Discrete Mathematics and Theoretical Computer Science DMTCS vol. **15:2**, 2013, 1–12

Descent variation of samples of geometric random variables

Charlotte Brennan† Arnold Knopfmacher‡

The John Knopfmacher Centre for Applicable Analysis and Number Theory, School of Mathematics, University of the Witwatersrand, Johannesburg, South Africa.

received 24th February 2011, revised 15th February 2013, 5th March 2013, accepted 21st March 2013.

In this paper, we consider random words $\omega_1\omega_2\omega_3\cdots\omega_n$ of length n, where the letters $\omega_i \in \mathbb{N}$ are independently generated with a geometric probability such that $\mathbb{P}\{\omega_i = k\} = pq^{k-1}$ where $p + q = 1$. We have a descent at position i whenever $\omega_{i+1} < \omega_i$. The size of such a descent is $\omega_i - \omega_{i+1}$ and the descent variation is the sum of all the descent sizes for that word. We study various types of random words over the infinite alphabet $\mathbb{N}$, where the letters have geometric probabilities, and find the probability generating functions for descent variation of such words.

Keywords: words, q-enumeration, generating functions, up-down words, Carlitz words, descents

1 Introduction

Recently Cristea and Prodinger [3] considered words $\omega = \omega_1\omega_2\ldots\omega_n$ over the infinite alphabet $\mathbb{N}$, where the letters occur independently with the geometric probability

$$\mathbb{P}(\omega_i = j) = pq^{j-1},$$

for all $n, i, j \in \mathbb{N}$ with $i \leq n$, where $0 < p, q < 1$ and $p + q = 1$. They defined the total variation to be

$$u(\omega) = \sum_{i=1}^{n-1} |\omega_{i+1} - \omega_i|.$$

They investigated the probability generating function of a word of given length n with a total variation $u(\omega) = r$.

Three different families of words over $\mathbb{N}$ were studied:

- the general case, namely all words where the letters occur with a geometric probability with no restriction,

†Email: `Charlotte.Brennan@wits.ac.za`
‡Email: `Arnold.Knopfmacher@wits.ac.za`

- words with an up-down pattern $<\geq<\geq\cdots<\geq$,
- Carlitz words where any two adjacent letters must be distinct.

Previously, several authors had studied the statistic total variation in other contexts. For more information, see the paper [7] by Mansour.

In our paper, we also consider words over $\mathbb{N}$ where the letters are independently generated with the same geometric probability. However, we study the statistic "descent variation". We say we have a descent at position i whenever $\omega_{i+1} < \omega_i$. The size of such a descent is the difference $\omega_i - \omega_{i+1}$. We study the sum of all such descent sizes for the whole word. We call this sum, the descent variation. We concentrate our study on the same three families of words as described above. In Section 2, we look at the general case. In Section 3, we look at the word with an up-down pattern $<\geq<\geq\cdots<\geq$ by first considering word of odd length then words of even length. Finally, in Section 4, we look at the descent variation of Carlitz words, where no two adjacent letters have equal value.

We use the "adding a slice" method to find the generating functions, this method has been used in [5]. The methodology is that we start with a word of length n ending with a letter with a specified value i and add an extra letter of value j at the end of the word. This leads to functional equations which can be solved to deduce the generating functions.

2 The general case: geometrically distributed words over $\mathbb{N}$

Let $\mathcal{W}_n$ denote the set of words of length n with alphabet $\mathbb{N}$ and where $n \geq 1$. For $\omega \in \mathcal{W}_n$, $n \geq 1$, $\omega = \omega_1\omega_2\ldots\omega_n$, we define

i) *a descent* at position i whenever $\omega_i > \omega_{i+1}$,

ii) *a descent size*, as the difference $\omega_i - \omega_{i+1}$ whenever there is a descent and

iii) *the descent variation* of the word ω, as the sum of all the descent sizes.

Our aim is to find, for given $n \geq 1$ and $r \geq 0$, the probability that a word $\omega \in \mathcal{W}_n$ has descent variation $u(\omega) = r$.

We introduce the following generating functions. Let

$$f(z,x) = \sum_{n\geq 1, r\geq 0} p(n,r) z^n x^r, \tag{2.1}$$

where $p(n,r)$ is the probability that a word of length n has descent variation r. Moreover, let

$$F(z,x,t) = \sum_{n\geq 1, r\geq 0, i\geq 1} g(n,r,i) z^n x^r t^i, \tag{2.2}$$

where $g(n,r,i)$ is the occurrence probability of a word of length n with last letter i ($\omega_n = i$) and descent variation r. We have $f(z,x) = F(z,x,1)$.

