Counting Subwords Occurrences in Base-b Expansions

Julien Leroy, Michel Rigo and Manon Stipulanti
University of Liège
Department of Mathematics
Allée de la Découverte 12 (B37)
4000 Liège, Belgium
J.Leroy@ulg.ac.be
M.Rigo@ulg.ac.be
M.Stipulanti@ulg.ac.be

Abstract

We count the number of distinct (scattered) subwords occurring in the base-b expansion of the non-negative integers. More precisely, we consider the sequence $(S_b(n))_{n\geq 0}$ counting the number of positive entries on each row of a generalization of the Pascal triangle to binomial coefficients of base-b expansions. By using a convenient tree structure, we provide recurrence relations for $(S_b(n))_{n\geq 0}$ leading to the b-regularity of the latter sequence. Then we deduce the asymptotics of the summatory function of the sequence $(S_b(n))_{n\geq 0}$.

1 Introduction

A *finite word* is a finite sequence of letters belonging to a finite set called the *alphabet*. The *binomial coefficient* $\binom{u}{v}$ of two finite words u and v is the number of times v occurs as a subsequence of u (meaning as a "scattered" subword). All along the paper, we let b denote an integer greater than 1. We let $\mathrm{rep}_b(n)$ denote the (greedy) base-b expansion of $n \in \mathbb{N}\setminus\{0\}$ starting with a non-zero digit. We set $\mathrm{rep}_b(0)$ to be the empty word denoted by ε. We let

$$L_b = \{1,\dots,b-1\}\{0,\dots,b-1\}^* \cup \{\varepsilon\}$$

be the set of base-b expansions of the non-negative integers. For all $w \in \{0,\dots,b-1\}^*$, we also define $\mathrm{val}_b(w)$ to be the value of w in base b, i.e., if $w = w_n \cdots w_0$ with $w_i \in \{0,\dots,b-1\}$ for all i, then $\mathrm{val}_b(w) = \sum_{i=0}^{n} w_i b^i$.

Several generalizations and variations of the Pascal triangle exist and lead to interesting combinatorial, geometrical or dynamical properties [5, 6, 13, 14, 15]. Ordering the words of L_b by increasing genealogical order, we introduced Pascal-like triangles P_b [15] where the entry $\mathrm{P}_b(m,n)$ is $\binom{\mathrm{rep}_b(m)}{\mathrm{rep}_b(n)}$. Clearly P_b contains $(b-1)$ copies of the usual Pascal triangle when only considering words of the form a^m with $a \in \{1,\dots,b-1\}$ and $m \geq 0$. In Figure 1, we depict the first few elements of P_3 A284441 and its compressed version highlighting the number of positive elements on each line. The data provided by this compressed version is summed up in Definition 1.

Definition 1. For $n \geq 0$, we define the sequence $(S_b(n))_{n\geq 0}$ by setting

$$S_b(n) := \#\left\{ v \in L_b \mid \binom{\mathrm{rep}_b(n)}{v} > 0 \right\}. \tag{1}$$

We also consider the summatory function $(A_b(n))_{n\geq 0}$ of the sequence $(S_b(n))_{n\geq 0}$ defined by $A_b(0) = 0$ and for all $n \geq 1$,

$$A_b(n) := \sum_{j=0}^{n-1} S_b(j).$$

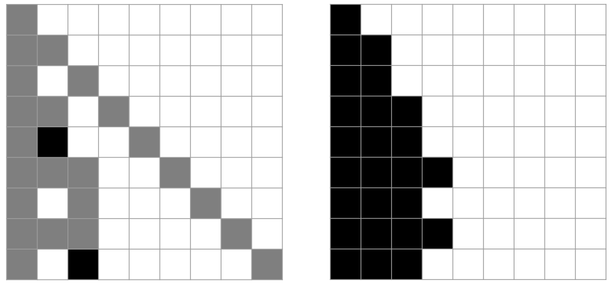

Figure 1: On the left, the first few rows of the generalized Pascal triangle P_3 (a white (resp., gray; resp., black) square corresponds to 0 (resp., 1; resp., 2)) and on the right, its compressed version.

The quantity $A_b(n)$ can be thought of as the total number of base-b expansions occurring as subwords in the base-b expansion of integers less than n (the same subword is counted k times if it occurs in the base-b expansion of k distinct integers).

In some sense, the sequences $(S_b(n))_{n\geq 0}$ and $(A_b(n))_{n\geq 0}$ measure the sparseness of P_b.

Example 2. If $b = 3$, then the first few terms of the sequence $(S_3(n))_{n\geq 0}$ A282715 are

$$1, 2, 2, 3, 3, 4, 3, 4, 3, 4, 5, 6, 5, 4, 6, 7, 7, 6, 4, 6, 5, 7, 6, 7, 5, 6, 4, 5, 7, 8, 8, 7, 10, \ldots$$

For instance, the subwords of the word 121 are $\varepsilon, 1, 2, 11, 12, 21, 121$. Thus, $S_3(\mathrm{val}_3(121)) = S_3(16) = 7$. The first few terms of $(A_3(n))_{n\geq 0}$ A284442 are

$$0, 1, 3, 5, 8, 11, 15, 18, 22, 25, 29, 34, 40, 45, 49, 55, \ldots$$

We studied [16] the triangle P_2 A282714 and the sequence $(S_2(n))_{n\geq 0}$ A007306, which turns out to be the subsequence with odd indices of the Stern–Brocot sequence. The sequence $(S_2(n))_{n\geq 0}$ is 2-regular in the sense of Allouche and Shallit [1]. We studied [17] the behavior of $(A_2(n))_{n\geq 0}$ A282720. To this aim, we exploited a particular decomposition of $A_2(2^\ell + r)$, for all $\ell \geq 1$ and all $0 \leq r < 2^\ell$, using powers of 3.

1.1 Our contribution

We conjectured six recurrence relations for $(S_3(n))_{n\geq 0}$ depending on the position of n between two consecutive powers of 3; see [16]. Using the heuristic from [3] suggesting recurrence relations, the sequence $(S_3(n))_{n\geq 0}$ was expected to be 3-regular. It was not obvious that we could derive general recurrence relations for $(S_b(n))_{n\geq 0}$ from the form of those satisfied by $(S_2(n))_{n\geq 0}$. We thought that $(b-1)b$ recurrence relations should be needed in the general case, leading to a cumbersome statement. Moreover it was computationally challenging to obtain many terms of $(S_b(n))_{n\geq 0}$ for large b because the number of words of length n in L_b grows like b^n. Therefore we lack data to conjecture the b-regularity of $(S_b(n))_{n\geq 0}$.

When studying $(A_2(n))_{n\geq 0}$, a possible extension seemed to emerge [17]. In particular, we prove that $A_2(2n) = 3A_2(n)$ and, sustained by computer experiments, we conjectured that $A_b(nb) = (2b-1)A_b(n)$.

Surprisingly, for all $b \geq 2$, we show in Section 2 that the recurrence relations satisfied by $(S_b(n))_{n\geq 0}$ reduce to three forms; see Proposition 3. In particular, this proves the conjecture stated in [16]. Therefore, in Section 3, we deduce the b-regularity of $(S_b(n))_{n\geq 0}$; see Theorem 16. Moreover we obtain a linear representation of the sequence with $b \times b$ matrices. We also show that $(S_b(n))_{n\geq 0}$ is palindromic over $[(b-1)b^\ell, b^{\ell+1}]$.

The key to study the asymptotics of $(A_b(n))_{n\geq 0}$ is to obtain specific recurrence relations for this sequence. In Proposition 26, we show that theses relations involve powers of $(2b-1)$. Therefore, we prove the conjecture

$\operatorname{rep}_b(n)$	ε	x	$x0$	xx	xy	$x00$	$x0x$	$x0y$	$xx0$	xxx	xxy	$xy0$	xyx	xyy	xyz
$S_b(n)$	1	2	3	3	4	4	5	6	5	4	6	7	7	6	8

Table 1: The first few values of $S_b(n)$ for $0 \le n < b^3$, with pairwise distinct $x, y, z \in \{1, \dots, b-1\}$.

about $A_b(nb)$. In Section 4, using the so-called $(2b-1)$-decompositions, we may apply the method introduced in [17].

We think that this paper motivates the quest for generalized Stern–Brocot sequences and analogues of the Farey tree [4, 7, 8, 11, 12, 18]. Namely can one reasonably define a tree structure, or some other combinatorial structure, in which the sequence $(S_b(n))_{n\ge 0}$ naturally appears?

Most of the results are proved by induction and the base case usually takes into account the values of $S_b(n)$ for $0 \le n < b^2$. These values are easily obtained from Definition 1 and summarized in Table 1.

2 General recurrence relations in base b

The aim of this section is to prove the following result exhibiting recurrence relations satisfied by the sequence $(S_b(n))_{n\ge 0}$. This result is useful to prove that the summatory function of the latter sequence also satisfies recurrence relations; see Section 4.

Proposition 3. *The sequence $(S_b(n))_{n\ge 0}$ satisfies $S_b(0) = 1$, $S_b(1) = \cdots = S_b(b-1) = 2$, and, for all $x, y \in \{1, \dots, b-1\}$ with $x \ne y$, all $\ell \ge 1$ and all $r \in \{0, \dots, b^{\ell-1} - 1\}$,*

$$S_b(xb^\ell + r) = S_b(xb^{\ell-1} + r) + S_b(r); \tag{2}$$

$$S_b(xb^\ell + xb^{\ell-1} + r) = 2S_b(xb^{\ell-1} + r) - S_b(r); \tag{3}$$

$$S_b(xb^\ell + yb^{\ell-1} + r) = S_b(xb^{\ell-1} + r) + 2S_b(yb^{\ell-1} + r) - 2S_b(r). \tag{4}$$

For the sake of completeness, we recall the definition of a particularly useful tool called *the trie of subwords* to prove Proposition 3. This tool is also useful to prove the b-regularity of the sequence $(S_b(n))_{n\ge 0}$; see Section 3.

Definition 4. Let w be a finite word over $\{0, \dots, b-1\}$. The language of its subwords is factorial, i.e., if xyz is a subword of w, then y is also a subword of w. Thus we may associate with w, the *trie*[1] *of its subwords*. The root is ε and if u and ua are two subwords of w with $a \in \{0, \dots, b-1\}$, then ua is a child of u. We let $\mathcal{T}(w)$ denote the subtree in which we only consider the children $1, \dots, b-1$ of the root ε and their successors, if they exist.

Remark 5. The number of nodes on level $\ell \ge 0$ in $\mathcal{T}(w)$ counts the number of subwords of length ℓ in L_b occurring in w. In particular, the number of nodes of the trie $\mathcal{T}(\operatorname{rep}_b(n))$ is exactly $S_b(n)$ for all $n \ge 0$.

Definition 6. For each non-empty word $w \in L_b$, we consider a factorization of w into maximal blocks of consecutively distinct letters (i.e., $a_i \ne a_{i+1}$ for all i) of the form

$$w = a_1^{n_1} \cdots a_M^{n_M},$$

with $n_\ell \ge 1$ for all ℓ. For each $\ell \in \{0, \dots, M-1\}$, we consider the subtree T_ℓ of $\mathcal{T}(w)$ whose root is the node $a_1^{n_1} \cdots a_\ell^{n_\ell} a_{\ell+1}$. For convenience, we set T_M to be an empty tree with no node. Roughly speaking, we have a root of a new subtree T_ℓ for each new variation of digits in w. For each $\ell \in \{0, \dots, M-1\}$, we also let $\#T_\ell$ denote the number of nodes of the tree T_ℓ.

