Communications in Combinatorics and Optimization
Vol. xx, No. x (xxxx), pp. 1-11
https://doi.org/10.22049/cco.2025.30296.2405

CCO
Commun. Comb. Optim.

*Research Article*

# Properties of subwords of binary words under Dejean morphism

**Somnath Bera[1], Sastha Sriram[2], A.S. Prasanna Venkatesan[3], Atulya K. Nagar[4,†], K.G. Subramanian[4,*]**

[1]Vellore Institute of Technology Chennai, Department of Mathematics, School of Advanced Sciences, Chennai 600036 India
somnathbera89@gmail.com

[2]Department of Mathematics, Rajalakshmi Engineering College, Thandalam, Chennai 602 105 India
sriram.discrete@gmail.com

[3]Department of Mathematics, Government Arts College, Dharmapuri, Tamil Nadu, 636 705 India
pacchu79@gacdpi.ac.in

[4]Department of Mathematics, Computer Science and Engineering, Liverpool Hope University, Liverpool L16 9JD U.K.
[†]nagara@hope.ac.uk
[*]kgsmani1948@gmail.com

*Received: 23 January 2025; Accepted: 12 November 2025*
*Published Online: 23 November 2025*

**Abstract:** A word $w$ is a finite sequence of symbols belonging to a finite set, called an alphabet. A scattered subword of a word $w$ is a subsequence of $w$. The Parikh matrix of a word $w$ over an ordered alphabet with an ordering on its elements, is an upper triangular matrix with its entries giving the counts of different occurrences of certain scattered subwords in the word $w$. Based on the notions of scattered subword and Parikh matrix, several properties of images of words under morphisms have been established. Here we consider Dejean morphism on three letters and derive several properties for images of binary words under this morphism in the context of Parikh matrices.

**Keywords:** Scattered subword, Parikh matrix, Dejean morphism.

**AMS Subject classification:** 68R15

* *Corresponding Author*

## 1. Introduction

The topic of "Combinatorics on words" [5, 11] is a subfield of the field of Discrete Mathematics having connections with concepts in different fields, especially in theoretical computer science. Research in this area has dealt with many challenging problems related to sequences of symbols, which are called words. As early as in the beginning years of 20th century, Axel Thue [4] has done pioneering work on various combinatorial problems related to words. There has been continued research on various combinatorial problems and properties of words.
Mateescu et al. [14] introduced the notion of Parikh matrix of a word which is a new research area in this field of combinatorics on words. Parikh matrix of a word extends the classical notion of Parikh vector [15] of a word which counts the number of times a symbol occurs in the word. The Parikh matrix of a word w over an alphabet $\{a_1, a_2, \ldots, a_k\}$ with an ordering $a_1 < a_2 < \cdots < a_k$, is an upper triangular matrix, with 1's on the main diagonal and 0's below it but the entries above the main diagonal provide information on the number of certain scattered subwords (also called subwords) in $w$. In fact the Parikh matrix has the Parikh vector in the second diagonal above the main diagonal. Several investigations establishing various properties of words based on the Parikh matrix have been done.
Establishing properties of Parikh matrices of words under certain mappings, called morphisms on words, is an area of fairly recent investigation. Atanasiu [2] considers Istrail morphisms and investigates an important notion called amiability (also called ambiguity) of words in the study of Parikh matrices while in [9, 13, 21] certain properties of Parikh matrices of words on two or three letters based on extensions of Thue morphism [17, 18] and the Fibonacci morphism [16] are obtained. In fact several studies investigating various properties of words based on the Parikh matrix have taken place (see, for example [1–3, 7, 9, 19, 20, 22, 23] and the references therein).
Here we consider Dejean morphism [8] which can be considered as an extension of Thue morphism. We derive properties of subwords of images of words under Dejean morphism and Parikh matrix equivalence of certain special kinds of words over an ordered binary/ternary alphabet.

