AUSTRALASIAN JOURNAL OF COMBINATORICS
Volume **39** (2007), Pages 49–72

Some strings in Dyck paths

A. Sapounakis I. Tasoulas P. Tsikouras

Department of Informatics
University of Piraeus
18534 Piraeus
Greece
{arissap, jtas, pgtsik}@unipi.gr

Abstract

The enumeration of Dyck paths according to various statistics has been studied in several papers. This paper deals with the statistics "number of τ's" for several strings τ. More precisely, the strings $(az\bar{a})^j$, $(az\bar{a})^j a$ and $a^i\bar{a}a^j$, where z is a fixed Dyck path, are considered and several known results, as well as many new ones, are derived.

1 Introduction

A *Dyck path* of semilength n is a lattice path of $\mathbb{N}^2$ running from $(0,0)$ to $(2n,0)$, whose allowed steps are the up diagonal step $(1,1)$ and the down diagonal step $(1,-1)$. These steps are called *rise* and *fall* respectively.

It is clear that every Dyck path of semilength n is coded by a word $u = u_1u_2\cdots u_{2n} \in \{a,\bar{a}\}^*$, called *Dyck word*, so that every rise (respectively fall) corresponds to the letter a (respectively $\bar{a}$).

Throughout this paper we denote with $\mathcal{D}$ the set of all Dyck paths (or equivalently Dyck words). Furthermore the subset of $\mathcal{D}$ that contains all the paths u of semilength $l(u) = n$ is denoted with $\mathcal{D}_n$. It is well-known that $|\mathcal{D}_n| = C_n$, where $C_n = \frac{1}{n+1}\binom{2n}{n}$ is the n-th Catalan number (A000108 of [11]), with generating function $C(x)$ which satisfies the relation

$$xC^2(x) - C(x) + 1 = 0.$$

Furthermore, the powers of $C(x)$ are given (see [3]) by the following relation:

$$[x^n]C^s = \frac{s}{2n+s}\binom{2n+s}{n},$$

for $(n,s) \neq (0,0)$.

The Catalan numbers are closely related to the Motzkin numbers M_n (A001006 of [11]) according to the following well-known identities ([4], [1]):

$$M_n = \sum_{k=0}^{[\frac{n}{2}]} \binom{n}{2k} C_k = \sum_{k=0}^{n} (-1)^k \binom{n}{k} C_{n+1-k}.$$

A word $\tau \in \{a, \bar{a}\}^*$, called in this context *string*, *occurs* in a Dyck path u if $u = w\tau v$, where $w, v \in \{a, \bar{a}\}^*$. If a string τ does not occur in u, we say that u *avoids* τ.

The statistic "number of occurrences of τ" (or simply "number of τ's") has been studied for various strings τ. This statistic has been studied for strings of length 2 in [3] and it has been shown that it follows the Narayana distribution (A001263 of [11]). Generalizations for strings of length 2 in k-colored Motzkin paths are given in [9].

The strings of length 3 have also been studied extensively. For instance, if $\tau = \bar{a}aa$ the corresponding statistic follows the Touchard distribution (see [3]), whereas if $\tau = a\bar{a}a$ it follows the Donaghey distribution (see [8, 12]). Some interesting bijective proofs for the above results are given in [2].

Finally, in [10] a systematic study of the statistic "number of τ's" for every string τ of length up to 4 is given.

We say that a string τ *occurs at height* j in a Dyck path, where $j \in \mathbb{N}$, if the minimum height of the points of τ in this occurrence is equal to j. An occurrence of a string τ at height equal to (respectively greater than) 0 is usually referred to as a *low* (respectively *high*) occurrence of τ.

The statistic "number of τ's at height j" has been first considered in [6] for $\tau = ud$ and $\tau = du$, and it is has been proved that the corresponding generating function can be expressed via Chebyshev polynomials and the Catalan generating function. In the same direction, in [10] it has been proved that for an arbitrary string τ the generating function corresponding to the statistic "number of τ's at height j" is evaluated with the aid of the generating function corresponding to the statistic "number of low τ's".

Recently, the statistics "number of τ's", "number of low τ's" and "number of high τ's" have been studied in [6] for the strings $\tau = (a\bar{a})^j a$ and $\tau = a^j \bar{a} a$, where $j \in \mathbb{N}^*$.

In this paper we study the above statistics for strings of a more general form.

In Section 2 we study the strings $(az\bar{a})^j$ and $(az\bar{a})^j a$, where z is a fixed Dyck path. This is accomplished with the use of multivariable generating functions which in addition allow us to obtain the corresponding statistics for the string $(a\bar{a})^j a^2$.

Finally, in Section 3 we deal with the string $a^i \bar{a} a^j$, showing bijectively that the statistics "number of $a^i \bar{a} a^j$'s" and "number of $a^j \bar{a} a^i$'s" are equidistributed. For the study of this string we use again multivariable generating functions which in addition allow us to obtain the corresponding statistics for the string $a^i \bar{a} a^j \bar{a}$.

2 The strings $(az\bar{a})^j$ and $(az\bar{a})^j a$

Throughout this section we denote with z a fixed Dyck path of semilength r, and we deal with the statistics "number of τ's", "number of low τ's" and "number of high τ's", where τ is either of the strings $(az\bar{a})^j$ or $(az\bar{a})^j a$, for $j \in \mathbb{N}^*$.

2.1 The statistics "number of $(az\bar{a})^j$'s" and "number of $(az\bar{a})^j a$'s"

For every $j \in \mathbb{N}^*$ and $u \in \mathcal{D}$, we denote with $\gamma_j(u)$ (respectively $\delta_j(u)$) the number of $(az\bar{a})^j$'s (respectively $(az\bar{a})^j a$'s) in u.

We consider the generating function of the set of Dyck paths according to the semilength and to the above parameters:

$$F(x;\mathbf{s},\mathbf{t}) = \sum_{u\in\mathcal{D}} x^{l(u)} \prod_{j\geq 1} s_j^{\gamma_j(u)} \prod_{j\geq 1} t_j^{\delta_j(u)},$$

where $\mathbf{s} = (s_j)$ and $\mathbf{t} = (t_j)$, for $j \in N^*$.

We have the following result.

Proposition 2.1 *The generating function $F = F(x;\mathbf{s},\mathbf{t})$ satisfies the equation*

$$x\left(\sum_{i=0}^{\infty} x^{(r+1)i} \prod_{j=1}^{i+1} (s_j t_j)^{i-j+1}\right) F^2 - \left(1 + x^{r+1} \sum_{i=0}^{\infty} x^{(r+1)i} \prod_{j=1}^{i+1} (s_j t_j)^{i-j+1}\right) F + \left(1 + x^{r+1} \sum_{i=0}^{\infty} x^{(r+1)i} \prod_{j=1}^{i+1} (s_j t_j)^{i-j+1} \prod_{j=1}^{i+1} s_j\right) = 0. \quad (1)$$

Proof. We consider the partition $\{\mathcal{B}_i\}_{i\in\mathbb{N}}$ of $\mathcal{D}$ defined by

$$\mathcal{B}_0 = \{\epsilon\} \cup \{aw\bar{a}v : w \in \mathcal{D} \setminus \{z\}, v \in \mathcal{D}\}$$

and

$$\mathcal{B}_i = \{(az\bar{a})^i v : v \in \mathcal{B}_0\},$$

for $i \geq 1$; (see Fig. 1a and 1b respectively).