In order to find the generating function of words with n letters and descent variation r, we proceed as follows.

Suppose a random word ω of length n has an arbitrary last letter with fixed value i and descent variation r,

$$\omega = \omega_1\omega_2\ldots\omega_{n-1}i,$$

where $r \geq 0$ and $i \geq 1$. We add a letter, j, at the end of the word ω, after the last letter i. A new word ω' of length $n+1$ is obtained,

$$\omega' = \omega_1\omega_2\ldots\omega_{n-1}ij.$$

We have two cases depending on the value of j relative to i:

- if $j > i$, there is no new descent and therefore the descent variation remains the same r, or
- if $j \leq i$, we have one more descent, thus the descent variation is now $r+i-j$.

The following two cases with the expression (2.3) represent the changes that adding the extra letter induces in the generating function $F(z,x,t)$.

$$\sum_{j=1}^{i-1} pq^{j-1}\,x^{i-j}\,t^j + \sum_{j\geq i} pq^{j-1}\,t^j = \frac{pt}{x-qt}x^i + \frac{pt\,(x-1)}{(x-qt)\,(1-qt)}(qt)^i. \tag{2.3}$$

According to the "adding a slice" method, see [5], the process of going from a word of length n with last letter i to a word of length $n+1$ with last letter j is therefore achieved by making the following substitution in our generating function

$$t^i \longrightarrow \frac{pt}{x-qt}x^i + \frac{pt\,(x-1)}{(x-qt)\,(1-qt)}(qt)^i. \tag{2.4}$$

Let $a_n(x,t)$, $n \geq 1$, be the function with the property

$$F(z,x,t) = \sum_{n\geq 1} a_n(x,t)z^n. \tag{2.5}$$

For given $n \geq 1$, using the substitution given in (2.4) we obtain

$$a_{n+1}(x,t) = \frac{pt}{x-qt}a_n(x,x) + \frac{pt\,(x-1)}{(x-qt)\,(1-qt)}a_n(x,qt), \tag{2.6}$$

which, after multiplication by z^{n+1} and then summing up for $n \geq 1$, leads to

$$F(z,x,t) = a_1(x,t)z + z\frac{pt}{x-qt}F(z,x,x) + z\frac{pt\,(x-1)}{(x-qt)\,(1-qt)}F(z,x,qt), \tag{2.7}$$

where $a_1(x,t) = \sum_{k\geq 1} pq^{k-1}t^k = \frac{pt}{1-qt}$. This can now be solved using the following elementary lemma:

Lemma 1 *A recursion of the form*

$$F(z,t)=f(z,t)+g(z,t)F(z,tq)$$

has solution

$$F(z,t)=\sum_{j\geq 1} f(z,tq^{j-1})\prod_{i=1}^{j-1} g(z,tq^{i-1}).$$

If we use Lemma 1 with

$$f(z,t)=\frac{ptz}{1-qt}+\frac{ptz}{x-qt}F(z,x,x)\quad \text{and} \quad g(z,t)=\frac{ptz\,(x-1)}{(x-qt)\,(1-qt)},$$

we obtain

$$F(z,x,t)=\sum_{j\geq 1}\left(\frac{ptzq^{j-1}}{1-tq^j}+\frac{ptzq^{j-1}}{x-tq^j}F(z,x,x)\right)\cdot\prod_{i=1}^{j-1}\frac{ptzq^{i-1}(x-1)}{(x-tq^i)(1-tq^i)}. \tag{2.8}$$

For $t=x$, we get

$$F(z,x,x)=\sum_{j\geq 1}\left(\frac{pxzq^{j-1}}{1-xq^j}+\frac{pzq^{j-1}}{1-q^j}F(z,x,x)\right)\cdot\prod_{i=1}^{j-1}\frac{pzq^{i-1}(x-1)}{(1-q^i)(1-xq^i)},$$

and thus

$$F(z,x,x)=\frac{\sum_{j\geq 1}\frac{(zp)^j x(x-1)^{j-1}q^{\binom{j}{2}}}{(q;q)_{j-1}(xq;q)_j}}{1-\sum_{j\geq 1}\frac{(zp)^j(x-1)^{j-1}q^{\binom{j}{2}}}{(q;q)_j(xq;q)_{j-1}}},$$

