COUNTING SUBWORD OCCURRENCES IN BASE-B EXPANSIONS

Julien Leroy[1]
Department of Mathematics, University of Liège, Allée de la Découverte 12, 4000 Liège, Belgium
J.Leroy@uliege.be

Michel Rigo
Department of Mathematics, University of Liège, Allée de la Découverte 12, 4000 Liège, Belgium
M.Rigo@uliege.be

Manon Stipulanti[2]
Department of Mathematics, University of Liège, Allée de la Découverte 12, 4000 Liège, Belgium
M.Stipulanti@uliege.be

Received: , Revised: , Accepted: , Published:

Abstract

We consider the sequence $(S_b(n))_{n\geq 0}$ counting the number of distinct (scattered) subwords occurring in the base-b expansion of the non-negative integers. 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 $(S_b(n))_{n\geq 0}$.

Jeff Shallit's influence in combinatorics on words cannot be underestimated. We are therefore very happy to contribute to this special issue dedicated to his birthday. This paper contains small bits of the recurrent topics he has been working on.

[1] J. Leroy is an FNRS post-doctoral fellow.
[2] M. Stipulanti is supported by a FRIA grant 1.E030.16.

1. Introduction

A *finite word* is a finite sequence of letters belonging to a finite set called the *alphabet*. Let u, v be two finite words. We say that v is a *scattered subword of* u and we write $v \prec u$, if v is a subsequence of u. All along the paper, we let b denote an integer greater than 1. We let $\operatorname{rep}_b(n)$ denote the (greedy) base-b expansion of $n \in \mathbb{N} \setminus \{0\}$ starting with a non-zero digit. We set $\operatorname{rep}_b(0)$ to be the empty word denoted by ε. We let

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

be the set of base-b expansions of the non-negative integers. For all $w \in \{0, \ldots, b-1\}^*$, we also define $\operatorname{val}_b(w)$ to be the value of w in base b, i.e., if $w = w_n \cdots w_0$ with $w_i \in \{0, \ldots, b-1\}$ for all i, then $\operatorname{val}_b(w) = \sum_{i=0}^{n} w_i b^i$. In this paper, we also make extensive use of the *genealogical order* defined as follows. If $u, v \in \{0, \ldots, b-1\}^*$ are two words, we say *u is less than v in the genealogical order*, and we write $u < v$, if either $|u| < |v|$, or if $|u| = |v|$ and there exist words $p, q, r \in \{0, \ldots, b-1\}^*$ and letters $a, a' \in \{0, \ldots, b-1\}$ with $u = paq$, $v = pa'r$ and $a < a'$. By $u \leq v$, we mean that either $u < v$, or $u = v$.

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

$$S_b(n) := \#\,\{v \in L_b \mid v \prec \operatorname{rep}_b(n)\}\,. \tag{1}$$

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 $n \geq 1$,

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

The quantity $A_b(n)$ can be thought of as the total number of base-b expansions occurring as scattered 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).

Example 1. If $b = 3$, then the first few terms of the sequence[3] $(S_3(n))_{n\geq 0}$ 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, we have $S_3(\operatorname{val}_3(121)) = S_3(16) = 7$. The first few terms of $(A_3(n))_{n\geq 0}$ A284442 are

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

[3]Some of the sequences of this paper are uploaded in [16].

Motivated by generalizations of the Pascal triangle [12], we dealt with the case $b = 2$, considering the sequences $(S_2(n))_{n\geq 0}$ [13] and $(A_2(n))_{n\geq 0}$ [14].

Firstly, using recurrence relations, we showed that the sequence $(S_2(n))_{n\geq 0}$ A007306 is 2-regular in the sense of Allouche and Shallit [1]. We also conjectured six recurrence relations for $(S_3(n))_{n\geq 0}$ depending on the position of n between two consecutive powers of 3. Using the heuristic from [3] suggesting recurrence relations, the sequence $(S_3(n))_{n\geq 0}$ was expected to be 3-regular. In Section 2 of this paper, we show that for all $b \geq 2$, the recurrence relations satisfied by $(S_b(n))_{n\geq 0}$ reduce to three forms; see Proposition 1. In particular, this proves the conjecture stated in [13] for $b = 3$. Then, in Section 3, we deduce the b-regularity of $(S_b(n))_{n\geq 0}$; see Theorem 1. 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 $[(b-1)b^\ell, b^{\ell+1}]$.

Secondly, we studied the behavior of $(A_2(n))_{n\geq 0}$ A282720, exhibiting a continuous periodic function $\mathcal{H}$ of period 1 such that $A_2(n) = 3^{\log_2(n)}\mathcal{H}(\log_2(n))$ for all n. To this aim, from the 2-regularity of $(S_2(n))_{n\geq 0}$, we derived recurrence relations for $(A_2(n))_{n\geq 0}$ involving powers of 3, leading to a particular decomposition of $A_2(n)$. Sustained by computer experiments, we also conjectured that $A_b(nb) = (2b-1)A_b(n)$. In Section 4 of this paper, generalizing the previous approach, we obtain specific recurrence relations for $(A_b(n))_{n\geq 0}$ involving powers of $2b-1$ (Proposition 6) and proving the above conjecture about $A_b(nb)$ (Corollary 1). Using these relations, we consider the so-called $(2b-1)$-decompositions of $A_b(n)$ and exhibit a continuous periodic function $\mathcal{H}_b$ of period 1 such that $A_b(n) = (2b-1)^{\log_b(n)}\mathcal{H}_b(\log_b(n))$ for all n (Theorem 2).