Let $u = aw\bar{a}v$, where $w, v \in \mathcal{D}$ and $w \neq z$; since for every $j \in \mathbb{N}$ we have

$$\gamma_j(u) = \gamma_j(w) + \gamma_j(v) \text{ and } \delta_j(u) = \delta_j(w) + \delta_j(v),$$

we deduce easily that the generating function $B_0 = B_0(x;\mathbf{s},\mathbf{t})$ of the set $\mathcal{B}_0$ is given by

$$B_0 = 1 + xF(F - x^r). \quad (2)$$

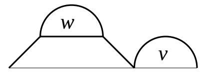

Figure 1a : The non-empty elements of $\mathcal{B}_0$

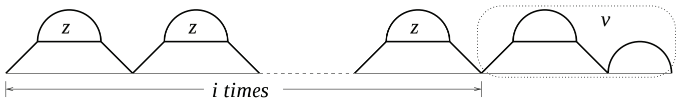

Figure 1b : The elements of $\mathcal{B}_i$

Furthermore, for every $v \in \mathcal{B}_0$ and $i, j \in \mathbb{N}^*$ we have

$$\gamma_j((az\bar{a})^i v) = \begin{cases} \gamma_j(v), & \text{if } i < j \\ \gamma_j(v) + i - j + 1, & \text{if } j \leq i \end{cases}$$

and

$$\delta_j((az\bar{a})^i v) = \begin{cases} \delta_j(v), & \text{if } i < j \\ i - j, & \text{if } j \leq i \text{ and } v = \epsilon \\ \delta_j(v) + i - j + 1, & \text{if } j \leq i \text{ and } v \neq \epsilon. \end{cases}$$

Hence, we obtain that the generating functions $B_i = B_i(x; \mathbf{s}, \mathbf{t})$ of the sets $\mathcal{B}_i$, for $i \geq 1$, are given by the relations

$$\begin{aligned} B_i &= \sum_{v \in \mathcal{B}_0} x^{l((az\bar{a})^i) + l(v)} \prod_{j \geq 1} s_j^{\gamma_j((az\bar{a})^i v)} \prod_{j \geq 1} t_j^{\delta_j((az\bar{a})^i v)} \\ &= x^{(r+1)i} \left(\left(\sum_{v \in \mathcal{B}_0 \setminus \{\epsilon\}} x^{l(v)} \prod_{j \geq 1} s_j^{\gamma_j(v)} \prod_{j \geq 1} t_j^{\delta_j(v)} \right) \prod_{j=1}^{i+1} s_j^{i-j+1} \prod_{j=1}^{i+1} t_j^{i-j+1} \right. \\ &\qquad \left. + \prod_{j=1}^{i+1} s_j^{i-j+1} \prod_{j=1}^{i} t_j^{i-j} \right) \\ &= x^{(r+1)i} \prod_{j=1}^{i+1} s_j^{i-j+1} \left((B_0 - 1) \prod_{j=1}^{i+1} t_j^{i-j+1} + \prod_{j=1}^{i} t_j^{i-j} \right). \end{aligned}$$

Using relation (2), we have

$$\begin{aligned} F &= \sum_{i=0}^{\infty} B_i \\ &= 1 + xF^2 - x^{r+1}F + \sum_{i=1}^{\infty} x^{(r+1)i} \prod_{j=1}^{i+1} s_j^{i-j+1} \left((xF^2 - x^{r+1}F) \prod_{j=1}^{i+1} t_j^{i-j+1} + \prod_{j=1}^{i} t_j^{i-j} \right) \end{aligned}$$

giving the required result. □

Remark From relation (1) it is clear that the generating function F depends only on the semilength of the Dyck path z and not on z itself.

To illustrate this, consider another Dyck path z^* with $l(z) = l(z^*)$ and let γ_j^*, δ_j^*, where $j \in \mathbb{N}^*$, be the corresponding parameters. Then, with the aid of a simple involution h of $\mathcal{D}$ defined inductively, it is shown that the parameters γ_j, γ_j^* (respectively δ_j, δ_j^*) are equidistributed for every $j \in \mathbb{N}^*$.

Indeed, if we set $h(\epsilon) = \epsilon$ and for a non-empty Dyck path $u = aw\bar{a}v$ we define

$$h(u) = \begin{cases} ah(w)\bar{a}h(v), & \text{if } w \neq z, z^* \\ az^*\bar{a}h(v), & \text{if } w = z \\ az\bar{a}h(v), & \text{if } w = z^*, \end{cases}$$

we can easily show by induction on the semilength of the path u, that h is an involution of $\mathcal{D}$, $\gamma_j^*(h(u)) = \gamma_j(u)$ and $\delta_j^*(h(u)) = \delta_j(u)$, for every $u \in \mathcal{D}$, $j \in \mathbb{N}^*$.

In the sequel we deal with the statistic "number of $(a\bar{a})^j a^2$'s".

Corollary 2.2 *The generating function $E = E(x; \mathbf{t})$ which counts the Dyck paths of prescribed semilength according to the statistics "number of $(a\bar{a})^j a^2$'s", where $j \in \mathbb{N}^*$, satisfies the equation*

$$x(1-x)\left(1+\sum_{i=1}^{\infty} x^i \prod_{j=1}^{i} t_j\right) E^2 - (1-x)\left(1+x\left(1+\sum_{i=1}^{\infty} x^i \prod_{j=1}^{i} t_j\right)\right) E + 1 = 0. \quad (3)$$

Proof. We can easily check that the number of all $(a\bar{a})^j a^2$'s in u is equal to $\delta_j(u) - \gamma_{j+1}(u)$ for every $j \in \mathbb{N}^*$ and $u \in \mathcal{D}$.

Thus, for $r = 0$, $s_1 = 1$ and $s_j = t_{j-1}^{-1}$ for $j \geq 2$, we obtain that

$$E(x; \mathbf{t}) = F(x; \mathbf{s}, \mathbf{t})$$

so that, by equation (1), we deduce (3). □

For fixed $k \in \mathbb{N}^*$, let $E_k = E_k(x; \mathbf{t})$ be the generating function of $\mathcal{D}$ according to the semilength and to the number of $(a\bar{a})^k a^2$'s; if we apply equation (3) for $t_k = t$ and $t_j = 1$ for every $j \neq k$, we obtain that

$$x(1-(1-t)x^k)E_k^2 - (1-(1-t)x^{k+1})E_k + 1 = 0.$$

Furthermore, using a version of the Lagrange inversion formula (see [3]) we obtain the following relation:

$$E_k^{\mu}(x,t) = 1 + \sum_{n=1}^{\infty} \sum_{j=0}^{[\frac{n-1}{k+1}]} \sum_{i=0}^{[\frac{n-1}{k+1}]-j} (-1)^i \frac{\binom{i+j}{j}\binom{n-k(i+j)}{i+j}\binom{2n-(2k+1)(i+j)+\mu-1}{n-k(i+j)+\mu}}{n-k(i+j)} x^n t^k \quad (4)$$

for every $\mu \in \mathbb{N}^*$. In particular, we have the following result.

Corollary 2.3 *The number of all Dyck paths of semilength n with j $(a\bar{a})^k a^2$'s is equal to*

$$[x^n t^j]E_k = \sum_{i=0}^{[\frac{n-1}{k+1}]-j} \frac{(-1)^i}{n-k(i+j)} \binom{i+j}{j} \binom{n-k(i+j)}{i+j} \binom{2n-(2k+1)(i+j)}{n-k(i+j)+1}.$$

We now come to study the statistics γ_j and δ_j separately.
For this, we apply equation (1) twice, for $t_j = 1$ for every $j \in \mathbb{N}^*$, and for $s_j = 1$ for every $j \in \mathbb{N}^*$. It follows that the corresponding generating functions

$$\Gamma(x; \mathbf{s}) = \sum_{u \in \mathcal{D}} x^{l(u)} \prod_{j \geq 1} s_j^{\gamma_j(u)} \text{ and } \Delta(x; \mathbf{t}) = \sum_{u \in \mathcal{D}} x^{l(u)} \prod_{j \geq 1} t_j^{\delta_j(u)}$$

satisfy the equations

$$xA(x; \mathbf{s})\Gamma^2(x; \mathbf{s}) - (1 + x^{r+1}A(x; \mathbf{s}))\Gamma(x; \mathbf{s}) + A(x; \mathbf{s}) = 0$$

and

$$xA(x; \mathbf{t})\Delta^2(x; \mathbf{t}) - (1 + x^{r+1}A(x; \mathbf{t}))\Delta(x; \mathbf{t}) + (1 + x^{r+1}A(x; \mathbf{t})) = 0$$

respectively, where

$$A(x; \mathbf{s}) = \sum_{i=0}^{\infty} x^{(r+1)i} \prod_{j=1}^{i+1} s_j^{i-j+1}.$$

Then, we can easily deduce the following result.

Proposition 2.4 *The generating functions $\Gamma(x; \mathbf{s})$ and $\Delta(x; \mathbf{s})$ are given by the formulas*

$$\Gamma(x; \mathbf{s}) = \frac{A(x; \mathbf{s})}{1 + x^{r+1}A(x; \mathbf{s})} C\left(x \left(\frac{A(x; \mathbf{s})}{1 + x^{r+1}A(x; \mathbf{s})}\right)^2\right)$$

and

$$\Delta(x; \mathbf{s}) = C\left(x \frac{A(x; \mathbf{s})}{1 + x^{r+1}A(x; \mathbf{s})}\right).$$

We note that the second formula of the above Proposition has been proved in [7] for $r = 0$.

