Journal of Mathematical Research with Applications
Nov., 2019, Vol. 39, No. 6, pp. 718–732
DOI:10.3770/j.issn:2095-2651.2019.06.015
Http://jmre.dlut.edu.cn

The Factor Spectrum and Derived Sequence

Yuke HUANG[1,*], **Zhiying WEN**[2]

1. *School of Science, Beijing University of Posts and Telecommunications, Beijing* 100876*, P. R. China;*
2. *Department of Mathematical Sciences, Tsinghua University, Beijing* 100084*, P. R. China*

Dedicated to the Memory of Professor L. C. HSU on the Occasion of His 100th Birthday

Abstract Given a sequence ρ over a finite alphabet $\mathcal{A}$, an important topic in combinatorics on words is to find out all factors ω of ρ and positive integers p such that ω_p (the p-th occurrence of ω) fulfills property $\mathcal{P}$. This problem is equivalent to determining a notion called the factor spectrum. Determining the factor spectrum is a difficult problem. To this aim, we introduce several notions, such as: kernel word, envelope word, return word and derived sequence of each factor ω. Using the factor spectrum and derived sequence, we can solve some enumerations of factors, such as the numbers of palindromes, fractional powers, etc. We will show some results for several sequences, such as the Fibonacci sequence, the Tribonacci sequence, the Period-doubling sequence, etc. And we think that these notions and methods are suitable for all recurrent sequences.

Keywords kernel word; envelope word; return word; derived sequence; the factor spectrum

MR(2010) Subject Classification 68R15

1. Introduction

A sequence ρ is said to be recurrent if every factor occurs infinitely often [1]. We arrange all occurrences of factor ω in ρ as the sequence $\{\omega_p\}_{p\geq 1}$, where ω_p denotes the p-th occurrence of ω. Let $\mathcal{P}$ be a property about factor, and if the p-th occurrence of ω has property $\mathcal{P}$, we denote $\omega_p\in\mathcal{P}$. We denote $\omega\in\mathcal{P}$, if there exists an integer $p\geq 1$ such that $\omega_p\in\mathcal{P}$. When $\omega\in\mathcal{P}$, maybe not all $p\in\mathbb{N}$ such that $\omega_p\in\mathcal{P}$. An important topic in combinatorics on words is to find out all ω and p such that ω_p fulfills property $\mathcal{P}$. More precisely, let ρ be a sequence and $\mathcal{P}$ be a property, we introduce the following notion factor spectrum of a sequence ρ and a property $\mathcal{P}$ that

$$Spt(\rho,\mathcal{P}):=\{(\omega,p)\mid \omega\prec\rho,\ p\geq 1,\ \omega_p\in\mathcal{P}\}\subset(\Omega_\rho,\mathbb{N}), \tag{1.1}$$

where Ω_ρ is the set of all factors in ρ. When we write $(\omega,p)\in Spt(\rho,\mathcal{P})$, we regard ω and p as two variable function over $(\Omega_\rho,\mathbb{N})$. By the definition in Eq. (1.1), the above problem is equivalent to determining the factor spectra.

Received August 15, 2019; Accepted October 10, 2019
Supported by the National Natural Science Foundation of China (Grant Nos. 11701024; 11431007).
* Corresponding author
E-mail address: hyk@bupt.edu.cn, hyg03ster@163.com (Yuke HUANG); wenzy@tsinghua.edu.cn (Zhiying WEN)

Determining the factor spectra is a difficult problem. For this aim, we use the return word and derived sequence of each factor ω, which are introduced in Durand [2], see Section 2 for details. From them, we may know the local structure near ω_p for each p. In Huang-Wen [3, 4], we first study the structure of derived sequences of the Fibonacci sequence and the Tribonacci sequence, then we determine the factor spectra for some combinatorial properties.

2. The return word and derived sequence

Let $\omega = x_1x_2\cdots x_N$ be a finite word. We denote by $|\omega|$ the number of letters in ω, called the length of ω. For $1 \le i \le j \le N$, we define $\omega[i,j] = x_ix_{i+1}\cdots x_{j-1}x_j$, by convention, $\omega[i] = \omega[i,i] = x_i$ and $\omega[i,i-1] = \varepsilon$ (empty word). We call $\omega[i,j]$ a factor of ω, denoted by $\omega[i,j] \prec \omega$. The position of factor $\omega[i,j]$ in ω is defined by i. We call a non-empty word ν a prefix (resp., suffix) of a word ω if there exists a word u such that $\omega = \nu u$ (resp., $\omega = u\nu$), denoted by $\nu \triangleleft \omega$ (resp., $\nu \triangleright \omega$). In this case, we write $\nu^{-1}\omega = u$ (resp., $\omega\nu^{-1} = u$), where ν^{-1} is the inverse word of ν such that $\nu\nu^{-1} = \nu^{-1}\nu = \varepsilon$. A palindrome ω is a finite word that reads the same backwards as forwards, i.e., $\overleftarrow{\omega} := x_N\cdots x_2x_1 = \omega$.

The definitions of both the return word and derived sequence are from Durand [2]. Let us recall them as below. Notice that Durand gave these notions for all prefixes of any recurrent sequence, and we extend these notions to all non-empty factors. Recall that ω_p denotes the p-th occurrence of ω. Denote by $\mathrm{occ}(\omega,p)$ the position of ω_p. For $p,q \ge 1$, $\omega_p \prec W_q$ means $\omega \prec W$ and $\mathrm{occ}(W,q) \le \mathrm{occ}(\omega,p) < \mathrm{occ}(\omega,p) + |\omega| - 1 \le \mathrm{occ}(W,q) + |W| - 1$.

We call $\rho[i,j-1]$ the p-th return word over ω where i and j are the positions of the p-th and $(p+1)$-th occurrences of ω in ρ, denoted by $R_{\rho,p}(\omega)$. Denote by $\mathcal{H}_{\rho,\omega}$ the set of return words over factor $\omega \prec \rho$. Then the sequence ρ can be written in a unique way as a concatenation $\rho = \rho[1,h-1]R_{\rho,1}(\omega)R_{\rho,2}(\omega)\cdots$ where $R_{\rho,p}(\omega) \in \mathcal{H}_{\rho,\omega}$ and $\rho[1,h-1]$ is the prefix of ρ occurring before the first occurrence of ω. Let us give to $\mathcal{H}_{\rho,\omega}$ the linear order defined by the rank of the first occurrence in ρ. This defines a one to one and onto map $\Lambda_{\rho,\omega} : \mathcal{H}_{\rho,\omega} \to \{1,\ldots,\mathrm{Card}(\mathcal{H}_{\rho,\omega})\} = \mathcal{N}_{\rho,\omega} \subset \{\alpha,\beta,\ldots\}$, and the sequence $\mathcal{D}_\omega(\rho) := \Lambda_{\rho,\omega}(R_{\rho,1}(\omega))\Lambda_{\rho,\omega}(R_{\rho,2}(\omega))\Lambda_{\rho,\omega}(R_{\rho,3}(\omega))\cdots$. This sequence of alphabet $\mathcal{N}_{\rho,\omega}$ is called a derived sequence of ρ. Notice that we omit the prefix $\rho[1,h-1]$. Moreover, we denote the reciprocal map of $\Lambda_{\rho,\omega}$ by $\Theta_{\rho,\omega} : \mathcal{N}_{\rho,\omega} \to \mathcal{H}_{\rho,\omega}$.

The main result of Durand [2] is: a sequence ρ is substitutive primitive if and only if the number of its different derived sequences is finite. The other property is: for any $\omega \prec \rho$ and $v \prec \mathcal{D}_\omega(\rho)$, there exists a factor $\mu \prec \rho$ such that derived sequence $\mathcal{D}_v(\mathcal{D}_\omega(\rho)) = \mathcal{D}_\mu(\rho)$ (see [2, Proposition 6 (5)]). By the two properties, if ρ is substitutive primitive, then for any $\omega \prec \rho$, derived sequence $\mathcal{D}_\omega(\rho)$ is still substitutive primitive.