## 2. Basic Definitions and Results

We recall certain basic concepts and results [11, 14] which are needed in the subsequent sections.
An ordered alphabet is a finite set of symbols, called an alphabet, having an ordering on the symbols. We denote the ordering by $<$. As an example, the binary alphabet $\{a, b\}$ having an ordering $a < b$, is an ordered alphabet and we write this alphabet as $\{a < b\}$. Likewise the ternary alphabet $\{a, b, c\}$ having an ordering $a < b < c$ is written as $\{a < b < c\}$. A word over an alphabet $\Sigma$, is a finite sequence of symbols taken from $\Sigma$. For example, *bcacbab* is a word over the alphabet $\{a, b, c\}$. A scattered subword or simply a subword $u$ of a word $w$ is a subsequence of the word $w$. We denote

the number of subwords $u$ in a word $w$ by $|w|_u$. As an illustration, in the word *abbbaba* over $\{a < b\}$, the first letter $a$ followed by the fourth letter $b$ and then followed by the last letter $a$ gives the word *aba* which is a subword of *abbbaba*. In fact the number of such subwords *aba* in *abbbaba* over $\{a < b\}$, is $|abbbaba|_{aba} = 8$.
The Parikh vector [15] of a word $w$ over an alphabet $\Sigma = \{a < b\}$ is given by $(n_a(w), n_b(w))$ where $n_x(w)$ is the number of occurrences of the symbol $x$ in the word $w$. If the alphabet is $\{a < b < c\}$, then the Parikh vector of $w$ is $(n_a(w), n_b(w), n_c(w))$. For example, the Parikh vector of the word *abbaabbb* over the alphabet $\{a < b\}$ is $(3, 5)$ while the Parikh vector of the word *acbccba* over $\{a < b < c\}$ is $(2, 2, 3)$. The notion of *Parikh matrix* of a word $w$ [14], is an extension of the notion of Parikh vector of $w$. The notion of Parikh matrix of a word is now recalled. But we mainly deal with the binary ordered alphabet $\{a < b\}$ and the ternary ordered alphabet $\{a < b < c\}$. Throughout this paper, we denote them by $\Sigma_2$ and $\Sigma_3$ respectively. For a formal definition of the Parikh matrix of a word over any ordered alphabet, we refer to [14]. Let $\Sigma_2 = \{a < b\}$ and $\Sigma_3 = \{a < b < c\}$. The Parikh matrices $M(\alpha)$ of a word $\alpha$ over $\Sigma_2$ and $M(\beta)$ of a word $\beta$ over $\Sigma_3$ are given by

$$M(\alpha) = \begin{pmatrix} 1 & |\alpha|_a & |\alpha|_{ab} \\ 0 & 1 & |\alpha|_b \\ 0 & 0 & 1 \end{pmatrix} \text{ and } M(\beta) = \begin{pmatrix} 1 & |\beta|_a & |\beta|_{ab} & |\beta|_{abc} \\ 0 & 1 & |\beta|_b & |\beta|_{bc} \\ 0 & 0 & 1 & |\beta|_c \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

We recall the technique [14] of obtaining the Parikh matrix of a word over the binary ordered alphabet $\Sigma_2 = \{a < b\}$. A $3 \times 3$ triangular matrix is associated with each of $a$ and $b$ as follows:

$$a \mapsto M(a) = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, b \mapsto M(b) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$

For a binary word $\alpha = \alpha_1\alpha_2 \cdots \alpha_n$ where $\alpha_i \in \Sigma_2$, the Parikh matrix $M(\alpha)$ of the word $\alpha$ is the matrix product $M(\alpha) = M(\alpha_1)M(\alpha_2) \cdots M(\alpha_n)$. If the alphabet is a ternary ordered alphabet $\Sigma_3 = \{a < b < c\}$, the technique is similar except that with each of $a, b$ and $c$, a $4 \times 4$ triangular matrix is associated as follows:

$$a \mapsto M(a) = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, b \mapsto M(b) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

$$c \mapsto M(c) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

**Example 1.** Consider the word $u = abab$ over $\Sigma_2$. The Parikh matrix $M(u)$ is the matrix product

$$\begin{aligned} M(u) &= M(a)M(b)M(a)M(b) \\ &= \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}. \end{aligned}$$

Two words $w_1, w_2$ over $\Sigma_2$ or $\Sigma_3$ are said to be $M$-equivalent, if the Parikh matrices of $w_1$ and $w_2$ are the same i.e. $M(w_1) = M(w_2)$. In this case, $w_1$ (as well as $w_2$) is said to be $M-$ambiguous.
For example, the words $u = ababaab, v = abaabba$ over $\{a < b\}$ are $M$-equivalent since they have the same Parikh matrix

$$M(u) = M(v) = \begin{pmatrix} 1 & 4 & 7 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{pmatrix}$$

and hence $u$ and $v$ are $M-$ambiguous.
We next recall the notion of a morphism [11] on words.
Let $\Sigma$ and $\Gamma$ be two alphabets. A morphism on $\Sigma^*$ is a mapping $\phi : \Sigma^* \to \Gamma^*$ such that $\phi(uv) = \phi(u)\phi(v)$, for words $u, v \in \Sigma^*$.