Note that for $k - i \ge 2$, one could possibly have $a_k = a_i$. For each $\ell \in \{0, \dots, M-1\}$, we let $\operatorname{Alph}(\ell)$ denote the set of letters occurring in $a_{\ell+1} \cdots a_M$. Then for each letter $a \in \operatorname{Alph}(\ell)$, we let $j(a, \ell)$ denote the smallest index in $\{\ell+1, \dots, M\}$ such that $a_{j(a,\ell)} = a$.

[1] This tree is also called prefix tree or radix tree. All successors of a node have a common prefix and the root is the empty word.

Example 7. In this example, we set $b = 3$ and $w = 22000112 \in L_3$. Using the previous notation, we have $M = 4$, $a_1 = 2$, $a_2 = 0$, $a_3 = 1$ and $a_4 = 2$. For instance, $\mathrm{Alph}(0) = \{0, 1, 2\}$, $\mathrm{Alph}(2) = \{1, 2\}$ and $j(0, 0) = 2$, $j(1, 0) = 3$, $j(2, 0) = 1$ and $j(2, 1) = 4$.

The following result describes the structure of the tree $\mathcal{T}(w)$. It directly follows from the definition.

Proposition 8 ([16, Proposition 27])**.** *Let w be a finite word in L_b. With the above notation about M and the subtrees T_ℓ, the tree $\mathcal{T}(w)$ has the following properties.*

1. *The node of label ε has $\#(\mathrm{Alph}(0) \setminus \{0\})$ children that are a for $a \in \mathrm{Alph}(0) \setminus \{0\}$. Each child a is the root of a tree isomorphic $T_{j(a,0)-1}$.*

2. *For each $\ell \in \{0, \dots, M-1\}$ and each $i \in \{0, \dots, n_{\ell+1} - 1\}$ with $(\ell, i) \neq (0, 0)$, the node of label $x = a_1^{n_1} \cdots a_\ell^{n_\ell} a_{\ell+1}^i$ has $\#(\mathrm{Alph}(\ell))$ children that are xa for $a \in \mathrm{Alph}(\ell)$. Each child xa with $a \neq a_{\ell+1}$ is the root of a tree isomorphic to $T_{j(a,\ell)-1}$.*

Example 9. Let us continue Example 7. The tree $\mathcal{T}(22000112)$ is depicted in Figure 2. We use three different colors to represent the letters $0, 1, 2$. The tree T_0 (resp., T_1; resp., T_2; resp., T_3) is the subtree of

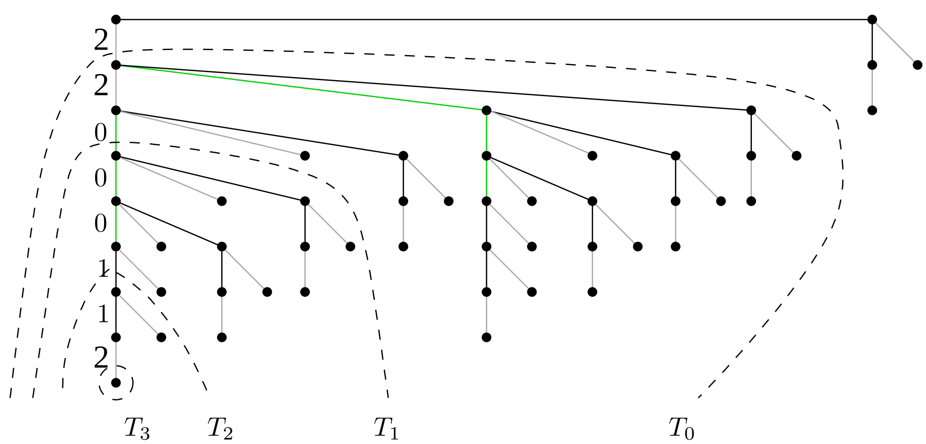

Figure 2: The trie $\mathcal{T}(22000112)$.

$\mathcal{T}(w)$ with root 2 (resp., $2^2 0$; resp., $2^2 0^3 1$; resp., $2^2 0^3 1^2 2$). These subtrees are represented in Figure 2 using dashed lines. The tree T_3 is limited to a single node since the number of nodes of T_{M-1} is n_M, which is equal to 1 in this example.

Using tries of subwords, we prove the following five lemmas. Their proofs are essentially the same, so we only prove two of them.

Lemma 10. *For each letter $x \in \{1, \dots, b-1\}$ and each word $u \in \{0, \dots, b-1\}^*$, we have*

$$\# \left\{ v \in L_b \mid \binom{x00u}{v} > 0 \right\} = 2 \cdot \# \left\{ v \in L_b \mid \binom{x0u}{v} > 0 \right\} - \# \left\{ v \in L_b \mid \binom{xu}{v} > 0 \right\}.$$

Proof. Recall that from Remark 5, we need to prove that $\#\mathcal{T}(x00u) = 2\#\mathcal{T}(x0u) - \#\mathcal{T}(xu)$.

Assume first that u is of the form $u = 0^n$, $n \geq 0$. The tree $\mathcal{T}(xu)$ is linear and has $n + 2$ nodes, $\mathcal{T}(x0u)$ has $n + 3$ nodes and $\mathcal{T}(x00u)$ has $n + 4$ nodes. The formula holds.

Now suppose that u contains other letters than 0. We let $a_1, \dots, a_m$ denote all the pairwise distinct letters of u different from 0. They are implicitly ordered with respect to their first appearance in u. If

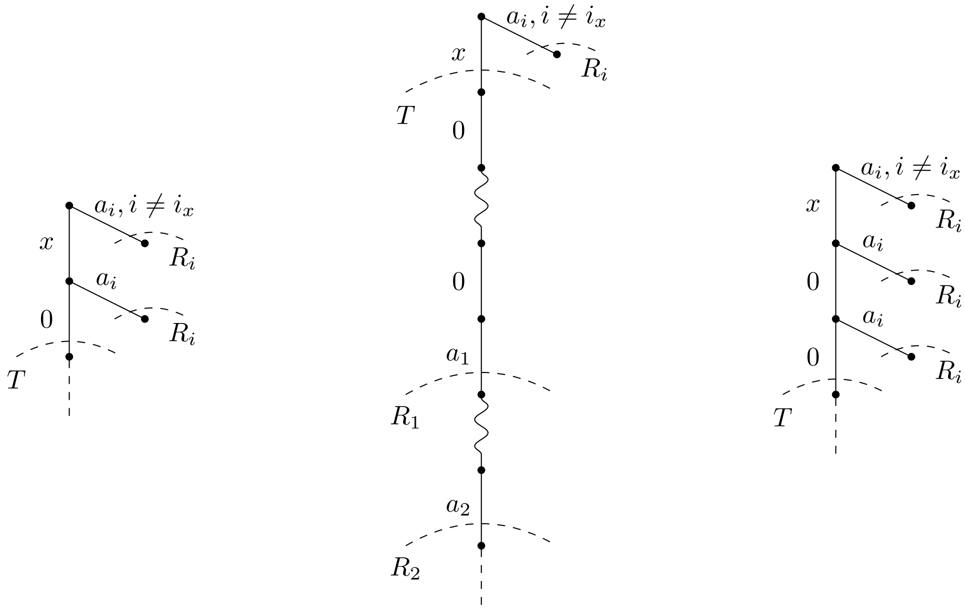

Figure 3: Schematic structure of the trees $\mathcal{T}(x0u)$, $\mathcal{T}(xu)$ and $\mathcal{T}(x00u)$.

$x \in \{a_1, \dots, a_m\}$, we let $i_x \in \{1, \dots, m\}$ denote the index such that $a_{i_x} = x$. For all $i \in \{1, \dots, m\}$, we let $u_i a_i$ denote the prefix of u that ends with the first occurrence of the letter a_i in u, and we let R_i denote the subtree of $\mathcal{T}(xu)$ with root xu_ia_i.

First, observe that the subtree T of $\mathcal{T}(xu)$ with root x is equal to the subtree of $\mathcal{T}(x0u)$ with root $x0$ and also to the subtree of $\mathcal{T}(x00u)$ with root $x00$.

Secondly, for all $i \in \{1, \dots, m\}$, the subtree of $\mathcal{T}(x0u)$ with root xa_i is R_i. Similarly, $\mathcal{T}(x00u)$ contains two copies of R_i: the subtrees of root xa_i and $x0a_i$.

Finally, for all $i \in \{1, \dots, m\}$ with $i \neq i_x$, the subtree of $\mathcal{T}(x0u)$ with root a_i is R_i and the subtree of $\mathcal{T}(x00u)$ with root a_i is R_i.

The situation is depicted in Figure 3 where we put a unique edge for several indices when necessary, e.g., the edge labeled by a_i stands for m edges labeled by $a_1, \dots, a_m$. The claimed formula holds since

$$2 \cdot (2 + \#T + 2 \sum_{\substack{1 \leq i \leq m \\ i \neq i_x}} \#R_i + \#R_{i_x}) - (1 + \#T + \sum_{\substack{1 \leq i \leq m \\ i \neq i_x}} \#R_i) = 3 + \#T + 3 \sum_{\substack{1 \leq i \leq m \\ i \neq i_x}} \#R_i + 2\#R_{i_x}.$$

□

Lemma 11. *For each letter $x \in \{1, \dots, b-1\}$ and each word $u \in \{0, \dots, b-1\}^*$, we have*

$$\#\left\{v \in L_b \mid \binom{xx0u}{v} > 0\right\} = \#\left\{v \in L_b \mid \binom{x0u}{v} > 0\right\} + \#\left\{v \in L_b \mid \binom{xu}{v} > 0\right\}.$$

Proof. The proof is similar to the proof of Lemma 10. □

Lemma 12. *For all letters $x, y \in \{1, \dots, b-1\}$ and each word $u \in \{0, \dots, b-1\}^*$, we have*

$$\#\left\{v \in L_b \mid \binom{x0yu}{v} > 0\right\} = \#\left\{v \in L_b \mid \binom{xyu}{v} > 0\right\} + \#\left\{v \in L_b \mid \binom{yu}{v} > 0\right\}.$$

Proof. The proof is similar to the proof of Lemma 10. Observe that one needs to divide the proof into two cases according to whether x is equal to y or not. As a first case, also consider $u = y^n$ with $n \geq 0$ instead of $u = 0^n$ with $n \geq 0$. $\square$

Lemma 13. *For all letters $x, y \in \{1, \ldots, b-1\}$ and each word $u \in \{0, \ldots, b-1\}^*$, we have*

$$\#\left\{v \in L_b \mid \binom{xxyu}{v} > 0\right\} = 2 \cdot \#\left\{v \in L_b \mid \binom{xyu}{v} > 0\right\} - \#\left\{v \in L_b \mid \binom{yu}{v} > 0\right\}.$$

Proof. The proof is similar to the proof of Lemma 12. $\square$

The next lemma having a slightly more technical proof, we present it.

Lemma 14. *For all letters $x, y \in \{1, \ldots, b-1\}$ with $x \neq y$, $z \in \{0, \ldots, b-1\}$ and each word $u \in \{0, \ldots, b-1\}^*$, we have*

$$\begin{aligned}\#\left\{v \in L_b \mid \binom{xyzu}{v} > 0\right\} =& \#\left\{v \in L_b \mid \binom{xzu}{v} > 0\right\} + 2 \cdot \#\left\{v \in L_b \mid \binom{yzu}{v} > 0\right\} \\ & - 2 \cdot \#\left\{v \in L_b \mid \binom{zu}{v} > 0\right\}.\end{aligned}$$

Proof. Let $x, y \in \{1, \ldots, b-1\}$ with $x \neq y$, $z \in \{0, \ldots, b-1\}$, and let $u \in \{0, \ldots, b-1\}^*$. Our reasoning is again based on the structure of the associated trees. The proof is divided into two cases depending on the fact that $z = 0$ or not.