2.1. Derived sequences of the Fibonacci and Tribonacci sequences

The Fibonacci morphism σ_1 over alphabet $\{a,b\}$ is a substitution defined by $\sigma_1(a) = ab$ and $\sigma_1(b) = a$. The Fibonacci sequence $\mathbb{F}$ is defined to be the fixed point beginning with the letter a of the Fibonacci morphism. It is a recurrent infinite word [1]. Define $F_m = \sigma_1^m(a)$ for $m \ge 0$, by convention, $F_{-1}=b$ and $F_{-2}=\varepsilon$. The m-th Fibonacci number f_m is equal to the length of F_m. As a classical example of sequences over the binary alphabet, the Fibonacci sequence $\mathbb{F}$ has many remarkable properties. We refer to Lothaire [5,6], Allouche-Shallit [1] and Berstel [7,8]. The Tribonacci sequence $\mathbb{T}$ is a natural generalization of the Fibonacci sequence, which is the fixed point beginning with the letter a of morphism $\sigma_2(a)=ab$, $\sigma_2(b)=ac$, $\sigma_2(c)=a$ defined over the alphabet $\{a,b,c\}$. Define $T_m=\sigma_2^m(a)$ for $m\geq 0$.

The m-th Tribonacci number t_m is equal to the length of T_m.

2.1.1. Kernel word

The main tool of Huang-Wen [3,4] is kernel word, which is a set of factors in the sequence ρ. For a factor ω, we first introduce the kernel of a factor ω as follows, denoted by $\mathrm{Ker}(\omega)$.

$$\begin{aligned}\mathrm{Ker}(\omega):=\min_{|\cdot|}\{W\mid\ &\text{the difference } \mathrm{occ}(W,p)-\mathrm{occ}(\omega,p)\\ &\text{is independent of } p \text{ for } p\geq 1,\ W\prec\omega\}.\end{aligned}\tag{2.1}$$

We can prove that the kernel of any factor ω is unique when ρ is the Fibonacci or Tribonacci sequence. If $\mathrm{Ker}(\omega)$ is unique for all ω in ρ, we define the set of kernel words by that $\mathcal{K}:=\{\mathrm{Ker}(\omega)\mid\omega\prec\rho\}$. More precisely, we give the definitions of kernel words in the Fibonacci and Tribonacci sequences as below. In fact, the kernel word in the Fibonacci sequence is just the singular words defined in Wen-Wen [9]. But unfortunately, the singular words of the Tribonacci sequence defined in Tan-Wen [10] is not a set satisfying Eq. (2.1), see Example 2.2(b).

Remark 2.1 The two notions are suitable not only for the Fibonacci and Tribonacci sequences, but also for some other sequences. Notice that the notions are not suitable for the Period-doubling sequence $\mathbb{D}$, for details see Example 2.2(a). For this kind of sequences, we need other tools.

Example 2.2 (a) Take $\omega=baaabababaaab\prec\mathbb{D}$, $U=aabab$ and $V=babaa$. We can check that $|U|=|V|$, $\mathrm{occ}(U,p)-\mathrm{occ}(\omega,p)\equiv 2$ and $\mathrm{occ}(V,p)-\mathrm{occ}(\omega,p)\equiv 6$ for all $p\geq 1$. But for any factor W in $\mathbb{D}$ of length $1\leq n\leq 4$, it dose not satisfy that $\mathrm{occ}(W,p)-\mathrm{occ}(\omega,p)$ is independent of p. This means both U and V satisfy Eq. (2.1).

(b) Tan-Wen [10] defined two kinds of singular words: Ω_m^1 and Ω_m^2 for $m\geq 1$.

$$\begin{cases}\Omega_m^1=\{\text{factor of length } f_m \text{ of the word } \alpha^{-1}\overleftarrow{E_m}D_{m-1}E_m\alpha^{-1}\},\\ \Omega_m^2=\{\text{factor of length } f_m \text{ of the word } \beta^{-1}E_{m+1}D_{m-2}\overleftarrow{E_{m+1}}\beta^{-1}\},\end{cases}$$

where $D_m=T_{m-1}T_{m-2}\cdots T_2T_1T_0$, $E_m=D_{m-1}^{-1}T_m$, α is the last letter of E_m and β is the first letter of E_{m+1}. By convention, $D_0=\varepsilon$. For instance, $\Omega_1^1=\{aa\}$, $\Omega_1^2=\{ac,ca\}$, $\Omega_2^1=\{abab,baba\}$ and $\Omega_2^2=\{abaa,baab,aaba\}$. We can check that the set of all singular words in $\mathbb{T}$ is not satisfying Eq. (2.1).

Definition 2.3 (Kernel words of the Fibonacci sequence) *(1) Let $\{K_m\}_{m\geq -1}$ be the sequence* *of factors with*

$$K_{-1}=a \text{ and } K_m=\delta_{m+1}F_m[1,f_m-1] \text{ for } m\geq 0,$$

where $\delta_m \triangleright F_m$. *We call* K_m *the* m*-th kernel word. More precisely,* $\delta_m=a$ *if* $m\equiv 0 \pmod 2$, $\delta_m=b$ *otherwise. Obviously,* $|K_m|=f_m$ *for* $m\geq -1$.

(2) Define $\mathcal{K}:=\{K_m \mid m\geq -1\}$, *which is called the kernel set of the Fibonacci sequence.*

(3) We define the order of kernel words that $K_m \sqsubset K_{m+1}$ *for all* $m\geq -1$. *For any factor* $\omega \prec \mathbb{F}$, *we define*

$$\mathrm{Ker}(\omega)=\max_{\sqsubset}\{K_m|K_m\prec\omega,\ m\geq -1\},$$

which is called the kernel of factor ω.

(4) We define the order $K_{m,p}\sqsubset K_{m+1,q}$ *for all* $p,q\geq 1$. *For any factor* ω_p *where* $(\omega,p)\in(\Omega_{\mathbb{F}},\mathbb{N})$, *we define*

$$\mathrm{Ker}(\omega_p)=\max_{\sqsubset}\{K_{m,q}|K_{m,q}\prec\omega_p,\ m\geq -1,\ q\geq 1\},$$

which is called the kernel of factor ω_p.

Here are the first few values of F_m, f_m and K_m:

- $\{F_m\}_{m\geq -1}=b,a,ab,aba,abaab,abaababa,abaababaabaab,abaababaabaababaababa,\ldots$.
- $\{f_m\}_{m\geq -1}=1,1,2,3,5,8,13,21,34,55,89,144,233,377,610,987,1597,2584,4181,\ldots$.
- $\{K_m\}_{m\geq -1}=a,b,aa,bab,aabaa,babaabab,aabaababaabaa,babaababaabaababaabab,\ldots$.

We have proved that all kernel words of $\mathbb{F}$ are palindrome [3].

Definition 2.4 (Kernel words of the Tribonacci sequence) *(1) Let* $\{k_m\}_{m\geq 1}$ *be the sequence of positive integers with*

$$k_1=k_2=k_3=1 \text{ and } k_m=k_{m-1}+k_{m-2}+k_{m-3}-1 \text{ for } m\geq 4.$$

The number k_m *is called the* m*-th kernel number.*

(2) Let $\{K_m\}_{m\geq 1}$ *be the sequence of factors with*

$$K_1=a,\ K_2=b,\ K_3=c \text{ and } K_m=\delta_m T_{m-3}[1,k_m-1] \text{ for } m\geq 4,$$

where $\delta_m \triangleright T_m$. *We call* K_m *the* m*-th kernel word. More precisely,* $\delta_m=a$ *if* $m\equiv 0 \pmod 3$, $\delta_m=b$ *if* $m\equiv 1 \pmod 3$, $\delta_m=c$ *otherwise. Obviously,* $|K_m|=k_m$ *for* $m\geq 1$.

(3) Define $\mathcal{K}:=\{K_m \mid m\geq 1\}$, *which is called the kernel set of the Tribonacci sequence.*

(4) We define the order of kernel words that $K_m\sqsubset K_{m+1}$ *for all* $m\geq 1$. *For any factor* $\omega\prec\mathbb{T}$, *we define* $\mathrm{Ker}(\omega)=\max_{\sqsubset}\{K_m \mid K_m\prec\omega,\ m\geq 1\}$, *which is called the kernel of factor* ω.

(5) We define the order $K_{m,p}\sqsubset K_{m+1,q}$ *for* $m,p,q\geq 1$. *For any factor* ω_p *where* $(\omega,p)\in(\Omega_{\mathbb{T}},\mathbb{N})$, *we define* $\mathrm{Ker}(\omega_p)=\max_{\sqsubset}\{K_{m,q} \mid K_{m,q}\prec\omega_p,\ m,q\geq 1\}$, *which is called the kernel of factor* ω_p.

