Combinatorics of Lattice Paths

Thokozani Paxwell Ncambalala

Supervisor: Professor Charlotte. Brennan

Co-supervisor: Professor Arnold. Knopfmacher

A Dissertation submitted to the Faculty of Science, University of the Witwatersrand, Johannesburg, in fulfillment of the requirements for the degree of Master of Science.

Johannesburg.

Declaration

I declare that this dissertation is my own. It is being submitted for the degree of Master of Science in the University of the Witwatersrand, Johannesburg.

Thokozani Ncambalala

day of 2014

Acknowledgments

I gratefully acknowledge the continued support and advice of my supervisors, Professor Charlotte Brennan and Professor Arnold Knopfmacher. I am extremely obliged for their guidance, support and encouragement.

Abstract

This dissertation consists of five chapters which deal with lattice paths such as Dyck paths, skew Dyck paths and generalized Motzkin paths. These lattice paths are discrete paths in the Cartesian plane made up of a finite sequence of steps that may include up steps $u = (1, 1)$, down steps $d = (1, -1)$, horizontal steps $q = (1, 0)$ and left steps $l = (-1, -1)$. They never go below the horizontal axis. We derive the generating functions to enumerate lattice paths according to different parameters. These parameters include strings of length 2, 3, 4 and r for all $r \in \{2, 3, 4, \cdots\}$, area and semi-base, area and semi-length, and semi-base and semi-perimeter. The coefficients in the series expansion of these generating functions give us the number of combinatorial objects we are interested to count. In particular

1. Chapter 1 is an introduction, here we derive some tools that we are going to use in the subsequent Chapters. We first state the Lagrange inversion formula which is a remarkable tool widely use to extract coefficients in generating functions, then we derive some generating functions for Dyck paths, skew Dyck paths and Motzkin paths.

2. In Chapter 2 we use generating functions to count the number of occurrences of strings in a Dyck path. We first derive generating functions for strings of length 2, 3, 4 and r for all $r \in \{2, 3, 4, \cdots\}$, we then extract the coefficients in the generating functions to get the number of occurrences of strings in the Dyck paths of semi-length n.

3. In Chapter 3, Sections 3.1 and 3.2 we derive generating functions for the relationship between strings of lengths 2 and 3 and the relationship between strings of lengths 3 and 4 respectively. In Section 3.3 we derive generating functions for the low occurrences of the strings of lengths 2, 3 and 4 and lastly Section 3.4 deals with derivations of generating functions for the high occurrences of some strings .

4. Chapter 4, Subsection 4.1.1 deals with the derivation of the generating functions for skew paths according to semi-base and area, we then derive the generating functions according to area. In Subsection 4.1.2, we consider the same as in Section 4.1.1, but here instead of semi-base we use semi-length. The last section 4.2, we use skew paths to enumerate the number of super-diagonal bar graphs according to perimeter.

5. Chapter 5 deals with the derivation of recurrences for the moments of generalized Motzkin paths, in particular we consider those Motzkin paths that never touch the x-axis except at (0,0) and at the end of the path.

Contents

List of Tables

Chapter 1

Introduction

We divide this introduction into two sections, Section 1.1 considers various authors contribution to the study of lattice paths. In Section 1.2 we derive some results from [5], [6] and [19] for Dyck paths, skew paths and Motzkin paths, that we are going to use in subsequent chapters.

1.1 Authors contribution to the study of enumeration of lattice paths according to different parameters

In this section we look at the contributions of several researchers on lattice paths i.e Dyck paths, skew Dyck paths and generalized Motzkin paths. We start with Dyck paths. The enumeration of Dyck paths according to different parameters has been studied by many researchers in the last three decades. In [3], [4], [5], [10], [11], [13], [16], [17], and [19] enumeration of Dyck paths according to different parameters has been studied extensively. These studies has shown that there are many ways of counting Dyck paths using some combination of different parameters.

Every Dyck path can be encoded by a word in a language $\mathcal{D}$ on the alphabet $\{u, d\}$ and let $\{u, d\}^*$ be the set of words made up by letters in $\{u, d\}$. We define a string τ to be a word in $\{u, d\}^*$, this word occurs in a Dyck path α if $\alpha = \beta\tau\delta$, where $\beta, \delta \in \{u, d\}^*$. A semi-length is the number of up steps in a Dyck path. In Chapter 2 we study the enumeration of Dyck paths according to occurrences of strings of length 2, 3, 4 and r by authors in [5], [16] and [19]. They use a string r and semi-length n as parameters. The occurrences of strings of length 2 i.e du, ud, uu, and dd in Dyck paths have been studied extensively and proved in [5] that their statistics follow the Narayana distribution (A001263 of [18]). In [17], general results were studied for occurrences of strings of length 2 for k-colored Motzkin paths.

The occurrences of strings of length 3 i.e udu, uud, uuu, dud, ddu, ddd, udd, and duu have been studied by several authors. The occurrence of the string uud in Dyck paths has been studied in [15]; and it has been proved that the corresponding generating function is

$$F(x,z) - 1 = z(1 + (x-1)z)F^2(x,z),$$

and that

$$b_{n,k} = [x^k z^n]F = \frac{1}{n+1}\binom{n+1}{k}\sum_{j=2k}^{n}\binom{j-k-1}{k-1}\binom{n+1-k}{n-j},$$

where x marks the number of occurrences of the string uud and z marks the semi-length.

The occurrence of the string udu in Dyck paths has been studied independently in [13] and [19] and it has been proved that the corresponding statistic follows the Donaghey distribution, i.e

$$zF^2(x,z) = (1-(x-1)z)F(x,z) - (1-(x-1)z)$$

and that

$$a_{n,k} = \binom{n-1}{k}M_{n-k-1},$$

where x marks the number of occurrences of the string udu and z marks the semi-length.

The occurrence of the string duu in Dyck paths has been studied in [5]; and it has been proved that the corresponding statistic follows the Touchard distribution, i.e

$$xzF^2(x,z) - (1 + 2(x-1)z)F(x,z) + 1 + (x-1)z = 0,$$

and that

$$a_{n,k} = [x^k z^n]F(x,z) = 2^{n-2k-1} C_k \binom{n-1}{2k}.$$

The generating function for the occurrence of the string uuu in Dyck paths is know to satisfy the equation

$$z(t + (1-t)z)F^2 - (1 - (1-t)z)F + 1 = 0.$$

If we consider the symmetry according to the vertical axis, the generating functions of the following pairs of strings $\{uuu, ddd\}$, $\{duu, ddu\}$, $\{uud, udd\}$ and $\{udu, dud\}$ are the same.

The generating functions for the occurrences of all 16 strings of length 4 have been studied in [16]. We notice that among these strings there is a symmetry with respect to a vertical axis, the generating functions for occurrences for some of them (given here in pairs) are equidistant:

$\{uuud, uddd\}$, $\{uuuu, dddd\}$, $\{uddu, duud\}$, $\{duuu, dddu\}$, $\{uudd\}$, $\{uudu, dudd\}$, $\{uduu, ddud\}$, $\{udud\}$,$\{dduu\}$, $\{dudu\}$.

In Chapter 3, enumeration of Dyck paths according to occurrences of strings of length 2, 3, and 4 at low level, even and odd levels, and high levels has been studied in [15], [16] and [19]. In [19] only the string udu is considered.

Finally we look at Chapter 4 and 5. In Chapter 4 enumeration of skew Dyck paths according to different parameters has been studied in [7], these approach is the same with that of Dyck paths in Chapter 2 and 3. There are various authors who contributed in the study of the area under the Dyck paths using different methods [8], [9] and [12]. These methods are related to models in queuing theory and Statistical physics [14]. In [1] a more general approach has been considered, both the exact and asymptotic enumeration of area and average area below directed lattice paths has been studied. In Chapter 5 the area and *moments* of Dyck paths and generalized Motzkin paths has been studied extensively in [3], [20] and [21].

1.2 Definitions and Notations

The Lagrange Inversion Formula (LIF) is a widely used formula for solving functional equations and can sometimes give explicit formulas. To apply LIF, our functional equation must be of the form,

$$H(z) = z\Phi(H(z)).$$

Here Φ is a function of $H(z)$ and we are solving for $H(z)$ in terms of z.

Theorem 1.2.1 *(The Lagrange Inversion Formula)[5](Appendix A). Let $A(z)$ be a generating function satisfying the equation*

$$A(z) = 1 + zH(A(z))$$

where $H(\lambda)$ is a polynomial in λ. The above equation has a unique solution $A(z)$ and if $G(\lambda)$ is a polynomial in λ, then

$$[z^n]G(A(z)) = \frac{1}{n}[\lambda^{n-1}]G'(1+\lambda)(H(1+\lambda))^n \quad \textit{for} \quad n \geq 1.$$

.

Where $[z^n]G(A(z))$ is the coefficient of z^n in the series expansion of $G(A(z))$.

Definition 1.2.2 *[6]A Dyck path is a lattice path in the first quadrant, which begins at the origin (0,0), ends at (2n,0) and consists of steps (1,1) (called rises), (1,-1) (called falls) (see also Figure 1.1).* We say n is the semi-length of the path, which is the number of up steps.

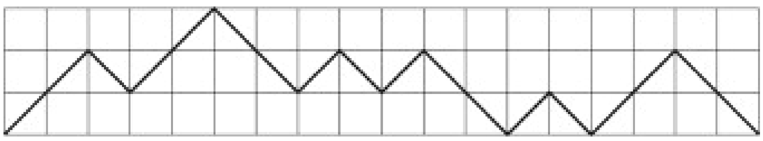

Figure 1.1: The Dyck path *uuduuddududduduudd*.

Let $\mathcal{D}_n$ be all the sets of Dyck paths with semi-length n and let $\mathcal{D} = \bigcup_{n=0}^{\infty} \mathcal{D}_n$, we denote an empty path by ε or by $\bullet$ with $\mid \mathcal{D}_0 \mid = 1$.

A Dyck path can be an empty path or it can be uniquely decomposed as the first return to the horizontal axis (see also Figure 1.2), $\gamma = u\alpha d\beta$ (with $\alpha, \beta \in \mathcal{D}$).

From the decomposition in Figure 1.2, we get the generating function $C(z)$ for Dyck paths as follows,

$$C(z) = 1 + zC(z)^2, \tag{1.1}$$

where z marks the number of up steps. We now want to show that the coefficient of z^n in the series expansion of $C(z)$ is $C_n = \frac{1}{n+1}\binom{2n}{n}$.

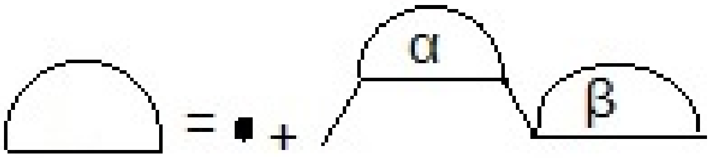

Figure 1.2: The first return decomposition of the Dyck path.

Using the quadratic formula in (1.1) we will show in lemma 1.3 that

$$\begin{aligned} C(z) &= \frac{1-\sqrt{1-4z}}{2z} \\ &= \sum_{n\geq 0} \frac{1}{n+1}\binom{2n}{n} z^n \\ &= 1 + z + 2z^2 + 5z^3 + 14z^4 + \cdots . \end{aligned} \tag{1.2}$$

In the above we use the negative square root since we get $C(0+) = lim_{z\downarrow 0} C(z) = 0$. The positive sign produces $C(0) = \frac{2}{0} = \infty$.

C_n is called the n^{th} Catalan number and $\mid \mathcal{D}_n \mid= C_n$.

The first few numbers of the sequence given by the coefficient C_n for all $n \in \{0, 1, 2, 3, \cdots\}$ are $1, 1, 2, 5, 14, \cdots$ (sequence A000108 in [18])

If $\mathcal{R}$ and $\mathcal{Q}$ are finite sets of Dyck paths, then we define the *concatenation* $\mathcal{RQ}$ of $\mathcal{R}$ and $\mathcal{Q}$ by

$\mathcal{RQ} = \{\alpha\beta : \alpha \in \mathcal{R}, \beta \in \mathcal{Q}\}$.

Lemma 1.2.3 *[6]* $[z^n]C(z) = \dfrac{1}{n+1}\binom{2n}{n}$ *as follows*

Proof

$$\begin{aligned} [z^n]C(z) &= [z^n]\frac{1-\sqrt{1-4z}}{2z} \\ &= [z^{n+1}]\frac{1-\sqrt{1-4z}}{2} \\ &= [z^{n+1}]\frac{1-\sum_{k\geq 0}\binom{\frac{1}{2}}{k}(-4z)^k}{2} \\ &= [z^{n+1}](-1)\frac{\sum_{k\geq 1}\binom{\frac{1}{2}}{k}(-4z)^k}{2} \end{aligned}$$

$$
\begin{aligned}
&= -\frac{1}{2}[z^{n+1}]\sum_{k\geq 1}\frac{(-1)^{k-1}(2k-2)!}{2^{2k-1}k!(k-1)!}(-4z)^k \\
&= -\frac{1}{2}[z^{n+1}]\sum_{k\geq 1}\frac{(-1)(2k-2)!}{2^{2k-1}k!(k-1)!}4^k z^k \\
&= \frac{1}{2}\frac{(2n)!}{2^{2n+1}(n+1)!n!}4^{n+1} \\
&= \frac{1}{n+1}\binom{2n}{n}, \qquad (1.3)
\end{aligned}
$$

as required.

Lemma 1.2.4 *[15] The powers for the Catalan number generating function $C(z)$ are.*

$$[z^n]C^s(z) = \frac{s}{2n+s}\binom{2n+s}{n}.$$

Proof

We prove the lemma 1.2.4 by using the Lagrange inversion formula
$[z^n]G(A(z)) = \frac{1}{n}[\lambda^{n-1}]G'(1+\lambda)(H(1+\lambda))^n$ for $n \geq 1$.
where $G(C(z)) = C(z)^s$.

$$
\begin{aligned}
[z^n]C^s(z) &= \frac{1}{n}[y^{n-1}]s(1+y)^{s-1}(1+y)^{2n} \\
&= \frac{s}{n}[y^{n-1}](1+y)^{2n+s-1} \\
&= \frac{s}{n}\frac{(2n+s-1)!}{(n-1)!(n+s)!} \\
&= \frac{s}{2n+s}\frac{(2n+s)!}{(n)!(n+s)!} \\
&= \frac{s}{2n+s}\binom{2n+s}{n},
\end{aligned}
$$

as required.

Definition 1.2.5 *[5]A skew Dyck path (skew path) is a lattice path, which lies in the first quadrant it begins at the origin, ends on the x-axis and consists of up steps $u = (1,1)$, down steps $d = (1,-1)$ and left steps $l = (-1,-1)$, such that left steps never overlap with up steps. It does not go below the x-axis.*

The length of skew path refers to the number of its steps and semi-length of skew path refers to half the number of its steps. We notice that a skew path of length $2n$ does not necessarily stop at $(2n,0)$. Figure 1.3 shows an example of a skew Dyck path and Dyck path.

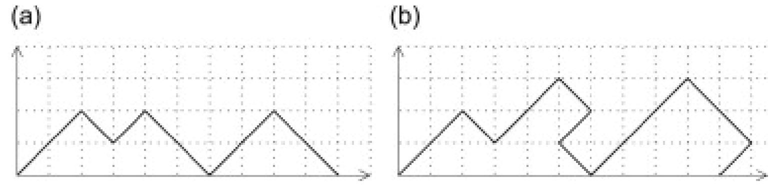

Figure 1.3: (a) a Dyck path and (b) a skew Dyck path

We denote an empty skew path by ε or by $\bullet$. Let $\mathcal{S}_n$ be the set of all skew paths of semi-length n and let $\mathcal{S} = \bigcup_{n=0}^{\infty} \mathcal{S}_n$.

Each skew path can be either empty or it can be uniquely decomposed as $u\gamma' d\gamma''$ (with $\gamma',\ \gamma'' \in S$) *or* u$\gamma' l$ (with $\gamma' \in S, \gamma' \neq \bullet$).

This decomposition can also be shown with a picture as follows

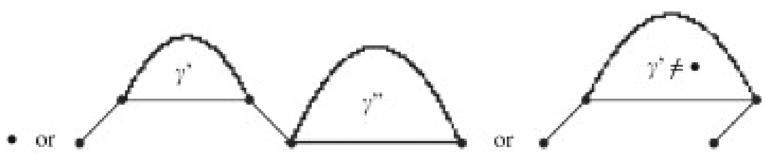

Figure 1.4: Main decomposition of skew Dyck paths.

From figure 1.4 we get the generating function $s(z)$ for skew paths as follows

$$s(z) = 1 + zs^2(z) + z(s(z) - 1), \tag{1.4}$$

where z marks up steps u's.

Solving (1.4) for $s(z)$ by using the quadratic formula we get

$$\begin{aligned} s(z) &= \frac{-(z-1) - \sqrt{z^2 - 2z + 1 - 4(z)(1-z)}}{2z} \\ &= \frac{1 - z - \sqrt{1 - 6z + 5z^2}}{2z}. \\ &= 1 + z + 3z^2 + 10z^3 + 36z^4 + 137z^5 + 543z^6 + \cdots \end{aligned} \tag{1.5}$$

In the above we use the negative square root since we get $s(0+) = lim_{z\downarrow 0} s(z) = 0$. The positive sign produces $s(0) = \frac{2}{0} = \infty$.

To get the coefficient of z^n from the series expansion of $s(z)$, we proceed as follows,

$$s_n = [z^n]s(z) = \frac{[y^{n-1}](1+3y+y^2)^n}{n}. \tag{1.6}$$

for $n \geq 0$. We expand $(1+3y+y^2)^n$ and after writing the trinomial in the three forms $(1+y)^2+y$, $(1-y)^2+5y^2$ and $(1+3y)+y^2$, then after applying the inversion Lagrange formula we get,

Theorem 1.2.6 [5]

$$s_n = \sum_{k=1}^{n} \binom{n-1}{k-1} C_k, \tag{1.7}$$

Theorem 1.2.7 [5]

$$s_n = \sum_{k=0}^{n} \binom{n-1}{k-1} (-1)^{k-1} 5^{n-k} C_k \tag{1.8}$$

and

Theorem 1.2.8 [5]

$$s_n = \frac{1}{n} \sum_{k=0}^{\lfloor \frac{n-1}{2} \rfloor} \binom{n}{k} \binom{n-k}{k+1} 3^{n-2k-1} \tag{1.9}$$

respectively.

The first few numbers of the sequence given by the coefficient s_n are

$1, 1, 3, 10, 36, 137, 543, \cdots$ (sequence A002212 in [18]).

Now we give proofs for theorem 1.2.6, theorem 1.2.7 and theorem 1.2.8, we start with theorem 1.2.6,

$$\begin{aligned}
[z^n]s(z) &= \frac{1}{n}[y^{n-1}]((1+y)^2+y)^n \\
&= \frac{1}{n}[y^{n-1}] \sum_{k=0}^{n} \binom{n}{k} y^{n-k}(y+1)^{2k} \\
&= \frac{1}{n} \sum_{k=0}^{n} \binom{n}{k} [y^{k-1}] \sum_{j=0}^{2k} \binom{2k}{j} y^j \\
&= \frac{1}{n} \sum_{k=0}^{n} \binom{n}{k} \binom{2k}{k-1}
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{k=0}^{n} \frac{(n-1)!}{k!(n-k)!} \frac{(2k)!}{(k-1)!(k+1)!} \\
&= \sum_{k=1}^{n} \binom{n-1}{k-1} C_k. \qquad (1.10)
\end{aligned}
$$

Now for theorem 1.2.7

$$
\begin{aligned}
[z^n]s(z) &= \frac{1}{n}[y^{n-1}]((1-y)^2+5y)^n \\
&= \frac{1}{n}[y^{n-1}] \sum_{k=0}^{n} \binom{n}{k} (1-y)^{2k} y^{n-k} 5^{n-k} \\
&= \frac{1}{n} \sum_{k=0}^{n} \binom{n}{k} [y^{k-1}] \sum_{j=0}^{2k} (-1)^j \binom{2k}{j} y^j 5^{n-k} \\
&= \frac{1}{n} \sum_{k=1}^{n} \binom{n}{k} (-1)^{k-1} \binom{2k}{k-1} 5^{n-k} \\
&= \sum_{k=1}^{n} \frac{(n-1)!}{k!(n-k)!} \frac{(2k)!}{(k-1)!(k+1)!} (-1)^{k-1} 5^{n-k} \\
&= \sum_{k=1}^{n} \frac{(n-1)!}{k!(n-k)!} \frac{1}{k+1} \frac{(2k)!}{k!k!} (-1)^{k-1} 5^{n-k} \\
&= \sum_{k=0}^{n} \binom{n-1}{k-1} (-1)^{k-1} 5^{n-k} C_k.
\end{aligned} \qquad (1.11)
$$

Lastly theorem 1.2.8.

$$
\begin{aligned}
[z^n]s(z) &= \frac{1}{n}[y^{n-1}]((1+3y)+y^2)^n \\
&= \frac{1}{n}[y^{n-1}] \sum_{k=0}^{n} \binom{n}{k} y^{2k} (1+3y)^{n-k} \\
&= \frac{1}{n} \sum_{k=0}^{\lfloor \frac{n-1}{2} \rfloor} \binom{n}{k} \binom{n-k}{n-2k-1} 3^{n-2k-1} \\
&= \frac{1}{n} \sum_{k=0}^{\lfloor \frac{n-1}{2} \rfloor} \binom{n}{k} \binom{n-k}{k+1} 3^{n-2k-1}.
\end{aligned}
$$

$$(1.12)$$

The relationship between the skew path generating function $s(z)$ and the Dyck path generating function $C(z)$ is as follows,

$$s(z) = C\left(\frac{z}{1-z}\right). \tag{1.13}$$

We prove the above relation as follows;

$$\begin{aligned} C\left(\frac{z}{1-z}\right) &= \frac{1-\sqrt{1-\frac{4z}{1-z}}}{2\frac{z}{1-z}} \\ &= \frac{1-\sqrt{\frac{1-5z}{1-z}}}{\frac{2z}{1-z}} \\ &= \frac{1-z-\sqrt{1-6z+5z^2}}{2z} \end{aligned} \tag{1.14}$$

as from equation (1.5) $s(z) = \dfrac{1-z-\sqrt{1-6z+5z^2}}{2z}$ therefore $s(z) = C(\frac{z}{1-z})$.

Definition 1.2.9 *[6]We define a Motzkin path to be a lattice path starting at the point (0,0) and ending at the point (n,0), it never goes below the horizontal axis. The steps are the up steps* $u = (1,1)$*, the horizontal steps* $q = (1,0)$ *and then down steps* $d = (1,-1)$ (see Figure 1.5).

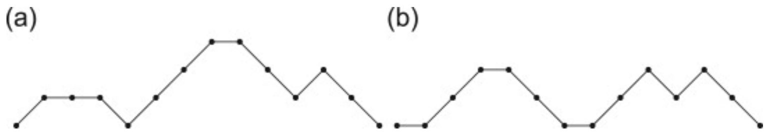

Figure 1.5: Two different Motzkin paths.

We define the length of a Motzkin path by the number of its steps. Let $\mathcal{M}_n$ be the set of all Motzkin paths with length n, then $\mathcal{M} = \bigcup_{n=0}^{\infty} \mathcal{M}_n$.

We now define the first return decomposition for Motzkin paths as follows, a Motzkin path can be empty or it can start with a horizontal step or it can start with an up step and return to the horizontal axis for the first time. Thus we get the following decomposition,

$$\mathcal{M} = \{\varepsilon\} \cup q\mathcal{M} \cup u\mathcal{M}d\mathcal{M},$$

this translate to the following , where $M(z)$ is the generating function for Motzkin paths

$$M(z) = 1 + zM(z) + z^2M(z)^2,$$

where z marks each of the horizontal, up and down steps.

Now from this generating function we can derive the recurrence for Motzkin paths i.e

Lemma 1.2.10 [19] $M_n = M_{n-1} + \sum_{k=0}^{n-2} M_{n-2}M_{n-2-k}$ where M_n is the n^{th} Motzkin number.

Proof

We consider the fact that by convolution rule $M^2(z) = \sum_{j\geq 0}\sum_{k=0}^{j} M_kM_{j-k}z^j$,

$$\begin{aligned}
M(z) &= 1 + zM(z) + z^2M(z)^2 \\
\sum_{k\geq 0} M_kz^k &= 1 + z\sum_{k\geq 0} M_kz^k + z^2\sum_{j\geq 0}\sum_{k=0}^{j} M_kM_{j-k}z^j \\
[z^n]\sum_{k\geq 0} M_kz^k &= [z^n]\left(1 + z\sum_{k\geq 0} M_kz^k + z^2\sum_{j\geq 0}\sum_{k=0}^{j} M_kM_{j-k}z^j\right) \\
&= [z^{n-1}]\sum_{k\geq 0} M_kz^k + [z^{n-2}]\sum_{j\geq 0}\sum_{k=0}^{j} M_kM_{j-k}z^j \\
M_n &= M_{n-1} + \sum_{k=0}^{n-2} M_{n-2}M_{n-2-k},
\end{aligned}$$

as required.

We know that there is only one empty Motzkin path, that is $M_0 = 1$, now from the above recurrence we see that $M_1 = 1$ and $M_2 = 2$. Now we want to show that

Lemma 1.2.11 [19] $M_n = \sum_{j=0}^{[n/2]} \binom{n}{2j} C_j$.

Proof: from above we have

$$M(z) = 1 + zM(z) + z^2M(z)^2$$

$$M(z) - 1 = zM(z) + z^2 M(z)^2.$$

We use the following equation which is in the suitable form to apply Lagrange inversion formula.

$$A(z) = 1 + zH(A(z)),$$

where $H(\gamma) = \gamma + z\gamma^2$ and $A(z) = M(z)$. Now using the Lagrange inversion formula we get

$$[z^\sigma]A(z) = \frac{1}{\sigma}[\gamma^{\sigma-1}](H(1+\gamma))^\sigma. \tag{1.15}$$

This gives us

$$\begin{aligned}
(H(1+\gamma))^\sigma &= \sum_{j=0}^{\sigma} \binom{\sigma}{j} z^j (1+\gamma)^{2j} (1+\gamma)^{\sigma-j} \\
&= \sum_{j=0}^{\sigma} \binom{\sigma}{j} z^j (1+\gamma)^{\sigma+j} \\
&= \sum_{j=0}^{\sigma} \binom{\sigma}{j} z^j \sum_{v=0}^{\sigma+j} \binom{\sigma+j}{v} \gamma^v \\
&= \sum_{v=0}^{\sigma} \sum_{j=(\sigma-v)^+}^{\sigma} \binom{\sigma+j}{v} \binom{\sigma}{j} z^j \gamma^v.
\end{aligned}$$

From (1.15) with $v = \sigma - 1$ we get

$$\begin{aligned}
[z^\sigma]M(z) &= \frac{1}{\sigma} \sum_{j=0}^{\sigma} \binom{\sigma+j}{\sigma-1} \binom{\sigma}{j} z^j \\
M(z) &= \sum_{\sigma \geq 1} \frac{1}{\sigma} \sum_{j=0}^{\sigma} \binom{\sigma+j}{\sigma-1} \binom{\sigma}{j} z^j z^\sigma.
\end{aligned}$$

Since we want z^n, let $n = \sigma + j$ then

$$\begin{aligned} M(z) &= \sum_{n\geq 1}\sum_{j=0}^{n-j} \frac{1}{n-j}\binom{n}{n-j-1}\binom{n-j}{j} z^n \\ &= \sum_{n\geq 1}\sum_{j=0}^{n-j} \frac{1}{n-j}\frac{n!}{(n-j-1)!(j+1)!}\frac{(n-j)!}{(n-2j)!j!}\frac{(2j)!}{(2j)!} z^n \\ &= \sum_{n\geq 1}\sum_{j=0}^{[n/2]} \binom{n}{2j} C_j z^n. \end{aligned}$$

Now we see that

$$[z^n]M(z) = M_n = \sum_{j=0}^{[n/2]} \binom{n}{2j} C_j,$$

as required.

The first few numbers of the sequence given by the coefficient M_n are $1, 1, 2, 5, 15, 52, \cdots$ (sequence A000110 in [18]).

It is known that

$$C_{n+1} = \sum_{j=0}^{n} \binom{n}{j} M_{n-j}.$$

We are now ready to apply these results of Chapter 1 into the following four chapters.

Chapter 2

Counting strings in Dyck paths

In this chapter we study the paper titled Counting strings in Dyck paths in [16]. We take into account the number of occurrences of the string τ.

For the basic definitions we again refer the reader to Deutch [5].

In a Dyck path ($\mathcal{D}_n$) we define a peak (valley) to be a point between an up (down) step and a down (resp. up) step. A double-up (double-down) step is a point preceded and followed by an up (down) step. The height of a point is defined to be its y-coordinate. We call a peak (valley) low if it is of height 1 (0). We call a peak (valley) small if it is not immediately preceded by a double-up (double-down) step and immediately followed by a double-down (double-up). From the above definition a low peak is a small peak. An ascent (descent) of a Dyck path is a maximal sequence of consecutive up (down) steps. Every Dyck path can be encoded by a word in a language $\mathcal{D}$ on the alphabet $\{u, d\}$ and let $\{u, d\}^*$ be the set of words made up by letters in $\{u, d\}$. For example the Dyck path α in Figure 2.1 is encoded by the word $\alpha = uuduudddududduduudd \in \{u, d\}^*$. We define a string τ to be a word in $\{u, d\}^*$, this word occurs in a Dyck path α if $\alpha = \beta\tau\delta$, where $\beta, \delta \in \{u, d\}^*$. For example there are three occurrences of each of the strings uu and dd in Figure 2.1.

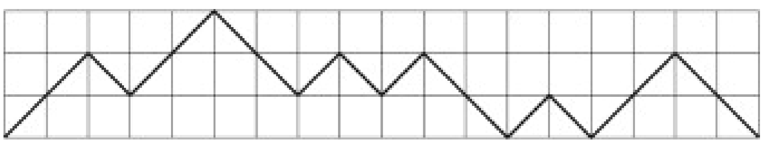

Figure 2.1: The Dyck path $uuduudddududduduudd$.

In this chapter we consider the number of occurrences of the string τ. We use the generating function $F_\tau(t,z) := F(t,z) := F$ where t counts the number of occurrences of τ and z marks the semi-length. We define $F_\tau(t,z)$ as follows,

$$F_\tau(t,z) = \sum_{n=0}^{\infty}\sum_{k=0}^{n} a_{n,k} t^k z^n. \tag{2.1}$$

Here a_{nk} is the number of all Dyck paths of semi-length n with k occurrences of τ. We define $a_{n,0} = a_n$ to be an avoiding sequence of the string τ.

2.1 The strings of length 2.

We first consider strings of length 2, namely uu, ud, du and dd.