## 3. Dejean Morphism

Thue morphism [11, 16] (also called Thue-Morse morphism), is a special type of morphism, well-known and widely used and investigated in different studies on combinatorics on words (see, for example, [16–18]). It is a mapping $\mu$ on $\Sigma^*$, where $\Sigma = \{a, b\}$ and is given by $\mu(a) = ab, \mu(b) = ba$. Dejean [8] has introduced a morphism which is an extension of Thue morphism as mentioned in [6]. We recall the definition of this morphism restricting the alphabet to three symbols.

**Definition 1.** Let $\Sigma_3 = \{a < b < c\}$. The Dejean morphism $d : \Sigma_3^* \mapsto \Sigma_3^*$ is defined by

$$d(a) = abcacbcabcbacbcacba,$$

$$d(b) = bcabacabcacbacabacb,$$

$$d(c) = cabcbabcabacbabcbac.$$

Note that $|d(a)|_a = 6; |d(a)|_b = 6; |d(a)|_c = 7;$ $|d(b)|_a = 7; |d(b)|_b = 6; |d(b)|_c = 6;$ $|d(c)|_a = 6; |d(c)|_b = 7; |d(c)|_c = 6;$ $|d(a)|_{ab} = 18; |d(a)|_{bc} = 21;$ $|d(b)|_{ab} = 21; |d(b)|_{bc} = 18;$ $|d(c)|_{ab} = 21; |d(c)|_{bc} = 21;$ $|d(a)|_{abc} = 46; |d(b)|_{abc} = 31;$
$|d(c)|_{abc} = 49$.
Formulae to count the number of each of the symbols $a, b, c$ and subwords $ab, bc, abc$ in the image word $d(w)$ of a word $w \in \Sigma_2^*$ were derived in [24]. We state these formulae here and provide proofs for completeness. We need the known results in Lemma 1 and Lemma 2.

**Lemma 1.** *[19] Let $\Gamma_1, \Gamma_2$ be two ordered alphabets and $f : \Gamma_1^* \longrightarrow \Gamma_2^*$ be a morphism. Then, for all $w \in \Gamma_1^*$, and for all $a \in \Gamma_2$, we have*

$$|f(w)|_a = \sum_{r \in \Gamma_1} |w|_r \cdot |f(r)|_a.$$

**Lemma 2.** *[19] Let $f : \Sigma_2^* \longrightarrow \Sigma_3^*$ be a morphism. For $x, y \in \Sigma_3$, $w \in \Sigma_2^+$, we have*

$$|f(w)|_{xy} = \sum_{r \in \Sigma_2} |w|_r |f(r)|_{xy} + \sum_{r,t \in \Sigma_2} |w|_{rt} |f(r)|_x |f(t)|_y.$$

**Theorem 1.** *For any word $w \in \Sigma_2^*$, we have*

*(i)* $|d(w)|_a = 6|w| + |w|_b, |d(w)|_b = 6|w|, |d(w)|_c = 6|w| + |w|_a$

*(ii)* $|d(w)|_{ab} = 21|w|^2 - 3|w|_a^2 - 6|w|_{ab}$

*(iii)* $|d(w)|_{bc} = 21|w|^2 - 3|w|_b^2 - 6|w|_{ab}$

*(iv)* $|d(w)|_{abc} = 126(|w|_a^2 + |w|_b^2) - 80|w|_a - 95|w|_b + 294[|w|_{ba} + |w|_{baa} + |w|_{bba}] + 216[|w|_{ab} + |w|_{aab} + |w|_{abb}] + 252[|w|_{aaa} + |w|_{bab} + |w|_{aba} + |w|_{bbb}]$

*Proof.* Using Lemma 1, we have

$$|d(w)|_a = \sum_{r \in \Sigma_2} |w|_r \cdot |d(r)|_a = 6|w|_a + 7|w|_b = 6|w| + |w|_b,$$

since $|d(a)|_a = 6, |d(b)|_a = 7$ and $|w|_a + |w|_b = |w|$. Likewise the formulae for $|d(w)|_b$ and $|d(w)|_c$ can be derived. This proves $(i)$.

