Mathematical Methods in the Applied Sciences

WILEY

RESEARCH ARTICLE OPEN ACCESS

Construction of Bernstein-Based Words and Their Patterns

Irem Kucukoglu[1] | Yilmaz Simsek[2]

[1]Department of Engineering Fundamental Sciences, Faculty of Engineering, Alanya Alaaddin Keykubat University, Antalya, Turkey | [2]Department of Mathematics, Faculty of Science, University of Akdeniz, Antalya, Turkey

Correspondence: Irem Kucukoglu (irem.kucukoglu@alanya.edu.tr)

Received: 22 November 2023 | **Revised:** 24 March 2025 | **Accepted:** 4 May 2025

Funding: The authors received no specific funding for this work.

Keywords: Bernstein basis functions | combinatorics on words | computational implementation | Farey fractions | recurrence relation | generating functions | tree diagrams | special functions

ABSTRACT

In this paper, with inspiration of the definition of Bernstein basis functions and their recurrence relation, we give construction of a new word family that we refer Bernstein-based words. By classifying these special words as the first and second kinds, we investigate their some fundamental properties involving periodicity and symmetricity. Providing schematic algorithms based on tree diagrams, we also illustrate the construction of the Bernstein-based words. For their symbolic computation, we also give computational implementations of Bernstein-based words in the Wolfram Language. By executing these implementations, we present some tables of Bernstein-based words and their decimal equivalents. In addition, we present black–white and four-colored patterns arising from the Bernstein-based words with their potential applications in computational science and engineering. We also give some finite sums and generating functions for the lengths of the Bernstein-based words. We show that these functions are of relationships with the Catalan numbers, the centered m-gonal numbers, the Laguerre polynomials, certain finite sums, and hypergeometric functions. We also raise some open questions and provide some comments on our results. Finally, we investigate relationships between the slopes of the Bernstein-based words and the Farey fractions.

MSC2020 Classification: 68R15, 05B05, 05A15, 11B37, 11B57

1 | Introduction, Definitions, and Preliminaries

The field of combinatorics on words is a quite new field that has been started to be studied in recent years by the researchers working on multifarious branches of mathematics such as number theory, group theory, theoretical computer science dealing with automata, and formal languages. Combinatorics on words concentrates on the study of formal languages, words, and strings formed by letters or symbols. In this aspect, the field of combinatorics on words is in essence to differ from combinatorics. The main idea behind the field of combinatorics on words is to make an investigation on words in algebraic, combinatorial, or algorithmic way. With the emergence of the book of Lothaire [1], providing a terminological and well-defined theory on combinatorics on words, this field has started to develop and grow even more. These developments encourage many researchers to define new word classes and still find their interesting and useful applications. Based upon the consequence of these developments, the source of our motivation in this paper is to construct new words, called Bernstein-based words, and present some their fundamental properties.

We first start with reminding terminology of the combinatorics on words, which are reminded by blending them from the books of Lothaire [1–3], and Shallit [4].

Let Σ be a nonempty set called alphabet, each element of which is called a letter. A finite sequence of letters, in the following form:

$$w = (a_1, a_2, \dots, a_n), \quad \forall a_i \in \Sigma; \quad i = 1, 2, \dots, n,$$

is called a finite word of length n over the alphabet Σ. If we use Σ^* to denote the set of all words over the alphabet Σ, then $w \in \Sigma^*$.

The concatenation, or so-called juxtaposition, is a binary operation on the set Σ^* with the following mapping:

$$\Sigma^* \times \Sigma^* \mapsto \Sigma^*$$

concatenating two words as in the following way:

$$w_1 w_2 = (a_1, a_2, \dots, a_n, b_1, b_2, \dots, b_m). \tag{1}$$

where $w_1 = (a_1, a_2, \dots, a_n) \in \Sigma^*$ and $w_2 = (b_1, b_2, \dots, b_m) \in \Sigma^*$.

It is clear that the concatenation is well-defined, closed, and an associative binary operation which is not commutative. Moreover, the set Σ^* has an identity element ϵ denoting the empty word which is a neutral element for concatenation. Due to these features, when it is equipped with the concatenation of two words, Σ^* forms an algebraic structure called free monoid over the alphabet Σ. Therefore, a word in the form of

$$w = (a_1, a_2, \dots, a_n)$$

can be expressed as follows:

$$w = a_1 a_2 \cdots a_n \tag{2}$$

(*cf*. [1–5]).

In addition, recall that the nth power of a word w is obtained by concatenating itself side by side as follows:

$$w^n = \begin{cases} \underbrace{ww\cdots w}_{n-\text{times}} & \text{if } n \in \mathbb{N}, \\ \epsilon & \text{if } n = 0, \end{cases}$$

(*cf*. [4, p. 3]). For example, $(011)^2 = 011011$.

We also recall the length of the word w which is the number of letters that forms the word w and denoted by $|w|$. Thus,

$$|a_1 a_2 \cdots a_n| = n$$

(see, for details, [1–3, 5]). For example, the word $w = 011011$ has length 6. Thus, we have $|011011| = 6$.

As for the Bernstein basis functions, $B_k^n(x)$, these functions are given by the following explicit formula involving the classical binomial coefficient:

$$B_k^n(x) = \binom{n}{k} x^k (1-x)^{n-k}, \quad (k = 0, 1, \dots, n; \ n \in \mathbb{N}_0), \tag{3}$$

which have relationships with a large number of concepts including the Catalan numbers, the binomial distribution, the proof of the Weierstrass approximation theorem, the Poisson distribution, and computer-aided geometric design (CAGD) involving Bezier curves and surfaces, splines, and so on. Moreover, these functions have found a wide variety of applications to themselves in areas of mathematics (especially in generating functions theory, probability theory, and approximation theory), engineering (especially in automobile engineering, machine learning, human-computer interaction systems and etc.), and almost all areas in recent years. Furthermore, the Bernstein basis functions are also used in some analytical methods such as the homotopy perturbation method and the variational iteration method (e.g., see [6], and the references cited therein). For other details regarding the Bernstein basis functions, also see [7–13], and also the cited references therein.

The recurrence relation for the Bernstein basis functions is given by the following:

$$B_k^n(x) = (1-x) B_k^{n-1}(x) + x B_{k-1}^{n-1}(x) \tag{4}$$

such that $B_0^0(x) = 1$ and $B_k^n(x) = 0$ for $k < 0$ and $k > n$ (*cf*. [9, 10, 13, 14]).

The Bernstein basis functions satisfy the following symmetry identity (*cf*. [9, 10, 13, 14]):

$$B_{n-k}^n(1-x) = B_k^n(x). \tag{5}$$

As stated in Section 2, the reason why we named our words as *Bernstein-based words* is that they are constructed by the inspiration arising from the combinations of Equations (1–4).

Before presenting our main results in the next sections, we shall briefly summarize other auxiliary concepts and their definitions needed to obtain the findings of this paper as follows:

The Catalan numbers are defined by the following:

$$C_m = \frac{1}{m+1}\binom{2m}{m}; \quad (m \in \mathbb{N}_0) \quad \text{and} \quad C_m = \prod_{k=2}^{m} \frac{m+k}{k}; \quad (m \geq 3) \tag{6}$$

which is also given by the following ordinary generating function:

$$\sum_{m=0}^{\infty} C_m t^m = \frac{1 - \sqrt{1-4t}}{2t} \tag{7}$$

where $0 < |t| \leq \frac{1}{4}$ (*cf*. [15]).

The Catalan numbers arise in the solution of many kinds of combinatorial and real-world problems such as the Euler's polygon problem and polygon triangulations, ballot sequences, parenthesizations, Dyck paths, Motzkin paths, binary trees, plane trees, and various kinds of enumeration problems. For some other applications in detail, see the book of Koshy [15].

The generalized hypergeometric series ${}_kF_r(\alpha_1, \dots, \alpha_k; \beta_1, \dots, \beta_r; z)$ is defined by

$${}_kF_r(\alpha_1, \dots, \alpha_k; \beta_1, \dots, \beta_r; z) = \sum_{n=0}^{\infty} \left(\frac{\prod_{j=1}^{k} (\alpha_j)_n}{\prod_{j=1}^{r} (\beta_j)_n} \right) \frac{z^n}{n!}, \tag{8}$$

where the above series converges for all z if $k < r+1$ and for $|z| < 1$ if $k = r+1$. Assuming that all parameters have general values, real or complex, except for the $\beta_j; (j = 1, 2, \ldots, r)$, none of which is equal to zero or a negative integer such that $(\beta)_v$ denotes Pochhammer's symbol, defined by

$$(\beta)_v = \prod_{j=0}^{v-1}(\beta + j),$$

and $(\beta)_0 = 1$ for $\beta \neq 1$, $v \in \mathbb{N}$, and $\beta \in \mathbb{C}$ (*cf*. [16, 17]).

Considering

$$\binom{\omega}{m} = \frac{\prod_{j=0}^{m-1}(\omega - j)}{m!} \quad \text{and} \quad \binom{\omega}{0} = 1,$$

the second author [16], introduced the sum $B_v(\omega; \lambda, p)$, involving higher powers of inverse binomial coefficients, by the following formula:

$$B_v(\omega; \lambda, p) = \sum_{m=0}^{\infty} \frac{m^v \lambda^m}{\binom{\omega}{m}^p}, \tag{9}$$

whose generating function is given by the following hypergeometric series:

$${}_{p+1}F_p(1, \ldots, 1; -\omega, \ldots, -\omega; (-1)^p \lambda e^z) = \sum_{v=0}^{\infty} B_v(\omega; \lambda, p)\frac{z^v}{v!} \tag{10}$$

where $v,\ p \in \mathbb{N}_0,\ -\omega \notin \{0, -1, -2, -3, \ldots\}$ and $\lambda \in \mathbb{R}$ (or $\mathbb{C}$) with $|\lambda| < 1$ (*cf*. [16]).

In [18], the second author also introduced the combinatorial numbers $y_6(n, k; \lambda, p)$, involving higher powers of binomial coefficients, by the following formula, for $n, m, p \in \mathbb{N}_0$:

$$y_6(m, n; \lambda, p) = \frac{1}{n!}\sum_{k=0}^{n}\binom{n}{k}^p k^m \lambda^k, \tag{11}$$

and constructed the following generating functions for these numbers in terms of the hypergeometric series as follows:

$$\frac{1}{n!}\,{}_pF_{p-1}(-n, \ldots, -n; 1, \ldots, 1; (-1)^p \lambda e^z) = \sum_{m=0}^{\infty} y_6(m, n; \lambda, p)\frac{z^m}{m!}, \tag{12}$$

where $n, p \in \mathbb{N}$ and $\lambda \in \mathbb{R}$ (or $\mathbb{C}$).

The numbers $y_6(m, n; \lambda, p)$ are referred to in the literature as the combinatorial Simsek numbers of the sixth kind (for details, see [11], and the references cited therein).

Now we briefly summarize our results in the next sections as follows.

In Section 2, we introduce Bernstein-based words and investigate their fundamental properties with examples and tables. We also give schematic algorithms of these words.

In Section 3, we provide computational implementations for evaluating the Bernstein-based words in the Wolfram language.

In Section 4, we construct some finite sums and generating functions for the lengths of the Bernstein-based words. We also derive some relations and results pertaining to the length of the Bernstein-based words.

In Section 5, we give relations between the slopes of the Bernstein-based words and the Farey fractions.

In the final section, we conclude the paper by providing some brief summary and observations on the results of this paper and future plans.