• As a first case, suppose that $z \neq 0$. Now assume that u is of the form $u = z^n$, $n \geq 0$. If $x \neq z$ and $y \neq z$, the tree $\mathcal{T}(zu)$ is linear and has $n+2$ nodes, $\mathcal{T}(xzu)$ and $\mathcal{T}(yzu)$ have $2(n+2)$ nodes and $\mathcal{T}(xyzu)$ has $4(n+2)$ nodes and the claimed formula holds. If $x \neq z$ and $y = z$, the tree $\mathcal{T}(zu)$ is linear and has $n+2$ nodes, $\mathcal{T}(xzu)$ has $2(n+2)$ nodes, $\mathcal{T}(yzu)$ has $n+3$ nodes and $\mathcal{T}(xyzu)$ has $2(n+3)$ nodes and the claimed formula holds. If $x = z$ and $y \neq z$, the tree $\mathcal{T}(zu)$ is linear and has $n+2$ nodes, $\mathcal{T}(xzu)$ has $n+3$ nodes, $\mathcal{T}(yzu)$ has $2(n+2)$ nodes and $\mathcal{T}(xyzu)$ has $3(n+2)+1$ nodes and the claimed formula holds.

Now suppose that u contains other letters than z. We let $a_1, \ldots, a_m$ denote all the pairwise distinct letters of u different from z. They are implicitly ordered with respect to their first appearance in u. If $x, y, 0 \in \{a_1, \ldots, a_m\}$, we let $i_x, i_y, i_0 \in \{1, \ldots, m\}$ respectively denote the indices such that $a_{i_x} = x$, $a_{i_y} = y$ and $a_{i_0} = 0$. For all $i \in \{1, \ldots, m\}$, we let $u_i a_i$ denote the prefix of u that ends with the first occurrence of the letter a_i in u, and we let R_i denote the subtree of $\mathcal{T}(zu)$ with root $zu_i a_i$.

First, observe that the subtree T of $\mathcal{T}(zu)$ with root z is equal to the subtree of $\mathcal{T}(xzu)$ with root xz, to the subtree of $\mathcal{T}(yzu)$ with root yz and also to the subtree of $\mathcal{T}(xyzu)$ with root xyz.

Suppose that $x \neq z$ and $y \neq z$. Using the same reasoning as in the proof of Lemma 10, the situation is depicted in Figure 4. The claimed formula holds since

$$\begin{aligned}&(2 + 2\#T + 2 \sum_{\substack{1 \leq i \leq m \\ i \neq i_x, i_y, i_0}} \#R_i + \#R_{i_x} + 2\#R_{i_y} + \#R_{i_0}) \\ &+ 2 \cdot (2 + 2\#T + 2 \sum_{\substack{1 \leq i \leq m \\ i \neq i_x, i_y, i_0}} \#R_i + 2\#R_{i_x} + \#R_{i_y} + \#R_{i_0}) \\ &- 2 \cdot (1 + \#T + \sum_{\substack{1 \leq i \leq m \\ i \neq i_x, i_y, i_0}} \#R_i + \#R_{i_x} + \#R_{i_y}) \\ &= 4 + 4\#T + 4 \sum_{\substack{1 \leq i \leq m \\ i \neq i_x, i_y, i_0}} \#R_i + 3\#R_{i_x} + 2\#R_{i_y} + 3\#R_{i_0}.\end{aligned}$$

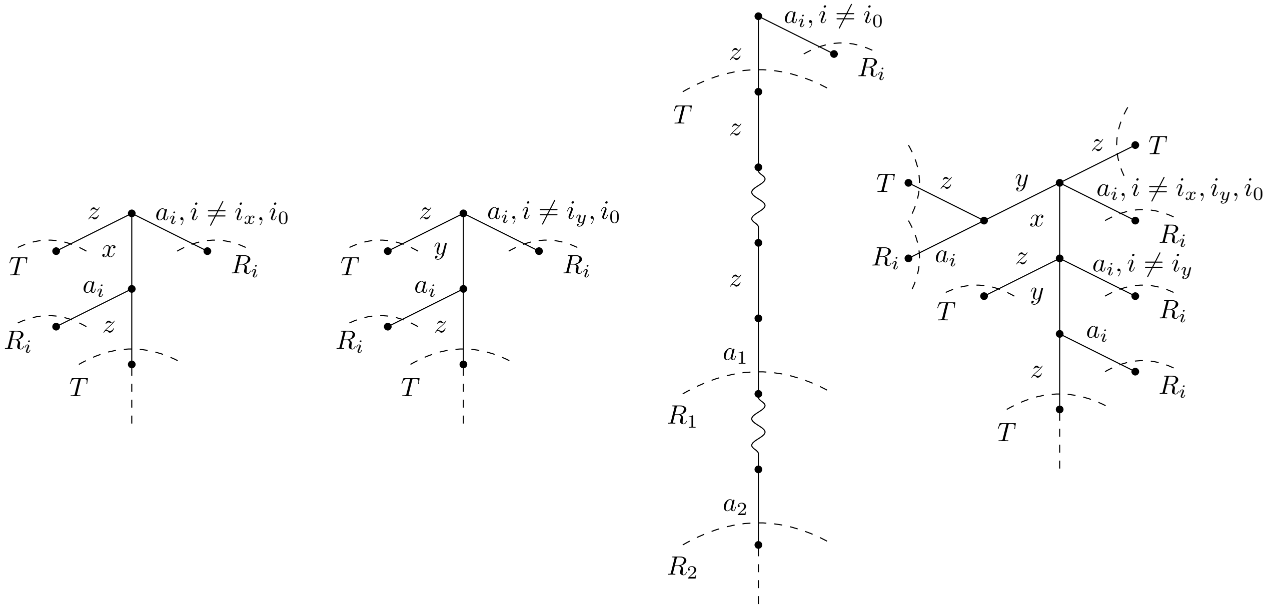

Figure 4: Schematic structure of the trees $\mathcal{T}(xzu)$, $\mathcal{T}(yzu)$, $\mathcal{T}(zu)$ and $\mathcal{T}(xyzu)$ when $x \neq z$, $y \neq z$ and $z \neq 0$.

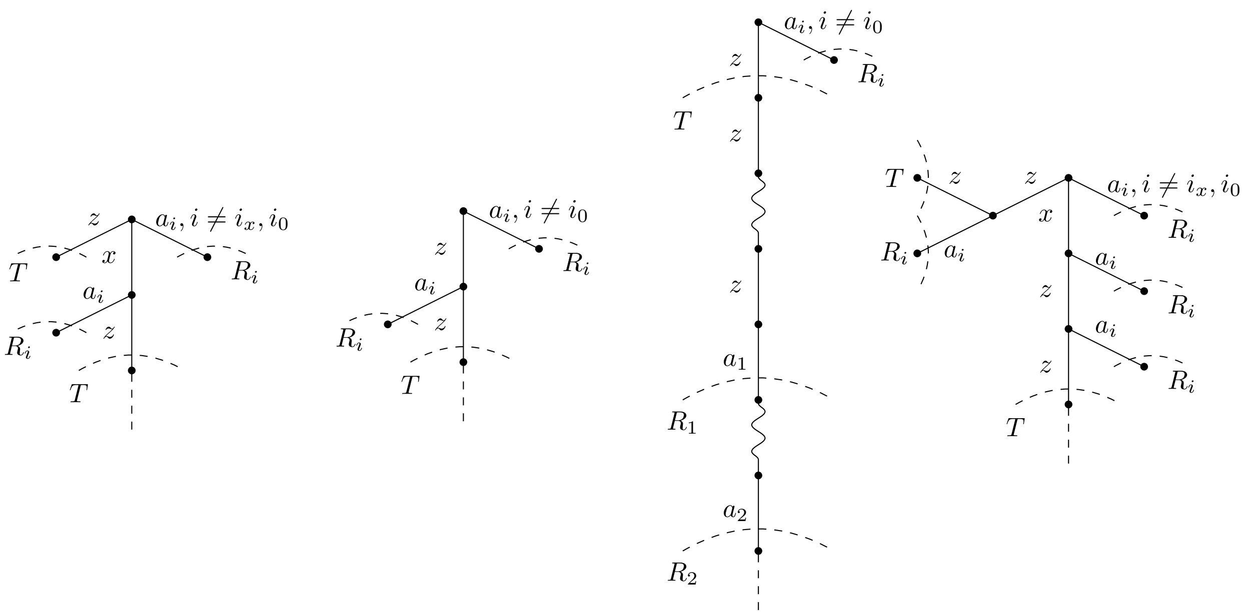

Figure 5: Schematic structure of the trees $\mathcal{T}(xzu)$, $\mathcal{T}(yzu)$, $\mathcal{T}(zu)$ and $\mathcal{T}(xyzu)$ when $x \neq z$, $y = z$ and $z \neq 0$.

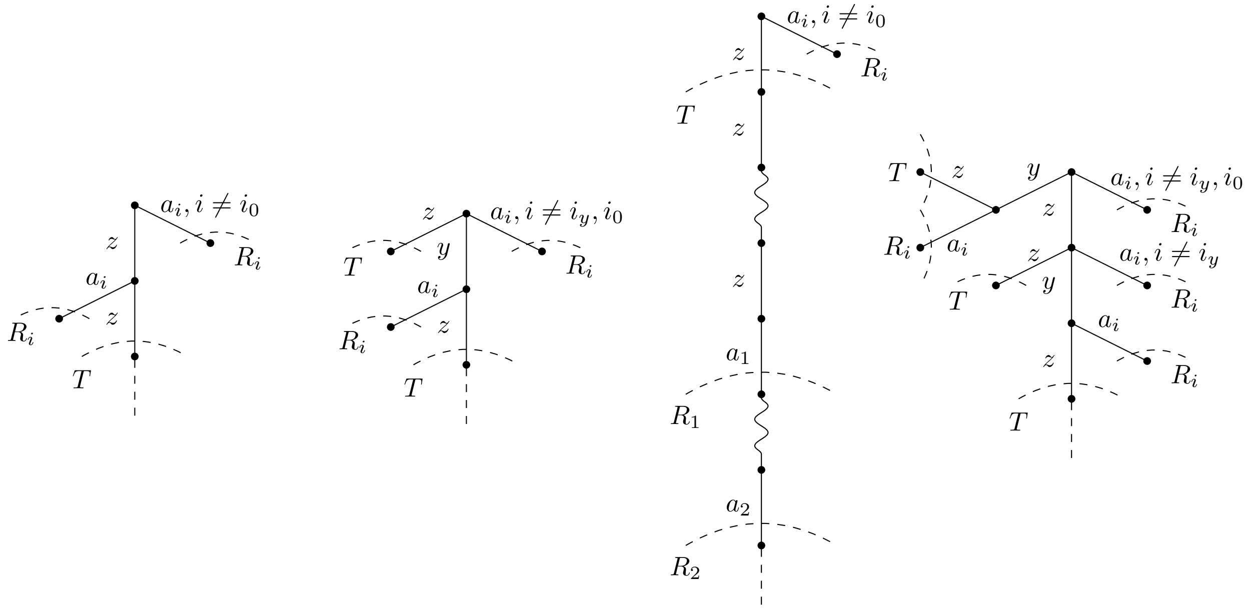

Figure 6: Schematic structure of the trees $\mathcal{T}(xzu)$, $\mathcal{T}(yzu)$, $\mathcal{T}(zu)$ and $\mathcal{T}(xyzu)$ when $x = z$, $y \neq z$ and $z \neq 0$.