Using Lemma 2 and the known identity $|w|_{ab} + |w|_{ba} = |w|_a \times |w|_b$, we have

$$\begin{aligned}
|d(w)|_{ab} &= |w|_a|d(a)|_{ab} + |w|_b|d(b)|_{ab} + |w|_{aa}|d(a)|_a|d(a)|_b + |w|_{ab}|d(a)|_a|d(b)|_b \\
&\quad +|w|_{ba}|d(b)|_a|d(a)|_b + |w|_{bb}|d(b)|_a|d(b)|_b \\
&= 18|w|_a + 21|w|_b + \frac{1}{2} \times 36|w|_a(|w|_a - 1) + 36|w|_{ab} + 42|w|_{ba} \\
&\quad +\frac{1}{2} \times 42|w|_b(|w|_b - 1) \\
&= 21(|w|_a^2 + |w|_b^2) + 42(|w|_{ab} + |w|_{ba}) - 3|w|_a^2 - 6|w|_{ab} \\
&= 21|w|^2 - 3|w|_a^2 - 6|w|_{ab} \text{ , since } |w|_{ab} + |w|_{ba} = |w|_a|w|_b, \\
&\qquad |w|_a + |w|_b = |w|,\ |d(a)|_a = |d(a)|_b = 6,\ |d(b)|_a = 7, \\
&\qquad |d(b)|_b = 6,\ |d(a)|_{ab} = 18, |d(b)|_{ab} = 21.
\end{aligned}$$

This proves statement $(ii)$.
The statement $(iii)$ can be proved in a similar manner.
In order to prove statement $(iv)$, we have

$$\begin{aligned}
|d(w)|_{abc} &= |w|_a|d(a)|_{abc} + |w|_b|d(b)|_{abc} + |w|_{aa}(|d(a)|_a|d(a)|_{bc} + |d(a)|_{ab}|d(a)|_c) \\
&\quad +|w|_{ab}(|d(a)|_a|d(b)|_{bc} + |d(a)|_{ab}|d(b)|_c) + |w|_{ba}(|d(b)|_a|d(a)|_{bc} \\
&\qquad +|d(b)|_{ab}|d(a)|_c) + |w|_{bb}(|d(b)|_a|d(b)|_{bc} + |d(b)|_{ab}|d(b)|_c) + \eta \\
&= 46|w|_a + 31|w|_b + \frac{1}{2}|w|_a(|w|_a - 1)(6 \times 21 + 18 \times 7) \\
&\quad +|w|_{ab}(6 \times 18 + 18 \times 6) + |w|_{ba}(7 \times 21 + 21 \times 7) \\
&\qquad +\frac{1}{2}|w|_b(|w|_b - 1)(7 \times 18 + 21 \times 6) + \eta
\end{aligned}$$

where

$$\begin{aligned}
\eta &= |w|_{aaa}(|d(a)|_a|d(a)|_b|d(a)|_c) + |w|_{aab}(|d(a)|_a|d(a)|_b|d(b)|_c) \\
&\quad +|w|_{aba}(|d(a)|_a|d(b)|_b|d(a)|_c) + |w|_{abb}(|d(a)|_a|d(b)|_b|d(b)|_c) \\
&\qquad +|w|_{bba}(|d(b)|_a|d(b)|_b|d(a)|_c) + |w|_{bbb}(|d(b)|_a|d(b)|_b|d(b)|_c) \\
&= 252(|w|_{aaa} + |w|_{aba} + |w|_{bab} + |w|_{bbb}) + 216(|w|_{aab} + |w|_{abb}) \\
&\qquad +294(|w|_{baa} + |w|_{bba})
\end{aligned}$$

Thus

$$|d(w)|_{abc} = 126((|w|_a)^2 + (|w|_b)^2) - 80|w|_a - 95|w|_b + 216|w|_{ab} + 294|w|_{ba}$$

$$+252(|w|_{aaa} + |w|_{aba} + |w|_{bab} + |w|_{bbb}) + 216(|w|_{aab} + |w|_{abb}) + 294(|w|_{baa} + |w|_{bba}).$$

This proves statement $(iv)$. □

Now we consider images of binary words under Dejean morphism and deal with the problem of $M$-equivalence of words. We need the following Lemma 3.