2 | Bernstein-Based Words

In this section, inspired by the explicit formula (3) and the recurrence relation (4) of the Bernstein basis functions, we introduce two kinds of Bernstein-based words over the alphabet $\Sigma = \{0, 1\}$.

2.1 | Bernstein Words of the First Kind

Here, by the following definition, inspired by the explicit formula (3) of the Bernstein basis functions, we first define so-called Bernstein words of the first kind as in the following definition.

Definition 1. Let $n, k \in \mathbb{N}_0$. Let $x \in \Sigma = \{0, 1\}$. By considering the product between letters or words as their concatenation, the Bernstein words of the first kind $w_B(x; n, k)$ over the alphabet $\Sigma = \{0, 1\}$ are defined by the following:

$$w_B(x; n, k) = \begin{cases} (0^k \cdot 1^{n-k})^{\binom{n}{k}} & \text{if} \quad x = 0, \\ (1^k \cdot 0^{n-k})^{\binom{n}{k}} & \text{if} \quad x = 1 \end{cases} \tag{13}$$

with $w_B(x; 0, 0) = \epsilon$ and $w_B(x; n, k) = \epsilon$ when $k < 0$ or $k > n$.

To illustrate Definition 1, a few Bernstein words of the first kind are exemplified as follows:

$$w_B(0; 3, 0) = 111 = (1)^3$$
$$w_B(0; 3, 1) = 011011011 = (011)^3$$
$$w_B(0; 3, 2) = 001001001 = (001)^3$$
$$w_B(0; 3, 3) = 000 = (0)^3$$

and

$$w_B(1; 3, 0) = 000 = (0)^3$$
$$w_B(1; 3, 1) = 100100100 = (100)^3$$
$$w_B(1; 3, 2) = 110110110 = (110)^3$$
$$w_B(1; 3, 3) = 111 = (1)^3$$

Next, by using (13), we provide some properties of the words $w_B(x; n, k)$ as follows.

Periodicity property: It is known that a periodic word can be expressed as a positive power of a shorter word (*cf*. [19, 20], and see also cited references therein). The definition, given by (13), means that we first juxtapose

k-times 0's (or 1's) with $(n-k)$-times 1's (or 0's). Then, the string obtained from the first process is brought side by side $\binom{n}{k}$ times to obtain the word $w_B(x; n, k)$. Here, $\binom{n}{k}$ times juxtaposition means that the words $w_B(x; n, k)$ can be expressed as a positive power of a shorter word. That is, the words $w_B(x; n, k)$ are all periodic.

Symmetry property with respect to vertical reflection: Let $a_1, \dots, a_n$ be letters of an alphabet Σ. Then, recall that the reversal (or so-called mirror image) of a word $w = a_1a_2 \dots a_n$ is defined by the word

$$w^R = a_na_{n-1} \dots a_1$$

(*cf*. [4, p. 11], [3, p. 4]). By applying the above definition to (13), it can be concluded that the words $w_B(x; n, k)$ satisfy the following symmetry properties:

$$\left(w_B(0; n, n-k)\right)^R = w_B(1; n, k)$$

and

$$\left(w_B(1; n, n-k)\right)^R = w_B(0; n, k).$$

Remark 1. Observe that the above symmetry properties are word analogs of (5).

These symmetry properties also mean that a concatenation of the words

$$w_B(0; n, n-k) \quad \text{and} \quad w_B(1; n, k)$$

or

$$w_B(1; n, n-k) \quad \text{and} \quad w_B(0; n, k)$$

generates a palindrome word spelling that is the same backward as forward. For some applications of palindrome words, see also [21].

For instance, let us consider the following words, which are reversal of each other as follows:

$$w_B(0; 3, 2) = 001001001 \quad \text{and} \quad w_B(1; 3, 1) = 100100100.$$

The spelling or pronunciation of any of the above forwards is the same as the spelling or pronunciation of the other backwards. The concatenation of them is given as follows:

$$w_B(0; 3, 2)w_B(1; 3, 1) = 001001001100100100,$$

which is a member of palindrome words.

Remark 2. There are many other applications of (13). For instance, Ruskey et al. [22] used word analogs associated with (13) as factors of gray codes while investigating the binary bubble languages and cool-lex order.

2.2 | Bernstein Words of the Second Kind

Here, inspired by the recurrence relation (4) of the Bernstein basis functions, secondly we define Bernstein words of the second kind as in the following definition.

Definition 2. Let $n, k \in \mathbb{N}_0$. Let $x \in \Sigma = \{0, 1\}$. By considering the product between letters or words as their concatenation, the Bernstein words of the second kind $\mathcal{W}_B(x; n, k)$ over the alphabet $\Sigma = \{0, 1\}$ are defined by the following recurrence relation:

$$\mathcal{W}_B(x; n, k) = \begin{cases} 1 \cdot \mathcal{W}_B(x; n-1, k) \cdot 0 \cdot \mathcal{W}_B(x; n-1, k-1) & \text{if } x = 0, \\ 0 \cdot \mathcal{W}_B(x; n-1, k) \cdot 1 \cdot \mathcal{W}_B(x; n-1, k-1) & \text{if } x = 1 \end{cases} \tag{14}$$

with $\mathcal{W}_B(x; 0, 0) = 1$ and $\mathcal{W}_B(x; n, k) = 0$ when $k < 0$ or $k > n$.

To illustrate Definition 2, by substituting $x = 0$, $k = 1$, and $n = 1$ into (14), we get

$$\mathcal{W}_B(0; 1, 1) = 1 \cdot \mathcal{W}_B(0; 0, 1) \cdot 0 \cdot \mathcal{W}_B(0; 0, 0) = 1001.$$

In the case when $x = 1$, $k = 1$, and $n = 1$, Equation (14) implies

$$\mathcal{W}_B(1; 1, 1) = 0 \cdot \mathcal{W}_B(1; 0, 1) \cdot 1 \cdot \mathcal{W}_B(1; 0, 0) = 0011.$$

Next, by using (14), we provide some properties of the words $\mathcal{W}_B(x; n, k)$ as follows.

As can be seen from the two examples above, the Bernstein words of the second kind are *not periodic* since they cannot be expressed as a positive power of a shorter word.

Observe also that unlike the Bernstein words of the first kind, the Bernstein words of the second kind do not satisfy the symmetry property with respect to vertical reflection. However, in this study, we may raise the following open question:

Open Question 1: *When we consider the set of all Bernstein words of the second kind, which subclasses of this set can be symmetric with respect to vertical reflection or periodic, or none?*

2.3 | Tree Diagram for Construction of the Bernstein Words of the Second Kind

To illustrate the construction of the Bernstein words of the second kind in a schematic way, in Figure 1, we give a tree diagram which shows the construction of the words $\mathcal{W}_B(0; n, k)$.

In Figure 1, blue directed edges (left) of the tree correspond to the concatenation by 1 from left (namely, juxtapose with the prefix 1) and red directed edges (right) of the tree correspond to the concatenation by 0 from left (namely, juxtapose with the prefix 0). Let the letter 1 be the root of the tree. In order to generate words in any next level of the trees, we concatenate two new words derived from the rule on the directed edges out of the previous nodes connecting to the corresponding node.

To show the construction of the words $\mathcal{W}_B(1; n, k)$, we also give another tree diagram by Figure 2.

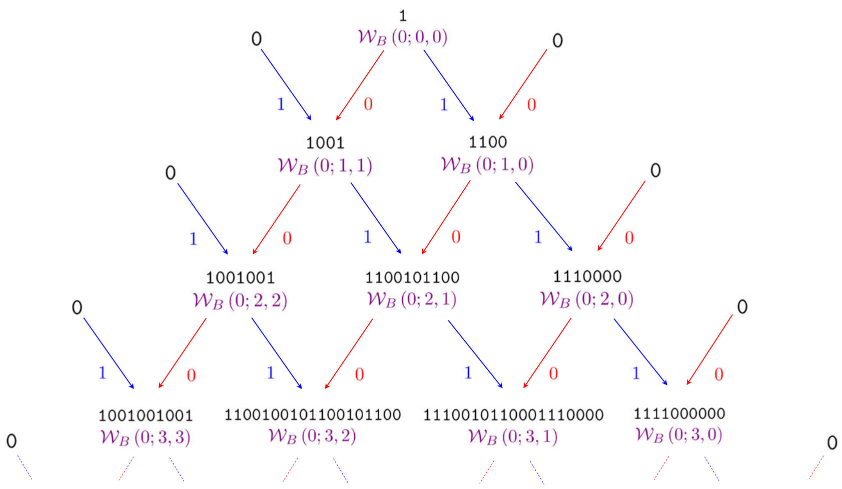

FIGURE 1 | Tree diagram which shows the construction of the Bernstein words $\mathcal{W}_B(0;n,k)$ of the second kind. [Colour figure can be viewed at wileyonlinelibrary.com]

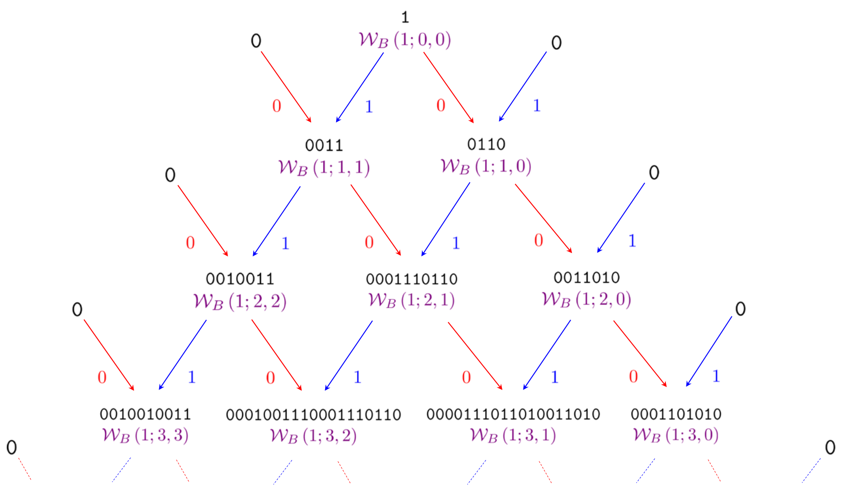

FIGURE 2 | Tree diagram which shows the construction of the Bernstein words $\mathcal{W}_B(1;n,k)$ of the second kind. [Colour figure can be viewed at wileyonlinelibrary.com]

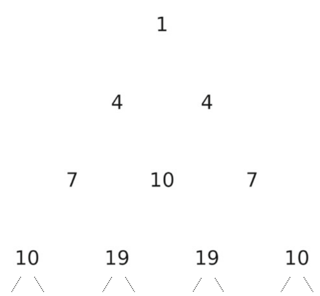

FIGURE 3 | Lengths of the Bernstein words of the second kind, appeared in the given tree diagrams by Figure 1 and Figure 2, with the same geometric pattern.

In Figure 2, red directed edges (left) of the tree correspond to the concatenation by 0 from left (namely, juxtapose with the prefix 0) and blue directed edges (right) of the tree correspond to the concatenation by 1 from left (namely, juxtapose with the prefix 1). Similarly, let the letter 1 be the root of the tree. To generate words in any next level of the trees, we concatenate two new words derived from the rule on the directed edges out of the previous nodes connecting to the corresponding node.

Remark 3. The tree diagrams, given in Figures 1 and 2, are provided to be help the researchers make some constructions and algorithmic applications in fields of graph theory, automata theory, and cryptology.