Example Let γ, δ be the parameters on $\mathcal{D}$ defined by

$$\gamma(u) = \sum_{i=1}^{\infty} (-1)^{i+1} \gamma_i(u) \text{ and } \delta(u) = \sum_{i=1}^{\infty} (-1)^{i+1} \delta_i(u).$$

Then, for $s_{2i-1} = s$ and $s_{2i} = s^{-1}$ for every $i \geq 1$, Proposition 2.4 gives

$$A(x; \mathbf{s}) = \frac{1 + sx^{r+1}}{1 - sx^{2(r+1)}}$$

and hence

$$\Gamma(x,s)=\sum_{u\in\mathcal{D}}x^{l(u)}s^{\gamma(u)}=\frac{1+sx^{r+1}}{1+x^{r+1}}C\left(x\left(\frac{1+sx^{r+1}}{1+x^{r+1}}\right)^2\right)$$

and

$$\Delta(x,s)=\sum_{u\in\mathcal{D}}x^{l(u)}s^{\delta(u)}=C\left(x\frac{1+sx^{r+1}}{1+x^{r+1}}\right).$$

It follows easily that the numbers of all $u\in\mathcal{D}_n$ with $\gamma(u)=j$ and $\delta(u)=j$ respectively, are given by the formulas

$$[x^n s^j]\Gamma=\sum_{i=0}^{\left[\frac{2n+1-(2r+3)j}{2(r+1)}\right]}(-1)^i\binom{2n-2(r+1)(i+j)+1}{j}\binom{2n-(2r+1)i-2(r+1)j}{i}C_{n-(r+1)(i+j)}$$

for $0\leq j\leq\left[\frac{2n+1}{2l+3}\right]$, and

$$[x^n s^j]\Delta=\sum_{i=0}^{\left[\frac{n-(r+2)j}{r+1}\right]}(-1)^i\binom{n-(r+1)(i+j)}{j}\binom{n-(r+1)j-ri-1}{i}C_{n-(r+1)(i+j)}$$

for $0\leq j\leq\left[\frac{n}{r+1}\right]$.
In particular for $r=1$ and $j=0$ we obtain the sequence 1, 1, 1, 2, 5, 13, 35, 97, 275, 794, 2327, 6905, 20705, ... (A082582 of [11]) (respectively 1, 1, 2, 4, 10, 28, 82, 248, 770, 2440, 7858, 25644, ...) which counts the number of all $u\in\mathcal{D}$ such that the total number of $(aa\bar{a}\bar{a})^{2i-1}$'s (respectively $(aa\bar{a}\bar{a})^{2i-1}a$'s) in u is equal to the total number of $(aa\bar{a}\bar{a})^{2i}$'s (respectively $(aa\bar{a}\bar{a})^{2i}a$'s) in u, for $i\in\mathbb{N}^*$.

In the sequel, consider a fixed $k\in\mathbb{N}^*$ and let $\Gamma_k(x,s)$ and $\Delta_k(x,s)$ be the generating functions that we obtain by setting $s_k=s$ and $s_j=1$ for every $j\neq k$, in $\Gamma(x;\mathbf{s})$ and $\Delta(x;\mathbf{s})$ respectively.
In this case we have

$$\begin{aligned}A(x;\mathbf{s})&=\sum_{i=0}^{k-1}x^{(r+1)i}+\sum_{i=k}^{\infty}x^{(r+1)i}s^{i-k+1}\\&=\frac{1-x^{(r+1)k}}{1-x^{r+1}}+\frac{sx^{(r+1)k}}{1-sx^{r+1}}\\&=\frac{1-sx^{r+1}-(1-s)x^{(r+1)k}}{(1-x^{r+1})(1-sx^{r+1})}.\end{aligned}$$

Then, applying Proposition 2.4, we obtain the following result.

Corollary 2.5 *The generating functions $\Gamma_k(x,s)$ and $\Delta_k(x,s)$ are given by*

$$\Gamma_k(x,s)=\frac{1-sx^{r+1}-(1-s)x^{(r+1)k}}{1-sx^{r+1}-(1-s)x^{(r+1)(k+1)}}C\left(x\left(\frac{1-sx^{r+1}-(1-s)x^{(r+1)k}}{1-sx^{r+1}-(1-s)x^{(r+1)(k+1)}}\right)^2\right)$$

and

$$\Delta_k(x,s) = C\left(x\frac{1-sx^{r+1}-(1-s)x^{(r+1)k}}{1-sx^{r+1}-(1-s)x^{(r+1)(k+1)}}\right).$$

We note that the formula for Δ_k has been proved in [7] for $r=0$.

Examples

1. For $k=1$ we obtain that

$$\Gamma_1(x,s) = \frac{1}{1+x^{r+1}(1-s)}C\left(\frac{x}{(1+(1-s)x^{r+1})^2}\right) \tag{5}$$

and

$$\Delta_1(x,s) = C\left(\frac{x}{1+(1-s)x^{r+1}}\right). \tag{6}$$

Using the above relations we can easily show that the number of all $u \in \mathcal{D}_n$ with j $az\bar{a}$'s and the number of all $u \in \mathcal{D}_n$ with j $az\bar{a}a$'s are given respectively by the formulas

$$[x^n s^j]\Gamma_1 = \sum_{i=0}^{[\frac{n}{r+1}]-j} \frac{(-1)^i}{n-r(i+j)}\binom{i+j}{j}\binom{n-r(i+j)}{i+j}\binom{2n-(2r+1)(i+j)}{n-r(i+j)-1}, \tag{7}$$

for $0 \le j \le [\frac{n}{r+1}]$, and

$$[x^n s^j]\Delta_1 = \sum_{i=0}^{[\frac{n-1}{r+1}]-j} (-1)^i\binom{i+j}{j}\binom{n-1-r(i+j)}{i+j}C_{n-(r+1)(i+j)}, \tag{8}$$

for $0 \le j \le [\frac{n-1}{r+1}]$.

Applying formulas (7) and (8) for certain values of r, we obtain some well known results, as well as some new results.

For instance, for $r=0$ we obtain the number of all $u \in \mathcal{D}_n$ with j $a\bar{a}$'s (see [3] and A001263 of [11]) and the number of all $u \in \mathcal{D}_n$ with j $a\bar{a}a$'s (see [12, 8] and A091869 of [11]).

Indeed, using relations 3.49 in [5] we have

$$\begin{aligned}[x^n s^j]\Gamma_1 &= \sum_{i=0}^{n-j}\frac{(-1)^i}{n}\binom{i+j}{j}\binom{n}{i+j}\binom{2n-(i+j)}{n-1}\\ &= \frac{1}{n}\binom{n}{j}\sum_{i=0}^{n-j}(-1)^i\binom{n-j}{i}\binom{2n-(i+j)}{n-1}\\ &= \frac{1}{n}\binom{n}{j}\binom{n}{j-1}\end{aligned}$$

and

$$\begin{aligned}[x^n s^j]\Delta_1 &= \sum_{i=0}^{n-1-j}(-1)^i\binom{i+j}{j}\binom{n-1}{i+j}C_{n-(i+j)}\\ &= \binom{n-1}{j}\sum_{i=0}^{n-1-j}(-1)^i\binom{n-1-j}{i}C_{n-(i+j)}\\ &= \binom{n-1}{j}M_{n-1-j}.\end{aligned}$$

For $r=1$ we find the following triangles whose elements, read by rows, count the number of all $u\in\mathcal{D}_n$ with j $aa\bar{a}\bar{a}$'s (see A098978 of [11]) and the number of all $u\in\mathcal{D}_n$ with j $aa\bar{a}\bar{a}a$'s (see A114848 of [11]) respectively:
1; 1; 1, 1; 2, 3; 5, 8, 1; 13, 23, 6; 35, 69, 27, 1; 97, 212, 110, 10; ... and
1; 1; 2; 4, 1; 10, 4; 28, 13, 1; 82, 44, 6; 248, 153, 27, 1; 770, 536, 116, 8; ...

2. A straightforward application of Corollary 2.5 for $k=2$, $r=0$ and $s=0$ gives the following formulas:

$$[x^n]\Gamma_2=\sum_{i=0}^{n}\sum_{j=0}^{n-i}(-1)^i\binom{2n-2j-i}{i}\binom{2n-2j-i+1}{j}C_{n-i-j}$$

and

$$[x^n]\Delta_2=\sum_{i=0}^{n-1}\binom{n-i}{i}M_{n-i-1}.$$

The first formula gives the sequence 1, 1, 1, 3, 7, 19, 53, 153, 453, 1367, 4191, 13015 ..., which counts the number of all $u\in\mathcal{D}_n$ that avoid $a\bar{a}a\bar{a}$ (A078481 of [11]), whereas the second gives the sequence 1, 1, 2, 4, 11, 31, 92, 283, 893, 2875, ..., which counts the number of all $u\in\mathcal{D}_n$ that avoid $a\bar{a}a\bar{a}a$.

