Bijections on Catalan Structures

OFIR AMMAR
Bijections on Catalan Structures

OFIR AMMAR

Master’s Thesis in Mathematics (30 ECTS credits)
Master Programme in Mathematics (120 credits)
Royal Institute of Technology year 2015
Supervisor at KTH was Svante Linusson
Examiner was Svante Linusson

TRITA-MAT-E 2015:34
ISRN-KTH/MAT/E--15/34--SE

Royal Institute of Technology
School of Engineering Sciences

KTH SCI
SE-100 44 Stockholm, Sweden

URL: www.kth.se/sci
"We are usually convinced more easily by reasons we have found ourselves than by those which have occurred to others."

Blaise Pascal, Pensées (1670)
An open problem introduced by J. Haglund was to find a bijective proof over Dyck paths that would interchange two of its statistics. This problem was known to be The Symmetry Problem of the q,t-Catalan polynomial and was proven by other means to be true. This project is an attempt to find a bijection, where we provide the bijection’s behaviour under certain constrains. Then, we introduce an attempt to translate the problem from Dyck paths to other combinatorial structures. Finally we try to solve a related conjecture, called The Symmetry Problem of parking functions, which generalizes the previous problem. Some results we obtained from The Symmetry Problem of parking functions helped us characterize part of a bijective proof for Dyck paths.
Acknowledgements

This paper constitutes a master thesis of 30 ECTS in the field of mathematics at the Department of Mathematics of KTH Royal Institute of Technology. This project is the very last step for acquiring the joint master degree offered by both KTH and Stockholm University.

This project would not have been possible without the generous help offered by my supervisor Prof. Svante Linusson. The questions, comments, suggestions and criticism I received from him have been of grate value. I want to express my vast gratitude to Prof. Linusson for his investment, guidance and contribution to this work.

Moreover, I owe my loving gratitude to my family, friends and colleges. Thank you for your encouragement and unconditional support.
## Contents

### Abstract

### Acknowledgements

### Contents

#### 1 Dyck Paths and The Symmetry Problem

1.1 Dyck Paths .............................................................. 1
   1.1.1 Statistics for Dyck Paths ........................................ 2
      1.1.1.1 The Area Statistic ...................................... 2
      1.1.1.2 The Dinv Statistic .................................... 3
      1.1.1.3 The Bounce Statistic .................................. 5

1.2 The Symmetry Problem .................................................. 6
   1.2.1 The bijection \( f \) for paths \( \pi \in L_{n,n+} \) with \( \text{area}(\pi) = 1 \) ........................................ 7
   1.2.2 The bijection \( f \) for paths \( \pi \in L_{n,n+} \) with \( \text{area}(\pi) = 2 \) ........................................ 10

#### 2 Representations of other Catalan Structures

2.1 Motivation ............................................................. 16

2.2 Murasaki Diagrams ..................................................... 18
   2.2.1 Catalan Structure ............................................. 18
   2.2.2 Representation of the \( \text{area}() \) statistic ...................... 19
   2.2.3 Representation of the \( \text{dinv}() \) statistic ..................... 20

2.3 Plane Trees ............................................................ 20
   2.3.1 Catalan Structure ............................................. 20
   2.3.2 Representation of \( \text{area}() \) statistic ...................... 22
   2.3.3 Representation of \( \text{dinv}() \) statistic ..................... 22

2.4 Complete Binary Trees ............................................... 22
   2.4.1 Catalan Structure ............................................. 22
   2.4.2 Representation of \( \text{area}() \) statistic ...................... 23
   2.4.3 Representation of \( \text{dinv}() \) statistic ..................... 23

2.5 Triangulations of \((n+2)\)-gons ...................................... 24
   2.5.1 Catalan Structure ............................................. 24

2.6 Non crossing matching of \(2n\) nodes on a circle ..................... 25
   2.6.1 Catalan Structure ............................................. 26

2.7 Further Representations .............................................. 26
3 Permutations ................................................. 27
  3.1 Motivation .............................................. 27
  3.2 Statistics of Permutations .............................. 28
    3.2.1 Inversions ........................................ 28
    3.2.2 Major Index ...................................... 28
    3.2.3 Pseudo-Mahonian Statistics ....................... 28
  3.3 312-avoiding Permutations ............................ 30
  3.4 231-avoiding Permutations ............................ 32
  3.5 Further approaches ..................................... 33

4 Parking Functions .......................................... 34
  4.1 Motivation .............................................. 34
  4.2 Definition .............................................. 35
  4.3 Statistics for Parking Functions ...................... 36
    4.3.1 Area .............................................. 36
    4.3.2 Dinv .............................................. 36
    4.3.3 Pmaj .............................................. 37
  4.4 The Symmetry Problem for Parking Functions ........... 37
  4.5 Parking Functions with Dinv = 0 ....................... 38
  4.6 Bijection for $p \in P_n$ with $area(p) = 0$ ............ 41
  4.7 Bijection for $p \in P_n$ with $dinv(p) + area(p) = \binom{n}{2}$ . 44
  4.8 Bijection for $p \in P_n$ with $dinv(p) + area(p) = \binom{n}{2} - 1$ . 47

5 Conclusions .................................................. 53

A INV to bounce map for $n = 4$ ................................ 54

B $\phi : S_4(231) \to L_{4,4}^+$ table .......................... 55

C $\sum_{\text{occupant}(1)=1} q^{dinv(p) + area(p)}$ is not symmetric .... 56

Bibliography .................................................... 57
Chapter 1

Dyck Paths and The Symmetry Problem

1.1 Dyck Paths

Consider a rectangular lattice of size $m \times n$ consists of all points in $(x, y) \in \mathbb{Z}^2$ such that $0 \leq x \leq n$ and $0 \leq y \leq m$. A lattice path is a sequence of North $N(0, 1)$ and East $E(1, 0)$ steps beginning at the origin $(0, 0)$ and ending at $(m, n)$. We denote by $L_{m,n}$ the set of all lattice paths beginning at $(0,0)$ and ending at $(m,n)$.

**Proposition 1.1.** $|\{ \pi : \pi \in L_{m,n} \}| = \binom{m+n}{m}$

**Proof.** All paths $\pi \in L_{m,n}$ must connect $(0,0)$ and $(m,n)$ using North and East steps. The path contains precisely $m+n$ steps, out of which we can choose which ones would be the $m$-Eastern ones. This choice would force the rest of the steps to be North steps. Thus, $|\{ \pi \in L_{m,n} \}| = \binom{m+n}{m} = \binom{m+n}{n}$.

\[\square\]

A square lattice is a lattice whose vertical and horizontal dimensions are similar ($m = n$). In the following, we will consider only square lattices (which we might call square grids), unless stated otherwise.

**Definition 1.2.** A Dyck Path is a lattice path on a square lattice $\pi \in L_{m,n}$, such that the path of $\pi$ starts with $N$-step and never crosses the diagonal line that connects $(0,0)$ with $(n,n)$. The path $\pi$ may indeed intersect the line $y = x$ in several points, but never go below it.
Chapter 1. Dyck Paths and The Symmetry Problem

2

area(\pi) = 3 \quad \text{dinv}(\pi) = 0
area(\pi) = 2 \quad \text{dinv}(\pi) = 1
area(\pi) = 1 \quad \text{dinv}(\pi) = 2
area(\pi) = 0 \quad \text{dinv}(\pi) = 3

FIGURE 1.1: All Dyck paths in $L_{3,3}^+$, $|\{\pi : \pi \in L_{3,3}^+\}| = \frac{1}{4} \binom{6}{3} = 5$

Let $L_{n,n}^+$ denote the set of all such Dyck paths. The following proposition is a well-known fact [8].

Proposition 1.3. $|\{\pi : \pi \in L_{n,n}^+\}| = \frac{1}{n+1} \binom{2n}{n}$

The number of elements of $L_{n,n}^+$ that was given by the formula above is also called Catalan number. The $n$th Catalan number is usually denoted by $C_n$, and counted by directly by the formula $C_n = \frac{1}{n+1} \binom{2n}{n}$. It is useful to let $C_0 = C_1 = 0$. An equivalent way to compute the value of $C_n$ could be given recursively by the formula $C_n = \sum_{k=0}^{n-1} C_k C_{n-1-k}$.

The initial value of the sequence of Catalan numbers is 1, 1, 2, 5, 14, 42, 132, 429, 1430, 4862, 16796, 58786, 208012 . . .

1.1.1 Statistics for Dyck Paths

A statistic is an element of $\mathbb{N}$ that is assigned to each element of a certain set. In the following we will present three statistics that are assigned to elements of the set $L_{n,n}^+$.

1.1.1.1 The Area Statistic

Let $\pi \in L_{n,n}^+$. By $\text{area}(\pi)$ we mean the area statistic, which is the sum of all complete unit squares that are located between the path of $\pi$ and the diagonal line $y = x$, where we disregard all partial (half) squares in our counting.

Given $\pi \in L_{n,n}^+$, we will denote by $a_k$ the area in complete squares the are located between the path $\pi$ and the diagonal $y = x$ in the $k$th row, where the bottom row is the first one.

Notice that $a_1 = 0$ always, and for all $1 \leq k \leq n$, $0 \leq a_k \leq k - 1$. The area vector of $\pi$ is denoted by $\vec{\text{area}}(\pi)$ and is equal to $\vec{\text{area}}(\pi) = (a_1, a_2, \ldots, a_n)$. The reader should convince herself that given an area vector of a Dyck path, she could uniquely draw the path on the grid. So there is a 1-to-1 correspondence between $\vec{\text{area}}(\pi)$ and the picture of $\pi$ on the $n \times n$ grid. The values of $\text{area}(\pi)$ for $\pi \in L_{3,3}^+$ are presented in figure 1.1.

Obviously,

$$\text{area}(\pi) = \sum_{k=1}^{n} a_k$$
Proposition 1.4. For any \( n \), there is exactly one \( \pi \in L_{n,n}^+ \) such that \( \text{area}(\pi) = 0 \). That path \( \pi \) is exactly \( \overline{\text{are}a}(\pi) = (0,0,\ldots,0) \).

Proof. First see that if \( \pi \in L_{n,n}^+ \) such that \( \overline{\text{are}a}(\pi) = (0,0,\ldots,0) \) then we can compute \( \text{area}(\pi) = \sum_{i=1}^{n} a_i = \sum_{i=1}^{n} 0 = 0 \). Now consider any \( \pi' \in L_{n,n}^+ \) such that \( \pi' \neq \pi \), then we can compute \( \text{area}(\pi') = \sum_{i=1}^{n} a_i \geq 1 \) since all \( a_i \geq 0 \) and at least one \( j \) satisfied \( a_j \geq 1 \).

\[ \square \]

Proposition 1.5. For any \( n \), there is exactly one \( \pi \in L_{n,n}^+ \) such that \( \text{area}(\pi) = \binom{n}{2} \). That path \( \pi \) is exactly \( \overline{\text{are}a}(\pi) = (0,1,2,\ldots,n-1) \).

Proof. For any \( 1 \leq i \leq n \) we have \( a_i \leq i-1 \), so if a given Dyck path \( \pi \in L_{n,n}^+ \) satisfies \( \overline{\text{are}a}(\pi) = (0,1,2,\ldots,n-1) \) it means that \( \pi \) has the maximal \( \text{area} \) possible. When we sum \( \sum_{i=1}^{n} a_i = 0+1+2+\ldots+(n-1) = \binom{n}{2} \). Any other path \( \pi' \in L_{n,n}^+ \) must have \( \text{area}(\pi') < \binom{n}{2} \).

\[ \square \]

Figure 1.2: The two polar paths \( \pi_1, \pi_2 \in L_{8,8}^+ \). \( \pi_1 \) on the left satisfies that the \( \text{area}(\pi_1) = \binom{8}{2} = 28 \), while \( \pi_2 \) on the right satisfies \( \text{area}(\pi_2) = 0 \).

Paths of the shape of \( \pi_1 \) from figure 1.2 will be called trivial paths (of size \( n \)), and paths similar to \( \pi_2 \) from figure 1.2 are called staircase paths (of size \( n \)).

Remark 1.6. The reader should convince herself that there are unique paths not only to paths \( \pi \in L_{n,n}^+ \) that satisfy \( \text{area}(\pi) = 0 \) or \( \text{area}(\pi) = \binom{n}{2} \), but also to \( \text{area}(\pi) = \binom{n}{2} - 1 \). Moreover, when we consider the elements of the area vector of a path \( \pi \in L_{n,n}^+ \), then for any row \( 2 \leq k \leq n \), \( 0 \leq a_k \leq a_{k-1} + 1 \leq k - 1 \).

1.1.1.2 The Dinv Statistic

Let \( \pi \in L_{n,n}^+ \) be a path with \( \overline{\text{are}a}(\pi) = (a_1,a_2,\ldots,a_n) \). By the Diagonal Inversion of \( \pi \), or shortly \( \text{dinv}(\pi) \), we denote the number of ordered pairs \((i,j)\), such that \( i < j \) and \( a_i = a_j \) or \( a_i = a_j + 1 \). The \( \text{dinv} \) statistic was discovered by M. Haiman, but first appeared
Lemma 1.7. Let $\pi \in L^*_{n,n}$ with $\text{dinv}(\pi) > 0$. If $(i, j)$ is the pair with the minimal $j$ such that the pair $(i, j)$ contributes to $\text{dinv}(\pi)$, then $a_k = k - 1$ for all $1 \leq k < j$.

Proof. We know that $a_1 = 0$ always, so if $j = 2$ and $a_j = 0$ then it is trivial. Let $3 \leq j$ and assume for contradiction that there is a minimal $k$, such that $2 \leq k < j$, and $a_k \neq k - 1$. Then, by the pigeonhole principle all of $a_1, \ldots, a_k$ get values from $\{0, 1, \ldots, k-2\}$ there is $k' < k$ such that $a_k = a_{k'}$, and hence the pair $(k', k)$ contributes to $\text{dinv}(\pi)$, contradicting the minimality of $j$.

Lemma 1.8. Given $\pi \in L^*_{n,n}$ If $a_i > a_{i+1}$ then there are two pairs $(j, i + 1)$ and $(k, i + 1)$ that contributes to $\text{dinv}(\pi)$.

Proof. For $\pi \in L^*_{n,n}$ assume $a_i > a_{i+1}$. Since for all $2 \leq k \leq i$ $a_k \leq a_{k-1} + 1$, then all area values $0, 1, 2, \ldots, a_i$ must appear at least once between the first row to the ith row. Necessarily, if $a_{i+1} < a_i$ there must be $j, k < i$ such that $a_k = a_{i+1}$ and $a_j = a_{i+1} + 1$ (as $a_{i+1} + 1 \leq a_i$), so $(j, i + 1)$ and $(k, i + 1)$ are counted in $\text{dinv}(\pi)$.

Proposition 1.9. For all $n$ and for all $\pi \in L^*_{n,n}$, $0 \leq \text{dinv}(\pi) \leq \binom{n}{2}$. Also, for all $n$ the only $\pi_1 \in L^*_{n,n}$ such that $\text{dinv}(\pi_1) = 0$ is $\overrightarrow{\text{area}}(\pi_1) = (0, 1, 2, \ldots, n-1)$, the only path $\pi_2 \in L^*_{n,n}$ such that $\text{dinv}(\pi_2) = \binom{n}{2}$ is $\overrightarrow{\text{area}}(\pi_2) = (0, 0, \ldots, 0)$.

Proof. Let $\pi \in L^*_{n,n}$. Since $\text{dinv}(\pi)$ counts the number of pairs $(i, j)$ with $i < j$ that would satisfy a certain condition, then obviously the minimal amount of pairs that would satisfy those conditions is 0, and the maximal value is given we pick all possible pairs of $(i, j)$ where $q \leq i < j \leq n$, which combinatorially corresponds to $\binom{n}{2}$.

Now consider $\pi_1 \in L^*_{n,n}$ with $\text{dinv}(\pi_1) = 0$. We will now explain why it must be the case that $\overrightarrow{\text{area}}(\pi_1) = (a_1, a_2, a_3, \ldots, a_n) = (0, 1, 2, \ldots, n-1)$ which will prove that $\pi_1$ is unique. We know that $a_1 = 0$. Now since $0 \leq a_2 \leq 1$, we get that either $a_2 = 0$ or $a_2 = 1$. If $a_2 = 0$, then $\text{dinv}(\pi_1) \geq 1$, since the pair $(1, 2)$ would be counted. So $a_2 = 1$. Inductively, for any $3 \leq k \leq n$, we know that $0 \leq a_k \leq k - 1$, and if $a_k \neq k - 1$ then $a_k = i < k - 1$ which would mean that the row $(i + 1, k)$ would be counted in $\text{dinv}(\pi_1)$ so $\text{dinv}(\pi_1) \geq 1$ (which contradicts $\text{dinv}(\pi_1) = 0$). Thus, $a_k = k - 1$ for all such $0 \leq k \leq n - 1$, and $\overrightarrow{\text{area}}(\pi_1) = (0, 1, 2, \ldots, n-1)$.
Let $\pi_2 \in L_{n,n}^+$ be a Dyck path such that $\operatorname{dinv}(\pi_2) = \binom{n}{2}$. We want to show that it must be the case that $\operatorname{area}(\pi_2) = (a_1, a_2, \ldots, a_n) = (0, 0, \ldots, 0)$. We know that $\operatorname{dinv}(\pi_2) = \binom{n}{2}$ means that every pair $(i, j)$ must be included in the counting of $\operatorname{dinv}(\pi_2)$ for $i < j$. But since $a_1 = 0$, and since $(1, j)$ would be counted for all $2 \leq j \leq n$, then it means that $a_j = 0$ (since it is impossible to have $a_j = -1$) for all such $j$. Hence, $\operatorname{area}(\pi_2) = (0, 0, \ldots, 0).

\[ \square \]

Remark 1.10. The previous proposition shows us that for any given $n \in \mathbb{N}$ we can uniquely find two Dyck paths $\pi_1, \pi_2 \in L_{n,n}^+$ such that $(\operatorname{area}(\pi_1), \operatorname{dinv}(\pi_1)) = (0, \binom{n}{2})$ and $(\operatorname{area}(\pi_2), \operatorname{dinv}(\pi_2)) = ((\binom{n}{2}), 0)$.