In Figure 3, we give the lengths of the Bernstein words of the second kind, appeared in the given tree diagrams in Figures 1 and 2, with the same geometric pattern. Note that the sequences arising from these lengths will be discussed later in Section 4.

3 | Computational Implementations of Bernstein-Based Words

In this section, we provide a procedure `BernsteinWordsKind1` (see: Implementation 1) by implementing (13), and also, we provide another procedure `BernsteinWordsKind2` (see: Implementation 2) by implementing the recurrence relation (14) in the Wolfram Language. By executing the procedures `BernsteinWordsKind1` and `BernsteinWordsKind2` in the Wolfram Mathematica version 12.0 and using the command `TableForm`, we present tables of the Bernstein words of the first kind $w_B(x;n,k)$ and Bernstein words of the second kind $\mathcal{W}_B(x;n,k)$ obtained just for a few special cases (among others).

3.1 | Computational Implementations for the Bernstein Words of the First Kind

Here, we provide computational implementations for the Bernstein words of the first kind in the Wolfram language.

Implementation 1 The following code, involving the procedure `BernsteinWordsKind1` written in the Wolfram Language, returns the words $w_B(x;n,k)$ for $x \in \Sigma = \{0,1\}$.

```
BernsteinWordsKind1[x_?StringQ, n_?IntegerQ,
 k_?IntegerQ] /; k < 0 || k > n:= "\[ Epsilon ]"
BernsteinWordsKind1[x_?StringQ, 0, 0] := "\[ Epsilon ]"
BernsteinWordsKind1[x_?StringQ, n_?IntegerQ,
 k_?IntegerQ] := First[{ Factor1CaseZero="" <>Table
 ["0",{ j ,1, k }]; Factor2CaseZero="" <>Table
 ["1",{ j ,1, n−k}]; Factor1CaseOne="" <>Table["1",{j ,1, k }];
Factor2CaseOne="" <>Table["0",{j ,1, n−k}];
Which[x=="0",result=Factor1CaseZero<>Factor2CaseZero,
x =="1", result =Factor1CaseOne<>Factor2CaseOne];
 result =""<>Table[""<>result,{j,1, Binomial[n,k]}]}]
```

By using Implementation 1 and the auxiliary commands of Wolfram language, we provide the following code written in Wolfram language:

```
TableForm[Evaluate[Table[BernsteinWordsKind1["0", n, k],
 {n, 0, 5}, {k, 0, 2}]], TableHeadings −>{{"n=0", "n=1
", "n=2", "n=3", "n=4", "n=5"}, {"k=0", "
k=1", "k=2"}}],
```

which returns Table 1, whose entries are the Bernstein words of the first kind $w_B(x;n,k)$, in the case when $x=0$, $n \in \{0,1,2,3,4,5\}$ and $k \in \{0,1,2\}$.

In addition, by the following code written in Wolfram language,

```
TableForm[Evaluate[Table[BernsteinWordsKind1["1", n, k],
  {n, 0, 5}, {k, 0, 2}]], TableHeadings −>{{"n=0", "n=1
", "n=2", "n=3", "n=4", "n=5"}, {"k=0", "k=1", "
k=2"}}]
```

we get Table 2, whose entries are the Bernstein words of the first kind $w_B(x;n,k)$, in the case when $x=1$, $n \in \{0,1,2,3,4,5\}$ and $k \in \{0,1,2\}$.

Note that the entries ϵ of Tables 1 and 2 denote the empty word.

3.2 | Computational Implementations for the Bernstein Words of the Second Kind

Here, we provide computational implementations for the Bernstein words of the second kind in the Wolfram language.

Implementation 2 The following code, involving the procedure `BernsteinWordsKind2` written in the Wolfram Language, returns the words $\mathcal{W}_B(x;n,k)$ for $x \in \Sigma = \{0,1\}$.

```
BernsteinWordsKind2[x_?StringQ, n_?IntegerQ,
 k_?IntegerQ] /; k < 0 || k > n:= "0";
BernsteinWordsKind2[x_?StringQ, 0, 0 ] := "1";
BernsteinWordsKind2[x_?StringQ, n_?IntegerQ,
 k_?IntegerQ] := Which[x =="0", "1" <> BernsteinWords
 Kind2[x, n−1, k] <> "0" <> BernsteinWordsKind2[x, n−1,
 k−1], x == "1", "0" <> BernsteinWordsKind2[x, n−1, k]
 <> "1" <> BernsteinWordsKind2[x, n−1, k−1]];
```

By using Implementation 2 and the auxiliary commands of Wolfram language, we also provide the following code written in Wolfram language:

```
TableForm[Evaluate[Table[BernsteinWordsKind2["0", n, k],
 {n, 0, 4}, {k, 0, 2}]], TableHeadings −>{{"n=0","n=1",
 "n=2", "n=3", "n =4"}, {" k =0"," k=1", " k =2"}}],
```

which returns Table 3, whose entries are the Bernstein words of the second kind $\mathcal{W}_B(x;n,k)$ in the case when $x=0$, $n \in \{0,1,2,3,4\}$ and $k \in \{0,1,2\}$.

In addition, by the following code written in Wolfram language,

```
TableForm[Evaluate[Table[BernsteinWordsKind2["1", n, k],
 {n, 0, 4}, {k, 0, 2}]], TableHeadings −>{{"n=0", "n=1",
 "n=2", "n=3", "n =4"}, {" k=0", " k=1", " k =2"}}]
```

we get Table 4, whose entries are the Bernstein words of the second kind $\mathcal{W}_B(x;n,k)$ in the case when $x=1$, $n \in \{0,1,2,3,4\}$ and $k \in \{0,1,2\}$.

In Table 5, we present decimal equivalent of the Bernstein words of the first kind $w_B(x;n,k)$ for the cases when $x=0$, $n \in \{0,1,\dots,15\}$ and $k=1$.

In Table 6, we present decimal equivalent of the Bernstein words of the first kind $w_B(x;n,k)$ for the cases when $x=1$, $n \in \{0,1,\dots,15\}$ and $k=1$.

In Table 7, we present decimal equivalent of the Bernstein words of the second kind $\mathcal{W}_B(x;n,k)$ for the cases when $x=0$, $n \in \{0,1,\dots,15\}$ and $k=1$.

In Table 8, we present decimal equivalent of the Bernstein words of the second kind $\mathcal{W}_B(x;n,k)$ for the cases when $x=1$, $n \in \{0,1,\dots,15\}$ and $k=1$.

3.3 | Patterns Arising From the Bernstein-Based Words

Here, by representing each successive letter of the Bernstein-based words as a square block with 1s colored black and 0s colored white, then by placing the corresponding square blocks side-by-side to be a row of colored squares, we present some patterns of the Bernstein-based words (see Figure 4).

TABLE 1 | The Bernstein words of the first kind $w_B(x; n, k)$ in the case when $x = 0, n \in \{0, 1, 2, 3, 4, 5\}$ and $k \in \{0, 1, 2\}$.

	k=0	k=1	k=2
n=0	ϵ	ϵ	ϵ
n=1	1	0	ϵ
n=2	11	0101	00
n=3	111	011011011	001001001
n=4	1111	0111011101110111	00110011001100110011
n=5	11111	0111101111011110111101111	00111001110011100111001110011100111001110011100111

TABLE 2 | The Bernstein words of the first kind $w_B(x; n, k)$ in the case when $x = 1, n \in \{0, 1, 2, 3, 4, 5\}$ and $k \in \{0, 1, 2\}$.

	k=0	k=1	k=2
n=0	ϵ	ϵ	ϵ
n=1	0	1	ϵ
n=2	00	1010	11
n=3	000	100100100	110110110
n=4	0000	1000100010001000	11001100110011001100
n=5	00000	1000010000100001000010000	11000110001100011000110001100011000110001100011000

TABLE 3 | The Bernstein words of the second kind $\mathcal{W}_B(x; n, k)$ in the case when $x = 0, n \in \{0, 1, 2, 3, 4\}$ and $k \in \{0, 1, 2\}$.

	k=0	k=1	k=2
n=0	1	0	0
n=1	1100	1001	0
n=2	1110000	1100101100	1001001
n=3	1111000000	1110010110001110000	1100100101100101100
n=4	1111100000000	1111001011000111000001111000000	1110010010110010110001110010110001110000

FIGURE 4 | The row of square blocks corresponding to the Bernstein word of the second kind $\mathcal{W}_B(1; 3, 2) = 0001001110001110110$.

By stacking up the row of square block representation of the first few Bernstein-based words, we obtain some patterns which are given in Figures 5–8.

Remark 4. It is well-known that the logical complement $\neg w$ (namely, so-called ones' complement or the Boolean complement in Boolean algebra) of a binary word w is obtained by changing each 0 in w to 1 and vice versa. Observe that Figures 5 and 6 are logical complement of each other since we draw them by representing zeros in the words with the white square blocks and ones in the words with the black square blocks.

Remark 5. Observe that Figures 7 and 8 are not logical complement of each other as opposed to the figures arising from the Bernstein words of the first kind.

Moreover, after associating 4-ary representations of the Bernstein-based words by the following morphism mapping letters 0, 1, 2, and 3, respectively, to red, green, blue, and yellow colored square blocks:

$0 \mapsto$ Red colored square block,

$1 \mapsto$ Green colored square block,

$2 \mapsto$ Blue colored square block,

$3 \mapsto$ Yellow colored square block,

we get a row of square blocks for the first few Bernstein-based words and then by stacking up these rows, we also obtain some patterns which are given in Figures 9–12.

Remark 6. The DNA (*deoxyribonucleic acid*) is a nucleic acid that contains the genetic instructions and information used in the development and functioning of all known living organisms. The DNA is a strand composed of four nucleotides or bases called adenine, cytosine, guanine and thymine, abbreviated by A, C, G and T, respectively (*cf*. [23]). In this context, we now consider 4-ary representations as well as their patterns of the obtained binary words to give an idea of combinatorial modeling on DNA sequencing. DNA sequences to be produced by this modeling may find application in not only DNA sequencing but also pharmaceutical technologies, biotechnology, microbiology, and so on. In fact, some of the DNA sequences to be detected here may correspond to the DNA sequence of a living species that exists in the literature or has not yet been discovered. This model is only suggested as an idea, and the species with which the model will overlap have not been investigated within the scope of this study.