where $(x;q)_n:=\prod_{j=0}^{n-1}(1-xq^j)$ for $n\geq 1$, and $(x;q)_0=1$.
Substituting the above expression for $F(z,x,x)$ into (2.8) we obtain

$$F(z,x,t)=\sum_{j\geq 1}\left(\frac{pztq^{j-1}}{1-tq^j}+\frac{zptq^{j-1}}{x-tq^j}\;\frac{\sum_{k\geq 1}\frac{(zp)^k x(x-1)^{k-1}q^{\binom{k}{2}}}{(q;q)_{k-1}(xq;q)_k}}{1-\sum_{k\geq 1}\frac{(zp)^k(x-1)^{k-1}q^{\binom{k}{2}}}{(q;q)_k(xq;q)_{k-1}}}\right)\cdot\frac{\Big(zpt(x-1)\Big)^{j-1}q^{\binom{j-1}{2}}}{x^{j-1}(\frac{tq}{x};q)_{j-1}(tq;q)_{j-1}}.$$

For $t=1$ we obtain

$$\begin{aligned} f(z,x)&=F(z,x,1)\\ &=\sum_{j\geq 1}\left(\frac{pzq^{j-1}}{1-q^j}+\frac{zpq^{j-1}}{x-q^j}\;\frac{\sum_{k\geq 1}\frac{(zp)^k x(x-1)^{k-1}q^{\binom{k}{2}}}{(q;q)_{k-1}(xq;q)_k}}{1-\sum_{k\geq 1}\frac{(zp)^k(x-1)^{k-1}q^{\binom{k}{2}}}{(q;q)_k(xq;q)_{k-1}}}\right)\cdot\frac{\Big(zp(x-1)\Big)^{j-1}q^{\binom{j-1}{2}}}{x^{j-1}(\frac{q}{x};q)_{j-1}(q;q)_{j-1}}. \end{aligned} \tag{2.9}$$

We can test our formula, by putting $x = 1$:

$$f(z,1) = F(z,1,1) = \frac{z}{1-z},$$

which corresponds to the case when neither the last letter nor the descent variation of the words are considered.

Furthermore,

$$\left.\frac{\partial f(x,z)}{\partial x}\right|_{x=1} = \frac{qz^2}{(1-q^2)(1-z)^2}.$$

Thus the mean descent variation of a word with n letters is

$$[z^n]\left.\frac{\partial f(x,z)}{\partial x}\right|_{x=1} = [z^n]\frac{qz^2}{(1-q^2)(1-z)^2} = \frac{q}{1-q^2}(n-1).$$

Cristea and Prodinger in [3] found twice that value for the mean total variation, which is to be expected by symmetry. However, the variance of the descent variation and of the total variation are quite different. We obtain

$$\left.\frac{\partial^2 f(x,z)}{\partial x^2}\right|_{x=1} = \frac{2q^2z^2\Big(-1+q(-2+z)-q^2(2-2z+z^2)+q^3(-1+z)+z-z^2\Big)}{(1-q^2)^2(1+q+q^2)(-1+z)^3},$$

where

$$[z^n]\left.\frac{\partial^2 f(x,z)}{\partial x^2}\right|_{x=1} = \begin{cases} \frac{q^2\left(n^2(1+q+q^2)+n(-3+q-q^2+2q^3)-2(-2+q-q^2+q^3)\right)}{(1-q^2)^2(1+q+q^2)}, & \text{for } n>1 \\ 0, & \text{for } n=1. \end{cases}$$

The variance for the descent variation is therefore

$$\frac{nq\left(q^4+3q^2+1\right)}{\left(q^2-1\right)^2\left(q^2+q+1\right)} - \frac{q\left(q^2-q+1\right)^2}{\left(q^2-1\right)^2\left(q^2+q+1\right)}, \quad \text{for } n>1.$$

3 Up-down words over $\mathbb{N}$ following the pattern $<\geq<\geq\cdots<\geq$

3.1 Up-down words of odd length

The aim of this subsection is to find the generating function of up-down words with pattern $<\geq<\geq\cdots<\geq$ with odd length and given descent variation. Corresponding generating functions for such words by total variation and by the number of rises have been found in [3, 2]. Note, these are not probability generating functions as we still need to normalize the coefficients of z^{2n+1} by dividing by the total probability of having alternating words of length $2n+1$.