1.1.1.3 The Bounce Statistic

Let $\pi \in L_{n,n}^+$. We define the bounce path of $\pi$ to be the path $\pi'$ which is obtained from $\pi$ by starting at the origin $(0, 0)$ and traveling $N$ until we hit the first $E$ step of $\pi$, and then $\pi'$ continue going $E$ until it hits the diagonal $y = x$. Once $\pi'$ hits the diagonal $y = x$, it starts going $N$ until it hits an $E$ step of $\pi$, then $\pi'$ turns $E$ until it hits the diagonal $y = x$ again. This procedure repeats itself until $\pi'$ reaches the point $(n, n)$.

Now let $(0, 0), (b_1, b_1), (b_2, b_2), \ldots, (b_k, b_k), (n, n)$ be the points where the bounce path $\pi'$ intersects the diagonal $y = x$. Then we define the bounce statistic to be

$$\operatorname{bounce}(\pi) = \operatorname{bounce}(\pi') = \sum_{i=1}^{k} (n - b_i)$$

\[ \text{Figure 1.3: The path } \pi \in L_{8,8}^+ \text{ (thick black) and the bounce path } \pi' \in L_{8,8}^+ \text{ (dashed).} \]

The path $\pi'$ intersects $y = x$ at $(0, 0), (1, 1), (3, 3), (5, 5)$ and $(8, 8)$, thus $\operatorname{bounce}(\pi) = (8 - 1) + (8 - 3) + (8 - 5) = 15$.

**Proposition 1.11.** For any $n \in \mathbb{N}$ and any $\pi \in L_{n,n}^+$, we get $0 \leq \operatorname{bounce}(\pi) \leq \binom{n}{2}$. Also, for all $n$ the only path $\pi_1 \in L_{n,n}^+$ such that $\operatorname{bounce}(\pi_1) = 0$ is $\operatorname{area}(\pi_1) = (0, 1, 2, \ldots, n-1)$, and the only path $\pi_2 \in L_{n,n}^+$ such that $\operatorname{bounce}(\pi_2) = \binom{n}{2}$ is $\operatorname{area}(\pi_2) = (0, 0, \ldots, 0)$. 
Proof. Consider the two polar paths \(\pi_1, \pi_2\) from figure 1.2. \(\text{area}(\pi_1) = (0, 1, 2, \ldots, n-1)\) and \(\text{area}(\pi_2) = (0, 0, \ldots, 0)\). We can see that these are the two extreme cases of the bounce statistic (since the bounce path \(\pi'_1\) of \(\pi_1\) doesn’t have any points intersecting the line \(y = x\), while the bounce path \(\pi'_2\) of \(\pi_2\) has all possible points). We count and see \(\text{bounce}(\pi_1) = 0\), and \(\text{bounce}(\pi_2) = (n-1) + (n-2) + \ldots + 1 = \binom{n}{2}\).

Obviously for any \(n \in \mathbb{N}\) there is exactly one \(\pi_1 \in L_{n,n}\) such that \(\text{bounce}(\pi_1) = 0\). This is because in order to get \(\text{bounce}(\pi_1) = 0\), we need that the first \(E\) step of \(\pi_1\) would be after \(n-N\) steps, otherwise we will have a bounce point \((b_i, b_i)\) somewhere and \(\text{bounce}(\pi_1) > 0\).

Also, for any \(n \in \mathbb{N}\) there is exactly one path \(\pi_2 \in L_{n,n}\) such that \(\text{bounce}(\pi_2) = \binom{n}{2}\). This is because in order to get \(\text{bounce}(\pi_2) = \binom{n}{2}\), we need that the bounce path \(\pi'_2\) of \(\pi_2\) would intersect the diagonal in all the points \((0, 0), (1, 1), (2, 2), \ldots, (n, n)\). But since in the case of \(\pi_2\) we get that \(\pi_2 = \pi'_2\), there could be no more paths between \(\pi_2\) and \(\pi'_2\), so \(\pi_2\) is the only path with the bounce path of \(\pi'_2\).

\[\Box\]

1.2 The Symmetry Problem

Using the same notation as Haglund [4, p. 41], let us define the polynomial

\[F_n(q, t) = \sum_{\pi \in L_{n,n}} q^\text{area}(\pi) t^\text{bounce}(\pi)\]

The Symmetry Problem deals with finding a bijective proof in order to show that \(F_n(q, t)\) is a symmetric function with respect to the variables \(q\) and \(t\). In other words, the problem would be to show that:

**Theorem 1.12. The Symmetry Problem**

\[F_n(q, t) = \sum_{\pi \in L_{n,n}} q^\text{area}(\pi) t^\text{bounce}(\pi) = \sum_{\pi \in L_{n,n}} q^\text{bounce}(\pi) t^\text{area}(\pi) = F_n(t, q)\]

As for the moment there is no combinatorial proof to Theorem 1.12. However, Haglund proved Theorem 1.12 with advanced algebraic tools, so we know that such a bijection must exist [4, pp. 48-49].

Haglund provided as well a bijective proof, where he used the map \(\zeta\) [4, p. 41], to shows that

\[\sum_{\pi \in L_{n,n}} q^\text{dinv}(\pi) t^\text{area}(\pi) = \sum_{\pi \in L_{n,n}} q^\text{area}(\pi) t^\text{bounce}(\pi)\]
Thus, if we set \( t = 1 \) in the last equation (and later, if we set \( q = 1 \)), we get the equidistribution:

\[
\sum_{\pi \in L^{+}_{n,n}} q^{\text{dinv}(\pi)} = \sum_{\pi \in L^{+}_{n,n}} q^{\text{area}(\pi)} = \sum_{\pi \in L^{+}_{n,n}} q^{\text{bounce}(\pi)}
\]

The bijections \( \zeta \) and \( \zeta^{-1} \) convert the symmetry problem with the problem of proving the equation:

\[
\sum_{\pi \in L^{+}_{n,n}} q^{\text{area}(\pi)} t^{\text{dinv}(\pi)} = \sum_{\pi \in L^{+}_{n,n}} q^{\text{dinv}(\pi)} t^{\text{area}(\pi)}
\]

by finding a bijection \( f \) between all ordered pairs of \((\text{area}, \text{dinv})\) of all Dyck paths of a fixed \( n \), to all the ordered pairs \((\text{dinv}, \text{area})\). The main goal of this project would be to find such \( f : L^{+}_{n,n} \rightarrow L^{+}_{n,n} \) such that \( \text{area}(\pi) = \text{dinv}(f(\pi)) \) and at the same time \( \text{dinv}(\pi) = \text{area}(f(\pi)) \).

Remark 1.13. Propositions 1.4, 1.5 and 1.9 show that such an \( f \) must pair the two polar paths described in Figure 1.2 together. This is because the paths described in Figure 1.2, the trivial one and the staircase one, are the only ones with values \((0, (\binom{n}{2}))\) and \((((\binom{n}{2}), 0))\) of \((\text{area}, \text{dinv})\), for any \( n \).

Finding the bijection \( f \) would solve Haglund’s open problem (3.11) [4, p. 49]. We could not find the function \( f \) itself, but in the following sections we will describe how must \( f \) match Dyck paths under certain constrains. The following two sections offer a description of how \( f \) must map those Dyck paths \( \pi \) with \( \text{area}(\pi) = 1 \) (See Section 1.2.1), and those Dyck paths \( \pi \) with \( \text{area}(\pi) = 2 \) (See Section 1.2.2). Moreover, Remark 4.20 in Chapter 4 offers the bijection for the case where \( \text{area}(\pi) + \text{dinv}(\pi) = \binom{n}{2} \).

### 1.2.1 The bijection \( f \) for paths \( \pi \in L^{+}_{n,n} \) with \( \text{area}(\pi) = 1 \)

In the following section we will provide a bijection

\[
f_{\text{area}=1} : \{ \pi \in L^{+}_{n,n} : \text{area}(\pi) = 1 \} \rightarrow \{ \pi \in L^{+}_{n,n} : \text{dinv}(\pi) = 1 \}
\]

that also satisfies \( \text{dinv}(\pi) = \text{area}(f_{\text{area}=1}(\pi)) \).

**Lemma 1.14.** There are exactly \( n - 1 \) different paths in \( L^{+}_{n,n} \) with \( \text{area}(\pi) = 1 \), and they are of the form \( \text{area}(\pi) = (0, 0, \ldots, 0, 1, 0, \ldots, 0) \) where \( a_{i} = 1 \) for exactly one \( 2 \leq i \leq n \) and all other \( k \neq i \) \( a_{k} = 0 \).
Proof. Consider $\pi \in L_{n,n}^+$ such that $\text{area}(\pi) = 1$. Since $\text{area}(\pi) = \sum_{k=1}^{n} a_k$ and all $0 \leq a_k$, then exactly one $a_i = 1$ while all other $a_j = 0$ for $j \neq i$. Since it is always the case that $a_1 = 0$, then we have $n - 1$ choices for such $a_i$.

Lemma 1.15. There is a unique path $\pi \in L_{n,n}^+$ such that $\text{area}(\pi) = 1$ and $\text{dinv}(\pi) = k$ for $(\frac{n-1}{2}) \leq k \leq (\frac{n}{2}) - 1$ and that path is $\overline{\text{area}}(\pi) = (0,0,\ldots,0,1,0,\ldots,0)$ where the 1 appears only at the $\left[\frac{n}{2} + 1 - k\right]$-st index of the area vector.

Proof. Consider $\pi_i \in L_{n,n}^+$ such that $\text{area}(\pi_i) = 1$ and $\overline{\text{area}}(\pi_i) = (0,\ldots,0,1,0,\ldots,0)$, where the value 1 appears only at the i-th position of the area vector. In that area vector we have $n - 1$ zeros, that add $(\frac{n-1}{2})$ to $\text{dinv}(\pi_i)$. In addition all $(i,j)$-pairs need to be counted for $i < j \leq n$ because $a_i = 1$ and for all $i < j$, $a_j = 0$. This adds $n - i$ to $\text{dinv}(\pi_i)$, which turns $\text{dinv}(\pi_i) = \left(\frac{n-1}{2}\right) + n - i$. Thus every $\pi_i$ have a unique $\text{dinv}$ value from the range $(\frac{n-1}{2})$ to $(\frac{n}{2}) - 1$.

Lemma 1.16. There are precisely $n - 1$ paths in $L_{n,n}^+$ with $\text{dinv}(\pi) = 1$.

Proof. Consider $\pi \in L_{n,n}^+$ such that $\text{dinv}(\pi) = 1$. This means that there is exactly one $(i,j)$ pair being counted, where $i < j$. If $a_i = a_j + 1$ then there must be at least one $k < i$ such that $a_k = a_j$ (by Lemma 1.8) which contradicts $\text{dinv}(\pi) = 1$. So it must be the case that $a_i = a_j$. If $j \neq i + 1$ then by Lemma 1.7 for all $j' < j$, $a_{j'} = j' - 1$, and more specifically $a_{i+1} = i = a_i + 1 = a_j + 1$ so both $(i,j)$ and $(i+1,j)$ are counted in $\text{dinv}(\pi)$ as a contradiction. So we get that $j = i + 1$ (so $(i,i+1)$ is the pair counted in $\text{dinv}(\pi)$). By Lemma 1.7, in order to have $\text{dinv}(\pi) = 1$ we must have as well that for all $k \neq i$ $a_k < a_{k+1}$. The index $i$ ranges between $1 \leq i \leq n - 1$ (since if $i = n$ then $i + 1$ would be a meaningless index).

Lemma 1.17. There is a unique path $\pi \in L_{n,n}^+$ such that $\text{dinv}(\pi) = 1$ and $\text{area}(\pi) = k$ for $(\frac{n-1}{2}) \leq k \leq (\frac{n}{2}) - 1$ and that path is $\overline{\text{area}}(\pi) = (0,1,2,\ldots,j-1,j-1,j,j+1,\ldots,n-1)$ where $j = \left[k - (\frac{n-1}{2}) + 1\right]$.

Proof. Consider $\pi'_i \in L_{n,n}^+$ such that $\text{dinv}(\pi'_i) = 1$. Since only $a_i = a_{i+1}$ for $1 \leq i \leq n - 1$, and for $k \neq i$, $a_k < a_{k+1}$, then we must have for all $k < i$ $a_k = k - 1$, and for all $i < k$ $a_k = k - 2$. This yields that $\overline{\text{area}}(\pi'_i) = (0,1,2,\ldots,i-1,i-1,i,\ldots,n-2)$, where $a_i = a_{i+1} = i - 1$. 

Chapter 1. Dyck Paths and The Symmetry Problem
Summing the elements of the area vector we get that \( \text{area}(\pi'_i) = 0 + 1 + 2 + \ldots + (n-2) + (i-1) = \binom{n-1}{2} + (i-1) \), which means \( \pi'_i \) has a unique area value for \( 1 \leq i \leq n-1 \).

\[ \square \]

**Definition 1.18.** Let \( \{ \pi \in L_{n,n}^+ : \text{area}(\pi) = 1 \} = \{ \pi_2, \pi_3, \ldots, \pi_{n-1} \} \), where \( \overline{\text{area}}(\pi_i) = (0,0,\ldots,0,1,0,\ldots,0) \) and \( a_i = 1 \) and \( \forall j \neq i a_j = 0 \). By the proof of Lemma 1.15 we have that \( \text{dinv}(\pi_i) = \binom{n-1}{2} + (n-i) \). Thanks to proof of Lemma 1.17 we are capable to denote the set \( \{ \pi \in L_{n,n}^+ : \text{dinv}(\pi) = 1 \} = \{ \pi'_1, \pi'_2, \ldots, \pi'_{n-1} \} \), where \( \overline{\text{area}}(\pi'_i) = (0,1,2,\ldots,i-1,i-1,i,\ldots,n-2) \) and \( \text{area}(\pi'_i) = \binom{n-1}{2} + (i-1) \). Now let us define,

\[
\text{f}_{\text{area}=1}(\pi_i) = \pi'_{n+1-i} \quad \text{for} \quad 2 \leq i \leq n-1
\]

**Theorem 1.19.** The map

\[
\text{f}_{\text{area}=1} : \{ \pi \in L_{n,n}^+ : \text{area}(\pi) = 1 \} \to \{ \pi \in L_{n,n}^+ : \text{dinv}(\pi) = 1 \}
\]

described above is a bijection that satisfies \( \text{dinv}(\pi) = \text{area}(\text{f}_{\text{area}=1}(\pi)) \).

**Proof.** According to Lemmata 1.14 and 1.16,

\[
|\{ \pi \in L_{n,n}^+ : \text{area}(\pi) = 1 \}| = |\{ \pi \in L_{n,n}^+ : \text{dinv}(\pi) = 1 \}| = n-1
\]

Which means that we need \( \text{f}_{\text{area}=1} \) to match the elements of those sets so that \( \text{area}(\pi) = \text{dinv}(\text{f}_{\text{area}=1}(\pi)) \) and \( \text{dinv}(\pi) = \text{area}(\text{f}_{\text{area}=1}(\pi)) \).

Notice that, \( (\text{area}(\pi_i), \text{dinv}(\pi_i)) = (1, \binom{n-1}{2} + n-i) = (\text{dinv}(\pi'_{n+1-i}), \text{area}(\pi'_{n+1-i})) \). Since \( \text{area}(\pi'_{n+1-i}) = \binom{n-1}{2} + (n+1-i) - 1 = \binom{n-1}{2} + n-i \), we get a unique matching.

\[ \square \]

**Remark 1.20.** We can describe the function \( \text{f}_{\text{area}=1} \) graphically:

![Diagram](image)

**Figure 1.4:** Dyck paths \( \pi \in L_{n,n}^+ \) with \( a_i = 1 \) are pushed by \( \text{f}_{\text{area}=1} \) to those where only \( a_{n+1-i} = a_{n+2-i} \).
Note as well that all the paths of the set \( \{ \pi \in L_{n,n}^+ : dinv(\pi) = 1 \} \) are those where we begin by \( i \) N-steps, followed by one E-step, \((n-i)\) additional N-steps and ended finally with \((n-1)\) E-steps.

1.2.2 The bijection \( f \) for paths \( \pi \in L_{n,n}^+ \) with \( area(\pi) = 2 \)

The following section will provide the bijection

\[
f_{area=2} : \{ \pi \in L_{n,n}^+ : area(\pi) = 2 \} \rightarrow \{ \pi \in L_{n,n}^+ : dinv(\pi) = 2 \}
\]

which satisfies \( dinv(\pi) = area(f_{area=2}(\pi)) \). This is a partial solution to The Symmetry Problem, where the general bijection \( f \) is restricted only to the cases where \( area \) or \( dinv \) statistics are 2.

**Lemma 1.21.** \( |\{ \pi \in L_{n,n}^+ : area(\pi) = 2 \}| = \binom{n-1}{2} = \frac{(n-1)(n-2)}{2} \).

**Proof.** Let \( \pi \in L_{n,n}^+ \) such that \( area(\pi) = 2 \). Since \( \sum_{k=1}^{n} a_k = 2 \), and since \( a_k \geq 0 \) then one of the two cases are possible: (a) Either there are \( 2 \leq i < j \leq n \) such that \( a_i = a_j = 1 \) and all other \( a_k = 0 \) for \( k \neq i,j \); or, (b) There is exactly one \( 2 \leq j \leq n \) such that \( a_j = 2 \) and for all other \( a_k = 0 \) for \( k \neq j \).