If we consider the symmetry according to the vertical axis, the statistics of the following pairs of strings $\{uu, dd\}$ are equidistant.

Theorem 2.1.1 *[5] The generating function for occurrences of the strings uu and dd is* $1 + tzF^2 - (1 - z + tz)F = 0$.

Proof

To derive the generating function $F(t,z) := F$ for the string $\tau = uu$, we partition $\mathcal{D}$ into $\{\Omega_i\}$, where Ω_i is the set of all Dyck paths with the length of the first ascent is equal to i, for all $i \geq 1$. We define $A_i(t,z) := A_i$ to be a generating function for Ω_i, where t counts the number of occurrences of the string $\tau = uu$. All elements α of Ω_i can be written uniquely as $\alpha = u^i d\alpha_1 d\alpha_2 d\alpha_3 \ldots . d\alpha_i$, where $\alpha_m \in \mathcal{D}$ for all $m \in [i]$. We conclude that there are $(i-1)$ new strings uu (in addition to those contributed by α_i's) if and only if $i \geq 2$, since there are i (a_m's) and also there are i ($u's$). These produce $A_i = t^{i-1} z^i F^i$ and if $i \leq 1$ we get $A_i = z^i F^i$. Combining these two possibilities we get

$\sum_{i=2}^{\infty} t^{i-1} z^i F^i$ and $\sum_{i=0}^{1} z^i F^i$.

Hence we get

$$\begin{aligned} F &= \sum_{i=0}^{1} z^i F^i + \sum_{i=2}^{\infty} t^{i-1} z^i F^i \\ &= 1 + zF + \sum_{i=2}^{\infty} t^{i-1} z^i F^i. \end{aligned}$$

$$(2.2)$$

Now we do some manipulations in the above equation as follows:

$$\begin{aligned} F &= 1 + zF + \sum_{i=2}^{\infty} t^{i-1} z^i F^i \\ &= 1 + zF + \sum_{i=0}^{\infty} t^{i+2-1} z^{i+2} F^{i+2} \\ &= 1 + zF + tz^2F^2 \sum_{i=0}^{\infty} t^i z^i F^i \\ &= 1 + zF + \frac{tz^2F^2}{1 - tzF} \\ F - tzF^2 &= 1 - tzF + zF - tz^2F^2 + tz^2F^2 \\ 1 + tzF^2 &- (1 - z + tz)F = 0. \end{aligned}$$

Then the generating function for the string uu satisfies,

$$tzF^2 - (1 - (1 - t)z)F + 1 = 0. \tag{2.3}$$

Theorem 2.1.2 *[5]The number of Dyck paths with semi-length n, having k occurrences of the strings uu and dd is equal to.*

$$a_{n,k} = [t^k z^n]F(t, z) = \frac{1}{n} \sum_{j=0}^{n} (-1)^{k-j} \binom{n}{j} \binom{n+j}{n-1} \binom{n-j}{k-j}. \tag{2.4}$$

Proof of (2.4)
Let $w(t, z) := w = F(t, z) - 1$

$$w = z(t(w + 1)^2 + (1 - t)(w + 1)). \tag{2.5}$$

We now apply Lagrange inversion formula to get

$$[z^n]w = \frac{1}{n}[y^{n-1}](t(y + 1)^2 + (1 - t)(y + 1))^n$$

$$= \frac{1}{n}[y^{n-1}] \sum_{j=0}^{n} \binom{n}{j} t^j (y+1)^{2j} (1-t)^{n-j} (y+1)^{n-j}$$
$$= \frac{1}{n}[y^{n-1}] \sum_{j=0}^{n} \binom{n}{j} t^j (y+1)^{n+j} \sum_{i=0}^{n-j} \binom{n-j}{i} (-1)^i t^i. \tag{2.6}$$

Then

$$[t^k z^n]w = \frac{1}{n}[y^{n-1}] \sum_{j=0}^{n} \binom{n}{j} (y+1)^{n+j} \binom{n-j}{k-j} (-1)^{k-j}$$
$$= \frac{1}{n} \sum_{j=0}^{n} (-1)^{k-j} \binom{n}{j} \binom{n+j}{n-1} \binom{n-j}{k-j}, \tag{2.7}$$

as required.

To derive the generating function $F(t,z) := F$ for the string $\tau = dd$, we partition $\mathcal{D}$ into $\{\Omega_i\}$ Where Ω_i is the set of all Dyck paths with the length of the last descent is equal to i, for all $i \geq 1$. We define $A_i(t,z) := A_i$ to be a generating function for Ω_i, where t counts the number of occurrences of the string $\tau = dd$. All elements α of Ω_i can be written uniquely as $\alpha = \alpha_1 u \alpha_2 u \alpha_3 .... u \alpha_i u d^i$, where $\alpha_m \in \mathcal{D}$ for all $m \in [i]$. We conclude that there are $(i-1)$ new strings dd (in addition to those contributed by α_i's) if and only if $i \geq 2$, since there are i (a_m's) and also there are i ($u's$). This produces $A_i = t^{i-1} z^i F^i$ and if $i \leq 1$ we get $A_i = z^i F^i$. Combining these two possibilities we get
$\sum_{i=2}^{\infty} t^{i-1} z^i F^i$ and $\sum_{i=0}^{1} z^i F^i$.
Then similar to the case uu the generating function for the string dd satisfies,

$$tzF^2 - (1-(1-t)z)F + 1 = 0. \tag{2.8}$$

This quadratic equation for the generating function for dd is the same as for the generating function for the string uu. We say that two or more strings are same if the have the same generating function, thus dd and uu have the same generating function.

Theorem 2.1.3 *[5]The generating function for occurrences of the strings ud is $zF^2 - (1+(1-t)z)F + 1 = 0$.*

Proof

We now consider the string $\tau = ud$, we derive its generating function $F(t,z) := F$ (where t counts the number of occurrences of the string $\tau = ud$) by using the first return decomposition of the non-empty Dyck path $\alpha = u\beta d\gamma$ where $\alpha, \beta, \gamma \in \mathcal{D}$. A new occurrence of ud appears in α (in addition to those contributed by β and γ) if and only if $\beta = \varepsilon$.

The path α can start with the string ud right from the start where $\beta = \varepsilon$, producing $z(tF)$ or α can start with uu where $\beta \neq \varepsilon$ producing $z(F-1)F$. Combining these two possibilities we get,

$$F = 1 + z(tF + (F-1)F)$$
$$F = 1 + ztF + zF^2 - zF.$$

Then the generating function for the occurrence of the string ud is given by,

$$zF^2 - (1+(1-t)z)F + 1 = 0. \tag{2.9}$$

Theorem 2.1.4 *[5]The number of Dyck paths with semilength n, having k occurrences of the string ud is equal to.*

$$a_{n,k} = [t^k z^n]F(t,z) = \frac{1}{n}\sum_{j=0}^{n}(-1)^{n-k-j}\binom{n}{j}\binom{n+j}{n-1}\binom{n-j}{k}. \tag{2.10}$$

Proof of (2.10)

Let $w(t,z) := w = F(t,z) - 1$.

$$w = z((w+1)^2 + (t-1)(w+1)) \tag{2.11}$$

We now apply Lagrange inversion formula to get

$$[z^n]w = \frac{1}{n}[y^{n-1}]((y+1)^2 + (t-1)(y+1))^n$$

$$
\begin{aligned}
&= \frac{1}{n}[y^{n-1}]\sum_{j=0}^{n}\binom{n}{j}(y+1)^{2j}(t-1)^{n-j}(y+1)^{n-j} \\
&= \frac{1}{n}[y^{n-1}]\sum_{j=0}^{n}\binom{n}{j}(y+1)^{n+j}\sum_{i=0}^{n-j}\binom{n-j}{i}(-1)^{n-j-i}t^{i}
\end{aligned}
\tag{2.12}
$$

Then

$$
\begin{aligned}
[t^k z^n]w &= \frac{1}{n}[y^{n-1}]\sum_{j=0}^{n}\binom{n}{j}(y+1)^{n+j}\binom{n-j}{k}(-1)^{n-k-j} \\
&= \frac{1}{n}\sum_{j=0}^{n}(-1)^{n-k-j}\binom{n}{j}\binom{n+j}{n-1}\binom{n-j}{k},
\end{aligned}
\tag{2.13}
$$

as required.

Theorem 2.1.5 *[5]The generating function for occurrences of the strings du is $ztF^2-(1-(1-t)z)F+1=0$.*

Proof

We now consider the string $\tau = du$, we derive its generating function $F(t,z) := F$ (where t counts the number of occurrence of the string $\tau = du$) by using the first return decomposition of the non-empty Dyck path $\alpha = \beta u \gamma d$ where $\alpha, \beta, \gamma \in \mathcal{D}$. A new occurrence of du appears in α (in addition to those contributed by β and γ) if and only if $\beta \neq \varepsilon$ producing $z(t(F-1)F)$ or α can be such that $\beta = \varepsilon$ producing zF. Combining these two possibilities we get,

$$
\begin{aligned}
F &= 1 + z(t(F-1)F + F) \\
F &= 1 + ztF^2 - tzF + zF \\
& ztF^2 - (1-(1-t)z)F + 1 = 0.
\end{aligned}
$$

Then the generating function for the occurrence of the string du is given by,

$$
ztF^2(t,z) - (1-(1-t)z)F(t,z) + 1 = 0. \tag{2.14}
$$

We see that the strings du and uu are the same, then for du

$$a_{n,k} = [t^k z^n]F(t,z) = \frac{1}{n}\sum_{j=0}^{n}(-1)^{k-j}\binom{n}{j}\binom{n+j}{n-1}\binom{n-j}{k-j}, \tag{2.15}$$

as in uu.

2.2 The strings of length 3.

We now consider the strings of length 3, which are udu, udd, duu, uuu, uud, dud, ddd and ddu.

If we consider the symmetry according to the vertical axis, the statistics of the following pairs of strings $\{uuu, ddd\}$, $\{duu, ddu\}$, $\{uud, udd\}$ and $\{udu, dud\}$ have the same generating function.

Theorem 2.2.1 *[16]The generating function for occurrences of the strings duu and ddu is* $xzF^2(x,z) - (1 + 2(x-1)z)F(x,z) + 1 + (x-1)z = 0$.

We will derive the generating function for $\tau = duu$ in two different ways

The first way is as follows:

Proof

Let Ω be the set of all Dyck paths where the first ascent is of size at least 2, with its generating function $A(x,z) = A$, where x counts the number of occurrences of the string $\tau = duu$. All elements α of Ω can be written uniquely as $\alpha = u\alpha_1 d\alpha_2$, where $\alpha_1, \alpha_2 \in D$ and $\alpha_1 \neq \varepsilon$. The new duu occurs if and only if $\alpha_2 \in \Omega$, this produces $xz(F-1)A$ and if α_2 does not belong to Ω then we get $z(F-1)(F-A)$. Combining these two possibilities we get

$$A = z(F-1)(F - A + xA).$$

We know that every non-empty Dyck path $\alpha \in D$ can be written uniquely as a first return decomposition, that is $\alpha = u\beta d\gamma$, where $\beta, \gamma \in D$. We see that a new string duu (in addition to those contributed by β and γ) occurs if and only if $\gamma \in \Omega$, producing $xzAF$, and if γ does not belong to Ω then we get $z(F-A)F$. Combining these two possibilities we now get

$$F = 1 + z(xA + (F - A))F.$$

In the equation $F = 1 + xzAF + z(F - A)F$ we solve for A and get
$A = \dfrac{F - 1 - zF^2}{xzF - zF}$.

We then substitute the expression for A into $A = z(F - 1)(F - A + tA)$, then we proceed as follows,

$$\begin{aligned}
\frac{F - 1 - zF^2}{xzF - zF} &= z(F-1)\left(F - \left(\frac{F - 1 - zF^2}{xzF - zF}\right) + x\left(\frac{F - 1 - zF^2}{xzF - zF}\right)\right)\\
F - 1 - zF^2 &= z(F-1)(xzF^2 - zF^2 - F + 1 + zF^2 + xF - x - xzF^2)\\
F - 1 - zF^2 &= z(F-1)(-F + 1 + xF - x)\\
F - 1 &= 2zF - 2xzF + xzF^2 - (1-x)z\\
xzF^2 - (1 + 2(x-1)z)F &+ 1 + (x-1)z = 0.
\end{aligned}$$

Alternatively we can derive the generating function for *duu* as follows: We derive its generating function by using the first return decomposition of non-empty Dyck paths $\alpha = \beta u \gamma d$ where $\alpha, \beta, \gamma \in D$, a new occurrence of *duu* appears in α (in addition to those contributed by β and γ) if and only if both $\beta, \gamma \neq \varepsilon$.

Now if $\beta, \gamma \neq \varepsilon$ we get $z(x(F-1)^2)$ or if both $\beta, \gamma = \varepsilon$ we get z and if one of β and γ is empty and the other one is not empty we get $2z(F-1)$. Combining these three possibilities we get

$$\begin{aligned}
F &= 1 + z(x(F-1)^2) + 2z(F-1) + z\\
F &= 1 + xzF^2 - 2xzF + xz + 2zF - z\\
xzF^2 &- (1 + 2(x-1)z)F + 1 + (x-1)z = 0.
\end{aligned}$$

Now the generating function of the string *duu* is

$$xzF^2(x,z) - (1 + 2(x-1)z)F(x,z) + 1 + (x-1)z = 0,$$

as required.

The coefficient $a_{n,k}$ of the series expansion for $F(x,z)$ is,

$$a_{n,k} = [x^k z^n]F(x,z) = \begin{cases} 1, & \text{if } n = k = 0 \\ 2^{n-2k-1} C_k \binom{n-1}{2k} & \text{if } n \geq 1, \end{cases} \tag{2.16}$$

Proof of (2.16)

$$z\left(xF(x,z)^2 - (x-1)(2F(x,z)-1)\right) = F(x,z) - 1. \tag{2.17}$$

Let $w(x,z) = F(x,z) - 1$ then

$$\begin{aligned} w &= z\left(x(w+1)^2 - (x-1)(2w+1)\right) \\ &= z(xw^2 + 2w + 1). \end{aligned} \tag{2.18}$$

We now apply Lagrange inversion formula to get

$$\begin{aligned} [z^n]w &= \frac{1}{n}[y^{n-1}](xy^2 + 2y + 1)^n \\ &= \frac{1}{n}[y^{n-1}] \sum_{j=0}^{n} \binom{n}{j} x^j y^{2j} (2y+1)^{n-j}. \end{aligned} \tag{2.19}$$

Then

$$\begin{aligned} [x^k z^n]w &= \frac{1}{n}[y^{n-1}] \binom{n}{k} y^{2k} (2y+1)^{n-k} \\ &= \frac{1}{n}\binom{n}{k} [y^{n-2k-1}] \sum_{m=0}^{n-k} \binom{n-k}{m} 2^m y^m \\ &= \frac{1}{n}\binom{n}{k}\binom{n-k}{n-2k-1} 2^{n-2k-1} \\ &= 2^{n-2k-1} \frac{(n-1)!}{k!(n-k)!} \frac{(n-k)!}{(n-2k-1)!(k+1)!} \end{aligned}$$

$$
\begin{aligned}
&= 2^{n-2k-1}\frac{(n-1)!}{(2k)!(n-2k-1)!}\frac{(2k)!}{k!k!(k+1)} \\
&= 2^{n-2k-1}\binom{n-1}{2k}C_k,
\end{aligned}
\tag{2.20}
$$

as required.

We now derive the generating function $F(t,z) := F$ for ddu in two ways.

Using the first return decomposition of a non-empty Dyck path $\alpha = u\beta d\gamma$ where $\alpha, \beta, \gamma \in \mathcal{D}$, a new occurrence of ddu appears in α (in addition to those contributed by β and γ) if and only if both $\beta, \gamma \neq \varepsilon$.

Now if $\beta, \gamma \neq \varepsilon$ we get $z(x(F-1)^2)$ or if both $\beta, \gamma = \varepsilon$ we get z and if exactly one of β and γ is empty and the other one is not empty we get $2z(F-1)$. Combining these three possibilities we get

$$
\begin{aligned}
F &= 1 + z(x(F-1)^2) + 2z(F-1) + z \\
F &= 1 + xzF^2 - 2xzF + xz + 2zF - z \\
xzF^2 &- (1 + 2(x-1)z)F + 1 + (x-1)z = 0.
\end{aligned}
$$

Alternatively we can derive the generating function for ddu as follows:

Let Ω be the set of all Dyck paths where the last descent is at least 2, with its generating function $A(x,z) = A$, where x counts the number of occurrence of the string $\tau = ddu$. All elements α of Ω can be written uniquely as $\alpha = \alpha_1 u \alpha_2 d$, where $\alpha_1, \alpha_2 \in \mathcal{D}$ and $\alpha_2 \neq \varepsilon$. The new ddu occurs if and only if $\alpha_1 \in \Omega$, this produces $xz(F-1)A$ and if α_1 does not belong to Ω then we get $z(F-1)(F-A)$. Combining these two possibilities we get

$$
A = z(F-1)(F - A + tA).
$$

We know that every non-empty Dyck path $\alpha \in D$ can be written uniquely as a first return decomposition, that is $\alpha = \beta u \gamma d$, where $\beta, \gamma \in \mathcal{D}$. We see that a new string ddu occurs (in addition to those contributed by β and γ) if and only if $\beta \in \Omega$, producing $xzAF$, and if β does not belong to Ω then we get $z(F-A)F$. Combining these two possibilities we now get

$$F = 1 + z(tA + (F - A))F.$$

Now as for duu the generating function of the string ddu is

$$xzF^2 - (1 + 2(x-1)z)F + 1 + (x-1)z = 0.$$

This generating function for ddu is the same as the generating function for the string duu, thus ddu and duu have the same generating function.

$$xzF^2(x,z) - (1 + 2(x-1)z)F(x,z) + 1 + (x-1)z = 0.$$

Theorem 2.2.2 *[16]The generating function for occurrences of the strings udu and dud is* $zF^2(x,z) = (1-(x-1)z)F(x,z) - (1-(x-1)z)$.

For the string udu we refer to the paper titled The statistic number of udu's in Dyck path by Yidong Sun [19].

Proof

We now consider the string $\tau = udu$, We derive its generating function by using the first return decomposition of non-empty Dyck path $\alpha = u\beta d\gamma$ where $\alpha, \beta, \gamma \in \mathcal{D}$. A new occurrence of udu appears in α (in addition to those contributed by β and γ) if and only if $\beta = \varepsilon$. Where $\gamma \neq \varepsilon$ producing $z(x(F(x,z)-1))$ or α can start with uu where $\beta \neq \varepsilon$ producing $z(F(x,z)-1)F(x,z)$ or we can have $\alpha = ud$ where $\beta = \gamma = \varepsilon$ producing z (where z marks up step u).

Now the non-empty Dyck path α have the generating function $F(x,z) - 1$ as follows.

$$\begin{gathered} F(x,z) - 1 = z(x(F(x,z)-1)) + z(F(x,z)-1)F(x,z) + z \\ zF^2(x,z) = (1-(x-1)z)F(x,z) - (1-(x-1)z). \end{gathered}$$

We now prove the following using two different methods,

$$a_{n,k} = [x^k z^n]F(x,z) = \begin{cases} 1 & \text{if } n = k = 0 \\ \binom{n-1}{k} M_{n-k-1} & \text{if } k \in [0, n-1]. \end{cases} \tag{2.21}$$

The first method.

Proof

We define $S_{n,k}$ to be the set of all Dyck paths denoted by $\mathcal{D}_n$ of semi-length $n \geq 2$, with k udu's, we now take $S_{n+1,0}$, that is a Dyck path of semi-length $n+1$ with no udu's. Let $a_{n+1,k}$ be the cardinality of $S_{n+1,k}$. Let $\alpha = u\beta d\gamma$ be the first return decomposition, where $\alpha, \beta, \gamma \in D$, then both β and γ are udu-avoiding and also, if $\beta = \varepsilon$, then $\gamma = \varepsilon$, otherwise if $\gamma \neq \varepsilon$ then we have udu. Therefore we see that for $n \geq 2$ in α, we need to have $\beta \neq \varepsilon$. From the generating function

$$F(x,z) - 1 = z(x(F(x,z)-1)) + z(F(x,z)-1)F(x,z) + z,$$

for udu above where x marks the number of occurrences of the string udu. let $x = 1$ then

$$\begin{aligned} F(1,z) - 1 &= z(F(1,z)-1) + z(F(1,z)-1)F(1,z) + z \\ F(1,z) - 1 &= zF^2(1,z). \end{aligned}$$

Now we consider $F(1,z) = \sum_{i\geq 0} a_{i,0} z^i$,we also know that
$zF^2(1,z) = z\sum_{k\geq 0}\sum_{i=0}^{k} a_{i,0}a_{k-i,0}z^k$ by the convolution rule, now we get

$$\sum_{i\geq 0} a_{i,0}z^i - 1 = \sum_{k\geq 0}\sum_{i=0}^{k} a_{i,0}a_{k-i,0}z^{i+1}.$$

Thus

$$[z^{n+1}]\left(\sum_{i\geq 0} a_{i,0}z^i - 1\right) = [z^{n+1}]\sum_{k\geq 0}\sum_{i=0}^{k} a_{i,0}a_{k-i,0}z^{k+1}$$

Therefore $a_{n+1,0} = \sum_{i=0}^{n} a_{i,0}a_{n-i,0}$.
and since $a_{0,0} = 1$

$$a_{n+1,0} = a_{n,0} + \sum_{i=1}^{n-1} a_{i,0}a_{n-i,0}.$$

We now use the above recurrence

$$a_{n+1,0} = a_{n,0} + \sum_{i=1}^{n-1} a_{i,0} a_{n-i,0}, \tag{2.22}$$

to apply a bijection with the Motzkin numbers M_n.

The number of Dyck paths with no udu's and with semi-length 0, 1, or 2 is only one, that is $a_{0,0} = a_{1,0} = a_{2,0} = 1$, thus from (2.22) $a_{3,0} = 2$.

In Chapter 1 we derived the recurrence for the Motzkin numbers M_n which gave us $M_0 = M_1 = 1$, with a recurrence as follows:

$$M_n = M_{n-1} + \sum_{i=0}^{n-2} M_i M_{n-2-i}. \tag{2.23}$$

We see from (2.23) that $M_2 = 2$.

The two recurrences (2.22) and (2.23) have the same initial conditions therefore

$$a_{n+1,0} = M_n.$$

For a Dyck path $\alpha \in S_{n-k+1,0}$ we look at $n-k+1$ endpoints of its $n-k+1$ up steps. In these $n-k+1$ up steps, we choose k points, we can choose any point with repetition allowed, and at each of the chosen points we insert a valley du. We then obtain k udu's and the semi-length increase from $n-k+1$ to $n+1$. This produces a Dyck path of semi-length $n+1$ with k occurrences of udu's that is $a_{n+1,k}$. Alternatively we can cancel the k valleys du that follow immediately after an up step u, this will result into k udu's and again we get $a_{n+1,k}$. Since the number of i-subsets of $[n]$ with repetitions allowed is $\binom{n+i-1}{i}$, we get $a_{n+1,k} = \binom{n}{k} a_{n-k+1,0}$. If we consider $a_{n+1,0} = M_n$, then $a_{n-k+1,0} = M_{n-k}$, this give us

$$a_{n+1,k} = \binom{n}{k} M_{n-k}. \tag{2.24}$$

Thus

$$a_{n,k} = \binom{n-1}{k} M_{n-k-1}, \tag{2.25}$$

as required.
Second method.

Proof

$$z(F^2 - F + xF + 1 - x) = F - 1$$

Let $w(x,z) = F(x,z) - 1$. where $w(x,z) := w$ then

$$z((w+1)^2 + (x-1)w) = w.$$

Now we apply Lagrange inversion formula to get $a_{n,k} = [x^k z^n]w$ as follows

$$\begin{aligned}
[z^n]w &= \frac{1}{n}[y^{n-1}]((y+1)^2 + (x-1)y)^n \\
&= \frac{1}{n}[y^{n-1}]\sum_{j=0}^{n}\binom{n}{i}(y+1)^{2i}(x-1)^{n-i}y^{n-i} \\
[x^k z^n]w &= [x^k]\frac{1}{n}[y^{n-1}]\sum_{i=0}^{n}\binom{n}{i}(y+1)^{2i}y^{n-i}\sum_{j=0}^{n-i}\binom{n-i}{j}(-1)^{n-j-i}x^j \\
&= \frac{1}{n}[y^{n-1}]\sum_{i=0}^{n}\binom{n}{i}(y+1)^{2i}y^{n-i}\binom{n-i}{k}(-1)^{n-k-i} \\
&= \frac{1}{n}[y^{n-1}]\sum_{i=0}^{n}\binom{n}{i}\binom{n-i}{k}(-1)^{n-k-i}\sum_{m=0}^{n}\binom{2i}{m}y^{n-i+m} \\
&= \frac{1}{n}\sum_{i=1}^{n}\binom{n}{i}\binom{n-i}{k}(-1)^{n-k-i}\binom{2i}{i-1} \\
&= \sum_{i=1}^{n}(-1)^{n-k-i}\frac{(n-1)!}{i!(n-i)!}\frac{(n-i)!}{k!(n-i-k)!}\frac{(2i)!}{(i-1)!(i+1)!}\frac{(n-k-1)!}{(n-k-1)!} \\
&= \binom{n-1}{k}\sum_{i=1}^{n}(-1)^{n-k-i}\binom{n-k-1}{i-1}C_i \\
&= \binom{n-1}{k}\sum_{i=0}^{n-k-1}(-1)^{n-k-1-i}\binom{n-k-1}{i}C_{i+1} \\
&= \binom{n-1}{k}M_{n-k-1}.
\end{aligned}$$

where the last equality follows by applying the Möbius inversion formula to the equation $C_{n+1} = \sum_{j=0}^{n} \binom{n}{j} M_{n-j}$ in Chapter one.

We now consider the string $\tau = dud$. We derive its generating function by using the first return decomposition of non-empty Dyck path $\alpha = \beta u \gamma d$ where $\alpha, \beta, \gamma \in \mathcal{D}$. A new occurrence of dud appears in α (in addition to those contributed by β and γ) if and only if $\gamma = \varepsilon$ and $\beta \neq \varepsilon$ producing $z(x(F-1))$ or α can be such that $\gamma \neq \varepsilon$ producing $z(F-1)F$ or we can have $\alpha = ud$ where $\beta = \gamma = \varepsilon$ producing z. Combining these three possibilities we get

$$\begin{aligned} F - 1 &= z(x(F-1)) + z(F-1)F + z \\ zF^2 &= (1-(x-1)z)F - (1-(x-1)z). \end{aligned}$$

This generating function for dud is the same as the generating function for the string udu, thus dud and udu are equidistant.

The table below shows some values for $a_{n,k}$.

n\\k	0	1	2	3	4	5	6
1	1						
2	1	1					
3	2	2	1				
4	4	6	3	1			
5	9	16	12	4	1		
6	21	45	40	20	5	1	
7	51	126	135	80	30	6	1

Table 2.1: The numbers $a_{n,k}$ of the string $\tau = udu$.

Theorem 2.2.3 *[16]The generating function for occurrences of the strings uuu and ddd is* $z(t+(1-t)z)F^2 - (1-(1-t)z)F + 1 = 0$.

Proof

To derive the generating function $F(t,z) = F$ for the string $\tau = uuu$, we partition $\mathcal{D}$ into $\{\Omega_i\}$, where Ω_i is the set of all Dyck paths with length of first ascent equal to i, for all $i \geq 1$, we define $A_i(t,z) = A_i$ to be a generating function for Ω_i, where

t counts the number of occurrences the string $\tau = uuu$. All elements α of Ω_i can be written uniquely as $\alpha = u^i d\alpha_1 d\alpha_2 d\alpha_3 \ldots . d\alpha_i$, where $\alpha_m \in \mathcal{D}$ for all $m \in [i]$. We conclude that there are $(i-2)$ new strings uuu (in addition to those contributed by α_i's) if and only if $i \geq 3$, since there are i (a_m's) and also there are i ($u's$), these produce $A_i = t^{i-2} z^i F^i$ and if $i \leq 2$ we get $A_i = z^i F^i$. Combining these two possibilities we get

$\sum_{i=3}^{\infty} t^{i-2} z^i F^i$ and $\sum_{i=0}^{2} z^i F^i$.

Hence we get

$$
\begin{aligned}
F &= \sum_{i=0}^{2} z^i F^i + \sum_{i=3}^{\infty} t^{i-2} z^i F^i \\
&= 1 + \sum_{i=1}^{2} z^i F^i + \sum_{i=3}^{\infty} t^{i-2} z^i F^i.
\end{aligned}
\tag{2.26}
$$

Now we do some manipulations in the above equation as follows,

$$
\begin{aligned}
F &= 1 + \sum_{i=1}^{2} z^i F^i + \sum_{i=3}^{\infty} t^{i-2} z^i F^i \\
&= 1 + \sum_{i=1}^{2} z^i F^i + \sum_{i=0}^{\infty} t^{i+3-2} z^{i+3} F^{i+3} \\
&= 1 + \sum_{i=1}^{2} z^i F^i + tz^3F^3 \sum_{i=0}^{\infty} t^i z^i F^i \\
&= 1 + zF + z^2F^2 + \frac{tz^3F^3}{1 - tzF} \\
F - tzF^2 &= 1 - tzF + zF - tz^2F^2 + z^2F^2 - tz^3F^3 + tz^3F^3 \\
z(t + (1-t)z)F^2 &- (1 - (1-t)z)F + 1 = 0.
\end{aligned}
$$

Then the generating function for the string uuu is given by,

$$
z(t + (1-t)z)F^2 - (1 - (1-t)z)F + 1 = 0. \tag{2.27}
$$

Now we derive the generating function $F(t,z) = F$ for the string $\tau = ddd$, we partition $\mathcal{D}$ into $\{\Omega_i\}$, where Ω_i is the set of all Dyck paths with length of last descent equal to i, for all $i \geq 1$, we define $A_i(t,z) = A_i$ to be a generating function for Ω_i, where t counts the number of occurrences of the string $\tau = ddd$. All elements α of Ω_i can be written uniquely as $\alpha = \alpha_1 u \alpha_2 u \alpha_3 .... u \alpha_i u d^i$, where $\alpha_m \in \mathcal{D}$ for all $m \in [i]$. We conclude that there are $(i-2)$ new strings ddd (in addition to those contributed by α_i's) if and only if $i \geq 3$, since there are i (a_m's) and also there are i ($u's$). This produces $A_i = t^{i-2} z^i F^i$ and if $i \leq 2$ we get $A_i = z^i F^i$. Combining these two possibilities we get

$\sum_{i=3}^{\infty} t^{i-2} z^i F^i$ and $\sum_{i=0}^{2} z^i F^i$.