Here are the first few values of T_m, t_m, K_m and k_m:

- $\{T_m\}_{m\geq 0}=a,ab,abac,abacaba,abacabaabacab,abacabaabacababacabaabac,\ldots$.
- $\{t_m\}_{m\geq 0}=1,2,4,7,13,24,44,81,149,274,504,927,1705,3136,\ldots$.
- $\{K_m\}_{m\geq 1}=a,b,c,aa,bab,cabac,aabacabaa,babacabaabacabab,\ldots$.

- $\{k_m\}_{m\geq 1} = 1, 1, 1, 2, 3, 5, 9, 16, 29, 53, 97, 178, 327, 601, \ldots$.

As the case of the Fibonacci sequence, all kernel words of $\mathbb{T}$ are also palindrome [4].

Remark 2.5 Notice that, the definitions of $\mathrm{Ker}(\omega)$ in Eq. (2.1) and Definitions 2.3 and 2.4 look different. In fact, they are equivalent. More precisely, Eq. (2.1) is more essential, and Definitions 2.3 and 2.4 are more applicable and more convenient.

2.1.2. Derived sequences

Huang-Wen proved the properties about "derived sequence" below, see [3, Theorem 2.11] and [4, Theorem 5.1], respectively.

Theorem 2.6 *(1) For any factor $\omega \prec \mathbb{F}$, derived sequence $\mathcal{D}_\omega(\mathbb{F})$ is still $\mathbb{F}$ itself.*

(2) For any factor $\omega \prec \mathbb{T}$, derived sequence $\mathcal{D}_\omega(\mathbb{T})$ is still $\mathbb{T}$ itself.

Using Theorem 2.6, we determined the numbers of palindromes, squares and cubes occurring in $\mathbb{F}[i, j]$ and $\mathbb{T}[i, j]$ for $1 \leq i \leq j$, see Huang-Wen [11, 12] for instance. These topics are of great importance in computer science.

Example 2.7 (Derived sequence $\mathcal{D}_\omega(\mathbb{F})$ for $\omega = baab$)

$$\mathbb{F} = a\underbrace{baaba}_{\alpha}\underbrace{baa}_{\beta}\underbrace{baaba}_{\alpha}\underbrace{baaba}_{\alpha}\underbrace{baa}_{\beta}\underbrace{baaba}_{\alpha}\underbrace{baa}_{\beta}\underbrace{baaba}_{\alpha}\underbrace{baaba}_{\alpha}\underbrace{baa}_{\beta}\underbrace{baaba}_{\alpha}\underbrace{baaba}_{\alpha}\underbrace{baa}_{\beta}\cdots,$$

where

$$\alpha = \Lambda_{\mathbb{F},\omega}(R_{\mathbb{F},1}(\omega)) = \Lambda_{\mathbb{F},\omega}(baaba) \text{ and } \beta = \Lambda_{\mathbb{F},\omega}(R_{\mathbb{F},2}(\omega)) = \Lambda_{\mathbb{F},\omega}(baa).$$

It is easy to see that the derived sequence $\mathcal{D}_\omega(\mathbb{F}) = \alpha\beta\alpha\alpha\beta\alpha\beta\alpha\alpha\beta\alpha\alpha\beta\cdots$ is a Fibonacci sequence over $\{\alpha, \beta\}$.

2.1.3. Three steps to prove Theorem 2.6

How to prove Theorem 2.6 by the kernel word? We need three steps:

Step 1. Determine derived sequence $\mathcal{D}_\omega(\rho)$ for any kernel words $\omega \prec \mathcal{K}$.

Step 2. For all $\omega \prec \rho$, prove that the difference $\mathrm{occ}(\mathrm{Ker}(\omega), p) - \mathrm{occ}(\omega, p)$ is independent of p.

Step 3. Prove derived sequence $\mathcal{D}_\omega(\rho)$ is exactly $\mathcal{D}_{\mathrm{Ker}(\omega)}(\rho)$.

For details, see Huang-Wen [3, 4] for instance.

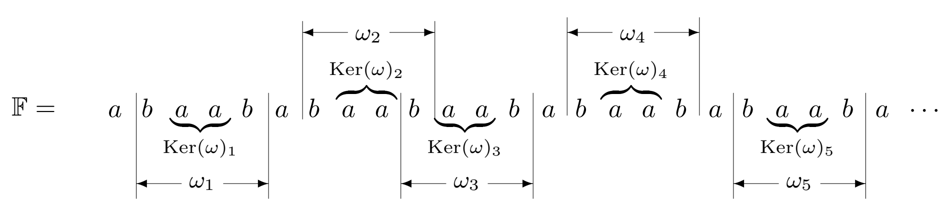

Figure 1 The relation between ω_p and $\mathrm{Ker}(\omega)_p$ for $\omega = baab$ and $1 \leq p \leq 5$.

2.2. Derived sequences of the Period-doubling sequence

Huang-Wen [3,4] determine the structure of derived sequences of the Fibonacci sequence and the Tribonacci sequence, respectively. But for the Period-doubling sequence $\mathbb{D}$, there are two differences. Firstly, we find that for different factors, they may have different derived sequences. But fortunately, we can divide set $\Omega_{\mathbb{D}}$ into two types, and each type corresponds to one derived sequence. Secondly, the main tool of [3, 4] is kernel word. But the kernel word technique is not valid for the Period-doubling sequence, thus we introduce the envelope words and corresponding techniques. In this way, we determine the factor spectra for some combinatorics properties.

Let $\mathcal{A} = \{a, b\}$ be a binary alphabet. The Period-doubling sequence $\mathbb{D}$ is the fixed point beginning with the letter a of substitution $\sigma_3(a) = ab$ and $\sigma_3(b) = aa$. It is also the first difference of the Thue-Morse sequence, where we use an equivalent substitution $\hat{\sigma}(1) = 10$, $\hat{\sigma}(0) = 11$, and the definition of the difference of an integer sequence is natural. It has been heavily studied in mathematics and computer science. Damanik [13] determined the numbers of palindromes, squares and cubes of length n occurring in $\mathbb{D}$. Allouche-Peyrière-Wen-Wen [14] proved that all the Hankel determinants of $\mathbb{D}$ are odd integers [15–20, 22–28]. We denote $A_m = \sigma_3^m(a)$ and $B_m = \sigma_3^m(b)$ for $m \geq 0$. Then $|A_m| = |B_m| = 2^m$. Let $\delta_m \in \{a, b\}$ be the last letter of A_m. Obviously, $\delta_m = a$ if and only if m is even; and δ_{m+1} is the last letter of B_m.

Recall that for any factor ω of $\mathbb{F}$, derived sequence $\mathcal{D}_\omega(\mathbb{F})$ is still $\mathbb{F}$ itself. But Huang-Wen [29] proved that for different factors in $\mathbb{D}$, there will be different derived sequences. In [29], we determined two types of derived sequences, see Theorem 2.9.

2.2.1. Envelope word

The main tool of Huang-Wen [29] is envelope word, which is a set of factors in sequence. We first introduce the envelope of a factor ω for the factor in sequence ρ, denoted by $\mathrm{Env}(\omega)$.

$$\begin{aligned}\mathrm{Env}(\omega) := \max_{|\cdot|}\{W \mid\ & \text{the difference } \mathrm{occ}(\omega, p) - \mathrm{occ}(W, p) \\ & \text{is independent of } p \text{ for } p \geq 1,\ \omega \prec W\}.\end{aligned} \tag{2.2}$$

We can prove that the envelope of any factor ω is unique. In fact, if U and V are two words satisfying Eq. (2.2) and $\mathrm{occ}(U, 1) < \mathrm{occ}(V, 1)$, then $W = \rho[\mathrm{occ}(U, 1), \mathrm{occ}(V, 1) + |V| - 1]$ also satisfies Eq. (2.2). Since $|W| > |U| = |V|$, a contradiction. We define the set of envelope words by that $\mathcal{E} := \{\mathrm{Env}(\omega) \mid \omega \prec \rho\}$.

More precisely, we give the definitions of envelope words in the Period-doubling sequence as below. Notice that, the definitions of $\mathrm{Env}(\omega)$ in Eq. (2.2) and Definition 2.8 look different. In fact, they are equivalent. More precisely, Eq. (2.2) is more essential, and Definition 2.8 is more applicable and more convenient.

