Statistics in words and partitions of a set

Walaa Asakly

Department of Mathematics, University of Haifa, Haifa, Israel walaa_asakly@hotmail.com

Submitted December 30, 2015 — Accepted September 12, 2016

Abstract

Let $[k] = \{1, 2, ..., k\}$ be an alphabet over k letters. A word ω of length n over alphabet [k] is an element of $[k]^n$ and is also called a k-ary word of length n. We say that ω contains an ℓ -peak, if it exists an i such that $2 \le i \le n - \ell$ where $\omega_i = \omega_{i+1} = \cdots = \omega_{i+\ell-1}$ and $\omega_{i-1} < \omega_i$ and $\omega_{i+\ell-1} > \omega_{i+\ell}$. A partition Π of set [n] of size k is a collection $\{B_1, B_2, ..., B_k\}$ of non empty disjoint subsets of [n], called blocks, whose union equals [n]. In this paper, we find an explicit formula for the generating function for the number of words of length n over alphabet [k] according to the number of ℓ -peaks in terms of Chebyshev polynomials of the second kind. As a consequence of the results obtained for words, we finally find the number of ℓ -peaks in set partitions of [n] with exactly k blocks.

Keywords: Set partitions, words, ℓ -peak, Chebyshev polynomials of the second kind

MSC: 05A05

1. Introduction

Words

Let $[k] = \{1, 2, ..., k\}$ be an alphabet over k letters. A word ω of length n over alphabet [k] is an element of $[k]^n$ and is also called a word of length n on k letters or a k-ary word of length n. The number of the words of length n over alphabet [k] is k^n . Similar statistics in patterns of subwords have been widely studied in the literature (see [2]). For example, Kitaev, Mansour and Remmel [3] enumerated the number of rises (respectively, levels and falls) which are subword patterns 12,

(respectively, 11 and 21) in words that have a prescribe first element. Heubach and Mansour [2] enumerated the number of words of length n over alphabet [k] that contain the subword pattern 111 and the subword pattern 112 exactly r times. Burstein and Mansour [1] generalized the result to subword pattern of length ℓ . More recently, Mansour [4] enumerated the number of peaks (subword patterns 121, 132 or 231) and valleys (subword patterns 212, 213 or 312) in words of length n over alphabet [k]. Our aim is to extend this result to patterns of arbitrary length. We say that ω contains an ℓ -peak, if exists $2 \le i \le n - \ell$ such that $\omega_i = \omega_{i+1} = \cdots = \omega_{i+\ell-1}$ and $\omega_{i-1} < \omega_i$ and $\omega_{i+\ell-1} > \omega_{i+\ell}$. For example, the word $12^413^42 = 12222133332$ in $[3]^{11}$ contains two 4-peaks, namely 122221 and 133332.

Set partitions

A partition Π of set [n] with exactly k blocks is a collection $\{B_1, B_2, \ldots, B_k\}$ of non empty disjoint subsets of [n] whose union is equal to [n]. We assume that blocks are listed in increasing order of their minimal elements, that is, $minB_1$ $minB_2 < \cdots < minB_k$. We denote the set of all partitions of [n] with exactly k blocks to be $P_{n,k}$. The number of all partitions of [n] with k blocks is S(n,k), these are the Stirling numbers of the second kind [9]. We denote the set of all partitions of [n] to be P_n , namely $P_n = \bigcup_{k=0}^n P_{n,k}$. The number of all partitions of [n] is $B_n = \sum_{k=0}^n S_{n,k}$, which is the *n*-th Bell number. Any partition Π can be written as $\pi_1 \pi_2 \cdots \pi_n$, where $i \in B_{\pi_i}$ for all i, and this form is called the *canonical* sequential form. For example $\Pi = \{\{12\}, \{3\}, \{4\}\}\$ is a partition of [4], the canonical sequential form is $\pi = 1123$. Several authors have studied different statistics on P_n (see [4]). For instance, Mansour and Munagi [6] found the generating function for the number of partitions of [n] according to rises, descents and levels, they also computed the total number of t-rises (respectively, t-descents and t-levels), this is a increasing subword pattern of size t (respectively, decreasing subword pattern of size t, fixed subword pattern of size t), see [5]. A lot of attention has been given to the statistics on $P_{n,k}$ (see [4]). For example, Shattuck [8] counted the rises, descents and levels in the set partition of [n] with exactly k blocks. In addition, Mansour [4] found an explicit formula for the generating functions for the number of set partition of [n] with exactly k blocks according to the statistics ℓ -rise (respectively, ℓ -descent and ℓ -level). Mansour and Shattuck [7] found an explicit formula for the generating function of set partitions of n with exactly [k] blocks according to the number of peaks (valleys). Our aim is to extend this result for the set $P_{n,k}$ according to the number of ℓ -peaks.