Then as for uuu the generating function for the string ddd is,

$$z(t + (1-t)z)F^2 - (1 - (1-t)z)F + 1 = 0. \tag{2.28}$$

This generating function for ddd is the same as the generating function for the string uuu, thus ddd and uuu have the same generating function.

If we set $t = 0$ in the above equation we obtain the generating function for the Motzkin numbers in Chapter one which is given by,

$$M(z) = 1 + zM(z) + z^2 M(z)^2. \tag{2.29}$$

Therefore the sequence for the path avoiding the strings uuu and ddd is the sequence of Motzkin number $M_n = \sum_{j=0}^{[n/2]} \binom{n}{2j} C_j$.

Theorem 2.2.4 *[16]The generating function for occurrences of the strings uud and udd is $F(x,z) - 1 = z(1 + (x-1)z)F^2(x,z)$.*

In [15] it is shown how to derive the generating function for the string $\bar{\tau} := uud$, we know that the strings udd and uud are equidistant with respect to the vertical axis.

Now we show the generating function $F(x,z)$ for the string uud is given by

$$F(x,z) - 1 = z(1 + (x-1)z)F^2(x,z).$$

Proof

Now let $\alpha = u\beta d\gamma$ be the first return decomposition of the non-empty Dyck path D, where β, $\gamma \in D$. We see that the new occurrence of uud appears in α

(in addition to those from β and γ) if and only if $\beta = ud\phi$ where $\phi \in \mathcal{D}$ producing $z(xzF(x,z))F(x,z)$ or α can be such that $\beta \neq ud\phi$ producing $z(F(x,z) - zF(x,z))F(x,z)$. Combining these two possibilities we get

$$F(x,z) - 1 = z(xzF(x,z))F(x,z) + z(F(x,z) - zF(x,z))F(x,z),$$

then we have

$$\begin{aligned} z(1+(x-1)z)F^2(x,z) - F(x,z) + 1 = 0 \\ F(x,z) - 1 = z(1+(x-1)z)F^2(x,z), \end{aligned} \tag{2.30}$$

as required.

In the paper [16], it is stated that the coefficient $b_{n,k}$ of the series expansion for $F(x,z)$ is

$$b_{n,k} = [x^k z^n]F = \frac{1}{n+1}\binom{n+1}{k}\sum_{j=2k}^{n}\binom{j-k-1}{k-1}\binom{n+1-k}{n-j} \tag{2.31}$$

We now consider the string $\tau = udd$. We derive its generating function by using the first return decomposition of non-empty Dyck path $\alpha = \beta u\gamma d$ where $\alpha, \beta, \gamma \in \mathcal{D}$. A new occurrence of udd appears in α (in addition to those contributed by β and γ) if and only if $\gamma = \delta ud$ producing $z(xzF(x,z))F(x,z)$ or α can be such that $\gamma \neq \delta ud$ producing $z(F(x,z) - zF(x,z))F(x,z)$. Combining these two possibilities we get

$$\begin{aligned} F(x,z) - 1 &= z(xzF(x,z))F(x,z) + z(F(x,z) - zF(x,z))F(x,z) \\ F(x,z) - 1 &= z(1+(x-1)z)F^2(x,z). \end{aligned}$$

This generating function for udd is the same as the generating function for the string uud, thus udd and uud are equidistant.

2.3 The strings of length 4.

In this section we will study the strings of length four, there are sixteen strings τ of length four, we notice that among these strings there is a symmetry with respect to a vertical axis, the statistic "number of τ's" for some of them (given here in pairs) are equidistant:

$\{uuud, uddd\}$, $\{uuuu, dddd\}$, $\{uddu, duud\}$, $\{duuu, dddu\}$, $\{uudd\}$, $\{uudu, dudd\}$, $\{uduu, ddud\}$, $\{udud\}$,$\{dduu\}$, $\{dudu\}$.

We will derive some generating functions $F(t,z) = F$ (where t marks the number of occurrences of a string τ and z marks the semi-length n) for several strings using first return decomposition.

Theorem 2.3.1 *[16]The generating function for occurrences of the string uudd is* $zF^2 + z^2(t-1)F - F + 1 = 0$.

Proof

We derive the generating function by using the first return decomposition of the non-empty Dyck path $\alpha = u\beta d\gamma$ where $\beta, \gamma \in \mathcal{D}$, a new occurrence of $uudd$ appears in α (in addition to those contributed by β and γ) if and only if $\beta = ud$.

The path α can have a new occurrence of $uudd$ where $\beta = ud$ producing $z(tz)F$ or α can be such that $\beta \neq ud$ producing $z(F-z)F$. Combining these two cases we get

$$F - 1 = z(tz + F - z)F. \tag{2.32}$$

Then we have

$$zF^2 + z^2(t-1)F - F + 1 = 0. \tag{2.33}$$

It is known that

$$a_{n,k} = [t^k z^n]F(t,z) = \sum_{j=k}^{[n/2]} \frac{(-1)^{j-k}}{n-j} \binom{j}{k} \binom{n-j}{j} \binom{2n-3j}{n-j-1}. \tag{2.34}$$

Proof of (2.34)

We use the following equation which is in the suitable form to apply Lagrange inversion formula

$$A(z) = 1 + zH(A(z)), \tag{2.35}$$

where $H(\gamma) = (t-1)z\gamma + \gamma^2$ and $A(z) = F(t,z)$ is taken at be a function of z only. Now using the Lagrange inversion formula we get

$$[z^\sigma]A(z) = \frac{1}{\sigma}[\gamma^{\sigma-1}](H(1+\gamma))^\sigma, \tag{2.36}$$

which gives us

$$\begin{aligned}
(H(1+\gamma))^\sigma &= \sum_{j=0}^{\sigma} \binom{\sigma}{j} (t-1)^j z^j (1+\gamma)^j (1+\gamma)^{2\sigma-2j} \\
&= \sum_{j=0}^{\sigma} \binom{\sigma}{j} (t-1)^j z^j (1+\gamma)^{2\sigma-j} \\
&= \sum_{j=0}^{\sigma} \binom{\sigma}{j} (t-1)^j z^j \sum_{v=0}^{2\sigma-j} \binom{2\sigma-j}{v} \gamma^v \\
&= \sum_{v=0}^{\sigma} \sum_{j=(2\sigma-v)^+}^{\sigma} \binom{2\sigma-j}{v} \binom{\sigma}{j} (t-1)^j z^j \gamma^v,
\end{aligned} \tag{2.37}$$

where $(2\sigma - v)^+$ means we only consider non-negative integers.

We substitute (2.37) into (2.36)

$$\begin{aligned}
[z^\sigma]F(t,z) &= \frac{1}{\sigma}[\gamma^{\sigma-1}] \sum_{v=0}^{\sigma} \sum_{j=(2\sigma-v)^+}^{\sigma} \binom{2\sigma-j}{v} \binom{\sigma}{j} (t-1)^j z^j \gamma^v \\
&= \frac{1}{\sigma} \sum_{j=0}^{\sigma} \binom{2\sigma-j}{\sigma-1} \binom{\sigma}{j} (t-1)^j z^j \\
F(t,z) &= \sum_{\sigma=1}^{\infty} \frac{1}{\sigma} \sum_{j=0}^{\sigma} \binom{2\sigma-j}{\sigma-1} \binom{\sigma}{j} (t-1)^j z^{\sigma+j}.
\end{aligned} \tag{2.38}$$

Since we need z^n, let $n = \sigma + j$

$$F(t,z) = \sum_{n=0}^{\infty} \sum_{j=0}^{n-j} \frac{1}{n-j} \binom{n-j}{j} \binom{2n-3j}{n-j-1} (t-1)^j z^n.$$

Now with $n \geq 0$, $0 \leq k \leq j$ and let $j \leq \sigma = n - j$ then

$$\begin{aligned} F(t,z) &= \sum_{n=0}^{\infty} \sum_{k=0}^{j} \sum_{j=k}^{n-j} (-1)^{j-k} \frac{1}{n-j} \binom{n-j}{j} \binom{2n-3j}{n-j-1} \binom{j}{k} t^k z^n \\ &= 1 + \sum_{n=1}^{\infty} \sum_{k=0}^{j} \sum_{j=k}^{n-j} (-1)^{j-k} \frac{1}{n-j} \binom{n-j}{j} \binom{2n-3j}{n-j-1} \binom{j}{k} t^k z^n. \end{aligned} \tag{2.39}$$

For the maximum value of j we have $n - j = j$ which implies $j = [n/2]$, hence we have

$$F(t,z) = 1 + \sum_{n=1}^{\infty} \sum_{k=0}^{[n/2]} \sum_{j=k}^{[n/2]} (-1)^{j-k} \frac{1}{n-j} \binom{n-j}{j} \binom{2n-3j}{n-j-1} \binom{j}{k} t^k z^n.$$

Then

$$a_{n,k} = [t^k z^n] F(t,z) = \sum_{j=k}^{[n/2]} (-1)^{j-k} \frac{1}{n-j} \binom{n-j}{j} \binom{2n-3j}{n-j-1} \binom{j}{k}, \tag{2.40}$$

as required.

Theorem 2.3.2 *[16]The generating function for occurrences of the strings $uudu$, $uduu$, $ddud$ and $dudd$ is $z(1-(1-t)z)F^2 + ((1-t)z^2 - 1)F + 1 = 0$.*

Proof

We first consider the string $\tau = uudu$. We derive its generating function by using the first return decomposition of the non-empty Dyck path $\alpha = u\beta d\gamma$ where $\beta, \gamma \in \mathcal{D}$. A new occurrence of $uudu$ appears in α (in addition to those contributed by β and γ) if and only if $\beta = ud\phi$, where $\phi \in \mathcal{D}$ and $\phi \neq \varepsilon$.

The path α can start with the string $uudu$ right from the start where $\beta = ud\phi$ producing $z(tz(F-1))F$ or α can start without the string $uudu$ in the beginning producing $z(F - z(F-1))F$.

Now we get the generating function $F(t,z) = F$ (where t counts the number of occurrences of $uudu$'s) as follows.

$$F - 1 = z(tz(F-1) + F - z(F-1))F, \tag{2.41}$$

which gives

$$z(1-(1-t)z)F^2+((1-t)z^2-1)F+1=0. \tag{2.42}$$

We now prove that

$$a_{n,k}=[t^kz^n]F(t,z)=\sum_{j=k}^{[(n-1)/2]}(-1)^{j-k}\frac{1}{n-j}\binom{n-j}{j}\binom{2n-3j}{n-j+1}\binom{j}{k}. \tag{2.43}$$

Proof of (2.43)

As for the string *uudd* we use the following equation which is in the suitable form to apply Lagrange inversion formula

$$A(z)=1+zH(A(z)).$$

Here $H(\gamma)=(t-1)z\gamma^2-(t-1)z\gamma+\gamma^2=(t-1)z\gamma(\gamma-1)+\gamma^2$ and $A(z)=F(t,z)$ considered as a function of z only. By the Lagrange inversion formula it follows that:

$$[z^\sigma]A(z)=\frac{1}{\sigma}[\gamma^{\sigma-1}](H(1+\gamma))^\sigma, \tag{2.44}$$

which give us

$$\begin{aligned}
(H(1+\gamma))^\sigma &= \sum_{j=0}^{\sigma}\binom{\sigma}{j}(t-1)^jz^j\gamma^j(1+\gamma)^j(1+\gamma)^{2\sigma-2j}\\
&= \sum_{j=0}^{\sigma}\binom{\sigma}{j}(t-1)^jz^j\gamma^j(1+\gamma)^{2\sigma-j}\\
&= \sum_{j=0}^{\sigma}\binom{\sigma}{j}(t-1)^jz^j\gamma^j\sum_{v=0}^{2\sigma-j}\binom{2\sigma-j}{v}\gamma^v\\
&= \sum_{v=0}^{\sigma}\sum_{j=(2\sigma-v)^+}^{\sigma}\binom{2\sigma-j}{v}\binom{\sigma}{j}(t-1)^jz^j\gamma^{v+j}.
\end{aligned} \tag{2.45}$$

Again we substitute (2.45) into (2.44)

$$
\begin{aligned}
[z^\sigma]F(t,z) &= \frac{1}{\sigma}[\gamma^{\sigma-1}]\sum_{v=0}^{\sigma}\sum_{j=(2\sigma-v)^+}^{\sigma}\binom{2\sigma-j}{v}\binom{\sigma}{j}(t-1)^j z^j\gamma^{v+j} \\
&= \frac{1}{\sigma}\sum_{j=0}^{\sigma}\binom{2\sigma-j}{\sigma-j-1}\binom{\sigma}{j}(t-1)^j z^j \\
F(t,z) &= \sum_{\sigma=1}^{\infty}\frac{1}{\sigma}\sum_{j=0}^{\sigma}\binom{2\sigma-j}{\sigma-1}\binom{\sigma}{j}(t-1)^j z^{\sigma+j}
\end{aligned}
\tag{2.46}
$$

Since we need z^n let $n = \sigma + j$

$$
F(t,z) = \sum_{n=0}^{\infty}\sum_{j=0}^{n-j}\frac{1}{n-j}\binom{n-j}{j}\binom{2n-3j}{n-j+1}(t-1)^j z^n. \tag{2.47}
$$

Now with $n \geq 0$, $0 \leq k \leq j$ and let $j \leq \sigma = n - j$ then

$$
\begin{aligned}
F(t,z) &= \sum_{n=0}^{\infty}\sum_{k=0}^{j}\sum_{j=k}^{n-j}(-1)^{j-k}\frac{1}{n-j}\binom{n-j}{j}\binom{2n-3j}{n-j+1}\binom{j}{k}t^k z^n \\
&= 1 + \sum_{n=1}^{\infty}\sum_{k=0}^{j}\sum_{j=k}^{n-j}(-1)^{j-k}\frac{1}{n-j}\binom{n-j}{j}\binom{2n-3j}{n-j+1}\binom{j}{k}t^k z^n.
\end{aligned}
\tag{2.48}
$$

For the maximum value of j we have $2n - 3j = n - j + 1$ which implies $j = \frac{n-1}{2}$, hence we have

$$
F(t,z) = 1 + \sum_{n=1}^{\infty}\sum_{k=0}^{(n-1)/2}\sum_{j=k}^{(n-1)/2}(-1)^{j-k}\frac{1}{n-j}\binom{n-j}{j}\binom{2n-2j}{n-j+1}\binom{j}{k}t^k z^n.
$$

We now get

$$
a_{n,k} = [t^k z^n]F(t,z) = \sum_{j=k}^{(n-1)/2}(-1)^{j-k}\frac{1}{n-j}\binom{n-j}{j}\binom{2n-3j}{n-j+1}\binom{j}{k}. \tag{2.49}
$$

We now consider the string $\tau = dudd$. We derive its generating function by using the first return decomposition of non-empty Dyck path $\alpha = \beta u \gamma d$ where $\beta, \gamma \in \mathcal{D}$. A new occurrence of $dudd$ appears in α (in addition to those contributed by β and γ) if and only if $\gamma = \phi ud$, where $\phi \in \mathcal{D}$ and $\phi \neq \varepsilon$.

The path α can have a new occurrence of $dudd$ where $\gamma = \phi ud$ producing $z(tz(F-1))F$ or α can be such that $\gamma \neq \phi ud$ producing $z(F - z(F-1))F$. Combining these two possibilities we get

$$F - 1 = z(tz(F-1) + F - z(F-1))F. \tag{2.50}$$

The generating function for $dudd$ is the same as the generating function for the string $uudu$. Thus $dudd$ and $uudu$ are equidistributed.

We now study the string $\tau = ddud$ we derive its generating function by using the first return decomposition of non-empty Dyck path $\alpha = u\beta d\gamma$ where $\beta, \gamma \in \mathcal{D}$. A new occurrence of $ddud$ appears in α (in addition to those contributed by β and γ) if and only if $\beta \neq \varepsilon$ and $\gamma = ud\phi$.

The path α can have a new occurrence of $ddud$ where $\beta \neq \varepsilon$ and $\gamma = ud\phi$ producing $z(tz(F-1))F$ or α can be such that $\gamma \neq ud\phi$ producing $z(F-z(F-1))F$. Combining these two cases we get

$$F - 1 = z(tz(F-1) + F - z(F-1))F. \tag{2.51}$$

This generating function for $ddud$ is the same as the generating function for the string $uudu$. Thus $ddud$ and $uudu$ are equidistributed.

Consider the string $\tau = uduu$. We derive its generating function by using first return decomposition of non-empty Dyck path $\alpha = \beta u \gamma d$ where $\beta, \gamma \in \mathcal{D}$. A new occurrence of $uduu$ appears in α (in addition to those contributed by β and γ) if and only if $\beta = \delta ud$ and $\gamma \neq \varepsilon$.

The path α can have a new occurrence of $uduu$ where $\beta = \delta ud$ and $\gamma \neq \varepsilon$ producing $z(tz(F-1))F$ or α can be such that $\gamma \neq ud\phi$ producing $z(F-z(F-1))F$. Combining these two possibilities we get

$$F - 1 = z(tz(F-1) + F - z(F-1))F. \tag{2.52}$$

This generating function for $uduu$ is the same as the generating function for the string $uudu$. Thus $uduu$ and $uudu$ are equidistributed.

We see that the four strings $uudu, uduu, ddud$ and $dudd$ are all equidistributed.

Theorem 2.3.3 *[16]The generating function for occurrances of the strings uuud and uddd is* $(t-1)z^3F^3 + zF^2 - F + 1 = 0$.

Proof
We show that the string $\tau = uuud$ have the same generating function to the string $uddd$. We derive the generating function of $\tau = uuud$ using the first return decomposition of non-empty Dyck path $\alpha = u\beta d\gamma$ where $\beta, \gamma \in \mathcal{D}$. A new occurrence of $uuud$ appears in α (in addition to those contributed by β and γ) if and only if $\beta = u\varphi d\delta$ with $\varphi = ud\phi$, where $\varphi, \delta, \phi \in \mathcal{D}$. This produces $z(z(tzF)F)F$. If $\varphi \neq ud\phi$ this produces $z(F - z^2F^2)F$.

So the generating function $F(t,z) = F$ (where t counts the number of occurrences of $uuud$'s) satisfies

$$F - 1 = z(z(tzF)F + F - z^2F^2)F. \tag{2.53}$$

Thus

$$(t-1)z^3F^3 + zF^2 - F + 1 = 0. \tag{2.54}$$

In order to derive a formula for $a_{n,k}$ from the above equation, we write

$$A(z) = 1 + zH(A(z)), \tag{2.55}$$

where $H(\gamma) = (t-1)z^2\gamma^3 + \gamma^2$ and $A(z)$ where $A(z) = F(t,z)$ is a function with variable z only. By using the Lagrange inversion formula it follows that

$$[z^\sigma]A(z) = \frac{1}{\sigma}[\gamma^{\sigma-1}](H(1+\gamma))^\sigma \tag{2.56}$$

We have

$$(H(1+\gamma))^\sigma = \sum_{j=0}^{\sigma} \binom{\sigma}{j} (t-1)^j z^{2j} (1+\gamma)^{3j} (1+\gamma)^{2\sigma-2j}$$

$$
\begin{aligned}
&= \sum_{j=0}^{\sigma} \binom{\sigma}{j} (t-1)^j z^{2j} (1+\gamma)^{2\sigma+j} \\
&= \sum_{j=0}^{\sigma} \binom{\sigma}{j} (t-1)^j z^{2j} \sum_{v=0}^{2\sigma+j} \binom{2\sigma+j}{v} \gamma^v \\
&= \sum_{v=0}^{3\sigma} \sum_{j=(v-2\sigma)^+}^{\sigma} \binom{2\sigma+j}{v} \binom{\sigma}{j} (t-1)^j z^{2j} \gamma^v,
\end{aligned}
\tag{2.57}
$$

we substitute (2.57) into (2.56)

$$
\begin{aligned}
[z^\sigma]F(t,z) &= \frac{1}{\sigma}[\gamma^{\sigma-1}] \sum_{v=0}^{3\sigma} \sum_{j=(v-2\sigma)^+}^{\sigma} \binom{2\sigma+j}{v} \binom{\sigma}{j} (t-1)^j z^{2j} \gamma^v \\
&= \frac{1}{\sigma} \sum_{j=0}^{\sigma} \binom{2\sigma+j}{\sigma-1} \binom{\sigma}{j} (t-1)^j z^{2j} \\
F(t,z) &= \sum_{\sigma=1}^{\infty} \frac{1}{\sigma} \sum_{j=0}^{\sigma} \binom{2\sigma-j}{\sigma-1} \binom{\sigma}{j} (t-1)^j z^{\sigma+2j}.
\end{aligned}
\tag{2.58}
$$

Since we need $[z^n]$ which is the coefficient of z^n in the series expansion of $F(t,z)$, let $n = \sigma + 2j$, then

$$
\begin{aligned}
F(t,z) &= \sum_{n=0}^{\infty} \sum_{k=0}^{j} \sum_{j=k}^{n-2j} (-1)^{j-k} \frac{1}{n-2j} \binom{n-2j}{j} \binom{2n-3j}{n-j+1} \binom{j}{k} t^k z^n \\
&= 1 + \sum_{n=1}^{\infty} \sum_{k=0}^{j} \sum_{j=k}^{n-2j} (-1)^{j-k} \frac{1}{n-2j} \binom{n-2j}{j} \binom{2n-3j}{n-j+1} \binom{j}{k} t^k z^n.
\end{aligned}
\tag{2.59}
$$

From the above generating function we do the following manipulations

$$
\begin{aligned}
&\frac{1}{n-2j} \binom{n-2j}{j} \binom{2n-3j}{n-j+1} \binom{j}{k} \\
&= \frac{1}{n-2j} \frac{(n-2j)!}{(n-3j)!j!} \frac{(2n-3j)!}{(n-2j-1)!(n-j+1)!} \frac{j!}{k!(j-k)!}
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{(n-2j)!}{n-2j}\frac{(n+1-k)!}{(j-k)!(n+1-j)!}\frac{(2n-3j)!}{(n-3j)!(n-2j-1)!}\frac{n!}{k!(n+1-k)!n!} \\
&= \frac{1}{n+1}\binom{n+1}{k}\binom{n+1-k}{j-k}\frac{(2n-3j)!}{(n-3j)!(n)!} \\
&= \frac{1}{n+1}\binom{n+1}{k}\binom{n+1-k}{j-k}\binom{2n-3j}{n}.
\end{aligned} \tag{2.60}
$$

Therefore

$$
F(t,z) = 1 + \sum_{n=1}^{\infty}\sum_{k=0}^{j}\sum_{j=k}^{n-2j}\frac{(-1)^{j-k}}{n+1}\binom{n+1}{k}\binom{n+1-k}{j-k}\binom{2n-3j}{n}t^k z^n. \tag{2.61}
$$

For the maximum value of j, $2n-3j=n$ which implies $j=\frac{n}{3}$. Hence we have

$$
\begin{aligned}
F(t,z) &= 1 + \sum_{n=1}^{\infty}\sum_{k=0}^{n/3}\sum_{j=k}^{n/3}\frac{(-1)^{j-k}}{n+1}\binom{n+1}{k}\binom{n+1-k}{j-k}\binom{2n-3j}{n}t^k z^n \\
a_{n,k} = [t^k z^n]F(t,z) &= \frac{1}{n+1}\binom{n+1}{k}\sum_{j=k}^{[n/3]}(-1)^{j-k}\binom{n+1-k}{j-k}\binom{2n-3j}{n},
\end{aligned} \tag{2.62}
$$

as required.

We now derive the generating function for the string $uddd$ by using the first return decomposition of non-empty Dyck path $\alpha = \beta u\gamma d$ where $\beta, \gamma \in \mathcal{D}$. A new occurrence of $uddd$ appears in α (in addition to those contributed by β and γ) if and only if $\gamma = \varphi u\delta d$ with $\delta = \phi ud$. Where $\varphi, \delta, \phi \in \mathcal{D}$ producing $z(z(tzF)F)F$ or α can be such that $\delta \neq \phi ud$ producing $z(F - z^2F^2)F$. Combining these possibilities we get

$$
F - 1 = z(z(tzF)F + F - z^2F^2)F. \tag{2.63}
$$

This generating function is the same as that of $uuud$, therefore $uuud$ and $uddd$ are equidistributed.

Theorem 2.3.4 *[16]The generating function for occurrances of the strings dddd and uuuu is* $(1-t)z^3F^3 + z(t - tz + z)F^2 + (z - tz - 1)F + 1 = 0.$

Proof

To derive the generating function $F(t, z) = F$ for the string $\tau = dddd$, we partition $\mathcal{D}$ into $\{\Omega_i\}$. Where Ω_i is the set of all Dyck paths according to length of last descent equal to i, for all $i \geq 1$. We define $A_i(t, z) = A_i$ to be a generating function for Ω_i, where t counts the number of occurrences of the string $\tau = dddd$. All elements α of Ω_i can be written uniquely as $\alpha = \alpha_1 u\alpha_2 u\alpha_3....u\alpha_i ud^i$. Where $\alpha_m \in \mathcal{D}$ for all $m \in [i]$. We conclude that there are $(i-3)$ new strings $dddd$ (in addition to those contributed by α_i's) if and only if $i \geq 4$, since there are i (a_m's) and also there are i ($u's$). These produce $A_i = t^{i-3}z^iF^i$ and if $i \leq 3$ we get $A_i = z^iF^i$. Combining these two possibilities we get

$$\sum_{i=4}^{\infty} t^{i-3}z^iF^i$$

and

$$\sum_{i=0}^{3} z^iF^i.$$

Hence we have

$$\begin{aligned} F &= \sum_{i=0}^{3} z^iF^i + \sum_{i=4}^{\infty} t^{i-3}z^iF^i \\ &= 1 + \sum_{i=1}^{3} z^iF^i + \sum_{i=4}^{\infty} t^{i-3}z^iF^i. \end{aligned} \tag{2.64}$$

Now we do some manipulations in the above equation as follows,

$$\begin{aligned} F &= 1 + \sum_{i=1}^{3} z^iF^i + \sum_{i=4}^{\infty} t^{i-3}z^iF^i \\ &= 1 + \sum_{i=1}^{3} z^iF^i + \sum_{i=0}^{\infty} t^{i+4-3}z^{i+4}F^{i+4} \end{aligned}$$

$$
\begin{aligned}
&= 1 + \sum_{i=1}^{3} z^i F^i + tz^4F^4 \sum_{i=0}^{\infty} t^i z^i F^i \\
&= 1 + zF + z^2F^2 + z^3F^3 + \frac{tz^4F^4}{1-tzF} \\
F - tzF^2 &= 1 - tzF + zF - tz^2F^2 + z^2F^2 - tz^3F^3 + z^3F^3 - tz^4F^4 + tz^4F^4 \\
&(1-t)z^3F^3 + z(t - tz + z)F^2 + (z - tz - 1)F + 1 = 0.
\end{aligned}
$$

Then the generating function for the string $dddd$ is,

$$
(1-t)z^3F^3 + z(t - tz + z)F^2 + (z - tz - 1)F + 1 = 0. \tag{2.65}
$$

This generating function is the same as for $uuuu$ in [16]. Therefore $dddd$ and $uuuu$ are equidistant.

Now we derive the generating function for the general string d^r for $r \geq 2$.

We proceed as in the case of $r = 4$ above.

To derive the generating function $F(t,z) = F$ for the string $\tau = d^r$, we partition $\mathcal{D}$ into $\{\Omega_i\}$. Where Ω_i is the set of all Dyck paths according to length of the last descent equal to i, for all $i \geq 1$. We define $A_i(t,z) = A_i$ to be a generating function for Ω_i, where t counts the number of occurrences of the string $\tau = d^r$. All elements α of Ω_i can be written uniquely as $\alpha = \alpha_1 u \alpha_2 u \alpha_3 .... u \alpha_i u d^i$. Where $\alpha_m \in \mathcal{D}$ for all $m \in [i]$. We conclude that there are $(i - r + 1)$ new strings d^r (in addition to those contributed by α_i's) if and only if $i \geq r$. Since there are i (a_m's) and also there are i $(u's)$ this produces $A_i = t^{i-r+1} z^i F^i$ and if $i \leq r-1$ we get $z^i F^i$. Combining these two possibilities we get

$A_i = t^{i-r+1} z^i F^i$ for $i \geq r$ and $z^i F^i$ for $i \leq r-1$.

Hence we get

$$
\begin{aligned}
F &= \sum_{i=0}^{r-1} z^i F^i + \sum_{i=r}^{\infty} t^{i-r+1} z^i F^i \\
&= 1 + \sum_{i=1}^{r-1} z^i F^i + \sum_{i=r}^{\infty} t^{i-r+1} z^i F^i.
\end{aligned} \tag{2.66}
$$

Now we do some manipulations in the above equation as follows:

$$
\begin{aligned}
F &= 1 + \sum_{i=1}^{r-1} z^i F^i + \sum_{i=r}^{\infty} t^{i-r+1} z^i F^i \\
&= 1 + \sum_{i=1}^{r-1} z^i F^i + \sum_{i=0}^{\infty} t^{i+r-r+1} z^{i+r} F^{i+r} \\
&= 1 + \sum_{i=1}^{r-1} z^i F^i + tz^r F^r \sum_{i=0}^{\infty} t^i z^i F^i \\
&= 1 + zF + z^2F^2 + z^3F^3 + ... + z^{r-1}F^{r-1} + \frac{tz^rF^r}{1-tzF} \\
F - tzF^2 &= 1 - tzF + zF - tz^2F^2 + z^2F^2 ... - tz^{r-1}F^{r-1} + z^{r-1}F^{r-1} - tz^rF^r + tz^rF^r \\
&= 1 + tzF^2 + \sum_{i=1}^{r-1} z^i F^i - t \sum_{i=1}^{r-1} z^i F^i \\
&= 1 + tzF^2 + (1-t) \sum_{i=1}^{r-1} z^i F^i .
\end{aligned}
$$

Thus the generating function for the string d^r for all $r \geq 2$ is,

$$
F = 1 + tzF^2 + (1-t) \sum_{i=1}^{r-1} z^i F^i . \tag{2.67}
$$

This generating function is the same as the one for u^r in [16], therefore d^r and u^r are equidistant.

Theorem 2.3.5 *[16]The generating function for occurrences of the string udud is* $z(1+(1-t)z)F^2 - (1+(1-t)z(z+1))F + 1 + (1-t)z = 0.$

Proof

To derive the generating function $F(t,z) = F$ for the string $\tau = udud$. We let Ω to be the set of all Dyck paths that starts with a low peak (ud), with its generating function $A(t,z) = A$. Where t counts the number of occurrences of the string $\tau = udud$. All elements α of Ω can be written uniquely as $\alpha = ud\beta$, where

$\beta \in \mathcal{D}$. The new string *udud* occurs if and only if $\beta \in \Omega$, this produces ztA and if β does not belong to Ω then we get $z(F - A)$. Combining these two possibilities we get

$$A = ztA + z(F - A). \tag{2.68}$$

In addition, we form the non-empty first return decomposition $\alpha = u\beta d\gamma$, where $\beta, \gamma \in \mathcal{D}$. The new occurrence of *udud* (in addition to those contributed by β and γ) is possible if and only if $\beta = \varepsilon$ and $\gamma \in \Omega$, this produces ztA. If $\beta = \varepsilon$ and γ does not belong to Ω, then we have $z(F - A)$ and if we consider the case where $\beta \neq \varepsilon$, we then obtain $z(F - 1)F$. Combining these three cases we get

$$F = 1 + ztA + z(F - A) + z(F - 1)F. \tag{2.69}$$

From the above two equations we eliminate A and get the generating function F as follows.