2.2 The statistics "number of low $(az\bar{a})^j$'s" and "number of low $(az\bar{a})^j a$'s"

For every $j\in\mathbb{N}^*$ and $u\in\mathcal{D}$, we denote with $\gamma'_j(u)$ (respectively $\delta'_j(u)$) the number of low $(az\bar{a})^j$'s (respectively low $(az\bar{a})^j a$'s) in u.

We consider the generating function of the set of Dyck paths according to the semilength and to the above parameters:

$$F'(x;\mathbf{s},\mathbf{t})=\sum_{u\in\mathcal{D}}x^{l(u)}\prod_{j\geq1}s_j^{\gamma'_j(u)}\prod_{j\geq1}t_j^{\delta'_j(u)},$$

where $\mathbf{s}=(s_j)$ and $\mathbf{t}=(t_j)$, for $j\in N^*$.

Let $\{\mathcal{B}_i\}_{i\in\mathbb{N}}$ be the partition of $\mathcal{D}$ used in the proof of Proposition 2.1 and $B'_i=B'_i(x;\mathbf{s},\mathbf{t})$ the generating function of $\mathcal{B}_i$ according to the parameters l, γ'_j and δ'_j, for every $j\in\mathbb{N}^*$.

Using similar arguments as in the proof of Proposition 2.1, we deduce the following relations:

$$B_0' = 1 + xF'(C(x) - x^r)$$

and

$$B_i' = x^{(r+1)i} \prod_{j=1}^{i+1} s_j^{i-j+1} \left((B_0' - 1) \prod_{j=1}^{i+1} t_j^{i-j+1} + \prod_{j=1}^{i} t_j^{i-j} \right),$$

for $i \geq 1$, from which we easily deduce the following result.

Proposition 2.6 *The generating function $F' = F'(x; \mathbf{s}, \mathbf{t})$ is given by the formula*

$$F' = \frac{1 + x^{r+1} \sum_{i=0}^{\infty} x^{(r+1)i} \prod_{j=1}^{i+1} (s_j t_j)^{i-j+1} \prod_{j=1}^{i+1} s_j}{1 - x(C(x) - x^r) \sum_{i=0}^{\infty} x^{(r+1)i} \prod_{j=1}^{i+1} (s_j t_j)^{i-j+1}}. \tag{9}$$

If we consider the generating function F' for $r = 0$, $s_1 = 1$ and $s_j = t_{j-1}^{-1}$ for $j \geq 2$, we obtain that the generating function $E' = E'(x; \mathbf{t})$, which counts the number of Dyck paths of prescribed semilength according to the statistics "number of low $(a\bar{a})^j a^2$'s", with $j \in \mathbb{N}^*$, is given by the formula

$$E' = \frac{1}{(1-x)\left(1 - x(C(x) - 1)\left(1 + \sum_{i=1}^{\infty} x^i \prod_{j=1}^{i} t_j\right)\right)}.$$

For fixed $k \in \mathbb{N}^*$, if we apply the above formula for $t_k = t$ and $t_j = 1$ for every $j \neq k$, we obtain the generating function of $\mathcal{D}$ according to the semilength and to the number of low $(a\bar{a})^k a^2$'s :

$$E_k'(x, t) = \frac{C(x)}{1 + (1-t)x^{k+2} C^3(x)}.$$

Furthermore, using some simple manipulations we obtain the following result.

Corollary 2.7 *The number of all Dyck paths of semilength n with j low $(a\bar{a})^k a^2$'s is equal to*

$$[x^n t^j] E_k' = \sum_{i=0}^{[\frac{n}{k+2}]-j} (-1)^i \frac{3(i+j)+1}{2n - (2k+1)(i+j) + 1} \binom{i+j}{j} \binom{2n - (2k+1)(i+j) + 1}{n - (k-1)(i+j) + 1}.$$

We now come to study the statistics γ'_j and δ'_j separately. For this, we apply equation (9) twice, for $t_j = 1$ for every $j \in \mathbb{N}^*$, and for $s_j = 1$ for every $j \in \mathbb{N}^*$. For the corresponding generating functions

$$\Gamma'(x;\mathbf{s}) = \sum_{u\in\mathcal{D}} x^{l(u)} \prod_{j\geq 1} s_j^{\gamma'_j(u)} \text{ and } \Delta'(x;\mathbf{t}) = \sum_{u\in\mathcal{D}} x^{l(u)} \prod_{j\geq 1} t_j^{\delta'_j(u)}$$

we have the following result.

Proposition 2.8 *The generating functions* $\Gamma' = \Gamma'(x;\mathbf{s})$ *and* $\Delta' = \Delta'(x;\mathbf{s})$ *are given by the formulas*

$$\Gamma' = \frac{A}{1 - x(C(x) - x^r)A} \quad \text{and } \Delta' = \frac{1 + x^{r+1}A}{1 - x(C(x) - x^r)A}$$

where

$$A = A(x;\mathbf{s}) = \sum_{i=0}^{\infty} x^{(r+1)i} \prod_{j=1}^{i+1} s_j^{i-j+1}.$$

For fixed $k \in \mathbb{N}^*$, using an argument similar to that of Corollary 2.5, we obtain the following result.

Corollary 2.9 *The generating functions* $\Gamma'_k = \Gamma'_k(x,s)$ *and* $\Delta'_k = \Delta'_k(x,s)$ *are given by*

$$\Gamma'_k = \frac{1 - sx^{r+1} - (1-s)x^{(r+1)k}}{1 - x^{(r+1)(k+1)} - x(1 - x^{(r+1)k})C(x) - sx^{r+1}\left(1 - x^{(r+1)k} - x(1 - x^{(r+1)(k-1)})C(x)\right)}$$

and

$$\Delta'_k = \frac{1 - x^{(r+1)(k+1)} - sx^{r+1}(1 - x^{(r+1)k})}{1 - x^{(r+1)(k+1)} - x(1 - x^{(r+1)k})C(x) - sx^{r+1}\left(1 - x^{(r+1)k} - x(1 - x^{(r+1)(k-1)})C(x)\right)}.$$

We note that the second of the above formulas has been proved in [7] for $r = 0$.

Example

For $k = 1$ we obtain that

$$\Gamma'_1(x;s) = \frac{C(x)}{1 + (1-s)x^{r+1}C(x)} \text{ and } \Delta'_1(x;s) = \frac{\left(1 + (1-s)x^{r+1}\right)C(x)}{1 + (1-s)x^{r+1}C(x)}.$$

Using the above relations we can easily show that the number of all $u \in \mathcal{D}_n$ with j low $az\bar{a}$'s and the number of all $u \in D_n$ with j low $az\bar{a}a$'s are given respectively by the formulas

$$[x^n s^j]\Gamma'_1 = \sum_{i=0}^{[\frac{n}{r+1}]-j} (-1)^i \frac{(i+j+1)\binom{i+j}{j}\binom{2n-(2r+1)(i+j)+1}{n+1-r(i+j)}}{2n - (2r+1)(i+j) + 1}, \tag{10}$$

for $0 \le j \le [\frac{n}{r+1}]$, and

$$[x^n s^j]\Delta'_1 = \sum_{i=0}^{[\frac{n-1}{r+1}]-j} (-1)^i \frac{(i+j+2)\binom{i+j}{j}\binom{2n-(2r+1)(i+j)-1}{n-r(i+j)}}{n+1-(i+j)r}, \tag{11}$$

for $0 \le j \le [\frac{n-1}{r+1}]$.

Notice that for $r = 0$, relation (10) (respectively (11)) gives that the number of all $u \in \mathcal{D}_n$ with j low peaks (respectively low $a\bar{a}a$'s) is equal to
$\frac{j+1}{n+1}\sum\limits_{i=0}^{n-j}(-1)^{n-j-i}\binom{n-i+1}{j+1}\binom{n+i}{n}$, (respectively $\frac{1}{n+1}\sum\limits_{i=0}^{n-j-1}(-1)^{n-j-i}(n-i+1)\binom{n-i-1}{j}\binom{n+i}{n}$)
thus obtaining a formula equivalent to (6.16) of [3] (respectively to that of Theorem 3.1 in [12]).

