Parikh q−Matrices for Circular Words

K. Janaki, *Member, IAENG,* M. Sivaji, R. Suresh, *Member, IAENG,* R. Elayaraja and S. Felixia

***Abstract*—Word combinatorics is a significant branch of discrete mathematics with wide-ranging applications. The introduction of Parikh vectors enabled the arithmetization of words, but this process often leads to the loss of crucial information. To address this, Parikh matrix mappings (Parikh matrices) were developed, offering a more detailed yet computationally feasible method for word characterization. Despite their advantages, Parikh matrices do not always uniquely determine words, raising challenges related to injectivity and M-equivalence. Recent research has extended these concepts to circular words—structures with no defined start or end, common in biological sequences like viral DNA. This paper further advances the field by introducing the Parikh q−matrix, which refines the distinction between words that share the same Parikh matrix but differ structurally. We explore the q−counting of scattered subwords in circular words and investigate the ambiguity and unambiguity of these structures. Our results explore the existence of q−unambiguous circular words with prints of unbounded length, providing new insights into the complexity and behavior of circular word structures. This study enhances the understanding of word combinatorics and extends the applicability of Parikh matrices to more intricate word forms.**

***Index Terms*—Subword, Circular words, Parikh matrix, Ambiguity.**

I. Introduction

DISCRETE mathematics with applications in several fields includes word combinatorics. In this context, the Parikh vector [20] have been introduced due to their applicability in arithmetizing words by vectors. In the process of converting words to the vectors, much of the information about that words were lost. In that sense, an extension to a special kind of matrix called the Parikh matrix mapping (Parikh matrices) [29] would provide more information while also remain computationally feasible. Under this circumstance, words become more characterized through numerical quantities. Words are generally not determined by Parikh matrices. The injectivity problem of Parikh matrices and M-equivalent classes of words have been investigated in this area. Nevertheless, problems such as injectivity and characterization of M-unambiguity have been elusive. Some properties related to the injectivity of the Parikh matrix over a binary alphabet are analyzed in [1], [24]. A special property related to the injectivity with respect to the sum of the positions of symbols are investigated in [12].

In [2], [3], Atanasiu analyzed the properties of class of amiable words and investigated the characterization theorem concerning a graph with respect to this class of amible words. Further Arto Salomaa investigated the properties of Lyndon image interms of Parikh matrices and showed that the ambiguity can be resolved by this Lyndon image in [25]. K.G. Subramanian et al. [31] introduced weak ratio property and analyzed the characterization of M-ambiguous words. Using the Parikh matrices, Teh et al. characterized core M-unambiguous words in [21], [23]. The set of Parikh matrices is shown to be a non-commutative semi-ring with a unit element in [30] as well as closed under the operation of shuffle on trajectory. In [17], Kalpana Mahalingam et al. examined the various theoretical aspects of Parikh matrices and analyzed properties of words that lead to their Parikh matrices mutually commute. Counting subwords of a word with repeated letters is not possible with the Parikh matrix. To facilitate this, in [27], the concept of extending Parikh matrix induced by a word was established and some of its properties were investigated. From the study of Parikh matrices, Salomaa introduced the notion of Parikh-friendly permutations and showed that every permutation is Parikh-friendly in [26]. Besides the fact that every permutation of an ordered alphabet is Parikh-friendly, Teh showed that every permutation should have a unique word that witnesses its Parikh-friendliness in [32]. In [4], Atansiau et al. derived a general formula to obtain the Parikh matrix of any power of a given word. In [15], [16], the various concepts and properties of extending Parikh matrix on partial words are examined and discuss their properties.

By transforming words into polynomial matrices, Egecioglu et al. [10], [11] introduced the Parikh q−matrix. Words that have the same Parikh q−matrix will also have the same Parikh matrix and Parikh vector. Interestingly, words with the same Parikh matrix have different Parikh q−matrix. Moreover, it was shown in [10] that the adjoint matrix of the Parikh q−matrix of the word corresponds to its mirror image of the alternate Parikh q−matrix. By using q−counting, Parikh q−matrix counts scattered subwords of certain words. According to [5], an alternating Parikh matrix can be expressed using a q−counting polynomial. In [6], [7], the various concepts and properties of Parikh q−matrix are examined and discuss the properties of words that allow their q−matrices to commute.

Circular words also referred to as necklaces or cyclic words in the literature differ from traditional linear words in that they lack a defined beginning or end. These circular sequences are not just theoretical constructs; they naturally occur in the DNA strands of certain viruses and bacteria [13] which is shown in $Fig1$. Despite their natural existence, circular words have not been as extensively studied as

Manuscript received June 3, 2024; revised January 30, 2025.

K. Janaki is an Assistant Professor in the Department of Mathematics, Chennai Institute of Technology, Kundrathur, Chennai-600069, Tamilnadu, India E-mail: (janu89lava@gmail.com).

M. Sivaji is an Assistant Professor in the Department of Mathematics, SRM Institute of Science and Technology Tiruchirappalli, Tiruchirappalli - 621105, Tamilnadu, India (Corresponding author E-mail: (sivaji.m@ist.srmtrichy.edu.in).

R. Suresh is an Assistant Professor in the Department of Mathematics, SRM Institute of Science and Technology Tiruchirappalli, Tiruchirappalli - 621105, Tamilnadu, India E-mail: (sureshnational@gmail.com).

R. Elayaraja is an Assistant Professor in the Department of Mathematics, SRM Institute of Science and Technology Tiruchirappalli, Tiruchirappalli - 621105, Tamilnadu, India E-mail: (elaya.ams@gmail.com).