In the equation $A = ztA + z(F - A)$ we solve for A and get
$A = \dfrac{zF}{1 - tz + z}$.
We then substitute the expression of A into
$F = 1 + ztA + z(F - A) + z(F - 1)F$, then we proceed as follows:

$$\begin{aligned} F &= 1 + zt(\frac{zF}{1 - tz + z}) + z(F - \frac{zF}{1 - tz + z}) + z(F - 1)F \\ F - ztF + zF &= 1 - zt + z + z^2tF + zF - z^2tF + z^2F - z^2F + (zF^2 - zF)(1 - zt + z) \\ (z - z^2t + z^2)F^2 &- (1 + z - zt - z^2t + z^2)F + 1 + (1 - t)z = 0. \end{aligned}$$

Now the generating function of the string *udud* is

$$z(1 + (1 - t)z)F^2 - (1 + (1 - t)z(z + 1))F + 1 + (1 - t)z = 0. \tag{2.70}$$

Theorem 2.3.6 *[16]The generating function for occurrences of the strings dudu is* $zF^2 + ((1 - t)(z - 1)z - 1)F + (1 - t)z + 1 = 0.$

Proof

To derive the generating function $F(t,z) = F$ for the string $\tau = dudu$. We let Ω to be the set of all Dyck paths with semi-length at least 2, that starts with a low peak, having the generating function $A(t,z) = A$. Where t counts the number of occurrences of the string $\tau = dudu$. All elements α of Ω can be written uniquely as $\alpha = ud\beta$, where $\beta \in \mathcal{D}\backslash\{\varepsilon\}$. The new string $dudu$ occurs if and only if $\beta \in \Omega$, this produces ztA and if β does not belong to Ω then we get $z(F-1-A)$. Combining these two possibilities we get

$$A = ztA + z(F-1-A). \tag{2.71}$$

In addition we form the first return decomposition $\alpha = u\beta d\gamma$. Where $\beta, \gamma \in \mathcal{D}$, The new occurrence of $dudu$ (in addition to those contributed by β and γ) is possible if and only if $\gamma \in \Omega$. This produces $ztFA$, if is not true that $\gamma \in \Omega$, then we have $zF(F-A)$. Combining these two possibilities we get

$$F = 1 + ztFA + z(F-A)F. \tag{2.72}$$

From the above two equations we eliminate A and get the generating function F as follows.

In the equation $A = ztA + z(F-1-A)$ we solve for A and get
$A = \dfrac{zF - z}{1 - tz + z}$.

We then substitute the expression of A into $F = 1 + ztFA + zF^2 - zFA$ then we proceed as follows:

$$\begin{aligned} F &= 1 + ztF\frac{zF-z}{1-tz+z} + zF^2 - zF\frac{zF-z}{1-tz+z} \\ F - ztF + zF &= 1 - tz + z + z^2tF^2 - z^2tF + zF^2 - z^2tF^2 + z^2F^2 - z^2F^2 + z^2F \\ zF^2 &+ (zt - z - z^2t + z^2 - 1)F + (1-t)z + 1 = 0. \end{aligned} \tag{2.73}$$

Now the generating function of the string $dudu$ is

$$zF^2 + ((1-t)(z-1)z - 1)F + (1-t)z + 1 = 0. \tag{2.74}$$

Theorem 2.3.7 *[16]The generating function for occurrences of the strings duud and uddu is* $zF^3 - ((1-t)z+1)F^2 + (1+2(1-t)z)F - (1-t)z = 0$.

Proof

To derive the generating function $F(t,z) = F$ for the string $\tau = duud$. We let Ω to be the set of all Dyck paths where the size of the first ascent is equal to 2, with its generating function $A(t,z) = A$. Where t counts the number of occurrences of the string $\tau = duud$. All elements α of Ω can be written uniquely as $\alpha = uud\alpha_1 d\alpha_2$, where $\alpha_1, \alpha_2 \in \mathcal{D}$. If both α_1 and α_2 do not belong to Ω then we get $z^2(F-A)^2$. If both α_1 and α_2 belong to Ω we get $z^2t^2A^2$. If only α_1 belongs to Ω we get $z^2t(F-A)A$ and if only α_2 belongs to Ω we get $z^2t(F-A)A$. Combining these four possibilities we get

$$A = z^2(F-A)^2 + 2z^2t(F-A)A + z^2t^2A^2. \tag{2.75}$$

We know that every non-empty $\alpha \in \mathcal{D}$ can be written uniquely as a first return decomposition, that is $\alpha = u\beta d\gamma$. Where $\beta, \gamma \in \mathcal{D}$, we see that a new string $duud$ (in addition to those contributed by β and γ) appears if and only if $\gamma \in \Omega$, producing $ztAF$. If γ do not belong to Ω then we get $z(F-A)F$. Combining these two possibilities we now get

$$F - 1 = ztAF + z(F-A)F. \tag{2.76}$$

From the above two equations we get the results similar to those of the string *uddu* where its generating function is derived in paper [16].

Therefore the strings *duud* and *uddu* are equidistant.

Chapter 3

Counting strings at even, odd, low and high levels

In this Section 3.1 we continue to study the paper titled Counting strings in Dyck paths by A. Sapounakis, I. Tasoulas, and P. Tsikouras [16]. Here we take into account the number of occurrences of the string τ at even and odd height.

We define the height m of the string τ, where $m \in \{0, 1, 2, 3, ...\}$, to be the minimum height of the point(s) in which τ occurs. For example in Figure 3.1 there is only one string uuu at height zero and one ddu at height zero and one.

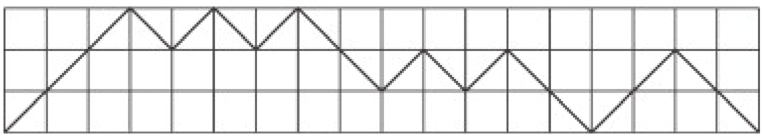

Figure 3.1: The Dyck path *uuudududdudududduudd*.

3.1 The relationship between strings of lengths 2 and 3

We now use the first return decomposition of the non-Dyck path $\alpha = u\beta d\gamma$ where (α, β, $\gamma \in \mathcal{D}$) to derive the generating functions for the occurrence of the string τ at an even height $E_\tau(t, z) := E$ (where t counts the number of occurrences of the string τ at even height) and at odd height $O_\tau(t, z) := O$ (where t counts the number of occurrences of the string τ at odd height). In this section we will show that $E_\tau(t, z) = F_{\tau_1}(t, z)$ and $O_\tau(t, z) = F_{\tau_2}(t, z)$, where $F_{\tau_1}(t, z)$ and $F_{\tau_2}(t, z)$ are generating functions for the strings τ_1 and τ_2 respectively defined in Chapter one.

Now the number statistic is the number of occurrences of τ at odd height can occur at odd height in β and even height γ thus:

$$\begin{aligned} O_\tau(t,z) - 1 &= zO_\tau(t,z)E_\tau(t,z) \\ O_\tau(t,z) &= \frac{1}{1 - zE_\tau(t,z)}. \end{aligned} \tag{3.1}$$

Theorem 3.1.1 *[16]The generating functions for occurrences of the string ud at even and odd height are* $z(z - zt + 1)E^2 + (zt - z - 1)E + 1 = 0$ *and* $zO^2 = (1 - (t-1)z)O - (1 - (t-1)z)$ *respectively.*

Proof

We now consider the string $\tau = ud$. We derive its generating function by using the first return decomposition of non-empty Dyck path $\alpha = u\beta d\gamma$ where $\alpha, \beta, \gamma \in \mathcal{D}$. The occurrences of ud at even height in α consist of the ones at odd height in β, as well as the ones at even height in γ. A new occurrence of ud appears in α (in addition to those contributed by β and γ) if and only if $\beta = \varepsilon$ producing $z(tE)$ or α starts with uu where $\beta \neq \varepsilon$ producing $z(O-1)E$. Combining these two possibilities we get,

$$E = 1 + z(tE + (O-1)E).$$

Now using $O_\tau(t,z) = \dfrac{1}{1 - zE_\tau(t,z)}$ from (3.1) we get

$$\begin{aligned} E &= 1 + ztE + zOE - zE \\ E &= 1 + ztE + zE\frac{1}{1 - zE} - zE \\ E - zE^2 &= 1 - zE + ztE - z^2tE^2 + zE - zE + z^2E^2 \\ 0 &= z(z - zt + 1)E^2 + (zt - z - 1)E + 1. \end{aligned}$$

From (3.1) let $E = \dfrac{O-1}{zO}$ substituting this into $E = 1 + ztE + zOE - zE$ we get

$$\begin{aligned}
\frac{O-1}{zO} &= 1 + zt\frac{O-1}{zO} + zO\frac{O-1}{zO} - z\frac{O-1}{zO} \\
O - 1 &= zO + zt(O-1) + z(O-1)O - z(O-1) \\
O - 1 &= (zt - z)O + zO^2 + (1-t)z \\
zO^2 &= (1-(t-1)z)O - (1-(t-1)z).
\end{aligned}$$

We see that $zO^2 = (1-(t-1)z)O - (1-(t-1)z)$ is the same as the generating function for the string udu which is

$zF^2(x,z) = (1-(x-1)z)F(x,z) - (1-(x-1)z)$.

We conclude that the generating function for ud at odd height is equidistant to the generating function for udu.

Theorem 3.1.2 *[16]The generating functions for occurrences of the string du at even and odd height are* $zE^2 - (1+(1-t)z)E + (1-t)z + 1 = 0$ *and* $z(t+z-tz)O^2 - (1-z+tz)O + 1 = 0$ *respectively.*

Proof

We now consider $\tau = du$. We derive its generating function by using the first return decomposition of non-empty Dyck path $\alpha = \beta u \gamma d$ where $\alpha, \beta, \gamma \in \mathcal{D}$. The occurrences of du at even height in α consist of the ones at even height in β, as well as the ones at odd height in γ. A new occurrence of du appears in α (in addition to those contributed by β and γ) if and only if $\beta \neq \varepsilon$ producing $z(t(E-1)O)$ or α can be such that $\beta = \varepsilon$ producing zO. Combining these two possibilities we get,

$$E = 1 + z(t(E-1)O + O).$$

Now using $O_\tau(t,z) = \dfrac{1}{1 - zE_\tau(t,z)}$ we get

$$E = 1 + ztE\frac{1}{1-zE} - tz\frac{1}{1-zE} + z\frac{1}{1-zE}$$

$$E - zE^2 = 1 - zE + ztE - zt + z$$
$$zE^2 - (1 + (1-t)z)E + (1-t)z + 1 = 0.$$

We see that the generating function at even height for the string du i.e $\bar{du}$ which is

$zE^2 = (1-(t-1)z)E - (1-(t-1)z)$ is the same as the generating function for the string udu and dud which is

$zF^2(x,z) = (1-(x-1)z)F(x,z) - (1-(x-1)z)$.

From (3.1) let $E = \dfrac{O-1}{zO}$ substituting this into $E = 1 + ztEO - tzO + zO$ we get

$$\frac{O-1}{zO} = 1 + ztO\frac{O-1}{zO} - tzO + zO$$
$$O - 1 = zO + ztO^2 - ztO - z^2tO^2 + z^2O^2$$
$$z(t + z - tz)O^2 - (1 - z + tz)O + 1 = 0.$$

We see that the generating function at odd height for the string $\bar{du}$ which is

$z(t+z-tz)O^2 - (1-z+tz)O + 1 = 0$ is the same as the generating function for the string uuu and ddd which is $z(t+z-tz)F^2 - (1-z+tz)F + 1 = 0$.

From Theorem 3.1.1 and 3.1.2 we see that $zO_{ud}^2 = (1-(t-1)z)O_{ud} - (1-(t-1)z)$ and $zE^2 = (1-(t-1)z)E - (1-(t-1)z)$ are the same generating functions respectively.

We conclude that the generating function for $\bar{du}$ is equidistant to the generating function for udu.

Theorem 3.1.3 *[16]The generating functions for occurrences of the string dd at even and odd height are $E - 1 = z(1+(t-1)z)E^2$ and $tzO^2(t,z) - (1+2(t-1)z)O(t,z) + 1 + (t-1)z = 0$ respectively.*

Proof

We consider the string $\tau = dd$. We want to show that the generating function at even height ($E_{dd}(t,z)$) for dd is the same as the generating function for the string $\tau = uud$ similarly the generating function for dd at odd height ($O_{dd}(t,z)$) is the same as the generating function for the string $\tau = duu$.

To derive the generating function $E(t,z) = E$ for the string $\tau = dd$, we partition $\mathcal{D}$ into $\{\Omega_i\}$, where Ω_i is the partitioning of all Dyck paths according the length of the last descent equal to i, for all $i \geq 1$. We define $A_i(t,z) = A_i$ to be the generating function for Ω_i. The number of dd at even height in all elements α of Ω_i, for $\alpha = \alpha_1 u \alpha_2 u \alpha_3 .... u \alpha_i u d^i$, include those at even height in every α_m for $i - m$ even. Where $\alpha_m \in \mathcal{D}$ for all $m \in [i]$, as well as those at odd height in every α_m for $i - m$ odd, where $\alpha_m \in \mathcal{D}$ for all $m \in [i]$, together with those that occur in the last descent d^i. Thus

$$\begin{aligned}
E &= 1 + \sum_{i=1}^{\infty} z^{2i} t^i E^i O^i + \sum_{i=1}^{\infty} z^{2i-1} t^{i-1} E^i O^{i-1} \\
&= 1 + z^2 tEO \sum_{i=1}^{\infty} z^{2i-2} t^{i-1} E^{i-1} O^{i-1} + zE \sum_{i=1}^{\infty} z^{2i-2} t^{i-1} E^{i-1} O^{i-1} \\
&= 1 + \frac{z^2 tEO + zE}{1 - z^2 tEO} \\
E - z^2 tE^2 O &= 1 - z^2 tEO + z^2 tEO + zE
\end{aligned}$$

$$\text{From (3.1) we have } O_\tau(t,z) = \frac{1}{1 - zE_\tau(t,z)}$$

$$\begin{aligned}
E - z^2 tE^2 \frac{1}{1 - zE} &= 1 + zE \\
E - zE^2 - z^2 tE^2 &= 1 - zE + zE - z^2 E^2 \\
E - 1 &= z(1 + (t-1)z)E^2.
\end{aligned}$$

Now the generating function at even height ($E_{dd}(t,z)$) for the string dd is,

$$E_{dd}(t,z) - 1 = z(1 + (t-1)z)E_{dd}^2(t,z).$$

The generating function for uud, derived from Chapter 2 is

$$F(x,z) - 1 = z(1 + (x-1)z)F^2(x,z).$$

These two generating functions are the same. Therefore the string dd at even height is equidistant to the strings uud and udd.

From (3.1) let $E = \dfrac{O-1}{zO}$ substituting this into

$E - 1 = z(1 + (t-1)z)E^2$ we obtain
$\dfrac{O-1}{zO} - 1 = z(1+(t-1)z)\left(\dfrac{O-1}{zO}\right)^2$ then

$$tzO^2_{dd}(t,z) - (1 + 2(t-1)z)O_{dd}(t,z) + 1 + (t-1)z = 0.$$

We see that this generating function of dd at odd height is the same as the generating function for duu and ddu which is
$xzF^2(x,z) - (1 + 2(x-1)z)F(x,z) + 1 + (x-1)z = 0$,
therefore the string dd at odd height is equidistant to the string duu.

3.2 The relationship between strings of lengths 3 and 4

Theorem 3.2.1 *[16]The generating functions for occurrences of the string duu at even and odd height are* $(z + z^2t - z^2)E^2 + (z^2 - z^2t - 1)E + 1 = 0$ *and* $z(t + (1-t)z)O - (1 + (1-t)(z-2)z)O + (t-1)z + 1 = 0$ *respectively.*

Proof
We consider the string $\tau = duu$. We derive its generating function by using the first return decomposition of non-empty Dyck path $\alpha = \beta u \gamma d$, where $\alpha, \beta, \gamma \in \mathcal{D}$. The occurrences of duu at even height in α consist of the ones at even height in β, as well as the ones at odd height in γ. A new occurrence of duu appears in α (in addition to those contributed by β and γ) if and only if $\beta, \gamma \neq \varepsilon$ producing $z(t(E-1)(O-1))$ or α can be such that $\beta = \varepsilon$ producing zO or $\gamma = \varepsilon$ producing zE or $\beta = \gamma = \varepsilon$ producing z. Combining these four possibilities we get,

$$E = 1 + zt(E-1)(O-1) + z(O-1) + z(E-1) + z.$$

Now substituting $O_\tau(t,z) = \dfrac{1}{1 - zE_\tau(t,z)}$ into
$E = 1 + z(t(E-1)(O-1)) + zO + zE + z$ we get

$$E = 1 + ztE\frac{1}{1-zE} - zEt - zt\frac{1}{1-zE} + zt + z\frac{1}{1-zE} - z + zE$$
$$E - zE^2 = 1 - zE + ztE - ztE + z^2tE^2 - zt + zt - z^2tE + z - z + z^2E + zE - z^2E^2$$

$$(z + z^2t - z^2)E^2 + (z^2 - z^2t - 1)E + 1 = 0.$$

Therefore the generating function at even height for duu is

$$z(1 - (1 - t)z)E^2 + ((1 - t)z^2 - 1)E + 1 = 0.$$

Therefore the generating function at even height for duu is equidistant with generating function for $uudu$, which is

$$z(1 - (1 - t)z)F^2 + ((1 - t)z^2 - 1)F + 1 = 0.$$

Let $E = \dfrac{O-1}{zO}$, substituting this into $E = 1+zEOt-zEt-zOt+zt+zO-z+zE$ we get

$$\frac{O-1}{zO} = 1 + ztO\frac{O-1}{zO} - zt\frac{O-1}{zO} - zOt + zt + zO - z + z\frac{O-1}{zO}$$
$$O - 1 = zO + ztO^2 - ztO - ztO + zt - z^2tO^2 + z^2tO + z^2O^2 - z^2O + zO - z$$
$$z(t + (1 - t)z)O - (1 + (1 - t)(z - 2)z)O + (t - 1)z + 1 = 0.$$

Therefore the generating function at odd height for duu is

$$z(t + (1 - t)z)O - (1 + (1 - t)(z - 2)z)O + (t - 1)z + 1 = 0.$$

In paper [16] it is stated that the generating function for $dduu$ is,

$$z(t + (1 - t)z)F^2 - (1 + (1 - t)(z - 2)z)F + (t - 1)z + 1 = 0.$$

Therefore the generating function at odd height for duu is equidistant with generating function for $dduu$.

Theorem 3.2.2 *[16]The generating functions for occurrences of the string udu at even and odd height are* $(z - z^2t + z^2)E^2 + (zt - z + z^2t - z^2 - 1)E + (1 - t)z + 1 = 0$ *and* $zO^2 + (zt - z^2t - z + z^2 - 1)O + (1 - t)z + 1 = 0$ *respectively.*

Proof

We now consider the string $\tau = udu$. We derive its generating function at even height by using the first return decomposition of non-empty Dyck path $\alpha = u\beta d\gamma$, where $\alpha, \beta, \gamma \in \mathcal{D}$, the occurrences of duu at even height in α consist of the ones at even height in γ, as well as the ones at odd height in β. A new occurrence of udu appears in α (in addition to those contributed by β and γ) if and only if $\beta = \varepsilon$. Where $\gamma \neq \varepsilon$ we have $z(t(E-1))$. If α starts with uu where $\beta \neq \varepsilon$ we have $z(O-1)E$. If $\alpha = ud$ where $\beta = \gamma = \varepsilon$ we have z. Combining these three possibilities we get

$$E - 1 = z(t(E-1)) + z(O-1)E + z.$$

Now substituting $O_\tau(t,z) = \dfrac{1}{1 - zE_\tau(t,z)}$ into $E = 1 + ztE - zt + zOE - zE + z$ we get

$$\begin{aligned}
E &= 1 + ztE - zt + zE\frac{1}{1-zE} - zE + z \\
E - zE^2 &= 1 - zE + ztE - z^2tE^2 - zt + z^2tE + zE - zE + z^2E^2 + z - z^2E \\
(z - z^2t + z^2)E^2 &+ (zt - z + z^2t - z^2 - 1)E + (1-t)z + 1 = 0.
\end{aligned}$$

Thus the generating function at even height for udu is

$$z(1 + (1-t)z)E^2 - (1 + (1-t)z(z+1))E + 1 + (1-t)z = 0.$$

The generating function at even height for udu is equidistant with generating function for $udud$, which is

$$z(1 + (1-t)z)F^2 - (1 + (1-t)z(z+1))F + 1 + (1-t)z = 0.$$

From (3.1) let $E = \dfrac{O-1}{zO}$, substituting this into $E = 1 + ztE - zt + zOE - zE + z$ we get

$$\begin{aligned}
\frac{O-1}{zO} &= 1 + zt\frac{O-1}{zO} - zt + zO\frac{O-1}{zO} - z\frac{O-1}{zO} + z \\
O - 1 &= zO + ztO - zt - z^2tO + zO^2 - zO - zO + z + z^2O
\end{aligned}$$

$$zO^2 + (zt - z^2t - z + z^2 - 1)O + (1 - t)z + 1 = 0.$$

Thus the generating function at odd height for udu is

$$zO^2 + ((1 - t)(z - 1)z - 1)O + (1 - t)z + 1 = 0.$$

The generating function at odd height for udu is equidistant with generating function for $dudu$ which is,

$$zF^2 + ((1 - t)(z - 1)z - 1)F + (1 - t)z + 1 = 0.$$

Theorem 3.2.3 *[16]The generating functions for occurrences of the string uud at even and odd height are* $zE^2 + (z^2t - z^2 - 1)E + 1 = 0$ *and* $z(1 - (1 - t)z)O^2 + ((1 - t)z^2 - 1)O + 1 = 0$ *respectively.*

Proof

We now consider the string $\tau = uud$. We derive the generating function at even height by using the first return decomposition of non-empty Dyck path $\alpha = u\beta d\gamma$. Where $\alpha, \beta, \gamma \in D$, the occurrences of uud at even height in α consist of the ones at even height in γ, as well as the ones at odd height in β. A new occurrence of udu at appears in α (in addition to those contributed by β and γ) if and only if $\beta = ud\delta$. Where $\delta \in \mathcal{D}$ producing $z(t(zO)E)$ or α can be such that $\beta \neq ud\delta$ producing $z(O - zO)E$. Combining these two possibilities we get,

$$E = 1 + z(t(zO)E) + z(O - zO)E.$$

Now substituting $O_\tau(t, z) = \dfrac{1}{1 - zE_\tau(t, z)}$ into $E = 1 + ztOE + zOE - z^2OE$ we get

$$\begin{aligned}
E &= 1 + z^2tE\frac{1}{1 - zE} + zE\frac{1}{1 - zE} - z^2E\frac{1}{1 - zE} \\
E - zE^2 &= 1 - zE + z^2tE + zE - z^2E \\
zE^2 &+ (z^2t - z^2 - 1)E + 1 = 0.
\end{aligned}$$

Thus the generating function at even height for *uud* is

$$zE^2 + (z^2(t-1) - 1)E + 1 = 0.$$

The generating function at even height for *uud* is equidistant with generating function for *uudd* which is,

$$zF^2 + (z^2(t-1) - 1)F + 1 = 0.$$

Let $E = \dfrac{O-1}{zO}$, substituting this into $E = 1 + z^2tOE + zOE - z^2OE$ we get

$$\begin{aligned}
\frac{O-1}{zO} &= 1 + z^2tO\frac{O-1}{zO} + zO\frac{O-1}{zO} - z^2O\frac{O-1}{zO} \\
O - 1 &= zO + z^2tO^2 - z^2tO + zO^2 - zO - z^2O^2 + z^2O \\
z(1-(1-t)z)O^2 &+ ((1-t)z^2 - 1)O + 1 = 0.
\end{aligned}$$

Thus the generating function at odd height for *uud* is

$$z(1-(1-t)z)O^2 + ((1-t)z^2 - 1)O + 1 = 0.$$

The generating function at odd height for *uud* is equidistant with generating function for *uudu* which is,

$$z(1-(1-t)z)F^2 + ((1-t)z^2 - 1)F + 1 = 0.$$

Theorem 3.2.4 *[16]The generating functions for occurrences of the string ddd at even and odd height are* $z(1-(1-t)z)E^2 + ((1-t)z^2 - 1)E + 1 = 0$ *and* $z(t+(1-t)z)O^2 - (1+(1-t)(z-2)z)O + 1 - (1-t)z$ *respectively.*

Proof

Lastly we consider the string $\tau = ddd$. We want to show that the generating function for *ddd* at even height ($E_{ddd}(t,z)$) is the same as the generating function for the string $\tau = uudu$ similarly the generating function for *ddd* at odd height ($O_{ddd}(t,z)$) is the same as the generating function for the string $\tau = dduu$.

To derive the generating function $E(t,z) = E$ for the string $\tau = ddd$, we partition $\mathcal{D}$ into $\{\Omega_i\}$. where Ω_i is partitioning of all Dyck paths according to length of the last descent equal to i, for all $i \geq 1$. We define $A_i(t,z) = A_i$ to be a generating function for Ω_i, where t counts the number of occurrence of the strings $\tau = ddd$ at even height. The number of ddd at even height in all elements α of Ω_i, for $\alpha = \alpha_1 u \alpha_2 u \alpha_3 .... u \alpha_i u d^i$, include those at even height in every α_m for $i - m$ even. Where $\alpha_m \in \mathcal{D}$ for all $m \in [i]$, as well as those at odd height in every α_m for $i - m$ odd, where $\alpha_m \in \mathcal{D}$ for all $m \in [i]$. Together with those that occur in the last descent d^i. Thus

$$
\begin{aligned}
E &= 1 + \sum_{i=1}^{\infty} z^{2i} t^{i-1} E^i O^i + \sum_{i=1}^{\infty} z^{2i-1} t^{i-1} E^i O^{i-1} \\
&= 1 + z^2 OE \sum_{i=1}^{\infty} z^{2i-2} t^{i-1} E^{i-1} O^{i-1} + zE \sum_{i=1}^{\infty} z^{2i-2} t^{i-1} E^{i-1} O^{i-1} \\
&= 1 + \frac{z^2 EO + zE}{1 - z^2 tEO} \\
E - z^2 t E^2 O &= 1 - z^2 tEO + z^2 EO + zE \\
E - z^2 t E^2 \frac{1}{1 - zE} &= 1 - z^2 tE \frac{1}{1 - zE} + z^2 E \frac{1}{1 - zE} + zE \\
E - zE^2 - z^2 t E^2 &= 1 - zE - z^2 tE + z^2 E + zE - z^2 E^2 \\
& z(1 - (1-t)z)E^2 + ((1-t)z^2 - 1)E + 1 = 0.
\end{aligned}
$$

The generating function at even height for ddd is equidistant with generating function for $uudu$ which is

$$
z(1 - (1-t)z)F^2 + ((1-t)z^2 - 1)F + 1 = 0.
$$

Let $E = \dfrac{O - 1}{zO}$, substituting this into $E - z^2 t E^2 O = 1 - z^2 tEO + z^2 EO + zE$ we get

$$
\begin{aligned}
\frac{O-1}{zO} - z^2 t \left(\frac{O-1}{zO} \right)^2 O = 1 - z^2 tO \frac{O-1}{zO} + z^2 O \frac{O-1}{zO} + z \frac{O-1}{zO} \\
z(t + (1-t)z)O^2 - (1 + (1-t)(z-2)z)O + 1 - (1-t)z = 0.
\end{aligned}
$$

Thus the generating function at odd height for *ddd* is

$$z(t+(1-t)z)O^2-(1+(1-t)(z-2)z)O+1-(1-t)z=0.$$

In paper [16] it is stated that the generating function for *dduu* is,

$$z(t+(1-t)z)F^2-(1+(1-t)(z-2)z)F+(t-1)z+1=0.$$

Therefore the generating function at odd height for *ddd* is equidistant with generating function for *dduu* which is,

$$z(t+(1-t)z)F^2-(1+(1-t)(z-2)z)F+1-(1-t)z=0.$$

We summarize our results in the following table

τ	$E_\tau(t,z)$	$O_\tau(t,z)$
dd	uud	duu
du	udu	ddd
ud	$\bar{d}u$	udu
duu	$uudu$	$dduu$
udu	$udud$	$dudu$
uud	$uudd$	$uduu$
ddd	$uudu$	$dduu$

Table 3.1: The relationship between strings of different lengths

3.3 Counting strings at low level

In this section we study the paper titled Dyck paths statistics by A. Sapounakis, I. Tasoulas, P. Tsikouras [15]. Here we take into account the number of low occurrences of the string τ.

We say that a string τ is at low height (level) if its minimum point(s) occurs on the horizontal axis. For example the string uuu in Figure 3.1 (page 51) is at low level. We define the generating function for low occurrences of the string τ as follows

$$L(t,z) = \sum_{n=0}^{\infty}\sum_{k=0}^{n} l_{n,k}t^k z^n, \tag{3.2}$$

where t counts the number of low occurrences of the string τ with z marking the semi-length.

In this section we are going to use the Dyck path generating function equation

$$C(z) = 1 + zC^2(z), \tag{3.3}$$

to simplify some manipulations, we are also going to use

$$[z^n]C^s(z) = \frac{s}{2n+s}\binom{2n+s}{n}, \tag{3.4}$$

which was derived in Chapter one.

We use the first return decomposition of non-empty Dyck path $\alpha = u\beta d\gamma$ to derive the generating function $L(t,z) := L$ for low occurrence of the string τ. In the decomposition $\alpha = u\beta d\gamma$, β is not on the horizontal axis therefore it cannot generate the low τ thus it is only a Dyck path $C(z) := C$. The path γ is at the horizontal axis, thus it can generate the low τ producing L. Combining these cases we obtain the generating function for the low occurrence of τ as follows,

$$L = 1 + zCL$$
$$L = \frac{1}{1-zC}.$$

(3.5)

In the following Subsections we shall consider several strings of lengths 2, 3, 4 and r such as: $ud, du, dd, duu, ddd, uudd, uudu, uduu, uuud, udud$, $dudu$ and d^r.