We think notions of envelop words will be suitable for all recurrent sequence.

Definition 2.8 (Envelope word of the Period-doubling sequence) *(1) Let* $\{E_m^i \mid i = 1, 2,\ m \geq 1\}$

be a set of factors with

$$E_m^1 = A_m\delta_m^{-1} \text{ and } E_m^2 = B_m B_{m-1}\delta_m^{-1}.$$

We call E_m^i the m-th envelope word of type i. Moreover $|E_m^1| = 2^m - 1$ and $|E_m^2| = 3 \times 2^{m-1} - 1$.

(2) Define $\mathcal{E} := \{E_m^i \mid i = 1, 2,\ m \geq 1\}$, which is called the envelope set of the Period-doubling sequence.

(3) We define the order of envelope words that $E_m^1 \sqsubset E_m^2$ and $E_m^i \sqsubset E_{m+1}^j$ for $i, j \in \{1, 2\}$, $m \geq 1$. For any factor $\omega \prec \mathbb{D}$, we define $\mathrm{Env}(\omega) = \min_{\sqsubset}\{E_m^i \mid \omega \prec E_m^i,\ i = 1, 2,\ m \geq 1\}$, which is called the envelope of factor ω.

(4) We define the order $E_{m,p}^i \sqsubset E_{n,q}^j$ if $E_m^i \sqsubset E_n^j$ for $i, j \in \{1, 2\}$, $m, n, p, q \geq 1$. For any factor ω_p where $(\omega, p) \in (\Omega_{\mathbb{D}}, \mathbb{N})$, we define $\mathrm{Env}(\omega_p) = \min\limits_{\sqsubset}\{E_{m,q}^i \mid \omega_p \prec E_{m,q}^i,\ i = 1, 2,\ m, q \geq 1\}$, which is called the envelope of factor ω_p.

Obviously, the lengths of all envelope words are odd, except $E_1^2 = aa$. Moreover, by the definition of E_m^i and $A_m\delta_m^{-1} = B_m\delta_{m+1}^{-1}$, we have $E_{m+1}^1 = E_m^1\delta_m E_m^1$ and $E_{m+1}^2 = E_m^1\delta_m E_m^1\delta_m E_m^1$ for $m \geq 1$. By induction, all envelope words are palindromes [29].

Here are the first few values of A_m, B_m, E_m^1 and E_m^2:

- $\{A_m\}_{m\geq 0} = a, ab, abaa, abaaabab, abaaabababaaabaa, \ldots.$
- $\{B_m\}_{m\geq 0} = b, aa, abab, abaaabaa, abaaabababaaabab, \ldots.$
- $\{E_m^1\}_{m\geq 1} = a, aba, abaaaba, abaaabababaaaba, \ldots.$
- $\{E_m^2\}_{m\geq 1} = aa, ababa, abaaabaaaba, abaaabababaaabababaaaba, \ldots.$

By this definition, there exists a unique envelope for each factor. For instance, $\mathrm{Env}(abab) = ababa = E_2^2$ and $\mathrm{Env}(abaaabab) = E_4^1$. We can check them by the expressions of E_m^i. More precisely, there exists an integer j and two words u and ν, such that

$$\omega = \mathrm{Env}(\omega)[j+1, j+|\omega|] \text{ and } \mathrm{Env}(\omega) = u \cdot \omega \cdot \nu, \tag{2.3}$$

where $0 \leq j \leq |\mathrm{Env}(\omega)| - |\omega|$, $|u| = j$, $u \triangleleft \mathrm{Env}(\omega)$ and $\nu \triangleright \mathrm{Env}(\omega)$. We proved that the integer j is unique for any fixed ω in [29].

2.2.2. Derived sequences

In Huang-Wen [29, Theorem 2.6], we determined two types of derived sequences.

Theorem 2.9 *Let factor $\omega \prec \mathbb{D}$ have expression in Eq. (2.3).*

(1) If there exists an integer $m \geq 1$ such that $\mathrm{Env}(\omega) = E_m^1$, derived sequence $\mathcal{D}_\omega(\mathbb{D}) = \mathbb{D}(\alpha, \beta\beta)$. More precisely,

$$\begin{cases} \alpha = \Lambda_{\mathbb{D},\omega}(R_{\mathbb{D},1}(\omega)), & R_{\mathbb{D},1}(\omega) = A_m, \quad |R_{\mathbb{D},1}(\omega)| = 2^m; \\ \beta = \Lambda_{\mathbb{D},\omega}(R_{\mathbb{D},2}(\omega)), & R_{\mathbb{D},2}(\omega) = A_{m-1}, \quad |R_{\mathbb{D},2}(\omega)| = 2^{m-1}. \end{cases} \tag{2.4}$$

Moreover $R_{\mathbb{D},0}(\omega) = \mathbb{D}[1, (\omega)_1 - 1] = u$ and $|R_{\mathbb{D},0}(\omega)| = j$.

(2) If there exists an integer $m \geq 1$ such that $\mathrm{Env}(\omega) = E_m^2$, derived sequence $\mathcal{D}_\omega(\mathbb{D}) =$

$\mathbb{D}(\alpha\beta, \alpha\gamma\alpha\gamma)$. *More precisely,*

$$\begin{cases} \alpha = \Lambda_{\mathbb{D},\omega}(R_{\mathbb{D},1}(\omega)), & R_{\mathbb{D},1}(\omega) = A_{m-1}, & |R_{\mathbb{D},1}(\omega)| = 2^{m-1}; \\ \beta = \Lambda_{\mathbb{D},\omega}(R_{\mathbb{D},2}(\omega)), & R_{\mathbb{D},2}(\omega) = A_{m-1}A_mB_{m+1}, & |R_{\mathbb{D},2}(\omega)| = 7 \times 2^{m-1}; \\ \gamma = \Lambda_{\mathbb{D},\omega}(R_{\mathbb{D},4}(\omega)), & R_{\mathbb{D},4}(\omega) = B_mB_{m-1}, & |R_{\mathbb{D},4}(\omega)| = 3 \times 2^{m-1}. \end{cases} \tag{2.5}$$

Moreover $R_{\mathbb{D},0}(\omega) = \mathbb{D}[1, (\omega)_1 - 1] = A_m u$ *and* $|R_{\mathbb{D},0}(\omega)| = 2^m + j$.

$\mathbb{D}$	=	a	b	a	a	a	b	a	b	a	b	a	a	a	b	a	$\cdots$
$\mathbb{D}(\alpha, \beta\beta)$	=	α	$\beta\beta$	α	α	α	$\beta\beta$	α	$\beta\beta$	α	$\beta\beta$	α	α	α	$\beta\beta$	α	$\cdots$
$\mathbb{D}(\alpha\beta, \alpha\gamma\alpha\gamma)$	=	$\alpha\beta$	$\alpha\gamma\alpha\gamma$	$\alpha\beta$	$\alpha\beta$	$\alpha\beta$	$\alpha\gamma\alpha\gamma$	$\alpha\beta$	$\alpha\gamma\alpha\gamma$	$\alpha\beta$	$\alpha\gamma\alpha\gamma$	$\alpha\beta$	$\alpha\beta$	$\alpha\beta$	$\alpha\gamma\alpha\gamma$	$\alpha\beta$	$\cdots$

Figure 2 The first few letters of $\mathbb{D}$, $\mathbb{D}(\alpha, \beta\beta)$ and $\mathbb{D}(\alpha\beta, \alpha\gamma\alpha\gamma)$

Example 2.10 (Derived sequence $\mathcal{D}_\omega(\mathbb{D})$) Let $\omega = aba$, $\mathrm{Env}(aba) = aba = E_2^1$ and $\mathcal{D}_{aba}(\mathbb{D}) = \mathbb{D}(\alpha, \beta\beta)$ where $\Theta_{\mathbb{D},aba}(\alpha) = R_{\mathbb{D},1}(aba) = abaa$ and $\Theta_{\mathbb{D},aba}(\beta) = R_{\mathbb{D},2}(aba) = ab$.

$$\mathbb{D} = \underbrace{abaa}_{\alpha}\underbrace{ab}_{\beta}\underbrace{ab}_{\beta}\underbrace{abaa}_{\alpha}\underbrace{abaa}_{\alpha}\underbrace{abaa}_{\alpha}\underbrace{ab}_{\beta}\underbrace{ab}_{\beta}\underbrace{abaa}_{\alpha}\underbrace{ab}_{\beta}\underbrace{ab}_{\beta}\cdots \tag{2.6}$$