S. Felixia is an Assistant Professor in the Department of Mathematics, Panimalar Engineering College, Poonamallee, Chennai-600123, Tamilnadu, India E-mail: (felixiaraj1994@gmail.com).

linear words. Current active research directions in the field

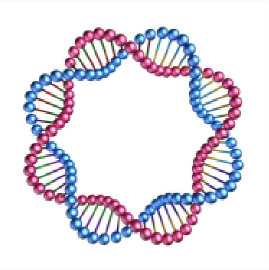

Fig. 1. Circular DNA

of circular words include pattern avoidance [9], [28] and splicing systems [8], [30]. To date, the work most closely related to the study of subword occurrences in circular words is [29]. In [22], the concept of Parikh matrices is extended to the context of circular words and studied the ambiguity in the identification of a circular word by its Parikh matrix. This motivates us to extend circular Parikh matrix to Parikh q−matrix and investigate ambiguity in the identification of a circular word by its Parikh q−matrix. The remainder of this paper is structured as follows. Section 2 provides basic definitions of Parikh matrix and Parikh q−matrix of words which are used in subsequent sections. In Section 3, we define q−counting scattered subwords of circular words and study its characterization. Section 4 studies q−ambiguous circular words and showed that there exists q−unambiguous words with print of unbounded length.

II. Preliminaries

In this section we recollect certain notions of Parikh matrix and Parikh q−matrix.

A. Subwords

Consider an alphabet $\Sigma = \{a_1, a_2, \cdots, a_k\}$ and the set of all words over Σ is Σ^*. For any word $x \in \Sigma^*$, the length of x is denoted by $|x|$. An ordered alphabet is an alphabet $\Sigma = \{a_1, a_2, \cdots, a_k\}$ with the total order relation $a_1 < a_2 < \cdots < a_k$ and it is denoted by Σ_k. The empty word is denoted by λ. A word $y \in \Sigma^*$ is called a *scattered subword* of x if there exist words $y_1, y_2, \cdots, y_n$ and $x_0, x_1, x_2, \cdots, x_n$ over Σ such that $y = y_1 y_2 \cdots y_n$ and $x = x_0 y_1 x_1 y_2 \cdots y_n x_n$. The number of occurrences of the word y as a scattered subword of the word x is denoted by $|x|_y$. For instance $|abbbaaab|_{aab} = 6$. Let a_{ij} be the word $a_i a_{i+1} \cdots a_j$ for $1 \le i < j \le k$ and if $i = j$ then $a_{ij} = a_i$.

B. Parikh matrix

Let $\mathcal{M}_k$ denote the set of all $k \times k$ upper triangular matrices with entries $\mathbb{N}$ and unit diagonal where $\mathbb{N}$ is the set of all non-negative integers.

Definition 1. *Let $\Sigma_k = \{a_1, a_2, \cdots, a_k\}$ be an ordered alphabet where $k \ge 1$. The Parikh matrix mapping denoted by ψ_k is the morphism $\psi_k : \Sigma_k^* \to \mathcal{M}_{k+1}$ defined as $\psi_k(a_l) = (m_{ij})_{1 \le i, j \le k+1}$ where*

- $m_{ii} = 1$ *for* $1 \le i \le k+1$
- $m_{l,(l+1)} = 1$

and all other entries are zero.

Two words $x, y \in \Sigma_k^*$ are said to be M*–equivalent* denoted by $x \sim_M y$ if and only if $\psi_k(x) = \psi_k(y)$. A word $z \in \Sigma_k^*$ is said to be M*–ambiguous* if there exists a word $w \ne z$ such that $z \sim_M w$. Otherwise z is called M*–unambiguous*.

C. Parikh q−matrix

The notion of Parikh matrices is extended to a mapping called Parikh q−matrix mapping which takes its values in matrices with polynomial entries. The entries of the Parikh q−matrices are obtained by q−counting the number of occurrences of certain words as scattered subwords of a given word. The *q−counting of a scattered subword a_{ij}* of a word x represented by $S_{x,aij}$ is defined as follows:

Definition 2. *Let $\Sigma_k = \{a_1, a_2, \cdots, a_k\}$ be an ordered alphabet where $k \ge 1$, $x \in \Sigma_k^*$ and a_{ij} be a scattered subword of x for $1 \le i \le j < k$. Then $S_{x,a_{ij}}(q) = \sum_{x=u_i a_i u_{i+1} \cdots u_j a_j u_{j+1}} q^{|u_i|_{a_i} + |u_{i+1}|_{a_{i+1}} + \cdots + |u_j|_{a_j} + |u_{j+1}|_{a_{j+1}}}$.*

Example 1. *Let $x = baaabb$ be a word over Σ_2. Considering x as a word over Σ_3. Then*

- *For $i = 1$ and $j = 1$ we get $a_{ij} = a$ and $S_{x,a}(q) = q^{0+2} + q^{1+2} + q^{2+2} = q^2 + q^3 + q^4$*
- *For $i = 2$ and $j = 2$ we get $a_{ij} = b$ and $S_{x,b}(q) = q^{0+0} + q^{1+0} + q^{2+0} = 1 + q + q^2$*
- *For $i = 1$ and $j = 2$ we get $a_{ij} = ab$ and $S_{x,ab}(q) = q^{0+0+0} + q^{0+1+0} + q^{1+0+0} + q^{1+1+0} + q^{2+0+0} + q^{2+1+0} = 1 + 2q + 2q^2 + q^3$.*

For any word $x \in \Sigma_k^*$, $S_{x,a_{ij}}(1) = |x|_{a_{ij}}$ for $1 \le i \le j \le k$. Let $\mathcal{M}_k(q)$ denote the set of all $k \times k$ upper triangular matrices with entries $\mathbb{N}(q)$ and unit diagonal where $\mathbb{N}(q)$ is the set of all polynomials in the variable q with coefficients from $\mathbb{N}$.

Definition 3. *Let $\Sigma_k = \{a_1, a_2, \cdots, a_k\}$ be an ordered alphabet and $x \in \Sigma_k^*$ then the Parikh q−matrix mapping denoted by ψ_q is the morphism $\psi_q : \Sigma_k^* \to \mathcal{M}_k(q)$ defined as $\psi_q(a_l) = (m_{ij})_{1 \le i, j \le k+1}$ where*

- $m_{ll} = q$
- $m_{ii} = 1$ *for* $1 \le i \le k$, $i \ne l$
- $m_{l(l+1)} = 1$ *if* $l < k$

and all other entries are zero.

