ROMANIAN JOURNAL OF INFORMATION SCIENCE AND TECHNOLOGY
Volume X, Number X, XXXX, XXX–XXX

On a special variant of Rauzy graphs

K.Mahalingam[1], R.Praveen[2], and R.Rama[3]

[1,2,3]Department of Mathematics, Indian Institute of Technology, Madras, Guindy, Chennai-600036, India
[1,2,3]Email: kmahalingam@iitm.ac.in, ma16d039@smail.iitm.ac.in, ramar@iitm.ac.in

Abstract. Representation of a word by a graph is an important branch of study in combinatorics of words. The vertices and the edges are defined using certain properties of words. Such a representation is useful in solving some problems in graph theory. One such representation is the idea of Rauzy graphs that was used to study Arnoux-Rauzy sequences. In this paper, we define half range Rauzy graphs and study the structural properties of such a variant for some special words.

Keywords: Half range Rauzy graphs, Primitive words, Directed graphs, Strongly connected graphs

1 Introduction

Theory of word representation graph lies in the intersection of combinatorics on words, Graph theory and Computer Science. One of the earliest study of word representable graphs is the well known De Bruijn graphs. It was later adapted by Gerard Rauzy in [14] to a graph in which each vertex is a subword of a given word. Rauzy graphs are mainly used in the study of words with low complexity [5, 8, 15]. A new class of word representable graphs has been first used for the free spectrum study of the famous Perkins semigroup [11]. A word w over some alphabet Σ, is given a graph description exploiting a particular property of the word. Hence, any word may not be suitable for its representation as a graph. For example, in [10] it is shown that Peterson graph cannot be represented by 2-uniform words. Also, all such words that are representable by graphs have many attractive properties [4]. In [9], Graham et.al., bring out the connections of word representable graphs to robot movement. In [2], Bera et.al., construct a graph for a word, based on the arrangements of scattered subwords. A complete characterization of Parikh word representable graphs over the binary alphabet in terms of chordal bipartite graph is also given in [2].

Rauzy graphs represent a set of words. The vertices of the Rauzy graphs are words of length n. Any two words w_1, w_2 have an edge between them if $w_1 = xv$, $w_2 = vy$, where x, y are symbols. An estimation of finite word complexities of infinite sequences using Rauzy graphs is given in [7]. Salimov in [16] has shown that every sequence of finite strongly connected directed graphs with bounded in-degree and out-degree, will contain a subsequence. A conjecture on automaticity function of a unary language is solved using the Rauzy graph structure in [6].

In 1994, Leonard Adleman [1] showed that DNA could be used to solve a graph theory problem. The seven node Hamiltonian path problem is solved by using single strand DNA molecule over A, G, C, T for each vertex of the graph. These strands are of length 20 out of which the suffix of 10 nucleotides of one strand is the complement of the prefix of 10 nucleotides of another strand indicating an edge between them. Mixing the DNA strands thus formed with DNA ligase and ATP (Adenosine Triphosphate) results in all paths through vertices.

A word, finite or infinite is a sequence of symbols taken from a finite alphabet. Mathematical research of words is well known as 'Combinatorics on words' and is connected to many modern as well as classical fields of mathematics. A word is primitive if it is not a proper power of a word of shorter length. The concept of primitivity of words plays a vital role in algebraic coding theory [17] and combinatorial theory of words [13].

In this paper, we introduce a new variant of Rauzy graphs, called as the half range Rauzy graphs. For any two vertices u and v, there is an edge from u to v if prefix of v is the same as that of the suffix of u of length k if $|u| = |v| = 2k$ or of length either $k + 1$ or k if $|u| = |v| = 2k + 1$. We study the graphs for a special type of infinite word of the form x^ω when x is primitive. We investigate the properties of such graphs. We initially show that Rauzy and half range Rauzy graphs are different for such type of words. The main result of this paper is the property that every component of such a half range Rauzy graph is a cycle. We also give conditions under which such a graph is strongly connected.

2 Preliminaries

In this section we recall some basic notions. For more information the reader can refer to [3,12,16]. Let $\Sigma = \{ a_1, a_2, \cdots, a_l \}$ be a finite alphabet. A word w over an alphabet Σ is a sequence (finite or infinite) of elements of Σ. The set of all finite words over an alphabet Σ is $\Sigma^+ \cup \{\lambda\}$ and is denoted by Σ^*, where Σ^+ is the set of all non-empty words and λ denotes the empty word. For $w = a_1 a_2 \cdots a_k$, $a_i \in \Sigma$, the number k is the length of the word w and it is denoted by $|w|$. A word of finite length is called a finite word, otherwise, it is infinite. We denote the set of all (right) infinite words by $\Sigma^{\mathbb{N}}$. It is the set of sequence of symbols in Σ indexed by non-negative integers. We denote $\Sigma^* \cup \Sigma^{\mathbb{N}}$ by Σ^∞, the set of finite or infinite words. A finite word x is a factor or subword of an infinite (finite) word w if

$$w = uxv, \qquad u \in \Sigma^*, v \in \Sigma^\infty$$

If $u = \lambda$ ($v = \lambda$) , then x is a prefix (suffix, respectively) of w. A set $F \subseteq \Sigma^*$ is said to be a factorial language if for any word $w = uxv \in F \Rightarrow x \in F$, where $u \in \Sigma^*, v \in \Sigma^\infty$ and $x \in \Sigma^*$. The set of all factors of length l in F is denoted by $F(l)$.