In this paper, we find the generating function of the words of length n over alphabet [k] according to the number of ℓ -peaks. We also compute the total number of ℓ -peaks in the words of length n over alphabet [k]. As a consequence of these results, we find the number of ℓ -peaks in set partitions of [n] with exactly k blocks.

2. Words and partitions of a set according to multi statistics ℓ -peaks

Let $W_k(x, q_1, ..., q_\ell)$ be the generating function for the number of words of length n over alphabet [k] according to the number of ℓ -peaks, namely,

$$W_k(x, q_1, \dots, q_\ell) = \sum_{n \ge 0} x^n \sum_{\omega \in [k]^n} \prod_{i=1}^\ell q_i^{i-peak(\omega)}.$$

Lemma 2.1. The generating function $W_k(x, q_1, ..., q_\ell)$ satisfies the recurrence relation

$$W_k(x, q_1, \dots, q_{\ell}) = \frac{A_{\ell} - xB_{\ell} + W_{k-1}(x, q_1, \dots, q_{\ell})(B_{\ell+1} - A_{\ell})}{(1 - x)(1 + A_{\ell}) + x^{\ell+1} - W_{k-1}(x, q_1, \dots, q_{\ell})((1 - x)A_{\ell} + x^{\ell+1})},$$

where $A_{\ell} = \sum_{i=1}^{\ell} x^i q_i$ and $B_{\ell} = \frac{1-x^{\ell}}{1-x}$.

Proof. It is obvious

$$W_k(x, q_1, \dots, q_\ell) = W_{k-1}(x, q_1, \dots, q_\ell) + W_k^{\dagger}(x, q_1, \dots, q_\ell), \tag{2.1}$$

where $W_k^{\dagger}(x, q_1, \ldots, q_{\ell})$ is the generating function for the number of words ω of length n over alphabet [k] according to the number of ℓ -peaks such that ω contains at least one occurrence of the letter k. A word ω that contains a letter k can be decomposed as either

- (1) k;
- (2) $k\omega'$, where ω' is a non empty word over [k];
- (3) $\omega''k^i\omega'''$, where k^i denotes a word $kk\cdots k$ with exactly i letters, ω'' is a non empty word over [k-1] and ω''' is a non empty word over [k] which starts with a letter $a \neq k$, for $1 \leq i \leq \ell$;
- (4) $\omega''k^i$, for $1 \le i \le \ell$; or
- (5) $\omega'' k^{\ell+1} \omega''''$, where ω'''' is a word over [k].

The corresponding generating functions of these decomposition are

- (1) x:
- (2) $x(W_k(x, q_1, \ldots, q_\ell) 1);$
- (3) $q_i x^i (W_{k-1}(x, q_1, \dots, q_\ell) 1) (W_k(x, q_1, \dots, q_\ell)(1-x) 1)$, for $1 \le i \le \ell$;
- (4) $x^{i}(W_{k-1}(x, q_{1}, \dots, q_{\ell}) 1)$, for $1 \leq i \leq \ell$; or