Definition 4. *Let $\Sigma_k = \{a_1, a_2, \cdots, a_k\}$ be an ordered alphabet and $x \in \Sigma_k^*$ then the principal diagonal entries of the matrix $\psi_q(x)$ is $(q^{|x|_{a_1}}, q^{|x|_{a_2}}, \cdots, q^{|x|_{a_k}})$.*

Note that the Parikh vector of x is given by the formal derivative of

$$(q^{|x|_{a_1}}, q^{|x|_{a_2}}, \cdots, q^{|x|_{a_k}})$$

with respect to q at q = 1. The entries of the q−matrices are obtained by q−counting the number of occurrences of certain words as scattered subwords of a given word.

Theorem 1. *[10] Let $\Sigma_k = \{a_1, a_2, \cdots, a_k\}$ be an ordered alphabet and $x \in \Sigma_k^*$. Then the Parikh q−matrix has the following properties*

- $m_{ij} = 0$ *for all* $1 \le j < i \le k$
- $m_{ii} = q^{|x|_{a_i}}$ *for* $1 \le i \le k$
- $m_{i(j+1)} = S_{x,a_{ij}}(q)$ *for all* $1 \le i \le j < k$.

Example 2. *Let $x = acbc$ over Σ_3 then the Parikh $q-$matrix of x is*

$$\begin{aligned}\psi_q(acbc) &= \psi_q(a)\psi_q(c)\psi_q(b)\psi_q(c)\\ &= \begin{bmatrix} q & 1 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}\begin{bmatrix} 1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & q\end{bmatrix}\\ &\quad \begin{bmatrix} 1 & 0 & 0\\ 0 & q & 1\\ 0 & 0 & 1\end{bmatrix}\begin{bmatrix} 1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & q\end{bmatrix}\\ &= \begin{bmatrix} q & q & q\\ 0 & q & q\\ 0 & 0 & q^2\end{bmatrix}\\ &= \begin{bmatrix} q^{|x|_a} & S_{x,a}(q) & S_{x,ab}(q)\\ 0 & q^{|x|_b} & S_{x,b}(q)\\ 0 & 0 & q^{|x|_c}\end{bmatrix}.\end{aligned}$$

The Parikh q−matrix of a word x over $\Sigma_k = \{a_1, a_2, \cdots, a_k\}$ coincides with the usual Parikh matrix, when the q−matrix is evaluated at q = 1 treating the word x as a word over $\Sigma_{k+1} = \{a_1, a_2, \cdots, a_{k+1}\}$. The Parikh matrix of the word $x = acbc$ over Σ_3 is a 4×4 upper triangular matrix given by

$$\psi_3(acbc) = \begin{bmatrix} 1 & 1 & 1 & 1\\ 0 & 1 & 1 & 1\\ 0 & 0 & 1 & 2\\ 0 & 0 & 0 & 1\end{bmatrix}.$$

By comparing Parikh matrix with Parikh q−matrix, add a new symbol d to Σ_3 to get $\Sigma_4 = \{a, b, c, d\}$ and compute the Parikh q−matrix of the word x treating it as a word over Σ_4. For example,

$$\begin{aligned}\psi_q(acbc) &= \psi_q(a)\psi_q(c)\psi_q(b)\psi_q(c)\\ &= \begin{bmatrix} q & 1 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{bmatrix}\begin{bmatrix} 1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & q & 1\\ 0 & 0 & 0 & 1\end{bmatrix}\\ &\quad \begin{bmatrix} 1 & 0 & 0 & 0\\ 0 & q & 1 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{bmatrix}\begin{bmatrix} 1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & q & 1\\ 0 & 0 & 0 & 1\end{bmatrix}\\ &= \begin{bmatrix} q & q & q & 1\\ 0 & q & q & 1\\ 0 & 0 & q^2 & q+1\\ 0 & 0 & 0 & 1\end{bmatrix}\end{aligned}$$

$$\psi_q(acbc) = \begin{bmatrix} q^{|x|_a} & S_{x,a}(q) & S_{x,ab}(q) & S_{x,abc}(q)\\ 0 & q^{|x|_b} & S_{x,b}(q) & S_{x,bc}(q)\\ 0 & 0 & q^{|x|_c} & S_{x,c}(q)\\ 0 & 0 & 0 & 1\end{bmatrix}.$$

Throughout the paper if there is a word x from Σ_k, we assume x to be a word from Σ_{k+1} and compute the Parikh q−matrix of x in $\mathcal{M}_{k+1}(q)$.

Two words $x, y \in \Sigma_k^*$ are said to be *q−equivalent* denoted by $x \sim_q y$ if and only if $\psi_q(x) = \psi_q(y)$. A word $z \in \Sigma_k^*$ is said to be *q−ambiguous* if there exists a word $w \neq z$ such that $z \sim_q w$. Otherwise z is called *q−unambiguous*. Note that if two words x, y are q−equivalent then they have same Parikh vector.

In this work, we will be dealing mostly with Σ_3 and without loss of generality, we let $\Sigma_3 = \{a, b, c\}$ and $a < b < c$ be the corresponding total order.

III. Q−COUNTING SCATTERED SUBWORD OF THE CIRCULAR WORD

In contrast to classical linear words, a circular word has neither a beginning nor an end. The circular word obtained from a word x will be denoted by $[x]$.