For $r = 1$, we find the following triangles whose elements, read by rows, count the number of all $u \in \mathcal{D}_n$ with j low $aa\bar{a}\bar{a}$'s (see 114486 of [11]) and the number of all $u \in \mathcal{D}_n$ with j low $aa\bar{a}\bar{a}a$'s respectively:
1; 1; 1, 1; 3, 2; 10, 3, 1; 31, 8, 3; 98, 27, 6, 1; 321, 88, 16, 4; ... and
1; 1; 2; 4, 1; 11, 3; 34, 7, 1; 108, 20, 4; 352, 65, 11, 1; 1176, 216, 33, 5; ...

2.3 The statistics "number of high $(az\bar{a})^j$'s" and "number of high $(az\bar{a})^j a$'s"

For every $j \in \mathbb{N}^*$ and $u \in \mathcal{D}$, we denote with $\gamma''_j(u)$ (respectively $\delta''_j(u)$) the number of high $(az\bar{a})^j$'s (respectively high $(az\bar{a})^j a$'s) in u.

We consider the generating function of the set of Dyck paths according to the semilength and to the above parameters:

$$F''(x;\mathbf{s},\mathbf{t}) = \sum_{u\in\mathcal{D}} x^{l(u)} \prod_{j\ge1} s_j^{\gamma''_j(u)} \prod_{j\ge1} t_j^{\delta''_j(u)},$$

where $\mathbf{s} = (s_j)$ and $\mathbf{t} = (t_j)$, for $j \in N^*$.

Using the first return decomposition of a non-empty Dyck path $u = aw\bar{a}v$, where $w, v \in \mathcal{D}$, it is easy to see that

$$\gamma''_j(u) = \gamma_j(w) + \gamma''_j(v) \text{ and } \delta''_j(u) = \delta_j(w) + \delta''_j(v).$$

It follows that

$$F''(x;\mathbf{s},\mathbf{t}) = 1 + xF(x;\mathbf{s},\mathbf{t})F''(x;\mathbf{s},\mathbf{t})$$

and thus we have the following result.

Proposition 2.10 *The generating function $F'' = F''(x;\mathbf{s},\mathbf{t})$ is given by*

$$F'' = \frac{1}{1-xF}$$

where F satisfies equation (1).

Using this result, we deduce that the generating function $E''_k(x,t)$ which counts the Dyck paths of prescribed semilength according to the statistics "number of high $(a\bar{a})^k a^2$'s" is given by

$$E''_k(x,t) = \frac{1}{1 - xE_k(x,t)}. \tag{12}$$

Thus, using relation (12) we can expand $E''_k(x,t)$ to a geometric series so that by relation (4) we obtain the following result.

Corollary 2.11 *The number of all Dyck paths of semilength n with j high $(a\bar{a})^k a^2$'s is equal to*

$$[x^n t^j]E''_k = \delta_{0j} + \sum_{m=j}^{\left[\frac{n-2}{k+1}\right]} \sum_{i=(k+1)m+1}^{n-1} (-1)^{m+j} \frac{(n-i)\binom{m}{j}\binom{i-km}{m}\binom{i+n-(2k+1)m-1}{n-km}}{i-km},$$

where δ_{0j} is the Kronecker symbol.

We note that for $k=1$ and $j=0$ we obtain the sequence A086581 of [11], which counts the Dyck paths that avoid $a\bar{a}a^2$ at high level.

For fixed $k \in \mathbb{N}^*$, let $\Gamma''_k(x,s)$ and $\Delta''_k(x,s)$ be the generating functions that count the Dyck paths of prescribed semilength according to the statistics γ''_k and δ''_k respectively. Then, by Proposition 2.10 we deduce that the generating functions $\Gamma''_k = \Gamma''(x,s)$ and $\Delta''_k = \Delta''_k(x,s)$ are given by the relations

$$\Gamma''_k = \frac{1}{1 - x\Gamma_k} \text{ and } \Delta''_k = \frac{1}{1 - x\Delta_k} \tag{13}$$

where Γ_k and Δ_k are given in Corollary 2.5.

Examples

1. For $k=1$ and $r=0$, using relations (5) and (6) we deduce that

$$[x^n s^j]\Gamma''_1 = \sum_{m=j}^{\left[\frac{n-1}{r+1}\right]} \frac{(-1)^{m+j}}{n-rm+1} \binom{m}{j} \binom{n-rm-1}{m} \binom{2n-(2r+1)m}{n-rm}$$

and

$$[x^n s^j]\Delta''_1 = \delta_{0j} + \sum_{m=j}^{\left[\frac{n-2}{r+1}\right]} \sum_{i=0}^{n-2-(r+1)m} (-1)^{m+j} \binom{m}{j} \binom{i+m}{m} B_{n-(r+1)m-1,i+1},$$

where $B_{k,l} = \frac{k-l+1}{k+1}\binom{k+l}{l}$ stands for the well-known double sequence of ballot numbers.

Using formula 3.49 of [5], the first of the above formulas, for $r = 0$ gives that

$$\begin{aligned}[x^n s^j]\Gamma_1'' &= \frac{1}{n+1}\sum_{m=j}^{n-1}(-1)^{m+j}\binom{n-1}{j}\binom{n-1-j}{m-j}\binom{2n-m}{n} \\ &= \frac{1}{n+1}\binom{n-1}{j}\binom{n+1}{j+1} \\ &= \frac{1}{n}\binom{n}{j}\binom{n}{j+1}.\end{aligned}$$

Thus, we obtain the well-known result (e.g. see [3]) that the number of high peaks is counted by the Narayana numbers.

The second of the above formulas has been proved in [12] for $r = 0$.

2. For $k = 2$, $r = 0$ and $s = 0$, we obtain the sequences 1, 1, 2, 4, 10, 26, 72, 206, 606, 1820, 5558, 17206, ... and 1, 1, 2, 5, 13, 37, 110, 338, 1066, 3430, ..., which count the number of all $u \in \mathcal{D}_n$ that avoid high $a\bar{a}a\bar{a}$ and high $a\bar{a}a\bar{a}a$ respectively.

3 The string $a^i\bar{a}a^j$

Throughout this section we deal with the string $\tau = a^i\bar{a}a^j$, where $i, j \in \mathbb{N}^*$.

3.1 The statistic "number of $a^i\bar{a}a^j$'s"

For every $i, j \in \mathbb{N}^*$, let $\rho_{ij}(u)$ denote the number of $a^i\bar{a}a^j$'s in u, and let the corresponding generating function

$$G(x;\mathbf{q}) = \sum_{u\in\mathcal{D}} x^{l(u)} \prod_{i,j\geq 1} q_{ij}^{\rho_{ij}(u)},$$

where $\mathbf{q} = (q_{ij})_{i,j\in\mathbb{N}^*}$.

We have the following result.

Proposition 3.1 *The generating function G satisfies the formula*

$$G(x;\mathbf{q}) = G(x;\mathbf{q}^T),$$

where $\mathbf{q} = (q_{ij})$ and $\mathbf{q}^T = (q_{ji})$.

Proof. It is enough to define an involution ϕ of $\mathcal{D}$ such that

$$l(\phi(u)) = l(u) \text{ and } \rho_{ij}(\phi(u)) = \rho_{ji}(u),$$

for every $i, j \in \mathbb{N}^*$ and $u \in \mathcal{D}$.

For this, we consider the set $\mathcal{P}$ of all paths with no double falls, that start and end with a rise. Every element of $\mathcal{P}$ can be written uniquely in the form

$$p = a^{\xi_1}\bar{a}a^{\xi_2}\bar{a}\cdots a^{\xi_{\rho-1}}\bar{a}a^{\xi_\rho},$$

where $\xi_i \in \mathbb{N}^*$, $\rho \in \mathbb{N}^*$.

We first define an involution ψ of $\mathcal{P}$, by setting

$$\psi(p) = a^{\xi_\rho}\bar{a}a^{\xi_{\rho-1}}\bar{a}\cdots a^{\xi_2}\bar{a}a^{\xi_1}.$$

Clearly, every Dyck path u can be uniquely decomposed in the form

$$u = p_1v_1p_2v_2\cdots p_tv_t,$$

where $p_i \in \mathcal{P}$, $v_i \in \{\bar{a}\}^* \setminus \{\epsilon\}$ and v_i has length at least 2 for every $i \neq t$.

We define $\phi(\epsilon) = \epsilon$, and for a non-empty Dyck path $u = p_1v_1p_2v_2\cdots p_tv_t$,

$$\phi(u) = \psi(p_1)v_1\psi(p_2)v_2\cdots\psi(p_t)v_t.$$

It is easy to check that ϕ is an involution satisfying the required properties. □

Remark From the above Proposition, it follows that the statistics "number of $a^i\bar{a}a^j$'s" and "number of $a^j\bar{a}a^i$'s" are equidistributed, for every $i, j \in \mathbb{N}^*$.