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 $(S_b(n))_{n\geq 0}$. This result is useful to prove that the summatory function of the latter sequence also satisfies recurrence relations; see Section 4.

Proposition 1. *The sequence $(S_b(n))_{n\geq 0}$ satisfies $S_b(0) = 1$,*

$$S_b(1) = \cdots = S_b(b-1) = 2,$$

and, for all $x, y \in \{1, \ldots, b-1\}$ with $x \neq y$, all $\ell \geq 1$ and all $r \in \{0, \ldots, b^{\ell-1}-1\}$,

$$\begin{aligned} S_b(xb^\ell + r) &= S_b(xb^{\ell-1} + r) + S_b(r); && (2)\\ S_b(xb^\ell + xb^{\ell-1} + r) &= 2S_b(xb^{\ell-1} + r) - S_b(r); && (3)\\ S_b(xb^\ell + yb^{\ell-1} + r) &= S_b(xb^{\ell-1} + r) + 2S_b(yb^{\ell-1} + r) - 2S_b(r). && (4) \end{aligned}$$

$\mathrm{rep}_b(n)$	ε	x	$x0$	xx	xy	$x00$	$x0x$	$x0y$
$S_b(n)$	1	2	3	3	4	4	5	6

$\mathrm{rep}_b(n)$	$xx0$	xxx	xxy	$xy0$	xyx	xyy	xyz
$S_b(n)$	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 $x, y, z \in \{1, \dots, b-1\}$.

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.

For the sake of completeness, we recall the definition of a particularly useful tool called *the trie of subwords* to prove Proposition 1. This tool is also useful to prove the b-regularity of the sequence $(S_b(n))_{n\ge 0}$; see Section 3.

Definition 2. 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*[4] *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 1. 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 $\mathcal{T}(\mathrm{rep}_b(n))$ is exactly $S_b(n)$ for all $n \ge 0$.

Definition 3. 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 alternation 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 $\mathrm{Alph}(\ell)$ denote the set of letters occurring in $a_{\ell+1} \cdots a_M$. Then for each letter $a \in \mathrm{Alph}(\ell)$, we let $j(a, \ell)$ denote the smallest index in $\{\ell+1, \dots, M\}$ such that $a_{j(a,\ell)} = a$.

[4]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 2. 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 2 ([13, 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. *Every letter $a \in \mathrm{Alph}(0) \setminus \{0\}$ is a child of the the node of label ε. This node has thus $\#(\mathrm{Alph}(0) \setminus \{0\})$ children. Each child a is the root of a tree isomorphic $T_{j(a,0)-1}$.*

2. *For each $\ell \in \{0, \ldots, M-1\}$ and each $i \in \{0, \ldots, n_{\ell+1}-1\}$ 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 3. Let us continue Example 2. The tree $\mathcal{T}(22000112)$ is depicted in Figure 1. We use three different colors to represent the letters $0, 1, 2$. The tree T_0

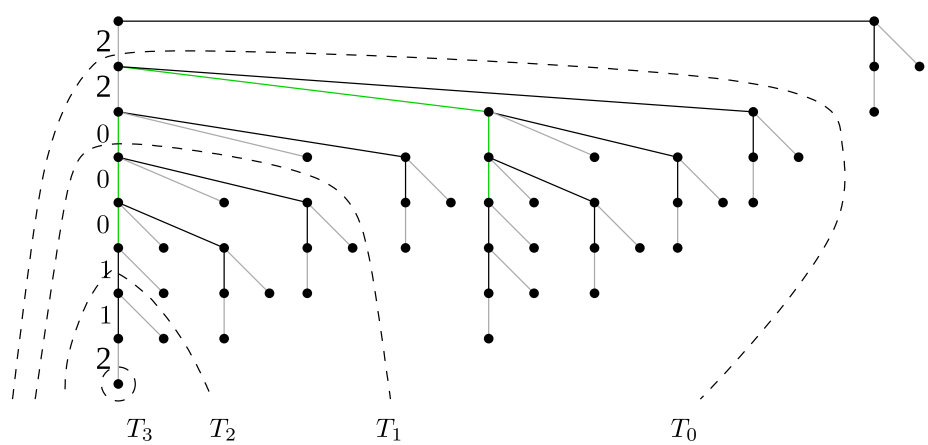


Figure 1: The trie $\mathcal{T}(22000112)$.

(resp., T_1; resp., T_2; resp., T_3) is the subtree of $\mathcal{T}(w)$ with root 2 (resp., 2^20; resp., 2^20^31; resp., $2^20^31^22$). These subtrees are represented in Figure 1 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 1. *For each letter $x \in \{1, \dots, b-1\}$ and each word $u \in \{0, \dots, b-1\}^*$, we have*

$$\# \{v \in L_b \mid v \prec x00u\} = 2 \cdot \# \{v \in L_b \mid v \prec x0u\} - \# \{v \in L_b \mid v \prec xu\}.$$

Proof. Recall that from Remark 1, 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 $x \in \{a_1, \dots, a_m\}$, 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 $xu_i a_i$.