The methodology is the same as for the general case in the previous section, we introduce the following generating functions. Let

$$f(z,x) = \sum_{n\geq 0, r\geq 0} p(2n+1,r) z^{2n+1} x^r, \tag{3.1}$$

where $p(2n+1,r)$ is the probability that a word of length $2n+1$ has descent variation r, and let

$$F(z,x,t) = \sum_{n\geq 0, r\geq 0, i\geq 1} g(2n+1,r,i) z^{2n+1} x^r t^i = \sum_{n\geq 0} a_{2n+1}(x,t) z^{2n+1} \tag{3.2}$$

where $g(2n+1,r,i)$ is the occurrence probability of a word of length $2n+1$, with last letter i ($\omega_{2n+1} = i$) and descent variation r. Again, $f(z,x) = F(z,x,1)$.

We use the "adding a slice" method but this time we need to add two letters at a time. We proceed inductively. Assume that we have a random word $\omega = \omega_1 \dots \omega_{2n} i$, of odd length $2n+1$ with last letter i, i.e., $w_{2n+1} = i$ and $w_{2n} \geq i$ and descent variation r, with arbitrarily fixed $i \geq 1$ and $r > 0$. We add two letters, k and j, after the last letter i. A new word, w' of odd length $2n+3$ is obtained,

$$\omega' = \omega_1 \dots \omega_{2n} ikj.$$

To carry on the up-down pattern, we need $i < k$, and $k \geq j$. The descent variation of ω' is now $r + k - j$. The following expression summarises the situation

$$\sum_{k>i} pq^{k-1} \sum_{j\leq k} pq^{j-1} \cdot t^j \cdot x^{k-j} = \frac{p^2 xt(xq)^i}{(x-qt)(1-qx)} - \frac{p^2t^2q(q^2t)^i}{(x-qt)(1-q^2t)}. \tag{3.3}$$

Again, using the "adding a slice" method, this leads to the following substitution

$$t^i \longrightarrow \frac{p^2 t}{x-qt} \cdot \left(\frac{x}{1-qx}(qx)^i - \frac{qt}{1-q^2t}(q^2t)^i \right), \tag{3.4}$$

or, equivalently,

$$t^i \longrightarrow A(x,t) \left(\frac{x}{(1-qx)}(qx)^i - B(t) \cdot (q^2t)^i \right), \tag{3.5}$$

where $A(x,t) = \frac{p^2 t}{x-qt}$ and $B(t) = \frac{qt}{1-q^2t}$. We therefore obtain, for $n \geq 0$,

$$a_{2n+3}(x,t) = A(x,t) \left(\frac{x}{1-qx}\, a_{2n+1}(x,qx) - B(t)\, a_{2n+1}(x,q^2t) \right), \tag{3.6}$$

which, after multiplication by z^{2n+3} and then summing up for $n \geq 0$, leads to

$$F(z,x,t) = a_1(x,t)z + A(x,t)z^2 \left(\frac{x}{1-qx} F(z,x,qx) - B(t) F(z,x,q^2t) \right), \tag{3.7}$$

with $a_1(x,t) = \sum_{k\geq 1} pq^{k-1}t^k = \frac{pt}{1-qt}$. We solve the functional equation using Lemma 1, where q is replaced by q^2. Thus

$$F(z,x,t)$$
$$= \sum_{k\geq 1} \left(\frac{ptq^{2k-2}z}{1-tq^{2k-1}} + A(x,tq^{2k-2}) \frac{z^2 x}{1-qx} F(z,x,qx) \right) (-1)^{k-1} z^{2(k-1)} \prod_{i=1}^{k-1} A(x,tq^{2i-2})\, B(tq^{2i-2}).$$

In order to obtain $F(z,x,qx)$, let $t=qx$

$$F(z,x,qx)$$
$$=\sum_{k\geq1}\left(\frac{pxq^{2k-1}z}{1-xq^{2k}}+A(x,xq^{2k-1})\frac{z^2x}{1-qx}F(z,x,xq)\right)(-1)^{k-1}z^{2(k-1)}\prod_{i=1}^{k-1}A(x,xq^{2i-1})B(xq^{2i-1})$$

which yields

$$F(z,x,qx)=\sum_{k\geq1}\left(\frac{pxq^{2k-1}z^{2k-1}(-1)^{k-1}}{1-xq^{2k}}\prod_{i=1}^{k-1}A(x,xq^{2i-1})B(xq^{2i-1})\right)\Big/$$
$$\left(1-\sum_{k\geq1}\frac{A(x,xq^{2k-1})x(-1)^{k-1}z^{2k}}{1-qx}\prod_{i=1}^{k-1}A(x,xq^{2i-1})B(xq^{2i-1})\right).$$