Case (b) is impossible, because if there is only one \( a_j = 2 \), then we know necessarily that \( a_j \leq a_{j-1} + 1 \), then if we flip the inequality we get \( a_{j-1} \geq a_j - 1 = 1 \), so at \( area(\pi) \geq a_{j-1} + a_j = 1 + 2 = 3 \) which contradicts \( area(\pi) = 2 \).

Only case (a) is possible, which is equivalent of choosing 2 different places \( i,j \) among the \( n-1 \) possible places in the area vector \( (a_1 = 0 \text{ always}) \) to place \( a_i = a_j = 1 \) while all other values are null. The number of ways to do so is \( \binom{n-1}{2} \).

\[ \square \]

**Remark 1.22.** In the case where \( area(\pi) = 1 \), we saw that each element of the set \( \{ \pi \in L_{n,n}^+ : area(\pi) = 1 \} \) had a distinguished \( dinv(\pi) \) value (see Lemma 1.15). This is no longer the case for the set \( \{ \pi \in L_{n,n}^+ : area(\pi) = 1 \} \). For example, one can consider the two paths that appear on figure 1.5.

**Definition 1.23.** Define the set \( A_1 = \{ \pi \in L_{n,n}^+ : area(\pi) = 2 , a_2 = 1 \} \), and the set \( A_2 = \{ \pi \in L_{n,n}^+ : area(\pi) = 2 , a_2 = 0 \} \).

Notice that we can express

\[ \{ \pi \in L_{n,n}^+ : area(\pi) = 2 \} = A_1 \cup A_2 \]
Lemma 1.24. \( |A_1| = n - 2 \).

Proof. We claim that \(|A_1| = n - 2\). This is because for any \( \pi \in A_1 \), we need to choose exactly one additional \( a_j = 1 \), where \( j \neq 1, 2 \) (because \( a_1 = 0 \) always and \( \pi \) satisfies already \( a_2 = 1 \)).

Now we let \( A_1 = \{ \pi_3, \pi_4, \ldots, \pi_n \} \), where for any \( 3 \leq j \leq n \) we get the path where \( a_j = a_2 = 1 \). Notice that \( \text{dinv}(\pi_j) = \binom{n-2}{2} + (n - 3) + 1 + (n - j) = \binom{n-1}{2} + n - j \) (since we first count all pairs \( (k, l) \) such that \( a_k = a_l = 0 \), add all pairs \( (2, l) \) and all pairs \( (j, l) \) for \( j < l \leq n \).

Lemma 1.25. \( |A_2| = \binom{n-2}{2} \).

Proof. We claim that \( A_2 = \binom{n-2}{2} \) because we need to choose exactly two indices \( 3 \leq i, j \leq n \) for which \( a_i = a_j = 1 \) (note that \( i, j \neq 1 \) since \( a_1 = 0 \) and \( i, j \neq 2 \) since \( a_2 = 0 \) for \( \pi \in A_2 \)).

We denote \( A_2 = \{ \pi_{i, j} \in L_{n,n}^+ : 3 \leq i < j \leq n, a_i = a_j = 1, \text{and all other } k \neq i, j, a_k = 0 \} \). Notice that \( \text{dinv}(\pi_{i, j}) = \binom{n-2}{2} + (n - i) + (n - j) \), since we begin counting the number of pairs \( (k, l) \) that satisfy \( k < l \) and \( a_k = a_l = 0 \), to this we add all pairs \( (i, k') \) where \( a_{k'} = 0 \) and \( i < k' \), and to this we add all pairs \( (j, k) \) where \( a_k = 0 \) and \( j < k \).

Definition 1.26. Let \( D_1 \) denote the subset of \( \{ \pi \in L_{n,n}^+ : \text{dinv}(\pi) = 2 \} \) where \( \text{dinv}(\pi) \) consists of the two pairs \( (i, k), (j, k) \) where \( i < j < k \), and \( D_2 \) denote the subset of \( \{ \pi \in L_{n,n}^+ : \text{dinv}(\pi) = 2 \} \) where \( \text{dinv}(\pi) \) consists of the two pairs \( (i, j), (k, l) \) where \( i < j < k < l \).

Lemma 1.27. \( |{\pi \in L_{n,n}^+ : \text{dinv}(\pi) = 2}| = \binom{n-1}{2} \).
Lemma 1.29. \( |D_1| = n - 2 \).

**Proof.** We claim that \(|\{ \pi \in L_{n,n}^+ : \text{dinv}(\pi) = 2\}| = \binom{n-2}{2}\), and this is because if we consider \(\pi \in L_{n,n}^+\) with \(\text{dinv}(\pi) = 2\), then \(\text{dinv}(\pi)\) is composed of the pairs \((i_1,j_1)\) and \((i_2,j_2)\) where all we know is that \(i_1 < j_1\) and \(i_2 < j_2\). Since there must be numbers bigger than \(i_1, i_2\) (i.e. \(\max\{j_1, j_2\}\)) then we just need to pick the two indices \(i_1, i_2\) from the set \(\{1, 2, \ldots, n-1\}\) since \(\max\{i_1, i_2\} < n\).

Another argument would be to deduce from the equidistributivity of \(\text{area}(\pi)\) and \(\text{dinv}(\pi)\), namely by considering the coefficient of \(q^2\) of \(\sum_{\pi \in L_{n,n}^+} q^{\text{area}(\pi)} = \sum_{\pi \in L_{n,n}^+} q^{\text{dinv}(\pi)}\). The fact that there are \(\binom{n-1}{2}\) different paths that satisfy \(\text{area}(\pi) = 2\), entails that there are \(\binom{n-1}{2}\) different paths who satisfy \(\text{dinv}(\pi) = 2\).

\(\square\)

**Lemma 1.28.** \(|D_1| = n - 2\).

**Proof.** We claim that \(|D_1| = n - 2\). This is because given \(\pi \in D_1\), and say that \(\text{dinv}(\pi)\) is counted by the pairs \((i,k), (j,k)\) (without loss of generality \(i < j < k\)), then Lemma 1.7 we get that for all \(k' < k\), \(a_{k'} = k' - 1\). Specifically, \(a_i = i - 1\) and \(a_j = j - 1\). Note that if \(a_k = k - 1\) then there can’t be a row below \(k\) with area of \(k - 1\) or \(k\) (so there will be no \(\text{dinv}\) pairs, as a contradiction). Also, if \(a_k = k - 2\) then the only row below \(k\) with area \(k - 2\) or \(k - 1\) would be \(a_{k-1}\), so we will get only one \(\text{dinv}\) pair \((k - 1, k)\) as a contradiction. So \(0 \leq a_k \leq k - 3\). We show that \(k = n\): assume for contradiction \(k \neq n\), then consider \(k + 1 \leq n\), with \(0 \leq a_{k+1} \leq a_k + 1 \leq k - 2\), but for any such choice of \(a_{k+1}\) we have \(k \leq k\) with \(a_k = a_{k+1}\) which adds the pair \((k, k + 1)\) to \(\text{dinv}(\pi)\) as a contradiction.

From the above we conclude that for all \(l < n\), \(a_l = l - 1\) and \(0 \leq a_n \leq n - 3\), so there are exactly \(n - 2\) different ways to chose \(a_n\) while all other rows are uniquely determined.

\(\square\)

Lemma 1.28 shows \(|D_1| = n - 2\), and for any \(\pi \in D_1\) we have \(\text{area}(\pi) = (0, 1, 2, \ldots, n-2, j)\), where \(0 \leq j \leq n - 3\). Denote \(D_1 = \{ \pi'_j : \text{area}(\pi') = (0, 1, 2, \ldots, n-2, j)\}, 0 \leq j \leq n - 3\} = \{\pi'_0, \pi'_1, \ldots, \pi'_{n-3}\}\). Notice that for \(0 \leq j \leq n - 3\) we have \(\text{area}(\pi'_j) = \binom{n-1}{2} + j\).

**Lemma 1.29.** \(|D_2| = \binom{n-2}{2}\).

**Proof.** Let \(\pi \in D_2\) to be a path where \(\text{dinv}(\pi)\) is counted by the pairs \((i,k), (j,l)\) where \(i < k < j < l\).

We claim that \(k = i + 1\) and \(l = j + 1\). Since, without loss of generality, if \(k > i + 1\) then \(a_{i+1} = i\) from Lemma 1.7, but then since \((i,k)\) is a pair, then it must be the case as well that \((i+1,k)\) or \((i-1,k)\) are counted in \(\text{dinv}(\pi)\), which is a contradiction. So the pairs
of $\text{dinv}(\pi)$ are $(i, i+1), (j, j+1)$ where $1 \leq i < i+1 < j < j+1 \leq n$ (note as well that it can not be the case that $j = i+2$ since otherwise we will have the extra pair $(i, i+2)$).

We regard $\pi'_{i,j} \in D_2$ as $\overline{\text{area}}(\pi'_{i,j}) = (0, 1, \ldots, i-1, i, \ldots, j-3, j-2, j-2, j-1, \ldots, n-3)$, where $a_i = a_{i+1} = i-1$, $a_j = a_{j+1} = j-2$, and for all $k_1 < i$ $a_{k_1} = k_1-1$, and for all $i+1 < k_2 < j$ $a_{k_2} = k_2-2$, and for all $j+1 < k_3$ we have $a_{k_3} = k_3-3$. Since the values of $a_k$s are uniquely defined for all $k$, by a certain selection of $i, j$, then we know that the number of elements of $D_2$ is equal to the number of ways we can choose $i, j$ such that $1 \leq i < i+1 < j < j+1 \leq n$, which is precisely $\binom{n-2}{2}$ different ways (the combinatorial proof goes as follows: assume we need to count all different ways to organize $n-4$ bricks and two pairs of bricks in a row, in total we have $n-2$ elements and we need to choose only the two positions of the two pairs, hence $\binom{n-2}{2}$).

Lemma 1.29 suggests that we may denote the elements of $D_2$ as $\pi'_{i,j}$ where only $a_i = a_{i+1}$ and $a_j = a_{j+1}$ and $0 \leq i < i+1 < j \leq n-1$. This implies $\overline{\text{area}}(\pi'_{i,j}) = (0, 1, \ldots, i-1, i-1, i, \ldots, j-3, j-2, j-2, j-2, j-1, \ldots, n-3)$. We then find that $\text{area}(\pi'_{i,j}) = \binom{n-2}{2} + (i-1) + (j-2) = \binom{n-2}{2} + i + j - 3$.

**Proposition 1.30.** $\{ \pi \in L^+_n : \text{dinv}(\pi) = 2 \} = D_1 \cup D_2$.

**Proof.** By the definition of $D_1$ and $D_2$, it is the case that $D_1 \cap D_2 = \emptyset$. Since according to the definition $D_1, D_2 \subseteq \{ \pi \in L^+_n : \text{dinv}(\pi) = 2 \}$, it would be enough to prove is that $|\{ \pi \in L^+_n : \text{dinv}(\pi) = 2 \}| = |D_1| + |D_2|$, but this is an immediate result of Lemmata 1.27-1.29, since $\binom{n-1}{2} = n-2 + \binom{n-2}{2}$.

Define $f^1_{\text{area} = 2} : A_1 \to D_1$, such that

$$f^1_{\text{area} = 2}(\pi_j) = \pi'_{n-j} \quad \text{for } 3 \leq j \leq n$$

(where $\pi_j$ and $\pi'_{n-j}$ are both described after Lemmata 1.24 and 1.28 respectively).

**Proposition 1.31.** The map $f^1_{\text{area} = 2} : A_1 \to D_1$ described above is a bijection that satisfies $\text{area}(\pi) = \text{dinv}(f^1_{\text{area} = 2}(\pi))$ and $\text{dinv}(\pi) = \text{area}(f^1_{\text{area} = 2}(\pi))$.

**Proof.** Lemma 1.24 suggests $|A_1| = n-2$, and Lemma 1.28 suggests $|D_1| = n-2$. The fact that $\text{area}(\pi) = \text{dinv}(f^1_{\text{area} = 2}(\pi))$ is obvious from the definitions of $A_1$ and $D_1$.

Notice that $\text{dinv}(\pi_j) = \binom{n-1}{2} + n - j = \text{area}(\pi'_{n-j})$ (the general equalities were presented after Lemma 1.24 and 1.28 respectively).

\[\Box\]
Let us define \( f^2_{\text{area}=2} : A_2 \to D_2 \), such that
\[
f^2_{\text{area}=2}(\pi_{i,j}) = \pi'_{(n-j+1),(n-i+2)} \text{ for } 3 \leq i < j \leq n
\]
where \( \pi_{i,j} \) and \( \pi'_{i,j} \) are both described after Lemmata 1.25 and 1.29 respectively.

**Proposition 1.32.** The map \( f^2_{\text{area}=2} : A_2 \to D_2 \) described above is a bijection that satisfies \( \text{area}(\pi) = \text{dinv}(f^2_{\text{area}=2}(\pi)) \) and \( \text{dinv}(\pi) = \text{area}(f^2_{\text{area}=2}(\pi)) \).

**Proof.** According to Lemma 1.25 \( |A_2| = \binom{n-2}{2} \), and by Lemma 1.29 \( |D_2| = \binom{n-2}{2} \). From the definition of \( A_2 \) and \( D_2 \) it is obvious that \( \text{area}(\pi) = \text{dinv}(f^2_{\text{area}=2}(\pi)) \).

From the comments after Lemmata 1.25 and 1.29 we can see that
\[
(\text{area}(\pi_{i,j}), \text{dinv}(\pi_{i,j})) = (2, \left(\frac{n-1}{2}\right)+2n-i-j)
\]
\[
= (\text{dinv}(\pi'_{(n-j+1),(n-i+2)}), \text{area}(\pi'_{(n-j+1),(n-i+2)})),
\]
since \( \text{area}(\pi'_{(n-j+1),(n-i+2)}) = \left(\frac{n-2}{2}\right) + (n-j+1) + (n-i+2) - 3 = \left(\frac{n-2}{2}\right) + 2n-i-j. \)

Finally, let us define \( f_{\text{area}=2} : \{ \pi \in L^+_{n,n} : \text{area}(\pi) = 2 \} \to \{ \pi \in L^+_{n,n} : \text{dinv}(\pi) = 2 \} \), such that
\[
f_{\text{area}=2}(\pi) = \begin{cases} f^1_{\text{area}=2}(\pi) & \text{if } \pi \in A_1 \\ f^2_{\text{area}=2}(\pi) & \text{if } \pi \in A_2 \end{cases}
\]

**Theorem 1.33.** The map \( f_{\text{area}=2} : \{ \pi \in L^+_{n,n} : \text{area}(\pi) = 2 \} \to \{ \pi \in L^+_{n,n} : \text{dinv}(\pi) = 2 \} \) described above satisfies both \( \text{area}(\pi) = \text{dinv}(f_{\text{area}=2}(\pi)) \) and \( \text{dinv}(\pi) = \text{area}(f_{\text{area}=2}(\pi)) \).

**Proof.** Thanks to propositions 1.31 and 1.32 we know that both functions \( f^1_{\text{area}=2}(\pi) \) and \( f^2_{\text{area}=2}(\pi) \) satisfy \( \text{area}(\pi) = \text{dinv}(f^i_{\text{area}=2}(\pi)) \) and \( \text{dinv}(\pi) = \text{area}(f^i_{\text{area}=2}(\pi)) \) (for \( i = 1, 2 \)). This entails that \( f_{\text{area}=2} \) will satisfy \( \text{area}(\pi) = \text{dinv}(f_{\text{area}=2}(\pi)) \) and \( \text{dinv}(\pi) = \text{area}(f_{\text{area}=2}(\pi)) \).

Since \( \text{Im}(f^1_{\text{area}=2}) = D_1 \) and \( \text{Im}(f^2_{\text{area}=2}) = D_2 \), and \( \{ \pi \in L^+_{n,n} : \text{dinv}(\pi) = 2 \} = D_1 \cup D_2 \), we are guaranteed that \( f_{\text{area}=2} \) is a well-defined bijection.
We may describe the function $f_{\text{area}=2}$ graphically using the following figures (notice how $f_{\text{area}=2}$ acts differently on paths from $A_1$ and $A_2$):

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure1.pdf}
\caption{Here we can see how the function $f_{\text{area}=2}$ maps differently Dyck paths from $\{\pi \in L_{n,n}^+ : \text{area}(\pi) = 2, a_2 = 1\}$ and Dyck paths from $\{\pi \in L_{n,n}^+ : \text{area}(\pi) = 2, a_2 = 0\}$.}
\end{figure}
Chapter 2

Representations of other Catalan Structures

2.1 Motivation

In the previous chapter we tried to solve The Symmetry Problem

\[
\sum_{\pi \in L_{n,n}^+} q^{area(\pi)} t^{dinv(\pi)} = \sum_{\pi \in L_{n,n}^+} q^{dinv(\pi)} t^{area(\pi)}
\]

by trying to find a bijection \( f : L_{n,n}^+ \to L_{n,n}^+ \) that would interchange the area and the \textit{dinv} statistics. Such a function \( f \) said to alter a certain path \( \pi \in L_{n,n}^+ \) and provide us a new path \( f(\pi) \in L_{n,n}^+ \), so that \( area(\pi) = inv(f(\pi)) \) and \( inv(\pi) = area(f(\pi)) \). Notice as well that we could have provided instead a bijection that interchanges the \textit{bounce} and the \textit{area} statistics. In any case, such a function \( f \) would be a certain algorithm, or a certain process, by which we alter one path in order to get another. The problem with such approach, is that it is sometimes difficult to read some of the statistics from Dyck Paths.

Obviously, for \( \pi \in L_{n,n}^+ \), one can consider the \textit{area}(\pi) statistic as a sum of area blocks that are locked between \( \pi \) and \( y = x \), and read those area blocks one row by another. We used this technique before, when we introduced \( \text{area}^\downarrow(\pi) \) and considered \( area(\pi) = \sum_{k=1}^{n} a_k \).

But the decision to read the area in terms of rows was rather comfortable, because we could use the inequalities \( a_k \leq k - 1 \) and \( a_{k+1} \leq a_k + 1 \) for all \( k \).