Let the i^{th} letter in a word w be denoted by w_i. A factor of $w = w_1 w_2 \cdots$ can also be represented as $w[i, i+k] = w_i w_{i+1} \cdots w_{i+k}$ for any $i, k \in \mathbb{N}$. Let $x \in \Sigma^+$ and x^ω is defined as the concatenation of infinite copies of x, i.e, $x^\omega = xxx\cdots$. An infinite word $w = w_1 w_2 w_3 \cdots$ is said to be periodic if there exists an integer p such that $w_{n+p} = w_n \ \forall \ n \in \mathbb{N}$ and w can also be written as $(w_1 w_2 \cdots w_p)^\omega$. The least number p which satisfies the above condition is called as the period of w. A non empty word $w \in \Sigma^\infty$ is said to be primitive if there is no word $x \in \Sigma^*$ such that $w = x^n$

for $n > 1$. Any two words $u, v \in \Sigma^*$ are said to be conjugates if there exist words $x, y \in \Sigma^*$ such that $u = xy$ and $v = yx$. The set of all conjugates of a word w is called as the conjugacy class of w and is denoted by $c(w)$.

Let $\Sigma = \{a, b\}$ be a binary alphabet. For any $x \in \Sigma^\infty$, x^c is defined as the word obtained by replacing $a's$ with $b's$ and vice-versa in x. For example, for a given $x = abba$, $x^c = baab$.

A graph G is an ordered pair $(V(G), E(G))$ consisting of a non empty set $V(G)$ of vertices of G, $E(G)$ is the set of edges, where each edge is a pair of vertices. A walk of length k in a graph is an alternating sequence of vertices and edges,

$$v_0, e_0, v_1, e_1, v_2, \ldots, v_{k-1}, e_{k-1}, v_k,$$

which begins and ends with vertices. A trail is a walk in which all edges are distinct. A path is a trail in which all vertices (except possibly the first and last) are distinct. A cycle is a closed trail in which all the vertices are distinct except for the first and the last vertex.

A graph is said to be connected if there is a path between every pair of vertices. A graph on n vertices is said to be complete if every pair of vertices in a graph is connected by a unique edge and is denoted by K_n. Complement of graph G is denoted by $\overline{G}$ is defined by $|V(G)| = |V(\overline{G})|$ and $(u, v) \in E(\overline{G}) \iff (u, v) \notin E(G)$. A directed graph is a graph whose edges are ordered pairs of vertices. In a directed graph edges are often called as arcs. For an arc (u, v) the first vertex u is called as the tail of the arc and the second vertex v is called as the head of the arc. We also say that the arc (u, v) leaves u and enters v. Indegree of a vertex v is defined as the number of arcs entering v and is denoted by $deg_{in}(v)$. Outdegree of a vertex v is defined as the number of arcs leaving v and is denoted by $deg_{out}(v)$. A directed graph is said to be strongly connected graph if it has a path from each vertex to every other vertex. A loop (or self-loop) is an edge from a vertex to itself.

Definition 1. *A Rauzy graph of order k for a factorial language F is a directed graph (V, E) where $V = F(k)$ and $(u, v) \in E$ iff*

$$u_2u_3 \cdots u_k = v_1v_2 \cdots v_{k-1} \qquad and \qquad u_1u_2 \cdots u_kv_k \in F(k+1)$$

A Rauzy graph of order k for an infinite word w is the Rauzy graph of order k for the language of subwords of w. We denote a Rauzy graph of order k for a factorial language F (for an infinite word w) by $R_F(k)$ (correspondingly, $R_w(k)$).

3 Half range Rauzy graphs

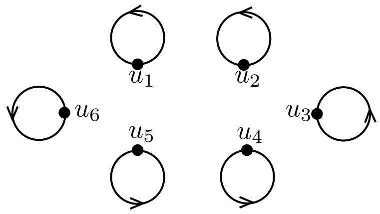
In this section we introduce a special variant of Rauzy graphs called as half range Rauzy graphs. In [1], Adelman solved an instance of a Travelling salesman problem using DNA strands. The encoding was done in such a way that vertices and edges are encoded in to a 20 length DNA-strand and an edge $e = xy$ between any two given vertices $x = x_1x_2$ and $y = y_1y_2$ is encoded into a DNA strand x_2y_1 where $|x_2| = |y_1| = 10$. Motivated by this encoding procedure of vertices and edges in to DNA-strands we define a special class of graphs.