We have

$$\prod_{i=1}^{k-1}A(x,xq^{2i-1})=\prod_{i=1}^{k-1}\frac{p^2q^{2i-1}}{1-q^{2i}}=\frac{p^{2(k-1)}q^{(k-1)^2}}{(q^2;q^2)_{k-1}}$$

and

$$\prod_{i=1}^{k-1}B(xq^{2i-1})=\prod_{i=1}^{k-1}\frac{xq^{2i}}{1-xq^{2i+1}}=\frac{x^{k-1}q^{k(k-1)}}{(xq^3;q^2)_{k-1}},$$

and thus

$$\prod_{i=1}^{k-1}A(x,xq^{2i-1})B(xq^{2i-1})=\frac{p^{2(k-1)}q^{(2k-1)(k-1)}x^{k-1}}{(q^2;q^2)_{k-1}(xq^3;q^2)_{k-1}}.$$

Moreover, since

$$\prod_{i=1}^{k-1}A(x,tq^{2(i-1)})=\prod_{i=1}^{k-1}\frac{p^2tq^{2(i-1)}}{(x-tq^{2i-1})}=\frac{(p^2t)^{k-1}q^{(k-2)(k-1)}}{x^{k-1}(\frac{tq}{x};q^2)_{k-1}}$$

and

$$\prod_{i=1}^{k-1}B(tq^{2(i-1)})=\prod_{i=1}^{k-1}\frac{tq^{2i-1}}{1-tq^{2i}}=\frac{t^{k-1}q^{(k-1)^2}}{(tq^2;q^2)_{k-1}},$$

we have

$$\prod_{i=1}^{k-1}A(x,tq^{2(i-1)})B(tq^{2(i-1)})=\frac{(p^2t^2)^{k-1}q^{(2k-3)(k-1)}}{x^{k-1}(\frac{tq}{x};q^2)_{k-1}(tq^2;q^2)_{k-1}}.$$

These finally yield an expression for $F(z,x,t)$:

$$\begin{aligned}
&F(z,x,t)\\
&=\sum_{n\geq 1}\left(\frac{p^2tq^{2n-2}z^2x}{(x-tq^{2n-1})(1-qx)}\cdot\frac{\sum_{k\geq 1}\frac{(-1)^{k-1}p^{2k-1}q^{k(2k-1)}x^kz^{2k-1}}{(1-xq^{2k})(q^2;q^2)_{k-1}(xq^3;q^2)_{k-1}}}{1-\sum_{k\geq 1}\frac{(-1)^{k-1}p^{2k}q^{k(2k-1)}x^kz^{2k}}{(q^2;q^2)_k(qx;q^2)_k}}+\frac{ptq^{2(n-1)}z}{1-tq^{2n-1}}\right)\\
&\quad\times\frac{(-1)^{n-1}(p^2t^2z^2)^{n-1}q^{(2n-3)(n-1)}}{x^{n-1}(\frac{qt}{x};q^2)_{n-1}(tq^2;q^2)_{n-1}}.
\end{aligned}$$

Setting $t=1$ gives the sought generating function for the descent variation,

$$\begin{aligned}
f(z,x)&=F(z,x,1)\\
&=\sum_{n\geq 1}\left(\frac{p^2q^{2n-2}z^2x}{(x-q^{2n-1})(1-qx)}\cdot\frac{\sum_{k\geq 1}\frac{(-1)^{k-1}p^{2k-1}q^{k(2k-1)}x^kz^{2k-1}}{(1-xq^{2k})(q^2;q^2)_{k-1}(xq^3;q^2)_{k-1}}}{1-\sum_{k\geq 1}\frac{(-1)^{k-1}p^{2k}q^{k(2k-1)}x^kz^{2k}}{(q^2;q^2)_k(qx;q^2)_k}}+\frac{pq^{2(n-1)}z}{1-q^{2n-1}}\right)\\
&\quad\times\frac{(-1)^{n-1}(p^2z^2)^{n-1}q^{(2n-3)(n-1)}}{x^{n-1}(\frac{q}{x};q^2)_{n-1}(q^2;q^2)_{n-1}}.
\end{aligned}\tag{3.8}$$

In order to compute probabilities, the coefficients of z^{2n+1} in $f(z,x)$ must be normalized by dividing by the coefficients of z^{2n+1} in $f(z,1)$.