Instead we could easily have chosen to read the area blocks of \( \pi \) one column by another, where the first column is the right-most one. Following that, we could have instead defined \( area^\uparrow(\pi) = (a_1^\uparrow, a_2^\uparrow, \ldots, a_n^\uparrow) \), where \( a_i^\uparrow \) is the area units of \( \pi \) locked between \( y = x \) and...
and $\pi$ itself at the $i$th column, where the first column is the one on the left. We would then get the inequality $0 \leq a_i^i \leq n - i$, where $a_n^i = 0$ always and $a_i^i \leq a_{i+1}^i - 1$.

Another natural way to read the $\text{area}(\pi)$ statistic would have been to consider how many area units are locked between $y = x$ and the path $\pi$, where we consider the $\text{area}(\pi)$ as a sum of diagonals. That is, let $\text{area}'(\pi) = (a_0', a_1', \ldots, a_{n-1}')$, and $\text{area}(\pi) = \sum_{k=0}^{n-1} a_k'$, where $a_k'$ is the number of area units located between $\pi$ and $y = x$ that the line $y = x + k$ intersects. Obviously $a_0' = 0$ always, and for all other $k$, $a_k' \leq n - k$. But also we get the inequality $a_{k+1}' \leq a_k' - 1$ since each elements of $a_{k+1}'$ could be identified by the squares beneath it and to its’ right, and those are elements of the diagonal $a_k'$.

Similarly to what we did with the $\text{area}(\pi)$ statistic, we might as well want to consider $\text{dinv}(\pi)$ as a sum of non-negative entries of vector in $\mathbb{Z}^n$. One way to do it would be to introduce $\text{dinv}^{\text{OUT}}(\pi) = (d_1^{\text{OUT}}, d_2^{\text{OUT}}, \ldots, d_n^{\text{OUT}})$, where $d_k^{\text{OUT}}$ is the number of $i$s such that $i < k$, and $a_i = a_k$ or $a_i - 1 = a_k$. By our definition, $d_1^{\text{OUT}} = 0$, and $0 \leq d_k^{\text{OUT}} \leq k - 1$. We get that $\text{dinv}(\pi) = \sum_{k=1}^{n} d_k^{\text{OUT}}$.

Alternatively, we may introduce $\text{dinv}^{\text{IN}}(\pi) = (d_1^{\text{IN}}, d_2^{\text{IN}}, \ldots, d_n^{\text{IN}})$, where $d_k^{\text{IN}}$ is the number of $i$s such that $k < i$, and $a_i = a_k$ or $a_i - 1 = a_k$. Obviously $d_1^{\text{IN}} = 0$, and for all $k$, $0 \leq d_k^{\text{IN}} \leq n - k$. This yields as well that $\text{dinv}(\pi) = \sum_{k=1}^{n} d_k^{\text{IN}}$.

One could have hoped to find a certain correspondence between those vectors, so that it would be easy to solve The Symmetry Problem using a bijection $f : L_{n,n}^+ \to L_{n,n}^+$. Unfortunately, non of those vectors seem to have a visible pattern for a general $n$. This could be the case because Dyck paths are only one example of a more general structure called Catalan Structure. A Catalan Structure is a collection of sets $\bigcup_{n=1}^{\infty} S_n$ of cardinality $|S_n| = C_n$, where the elements each set $S_n$ follows a certain rule and there is a bijection between $g : L_{n,n}^+ \to S_n$. We may regard a Catalan Structure as a combinatorial interpretation of $C_n$. Stanley published over 207 such interpretations [7] [8, p. 20].

Some of the combinatorial interpretations known today could be less intuitive than Dyck Paths. Yet one can hope that given a certain bijection from Dyck paths to other Catalan Structure (i.e. trees or permutations), one could solve The Symmetry Problem if it would be possible to translate the $\text{area}(\pi)$, $\text{dinv}(\pi)$ or $\text{bounce}(\pi)$ statistics to a known statistic in the new Catalan Structure. Since many of the Catalan Structures that appears in Stanley’s list [8, pp. 219-229] are similar, the following sections would deal with distinct representations of Catalan Structures. In each structure we examine, we will describe a classical known bijection from Dyck paths to that structure [9], together with our personal attempts to read some of the new structure’s statistics.
2.2 Murasaki Diagrams

Given $n$ vertical lines, we may construct the diagram obtained by joining some of these $n$ lines with non-intersecting horizontal lines.

![Diagram](image)

For $n = 1$ there is only one such diagram, there are precisely two diagrams when $n = 2$, and for $n = 3$ there are 5 possible diagrams. For a general $n$ there would be $C_n$ different diagrams.

We call such diagrams **Murasaki diagrams**.

2.2.1 Catalan Structure

In fact one can show that the number of Murasaki diagrams with $n$ vertical lines is precisely $C_n$, which’s the $n$-th Catalan number. The bijection goes as follows: we begin by labeling the vertical lines by $1 \ldots n$ from left to right and then the vertical lines denote a certain partitioning of the $n$ lines. Every vertical line that is not connected to any vertical line correspond to the $NE$ steps pair (of an ordinary Dyck path), while every group of vertical lines connected with a horizontal line correspond to $N \square E$ where by $\square$ we mean taking the subdiagram that starts from the second vertical line in the group to the last one.

![Diagram](image)

Figure 2.2: The process of translating a Murasaki diagram to a Dyck path

Alternatively we can view the noncrossing Murasaki diagrams as a noncrossing partitioning of $\{1, 2, \ldots, n\}$. 
Chapter 2. Representations of other Catalan Structures

2.2.2 Represetation of the \textit{area()} statistic

We are capable of reading out the \textit{area()} statistic of a certain Murasaki diagram. This is done by the following recursive algorithm: begin reading the diagram from left to right, whenever you reach a horizontal line, count the number of vertical lines that lie underneath it and subtract 1, add it to your sum. Now consider all the lines underneath that vertical lines as another Murasaki diagram and run the same operation again (where the first vertical line is excluded).

For instance, given the Murasaki diagram in figure 2.3, we can count the area as -

\[
\text{area()} = 1 + 0 + 0 + 0 + 3 + 2 + 0 + 0 = 6
\]

A different way to count the \textit{area()} would be to read the Murasaki diagram from right to left, and in every step we count the number of horizontal lines that lie above or connected to that vertical line from the left, whereas if one horizontal line is connected to \(k\) vertical lines, then we consider this as \(k\) different horizontal lines. We can again compute the area of the diagram in figure 2.3 by -

\[
\text{area()} = 2 + 2 + 1 + 0 + 0 + 1 + 0 + 0 = 6
\]

Where it is easier to see the different \(k\) vertical lines in figure 2.4:

\[
\text{Figure 2.4: This is the same diagram as the one in figure 2.3, with an emphasis of different vertical lines.}
\]

In fact, the first method reads the area statistic as a collection of area unit squares that lie in the diagonal of a Dyck path (what was previously introduced as \(\sum_{k=0}^{n-1} a'_k\)), while the second method reads the area as rows (or, \(\sum_{k=1}^{n} a_k\)). This is because the first method counts a subdiagram that is locked between \(N\) and \(E\) steps, which means that all the diagonal between the two necessarily exists. The second method, on the other
2.2.3 Representation of the \textit{dinv()} statistic

In order to count the \textit{dinv()} statistic we need to run the following procedure: start from right to left, for every vertical line you reach, count how many horizontal lines are located above it (and in this sense we consider horizontal lines that are connected to a vertical line from the left side as if they were located above), while keeping in mind that if see a horizontal line that is connected to several vertical lines, then we consider it as several, distinct, "bridges" that connect all other lines to the right-most with "bridges". We disregard any "bridge" that is connected to the vertical line in question, if it is connected to it and the "bridge" is located to its right. See figure 2.4.

Say that our vertical line has $k$ horizontal lines above it ("bridges"), in that sense, then we count how many vertical lines appears to the left of that vertical line, with either $k$ or $k + 1$ "bridges" above them.

For instance the \textit{dinv} statistic of the example given in figure 2.4 we get that (summing from right to left) -

$$\text{dinv}() = 1 + 0 + 1 + 4 + 3 + 2 + 0 + 0 = 11$$

This is because we compare area lines of the same length or of different length, where the bigger area value differs from the smaller one by one area unit and located below the smaller one.

2.3 Plane Trees

A plane tree is a rooted tree for which an ordering is specified for the children of each vertex. We denote with $T_n$ the collection of all plane trees over $n$ vertices.

2.3.1 Catalan Structure

We can describe a bijection between $T_{n+1}$ to $L_{n,n}^*$ in order to prove that $T_{n+1}$ is indeed a Catalan Structure of order $n$. The Dyck path that corresponds to a certain planted
Figure 2.5: $T_2$ contains only one possible plane tree, $T_3$ has two different trees and $T_4$ has 5 different trees. In general, $|T_{n+1}| = C_n$.

A tree is given by the following algorithm: begin at the root of the tree, and do a left-depth-search\(^1\) where in every step we go deeper in the graph we add a $N$-step to our Dyck path, and every step we go up a level we add $E$-step to our Dyck path. At the end of the left depth-search, return to the root node with as many up steps needed (which corresponds to as many $N$s that need to be added to the Dyck path). This is indeed a Dyck path, because at every step, there would be at least as many $N$s than $E$s, as the level of the every node is non-negative. Notice that depth-search of a planted graph of $n + 1$ nodes contains $2n$ steps (together with returning to the root node), which means that the Dyck path would be in $L_{n,n}$.

Figure 2.6: The left depth-search of the tree corresponds to the sequence $NNENNEEE$.

---

\(^1\)Depth-search, or depth-first-search, is an algorithm for searching within a tree, where one starts at the root and explores as far as possible along each branch before backtracking, preferring always to search down on the left most possible node.
2.3.2 Representation of area() statistic

Let $T$ be our rooted tree, and let $V(T)$ be the set of all the nodes of the tree, where $v_0$ is the root. Define $\text{depth}(v)$ to be the number of vertices that are located above $v$, where the root $v_0$ is not counted. Then

$$\text{area}(T) = \sum_{v \in V(T)} \text{depth}(v)$$

This is because $\text{depth}(v)$ is translated in the Dyck path to the difference between $N$-steps to $E$-steps that we done so far. This is precisely the area value of every row, so $\text{area}(\pi) = (\text{depth}(v_1), \ldots, \text{depth}(v_n))$.

2.3.3 Representation of dinv() statistic

In order to read off the dinv() statistic from a rooted tree, we need to run the following algorithm: for any $v \in V(T)$, count how many vertices $u \in V(T)$ satisfy $\text{depth}(v) = \text{depth}(u)$ or $\text{depth}(v) = \text{depth}(u) + 1$ and the vertex $u$ is located on a branch which is to the right of the branch of $v$.

For instance, consider the following tree:

![Figure 2.7](image)

**Figure 2.7:** Counting the dinv() of the tree consists of counting the number of nodes to the right, at the same depth, or one level above, of each node.

2.4 Complete Binary Trees

Complete binary trees are rooted trees where every node has either no children, or exactly two of them. We denote the set of all complete binary trees of $n$ nodes with $B_n$.

2.4.1 Catalan Structure

We claim that there is a bijection between $B_{2n+1}$ to $\{ \pi : \pi \in L_{n,n}^+ \}$, given by a left depth search (see section 2.3): Run a left depth search on the tree, whenever the search goes down through the left-child, write $N$-step, and whenever the search goes down through
Chapter 2. Representations of other Catalan Structures

2.4.2 Representation of \( \text{area}() \) statistic

let \( T \) be our binary tree, and let \( V' := \{v_1, v_2, \ldots, v_n\} \subset V \) be the set of all non-leaf and non-root vertices. In other words any vertex in \( V' \) contains two children, and the root is not in \( V' \). Let \( v_0 \) be the root of \( T \). For any \( v \in V' \) there is a unique path from \( v_0 \) to \( v \) consisting of left or right down-steps. We denote with \( L(v) \) the total number of left down-steps needed to walk from \( v_0 \) to \( v \). Hence, if \( \pi \) is the corresponding Dyck path of \( T \), then

\[
\text{area}(\pi) = \sum_{v \in V'} (L(v) - 1)
\]

This is because every \( a_i \) which is not \( a_1 \) is represented by a vertex of \( V' \) (\( a_1 \) is represented by the root \( v_0 \)), and when we count \( L(v) \) we in fact count the number of \( N \)s subtracted by the number of \( E \)s in the corresponding row of \( \pi \).

2.4.3 Representation of \( \text{dinv}() \) statistic

Unfortunately, I could not find an elegant method of extracting the \( \text{dinv}() \) statistic from those binary trees. The easiest way to count the \( \text{dinv}() \) would be to generate a list of
\((a_1, a_2, \ldots, a_n)\) using the previous algorithm in the right order, and then to calculate \(\text{dinv}\) from the list.

The lack of elegant statistics to this representation made it rather cumbersome to work with. However, one could still expect that there might be a simple rule regarding the matching of those graphs (matching that corresponds to the desired function from the symmetry problem).

Label the trees of figure 2.9 from left to right with \(a, b, c, d, e\) respectively. Then, while we will get that \(a - e\) is one match (which is a symmetric reflection), we will also get that \(b - c\) is also a required match, and \(d - d\) is one too. When we extended the problem to higher \(n\) two problems arose: first, the element of choice appeared - there was no 1-to-1 correspondence between an ordered pair \((x, y)\) of \((\text{area}, \text{dinv})\) values and a single tree (or Dyck path, in general). This means that even if we tried to find a pattern through examples, then it was no longer possible for us to know for certainty that we are trying to follow the right rule (already for the case of \(n = 5\)). Secondly, that matchings of those trees (like those of Dyck paths) seemed rather unexpected - sometimes the matching was based on pure left to right symmetry, some other times it was based on a right-deep search instead for left, while in most cases we could not tell.

### 2.5 Triangulations of \((n + 2)\)-gons

The first problem given by Stanley in his list is to try prove why the number of different ways to triangulate an \((n + 2)\)-gon is \(C_n\)\(^8\), p. 220].

![Figure 2.9: All possible triangulations of a pentagon (\(C_3 = 5\)).](image)

#### 2.5.1 Catalan Structure

There is a bijection from these triangulations to binary trees of \(n\) vertices (and by remark 2.1 also to complete binary trees of \(2n + 1\) vertices, see section 2.4). Observe that every triangle in our triangulation has at most 3 neighbour triangles, and at most 2 neighbour triangles if one of its’ edges is a boundary edge of the polygon. We begin by choosing one edge of the polygon, which we call the \textit{rooted edge}, and we place a node at the center
of the triangle that contains the rooted edge. If there is a triangle that is adjacent to the edge to the left of the rooted edge, then we draw a left child to that node and, enter the new triangle through the common edge. Similarly, we draw right child to the node if there is a triangle that share the same edge as the edge to the right of the rooted edge. To any new triangle we enter, we draw a left child to the relevant node for any adjacent unvisited triangle that appears to the left of the edge we entered the triangle from. Similarly we do with right children. We continue in this manner until we cover all triangles.

This representation was slightly inconvenient to work with. First, even if we always pick the same edge of the polygon, it seemed difficult to be able to find an alternative way to read the area() statistic, besides applying the bijection and reading it off the complete tree. Secondly, the choice we make of selecting the initial edge changes completely the \((\text{area}, \text{dinv})\) values we should assign to the triangulation. In other words, a certain triangulation has no specific \((\text{area}, \text{dinv})\) values unless we also mention an initial edge.

### 2.6 Non crossing matching of \(2n\) nodes on a circle

Let \(v_1, v_2, \ldots, v_{2n}\) be \(2n\) nodes located at the boundary of a circle with equal distance between every two adjacent nodes. Let \(R_n\) be the set of different ways to pair the \(2n\) nodes with edges (as straight lines) that do no intersect. Then, \(|R_n| = \binom{2n}{n}\).
2.6.1 Catalan Structure

In order to show that $|R_n| = C_n$ we will provide a bijection between each and every element of $R_n$ to a Dyck path $\pi \in L^+_{n,n}$. The bijection goes as follows: Choose an initial node, called it the rooted node, and travel clockwise along the circle. Every time you encounter a line for the first time, write $N$-step in $\pi$, and every time you encounter a line for the second time, write $E$-step in $\pi$. Stop when you cover all nodes. There are in total $n$ lines so we are guaranteed to have $n$-$N$s and $n$-$E$s.

Similar to the case of triangulation of polygons, we made a choice here regarding which node would be the rooted node. This means that every statistic we try to find should take into account the location of the rooted node.

![Figure 2.12: The bijection to/from non-crossing matching of $2n$ nodes on the circle to Dyck paths](image)

We were not able to come up with a simpler way to read the area, bounce or dinv statistics other than applying the bijection above.

2.7 Further Representations

In this section we covered several Catalan Structures that appeared on Stanley’s list[8, pp. 219-229]. We hoped that by translating the problem from Dyck paths with the statistics we knew to new structures with new statistics, we would be able to find new connection between the new statistics. Unfortunately, it seemed not to work with the bijections we tested.

Although some of Stanley’s examples are variations of the bijections given above, there is still a very important structure that we did not consider: permutations.
Chapter 3

Permutations

3.1 Motivation

The study of permutations is one of the most developed fields in combinatorics. Vast work was dedicated to the development of statistics of permutations, and we hope we would be able to come up with interesting results when we convert The Symmetry Problem to a symmetry problem regarding different statistics of permutations.

As the size of all permutations of size $n$ is $|S_n| = n!$, we need to restrict somehow the number of permutations we would like to consider so that we would be able to come out with a 1-to-1 correspondence with those permutations and Dyck paths. To do so, we introduce the concept of pattern-avoidance. Given $\sigma \in S_n$ and $\tau \in S_k$, where $2 \leq \tau \leq n$, we say that $\sigma$ avoids $\tau$, or $\sigma$ is $\tau$-avoiding, if $\sigma$ does not contain a subword of length $k$ that have the same relative order as $\tau$. For instance, $52314 \in S_5$ is not 312-avoiding, since the subword 524 exists in 52341, but it is 132-avoiding. Let $\tau \in S_k$, where $2 \leq k \leq n$, we denote the set of all permutations of size $n$ that are $\tau$-avoiding by $S_n(\tau)$. Knuth discovered that for any $\tau \in S_3$, $S_n(\tau)$ is a Catalan Structure [5].