Definition 5. *Suppose Σ is an alphabet and $x = x_1x_2\cdots x_n \in \Sigma^*$. The circular word $[x]$ over Σ represented by a word $x \in \Sigma^*$ is the equivalence class of x under the conjugacy relation. The set of all circular words over Σ by Σ_c^*.*

Example 3. *Suppose $\Sigma = \{a, b, c\}$ and consider the word $x = abcabacb$. The circular word $[w]$ as follows*

$$\begin{aligned}[w] &= \{abcabacb, bcabacba, cabacbab, abacbabc,\\ &\quad bacbabca, acbabcab, cbabcaba, babcabac\}\end{aligned}$$

Definition 6. *Suppose Σ is an ordered alphabet and $[x] \in \Sigma_c^*$. The q−counting scattered subword w of the circular word $[x]$ with respect to Σ denoted by $S_{[x],w}(q)$ is defined by*

$$S_{[x],w}(q) = \frac{1}{|[x]|}\sum_{v\in[x]} S_{v,w}(q).$$

Example 4. *Consider $[x] = [abaa]$ over $\Sigma = \{a, b\}$ then the q−counting scattered subword ab of the circular word $[abaa]$ is*

$$\begin{aligned}S_{[abaa],ab}(q) &= \frac{1}{4}\left[S_{abaa,ab}(q) + S_{baaa,ab}(q) +\right.\\ &\quad \left. S_{aaab,ab}(q) + S_{aaba,ab}(q)\right]\\ &= \frac{1}{4}\left[1 + 0 + 1 + q + q^2 + 1 + q\right]\\ &= \frac{1}{4}\left[3 + 2q + q^2\right].\end{aligned}$$

Definition 7. *Suppose Σ is an ordered alphabet and $[x] \in \Sigma_c^*$. The Parikh q−matrix of the circular word $[x]$ with respect to Σ denoted by $\psi_q([x])$ is defined by*

$$\psi_q([x]) = \frac{1}{|[x]|}\sum_{v\in[x]} \psi_q(v).$$

Example 5. *Consider $[x] = [abaa]$ over $\Sigma = \{a, b\}$ then the Parikh q−matrix of the circular word $[abaa]$ is*

$$\begin{aligned}\psi_q([abaa]) &= \frac{1}{4}\left[\psi_q(abaa) + \psi_q(baaa) +\right.\\ &\quad \left.\psi_q(aaab) + \psi_q(aaba)\right]\\ &= \frac{1}{4}\left(\begin{bmatrix} q^3 & q^2+2q & 1\\ 0 & q & 1\\ 0 & 0 & 1\end{bmatrix} +\right.\\ &\quad \begin{bmatrix} q^3 & q^2+q+1 & 0\\ 0 & q & 1\\ 0 & 0 & 1\end{bmatrix} +\\ &\quad \begin{bmatrix} q^3 & q^3+q^2+q & 1\\ 0 & q & 1\\ 0 & 0 & 1\end{bmatrix} +\\ &\quad \left.\begin{bmatrix} q^3 & 2q^2+q & q+1\\ 0 & q & 1\\ 0 & 0 & 1\end{bmatrix}\right)\end{aligned}$$

$$\psi_{\mathrm{q}}([abaa]) = \frac{1}{4}\begin{bmatrix} \mathrm{q}^{12} & 1+5\mathrm{q}+5\mathrm{q}^2+\mathrm{q}^3 & 3+2\mathrm{q}+\mathrm{q}^2 \\ 0 & \mathrm{q}^4 & 4 \\ 0 & 0 & 1 \end{bmatrix}.$$

ALGORITHM FOR COMPUTING PARIKH q-MATRIX

Input: Circular word $[x]$, ordered alphabet Σ
Output: Parikh q-matrix $\psi_{\mathrm{q}}([x])$

1) Define $\psi_{\mathrm{q}}(v)$ for $v \in [x]$ as follows:
   - $m_{ll} = \mathrm{q}$
   - $m_{ii} = 1$ for $1 \le i \le k$, $i \ne l$
   - $m_{l(l+1)} = 1$ if $l < k$
   - All other entries are zero.
2) Compute $\psi_{\mathrm{q}}([x])$ using the following formula:

$$\psi_{\mathrm{q}}([x]) = \frac{1}{|[x]|}\sum_{v\in[x]}\psi_{\mathrm{q}}(v) = \frac{1}{|[x]|}\left(\psi_{\mathrm{q}}(v_1)+\psi_{\mathrm{q}}(v_2)+\cdots+\psi_{\mathrm{q}}(v_{|[x]|})\right)$$

where $|[x]|$ is the length of the circular word $[x]$ and $v_1, v_2, \ldots, v_{|[x]|}$ are its cyclic permutations.

Theorem 2. *Let $\Sigma = \{a_1, a_2, \cdots, a_k\}$ be an ordered alphabet and $[x] \in \Sigma_c^*$. Then the Parikh q–matrix $\psi_q([x]) = M$ has the following properties:*

(i) $M_{ij} = 0$ *for all* $1 \le j < i \le k+1$
(ii) $M_{ii} = q^{|[x]|_{a_i}}$ *for* $1 \le i \le k+1$
(iii) $M_{i(j+1)} = S_{[x],a_{ij}}(q)$ *for all* $1 \le i \le j < k$.

Proof: Obviously the first two properties (i) and (ii) are true. Now we prove the property (iii). By induction on n, assume that $|[x]| = n$. If $n \le 1$, the assertion holds. Assume now that the assertion (iii) is true for all words of length at most n and let $[x]$ be the length $n+1$. Hence $[x] = [x']a_i$ where $|[x']| = n$ and $a_j \in \Sigma$. Then

$$\psi_{\mathrm{q}}([x]) = \psi_{\mathrm{q}}([x'])\psi_{\mathrm{q}}(a_j).$$

Assume that

$$\psi_{\mathrm{q}}([x']) = \begin{bmatrix} \mathrm{q}^{|[x']|_{a_1}} & m'_{1,2} & \cdots & \cdots & m'_{1,k} \\ 0 & \mathrm{q}^{|[x']|_{a_2}} & \cdots & \cdots & m'_{2,k} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \cdots & \mathrm{q}^{|[x']|_{a_k}} \\ 0 & 0 & \cdots & \cdots & 1 \end{bmatrix} = M'.$$

By the inductive hypothesis the matrix $\psi_{\mathrm{q}}([x'])$ has property (iii). By Definition 7,

$$\psi_{\mathrm{q}}(a_j) = \begin{bmatrix} 1 & 0 & \cdots & \cdots & \cdots & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \cdots & \cdots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \cdots & \cdots & \vdots \\ 0 & \cdots & q & 1 & \cdots & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \cdots & \cdots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \cdots & \cdots & \vdots \\ 0 & 0 & \cdots & \cdots & \cdots & \cdots & 1 \end{bmatrix}$$

where the matrix differs from I_{k+1} only in two entries, there are

(i) the entry in position (j,j) is q
(ii) the entry in position $(j, j+1)$ is 1.