Remarks As a simple check of (3.8), a series expansion of $f(z,1)$ and $f(z,x)$, for $p=q=1/2$ give $2/7$ as the probability that a three letter word is up-down, and $34/63$ for the mean descent variation. These values are easily verified by direct calculation. For example, the mean descent variation in the case of a three letter word is evaluated by the sum

$$\sum_{i=1}^{\infty}pq^{i-1}\sum_{k=i+1}^{\infty}pq^{k-1}\sum_{j=1}^{k}(k-j)pq^{j-1}.$$

By contrast, direct calculation gives a value of $8/7$ for the mean total variation in this case, for $p=q=1/2$. However, this does not correspond to a series expansion of the generating function in [3], indicating that the result there contains errors. Similarly, we find that the total variation generating function for up-down words of even length in [3] is incorrect.

3.2 Up-down words of even length

In this subsection, we are finding the generating functions of up-down words of even length with a given descent variation.

Assume we have a random word $\omega'=\omega_1\ldots\omega_{2n+1}l$ of even length $2n+2$, with last letter l, and descent variation r, with arbitrarily fixed $l\geq 1$ and $r>0$. We take a word of odd length $2n+1$ ending with j and add the letter l. We use and adapt the result we had for a word of odd length in (3.3). Then according to the pattern, we need $l>j$, and therefore the descent variation of ω' is $r+l-j$. Thus we have the following expression

$$\left(\sum_{k>i}p\,q^{k-1}\sum_{j\leq k}p\,q^{j-1}\,x^{k-j}\right)\sum_{l>j}p\,q^{l-1}z\,t^l.\tag{3.9}$$

We introduce the generating function

$$g(z,x) := \sum_{n\geq 0, r\geq 0} p(2n+2,r) z^{2n+2} x^r, \tag{3.10}$$

where $p(2n+2,r)$ is the probability that an up-down word of length $2n+2$ has descent variation r.

Since $\sum_{l>j} pq^{l-1} zt^l = \frac{pzt}{1-qt}(qt)^j$, we can obtain the sum (3.9) by replacing t^j in the sum (3.3) by $\frac{pzt}{1-qt}(qt)^j$.

Therefore, for $t=1$ we simply use $F(z,x,t)$ in the case of up-down words of odd length, replace t by q and multiply by z to get

$$g(z,x) = zF(z,x,q),$$

i.e., the generating function for descent variation in up-down words of even length is

$$g(z,x) = \sum_{n\geq 1} \left(\frac{p^2 q^{2n-1} z^3 x}{(x-q^{2n})(1-qx)} \cdot \frac{\sum_{k\geq 1} \frac{(-1)^{k-1} p^{2k-1} q^{k(2k-1)} x^k z^{2k-1}}{(1-xq^{2k})(q^2;q^2)_{k-1}(xq^3;q^2)_{k-1}}}{1-\sum_{k\geq 1} \frac{(-1)^{k-1} p^{2k} q^{k(2k-1)} x^k z^{2k}}{(q^2;q^2)_k (qx;q^2)_k}} + \frac{pq^{2n-1} z^2}{1-q^{2n}} \right)$$
$$\times \frac{(-1)^{n-1} (p^2 q^2 z^2)^{n-1} q^{(2n-3)(n-1)}}{x^{n-1} (\frac{q^2}{x}; q^2)_{n-1} (q^3;q^2)_{n-1}}.$$

Once again, to compute the probabilities we need to divide $[z^{2n}]f(z,x)$ by $[z^{2n}]f(z,1)$.

4 Carlitz words

In this section, we consider a special type of words called *Carlitz words*. These are are words over the alphabet $\mathbb{N}$, where the letters again have geometric probabilities and with the property that any two adjacent letters are distinct. In this section, we find the generating function for a Carlitz word with given descent variation. Various people have studied the related topic of Carlitz compositions of natural numbers, namely, Carlitz [1], Knopfmacher and Prodinger [5], Goh and Hitczenko [4], and Louchard and Prodinger [6].