(5)
$$x^{\ell+1}(W_{k-1}(x,q_1,\ldots,q_\ell)-1)W_k(x,q_1,\ldots,q_\ell),$$

respectively. Hence, by (2.1), we obtain

$$\begin{split} W_k(x,q_1,\ldots,q_\ell) &= W_{k-1}(x,q_1,\ldots,q_\ell) + x + x(W_k(x,q_1,\ldots,q_\ell) - 1) \\ &+ \sum_{i=1}^{\ell} q_i x^i (W_{k-1}(x,q_1,\ldots,q_\ell) - 1) (W_k(x,q_1,\ldots,q_\ell)(1-x) - 1) \\ &+ \sum_{i=1}^{\ell} x^i (W_{k-1}(x,q_1,\ldots,q_\ell) - 1) \\ &+ x^{\ell+1} (W_{k-1}(x,q_1,\ldots,q_\ell) - 1) W_k(x,q_1,\ldots,q_\ell), \end{split}$$

which equivalent to

$$W_k(x, q_1, \dots, q_{\ell}) = \frac{A_{\ell} - xB_{\ell} + W_{k-1}(x, q_1, \dots, q_{\ell})(B_{\ell+1} - A_{\ell})}{(1 - x)(1 + A_{\ell}) + x^{\ell+1} - W_{k-1}(x, q_1, \dots, q_{\ell})\left((1 - x)A_{\ell} + x^{\ell+1}\right)},$$
(2.2)

where
$$A_{\ell} = \sum_{i=1}^{\ell} x^i q_i$$
 and $B_{\ell} = \frac{1-x^{\ell}}{1-x}$.

We plan to find an explicit formula for the generating function $P_k(x, q_1, \dots, q_\ell)$ for the number of partitions of n with exactly k blocks according to the number of ℓ -peaks.

$$P_k(x, q_1, \dots, q_\ell) = \sum_{n>0} x^n \sum_{\pi \in P_{n,k}} \prod_{i=1}^{\ell} q^{i-peak(\pi)}.$$

To do that we will use Lemma 2.1.

Theorem 2.2. For all $k \geq 1$,

$$P_k(x, q_1, \dots, q_\ell) = \prod_{j=1}^k \left(\sum_{i=1}^\ell x^i (q_i(W_j(x, q_1, \dots, q_\ell)(1-x) - 1) + 1) + x^{\ell+1} W_j(x, q_1, \dots, q_\ell) \right).$$

Proof. Any partition π of [n] with exactly k blocks can be decomposed either

- (1) $\pi k^i \pi'$, πk^i , for $1 \le i \le \ell$, where π is a set partition with exactly k-1 blocks, π' is a non empty word over alphabet [k] which starts with a letter a < k; or
- (2) $\pi k^{\ell+1} \pi''$, where π'' is a word over alphabet [k].

The corresponding generating functions are

$$q_i x^i P_{k-1}(x, q_1, \dots, q_\ell) (W_k(x, q_1, \dots, q_\ell) - x W_k(x, q_1, \dots, q_\ell) - 1) + x^i P_{k-1}(x, q_1, \dots, q_\ell),$$

for $1 < i < \ell$;

(2)
$$x^{\ell+1}P_{k-1}(x, q_1, \dots, q_{\ell})W_k(x, q_1, \dots, q_{\ell}),$$

respectively. By summing all the last terms we obtain

$$\begin{split} &P_k(x,q_1,\ldots,q_\ell) \\ &= \sum_{i=1}^\ell q_i x^i P_{k-1}(x,q_1,\ldots,q_\ell) (W_k(x,q_1,\ldots,q_\ell) - x W_k(x,q_1,\ldots,q_\ell) - 1) \\ &+ \sum_{i=1}^\ell x^i P_{k-1}(x,q_1,\ldots,q_\ell) + x^{\ell+1} P_{k-1}(x,q_1,\ldots,q_\ell) W_k(x,q_1,\ldots,q_\ell) \\ &= P_{k-1}(x,q_1,\ldots,q_\ell) \cdot \\ &\cdot \left(\sum_{i=1}^\ell x^i (q_i(W_j(x,q_1,\ldots,q_\ell)(1-x)-1) + 1) + x^{\ell+1} W_j(x,q_1,\ldots,q_\ell) \right). \end{split}$$

Thus, by induction on k together with the initial condition $P_0(x,q) = 1$, we complete the proof.

Example 2.3. Using the recursion given in Theorem 2.2, we may obtain the generating function for the number of partitions of [n] with exactly k blocks,

$$P_k(x,1,\ldots,1) = \prod_{j=1}^k \sum_{i=1}^\ell x^i W_j(x,1,\ldots,1)(1-x) + x^{\ell+1} W_j(x,1,\ldots,1)$$

$$= \prod_{j=1}^k \left(x \frac{1-x^\ell}{1-x} \frac{1}{1-jx} (1-x) + x^{\ell+1} \frac{1}{1-jx} \right)$$

$$= x^k \prod_{j=1}^k \frac{1}{1-jx},$$

which is in accord with the well-known the generating function for the number of partitions of [n] with exactly k blocks.