**Lemma 3.** *[7] Let $\alpha, \beta$ be any two $M$-equivalent words over $\Sigma_2$ and $\phi : \Sigma_2^* \mapsto \Sigma_3^*$ be a morphism Then $M(\phi(\alpha)) - M(\phi(\beta))$ is a $4 \times 4$ matrix with all entries zero except the entry in the first row and last column, which is $n = |\phi(\alpha)|_{abc} - |\phi(\beta)|_{abc}$, an integer.*

If $u$ and $v$ are two $M$-equivalent words over $\{a < b\}$, then $d(u)$ and $d(v)$ need not be equivalent. For example, the binary words $aabba$ and $abaab$ are $M$-equivalent while $d(aabba)$ and $d(abaab)$ are not $M$-equivalent by Lemma 3, since $|d(aabba)|_{abc} \neq |d(abaab)|_{abc}$. We show that the converse, namely $M(d(u)) = M(d(v))$ implies $M(u) = M(v)$ holds whenever the binary words have the same Parikh vector.

**Theorem 2.** *If $u$ and $v$ are two words over $\Sigma_2^*$ such that (i) $|u|_a = |v|_a$, $|u|_b = |v|_b$ and (ii) $d(u)$ and $d(v)$ are $M$-equivalent, then $u$ and $v$ are $M$-equivalent.*

*Proof.* Since $u$ and $v$ are two words having same Parikh vector over $\Sigma_2^*$, we have $|u|_a = |v|_a$, $|u|_b = |v|_b$ and $|u| = |v|$. Using Theorem 1 we see that

$$\begin{aligned} |u|_{ab} &= \frac{1}{6}(21|u|^2 - 3|u|_a^2 - |d(u)|_{ab}) \\ &= \frac{1}{6}(21|v|^2 - 3|v|_a^2 - |d(v)|_{ab}) \text{ , since } d(u) \text{ and } d(v) \text{ are } M\text{-equivalent} \\ &= |v|_{ab} \end{aligned}$$

□

A sufficient condition on two binary words $u, v$ is now derived in order that the Parikh matrices of the words $d(u)$ and $d(v)$ commute. We recall a property, called weak ratio property of words, considered in [20] and state a known result [12] needed to obtain this condition.

Two words $u, v$ over $\{a < b < c\}$ satisfy a weak-ratio property [20], written $u \sim_{wr} v$, if there is a constant $k > 0$, such that $|u|_a = k|v|_a$, $|u|_b = k|v|_b$ and $|u|_c = k|v|_c$. If the alphabet is $\{a < b\}$, $u \sim_{wr} v$, if there is a constant $k > 0$, such that $|u|_a = k|v|_a$ and $|u|_b = k|v|_b$.

**Lemma 4.** *[12] Let $u, v$ be two words over $\Sigma_3^*$ satisfying (i) the weak-ratio property, namely, $u \sim_{wr} v$ and (ii) $|u|_a|u|_{bc} = |u|_{ab}|u|_c$ and $|v|_a|v|_{bc} = |v|_{ab}|v|_c$. Then the words $uv, vu$ are $M$-equivalent i.e. $M(uv) = M(vu)$.*

**Theorem 3.** *Let the words $u, v \in \Sigma_2^*$ be such that $|u|_a = |u|_b$ and $|v|_a = |v|_b$ Then $M(d(uv)) = M(d(vu))$ so that $d(uv), d(vu)$ are $M$-equivalent, where $d$ is the Dejean morphism.*

*Proof.* Since $|u|_a = |u|_b$ and $|v|_a = |v|_b$, the words $u, v$ satisfy the weak-ratio property. Also, using Theorem 1 , we have

$$\begin{aligned} |d(u)|_a|d(u)|_{bc} &= (6|u| + |u|_b)\left[21|w|^2 - 3|w|_a^2 - 6|w|_{ab})\right] \\ &= \left[21|w|^2 - 3|w|_b^2 - 6|w|_{ab})\right](6|u| + |u|_a) \\ &= |d(u)|_{ab}|d(u)|_c \end{aligned}$$