We use the same notation as for the general case, in Section 2, namely

$$f(z,x) = \sum_{n\geq 1, r\geq 0} p(n,r) z^n x^r, \tag{4.1}$$

where $p(n,r)$ is in this case the probability that a Carlitz word of length n has descent variation r, and

$$F(z,x,t) = \sum_{n\geq 1, r\geq 0, i\geq 1} g(n,r,i) z^n x^r t^i, \tag{4.2}$$

where $g(n,r,i)$ is the probability that a Carlitz word of length n, with last letter i ($\omega_n = i$) has descent variation r. We have $f(z,x) = F(z,x,1)$.

Again, to obtain probability generating functions we need to normalize the coefficients of z^n by dividing by the probability of having a Carlitz word of length n.

We perform the "adding a slice" technique again. Suppose we have a random Carlitz word of length n with last letter i, $\omega = \omega_1 \dots \omega_{n-1} i$, with descent variation r, with arbitrarily fixed $i \geq 1$ and $r \geq 0$. We add a letter, j, after the last letter i, and thus get a new word of length $n+1$, $\omega' = \omega_1 \dots \omega_{n-1} ij$.

The letter j can be either greater than i or less than i, but not equal to i.

- If $j > i$ then the descent variation remains the same.
- If $j < i$ then the new word ω' has descent variation $r + i - j$.

This is reflected by the following expression

$$\sum_{j=1}^{i-1} p\, q^{j-1}\, x^{i-j}\, t^j + \sum_{j \geq i+1} p\, q^{j-1}\, t^j. \tag{4.3}$$

Thus, we have the substitution

$$t^i \longrightarrow \frac{pt}{(x-qt)} x^i + \left(\frac{pt}{1-qt} - \frac{px}{q(x-qt)} \right) (qt)^i. \tag{4.4}$$

Again, we let $F(z,x,t) = \sum_{n \geq 1} a_n(x,t) z^n$. For $n \geq 1$, we get from (4.4),

$$a_{n+1}(x,t) = \frac{pt}{(x-qt)} a_n(x,x) + \left(\frac{pt}{1-qt} - \frac{px}{q(x-qt)} \right) a_n(x,qt) \tag{4.5}$$

which, after multiplication by z^{n+1} and then summing over $n \geq 1$, we have

$$F(z,x,t) = a_1(x,t)z + \frac{pt}{(x-qt)} zF(z,x,x) + \left(\frac{pt}{1-qt} - \frac{px}{q(x-qt)} \right) zF(z,x,qt), \tag{4.6}$$

where $a_1(x,t) = \sum_{k \geq 1} pq^{k-1} t^k = \frac{pt}{1-qt}$.

Using Lemma 1, with

$$f(z,x,t) = \frac{ptz}{1-qt} + \frac{ptz}{x-qt} F(z,x,x),$$

and

$$g(z,x,t) = \frac{ptz}{1-qt} - \frac{pxz}{q(x-qt)}$$

we obtain

$$F(z,x,t) = \sum_{j \geq 1} \left(\frac{ptq^{j-1}z}{1-tq^j} + \frac{ptq^{j-1}z}{x-tq^j} F(z,x,x) \right) (pz)^{j-1} \cdot \prod_{i=1}^{j-1} \left(\frac{tq^{i-1}}{1-tq^i} - \frac{x}{q(x-tq^i)} \right). \tag{4.7}$$

For $t = x$, we get

$$F(z,x,x) = \sum_{j \geq 1} \left(\frac{pxq^{j-1}z}{1-xq^j} + \frac{pq^{j-1}z}{1-q^j} F(z,x,x) \right) (pz)^{j-1} \cdot \prod_{i=1}^{j-1} \left(\frac{xq^{i-1}}{1-xq^i} - \frac{1}{q(1-q^i)} \right) \tag{4.8}$$

and thus

$$F(z,x,x)=\frac{\sum_{j\geq 1}\frac{xq^{j-1}(pz)^j}{1-xq^j}\prod_{i=1}^{j-1}\left(\frac{xq^{i-1}}{1-xq^i}-\frac{1}{q(1-q^i)}\right)}{1-\sum_{j\geq 1}\frac{q^{j-1}}{1-q^j}(pz)^j\prod_{i=1}^{j-1}\left(\frac{xq^{i-1}}{1-xq^i}-\frac{1}{q(1-q^i)}\right)}.$$