TABLE 4 | The Bernstein words of the second kind $\mathcal{W}_B(x; n, k)$ in the case when $x = 1, n \in \{0, 1, 2, 3, 4\}$ and $k \in \{0, 1, 2\}$.

	k=0	k=1	k=2
n=0	1	0	0
n=1	0110	0011	0
n=2	0011010	0001110110	0010011
n=3	0001101010	0000111011010011010	0001001110001110110
n=4	0000110101010	0000011101101001101010001101010	0000100111000111011010000111011010011010

TABLE 5 | Numbers obtained by converting the Bernstein words of the first kind $w_B(x; n, k)$ to decimal in the cases when $x = 0, n \in \{0, 1, ..., 15\}$ and $k = 1$.

	k=1
n=0	0
n=1	0
n=2	5
n=3	219
n=4	30 583
n=5	16 236 015
n=6	33 814 345 695
n=7	279 258 638 311 359
n=8	9 187 201 950 435 737 471
n=9	1 206 560 015 662 350 056 947 455
n=10	633 205 725 040 689 368 685 058 981 375
n=11	1 328 578 641 610 130 862 706 980 579 058 908 159
n=12	11 147 649 675 553 647 270 017 976 875 240 829 304 698 879
n=13	374 098 741 654 677 608 890 559 610 263 248 398 282 433 696 362 495
n=14	50 213 748 704 928 086 076 131 552 136 232 920 089 648 434 055 403 681 079 295
n=15	26 959 123 889 762 805 978 944 041 759 736 479 343 619 943 057 007 489 178 619 980 267 519

TABLE 6 | Numbers obtained by converting the Bernstein words of the first kind $w_B(x; n, k)$ to decimal in the cases when $x = 1, n \in \{0, 1, ..., 15\}$ and $k = 1$.

	k=1
n=0	0
n=1	1
n=2	10
n=3	292
n=4	34 952
n=5	17 318 416
n=6	34 905 131 040
n=7	283 691 315 109 952
n=8	9 259 542 123 273 814 144
n=9	1 211 291 623 566 908 292 464 896
n=10	634 444 875 187 540 032 811 644 224 000
n=11	1 329 877 349 959 700 883 100 633 541 501 780 992
n=12	11 153 095 522 976 975 871 517 741 397 407 532 201 281 536
n=13	374 190 096 658 744 685 229 727 024 087 488 507 781 403 765 641 216
n=14	50 219 879 061 258 806 145 241 078 635 089 742 567 989 253 056 020 871 127 040
n=15	26 960 769 444 538 473 610 389 988 414 302 782 003 654 345 788 073 655 783 587 240 230 912

It is time to give the modeling idea mentioned above: Associate 4-ary representations of the Bernstein-based words by the following morphism mapping letters 0, 1, 2, and 3, respectively, to A, C, G and T:

$$0 \mapsto \text{A}, \quad 1 \mapsto \text{C}, \quad 2 \mapsto \text{G}, \quad 3 \mapsto \text{T};$$

it is also possible to determine of which cell gives the nucleotide base (nucleobase) sequence in the DNA molecule and which biological information this sequence encodes, this type of studies also reveals an area of potential application of the Bernstein-based words. For nucleotide base (nucleobase) sequences corresponding to the Bernstein-based words, see Tables 9–12.

We also note that the above observation makes an allusion to DNA but transforming a sequence of numbers into a sequence of ACGT makes sense only if one obtains existing significant sequences in biology. In this context, in this paper, we may also raise the following open question.

TABLE 7 | Numbers obtained by converting the Bernstein words of the second kind $\mathcal{W}_B(x;n,k)$ to decimal in the cases when $x=0, n \in \{0,1,...,15\}$ and $k=1$.

	k=1
n=0	0
n=1	9
n=2	812
n=3	470 128
n=4	2 036 564 928
n=5	68 551 451 877 120
n=6	18 208 547 937 292 712 960
n=7	38 435 859 475 728 710 580 563 968
n=8	646 941 911 943 400 394 188 959 571 230 720
n=9	86 971 679 750 389 756 074 485 227 918 487 065 657 344
n=10	93 460 617 420 352 574 081 684 338 890 047 069 228 652 262 326 272
n=11	803 144 806 349 129 759 355 741 991 213 423 752 868 260 161 293 451 476 860 928
n=12	55 202 830 804 936 378 945 685 118 505 712 696 807 874 544 602 082 019 437 294 806 072 557 568
n=13	30 351 139 309 558 186 954 230 650 981 997 648 463 698 028 347 652 318 857 320 105 397 792 195 154 374 819 840
n=14	133 492 456 046 745 365 861 711 369 659 735 238 250 384 205 909 743 384 530 671 069 002 481 127 412 194 852 729 127 289 487 360
n=15	4 696 966 705 074 203 326 538 999 460 271 928 244 237 556 412 863 889 593 327 541 289 556 977 680 853 840 716 057 160 471 539 176 957 687 103 488

TABLE 8 | Numbers obtained by converting the Bernstein words of the second kind $\mathcal{W}_B(x;n,k)$ to decimal in the cases when $x=1, n \in \{0,1,...,15\}$ and $k=1$.

	k=1
n=0	0
n=1	3
n=2	118
n=3	30 362
n=4	62 182 506
n=5	1 018 798 186 922
n=6	133 535 915 956 307 626
n=7	140 022 556 609 801 225 771 690
n=8	1 174 594 338 557 431 440 918 209 129 130
n=9	78 825 691 721 420 622 757 904 131 570 377 271 978
n=10	42 319 221 003 509 939 675 643 946 324 756 438 191 169 776 298
n=11	181 759 670 202 271 492 117 823 597 205 208 297 196 519 585 185 695 181 482
n=12	6 245 214 714 004 014 108 473 393 029 547 573 098 615 200 827 168 375 131 936 985 361 066
n=13	1 716 671 549 001 294 963 751 412 451 075 916 040 622 186 908 234 260 446 536 258 384 551 127 655 557 802
n=14	3 775 000 658 398 322 345 452 452 839 003 268 816 749 619 334 587 136 883 005 228 154 804 677 079 317 460 160 176 892 586
n=15	66 410 513 900 336 178 032 575 689 795 589 704 929 428 406 528 004 237 618 310 484 467 320 886 573 094 049 340 385 997 758 146 397 514 410

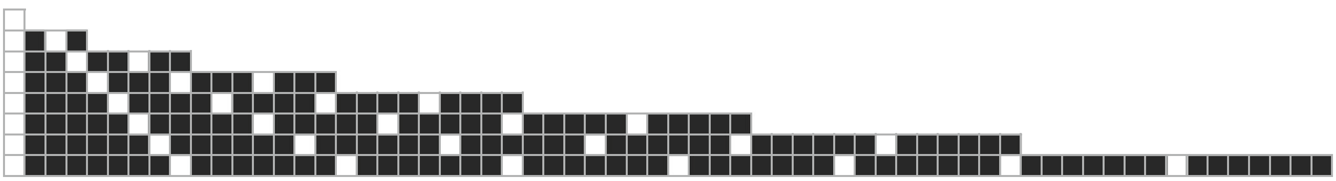

FIGURE 5 | Pattern obtained by the Bernstein words of the first kind $w_B(x;n,k)$ in the case when $x=0,\ n \in \{1,2,\ \dots\ ,8\}$ and $k=1$.

Open Question 2: *Is there any significant sequence in biology which matches the nucleotide base (nucleobase) sequence corresponding to the Bernstein-based words? Moreover, what kind of applications of a DNA sequence produced with the nucleotide base (nucleobase) sequence corresponding to the Bernstein-based words might have in pharmaceutical technologies or biotechnology?*

4 | Relations Arising From Finite Sums and Generating Functions for the Lengths of the Bernstein-Based Words

In this section, we give some finite sums and generating functions for the lengths of the Bernstein-based words. Moreover, we give some relations and results derived from the length of the Bernstein words of the first and second kinds.

4.1 | Generating Functions for the Lengths of the Bernstein Words of the First Kind

Here, we give some formulas, finite sums, and generating functions for the lengths $|w_B(x;n,k)|$ of the Bernstein words of the first kind.

The definition, given by (13), means that we first juxtapose k-times 0's or 1's with $(n-k)$-times 0's or 1's. Then, the words obtained from the first process is brought side by side $\binom{n}{k}$ times

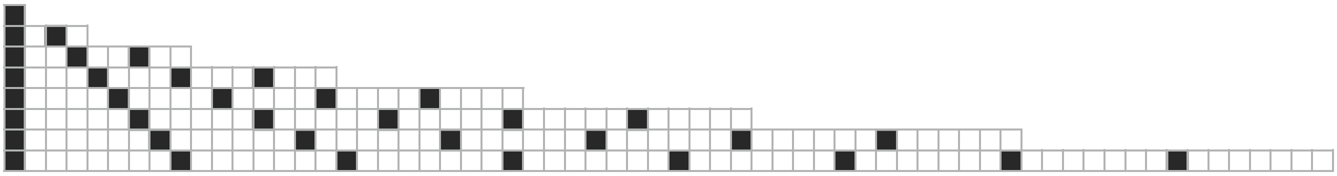

FIGURE 6 | Pattern obtained by the Bernstein words of the first kind $w_B(x;n,k)$ in the case when $x=1$, $n \in \{1,2,\ldots,8\}$ and $k=1$.

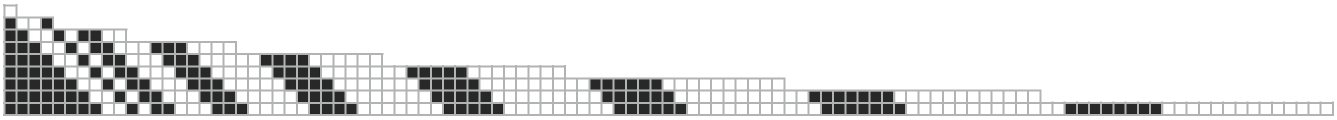

FIGURE 7 | Pattern obtained by the Bernstein words of the second kind $\mathcal{W}_B(x;n,k)$ in the case when $x=0$, $n \in \{0,1,\ldots,8\}$ and $k=1$.

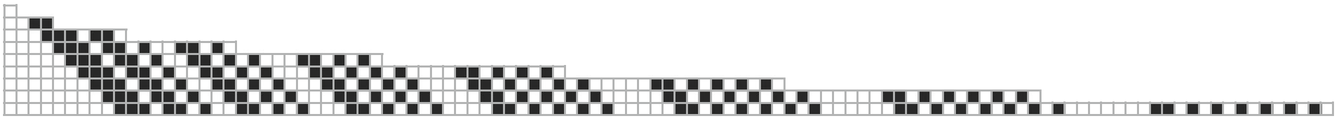

FIGURE 8 | Pattern obtained by the Bernstein words of the second kind $\mathcal{W}_B(x;n,k)$ in the case when $x=1$, $n \in \{0,1,\ldots,8\}$ and $k=1$.

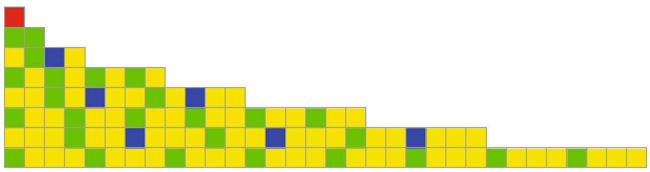

FIGURE 9 | Pattern obtained by 4-ary representations of the Bernstein words of the first kind $w_B(x;n,k)$ in the case when $x=0$, $n \in \{1,\ldots,8\}$ and $k=1$. [Colour figure can be viewed at wileyonlinelibrary.com]

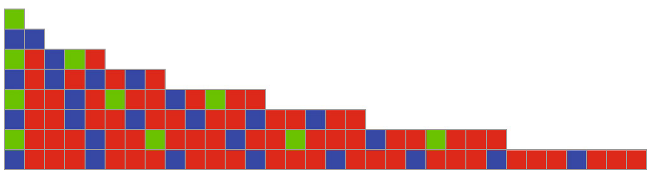

FIGURE 10 | Pattern obtained by 4-ary representations of the Bernstein words of the second kind $w_B(x;n,k)$ in the case when $x=1$, $n \in \{1,\ldots,8\}$ and $k=1$. [Colour figure can be viewed at wileyonlinelibrary.com]

to obtain the word $w_B(x;n,k)$. Therefore, the length of the word $w_B(x;n,k)$ is equal to the product of $(k+n-k)$ and $\binom{n}{k}$ which yields the assertion of the following theorem:

Theorem 1. *Let* $x \in \Sigma = \{0,1\}$ *and* $n,k \in \mathbb{N}_0$ *. Then, the length of the Bernstein words of the first kind* $w_B(x;n,k)$ *is given by*

$$|w_B(x;n,k)| = n\binom{n}{k}. \tag{15}$$

Using (15), we get Tables 13 and 14 involving the lengths of the Bernstein words of the first kind; $w_B(x;n,k)$ are provided as tables for the cases of $x \in \Sigma = \{0,1\}$, $n \in \{0,1,2,\ldots,15\}$ and $k \in \{0,1,2,\ldots,10\}$.