Theorem 3.3.1 *[15]The generating function for low occurrences of the string ud is* $L = \dfrac{1}{1-zt-z(C-1)}$.

Proof

We consider the string $\tau = ud$ we first derive its generating function $L(t,z) = L$ (where t counts the number of occurrences of ud at low level). We use the first return decomposition of non-empty Dyck path $\alpha = u\beta d\gamma$ where $\alpha, \beta, \gamma \in \mathcal{D}$. The path α can start with the string ud right from the start where $\beta = \varepsilon$ producing $z(tL)$ or α can start with uu where $\beta \neq \varepsilon$ producing $z(C-1)L$. Combining these two possibilities we get

$$L = 1 + z(tL + (C-1)L).$$

Thus the generating function for the low occurrence of ud is,

$$L = \frac{1}{1-zt-z(C-1)}.$$

Now we get $[z^n t^k]L = l_{n,k}$ as follows,

$$L = \frac{1}{1-zt-z(C-1)}$$

Using equation (3.3) we see that $C - 1 = zC^2$ then we get

$$\begin{aligned} L &= \frac{1}{1-zt-z(zC^2)} \\ &= \sum_{m=0}^{\infty} (t+zC^2)^m z^m \\ &= \sum_{m=0}^{\infty} \sum_{j=0}^{m} \binom{m}{j} z^j C^{2j} t^{m-j} z^m \end{aligned}$$

$$[z^n]L = \sum_{m=0}^{\infty}\sum_{j=0}^{m}\binom{m}{j}[z^{n-j-m}]C^{2j}t^{m-j}.$$

Now from equation (3.4) $[z^n]C^s(z) = \dfrac{s}{2n+s}\binom{2n+s}{n}$ we get $[z^{n-j-m}]C^{2j} = \dfrac{2j}{2(n-j-m)+2j}\binom{2(n-j-m)+2j}{n-j-m}$, then

$$\begin{aligned}[z^n]L &= \sum_{m=0}^{\infty}\sum_{j=0}^{m}\binom{m}{j}\frac{2j}{2(n-j-m)+2j}\binom{2(n-j-m)+2j}{n-j-m}t^{m-j}\\ &= \sum_{m=0}^{\infty}\sum_{j=0}^{m}\binom{m}{j}\frac{j}{n-m}\binom{2n-2m}{n+j-m}t^{m-j}.\end{aligned}$$

Now let $k = m-j$, then

$$[t^kz^n]L = l_{n,k} = \sum_{j=0}^{m}\binom{j+k}{k}\frac{j}{n-k-j}\binom{2n-2k-2j}{n-k}.$$

For the maximum value of j in terms of n and k we have $2n-2k-2j = n-k$, then $j = \dfrac{n-k}{2}$ thus,

$$[t^kz^n]L = l_{n,k} = \sum_{j=0}^{[\frac{n-k}{2}]}\binom{j+k}{k}\frac{j}{n-k-j}\binom{2n-2k-2j}{n-k}.$$

Theorem 3.3.2 *[15]The generating function for low occurrences of the string du is* $L = 1 + \dfrac{zC}{1-ztC}$.

Proof

We now do the string $\tau = du$, we derive its generating function $L(t,z) := L$ (where t counts the number of low occurrence of the string $\tau = du$) by using the first return decomposition of non-empty Dyck path $\alpha = \beta u\gamma d$ where $\alpha, \beta, \gamma \in \mathcal{D}$. A new low occurrence of du appears in α (in addition to those contributed by β and γ) if and only if $\beta \neq \varepsilon$ producing $z(t(L-1)C)$ or α can be such that $\beta = \varepsilon$ producing zC. Combining these two possibilities we get,

$$
\begin{aligned}
L &= 1 + z(t(L-1)C + C) \\
L - ztLC &= 1 - ztC + zC \\
L &= \frac{1 - ztC + zC}{1 - ztC}.
\end{aligned}
$$

The generating function for the low occurrence of du is,

$$L = 1 + \frac{zC}{1 - ztC}.$$

Now we get $[z^n t^k]L = l_{n,k}$ as follows,

$$
\begin{aligned}
L &= 1 + \frac{zC}{1 - ztC} \\
&= 1 + zC \sum_{m=0}^{\infty} C^m z^m t^m \\
&= 1 + \sum_{m=0}^{\infty} C^{m+1} z^{m+1} t^m \\
[z^n]L &= [z^n] \sum_{m=0}^{\infty} C^{m+1} z^{m+1} t^m \\
&= \sum_{m=0}^{\infty} [z^{n-m-1}] C^{m+1} t^m \\
&= \sum_{m=0}^{\infty} \frac{m+1}{2(n-m-1)+m+1} \binom{2(n-m-1)+m+1}{n-m-1} t^m \\
&= \sum_{m=0}^{\infty} \frac{m+1}{2n-m-1} \binom{2n-m-1}{n-m-1} t^m \\
[t^k z^n]L &= \frac{k+1}{2n-k-1} \binom{2n-k-1}{n-k-1}.
\end{aligned}
$$

Therefore

$$l_{n,k} = \frac{k+1}{2n-k-1} \binom{2n-k-1}{n}.$$

Theorem 3.3.3 *[15]The generating function for low occurrences of the strings d^2 and u^2 is* $L = \dfrac{C}{1 + (1-t)z^2C^3}$.

Proof

To derive the generating function $L(t,z) = L$ for the low occurrence of the string $\tau = d^2$. We define partition of $\mathcal{D}$ as $\{\Omega_i\}$, where Ω_i is set of all Dyck paths according to length of last descent equal to i, for all $i \geq 1$. We define $A_i(t,z) = A_i$ to be a generating function for Ω_i, where t counts the number of low occurrences the string $\tau = d^2$. All elements α of Ω_i can be written uniquely as $\alpha = \alpha_1 u \alpha_2 u \alpha_3 .... u \alpha_i u d^i$, where $\alpha_m \in \mathcal{D}$ for all $m \in [i]$. A low occurrence for $i \leq 1$ that is for $\alpha_1 ud$ produces zL, since α_1 is in the horizontal axis it produces L. The new low occurrence of d^2 appears in α (in addition to those contributed by a_m's) if and only if $i \geq 2$, since there are i (a_m's) and also there are i (u's), these produces $\sum_{i=2}^{\infty} tz^iC^{i-1}L$ and if $i \leq 1$ we get $zL = \sum_{i=1}^{1} z^iC^{i-1}L$. Combining these two possibilities taking into account of an empty path we get,

$$
\begin{aligned}
L &= 1 + \sum_{i=1}^{1} z^iC^{i-1}L + \sum_{i=2}^{\infty} tz^iC^{i-1}L \\
&= 1 + \sum_{i=1}^{1} z^iC^{i-1}L + \sum_{i=0}^{\infty} tz^{i+2}C^{i+2-1}L \\
&= 1 + \sum_{i=1}^{1} z^iC^{i-1}L + t\frac{z^2CL}{1-zC}
\end{aligned}
$$

From $\sum_{i=1}^{1} z^iC^{i-1}L$ we get the following

$$
\sum_{i=1}^{1} z^iC^{i-1}L = \sum_{i=1}^{\infty} z^iC^{i-1}L - \sum_{i=2}^{\infty} z^iC^{i-1}L = \frac{zL}{1-zC} - \frac{z^2CL}{1-zC}.
$$

Thus we have

$$
L = 1 + \frac{zL}{1-zC} - \frac{z^2C^1L}{1-zC} + t\frac{z^2CL}{1-zC}
$$

$$= 1 + \frac{zCL}{C - zC^2} - (1-t)\frac{z^2C^2L}{C - zC^2}.$$

We know from $C = 1 + zC^2$ that $C - zC^2 = 1$, then

$$\begin{aligned}
&= 1 + zCL - (1-t)z^2C^2L \\
&= \frac{1}{1 - zC + (1-t)z^2C^2} \\
&= \frac{C}{C - zC^2 + (1-t)z^2C^3} \\
&= \frac{C}{1 + (1-t)z^2C^3}.
\end{aligned}$$

Thus the generating function for the low occurrence of d^2 is,

$$L = \frac{C}{1 + (1-t)z^2C^3}.$$

Now we get $[z^nt^k]L = l_{n,k}$ as follows,

$$\begin{aligned}
L &= \frac{C}{1 + (1-t)z^2C^3} \\
&= \sum_{m=0}^{\infty} C^{3m+1}z^{2m}(t-1)^m \\
[z^n]L &= \sum_{m=0}^{\infty} [z^{n-2m}]C^{3m+1}(t-1)^m \\
&= \sum_{m=0}^{\infty} (t-1)^m \frac{3m+1}{2(n-2m)+3m+1}\binom{2(n-2m)+3m+1}{n-2m} \\
&= \sum_{m=0}^{\infty}\sum_{j=0}^{m} \binom{m}{j}(-1)^j t^{m-j}\frac{3m+1}{2n-m+1}\binom{2n-m+1}{n-2m} \\
&= \sum_{m=0}^{\infty}\sum_{j=0}^{m} (-1)^j t^{m-j}\binom{m}{j}\frac{3m+1}{2n-m+1}\frac{(2n-m+1)!}{(n-2m)!(n+m+1)!} \\
&= \sum_{m=0}^{\infty}\sum_{j=0}^{m} (-1)^j t^{m-j}\binom{m}{j}\frac{3m+1}{n+m+1}\frac{(2n-m)!}{(n-2m)!(n+m)!}
\end{aligned}$$

$$= \sum_{m=0}^{\infty} \sum_{j=0}^{m} (-1)^j t^{m-j} \binom{m}{j} \frac{3m+1}{n+m+1} \binom{2n-m}{n+m}.$$

Now let $k = m - j$ for the maximum value of m let $2n - m = n + m$ then $m = \frac{n}{2}$, from $k = m - j$ the maximum value of j is $j = \frac{n}{2} - k$ thus,

$$l_{n,k} = [t^k z^n]L = \sum_{j=0}^{[\frac{n}{2}]-k} (-1)^j \binom{j+k}{j} \frac{3(j+k)+1}{n+j+k+1} \binom{2n-(j+k)}{n+j+k}.$$

Theorem 3.3.4 *[15]The generating function for low occurrences of the string duu is* $L = 1 + \frac{zC^2}{1 + (1-x)z^2C^3}$.

We derive the generating function $L(x, z) := L$ for the low occurrence of the string *duu*. Let Ω be the set of all Dyck paths where the first ascent is of size at least 2, with its generating function $A(x, z) = A$, where x counts the number of low occurrence of the string $\tau = duu$. All elements α of Ω can be written uniquely as $\alpha = u\alpha_1 d\alpha_2$, where $\alpha_1, \alpha_2 \in \mathcal{D}$ and $\alpha_1 \neq \varepsilon$. The new low occurrence of *duu* occurs if and only if $\alpha_2 \in \Omega$, this produces $xz(C-1)A$ and if α_2 does not belong to Ω then we get $z(C-1)(L-A)$. Combining these two possibilities we get

$$A = z(C-1)(L - A + xA).$$

We know that every non-empty Dyck path $\alpha \in \mathcal{D}$ can be written uniquely as a first return decomposition, that is $\alpha = u\beta d\gamma$, where $\beta, \gamma \in \mathcal{D}$. We see that a new occurrence of the string *duu* (in addition to those contributed by γ) appears in α if and only if $\gamma \in \Omega$, producing $xzAC$, and if γ does not belong to Ω then we get $z(L - A)C$. Combining these two possibilities we now get

$$L = 1 + z(xA + L - A)C.$$

In the equation $L = 1 + z(xA + L - A)C$ we solve for A and get $A = \frac{L - 1 - zLC}{zxC - zC}$.

We then substitute the expression of A into
$A = z(C-1)(L - A + xA)$, then we proceed as follows:

$$\begin{aligned}
\frac{L-1-zLC}{ztC-zC} &= z(C-1)\left(L - \frac{L-1-zLC}{ztC-zC} + x\left(\frac{L-1-zLC}{ztC-zC}\right)\right) \\
L-1-zLC &= z(C-1)(zxCL - zCL - L + 1 + zLC + xL - x - zxLC) \\
L-1-zLC &= z(C-1)(-L+1+xL-x) \\
L - zLC + z(C-1)L - zx(C-1)L &= zC - z - xzC + xz \\
L &= \frac{1+zC-z-xzC+xz}{1-z-zx(C-1)} \\
&= \frac{zC+1-z-zx(C-1)}{1-z-zx(C-1)} \\
&= \frac{zC+1-z(C-zC^2)-zx(zC^2)}{1-z(C-zC^2)-zx(zC^2)} \\
&= 1 + \frac{zC}{1-z(C-zC^2)-zx(zC^2)} \\
&= 1 + \frac{zC^2}{C-z(C^2-zC^3)-zx(zC^3)} \\
&= 1 + \frac{zC^2}{C-zC^2+z^2C^3-xz^2C^3} \\
&= 1 + \frac{zC^2}{1+(1-x)z^2C^3}.
\end{aligned}$$

Thus the generating function for the low occurrence of duu is,

$$L = 1 + \frac{zC^2}{1+(1-x)z^2C^3}.$$

Now we get $[z^n x^k]L = l_{n,k}$ as follows,

$$\begin{aligned}
L &= 1 + \frac{zC^2}{1+(1-x)z^2C^3)} \\
&= 1 + zC^2 \sum_{m=0}^{\infty} C^{3m} z^{2m} (x-1)^m
\end{aligned}$$

$$
\begin{aligned}
[z^n]L &= \sum_{m=0}^{\infty}[z^{n-2m-1}]C^{3m+2}(x-1)^m \\
&= \sum_{m=0}^{\infty}(x-1)^m \frac{3m+2}{2(n-2m-1)+3m+2}\binom{2(n-2m-1)+3m+2}{n-2m-1} \\
&= \sum_{m=0}^{\infty}\sum_{j=0}^{m}\binom{m}{j}(-1)^j x^{m-j}\frac{3m+2}{2n-m}\binom{2n-m}{n+m+1}
\end{aligned}
$$

Now let $k = m - j$ for the maximum value of m let $2n - m - 1 = n + m$ then $m = \dfrac{n-1}{2}$, from $k = m - j$ the maximum value of j is $j = \dfrac{n-1}{2} - k$ thus,

$$
[x^k z^n]L = l_{n,k} = \sum_{j=0}^{[\frac{n-1}{2}]-k}(-1)^j\binom{j+k}{k}\frac{3j+3k+2}{2n-j-k}\binom{2n-j-k}{n+j+k+1}.
$$

Theorem 3.3.5 *[15]The generating function for low occurrences of the strings d^3 and u^3 is $L = \dfrac{C}{1+(1-t)z^3C^4}$.*

Proof

We derive the generating function $L(t,z) = L$ for the low occurrence of the string $\tau = d^3$, We partition $\mathcal{D}$ into $\{\Omega_i\}$, where Ω_i is partitioning of all Dyck paths according to length of last descent equal to i, for all $i \geq 1$. We define $A_i(t,z) = A_i$ to be a generating function for Ω_i, where t counts the number of low occurrences the string $\tau = d^3$. All elements α of Ω_i can be written uniquely as $\alpha = \alpha_1 u\alpha_2 u\alpha_3....u\alpha_i ud^i$, where $\alpha_m \in \mathcal{D}$ for all $m \in [i]$. A low occurrence for $i \leq 2$ that is for $\alpha_1 ud$ and $\alpha_1 u\alpha_2 ud^2$ produces zL and z^2CL respectively, since α_1 is in the horizontal axis it produces L. The same happens for $i \geq 3$, thus a new low occurrence of d^3 appears in α (in addition to those contributed by a_m's) if and only if $i \geq 3$, since there are i (a_m's) and also there are i (u's), these produces $\sum_{i=3}^{\infty} tz^iC^{i-1}L$ and if $i \leq 2$ we get $\sum_{i=1}^{2} z^iC^{i-1}L$. Combining these two possibilities taking into account an empty path we get,

$$
L = 1 + \sum_{i=1}^{2} z^iC^{i-1}L + \sum_{i=3}^{\infty} tz^iC^{i-1}L
$$

$$= 1 + \sum_{i=1}^{r-1} z^i C^{i-1} L + \sum_{i=0}^{\infty} t z^{i+3} C^{i+2} L$$
$$= 1 + \sum_{i=1}^{2} z^i C^{i-1} L + t \frac{z^3 C^3 L}{1 - zC}.$$

From $\sum_{i=1}^{2} z^i C^{i-1} L$ we get the following

$$\sum_{i=1}^{2} z^i C^{i-1} L = \sum_{i=1}^{\infty} z^i C^{i-1} L - \sum_{i=3}^{\infty} z^i C^{i-1} L = \frac{zL}{1 - zC} - \frac{z^3 C^2 L}{1 - zC}.$$

Thus we have

$$L = 1 + \frac{zL}{1 - zC} - \frac{z^3 C^2 L}{1 - zC} + t \frac{z^3 C^2 L}{1 - zC}$$
$$= 1 + \frac{zCL}{C - zC^2} - (1 - t) \frac{z^3 C^3 L}{C - zC^2}.$$

We know from $C = 1 + zC^2$ then $C - zC^2 = 1$ thus

$$= 1 + zCL - (1 - t) z^3 C^3 L$$
$$= \frac{1}{1 - zC + (1 - t) z^3 C^3}$$
$$= \frac{C}{C - zC^2 + (1 - t) z^3 C^4}$$
$$= \frac{C}{1 + (1 - t) z^3 C^4}.$$

Thus the generating function for the low occurrence of ddd is,

$$L = \frac{C}{1 + (1 - t) z^3 C^4}.$$

Now we get $[z^n t^k] L = l_{n,k}$ as follows,

$$L = \frac{C}{1 + (1 - t) z^3 C^4}$$

$$
\begin{aligned}
&= \sum_{m=0}^{\infty} C^{4m+1} z^{3m} (t-1)^m \\
[z^n]L &= \sum_{m=0}^{\infty} [z^{n-3m}] C^{4m+1} (t-1)^m \\
&= \sum_{m=0}^{\infty} (t-1)^m \frac{4m+1}{2(n-3m)+4m+1} \binom{2(n-3m)+4m+1}{n-3m} \\
&= \sum_{m=0}^{\infty} \sum_{j=0}^{m} \binom{m}{j} (-1)^j t^{m-j} \frac{(4m+1}{2n-2m+1} \binom{2n-2m+1}{n-3m} \\
&= \sum_{m=0}^{\infty} \sum_{j=0}^{m} (-1)^j t^{m-j} \binom{m}{j} \frac{4m+1}{2n-2m+1} \frac{(2n-2m+1)!}{(n-3m)!(n+m+1)!} \\
&= \sum_{m=0}^{\infty} \sum_{j=0}^{m} (-1)^j t^{m-j} \binom{m}{j} \frac{4m+1}{n+m+1} \frac{(2n-2m)!}{(n-2m)!(n+m)!} \\
&= \sum_{m=0}^{\infty} \sum_{j=0}^{m} (-1)^j t^{m-j} \binom{m}{j} \frac{4m+1}{n+m+1} \binom{2n-2m}{n+m}.
\end{aligned}
$$

Now let $k = m - j$ for the maximum value of m let $2n - 2m = n + m$ then $m = \frac{n}{3}$, from $k = m - j$ the maximum value of j is $j = \frac{n}{3} - k$ thus,

$$
l_{n,k} = [t^k z^n]L = \sum_{j=0}^{[\frac{n}{3}]-k} (-1)^j \binom{j+k}{j} \frac{4(j+k)+1}{n+j+k+1} \binom{2n-2(j+k)}{n+j+k}.
$$

Theorem 3.3.6 *[15]The generating function for low occurrences of the string uudu is* $L = \dfrac{C}{1+(1-t)z^3C^3}$.

Proof

We derive the string $\tau = uudu$. We derive its generating function by using the first return decomposition of non-empty Dyck path $\alpha = u\beta d\gamma$ where $\beta, \gamma \in \mathcal{D}$, a new low occurrence of *uudu* appears in α (in addition to those contributed by β and γ) if and only if $\beta = ud\phi$, where $\phi \in \mathcal{D}$ and $\phi \neq \varepsilon$.

The path α can start with the string *uudu* right from the start where $\beta = ud\phi$ producing $z(tz(C-1))L$ or α can start without the string *uudu* in the beginning producing $z(C - z(C-1))L$.

Now we get the generating function $L(t,z) = L$ (where t counts the number of low occurrences of $uudu$'s) as follows,

$$
\begin{aligned}
L - 1 &= z(tz(C-1))L + z(C - z(C-1))L \\
1 &= L - tz^2(C-1))L - z(C - z(C-1))L \\
L &= \frac{1}{1 - tz^2(C-1)) - z(C - z(C-1))}.
\end{aligned}
$$

We now use the Catalan number generating function C where $C = 1 + zC^2$ to produce

$$
\begin{aligned}
L &= \frac{1}{1 - tz^3C^2 - (z + z^2C^2 - z^3C^2)} \\
&= \frac{C}{C + (1-t)z^3C^3 - zC - z^2C^3)} \\
&= \frac{C}{C + (1-t)z^3C^3 - zC(1 + zC^2)} \\
&= \frac{C}{C + (1-t)z^3C^3 - zC^2} \\
&= \frac{C}{1 + (1-t)z^3C^3}.
\end{aligned}
$$

Thus the generating function for the low occurrence of $uudu$ is,

$$
L = \frac{C}{1 + (1-t)z^3C^3}.
$$

Now we get $[z^n t^k]L = l_{n,k}$ as follows,

$$
\begin{aligned}
L &= \frac{C}{1 + (1-t)z^3C^3} \\
&= \sum_{m=0}^{\infty} C^{3m+1} z^{3m} (t-1)^m
\end{aligned}
$$

$$
\begin{aligned}
[z^n]L &= \sum_{m=0}^{\infty}[z^{n-3m}]C^{3m+1}(t-1)^m \\
&= \sum_{m=0}^{\infty}(t-1)^m \frac{3m+1}{2(n-3m)+3m+1}\binom{2(n-3m)+3m+1}{n-3m} \\
&= \sum_{m=0}^{\infty}\sum_{j=0}^{m}\binom{m}{j}(-1)^j t^{m-j}\frac{3m+1}{2n-3m+1}\binom{2n-3m+1}{n-3m} \\
&= \sum_{m=0}^{\infty}\sum_{j=0}^{m}(-1)^j t^{m-j}\binom{m}{j}\frac{3m+1}{2n-3m+1}\frac{(2n-3m+1)!}{(n-3m)!(n+1)!} \\
&= \sum_{m=0}^{\infty}\sum_{j=0}^{m}(-1)^j t^{m-j}\binom{m}{j}\frac{3m+1}{n+1}\frac{(2n-3m)!}{(n-3m)!(n)!} \\
&= \sum_{m=0}^{\infty}\sum_{j=0}^{m}(-1)^j t^{m-j}\binom{m}{j}\frac{3m+1}{n+1}\binom{2n-3m}{n}.
\end{aligned}
$$

Now let $k = m - j$ for the maximum value of m let $2n - 3m = n$ then $m = \dfrac{n}{3}$, from $k = m - j$ the maximum value of j is $j = \dfrac{n}{3} - k$ thus,

$$
[t^k z^n]L = l_{n,k} = \frac{1}{n+1}\sum_{j=0}^{[\frac{n}{3}]-k}(-1)^j\binom{j+k}{k}(3j+3k+1)\binom{2n-3k-3j}{n}.
$$

Theorem 3.3.7 *[15]The generating function for low occurrences of the string uduu is* $L - 1 = z(tz(C-1))L + z(C - z(C-1))L$*.*

Proof

We derive its generating function for $\tau = uduu$ by using the first return decomposition of non-empty Dyck path $\alpha = \beta u\gamma d$ where $\beta, \gamma \in \mathcal{D}$. A new low occurrence of *uduu* appears in α (in addition to those contributed by β and γ) if and only if $\beta = \delta ud$ and $\gamma \neq \varepsilon$. Where $\delta \in \mathcal{D}$ producing $z(tz(C-1))L$ or α can be such that $\beta \neq \delta ud$ producing $z(C - z(C-1))L$. Combining these two possibilities we get

$$
L - 1 = z(tz(C-1))L + z(C - z(C-1))L. \tag{3.6}
$$

This generating function for low occurrences of *uduu* is the same as the generating function for low occurrences of *uudu*, thus the number of low occurrences for *uduu* is equidistant to the number of low occurrences for *uudu*, thus as for *uudu*;

$$[t^k z^n]L = l_{n,k} = \frac{1}{n+1}\sum_{j=0}^{[\frac{n}{3}]-k}(-1)^j\binom{j+k}{k}(3j+3k+1)\binom{2n-3k-3j}{n}.$$

Theorem 3.3.8 *[15]The generating function for low occurrences of the strings uuud is* $L = \dfrac{C}{1+(1-t)z^3C^3}$.

Proof

We study the low occurrence of the string $\tau = uuud$. We derive the generating function for the low occurrence of $\tau = uuud$ using the first return decomposition of non-empty Dyck path $\alpha = u\beta d\gamma$ where $\beta, \gamma \in \mathcal{D}$. A new low occurrence of $uuud$ appears in α (in addition to those contributed by β and γ) if and only if $\beta = u\varphi d\delta$ with $\varphi = ud\phi$. Where $\varphi, \delta, \phi \in \mathcal{D}$ producing $z(z(tzC)C)L$ or α can be such that $\varphi \neq ud\phi$ producing $z(C - z^2C^2)L$.

Now we get the generating function $L(t, z) = L$ (where t counts the number of low occurrences of $uuud$) as follows,

$$\begin{aligned} L - 1 &= z(z(tzC)C + C - z^2C^2)L \\ L - 1 &= z^3tC^2L + zCL - z^3C^2L \\ L &= \frac{1}{1 - z^3tC^2 - zC + z^3C^2} \\ &= \frac{C}{C - z^3tC^3 - zC^2 + z^3C^3} \end{aligned}$$

Since $C - zC^2 = 1$, then

$$= \frac{C}{1+(1-t)z^3C^3}. \tag{3.7}$$

This generating function for low occurrences of $uuud$ is the same as the generating functions for low occurrences of $uudu$ and $uduu$. Thus the number of low occurrences for $uduu$ is equidistant to the number of low occurrences for $uudu$ and $uduu$, thus as for $uudu$,

$$[t^k z^n]L = l_{n,k} = \frac{1}{n+1} \sum_{j=0}^{[\frac{n}{3}]-k} (-1)^j \binom{j+k}{k} (3j+3k+1) \binom{2n-3k-3j}{n}.$$

Theorem 3.3.9 *[15]The generating function for low occurrences of the strings uudd is* $L = \dfrac{C}{1+(1-t)z^2C}$.

Proof

We study the generating function for low occurrence of $\tau = uudd$ by using the first return decomposition of non-empty Dyck path $\alpha = u\beta d\gamma$ where $\beta, \gamma \in \mathcal{D}$. A new low occurrence of $uudd$ appears in α (in addition to those contributed by β and γ) if and only if $\beta = ud$ producing $z(tz)L$ or α can be such that $\beta \neq ud$ producing $z(C-z)L$.

Now we get the generating function $L(t,z) = L$ (where by t counts the number of low occurrences of $uudd$) as follows.

$$\begin{aligned} L - 1 &= z(tz + C - z)L \\ 1 &= L - z^2tL - zCL + z^2L \\ L &= \frac{1}{1 - z^2t - zC + z^2} \\ &= \frac{C}{C - z^2tC - zC^2 + z^2C} \\ &= \frac{C}{1+(1-t)z^2C}. \end{aligned} \tag{3.8}$$

The generating function for the low occurrence of $uudd$ is,

$$L = \frac{C}{1+(1-t)z^2C}.$$

Now we get $[z^n t^k]L = l_{n,k}$ as follows,

$$\begin{aligned} L &= \frac{C}{1+(1-t)z^2C} \\ &= C \sum_{m=0}^{\infty} C^m z^{2m} (t-1)^m \end{aligned}$$

$$
\begin{aligned}
[z^n]L &= \sum_{m=0}^{\infty}[z^{n-2m}]C^{m+1}(t-1)^m \\
&= \sum_{m=0}^{\infty}[z^{n-2m}]C^{m+1}(t-1)^m \\
&= \sum_{m=0}^{\infty}(t-1)^m \frac{m+1}{2(n-2m)+m+1}\binom{2(n-2m)+m+1}{n-2m} \\
&= \sum_{m=0}^{\infty}\sum_{j=0}^{m}\binom{m}{j}(-1)^j t^{m-j}\frac{m+1}{2n-3m+1}\binom{2n-3m+1}{n-2m} \\
&= \sum_{m=0}^{\infty}\sum_{j=0}^{m}(-1)^j t^{m-j}\binom{m}{j}\frac{m+1}{2n-3m+1}\frac{(2n-3m+1)!}{(n-2m)!(n-m+1)!} \\
&= \sum_{m=0}^{\infty}\sum_{j=0}^{m}(-1)^j t^{m-j}\binom{m}{j}\frac{m+1}{n-m+1}\frac{(2n-3m)!}{(n-2m)!(n-m)!} \\
&= \sum_{m=0}^{\infty}\sum_{j=0}^{m}(-1)^j t^{m-j}\binom{m}{j}\frac{m+1}{n-m+1}\binom{2n-3m}{n-m}.
\end{aligned}
$$

Now let $k = m - j$ for the maximum value of m let $2n - 3m = n - m$ then $m = \frac{n}{2}$, from $k = m - j$ the maximum value of j is $j = \frac{n}{2} - k$ thus,

$$
[t^k z^n]L = l_{n,k} = \sum_{j=0}^{[\frac{n}{2}]-k}(-1)^j\binom{j+k}{k}\frac{j+k+1}{n-k-j+1}\binom{2n-3k-3j}{n-k-j}.
$$

Theorem 3.3.10 *[15]The generating function for low occurrences of the strings dudu is* $L = 1 + zC + \dfrac{z^2C^3}{1+(1-t)zC}$.

Proof

To derive the generating function $L(t, z) = L$ for the low occurrence of the string $\tau = dudu$. We let Ω to be the set of all Dyck paths with semi-length at least 2, which starts with the low peak, with its generating function $A(t, z) = A$. Where t counts the number of low occurrences of the string $\tau = dudu$. All elements α of Ω can be written uniquely as $\alpha = ud\beta$. Where $\beta \in \mathcal{D}\backslash\{\varepsilon\}$. The new low occurrence of *dudu* occurs if and only if $\beta \in \Omega$. This produces ztA and if β does not belong to Ω then we get $z(L - 1 - A)$. Combining these two possibilities we get

$$A = ztA + z(L - 1 - A). \tag{3.9}$$

In addition we form the first return decomposition $\alpha = u\beta d\gamma$, where $\beta, \gamma \in \mathcal{D}$, the new low occurrence of $dudu$ is possible if and only if $\gamma \in \Omega$ this produces $ztCA$. if γ does not belong to Ω, then we have $zC(L-A)$. Combining these two possibilities we get

$$L = 1 + ztCA + z(L - A)C. \tag{3.10}$$

From the above two equations we eliminate A and get the generating function L as follows.