Therefore

$$\psi_{\mathrm{q}}([x']) = \begin{bmatrix} \mathrm{q}^{|[x']|_{a_1}} & m'_{1,2} & \cdots & \cdots & m'_{1,k} \\ 0 & \mathrm{q}^{|[x']|_{a_2}} & \cdots & \cdots & m'_{2,k} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \cdots & \mathrm{q}^{|[x']|_{a_k}} \\ 0 & 0 & \cdots & \cdots & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & \cdots & \cdots & \cdots & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \cdots & \cdots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \cdots & \cdots & \vdots \\ 0 & \cdots & q & 1 & \cdots & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \cdots & \cdots & \vdots \\ 0 & 0 & \cdots & \cdots & \cdots & \cdots & 1 \end{bmatrix}$$

The resulting matrix $M = (m_{p,q})_{1\le p,q\le k+1}$ then

$$m_{i,j} = \mathrm{q}m'_{i,j}, \forall 1 \le i \le j$$
$$m_{i,j+1} = m'_{i,j} + m'_{i,j+1}, \forall 1 \le i \le j$$

and for all other entries $m_{p,q} = m'_{p,q}$. But these are immediate from the definition of the polynomials $S_{[x],a_{i,j}}(\mathrm{q})$ which satisfy

$$S_{[x']a_j,a_{i,j-1}}(\mathrm{q}) = \mathrm{q}S_{[x],a_{i,j-1}}(\mathrm{q}), \forall 1 \le i \le j$$
$$S_{[x']a_j,a_{i,j}}(\mathrm{q}) = S_{[x'],a_{i,j-1}}(\mathrm{q}) + S_{[x'],a_{i,j}}(\mathrm{q}), \forall 1 \le i \le j$$

and are unchanged otherwise. ■

The subword history of any word x is given by the equation

$$|x|_{ab} + |x|_{ba} = |x|_a . |x|_b.$$

For q–counting scattered subwords, the following analogy can be drawn.

Theorem 3. *[11] Suppose Σ is an ordered alphabet and $x \in \Sigma^*$. Then*

$$Adj(\psi_q(x)) = \overline{\psi}_q(mi(x)).$$

We observe that Theorem 3 is not satisfied for the circular words. For example consider the circular word $[x] = [aab]$ then

$$Adj(\psi_q([aab])) = \frac{1}{3}\begin{bmatrix} \mathrm{q}^3 & -(1+4\mathrm{q}+\mathrm{q}^2) & 3+12\mathrm{q}+3\mathrm{q}^2+2\mathrm{q}^3+\mathrm{q}^4 \\ 0 & \mathrm{q}^6 & -3\mathrm{q}^6 \\ 0 & 0 & \mathrm{q}^9 \end{bmatrix}$$

$$\overline{\psi}_q([mi(aab)]) = \frac{1}{3}\begin{bmatrix} \mathrm{q}^3 & -(1+\mathrm{q}) & 1 \\ 0 & \mathrm{q}^6 & -3\mathrm{q}^2 \\ 0 & 0 & \mathrm{q}^9 \end{bmatrix}$$

Theorem 4. *[6] Let x be a word over Σ_2 then $S_{x,a}(q)S_{x,b}(q) - q^{|x|_b}S_{x,ab}(q) = \overline{S}_{mi(x),ab}(q)$.*

We extend Theorem 4 for circular word as follows.

Theorem 5. *Let x be a word over Σ_2 then*

$$S_{[x],a}(q)S_{[x],b}(q) - q^{|[x]|_b}S_{[x],ab}(q) = \overline{S}_{[mi(x)],ab}(q).$$

Proof: Let $\Sigma = \{a, b\}$. As in Theorem 4, $\overline{\mathrm{S}}_{mi(x),ab}(\mathrm{q}) = \mathrm{S}_{x,a}(\mathrm{q})\mathrm{S}_{x,b}(\mathrm{q}) - \mathrm{q}^{|x|_b}\mathrm{S}_{x,ab}(\mathrm{q})$. Thus it is enough to show that

$$\begin{aligned}\overline{\mathrm{S}}_{[mi(x)],ab}(\mathrm{q}) &= \frac{1}{[mi(x)]}\sum_{y\in[mi(x)]}\mathrm{S}_{y,ab}(\mathrm{q})\\ &= \frac{1}{|[x]|}\sum_{y\in[x]}\mathrm{S}_{mi(y),ab}(\mathrm{q})\\ &= \frac{1}{|[x]|}\sum_{y\in[x]}\big[\mathrm{S}_{y,a}(\mathrm{q})\mathrm{S}_{y,b}(\mathrm{q})-\\ &\quad \mathrm{q}^{|y|_b}\mathrm{S}_{y,ab}(\mathrm{q})\big]\\ &= \frac{1}{|[x]|}\Big[\mathrm{S}_{y,a}(\mathrm{q})\mathrm{S}_{y,b}(\mathrm{q})\mathrm{q}^{|[x]|}-\\ &\quad \mathrm{q}^{|[x]|_b}\sum_{y\in[x]}\mathrm{S}_{y,ab}(\mathrm{q})\Big]\\ &= \mathrm{S}_{[x],a}(\mathrm{q})\mathrm{S}_{[x],b}(\mathrm{q})-\mathrm{q}^{|[x]|_b}\mathrm{S}_{[x],ab}(\mathrm{q}).\end{aligned}$$

■

Theorem 6. *If $S_{x,a}(q)=S_{y,a}(q)$ where $x,\ y$ be the words over Σ_2 then*

$$S_{[x],a}(q)=S_{[y],a}(q).$$

Proof: Since $\mathrm{S}_{x,a}(\mathrm{q})=\mathrm{S}_{y,a}(\mathrm{q})$, we have $|x|_a=|y|_a$ which implies that $|[x]|_a=|[y]|_a$. Hence

$$\begin{aligned}\mathrm{S}_{[x],a}(\mathrm{q}) &= \frac{1}{|[x]|}\sum_{v\in[x],a\in v}\mathrm{S}_{v,a}(\mathrm{q})\\ &= \frac{1}{|[y]|}\sum_{w\in[y],a\in w}\mathrm{S}_{w,a}(\mathrm{q})\\ &= \mathrm{S}_{[y],a}(\mathrm{q}).\end{aligned}$$

Therefore $\mathrm{S}_{x,a}(\mathrm{q})=\mathrm{S}_{y,a}(\mathrm{q})\Leftrightarrow\mathrm{S}_{[x],a}(\mathrm{q})=\mathrm{S}_{[y],a}(\mathrm{q})$. ■

IV. q−AMBIGUOUS CIRCULAR WORDS

In this section, we introduce the concept of C_{q}−ambiguous words. While the Parikh q-matrix of a word provides additional information compared to the Parikh matrix, it has been demonstrated in [11] that this mapping is also not one-to-one.