Definition 2. *A half range Rauzy graph (HRR graph in short) of order $k > 1$, for a factorial language F is a directed graph (V, E), where $V = F(k)$ and edge set is defined as follows:*

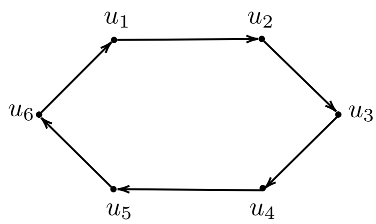
1. *For an even k, $(u,v) \in E$ iff*
   $u_{\frac{k}{2}+1}u_{\frac{k}{2}+2}\cdots u_k = v_1v_2\cdots v_{\frac{k}{2}}$ *and* $u_1u_2\cdots u_kv_{\frac{k}{2}+1}\cdots v_k \in F(\frac{3k}{2})$
2. *For an odd k, there are two types of graphs.* $(u,v) \in E$ *iff*
   Type I: $u_{\frac{k+1}{2}}u_{\frac{k+3}{2}}\cdots u_k = v_1v_2\cdots v_{\frac{k+1}{2}}$ *and* $u_1u_2\cdots u_kv_{\frac{k+3}{2}}\cdots v_k \in F(\frac{3k-1}{2})$
   Type II: $u_{\frac{k+3}{2}}u_{\frac{k+5}{2}}\cdots u_k = v_1v_2\cdots v_{\frac{k-1}{2}}$ *and* $u_1u_2\cdots u_kv_{\frac{k+1}{2}}\cdots v_k \in F(\frac{3k+1}{2})$

We denote an HRR graph for F as, $\mathbb{HR}_F(k,*) = \begin{cases} \mathbb{HR}_F(k) : k \text{ is even} \\ \mathbb{HR}_F(k,I) : k \text{ is odd and of Type} I \\ \mathbb{HR}_F(k,II) : k \text{ is odd and of Type} II \end{cases}$

If the underlying language is a set of all factors of a given word w, then $\mathbb{HR}_F(k,*)$ is simply represented as $\mathbb{HR}_w(k,*)$.



(a) $\mathbb{HR}_w(12)$



(b) $R_w(12)$

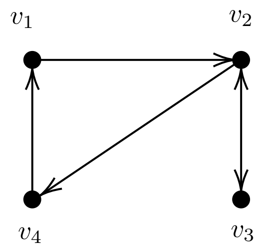
Fig. 1: Rauzy and half range Rauzy graphs of order 12

Example 1. *Let $w = (abbaaa)^\omega$ be an infinite word. half range Rauzy graph of order 12 for an infinite word w is $\mathbb{HR}_w(12) = (U,E)$, where $U : \{u_1 = abbaaaabbaaa$, $u_2 = bbaaaabbaaaa$, $u_3 = baaaabbaaaab$, $u_4 = aaaabbaaaabb$, $u_5 = aaabbaaaabba$, $u_6 = aabbaaaabbaa\}$ and arcs in E are shown in Fig. 1a. Rauzy graph of order 12 for an infinite word w is $R_w(12) = (U,E')$ and arcs in E' are shown in Fig. 1b.*

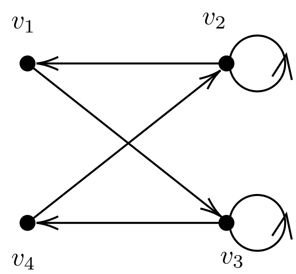
Example 2. *Let $w = (aabababaabab)^\omega$ be an infinite word. Half range Rauzy graph of order 3 has two types of directed graph $\mathbb{HR}_w(3,\ I) = (V,E_1)$ and $\mathbb{HR}_w(3,\ II) = (V,E_2)$, where $V = \{v_1 = aab,\ v_2 = aba,\ v_3 = bab,\ v_4 = baa\}$. Arcs in E_1 and E_2 are shown in Fig. 2a and Fig. 2b respectively.*

We observe that $\mathbb{HR}_w(6) = (V',E_3)$, where $V' = \{v_1' = aababa,\ v_2' = ababab,\ v_3' = bababa,\ v_4' = ababaa,\ v_5' = babaab,\ v_6' = abaaba,\ v_7' = baabab\}$ and arcs in E_3 is shown in Fig. 3a.

Note that $R_w(3) = \mathbb{HR}_w(3,\ I)$, since $|V(R_w(3))| = |V(\mathbb{HR}_w(3,\ I))|$ and $(u,v) \in F_w(4)$ for both the graphs. So, Fig. 2a is $R_w(3)$. Also, the Rauzy graph of order 6 for w, $R_w(6) = (V',E_4)$, where E_4 is shown in Fig. 3b