TABLE 9 | The nucleotide base (nucleobase) sequence corresponding to the Bernstein words of the first kind $w_B(x;n,k)$ in the case when $x=0, n \in \{0,1,\ldots,8\}$ and $k=1$.

	k=1
n=1	A
n=2	CC
n=3	TCGT
n=4	CTCTCTCT
n=5	TTCTGTTCTGTT
n=6	CTTCTTCTTCTTCTTCTT
n=7	TTTCTTGTTTCTTGTTTCTTGTTT
n=8	CTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT

TABLE 10 | The nucleotide base (nucleobase) sequence corresponding to the Bernstein words of the first kind $w_B(x;n,k)$ in the case when $x=1, n \in \{0,1,\ldots,8\}$ and $k=1$.

	k=1
n=1	C
n=2	GG
n=3	CAGCA
n=4	GAGAGAGA
n=5	CAAGACAAGACAA
n=6	GAAGAAGAAGAAGAAGAA
n=7	CAAAGAACAAAGAACAAAGAACAAA
n=8	GAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAA

In Table 14, the second and third columns, respectively, show the first terms of the sequences $\left\{|w_B(x;n,k)|\right\}_{n=k}^{\infty}$ for $k \in \{0,\ldots,5\}$ and the symbolic notations of the corresponding sequences. As for the last column, it provides the IDs of the corresponding

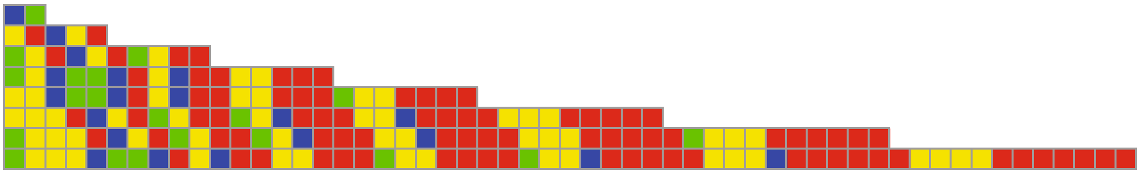

FIGURE 11 | Pattern obtained by 4-ary representations of the Bernstein words of the second kind $\mathcal{W}_B(x;n,k)$ in the case when $x = 0,\ n \in \{1, \ldots, 8\}$ and $k = 1$. [Colour figure can be viewed at wileyonlinelibrary.com]

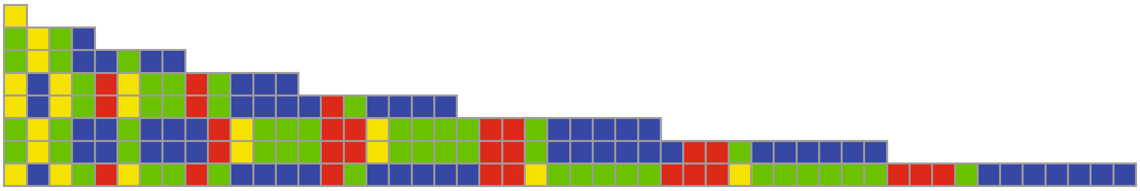

FIGURE 12 | Pattern obtained by 4-ary representations of the Bernstein words of the second kind $\mathcal{W}_B(x;n,k)$ in the case when $x = 1,\ n \in \{1, \ldots, 8\}$ and $k = 1$. [Colour figure can be viewed at wileyonlinelibrary.com]

TABLE 11 | The nucleotide base (nucleobase) sequence corresponding to the Bernstein words of the second kind $\mathcal{W}_B(x;n,k)$ in the case when $x = 0, n \in \{0, 1, \ldots, 8\}$ and $k = 1$.

	k=1
n=1	GC
n=2	TAGTA
n=3	CTAGTACTAA
n=4	CTGCCGATGAATTAAA
n=5	TTGCCGATGAATTAAACTTAAAA
n=6	TTTAGTACTAACTGAAATTGAAAATTTAAAAA
n=7	CTTTAGTACTAACTGAAATTGAAAATTTAAAAACTTTAAAAAA
n=8	CTTTGCCGATGAATTAAACTTAAAACTTGAAAAATTTGAAAAAATTTTAAAAAAA

sequences in Sloane's *On-Line Encyclopedia of Integer Sequences*(OEIS).

Some other applications of (15) are given as follows.

xSubstituting $n = 2m$ and $k = m$ into (15), we get

$$|w_B(x;2m,m)| = 2m\binom{2m}{m}. \tag{16}$$

Combining (6) with (16) gives a relation, between the length of the words $w_B(x;2m,m)$ and the Catalan numbers C_m, given the following theorem.

Theorem 2. *Let $x \in \Sigma = \{0,1\}$ and $m \in \mathbb{N}_0$. Then, we have*

$$|w_B(x;2m,m)| = 2m(m+1)C_m \tag{17}$$

or equivalently, $m \geq 3$ for $m \geq 2$

$$|w_B(x;2m,m)| = 2m(m+1)\prod_{k=2}^{m}\frac{m+k}{k}. \tag{18}$$

The combination of (17) with (7) also yields the following corollary.

Corollary 1. *Let $x \in \Sigma = \{0,1\}$ and $0 < |t| \leq \frac{1}{4}$. Then we have*

$$\sum_{m=1}^{\infty}\frac{|w_B(x;2m,m)|}{m(m+1)}t^m = \frac{4}{1+\sqrt{1-4t}}. \tag{19}$$

Summing Equation (15) over all $0 \leq k \leq n$, we get

$$\sum_{k=0}^{n}|w_B(x;n,k)| = \sum_{k=0}^{n} n\binom{n}{k} \tag{20}$$

by which and by the well-known formula of the sum of the binomial coefficients, we have the following:

$$\sum_{k=0}^{n}|w_B(x;n,k)| = 2\sum_{j=0}^{n} j\binom{n}{j}. \tag{21}$$

Combining the above equation with Equation (1) of [18, p. 1329], we deduce to the following corollary.

Corollary 2. *Let $x \in \Sigma = \{0,1\}$ and $n \in \mathbb{N}_0$. Then, we have*

$$\sum_{k=0}^{n}|w_B(x;n,k)| = n2^n. \tag{22}$$

TABLE 12 | The nucleotide base (nucleobase) sequence corresponding to the Bernstein words of the second kind $\mathcal{W}_B(x; n, k)$ in the case when $x = 1, n \in \{0, 1, ..., 8\}$ and $k = 1$.

	k=1
n=1	T
n=2	CTCG
n=3	CTCGGCGG
n=4	TGTCATCCACGGG
n=5	TGTCATCCACGGGGACGGGG
n=6	CTCGGCGGGATCCCAATCCCCAACGGGGG
n=7	CTCGGCGGGATCCCAATCCCCAACGGGGGGAACGGGGGG
n=8	TGTCATCCACGGGGACGGGGGAATCCCCCAAATCCCCCCAAACGGGGGGG

TABLE 13 | For $x \in \Sigma = \{0, 1\}, n \in \{0, 1, 2, ..., 15\}$ and $k \in \{0, 1, 2, ..., 10\}$, the lengths of the words $w_B(x; n, k)$, that is, $|w_B(x; n, k)|$.

	k=0	k=1	k=2	k=3	k=4	k=5	k=6	k=7	k=8	k=9	k=10
n=0	0	0	0	0	0	0	0	0	0	0	0
n=1	1	1	0	0	0	0	0	0	0	0	0
n=2	2	4	2	0	0	0	0	0	0	0	0
n=3	3	9	9	3	0	0	0	0	0	0	0
n=4	4	16	24	16	4	0	0	0	0	0	0
n=5	5	25	50	50	25	5	0	0	0	0	0
n=6	6	36	90	120	90	36	6	0	0	0	0
n=7	7	49	147	245	245	147	49	7	0	0	0
n=8	8	64	224	448	560	448	224	64	8	0	0
n=9	9	81	324	756	1134	1134	756	324	81	9	0
n=10	10	100	450	1200	2100	2520	2100	1200	450	100	10
n=11	11	121	605	1815	3630	5082	5082	3630	1815	605	121
n=12	12	144	792	2640	5940	9504	11 088	9504	5940	2640	792
n=13	13	169	1014	3718	9295	16 731	22 308	22 308	16 731	9295	3718
n=14	14	196	1274	5096	14 014	28 028	42 042	48 048	42 042	28 028	14 014
n=15	15	225	1575	6825	20 475	45 045	75 075	96 525	96 525	75 075	45 045

TABLE 14 | Table of the lengths of the words $w_B(x; n, k)$, i.e. $|w_B(x; n, k)|$.

k	$\{\lvert w_B(x; n, k)\rvert\}_{n=k}^{\infty}$	Corresponding sequence	Also, see OEIS
$k = 0$	$\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, \ldots\}$	$\{n\}_{n=0}^{\infty}$	A001477
$k = 1$	$\{1, 4, 9, 16, 25, 36, 49, 64, 81, 100, \ldots\}$	$\{n^2\}_{n=1}^{\infty}$	A000290
$k = 2$	$\{2, 9, 24, 50, 90, 147, 224, 324, 450, \ldots\}$	$\left\{\frac{(n-1)n^2}{2}\right\}_{n=2}^{\infty}$	A006002
$k = 3$	$\{3, 16, 50, 120, 245, 448, 756, 1200, \ldots\}$	$\left\{\frac{(n-2)(n-1)n^2}{6}\right\}_{n=3}^{\infty}$	A004320
$k = 4$	$\{4, 25, 90, 245, 560, 1134, 2100, \ldots\}$	$\left\{n\binom{n}{4}\right\}_{n=4}^{\infty}$	A027764
$k = 5$	$\{5, 36, 147, 448, 1134, 2520, 5082, \ldots\}$	$\left\{n\binom{n}{5}\right\}_{n=5}^{\infty}$	A027765

Combining (22) and (11), we obtain a relation, between the numbers $y_6(m, n; \lambda, p)$ and the finite sums of the lengths $|w_B(x; n, k)|$, as in the following corollary:

Corollary 3. *Let $x \in \Sigma = \{0, 1\}$ and $n \in \mathbb{N}_0$. Then, we have*

$$\sum_{k=0}^{n} |w_B(x; n, k)| = nn! y_6(0, n; 1, 1). \tag{23}$$

Using (15), we get the ordinary generating functions for the lengths $|w_B(x; n, k)|$, given in the following theorem.

Theorem 3. *Let $x \in \Sigma = \{0, 1\}$ and $n \in \mathbb{N}_0$. Then we have*

$$\sum_{k=0}^{\infty} |w_B(x; n, k)| t^k = (1 + t)\frac{d}{dt}\{(1 + t)^n\}. \tag{24}$$

By combining (15) with the following well-known formula of the Laguerre polynomials $L_n(t)$:

$$L_n(t) = \sum_{k=0}^{n} (-1)^k \binom{n}{k} \frac{t^k}{k!},$$

and after some elementary calculations, we get the exponential generating function for the lengths $|w_B(x; n, k)|$, given the following theorem.

