (0, 1, 1)$, as in Fig. 44, left. Then the number of perfect matchings $\Omega \in \mathrm{PM}_{(\mathbf{a})}$ is

$$\prod_{i=1}^{n}(a_i + 1) = 1 \cdot 2 \cdot 2 = 4. \tag{25}$$

We refine the set cardinality by q-counting this set with respect to the number of crossings in each perfect matching. Using (23) we have that

$$|\mathrm{PM}_{(\mathbf{a})}|_q = \prod_{i=1}^{n}[a_i + 1]_q = (q^0)(q^0 + q^1)(q^0 + q^1) = 1q^0 + 2q^1 + 1q^2. \tag{26}$$

Hence, in this set we have 1 Ω with no crossings, 2 Ω with one crossing and 1 Ω with two crossings. We encourage the reader to verify this result (or just take a look at Fig. 41). For $q = 1$, the sum of (26) equals the cardinality of the set, which is 4.

According to Corollary 30, this refinement mirrors the refinement on the set of acyclic orientations $\mathrm{AO}_{(\mathbf{a})}$ on the unit interval graph $\mathrm{UIG}_{\mathbf{a}}$, with respect to the number of ascending edges in each orientation $\vartheta \in \mathrm{AO}_{(\mathbf{a})}$, and the refinement on the set of rook placements $\mathrm{RP}_{(\mathbf{a})}$ on the Ferrers board $\mathrm{FB}_{\mathbf{a}}$, with respect to the number of inversions in each rook placement $\Psi \in \mathrm{RP}_{(\mathbf{a})}$, see Fig. 44.

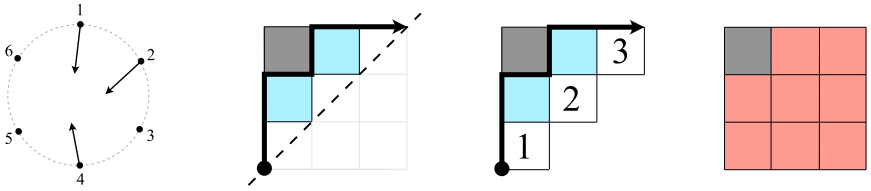


Figure 44: From left to right: 1) The set-up of starting points $\mathrm{PMSP}_{(\mathbf{a})}$. 2) The corresponding area sequence $\mathbf{a} = (0, 1, 1)$. 3) The unit interval graph $\mathrm{UIG}_{\mathbf{a}}$. 4) The Ferrers board $\mathrm{FB}_{\mathbf{a}}$.

5 Conclusions and Possible Applications

The main conclusion of this thesis is summarised in the following theorem.

Theorem 32. *Let* $\mathrm{AO}(\mathbf{a})$ *be the set of distinct acyclic orientations,* $\mathrm{RP}(\mathbf{a})$ *the set of distinct rook placements, and* $\mathrm{PM}(\mathbf{a})$ *the set of distinct perfect matchings, all corresponding to the same area sequence* $\mathbf{a} = (a_1, a_2, a_3, \ldots, a_n)$*. Then the q-analog of acyclic orientations* $\vartheta \in \mathrm{AO}(\mathbf{a})$*, using ascending edges as statistic, coincides with the q-analog of rook placements* $\Psi \in \mathrm{RP}(\mathbf{a})$*, using inversions as statistic, which coincides with the q-analog of perfect matchings* $\Omega \in \mathrm{AO}(\mathbf{a})$*, using crossings as statistic. Hence, for any area sequence* $\mathbf{a}$*,*

$$\begin{aligned} |\mathrm{AO}(\mathbf{a})|_q &= \sum_{\vartheta \in \mathrm{AO}(\mathbf{a})} q^{asc(\vartheta)} \\ = |\mathrm{RP}(\mathbf{a})|_q &= \sum_{\Psi \in \mathrm{RP}(\mathbf{a})} q^{inv(\Psi)} \\ = |\mathrm{PM}(\mathbf{a})|_q &= \sum_{\Omega \in \mathrm{PM}(\mathbf{a})} q^{cr(\Omega)} = \prod_{i=1}^{n} [a_i + 1]_q \,. \end{aligned} \tag{27}$$

We also note that

$$\prod_{i=1}^{n} [a_i + 1]_q = \prod_{i=1}^{n} [b_i + 1]_q \,, \tag{28}$$

where a_i *is the* row *area at row* i *and* b_i *the* column *area at column* i*.*

Proof. Theorem 32 summarises Theorem 14, Theorem 18 and Theorem 29. □

Theorem 32 states that the cardinality of $\mathrm{AO}(\mathbf{a})$, $\mathrm{RP}(\mathbf{a})$ and $\mathrm{PM}(\mathbf{a})$ coincide, for any area sequence $\mathbf{a}$. More importantly, the theorem also provides a *refinement* of this statement, with respect to our statistic. The refinement provides a strong indication of a more general bijective correspondence between these three C-Indexed families. As mentioned in the introduction, such bijections might be of interest for solving some of the classic open problems within the field of algebraic combinatorics.