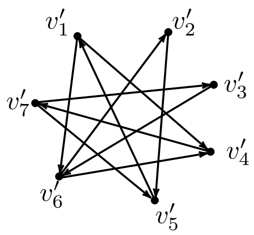


(a) $\mathbb{HR}_w(3,\ I)$

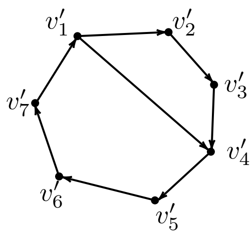


(b) $\mathbb{HR}_w(3,\ II)$

Fig. 2: $\mathbb{HR}_w(3)$



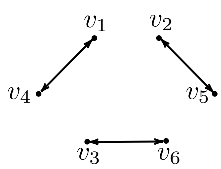
(a) $\mathbb{HR}_w(6)$



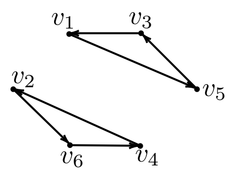
(b) $R_w(6)$

Fig. 3: $\mathbb{HR}_w(6)$ and $R_w(6)$

Example 3. *Let $w = (abbaaa)^\omega$ be an infinite word. Then the HRR graph of order 7 for an infinite word w has two types of directed graph $\mathbb{HR}_w(7,\ I) = (V, E_1)$ and $\mathbb{HR}_w(7,\ II) = (V, E_2)$, where $V :\ \{v_1 = abbaaaa\ ,\ v_2 = bbaaaab,\ v_3 = baaaabb,$ $v_4 = aaaabba,\ v_5 = aaabbaa,\ v_6 = aabbaaa\}$ and the arcs in the set E_1 and E_2 are shown in Fig. 4a and Fig. 4b respectively.*



(a) $\mathbb{HR}_w(7,\ I)$



(b) $\mathbb{HR}_w(7,\ II)$

Fig. 4: $\mathbb{HR}_w(7)$

In the next section we consider properties of HRR graphs for words of the form x^ω when x is a binary primitive word.

4 Properties of Half Range Rauzy graphs

Through out this section, we consider words $x \in \Sigma^*$ that are primitive over a binary alphabet $\Sigma = \{a, b\}$, with $|x| = n$. If x is not primitive, then there exists a primitive root u such that $x = u^i$ for some i and $x^\omega = u^\omega$. So, it is enough to consider words of the form x^ω when x is primitive.

Example 4. *Let $x = b^4a$ and $w = x^\omega$.*

$$w[1,3] = b^3 = w[2,4], \qquad w[3,5] = b^2a, \qquad w[4,6] = bab, \qquad w[5,7] = ab^2.$$

There are only 4 subwords of length 3 in x^ω i.e, $|V(\mathbb{HR}_w(3))| = 4 < |x|$.

We have the following observations.

Lemma 1. *For a given $w = x^\omega$, $\mathbb{HR}_w(k) \simeq \mathbb{HR}_{\overline{w}}(k)\ \forall\ k \in \mathbb{N}$.*

Proof. *Let $w = x^\omega$ and $\overline{w} = (x^c)^\omega$. A morphism $\phi : \mathbb{HR}_w(k) \to \mathbb{HR}_{\overline{w}}(k)$ is given by $\phi(y) = y^c$, where $y \in V(\mathbb{HR}_w(k))$. Also the arcs $(u,v) \in E(\mathbb{HR}_w(k)) \Leftrightarrow (u^c, v^c) \in E(\mathbb{HR}_{\overline{w}}(k))$. Hence ϕ is an isomorphism and $\mathbb{HR}_w(k) \simeq \mathbb{HR}_{\overline{w}}(k)$ for all $k \in \mathbb{N}$.* $\square$

Lemma 2. *Let $w = x^\omega$. If $y \in c(x)$, the conjugacy class of x, then $\mathbb{HR}_{y^\omega}(k) = \mathbb{HR}_{x^\omega}(k)$.*

Proof. *Let $x = x_1x_2\cdots x_n$ be a primitive word and $y^{(i)} = x_ix_{i+1}\cdots x_{n+i-1}$ for $1 \le i \le n$ are the conjugates of x. For each $w^{(i)} = (y^{(i)})^\omega$ for $1 \le i \le n$ we have $V(\mathbb{HR}_{w^{(i)}}(k)) = \{w^{(1)}[1,k], w^{(1)}[2,k+1], \cdots, w^{(1)}[j,k+j-1]\}$, $1 \le i \le n$, for some $j \le n$ and hence, $\mathbb{HR}_{w^{(i)}}(k) = \mathbb{HR}_{w^{(1)}}(k)$, $2 \le i \le n$.* $\square$

In the following we show that the maximum number of vertices in any HRR graph of order k is atmost n.

Proposition 1. *Let $w = x^\omega$ be an infinite word. Then $|V(\mathbb{HR}_w(k))| \le n$, in particular $|V(\mathbb{HR}_w(k))| = n$, $\forall\ k \ge n-1$.*

Proof. *We note that $|V(\mathbb{HR}_w(k))|$ is the number of distinct k-length subwords of x^ω, where $x = x_1x_2\cdots x_n$. If a word of length k starting with x_i is different from the one starting with x_j for $i \ne j$ for any $i, j \le n$, then there are n distinct subwords of length k. Thus if for any k, subwords of length k are repeated in the prefix of x^ω of length $(n+k-1)$, then $|V(\mathbb{HR}_w(k))| < n$.*

It is well known that for a primitive word x, the number of distinct subwords of length k in x^ω is n $\forall k \ge n-1$ and hence $|V(\mathbb{HR}_w(k))| = n$, $\forall\ k \ge n-1$. $\square$

We also count the number of edges any $\mathbb{HR}_w(k,*)$ can have in the following.