Suppose that $x \neq z$ and $y = z$. The situation is depicted in Figure 5. The claimed formula holds since

$$\begin{aligned}
&(2 + 2\#T + 2 \sum_{\substack{1 \le i \le m \\ i \neq i_x, i_0}} \#R_i + \#R_{i_x} + \#R_{i_0}) \\
&+ 2 \cdot (2 + \#T + 2 \sum_{\substack{1 \le i \le m \\ i \neq i_x, i_0}} \#R_i + 2\#R_{i_x} + \#R_{i_0}) \\
&- 2 \cdot (1 + \#T + \sum_{\substack{1 \le i \le m \\ i \neq i_x, i_0}} \#R_i + \#R_{i_x}) \\
&= 4 + 2\#T + 4 \sum_{\substack{1 \le i \le m \\ i \neq i_x, i_0}} \#R_i + 3\#R_{i_x} + 3\#R_{i_0}.
\end{aligned}$$

Suppose that $x = z$ and $y \neq z$. The situation is depicted in Figure 6. The claimed formula holds since

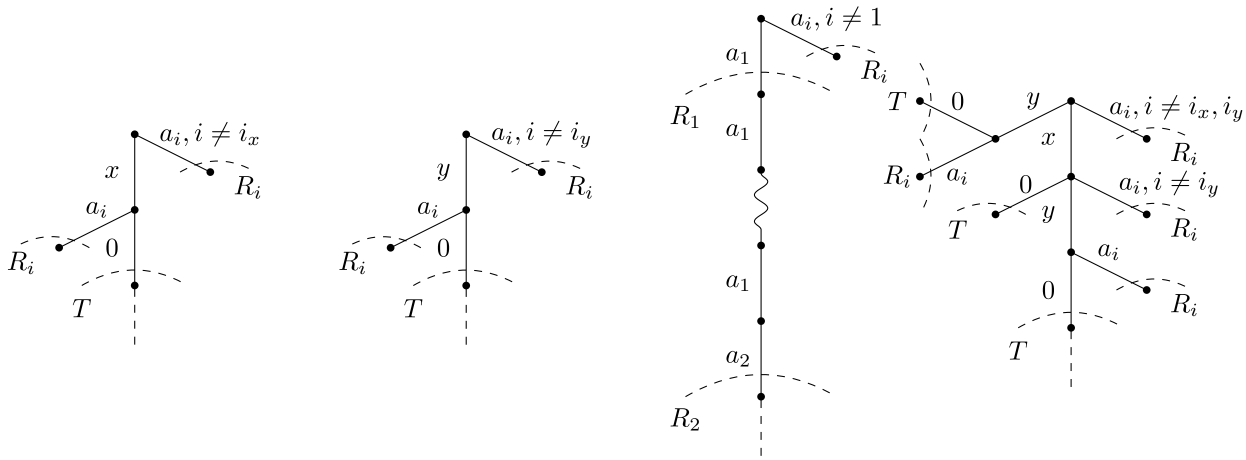

Figure 7: Schematic structure of the trees $\mathcal{T}(x0u)$, $\mathcal{T}(y0u)$, $\mathcal{T}(\text{rep}_b(\text{val}_b(u)))$ and $\mathcal{T}(xy0u)$.

$$\begin{aligned}
&(2 + \#T + 2 \sum_{\substack{1\le i\le m\\ i\neq i_y, i_0}} \#R_i + 2\#R_{i_y} + \#R_{i_0}) \\
&+ 2\cdot(2 + 2\#T + 2 \sum_{\substack{1\le i\le m\\ i\neq i_y, i_0}} \#R_i + \#R_{i_y} + \#R_{i_0}) \\
&- 2\cdot(1 + \#T + \sum_{\substack{1\le i\le m\\ i\neq i_y, i_0}} \#R_i + \#R_{i_y}) \\
&= 4 + 3\#T + 4 \sum_{\substack{1\le i\le m\\ i\neq i_y, i_0}} \#R_i + 2\#R_{i_y} + 3\#R_{i_0}.
\end{aligned}$$

• As a second case, suppose that $z = 0$. Then, by convention, leading zeroes are not allowed in base-b expansions and we must prove that the following formula holds

$$\begin{aligned}
\#\left\{v \in L_b \mid \binom{xy0u}{v} > 0\right\} =& \#\left\{v \in L_b \mid \binom{x0u}{v} > 0\right\} + 2\cdot\#\left\{v \in L_b \mid \binom{y0u}{v} > 0\right\} \\
&- 2\cdot\#\left\{v \in L_b \mid \binom{\text{rep}_b(\text{val}_b(u))}{v} > 0\right\}.
\end{aligned}$$

It is useful to note that $\text{rep}_b(\text{val}_b(\cdot)) : \{0, \dots, b-1\}^* \mapsto L_b$ plays a normalization role. It removes leading zeroes.

If $u = 0^n$, with $n \geq 0$, then $\text{rep}_b(\text{val}_b(u)) = \varepsilon$ and the tree $\mathcal{T}(\text{rep}_b(\text{val}_b(u)))$ has only one node. The trees $\mathcal{T}(x0u)$ and $\mathcal{T}(y0u)$ both have $n+3$ nodes and the tree $\mathcal{T}(xy0u)$ has $3(n+2)+1$ nodes and the claimed formula holds.

Now suppose that u contains other letters than 0. We let $a_1, \dots, a_m$ denote all the pairwise distinct letters of u different from 0. They are implicitly ordered with respect to their first appearance in u. If $x, y \in \{a_1, \dots, a_m\}$, we let $i_x, i_y \in \{1, \dots, m\}$ respectively denote the indices such that $a_{i_x} = x$ and $a_{i_y} = y$. For all $i \in \{1, \dots, m\}$, we let $u_i' a_i$ denote the prefix of $\text{rep}_b(\text{val}_b(u))$ that ends with the first occurrence of the letter a_i in $\text{rep}_b(\text{val}_b(u))$, and we let R_i denote the subtree of $\mathcal{T}(\text{rep}_b(\text{val}_b(u)))$ with root $u_i' a_i$.

The situation is depicted in Figure 7. Observe that the subtree T of $\mathcal{T}(y0u)$ with root $y0$ is equal to the subtree of $\mathcal{T}(x0u)$ with root $x0$ and to the subtree of $\mathcal{T}(xy0u)$ with root $xy0$. The claimed formula holds since

$$\begin{aligned}
&(2+\#T+2\sum_{\substack{1\le i\le m\\ i\ne i_x,i_y}}\#R_i+\#R_{i_x}+2\#R_{i_y})\\
&+2\cdot(2+\#T+2\sum_{\substack{1\le i\le m\\ i\ne i_x,i_y}}\#R_i+2\#R_{i_x}+\#R_{i_y})\\
&-2\cdot(1+\sum_{\substack{1\le i\le m\\ i\ne i_x,i_y}}\#R_i+\#R_{i_x}+\#R_{i_y})\\
&=4+3\#T+4\sum_{\substack{1\le i\le m\\ i\ne i_x,i_y}}\#R_i+3\#R_{i_x}+2\#R_{i_y}.
\end{aligned}$$

□

Those five lemmas can be translated into recurrence relations satisfied by the sequence $(S_b(n))_{n\ge 0}$ using Definition 1.

Proof of Proposition 3. The first part is clear using Table 1. Let $x,y\in\{1,\dots,b-1\}$ with $x\ne y$. Proceed by induction on $\ell\ge 1$.

Let us first prove (2). If $\ell=1$, then $r=0$ and (2) follows from Table 1. Now suppose that $\ell\ge 2$ and assume that (2) holds for all $\ell'<\ell$. Let $r\in\{0,\dots,b^{\ell-1}-1\}$, and let u be a word in $\{0,\dots,b-1\}^*$ such that $|u|\ge 1$ and $\mathrm{rep}_b(xb^\ell+r)=x0u$. The proof is divided into two parts according to the first letter of u. If $u=0u'$ with $u'\in\{0,\dots,b-1\}^*$, then

$$\begin{aligned}
S_b(xb^\ell+r) &= 2S_b(xb^{\ell-1}+r)-S_b(xb^{\ell-2}+r) && \text{(by Lemma 10)}\\
&= 2(S_b(xb^{\ell-2}+r)+S_b(r))-S_b(xb^{\ell-2}+r) && \text{(by induction hypothesis)}\\
&= S_b(xb^{\ell-2}+r)+S_b(r)+S_b(r)\\
&= S_b(xb^{\ell-1}+r)+S_b(r), && \text{(by induction hypothesis)}
\end{aligned}$$

which proves (2). Now if $u=zu'$ with $z\in\{1,\dots,b-1\}$ and $u'\in\{0,\dots,b-1\}^*$, then (2) directly follows from Definition 1 and Lemma 12.

Let us prove (3). If $\ell=1$, then $r=0$ and (2) follows from Table 1. Now suppose that $\ell\ge 2$ and assume that (3) holds for all $\ell'<\ell$. Let $r\in\{0,\dots,b^{\ell-1}-1\}$, and let u be a word in $\{0,\dots,b-1\}^*$ such that $|u|\ge 1$ and $\mathrm{rep}_b(xb^\ell+xb^{\ell-1}+r)=xxu$. The proof is divided into two parts according to the first letter of u. If $u=0u'$ with $u'\in\{0,\dots,b-1\}^*$, then

$$\begin{aligned}
S_b(xb^\ell+xb^{\ell-1}+r) &= S_b(xb^{\ell-1}+r)+S_b(xb^{\ell-2}+r) && \text{(by Lemma 11)}\\
&= S_b(xb^{\ell-2}+r)+S_b(r))+S_b(xb^{\ell-2}+r) && \text{(using (2))}\\
&= 2(S_b(xb^{\ell-2}+r)+S_b(r))-S_b(r)\\
&= 2S_b(xb^{\ell-1}+r)-S_b(r), && \text{(using (2))}
\end{aligned}$$

which proves (3). Now if $u=zu'$ with $z\in\{1,\dots,b-1\}$ and $u'\in\{0,\dots,b-1\}^*$, then (3) directly follows from Definition 1 and Lemma 13.

Let us finally prove (4). If $\ell=1$, then $r=0$ and (2) follows from Table 1. Now suppose that $\ell\ge 2$ and assume that (4) holds for all $\ell'<\ell$. Let $r\in\{0,\dots,b^{\ell-1}-1\}$, let z be a letter of $\{1,\dots,b-1\}$ and let u be a word in $\{0,\dots,b-1\}^*$ such that $\mathrm{rep}_b(xb^\ell+yb^{\ell-1}+r)=xyzu$. Using Definition 1 and Lemma 14, we directly have that

$$S_b(xb^\ell+yb^{\ell-1}+r)=S_b(xb^{\ell-1}+r)+S_b(yb^{\ell-1}+r)-2S_b(r),$$

which proves (4). □

3 Regularity of the sequence $(S_b(n))_{n\geq 0}$

The sequence $(S_2(n))_{n\geq 0}$ is shown to be 2-regular; see [16]. We recall that the *b-kernel* of a sequence $s = (s(n))_{n\geq 0}$ is the set

$$\mathcal{K}_b(s) = \{(s(b^i n + j))_{n\geq 0} | \; i \geq 0 \text{ and } 0 \leq j < b^i\}.$$

A sequence $s = (s(n))_{n\geq 0} \in \mathbb{Z}^{\mathbb{N}}$ is *b-regular* if there exists a finite number of sequences $(t_1(n))_{n\geq 0}, \ldots, (t_\ell(n))_{n\geq 0}$ such that every sequence in the $\mathbb{Z}$-module $\langle \mathcal{K}_b(s) \rangle$ generated by the b-kernel $\mathcal{K}_b(s)$ is a $\mathbb{Z}$-linear combination of the t_r's. In this section, we prove that the sequence $(S_b(n))_{n\geq 0}$ is b-regular. As a consequence, one can get matrices to compute $S_b(n)$ in a number of matrix multiplications proportional to $\log_b(n)$. To prove the b-regularity of the sequence $(S_b(n))_{n\geq 0}$ for any base b, we first need a lemma involving some matrix manipulations.