Example 2.4. By substituting $\ell = 1$ and $q_1 = q$ in Lemma 2.1, we get $W_k(x,q)$ the generating function for the number of words of length n over the alphabet [k] according to the number of peaks (peak of length one), which gives the following recursion

$$W_k(x,q) = \frac{x(q-1) + (1 - x(q-1))W_{k-1}(x,q)}{1 - x(1-q)(1-x) - x(x+q(1-x))W_{k-1}(x,q)}.$$

By using the same substitution in Theorem 2.2, we obtain the recurrence relation for the generating function for the number of set partitions $P_{n,k}$ according to the number of peaks (peak of length one), which gives the following recursion

$$P_k(x,q) = x^k \prod_{j=1}^k (1 + x(1-q)W_j(x,q) + q(W_j(x,q) - 1)),$$

where the two above results agree with the results of Mansour and Shattuck (see [7]).

2.1. Counting ℓ -peaks in words and partitions of a set

Let $W_k(x,q)$ be the generating function for the number of words of length n over alphabet [k] according to the number of ℓ -peaks.

$$W_k(x,q) = \sum_{n\geq 0} x^n \left(\sum_{\omega \in [k]^n} q^{\ell-peak(\omega)} \right).$$

Corollary 2.5. The generating function $W_k(x,q)$ for the number of words of length n over alphabet [k] according to the number of ℓ -peaks is

$$W_k(x,q) = \frac{A + (1-A)W_{k-1}(x,q)}{a_{\ell} - xW_{k-1}(x,q)a_{\ell-1}}$$
(2.3)

where $A = x^{\ell}(q-1)$ and $a_{\ell} = 1 + x^{\ell}(q-1)(1-x)$, which is equivalent to

$$W_k(x,q) = \frac{x^{\ell}(q-1)(U_{k-1}(t) - U_{k-2}(t))}{U_k(t) - U_{k-1}(t) - (1 - x^{\ell}(q-1))(U_{k-1}(t) - U_{k-2}(t))},$$
 (2.4)

where $t = 1 + \frac{x^{\ell+1}}{2}(1-q)$ and U_m is the m-th Chebyshev polynomial of the second kind.

Proof. By substituting $q_i = 1$ for $i \neq \ell$, and $q_\ell = q$ in (2.2) we obtain (2.3). Then, by applying [Appendix D] [4] for (2.2), we obtain (2.4).

Now, our aim is to find the total number of ℓ -peaks in all words of length n over alphabet [k].

Lemma 2.6. For all $k \geq 1$,

$$\frac{d}{dq}W_k(x,q)\mid_{q=1} = \frac{x^{\ell+2}}{(1-kx)^2} \left(2\binom{k}{3} + \binom{k}{2}\right).$$

Proof. We compute the number of ℓ -peaks in all the words of length n over alphabet [k]. By differentiating (2.3) with respect to q, we obtain

$$V_k(x) = \frac{d}{dq} W_k(x,q) \mid_{q=1}$$

$$=\frac{(x^{\ell}(1-W_{k-1}(x,1))+V_{k-1}(x))(1-xW_{k-1}(x,1))}{(1-xW_{k-1}(x,1))^{2}} - \frac{W_{k-1}(x,1)(x^{\ell}(1-x)(1-W_{k-1}(x,1))-xV_{k-1}(x))}{(1-xW_{k-1}(x,1))^{2}},$$

and using $W_k(x,1) = \frac{1}{1-kx}$ (easy to prove by induction), we obtain

$$\frac{d}{dq}W_k(x,q)\mid_{q=1} = \frac{x^{\ell+2}}{(1-kx)^2} \left(2\binom{k}{3} + \binom{k}{2}\right),\tag{2.5}$$

as claimed. \Box

By finding the coefficient of x^n in (2.5) we get the following result

Corollary 2.7. The total number of ℓ -peaks in all the words of length n over alphabet [k] is given by

$$(n-1-\ell)k^{n-2-\ell}\left(2\binom{k}{3}+\binom{k}{2}\right).$$

We plan to find the explicit formula for the generating function $P_k(x,q)$ for the number of $P_{n,k}$ according to the number of ℓ -peaks.