We now come to evaluate the generating function $G = G(x;\mathbf{q})$. For this, we consider the partition $(\mathcal{A}_\nu)_{\nu\in\mathbb{N}}$ of $\mathcal{D}$, where $\mathcal{A}_0 = \{\epsilon\}$, $\mathcal{A}_\nu$ is the set of all Dyck paths with length of the first ascent equal to ν (for $\nu \geq 1$), and the generating function $A_\nu = A_\nu(x;\mathbf{q})$ of the sets $\mathcal{A}_\nu$ according to the parameters l and ρ_{ij}, for every $i, j \in \mathbb{N}^*$.

Obviously,

$$G = \sum_{\nu=0}^{\infty} A_\nu. \tag{14}$$

For every $\nu \in \mathbb{N}^*$ and $\mu \in \mathbb{N}$ we define the function $f_{\nu,\mu} = f_{\nu,\mu}(\mathbf{q})$ with

$$f_{\nu,\mu} = \prod_{\substack{1\leq i\leq \nu\\ 1\leq j\leq \mu}} q_{ij},$$

where $\mathbf{q} = (q_{ij})_{i,j\in\mathbb{N}}$.

Clearly $f_{\nu,0} = 1$, for every $\nu \in \mathbb{N}^*$.

We then have the following result.

Proposition 3.2 *The generating function $A_\nu = A_\nu(x;\mathbf{q})$, where $\nu \in \mathbb{N}^*$, satisfies the following relation*

$$A_\nu = x^\nu G^{\nu-1}\sum_{\mu=0}^{\infty} f_{\nu,\mu}A_\mu. \tag{15}$$

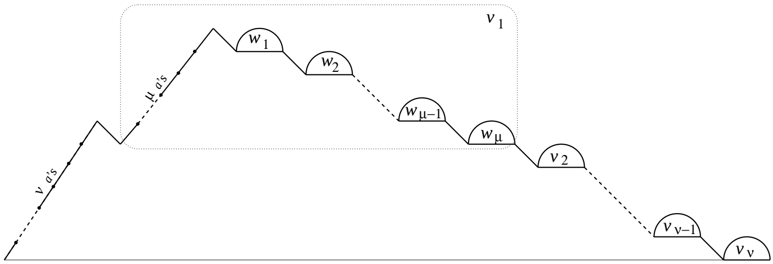

Figure 2 : The elements of $\mathcal{A}_{\mu,\nu}$

Proof. Every set $\mathcal{A}_\nu$ is decomposed into the sets

$$\mathcal{A}_{\nu,\mu} = \{a^\nu \bar{a} v_1 \bar{a} v_2 \bar{a} \cdots v_{\nu-1} \bar{a} v_\nu : v_1 \in \mathcal{A}_\mu, v_i \in \mathcal{D} \text{ for } i \in [2,\nu]\}, \text{ where } \mu \in \mathbb{N};$$

(see Fig. 2).
Hence,

$$\begin{aligned} A_\nu &= \sum_{\mu=0}^{\infty} \sum_{u \in \mathcal{A}_{\nu,\mu}} x^{l(u)} \prod_{i,j \geq 1} q_{ij}^{\rho_{ij}(u)} \\ &= \sum_{\mu=0}^{\infty} \sum_{\substack{v_1 \in \mathcal{A}_\mu \\ v_i \in \mathcal{D}, i \neq 1}} x^{l(v_1)+l(v_2)+\cdots+l(v_\nu)+\nu} \prod_{i,j \geq 1} q_{ij}^{\rho_{ij}(v_1)+\rho_{ij}(v_2)+\cdots+\rho_{ij}(v_\nu)} \prod_{\substack{1 \leq i \leq \nu \\ 1 \leq j \leq \mu}} q_{ij} \\ &= \sum_{\mu=0}^{\infty} \prod_{\substack{1 \leq i \leq \nu \\ 1 \leq j \leq \mu}} q_{ij} x^\nu G^{\nu-1} A_\mu \\ &= x^\nu G^{\nu-1} \sum_{\mu=0}^{\infty} f_{\nu,\mu} A_\mu. \end{aligned}$$

Examples

1. If we set

$$q_{ij} = \begin{cases} 1, & \text{if } j \geq 2 \\ q_i, & \text{if } j = 1, \end{cases}$$

we obtain the equation of Theorem 2.1 in [7], satisfied by the generating function of $\mathcal{D}$ according to the statistics "number of $a^i \bar{a} a$'s", where $i \in \mathbb{N}^*$.

Indeed, in this case we have

$$f_{\nu,\mu} = \prod_{i=1}^{\nu} q_i, \quad \text{for every } \nu, \mu \in \mathbb{N}^*.$$

Hence, by relations (14) and (15) it follows that

$$\begin{aligned} G &= 1 + \sum_{\nu=1}^{\infty} x^{\nu} G^{\nu-1} \left(1 - \prod_{i=1}^{\nu} q_i (1-G)\right) \\ &= \sum_{\nu=0}^{\infty} x^{\nu} G^{\nu} \left(1 + x(1-G)\right) - \sum_{\nu=0}^{\infty} x^{\nu} G^{\nu} x \prod_{i=1}^{\nu+1} q_i (1-G) \\ &= \sum_{\nu=0}^{\infty} x^{\nu} G^{\nu} \left(1 + x\,(1-G)\left(1 - \prod_{i=1}^{\nu+1} q_i\right)\right). \end{aligned}$$

2. For fixed $k, \lambda \in \mathbb{N}^*$, we denote with $T_{k,\lambda} = T_{k,\lambda}(x,q)$ the generating function that counts the Dyck paths of prescribed semilength according to the statistic "number of $a^k \bar{a} a^{\lambda} \bar{a}$'s".

It is clear that $T_{k,\lambda}(x,q) = G(x;\mathbf{q})$ for $q_{k\lambda} = q$, $q_{k\,\lambda+1} = q^{-1}$ and $q_{ij} = 1$ otherwise.

In this case we have that $f_{\nu,\mu} = q$ for $\nu \geq k$ and $\mu = \lambda$, and $f_{\nu,\mu} = 1$ otherwise.

Furthermore, by Proposition 3.2 we obtain that

$$A_{\nu} = x^{\nu} T_{k,\lambda}^{\nu}, \text{ for } \nu < k \tag{16}$$

and

$$A_{\nu} = x^{\nu} T_{k,\lambda}^{\nu} - (1-q) x^{\nu} T_{k,\lambda}^{\nu-1} A_{\lambda}, \text{ for } \nu \geq k. \tag{17}$$

It follows that

$$\begin{aligned} T_{k,\lambda} &= \sum_{\nu=1}^{\infty} x^{\nu} T_{k,\lambda}^{\nu} - (1-q) A_{\lambda} \sum_{\nu=k}^{\infty} x^{\nu} T_{k,\lambda}^{\nu-1} \\ &= \frac{1}{1 - x T_{k,\lambda}} - \frac{(1-q) A_{\lambda} x^k T_{k,\lambda}^{k-1}}{1 - x T_{k,\lambda}} \end{aligned}$$

and hence

$$T_{k,\lambda} = 1 + x T_{k,\lambda}^2 - (1-q) A_{\lambda} x^k T_{k,\lambda}^{k-1}. \tag{18}$$

We consider two cases:

For $\lambda < k$, from relation (16) we have that

$$A_{\lambda} = x^{\lambda} T_{k,\lambda}^{\lambda},$$

so that relation (18) gives

$$T_{k,\lambda} = 1 + x T_{k,\lambda}^2 - (1-q) x^{k+\lambda} T_{k,\lambda}^{k+\lambda-1}. \tag{19}$$

For $\lambda \geq k$, from relation (17) we easily obtain that

$$A_\lambda = \frac{x^\lambda T_{k,\lambda}^\lambda}{1 + (1-q)x^\lambda T_{k,\lambda}^{\lambda-1}},$$

so that relation (18) gives

$$T_{k,\lambda} = 1 + xT_{k,\lambda}^2 - (1-q)\frac{x^{k+\lambda} T_{k,\lambda}^{k+\lambda-1}}{1 + (1-q)x^\lambda T_{k,\lambda}^{\lambda-1}}. \tag{20}$$

For $k = 2$ and $\lambda = 1$ (respectively $k = 1$ and $\lambda = 2$) relation (19) (respectively (20)) gives the triangle, read by rows,
1; 1; 2; 4, 1; 10, 4; 27, 15; 78, 52 2; 234, 180, 15; 722, 624, 84; 2274, 2178, 405, 5; ...
(respectively 1; 1; 2; 4, 1; 10, 4; 28, 13, 1; 83, 42, 7; 254, 141, 33, 1; 795, 489, 135, 11; ...),
which counts the number of $a^2\bar{a}a\bar{a}$'s (respectively $a\bar{a}a^2\bar{a}$'s). Thus, we realize that the statistics "number of $a^k\bar{a}a^\lambda\bar{a}$'s" and "number of $a^\lambda\bar{a}a^k\bar{a}$'s" are not in general equidistributed, as opposed to the statistics "number of $a^k\bar{a}a^\lambda$'s" and "number of $a^\lambda\bar{a}a^k$'s".