In the equation $A = ztA + z(L - 1 - A)$ we solve for A and get $A = \dfrac{zL - z}{1 - tz + z}$.

We then substitute the expression of A into $L = 1 + ztCA + z(L - A)C$ and proceed as follows:

$$\begin{aligned}
L &= 1 + ztC(\frac{zL - z}{1 - tz + z}) + z(L - (\frac{zL - z}{1 - tz + z}))C \\
L - tzL + zL &= 1 - tz + z + z^2tCL - z^2tC + zCL - tz^2LC + z^2LC - z^2CL + z^2C \\
L &= \frac{1 - tz + z - z^2tC + z^2C}{1 - tz + z - zC} \\
&= \frac{C - tzC + zC - z^2tC^2 + z^2C^2}{C - tzC + zC - zC^2} \\
&= \frac{C - tzC + zC - z^2tC^2 + z^2C^2}{1 + (1 - t)zC}.
\end{aligned}$$

We now multiply $C = 1 + zC^2$ by zC and get $zC^2 = zC + z^2C^3$ so $zC^2 - zC = z^2C^3$, then we get

$$\begin{aligned}
&= \frac{1 + zC^2 - zC + zC - tzC + zC - z^2tC^2 + z^2C^2}{1 + (1 - t)zC} \\
&= \frac{1 + z^2C^3 + zC - tzC + zC - z^2tC^2 + z^2C^2}{1 + (1 - t)zC}
\end{aligned}$$

$$
\begin{aligned}
&= \frac{1+(1-t)zC+zC(1+(1-t)zC)+z^2C^3}{1+(1-t)zC} \\
&= 1+zC+\frac{z^2C^3}{1+(1-t)zC}.
\end{aligned}
\tag{3.11}
$$

The generating function for the low occurrence of *dudu* is,

$$L = 1+zC+\frac{z^2C^3}{1+(1-t)zC}.$$

Now we get $[z^nt^k]L = l_{n,k}$ as follows,

$$
\begin{aligned}
[z^n]L &= [z^{n-1}]C+\sum_{m=0}^{\infty}[z^{n-m-2}]C^{m+3}(t-1)^m \\
&= \frac{1}{n}\binom{2(n-1)}{n-1}+\sum_{m=0}^{\infty}(t-1)^m\frac{m+3}{2(n-m-2)+m+3}\binom{2(n-m-2)+m+3}{n-m-2} \\
&= C_{n-1}+\sum_{m=0}^{\infty}\sum_{j=0}^{m}\binom{m}{j}(-1)^jt^{m-j}\frac{m+3}{2n-m-1}\binom{2n-m-1}{n-m-2} \\
&= C_{n-1}+\sum_{m=0}^{\infty}\sum_{j=0}^{m}(-1)^jt^{m-j}\binom{m}{j}\frac{m+3}{2n-m-1}\frac{(2n-m-1)!}{(n-m-2)!(n+1)!} \\
&= C_{n-1}+\sum_{m=0}^{\infty}\sum_{j=0}^{m}(-1)^jt^{m-j}\binom{m}{j}\frac{m+3}{n+1}\frac{(2n-m-2)!}{(n-m-2)!n!} \\
&= C_{n-1}+\sum_{m=0}^{\infty}\sum_{j=0}^{m}(-1)^jt^{m-j}\binom{m}{j}\frac{m+3}{n+1}\binom{2n-m-2}{n}
\end{aligned}
$$

Now let $k=m-j$ for the maximum value of m let $2n-m-2=n$ then $m=n-2$, from $k=m-j$ the maximum value of j is $j=n-2-k$ thus,

$$[t^kz^n]L = \delta_{0k}C_{n-1}+\sum_{j=0}^{n-2-k}(-1)^j\binom{j+k}{j}\frac{j+k+3}{n+1}\binom{2n-j-k-2}{n},$$

where δ_{0k} is the Kronecker function, that is

$$\delta_{nk} = \begin{cases} 1 & \text{if } n = k \\ 0 & \text{if } n \neq k. \end{cases} \tag{3.12}$$

Theorem 3.3.11 *[15]The generating function for low occurrences of the strings d^r and u^r, where $r \geq 2$ is $L = \dfrac{C}{1 + (1-t)z^r C^{r+1}}$.*

Proof

We derive the generating function $L(t,z) = L$ for the low occurrence of the string $\tau = d^r$. We partition $\mathcal{D}$ into $\{\Omega_i\}$. Where Ω_i is set of all Dyck paths according to length of last descent equal to i, for all $i \geq 1$. We define $A_i(t,z) = A_i$ to be a generating function for Ω_i, where t counts the number of low occurrences the string $\tau = d^r$. All elements α of Ω_i can be written uniquely as $\alpha = \alpha_1 u \alpha_2 u \alpha_3 \ldots . u \alpha_i u d^i$, where $\alpha_m \in \mathcal{D}$ for all $m \in [i]$. A low occurrence for $i \leq 2$ that is for $\alpha_1 ud$ and $\alpha_1 u \alpha_2 u d^2$ produces zL and $z^2 CL$ respectively, since α_1 is in the horizontal axis it produces L. The same happens for $i \geq r$. Thus a new low occurrence of d^r appears in α (in addition to those contributed by a_m's) if and only if $i \geq r$, since there are i (a_m's) and also there are i (u's). This produces $\sum_{i=r}^{\infty} t z^i C^{i-1} L$ and if $i \leq r-1$ we get $\sum_{i=1}^{r-1} z^i C^{i-1} L$. Combining these two possibilities taking into account an empty path we get,

$$\begin{aligned} L &= 1 + \sum_{i=1}^{r-1} z^i C^{i-1} L + \sum_{i=r}^{\infty} t z^i C^{i-1} L \\ &= 1 + \sum_{i=1}^{r-1} z^i C^{i-1} L + \sum_{i=0}^{\infty} t z^{i+r} C^{i+r-1} L \\ &= 1 + \sum_{i=1}^{r-1} z^i C^{i-1} L + t \frac{z^r C^{r-1} L}{1 - zC} \end{aligned}$$

From $\sum_{i=1}^{r-1} z^i C^{i-1} L$ we get the following

$$\sum_{i=1}^{r-1} z^i C^{i-1} L = \sum_{i=1}^{\infty} z^i C^{i-1} L - \sum_{i=r}^{\infty} z^i C^{i-1} L = \frac{zL}{1 - zC} - \frac{z^r C^{r-1} L}{1 - zC}.$$

Thus we have

$$\begin{aligned} L &= 1 + \frac{zL}{1-zC} - \frac{z^r C^{r-1} L}{1-zC} + t\frac{z^r C^{r-1} L}{1-zC} \\ &= 1 + \frac{zCL}{C-zC^2} - (1-t)\frac{z^r C^r L}{C-zC^2}. \end{aligned}$$

We know from $C = 1 + zC^2$ that $C - zC^2 = 1$ thus

$$\begin{aligned} &= 1 + zCL - (1-t)z^r C^r L \\ &= \frac{1}{1 - zC + (1-t)z^r C^r} \\ &= \frac{C}{C - zC^2 + (1-t)z^r C^{r+1}} \\ &= \frac{C}{1 + (1-t)z^r C^{r+1}}. \end{aligned}$$

Thus the generating function for the low occurrence of d^r is,

$$L = \frac{C}{1 + (1-t)z^r C^{r+1}}.$$

Now we get $[z^n t^k]L = l_{n,k}$ as follows,

$$\begin{aligned} L &= \frac{C}{1 + (1-t)z^r C^{r+1}} \\ &= \sum_{m=0}^{\infty} C^{(r+1)m+1} z^{rm} (t-1)^m \\ [z^n]L &= \sum_{m=0}^{\infty} [z^{n-rm}] C^{(r+1)m+1} (t-1)^m \\ &= \sum_{m=0}^{\infty} (t-1)^m \frac{(r+1)m+1}{2(n-rm)+(r+1)m+1} \binom{2(n-rm)+(r+1)m+1}{n-rm} \\ &= \sum_{m=0}^{\infty} \sum_{j=0}^{m} \binom{m}{j} (-1)^j t^{m-j} \frac{(r+1)m+1}{2n-(r-1)m+1} \binom{2n-(r-1)m+1}{n-rm} \\ &= \sum_{m=0}^{\infty} \sum_{j=0}^{m} (-1)^j t^{m-j} \binom{m}{j} \frac{(r+1)m+1}{2n-(r-1)m+1} \frac{(2n-(r-1)m+1)!}{(n-rm)!(n+m+1)!} \end{aligned}$$

$$= \sum_{m=0}^{\infty} \sum_{j=0}^{m} (-1)^j t^{m-j} \binom{m}{j} \frac{(r+1)m+1}{n+m+1} \frac{(2n-(r-1)m)!}{(n-rm)!(n+m)!}$$
$$= \sum_{m=0}^{\infty} \sum_{j=0}^{m} (-1)^j t^{m-j} \binom{m}{j} \frac{(r+1)m+1}{n+m+1} \binom{2n-(r-1)m}{n+m}.$$

Now let $k = m - j$ for the maximum value of m let $2n - (r-1)m = n + m$ then $m = \dfrac{n}{r}$, from $k = m - j$ the maximum value of j is $j = \dfrac{n}{r} - k$ thus,

$$l_{n,k} = [t^k z^n]L = \sum_{j=0}^{[\frac{n}{r}]-k} (-1)^j \binom{j+k}{j} \frac{(r+1)(j+k)+1}{n+j+k+1} \binom{2n-(r-1)(j+k)}{n+j+k}.$$

3.4 Counting strings at high level

In this section we study the paper titled Counting strings in Dyck paths by [16]. Here we take into account the number of high occurrences of the string τ.

We say that a string τ is at high level if its minimum point(s) occurs above the horizontal axis. For example the two strings udu in Figure 3.1 (page 48) occur at high level. We define the generating function for the high occurrences of the string τ as follows

$$H(t,z) = \sum_{n=0}^{\infty} \sum_{k=0}^{n} h_{n,k} t^k z^n. \tag{3.13}$$

We use the first return decomposition of non-empty Dyck path $\alpha = u\beta d\gamma$ to derive the generating function $H(t,z) := H$ (where t counts the number of high occurrences the string τ) for high occurrences of the string τ. In the decomposition $\alpha = u\beta d\gamma$, β is not at the horizontal axis therefore it can generate the high string τ producing $H(t,z) := H$. The path γ is at the horizontal axis, thus it can generate both high and low τ producing $F(t,z)$ (defined in Chapter two as a generating function for τ). Combining these cases we obtain the generating function for the high occurrences of τ as follows,

$$H = 1 + zHF$$
$$H = \frac{1}{1 - zF}.$$

Now we consider the three strings τ, τ_1 and τ_2, from Section 3.1 we have $E_\tau(t, z) = F_{\tau_1}(t, z)$ and $O_\tau(t, z) = F_{\tau_2}(t, z)$
from $H = \dfrac{1}{1 - zF}$ we get

$$H_{\tau_1}(t, z) = \frac{1}{1 - zF_{\tau_1}(t, z)},$$

and from $O_\tau(t, z) = \dfrac{1}{1 - zE_\tau(t, z)}$ we have

$$H_{\tau_1}(t, z) = O_\tau(t, z) = F_{\tau_2}(t, z) = \frac{1}{1 - zF_{\tau_1}(t, z)},$$

thus we have

$$H_{\tau_1}(t, z) = F_{\tau_2}(t, z).$$

Now we find the generating functions at high level for the strings *uud*, *udu*, *uduu*, *uudd* and *udud* using the above results.

Theorem 3.4.1 *[16]The generating function for high occurrences of the string uud is* $tzH^2(t, z) - (1 + 2(t - 1)z)H(t, z) + 1 + (t - 1)z = 0$.

Proof

We derive the generating function $H(t, z)$ for the string *uud* by cosidering the relation $H_{\tau_1}(t, z) = O_\tau(t, z) = F_{\tau_2}(t, z)$. From Sectiom 3.1 $E_{dd}(t, z) = F_{uud}(t, z)$ and $O_{dd}(t, z) = F_{duu}(t, z)$, therefore $H_{uud}(t, z) = F_{duu}(t, z)$. Since the generating function $F_{duu}(t, z)$ for *duu* from Chapter two is $tzF^2(t, z) - (1 + 2(t - 1)z)F(t, z) + 1 + (t - 1)z = 0$, then generating function $H_{uud}(t, z)$ for *uud* is

$$tzH^2(t, z) - (1 + 2(t - 1)z)H(t, z) + 1 + (t - 1)z = 0$$

Theorem 3.4.2 *[16]The generating functions for high occurrences of the strings udu, uduu, uudd and udud are as follows.*

We know that from Section 3.1 $E_{du}(t,z) = F_{udu}(t,z)$ and $O_{du}(t,z) = F_{ddd}(t,z)$, then $H_{udu}(t,z) = F_{ddd}(t,z)$. We proceed as for the string *uud* above. Thus the generating function $H_{udu}(t,z)$ for *udu* is

$$z(t+z-tz)H^2 - (1-z+tz)H + 1 = 0.$$

If we do the same for the strings *uduu*, *uudd* and *udud* we obtain that,
(1) $E_{duu}(t,z) = F_{uduu}(t,z)$ and $O_{duu}(t,z) = F_{dduu}(t,z)$, thus
$H_{uduu}(t,z) = F_{dduu}(t,z)$, then the generating function $H_{uduu}(t,z)$ for *uduu* is

$$z(t+(1-t)z)H^2 - (1+(1-t)(z-2)z)H + (t-1)z + 1 = 0,$$

(2) $E_{uud}(t,z) = F_{uudd}(t,z)$ and $O_{uud}(t,z) = F_{uduu}(t,z)$, thus
$H_{uudd}(t,z) = F_{uduu}(t,z)$, then the generating function $H_{uudd}(t,z)$ for *uudd* is

$$z(1-(1-t)z)H^2 + ((1-t)z^2 - 1)H + 1 = 0,$$

(3) $E_{udu}(t,z) = F_{udud}(t,z)$ and $O_{uud}(t,z) = F_{dudu}(t,z)$, thus
$H_{udud}(t,z) = F_{dudu}(t,z)$, then the generating function $H_{udud}(t,z)$ for *udud* is

$$zH^2 + ((1-t)(z-1)z - 1)H + (1-t)z + 1 = 0,$$

respectively.

Chapter 4

Skew Dyck paths, and superdiagonal bargraphs

In this chapter we study the paper titled Skew Dyck paths, area, and superdiagonal bargraphs by [7]. We enumerate skew Dyck paths according to different parameters.

4.1 Enumeration of skew Dyck paths according to different parameters

In this section we enumerate skew paths according to area. We define the area for skew paths, but first we define the area for Dyck paths. The region between a path and the x-axis can be decomposed into right triangles with unit areas, we can conclude that the area of Dyck paths is the number of these unit triangles. In a similar way we can define the area of a skew paths as the region below the path and above the x-axis (see Figure 4.1). Here we are interested in an enumeration of skew paths according to area, semi-length (half of the skew path length) and semi-base. We show bijectively that the number of skew paths of area n is the Fibonacci number F_n. This provides a striking property about the occurrence of the Fibonacci numbers.

4.1.1 Enumeration of skew paths according to area and semi-base

Here we compute the number a_{nk} of all skew paths having area n and semi-base k, then we determine the number a_n of all skew paths with area n. In order for us

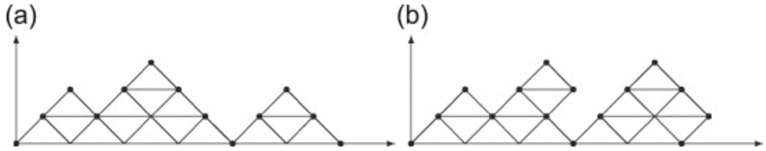

Figure 4.1: (a) A Dyck path of area 16 and (b) a skew path of area 17

to be able to obtain these numbers by a bijective argument, we first describe the following decomposition of skew paths according to the leftmost peak. The class of all skew paths can be split into a class of all skew paths with a low peak (i.e skew paths starting with ud) (see Figure 4.2(a)) and all skew paths with left most peak at height at least two, like $uu\gamma$, where γ is a skew path (see Figure 4.2(b)).

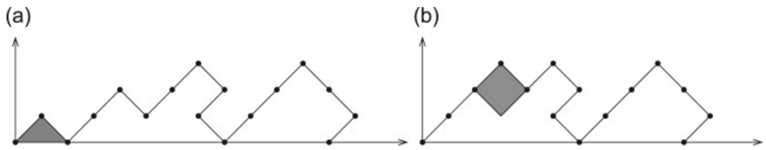

Figure 4.2: (a) A skew path beginning with a low peak and (b) a skew path whose left most peak is at least 2

If the low peak is deleted, the skew path in the first class turns out to be an arbitrary skew path (where the semi-base and area are decreased by 1). If we remove the square determined by the left most peak, the skew path in the second class turns out to be the arbitrary skew path (here the semi-base does not change but the area is decreased by 2).

We define $a(q;y)$ to be the generating function for the class of all skew paths with respect to their area, marked by q and to semi-base marked by y. The generating function $a(q;y)$ can be empty producing 1 or it can be of the class in Figure 4.2(a) producing $qya(q;y)$ or Figure 4.2(b) producing $q^2(a(q;y)-1)$. Combining the three cases we get

$$\begin{aligned}
a(q;y) &= 1 + qya(q;y) + q^2(a(q;y)-1) \\
a(q;y) - qya(q;y) - q^2a(q;y) &= 1 - q^2 \\
a(q;y) &= \frac{1-q^2}{1-qy-q^2}.
\end{aligned} \tag{4.1}$$

Therefore

$$a(q;y) = \sum_{k\geq 0}\sum_{n\geq 0} a_{nk}q^n y^k = \frac{1-q^2}{1-qy-q^2}. \tag{4.2}$$

Also if

$$a_k(q) = \sum_{n\geq 0} a_{nk}q^n, \tag{4.3}$$

then from Figure 4.2(a) $a_{k+1}(q) = qa_k(q)$ or from Figure 4.2(b) $a_{k+1}(q) = q^2a_{k+1}(q)$. Combining these two cases we obtain

$$\begin{aligned} a_{k+1}(q) &= qa_k(q) + q^2a_{k+1}(q) \\ a_{k+1}(q) &= \frac{q}{1-q^2}a_k(q). \end{aligned} \tag{4.4}$$

We know that there is only one skew path with semi-base zero that is the empty path, then

$$a_0(q) = 1. \tag{4.5}$$

It follows from equation (4.4) that

$$\begin{aligned} a_k(q) &= \frac{q}{1-q^2}a_{k-1}(q) \\ a_k(q) &= \frac{q^2}{(1-q^2)^2}a_{k-2}(q) \\ a_k(q) &= \frac{q^k}{(1-q^2)^k}a_0(q) \\ a_k(q) &= \frac{q^k}{(1-q^2)^k}, \end{aligned} \tag{4.6}$$

for every $k \in N$. If we do a series expansion in (4.6) we get the number a_{nk} for all skew paths with semi-base k and area n.

We know that there is one each skew path of area 0, 1, or 2 i.e $a_0 = a_1 = a_2 = 1$ respectively. These initial conditions give us the recurrence $a_{n+3} = a_{n+2} + a_{n+1}$.

Thus we obtain $a_n = F_n$, for every $n \geq 1$, where F_n are the well known Fibonacci numbers (defined such that $F_n = F_{n-1} + F_{n-2}$ for $F_0 = 0$ and $F_1 = 1$).

Theorem 4.1. *[6]The number of all skew paths having semi-base k and area n is*

$$a_{nk} = \begin{cases} \binom{\frac{n-k}{2}+k-1}{k-1} & \text{if } n \equiv k(mod\ 2), \\ 0 & \text{otherwise.} \end{cases} \tag{4.7}$$

Proof

$$\begin{aligned} a_{nk} &= [q^n]a_k(q) \\ [q^n]a_k(q) &= [q^n]q^k(1-q^2)^{-k} \\ &= [q^n]\sum_{i\geq 0}\binom{-k}{i}(-1)^i q^{2i+k} \\ &= [q^n]\sum_{i\geq 0}\frac{(-k)(-k-1)...(-k-i+1)}{i!}(-1)^i q^{2i+k} \\ &= [q^n]\sum_{i\geq 0}\frac{(k)(k+1)...(k+i-1)}{i!}(-1)^{2i} q^{2i+k} \\ &= [q^n]\sum_{i\geq 0}\frac{(k-1)!(k)(k+1)...(k+i-1)}{i!(k-1)!}(-1)^{2i} q^{2i+k} \\ &= [q^n]\sum_{i\geq 0}\binom{k+i-1}{k-1} q^{2i+k}. \\ &\text{Let } n = 2i+k \\ a_{nk} &= \binom{\frac{n-k}{2}+k-1}{k-1}. \end{aligned} \tag{4.8}$$

The number of skew paths with area n i.e $a_n = \sum_{k=1}^{n} a_{nk}$ is equal to the Fibonacci number $F_n (n \geq 1)$, therefore $a_n = \sum_{k=1}^{n} \binom{\frac{n-k}{2}+k-1}{k-1} = F_n$.

It is easy to show that $a_3 = 2$, $a_4 = 3$, $a_5 = 5$, $a_6 = 8$ and $a_7 = 13$, this means that there are 2 skew paths with area 3, 3 with area 4, 5 with area 5 (this is confirmed by Figure 4.3) etc.

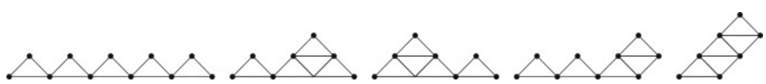

Figure 4.3: The 5 skew Dyck paths with area 5

4.1.2 Enumeration of skew paths according to area and semi-length

Let $f_n(q)$ be generating function for all skew paths with semi-length n, where q marks the area of the skew path.

We derive $f_n(q)$ by first obtaining the generating function $g_n(q, y)$, of all skew paths with semi-length n , where y marks the semi-base and q marks the area. This gives $f_n(q) = g_n(q, 1)$. Let $H(x, y, q)$ be the generating function of all skew paths, where x marks the semi-length, q marks the area and y marks the semi-base. From Figure 1.4 (from Chapter one) we get

$$H(q, y, x) = 1 + qyxH(q, q^2y, x)H(q, y, x) + x(H(q, q^2y, x) - 1). \tag{4.9}$$

Let

$$H(q, y, x) = \sum_{n\geq 0} g_n(q, y)x^n. \tag{4.10}$$

We know that there is only one skew path with area zero, then $g_0(q, y) = 1$ and also there is only one skew path of area 1 with both semi-length and semi-base equal to 1, i.e $g_1(q, y) = qy$.

From (4.9) and (4.10) we get

$$\begin{aligned} H(q, y, x) &= \sum_{n\geq 0} g_n(q, y)x^n \\ &= 1 + qyxH(q, q^2y, x)H(x, y, q) + x(H(q, q^2y, x) - 1) \\ &= 1 + qyx\sum_{n\geq 0} g_n(q, q^2y)x^n \sum_{n\geq 0} g_n(q, y)x^n + x\sum_{n\geq 0} g_n(q, q^2y)x^n - x. \end{aligned} \tag{4.11}$$

By convolution rule we have

$$\begin{aligned} H(q, q^2y, x)H(x, y, q) &= \sum_{n\geq 0} g_n(q, q^2y)x^n \sum_{n\geq 0} g_n(q, y)x^n \\ &= \sum_{k=0}^{\infty}\sum_{i=0}^{k} g_i(q, y)g_{k-i}(q, q^2y)x^k \end{aligned} \tag{4.12}$$

then

$$\sum_{n\geq 0} g_n(q, y)x^n = 1 + qyx\sum_{k=0}^{\infty}\sum_{i=0}^{k} g_i(q, y)g_{k-i}(q, q^2y)x^k + x\sum_{n\geq 0} g_n(q, q^2y)x^n - x$$

$$[x^n]\sum_{n\geq 0} g_n(q,y)x^n = [x^n]\left(1+qy\sum_{k=0}^{\infty}\sum_{i=0}^{k} g_i(q,y)g_{k-i}(q,q^2y)x^{k+1} + \sum_{n\geq 0} g_n(q,q^2y)x^{n+1} - x\right)$$

$$g_n(q,y) = qy\sum_{i=0}^{n-1} g_i(q,y)g_{n-1-i}(q,q^2y) + g_{n-1}(q,q^2y). \tag{4.13}$$

Using computer algebra such as Mathematica we get the following equations

$$g_2(q,y) = q^2y^2 + q^3y + q^4y^2, \tag{4.14}$$

$$g_3(q,y) = q^3y^3 + q^4y^2 + q^5y + 2q^5y^3 + 2q^6y^2 + q^7y^3 + q^8y^2 + q^9y^3. \tag{4.15}$$

Therefore to find $f_n(q)$ we use the fact that $f_n(q) = g_n(q,1)$ thus

$$f_2(q) = q^2 + q^3 + q^4, \tag{4.16}$$

$$f_3(q) = q^3 + q^4 + 3q^5 + 2q^6 + q^7 + q^8 + q^9. \tag{4.17}$$

Remark. We can interpret the two equations (4.16) and (4.17) as follows. In both $f_2(q)$ and $f_3(q)$ the coefficient of q^3 is 1. Summing up these coefficients we get 2. This means that there are only 2 skew paths with area 3. This corresponds to $a_3 = 2$ in the previous Section 4.1.1.

The coefficient of q^4 is 1 in each $f_2(q)$, $f_3(q)$ and $f_4(q)$, again this means that there are only 3 skew paths with area 4, this corresponds to $a_4 = 3$ in the previous Section 4.1.1.

If we continue in this way for q^5, q^6, q^7 etc we get 5, 8, 13 respectively. This shows that the two methods of counting skew paths using different combination of parameters produce the same results. Again this is the Fibonacci sequence.

Remark. We note that equations (4.16) and(4.17) are divisible by q^2, q^3, and they are polynomials of degrees 2^2, 3^2 respectively. In general $f_n(q)$ is divisible by q^n and is a polynomial of degree n^2.

Let $h_n(q)$ be generating function for all Dyck paths with semi-length n, where q marks the area of the Dyck path.

We derive $h_n(q)$ by first obtaining the generating function $t_n(q,y)$, of all Dyck paths with semi-length n , where y marks the semi-base and q marks the area, then $h_n(q) = t_n(q,1)$. Let $T(x,y,q)$ be the generating function of all Dyck paths, where

x marks the semi-length, q marks the area and y marks the semi-base. From Figure 1.2 (from Chapter one) we get

$$T(q, y, x) = 1 + qyxT(q, q^2y, x)T(q, y, x).$$

Let

$$T(q, y, x) = \sum_{n\geq 0} t_n(q, y)x^n.$$

We know that there is only one Dyck path with area zero, then $t_0(q, y) = 1$ and also there is only one Dyck path of area 1 with both semi-length and semi-base equal to 1, i.e $t_1(q, y) = qy$.

From the two above equations we get

$$\begin{aligned}
T(q, y, x) &= \sum_{n\geq 0} t_n(q, y)x^n \\
&= 1 + qyxT(q, q^2y, x)T(x, y, q) \\
&= 1 + qyx \sum_{n\geq 0} t_n(q, q^2y)x^n \sum_{n\geq 0} t_n(q, y)x^n.
\end{aligned}$$

then

$$\begin{aligned}
&\sum_{n\geq 0} t_n(q, y)x^n = 1 + qyx \sum_{k=0}^{\infty} \sum_{i=0}^{k} t_i(q, y)t_{k-i}(q, q^2y)x^k \\
&[x^n] \sum_{n\geq 0} t_n(q, y)x^n = [x^n] \left(1 + qy \sum_{k=0}^{\infty} \sum_{i=0}^{k} t_i(q, y)t_{k-i}(q, q^2y)x^{k+1}\right) \\
&t_n(q, y) = qy \sum_{i=0}^{n-1} t_i(q, y)t_{n-1-i}(q, q^2y).
\end{aligned}$$

Using computer algebra such as Mathematica we get the following equations

$$\begin{aligned}
t_2(q, y) &= q^2y^2 + q^4y^2, \\
t_3(q, y) &= q^3y^3 + 2q^5y^3 + q^7y^3 + q^9y^3.
\end{aligned}$$

Therefore to find $h_n(q)$ we use the fact that $h_n(q) = t_n(q, 1)$ thus

$$h_0(q) = 1,$$

$$h_1(q) = q,$$

$$h_2(q) = q^2 + q^4,$$

$$h_3(q) = q^3 + 2q^5 + q^7 + q^9.$$

We can explain the above four equations as follows,

(1) there is only one Dyck path with both semi-length and area equal to 0 ($4^0 - \frac{1}{2}\binom{2(0)+2}{0+1}$),

(2) there is only one Dyck path with both semi-length and area equal to 1 ($4^1 - \frac{1}{2}\binom{2(1)+2}{1+1}$),

(3) there are exatly two Dyck paths with each having the semi-length equal to 2 and the sum of their areas equal to 6 ($4^2 - \frac{1}{2}\binom{2(2)+2}{2+1}$), and

(4) there are exatly five Dyck paths with each having the semi-length equal to 3 and the sum of their areas equal to 29 ($4^3 - \frac{1}{2}\binom{2(3)+2}{3+1}$), respectively.

We now conclude by saying that given all Dyck paths with semi-lengths n the sum of their areas is equal to $4^n - \frac{1}{2}\binom{2n+2}{n+1}$ [12].

We now use an alternative method from the above, by defining skew paths with left most peak. Let $F_k(q, x)$ be a generating function where k is the left most peak, q marks the area and x marks the semi-length.

Let

$$f(q; x) = \sum_{k \geq 0} F_k(q; x) = \sum_{n \geq 0} f_n(q) x^n. \tag{4.18}$$

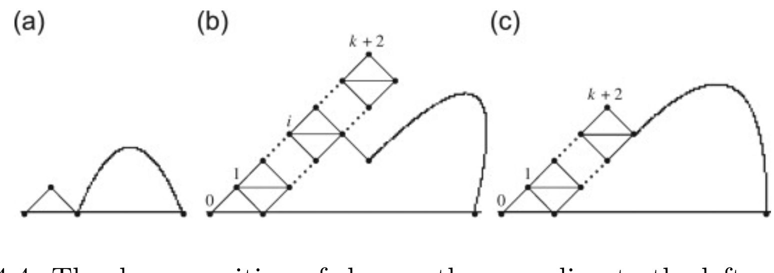

Figure 4.4: The decomposition of skew paths according to the left most peak

We know that if there is only one skew path with no left most peak, then $F_0(q;x) = 1$. We also know that a skew path with one left most peak can be decomposed as $ud\gamma'$, where $\gamma' \in S$ (see Figure 4.4(a)), then we get $F_1(q;x) = qxf(q;x)$. For all skew paths with left most peak at least two we get the two following cases.