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 2 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

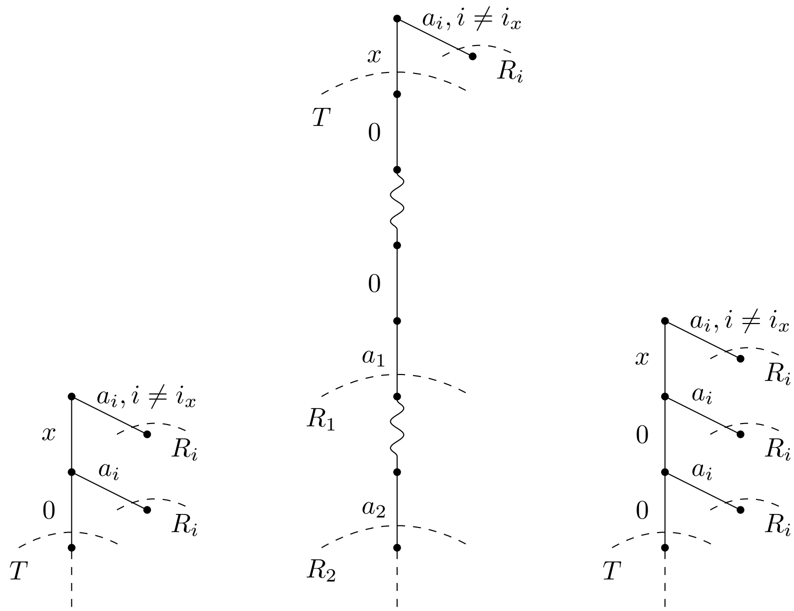
$$\begin{aligned} &2 \cdot \left(2 + \#T + 2 \sum_{\substack{1 \leq i \leq m \\ i \neq i_x}} \#R_i + \#R_{i_x} \right) - \left(1 + \#T + \sum_{\substack{1 \leq i \leq m \\ i \neq i_x}} \#R_i \right) \\ &= 3 + \#T + 3 \sum_{\substack{1 \leq i \leq m \\ i \neq i_x}} \#R_i + 2\#R_{i_x}. \end{aligned}$$

□

Lemma 2. *For each letter $x \in \{1, \dots, b-1\}$ and each word $u \in \{0, \dots, b-1\}^*$, we have*

$$\# \{v \in L_b \mid v \prec xx0u\} = \# \{v \in L_b \mid v \prec x0u\} + \# \{v \in L_b \mid v \prec xu\}.$$

Proof. The proof is similar to the proof of Lemma 1. □



(a) The tree $\mathcal{T}(x0u)$. (b) The tree $\mathcal{T}(xu)$. (c) The tree $\mathcal{T}(x00u)$.

Figure 2: Schematic structure of the trees $\mathcal{T}(x0u)$, $\mathcal{T}(xu)$ and $\mathcal{T}(x00u)$.

Lemma 3. *For all letters $x, y \in \{1, \ldots, b-1\}$ and each word $u \in \{0, \ldots, b-1\}^*$, we have*

$$\#\{v \in L_b \mid v \prec x0yu\} = \#\{v \in L_b \mid v \prec xyu\} + \#\{v \in L_b \mid v \prec yu\}.$$

Proof. The proof is similar to the proof of Lemma 1. 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 $n \geq 0$. □

Lemma 4. *For all letters $x, y \in \{1, \ldots, b-1\}$ and each word $u \in \{0, \ldots, b-1\}^*$, we have*

$$\#\{v \in L_b \mid v \prec xxyu\} = 2 \cdot \#\{v \in L_b \mid v \prec xyu\} - \#\{v \in L_b \mid v \prec yu\}.$$

Proof. The proof is similar to the proof of Lemma 3. □

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

Lemma 5. *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*

$$\begin{aligned} \#\{v \in L_b \mid v \prec xyzu\} \quad = \quad & \#\{v \in L_b \mid v \prec xzu\} + 2 \cdot \#\{v \in L_b \mid v \prec yzu\} \\ & -2 \cdot \#\{v \in L_b \mid v \prec \mathrm{rep}_b(\mathrm{val}_b(zu))\}. \end{aligned}$$

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 whether $z = 0$ or not.

• As a first case, suppose that $z \neq 0$. Then, observe that $\mathrm{rep}_b(\mathrm{val}_b(zu)) = zu$. 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.

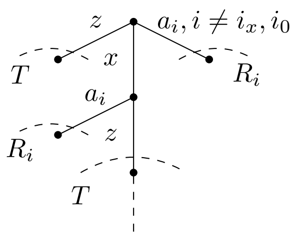
Suppose that $x \neq z$ and $y \neq z$. Using the same reasoning as in the proof of Lemma 1, the situation is depicted in Figure 3. The claimed formula holds since

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

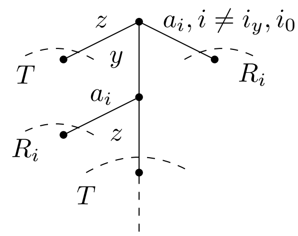
Suppose that $x \neq z$ and $y = z$. The situation is depicted in Figure 4. The claimed formula holds since

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

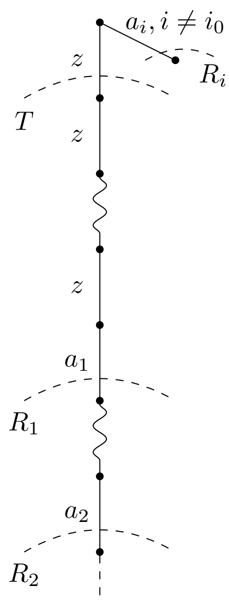
Suppose that $x = z$ and $y \neq z$. The situation is depicted in Figure 5. The



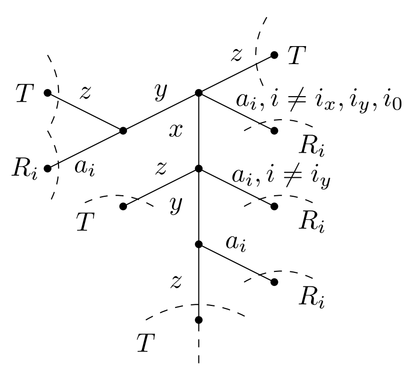
(a) The tree $\mathcal{T}(xzu)$.