Let $\omega = aa$, $\mathrm{Env}(aa) = aa = E_1^2$. In this case $\mathbb{D}(\alpha\beta, \alpha\gamma\alpha\gamma)$ where $\Theta_{\mathbb{D},aa}(\alpha) = R_{\mathbb{D},1}(aa) = a$, $\Theta_{\mathbb{D},aa}(\alpha) = R_{\mathbb{D},2}(aa) = aababab$ and $\Theta_{\mathbb{D},aa}(\alpha) = R_{\mathbb{D},4}(aa) = aab$.

$$\begin{aligned} \mathbb{D} = ab &\underbrace{a}_{\alpha}\underbrace{aababab}_{\beta}\underbrace{a}_{\alpha}\underbrace{aab}_{\gamma}\underbrace{a}_{\alpha}\underbrace{aab}_{\gamma}\underbrace{a}_{\alpha}\underbrace{aababab}_{\beta}\underbrace{a}_{\alpha}\underbrace{aababab}_{\beta} \\ &\underbrace{a}_{\alpha}\underbrace{aababab}_{\beta}\underbrace{a}_{\alpha}\underbrace{aab}_{\gamma}\underbrace{a}_{\alpha}\underbrace{aab}_{\gamma}\underbrace{a}_{\alpha}\underbrace{aababab}_{\beta}\underbrace{a}_{\alpha}\underbrace{aab}_{\gamma}\underbrace{a}_{\alpha}\underbrace{aab}_{\gamma}\cdots \end{aligned} \tag{2.7}$$

2.2.3. Three steps to prove Theorem 2.9

How to prove Theorem 2.9 by the envelope word? We need three steps:

Step 1. Determine derived sequence $\mathcal{D}_\omega(\rho)$ for any envelope word $\omega \prec \mathcal{E}$.

Step 2. For $\omega \prec \rho$, prove that the difference $\mathrm{occ}(\omega, p) - \mathrm{occ}(\mathrm{Env}(\omega), p)$ is independent of p.

Step 3. Prove derived sequence $\mathcal{D}_\omega(\rho)$ is exactly $\mathcal{D}_{\mathrm{Env}(\omega)}(\rho)$.

For details [29].

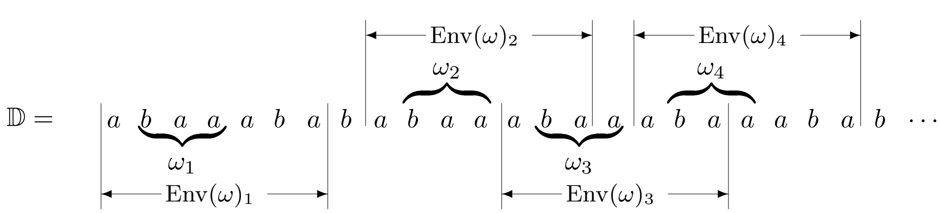

Figure 3 The relation between ω_p and $\mathrm{Env}(\omega)_p$ for $\omega = baa$ and $p = 1, 2, 3, 4$

2.2.4. The reflexivity of derived sequences

A known result in that for any $\omega \prec \rho$ and $v \prec \mathcal{D}_\omega(\rho)$, there exists a factor $\mu \prec \rho$ such that derived sequence $\mathcal{D}_v(\mathcal{D}_\omega(\rho))=\mathcal{D}_\mu(\rho)$ (see [2, Proposition 6(5)]). Thus for any $\omega\prec\mathbb{D}$ and $v\prec\mathcal{D}_\omega(\mathbb{D})$, derived sequence $\mathcal{D}_v(\mathcal{D}_\omega(\mathbb{D}))\in\{\mathcal{D}_\omega(\mathbb{D})\mid\omega\prec\mathbb{D}\}=\{\mathbb{D}(\alpha,\beta\beta),\mathbb{D}(\alpha\beta,\alpha\gamma\alpha\gamma)\}$.

Let $\rho\in\{\mathbb{D}(\alpha,\beta\beta),\mathbb{D}(\alpha\beta,\alpha\gamma\alpha\gamma)\}$. In Huang-Wen [29], we first define the envelope words in ρ. Then we divide Ω_ρ into two types: Ω_ρ^1 and Ω_ρ^2, according to their envelopes. At last, we prove that for all $v\in\Omega_\rho^1$, $\mathcal{D}_v(\rho)=\mathbb{D}(\alpha,\beta\beta)$; otherwise $\mathcal{D}_v(\rho)=\mathbb{D}(\alpha\beta,\alpha\gamma\alpha\gamma)$. We call it the reflexivity property of derived sequence, see Figure 4.

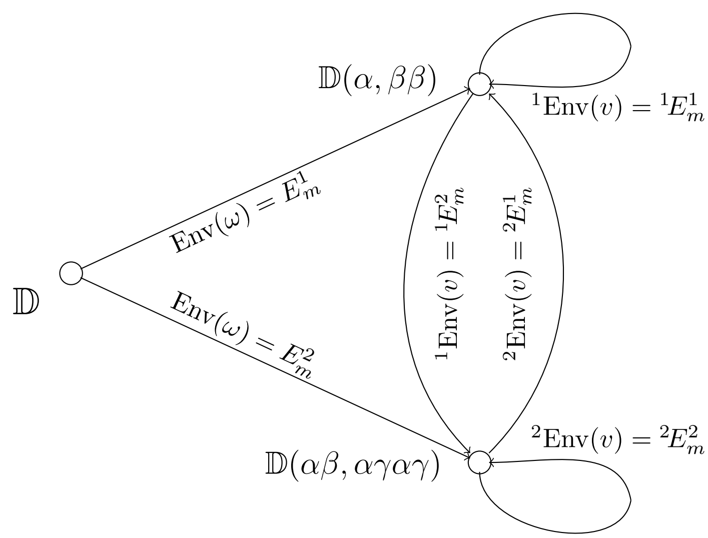

Figure 4 The reflexivity of derived sequences. For instance, the edge "$\mathbb{D}\xrightarrow{\mathrm{Env}(\omega)=E_m^1}\mathbb{D}(\alpha,\beta\beta)$" means that for any $\omega\prec\mathbb{D}$ if $\mathrm{Env}(\omega)=E_m^1$ then derived sequence $\mathcal{D}_\omega(\mathbb{D})=\mathbb{D}(\alpha,\beta\beta)$, which is given in Theorem 2.9 (1)

3. The factor spectrum

Recall the definition of factor spectrum given in the first paragraph in Section 1. Our aim in this section is to determine the factor spectra for some combinatorics properties, such as separated ($\mathcal{P}_1$), adjacent ($\mathcal{P}_2$) and overlapped ($\mathcal{P}_3$). The structure of derived sequences will play an important role in these studies.

For instance, $\omega_p\in\mathcal{P}_2$ means that there exists an integer q $(>p)$ such that the p-th and q-th occurrences of ω are adjacent. And $\omega\in\mathcal{P}_2$ means that there exist two integers p and q such that the p-th and q-th occurrences of ω are adjacent. In this case, $\omega\omega\prec\rho$. We call $\omega\omega$ a square in ρ. We consider $\omega=ab\in\mathbb{D}$ for example.

$$\mathbb{D}=\underbrace{ab}_{[1]}\,aa\,\underbrace{ab}_{[2]}\,\underbrace{ab}_{[3]}\,\underbrace{ab}_{[4]}\,aa\,\underbrace{ab}_{[5]}\,aa\ \cdots. \tag{3.1}$$

We denote the first five occurrences of ω by notations [1] to [5] in Eq. (3.1). Thus the first five occurrences of ω are $\mathbb{D}[1,2]$, $\mathbb{D}[5,6]$, $\mathbb{D}[7,8]$, $\mathbb{D}[9,10]$ and $\mathbb{D}[13,14]$.

$$\begin{cases} \text{The first (resp., 4-th, 5-th) occurrence of } \omega \text{ is not followed by another } \omega = ab; \\ \qquad\qquad \Longrightarrow \omega_1 \notin \mathcal{P}_2,\ \omega_4 \notin \mathcal{P}_2,\ \omega_5 \notin \mathcal{P}_2; \\ \qquad\qquad \Longrightarrow \text{square } \omega\omega = abab \text{ does not exist at these positions.} \\ \text{The second (resp., third) occurrence of } \omega \text{ is followed by another } \omega = ab; \\ \qquad\qquad \Longrightarrow \omega_2 \in \mathcal{P}_2,\ \omega_3 \in \mathcal{P}_2; \\ \qquad\qquad \Longrightarrow \text{square } \omega\omega = abab \text{ exists at these positions.} \end{cases}$$