For instance,

$$S_4(312) = \{1432, 1342, 1324, 1243, 1234, 2143, 2134, 3214, 2314, 4321, 3421, 3241, 2431, 2341\}$$
Chapter 3.  Permutations

\[ S_5(312) = \{12543, 12453, 12435, 12354, 13245, 14325, 45321, 43521, \\
13425, 15432, 14532, 14352, 13452, 21543, 21453, 21435, 21354, \\
21345, 32154, 32145, 23154, 23145, 43215, 34215, 32415, 24315, 23415, \\
25431, 24531, 24351, 23541, 23451, 32541, 32451, 43251, 34251, 54321, \\
35421, 34521\} \]

3.2 Statistics of Permutations

3.2.1 Inversions

Some statistics of permutations raised some more interest than others. Especially those called Mahonian statistics. A permutation statistic is called Mahonian if and only if it has the same distribution over \( S_n \) as the Inv statistic, which counts the number of inversions. A formal definition of Inv is given by -

\[ \text{Inv}(\sigma_1 \sigma_2 \ldots \sigma_n) = |\{ (\sigma_i, \sigma_j) : i < j \text{ and } \sigma_i > \sigma_j \}| \]

We say that an inversion in a permutation is an occurrence of \((21)\) in it. For instance, \(\text{Inv}(24315) = 4\), as the inversions are \((2, 1), (4, 3), (4, 1), (3, 1)\). An inversion of adjacent indices is called a descent.

3.2.2 Major Index

Another famous Mahonian statistic is the Major index, introduced first by MacMahon himself [6]. It counts the locations of the descents of the permutation.

\[ \text{maj}(\sigma_1 \sigma_2 \ldots \sigma_n) = \sum_{\sigma_i > \sigma_{i+1}} i \]

For instance, \(\text{maj}(24315) = 2 + 3 = 5\).

3.2.3 Pseudo-Mahonian Statistics

We say that two different statistics on a class of objects are equidistributed if they have the same generating function over that class of objects. In other words we may say that they distribute the same over the class of objects. We stated before that Inv and maj
are equidistributed as they have the same distribution over the class of permutations \([6]\).
The Symmetry Problem implies that \(\text{area}, \text{bounce}\) and \(\text{dinv}\) statistics are equidistributed as well over the class of all Dyck paths, and so on.

It would be convenient to say that any statistic that distributes the same as the \(\text{area}\) statistic if \textbf{pseudo-Mahonian}. Thus, we consider \(\text{bounce}\) and \(\text{dinv}\) to be pseudo-Mahonian statistics (see section 1.2).

Let \(k\) be a non-negative integer, and denote with \(F(k)\) be \(|\{\pi \in L_{n,n}^+ : \text{area}(\pi) = k\}|\). So we get the following distribution table of pseudo-Mahonian statistic for \(n = 4\):

\[
\begin{array}{ccccccc}
 k & 0 & 1 & 2 & 3 & 4 & 5 \\
 F(k) & 1 & 3 & 3 & 3 & 2 & 1 \\
\end{array}
\]

And for \(n = 5\):

\[
\begin{array}{ccccccccccc}
 k & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\
 F(k) & 1 & 4 & 6 & 7 & 7 & 5 & 5 & 3 & 2 & 1 & 1 \\
\end{array}
\]

**Proposition 3.1.** The \(\text{inv}\) statistic restricted on \(S_n(123)\) is not pseudo-Mahonian

**Proof.** Consider the distribution table of \(\text{inv}\) over \(S_5(123)\):

\[
\begin{array}{ccccccccccc}
 k & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\
 F(k) & 1 & 4 & 9 & 12 & 10 & 4 & 2 & 0 & 0 & 0 & 0 \\
\end{array}
\]

\[\square\]

**Proposition 3.2.** The \(\text{inv}\) statistic restricted on \(S_n(321)\) is not pseudo-Mahonian

**Proof.** Consider the distribution table of \(\text{inv}\) over \(S_5(321)\):

\[
\begin{array}{ccccccccccc}
 k & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\
 F(k) & 1 & 4 & 9 & 12 & 10 & 4 & 2 & 0 & 0 & 0 & 0 \\
\end{array}
\]

\[\square\]
Proposition 3.3. For all $\tau \in S_3$, the maj statistic restricted on $S_n(\tau)$ is not pseudo-Mahonian

Proof. Consider the distribution tables of maj for $S_5(\tau)$ for all $\tau \in S_3$:

\[
\begin{array}{cccccccccccc}
\tau = 123: & k & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\
& F(k) & 0 & 0 & 0 & 0 & 5 & 5 & 9 & 9 & 9 & 4 & 1 \\
\tau = 132: & k & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\
& F(k) & 1 & 4 & 3 & 8 & 6 & 6 & 4 & 2 & 1 & 1 \\
\tau = 213: & k & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\
& F(k) & 1 & 1 & 2 & 4 & 6 & 6 & 8 & 3 & 4 & 1 \\
\tau = 231: & k & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\
& F(k) & 1 & 4 & 3 & 8 & 6 & 6 & 4 & 2 & 1 & 1 \\
\tau = 312: & k & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\
& F(k) & 1 & 1 & 2 & 4 & 6 & 6 & 8 & 3 & 4 & 1 \\
\tau = 321: & k & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\
& F(k) & 1 & 4 & 9 & 9 & 9 & 5 & 5 & 0 & 0 & 0 & 0
\end{array}
\]

\[\square\]

3.3 312-avoiding Permutations

In the following section we introduce Bandlow’s and Killpatrick’s bijection from Dyck paths to 312-avoiding permutations that translates the area statistic to INV [2].

Let $\sigma \in S_n$ to be $\sigma = \sigma_1 \sigma_2 \ldots \sigma_n$. Let $\text{INV}_i(\sigma) = |\{\sigma_k : k > j \text{ and } \sigma_k < i \text{ where } \sigma_j = i\}|$, namely $\text{INV}_i(\sigma)$ counts how many numbers appear to the right of the number $i$ in $\sigma$ and are smaller than $i$. Obviously $\text{INV}(\sigma) = \sum_{i=1}^{n} \text{INV}_i(\sigma)$. 
The bijection $f : S_n(312) \to L^+_{n,n}$ introduced by Bandlow and Killpatrick, does the the following [2, pp. 9-11]:

$$f(\sigma) = \pi, \text{ such that } \overrightarrow{area}(\pi) = (\text{INV}_1(\sigma), \text{INV}_2(\sigma), \ldots, \text{INV}_n(\sigma))$$

**Proposition 3.4.** The function $f : S_n(312) \to L^+_{n,n}$, with $\sigma \mapsto \pi$ and

$$\overrightarrow{area}(\pi) = (\text{INV}_1(\sigma), \text{INV}_2(\sigma), \ldots, \text{INV}_n(\sigma))$$

satisfies that $\text{INV}(\sigma) = \text{area}(\pi)$.

**Proof.**

$\text{area}(\pi) = \sum_{i=1}^{n} a_i = \sum_{i=1}^{n} \text{INV}_i(\sigma) = \text{INV}(\sigma)$, so indeed the bijection translates the INV statistic to area.

We know that in order for $f$ to provide a legitimate $\overrightarrow{area}$ vector, each $a_i$ must satisfy both $0 \leq a_i \leq i - 1$ and $a_i \leq a_{i-1} + 1$. The first demand is satisfied because there are no numbers smaller than 1, at most one number smaller than 2 that appears to it’s right at the permutation, at most two numbers smaller than 3 that appear on it’s right, and so on. The second demand is satisfied because our demand is that $\sigma \in S_n(312)$, so for every $j$ there can’t be numbers that are both smaller than $j$ and $j + 1$ that appear between them. In other words, every area line can have at most one more area value than it’s previous.

![Figure 3.1: The corresponding Dyck path of the permutation 453621 ∈ S_n(312)](image)

When we consider $f^{-1}$, it would be helpful to know where do the other statistics go to under this bijection. Bandlow and Killpatrick mention that there was no Mahonian statistic that would correspond to $\text{bounce}(\pi)$ according to $f$ above. They considered all possible Mahonian statistics given in a list by Babson and Steingrimsson [1]. Unfortunately, none of these statistics seem to correspond well to the $\text{dinv}(\pi)$ statistic either.

Consider the function $g : S_n(312) \to L^+_{n,n}$ described by the following algorithm: let $\sigma \in S_n(312)$ and begin the path $\pi \in L^+_{n,n}$ with as many $N$-steps as $|\{i : \text{INV}_i(\sigma) = 0\}|$.
Then, for all $1 \leq i \leq n$ by order, check $\text{inv}_i(\sigma)$, if $\text{inv}_i(\sigma) = 0$ continue $\pi$ with one $E$-step, and if $\text{inv}_i(\sigma) = 1$ continue $\pi$ with one $N$-step. Then, for all $1 \leq i \leq n$ by order, check $\text{inv}_i(\sigma)$, if $\text{inv}_i(\sigma) = 1$ continue $\pi$ with one $E$-step, and if $\text{inv}_i(\sigma) = 2$ continue $\pi$ with one $N$-step. Then, for all $1 \leq i \leq n$ by order, check $\text{inv}_i(\sigma)$, if $\text{inv}_i(\sigma) = 2$ continue $\pi$ with one $E$-step, and if $\text{inv}_i(\sigma) = 3$ continue $\pi$ with one $N$-step. Continue in this manner until you place $N$-steps for every $\text{inv}_i(\sigma) = n$, and $E$-steps for every $\text{inv}_i(\sigma) = n - 1$.

Then, place as many $E$-steps in $\pi$ as divides.$_{\text{alt0}}$ $i : \text{inv}_i(\sigma)$ divides.$_{\text{alt0}}$ $n$.

**Theorem 3.5.** The function $g : S_n(312) \to L_{n,n}^+$, where $\sigma \mapsto \pi$, described above satisfies that $\text{inv}(\sigma) = \text{bounce}(\pi)$.

**Proof.** The function $g$ is a composition of these two functions $f : S_n(312) \to L_{n,n}^+$ described above, such that $\text{inv}(\sigma) = \text{area}(\pi)$, together with the bijection $\zeta : L_{n,n}^+ \to L_{n,n}^+$ by Haglund that satisfies $\text{dinv}(\pi) = \text{area}(\zeta(\pi))$ and $\text{area}(\pi) = \text{bounce}(\zeta(\pi))$ [4, p. 59]. □

Appendix A presents the bijection $g$ from $S_4(312)$ to $L_{4,4}^+$. We were not able to find the statistic over $S_n(312)$ that corresponds to $\text{area}(g(\sigma))$.

### 3.4 231-avoiding Permutations

Let $\pi \in L_{n,n}^+$ be a Dyck path that contains a total of $n$ $N$-steps and $n$ $E$-steps. We call a an adjacent pair of $EN$-steps a *valley* of $\pi$, and an adjacent pair of $NE$-steps a *peak* of $\pi$. Note that given the set of all valleys and all peaks, one can uniquely determine $\pi \in L_{n,n}^+$.

**Definition 3.6.** Let $\pi \in L_{n,n}^+$ be a Dyck path written as a sequence of $N$s and $E$s. Let $\text{Des}(\pi)$ be the set that assigns to every valley of $\pi$ the total number of letters that appear before $N$ in the $EN$ pair.

For instance, let $\pi \in L_{n,n}^+$ be $NENNENEENENE$. Then we denote the four valleys of $\pi$ with parentheses: $\pi = N(EN)N(EN)E(EN)(EN)E$, and sum the number of letters that precedes every $N$ step of every valley:

$$\text{Des}(\pi) = 2 + 5 + 8 + 10 = 25$$

**Definition 3.7.** Given $\sigma \in S_n$, let $\sigma^{-1}$ be the inverse of $\sigma$ (so that $\sigma\sigma^{-1} = \text{Id}$). We let the *inverse major index* be

$$\text{imaj}(\sigma) = \text{maj}(\sigma^{-1})$$
Definition 3.8. Given \( \sigma \in S_n \), such that \( \sigma = \sigma_1 \sigma_2 \ldots \sigma_n \), then we let the **descent set** of \( \sigma \) be -

\[
\text{des}(\sigma) = \{ j : \sigma_j > \sigma_{j+1} \}
\]

Similarly, we let the **inverse descent set** of \( \sigma \) be -

\[
\text{ides}(\sigma) = \text{des}(\sigma^{-1})
\]

where \( \sigma^{-1} \) is the inverse of \( \sigma \).

In his paper, Christian Stump provides a bijection \( \phi : S_n(231) \to L^+_{n,n} \), such that \( \text{Des}(\pi) = \text{maj}(\sigma) + \text{imaj}(\sigma) \) [11].

Stump’s bijection \( \phi \) works according to the following algorithm: let \( \sigma \in S_n(231) \) with \( \text{des}(\sigma) = \{i_1, i_2, \ldots, i_k\} \) and \( \text{ides}(\sigma) = \{i'_1, i'_2, \ldots, i'_k\} \). Then we begin writing \( \pi \in L^+_{n,n} \) with \( i'_1 \) N-steps, followed by \( i_1 \) E-steps. Then we add \( i'_2 - i'_1 \) N-steps in \( \pi \), followed by \( i_2 - i_1 \) E-steps, and so on. We continue in this manner until we end \( \pi \) with \( n - i'_k \) N-steps followed by \( n - i_k \) E-steps.

An example of the \( \phi \) bijection appears on Appendix B (for the case \( n = 4 \)). We were not able to find a corresponding pseudo-Mahonian statistic in \( S_n(231) \) that would correspond to \( \text{area}(\phi(\sigma)) \), to \( \text{dinv}(\phi(\sigma)) \) or to \( \text{bounce}(\phi(\sigma)) \).

3.5 Further approaches

In this section we tried to come up with statistics for permutations that would correspond well to one of the statistics we know for Dyck paths, under famous bijections. There are more bijections from \( S_n(\tau) \), for \( \tau \in S_3 \), to Dyck paths and we did not cover all of them.

In addition to that, we tried to see whether we could consider Mahonian statistics over \( S_n \), and see whether they turn to pseudo-Mahonian statistics over \( S_n(\tau) \).

We had few attempts as well to try to introduce new pseudo-Mahonian statistics (that are not restrictions of Mahonian ones), but our attempts failed.
Chapter 4

Parking Functions

4.1 Motivation

The Symmetry Problem deals with finding a bijection that could interchange $\text{area}(\pi)$ and $\text{dinv}(\pi)$ statistics of a Dyck path. Although this problem is solved, attempts to find a combinatorial proof failed [4, p. 49], even when the problem translated itself into other Catalan Structures.

The Symmetry Problem is extended to a similar problem regarding parking functions [4, p. 82]. A parking function is an extension of permutation, and we can view parking functions as labeled (or indexed) Dyck paths. Similarly, we introduce the two modified statistics $\text{area}(p), \text{dinv}(p)$ for parking functions and want to prove that

$$
\sum_{p \in \mathcal{P}_n} q^{\text{area}(p)} t^{\text{dinv}(p)}
$$

is a symmetric polynomial with respect to $q, t$.

Unlike the case of Dyck paths, this problem was not solved at all, so we have no algebraic background that would guarantee that the polynomial is indeed symmetric. However, the conjecture was affirmed to be true up to $n = 10$.

Proving the symmetry conjecture of parking functions combinatorially might shed some light on how the solution to The Symmetry Problem of Dyck paths look like, and hence of major importance to this work.
4.2 Definition

A parking function of size \( n \) is a word \( c_1c_2 \ldots c_n \) such that every letter \( 1 \leq c_i \leq n \), and in addition \( \{|k : c_k \leq i| \} \geq i \). Let \( P_n \) denote the collection of all parking functions of size \( n \).

The following is a well-known result [4, p. 77]:

**Proposition 4.1.** \(|P_n| = (n + 1)^{n-1}\).

One way to represent a parking function is by a labeled Dyck path on a \( n \times n \) grid, where the number of \( N \)-steps in the \( i \)-th column correspond to the number of \( k \)'s such that \( c_k = i \). Due to the fact that \( \{|k : c_k \leq i| \} \geq i \), we are guaranteed that the path on the grid would never go below the diagonal \( y = x \). Furthermore, we write next to each \( N \)-step in the \( i \)-th column the index \( k \) such that \( c_k = i \), where numbers in the same column are always ordered in descending order along that column.

**Definition 4.2.** Let \( p \in P_n \) such that \( p = c_1c_2 \ldots c_n \). We call \( c_i \) the \( i \)-th car of \( p \), and define \( \text{occupant}(j) = i \) if the car \( i \) is located in row \( j \) in the Dyck path representation [4, p. 78]. We understand \( \text{occupant}(i) \) to be the car index we assign to the \( N \)-step of the \( i \)-th row.

![Dyck path form of the parking function 57113515](image)

**Figure 4.1:** The Dyck path form of the parking function 57113515. We can see that \( \text{occupant}(1) = 3, \text{occupant}(2) = 4, \text{occupant}(3) = 7, \ldots, \text{occupant}(8) = 2 \).

We might refer to the parking functions as words, but more often we will consider them as Dyck paths with car indices written to the right of every \( N \)-step.

**Definition 4.3.** Given \( p \in P_n \), we define \( \text{read}(p) \), or the reading word of \( p \), to be the permutation obtained by reading the car’s indices along the diagonal in South-East direction, starting from the diagonal farthest from the line \( y = x \), and continuing inwords.