(b) The tree $\mathcal{T}(yzu)$.

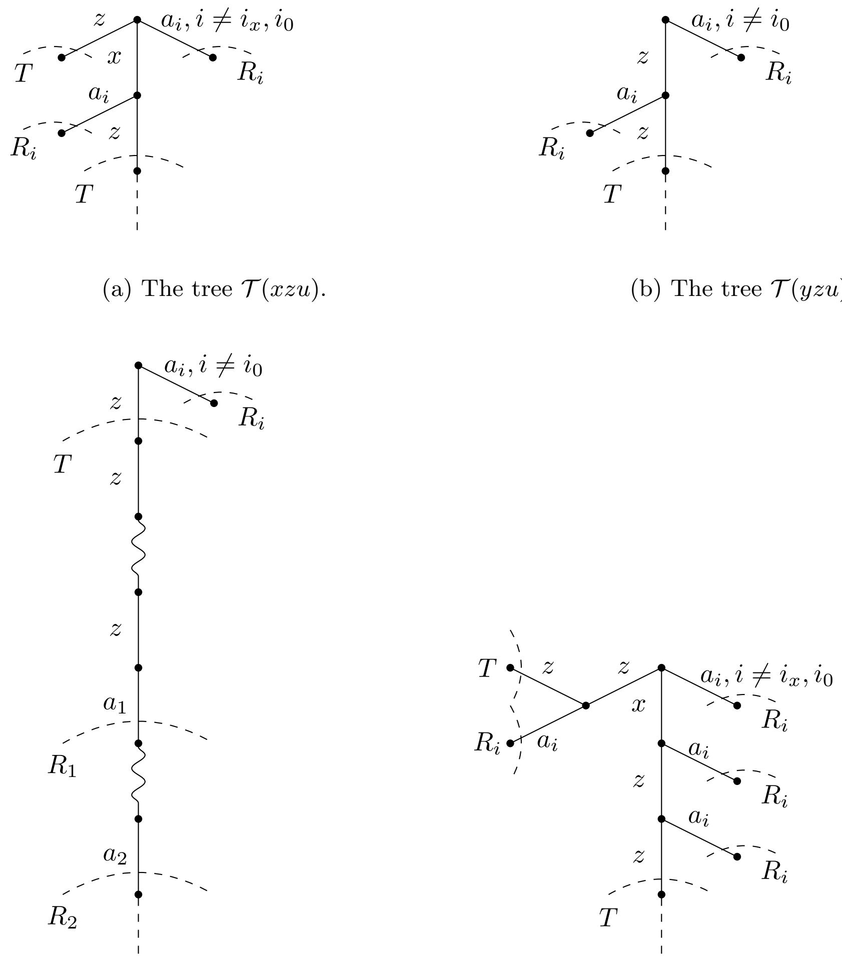


(c) The tree $\mathcal{T}(zu)$.



(d) The tree $\mathcal{T}(xyzu)$.

Figure 3: 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 $z \neq 0$.



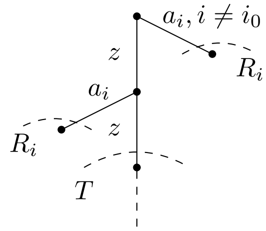
(a) The tree $\mathcal{T}(xzu)$.

(b) The tree $\mathcal{T}(yzu)$.

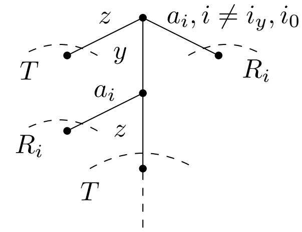
(c) The tree $\mathcal{T}(zu)$.

(d) The tree $\mathcal{T}(xyzu)$.

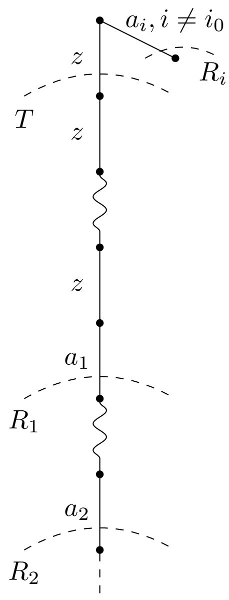
Figure 4: Schematic structure of the trees $\mathcal{T}(xzu)$, $\mathcal{T}(yzu)$, $\mathcal{T}(zu)$ and $\mathcal{T}(xyzu)$ when $x \neq z$, $y = z$ and $z \neq 0$.



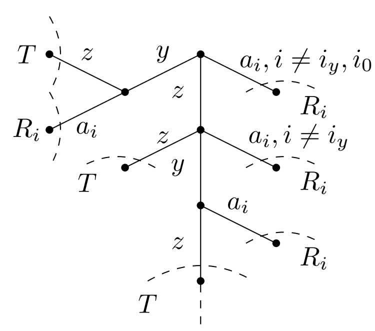
(a) The tree $\mathcal{T}(xzu)$.



(b) The tree $\mathcal{T}(yzu)$.



(c) The tree $\mathcal{T}(zu)$.