3.1. The factor spectra in the Fibonacci sequence

Huang-Wen [3] discussed the factor spectra for some combinatorial properties in the Fibonacci sequence. We first define several subsets of $(\Omega_{\mathbb{F}}, \mathbb{N})$ such that the disjoint union of them is $(\Omega_{\mathbb{F}}, \mathbb{N})$, see Eq. (3.2) and Figure 5.

$$\begin{cases} S_1 \ \ = \bigcup_{m=-1}\{(\omega,p) \mid \mathrm{Ker}(\omega) = K_m, |\omega| = f_m\}, \\ S_{2.1} = \bigcup_{m=-1}\{(\omega,p) \mid \mathrm{Ker}(\omega) = K_m, |\omega| = f_{m+1}, \mathbb{F}[p] = a\}, \\ S_{2.2} = \bigcup_{m=-1}\{(\omega,p) \mid \mathrm{Ker}(\omega) = K_m, |\omega| = f_{m+1}, \mathbb{F}[p] = b\}, \\ S_{3.1} = \bigcup_{m=-1}\{(\omega,p) \mid \mathrm{Ker}(\omega) = K_m, |\omega| = f_{m+2}, \mathbb{F}[p] = a\}, \\ S_{3.2} = \bigcup_{m=-1}\{(\omega,p) \mid \mathrm{Ker}(\omega) = K_m, |\omega| = f_{m+2}, \mathbb{F}[p] = b\}, \\ S_4 \ \ = \bigcup_{m=-1}\{(\omega,p) \mid \mathrm{Ker}(\omega) = K_m, f_m < |\omega| < f_{m+1}\}, \\ S_{5.1} = \bigcup_{m=-1}\{(\omega,p) \mid \mathrm{Ker}(\omega) = K_m, f_{m+1} < |\omega| < f_{m+2}, \mathbb{F}[p] = a\}, \\ S_{5.2} = \bigcup_{m=-1}\{(\omega,p) \mid \mathrm{Ker}(\omega) = K_m, f_{m+1} < |\omega| < f_{m+2}, \mathbb{F}[p] = b\}, \\ S_6 \ \ = \bigcup_{m=-1}\{(\omega,p) \mid \mathrm{Ker}(\omega) = K_m, f_{m+2} < |\omega| < f_{m+3}\}. \end{cases} \tag{3.2}$$

Obviously, the disjoint union of $S_{2.1}$ and $S_{2.2}$ is $\bigcup_{m=-1}\{(\omega,p) \mid \mathrm{Ker}(\omega) = K_m, |\omega| = f_{m+1}\}$.

	$\mathbb{F}[p] = a$	$\mathbb{F}[p] = b$
$\lvert\omega\rvert = f_m$	S_1	
$\lvert\omega\rvert = f_{m+1}$	$S_{2.1}$	$S_{2.2}$
$\lvert\omega\rvert = f_{m+2}$	$S_{3.1}$	$S_{3.2}$
$f_m < \lvert\omega\rvert < f_{m+1}$	S_4	
$f_{m+1} < \lvert\omega\rvert < f_{m+2}$	$S_{5.1}$	$S_{5.2}$
$f_{m+2} < \lvert\omega\rvert < f_{m+3}$	S_6	

Figure 5 Several subsets of $(\Omega_{\mathbb{F}}, \mathbb{N})$ in Eq. (3.2)

Property 3.1 $Spt(\mathbb{F}, \mathcal{P}_1) = S_1 \cup S_{2.1} \cup S_4 \cup S_{5.1}$, $Spt(\mathbb{F}, \mathcal{P}_2) = S_{2.2} \cup S_{3.1}$, $Spt(\mathbb{F}, \mathcal{P}_3) = S_{3.2} \cup S_{5.2} \cup S_6$.

Using the factor spectrum $Spt(\mathbb{F}, \mathcal{P}_2)$, we get a new proof of the conclusion that all squares in $\mathbb{F}$ are of length $2f_m$. Furthermore, we find that not all occurrences of all factors of length f_m belong to $Spt(\mathbb{F}, \mathcal{P}_2)$.

3.2. The factor spectra in the Period-doubling sequence

Huang-Wen [29] discussed the factor spectra for some combinatorial properties in the Period-doubling sequence. We first define several subsets of $(\Omega_{\mathbb{D}}, \mathbb{N})$ such that the disjoint union of them is $(\Omega_{\mathbb{D}}, \mathbb{N})$, see Eq. (3.3) and Figure 6.

$$\begin{cases} S_1 \;\; = \bigcup_{m=1}\{(\omega,p) \mid \mathrm{Env}(\omega)=E_m^1, \mathbb{D}(\alpha,\beta\beta)[p]=\alpha\}, \\ S_{2.1} = \bigcup_{m=1}\{(\omega,p) \mid \mathrm{Env}(\omega)=E_m^1, |\omega|<2^{m-1}, \mathbb{D}(\alpha,\beta\beta)[p]=\beta\}, \\ S_{2.2} = \bigcup_{m=1}\{(\omega,p) \mid \mathrm{Env}(\omega)=E_m^1, |\omega|=2^{m-1}, \mathbb{D}(\alpha,\beta\beta)[p]=\beta\}, \\ S_{2.3} = \bigcup_{m=1}\{(\omega,p) \mid \mathrm{Env}(\omega)=E_m^1, |\omega|>2^{m-1}, \mathbb{D}(\alpha,\beta\beta)[p]=\beta\}, \\ S_3 \;\; = \bigcup_{m=1}\{(\omega,p) \mid \mathrm{Env}(\omega)=E_m^2, \mathbb{D}(\alpha\beta,\alpha\gamma\alpha\gamma)[p]=\alpha\}, \\ S_4 \;\; = \bigcup_{m=1}\{(\omega,p) \mid \mathrm{Env}(\omega)=E_m^2, \mathbb{D}(\alpha\beta,\alpha\gamma\alpha\gamma)[p]\neq\alpha\}. \end{cases} \tag{3.3}$$

The disjoint union of $S_{2.1}$, $S_{2.2}$ and $S_{2.3}$ is $\bigcup_{m=1}\{(\omega,p) \mid \mathrm{Env}(\omega)=E_m^1, \mathbb{D}(\alpha,\beta\beta)[p]=\beta\}$.

	$\mathbb{D}(\alpha,\beta\beta)[p]=\alpha$	$\mathbb{D}(\alpha,\beta\beta)[p]=\beta$
$\mathrm{Env}(\omega)=E_m^1$	S_1	$S_{2.1}$
		$S_{2.2}$
		$S_{2.3}$
$\mathrm{Env}(\omega)=E_m^2$	S_3	S_4
	$\mathbb{D}(\alpha\beta,\alpha\gamma\alpha\gamma)[p]=\alpha$	$\mathbb{D}(\alpha\beta,\alpha\gamma\alpha\gamma)[p]\neq\alpha$

Figure 6 Several subsets of $(\Omega_{\mathbb{D}}, \mathbb{N})$ in Eq. (3.3)

Property 3.2 $Spt(\mathbb{D},\mathcal{P}_1)=S_1\cup S_4\cup S_{2.1}$, $Spt(\mathbb{D},\mathcal{P}_2)=S_{2.2}$, $Spt(\mathbb{D},\mathcal{P}_3)=S_3\cup S_{2.3}$.

Using $Spt(\mathbb{D},\mathcal{P}_2)$, we get new proofs of some conclusions in Damanik [13]:

(1) All squares in $\mathbb{D}$ are $\bigcup_{m\geq 1}\{A_{m-1}[j+1,2^{m-1}]A_{m-1}A_{m-1}[1,j] \mid 0\leq j<2^{m-1}\}$;

(2) There are 2^{m-1}'s distinct squares with length 2^m for $m\geq 1$.

4. Enumerations

Using Theorems 2.6 and 2.9, we determined the numbers of palindromes, squares and cubes occurring in each factor of $\mathbb{F}$, $\mathbb{T}$ and $\mathbb{D}$, see Huang-Wen [11,12,30] for instance. These topics are of great importance in computer science.