One such problem is the *Stanley–Stembridge Conjecture* from 1993, [SS93], [Sta95], which was generalised by Shareshian–Wachs in 2012, [SW12], [SW16]. Briefly summarised, the *Shareshian–Wachs Conjecture* states that there should be a combinatorial formula on the form

$$\sum_{\theta \in AO(a)} q^{asc(\theta)} e_{\mu(\theta)} \,, \tag{29}$$

for a certain type of symmetric function. Note that in (29), e is an elementary symmetric function, and not the mathematical constant e. Setting $q = 1$ gives the Stanley–Stembridge Conjecture. The problem is that, in (29), μ is not known, and the problem of finding μ is evidently very difficult. Although a similar conjecture was recently covered in [AP18] and [GHQR19], and later solved in [AS20], the Shareshian–Wachs Conjecture remains unsolved.

In the Stanley–Stembridge / Shareshian–Wachs Conjecture, the core of the problem relates to acyclic orientations. However, in the light of Theorem 32, it might be more fruitful to consider a corresponding family instead. Indeed, the identification of certain properties in a perfect matching might be far more intuitive than within the corresponding acyclic orientation, and vice versa. For example, in an acyclic orientation we can easily identify *sinks*, i.e. vertices where the out-degree is 0. However, in a rook placement, we do not know what sinks correspond to. This means that properties that remain hidden in one family, might be found in a corresponding family.

The bijective correspondence described in this thesis is limited to three families, using one statistic per family. However, there are indications of correspondence to other families as well. We trust that these results could be a part of a larger puzzle, covering multiple statistics within a wide range of families. If so, they could be important for advances in several open problems of the field.

References

[And87] Désiré André. Solution directe du probleme résolu par M. Bertrand. *CR Acad. Sci. Paris*, 105(436-437):7, 1887.

[AP18] Per Alexandersson and Greta Panova. LLT polynomials, chromatic quasisymmetric functions and graphs with cycles. *Discrete Mathematics*, 341(12):3453–3482, December 2018. doi:10.1016/j.disc.2018.09.001.

[AS20] Per Alexandersson and Robin Sulzgruber. A combinatorial expansion of vertical-strip LLT polynomials in the basis of elementary symmetric functions. *ArXiv e-prints*, 2020. arXiv:2004.09198.

[Che08] Y.-M. Chen. The Chung–Feller theorem revisited. *Discrete Mathematics*, 308(7):1328–1329, April 2008. doi:10.1016/j.disc.2007.03.068.

[CWF49] K.L. Chung and W-Feller. On fluctuations in coin-tossing. *Proc. Natl. Acad. Sci. USA*, 35:605–608, 1949.

[Dav10] Tom Davis. Catalan numbers. dec, 2010.

[GHQR19] Adriano M. Garsia, James Haglund, Dun Qiu, and Marino Romero. e-positivity results and conjectures. *ArXiv e-prints*, 2019. arXiv:1904.07912.

[SS93] Richard P. Stanley and John R. Stembridge. On immanants of Jacobi–Trudi matrices and permutations with restricted position. *Journal of Combinatorial Theory, Series A*, 62(2):261–279, March 1993. doi:10.1016/0097-3165(93)90048-d.

[Sta95] Richard P. Stanley. A symmetric function generalization of the chromatic polynomial of a graph. *Advances in Mathematics*, 111(1):166–194, 1995. doi:10.1006/aima.1995.1020.

[Sta11] Richard P. Stanley. *Enumerative Combinatorics: Volume 1*. Cambridge University Press, second edition, 2011. doi:10.1017/CBO9781139058520.

[Sta15] Richard P. Stanley. *Catalan Numbers*. Cambridge University Press, 2015. doi:10.1017/CBO9781139871495.

[SW12] John Shareshian and Michelle L. Wachs. Chromatic quasisymmetric functions and Hessenberg varieties. In *Configuration Spaces*, pages 433–460. Scuola Normale Superiore, 2012. doi:10.1007/978-88-7642-431-1_20.

[SW16] John Shareshian and Michelle L. Wachs. Chromatic quasisymmetric functions. *Advances in Mathematics*, 295(4):497–551, June 2016. doi:10.1016/j.aim.2015.12.018.