Theorem 4. *Let $x \in \Sigma = \{0, 1\}$ and $n \in \mathbb{N}_0$. Then, we have*

$$nL_n(-t) = \sum_{k=0}^{\infty} |w_B(x; n, k)| \frac{t^k}{k!} \tag{25}$$

Remark 7. By combining (8) with (25), we also write the exponential generating function for the lengths $|w_B(x; n, k)|$ in terms of the hypergeometric series as follows:

$$ {}_1F_1(-n;1;-t) = \frac{1}{n}\sum_{k=0}^{\infty}|w_B(x;n,k)|\frac{t^k}{k!}. $$

Summing the reciprocals of Equation (15) over all $0 \le k \le n$, we get

$$ \sum_{k=0}^{n}\frac{1}{|w_B(x;n,k)|} = \sum_{k=0}^{n}\frac{1}{n\binom{n}{k}}. \tag{26} $$

Since the following well-known equality holds true (*cf*. [7]; and see also the references cited therein)

$$ \sum_{k=0}^{n}\frac{1}{\binom{n}{k}} = \frac{n+1}{2^{n+1}}\sum_{k=0}^{n+1}\frac{2^k}{k}. \tag{27} $$

combining (26) with the above equation, we arrive at the following theorem.

Theorem 5. *Let $x \in \Sigma = \{0,1\}$ and $n \in \mathbb{N}_0$. Then, we have*

$$ \sum_{k=0}^{n}\frac{1}{|w_B(x;n,k)|} = \frac{n+1}{n2^{n+1}}\sum_{k=0}^{n+1}\frac{2^k}{k}. \tag{28} $$

By combining (8) and (10) with (15), we get the following theorem, which gives the ordinary generating functions for the reciprocal of the lengths $|w_B(x;n,k)|$:

Theorem 6. *Let $x \in \Sigma = \{0,1\}$ and $n \in \mathbb{N}_0$. Then we have*

$$ \frac{{}_2F_1(1,1;-n;-t)}{n} = \sum_{k=0}^{\infty}\frac{t^k}{|w_B(x;n,k)|}. \tag{29} $$

By combining (8) and (10) with (15), we get the following theorem, which gives the exponential generating functions for the reciprocal of the lengths $|w_B(x;n,k)|$:

Theorem 7. *Let $x \in \Sigma = \{0,1\}$ and $n \in \mathbb{N}_0$. Then we have*

$$ \frac{{}_1F_1(1;-n;-t)}{n} = \sum_{k=0}^{\infty}\frac{1}{|w_B(x;n,k)|}\frac{t^k}{k!}. \tag{30} $$

Remark 8. By using linear differential equations, the generating function Equations (29) and (30) may be represented by another special functions.

4.2 | Generating Functions for the Lengths of the Bernstein Words of the Second Kind

Here, we provide some tables involving the lengths of the Bernstein words of the second kind. We also give some observations and open questions regarding generating functions for these lengths.

By Tables 15 and 16, the lengths of the Bernstein words of the second kind $\mathcal{W}_B(x;n,k)$ are provided as tables for the cases of $x \in \Sigma = \{0,1\}$, $n \in \{0,1,2,\dots,15\}$ and $k \in \{0,1,2,\dots,10\}$.

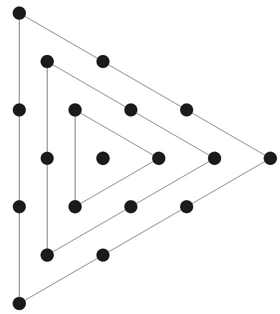

FIGURE 13 | The geometric origin of the centered triangular numbers (*cf.* [24, p. 48]).

By comparing Table 15 with the trees in Figures 1 and 2, we come up with some novel sequences of words with their lengths derived from the tree diagrams. Some of these were given in the rows of Table 16. For example, the sequence in the first row of Table 16 is obtained from the lengths of the words encountered on the nodes while traveling the first left branches of the trees in Figures 1 and 2.

In Table 16, the second and third columns, respectively, show the first terms of the sequences $\left\{|\mathcal{W}_B(x;n,k)|\right\}_{n=k}^{\infty}$ for $k \in \{0,\dots,5\}$ and the symbolic notations of the corresponding sequences (if exist). As for the last column, it provides the IDs (if exist) of the corresponding sequences in Sloane's *On-Line Encyclopedia of Integer Sequences*(OEIS).

Remark 9. Observe from the second column of Table 15 that the length of the words $\mathcal{W}_B(x;n,1)$ gives the following sequence, for $n \in \mathbb{N}_0$:

$$ \{1,4,10,19,31,46,64,85,109,136,166, $$
$$ 199,235,274,316,361,\dots\}, $$

which is overlapping with the $(n+1)$*-th centered 3-gonal numbers* or so-called *centered triangular numbers* that originate from a centered polygonal number consisting of a central dot with three dots around it and then additional dots in the gaps between adjacent dots; see Figure 13 (*cf*. [24, 25, OEIS: A005448]).

The n-th centered m-gonal number $CS_m(n)$ is given by the following explicit formulas:

$$ \begin{aligned} CS_m(n) &= 1 + m\binom{n}{2} \\ &= \frac{mn^2 - mn + 2}{2} \end{aligned} \tag{31} $$

which have the following generating function:

$$ \frac{t\left(1+(m-2)t+t^2\right)}{(1-t)^3} = \sum_{n=1}^{\infty} CS_m(n)t^n, $$

where $|t| < 1$ (*cf*. [24, p. 51]; see also [25, sequence A005448 in the OEIS]).

Thus, from Remark 9, we deduce that we have the following relation, for $n \in \mathbb{N}_0$:

$$ CS_3(n+1) = |\mathcal{W}_B(x;n,1)|, \tag{32} $$

TABLE 15 | For $x \in \Sigma = \{0,1\}$, $n \in \{0,1,2,\ldots,15\}$ and $k \in \{0,1,2,\ldots,10\}$, the lengths of the words $\mathcal{W}_B(x;n,k)$, that is, $|\mathcal{W}_B(x;n,k)|$.

	k=0	k=1	k=2	k=3	k=4	k=5	k=6	k=7	k=8	k=9	k=10
n=0	1	1	1	1	1	1	1	1	1	1	1
n=1	4	4	1	1	1	1	1	1	1	1	1
n=2	7	10	7	1	1	1	1	1	1	1	1
n=3	10	19	19	10	1	1	1	1	1	1	1
n=4	13	31	40	31	13	1	1	1	1	1	1
n=5	16	46	73	73	46	16	1	1	1	1	1
n=6	19	64	121	148	121	64	19	1	1	1	1
n=7	22	85	187	271	271	187	85	22	1	1	1
n=8	25	109	274	460	544	460	274	109	25	1	1
n=9	28	136	385	736	1006	1006	736	385	136	28	1
n=10	31	166	523	1123	1744	2014	1744	1123	523	166	31
n=11	34	199	691	1648	2869	3760	3760	2869	1648	691	199
n=12	37	235	892	2341	4519	6631	7522	6631	4519	2341	892
n=13	40	274	1129	3235	6862	11 152	14 155	14 155	11 152	6862	3235
n=14	43	316	1405	4366	10 099	18 016	25 309	28 312	25 309	18 016	10 099
n=15	46	361	1723	5773	14 467	28 117	43 327	53 623	53 623	43 327	28 117

TABLE 16 | Table of the lengths of the words $\mathcal{W}_B(x;n,k)$, i.e. $|\mathcal{W}_B(x;n,k)|$.

k	$\{\lvert\mathcal{W}_B(x;n,k)\rvert\}_{n=k}^{\infty}$	Corresponding sequence	Also, see OEIS
$k=0$	$\{1,4,7,10,13,16,19,22,25,28,31,\ldots\}$	$\{3n+1\}_{n=0}^{\infty}$	A016777
$k=1$	$\{4,10,19,31,46,64,85,109,136,166,\ldots\}$	$\left\{\frac{3n(n-1)}{2}+1\right\}_{n=0}^{\infty}$	A005448
$k=2$	$\{7,19,40,73,121,187,274,385,523,\ldots\}$	New Sequence	Does Not Exist
$k=3$	$\{10,31,73,148,271,460,736,1123,\ldots\}$	New Sequence	Does Not Exist
$k=4$	$\{13,46,121,271,544,1006,1744,\ldots\}$	New Sequence	Does Not Exist
$k=5$	$\{16,64,187,460,1006,2014,3760,\ldots\}$	New Sequence	Does Not Exist

and we thus conclude that the generating function for $|\mathcal{W}_B(x;n,1)|$ is given by the following:

$$\frac{t^2+t+1}{(1-t)^3} = \sum_{n=0}^{\infty} |\mathcal{W}_B(x;n,1)| t^n; \qquad |t|<1 \tag{33}$$

which is also related to the sequence A005448 in the OEIS [25], and [24, p. 51].

At this stage, the following questions come to mind for the case $k \neq 1$: Namely, we have the following open questions:

Open Question 3: *What is the explicit formula for the lengths* $|\mathcal{W}_B(x;n,k)|$*?*

Open Question 4: *How can we construct ordinary or exponential function for the lengths* $|\mathcal{W}_B(x;n,k)|$*?. That is, are there explicit formulas for the following generating functions, respectively:*

$$\mathcal{G}_1(t;k) = \sum_{n=0}^{\infty} |\mathcal{W}_B(x;n,k)| t^n = ?, \tag{34}$$

$$\mathcal{G}_2(t;n) = \sum_{k=0}^{\infty} |\mathcal{W}_B(x;n,k)| t^k = ?, \tag{35}$$

$$\mathcal{G}_3(t;k) = \sum_{n=0}^{\infty} |\mathcal{W}_B(x;n,k)| \frac{t^n}{n!} = ?, \tag{36}$$

and

$$\mathcal{G}_4(t;n) = \sum_{k=0}^{\infty} |\mathcal{W}_B(x;n,k)| \frac{t^k}{k!} = ? \tag{37}$$

by keeping in mind that the lengths $|\mathcal{W}_B(x;n,k)|$ *are the same for both selections of* $x=0$ *and* $x=1$ *with the fixed values of* n *and* k.

Open Question 5: The above question also brings to mind another question whether the formal power series $\mathcal{G}_j(t;k)$; $(j=1,3)$ and $\mathcal{G}_j(t;n)$; $(j=2,4)$ are DD-finite, D^n-finite, D-algebraic, D^{∞}-finite, or not. For a sample study that addresses these concepts, see [26].

5 | Slopes of the Bernstein-Based Words and Their Relations With Farey Fractions

In this section, we give relations between slopes of the Bernstein-based words and the Farey fractions.

It is known from the book of Lothaire [2] that the slope of a nonempty w is defined by

$$\text{slope}(w) = \frac{\text{height}(w)}{|w|} \tag{38}$$

where height(w) denotes the height of the word w which corresponds to the number of letters equal to 1 in the word w (*cf*. [2, pp. 42-45]; see also [27]).

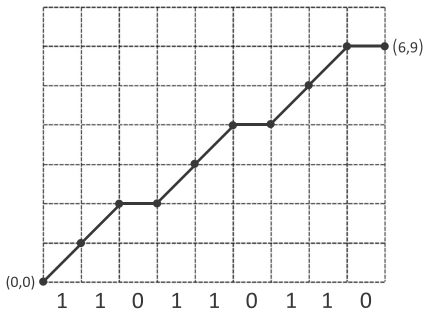

FIGURE 14 | Slope diagram the word $w_B(1;3,2) = 110110110$.

For instance, the height of the word $w_B(1;3,2) = 110110110$ is 6, and the length of the word $w_B(1;3,2)$ is 9. Therefore, the slope of the word $w_B(1;3,2)$ is $\frac{6}{9}$, namely $\frac{1}{3}$.