Lemma 15. *Let I and 0 respectively be the identity matrix of size $b^2 \times b^2$ and the zero matrix of size $b^2 \times b^2$. Let M_b be the block-matrix of size $b^3 \times b^3$*

$$M_b := \begin{pmatrix} I & I & 2I & \cdots & \cdots & \cdots & 2I \\ 2I & 3I & 3I & 4I & \cdots & \cdots & 4I \\ \vdots & \vdots & 4I & \ddots & \ddots & & \vdots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \vdots & & \ddots & \ddots & 4I \\ \vdots & \vdots & \vdots & & & \ddots & 3I \\ 2I & 3I & 4I & \cdots & \cdots & \cdots & 4I \end{pmatrix}.$$

This matrix is invertible and its inverse is given by

$$M_b^{-1} := \begin{pmatrix} 3I & 2I & \cdots & \cdots & 2I & -(2b-3)I \\ -2I & 0 & \cdots & \cdots & 0 & I \\ 0 & -I & \ddots & & \vdots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ \vdots & & \ddots & \ddots & 0 & \vdots \\ 0 & \cdots & \cdots & 0 & -I & I \end{pmatrix}.$$

For the proof of the previous lemma, simply proceed to the multiplication of the two matrices. Using this lemma, we prove that the sequence $(S_b(n))_{n\geq 0}$ is b-regular.

Theorem 16. *For all $r \in \{0, \ldots, b^2 - 1\}$, we have*

$$S_b(nb^2 + r) = a_r S_b(n) + \sum_{s=0}^{b-2} c_{r,s} S_b(nb + s) \quad \forall\, n \geq 0, \tag{5}$$

where the coefficients a_r and $c_{r,s}$ are unambiguously determined by the first few values $S_b(0)$, $S_b(1), \ldots,$ $S_b(b^3 - 1)$ and given in Table 2, Table 3 and Table 4. In particular, the sequence $(S_b(n))_{n\geq 0}$ is b-regular. Moreover, a choice of generators for $\langle \mathcal{K}_b(s) \rangle$ is given by the b sequences $(S_b(n))_{n\geq 0}$, $(S_b(bn))_{n\geq 0}$, $(S_b(bn + 1))_{n\geq 0}, \ldots, (S_b(bn + b - 2))_{n\geq 0}$.

Proof. We proceed by induction on $n \geq 0$. For the base case $n \in \{0, 1, \ldots, b^2 - 1\}$, we first compute the coefficients a_r and $c_{r,s}$ using the values of $S_b(nb^2 + r)$ for $n \in \{0, \ldots, b - 1\}$ and $r \in \{0, \ldots, b^2 - 1\}$. Then we show that (5) also holds with these coefficients for $n \in \{b, \ldots, b^2 - 1\}$.

$\mathrm{rep}_b(r)$	ε	x	$b-1$	$x0$	$(b-1)0$	xx	$(b-1)(b-1)$	xy	$(b-1)x$	$x(b-1)$
a_r	-1	-2	$2b-3$	-2	$4b-4$	-1	$4b-3$	-2	$4b-4$	$2b-3$

Table 2: Values of a_r for $0 \le r < b^2$ with $x, y \in \{1, \dots, b-2\}$ and $x \neq y$.

$\mathrm{rep}_b(r)$	ε	x	$b-1$	$x0$	$(b-1)0$	xx	$(b-1)(b-1)$	xy	$(b-1)x$	$x(b-1)$
$c_{r,0}$	2	2	1	1	-1	0	-2	0	-2	-1

Table 3: Values of $c_{r,0}$ for $0 \le r < b^2$ with $x, y \in \{1, \dots, b-2\}$ and $x \neq y$.

Base case. Let I denote the identity matrix of size $b^2 \times b^2$. The system of b^3 equations (5) when $n \in \{0, \dots, b-1\}$ and $r \in \{0, \dots, b^2-1\}$ can be written as $MX = V$ where the matrix $M \in \mathbb{Z}_{b^3}^{b^3}$ is equal to

$$\begin{pmatrix} S_b(0)I & S_b(0)I & S_b(1)I & S_b(2)I & \cdots & S_b(b-2)I \\ S_b(1)I & S_b(b)I & S_b(b+1)I & S_b(b+2)I & \cdots & S_b(2b-2)I \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ S_b(b-1)I & S_b(b(b-1))I & S_b(b(b-1)+1)I & S_b(b(b-1)+2)I & \cdots & S_b(b(b-1)+b-2)I \end{pmatrix}$$

and the vectors $X, V \in \mathbb{Z}^{b^3}$ are respectively given by

$$X^\mathsf{T} = \begin{pmatrix} a_0 & \cdots & a_{b^2-1} & c_{0,0} & c_{1,0} & \cdots & c_{b^2-1,0} & \cdots & c_{0,b-2} & c_{1,b-2} & \cdots & c_{b^2-1,b-2} \end{pmatrix},$$
$$V^\mathsf{T} = \begin{pmatrix} S_b(0) & S_b(1) & \cdots & S_b(b^3-1) \end{pmatrix}.$$

Observe that in the vector X, the coefficients $c_{r,s}$ are first sorted by s then by r. Using Table 1, the matrix M is equal to the matrix M_b of Lemma 15. By this lemma, the previous system has a unique solution given by $X = M_b^{-1}V$. Consequently, using Lemma 15, we have, for all $r \in \{0, \dots, b^2-1\}$ and all $s \in \{1, \dots, b-2\}$,

$$\begin{aligned} a_r &= 3S_b(r) + 2\sum_{j=1}^{b-2} S_b(jb^2 + r) - (2b-3)\, S_b((b-1)b^2 + r), \\ c_{r,0} &= -2S_b(r) + S_b((b-1)b^2 + r), \\ c_{r,s} &= -S_b(sb^2 + r) + S_b((b-1)b^2 + r). \end{aligned}$$

The values of the coefficients can then be computed using Table 1 and are stored in Table 2, Table 3 and Table 4.

For $n \in \{b, \dots, b^2-1\}$, the values of $S_b(nb^2 + r)$ are given in Table 5, Table 6 and Table 7 according to whether $\mathrm{rep}_b(n)$ is of the form $x0$, xx or xy with $x \neq y$. The proof that (5) holds for each $n \in \{b, \dots, b^2-1\}$ only requires easy computations that are left to the reader.

$\mathrm{rep}_b(r)$	ε	x		$b-1$	$x0$		$(b-1)0$	xx	
s	z	x	z	z	x	z	z	x	z
$c_{r,s}$	0	1	0	-1	2	0	-2	2	0

$\mathrm{rep}_b(r)$	$(b-1)(b-1)$	xy			$x(b-1)$		$(b-1)x$	
s	z	x	y	z	x	z	x	z
$c_{r,s}$	-2	2	1	0	1	-1	-1	-2

Table 4: Values of $c_{r,s}$ for $0 \le r < b^2$ and $1 \le s \le b-2$ with $x, y, z \in \{1, \dots, b-2\}$ pairwise distinct.

$\mathrm{rep}_b(r)$	ε	x	y	$x0$	$y0$	xx	yy	xy	yx	yz
$S_b(nb^2+r)$	5	7	8	8	10	7	9	10	11	12

Table 5: Values of $S_b(nb^2+r)$ for $b \le n < b^2$ with $\mathrm{rep}_b(n) = x0$ and $x, y, z \in \{1, \dots, b-1\}$ pairwise distinct.

$\mathrm{rep}_b(r)$	ε	x	y	$x0$	$y0$	xx	yy	xy	yx	yz
$S_b(nb^2+r)$	7	8	10	7	11	5	9	8	10	12

Table 6: Values of $S_b(nb^2+r)$ for $b \le n < b^2$ with $\mathrm{rep}_b(n) = xx$ and $x, y, z \in \{1, \dots, b-1\}$ pairwise distinct.

Inductive step. Consider $n \ge b^2$ and suppose that the relation (5) holds for all $m < n$. Then $|\mathrm{rep}_b(n)| \ge 3$. Like for the base case, we need to consider several cases according to the form of the base-b expansion of n. More precisely, we need to consider the following five forms, where $u \in \{0, \dots, b-1\}^*$, $x, y, z \in \{1, \dots, b-1\}$, $x \neq z$, and $t \in \{0, \dots, b-1\}$:

$$x00u \quad \text{or} \quad xx0u \quad \text{or} \quad x0yu \quad \text{or} \quad xxyu \quad \text{or} \quad xztu.$$

Let us focus on the first form of $\mathrm{rep}_b(n)$ since the same reasoning can be applied for the other ones. Assume that $\mathrm{rep}_b(n) = x00u$ where $x \in \{1, \dots, b-1\}$ and $u \in \{0, \dots, b-1\}^*$. For all $r \in \{0, \dots, b^2-1\}$, there exist $r_1, r_2 \in \{0, \dots, b-1\}$ such that $\mathrm{val}_b(r_1 r_2) = r$. We have

$$\begin{aligned}
S_b(nb^2+r) &= S_b(\mathrm{val}_b(x00ur_1r_2)) \\
&= 2S_b(\mathrm{val}_b(x0ur_1r_2)) - S_b(\mathrm{val}_b(xur_1r_2)) && \text{(by Lemma 10)} \\
&= a_r\, 2S_b(\mathrm{val}_b(x0u)) + \textstyle\sum_{s=0}^{b-2} c_{r,s}\, 2S_b(\mathrm{val}_b(x0us)) \\
&\quad - a_r S_b(\mathrm{val}_b(xu)) - \textstyle\sum_{s=0}^{b-2} c_{r,s} S_b(\mathrm{val}_b(xus)) && \text{(by induction hypothesis)} \\
&= a_r S_b(\mathrm{val}_b(x00u)) + \textstyle\sum_{s=0}^{b-2} c_{r,s} S_b(\mathrm{val}_b(x00us)) && \text{(by Lemma 10)} \\
&= a_r S_b(n) + \textstyle\sum_{s=0}^{b-2} c_{r,s} S_b(nb+s), && \text{(by Lemma 10)}
\end{aligned}$$

which proves (5).

***b*-regularity.** From the first part of the proof, we directly deduce that the $\mathbb{Z}$-module $\langle \mathcal{K}_b(S_b) \rangle$ is generated by the $(b+1)$ sequences

$$(S_b(n))_{n\ge0}, (S_b(bn))_{n\ge0}, (S_b(bn+1))_{n\ge0}, \dots, (S_b(bn+b-1))_{n\ge0}.$$

We now show that we can reduce the number of generators. To that aim, we prove that

$$S_b(nb+b-1) = (2b-1)S_b(n) - \sum_{s=0}^{b-2} S_b(nb+s) \quad \forall\, n \ge 0. \tag{6}$$

We proceed by induction on $n \ge 0$. As a base case, the proof that (6) holds for each $n \in \{b, \dots, b^2-1\}$ only requires easy computations that are left to the reader (using Table 1). Now consider $n \ge b^2$ and suppose that the relation (6) holds for all $m < n$. Then $|\mathrm{rep}_b(n)| \ge 3$. Mimicking the first induction step of this proof, we need to consider several cases according to the form of the base-b expansion of n. More precisely, we need to consider the following five forms, where $u \in \{0, \dots, b-1\}^*$, $x, y, z \in \{1, \dots, b-1\}$, $x \neq z$, and $t \in \{0, \dots, b-1\}$:

$$x00u \quad \text{or} \quad xx0u \quad \text{or} \quad x0yu \quad \text{or} \quad xxyu \quad \text{or} \quad xztu.$$

$\mathrm{rep}_b(r)$	ε	x	y	z	$x0$	$y0$	$z0$	xx	yy	zz	xy	xz	yx	yz	zx	zy	zt
$S_b(nb^2+r)$	10	13	12	14	13	11	15	10	8	12	12	14	11	12	15	14	16

Table 7: Values of $S_b(nb^2+r)$ for $b \le n < b^2$ with $\mathrm{rep}_b(n) = xy$ and $x, y, z, t \in \{1, \dots, b-1\}$ pairwise distinct.