1. $\gamma = u^{k+2}dl^{k-i+1}\gamma'$, with γ', not empty, starting with a down step and $i \le k$ (see Figure 4.4(b)). The class of skew paths γ is the same as the class of skew paths $\bar{\gamma} = u^i\gamma'$ where the left most peak is at level i, with semi-length reduced by $k+2-i$ and the area reduced by $2k+3$, this first case produces $q^{2k+3}\sum_{i=0}^{k} x^{k+2-i}F_i(q;x)$.

2. $\gamma = u^{k+2}d\gamma'$, with γ' starting with an up step or down step(see Figure 4.4(c)). Now the class of skew paths γ is the same as the class of skew paths $\bar{\gamma} = u^{k+1}\gamma'$ where the left most peak is at level $k+1$, with semi-length reduced by $k+2-i$ and the area reduced by $2k+3$, this second case produces $q^{2k+3}\sum_{i=k+1}^{\infty} xF_i(q;x)$. Combining these two cases we get

$$\begin{aligned}
F_{k+2}(q;x) &= q^{2k+3}\sum_{i=0}^{k} x^{k+2-i}F_i(q;x) + q^{2k+3}\sum_{i=k+1}^{\infty} xF_i(q;x) \\
&= q^{2k+3}\sum_{i=0}^{k} x^{k+2-i}F_i(q;x) + q^{2k+3}x\left[\sum_{i\ge 0} F_i(q;x) - \sum_{i=0}^{k} F_i(q;x)\right] \\
&= q^{2k+3}\sum_{i=0}^{k} x^{k+2-i}F_i(q;x) + q^{2k+3}x\left[f(q;x) - \sum_{i=0}^{k} F_i(q;x)\right] \\
&= \sum_{i=0}^{k}[x^{k+2-i} - x]F_i(q;x)q^{2k+3} + q^{2k+3}xf(q;x).
\end{aligned} \tag{4.19}$$

We know that $F_0(q;x) = 1$ and $F_1(q;x) = qxf(q;x)$ now

$$F(q;x,y) = \sum_{k\ge 0} F_k(q;x)y^k. \tag{4.20}$$

Here y marks the left most peak. We prove the following theorem by using (4.19) and (4.20).

Theorem 4.2 [6]

$$F(q;x,y) = 1 + \frac{qxy}{1-q^2y} f(q;x) - \frac{q^3(x-x^2)y^2}{(1-q^2xy)(1-q^2y)} F(q;x.q^2y). \tag{4.21}$$

Proof

$$\begin{aligned} F(q;x,y) &= F_0(q;x) + F_1(q;x)y + F_2(q;x)y^2 + ... \\ &= 1 + qxf(q;x)y + \sum_{k\geq 0} F_{k+2}(q;x)y^{k+2} \end{aligned} \tag{4.22}$$

from (4.19)

$$\begin{aligned} \sum_{k\geq 0} F_{k+2}(q;x)y^{k+2} &= \sum_{k\geq 0} \left(q^{2k+3}y^{k+2} \sum_{i=0}^{k} x^{k+2-i} F_i(q;x) + q^{2k+3}xy^{k+2} \left[f(q;x) - \sum_{i=0}^{k} F_i(q;x) \right] \right) \\ = \sum_{k\geq 0} q^{2k+3}y^{k+2} \sum_{i=0}^{k} x^{k+2-i} F_i(q;x) &+ \sum_{k\geq 0} q^{2k+3}xy^{k+2} f(q;x) - \sum_{k\geq 0} q^{2k+3}xy^{k+2} \sum_{i=0}^{k} F_i(q;x) \end{aligned} \tag{4.23}$$

then we get

$$\begin{aligned} &F(q;x,y) = 1 + qxf(q;x)y + \\ &\sum_{k\geq 0} \left(q^{2k+3}y^{k+2} \sum_{i=0}^{k} x^{k+2-i} F_i(q;x) + q^{2k+3}xy^{k+2} \left[f(q;x) - \sum_{i=0}^{k} F_i(q;x) \right] \right) \\ &= 1 + \frac{qxy}{1-q^2y} f(q;x) - \frac{q^3(x-x^2)y^2}{(1-q^2xy)(1-q^2y)} F(q;x.q^2y). \end{aligned} \tag{4.24}$$

We first derive the middle term $\frac{qxy}{1-q^2y} f(q;x)$ of (4.24), in (4.23) from $\sum_{k\geq 0} q^{2k+3}xf(q;x)y^{k+2}$ we get,

$$\sum_{k\geq 0} q^{2k+3}xf(q;x)y^{k+2} = \frac{f(q;x)xq^3y^2}{1-q^2y}.$$

Since $F_1(q;x) = qxf(q;x)$, then we have

$$\begin{aligned} F_1(q;x)y &= qxf(q;x)y \\ &= \frac{(1-q^2y)qxf(q;x)y}{1-q^2y} \\ &= \frac{f(q;x)qxy - f(q;x)xq^3y^2}{1-q^2y}. \end{aligned}$$

Now we get the middle term as follows

$$\begin{aligned} &qxf(q;x)y + \sum_{k\geq 0} q^{2k+3}xf(q;x)y^{k+2} \\ &= \frac{f(q;x)qxy - f(q;x)xq^3y^2}{1-q^2y} + \frac{f(q;x)xq^3y^2}{1-q^2y} \\ &= \frac{qxy}{1-q^2y}f(q;x). \end{aligned}$$

Having found the middle term of (4.24), we have to find the last term $-\frac{q^3(x-x^2)y^2}{(1-q^2xy)(1-q^2y)}F(q;x,q^2y)$, from (4.23) we get

$$\begin{aligned} &\sum_{k\geq 0} q^{2k+3}y^{k+2}\sum_{i=0}^{k} x^{k+2-i}F_i(q;x) - \sum_{k\geq 0} q^{2k+3}xy^{k+2}\sum_{i=0}^{k} F_i(q;x) \\ &= \sum_{k\geq 0}\left(\sum_{i=0}^{k}[x^{k+2-i} - x]F_i(q;x)q^{2k+3}y^{k+2}\right) \\ &= y^2q^3\sum_{i\geq 0}\left(F_i(q;x)\sum_{k\geq i} x^ky^kq^{2k}x^{2-i} - F_i(q;x)\sum_{k\geq i} xy^kq^{2k}\right) \\ &= y^2q^3\sum_{i\geq 0}\left[\frac{F_i(q;x)x^2y^iq^{2i}}{1-xq^2y} - \frac{xF_i(q;x)y^iq^{2i}}{1-yq^2}\right] \\ &= y^2q^3\left[\frac{x^2}{1-xq^2y}\sum_{i\geq 0} F_i(q;x)y^iq^{2i} - \frac{x}{1-yq^2}\sum_{i\geq 0} F_i(q;x)y^iq^{2i}\right] \\ &= y^2q^3\left[x^2\frac{F(q;x,yq^2)}{1-xq^2y} - \frac{xF(q;x;q^2y)}{1-yq^2}\right] \end{aligned}$$

$$= y^2q^3\left[\frac{x^2 - x^2yq^2 - x + x^2yq^2}{(1-xq^2y)(1-yq^2)}\right]F(q,x,q^2y)$$
$$= -\frac{q^3(x-x^2)y^2}{(1-q^2xy)(1-q^2y)}F(q;x,q^2y).$$

We now have prove that the last term is $-\frac{q^3(x-x^2)y^2}{(1-q^2xy)(1-q^2y)}F(q;x,q^2y)$, thus

$$F(q;x,y) = 1 + \frac{qxy}{1-q^2y}f(q;x) - \frac{q^3(x-x^2)y^2}{(1-q^2xy)(1-q^2y)}F(q;x,q^2y),$$

as required.

Theorem 4.3 [6]

$$f(q;x) = \frac{\sum_{k\geq o}\frac{(-1)^k q^{k(2k+1)}(x-x^2)^k}{(1-q^2x)...(1-q^{2k}x)(1-q^2)...(1-q^{2k})}}{\sum_{k\geq o}\frac{(-1)^k q^{k(2k-1)}}{(1-q^2)...(1-q^{2k})}\frac{(x-x^2)^k}{(1-x)(1-q^2x)...(1-q^{2k-2}x)}}.$$

Proof

We see that to get $F(q;x.q^2y)$, we need to replace y by q^2y in $F(q;x,y)$ from Theorem 4.2, then we get

$$F(q;x,q^2y) = 1 + \frac{qxq^2y}{1-q^4y}f(q;x) - \frac{q^3(x-x^2)q^4y^2}{(1-q^2xq^2y)(1-q^4y)}F(q;x,q^4y), \tag{4.25}$$

then we substitute (4.25) into Theorem 4.2 and get

$$\begin{aligned} F(q;x,y) = 1 &+ \frac{qxy}{1-q^2y}f(q;x) - \frac{q^3(x-x^2)y^2}{(1-q^2xy)(1-q^2y)} \\ &\times\left(1 + \frac{qxq^2y}{1-q^4y}f(q;x) - \frac{q^3(x-x^2)q^4y^2}{(1-q^2xq^2y)(1-q^4y)}F(q;x,q^4y)\right). \end{aligned} \tag{4.26}$$

We simplify (4.26) and get

$$\begin{aligned}F(q;x,y) = 1 + \frac{qxy}{1-q^2y} f(q;x) - \frac{q^3(x-x^2)y^2}{(1-q^2xy)(1-q^2y)} \\ - \frac{q^6(x-x^2)xy^3 f(q;x)}{(1-q^2xy)(1-q^2y)(1-q^4y)} \\ + \frac{q^{10}(x-x^2)^2y^4F(q;x,q^4y)}{(1-q^2xy)(1-q^4xy)(1-q^2y)(1-q^4y)}.\end{aligned} \tag{4.27}$$

We iterate and obtain

$$\begin{aligned}F(q;x,y) = 1 + \sum_{k\geq 1} \frac{(-1)^k q^{k(2k+1)}(x-x^2)^k y^{2k}}{(1-q^2xy)...(1-q^{2k}xy)(1-q^2y)...(1-q^{2k}y)} \\ + \sum_{k\geq 1} \frac{(-1)^{k-1}q^{k(2k-1)}x^k(1-x)^{k-1}y^{2k-1}f(q;x)}{(1-q^2xy)...(1-q^{2(k-1)}xy)(1-q^2y)...(1-q^{2k}y)}.\end{aligned}$$

Now we know that $F(q;x,1) = f(q;x)$, therefore

$$\begin{aligned}f(q;x) = \frac{1+\sum_{k\geq 1} \frac{(-1)^k q^{k(2k+1)}(x-x^2)^k}{(1-q^2x)...(1-q^{2k}x)(1-q^2)...(1-q^{2k})}}{1-\sum_{k\geq 1}\frac{(-1)^{k-1}q^{k(2k-1)}}{(1-q^2)...(1-q^{2k})}\frac{x^k(1-x)^{k-1}}{(1-q^2x)...(1-q^{2k-2}x)}} \\ f(q;x) = \frac{\sum_{k\geq 0} \frac{(-1)^k q^{k(2k+1)}(x-x^2)^k}{(1-q^2x)...(1-q^{2k}x)(1-q^2)...(1-q^{2k})}}{\sum_{k\geq o}\frac{(-1)^{k}q^{k(2k-1)}}{(1-q^2)...(1-q^{2k})}\frac{(x-x^2)^k}{(1-x)(1-q^2x)...(1-q^{2k-2}x)}},\end{aligned} \tag{4.28}$$

as required.
Using Mathematica to expand (4.28) we obtain the following polynomials

$$\begin{aligned}f_0(q) &= 1, \\ f_1(q) &= 1, \\ f_2(q) &= q^2+q^3+q^4, \\ f_3(q) &= q^3+q^4+3q^5+2q^6+q^7+q^8+q^9.\end{aligned}$$

4.2 Superdiagonal bargraphs

A bargraph is a colunm-convex polyomino where all the columns are bottom justified. Bargraphs are well know combinatorial objects, in this section we will study the enumeration of bargraphs according to semi-perimeter. We consider a special case of bargraphs called superdiagonal bargraphs, derived from skew paths in a natural way. In each skew path γ from (0,0) to ($2n$,0) we can form a superdiagonal bargraph $B(\gamma)$ whose boundary is defined by the path γ itself, rotated anti-clockwise by $\frac{\pi}{4}$, and by the lines $y = 0$ and $x = n$ (see Figure 4.5).

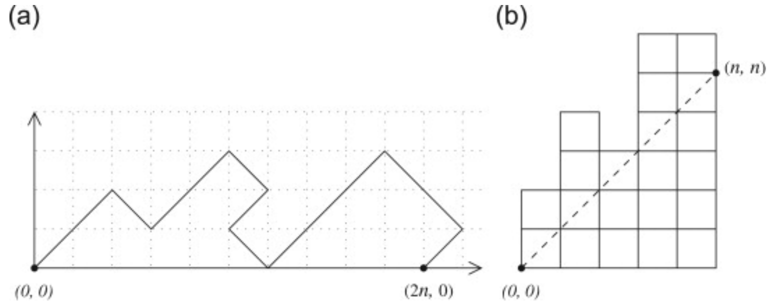

Figure 4.5: A skew Dyck path with corresponding superdiagonal bargraph

4.2.1 Enumeration of superdiagonal bargraphs according to semi-perimeter

In this Subsection we solve the problem of enumerating all superdiagonal bargraphs according to their semi-perimeter. We easily see that the perimeter of a superdiagonal bargraph $B(\gamma)$ is given by the skew path γ of semi-base n. Let $u(\gamma)$, $d(\gamma)$, and $l(\gamma)$ denote the numbers of up, down and left steps in γ respectively. Therefore we have $n = d(\gamma)$ and $l(\gamma)$+$d(\gamma)$=$u(\gamma)$. Thus the perimeter of $B(\gamma)$ is given by $l(\gamma)$+$d(\gamma)$+$u(\gamma)$+$2n$=$2(u(\gamma)$+$d(\gamma))$.

We define $S(x, y)$ to be the generating function for skew paths, where x marks the number of up steps and y marks the number of down steps. The generating function $b(x)$ for super-diagonal bargraphs according to semi-perimeter, can be derived from $S(x, y)$ by setting $x = y$, so that $b(x) = S(x, x)$. Using equation (1.4) for skew paths we get the equation:

$$S(x, y) = 1 + xyS^2(x, y) + x(S(x, y) - 1), \tag{4.29}$$

from which we get

$$
\begin{aligned}
S(x,y) &= \frac{-(x-1)-\sqrt{(x-1)^2-4(xy)(1-x)}}{2xy} \\
&= \frac{1-x-\sqrt{x^2-2x+1-4xy+4x^2y}}{2xy} \\
&= \frac{1-x-\sqrt{1-2x+x^2-4xy+4x^2y}}{2xy}.
\end{aligned}
\tag{4.30}
$$

We let $x=y$ in equations (4.29) and (4.30), this produces the generating function $b(x)$ as follows,

$$x^2b(x)^2-(1-x)b(x)+1-x=0. \tag{4.31}$$

From (4.29) we get

$$b(x)=\frac{1-x-\sqrt{1-2x-3x^2+4x^3}}{2x^2}. \tag{4.32}$$

We show that $b(x)=C(\frac{x^2}{1-x})$, from the generating function of the Dyck paths $C(z)=\dfrac{1-\sqrt{1-4z}}{2z}$, we have that

$$
\begin{aligned}
C\left(\frac{x^2}{1-x}\right) &= \frac{1-\sqrt{1-\frac{4x^2}{1-x}}}{2\frac{x^2}{1-x}} \\
&= \frac{1-\sqrt{\frac{1-x-4x^2}{1-x}}}{\frac{2x^2}{1-x}} \\
&= \frac{1-x-\sqrt{(1-x-4x^2)(1-x)}}{2x^2} \\
b(x) &= \frac{1-x-\sqrt{1-2x-3x^2+4x^3}}{2x^2}.
\end{aligned}
\tag{4.33}
$$

Now we want $[x^n]b(x)=b_n$, we use $b(x)=C(\frac{x^2}{1-x})$ and $C(z)=\sum_{k\geq 0}C_kz^k$.

$$b_n=[x^n]\sum_{k\geq 0}C_k\left(\frac{x^2}{1-x}\right)^k$$

$$= [x^n] \sum_{k\geq 0} C_k x^{2k} \sum_{j\geq 0} \binom{-k}{j} (-1)^j x^j$$

$$= [x^n] \sum_{k\geq 1} C_k \sum_{j\geq 0} \binom{k+j-1}{k-1} x^{j+2k}.$$

Let $n = j + 2k$ *then*

$$b_n = \sum_{k=1}^{n-1} \binom{n-k-1}{k-1} C_k. \tag{4.34}$$

Now we conclude that the number of superdiagonal bargraphs with n columns are counted by $\sum_{k=1}^{n-1} \binom{n-k-1}{k-1} C_k$.

Chapter 5

Moments of Generalized Motzkin Paths

In this chapter we study the paper titled Moments of Generalized Motzkin Paths by [20]. We consider the paths and their moments, the recurrences, enumerating restricted paths, factorial moments, area, and second moments.

5.1 The paths and their moments

Let w be a non-negative integer. We consider the lattice paths in the Cartesian plane with up steps $u = (1,1)$, down steps $d = (1,-1)$, and horizontal steps $q = (w,0)$ (where horizontal steps are a multiple of w). If we have $w = 0$, then only u steps and d steps are allowed. The steps u and d are each weighted by assigning 1 in each of them, the q step is weighted by assigning it with t. We have the t-weight of the path P, denoted by $|P|$, which is the product of the weights of its steps and we have the t-weight for the set of paths S, denoted by $|S|$, which is the sum of the t-weights of the paths in the set S.

We define $U(x,y)$ to be all unrestricted lattice paths using the allowed steps starting from (0,0) and ending at (x,y). Let $M(x,y)$ denote the generalized Motzkin paths which is the set of paths in $U(x,y)$ that starts at (0,0) and end on the x-axis. In this chapter we are interested in the set of elevated paths, which is denoted by $E(x,y)$, we define $E(x,y)$ to be those paths in $M(x,y)$ that never touch the x-axis except at (0,0) and at the end of the path. For an example, see Figure 5.1 and the left column of Table 5.1 which shows the seven *elevated Motzkin paths* in $E(6,0)$

when $w = 1$.

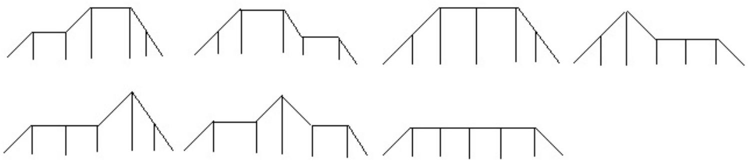

Figure 5.1: The seven elevated Motzkin paths of $E(6, 0)$

Let $f_n(w) = |E(n, 0)|$ with $n \geq 2$, and for all w, we define $f_0(w) = f_1(w) = 0$. If we let $t = 1$, we find that there are three well known sequences from this notation, we know that for $w = 0$, we only have the up u steps and down steps d. Then we have the paths $E(n, 0)$ which are called elevated Dyck paths the elevated Dyck sequence is $(f_n(0))_{n\geq 2} = (1, 0, 1, 0, 2, 0, 5, 0, 14, ...)$, which is the sequence of (aerated) Catalan numbers. If we let $w = 1$ and $t = 1$, then we get the sequence $(f_n(1))_{n\geq 2} = (1, 1, 2, 4, 9, 21, 51, 127, 323, ..)$, which is the sequence of Motzkin numbers. Finaly if we let $w = 2$ and $t = 1$, then we have the following sequence $(f_n(2))_{n\geq 2} = (1, 0, 2, 0, 6, 0, 22, 0, 90, 0, 394, ...)$, which is called (aerated) large Schr$\ddot{o}$der numbers.

Let the path P be the curve in $E(n, 0)$. Let $(j, P(j))$ for $j \in [0, 1, 2, ..., n]$, be coordinates on the path P. Let the r^{th} moment of the path P be $\frac{1}{n-1}\sum_{j=1}^{n-1} p(j)^r$.

We define the $zero^{th}$ moment of P to be equal to 1. It is clear that $\sum_{0\leq j\leq n} P(j)$ is the area bounded by the path P and the x-axis.

Path	Contribution	Contribution	Contribution	Contribution
	to $f_6(1)$	to $g_6(1)$	to Total Area	to $h_6(1)$
$uquqdd$	t^2	$7t^2/5$	$7t^2$	$11t^2/5$
$uuqdqd$	t^2	$7t^2/5$	$7t^2$	$11t^2/5$
$uuqqdd$	t^2	$8t^2/5$	$8t^2$	$14t^2/5$
$uudqqd$	t^2	$6t^2/5$	$6t^2$	$8t^2/5$
$uqqudd$	t^2	$6t^2/5$	$6t^2$	$8t^2/5$
$uqudqd$	t^2	$6t^2/5$	$6t^2$	$8t^2/5$
$uqqqqd$	t^4	$5t^4/5$	$5t^4$	$5t^4/5$

Table 5.1: $f_6(1) = 6t^2 + t^4$, $a_6(1) = 40t^2 + 5t^4$, $g_6(1) = 8t^2 + t^4$ and $h_6(1) = 12t^2 + t^4$

For $w \geq 0$ and $n \geq 2$, we first define the sums $f_n(w), a_n(w), g_n(w), h_n(w)$ of the t-weighted moments for the path set $E(n, 0)$ as follows,

$$f_n(w) = |E(n,0)| = \sum_{P\in E(n,0)} |P|, \tag{5.1}$$

$$a_n(w) = \sum_{P\in E(n,0)} |P| \sum_{j=1}^{n-1} P(j), \tag{5.2}$$

$$g_n(w) = \sum_{P\in E(n,0)} \frac{|P|}{n-1} \sum_{j=1}^{n-1} P(j), \tag{5.3}$$

$$h_n(w) = \sum_{P\in E(n,0)} \frac{|P|}{n-1} \sum_{j=1}^{n-1} P(j)^2. \tag{5.4}$$

We now do an example for the case $E(6,0)$ using Table 5.1, where there are only seven elevated Motzkin paths of $E(6,0)$ (see Figure 5.1).

Example 5.1

We start by finding $f_6(1)$. In Table 5.1 there is only one graph with 4 horizontal steps q's. The other 6 graphs all have 2 horizontal steps. Where each horizontal step is marked by q.As each q is weighted by a t we have six t^2's and one t^4 as shown in Table 5.1.

Thus $f_6(1) = \sum_{P\in E(6,0)} |P| = t^2+t^2+t^2+t^2+t^2+t^2+t^4 = 6t^2+t^4$.

For $a_6(1)$ we have

$$a_6(1) = \sum_{P\in E(6,0)} |P| \sum_{j=1}^{5} P(j).$$

Here we add all the unit heights in each graph in Figure 5.1. For example in the graph *uquqdd* we have the heights

$$P(1)+P(2)+P(3)+P(4)+P(5) = 1+1+2+2+1 = 7.$$

But there are two q's in *uquqdd* and thus we get t^2. Hence *uquqdd* contribute $7t^2$ to $a_6(1)$. We do the same for all the other graphs. The results are shown in Table 5.1. For all seven graphs in Table 5.1 we get

$a_6(1) = 7t^2+7t^2+6t^2+6t^2+6t^2+5t^4 = 40t^2+5t^4$.

We do the same for

$$g_6(1) = \sum_{P\in E(6,0)} \frac{|P|}{5} \sum_{j=1}^{5} P(j)$$

as for $a_6(1)$ except that we divide $a_6(1)$ by 5. Thus

$$g_6(1) = \frac{7t^2 + 7t^2 + 6t^2 + 6t^2 + 6t^2 + 5t^4}{5} = \frac{40t^2 + 5t^4}{5} = 8t^2 + t^4.$$

Now for $h_6(1)$, using the graph *uquqdd* and from the formula

$$h_6(1) = \sum_{P \in E(6,0)} \frac{|P|}{5} \sum_{j=1}^{5} P(j)^2. \tag{5.5}$$

We get

$$\begin{aligned} &\frac{t^2(P(1)^2 + P(2)^2 + P(3)^2 + P(4)^2 + P(5)^2)}{5} \\ &= \frac{t^2(1 + 1 + 4 + 4 + 1)}{5} \\ &= \frac{11t^2}{5}. \end{aligned}$$

Thus *uquqdd* contribute $\frac{11t^2}{5}$ to $h_6(1)$. Now adding the results from all the seven graphs we get $h_6(1) = \frac{11t^2 + 11t^2 + 14t^2 + 8t^2 + 8t^2 + 8t^2 + 5t^4}{5} = 12t^2 + t^4$.

If we let $t = 1$ from the above discusion we obtain the following results,

(1) there are only 7 ($f_6(1) = 7$) elevated Motzkin paths with $n = 6$,

(2) the total area for all elevated Motzkin paths with $n = 6$ is 45 ($a_6(1) = 45$),

(3) the total average area for all elevated Motzkin paths with $n = 6$ is ($g_6(1) = 9$), and

(4) the total second moment for all elevated Motzkin paths with $n = 6$ is 2.2 ($h_6(1) = \frac{11}{5}$).

5.2 The recurrences

Let $w \geq 0$. We now state three recurrences for the sequences $(f_n(w))_{n\geq 2}$, $(g_n(w))_{n\geq 2}$ and $(h_n(w))_{n\geq 2}$ which we are going to prove in Sections 5.3 and 5.5,

$$nf_n = 4(n-3)f_{n-2} + (2n-3w)tf_{n-w} - (n-3w)t^2 f_{n-2w}. \tag{5.6}$$

$$(n-1)g_n = 4(n-3)g_{n-2} + (2n-2w-2)tg_{n-w} - (n-2w-1)t^2 g_{n-2w}. \tag{5.7}$$

$$(n-2)h_n = 4(n-3)h_{n-2} + (2n-w-4)th_{n-w} - (n-w-2)t^2 h_{n-2w}. \tag{5.8}$$

Proposition 5.2.1 *[20] In the above three recurrences, let $w = 0$ and $t = 0$ then the three sequences $(f_n(0))_{n\geq 2}$, $(g_n(0))_{n\geq 2}$ and $(h_n(0))_{n\geq 2}$ produce the following three elevated Dyck paths recurrences;*

$$nf_n(0) = 4(n-3)f_{n-2}(0), \tag{5.9}$$

$$(n-1)g_n(0) = 4(n-3)g_{n-2}(0), \tag{5.10}$$

$$(n-2)h_n(0) = 4(n-3)h_{n-2}(0), \tag{5.11}$$

for any $n \geq 3$, given the initial conditions $f_n(0) = g_n(0) = h_n(0) = 0$ for all $n < 2$ and $f_2(0) = g_2(0) = h_2(0) = 1$.

From the above three equations (5.9), (5.10) and (5.11) we set $n = 2k+2$ and get

Proposition 5.2.2[20]

$$f_{2k+2}(0) = \frac{1}{k+1}\binom{2k}{k} = C_k,$$

which is all the number of elevated Dyck paths with $k+1$ up steps,

$$g_{2k+2}(0) = \frac{4^k}{2k+1},$$

which is the total average area for all elevated Dyck paths with $k+1$ up steps, and

$$h_{2k+2}(0) = \binom{2k}{k},$$

which is the total second moment for all elevated Dyck paths with $k+1$ up steps.

Proof of Proposition 5.2.2

We start with $f_{2k+2}(0) = \frac{1}{k+1}\binom{2k}{k}$, taking into account that $f_2(0) = 1$.

From equation (5.9) we have

$$\begin{aligned}
nf_n(0) &= 4(n-3)f_{n-2}(0) \\
&= 4(n-3)(4n-5)\frac{f_{n-4}(0)}{n-2} \\
&= \frac{4^2(n-3)(n-5)}{n-2}4(n-7)\frac{f_{n-6}(0)}{n-4}
\end{aligned}$$

$$
\begin{aligned}
&= \frac{4^{\frac{n}{2}-1}(n-3)(n-5)(n-7)...(3)(1)f_2}{(n-2)(n-4)...(6)(4)} \\
&= \frac{2 \times 4^{\frac{n}{2}-1}(n-2)!}{[(n-2)(n-4)...(6)(4)(2)]^2}.
\end{aligned} \tag{5.12}
$$

Now for $n = 2k+2$

$$
\begin{aligned}
f_n(0) &= \frac{2 \times 4^{\frac{n}{2}-1}(n-2)!}{n[(n-2)(n-4)...(6)(4)(2)]^2} \\
f_{2k+2}(0) &= \frac{2 \times 4^k (2k)!}{(2k+2)[(2k)(2k-2)...(6)(4)(2)]^2} \\
&= \frac{2 \times 4^k (2k)!}{2(k+1)2^{2k}(k!)^2} \\
&= \frac{1}{k+1}\binom{2k}{k}.
\end{aligned} \tag{5.13}
$$

Secondly we consider $g_{2k+2}(0) = \frac{4^k}{2k+1}$, taking into account that $g_2(0) = 1$.

From equation (5.10) we have

$$
\begin{aligned}
(n-1)g_n(0) &= 4(n-3)g_{n-2}(0) \\
&= 4(n-3)4(n-5)\frac{g_{n-4}(0)}{n-3} \\
&= \frac{4^2(n-3)(n-5)}{n-3}4(n-7)\frac{g_{n-6}(0)}{n-5} \\
&= \frac{4^{\frac{n}{2}-1}(n-3)(n-5)(n-7)...(3)(1)g_2(0)}{(n-3)(n-5)...(5)(3)} \\
&= 4^{\frac{n}{2}-1}.
\end{aligned} \tag{5.14}
$$

Now for $n = 2k+2$

$$
\begin{aligned}
g_n(0) &= \frac{4^{\frac{n}{2}-1}}{n-1} \\
g_{2k+2}(0) &= \frac{4^{k+1-1}}{2k+1} \\
&= \frac{4^k}{2k+1}.
\end{aligned}
$$

$$(5.15)$$

Lastly we consider $h_{2k+2}(0) = \binom{2k}{k}$, taking into account that $h_2(0) = 1$. From equation (5.11) we have

$$\begin{aligned}
(n-2)h_n(0) &= 4(n-3)h_{n-2}(0) \\
&= 4(n-3)4(n-5)\frac{h_{n-4}(0)}{n-4} \\
&= \frac{4^{\frac{n}{2}-1}(n-3)(n-5)(n-7)...(3)(1)h_2(0)}{(n-4)(n-6)...(4)(2)}.
\end{aligned} \tag{5.16}$$

Now for $n = 2k+2$

$$\begin{aligned}
h_n(0) &= \frac{4^{\frac{n}{2}-1}(n-2)!}{[(n-2)(n-4)...(6)(4)(2)]^2} \\
h_{2k+2}(0) &= \frac{4^k(2k)!}{[(2k)(2k-2)(2k-4)...(6)(4)(2)]^2} \\
&= \frac{4^k(2k)!}{2^{2k}(k!)^2} \\
&= \binom{2k}{k}
\end{aligned} \tag{5.17}$$

as required.