Proposition 2. *Let $w = x^\omega$ be an infinite word. Then $|E(\mathbb{HR}_w(k,*))| = n$, $\forall\ k \ge \frac{2n-1}{3}$.*

Proof. *We only prove for the case when k is even. The cases when k is odd is similar. Each arc $e \in E(\mathbb{HR}_w(k))$ implies that $e \in F_w(\frac{3k}{2})$. Given $k \ge \frac{2n-1}{3}$, length of each arc e is $\frac{3k}{2} \ge \frac{3}{2} \times \frac{2n-1}{3} > n-1$. By Proposition 1, there are n distinct subwords of length $\frac{3k}{2}$, if $\frac{3k}{2} \ge n-1$. Hence there are n distinct arcs $\forall\ k \ge \frac{2n-1}{3}$.* $\square$

In the following we show that there exists no isolated vertices in an $\mathbb{HR}_w(k,*)$ graph of $w = x^\omega$.

Proposition 3. *Let $v \in V(\mathbb{HR}_w(k,*))$, where $w = x^\omega$, $k \in \mathbb{N}$. Then,*

1. $deg_{out}(v) \ge 1$.

2. $deg_{in}(v) \geq 1$.

Proof. *In $\mathbb{HR}_w(k)$, each vertex $v \in F_w(k)$ and an arc $(u, v) \in E(\mathbb{HR}_w(k))$. By Proposition 1, there are atmost n-vertices*

$$w[1, k],\ w[2, k+1],\ w[3, k+2],\ \cdots,\ w[n, k+n-1].$$

We only show for the case when k is even as the proof is similar for an odd k.

1. *Since x is primitive, the infinite word w has period $|x|(= n)$ and hence*

$$w[ln + j, k + ln + j - 1] = w[j, k + j - 1]$$

for $1 \leq j \leq n$ and $l \in \mathbb{N}$. If k is even, then for each vertex $u = w[i, k + i - 1]$, there always exists a vertex $v = w[\frac{k}{2} + i, k + \frac{k}{2} + i - 1]$ such that $(u, v) \in E(\mathbb{HR}_w(k))$. Thus, $deg_{out}(u) \geq 1$.

2. *Let k be even and $k \leq n$. For each vertex $v = w[j, k + j - 1]$, $1 \leq j \leq \frac{k}{2}$ there exists a vertex $u = w[n - \frac{k}{2} + j, n + \frac{k}{2} + (j - 1)]$ such that $(u, v) \in E(\mathbb{HR}_w(k))$. Note that since, w is of period n, $u = w[n - \frac{k}{2} + j, n + \frac{k}{2} + (j - 1)] = w[n - \frac{k}{2} + j, n]w[1, \frac{k}{2} + j - 1]$ and hence $(u, v) \in E$. Similarly, for each vertex $v = w[\frac{k}{2} + j, k + \frac{k}{2} + j - 1]$, $1 \leq j \leq n - \frac{k}{2}$, there exists a vertex $u = w[j, k + j - 1]$ such that $(u, v) \in E(\mathbb{HR}_w(k))$. Now, let $k \geq n$ and $k = in + r,\ i > 0,\ 0 \leq r < n$. Then for each vertex $v = w[j, k + j - 1]$ there exist a vertex $u = w[(in - \frac{k}{2} + j \mod n), n + \frac{k}{2} + (j - 1)]$ such that $(u, v) \in E(\mathbb{HR}_w(k))$ for $1 \leq j \leq \frac{k}{2}$ and similarly for each vertex $v = w[\frac{k}{2} + j, k + \frac{k}{2} + (j - 1)]$ there exist a vertex $u = w[j, k + (j - 1)]$ such that $(u, v) \in E(\mathbb{HR}_w(k))$ for $1 \leq j \leq n - \frac{k}{2}$. Thus $deg_{in}(v) \geq 1$.* □

It is clear from Proposition 1 and Proposition 2 that for $k \geq n-1$ the corresponding HRR graph has exactly n vertices and n edges. Thus we have the following corollary.

Corollary 1. *Let $w = x^\omega$. If $k \geq n-1$, then for each $v \in V(\mathbb{HR}_w(k, *))$, $deg_{in}(v) = deg_{out}(v) = 1$.*

Corollary 2. *Let $w = x^\omega$. If $k \geq n - 1$, then each component of $\mathbb{HR}_w(k, *)$ is a directed cycle with atleast 2 vertices or a self loop.*

Proof. *By Corollary 1, none of the vertices in $\mathbb{HR}_w(k, *)$ are isolated. If any of its component is a tree, then there exist atleast two vertices u, v such that $deg_{in}(u) = 1 = deg_{in}(v)$ and $deg_{out}(u) = 0 = deg_{out}(v)$, which is a contradiction. If any of its component contains a cycle along with some more arcs (say (u, v)), then $deg_{out}(u) > 1$ or $deg_{in}(v) > 1$ or both, which is again a contradiction. Hence each component of $\mathbb{HR}_w(k, *)$ is a directed cycle with atleast 2 vertices or a self loop.* □

We know that for any w of the form x^ω, the HRR graph of w has only either a self loop or a directed cycle with atleast 2 vertices as its components. We investigate the total number of components in any given graph and the number of such components that are t-cycles for some t.