Definition 8. *Two circular words $[x],\ [y]\in\Sigma_k^*$ are said to be C_q−equivalent denoted by $[x]\sim_{C_q}[y]$ if and only if*

$$\psi_q([x])=\psi_q([y]).$$

A circular word $[x]\in\Sigma_k^*$ is said to be C_q−ambiguous if there exists a word $[z]\neq[x]$ such that $[z]\sim_{C_q}[x]$. We say that $[x]$ is C_q−unambiguous.

Example 6. *Let $[x]=[aabb]=\{aabb,abba,bbaa,baab\}$ and $[y]=[bbaa]=\{bbaa,baab,aabb,abba\}$ be two circular words over Σ_2. Then one can verify that the words x and y are not q−ambiguous but the circular words $[x]$ and $[y]$ are C_q−ambiguous, since*

$$\begin{aligned}\psi_q([aabb]) &= \frac{1}{4}\big[\psi_q(aabb)+\psi_q(abba)+\\ &\quad \psi_q(bbaa)+\psi_q(baab)\big]\\ &= \frac{1}{4}\left(\begin{bmatrix}q^2 & q^3+q^2 & q^2+2q+1\\ 0 & q^2 & q+1\\ 0 & 0 & 1\end{bmatrix}+\right.\\ &\quad \begin{bmatrix}q^2 & q+q^2 & q+1\\ 0 & q^2 & q+1\\ 0 & 0 & 1\end{bmatrix}+\\ &\quad \begin{bmatrix}q^2 & q+1 & 0\\ 0 & q^2 & q+1\\ 0 & 0 & 1\end{bmatrix}+\\ &\quad \left.\begin{bmatrix}q^2 & q^2+q & q+1\\ 0 & q^2 & q+1\\ 0 & 0 & 1\end{bmatrix}\right)\\ &= \frac{1}{4}\begin{bmatrix}q^8 & q^3+3q^2+3q+1 & q^2+4q+3\\ 0 & q^8 & 4q+4\\ 0 & 0 & 1\end{bmatrix}\\ &= \psi_q([bbaa])\end{aligned}$$

Theorem 7. *If the words x and y are q−equivalent over Σ_2 then*

$$[x]\sim_{C_q}[y].$$

Proof: Let x and y be the q−equivalent over Σ_2 such that

$$\mathrm{S}_{x,w}(\mathrm{q}) = \mathrm{S}_{y,w}(\mathrm{q})$$

for $w\in\{a,b,ab\}$ which implies that x and y have same Parikh vector. Therefore we have

$$\begin{aligned}\mathrm{S}_{[x],w}(\mathrm{q}) &= \frac{1}{|[x]|}\sum_{v\in[x]}\mathrm{S}_{v,w}(\mathrm{q})\\ &= \frac{1}{|[x]|}\big[\mathrm{S}_{(x_1x_2x_3\cdots x_n),w}(\mathrm{q})+\\ &\quad \mathrm{S}_{(x_2x_3x_4\cdots x_nx_1),w}(\mathrm{q})\\ &\quad +\cdots+\mathrm{S}_{(x_nx_1x_2\cdots x_{n-1}),w}(\mathrm{q})\big]\\ &= \frac{1}{|[y]|}\big[\mathrm{S}_{(x_1x_2x_3\cdots x_n),w}(\mathrm{q})+\\ &\quad \mathrm{S}_{(x_2x_3x_4\cdots x_nx_1),w}(\mathrm{q})\\ &\quad +\cdots+\mathrm{S}_{(x_nx_1x_2\cdots x_{n-1}),w}(\mathrm{q})\big]\\ &= \frac{1}{|[y]|}\big[\mathrm{S}_{(y_1y_2y_3\cdots y_n),w}(\mathrm{q})+\\ &\quad \mathrm{S}_{(y_2y_3y_4\cdots y_ny_1),w}(\mathrm{q})\\ &\quad +\cdots+\mathrm{S}_{(y_ny_1y_2\cdots y_{n-1}),w}(\mathrm{q})\big]\\ &= \mathrm{S}_{[y],w}(\mathrm{q})\end{aligned}$$

Therefore $[x]\sim_{C_q}[y]$. ■

Definition 9. *Two words $[x],[y]\in\Sigma_k^*$ are considered to fulfill C_q-weak ratio property termed as $[x]\sim_{C_qwr}[y]$ if for each $a_i\in\Sigma_k$, $|[x]|_{a_i}=m|[y]|_{a_i}$ where m be a nonzero rational number and for each $1\le i\le k-1$, any one of (R_1) and (R_2) be true, where*
$(R_1):|[x]|_{a_i}=|[x]|_{a_{i+1}}$ and $|[y]|_{a_i}=|[y]|_{a_{i+1}}$
$(R_2):\ \dfrac{S_{[x],a_i}(q)}{S_{[y],a_i}(q)}=\dfrac{q^{|[x]|_{a_{i+1}}}-q^{|[x]|_{a_i}}}{q^{|[y]|_{a_{i+1}}}-q^{|[y]|_{a_i}}}$, where $q\neq0$ and $|[x]|_{a_i}\neq|[x]|_{a_{i+1}}$, $|[y]|_{a_i}\neq|[y]|_{a_{i+1}}$.