$$P_k(x,q) = \sum_{n\geq 0} x^n \left(\sum_{\pi \in P_{n,k}} q^{\ell - peak(\pi)} \right).$$

Corollary 2.8. For all $k \geq 1$, the generating function $P_k(x,q)$ is given by

$$x^{k} \prod_{j=1}^{k} (W_{j}(x,q)(1+x^{\ell-1}(x-1)) + x^{\ell-1} + qx^{\ell-1}(W_{j}(x,q)(1-x) - 1)).$$

Proof. By substituting $q_i = 1$ for $i \neq \ell$, and $q_\ell = q$ in Theorem (2.2).

Lemma 2.9. For all $k \geq 3$,

$$\frac{d}{dq}P_k(x,q)\mid_{q=1} = \frac{x^{k+\ell}\binom{k}{2}}{(1-x)\cdots(1-kx)} + \frac{x^{k+\ell+2}}{(1-x)\cdots(1-kx)} \sum_{i=2}^k \frac{2\binom{i}{3}+\binom{j}{2}}{(1-jx)}.$$

Proof. By Corollary 2.8, we have

$$\frac{d}{dq} P_k(x,q) \mid_{q=1} = P_k(x,1) \sum_{j=1}^k \lim_{q \to 1} \left(\frac{\frac{d}{dq} L_j(q)}{L_j(q)} \right), \tag{2.6}$$

where

$$L_j(q) = (W_j(x,q)(1+x^{\ell-1}(x-1)) + x^{\ell-1} + qx^{\ell-1}(W_j(x,q)(1-x) - 1).$$

Note that

$$\lim_{q \to 1} \frac{d}{dq} L_j(q) = \lim_{q \to 1} \left(\frac{d}{dq} W_j(x, q) + x^{\ell - 1} (W_j(x, q)(1 - x) - 1) \right)$$

$$= \frac{x^{\ell} \left((j - 1) - (j - 1)jx + (2\binom{j}{3}) + \binom{j}{2})x^2 \right)}{(1 - jx)^2} = x^{\ell} \left(\frac{j - 1}{1 - jx} + \frac{(2\binom{j}{3}) + \binom{j}{2})x^2}{(1 - jx)^2} \right).$$

Hence, by using (2.6) we obtain

$$\frac{d}{dq}P_k(x,q)|_{q=1} = \frac{x^{k+\ell}}{(1-x)\cdots(1-kx)} \sum_{j=1}^k \left(j-1 + \frac{(2\binom{j}{3} + \binom{j}{2})x^2}{(1-jx)}\right)$$
$$= \frac{x^{k+\ell}\binom{k}{2}}{(1-x)\cdots(1-kx)} + \frac{x^{k+\ell+2}}{(1-x)\cdots(1-kx)} \sum_{j=3}^k \frac{2\binom{j}{3} + \binom{j}{2}}{(1-jx)},$$

as required.

By using the facts that $P_k(x,1) = \sum_{n\geq 1} S_{n,k} x^n$ and $\sum_{j=1}^k (j-1) x^{\ell} = x^{\ell} {k \choose 2}$, together with Lemma 2.9 we get the following corollary.

Corollary 2.10. The total number of the ℓ -peaks in all set partitions $P_{n,k}$ is given by

$$\binom{k}{2} S_{n-\ell,k} + \sum_{i=\ell+2}^{n-k} S_{n-i,k} \sum_{i=3}^{k} j^{i-\ell-2} \left(2 \binom{j}{3} + \binom{j}{2} \right).$$

2.2. Applications

By substituting $\ell = 2$ in Corollary 2.7, we obtain the following result

Corollary 2.11. The total number of the 2-peaks in all the words of length n over alphabet [k] is given by

$$(n-3)k^{n-4}\left(2\binom{k}{3}+\binom{k}{2}\right).$$

By substituting $\ell = 2$ in Corollary 2.10, this leads to

Corollary 2.12. The total number of the 2-peaks in all set partitions $P_{n,k}$ is given by

$$\binom{k}{2} S_{n-2,k} + \sum_{i=4}^{n-k} S_{n-i,k} \sum_{i=3}^{k} j^{i-4} \left(2 \binom{j}{3} + \binom{j}{2} \right).$$