Theorem 1. *Let $w = x^\omega$.*

1. *$\mathbb{HR}_w(2k)$, $\mathbb{HR}_w(2k+1,\ I)$ has only self loops $\iff$ $2k = rn$, where $r \in 2\mathbb{N}$.*

2. $\mathbb{HR}_w(2k-1,\ II)$ *has only self loops* $\iff$ $2k = rn > 2$, *where* $r \in 2\mathbb{N}$.

Proof. Case 1: $n = 1$. In $\mathbb{HR}_w(l,\ *)$ where $l \in \mathbb{N}\backslash\{1\}$, there is only one vertex as a word of length l. And this vertex has a self loop for all l. As $|x| = 1$, $2k = r$ is always even.

Case 2: $n > 1$. In $\mathbb{HR}_w(2k)$, every vertex is an r^{th} power of conjugate y of x. If r is even, then every vertex $y^r = y^{\frac{r}{2}}y^{\frac{r}{2}}$ has prefix $y^{\frac{r}{2}}$ of each vertex is same as suffix $y^{\frac{r}{2}}$ of the same vertex.

Conversely, if there are only self loops in $\mathbb{HR}_w(2k)$. Suppose r is odd, then any vertex is of the form $y^{\frac{r-1}{2}}y[1, \frac{n}{2}]y[\frac{n}{2}+1, n]y^{\frac{r-1}{2}}$ and $y[\frac{n}{2}+1, n]y^{\frac{r-1}{2}} = y^{\frac{r-1}{2}}y[1, \frac{n}{2}]$ which impies $y[1, \frac{n}{2}] = y[\frac{n}{2}+1, n]$. This is a contradiction to the fact that y is a primitive word (as it is conjugate to primitive word x). Hence r must be even.

In $\mathbb{HR}_w(2k-1,\ II)$, every vertex is of the form $y^{r-1}y[1, n-1]$ where y is the conjugate of x. If r is even, then every vertex $y^{r-1}y[1, n-1] = y^{\frac{r}{2}}y^{\frac{r}{2}-1}y[1, n-1]$ has the prefix and suffix word of length $k-1$ to be the same word $y^{\frac{r}{2}-1}y[1, n-1]$.

Conversely, suppose there are only self loops in $\mathbb{HR}_w(2k-1,\ II)$. If r is odd, then, any vertex is of the form $y^{\frac{r-1}{2}}y[1, \frac{n}{2}]y[\frac{n}{2}+1, n]y^{\frac{r-3}{2}}y[1, n-1]$ and the suffix $y[\frac{n}{2}+1, n]y^{\frac{r-3}{2}}y[1, n-1]$ is same as that of the prefix $y^{\frac{r-1}{2}}y[1, \frac{n}{2}-1]$ of length $\frac{n}{2}$ (Comparing the prefix words of lengths $\frac{n}{2}$). This is a contradiction to the fact that y is a primitive word (as it is conjugate to primitive word x). Hence r is even.

The case is similar for $\mathbb{HR}_w(2k+1,\ I)$ and hence we omit the proof.

□

Let $v_1v_2\cdots v_jv_1$ represent a cycle in $\mathbb{HR}_w(2k)$. We associate a word to this cycle as follows:

$$v_1[1, 2k]v_2[k+1, 2k]\cdots v_{j-1}[k+1, 2k]$$

Here we concatenate only suffix of length k of the vertices $v_2, \cdots, v_{j-1}$ as $v_i[1, k] = v_{i-1}[k+1, 2k]$ for $2 \leq i \leq j-1$. Also $v_j[1, k] = v_{j-1}[k+1, 2k]$ and $v_j[k+1, 2k] = v_1[1, k]$. And the length of the word formed by the cycle $v_1v_2\cdots v_jv_1$ is jk. Similarly, if $u_1u_2\cdots u_ju_1$ is a cycle in $\mathbb{HR}_w(2k-1,\ II)$, then the word associated to this cycle is

$$u_1[1, 2k-1]u_2[k, 2k-1]\cdots u_{j-1}[k, 2k-1]u_j[k, k]$$

Here we concatenate suffix of length k of the vertices $u_2, u_3, \cdots, u_{j-1}$ and k^{th} letter of u_j is taken neither in $u_j[1, k-1]$ nor in $u_j[k+1, 2k-1]$, so we concatenate k^{th} letter of u_j in the last. The length of the word formed by the cycle $u_1u_2\cdots u_ju_1$ is $2k-1+(j-2)k+1 = kj$.