Let us focus on the first form of $\text{rep}_b(n)$ since the same reasoning can be applied for the other ones. Assume that $\text{rep}_b(n) = x00u$ where $x \in \{1, \ldots, b-1\}$ and $u \in \{0, \ldots, b-1\}^*$. We have

$$\begin{aligned}
S_b(nb+b-1) &= S_b(\text{val}_b(x00u(b-1))) \\
&= 2S_b(\text{val}_b(x0u(b-1))) - S_b(\text{val}_b(xu(b-1))) && \text{(by Lemma 10)} \\
&= (2b-1)\,2S_b(\text{val}_b(x0u)) - \textstyle\sum_{s=0}^{b-2} 2S_b(\text{val}_b(x0us)) \\
&\quad -(2b-1)S_b(\text{val}_b(xu)) + \textstyle\sum_{s=0}^{b-2} S_b(\text{val}_b(xus)) && \text{(by induction hypothesis)} \\
&= (2b-1)S_b(\text{val}_b(x00u)) - \textstyle\sum_{s=0}^{b-2} S_b(\text{val}_b(x00us)) && \text{(by Lemma 10)} \\
&= (2b-1)S_b(n) - \textstyle\sum_{s=0}^{b-2} S_b(nb+s), && \text{(by Lemma 10)}
\end{aligned}$$

which proves (5).

The $\mathbb{Z}$-module $\langle \mathcal{K}_b(S_b) \rangle$ is thus generated by the b sequences

$$(S_b(n))_{n\geq 0}, (S_b(bn))_{n\geq 0}, (S_b(bn+1))_{n\geq 0}, \ldots, (S_b(bn+b-2))_{n\geq 0}.$$

□

Example 17. Let $b = 2$. Using Table 2, Table 3 and Table 4, we find that $a_0 = -1$, $a_1 = 1$, $a_2 = 4$, $a_3 = 5$, $c_{0,0} = 2$, $c_{1,0} = 1$, $c_{2,0} = -1$ and $c_{3,0} = -2$. In this case, there are no $c_{r,s}$ with $s > 0$. Applying Theorem 16 and from (6), we get

$$\begin{aligned}
S_2(2n+1) &= 3S_2(n) - S_2(2n), \\
S_2(4n) &= -S_2(n) + 2S_2(2n), \\
S_2(4n+1) &= S_2(n) + S_2(2n), \\
S_2(4n+2) &= 4S_2(n) - S_2(2n), \\
S_2(4n+3) &= 5S_2(n) - 2S_2(2n)
\end{aligned}$$

for all $n \geq 0$. This result is a rewriting of [16, Theorem 21]. Observe that the third and the fifth identities are redundant: they follow from the other ones.

Example 18. Let $b = 3$. Using Table 2, Table 3 and Table 4, the values of the coefficients a_r, $c_{r,0}$ and $c_{r,1}$ can be found in Table 8. Applying Theorem 16 and from (6), we get

r	0	1	2	3	4	5	6	7	8
a_r	-1	-2	3	-2	-1	3	8	8	9
$c_{r,0}$	2	2	1	1	0	-1	-1	-2	-2
$c_{r,1}$	0	1	-1	2	2	1	-2	-1	-2

Table 8: The values of $a_r, c_{r,0}, c_{r,1}$ when $b = 3$ and $r \in \{0, \ldots, 8\}$.

$$\begin{aligned}
S_3(3n+2) &= 5S_3(n) - S_3(3n) - S_3(3n+1), \\
S_3(9n) &= -S_3(n) + 2S_3(3n), \\
S_3(9n+1) &= -2S_3(n) + 2S_3(3n) + S_3(3n+1), \\
S_3(9n+2) &= 3S_3(n) + S_3(3n) - S_3(3n+1), \\
S_3(9n+3) &= -2S_3(n) + S_3(3n) + 2S_3(3n+1), \\
S_3(9n+4) &= -S_3(n) + 2S_3(3n+1), \\
S_3(9n+5) &= 3S_3(n) - S_3(3n) + S_3(3n+1), \\
S_3(9n+6) &= 8S_3(n) - S_3(3n) - 2S_3(3n+1), \\
S_3(9n+7) &= 8S_3(n) - 2S_3(3n) - S_3(3n+1), \\
S_3(9n+8) &= 9S_3(n) - 2S_3(3n) - 2S_3(3n+1)
\end{aligned}$$

for all $n \geq 0$. This result is a proof of [16, Conjecture 26]. Observe that the fourth, the seventh and the tenth identities are redundant.

Remark 19. Combining (5) and (6) yield b^2+1 identities to generate the $\mathbb{Z}$-module $\langle \mathcal{K}_b(S_b) \rangle$. However, as illustrated in Example 17 and Example 18, only b^2-b+1 identities are useful: the relations established for the sequences $(S_b(b^2n+br+b-1))_{n\geq 0}$, with $r \in \{0,\dots,b-1\}$, can be deduced from the other identities.

Remark 20. Using Theorem 16 and (6) and the set of b generators of the $\mathbb{Z}$-module $\langle \mathcal{K}_b(S_b) \rangle$ being

$$\{(S_b(n))_{n\geq 0}, (S_b(bn))_{n\geq 0}, (S_b(bn+1))_{n\geq 0}, \dots, (S_b(bn+b-2))_{n\geq 0}\},$$

we get matrices to compute $S_b(n)$ in a number of steps proportional to $\log_b(n)$. For all $n \geq 0$, let

$$V_b(n) = \begin{pmatrix} S_b(n) \\ S_b(bn) \\ S_b(bn+1) \\ \vdots \\ S_b(bn+b-2) \end{pmatrix} \in \mathbb{Z}^b.$$

Consider the matrix-valued morphism $\mu_b : \{0,1,\dots,b-1\}^* \to \mathbb{Z}_b^b$ defined, for all $s \in \{0,\dots,b-2\}$, by

$$\mu_b(s) = \begin{pmatrix} 0 & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ a_{bs} & c_{bs,0} & \cdots & c_{bs,s-1} & c_{bs,s} & c_{bs,s+1} & \cdots & c_{bs,b-2} \\ a_{bs+1} & c_{bs+1,0} & \cdots & c_{bs+1,s-1} & c_{bs+1,s} & c_{bs+1,s+1} & \cdots & c_{bs+1,b-2} \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{bs+b-2} & c_{bs+b-2,0} & \cdots & c_{bs+b-2,s-1} & c_{bs+b-2,s} & c_{bs+b-2,s+1} & \cdots & c_{bs+b-2,b-2} \end{pmatrix}$$

and

$$\mu_b(b-1) = \begin{pmatrix} (2b-1) & -1 & -1 & \cdots & -1 \\ a_{b(b-1)} & c_{b(b-1),0} & c_{b(b-1),1} & \cdots & c_{b(b-1),b-2} \\ a_{b(b-1)+1} & c_{b(b-1)+1,0} & c_{b(b-1)+1,1} & \cdots & c_{b(b-1)+1,b-2} \\ \vdots & \vdots & \vdots & & \vdots \\ a_{b(b-1)+b-2} & c_{b(b-1)+b-2,0} & c_{b(b-1)+b-2,1} & \cdots & c_{b(b-1)+b-2,b-2} \end{pmatrix}.$$

Observe that the number of generators explains the size of the matrices above. For each $s \in \{0,\dots,b-2\}$, exactly $b-1$ identities from Theorem 16 are used to define the matrix $\mu_b(s)$. If $s, s' \in \{0,\dots,b-2\}$ are such that $s \neq s'$, then the relations used to define the matrices $\mu_b(s)$ and $\mu_b(s')$ are pairwise distinct. Finally, the first row of the matrix $\mu_b(b-1)$ is (6) and the other rows are $b-1$ identities from Theorem 16, which are distinct from the previous relations. Consequently, $(b-1)(b-1)+b$ identities are used, which corroborates Remark 19.

Using the definition of the morphism μ, we can show that $V_b(bn+s) = \mu_b(s)V_b(n)$ for all $s \in \{0,\dots,b-1\}$ and $n \geq 0$. Consequently, if $\text{rep}_b(n) = n_k \cdots n_0$, then

$$S_b(n) = \begin{pmatrix} 1 & 0 & \cdots & 0 \end{pmatrix} \mu_b(n_0) \cdots \mu_b(n_k) V_b(0).$$

For example, when $b=2$, the matrices $\mu_2(0)$ and $\mu_2(1)$ are those given in [16, Corollary 22]. When $b=3$, we get

$$\mu_3(0) = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 2 & 0 \\ -2 & 2 & 1 \end{pmatrix}, \quad \mu_3(1) = \begin{pmatrix} 0 & 0 & 1 \\ -2 & 1 & 2 \\ -1 & 0 & 2 \end{pmatrix}, \quad \mu_3(2) = \begin{pmatrix} 5 & -1 & -1 \\ 8 & -1 & -2 \\ 8 & -2 & -1 \end{pmatrix}.$$

The class of b-synchronized sequences is intermediate between the classes of b-automatic sequences and b-regular sequences. These sequences were first introduced in [9].

Proposition 21. *The sequence* $(S_b(n))_{n\geq 0}$ *is not* b*-synchronized.*

Proof. The proof is exactly the same as [16, Proposition 24]. □

To conclude this section, the following result proves that the sequence $(S_b(n))_{n\geq 0}$ has a partial palindromic structure as the sequence $(S_2(n))_{n\geq 0}$; see [16]. For instance, the sequence $(S_3(n))_{n\geq 0}$ is depicted in Figure 8 inside the interval $[2\cdot 3^4, 3^5]$.

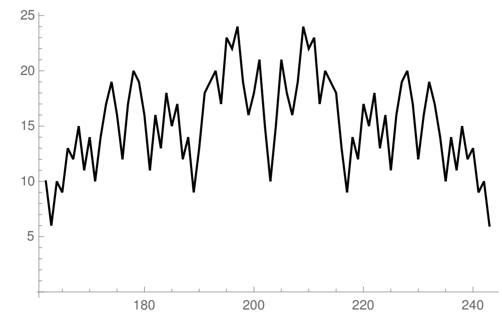

Figure 8: The sequence $(S_3(n))_{n\geq 0}$ inside the interval $[2\cdot 3^4, 3^5]$.

Proposition 22. *Let* u *be a word in* $\{0,1,\dots,b-1\}^*$*. Define* $\bar{u}$ *by replacing in* u *every letter* $a\in\{0,1,\dots,b-1\}$ *by the letter* $(b-1)-a\in\{0,1,\dots,b-1\}$*. Then*

$$\#\left\{v\in L_b \mid \binom{(b-1)u}{v}>0\right\}=\#\left\{v\in L_b \mid \binom{(b-1)\bar{u}}{v}>0\right\}.$$

In particular, there exists a palindromic substructure inside of the sequence $(S_b(n))_{n\geq 0}$*, i.e., for all* $\ell\geq 1$ *and* $0\leq r<b^\ell$*,*

$$S_b((b-1)\cdot b^\ell + r) = S_b((b-1)\cdot b^\ell + b^\ell - r - 1).$$

Proof. The trees $\mathcal{T}((b-1)u)$ and $\mathcal{T}((b-1)\bar{u})$ are isomorphic. Indeed, on the one hand, each node of the form $(b-1)x$ in the first tree corresponds to the node $(b-1)\bar{x}$ in the second one and conversely. On the other hand, if there exist letters $a\in\{1,\dots,b-2\}$ in the word $(b-1)u$, the position of the first letter a in the word $(b-1)u$ is equal to the position of the first letter $(b-1)-a$ in the word $(b-1)\bar{u}$ and conversely. Consequently, the node of the form ax in the first tree corresponds to the node of the form $((b-1)-a)\bar{x}$ in the second tree and conversely.