Likewise $|d(v)|_a|d(v)|_{bc} = |d(v)|_{ab}|d(v)|_c$. Hence by Lemma 4,

$$M(d(uv)) = M(d(vu)).$$

□

Atanasiu [2] has considered some specific binary words and has shown that the images of these words under a special morphism, known as Istrail morphism, are not $M$-equivalent. Here we consider the images of these binary words under Dejean morphism and note that they are $M$-equivalent under certain conditions. We also consider an antimorphic involution [10] $\theta$ with $\theta(a) = b$, $\theta(b) = a$ and alphabet $\Sigma_2$ satisfying $\theta(uv) = \theta(v)\theta(u)$ for words $u, v$ over $\Sigma_2$. We then prove that images under the Dejean morphism of a word $u$ over $\Sigma_2$ and $\theta(u)$ are $M$-equivalent whenever $u, \theta(u)$ are $M$-equivalent. First we recall some results from [7] and [19].

**Lemma 5.** *[19] (a) Let $\Sigma_2 = \{a < b\}$, $\Sigma_3 = \{a < b < c\}$. Let $w = ab\alpha ba$ and $w' = ba\alpha ab$ where $\alpha \in \Sigma_2^*$ and $|\alpha|_a = |\alpha|_b$. Let $\phi : \Sigma_2^* \to \Sigma_3^*$ be a morphism such that (i) $|\phi(a)|_a = |\phi(b)|_c; |\phi(b)|_a = |\phi(a)|_c$ and (ii) $M(\phi(abba)) = M(\phi(baab))$. Then $M(\phi(w)) = M(\phi(w'))$.*
*(b) Let $\Sigma$ and $\Sigma'$ be two finite ordered alphabets. Let $\phi : \Sigma^* \to \Sigma'^*$ be a morphism. Let $u, v$ be two words over $\Sigma$ satisfying the weak ratio property. Then the image words $\phi(u)$ and $\phi(v)$ also satisfy weak ratio property.*

**Theorem 4.** *Let $\delta \in \Sigma_2^*$ with an equal number of $a's$ and $b's$. Then the word $d(ab\delta ba)$ is $M$-equivalent to the corresponding word $d(ba\delta ab)$ where $d$ is the Dejean morphism.*

*Proof.* We have $|\delta|_a = |\delta|_b$ and from Theorem 1, $|d(a)|_a = 6 = |d(b)|_c$; $|d(b)|_a = 7 = |d(a)|_c$. Also, it can be seen that $M(d(abba) = M(d(baab))$, on taking $u = ab$ and $v = ba$ in Theorem 3. Thus by Lemma 5, we have $M(d(ba\delta ab)) = M(d(ab\delta ba))$. □

**Theorem 5.** *Let the words $u, v \in \Sigma_2^*$ be such that $\theta(u) = v$ and $M(u) = M(v)$, where $\theta$ is the antimorphic involution given by $\theta(a) = b, \theta(b) = a$. Then $M(d(u)) = M(d(v))$, where $d$ is the Dejean morphism.*

*Proof.* Since $M(u) = M(v)$, we have $|u|_a = |v|_a$, $|u|_b = |v|_b$ and so the words $u, v$ satisfy the weak-ratio property. Therefore $d(u)$ and $d(v)$ satisfy the weak ratio property by Lemma 5. Also $|u|_{ab} = |v|_{ab}$. Using Lemma 3, $d(u)$ and $d(v)$ are $M$-equivalent if $|d(u)|_{abc} - |d(v)|_{abc} = 0$. Now using the formulae (Theorem 1) for counting subword $abc$ under Dejean morphism [8] we obtain