In the following, we consider the statistic "number of $a^k\bar{a}a^\lambda$'s" for fixed $k, \lambda \in \mathbb{N}^*$. Clearly, it is enough to restrict ourselves to the case where $k \geq \lambda$.

Proposition 3.3 *The generating function $G_{k,\lambda} = G_{k,\lambda}(x,q)$ that counts the Dyck paths according to the semilength and to the number of $a^k\bar{a}a^\lambda$'s, for fixed $k \geq \lambda$, satisfies the equation*

$$G_{k,\lambda} = 1 + xG_{k,\lambda}^2 - (1-q)x^k G_{k,\lambda}^{k-1}\left(G_{k,\lambda} - \frac{1 - x^\lambda G_{k,\lambda}^\lambda}{1 - xG_{k,\lambda}}\right).$$

Proof. If we set

$$q_{ij} = \begin{cases} q, & \text{if } (i,j) = (k,\lambda) \\ 1, & \text{if } (i,j) \neq (k,\lambda), \end{cases}$$

then $G_{k,\lambda}(x,q) = G(x;\mathbf{q})$ and

$$f_{\nu,\mu} = \begin{cases} q, & \text{if } \nu \geq k \text{ and } \mu \geq \lambda \\ 1, & \text{otherwise.} \end{cases}$$

Hence, by relations (14) and (15), it follows that

$$\begin{aligned} G_{k,\lambda} &= 1 + \sum_{\nu=1}^{k-1} x^{\nu} G_{k,\lambda}^{\nu-1} \left(\sum_{\mu=0}^{\infty} A_{\mu} \right) + \sum_{\nu=k}^{\infty} x^{\nu} G_{k,\lambda}^{\nu-1} \left(\sum_{\mu=0}^{\lambda-1} A_{\mu} + q \sum_{\mu=\lambda}^{\infty} A_{\mu} \right) \\ &= 1 + \sum_{\nu=1}^{k-1} x^{\nu} G_{k,\lambda}^{\nu} + \sum_{\nu=k}^{\infty} x^{\nu} G_{k,\lambda}^{\nu-1} \left(G_{k,\lambda} - (1-q) \left(G_{k,\lambda} - \sum_{\mu=0}^{\lambda-1} A_{\mu} \right) \right) \\ &= \sum_{\nu=0}^{\infty} x^{\nu} G_{k,\lambda}^{\nu} - (1-q) \sum_{\nu=k}^{\infty} x^{\nu} G_{k,\lambda}^{\nu-1} \left(G_{k,\lambda} - \sum_{\mu=0}^{\lambda-1} A_{\mu} \right). \end{aligned}$$

Since $k \geq \lambda$, then

$$\sum_{\mu=0}^{\lambda-1} A_{\mu} = \sum_{\mu=0}^{\lambda-1} x^{\mu} G_{k,\lambda}^{\mu} = \frac{1 - x^{\lambda} G_{k,\lambda}^{\lambda}}{1 - x G_{k,\lambda}},$$

giving the required result. □

For example, if $k = \lambda = 2$, by the above Proposition we obtain that the generating function of $\mathcal{D}$ according to the statistic "number of $a^2\bar{a}a^2$'s" satisfies the equation

$$x(1 + (1-q)x(x-1))G^2 - (1 - (1-q)x^2)G + 1 = 0. \tag{21}$$

The first terms of the corresponding triangle, read by rows, are: 1; 1; 2; 5; 13, 1; 36, 6; 105, 26, 1; 317, 104, 8; 982, 402, 45, 1.

3.2 The statistic "number of low $a^i\bar{a}a^j$'s"

For every $i, j \in \mathbb{N}^*$, let $\rho'_{ij}(u)$ denote the number of low $a^i\bar{a}a^j$'s in u, and let the corresponding generating function

$$G'(x; \mathbf{q}) = \sum_{u \in D} x^{l(u)} \prod_{i,j \geq 1} q_{ij}^{\rho'_{ij}(u)}.$$

For every $\nu \in \mathbb{N}^*$ and $\mu \in \mathbb{N}$ we define the function $g_{\nu,\mu} = g_{\nu,\mu}(\mathbf{q})$ with

$$g_{\nu,\mu} = \prod_{1 \leq j \leq \mu} q_{\nu j}.$$

Clearly, $g_{\nu,0} = 1$, for every $\nu \in \mathbb{N}^*$.

We have the following result.

Proposition 3.4 *The generating function $G' = G'(x; \mathbf{q})$ is given by the formula*

$$G'(x; \mathbf{q}) = \frac{1 + x(1 - q_{11})}{1 - x q_{11} - \sum_{\nu=2}^{\infty} \sum_{\mu=0}^{\infty} g_{\nu,\mu} (x(g_{1,\nu} - q_{11}) + 1) x^{\nu+\mu} C^{\nu+\mu-2}(x)}.$$

Proof. For every $\nu \in \mathbb{N}^*$ and $u \in \mathcal{A}_\nu$, such that $u = a^\nu \bar{a} v_1 \bar{a} v_2 \cdots \bar{a} v_\nu$ we have that

$$\rho'_{ij}(u) = \begin{cases} \rho'_{ij}(v_\nu), & \text{if } i \neq \nu \text{ or } v_1 \in \bigcup_{\mu=0}^{j-1} \mathcal{A}_\mu \\ \rho'_{\nu j}(v_\nu) + 1, & \text{if } i = \nu \text{ and } v_1 \in \bigcup_{\mu=j}^{\infty} \mathcal{A}_\mu. \end{cases}$$

Using the above relations, we will find the generating functions $A'_\nu = A'_\nu(x, \mathbf{q})$ of the sets $\mathcal{A}_\nu$ according to the parameters l and ρ'_{ij}, for every $i, j \in \mathbb{N}^*$.
Indeed,

$$\begin{aligned} A'_1 &= x + \sum_{\nu=1}^{\infty} \sum_{v_1 \in \mathcal{A}_\nu} x^{l(v_1)+1} \prod_{i,j \geq 1} q_{ij}^{\rho'_{ij}(v_1)} \prod_{1 \leq j \leq \nu} q_{1j} \\ &= x + x \sum_{\nu=1}^{\infty} g_{1,\nu} A'_\nu \\ &= x + x q_{11} A'_1 + x \sum_{\nu=2}^{\infty} g_{1,\nu} A'_\nu \end{aligned}$$

and hence,

$$A'_1 = \frac{x + x \sum_{\nu=2}^{\infty} g_{1,\nu} A'_\nu}{1 - x q_{11}}. \tag{22}$$

Furthermore, for $\nu \geq 2$ we have

$$\begin{aligned} A'_\nu &= \sum_{\mu=0}^{\infty} \sum_{\substack{v_1 \in \mathcal{A}_\mu \\ v_i \in \mathcal{D}, i \geq 2}} x^{l(v_1)+l(v_2)+\cdots l(v_\nu)+\nu} \prod_{i,j \geq 1} q_{ij}^{\rho'_{ij}(v_\nu)} \prod_{1 \leq j \leq \mu} q_{\nu j} \\ &= x^\nu C^{\nu-2}(x) G' \sum_{\mu=0}^{\infty} g_{\nu,\mu} \sum_{v_1 \in \mathcal{A}_\mu} x^{l(v_1)} \\ &= x^\nu C^{\nu-2}(x) G' \sum_{\mu=0}^{\infty} g_{\nu,\mu} x^\mu C^\mu(x). \end{aligned} \tag{23}$$

Finally, since

$$G' = 1 + \sum_{\nu=1}^{\infty} A'_\nu,$$

from relations (22) and (23) we obtain the required result. □

Example

For fixed $k, \lambda \in \mathbb{N}^*$, we denote with $T'_{k,\lambda} = T'_{k,\lambda}(x, q)$ the generating function that counts the Dyck paths of prescribed semilength according to the statistic "number of low $a^k \bar{a} a^\lambda \bar{a}$'s".