For the special case, note that for every word z of length ℓ, there exists $r\in\{0,\dots,b^\ell-1\}$ such that $\mathrm{rep}_b((b-1)\cdot b^\ell + r)=(b-1)z$ and

$$\mathrm{val}_b(\bar{z}) = b^\ell - 1 - r \in \{0\dots, b^\ell - 1\}.$$

Hence, $(b-1)\bar{z}=\mathrm{rep}_b((b-1)\cdot b^\ell + b^\ell - 1 - r)$. Using (1), we obtain the desired result. □

4 Asymptotics of the summatory function $(A_b(n))_{n\geq 0}$

In this section, we consider the summatory function $(A_b(n))_{n\geq 0}$ of the sequence $(S_b(n))_{n\geq 0}$; see Definition 1. The aim of this section is to apply the method introduced in [17] to obtain the asymptotic behavior of $(A_b(n))_{n\geq 0}$. As an easy consequence of the b-regularity of $(S_b(n))_{n\geq 0}$, we have the following result.

Proposition 23. *For all* $b\geq 2$*, the sequence* $(A_b(n))_{n\geq 0}$ *is* b*-regular.*

Proof. This is a direct consequence of Theorem 16 and of the fact that the summatory function of a b-regular sequence is also b-regular; see [2, Theorem 16.4.1]. $\square$

From a linear representation with matrices of size $d \times d$ associated with a b-regular sequence, one can derive a linear representation with matrices of size $2d \times 2d$ associated with its summatory function; see [10, Lemma 1]. Consequently, using Remark 20, one can obtain a linear representation with matrices of size $2b\times 2b$ for the summatory function $(A_b(n))_{n\geq 0}$. The goal is to decompose $(A_b(n))_{n\geq 0}$ into linear combinations of powers of $(2b-1)$. We need the following two lemmas.

Lemma 24. *For all $\ell \geq 0$ and all $x \in \{1,\dots,b-1\}$, we have*

$$A_b(xb^\ell) = (2x-1)\cdot(2b-1)^\ell.$$

Proof. We proceed by induction on $\ell \geq 0$. If $\ell = 0$ and $x \in \{1,\dots,b-1\}$, then using Table 1, we have

$$A_b(x) = S_b(0) + \sum_{j=1}^{x-1} S_b(j) = 2x-1.$$

If $\ell = 1$ and $x \in \{1,\dots,b-1\}$, then we have

$$A_b(xb) = A_b(b) + \sum_{y=1}^{x-1}\sum_{j=0}^{b-1} S_b(yb+j).$$

Using Table 1, we get $A_b(xb) = (2x-1)(2b-1)$.

Now suppose that $\ell \geq 1$ and assume that the result holds for all $\ell' \leq \ell$. To prove the result, we again proceed by induction on $x \in \{1,\dots,b-1\}$. When $x = 1$, we must show that $A_b(b^{\ell+1}) = (2b-1)^{\ell+1}$. We have

$$A_b(b^{\ell+1}) = A_b(b^\ell) + \sum_{y=1}^{b-1}\sum_{j=0}^{b^\ell-1} S_b(yb^\ell + j).$$

By decomposing the sum into three parts accordingly to Proposition 3, we get

$$\begin{aligned} A_b(b^{\ell+1}) &= A_b(b^\ell) + \sum_{y=1}^{b-1}\sum_{j=0}^{b^{\ell-1}-1} S_b(yb^\ell + j) + \sum_{y=1}^{b-1}\sum_{j=0}^{b^{\ell-1}-1} S_b(yb^\ell + yb^{\ell-1} + j) \\ &\quad + \sum_{y=1}^{b-1}\sum_{\substack{1\leq z\leq b-1\\ z\neq y}}\sum_{j=0}^{b^{\ell-1}-1} S_b(yb^\ell + zb^{\ell-1} + j), \end{aligned}$$

and, using Proposition 3,

$$\begin{aligned} A_b(b^{\ell+1}) &= A_b(b^\ell) \\ &\quad + \sum_{y=1}^{b-1}\sum_{j=0}^{b^{\ell-1}-1} (S_b(yb^{\ell-1}+j) + S_b(j)) \qquad (7) \\ &\quad + \sum_{y=1}^{b-1}\sum_{j=0}^{b^{\ell-1}-1} (2S_b(yb^{\ell-1}+j) - S_b(j)) \qquad (8) \\ &\quad + \sum_{y=1}^{b-1}\sum_{\substack{1\leq z\leq b-1\\ z\neq y}}\sum_{j=0}^{b^{\ell-1}-1} (S_b(yb^{\ell-1}+j) + 2S_b(zb^{\ell-1}+j) - 2S_b(j)). \qquad (9) \end{aligned}$$

By observing that for all y,

$$\sum_{j=0}^{b^{\ell-1}-1} S_b(yb^{\ell-1}+j) = A_b((y+1)b^{\ell-1}) - A_b(yb^{\ell-1}) \qquad \text{and} \qquad \sum_{j=0}^{b^{\ell-1}-1} S_b(j)) = A_b(b^{\ell-1}), \tag{10}$$

and that

$$\sum_{y=1}^{b-1} \big(A_b((y+1)b^{\ell-1}) - A_b(yb^{\ell-1})\big) = A_b(b^{\ell}) - A_b(b^{\ell-1}), \tag{11}$$

we obtain

$$\begin{aligned}
(7) &= A_b(b^{\ell}) + (b-2)A_b(b^{\ell-1}),\\
(8) &= 2A_b(b^{\ell}) - (b+1)A_b(b^{\ell-1}),\\
(9) &= 3(b-2)(A_b(b^{\ell}) - A_b(b^{\ell-1})) - 2(b-1)(b-2)A_b(b^{\ell-1}) = 3(b-2)A_b(b^{\ell}) - (b-2)(2b+1)A_b(b^{\ell-1}),
\end{aligned}$$

and finally

$$A_b(b^{\ell+1}) = (3b-2)A_b(b^{\ell}) - (2b^2-3b+1)A_b(b^{\ell-1}).$$

Using the induction hypothesis, we obtain

$$A_b(b^{\ell+1}) = (3b-2)(2b-1)^{\ell} - (2b^2-3b+1)(2b-1)^{\ell-1} = (2b-1)^{\ell+1},$$

which ends the case where $x = 1$.

Now suppose that $x \in \{2, \dots, b-1\}$ and assume that the result holds for all $x' < x$. The proof follows the same lines as in the case $x = 1$ with the difference that we decompose the sum into

$$\begin{aligned}
A_b(xb^{\ell+1}) &= A_b((x-1)b^{\ell+1}) + \sum_{j=0}^{b^{\ell+1}-1} S_b((x-1)b^{\ell+1}+j)\\
&= A_b((x-1)b^{\ell+1}) + \sum_{j=0}^{b^{\ell}-1} S_b((x-1)b^{\ell+1}+j) + \sum_{j=0}^{b^{\ell}-1} S_b((x-1)b^{\ell+1}+(x-1)b^{\ell}+j)\\
&+ \sum_{\substack{1\le y\le b-1\\ y\neq x-1}} \sum_{j=0}^{b^{\ell}-1} S_b((x-1)b^{\ell+1}+yb^{\ell}+j).
\end{aligned}$$

Applying Proposition 3 and using (10) and (11) leads to the equality

$$A_b(xb^{\ell+1}) = A_b((x-1)b^{\ell+1}) + (b-1)A_b(xb^{\ell}) - (b-1)A_b((x-1)b^{\ell}) + 2A_b(b^{\ell+1}) - 2(b-1)A_b(b^{\ell}).$$

The induction hypothesis ends the computation. □

Lemma 25. *For all $\ell \ge 1$ and all $x, y \in \{1, \dots, b-1\}$, we have*

$$A_b(xb^{\ell} + yb^{\ell-1}) = \begin{cases} (4xb - 2x + 4y - 2b)\cdot(2b-1)^{\ell-1}, & \text{if } y \le x;\\ (4xb - 2x + 4y - 2b - 1)\cdot(2b-1)^{\ell-1}, & \text{if } y > x.\end{cases}$$

Proof. The proof of this lemma is similar to the proof of Lemma 24 so we only proof the formula for $A_b(xb^{\ell} + xb^{\ell-1})$, the other being similarly handled. We proceed by induction on $\ell \ge 1$. If $\ell = 1$, the result follows from Table 1. Assume that $\ell \ge 2$ and that the formulas hold for all $\ell' < \ell$. We have

$$A_b(xb^{\ell} + xb^{\ell-1}) = A_b(xb^{\ell}) + \sum_{j=0}^{b^{\ell-1}-1} S_b(xb^{\ell}+j) + \sum_{y=1}^{x-1}\sum_{j=0}^{b^{\ell-1}-1} S_b(xb^{\ell}+yb^{\ell-1}+j).$$

Applying Proposition 3 and using (10) and (11) leads to the equality

$$A_b(xb^\ell + xb^{\ell-1}) = A_b(xb^\ell) + xA_b((x+1)b^{\ell-1}) + (2-x)A_b(xb^{\ell-1}) + (1-2x)A_b(b^{\ell-1}).$$

Using Lemma 24 completes the computation. □

Lemma 24 and Lemma 25 give rise to recurrence relations satisfied by the summatory function $(A_b(n))_{n\geq 0}$ as stated below. This is a key result that permits us to introduce $(2b-1)$-decompositions (Definition 28 below) of the summatory function $(A_b(n))_{n\geq 0}$ and allows us to easily deduce Theorem 30; see [17] for similar results in base 2.

Proposition 26. *For all $x, y \in \{1, \ldots, b-1\}$ with $x \neq y$, all $\ell \geq 1$ and all $r \in \{0, \ldots, b^{\ell-1}\}$,*

$$A_b(xb^\ell + r) = (2b-2)\cdot(2x-1)\cdot(2b-1)^{\ell-1} + A_b(xb^{\ell-1}+r) + A_b(r); \quad (12)$$

$$A_b(xb^\ell + xb^{\ell-1} + r) = (4xb - 2x - 2b + 2)\cdot(2b-1)^{\ell-1} + 2A_b(xb^{\ell-1}+r) - A_b(r); \quad (13)$$

$$A_b(xb^\ell + yb^{\ell-1} + r) = \begin{cases} (4xb - 4x - 2b + 3)\cdot(2b-1)^{\ell-1} + A_b(xb^{\ell-1}+r) \\ +2A_b(yb^{\ell-1}+r) - 2A_b(r), & \text{if } y < x; \\ (4xb - 4x - 2b + 2)\cdot(2b-1)^{\ell-1} + A_b(xb^{\ell-1}+r) \\ +2A_b(yb^{\ell-1}+r) - 2A_b(r), & \text{if } y > x. \end{cases} \quad (14)$$

Proof. We first prove (12). Let $x \in \{1, \ldots, b-1\}$, $\ell \geq 1$ and $r \in \{0, \ldots, b^{\ell-1}\}$. If $r = 0$, then (12) holds using Lemma 24. Now suppose that $r \in \{1, \ldots, b^{\ell-1}\}$. Applying successively Proposition 3 and Lemma 24, we have

$$\begin{aligned} A_b(xb^\ell + r) &= A_b(xb^\ell) + \sum_{j=0}^{r-1} S_b(xb^\ell + j) \\ &= A_b(xb^\ell) + \sum_{j=0}^{r-1} (S_b(xb^{\ell-1}+j) + S_b(j)) \\ &= A_b(xb^\ell) + (A_b(xb^{\ell-1}+r) - A_b(xb^{\ell-1})) + A_b(r) \\ &= (2b-2)(2x-1)(2b-1)^{\ell-1} + A_b(xb^{\ell-1}+r) + A_b(r), \end{aligned}$$

which proves (12).