(d) The tree $\mathcal{T}(xyzu)$.

Figure 5: 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 $z \neq 0$.

claimed formula holds since

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

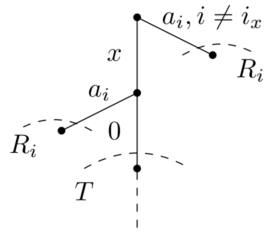
• As a second case, suppose that $z=0$. It is useful to note that $\mathrm{rep}_b(\mathrm{val}_b(\cdot)):\{0,\dots,b-1\}^*\mapsto L_b$ plays a normalization role and removes leading zeroes. Consequently, $\mathrm{rep}_b(\mathrm{val}_b(zu))=\mathrm{rep}_b(\mathrm{val}_b(u))$. Then we must prove that the following formula holds

$$\begin{aligned}
\#\{v\in L_b\mid v\prec xy0u\} &= \#\{v\in L_b\mid v\prec x0u\}+2\cdot\#\{v\in L_b\mid v\prec y0u\}\\
&\quad -2\cdot\#\{v\in L_b\mid v\prec \mathrm{rep}_b(\mathrm{val}_b(u))\}.
\end{aligned}$$

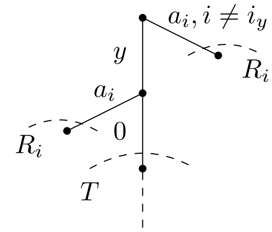
If $u=0^n$, with $n\ge 0$, then $\mathrm{rep}_b(\mathrm{val}_b(u))=\varepsilon$ and the tree $\mathcal{T}(\mathrm{rep}_b(\mathrm{val}_b(u)))$ has only one node. The trees $\mathcal{T}(x0u)$ and $\mathcal{T}(y0u)$ both have $n+3$ nodes and the tree $\mathcal{T}(xy0u)$ has $3(n+2)+1$ nodes and the claimed 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 $\mathrm{rep}_b(\mathrm{val}_b(u))$ that ends with the first occurrence of the letter a_i in $\mathrm{rep}_b(\mathrm{val}_b(u))$, and we let R_i denote the subtree of $\mathcal{T}(\mathrm{rep}_b(\mathrm{val}_b(u)))$ with root $u_i'a_i$.

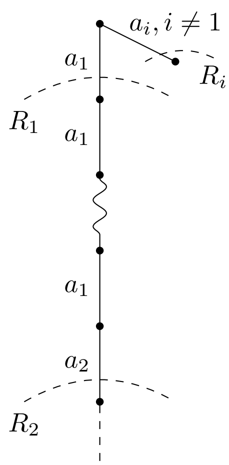
The situation is depicted in Figure 6. Observe that the subtree T of $\mathcal{T}(y0u)$ with root $y0$ is equal to the subtree of $\mathcal{T}(x0u)$ with root $x0$ and to the subtree of



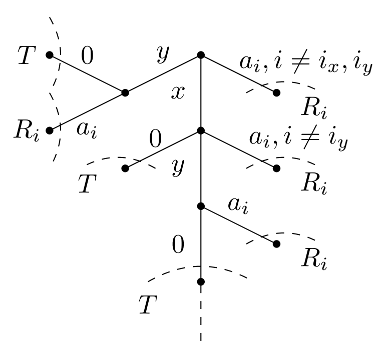
(a) The tree $\mathcal{T}(x0u)$.



(b) The tree $\mathcal{T}(y0u)$.



(c) The tree $\mathcal{T}(\mathrm{rep}_b(\mathrm{val}_b(u)))$.



(d) The tree $\mathcal{T}(xy0u)$.

Figure 6: Schematic structure of the trees $\mathcal{T}(x0u)$, $\mathcal{T}(y0u)$, $\mathcal{T}(\mathrm{rep}_b(\mathrm{val}_b(u)))$ and $\mathcal{T}(xy0u)$.

$\mathcal{T}(xy0u)$ with root $xy0$. The claimed formula holds since

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

□

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

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

Let us first prove (2). If $\ell = 1$, then $r = 0$ and (2) follows from Table 1. Now suppose that $\ell \geq 2$ and assume that (2) holds for all $\ell' < \ell$. Let $r \in \{0, \ldots, b^{\ell-1}-1\}$, and let u be a word in $\{0, \ldots, b-1\}^*$ such that $|u| \geq 1$ 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 $u' \in \{0, \ldots, b-1\}^*$, then, using Lemma 1 and the induction hypothesis twice,

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

which proves (2). Now if $u = zu'$ with $z \in \{1, \ldots, b-1\}$ and $u' \in \{0, \ldots, b-1\}^*$, then (2) directly follows from Definition 1 and Lemma 3.

Let us prove (3). If $\ell = 1$, then $r = 0$ and (2) follows from Table 1. Now suppose that $\ell \geq 2$ and assume that (3) holds for all $\ell' < \ell$. Let $r \in \{0, \ldots, b^{\ell-1} - 1\}$, and let u be a word in $\{0, \ldots, b-1\}^*$ such that $|u| \geq 1$ 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, using Lemma 2 and (2),

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

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 4.