The $C_{\mathtt{q}}$-weak ratio property is a necessary condition for the Parikh $\mathtt{q}$-matrices of two circular words $[x]$ and $[y]$ in Σ_k^* to commute.

Theorem 8. *Let $[x],[y]$ be the circular words over Σ_k^*. If $[x]$, $[y]$ satisfy C_q–weak ratio property then their Parikh $\mathtt{q}$–matrices commute, i.e.,*

$$\psi_{\mathtt{q}}([xy]) = \psi_{\mathtt{q}}([yx]).$$

Proof: Consider the circular words $[x],[y] \in \Sigma_2^*$ satisfied $C_{\mathtt{q}}$–weak ratio property such that $\psi_{\mathtt{q}}(xy) = \psi_{\mathtt{q}}(yx)$ then we have

(i) $xy \sim_{\mathtt{q}} yx \Rightarrow \mathtt{S}_{xy,w}(\mathtt{q}) = \mathtt{S}_{yx,w}(\mathtt{q})$ for $w \in \{a,b,ab\}$

(ii) xy and yx have same Parikh vector.

Therefore,

$$\begin{aligned}
\mathtt{S}_{[xy],w}(\mathtt{q}) &= \frac{1}{|[xy]|}\sum_{v\in[xy]}\mathtt{S}_{v,w}(\mathtt{q}) \\
&= \frac{1}{|[xy]|}\left(|[xy]|_a \mathtt{q}^{|[xy]|} + |[xy]|_b \mathtt{q}^{|[xy]|-1}\right) \\
&= \frac{1}{|[yx]|}\left(|[yx]|_a \mathtt{q}^{|[yx]|} + |[yx]|_b \mathtt{q}^{|[yx]|-1}\right) \\
&= \frac{1}{|[yx]|}\sum_{v\in[yx]}\mathtt{S}_{v,w}(\mathtt{q}) \\
&= \mathtt{S}_{[yx],w}(\mathtt{q}).
\end{aligned}$$

Hence $\psi_{\mathtt{q}}([xy]) = \psi_{\mathtt{q}}([yx])$. ■

In the following theorem we show that there exists $\mathtt{q}$-unambiguous words with print of unbounded length.

Theorem 9. *If $[x]$ and $[y]$ are words of the forms $[(ab)^m]$ and $[(ba)^m]$ respectively, where $m \geq 1$ and the words are over the alphabet Σ_2, then $[x]$ and $[y]$ are C_q-unambiguous.*

Proof: Let x and y be two words over Σ_2 that share the same Parikh $\mathtt{q}$-matrix. Consider the word y, where the letter a appears as the rightmost character, such that $y = uav$ with $|u|_a = m-1$. This implies $[y] = [uav]$. Since each monomial $\mathtt{S}_{[x],a}(\mathtt{q})$ is of degree m, it follows that $|v|_b = 1$. Thus, we can express:

$$\begin{aligned}
[y] &= [u_1 av] \\
&= [u_1]ab.
\end{aligned}$$

Next, considering the rightmost a in $[u_1]$, we can further express:

$$\begin{aligned}
[y] &= [u_2 av]ab \\
&= [u_2]abab.
\end{aligned}$$

Continuing this process, we obtain:

$$[y] = [(ababab)^m] = [x].$$

Therefore, $[x]$ and $[y]$ are indeed $C_{\mathtt{q}}$-unambiguous. ■

Theorem 10. *If $[x]$ and $[y]$ are words of the forms $[(ab)^m]$ and $[(ab)^{m-1}(ba)]$, where $m \geq 1$ over Σ_2, then $[x]$ and $[y]$ are C_q-unambiguous.*

Proof: Let $x = (ab)^m$ and $y = (ab)^{m-1}(ba)$ be two words over Σ_2 with the same Parikh $\mathtt{q}$-matrix. Consider the word y, where the letter a appears as the rightmost character, such that $y = uav$ with $|u|_a = m-2$. This implies $[y] = [uav]$. Since each monomial $\mathtt{S}_{[x],a}(\mathtt{q})$ is of degree m, it follows that $|v|_b = 1$. Thus, we can express:

$$\begin{aligned}
[y] &= [u_1 av] \\
&= [u_1]ab \\
&= [(ab)^{m-2}ab].
\end{aligned}$$

Next, considering the rightmost a in $[u_1]$, we can further express:

$$\begin{aligned}
[y] &= [u_2 av]ab \\
&= [u_2]abab \\
&= [(ab)^{m-2}abab].
\end{aligned}$$

Continuing this process, we obtain:

$$[y] = [(ababab)^m] = [x].$$

Thus, $[x]$ and $[y]$ are $C_{\mathtt{q}}$-unambiguous. ■

Theorem 11. *If $[x]$ and $[y]$ are words of the forms $[(ab)^m]$ and $[a(ba)^{m-1}b]$, where $m \geq 1$ over Σ_2, then $[x]$ and $[y]$ are C_q-unambiguous.*

Proof: Let $x = (ab)^m$ and $y = a(ba)^{m-1}b$ be two words over Σ_2 with the same Parikh $\mathtt{q}$-matrix. Consider the word y, where the letter a appears as the rightmost character, such that $y = uav$ with $|u|_a = m-2$. This implies $[y] = [uav]$. Since each monomial $\mathtt{S}_{[x],a}(\mathtt{q})$ is of degree m, it follows that $|v|_b = 1$. Thus, we can express:

$$\begin{aligned}
[y] &= [u_1 av] \\
&= [u_1]ab.
\end{aligned}$$

Next, considering the rightmost a in $[u_1]$, we can further express:

$$\begin{aligned}
[y] &= [u_2 av]ab \\
&= [u_2]abab.
\end{aligned}$$

Continuing this process, we obtain:

$$[y] = [(ababab)^m] = [x].$$

Thus, $[x]$ and $[y]$ are $C_{\mathtt{q}}$-unambiguous. ■

V. Conclusion

This paper introduces the concept of Parikh $\mathtt{q}$–matrices for circular words, advancing the field of word combinatorics by refining the characterization of circular sequences, particularly those encountered in biological structures such as viral DNA. By employing $\mathtt{q}$-counting techniques, we have extended the traditional Parikh matrix framework to account for scattered subwords within circular words, addressing the ambiguity and injectivity challenges that arise. Our study provides a new perspective on the structure and behavior of circular words, demonstrating the existence of $\mathtt{q}$–unambiguous words with unbounded lengths. These findings enhance the theoretical understanding of circular word structures and lay the groundwork for further applications in discrete mathematics and computational biology.