It is clear that $T'_{k,\lambda}(x,q) = G'(x;\mathbf{q})$ for $q_{k\lambda} = q$, $q_{k\ \lambda+1} = q^{-1}$ and $q_{ij} = 1$ otherwise. In this case we have $g_{\nu,\mu} = q$ for $\nu = k$ and $\mu = \lambda$, and $g_{\nu,\mu} = 1$ otherwise.
Furthermore, using Proposition 3.4 we evaluate the generating function $T'_{k,\lambda}$, considering the following cases:
For $k \geq 2$,

$$\begin{aligned} T'_{k,\lambda} &= \frac{1}{1 - x - \sum_{\nu=2}^{\infty}\sum_{\mu=0}^{\infty} x^{\nu+\mu}C^{\nu+\mu-2}(x) + (1-q)x^{k+\lambda}C^{k+\lambda-2}(x)} \\ &= \frac{1}{1 - x - x^2C^2(x) + (1-q)x^{k+\lambda}C^{k+\lambda-2}(x)} \\ &= \frac{C(x)}{1 + (1-q)x^{k+\lambda}C^{k+\lambda-1}(x)}. \end{aligned}$$

For $k = 1$ and $\lambda \geq 2$,

$$\begin{aligned} T'_{1,\lambda} &= \frac{1}{1 - x - \sum_{\nu=2}^{\infty}(x(g_{1,\nu} - 1) + 1)x^{\nu}C^{\nu-2}(x)\sum_{\mu=0}^{\infty} g_{\nu,\mu}x^{\mu}C^{\mu}(x)} \\ &= \frac{1}{1 - x - \sum_{\nu=2}^{\infty}(x(g_{1,\nu} - 1) + 1)x^{\nu}C^{\nu-1}(x)} \\ &= \frac{1}{1 - x - \sum_{\nu=2}^{\infty} x^{\nu}C^{\nu-1}(x) - x(q-1)x^{\lambda}C^{\lambda-1}(x)} \\ &= \frac{C(x)}{1 + (1-q)x^{\lambda+1}C^{\lambda}(x)}. \end{aligned}$$

Finally, for $k = \lambda = 1$,

$$T'_{1,1} = \frac{1 + x(1-q)}{1 - xq - (x(1-q) + 1)x^2C^2(x)}.$$

In the following, we consider the statistic "number of low $a^k\bar{a}a^{\lambda}$'s" for fixed $k, \lambda \in \mathbb{N}^*$, with generating function $G'_{k,\lambda} = G'_{k,\lambda}(x,q)$.
A simple application of Proposition 3.4 for $q_{11} = q$ and $q_{ij} = 1$ otherwise, gives

$$G'_{1,1}(x,q) = 1 + \frac{xC(x)}{1 + x(1-q) - xC(x)}$$

which has also been proved in [12].
For $k \geq 2$ and $\lambda = 1$ the generating function $G'_{k,1}$ has been evaluated in [7].
In the following result we evaluate $G'_{k,\lambda}$ for $(k,\lambda) \neq (1,1)$.

Proposition 3.5 *The generating function of $\mathcal{D}$ according to the semilength and to the number of low $a^k\bar{a}a^\lambda$'s, for $(k,\lambda) \neq (1,1)$, is given by the formula*

$$G'_{k,\lambda} = \frac{C(x)}{1+(1-q)(xC(x))^{k+\lambda}}.$$

Proof. We first consider the case where $k \geq 2$ and $\lambda \geq 1$; in this case, applying Proposition 3.4 for $q_{k\lambda} = q$ and $q_{ij} = 1$ otherwise, we have that $g_{\nu,\mu} = q$ if $\nu = k$ and $\mu \geq \lambda$, and $g_{\nu,\mu} = 1$ otherwise.
Thus we have

$$\begin{aligned} G'_{k,\lambda} &= \frac{1}{1-x-\sum\limits_{\nu=2}^{\infty}\sum\limits_{\mu=0}^{\infty} g_{\nu,\mu}x^{\nu+\mu}C^{\nu+\mu-2}(x)} \\ &= \frac{1}{1-x-\sum\limits_{\nu=2}^{\infty}\sum\limits_{\mu=0}^{\infty} x^{\nu+\mu}C^{\nu+\mu-2}(x)+(1-q)\sum\limits_{\mu=\lambda}^{\infty} x^{k+\mu}C^{k+\mu-2}(x)} \\ &= \frac{1}{1-x-x^2C^2(x)+(1-q)x^{k+\lambda}C^{k+\lambda-1}(x)} \\ &= \frac{C(x)}{1+(1-q)x^{k+\lambda}C^{k+\lambda}(x)}. \end{aligned}$$

The case $k = 1$ and $\lambda \geq 2$ is treated similarly. □

Furthermore, after some simple manipulations we obtain the following result.

Corollary 3.6 *The number of all Dyck paths of semilength n with j low $a^k\bar{a}a^\lambda$'s, for $(k,\lambda) \neq (1,1)$, is equal to*

$$[x^nq^j]G'_{k,\lambda} = \sum_{i=0}^{[\frac{n}{k+\lambda}]-j} (-1)^i \frac{(j+i)(k+\lambda)+1}{2n+1-(j+i)(k+\lambda)} \binom{j+i}{i} \binom{2n+1-(j+i)(k+\lambda)}{n+1},$$

where $0 \leq j \leq [\frac{n}{k+\lambda}]$.

3.3 The statistic "number of high $a^i\bar{a}a^j$'s"

For every $i, j \in \mathbb{N}^*$, let $\rho''_{ij}(u)$ denote the number of high $a^i\bar{a}a^j$'s in u, and let the corresponding generating function

$$G''(x;\mathbf{q}) = \sum_{u\in\mathcal{D}} x^{l(u)} \prod_{i,j\geq 1} q_{ij}^{\rho''_{ij}(u)}.$$

Using the first return decomposition $u = aw\bar{a}v$, where $w, v \in D$, it is easy to see that

$$\rho''_{ij}(u) = \rho_{ij}(w) + \rho''_{ij}(v).$$

It follows that

$$G''(x;\mathbf{q}) = 1 + xG(x;\mathbf{q})G''(x;\mathbf{q}).$$

Thus, we have the following result.

Proposition 3.7 *The generating function $G'' = G''(x;\mathbf{q})$ is given by*

$$G'' = \frac{1}{1 - xG}$$

where G satisfies relations (14) *and* (15).

Example

Using the above Proposition and relation (21), we obtain that the generating function $G'' = G''(x,q)$ of $\mathcal{D}$, according to the statistic "number of high $a^2\bar{a}a^2$'s" satisfies equation

$$(qx + 2x^2 - 2qx^2){G''}^2 - (1 - 2x + 2qx + 3x^2 - 3qx^2)G'' + 1 - x + qx + x^2 - qx^2 = 0.$$

The first terms of the corresponding triangle, read by rows, are: 1; 1; 2; 5; 14; 41, 1; 124, 8; 385, 43, 1; 1220, 200, 10; 3929, 866, 66, 1.

References

[1] M. Aigner, Motzkin numbers, *European J. Combin.* **19** (1998), 663–675.

[2] D. Callan, Two bijections for Dyck paths parameters, *Preprint* (2004), 4pp, `http://www.arxiv.org/abs/math.CO/0406381`.

[3] E. Deutsch, Dyck path enumeration, *Discrete Math.* **204** (1999), 167–202.

[4] R. Donaghey and L.W. Shapiro, Motzkin numbers, *J. Combin. Theory A* **23** (1977), 291–301.

[5] H. Gould, *Combinatorial Identities*, Morgantown, 1972.

[6] T. Mansour, Counting peaks at height k in a Dyck path, *J. Integer Seq.* **5** (2002), Article 02.1.1.

[7] T. Mansour, Statistics on Dyck paths, *J. Integer. Seq.* **9** (2006), Article 06.1.5.

[8] D. Merlini, R. Sprungoli and M. Verri, Some statistics on Dyck paths, *J. Statist. Plann. and Infer.* **101** (2002), 211–227.

[9] A. Sapounakis and P. Tsikouras, Counting peaks and valleys in k-colored Motzkin paths, *Electronic J. Combin.*, **12** (2005), R16.

[10] A. Sapounakis, I. Tasoulas and P. Tsikouras, Counting strings in Dyck paths, *Discrete Math.*, in press.

[11] N. J. A. Sloane, Online Encyclopedia of Integer Sequences, published electronically at `http://www.research.att.com/~njas/sequences/`.

[12] Y. Sun, The statistic "number of udu's" in Dyck paths, *Discrete Math.* **237** (2004), 177–186.

(Received 13 July 2006; revised 17 Apr 2007)