In $\mathbb{HR}_w(2k+1,\ I)$, the word associated with the cycle $y_1y_2\cdots y_jy_1$ is

$$y_1[1, 2k+1]y_2[k+2, 2k+1]\cdots y_{j-2}[k+2, 2k+1]y_{j-1}[k+2, 2k]$$

Here we concatenate suffix of length k of the vertices $y_2, y_3, \cdots, y_{j-2}$. Note that, $y_{j-2}[k+2, 2k+1] = y_{j-1}[2, k+1]$ and $y_j[k+1, 2k+1] = y_1[1, k+1]$, so it enough to consider the subword $y_{j-1}[k+2, 2k]$ of $y_{j-1}[k+2, 2k+1] = y_j[2, k+1]$ at last. The length of the word formed by the cycle $y_1y_2\cdots y_jy_1$ is $2k+1+(j-3)k+(k-1) = kj$.

Theorem 2. *Let* $w = x^\omega$. *If* $2 < n < 2k$ *and* $2k \neq rn$ *where* $r \in 2\mathbb{N}$ *, then* $\mathbb{HR}_w(2k-1,\ II)$, $\mathbb{HR}_w(2k)$ *and* $\mathbb{HR}_w(2k+1,\ I)$ *have* α*-components and each component is a* β*-cycle, where* $\alpha = g.c.d\ (k, n)$ *and* $\beta = \frac{l.c.m\ (k,n)}{k}$.

Proof. By corollary 2, each of $\mathbb{HR}_w(2k-1,\ II)$, $\mathbb{HR}_w(2k)$, $\mathbb{HR}_w(2k+1,\ I)$ contain a cycle with atleast two vertices or a loop. By Theorem 1, existence of loop is eliminated and so it contains only cycle with atleast two vertices.

In $\mathbb{HR}_w(2k)$, choose a vertex v_1 (say $v_1 = w[i, 2k+i-1]$) and it belongs to a component which is a cycle. Since every vertex is a word of length $2k$, the word formed by a cycle $v_1v_2\cdots v_jv_1$ is $v_1[1,2k]v_2[k+1,2k]\cdots v_{j-1}[k+1,2k] = w[i,2k+i-1]w[2k+i,3k+i-1]\cdots w[(j-1)k+i, jk+i-1] = w[i, jk+i-1]$ and the length of the word formed is jk. As the vertex v_j makes an arc with v_1, the length of the word formed by cycle is also a multiple of n since w is of period n.

In $\mathbb{HR}_w(2k-1,\ II)$, choose a vertex u_1 (say $u_1 = w[i, 2k-1+(i-1)]$) and it belongs to a component which is a cycle. Since every vertex is a word of length $2k-1$, the word formed by a cycle $u_1u_2\cdots u_ju_1$ is $u_1[1,2k-1]u_2[k,2k-1]\cdots u_{j-1}[k,2k-1]u_j[k,k] = w[i,2k-1+(i-1)]w[i+2k-1,3k-1+(i-1)]\cdots w[(j-1)k+i-1, jk-1+(i-1)]w_{i+jk-1} = w[i, i+jk-1]$. Here k^{th} letter of u_j is taken neither in $u_j[1,k-1]$ nor in $u_j[k+1,2k-1]$ and $u_j[k+1,2k-1] = u_1[1,k-1]$ because (v_j, v_1) is an arc in the cycle. And the length of the word formed by the cycle $u_1u_2\cdots u_ju_1$ is $2k-1+(j-2)k+1 = kj$. As the vertex u_j makes an arc with u_1, the length of the word formed by cycle is also a multiple of n since w is of period n. Similarly one can show for $\mathbb{HR}_w(2k+1,\ I)$.

Length of the word formed by a cycle is a multiple of k and also a multiple of n in all the cases. Thus least common multiple of k and n, $l.c.m(k,n)$ is the required length of the word formed by cycle which starts and ends with a vertex u_1 or v_1 or y_1. Since we are counting only suffix word of length k in each vertex, we need $j = \frac{l.c.m(k,n)}{k}$ number of vertices to form a cycle.

We have chosen a vertex arbitrarily and the component containing the vertex is $\frac{l.c.m(k,n)}{k}$-cycle. Thus each component in $\mathbb{HR}_w(2k-1,\ II)$, $\mathbb{HR}_w(2k)$, $\mathbb{HR}_w(2k+1,\ I)$ is a β- cycle, where $\beta = \frac{l.c.m\ (k,n)}{k}$.