The proof of (13) and (14) are similar, thus we only prove (13). Let $x \in \{1, \ldots, b-1\}$, $\ell \geq 1$ and $r \in \{0, \ldots, b^{\ell-1}\}$. If $r = 0$, then (13) holds using Lemma 25. Now suppose that $r \in \{1, \ldots, b^{\ell-1}\}$. Applying Proposition 3, we have

$$\begin{aligned} A_b(xb^\ell + xb^{\ell-1} + r) &= A_b(xb^\ell + xb^{\ell-1}) + \sum_{j=0}^{r-1} S_b(xb^\ell + xb^{\ell-1} + j) \\ &= A_b(xb^\ell + xb^{\ell-1}) + \sum_{j=0}^{r-1} (2S_b(xb^{\ell-1}+j) - S_b(j)) \\ &= A_b(xb^\ell + xb^{\ell-1}) + 2(A_b(xb^{\ell-1}+r) - A_b(xb^{\ell-1})) - A_b(r). \end{aligned}$$

Using Lemma 24 and Lemma 25, we get

$$\begin{aligned} A_b(xb^\ell + xb^{\ell-1} + r) &= (4xb + 2x - 2b)(2b-1)^{\ell-1} - 2(2x-1)(2b-1)^{\ell-1} + 2A_b(xb^{\ell-1}+r) - A_b(r) \\ &= (4xb - 2x - 2b + 2)(2b-1)^{\ell-1} + 2A_b(xb^{\ell-1}+r) - A_b(r), \end{aligned}$$

which proves (13). □

The following corollary was conjectured in [17].

Corollary 27. *For all $n \geq 0$, we have $A_b(nb) = (2b-1)A_b(n)$.*

Proof. Let us proceed by induction on $n \geq 0$. It is easy to check by hand that the result holds for $n \in \{0, \ldots, b-1\}$. Thus consider $n \geq b$ and suppose that the result holds for all $n' < n$. The reasoning is divided into three cases according to the form of the base-b expansion of n. As a first case, we write $n = xb^\ell + r$ with $x \in \{1, \ldots, b-1\}$, $\ell \geq 1$ and $0 \leq r < b^{\ell-1}$. By Proposition 26, we have

$$\begin{aligned}A_b(nb) - (2b-1)A_b(n) = {} & (2b-2)\cdot(2x-1)\cdot(2b-1)^\ell + A_b(xb^\ell + br) + A_b(br) - (2b-2)\cdot(2x-1)\cdot(2b-1)^\ell \\ & - (2b-1)A_b(xb^{\ell-1}+r) - (2b-1)A_b(r)\end{aligned}$$

We conclude this case by using the induction hypothesis. The other cases can be handled using the same technique. □

Using Proposition 26, we can define $(2b-1)$-decompositions as follows.

Definition 28. Let $n \geq b$. Applying iteratively Proposition 26 provides a unique decomposition of the form

$$A_b(n) = \sum_{i=0}^{\ell_b(n)} d_i(n)\,(2b-1)^{\ell_b(n)-i}$$

where $d_i(n)$ are integers, $d_0(n) \neq 0$ and $\ell_b(n)$ stands for $\lfloor \log_b n \rfloor - 1$. We say that the word

$$d_0(n)\cdots d_{\ell_b(n)}(n)$$

is the $(2b-1)$-*decomposition* of $A_b(n)$. For the sake of clarity, we also write $(d_0(n), \ldots, d_{\ell_b(n)}(n))$. Also notice that the notion of $(2b-1)$-decomposition is only valid for integers in the sequence $(A_b(n))_{n\geq 0}$.

Example 29. Let $b = 3$. Let us compute the 5-decomposition of $A_3(150) = 1665$. We have $\mathrm{rep}_3(150) = 12120$ and $\ell_3(150) = 3$. Applying once Proposition 26 leads to

$$A_3(150) = A_3(3^4 + 2\cdot 3^3 + 15) = 4\cdot 5^3 + A_3(3^3+15) + 2A_3(2\cdot 3^3 + 15) - 2A_3(15). \tag{15}$$

Applying again Proposition 26, we get

$$\begin{aligned}A_3(3^3+15) &= A_3(3^3+3^2+6) = 6\cdot 3^2 + 2A_3(3^2+6) - A_3(6),\\ A_3(2\cdot 3^3+15) &= A_3(2\cdot 3^3+3^2+6) = 13\cdot 3^2 + A_3(2\cdot 3^2+6) + 2A_3(3^2+6) - 2A_3(6),\\ A_3(15) &= A_3(3^2+2\cdot 3^1) = 4\cdot 5^1 + A_3(3^1) + 2A_3(2\cdot 3^1) - 2A_3(0).\end{aligned}$$

Using Proposition 26, we find

$$\begin{aligned}A_3(3^2+6) &= A_3(3^2+2\cdot 3^1) = 4\cdot 5^1 + A_3(3^1) + 2A_3(2\cdot 3^1) - 2A_3(0),\\ A_3(2\cdot 3^2+6) &= A_3(2\cdot 3^2+2\cdot 3^1) = 16\cdot 5^1 + 2A_3(2\cdot 3^1) - A_3(0),\\ A_3(6) &= A_3(2\cdot 3^1) = 12\cdot 5^0 + A_3(2\cdot 3^0) + A_3(0) = 15\cdot 5^0.\end{aligned}$$

Using Lemma 24, we have $A_3(3^1) = 5^1$ and $A_3(2\cdot 3^1) = 3\cdot 5^1$. Plugging all those values together in (15), we finally have

$$A_3(150) = 4\cdot 5^3 + 32\cdot 5^2 + 82\cdot 5^1 - 45\cdot 5^0.$$

The 5-decomposition of $A_3(150)$ is thus $(4, 32, 82, -45)$.

The proof of the next result follows the same lines as the proof of [17, Theorem 1]. Therefore we only sketch it.

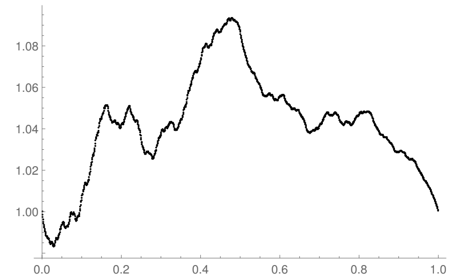

Figure 9: The function $\mathcal{H}_3$ over one period.

Theorem 30. *There exists a continuous and periodic function $\mathcal{H}_b$ of period 1 such that, for all large enough n,*

$$A_b(n) = (2b-1)^{\log_b n}\ \mathcal{H}_b(\log_b n).$$

As an example, when $b = 3$, the function $\mathcal{H}_3$ is depicted in Figure 9 over one period.

Sketch of the proof of Theorem 30. Let us start by defining the function $\mathcal{H}_b$. Given any integer $n \geq 1$, we let ϕ_n denote the function

$$\phi_n(\alpha) = \frac{A_b(e_n(\alpha))}{(2b-1)^{\log_b(e_n(\alpha))}}, \quad \alpha \in [0,1)$$

where $e_n(\alpha) = b^{n+1} + b\lfloor b^n\alpha\rfloor + 1$. With a proof analogous to the one of [17, Proposition 20], the sequence of functions $(\phi_n)_{n\geq 1}$ uniformly converges to a function Φ_b. As in [17, Theorem 5], this function is continuous on $[0,1]$ and such that $\Phi_b(0) = \Phi_b(1) = 1$. Furthermore, it satisfies

$$A_b(b^k + r) = (2b-1)^{\log_b(b^k+r)}\Phi_b\left(\frac{r}{b^k}\right) \quad k \geq 1, 0 \leq r < b^k;$$

see [17, Lemma 24]. Using Corollary 27, we get that, for all $n = b^j(b^k + r)$, $j, k \geq 0$ and $r \in \{0, \ldots, b^k - 1\}$,

$$A_b(n) = (2b-1)^{\log_b(n)}\Phi_b\left(\frac{r}{b^k}\right).$$

The function $\mathcal{H}_b$ is defined by $\mathcal{H}_b(x) = \Phi_b(b^{\{x\}} - 1)$ for all real x ($\{\cdot\}$ stands for the fractional part). □

References

[1] J.-P. Allouche and J. Shallit, The ring of k-regular sequences, *Theoret. Comput. Sci.* **98** (1992), 163–197.

[2] J.-P. Allouche and J. Shallit, *Automatic sequences. Theory, applications, generalizations*, Cambridge University Press, (2003).

[3] J.-P. Allouche and J. Shallit, The ring of k-regular sequences II, *Theoret. Comput. Sci.* **307** (2003), no. 1, 3–29.

[4] B. Bates, M. Bunder, and K. Tognetti, Locating terms in the Stern–Brocot tree, *European J. Combin.* **31** (2010), 1020–1033.

[5] H. Belbachir, L. Németh, and L. Szalay, Hyperbolic Pascal triangles, *Appl. Math. Comput.* **273** (2016), 453–464.

[6] H. Belbachir and L. Szalay, On the arithmetic triangles, *Šiauliai Math. Semin.* **9** (2014), no. 17, 15–26.

[7] N. J. Calkin and H.S. Wilf, Binary partitions of integers and Stern–Brocot-like trees (1998), unpublished. Updated version August 5, 2009, 19 pages, `http://cecas.clemson.edu/~janoski/reu/2010/latestversion.pdf`.

[8] I. Canakci and R. Schiffler, Cluster algebras and continued fractions, unpublished. Updated version September 26, 2016, 28 pages, `https://arxiv.org/pdf/1608.06568.pdf`.

[9] A. Carpi and C. Maggi, On synchronized sequences and their separators, *Theor. Inform. Appl.* **35** (2001), 513–524.

[10] P. Dumas, Joint spectral radius, dilation equations, and asymptotic behavior of radix-rational sequences, *Linear Algebra Appl.* **438** (2013), no. 5, 2107–2126.

[11] T. Garrity, A multidimensional continued fraction generalization of Stern's diatomic sequence, *J. Integer Seq.* **16** (2013), no. 7, Article 13.7.7, 24.

[12] S. P. Glasby, Enumerating the rationals from left to right, *Amer. Math. Monthly* **118** (2011), 830–835.

[13] F. von Haeseler, H.-O. Peitgen, and G. Skordev, Pascal's triangle, dynamical systems and attractors, *Ergod. Th. & Dynam. Sys.* **12** (1992), 479–486.

[14] É Janvresse, T. de la Rue, and Y. Velenik, Self-similar corrections to the ergodic theorem for the Pascal-adic transformation, *Stoch. Dyn.* **5** (2005), no. 1, 1–25.

[15] J. Leroy, M. Rigo, and M. Stipulanti, Generalized Pascal triangle for binomial coefficients of words, *Adv. in Appl. Math.* **80** (2016), 24–47.

[16] J. Leroy, M. Rigo, and M. Stipulanti, Counting the number of non-zero coefficients in rows of generalized Pascal triangles, *Discrete Math.* **340** (2017) 862–881.

[17] J. Leroy, M. Rigo, and M. Stipulanti, Behavior of digital sequences through exotic numeration systems, *Electron. J. Combin.* **24** (2017), no. 1, Paper 1.44, 36 pp.

[18] S. Morier-Genoud, V. Ovsienko, and S. Tabachnikov, $\mathrm{SL}_2(\mathbb{Z})$-tilings of the torus, Coxeter–Conway friezes and Farey triangulations, *Enseign. Math.* **61** (2015), 71–92.

[19] N. J. A. Sloane, The On-Line Encyclopedia of Integer Sequences. Published electronically at `http://oeis.org`, 2017.

2010 *Mathematics Subject Classification*: 11A63, 11B65, 11B85, 41A60, 68R15.
Keywords: Binomial coefficients, subwords, generalized Pascal triangles, base-b expansions, regular sequences, summatory function, asymptotic behavior

---

(Concerned with sequences A007306, A282714, A282715, A282720, A282728, A284441, and A284442.)