REFERENCES

[1] A. Atanasiu, C. Martin-Vide and A. Mateescu, "On the injectivity of the Parikh matrix mapping," *Fundamenta Informaticae* 49(4) (2002), 289–299.

[2] A. Atanasiu, "Binary amiable words," *International Journal of Foundations of Computer Science* 18(2) (2007), 387–400.

[3] A. Atanasiu, R. Atanasiu and I. Petre, "Parikh matrices and amiable words," *Theoretical Computer Science* 390(1) (2008), 102–109.

[4] A. Atanasiu, G. Poovanandran and W.C. Teh, "Parikh matrices for powers of words," *Acta Informatica* 56(6) (2019), 521–535.

[5] S. Bera, R. Ceterchi, K. Mahalingam, K.G. Subramanian, "Parikh q-matrix and q-ambiguous Words," *International Journal of Foundations of Computer Science* 31 (2020), 23–36.

[6] S. Bera, K. Mahaligam, "Some algebraic aspects of Parikh q-matrices," *International Journal of Foundations of Computer Science* 27 (2016), 479–499.

[7] S. Bera, K. Mahaligam, "On commuting Parikh q-matrix," *Fundamenta Informaticae* 172 (2020), 327–341.

[8] P. Bonizzoni, C. De Felice, G. Mauri and R. Zizza, "On the power of circular splicing," *Discrete Applied Mathematics* 150(1-3) (2005), 51–66.

[9] J.D. Currie and D.S. Fitzpatrick, "Circular words avoiding patterns," *Lecture Notes in Computer Science* 2450 (2003), 319–325.

[10] O. Egecioglu, O.H. Ibarra, "A matrix q-analogue of the Parikh map," *In Exploring New Frontiers of Theoretical Informatics, Springer, Boston, MA* (2004), 125–138.

[11] O. Egecioglu, O.H. Ibarra, "A q-Analogue of the Parikh matrix mapping," *Formal Models, Languages and Applications, Series in Machine perception and Artificial Intelligence, World scientific* 66 (2006), 97–111.

[12] S. Fosse and G. Richomme, "Some characterizations of Parikh matrix equivalent binary words," *Information Processing Letters* 92(2) (2004), 77–82.

[13] D.R. Helinski and D.B. Clewell, "Circular DNA," *Annual Review of Biochemistry* 40(1) (1971), 899–942.

[14] K. Janaki, R. Arulprakasam, P. Meenakshi and V.R. Dare, "Algebraic properties of Parikh $q-$matrices on two-dimensional words," *Lecture Notes in Computer Science* 13348 (2023), 171-188.

[15] K. Janaki, R. Krishna Kumari and R. Arulprakasam, "Algebraic aspects of extending Parikh matrices of picture arrays," *Mathematics in Science, Engineering, and Aerospace (MESA)* 14 (2023), 1139-1159.

[16] K. Janaki, R. Krishna Kumari, S. Marichamy, S. Felixia, R. Arulprakasam, "Algebraic Aspects of Generalized Parikh Matrices on Partial Words," *IAENG International Journal of Computer Science* 51(3) (2024), 232-242.

[17] K. Mahalingam and K.G. Subramanian, "Product of Parikh Matrices and Commutativity," *International Journal of Foundations of Computer Science* 23(1) (2012), 207–223.

[18] A. Mateescu, A. Salomaa, K. Salomaa and S. Yu, "A sharpening of the Parikh mapping," *Theoretical Informatics and Applications* 35(6) (2001), 551–564.

[19] A. Mateescu, "Algebraic aspects of Parikh matrices," *Theory is Forever* 10 (2004), 170–180.

[20] R.J. Parikh, "On context-free languages," *Journal of the ACM* 13(4) (1966), 570–581.

[21] G. Poovanandran and W.C. Teh, "On M-equivalence and strong M-equivalence for Parikh matrices," *International Journal of Foundations of Computer Science* 29(1) (2018), 123–137.

[22] G. Poovanandran, J. Simpson and W.C. Teh, "Counting subwords in circular words and their Parikh matrices," *Theoretical Computer Science* 985 (2024), 114344.

[23] G. Poovanandran and W.C. Teh, "Strong (2,t) and strong (3,t) transformations for strong M-equivalence," *International Journal of Foundations of Computer Science* 30(5) (2019), 719–733.

[24] A. Salomaa, "On the injectivity of Parikh matrix mappings," *Fundamenta Informaticae* 64(1-4) (2005), 391-404.

[25] A. Salomaa and S. Yu, "Subword occurrences, Parikh matrices and Lyndon images," *International Journal of Foundations of Computer Science* 21(1) (2010), 91-111.

[26] A. Salomaa, "Parikh matrices: subword indicators and degrees of ambiguity," *Lecture Notes in Computer Science* 11011 (2018), 100–112.

[27] T.F. Serbanuta, "Extending Parikh matrices," *Theoretical Computer Science* 310(1) (2004), 233–246.

[28] A.M. Shur, "On ternary square-free circular words," *Electronic Journal of Combinatorics* 17(1) (2010).

[29] J. Simpson, "Palindromes in circular words," *Theoretical Computer Science* 550 (2014), 66–78.

[30] R. Siromoney, K.G. Subramanian and V.R. Dare, "Circular DNA and splicing systems," *Lecture Notes in Computer Science* 654 (1992), 260–273.

[31] K.G. Subramanian, A.M. Huey and A.K. Nagar, "On Parikh matrices," *International Journal of Foundations of Computer Science* 20(2) (2009), 211–219.

[32] W.C. Teh, "Parikh-friendly permutations and uniformly Parikh-friendly words," *Australasian Journal of Combinatorics* 76 (2020), 208–219.