According to figure 4.1, we can see that \( \text{read}(57113515) = 87265413 \).
4.3 Statistics for Parking Functions

4.3.1 Area

For a given \( p \in P_n \), we define \( \text{area}(p) \) to be the number of complete unit squares locked between the Dyck path \( \pi \in L_{n,n}^\ast \) of \( p \) and the diagonal \( y = x \). Similarly, we define \( \overline{\text{area}}(p) = (a_1, a_2, \ldots, a_n) \) to be the area vector of the parking function \( p \), where \( a_i \) denoted the number of complete unit squares locked between \( p \) and the diagonal \( y = x \) at the \( i \)-th row (where the bottom row is the first). Thus, \( \text{area}(p) = \sum_{i=1}^{n} a_i \). In the example given in figure 4.1, \( \overline{\text{area}}(p) = (0, 1, 2, 1, 0, 1, 2, 1) \), so \( \text{area}(p) = 8 \).

4.3.2 dinv

For a given \( p \in P_n \), we define \( \text{dinv}(p) \) to be the number of pairs of rows of \( p \) of the same length, with the row above containing the larger occupant, or which differ by one in length, with the longer row below the shorter, and the longer row has the larger occupant. This yields the definition

\[
\text{dinv}(p) = \left| \left\{ (i, j) : 1 \leq i < j \leq n \quad a_i = a_j \text{ and } \text{occupant}(i) < \text{occupant}(j) \right\} \right|
\]

\[+ \left| \left\{ (i, j) : 1 \leq i < j \leq n \quad a_i = a_j + 1 \text{ and } \text{occupant}(i) > \text{occupant}(j) \right\} \right| \]

In the example in figure 4.1 we can see that \( \text{dinv}(p) = 7 \) since we have the consider the pairs \((3, 8), (7, 8), (3, 7), (4, 6), (2, 6), (3, 4) \) and \((2, 4) \).

**Proposition 4.4.** Let \( p \in P_n \) be a parking function with a Dyck path of the shape of \( \pi \in L_{n,n}^\ast \). Then, \( \text{dinv}(p) = \text{dinv}(\pi) \leftrightarrow \text{read}(p) = n\cdots21 \).

**Proof.** This is a remark by Haglund [4, p. 79]. Since \( \text{read}(p) = n\cdots21 \), then we are guaranteed that for any \( i, j \) if \( a_i = a_j + 1 \) then \( \text{occupant}(i) > \text{occupant}(j) \), meaning that the longer row among any \( i, j \) would always have a bigger \( \text{occupant} \). Also, for any \( i, k \) if \( i < k \) and \( a_i = a_k \), then \( \text{occupant}(i) < \text{occupant}(k) \). So \( \text{dinv}(\pi) = \text{dinv}(p) \).

First note that \( \text{dinv}(p) \leq \text{dinv}(\pi) \), since it has constrains over \( \text{occupant}(i) \) and \( \text{occupant}(j) \), and \( \text{dinv}(\pi) \) has not. If \( \text{read}(p) = \mu_1\mu_2\cdots\mu_n \neq n\cdots21 \), then there’s \( 1 \leq i \leq n - 1 \) for which \( \mu_i < \mu_{i+1} \). Let \( \text{occupant}(j_1) = \mu_i \) and \( \text{occupant}(j_2) = \mu_{i+1} \). Since \( \mu_i < \mu_{i+1} \) it can’t be the case that those indices are in the same column (car indices are ordered in a descending order in every column), so \( a_{j_1} = a_{j_2} \) or \( a_{j_1} = a_{j_2} + 1 \). If \( j_1 > j_2 \) then we get \( a_{j_1} = a_{j_2} \) but \( \text{occupant}(j_1) < \text{occupant}(j_2) \) so we have added to \( \text{dinv}(\pi) \) but not to \( \text{dinv}(p) \). If \( j_1 > j_2 \) then \( a_{j_1} = a_{j_2} + 1 \) but \( \text{occupant}(j_1) < \text{occupant}(j_2) \) and so we have the pair...
(j_2, j_1) in dinv(\pi) that is not in dinv(p). We get from this dinv(p) < dinv(\pi) and hence dinv(p) \neq dinv(\pi).

\[\square\]

### 4.3.3 Pmaj

Before we turn to define the \( p_{maj}(p) \) it would be helpful to begin by defining the \( pmaj \)-parking order of a parking function, that is denoted by \( \beta(p) \). Let \( C_i = C_i(p) \) denote the set of cars in the \( i \)-th column of \( p \). Let \( \beta_1 \) be the largest car in \( C_1 \), so we park \( \beta_1 \) at spot 1. Then, we let \( C_2' = C_1 \cup C_2 \setminus \{ \beta_1 \} \), and let \( \beta_2 \) be the largest car in \( C_2' \) that is smaller than \( \beta_1 \) (if such does not exist, let \( \beta_2 \) be the biggest car in \( C_2' \)), and park \( \beta_2 \) at spot 2. Then, let \( C_3' = C_2' \cup C_3 \setminus \{ \beta_2 \} \), and let \( \beta_3 \) be the largest car in \( C_3' \) that is smaller than \( \beta_2 \) (if such car doesn’t exist let \( \beta_3 \) be the largest car in \( C_3' \)), park \( \beta_3 \) at spot 3 and continue until you create the sequence \( \beta(p) = \beta_1 \beta_2 \beta_3 \ldots \beta_n \). Then, we define the statistic \( p_{maj}(p) = maj(\text{rev}(\beta(p))) = maj(\beta_n \beta_{n-1} \ldots \beta_2 \beta_1) \), where \( maj \) stands for the major index

\[
maj(\sigma) = \sum_{\sigma(k) > \sigma(k+1)} k
\]

### 4.4 The Symmetry Problem for Parking Functions

The task of the following sections would be to try to find a bijection \( \psi : P_n \to P_n \) such that \( area(p) = dinv(\psi(p)) \) and \( dinv(p) = area(\psi(p)) \).

Similar to the case of Dyck paths, the bijection \( \psi \) seems to be rather complex. However, we were capable of finding \( \psi \) for cases where \( area(p) = 0 \) (see section 4.6), where \( area(p) + dinv(p) = \binom{n}{2} \) (see section 4.7) and where \( area(p) + dinv(p) = \binom{n}{2} - 1 \) (see section 4.8). Contrary to the guess of Haglund [4, page 82, remark 5.5], proving the symmetry

\[
\sum_{p \in P_n} q^{dinv(p)} t^{area(p)}
\]

did involve a bijection over statistics of permutations, but did not involve a generalization of Foata’s bijective transformation (\( maj \) to \( INV \)) [3].

Moreover, one could expect to solve the following conjecture.

**Conjecture 4.5.** Let \( p \in P_n \). Then, for any \( 1 \leq i \leq n \), prove that

\[
\sum_{p \in P_n} q^{dinv(p)} t^{area(p)}
\]

occupant(1)=i

\[\square\]
is a symmetric polynomial in $q,t$.

Conjecture 4.5 was affirmed to be true for all $n$ up to $n = 9$. On the other hand, we could find a counter example that shows that it is not true that for all $i,j$

$$\sum_{p \in P_n} q^{\text{dinv}(p)} t^{\text{area}(p)}$$

is a symmetric polynomial in $q,t$ (one counter example appears on appendix C).

### 4.5 Parking Functions with $\text{Dinv} = 0$

Now we turn to discuss properties of parking functions $p \in P_n$ which satisfy $\text{dinv}(p) = 0$. According to a result by Haglund [4, p. 84], we know that

$$\sum_{p \in P_n} p^{\text{pmaj}(p)} = \sum_{p \in P_n} q^{\text{area}(p)} = \sum_{p \in P_n} p^{\text{dinv}(p)}$$

But when we consider the set of all parking functions $p$ that satisfy $\text{area}(p) = 0$ we get all possible permutations in $S_n$, so $|\{p \in P_n : \text{area}(p) = 0\}| = n!$. This is because it corresponds to labeling every single $N$-step in the diagram with one of the elements of $\{1, 2, \ldots, n\}$ without repeating the same number twice (see figure 4.2).

![Figure 4.2: The only path with $\text{area}(p) = 0$ must be the staircase path, and the parking functions that describe such a path are elements of $S_n$](image)

We now turn to consider the constrains that would allow a path of $p$ to have $\text{dinv}(p) = 0$.

**Lemma 4.6.** Let $p \in P_n$ such that $\text{dinv}(p) = 0$ and $\text{area}(p) = (a_1, a_2, \ldots, a_n)$. Then, $a_1 \leq a_2 \leq \ldots \leq a_n$.

**Proof.** Assume for contradiction that there’s a decrease somewhere in the sequence, namely integers $1 \leq i' < j' < k \leq n$ such that $a_{j'} > a_k$ and $a_{j'} > a_{i'}$. Thus we can find $i \leq i'$ and $j \leq j'$ so that $a_i = a_k = a_j - 1$ (and $1 \leq i < j < k \leq n$). Notice that we can
chose such $i,j$ that $\text{occupant}(i)$ and $\text{occupant}(j)$ would lie on the same column. Since $dinv(p) = 0$ and that $i < k$ and $a_i = a_k$ then we must have $\text{occupant}(k) < \text{occupant}(i)$ (otherwise $dinv(p) \geq 1$). Also, since $j < k$ and $a_j = a_k + 1$ then we must have that $\text{occupant}(j) < \text{occupant}(k)$ (also otherwise $dinv(p) \geq 1$). If we combine the two we get $\text{occupant}(j) < \text{occupant}(k) < \text{occupant}(i)$, so $\text{occupant}(j) < \text{occupant}(i)$ for $i < j$ which contradicts the fact that on every column the car indices are ordered in descending order.

\[ \square \]

Let $\pi \in L_{n,n}^+$ be the Dyck path that corresponds to a path of $p \in P_n$, where $dinv(p) = 0$. Lemma 4.5 implies that we the sequence of steps "EEN" can’t exist in $\pi$.

**Definition 4.7.** An $\text{upread}(p)$ of a parking function is the permutation given by reading the car indices of the parking function from the bottom line to the top one. So, $\text{upread}(p) = \text{occupant}(1) \text{occupant}(2)\ldots\text{occupant}(n)$.

For instance, in the following figure we can see that $\text{upread}(111341) = 123654$:

![Figure 4.3: upread(111341) = 123654](image)

If we know that for $p$ that satisfy $dinv(p) = 0$, the area vector $a_1a_2\ldots a_n$ is weakly increasing, then we know that there are at most $2^{n−1}$ possible paths (since for any $1 \leq i \leq n − 1$ $a_{i+1} = a_i$ or $a_{i+1} = a_i + 1$). I now turn to describe each of the possible $2^{n−1}$ has a legitimate ordering of cars that would satisfy that $dinv(p) = 0$. The main idea is to first force $\text{upread}(p) = 12\ldots n$. Then, for any chains where $i_1 < i_2 < \ldots < i_k$ and $a_{i_1} = a_{i_2} = \ldots = a_{i_k}$ we switch the labeling so that $\text{occupant}(i_k) = \text{occupant}(i_{k−b+1})$. See figure 4.4:

The labeling described above guarantees that $dinv(p) = 0$, since $a_1 \leq a_2 \leq \ldots \leq a_n$ then for $i < j$ we have $a_i \leq a_j$, and if there is $i < j$ such that $a_i = a_j$ then we will get $\text{occupant}(j) < \text{occupant}(i)$, which doesn’t contribute to $dinv(p)$.

We can define the $B_n := \{ p \in P_n : \text{dinv}(p) = 0 \}$ and $\varphi : B_n \to B_n$ as,

$\varphi(p) = q$ such that:

\[
\text{area}(p) = (a_1,a_2,a_3,\ldots,a_n) \rightarrow \text{area}(q) = (a_1,1-a_2,2-a_3,\ldots,n-1-a_n)
\]

$\text{occupant}(i)_{of\; q}$ is $n+1-\text{occupant}(i)_{of\; p}$.
Chapter 4. Parking Functions

Figure 4.4: a weakly decreasing Dyck path is labeled first by order and the mirroring equivalent area rows to generate \( \text{dinv}(p) = 0 \)

Figure 4.5: \( \varphi \) changes the path of \( p \) in such a way that a missing area square of every row becomes the area squares of \( q \), and mirrors the car indices.

Lemma 4.8. \( \varphi \) is an involution, so \( \varphi(\varphi(p)) = p \).

Proof. This is because \( \text{area}(\varphi(\varphi(p))) = \text{area}(p) \), since for all \( i \), we get that \( i - 1 - (i - 1 - a_i) = a_i \). But also because, \( \text{occupant}(i)_{p} = \text{occupant}(i)_{\varphi(\varphi(p))} \), since \( n + 1 - (n + 1 - \text{occupant}(i)) = \text{occupant}(i) \). These two entail that \( \varphi(\varphi(p)) = p \).

\( \Box \)

Let \( p, q \in P_n \) such that \( \varphi(p) = q \). Notice that given \( i_1 < i_2 < \ldots < i_k \) and such that \( a_{i_1} = a_{i_2} = \ldots = a_{i_k} \) in \( p \), would turn to \( a_{i_1} < a_{i_2} < \ldots < a_{i_k} \) in \( q \). Also, given \( i_1 < i_2 < \ldots < i_k \) and such that \( a_{i_1} < a_{i_2} < \ldots < a_{i_k} \) in \( p \), would turn to \( a_{i_1} = a_{i_2} = \ldots = a_{i_k} \) in \( q \).

This is why when we move \( \text{occupant}(i) \mapsto n + 1 - \text{occupant}(i) \) we make sure we turn an increasing sequence to decreasing sequence, and vice versa (so that if \( \text{dinv}(p) = 0 \), then \( \text{dinv}(q) = 0 \)).

From all of the above we can conclude that \( |B_n| = n! \), and that there are exactly \( 2^{n-1} \) distinct lattice paths. Also, given a certain \( \pi \in L^+_{n,n} \), if there are \( m \) different parking functions \( p_1, \ldots, p_m \in P_n \) such the path of each \( p_i \) is \( \pi \) and \( \text{dinv}(p_i) = 0 \), then it means that there are precisely \( m \) different parking functions \( \varphi(p_1), \varphi(p_2), \ldots, \varphi(p_m) \) such that they all have the same Dyck path \( \pi' \in L^+_{n,n} \) and \( \text{dinv}(\varphi(p_i)) = 0 \).
Let $\lambda_1, \lambda_2, \ldots, \lambda_{2n-1}$ be subsets of $B_n$ such that for all $i$, if $p, q \in \lambda_i$ then $p$ and $q$ have the same Dyck path $\pi \in L_{n,n}^+$. We basically group all parking functions of $B_n$ according to different paths $\pi \in L_{n,n}^+$ they have, to get a partitioning $\lambda_1, \lambda_2, \ldots, \lambda_{2n-1}$ of $B_n$, such that $\forall i \mid \lambda_i \geq 1$. This would mean that $\sum_{i=1}^{2n-1} |\lambda_i| = n!$ because for all $i \neq j$, $\lambda_i \cap \lambda_j = \emptyset$. From the fact that $\varphi$ is an involution we get that for every $i$, if $\lambda_i = k$, then there must be $j \neq i$ for which $|\lambda_j| = k$. Here are $|\lambda_1|, |\lambda_2|, \ldots, |\lambda_{2n-1}|$, for $1 \leq n \leq 5$:

<table>
<thead>
<tr>
<th>$n$</th>
<th>different shapes of parking functions in $P_n$ with $\text{dinv} = 0$</th>
<th>$\sum$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1!</td>
</tr>
<tr>
<td>2</td>
<td>1, 1</td>
<td>2!</td>
</tr>
<tr>
<td>3</td>
<td>1, 2, 2, 1</td>
<td>3!</td>
</tr>
<tr>
<td>4</td>
<td>1, 3, 3, 5, 5, 3, 3, 1</td>
<td>4!</td>
</tr>
<tr>
<td>5</td>
<td>1, 4, 4, 6, 9, 9, 11, 16, 16, 16, 11, 9, 6, 4, 4, 1</td>
<td>5!</td>
</tr>
</tbody>
</table>

### 4.6 Bijection for $p \in P_n$ with $\text{area}(p) = 0$

Given a parking function $p \in P_n$ with $\text{area}(p) = 0$ we can identify $\text{dinv}(p) = \text{INV}(\text{read}(p))$, which is the number of inversions of the permutation given from when we read the indices of the cars in South-East direction. For instance let $p = 51837426$,

![Figure 4.6: We can see that $\text{dinv}(p) = \text{INV}(35816472) = 14$](image)

In the following we will describe a function $f_{\text{area}=0}$ that transfers parking functions with $\text{area}(p) = 0$ to parking functions whose $\text{dinv}(p) = 0$. The function $f_{\text{area}=0}$ is in fact a composition of a function $\gamma : S_n \rightarrow S_n$ that permutes the car indices, together with a path alternating function. We begin by describing the function $\gamma$ and later will prove why the function $f_{\text{area}=0}$ could serve a bijection between the sets $\{p \in P_n : \text{area}(p) = 0\}$ and $\{p \in P_n : \text{dinv}(p) = 0\}$.

Let $\sigma = \sigma_1 \sigma_2 \ldots \sigma_n$ be a permutation in $S_n$, and let us denote $\gamma(\sigma) = \tau \in S_n$ where $\tau = \tau_1 \tau_2 \ldots \tau_n$. Set $\tau_1 = \sigma_1$, $T_1 = \{1, 2, \ldots, n\}$ and $\Sigma_1 = \{1, 2, \ldots, n\} \setminus \{\sigma_1\}$. In order to get $\tau_2$ we need to jump cyclically from $\tau_1$ in $T_1$ as many steps $k_1 = |\{\sigma_1 \in \Sigma_1 : \sigma_i \leq \sigma_2\}|$
(which is the absolute position of $\sigma_2$ in $\Sigma_1$). In other words, $\tau_2 = \tau_1 + k_1 \mod n$. Set $\Sigma_2 = \Sigma_1 \setminus \{\sigma_2\}$ and $T_2 = T_1 \setminus \{\tau_1\}$. In order to get $\tau_3$ we jump $k_2 = |\{\sigma_i \in \Sigma_2 : \sigma_i \leq \sigma_3\}|$ many steps from $\tau_2$ in $T_2$ (cyclically, i.e., modulo the number of elements in $T_2$). So $\tau_3 = \tau_2 + k_2$ (modulo the elements of $T_2$). We continue in this manner, and set $T_{j+1} = T_j \setminus \{\tau_j\}$. $\Sigma_{j+1} = \Sigma_j \setminus \{\sigma_{j+1}\}$ and $k_{j+1} = |\{\sigma_i \in \Sigma_{i+1} : \sigma_i \leq \sigma_{j+1}\}|$, in order to find all values of $\tau_4, \ldots, \tau_n$. Since when we pick $\tau_i$, we skip elements of $\{1, 2, \ldots, n\}$ we selected previously, we are guaranteed that $\tau \in S_n$.