$$\begin{aligned}
|d(u)|_{abc} &= 126(|u|_a^2 + |u|_b^2) - 80|u|_a - 95|u|_b + 294[|u|_{ba} + |u|_{baa} + |u|_{bba}] \\
&\quad +216[|u|_{ab} + |u|_{aab} + |u|_{abb}] + 252[|u|_{aaa} + |u|_{bab} + |u|_{aba} + |u|_{bbb}] \\
&= 126(|v|_a^2 + |v|_b^2) - 80|v|_a - 95|v|_b + 294[|v|_{ba} + |\theta(u)|_{\theta(baa)} \\
&\quad +|\theta(u)|_{\theta(bba)}] + 216[|v|_{ab} + |\theta(u)|_{\theta(aab)} + |\theta(u)|_{\theta(abb)}] \\
&\quad +252[|v|_{aaa} + |\theta(u)|_{\theta(bab)} + |\theta(u)|_{\theta(aba)} + |v|_{bbb}] \\
&= 126(|v|_a^2 + |v|_b^2) - 80|v|_a - 95|v|_b + 294[|v|_{ba} + |v|_{bba} + |v|_{baa}] \\
&\quad +216[|v|_{ab} + |v|_{abb} + |v|_{aab}] + 252[|v|_{aaa} + |v|_{aba} + |v|_{bab} + |v|_{bbb}] \\
&= |d(v)|_{abc}.
\end{aligned}$$

This proves the Theorem. □

In more general terms, we give the following necessary and sufficient condition for the morphic images of two binary words $u$ and $v$ to be $M$-equivalent.

**Theorem 6.** *Let the words $u$ and $v$ be two $M$-equivalent binary words over $\Sigma_2$ satisfying the condition* $49(|u|_{baa}+|u|_{bba})+36(|u|_{aab}+|u|_{abb})+43(|u|_{bab}+|u|_{aba}) = 49(|v|_{baa}+|v|_{bba})+36(|v|_{aab}+|v|_{abb})+43(|v|_{bab}+|v|_{aba})$. *Then $d(u)$ and $d(v)$ are $M$-equivalent.*

*Proof.* Let the words $u$ and $v$ be $M$-equivalent words satisfying the above condition, then we have $|u|_a = |v|_a$, $|u|_b = |v|_b$, $|u| = |v|$, $|u|_{ab} = |v|_{ab}$. Therefore, $|u|_{ba} = |v|_{ba}$, $|u|_{aaa} = |v|_{aaa}$ and $|u|_{bbb} = |v|_{bbb}$. □

Now using the formulae as in Theorem 1 for counting scattered subword $a, b, c, ab, bc$ and $abc$, it can be easily shown that $d(u)$ and $d(v)$ are $M$-equivalent.

**Remark 1.** The set of words satisfying the condition in Theorem 6 is not empty. For example, take the words *abaab* and *aabba*.

Atanasiu [2] has proved a necessary and sufficient condition ([2, Theorem 7]) relating to $M$-equivalence under Istrail morphism. Here we obtain an analogue of this result as a corollary of Theorem 6.

**Corollary 1.** *Let the words $u$ and $v$ be two words over $\{a < b\}$ satisfying $|u|_x = |v|_x$, for $x \in \{baaa, bba, aab, abb, bab, aba\}$. Then $d(u)$ and $d(v)$ are $M$-equivalent if and only if $u$ and $v$ are $M$-equivalent.*

## 4. Conclusions

We have derived certain properties related to subwords and Parikh matrices of the image words under Dejean morphism, which in some sense is an extension of Thue morphism. Construction of other kinds of $M$-equivalent binary words whose images under special morphisms on three or more letters are $M$-equivalent, is a problem of interest for future work.

**Statements and Declarations**

The authors declare that no funds, grants, or other support were received during the preparation of this manuscript. The authors have no relevant financial or non-financial interests to disclose. All authors contributed to the study conception and design and commented on previous versions of the manuscript. All authors read and approved the final manuscript. This research did not generate or analyze any datasets.

## References

[1] A. Atanasiu, *Morphisms on amiable words*, Annals of Bucharest University **59** (2010), 99–111.

[2] ———, *Parikh matrices, amiability and istrail morphism*, Int. J. Found. Comput. Sci. **21** (2010), no. 06, 1021–1033. https://doi.org/10.1142/S0129054110007702.

[3] A. Atanasiu, K. Mahalingam, and K.G. Subramanian, *Morphisms and weak-ratio property*, 11 2014, pp. 13–21.

[4] J. Berstel, *Axel Thue's Papers on Repetitions in Words: a Translation*, vol. 20, Départements de mathématiques et d'informatique, Université du Québec à Montréal, 1995.

[5] J. Berstel and D. Perrin, *The origins of combinatorics on words*, European J. Combin. **28** (2007), no. 3, 996–1022. https://doi.org/10.1016/j.ejc.2005.07.019.