This yields

$$\begin{aligned}F(z,x,t)=\sum_{n\geq 1}&\left(\frac{tq^{n-1}}{1-tq^n}+\frac{tq^{n-1}}{x-tq^n}\,\frac{\sum_{k\geq 1}\frac{xq^{k-1}(pz)^k}{1-xq^k}\prod_{i=1}^{k-1}\left(\frac{xq^{i-1}}{1-xq^i}-\frac{1}{q(1-q^i)}\right)}{1-\sum_{k\geq 1}\frac{q^{k-1}}{1-q^k}(pz)^k\prod_{i=1}^{k-1}\left(\frac{xq^{i-1}}{1-xq^i}-\frac{1}{q(1-q^i)}\right)}\right)\\&\times(pz)^n\cdot\prod_{i=1}^{n-1}\left(\frac{tq^{i-1}}{1-tq^i}-\frac{x}{q(x-tq^i)}\right).\end{aligned}\tag{4.9}$$

After substituting $t=1$ we get our desired probability generating function

$$\begin{aligned}f(z,x)=F(z,x,1)=\sum_{n\geq 1}&\left(\frac{q^{n-1}}{1-q^n}+\frac{q^{n-1}}{x-q^n}\,\frac{\sum_{k\geq 1}\frac{xq^{k-1}(pz)^k}{1-xq^k}\prod_{i=1}^{k-1}\left(\frac{xq^{i-1}}{1-xq^i}-\frac{1}{q(1-q^i)}\right)}{1-\sum_{k\geq 1}\frac{q^{k-1}}{1-q^k}(pz)^k\prod_{i=1}^{k-1}\left(\frac{xq^{i-1}}{1-xq^i}-\frac{1}{q(1-q^i)}\right)}\right)\\&\times(pz)^n\cdot\prod_{i=1}^{n-1}\left(\frac{q^{i-1}}{1-q^i}-\frac{x}{q(x-q^i)}\right).\end{aligned}\tag{4.10}$$

Furthermore, we obtain after substituting $x=1$ and $t=1$ in (4.7):

$$\begin{aligned}f(z,1)=F(z,1,1)&=\sum_{j\geq 1}\left(\frac{pq^{j-1}z}{1-q^j}+\frac{pq^{j-1}z}{1-q^j}\,F(z,1,1)\right)(pz)^{j-1}\cdot\prod_{i=1}^{j-1}\left(\frac{q^{i-1}}{1-q^i}-\frac{1}{q(1-q^i)}\right)\\&=\sum_{j\geq 1}\left(\frac{1}{1-q^j}+\frac{1}{1-q^j}F(z,1,1)\right)(-1)^{j-1}(pz)^j,\end{aligned}$$

which yields

$$f(z,1)+1=\frac{1}{1-\sum_{j\geq 1}\frac{(-1)^{j-1}(pz)^j}{1-q^j}},\tag{4.11}$$

which is the generating function for Carlitz words. [(i)]

In order to compute probabilities, the coefficients of z^n in $f(z,x)$ must be normalized by dividing by the coefficients of z^n in $f(z,1)$.

[(i)] The corresponding $f(z,1)$ for Carlitz words in [3] does not match (4.11), so is incorrect.

References

[1] L. Carlitz, *Restricted compositions*, The Fibonacci Quarterly, Vol. **14**, 1976, 254–264.

[2] L.L. Cristea and H. Prodinger, *q-enumeration of up-down words by number of rises*, *The Fibonacci Quarterly*, **46/47**, 2008/2009, no. 2, 126–134.

[3] L.L. Cristea and H. Prodinger, *q-enumeration of words by their total variation*, *Discrete Mathematics and Theoretical Computer Science*, **12:3**, 2010, 139–150.

[4] W.M.Y Goh and P. Hitczenko, *Average Number of Distinct Path Sizes in a Random Carlitz Composition*, Europ J. Combinatorics, 2002, **23**, 647–657.

[5] A. Knopfmacher and H. Prodinger, *On Carlitz compositions*, European Journal of Combinatorics, **19**, 1998, 579–589.

[6] G. Louchard and H. Prodinger, *Probabilistic Analysis of Carlitz Compositions*, Discrete Mathematics and Theoretical Computer Science, **5**, 2002, 71–96.

[7] T. Mansour, *Enumeration of words by the sum of differences between adjacent letters*, Discrete Mathematics and Theoretical Computer Science, Vol. **11:1**, 2009, 173–186.