ENUMERATING SYMMETRIC AND NON-SYMMETRIC PEAKS IN WORDS

WALAA ASAKLY

Abstract. Let $[k] = \{1, 2, \ldots, k\}$ 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 k-ary word of length n. We say that ω contains a *peak*, if exists $2 \le i \le n-1$ such that $\omega_i < \omega_{i+1}$, $\omega_{i+2} < \omega_{i+1}$. We say that ω contains a *symmetric peak*, if exists $2 \le i \le n-1$ such that $\omega_{i-1} = \omega_{i+1} < \omega_i$, and contains a *non-symmetric peak*, otherwise. In this paper, we find an explicit formula for the generating functions for the number of k-ary words of length n according to the number of symmetric peaks and non-symmetric peaks in terms of Chebyshev polynomials of the second kind. Moreover, we find the number of symmetric and non-symmetric peaks in k-ary word of length n in two ways by using generating functions techniques, and by applying probabilistic methods.

1. Introduction

Let $[k] = \{1, 2, \ldots, k\}$ 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 *word of length n on k letters* or *k-ary word of length n*. The number of k-ary words of length n is k^n. Kitaev, Mansour and Remmel [6] 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 [5] enumerated the number of k-ary words of length n that contain the subword patterns 111 and 112 exactly r times. Burstein and Mansour [1] generalized the result to subword patterns of length ℓ. Knopfmacher, Munagi and Wagner [4] found the mean and variance of the k-ary words of length n according to the number of *p-successions*, (p-succession in a k-ary word $\omega_1\omega_2 \cdots \omega_n$ of length n is two consecutive letters of the form $(\omega_i, \omega_i + p)$, where $i = 1, 2, \cdots, n-1$). Heubach and Mansour [5] found the number of rises, descents and levels in k-ary words of length n, after that Mansour [7] found the number of *peaks* (occurrence of a subword pattern either 121, 132 or 231) and *valleys* (occurence of a subword pattern either 212, 213 or 312) in k-ary words of length n by using generating function. Mansour and Shattuck [10] proved the last result by combinatorial tools. In this paper we restrict our attention in two kind of peaks in k-ary words of length n. We say that $\omega = \omega_1\omega_2 \ldots \omega_n$ contains *a symmetric peak*, if exists $2 \le i \le n-1$ such that $\omega_{i-1}\omega_i\omega_{i+1}$ is a peak, $\omega_i > \max(\omega_{i-1}, \omega_{i+1})$, and $\omega_{i-1} = \omega_i$, and we say that ω contains *a non-symmetric peak*,

Date: August 28, 2018.

1991 *Mathematics Subject Classification.* 05A05, 05A15.

Key words and phrases. Word, symmetric peak, non-symmetric peak, Chebyshev polynomial of the second kind.

if exists $2 \le i \le n-1$ such that $\omega_{i-1}\omega_i\omega_{i+1}$ is a peak and $\omega_{i-1} \neq \omega_{i+1}$. Let ω be any k-ary word of length n, we define $u(\omega) = u_{n,k}(\omega)$ to be the number of symmetric peaks in ω, and we define $\widetilde{u}(\omega) = \widetilde{u}_{n,k}(\omega)$ to be the number of non-symmetric peaks in ω. For example, if $\omega = 12^4 1213^3 132 = 12222121333132 \in [3]^{14}$, then it contains one symmetric peak, namely 121, and one non-symmetric peak, namely 132, so $u(\omega) = \widetilde{u}(\omega) = 1$. Our aim is to find the number of symmetric and non-symmetric peaks in k-ary words of length n. To achieve our goal, we use two different ways, by using generating functions, and by probabilistic approach.

2. Counting symmetric peaks

Let $W_k(x,q)$ be the generating function for the number of k-ary words of length n according to the number of symmetric peaks

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

Lemma 1. *(see [7, Proposition D.5]) Let a_n be any sequence given by*

$$a_n = \frac{A + Ba_{n-1}}{C + Da_{n-1}}$$

with $a_0 = 1$ such that $\alpha = BC - AD \neq 0$. Then for all $n \geq 0$,

$$a_n = \frac{A\left(\frac{A+B}{\sqrt{\alpha}}U_{n-1}(t) - U_{n-2}(t)\right)}{\sqrt{\alpha}\left(\frac{A+B}{\sqrt{\alpha}}U_n(t) - U_{n-1}(t)\right) - B\left(\frac{A+B}{\sqrt{\alpha}}U_{n-1}(t) - U_{n-2}(t)\right)},$$

where $t = \frac{B+C}{2\sqrt{\alpha}}$ and U_m is the m-th Chebyshev polynomial of the second kind.