4.1. Enumeration of palindromes in $\mathbb{F}$ and $\mathbb{T}$

A palindrome is a finite word that reads the same backwards as forwards. Let $\mathrm{Pal}_{\mathbb{F}}$ (resp., $\mathrm{Pal}_{\mathbb{T}}$) be all palindromes occurring in $\mathbb{F}$ (resp., $\mathbb{T}$). Some previous research has been done on "rich word", which is based on the number of distinct palindromes. A finite word ω is rich if and only if ω contains exactly $|\omega|+1$ distinct palindromes (including the empty word). An infinite word is rich if and only if all of its factors are rich. Droubay-Justin-Pirillo [31] proved that episturmian sequences are rich. Therefore, as special cases, $\mathbb{F}$ and $\mathbb{T}$ are rich. Thus the number of distinct palindromes in $\mathbb{F}[1,n]$ (resp., $\mathbb{T}[1,n]$) is $n+1$ for all n.

In Huang-Wen [11], we consider the numbers of repeated palindromes in $\mathbb{F}[1,n]$ and $\mathbb{T}[1,n]$. All results in this subsection are from this paper. Denote

$$A(n)=\#\{(\omega,p)\mid \omega\in \mathrm{Pal}_{\mathbb{F}},\omega_p\prec\mathbb{F}[1,n]\}\text{ and }B(n)=\#\{(\omega,p)\mid \omega\in \mathrm{Pal}_{\mathbb{T}},\omega_p\prec\mathbb{T}[1,n]\}.$$

The research on counting the repeated palindromes is not rich. From our knowledge, it seems the first work to study this problem. In related fields, the numbers of special types of factors have been investigated in recent years, such as squares, cubes, r-powers, palindromes, runs, Lyndon factors, etc [10, 21, 31–41].

The main difficulty of this problem is twofold:

(1) The positions of all occurrences for all palindromes are not easy to be determined. In [11], we overcome this difficulty by using the derived sequence properties of $\mathbb{F}$ and $\mathbb{T}$, which we introduced and studied in [3, 4], see also Theorem 2.6.

(2) Taking $\mathbb{F}$ for instance, by the derived sequence property of $\mathbb{F}$, we can find out all distinct palindromes in $\mathbb{F}[1,n]$. We can also count the number of occurrences of each palindrome. So the summation of these numbers are the numbers of repeated palindromes in $\mathbb{F}[1,n]$. But this method is complicated. We overcome this difficulty by studying the relations among positions of each ω_p, and establishing the recursive structure of $\mathrm{Pal}_{\mathbb{F}}$. Using the derived sequence properties and recursive structures, we give algorithms for counting $A(n)$ and $B(n)$, respectively.

Take $A(n)$ for instance. We have $A(n)=\sum_{i=1}^n a(i)$ where $a(i)$ is given in Property 4.1.

Property 4.1 *The vectors* $[a(1)]=[1]$, $[a(2),a(3)]=[1,2]$ *and for* $m\geq 3$

$$\begin{aligned}&[a(f_m-1),\dots,a(f_{m+1}-2)]\\&\quad=[a(f_{m-2}-1),\dots,a(f_{m-1}-2),a(f_{m-1}-1),\dots,a(f_m-2)]+[\underbrace{1,\dots,1}_{f_{m-1}}].\end{aligned}$$

The first few values of $a(n)$ are $[a(1)]=[1]$, $[a(2),a(3)]=[1,2]$,
$[a(4),a(5),a(6)]=[a(1),a(2),a(3)]+[1,1,1]=[2,2,3]$,
$[a(7),\dots,a(11)]=[a(2),\dots,a(6)]+[1,1,1,1,1]=[2,3,3,3,4]$,
$[a(12),\dots,a(19)]=[a(4),\dots,a(11)]+[\underbrace{1,\dots,1}_{8}]=[3,3,4,3,4,4,4,5]$.
$[a(20),\dots,a(32)]=[a(7),\dots,a(19)]+[\underbrace{1,\dots,1}_{13}]=[3,4,4,4,5,4,4,5,4,5,5,5,6]$.

We also get explicit expressions for some special n, such as: for $m\geq 0$,

$$\begin{cases}A(f_m)=\frac{m-3}{5}f_{m+2}+\frac{m-1}{5}f_m+m+3,\\ B(t_m)=\frac{m}{22}\big(10t_m+5t_{m-1}+3t_{m-2}\big)+\frac{1}{22}\big(-23t_m+12t_{m-1}-5t_{m-2}\big)+m+\frac{3}{2}.\end{cases}\tag{4.1}$$

We think this method for counting the repeated palindromes is valid for the m-bonacci word, and even valid for sturmian sequences, episturmian sequences etc. But now we only have the derived sequence properties of $\mathbb{F}$, $\mathbb{T}$ and $\mathbb{D}$. In the final remark in [11], we establish the cylinder structures and chain structures of $\mathrm{Pal}_{\mathbb{F}}$ and $\mathrm{Pal}_{\mathbb{T}}$. Using them, we prove some known results.

4.2. Enumeration of repetitions in $\mathbb{F}$

The fractional power is a topic dealing with repetitions in words. We say a (finite or infinite) word ω contains a r-power (real $r > 1$) if ω has a factor of the form $x^{\lfloor r \rfloor}x'$ where x' is a prefix of x and $|x^{\lfloor r \rfloor}x'| \geq r|x|$ see [1]. In this case, we call $x^{\lfloor r \rfloor}x'$ a r-power with *size* $|x|$. For instance, taking $x = ab$, then $\mathbb{F}[4,8] = ababa$ is a $\frac{5}{2}$-power of size $|ab| = 2$ in $\mathbb{F}$. Obviously the notion r-power is a generalization of square (2-power) and cube (3-power).

The study of power of a word has a long history. There are many significant contributions, for example [9, 10, 32–35, 39, 42–44]. In particular, Iliopoulos-Moore-Smyth [35] computed the positions of all squares in $\mathbb{F}$. Fraenkel-Simpson [32,33] obtained the number of squares in $F_m = \sigma^m(a)$. Damanik-Lenz [43] studied the index of Sturmian sequences. Glen [34] determined all of the squares (and subsequently higher powers) occurring in episturmian words. Using the results in [34], it was possible to determine the exact number of distinct squares in each building block (for instance, the building blocks in the Fibonacci sequence is $F_m = \sigma^m(a)$), which extends Fraenkel and Simpson's result [32,33]. Du-Mousavi-Schaeffer-Shallit [44,45] obtained the numbers of repeated squares and cubes in $\mathbb{F}[1,n]$ for all $n \geq 1$. All their numerations start from the first letter of the sequence.

In Huang-Wen [12], we count the number of distinct r-powers in $\mathbb{F}[i, n+i-1]$ for all $i, n \geq 1$ and $r \geq 2$, denoted by $\mathrm{D}(r,i,n)$. Our numeration can start from any letter of the sequence, comparing starting from the first letter, there are some difficulties. To overcome these difficulties, we introduce a new notion called the position sequence. In Huang-Wen [30], we count the number of repeated r-powers in $\mathbb{F}[i, n+i-1]$ for all $i, n \geq 1$ and $r \geq 2$, denoted by $\mathrm{R}(r,i,n)$. We give precise results for $r \in \{2, 2+\epsilon, 3\}$, where ϵ is a small positive number.

The methods of counting the distinct or repeated factors in $\mathbb{F}[i, n+i-1]$ is quite different. Take $r = 2$ (squares) for instance. Let

$$h = \lfloor \frac{\ln(\sqrt{5}n/2)}{\ln \alpha} + \frac{2}{n} \rfloor - 2 \quad \text{(i.e., } 2f_h \leq n < 2f_{h+1}\text{)}.$$

(1) A well known result is that the number of distinct squares of length $2f_m$ in the Fibonacci sequence is f_m (see [9]). By the arguments in [12, Subsection 5.2], $\mathbb{F}[i, n+i-1]$ contains all distinct squares of length $2f_m$ for any $m \geq 0$ and $n \geq f_{m+3} + 2f_m - 1$. Thus in order to count the number of distinct square in $\mathbb{F}[i, n+i-1]$ for $n \geq 6$, we only need to pay close attention to the number of the squares of length f_m where $m \in \{h-2, h-1, h\}$. Through careful observation and analysis, we get explicit expression of the number of distinct fractional powers in each factor of the Fibonacci sequence.