**Example 4.1.** Let us consider the output of $\gamma(351642)$. We know that $\tau_1 = 3$, as $\sigma_1 = 3$. Set $T_1 = \{1, 2, 3, 4, 5, 6\}$, and $\Sigma_1 = \{5, 1, 6, 4, 2\}$. Now $\sigma_2 = 5$ is the fourth largest number in $\Sigma_1$, hence $k_1 = |\{\sigma_i \in \Sigma_1 : \sigma_i \leq 5\}| = 4$, and we need to make 4 "jumps" from the element '3' in $T_1$, so $\tau_2 = 1$ (since we "jumped" over 4, 5, 6, 1 in $T_1$). Now, set $T_2 = \{1, 2, 4, 5, 6\}$ and $\Sigma_2 = \{1, 6, 4, 2\}$. In the next step, $\sigma_3 = 1$ is the smallest number in $\Sigma_2$, so $k_2 = |\{\sigma_i \in \Sigma_2 : \sigma_i \leq 1\}| = 1$, and we need to make one "jump" from the element $\tau_2 = 1$ in $T_2$, so we end up with $\tau_3 = 2$. Next, we set $\Sigma_3 = \{6, 4, 2\}$ and $T_3 = \{2, 4, 5, 6\}$. Now, $\sigma_4 = 6$ is the third largest number in $\Sigma_3$, so $k_3 = 3$, so we need to make 3 "jumps" from the element 2 in $T_3$ and we get that $\tau_4 = 6$ (as we "jumped" 4, 5, 6). Next, we have $\Sigma_4 = \{4, 2\}$ and $T_4 = \{4, 5, 6\}$, and $\sigma_5 = 4$ is the second largest in $\Sigma_4$, so $k_4 = 2$, and hence we need to "jump" twice from the element $\tau_4 = 6$ in $T_4$, and reach the number 5. Thus $\tau_5 = 5$. Finally $\tau_6 = 4$ since there are no more options left. Thus, $\gamma(351642) = 312654$.

![Figure 4.7: $\gamma(351642) = 312654$](image)

The function $f|_{\text{area}=0}$ itself is a composition of $\gamma$ together with an appropriate path. If $p \in P_n$, we let $f|_{\text{area}=0}(p) = q$ where $q \in P_n$ as well. In order to construct $q$, we first let $\sigma = \text{read}(p)$, and construct a $q$ such that $\text{up\!read}(q) = \gamma(\sigma)$. In order to assure $\text{dinv}(q) = 0$ we force the path of $q$ to be such that two indices would have the same column if the the next one is bigger than the previous, otherwise the next one would be exactly one column to the right.

**Remark 4.9.** We are guaranteed to have $\text{dinv}(f|_{\text{area}=0}(p)) = 0$ because $a_1, a_2, \ldots, a_n$ of $f|_{\text{area}=0}(p)$ is a monotonous increasing sequence, and if two rows have the same area value $a_i = a_j$ for some $i < j$, then $\text{occupant}(i) > \text{occupant}(j)$. 
Theorem 4.10. The function $\gamma : S_n \to S_n$ satisfies $\text{INV}(\text{rev}(\sigma)) = \text{maj}(\text{rev}(\gamma(\sigma)))$.

Proof. We claim that given the function $f_{\text{area}=0}$ described above, $\text{area}(f_{\text{area}=0}(p))$ is the same as $\text{maj}(\text{rev}(\text{upread}(f_{\text{area}=0}(p))))$. Let $\text{rev}(\text{upread}(f_{\text{area}=0}(p))) = \sigma_1 \sigma_2 \ldots \sigma_n$, and let $i_1, i_2, \ldots, i_m$ be the indices such that $\sigma_{i_k} > \sigma_{i_k+1}$. So, $\text{maj}(\sigma_1 \sigma_2 \ldots \sigma_n) = \sum_{k=1}^{m} i_k$.

On the other hand, for $1 \leq k \leq n$ let $j_k$ be such that $\text{occupant}(j_k) = \sigma_k$. Now, we know that for $i_1, i_2, \ldots, i_m$ we have $\sigma_{i_k} > \sigma_{i_k+1}$. This implies that $a_{i_k} + 1 = a_{i_k+1}$ (meaning that the numbers $\sigma_{j_k}$ and $\sigma_{j_{k+1}}$ lie in the same column). Since $a_1 \leq a_2 \leq \ldots \leq a_n$, we can conclude that $a_k^* = i_k$ (according to section 2.1), where $k = a_{j_k}$ (the distance of the index $\sigma_{i_k}$ from the diagonal) and hence $\text{area}(p) = \sum_{k=1}^{m} a_k^* = \sum_{k=1}^{m} i_k$.

Moreover, $\text{INV}(\text{rev}(\sigma))$ is precisely $\text{dinv}(p)$. We are left to prove that we have the equality $\text{dinv}(p) = \text{area}(f_{\text{area}=0}(p))$, but this would be proven in Theorem 4.11. ∎

Theorem 4.11. The function $f_{\text{area}=0} : \{p \in P_n : \text{area}(p) = 0\} \to \{p \in P_n : \text{dinv}(p) = 0\}$ described above satisfies $\text{dinv}(p) = \text{area}(f_{\text{area}=0}(p))$.

Proof. We let $\text{rev}(\text{read}(p)) = \sigma = \sigma_1 \sigma_2 \ldots \sigma_n$. From this we can see that $\text{INV}(\text{rev}(\sigma)) = \text{INV}(\text{rev}(\text{rev}(\text{read}(p)))) = \text{INV}(\text{read}(p)) = \text{dinv}(p)$. Let $\gamma(\sigma) = \tau = \tau_1 \tau_2 \ldots \tau_n$, and note that $\text{upread}(f_{\text{area}=0}(p)) = \tau$. At every step $i$ in our algorithm of the function $\gamma$, we consider $\sigma_i$, and check the size of $k_i = |\{j : \sigma_j \leq \sigma_i \text{ and } i < j\}|$ which provides us with the indication that there are $n - i - k_i$ number of $\sigma_j$ such that $i < j$ and $\sigma_i < \sigma_j$. In fact, this is the indication of the contribution of $\sigma_i$ to $\text{dinv}(p)$. Then, the $\tau_i$ that we pick at the $i$th stage, is the one that guarantees exactly additional $n - i - k_i$ area units at $f_{\text{area}=0}(p)$. This is because when we place $\tau_1$ we promised to have $n - \tau_1$ area units above the diagonal (the minimal area-contribution case is when $\tau_1$ is followed by $\tau_1 - 1, \tau_1 - 2$ etc, since they’ll all be places in the same diagonal). In a similar manner for all other $\tau_i$. These units are located always in the diagonal above the location of where we write $\tau_i$ in the path. So, if $\tau_{i+1} < \tau_i$ then $\tau_i$ wouldn’t have any contribution to area($f_{\text{area}=0}(p)$), but if $\tau_{i+1} > \tau_i$ then the indices $\tau_i$ and $\tau_{i+1}$ would be written under the same column and there would be exactly $n - i$ area units spread diagonally above $\tau_i$ (see figure 4.9).
Figure 4.9: if $\tau_{i+1} < \tau_i$ then we have no contribution in terms of area units in $f_{\text{area}=0}(p)$, but if $\tau_{i+1} > \tau_i$ then we will have new $n-i$ area units located diagonally over $\tau_i$.

This construction of $\tau = \tau_1\tau_2\ldots\tau_n$ guarantees that $\text{dinv}(p) = \text{area}(f_{\text{area}=0}(p))$.

Remark 4.12. Theorem 4.11 implies a certain bijection between the $\text{inv}$ statistic to the $\text{maj}$ statistic in permutations. Although we came with the proof of Theorem 4.11 ourselves, it occurred to us that Vincent Vajnovszki published the same bijection before us, using a different proof that does not use parking functions [10].

4.7 Bijection for $p \in P_n$ with $\text{dinv}(p) + \text{area}(p) = \binom{n}{2}$

In this section we will provide a bijection

$$\vartheta : \left\{ p \in P_n : \text{area}(p) + \text{dinv}(p) = \binom{n}{2} \right\} \rightarrow \left\{ p \in P_n : \text{area}(p) + \text{dinv}(p) = \binom{n}{2} \right\}$$

such that $\text{area}(p) = \text{dinv}(\vartheta(p))$ and $\text{dinv}(p) = \text{area}(\vartheta(p))$.

Lemma 4.13. For all $p \in P_n$, $\text{area}(p) + \text{dinv}(p) \leq \binom{n}{2}$

Proof. Let $p \in P_n$ such that $\overline{\text{area}}(p) = (a_1, a_2, \ldots, a_n)$. Obviously, for all $i$, $0 \leq a_i \leq i - 1$ and $a_1 = 0$. Fix $k$ and check how many rows below the $k$th row may possibly have area value of $a_k$ or $a_k + 1$ (in order to consider in how many rows could possibly the $k$th row contribute to the $\text{dinv}$ where the $k$th row is the upper one in the $\text{dinv}$-pair). Since $a_k \leq a_{k-1} + 1 \leq a_{k-2} + 2 \leq \ldots \leq a_k + i$, we can see that $1 \leq i \leq k$ and $a_{k-1} + i \leq k - 1 - a_k$ if we assume that all possible rows below $k$ have the same area value $a_k$. This yields at most $k - 1 - a_k$ different values of $i$, which means $k - 1 - a_k$ lines below $k$ that could possibly have area as $a_k$. Hence,

$$\text{area}(p) + \text{dinv}(p) \leq \sum_{k=1}^{n} a_k + \sum_{k=1}^{n} (k - 1 - a_k) = \sum_{k=1}^{n} (k - 1) = \binom{n}{2}$$
Remark 4.14. The graphical interpretation of the previous lemma in $L_{n,n,k}$, could be illustrated by a staircase path going down from the edge of the left most area square at row $k$. Such a staircase path would have exactly $k-1-ak$ stairs until it would reach the leftmost edge of the grid.

Definition 4.15. A lattice path $\pi \in L_{n,n}^+$ is called optimal, if there exists a parking function $p \in P_n$, such that $p$’s path is $\pi$, and $\text{area}(p) + \text{dinv}(p) = \binom{n}{2}$.

Lemma 4.16. For $p, q \in P_n$, if $p$ and $q$ are optimal and satisfy $\text{area}(p) = \text{area}(q) = c$ then $p$ has the same path as $q$.

Proof. If $a_k$ stands for the area of the path of $p$ at the $k$th row, where $0 \leq a_k \leq k-1$ and the first row is the bottom row, then we would like to assure that every row $k$ has exactly $k - 1 - a_k$ rows below it with area $a_k$ or $a_k + 1$. This means that either the a certain row $k$ has precisely $a_k = k - 1$ as area, and in such a case it means that all rows below it must have $\forall i \leq k \ (a_i = i - 1)$. Otherwise, if the row $k$ satisfy $a_k < k - 1$, then if a row $i < k$ has at some point $a_i = a_k + 1$, then all rows below $i$ must have $a_j = a_i$ or $a_j = a_i + 1$, and so on, until we reach a row that satisfy $a_s = s - 1$. So if a certain row $k$ satisfy $a_k = k - 1$ then immediately we know that all rows below it fill in all possible area slots (similar to Lemma 1.7). Now let $k$ be the highest row so that $a_k = k - 1$, and from this we can see that $\forall i \geq k \ : a_i = a_k$ or $a_i = a_k + 1$, so that the area is slowly decreasing over the diagonal. So there is a unique way to place area on the path of $p$ so that we will have $\text{area}(p) = c$.

Remark 4.17. The last lemma tells us that we could consider the process of placing the area unit squares along the diagonals (from bottom to top), in order to generate a path $p$ which is optimal with $\text{area}(p) = c$. According to that process, since $0 \leq c \leq \binom{n}{2}$, then if $c \geq n - 1$ then we will in the first diagonal, if $c - (n - 1) \geq n - 2$ then we fill in the second, if $c - (n - 1) - (n - 2) \geq n - 3$ then we will in the first diagonal, and so on. We continue filling the diagonal until we reach a certain $i$ for which $c - (n - 1) - \ldots - n + 1 < n - i$, and then we place the remaining $c - (n - 1) - \ldots - (n - i + 1)$ area slots along the $i$th diagonal, from bottom to top.

For instance, this is how we generate $p \in P_6$ with $\text{area}(p) = 11$

Theorem 4.18. For a fixed value of $c$, there exists a unique parking function $p \in P_n$ for which $\text{area}(p) = c$ and $\text{dinv}(p) = \binom{n}{2} - c$.

Proof. We construct the path of $p$, in such a way that $p$ would be optimal and $\text{area}(p) = c$ as described in Lemma 4.16 above. We set the indices of the cars of $p$ to be such that
Figure 4.10: We fill in the first two diagonals and the third is filled from bottom to top with the remaining unit squares

read(p) = n⋯21. This guarantees that p would have maximal dinv, which is dinv(p) = \( \binom{n}{2} - c \), since the sum of the dinv(p) + area(p) ≤ \( \binom{n}{2} \). Any other reading word of p would either be illegal (in the sense that the order of the cars sharing the same column would not be decreasing), or could lose some potential dinv. This is because every row c for which it’s area satisfy \( a_c < c - 1 \) would have always rows above it \( c < i \) with \( a_i = a_c \) or \( a_i = a_c - 1 \). This proves uniqueness.

Let \( p \in P_n \) be a parking function with an optimal path, and let \( \text{area}(p) = c \) and \( \text{read}(p) = n⋯21 \). Define \( \vartheta : \{ p \in P_n : \text{area}(p) + \text{dinv}(p) = \binom{n}{2} \} \rightarrow \{ p \in P_n : \text{area}(p) + \text{dinv}(p) = \binom{n}{2} \} \) such that \( \vartheta(p) = q \), where q’s path is optimal, \( \text{area}(q) = \binom{n}{2} - c \) and \( \text{read}(q) = n⋯21 \).

**Theorem 4.19.** The function \( \vartheta \) described above is a bijection that satisfies both \( \text{area}(p) = \text{dinv}(\vartheta(p)) \) and \( \text{dinv}(p) = \text{area}(\vartheta(p)) \).

**Proof.** Let \( \text{area}(p) = c \), and \( c' = \binom{n}{2} - c \). Then, by Theorem 4.18, there is exactly one \( q \) such that \( \text{area}(q) = c' \) and q’s path is optimal. If \( \text{read}(q) = n⋯21 \), then we get that \( \text{dinv}(q) = \binom{n}{2} - c' = \binom{n}{2} - \binom{n}{2} + c = c = \text{area}(p) \).

**Example 4.2.** Let \( p \in P_6 \) with \( \text{area}(p) = 10 \) and \( \text{dinv}(p) = \binom{6}{2} - 10 = 5 \). We can find the equivalent \( q \) with \( \text{area}(q) = 5 \) and \( \text{dinv}(q) = 10 \):

Figure 4.11: \( p,q \in P_n \). The parking function \( p \) is portrayed on the left and \( q \) on the right. \( \text{area}(p) = \text{dinv}(q) = 10 \) and \( \text{dinv}(p) = \text{area}(q) = 5 \).
Remark 4.20. We can consider the bijection given in this section \( \varphi \) to be also a bijection between Dyck paths \( \pi \in L_{n,n}^+ \) such that \( \text{area}(\pi) + \text{dinv}(\pi) = \binom{n}{2} \). The core idea is to map an optimal path \( \pi \in L_{n,n}^+ \) with \( \text{area}(\pi) = c \) to the unique optimal path \( \pi' \in L_{n,n}^+ \) with \( \text{area}(\pi') = \binom{n}{2} - c \). The reason we can do this is because the parking functions we considered in this section have \( \text{read}(p) = n \cdots 21 \), so \( \text{dinv}(p) = \text{dinv}(\pi) \) if \( \pi \) is the path of \( p \) (by Proposition 4.4).

4.8 Bijection for \( p \in P_n \) with \( \text{dinv}(p) + \text{area}(p) = \binom{n}{2} - 1 \)

Definition 4.21. A path \( \pi \in L_{n,n}^+ \) is called almost optimal if \( \pi \) is not optimal and there exists a parking function \( p \in P_n \), such that \( p \)'s path is \( \pi \), and \( \text{area}(p) + \text{dinv}(p) = \binom{n}{2} - 1 \).

Note that optimal paths are necessarily different from almost optimal paths as a demand.

There is a unique way to construct the path \( \pi \in L_{n,n}^+ \) of \( p \in P_n \), if \( \text{area}(p) = c \), \( \pi \) is almost optimal and the value of \( c \) is fixed. The idea is that exactly like in optimal paths, we need that all rows that have full area to have at most the same area or one value less above them. Unlike optimal paths, almost optimal paths require that there will be exactly one row \( i \) for which there is a line above it \( j \), namely \( i < j \), with either (a) \( a_i < a_j \) (\( a_i \) denotes the area at the \( i \)th row) ; or (b) \( a_i = a_j + 2 \).