Lemma 2. *The generating function $W_k(x,q)$ satisfies the recurrence relation*

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

where $W_0(x,q) = 1$, which is equivalent to

$$W_k(x,q) = \frac{x(q-1)A_{k-1}}{\alpha A_k - (1 - x(q-1))A_{k-1}}, \tag{2}$$

where $t = \frac{2+x^2(k-1)(1-q)}{2\alpha}$, $\alpha = \sqrt{1 + x^2(k-2)(1-q)}$, $A_k = \frac{U_k(t)}{\alpha} - U_{k-1}(t)$ and U_m is the m-th Chebyshev polynomial of the second kind.

Proof. We write an equation for $W_k(x,q)$. A k-ary word of length n may or may not contains the letter k, so it is obvious that

$$W_k(x,q) = W_{k-1}(x,q) + W_k^*(x,q), \tag{3}$$

where $W_k^*(x,q)$ is the generating function for the number of k-ary words of length n according to the number of symmetric peaks containing the letter k. A k-ary word

ω of length n that contains the letter k may be decomposed as either k, $\omega' k$, $k\omega''$, $\omega' k\omega'''$, or $\omega' kb\omega''''$, where ω' is a nonempty $(k-1)$-ary word, ω'' is a nonempty k-ary word, ω''' is a nonempty k-ary word, which first letter equals to the last letter in ω', ω'''' (could be empty) is a k-ary word and b is a letter that different from the last letter in ω'. The corresponding generating functions are given by x, $x(W_{k-1}(x,q)-1)$, $x(W_k(x,q)-1)$, $xq(W_{k-1}(x,q)-1)(W_k(x,q)-(k-1)xW_k(x,q)-1)$ and $(W_{k-1}(x,q)-1)x^2(k-1)W_k(x,q)$, respectively. By substituting the last terms in (3) we obtain

$$\begin{aligned} W_k^*(x,q) = x &+ x(W_{k-1}(x,q)-1) + x(W_k(x,q)-1) \\ &+ xq(W_{k-1}(x,q)-1)(W_k(x,q)(1-(k-1)x)-1) \\ &+ (W_{k-1}(x,q)-1)x^2(k-1)W_k(x,q). \end{aligned}$$

Thus

$$\begin{aligned} W_k(x,q) = W_k(x,q)&\left(x - x^2(k-1)\right) \\ &- xqW_k(x,q)(1-(k-1)x) + xW_{k-1}(x,q)W_k(x,q)(x(k-1)+q(1-(k-1)x)) \\ &+ x(q-1) + (1-x(q-1))W_{k-1}(x,q), \end{aligned}$$

which leads to (1). By applying Lemma 1 for (1) we obtain (2), which completes the proof. ∎

Now our plan is to find the total number of symmetric peaks in k-ary words of length n.

Theorem 3. *The total number of symmetric peaks in k-ary words of length n is*

$$(n-2)\binom{k}{2}k^{n-3}.$$

Proof. Define $V_k(x) = \frac{d}{dq}W_k(x,1)$. By differentiating (1) with respect to q and substituting $q=1$, we obtain

$$\begin{aligned} V_k(x) &= \frac{d}{dq}W_k(x,1) \\ &= \frac{x - xW_{k-1}(x,1) + V_{k-1}(x)}{(1-xW_{k-1}(x,1))} \\ &- \frac{W_{k-1}(x,1)(x(1-x(k-1)) - xV_{k-1}(x) - x(1-(k-1)x)W_{k-1}(x,1))}{(1-xW_{k-1}(x,1))^2}. \end{aligned}$$

By substituting $q=1$ in the last equation, and using $W_k(x,1) = \frac{1}{1-kx}$ (which it is followed from 1 and induction on k), we obtain

$$\frac{d}{dq}W_k(x,q)\mid_{q=1} = \frac{x^3\binom{k}{2}}{(1-kx)^2}, \tag{4}$$

and finally by finding the coefficient of x^n in (4) we get the result. ∎

Another proof for Theorem 3. Now we show alternative proof for Theorem 3, by using probability tools. In order to do that, we define $X_i = X_i(\omega)$, $i = 2, 3, \cdots, n-1$ and $\omega \in [k]^n$, to be the discrete random variable such that $\omega_{i-1} = \omega_{i+1} < \omega_i$. It is obvious $P(X_i = m) = \frac{m-1}{k^3}$, for $m = 2, 3, \cdots, k$, where $P(X = m)$ denote the probability that the discrete random variable X equals m. Then all X_i's, $i = 2, 3, \cdots, n-1$, have the same distribution and $u_{n,k} = \sum_{i=2}^{n-1} X_i$. The value of $u_{n,k}$ is given by $k^n E(u_{n,k}) = k^n \sum_{i=2}^{n-1} E(X_i) = k^n (n-2) E(X_2)$, where

$$E(X_2) = \sum_{m=2}^{k} P(X_2 = m) = \sum_{m=2}^{k} \frac{m-1}{k^3} = \frac{k(k-1)}{2k^3} = \frac{k-1}{2k^2}.$$

Therefore,

$$k^n E(u_{n,k}) = k^n \frac{(n-2)(k-1)}{2k^2} = (n-2)\binom{k}{2} k^{n-3},$$

which is accord with the result in Theorem 3.