Each component in $\mathbb{HR}_w(2k-1,\ II)$, $\mathbb{HR}_w(2k)$, $\mathbb{HR}_w(2k+1,\ I)$ have $\beta = \frac{l.c.m(k,n)}{k}$ vertices, so the number of components in $\mathbb{HR}_w(2k-1,\ II)$, $\mathbb{HR}_w(2k)$, $\mathbb{HR}_w(2k+1,\ I)$ is

$$\frac{Total\ no.\ of\ vertices}{No.\ of\ vertices\ in\ each\ component} = \frac{n}{\frac{l.c.m(k,n)}{k}} = \frac{n.k}{l.c.m(k,n)} = g.c.d(k,n).$$

Thus each of $\mathbb{HR}_w(2k-1,\ II)$, $\mathbb{HR}_w(2k)$, $\mathbb{HR}_w(2k+1,\ I)$ have α-components, where $\alpha = g.c.d(k,n)$. □

Corollary 3. *Let $w = x^\omega$. Then, for $2 < n < 2k$ and $(k,n) = 1$, $\mathbb{HR}_w(2k-1,\ II)$, $\mathbb{HR}_w(2k)$ and $\mathbb{HR}_w(2k+1,\ I)$ are strongly connected.*

Proof. If $(k,n) = 1$, then $\mathbb{HR}_w(2k)$ has 1-component and by Theorem 2, it is an $n-$ cycle. Hence it is strongly connected. □

Corollary 4. *Let $w = x^\omega$. If $2 < n < 2k$ and $2k \neq rn$ where $r \in 2\mathbb{N}$, then*

$$\mathbb{HR}_w(2k-1,\ II) \simeq \mathbb{HR}_w(2k) \simeq HR_w(2k+1,\ I).$$

Proof. $|V(\mathbb{HR}_w(2k-1,\ II))| = |V(\mathbb{HR}_w(2k))| = |V(\mathbb{HR}_w(2k+1,\ I))| = n$ by Proposition 1. Each component of $\mathbb{HR}_w(2k-1,\ II)$, $\mathbb{HR}_w(2k)$, $\mathbb{HR}_w(2k+1,\ I)$ is a β-cycle and there are α-components by Theorem 2. Hence they are isomorphic to each other. □

5 Conclusion

In this paper we define HRR and study several interesting properties of the structure of graphs that are formed. We consider one particular infinite sequence and study the properties. However, there are several special sequences that are needed to be analysed. Pattern avoidance in a set of words over an alphabet is an important combinatorial problem. A word w over an alphabet contains a pattern l if w contains sequence order-isomorphic to l. It will be interesting to see this concept and its properties reflected in HRR. Also one can determine all Wilf-equivalence classes of words patterns of length t, $t \leq 6$. Several problems with in the set of integer composition can be analogously studied.

References

[1] L. M. Adleman : Molecular computation of solutions to combinatorial problems, Science, Volume 266(5187) pp: 1021-1024, (1994)

[2] S. Bera and K. Mahalingam : Structural properties of word representable graphs, Mathematics in computer science, Volume 10(2), pp: 209–222, (2016).

[3] A. Bondy, M.R. Murthy : Graph Theory, Springer, (2008)

[4] A. Brandstädt, V. B. Le , J. P. Spinrad : Graph classes: A survey, SIAM ISBN-10-89871-432X, (1999).

[5] J. Cassaigne : Special factors of sequences with linear subword complexity, Developments in language theory II: At the cross roads of Mathematics, Computer science and Biology, pp: 25-34, (1995).

[6] J. Cassaigne : On a conjecture of J. Shalit, ICALP 1997, Lecture notes in computer science, Volume 1256, pp: 693-704, (1997).

[7] S. Ferenczi and Z. Kása : Complexity for finite factors of infinite sequences, Theoretical computer science, Volume 218, pp: 177-195, (1999).

[8] A.E. Frid : On factor graphs of DOL words, Discrete applied mathematics, Volume 114 (1-3), pp. 121-30, (2001).

[9] R. Graham, N. Zang : Enumerating split-pair arrangements, Journal of combinatorial theory, Series A, Volume 115(2), pp: 293-303, (2008).

[10] M. M. Halldórsson, S. Kitaev, A. Pyatkin : Graphs capturing alternations in words, Proceedings of 14th international conference, DLT 2010, London, ON, Canada, August 17-20, pp: 436-437 (2010).

[11] S. Kitaev, A. V. Pyatkin : On representable graphs, Journal of automata, languages and combinatorics, Volume 13(1), pp: 45-54, (2008).

[12] M. Lothaire : Algebraic combinatorics on words, Cambridge university press (2002)

[13] R. C. Lyndon, M. P. Schützenberger : The equation $a^m = b^n c^p$ in a free group, Michigan mathematical journal, Volume 9, pp: 289-298, (1962).

[14] G. Rauzy : Suites à termes dans un alphabet fini, Seminar on number theory (1982-1983), University of Bordeaux I, Talence, Volume 25, pp: 1-16, (1983).

[15] G.Rote : Sequences with subword complexity $2n$, Journal of number theory, Volume 46, pp: 196-213, (1993).

[16] P. V. Salimov : On Rauzy graph sequences of infinite words, Journal of applied and industrial mathematics, Volume 4(1), pp: 127-135, (2008).

[17] H. J. Shyr, G. Thierrin : Disjunctive languages and codes, Proceedings of FCT, LNCS, Volume 56, pp: 171-176, (1977).