(2) Obviously, this method above is not fit for counting the repeated squares. Comparing with the distinct case, we will introduce two completely different tools, which we call square tree and matrix decomposition. The former is used for counting the number of repeated r-powers in $\mathbb{F}[1, n+i-1]$, i.e., $\mathrm{R}(r,1,n)$. The latter is used for determining the difference between $\mathrm{R}(r,i,n)$ and $\mathrm{R}(r,1,n)$.

Acknowledgements We thank the referees for their time and comments.

References

[1] J. P. ALLOUCHE, J. SHALLIT. *Automatic sequences: Theory, applications, generalizations.* Cambridge University Press, Cambridge, 2003.

[2] F. DURAND. *A characterization of substitutive sequences using return words.* Discrete Math., 1998, **179**(1-3): 89–101.

[3] Yuke HUANG, Zhiying WEN. *The sequence of return words of the Fibonacci sequence.* Theoret. Comput. Sci., 2015, **593**: 106–116.

[4] Yuke HUANG, Zhiying WEN. *Kernel words and gap sequence of the Tribonacci sequence.* Acta Math. Sci. Ser. B (Engl. Ed.), 2016, **36**(1): 173–194.

[5] M. LOTHAIRE. *Combinatorics on Words.* Addison-Wesley Publishing Co., Reading, Mass., 1983.

[6] M. LOTHAIRE. *Algebraic Combinatorics on Words.* Cambridge Univ. Press, Cambridge, 2002.

[7] J. BERSTEL. *Recent Results in Sturmian Words.* World Scientific, Singapore, 1966.

[8] J. BERSTEL, *Mot de Fibonacci, Séminaire d'informatique thérique.* L.I.T.P., Paris, Année, 1980/1981, 57–78.

[9] Zhixiong WEN, Zhiying WEN. *Some properties of the singular words of the Fibonacci word.* European J. Combin., 1994, **15**(6): 587–598.

[10] Bo TAN, Zhiying WEN. *Some properties of the Tribonacci sequence.* European J. Combin., 2007, **28**(6): 1703–1719.

[11] Yuke HUANG, Zhiying WEN. *The numbers of repeated palindromes in the Fibonacci and tribonacci words.* Discrete Appl. Math., 2017, **230**: 78–90.

[12] Yuke HUANG, Zhiying WEN. *The number of fractional powers in the Fibonacci word.* preprint.

[13] D. DAMANIK. *Local symmetries in the period–doubling sequence.* Discrete Appl. Math., 2000, **100**(1-2): 115–121.

[14] J. P. ALLOUCHE, J. PEYRIÈRE, Zhixiong WEN, et al. *Hankel determinants of the thue-morse sequence.* Ann. Inst. Fourier (Grenoble), 1998, **1**(1): 1–27.

[15] J. P. ALLOUCHE, M. BAAKE, J. CASSAIGNE, et al. *Palindrome complexity.* Theoret. Comput. Sci., 2003, **292**(1): 9–31.

[16] S. BROWN. *Squares and overlaps in the Thue–Morse sequence and some variants.* Theor. Inform. Appl., 2006, **40**(3): 473–484.

[17] É. CHARLIER, J. LEROY, M. RIGO. *Asymptotic properties of free monoid morphisms.* Linear Algebra Appl., 2015, **500**: 119–148.

[18] É. CHARLIER, N. RAMPERSAD, J. SHALLIT. *Enumeration and Decidable Properties of Automatic Sequences.* Internat. J. Found. Comput. Sci., 2012, **23**(5): 1035–1066.

[19] D. DAMANIK. *Singular continuous spectrum for the period doubling Hamiltonian on a set of full measure.* Comm. Math. Phys., 1998, **196**(2): 477–483.

[20] D. DAMANIK. *Uniform singular continuous spectrum for the period doubling Hamiltonian.* Ann. Henri Poincaré, 2001, **2**(1): 101–108.

[21] X. DROUBAY. *Palindromes in the Fibonacci word.* Inform. Process. Lett., 1995, **55**(4): 217–221.

[22] Han FU, Guoniu HAN. *On t-extensions of the Hankel determinants of certain automatic sequences.* Theoret. Comput. Sci., 2015, **562**: 46–56.

[23] R. J. FOKKINK, C. KRAAIKAMP, J. SHALLIT. *Hankel matrices for the period-doubling sequence.* Indag. Math. (N.S.), 2017, **28**(1): 108–119.

[24] Yingjun GUO, Zhixiong WEN. *Automaticity of the Hankel determinants of difference sequences of the Thue-Morse sequence.* Theoret. Comput. Sci., 2014, **552**(4): 1–12.

[25] E. LIVIOTTI. *A study of the structure factor of Thue-Morse and period-doubling chains by wavelet analysis.* J. Phys. Condensed Matter, 1998, **8**(27): 5007.

[26] M. MAKAROV. *On the infinite permutation generated by the period doubling word.* European J. Combin., 2010, **31**(1): 368–378.

[27] A. PARREAU, M. RIGO, E. ROWLAND, et al. *A new approach to the 2-regularity of the l-abelian complexity of 2-automatic sequences.* Electron. J. Combin., 2015, **22**(1): 1–27.

[28] M. RIGO, P. SALIMOV, É. VANDOMME. *Some properties of abelian return words.* J. Integer Seq., 2013, **16**(2): Article 13.2.5, 21 pp.

[29] Yuke HUANG, Zhiying WEN. *Derived sequences and factor spectrum of the period-doubling sequence.* Preprint.

[30] Yuke HUANG, Zhiying WEN. *Enumeration of repetitions in the Fibonacci word.* Preprint.

[31] X. DROUBAY, J. JUSTIN, G. PIRILLO. *Episturmian words and some constructions of de Luca and Rauzy.* Theoret. Comput. Sci., 2001, **255**(1-2): 539–553.

[32] A. S. FRAENKEL, J. SIMPSON. *The exact number of squares in Fibonacci words.* Theoret. Comput. Sci., 1999, **218**(1): 95–106.

[33] A. S. FRAENKEL, J. SIMPSON. *Corrigendum to "The exact number of squares in Fibonacci words".* Theoret. Comput. Sci., 2014, **547**: 122.

[34] A. GLEN. *On Sturmian and Episturmian Words, and Related Topics.* PhD Thesis, The University of Adelaide, Australia. 2006.

[35] C. S. ILIOPOULOS, D. MOORE, W. F. SMYTH. *A characterization of the squares in a Fibonacci string.* Theoret. Comput. Sci., 1997, **172**(1-2): 281–291.

[36] J. KARHUMAKI. *On cube-free ω-words generated by binary morphism.* Discrete Appl. Math., 1983, **5**(3): 279–297.

[37] D. KRIEGER. *Critical Exponents and Stabilizers of Infinite Words.* Theoret. Comput. Sci., 2008, **400**(1-3): 169–181.

[38] F. MIGNOSI, G. PIRILLO. *Repetitions in the Fibonacci infinite word.* RAIRO Inform. Théor. Appl., 1992, **26**(3): 199–204.

[39] H. MOUSAVI, J. SHALLIT. *Mechanical Proofs of Properties of the Tribonacci Word.* Springer, Cham, 2015.

[40] H. MOUSAVI, L. SCHAEFFER, J. SHALLIT. *Decision algorithms for Fibonacci–automatic words, I: Basic results.* RAIRO Theor. Inform. Appl., 2016, **50**(1): 39–66.

[41] W. RYTTER. *The structure of subword graphs and suffix trees of Fibonacci words.* Theoret. Comput. Sci., 2006, **363**(2): 211–223.

[42] K. SAARI, *Everywhere α–repetitive sequences and Sturmian words.* European J. Combin., 2010, **31**(1): 177–192.

[43] D. DAMANIK, D. LENZ. *Powers in Sturmian sequences.* European J. Combin., 2003, **24**(4): 377–390.

[44] Chenfei DU, H. MOUSAVI, L. SCHAEFFER, et al. *Decision algorithms for Fibonacci-automatic words, III: enumeration and abelian properties.* Internat. J. Found. Comput. Sci., 2016, **27**(8): 943–963.

[45] Chenfei DU, H. MOUSAVI, L. SCHAEFFER, et al. *Decision Algorithms for Fibonacci-Automatic Words, with Applications to Pattern Avoidance.* Presented at 15^e Journées Montoises d'Informatique Théorique, 23–26 September 2014, Nancy, France.