Let us finally prove (4). If $\ell = 1$, then $r = 0$ and (2) follows from Table 1. Now suppose that $\ell \geq 2$ and assume that (4) holds for all $\ell' < \ell$. Let $r \in \{0,\dots,b^{\ell-1}-1\}$, let z be a letter in $\{0,\dots,b-1\}$ and let u be a word in $\{0,\dots,b-1\}^*$ such that $\operatorname{rep}_b(xb^\ell + yb^{\ell-1} + r) = xyzu$. Using Definition 1 and Lemma 5, we directly have that

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

since $\operatorname{rep}_b(r) = \operatorname{rep}_b(\operatorname{val}_b(zu))$, which proves (4). $\square$

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

The sequence $(S_2(n))_{n\geq 0}$ is known to be 2-regular; see [13]. 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 $(t_1(n))_{n\geq 0}, \dots, (t_\ell(n))_{n\geq 0}$ 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 6. *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 $b^3 \times b^3$*

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

More precisely, the matrix M_b is the block-matrix $(B_{i,j})_{1\le i,j\le b}$, where $B_{i,j}$ is the matrix of size $b^2\times b^2$ such that

$$B_{ij}=\begin{cases} I, & \text{if } i=1 \text{ and } j\in\{1,2\};\\ 2I, & \text{if } (i=1 \text{ and } j\ge 3) \text{ or } (j=1 \text{ and } i\ge 2);\\ 3I, & \text{if } (j=2 \text{ and } i\ge 2) \text{ or } (j=i+1\ge 3);\\ 4I, & \text{otherwise.}\end{cases}$$

This matrix is invertible and its inverse is given by

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

More precisely, the matrix M_b^{-1} is the block-matrix $(C_{i,j})_{1\le i,j\le b}$, where $C_{i,j}$ is the matrix of size $b^2\times b^2$ such that

$$C_{ij}=\begin{cases} 3I, & \text{if } i=j=1;\\ -(2b-3)I, & \text{if } i=1 \text{ and } j=b;\\ 2I, & \text{if } i=1 \text{ and } 2\le j<b;\\ -2I, & \text{if } i=2 \text{ and } j=1;\\ I, & \text{if } j=b \text{ and } i\ge 2;\\ -I, & \text{if } i=j+1\ge 2;\\ 0, & \text{otherwise.}\end{cases}$$

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

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

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

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

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

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

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

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

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

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

$\mathrm{rep}_b(r)$	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 $0 \leq r < b^2$ and $1 \leq s \leq b-2$ with $x, y, z \in \{1, \ldots, b-2\}$ pairwise distinct.

$\text{rep}_b(r)$	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 $\text{rep}_b(n) = x0$ and $x, y, z \in \{1, \ldots, b-1\}$ pairwise distinct.

Proof. We proceed by induction on $n \ge 0$. 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 $n \in \{b, \ldots, b^2-1\}$.

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

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

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

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

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 6. By this lemma, the previous system has a unique solution given by $X = M_b^{-1}V$. Consequently, using Lemma 6, we have, for all $r \in \{0, \ldots, b^2-1\}$ and all $s \in \{1, \ldots, b-2\}$,

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

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

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

Inductive step. Consider $n \ge b^2$ and suppose that the relation (5) holds for all $m < n$. Then $|\text{rep}_b(n)| \ge 3$. Like for the base case, we need to consider several

$\mathrm{rep}_b(r)$	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 $x, y, z \in \{1, \ldots, b-1\}$ pairwise distinct.

$\mathrm{rep}_b(r)$	ε	x	y	z	$x0$	$y0$	$z0$	xx	yy
$S_b(nb^2+r)$	10	13	12	14	13	11	15	10	8

$\mathrm{rep}_b(r)$	zz	xy	xz	yx	yz	zx	zy	zt
$S_b(nb^2+r)$	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 $x, y, z, t \in \{1, \ldots, b-1\}$ pairwise distinct.

cases according to the form of the base-b expansion of n. More precisely, we need to consider the following five forms, where $u \in \{0, \ldots, b-1\}^*$, $x, y, z \in \{1, \ldots, b-1\}$, $x \neq z$, and $t \in \{0, \ldots, b-1\}$:

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

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 $x \in \{1, \ldots, b-1\}$ and $u \in \{0, \ldots, b-1\}^*$. For all $r \in \{0, \ldots, b^2-1\}$, there exist $r_1, r_2 \in \{0, \ldots, b-1\}$ such that $\mathrm{val}_b(r_1 r_2) = r$. Using Lemma 1 and the induction hypothesis, we have

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

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

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

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

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

We proceed by induction on $n \geq 0$. As a base case, the proof that (6) holds for each $n \in \{b, \ldots, b^2-1\}$ only requires easy computations that are left to the reader (using Table 1). Now consider $n \geq b^2$ and suppose that the relation (6) holds for all $m < n$. Then $|\operatorname{rep}_b(n)| \geq 3$. 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 $u \in \{0, \ldots, b-1\}^*$, $x, y, z \in \{1, \ldots, b-1\}$, $x \neq z$, and $t \in \{0, \ldots, b-1\}$:

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

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

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

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 4. 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 1 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 [13, Theorem 21]. Observe that the third and the fifth identities are redundant: they follow from the other ones.

Example 5. 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 1 and

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 $r \in \{0, \dots, 8\}$.

from (6), we get

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

for all $n \geq 0$. This result is a proof of [13, Conjecture 26]. Observe that the fourth, the seventh and the tenth identities are redundant.