3. Counting non-symmetric peaks

We define $\widetilde{W}_k(x,q)$ to be the generating function for the number of k-ary words of length n according to the number of non-symmetric peaks

$$\widetilde{W}_k(x,q) = \sum_{n \geq 0} \left(x^n \sum_{\omega \in [k]^n} q^{\widetilde{u}(\omega)} \right).$$

Lemma 4. *The generating function $\widetilde{W}_k(x,q)$ satisfies the recurrence relation*

$$\widetilde{W}_k(x,q) = \frac{x(q-1) + (1 - x(q-1))\widetilde{W}_{k-1}(x,q)}{1 - x(1-q)(1-2x) - x\widetilde{W}_{k-1}(x,q)(2x + q(1-2x))} \tag{5}$$

where $\widetilde{W}_0(x,q) = 1$, which is equivalent to

$$\widetilde{W}_k(x,q) = \frac{x(q-1)A_{k-1}}{\alpha A_k - (1 - x(q-1))A_{k-1}}, \tag{6}$$

where $t = \frac{1+x^2(1-q)}{\alpha}$, $\alpha = \sqrt{1 + x^2(1-q)}$, $A_k = \frac{U_k(t)}{\alpha} - U_{k-1}(t)$ and U_m is the m-th Chebyshev polynomial of the second kind.

Proof. It is obvious that any k-ary word of length n may or may not contains the letter k, so it is leads that $\widetilde{W}_k(x,q)$ satisfies the following equation

$$\widetilde{W}_k(x,q) = \widetilde{W}_{k-1}(x,q) + \widetilde{W}^*_k(x,q), \tag{7}$$

where $\widetilde{W}^*_k(x,q)$ is the generating function for the number of k-ary words of length n according to the number of non-symmetric peaks containing the letter k. A k-ary word ω of length n that contains the letter k may be decomposed as either k, $\omega' k$, $k\omega''$, $\omega' k \omega'''$, or $\omega' k b \omega''''$, where ω' is a nonempty $(k-1)$-ary word, ω'' is a nonempty k-ary word, ω''' is a nonempty k-ary word such that the first letter in it different from the last letter in ω', ω'''' is a k-ary word and b is a letter that equals to the last letter in ω'. The corresponding generating functions are given by x, $x(\widetilde{W}_{k-1}(x,q)-1)$, $x(\widetilde{W}_k(x,q)-1)$, $xq(\widetilde{W}_{k-1}(x,q)-1)(\widetilde{W}_k(x,q)-2x\widetilde{W}_k(x,q)-1)$ and $(\widetilde{W}_{k-1}(x,q)-1)2x^2\widetilde{W}_k(x,q)$, respectively. By substituting the last terms in (7) we obtain

$$\begin{aligned}\widetilde{W}^*_k(x,q) &= x + x(\widetilde{W}_{k-1}(x,q)-1) + x(\widetilde{W}_k(x,q)-1)\\ &\quad + xq(\widetilde{W}_{k-1}(x,q)-1)(\widetilde{W}_k(x,q)(1-2x)-1)\\ &\quad + (\widetilde{W}_{k-1}(x,q)-1)2x^2\widetilde{W}_k(x,q).\end{aligned}$$

Therefore,

$$\begin{aligned}\widetilde{W}_k(x,q) &= \widetilde{W}_k(x,q)\Big(x - x(2x+q(1-2x)) + x\widetilde{W}_{k-1}(x,q)(2x+q(1-2x))\Big)\\ &\quad + x(q-1) + (1-x(q-1))\widetilde{W}_{k-1}(x,q),\end{aligned}$$

which is equivalent to (5). By applying [Appendix D] [7] for (5) we obtain (6), which completes the proof. ■

Now our aim is to find the total number of non-symmetric peaks in k-ary words of length n.