The word $w_B(1;3,2) = 110110110$ can be drawn on a grid by representing a 0 (resp. a 1) as horizontal (resp. a diagonal) unit segment. This gives a polygonal line from the origin to the point $(|w|, \text{slope}(w))$, and the line from the origin to this point has the slope $\text{slope}(w)$; see Figure 14.

In order to present some relations between the slopes of the Bernstein-based words and the Farey fractions, we briefly recall the definition of the set of consecutive Farey fractions of order n, denoted by F_n, as follows.

F_n is a set of reduced fractions in the closed interval $[0,1]$ with denominators $\leq n$.

The first 10 sets of consecutive Farey fractions are given as follows:

$$F_1 : \left\{\frac{0}{1}, \frac{1}{1}\right\}$$

$$F_2 : \left\{\frac{0}{1}, \frac{1}{2}, \frac{1}{1}\right\}$$

$$F_3 : \left\{\frac{0}{1}, \frac{1}{3}, \frac{1}{2}, \frac{2}{3}, \frac{1}{1}\right\}$$

$$F_4 : \left\{\frac{0}{1}, \frac{1}{4}, \frac{1}{3}, \frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \frac{1}{1}\right\}$$

$$F_5 : \left\{\frac{0}{1}, \frac{1}{5}, \frac{1}{4}, \frac{1}{3}, \frac{2}{5}, \frac{1}{2}, \frac{3}{5}, \frac{2}{3}, \frac{3}{4}, \frac{4}{5}, \frac{1}{1}\right\}$$

$$F_6 : \left\{\frac{0}{1}, \frac{1}{6}, \frac{1}{5}, \frac{1}{4}, \frac{1}{3}, \frac{2}{5}, \frac{1}{2}, \frac{3}{5}, \frac{2}{3}, \frac{3}{4}, \frac{4}{5}, \frac{5}{6}, \frac{1}{1}\right\}$$

$$F_7 : \left\{\frac{0}{1}, \frac{1}{7}, \frac{1}{6}, \frac{1}{5}, \frac{1}{4}, \frac{2}{7}, \frac{1}{3}, \frac{2}{5}, \frac{3}{7}, \frac{1}{2}, \frac{4}{7}, \frac{3}{5}, \frac{2}{3}, \frac{5}{7}, \frac{3}{4}, \frac{4}{5}, \frac{5}{6}, \frac{6}{7}, \frac{1}{1}\right\}$$

$$F_8 : \left\{\frac{0}{1}, \frac{1}{8}, \frac{1}{7}, \frac{1}{6}, \frac{1}{5}, \frac{1}{4}, \frac{2}{7}, \frac{1}{3}, \frac{3}{8}, \frac{2}{5}, \frac{3}{7}, \frac{1}{2}, \frac{4}{7}, \frac{3}{5}, \frac{5}{8}, \frac{2}{3}, \frac{5}{7}, \frac{3}{4}, \frac{4}{5}, \frac{5}{6}, \frac{6}{7}, \frac{7}{8}, \frac{1}{1}\right\}$$

$$F_9 : \left\{\frac{0}{1}, \frac{1}{9}, \frac{1}{8}, \frac{1}{7}, \frac{1}{6}, \frac{1}{5}, \frac{2}{9}, \frac{1}{4}, \frac{2}{7}, \frac{1}{3}, \frac{3}{8}, \frac{2}{5}, \frac{3}{7}, \frac{4}{9}, \frac{1}{2}, \frac{5}{9}, \frac{4}{7}, \frac{3}{5}, \frac{5}{8}, \frac{2}{3}, \frac{5}{7}, \frac{3}{4}, \frac{7}{9}, \frac{4}{5}, \frac{5}{6}, \frac{6}{7}, \frac{7}{8}, \frac{8}{9}, \frac{1}{1}\right\}$$

$$F_{10} : \left\{\frac{0}{1}, \frac{1}{10}, \frac{1}{9}, \frac{1}{8}, \frac{1}{7}, \frac{1}{6}, \frac{1}{5}, \frac{2}{9}, \frac{1}{4}, \frac{2}{7}, \frac{3}{10}, \frac{1}{3}, \frac{3}{8}, \frac{2}{5}, \frac{3}{7}, \frac{4}{9}, \frac{1}{2}, \frac{5}{9}, \frac{4}{7}, \frac{3}{5}, \frac{5}{8}, \frac{2}{3}, \frac{7}{10}, \frac{5}{7}, \frac{3}{4}, \frac{7}{9}, \frac{4}{5}, \frac{5}{6}, \frac{6}{7}, \frac{7}{8}, \frac{8}{9}, \frac{9}{10}, \frac{1}{1}\right\}$$

and so on.

TABLE 17 | Table of the slope of the words $w_B(0;n,k)$, that is, $\text{slope}\left(w_B(0;n,k)\right)$, for the cases when $k \in \{2,3,4,5\}$.

k	$\left\{\text{slope}\left(w_B(0;n,k)\right)\right\}_{n=k}^{\infty}$
$k=2$	$\left\{\frac{0}{1}, \frac{1}{3}, \frac{1}{2}, \frac{3}{5}, \frac{2}{3}, \frac{5}{7}, \frac{3}{4}, \frac{7}{9}, \frac{4}{5}, \frac{9}{11}, \frac{5}{6}, \frac{11}{13}, \frac{6}{7}, \frac{13}{15}, \dots\right\}$
$k=3$	$\left\{\frac{0}{1}, \frac{1}{4}, \frac{2}{5}, \frac{1}{2}, \frac{4}{7}, \frac{5}{8}, \frac{2}{3}, \frac{7}{10}, \frac{8}{11}, \frac{3}{4}, \frac{10}{13}, \frac{11}{14}, \frac{4}{5}, \frac{13}{16}, \dots\right\}$
$k=4$	$\left\{\frac{0}{1}, \frac{1}{5}, \frac{1}{3}, \frac{3}{7}, \frac{1}{2}, \frac{5}{9}, \frac{3}{5}, \frac{7}{11}, \frac{2}{3}, \frac{9}{13}, \frac{5}{7}, \frac{11}{15}, \frac{3}{4}, \frac{13}{17}, \dots\right\}$
$k=5$	$\left\{\frac{0}{1}, \frac{1}{6}, \frac{2}{7}, \frac{3}{8}, \frac{4}{9}, \frac{1}{2}, \frac{6}{11}, \frac{7}{12}, \frac{8}{13}, \frac{9}{14}, \frac{2}{3}, \frac{11}{16}, \frac{12}{17}, \frac{13}{18}, \dots\right\}$

TABLE 18 | Table of the slope of the words $w_B(1;n,k)$, that is, $\text{slope}\left(w_B(1;n,k)\right)$, for the cases when $k \in \{2,3,4,5\}$.

k	$\left\{\text{slope}\left(w_B(1;n,k)\right)\right\}_{n=k}^{\infty}$
$k=2$	$\left\{\frac{1}{1}, \frac{2}{3}, \frac{1}{2}, \frac{2}{5}, \frac{1}{3}, \frac{2}{7}, \frac{1}{4}, \frac{2}{9}, \frac{1}{5}, \frac{2}{11}, \frac{1}{6}, \frac{2}{13}, \frac{1}{7}, \frac{2}{15}, \frac{1}{8}, \dots\right\}$
$k=3$	$\left\{\frac{1}{1}, \frac{3}{4}, \frac{3}{5}, \frac{1}{2}, \frac{3}{7}, \frac{3}{8}, \frac{1}{3}, \frac{3}{10}, \frac{3}{11}, \frac{1}{4}, \frac{3}{13}, \frac{3}{14}, \frac{1}{5}, \frac{3}{16}, \dots\right\}$
$k=4$	$\left\{\frac{1}{1}, \frac{4}{5}, \frac{2}{3}, \frac{4}{7}, \frac{1}{2}, \frac{4}{9}, \frac{2}{5}, \frac{4}{11}, \frac{1}{3}, \frac{4}{13}, \frac{2}{7}, \frac{4}{15}, \frac{1}{4}, \frac{4}{17}, \frac{2}{9}, \dots\right\}$
$k=5$	$\left\{\frac{1}{1}, \frac{5}{6}, \frac{5}{7}, \frac{5}{8}, \frac{5}{9}, \frac{1}{2}, \frac{5}{11}, \frac{5}{12}, \frac{5}{13}, \frac{5}{14}, \frac{1}{3}, \frac{5}{16}, \frac{5}{17}, \frac{5}{18}, \dots\right\}$

Some properties of Farey fractions are given as follows.

It is easy to see that $F_n \subset F_{n+1}$ with $n \in \mathbb{N}$.

Let $\frac{a}{b} < \frac{c}{d}$ be consecutive Farey fractions. Then their mediant $\frac{a+b}{c+d}$ satisfies

$$\frac{a}{b} < \frac{a+b}{c+d} < \frac{c}{d}.$$

If $\frac{a}{b}$ and $\frac{c}{d}$ are consecutive Farey fractions, then the following equality holds true:

$$ad - bc = -1$$

(*cf*. [28, p. 98]).

Fractions that appear as neighbors in a set of consecutive Farey fractions have closely associated with the concept of the continued fraction expansions, and every fraction has two continued fraction expansions. For further properties on the Farey fractions and continued fractions, the interested reader may refer to the book of Apostol [28].

By choosing the penultimate element of each set of consecutive Farey fractions $\{F_1, F_2, \dots, F_{n-1}, F_n\}$, we obtain the following new set of Farey fractions:

$$\mathcal{F}_{0,n} := \left\{\frac{0}{1}, \frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \frac{4}{5}, \frac{5}{6}, \frac{6}{7}, \dots, \frac{n}{n+1}, \frac{n+1}{n+2}, \dots\right\} \tag{39}$$

so that $n \in \mathbb{N}_0$.

On the other hand, by choosing the second element of each set of consecutive Farey fractions $\{F_1, F_2, \dots, F_{n-1}, F_n\}$, we obtain the following another new set of Farey fractions:

$$\mathcal{F}_{1,n} := \left\{\frac{1}{1}, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5}, \frac{1}{6}, \frac{1}{7}, \dots, \frac{1}{n}, \frac{1}{n+1}, \dots\right\} \tag{40}$$

so that $n \in \mathbb{N}_0$.

TABLE 19 | Table of the slopes of the words $\mathcal{W}_B(x;n,k)$, that is, $\text{slope}\left(\mathcal{W}_B(x;n,k)\right)$.

k	$\left\{\text{slope}\left(\mathcal{W}_B(x;n,k)\right)\right\}_{n=0}^{\infty}$	Corresponding sequence
$k=0$	$\left\{\frac{1}{1},\frac{2}{4},\frac{3}{7},\frac{4}{10},\frac{5}{13},\frac{6}{16},\frac{7}{19},\frac{8}{22},\frac{9}{25},\frac{10}{28},\frac{11}{31},\dots\right\}$	$\left\{\frac{n+1}{3n+1}\right\}_{n=0}^{\infty}$
$k=1$	$\left\{\frac{0}{1},\frac{2}{4},\frac{5}{10},\frac{9}{19},\frac{14}{31},\frac{20}{46},\frac{27}{64},\frac{35}{85},\frac{44}{109},\frac{54}{136},\frac{65}{166},\dots\right\}$	$\left\{\frac{n(n+3)}{3n^2+3n+1}\right\}_{n=0}^{\infty}$
$k=2$	$\left\{\frac{0}{1},\frac{0}{1},\frac{3}{7},\frac{9}{19},\frac{19}{40},\frac{34}{73},\frac{55}{121},\frac{83}{187},\frac{119}{274},\frac{164}{385},\frac{219}{523},\dots\right\}$	`New Sequence`

By implementing Equation (38) in the Wolfram Language, we write the following procedure, `CalculateWordSlope`:

```
CalculateWordSlope[w_?StringQ]:=StringCount[w,
“1”]/StringLength[w]
```

for calculating the slope of the word w.