Remark 2. 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 4 and Example 5, 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 3. Using Theorem 1 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 map $\mu_b:\{0,1,\dots,b-1\}^*\to\mathbb{Z}^{b\times b}$ defined as follows. If $s\in\{0,\dots,b-2\}$, then we set

$$\mu_b(s):=\begin{pmatrix} A(s) & C_0(s) & \cdots & C_{s-1}(s) & C_s(s) & C_{s+1}(s) & \cdots & C_{b-2}(s)\end{pmatrix}$$

where the vectors $A(s),C_0(s),\dots,C_{b-2}(s)\in\mathbb{Z}^b$ are given by

$$\begin{aligned}
A(s)^{\mathsf{T}} &= \begin{pmatrix} 0 & a_{bs} & a_{bs+1} & \cdots & a_{bs+b-2}\end{pmatrix};\\
C_i(s)^{\mathsf{T}} &= \begin{pmatrix} 0 & c_{bs,i} & c_{bs+1,i} & \cdots & c_{bs+b-2,i}\end{pmatrix}\ \forall\, i\in\{0,\dots,b-2\}\setminus\{s\};\\
C_s(s)^{\mathsf{T}} &= \begin{pmatrix} 1 & c_{bs,s} & c_{bs+1,s} & \cdots & c_{bs+b-2,s}\end{pmatrix}.
\end{aligned}$$

We also set

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

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 1 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 1, which are distinct from the previous relations. Consequently, $(b-1)(b-1)+b$ identities are used, which corroborates Remark 2.

Using the definition of the map μ_b, 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 $\operatorname{rep}_b(n)=n_k\cdots n_0$, 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 [13, Corollary 22]. When $b = 3$, we get

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

The class of b-synchronized sequences is intermediate between the classes of b-automatic sequences and b-regular sequences. The map rep_b is extended to $\mathbb{N}\times\mathbb{N}$ as follows. For all $m, n \in \mathbb{N}$, $\operatorname{rep}_b(m, n) := (0^{M-|\operatorname{rep}_b(m)|}\operatorname{rep}_b(m), 0^{M-|\operatorname{rep}_b(n)|}\operatorname{rep}_b(n))$ where $M = \max\{|\operatorname{rep}_b(m)|, |\operatorname{rep}_b(n)|\}$. The idea is that the shortest word is padded with leading zeros to get two words of the same length. A sequence $s = (s(n))_{n\geq 0} \in \mathbb{Z}^{\mathbb{N}}$ is *b-synchronized* if the language $\{\operatorname{rep}_b(n, s(n)) \mid n \in \mathbb{N}\}$ is accepted by some finite automaton reading pairs of digits. These sequences were first introduced in [8].

Proposition 3. *The sequence $(S_b(n))_{n\geq 0}$ is not b-synchronized.*

Proof. The proof is exactly the same as [13, Proposition 24]. $\square$

To conclude this section, the following result shows that the sequence $(S_b(n))_{n\geq 0}$ has a partial palindromic structure. For instance, the sequence $(S_3(n))_{n\geq 0}$ is depicted in Figure 7 inside the interval $[2 \cdot 3^4, 3^5]$.

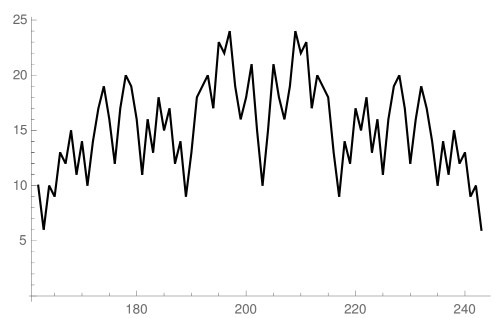


Figure 7: The sequence $(S_3(n))_{n\geq 0}$ inside the interval $[2 \cdot 3^4, 3^5]$.

Proposition 4. *Let u be a word in $\{0, 1, \ldots, b-1\}^*$. Define $\bar{u}$ by replacing in u every letter $a \in \{0, 1, \ldots, b-1\}$ by the letter $(b-1) - a \in \{0, 1, \ldots, b-1\}$. Then*

$$\#\{v \in L_b \mid v \prec (b-1)u\} = \#\{v \in L_b \mid v \prec (b-1)\bar{u}\}.$$

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 [14] 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 5. *For all $b \geq 2$, the sequence $(A_b(n))_{n\geq 0}$ is b-regular.*

Proof. This is a direct consequence of Theorem 1 and of the fact that the summatory function of a b-regular sequence is also b-regular; see [2, Theorem 16.4.1]. □

Remark 4. 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 [9, Lemma 1]. Consequently, using Remark 3, one can obtain a linear representation with matrices of size $2b \times 2b$ for the summatory function $(A_b(n))_{n\geq 0}$.

In order to prove Theorem 2, 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 7. *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 $x \in \{1, \ldots, b-1\}$, 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 $x \in \{1, \ldots, b-1\}$. 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 1, we get

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

and, using Proposition 1,

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

By observing that for all y,

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

$$\sum_{j=0}^{b^{\ell-1}-1} S_b(j)) = A_b(b^{\ell-1}), \qquad (11)$$

$$\sum_{y=1}^{b-1} \big(A_b((y+1)b^{\ell-1}) - A_b(yb^{\ell-1})\big) = A_b(b^\ell) - A_b(b^{\ell-1}), \qquad (12)$$

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

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

Applying Proposition 1 and using (10), (11) and (12) lead to the equality

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

The induction hypothesis ends the computation. □

Lemma 8. *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 7 so we only prove the formula for $A_b(xb^\ell + xb^{\ell-1})$, the others being similarly handled. We proceed by induction on $\ell \geq 1$. If $\ell = 1$, the result follows from Table 1. Assume that $\ell \geq 2$ 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 1 and using (10), (11) and (12) 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 7 completes the computation. $\square$

Lemma 7 and Lemma 8 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 4 below) of the summatory function $(A_b(n))_{n\geq 0}$ and allows us to easily deduce Theorem 2; see [14] for similar results in base 2. It permits us to express $A_b(n)$ as a linear combination of a power of $(2b-1)$ and elements of the form $A_b(m)$ with $m < n$.