By substituting q = 0 in (2.4), we obtain that the generating function for the number of words of length n over alphabet [k] without ℓ -peaks is given by

$$W_k(x,0) = \frac{-x^{\ell}(U_{k-1}(t) - U_{k-2}(t))}{U_k(t) - U_{k-1}(t) - (1 + x^{\ell})(U_{k-1}(t) - U_{k-2}(t))},$$
(2.7)

where $t = 1 + \frac{x^{\ell+1}}{2}$ and U_m is m-th Chebyshev polynomial of the second kind. By substituting q = 0 in Corollary 2.8, we get

$$P_k(x,0) = x^k \prod_{j=1}^k \left(W_j(x,0)(1+x^{\ell-1}(x-1)) + x^{\ell-1} \right), \tag{2.8}$$

by substituting (2.7) into (2.8), and using the relation $U_{j+1}(t) = 2tU_j(t) - U_{j-1}(t)$, we get

$$P_k(x,0) = x^k \prod_{j=1}^k \frac{U_{j-1}(t) - (1+x^\ell)U_{j-2}(t)}{(1-x)U_{j-1}(t) - U_{j-2}(t)},$$

where $t = 1 + \frac{x^{\ell+1}}{2}$, which is the generating function of $P_{n,k}$ without ℓ -peaks. By using the above result with $\ell = 1$, we obtain the same result of Mansour and Shattuck (see [7]).

Corollary 2.13. The generating function for the number of set partitions of P_n without ℓ -peaks is given by

$$1 + \sum_{k>1} P_k(x,0) = \sum_{k>0} x^k \prod_{j=1}^k \frac{U_{j-1}(t) - (1+x^\ell)U_{j-2}(t)}{(1-x)U_{j-1}(t) - U_{j-2}(t)},$$

where $t = 1 + \frac{x^{\ell+1}}{2}$ and U_m is the m-th Chebyshev polynomial of the second kind.

2.3. Conclusion

In the present paper, we determined the generating function for the number of k-ary words of length n according to the number of ℓ -peaks. Also, we determined the generating function for the number of set partitions of [n] with exactly k blocks according to the number of ℓ -peaks. Seems our techniques can be extended to the case of compositions of n (a composition of n is a word $\sigma_1 \sigma_2 \cdots \sigma_m$ such that $\sum_{i=1}^m \sigma_i = n$), where we leave it to the interest reader.

Acknowledgment. The author expresses her appreciation to the referee for his/her careful reading of the manuscript.

References

- [1] Burstein, A., Mansour, T., Counting occurrences of some subword patterns, *Discret Math. Theor. Comput. Sci.* 6(1) (2003) 1–11.
- [2] HEUBACH S., MANSOUR, T., Combinatorics of compositions and words (Boca Raton), CRC Press, Boca Raton, 2010.
- [3] KITAVE, S., MANSOUR, T., REMMEL, J.B., Counting descents, rises, and levels, with prescribe first element, in words, *Discrete Math. Theor. Comput. Sci.* 10(3) (2008) 1–22.

[4] Mansour, T., Combinatorics of set partitions (Boca Raton), CRC Press, Boca Raton, 2013.

- [5] Mansour, T., Munagi, A.O., Enumeration of partitions by long rises, levels, and descents, J. Integer Seq. 12 (2009) Art. 9.1.8.
- [6] Mansour, T., Munagi, A.O., Enumeration of partitions by rises, levels, and descents, in Permutation Patterns: London Mathematical Society, *Lect. Note Ser.* 376, Cambridge University Press, 2010.
- [7] MANSOUR, T., SHATTUCK, M., Counting peaks and valleys in a partition of a set, J. Integer Seq. 13 (2010) Art. 10.6.8.
- [8] Shattuck, M., Recounting the number of rises, levels and descentes in finite set partitions, *Integers* 10:2 (2010) 179–185.
- [9] STANLEY, R.P., Enumerative Combinatorics, Vol. 1, Cambridge University Press, Cambridge, UK, 1996.