By executing the procedure `CalculateWordSlope` with the input words $w_B(x;n,1)$ for $x \in \Sigma = \{0,1\}$, we obtain the same list respectively given in (39) and (40). Therefore, we arrive at the assertion of the following theorem:

Theorem 8. *Let $n \in \mathbb{N}$. Then, we have*

$$\mathcal{F}_{0,n} = \left\{\text{slope}\left(w_B(0;n,1)\right)\right\}_{n=1}^{\infty}. \tag{41}$$

Theorem 9. *Let $n \in \mathbb{N}$. Then, we have*

$$\mathcal{F}_{1,n} = \left\{\text{slope}\left(w_B(1;n,1)\right)\right\}_{n=1}^{\infty}. \tag{42}$$

Remark 10. Observe from Tables 17 and 18 that each sequence in the tables is a sequence of Farey type fractions. However, in case $k > 1$, it is an open problem how the sequence of the slopes of the words to be chosen and from which set of Farey fractions similar to the above methods.

Remark 11. Let $\frac{a}{b}$ be a rational number whose numerator and denominator are co-primes, that is, $(a,b)=1$. Then, the Ford circle $C(a,b)$ belonging to the fraction $\frac{a}{b}$ is defined as the circle in the complex plane with radius $\frac{1}{2b^2}$ and center at the point $\frac{a}{b}+\frac{i}{2b^2}$ so that $i^2=-1$ (*cf*. [28, p. 99]).

At this stage, another question comes to mind as follows:

If so, what are the relations between the Ford circles and the geometry arising from the sets $\mathcal{F}_{0,n}$ and $\mathcal{F}_{1,n}$, respectively?

Remark 12. Observe that the sets $\mathcal{F}_{0,n}$ and $\mathcal{F}_{1,n}$ form a convergent subsequences derived from the sequence of consecutive Farey fractions although each of $F_1, F_2, \dots, F_{n-1}, F_n, \dots$ is not convergent.

Since every convergent sequence is a Cauchy sequence, we also conclude that each of the sets $\mathcal{F}_{0,n}$ and $\mathcal{F}_{1,n}$ forms a Cauchy sequence.

In Table 19, the second and third columns, respectively, shows the first terms of the sequences $\left\{\text{slope}\left(\mathcal{W}_B(x;n,k)\right)\right\}_{n=0}^{\infty}$ for $k \in \{0,1,2\}$ and the symbolic notations of the corresponding sequences (if exist).

Remark 13. The fact that the Bernstein-based words have quite high potential to be a tool for establishing a relationship with diverse areas is an indication of that they may offer the opportunity to work in very different areas. For example, integrating the old Babylonian algorithm with the Bernstein-based words may potentially offer a new perspective on the study of special words. Indeed, the old Babylonian algorithm is one of the well-known iterative methods used for solving diverse mathematical problems oriented calculation of the square root of a number. When we represent the obtained root via this algorithm as a fraction, the fraction obtained with the old Babylonian algorithm can have a relationship with the Farey fractions resulting from the Bernstein-based words. Therefore, this potential relationship may also offer a new field of study to those working on the relevant algorithm. For further details on old Babylonian algorithm and its some applications, refer to [29], and the references therein.

6 | Conclusion

After recalling the generating functions for some special numbers and polynomials, we introduced a new family of words by calling them as the Bernstein-based words. We also investigated fundamental properties of these words and gave some examples and tables regarding them. In addition, we provided two schematic algorithms for these words. For evaluating the Bernstein-based words, we provided computational implementations in the Wolfram language associated with the newly defined words. For the lengths of the Bernstein-based words, we also constructed some finite sums and generating functions. Using these functions, we derived some relations and results pertaining to the length of the Bernstein-based words. We also derived some relations between the slopes of the Bernstein-based words and the Farey fractions.

As a conclusion, our paper contains lots of new results related to the newly defined concepts called Bernstein-bsed words. Moreover, our paper includes various open questions and remarks within its content. Due to the results and observations in this paper, we gave an idea of combinatorial modeling on DNA sequencing which have the potential to find an application in not only DNA sequencing and also pharmaceutical technologies, biotechnology, microbiology, and so on. We also assumed that DNA sequences some living species can be coded or decoded by our combinatorial model.

The new definitions, results, applications, and also open problems presented in this study can provide research potential from different perspectives for both mathematicians and those working in other applied sciences.

By using the constructed Bernstein-based words and the proposed models, we establish the following future plans:

In the near future, it is planned to investigate new DNA sequencing models and create new tools for biologists and bioinformaticians with their collaboration. When these plans are realized, new models that may emerge may be beneficial for the validation of real DNA sequencing data and exploring their applications in drug design.

To be noted, the algorithmic complexity of the given algorithm has not been investigated in this study. Only the outputs of the algorithms obtained have been focused on. Therefore, additionally to the above future plan, investigating the complexity of the algorithm and obtaining its applications are also among our near future plans.

The other near future plan is to optimize the existing computational implementations in the Wolfram Language and explore an alternative platform or parallel computing frameworks in order to improve the efficiency and scalability of working with the Bernstein-based words.

Author Contributions

Irem Kucukoglu: conceptualization, investigation, writing – original draft, methodology, validation, visualization, writing – review and editing, software, formal analysis. **Yilmaz Simsek:** conceptualization, investigation, writing – original draft, methodology, validation, visualization, writing – review and editing, software, formal analysis.

Acknowledgments

The authors would like to sincerely thank Professor Jeffrey O. Shallit for his valuable comments and suggestions. The authors would also like to thank the referees for their valuable comments and suggestions.

Conflicts of Interest

The authors declare no conflicts of interest.

References

1. M. Lothaire, *Combinatorics on Words* (Cambridge: Cambridge University Press, 1997).

2. M. Lothaire, *Algebraic Combinatorics on Words* (Cambridge, New York: Cambridge University Press, 2002).

3. M. Lothaire, *Applied Combinatorics on Words* (Cambridge: Cambridge University Press, 2005).

4. J. Shallit, *The Logical Approach to Automatic Sequences: Exploring Combinatorics on Words With Walnut, London Math. Soc. Lecture Note Series*, vol. 482 (Cambridge University Press, 2022).

5. I. Kucukoglu, 2018. *Generating Functions Containing Words Defined Over Lexicographical Ordered Finite Alphabet and De Bruijn Type Sequences and Their Applications*, PhD Thesis in Mathematics, Akdeniz University, Institute of Natural and Applied Sciences, Antalya.

6. J. Zhang, Y. Tang, F. Liu, Z. Jin, and Y. Lu, "Solving Fractional Differential Equation Using Block-Pulse Functions and Bernstein Polynomials," *Mathematical Methods in the Applied Sciences* 44, no. 7 (2021): 5501–5519.

7. Y. Simsek, "Combinatorial Sums and Binomial Identities Associated With the Beta-Type Polynomials," *Hacettepe Journal of Mathematics and Statistics* 47, no. 5 (2018): 1144–1155.

8. I. Kucukoglu, B. Simsek, and Y. Simsek, "Multidimensional Bernstein Polynomials and Bezier Curves: Analysis of Machine Learning Algorithm for Facial Expression Recognition Based on Curvature," *Applied Mathematics and Computation* 344–345 (2019): 150–162.

9. G. G. Lorentz, *Bernstein Polynomials* (Chelsea, New York: Chelsea Publishing Company, 1986).

10. Y. Simsek, "Analysis of the Bernstein Basis Functions: An Approach To Combinatorial Sums Involving Binomial Coefficients and Catalan Numbers," *Mathematical Methods in the Applied Sciences* 38, no. 14 (2015): 3007–3021.

11. I. Kucukoglu, "Generating functions for divisor sums and totative sums arising from combinatorial Simsek numbers," *Filomat* 38, no. 5 (2024): 1513–1529.

12. I. Kucukoglu and Y. Simsek, "A Note on Generating Functions for the Unification of the Bernstein Type Basis Functions," *Filomat* 30, no. 4 (2016): 985–992.

13. Y. Simsek, "Functional Equations From Generating Functions: A Novel Approach to Deriving Identities for the Bernstein Basis Functions," *Fixed Point Theory and Applications* 2013 (2013): 1–13.

14. Y. Simsek, "Generating Functions for the Bernstein Type Polynomials: A New Approach to Deriving Identities and Applications for the Polynomials," *Hacettepe Journal of Mathematics and Statistics* 43, no. 1 (2014): 1–14.

15. T. Koshy, *Catalan Numbers With Applications* (Oxford, UK: Oxford University Press, 2009).

16. Y. Simsek, "Generating Functions for Series Involving Higher Powers of Inverse Binomial Coefficients and Their Applications," *Mathematical Methods in the Applied Sciences* 46, no. 12 (2023): 12591–12617.

17. H. M. Srivastava and J. Choi, *Zeta and q-Zeta Functions and Associated Series and Integrals* (Amsterdam: Elsevier Science Publishers, 2012).

18. Y. Simsek, "Generating Functions for Finite Sums Involving Higher Powers of Binomial Coefficients: Analysis of Hypergeometric Functions Including New Families of Polynomials and Numbers," *Journal of Mathematical Analysis and Applications* 477, no. 2 (2019): 1328–1352.

19. I. Kucukoglu, A. Bayad, and Y. Simsek, "k-ary Lyndon Words and Necklaces Arising as Rational Arguments of Hurwitz-Lerch Zeta Function and Apostol-Bernoulli Polynomials," *Mediterranean Journal of Mathematics* 14 (2017): 223.

20. I. Kucukoglu, G. V. Milovanovic, and Y. Simsek, "Analysis of Generating Functions for Special Words and Numbers and Algorithms for Computation," *Mediterranean Journal of Mathematics* 19 (2022): 268.

21. A. E. Frid, S. Puzynina, and L. Q. Zamboni, "On Palindromic Factorization of Words," *Advances in Applied Mathematics* 50 (2013): 737–748.

22. F. Ruskey, J. Sawada, and A. Williams, "Binary Bubble Languages and Cool-Lex Order," *Journal of Combinatorial Theory* 119 (2012): 155–169.

23. P. Ligeti, 2007. *Combinatorics on words and its applications*, Doctoral Thesis, Institute of Mathematics, Etvs Lornd University.

24. E. Deza and M. M. Deza, *Figurate Numbers* (Hackensack, NJ: World Scientific Publishing, 2012).

25. N. J. A. Sloane, *The On-Line Encyclopedia of Integer Sequences* (Sequence: OEIS, 2023) A005448.

26. A. Bostan and A. Jiménez-Pastor, "On the Exponential Generating Function of Labelled Trees," *Comptes Rendus. Mathématique Académie des Sciences* 358, no. 9-10 (2020): 1005–1009.

27. J. Berstel and D. Perrin, "The Origins of Combinatorics on Words," *European Journal of Combinatorics* 28, no. 3 (2007): 996–1022.

28. T. M. Apostol, *Modular Functions and Dirichlet Series in Number Theory*, Second ed. (New York: Springer-Verlag, 1990).

29. J.-H. He, “An old Babylonian Algorithm and Its Modern Applications,” *Symmetry* 16, no. 11 (2024): 1467.