Theorem 5. *The total number of non-symmetric peaks in k-ary words of length n is*

$$2(n-2)\binom{k}{3}k^{n-3}.$$

Proof. By differentiating (5) with respect to q and substituting $q=1$, we obtain

$$\begin{aligned}&\widetilde{V}_k(x) = \frac{d}{dq}\widetilde{W}_k(x,1)\\ &= \frac{x - x\widetilde{W}_{k-1}(x,1) + \widetilde{V}_{k-1}(x)}{(1-x\widetilde{W}_{k-1}(x,1))} - \frac{\widetilde{W}_{k-1}(x,1)(x(1-2x) - x\widetilde{V}_{k-1}(x) - x(1-2x)\widetilde{W}_{k-1}(x,1))}{(1-x\widetilde{W}_{k-1}(x,1))^2}.\end{aligned}$$

By substituting $q=1$ in the last equation, and using $\widetilde{W}_k(x,1)=\frac{1}{1-kx}$ (easy to proof by induction), we obtain

$$\frac{d}{dq}\widetilde{W}_k(x,q)\mid_{q=1} = \frac{2x^3\binom{k}{3}}{(1-kx)^2}, \tag{8}$$

and finally by finding the coefficient of x^n in (8) we get the result. ■

Note that the total number of symmetric peaks in k-ary words of length n, and the total number of the non-symmetric peaks in k-ary words of length n are equal to the total number of all peaks in k-ary words of length n. Mansour and Shattuck see [10] found that the number of all peaks in k-ary words of length n which is $(n-2)k^{n-3}\left(2\binom{k}{3}+\binom{k}{2}\right)$, by using Theorem 3 and the above result we obtain that,

$$\widetilde{u}(t) = (n-2)k^{n-3}\left(2\binom{k}{3}+\binom{k}{2}\right) - (n-2)k^{n-3}\binom{k}{2} = 2(n-2)\binom{k}{3}k^{n-3}.$$

Another proof for Theorem 5. Now, we give a probabilistic proof for Theorem 5. For that, we define $\widetilde{X}_i = \widetilde{X}_i(\omega)$, $i = 2, 3, \cdots, n-1$ and $\omega \in [k]^n$, to be the discrete random variable such that $\omega_{i-1}, \omega_{i+1} < \omega_i$ and $\omega_{i-1} \neq \omega_{i+1}$. It is obvious $P(\widetilde{X}_i = m) = \frac{(m-1)(m-2)}{k^3}$, for $m = 2, 3, \cdots, k$, where $P(X = m)$ denote the probability that the discrete random variable X equals m. Then all $\widetilde{X}_i$'s, $i = 2, 3, \cdots, n-1$, have the same distribution and $\widetilde{u}_{n,k} = \sum_{i=2}^{n-1} \widetilde{X}_i$. The value of $\widetilde{u}_{n,k}$ is given by $k^n E(\widetilde{u}_{n,k}) = k^n \sum_{i=2}^{n-1} E(\widetilde{X}_i) = k^n (n-2) E(\widetilde{X}_2)$, where

$$\begin{aligned} E(\widetilde{X}_2) &= \sum_{m=2}^{k} P(\widetilde{X}_2 = m) = \sum_{m=2}^{k} \frac{(m-1)(m-2)}{k^3} = \sum_{m=1}^{k-1} \frac{m(m-1)}{k^3} \\ &= \sum_{m=1}^{k-1} \frac{2\binom{m}{2}}{k^3} = \frac{2\binom{k}{3}}{k^3}. \end{aligned}$$

Hence we have,

$$k^n E(\widetilde{u}_{n,k}) = k^n \frac{2(n-2)\binom{k}{3}}{k^3} = 2(n-2)\binom{k}{3} k^{n-3},$$

which is accord with the result in Theorem 5.

Acknowledgement. The research of the author was supported by the Ministry of Science and Technology, Israel.

References

[1] A. Burstein and T. Mansour, *Counting occurrences of some subword patterns*, Discrete Math. Theor. Comput. Sci., 6(1) (2003), 1–11.

[2] A. Burstein, *Enumeration of words with fordibben patterns*, Ph.D. thesis, University of Pennsylvania, (1998).

[3] A. Regev, *Asymptotics of number of k-words with an ℓ-descents*, Electron. J. Combin., 5 (1998), #R15.

[4] A. Knopfmacher, A. Munagi, and S. Wagner, *Successions in words and compositions*, Ann. Comb., 16 (2012), 277–287.

[5] S. Heubach and T. Mansour, *Combinatorics of compositions and words*, CRC Press, Boca Raton, (2010).

[6] S. Kitave, T. Mansour, and J. B. Remmel, *Counting descents, rises, and levels with prescribe first element in words*, Discrete Math. Theor. Comput. Sci., 10(3) (2008), 1–22.

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

[8] T. Mansour, *Restricted 132-avoiding k-ary words, Chebyshev polynomials, and continued fractions*, Adv. in Appl. Math., 36(2) (2006), 175–193.

[9] T. Mansor, *Longest alternating subsequences in pattern-restricted k-ary words*, Online J. Analytic Combin., 3 (2008).

[10] T. Mansour and M. Shattuck, *Counting peaks and valleys in a partition of a set*, J. Integer Seq., 13 (2010), Art. 10.6.8.

Department of Computer Science, University of Haifa, 3498838 Haifa, Israel
E-mail address: walaa_asakly@hotmail.com