From equations (5.2), (5.3) and (5.15) we get the total area for all elevated Dyck paths with $k+1$ up steps to be $a_{2k+2}(0) = 4^k$ see [21].

We illustrate the above discusion with the following two examples.

Example 5.2 We use the unit heights to calculate the areas of the Dyck paths as in Example 5.1.

(1) Figure 5.2 shows that for $k = 0$ (with one up step) there is exactly one elevated Dyck path ($f_2(0) = \frac{1}{0+1}\binom{0}{0} = 1$) with area equal to 1 ($a_2(0) = 4^0 = 1$)(there is 1 unit height on the graph) satisfying the initial condition $f_2(0) = g_2(0) = h_2(0) = 1$,

(2) for $k = 1$ (with two up steps) there is exactly one elevated Dyck path

($f_4(0) = \frac{1}{1+1}\binom{2}{1} = 1$) with area equal to 4 ($a_4(0) = 4$)(there are 4 unit heights on the graph), and

(3) for $k = 2$ (with three up steps) there are only two elevated Dyck paths ($f_6(0) = \frac{1}{2+1}\binom{4}{2} = 2$) with area equal to 16 ($a_6(0) = 4^2$)(there are 16 in total unit heights on the two graphs).

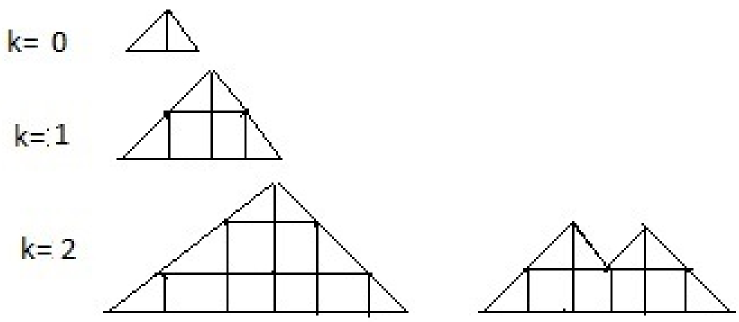

Figure 5.2: The illustation of elevated Dyck paths using unit heights.

Example 5.3 Now from Chapter four we used unit triangles to count the areas of the Dyck paths, using these unit triangles from Figure 5.3 we get the same results as on the Example 5.2 as follows,

(1) For $k = 0$ (with one up step) there is exactly one elevated Dyck path ($f_2(0) = \frac{1}{0+1}\binom{0}{0} = 1$) with area equal to 1 ($a_2(0) = 4^0$)(there is 1 unit triangle on the graph) satisfying the initial condition $f_2(0) = g_2(0) = h_2(0) = 1$,

(2) for $k = 1$ (with two up steps) there is exactly one elevated Dyck path ($f_4(0) = \frac{1}{1+1}\binom{2}{1} = 1$) with area equal to 4 ($a_4(0) = 4$)(there are 4 unit triangles on the graph), and

(3) for $k = 2$ (with three up steps) there are only two elevated Dyck paths ($f_6(0) = \frac{1}{2+1}\binom{4}{2} = 2$) with area equal to 16 ($a_6(0) = 4^2$)(there are 16 in total unit triangles on the two graphs).

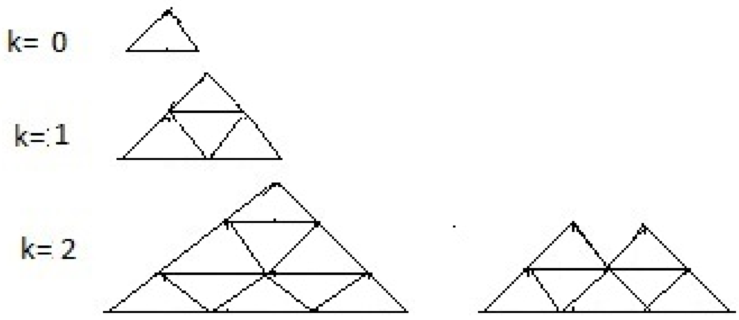

Figure 5.3: The illustation of elevated Dyck paths using unit triangles.

We conclude by saying that these two methods give us the same set of solutions when we use different parameters (unit heights and triangles) to enumerate the elevated Dyck paths, and their areas.

5.3 Enumerating restricted paths

In this section we will derive the restricted paths recurrence

$$nf_n = 4(n-3)f_{n-2} + (2n-3w)tf_{n-w} - (n-3w)t^2 f_{n-2w} \tag{5.18}$$

from the previous Section.

Let $M(z) := M = \sum_{n\geq 0} |M(n,0)|z^n$ be the generating function for the Motzkin paths.

We note that $M(n,0)$ can be an empty path or it can start with an q-step or it can leave the x-axis right from the beginning and return to the x-axis for the first time.

Let $L = \bigcup_{n=0}^{\infty} M(n,0)$. Then we have the decomposition shown in Figure 5.4.

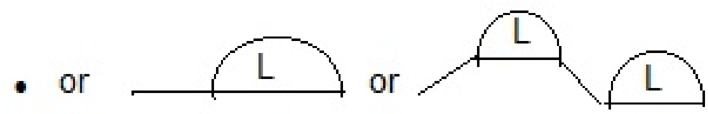

Figure 5.4: The first return decomposition of the Motzkin path

We now form the equation from the first return decomposition
$L = \bigcup_{n=0}^{\infty} M(n,0)$,

$$L = \{\varepsilon\} \cup qL \cup uLdL. \tag{5.19}$$

We now assign z to mark each of the horizontal, up, and down units steps. t weight each horizontal $H = (w,0)$-step, and 1 weight each up and down steps. We have

$$M(z) = 1 + tz^w M(z) + z^2 M(z)^2. \tag{5.20}$$

Solving the above equation for $M(z)$ by using the quadratic formula (in 5.21 we consider the negative square root since we get $M(0+) = lim_{z\downarrow 0} M(z) = 0$, and ignore the positive sign since it produces $M(0) = \frac{2}{0} = \infty$) we obtain

$$M(z) = \frac{1 - tz^w - \sqrt{1 - 4z^2 - 2tz^w + t^2 z^{2w}}}{2z^2}. \tag{5.21}$$

From the previous section we know that $f_n(0) = 0$ for all $n < 2$. Consider the generating function $F(z) = \sum_{n\geq 2} f_n z^n$. Then from $f_{n+2} = |M(n,0)|$. Now

$$F(z) = \frac{1 - tz^w - \sqrt{1 - 4z^2 - 2tz^w + t^2 z^{2w}}}{2}, \tag{5.22}$$

and we get $\Psi(z) = F(z) - \dfrac{1 - tz^w}{2}$. Then

$$\Psi(z) = \frac{-\sqrt{1 - 4z^2 - 2tz^w + t^2 z^{2w}}}{2}. \tag{5.23}$$

We note that the coefficient of z^n in the series expansion of both $\Psi(z)$ and $F(z)$ is the same. Therefore they should have the same coefficients f_n as $F(z)$, unless $n = 0$ or $n = w$.

Taking logarithms and differentiating we have $\ln \Psi(z) = \ln \left(\dfrac{-\sqrt{1 - 4z^2 - 2tz^w + t^2 z^{2w}}}{2} \right)$ then we differentiate as follows:

$$\begin{aligned}
&\frac{d}{dz} \ln \Psi(z) = \frac{d}{dz} \ln \left(\frac{-\sqrt{1 - 4z^2 - 2tz^w + t^2 z^{2w}}}{2} \right) \\
&\frac{\Psi'(z)}{\Psi(z)} = \frac{-8z - 2wtz^{w-1} + 2wt^2 z^{2w-1}}{2(1 - 4z^2 - 2tz^w + t^2 z^{2w})}. \\
&(1 - 4z^2 - 2tz^w + t^2 z^{2w})\Psi'(z) + (4z + wtz^{w-1} - wt^2 z^{2w-1})\Psi(z) = 0.
\end{aligned} \tag{5.24}$$

expressing (5.24) as powers of z we obtain (5.25).

$$(1 - 4z^2 - 2tz^w + t^2 z^{2w}) \sum_{n\geq 2} n f_n z^{n-1} + (4z + wtz^{w-1} - wt^2 z^{2w-1}) \sum_{n\geq 2} f_n z^n = 0 \tag{5.25}$$

We multiply both sides of (5.25) by z and extract coefficients of z^n

$$\begin{aligned}
&(1 - 4z^2 - 2tz^w + t^2 z^{2w}) \sum_{n\geq 2} n f_n z^n + (4z + wtz^{w-1} - wt^2 z^{2w-1}) \sum_{n\geq 2} f_n z^{n+1} = 0 \\
&n f_n - 4(n-2) f_{n-2} - 2t(n-w) f_{n-w} + t^2(n-2w) f_{n-2w} + 4 f_{n-2} \\
&\quad + wt f_{n-w} - wt^2 f_{n-2w} = 0
\end{aligned}$$

$$nf_n = 4(n-3)f_{n-2} + (2n-3w)tf_{n-w} - (n-3w)t^2 f_{n-2w}. \tag{5.26}$$

as required by the first recurrence (5.18).

5.4 Factorial moments

In this Section we introduce factorial moments. This will help us to derive the last two recurrences namely (5.27) and (5.28),

$$(n-1)g_n = 4(n-3)g_{n-2} + (2n-2w-2)tg_{n-w} - (n-2w-1)t^2 g_{n-2w} \tag{5.27}$$

and

$$(n-2)h_n = 4(n-3)h_{n-2} + (2n-w-4)th_{n-w} - (n-w-2)t^2 h_{n-2w}. \tag{5.28}$$

We consider the following falling factorial moment,

$$\mu(n,r) = \sum_{P \in E(n,0)} |P| \sum_{0<j<n} (P(j))_r. \tag{5.29}$$

Here $(n)_r = n(n-1)_{r-1}$, and $(n)_0 = 1$

We first prove three lemmas that we will need in order to prove Proposition 5.4.3.

Lemma 5.4.1 [20] $\dfrac{z^2M^2}{(1-z^2M^2)^2} = \dfrac{z^2}{(1-tz^w)^2 - 4z^2}$

Proof

From the generating function $M(z) = 1 + tz^w M(z) + z^2 M(z)^2$, we get

$$\begin{aligned}
(1-tz^w)M(z) &= 1 + z^2M(z)^2 \\
(1-tz^w)M(z) - 2 &= -1 + z^2M(z)^2 \\
((1-tz^w)M(z) - 2)^2 &= (1 - z^2M(z)^2)^2,
\end{aligned}$$

hence we have

$$\begin{aligned}
\frac{z^2M^2}{(1-z^2M^2)^2} &= \frac{z^2M^2}{((1-tz^w)M-2)^2} \\
&= \frac{z^2M^2}{(1-tz^w)^2M^2 + 4 - 4(1-tz^w)M} \\
&= \frac{z^2M^2}{(1-tz^w)^2M^2 + 4 - 4(1+z^2M(z)^2)}
\end{aligned}$$

$$
\begin{aligned}
&= \frac{z^2M^2}{(1-tz^w)^2M^2 - 4z^2M^2}\\
&= \frac{z^2}{(1-tz^w)^2 - 4z^2},
\end{aligned}
\tag{5.30}
$$

as required.

Lemma 5.4.2 [20] $\dfrac{z^2M^2}{1-z^2M^2} = \dfrac{(1-tz^w-\sqrt{(1-tz^w)^2-4z^2})}{2\sqrt{(1-tz^w)^2-4z^2}}$.

Proof

From the equation $M(z) = \dfrac{1-tz^w-\sqrt{1-4z^2-2tz^w+t^2z^{2w}}}{2z^2}$ we derive the following three equations.

Let

$$\triangle = 1-4z^2-2tz^w+t^2z^{2w} = (1-tz^w)^2-4z^2. \tag{5.31}$$

This gives

$$2z^2M(z) = 1-tz^w-\sqrt{\triangle}. \tag{5.32}$$

Therefore

$$
\begin{aligned}
\frac{z^2M^2}{1-z^2M^2} &= \frac{4z^2(z^2M^2)}{4z^2(1-z^2M^2)}\\
&= \frac{(2z^2M)^2}{4z^2-(2z^2M)^2}\\
&= \frac{(1-tz^w-\sqrt{\triangle})^2}{4z^2-(1-tz^w-\sqrt{\triangle})^2}\\
&= \frac{(1-tz^w-\sqrt{\triangle})^2}{4z^2-(1-tz^w)^2+2(1-tz^w)\sqrt{\triangle}-\triangle}\\
&= \frac{(1-tz^w-\sqrt{\triangle})^2}{-\triangle+2(1-tz^w)\sqrt{\triangle}-\triangle}\\
&= \frac{(1-tz^w-\sqrt{\triangle})^2}{-2\triangle+2(1-tz^w)\sqrt{\triangle}}\\
&= \frac{(1-tz^w-\sqrt{\triangle})^2}{2\sqrt{\triangle}(1-tz^w-\sqrt{\triangle})}\\
&= \frac{(1-tz^w-\sqrt{\triangle})}{2\sqrt{\triangle}}\\
&= \frac{(1-tz^w-\sqrt{(1-tz^w)^2-4z^2})}{2\sqrt{(1-tz^w)^2-4z^2}},
\end{aligned}
$$

$$(5.33)$$

as required.

We define an Iverson $\mathcal{I}(B)$, where $\mathcal{I}(B) = 1$ if B is a true and $\mathcal{I}(B) = 0$ if B is false. Let (i, k) be a point at the end of the path P or let (i, k) be a point inside the horizontal step of the path P. This horizontal step will start from (j, k) and stops at $(j + w, k)$. Let Q be any path in $E(j, k)$. Let R and R' be two paths such that R starts from (j, k) and stops at $(n, 0)$ and R' starts from $(j+w, k)$ and stops at $(n, 0)$: By symmetry, R can be matched with a path in $E(n-j, k)$. Let $m(j, k) = |E(j, k)|$ then $m(n-j, k) = |E(n-j, k)|$.

We now use Lemma 5.4.1, Lemma 5.4.2, $\mu(n, r)$ and the above definitions to prove Proposition 5.4.3.

Proposition 5.4.3 [20] *Let r be a positive integer then*

$$\sum_{n\geq 2} \mu(n,r) z^n = \frac{r! z^2 z^n (1 + (w-1) t z^w) \left(1 - t z^w - \sqrt{(1 - t z^w)^2 - 4z^2}\right)^{r-1}}{2^{r-1} (\sqrt{(1 - t z^w)^2 - 4z^2})^{r+1}}. \quad (5.34)$$

Proof

From equation (5.29) with $n \geq 2$, we have

$$\begin{aligned}
\mu(n,r) &= \sum_{P \in E(n,0)} |P| \sum_{0<j<n} (P(j))_r \\
&= \sum_{P} \sum_{k>0} \sum_{0<j<n} (k)_r |P| B(P(j) = k) \\
&= \sum_{k>0} \sum_{0<j<n} \sum_{P} (k)_r |P| B(P(j) = k) \\
&= \sum_{k>0} (k)_r \sum_{0<j<n} \sum_{P} |P| (B((j,k) \text{ is the end of } P) + B((j,k) \text{point inside of } P) \\
&= \sum_{k>0} (k)_r \sum_{i>0} [\sum_{Q,R} |Q||R| + \sum_{Q,R'} (w-1) t |Q||R'|] \\
&= \sum_{k} (k)_r [\sum_{i} m(i,k) m(n-i,k) + (w-1) t \sum_{i} m(i,k) m(n-i-w,k)].
\end{aligned} \quad (5.35)$$

Multiplying $\mu(n, r)$ by z^n and summing up n from two to infinity we get,

$$\sum_{n\geq 2} \mu(n,r) z^n$$

$$= \sum_k (k)_r [\sum_n \sum_i m(i,k)m(n-i,k)z^n + (w-1)t \sum_n \sum_i m(i,k)m(n-i-w,k)z^n]$$
$$= \sum_k (k)_r [\sum_n \sum_i m(i,k)m(n-i,k)z^n + (w-1)tz^w \sum_n \sum_i m(i,k)m(n-i-w,k)z^{n-w}]*$$
$= [1+(w-1)tz^w] \sum_k (k)_r (zM)^{2k}$** From * to ** see comment at end of the proof.

By the Binomial Theorem

$$\frac{r!y^r}{(1-y)^{r+1}} = \sum_{k \geq r} (k)_r y^k$$

From ** we get $[1+(w-1)tz^w] \dfrac{r!(z^2M^2)^r}{(1-z^2M^2)^{r+1}}$

$$= r![1+(w-1)tz^w] \frac{z^2M^2}{(1-z^2M^2)^2} \frac{(z^2M^2)^{r-1}}{(1-z^2M^2)^{r-1}}.R$$

Then by Lemmas 5.4.1 and 5.4.2 $\dfrac{z^2M^2}{(1-z^2M^2)^2} = \dfrac{z^2}{(1-tz^w)^2-4z^2}$

and $\dfrac{z^2M^2}{1-z^2M^2} = \dfrac{(1-tz^w-\sqrt{(1-tz^w)^2-4z^2})}{2\sqrt{(1-tz^w)^2-4z^2}}$.

From R we get

$$r![1+(w-1)tz^w] \frac{z^2}{(1-tz^w)^2-4z^2} \frac{(1-tz^w-\sqrt{(1-tz^w)^2-4z^2})^{r-1}}{(2\sqrt{(1-tz^w)^2-4z^2})^{r-1}}$$
$$= \frac{r!z^2(1+(w-1)tz^w)(1-tz^w-\sqrt{(1-tz^w)^2-4z^2})^{r-1}}{2^{r-1}(\sqrt{(1-tz^w)^2-4z^2})^{r+1}}, \tag{5.36}$$

as required.

To establish this string we first note that each path in $E(j,k)$ must depart from each line $y=d$, for integer d, $0 \leq d < k$, for a last time, hence a simple convolution argument shows that the generating function for $m(j,k)$ (see Figure 5.5) satisfies

$$\sum_j m(j,k)z^j = (zM(z))^k.$$

5.5 Area and second moments

In this Section we prove the total average area recurrence formula,

$$(n-1)g_n = 4(n-3)g_{n-2} + (2n-2w-2)tg_{n-w} - (n-2w-1)t^2g_{n-2w}, \tag{5.37}$$

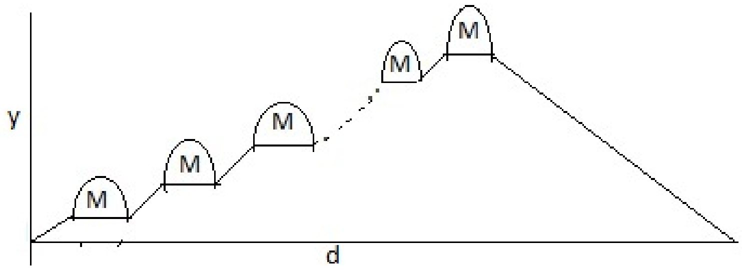

Figure 5.5: This Figure shows the k Motzkin paths departing from line $y = d$

and the total second moment recurrence formula,

$$(n-2)h_n = 4(n-3)h_{n-2} + (2n-w-4)th_{n-w} - (n-w-2)t^2h_{n-2w}. \tag{5.38}$$

5.5.1 The recurrence (5.37)

We start by proving the total average area recurrence formula

For $r = 1$ in $\mu(n,r) = \sum_{P\in E(n,0)} |P| \sum_{0<j<n}(P(j))_r$ we get

$$\mu(n,1) = \sum_{P\in E(n,0)} |P| \sum_{j=1}^{n-1} P(j).$$

We see that from equation (5.2) $a_n(w) = \sum_{P\in E(n,0)} |P| \sum_{j=1}^{n-1} P(j)$ we get $\mu(n,1) = a_n(w) = a_n$ therefore for $r = 1$ in

$$\sum_{n\geq 2} \mu(n,1)z^n = \frac{z^2(1+(w-1)tz^w)(1-tz^w-\sqrt{(1-tz^w)^2-4z^2})^{1-1}}{2^{1-1}(\sqrt{(1-tz^w)^2-4z^2})^{1+1}},$$

we have

$$\sum_{n\geq 2} a_n z^n = \frac{z^2(1+(w-1)tz^w)}{(1-tz^w)^2-4z^2}. \tag{5.39}$$

We now prove the recurrence (5.37) by using the two formulas from equations (5.2) and (5.3),

$$a_n(w) = \sum_{P\in E(n,0)} |P| \sum_{j=1}^{n-1} P(j),$$

and

$$g_n(w) = \sum_{P \in E(n,0)} \frac{|P|}{n-1} \sum_{j=1}^{n-1} P(j).$$

Relating the two equations we get

$$\begin{aligned}
&(n-1)g_n(w) = a_n(w) \\
&\sum_{n\geq 2}(n-1)g_n(w)z^n = \sum_{n\geq 2} a_n(w)z^n.
\end{aligned} \tag{5.40}$$

We proceed as follows

$$\begin{aligned}
&\sum_{n\geq 2} a_n z^n = \frac{z^2(1+(w-1)tz^w)}{(1-tz^w)^2-4z^2} \\
&\text{cross multiplying we get} \\
&((1-tz^w)^2-4z^2)\sum_{n\geq 2} a_n z^n = z^2(1+(w-1)tz^w) \\
&\text{substituting } a_n = (n-1)g_n \text{ we get} \\
&((1-tz^w)^2-4z^2)\sum_{n\geq 2}(n-1)g_n z^n = z^2(1+(w-1)tz^w).
\end{aligned} \tag{5.41}$$

Now we extract the coefficients of z^n as follows,

$$\begin{aligned}
&[z^n](1-4z^2-2tz^w+t^2z^{2w})\sum_{n\geq 2}(n-1)g_n z^n = [z^n](z^2(1+(w-1)tz^w)) \\
&(n-1)g_n - 4(n-3)g_{n-2} - 2t(n-w-1)g_{n-w} + t^2(n-2w-1)g_{n-2w} = 0 \\
&(n-1)g_n = 4(n-3)g_{n-2} + (2n-2w-2)tg_{n-w} - (n-2w-1)t^2 g_{n-2w},
\end{aligned} \tag{5.42}$$

as required in equation (5.37).

5.5.2 The recurrence (5.38)

We now derive the total second moment recurrence formula (5.38)

We use the total second moment formula (5.4) $h_n(w) = \sum_{P\in E(n,0)} \dfrac{|P|}{n-1} \sum_{j=1}^{n-1} P(j)^2$.

Let $h_n(w) := h_n$

We consider the following falling factorial moment

$$\mu(n,r) = \sum_{P\in E(n,0)} |P| \sum_{0<j<n} (P(j))_r.$$

Here $(n)_r = n(n-1)_{r-1}$, and $(n)_0 = 1$.

Now let $H(z) = \sum_{n\geq 2} h_n z^n$ be the generating function for the second moment for h_n. Then

$$H(z) = \sum_{n\geq 2} h_n z^n$$
$$= \sum_{n\geq 2} \frac{1}{n-1} \sum_{P\in E(n,0)} |P| \sum_{j=1}^{n-1} P(j)^2 z^n.$$

Now we use the following factorial moments

$(P(j))_1 + (P(j))_2 = P(j) + P(j)(P(j)-1) = P(j)^2$

to get

$$\sum_{n\geq 2} \sum_{P\in E(n,0)} \frac{|P|}{n-1} \sum_{j=1}^{n-1} ((P(j))_1 + (P(j))_2) z^n$$
$$= \sum_{n\geq 2} \frac{1}{n-1} \left(\sum_{P\in E(n,0)} |P| \sum_{j=1}^{n-1} (P(j))_1 + \sum_{P\in E(n,0)} |P| \sum_{j=1}^{n-1} (P(j))_2 \right) z^n.$$

We know that

$$\mu(n,1) = \sum_{P\in E(n,0)} |P| \sum_{0<j<n} (P(j))_1 \text{ and } \mu(n,2) = \sum_{P\in E(n,0)} |P| \sum_{0<j<n} (P(j))_2.$$

Therefore we get

$$H(z) = \sum_{n\geq 2} \frac{1}{n-1} (\mu(n,1) + \mu(n,2)) z^n. \tag{5.43}$$

In Proposition 5.4.3 let $r = 2$ and $r = 1$. Then

$$\sum_{n\geq 2} \frac{1}{n-1} \mu(n,2) z^n = \sum_{n\geq 2} \frac{1}{n-1} \frac{2! z^2 (1+(w-1)tz^w)(1-tz^w-\sqrt{(1-tz^w)^2-4z^2}) z^n}{2(\sqrt{(1-tz^w)^2-4z^2})^3}$$

and

$$\sum_{n\geq 2}\frac{1}{n-1}\mu(n,1)z^n = \sum_{n\geq 2}\frac{1}{n-1}\frac{z^2(1+(w-1)tz^w)z^n}{(1-tz^w)^2-4z^2}\left(\frac{\sqrt{(1-tz^w)^2-4z^2}}{\sqrt{(1-tz^w)^2-4z^2}}\right)$$

$$= \sum_{n\geq 2}\frac{1}{n-1}\frac{z^2(1+(w-1)tz^w)z^n(\sqrt{(1-tz^w)^2-4z^2})}{(\sqrt{(1-tz^w)^2-4z^2})^3}$$

respectively.

We have

$$H(z) = \sum_{n\geq 2}\frac{1}{n-1}(\mu(n,1)+\mu(n,2))z^n$$

$$= z\sum_{n\geq 2}\frac{1}{n-1}(\mu(n,1)+\mu(n,2))z^{n-1}$$

$$= z\int\sum_{n\geq 2}(\mu(n,1)+\mu(n,2))z^{n-2}dz$$

$$= z\int\sum_{n\geq 2}\frac{z^2(1+(w-1)tz^w)(1-tz^w)}{\sqrt{((1-tz^w)^2-4z^2)^3}}z^{n-2}dz$$

from paper [20] it is stated that we get $H(z) = \dfrac{z^2}{\sqrt{((1-tz^w)^2-4z^2)}}$. (5.44)

Let $\Psi(z) = \sum_{n\geq 0} h_{n+2}z^n = \dfrac{H(z)}{z^2} = \dfrac{1}{\sqrt{((1-tz^w)^2-4z^2)}}$ and $\ln(\Psi(z)) = \frac{-1}{2}\ln((1-tz^w)^2-4z^2)$.

Now we differentiate $\ln(\Psi(z)) = \frac{-1}{2}\ln((1-tz^w)^2-4z^2)$ with respect to z and get

$$\frac{\Psi'(z)}{\Psi(z)} = \frac{4z+twz^{w-1}-t^2wz^{2w-1}}{(1-tz^w)^2-4z^2}$$

$$((1-tz^w)^2-4z^2)\Psi'(z)-(4z+twz^{w-1}-t^2wz^{2w-1})\Psi(z)=0$$

$$((1-tz^w)^2-4z^2)\sum_{n\geq 1}nh_{n+2}z^{n-1}-(4z+twz^{w-1}-t^2wz^{2w-1})\sum_{n\geq 0}h_{n+2}z^n=0$$

$$[z^{n-2}](1-4z^2-2tz^w+t^2z^{2w})\sum_{n\geq 0} nh_{n+2}z^n - [z^{n-2}](4z+twz^{w-1}-t^2wz^{2w-1})\sum_{n\geq 0} h_{n+2}z^{n+1} = 0$$

$$(n-2)h_n - 4(n-4)h_{n-2} - 2t(n-w-2)h_{n-w} + t^2(n-2w-2)h_{n-2w} - 4h_{n-2} - twh_{n-w}$$
$$+ t^2wh_{n-2w} = 0$$
$$(n-2)h_n = 4(n-3)h_{n-2} + (2n-w-4)th_{n-w} - (n-w-2)t^2h_{n-2w}, \tag{5.45}$$

as required in (5.38).

Bibliography

[1] C. Banderier, B. Gittenberger, enumeration and asymptotics for the area, *Discrete Math.* (2006) 345-355

[2] D. Callan, Two bijections for Dyck path parameters, preprint, 4pp, (2004). {://www.arxiv.org/abs/math.CO/0406381}.

[3] R. Chapman, Moments of Dyck paths, *Discrete Math*, **204** (1999), 113-117.

[4] E. Deutsch, A bijection on Dyck paths and its consequences, *Discrete Math.* **179** (1998) 253-256.

[5] E. Deutsch, Dyck path emumeration, *Discrete Math.* **204**, (1999), 167-202.

[6] E. Deutsch, E. Munarini, S. Rinaldi, Skew Dyck paths, *Journal of Statistical Planning and Inference*, **140**, (2010), 2191-2203.

[7] E. Deutsch, E. Munarini, S. Rinaldi, Skew Dyck paths, area, and superdiagonal bargraphs, **140**, (2010), 1550-1562.

[8] J. Furlinger, J. Hofbauer q-Catalan numbers, J.Comb. Theory Ser. **40** (1985) 248-264

[9] G. Kreweras, Aires des chemins surdiagonaux a tapes obliques permises. *Cahier du* B.U.R.O. **24** (1976) 9-18

[10] T. Mansour, Counting peaks at height k in a Dyck path, *J. Integer Seq.* **5** (2002) Article 02.1.1.

[11] T. Mansour, Statistics on Dyck paths, *J. Integer Seq.* **9** (2006) Article 06.1.5.

[12] D. Merlini, R. Sprugnoli, and M. C. Verri, The area determined by underdiagonal lattice paths, Proceedings of CAAP'96, Lecture Notes in Computer Science **1059** (1996) 59-71.

[13] D. Merlini, R. Sprungoli, M. Verri, *Some statistics on Dyck paths*, J. Statist. Plann. Inference **101** (2002) 211-227.

[14] E.J.J Van Rensgurg, Statistical mechanics of directed models of polymers in the square lattice.J.Phys.A Math.Gen. **36** (2003) R11-R61.

[15] A. Sapounakis, I. Tasoulas, P. Tsikouras, Dyck path statistics, Proceedings of Ninth WSEAS Conference MATH06, Istanbul, Turkey, (2006).

[16] A. Sapounakis, I. Tasoulas and P. Tsikouras, Counting strings in Dyck paths, *Discrete Math*, **307** (2007), 2909-2924.

[17] A. Sapounakis, P. Tsikouras, Counting peaks and valleys in k-colored Motzkin paths, Electron. J. Combin. **12** (2005), R16.

[18] N.J.A Sloane, Online Encyclopedia of Integer Sequences, published electronically at {http://www.research.att.com/njas/sequences/}.

[19] Y. Sun, The statistic "number of *udu*'s" in Dyck paths, *Discrete Math.* **287** (2004), 177-186.

[20] R. Sulanke, Moments of Generalized Motzkin Paths, *Article* 00.1.1 **3** (2000).

[21] W.J Woan, L. Shapiro, and D. G. Rogers, The Catalan numbers, the Lebesgue integral and 4^{n-2}, Am. Math. Monthly, **104**, (1997) 926-931.