Proposition 6. *For all $x, y \in \{1, \ldots, b-1\}$ with $x \neq y$, all $\ell \geq 1$ and all $r \in \{0, \ldots, b^{\ell-1}\}$,*

$$\begin{aligned} A_b(xb^\ell + r) &= (2b-2)\cdot(2x-1)\cdot(2b-1)^{\ell-1} + A_b(xb^{\ell-1} + r) \\ &\quad + A_b(r); \qquad (13) \\ A_b(xb^\ell + xb^{\ell-1} + r) &= (4xb - 2x - 2b + 2)\cdot(2b-1)^{\ell-1} + 2A_b(xb^{\ell-1} + r) \\ &\quad - A_b(r); \qquad (14) \\ A_b(xb^\ell + yb^{\ell-1} + r) &= \begin{cases} (4xb - 4x - 2b + 3)\cdot(2b-1)^{\ell-1} & \\ +A_b(xb^{\ell-1} + r) & \\ +2A_b(yb^{\ell-1} + r) - 2A_b(r), & \text{if } y < x; \\ (4xb - 4x - 2b + 2)\cdot(2b-1)^{\ell-1} & \\ +A_b(xb^{\ell-1} + r) + 2A_b(yb^{\ell-1} + r) & \\ -2A_b(r), & \text{if } y > x. \end{cases} \qquad (15) \end{aligned}$$

Proof. We first 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 7. Now suppose that $r \in \{1, \ldots, b^{\ell-1}\}$.

Applying successively Proposition 1 and Lemma 7, 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 (13).

The proof of (14) and (15) are similar, thus we only prove (14). Let $x \in \{1, \ldots, b-1\}$, $\ell \geq 1$ and $r \in \{0, \ldots, b^{\ell-1}\}$. If $r = 0$, then (14) holds using Lemma 8. Now suppose that $r \in \{1, \ldots, b^{\ell-1}\}$. Applying Proposition 1, 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 7 and Lemma 8, we get

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

which proves (14). □

The following corollary was conjectured in [14].

Corollary 1. *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 6, we have

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

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

Using Proposition 6, we can define $(2b-1)$-decompositions as follows.

Definition 4. Let $n \geq b$. Iteratively applying Proposition 6 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 6. 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 6 leads to

$$A_3(150) = A_3(3^4+2\cdot 3^3+15) = 4\cdot 5^3+A_3(3^3+15)+2A_3(2\cdot 3^3+15)-2A_3(15). \quad (16)$$

Applying Proposition 6 on terms of the form $A_3(m)$ that have just appeared in the r.h.s. of (16), 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 again Proposition 6 on the new terms of the form $A_3(m)$, 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 7, we have $A_3(3^1) = 5^1$ and $A_3(2 \cdot 3^1) = 3 \cdot 5^1$, and the procedure halts. Plugging all those values together in (16), 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 [14, Theorem 1]. Therefore we only sketch it.

Theorem 2. *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 8 over one period.

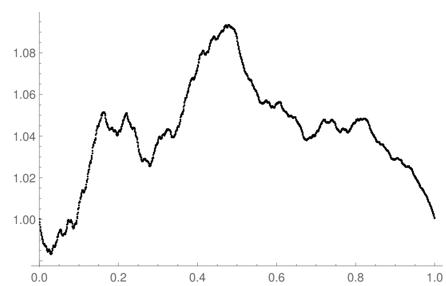


Figure 8: The function $\mathcal{H}_3$ over one period.

Sketch of the proof of Theorem 2. 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 [14, Proposition 20], the sequence of functions $(\phi_n)_{n\geq 1}$ uniformly converges to a function Φ_b. As in [14, Theorem 5], this function is continuous on $[0, 1]$ and such that $\Phi_b(0) = \Phi_b(1) = 1$. Furthermore, it satisfies

$$A_b(b^k + r) = (2b-1)^{\log_b(b^k+r)}\Phi_b\left(\frac{r}{b^k}\right) \quad k \geq 1, 0 \leq r < b^k; \tag{17}$$

see [14, Lemma 24]. Using Corollary 1, 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). $\square$

Remark 5. As stated in [5, Remark 9.2.2], observe that since the 1-periodic function Φ_b is continuous, it is completely defined in the interval $[0, 1]$ by the values taken on the dense set of points of the form r/b^k. Having no error term for these values, see (17), there is no error term in Theorem 2.

Remark 6. The sequence $(S_2(n))_{n\geq 0}$ turns out to be the subsequence with odd indices of the Stern–Brocot sequence [13]. Hence the present paper could motivate the quest for generalized Stern–Brocot sequences and analogues of the Farey tree [4, 6, 7, 10, 11, 15]. Namely can one reasonably define a tree structure, or some other combinatorial structure, in which the sequence $(S_b(n))_{n\geq 0}$ naturally appears?

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] V. Berthé, M. Rigo (Eds.), *Combinatorics, automata and number theory*, Encycl. of Math. and its Appl. **135**, Cambridge University Press, 2010.

[6] 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`.

[7] 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`.

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

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

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

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

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

[13] 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.

[14] 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.

[15] 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.

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