[6] F.J. Brandenburg, *Uniformly growing k-th power-free homomorphisms*, Theor. Comput. Sci. **23** (1988), no. 1, 69–82. https://doi.org/10.1016/0304-3975(88)90009-6.

[7] W. Chean Teh, *Separability of M-equivalent words by morphisms*, Int. J. Found. Comput. Sci. **27** (2016), no. 1, 39–52. https://doi.org/10.1142/S0129054116500039.

[8] F. Dejean, *Sur un théoreme de thue*, J. Comb. Theory Ser. A. **13** (1972), no. 1, 90–99. https://doi.org/10.1016/0097-3165(72)90011-8.

[9] P. Isawasan, I. Venkat, K.G. Subramanian, and N.H. Sarmin, *Morphic images of binary words and parikh matrices*, AIP Conf. Proc. **1605** (2014), no. 1, 708–713. https://doi.org/10.1063/1.4887676.

[10] L. Kari and K. Mahalingam, *Involutively bordered words*, Int. J. Found. Comput. Sci. **18** (2007), no. 05, 1089–1106.
https://doi.org/10.1142/S0129054107005145.

[11] M. Lothaire, *Combinatorics on Words*, 2 ed., Cambridge University Press, Cambridge, 1997.

[12] K. Mahalingam and K.G. Subramanian, *Product of parikh matrices and commutativity*, Int. J. Found. Comput. Sci. **23** (2012), no. 1, 207–223.
https://doi.org/10.1142/S0129054112500049.

[13] A. Mateescu, *Algebraic aspects of parikh matrices*, Theory Is Forever: Essays Dedicated to Arto Salomaa on the Occasion of His 70th Birthday (J. Karhumäki, H. Maurer, G. Păun, and G. Rozenberg, eds.), Springer Berlin Heidelberg, Berlin, Heidelberg, 2004, pp. 170–180.
https://doi.org/10.1007/978-3-540-27812-2_16.

[14] A. Mateescu, A. Salomaa, K. Salomaa, and S. Yu, *A sharpening of the Parikh mapping*, RAIRO - Theor. Inform. Appl. **35** (2001), no. 6, 551–564.
https://doi.org/10.1051/ita:2001131.

[15] G. Rozenberg and A. Salomaa, *Handbook of Formal Languages: Volume 3 Beyond Words*, Springer Science & Business Media, 2012.

[16] P. Séébold, *Sequences generated by infinitely iterated morphisms*, Discret. Appl. Math. **11** (1985), no. 3, 255–264.
https://doi.org/10.1016/0166-218X(85)90077-0.

[17] ———, *Lyndon factorization of the prouhet words*, Theor. Comput. Sci. **307** (2003), no. 1, 179–197.
https://doi.org/10.1016/S0304-3975(03)00100-2.

[18] ———, *On some generalizations of the Thue–Morse morphism*, Theor. Comput. Sci. **292** (2003), no. 1, 283–298.
https://doi.org/10.1016/S0304-3975(01)00228-6.

[19] K.G. Subramanian, *On prouhet morphisms and Parikh matrices*, The 4th Int. Conf. Math. Sci., Universiti Kebangsaan Malaysia, 2016.

[20] K.G. Subramanian, A.M. Huey, and A.K. Nagar, *On Parikh matrices*, Int. J. Found. Comput. Sci. **20** (2009), no. 2, 211–219.
https://doi.org/10.1142/S0129054109006528.

[21] K.G. Subramanian, P. Isawasan, I. Venkat, and A.T. Khader, *Parikh matrices of words under a generalized Thue morphism*, AIP Conf. Proc. **1558** (2013), no. 1, 2516–2519.
https://doi.org/10.1063/1.4826052.

[22] ———, *Parikh matrices of words under a generalized Thue morphism*, AIP Conf. Proc. **1558** (2013), no. 1, 2516–2519.
https://doi.org/10.1063/1.4826052.

[23] K.G. Subramanian, A. Nagar, and S. Sriram, *Parikh matrices of binary and Ternary words under prouhet morphism*, Int. J. Math. Comb. **1** (2018), 78–88.

[24] K.G. Subramanian, S. Sriram, and A.S. Prasanna Venkatesan, *Parikh matrix of a word under Dejean morphism*, 13th Annual ADMA Conference and Graph Theory Day-XIII, India, 2017.