Regarding the (a) case, since we want exactly one row \( i \) which has a line above it \( i < j \) with \( a_i < a_j \), then we claim that \( a_i = a_j - 1 \). This is because if \( a_i < a_j - 1 \) then there must have been a row \( i < k < j \) such that \( a_i = a_k - 1 \), since area could increase at most in 1 from one row to another. The existence of such \( k \) is a contradiction to the uniqueness of \( i \). So, we must have \( a_j = a_i + 1 \), now we let \( j \) be the smallest row for which \( a_j = a_i + 1 \). If \( j > i + 1 \) then the row \( i + 1 \) is a contradiction to the uniqueness of \( i \), because in such a case either both \( i \) and \( i + 1 \) would have a row \( i, i + 1 < j \) with \( a_i = a_{i+1} = a_j - 1 \).

Remark 4.22. Almost optimal paths of type (a) satisfy that their area vector is \( \overrightarrow{\text{area}}(a) = (0, 1, 2, \ldots, k, k, k, \ldots, k, k - 1, k, k - 1, k - 1, \ldots, k - 1) \) for some \( k \).

For instance, this is how we generate \( p \in P_6 \) with \( \text{area}(p) = 11 \)

An algorithm that describes the area placement of such a path is giving by the following:
Fill the area along the diagonal from bottom to top, where the biggest diagonal comes first. Once one diagonal is filled, moved to the next and begin again from the bottom-left most position. Instead of placing the last area unit square in it’s row, we place it one row above.
As for the (b) case, we need to make sure that there would be exactly one line \( i \) with a line above it \( i < j \) such that \( a_i = a_j + 2 \). All other lines needs to have either all possible area (namely that for a row \( k \), \( a_k = k - 1 \)), or that all lines above them have same or one less area value. This forces line \( i \) to have \( a_i = i - 1 \), and hence it is the highest row for which \( a_i = i - 1 \). This forces the row \( j \) that satisfies \( i < j \) and \( a_j + 2 = a_i \) to be \( j = n \), because if \( j \) is not the last row then we end up with a row \( j \) that has rows with higher area above it, or more than one row above \( j \) with the same area, and in both cases we lose more \( \text{dinv} \) value than 1.

The algorithm for place the area value for paths of type (b) is precisely the same as the ones of type (a). The only difference between the two is that in those of type (b) we get to skip the top most line in the grid. This means that if \( p \in P_n \) is an almost optimal parking function of type (b), then if \( \text{area}(p) = c \) for some \( c \), then there is a positive integer \( t \) for which \( c = (n - 1) + (n - 2) + \ldots + (n - t) \).

For instance, this is how we generate \( p \in P_6 \) with \( \text{area}(p) = 12 \)

**Lemma 4.23.** Let \( 1 \leq c \leq \binom{n}{2} - 2 \) and let \( \pi \in L_{n,n}^* \) be an almost optimal path with \( \text{area}(\pi) = c \). If \( p, q \in P_n \) are two parking function with the same almost optimal path \( \pi \), such that \( \text{area}(p) + \text{dinv}(p) = \text{area}(q) + \text{dinv}(q) = \binom{n}{2} - 1 \), then \( p = q \).

**Proof.** According to Proposition 4.4, \( \text{dinv}(p) = \text{dinv}(\pi) \) if and only if \( \text{read}(p) = n \ldots 21 \) [4, p. 79]. From this we deduce that \( \text{area}(p) = \text{area}(q) = \text{area}(\pi) = c \), and from
area(p) + dinv(p) = area(q) + dinv(q) = \binom{n}{2} - 1 we deduce that dinv(p) = dinv(q) = \binom{n}{2} - 1 - c = dinv(\pi). Then read(p) = read(q) = n \cdots 21, and since area(p) = area(q) we get that p = q.

Fix the value of c, and let us denote by p_n the parking function with area(p_n) = c and dinv(p_n) = \binom{n}{2} - 1 - c.

The previous lemma show that there is a unique way to maximize the dinv value of an almost optimal path. Since almost optimal paths are unique up to selection of area value, and they satisfy area + dinv \leq \binom{n}{2} - 1 then we necessarily created one path. We will now turn to generate the n − 1 additional paths p_1,p_2,\ldots,p_{n-1}, and then to show that there are no other possible paths that could be generated.

Since we saw that there was one way that would assure that dinv + area = \binom{n}{2} - 1 for the case of almost optimal paths, then in order to generate the other n − 1 desired parking functions we know that they must have optimal paths. This is because if they won’t have optimal paths, since they can’t be almost optimal, then they will satisfy area + dinv \leq \binom{n}{2} - 2 which is not what we want.

**Lemma 4.24.** Let 1 \leq c \leq \binom{n}{2} - 2 and let \pi \in L_{n,n}^+ be an optimal path with area(\pi) = c. Then, there are at least n − 1 different parking functions p_1,p_2,\ldots,p_{n-1} who all have the same path \pi, and all satisfy area(p_i) = c and area(p_i) + dinv(p_i) = \binom{n}{2} - 1.

**Proof.** Fix c, and construct an optimal path with area(p) + dinv(p) = \binom{n}{2}, as we did before. Now consider the indices of the cars, and for any 1 \leq i \leq n−1, replace the position of i the the smallest j, such that i < j \leq n and i and j are not both in the same column. Since it is never the case that the indices n − 1 and n share the same column (given our construction of the path), then we are always guaranteed to be able to replace the index i (at most with the indices n − 1 or n). Let p_i \in P_n denote the path created for every, where i is the swapped index (so we actually generate p_1,p_2\ldots p_{n-1}).

We claim that the swapping of the index i with the smallest possible that doesn’t lie in the same column as the index i itself should substruct exactly one from the value of the maximal dinv of p. This is because, leaving all other relations the same, when we swap i with j, we avoid the existed (i,j) − dinv; all other indices that were above and j with higher value than j would remain above i with higher value than i, and similarly, all indices that there below i that i had a bigger value from, would now turn to be below j, where j would have a bigger value from them.
**Example 4.3.** In the following we will illustrate how to construct 5 different paths \( p_1, p_2 \ldots p_5 \), such that \( \forall 1 \leq i \leq 5 \) we have \( \text{area}(p_i) = 10 \) and \( \text{dinv}(p_i) = \left( \binom{n}{2} \right) - 1 - 10 = 4 \)

![Diagram of parking functions](image)

Figure 4.14: All paths above share \( \text{area} = 10 \), but while the upper left one have \( \text{dinv} = 5 \), all the others are constructed from swaps of \( i = 1, 2, \ldots, 5 \) to create \( p_i \) with \( \text{dinv}(p_i) = 5 - 1 = 4 \)

So far we composed the \( n \) different parking functions \( p_1, p_2, \ldots, p_n \). In order to complete our theorem, we are required to explain why we could not compose any more paths. Since paths that satisfy \( \text{dinv} + \text{area} = \left( \binom{n}{2} \right) - 1 \) must be either optimal paths or almost optimal, and we saw that there was only one way to permute the cars so that almost optimal paths would have their maximal \( \text{dinv} \), we need only to show why we could not construct more than \( n - 1 \) parking functions with optimal paths as such.

**Lemma 4.25.** Let \( 1 \leq c \leq \left( \binom{n}{2} \right) - 2 \) and let \( \pi \in L_{n,n}^+ \) be an optimal path with \( \text{area}(\pi) = c \). Then, there are at most \( n - 1 \) different parking functions \( p_1, p_2, \ldots, p_{n-1} \) who all have the same path \( \pi \), and all satisfy \( \text{area}(p_i) = c \) and \( \text{area}(p_i) + \text{dinv}(p_i) = \left( \binom{n}{2} \right) - 1 \).

**Proof.** Consider the general schema on figure 4.15.

This schema represents the case of an optimal path, where \( \text{dinv} + \text{area} = \left( \binom{n}{2} \right) \), as the given parking function \( p \) is optimal and satisfies \( \text{read}(p) = n \ldots 21 \). Let \( A, B \) and \( C \) be the set of car indices of the schema, namely \( A = \{1, 2, 3, \ldots k\} \), \( B = \{n - r, \ldots, n - 2, n - 1, n\} \) and \( C = \{k + 1, k + 2, \ldots, n - r - 2, n - r - 1\} \), for some fixed \( k \) and \( r \) (as the path itself is fixed). We want to alter the indices of the optimal path \( p \) in order to get a new parking function with \( \text{area} + \text{dinv} = \left( \binom{n}{2} \right) - 1 \). Hence we will denote the new car indices of the current \( A, B, C \) with \( A', B', C' \) respectively.
There are 4 distinguished cases.

Case 1: Assume that the indices of $C'$ has been changed, namely that $C' \neq C$. Then the new indices must be $C' = \{n, n-1, \ldots, n-r+1, n-r-1\} = C\setminus\{n-1\} \cup \{n-r-1\}$, and this is because all the elements of $C'$ must be both bigger than the elements of $A'$ (since they least element of $C'$ lies in the same column as $A'$) and the elements of $B'$ (in order to subtract at most 1 from the $dinv$). In this case, the elements of $C'$ must be ordered in descending order, in order not to subtract more than 1 from $dinv$, and similarly the elements of $B'$ are sorted in descending order (obviously the elements of $A'$ too, since they share the same column). Hence, the elements of $B'$ must be $B' = \{n-r, n-r-1, n-r-2, \ldots, k+1\}$, because only in this case there is exactly one element in $B'$ which is bigger than exactly one element in $C'$. Thus, if $C' \neq C$, then we must get this unique described case.

Case 2: Now assume $C' = C$, as a set. If $B'$ remains unchanged (namely $B' = B$ as a set and the indices are not permuted), then we can create at most $r$ different parking functions that would take off 1 from the maximal $dinv$, by considering all swaps between the adjacent elements of $\{n, n-1, \ldots, n-r\}$, since that set contains $r+1$ elements and we can have $r$ different swaps between adjacent elements. Note that we can permute elements of $C'$ or of $B'$ but not both at the same time, and still subtract exactly 1 from the $dinv$.

Case 3: Now assume that the elements of $C'$ remain unchanged, namely that $C' = C$, not only as a set but also in the same order. For a similar argument as for case 2, if
Chapter 4. Parking Functions

$p' = B$ as a set, then we obtain at most $(n - r - 1) - (k + 1) + 1 = n - r - k - 2$ different swaps among the elements of $B'$ that would take away at most $1 \text{dinv}$.

Case 4: Assume again that $C'$ remains unchanged, and let $B' = B\times = B\{k + 1\} \cup \{x\}$, for $1 \leq x \leq k$. So $B' = \{n - r - 1, n - r - 2, \ldots, n - r - (n - r - k - 2), x\}$. The elements of $B'$ must be ordered in decreasing order as well, and note that this is the only case where all the elements of $B'$ are smaller than those of $C'$ but there exists exactly one element in $B'$ that is smaller than the biggest value of $A'$. The index $\{x\}$ is taken from the elements of $A$, so $A' = \{k, k - 1, \ldots, 2, 1\}/\{x\} \cup \{k + 1\}$. Note that there could be at most $k + 1$ different ways to select $x$ as such $(1 \leq x \leq k)$.

If we sum all the possible cases, we get that the number of different ways to create an almost optimal parking function with 1 less $\text{dinv}$ than the maximal possible, is $1 + r + (n - r - k - 2) + k = n - 1$

Proposition 4.26. Fix $1 \leq c \leq \binom{n}{2} - 1$. Then, there are exactly $n$ different parking functions $p_1, p_2, \ldots, p_n$ who satisfy $\text{area}(p_i) = c$ and $\text{area}(p_i) + \text{dinv}(p_i) = \binom{n}{2} - 2$

Proof. Each of those $p_i$ must have $\overline{\text{area}}(p_i) = c$ and either an optimal path or almost optimal path. The existence of $p_1, p_2, \ldots, p_{n-1}$ is an immediate result of Lemmata 2.24 and 2.25. Lemma 2.23 provides us with exactly one $p_n$ with $\overline{\text{area}}(p_n) = c$ and $p_n$ has an almost optimal path.

Let $1 \leq c \leq \binom{n}{2} - 1$, and let $p_1, p_2, \ldots, p_n$ be all parking functions provided by proposition 4.26, such that $\text{area}(p_i) = c$ and $\text{area}(p_i) + \text{dinv}(p_i) = \binom{n}{2} - 1$ (in our standard notation, only $p_n$ has an almost optimal path, and for all $1 \leq k \leq n - 1$ $p_k$ has an optimal path with a swap of index $k$). Denote by $q_1, q_2, \ldots, q_n$ all parking functions provided by proposition 4.26, such that $\text{area}(q_i) = \binom{n}{2} - 1 - c$ and $\text{area}(q_i) + \text{dinv}(q_i) = \binom{n}{2} - 1$ (same standard indexing). Then define

$\chi: \{ p \in P_n : \text{area}(p) + \text{dinv}(p) = \binom{n}{2} - 1 \} \rightarrow \{ p \in P_n : \text{area}(p) + \text{dinv}(p) = \binom{n}{2} - 1 \}$

by $\chi(p_i) = q_i$.

Theorem 4.27. The function $\chi$ described above is a bijection that satisfies both $\text{area}(p) = \text{dinv}(\chi(p))$ and $\text{dinv}(p) = \text{area}(\chi(p))$.

Proof. Proposition 4.28 provides us that for any $p_i$, $\text{area}(p_i) = c$ and $\text{dinv}(p_i) = \binom{n}{2} - 1 - c$, but also $\text{area}(q_i) = \binom{n}{2} - 1 - c$ and $\text{dinv}(q_i) = c$. So if $\chi(p_i) = q_i$, then $\text{area}(p) = \text{dinv}(\chi(p))$ and $\text{dinv}(p) = \text{area}(\chi(p))$. 

□
Chapter 5

Conclusions

This work was an attempt to find a bijective proof that would show that

$$\sum_{\pi \in L_{n,n}} q^{\text{area}(\pi)} t^{\text{dinv}(\pi)}$$

is a symmetric polynomial in $q, t$. Unfortunately, many of our attempts to find such a bijective proof failed, yet we were capable to characterize how must such a bijective proof behave under certain conditions. Later, while studying the behavior of parking functions we came up with additional results that shed some more light on how such a bijection should look like. Attempts to understand the bijection through some examples were difficult to follow since we were not always capable to tell how to match all $\pi \in L_{n,n}$ for a given $n \geq 5$.

I suppose that in order to continue this project one should try consider constrains for different parking functions (i.e, bijection for $p \in P_n$ such that $\text{dinv}(p) + \text{area}(p) = \binom{n}{2} - k$ for some $k \geq 2$). Under those new constrains one should describe an bijection that interchanges $\text{dinv}$ and $\text{area}$ (or possibly other statistics). Alternatively, one could hope to find pseudo-Mahonian statistics and classify them over all $S_n(\tau)$ for $\tau \in S_3$, in the very same manner done by Babson and Steingrimsson [1].
Appendix A

INV to bounce map for $n = 4$

The following describes the input and the output of the function $g : S_4(312) \to L_{n,n}^*$ such that $\text{INV}(\sigma) = \text{bounce}(\pi)$, as it was described in Theorem 3.5:

![Diagram](image-url)

**Figure A.1:** All paths $\pi \in L_{4,4}^*$ above have $\text{bounce}(\pi)$ corresponding to $\text{INV}(\sigma)$.
Appendix B

\[ \phi : S_4(231) \to L_{4,4}^+ \text{ table} \]

The following describes the input and the output of the function \( \phi : S_4(231) \to L_{4,4}^+ \) described by Stump [11].

![Table of bijection \( \phi \)](image)

**Figure B.1:** The bijection \( \phi : S_4(231) \to L_{4,4}^+ \) by Stump.
Appendix C

\[ \sum_{\text{occupant}(1) = i}^{\text{occupant}(2) = j} q^{\text{dinv}(p)} t^{\text{area}(p)} \text{ is not symmetric} \]

Conjecture 4.5 claimed that it might be possible to find an interchanging function of area\((p)\) with dinv\((p)\) for specifically all \( p \in P_n \) where occupant\((1) = i \) for some \( i \).

However we can see that

\[ \sum_{p \in P_n \atop \text{occupant}(1) = i \atop \text{occupant}(2) = j} q^{\text{dinv}(p)} t^{\text{area}(p)} \]

is not a symmetric polynomial in \( q, t \). Consider all \( p \in P_4 \) with \((\text{area}(p), \text{dinv}(p)) = (1, 2)\), with \( \text{occupant}(1) = 1 \):

\[
\begin{array}{c|c|c|c}
 & 1 & 2 & 4 \\
\hline
1 & &  & \\
2 & &  & \\
\end{array}
\begin{array}{c|c|c|c}
 & 1 & 2 & 4 \\
\hline
1 & &  & \\
2 & &  & \\
\end{array}
\begin{array}{c|c|c|c}
 & 1 & 2 & 4 \\
\hline
1 & &  & \\
2 & &  & \\
\end{array}
\]

**Figure C.1:** All those parking functions have area\((p) = 1, \text{ and dinv}(p) = 2.\)

With and all \( p \in P_4 \) with \((\text{area}(p), \text{dinv}(p)) = (2, 1)\) and occupant\((1) = 1 \):

\[
\begin{array}{c|c|c|c}
 & 1 & 2 & 4 \\
\hline
1 & &  & \\
2 & &  & \\
\end{array}
\begin{array}{c|c|c|c}
 & 1 & 2 & 4 \\
\hline
1 & &  & \\
2 & &  & \\
\end{array}
\begin{array}{c|c|c|c}
 & 1 & 2 & 4 \\
\hline
1 & &  & \\
2 & &  & \\
\end{array}
\]

**Figure C.2:** All those parking functions have area\((p) = 1, \text{ and dinv}(p) = 2.\)

Among the first set, occupant\((2)\) was \(2, 3, 4\), while at the second set, occupant\((2) = 2, 3, 3.\)

The leftmost parking function in figure C.1 and the leftmost parking function in figure C.2 shows us that the polynomial is not symmetric even if \( i, j \) lie in the same column.
Bibliography


