Uniform Distribution Theory **4** (2009), no.2, 81–100

uniform distribution theory

REGULARITIES OF THE DISTRIBUTION OF ABSTRACT VAN DER CORPUT SEQUENCES

Wolfgang Steiner

ABSTRACT. Similarly to β-adic van der Corput sequences, abstract van der Corput sequences can be defined by abstract numeration systems. Under some assumptions, these sequences are low discrepancy sequences. The discrepancy function is computed explicitly, and the bounded remainder sets of the form $[0, y)$ are characterized.

Communicated by Pierre Liardet

1. Introduction

Let $(x_n)_{n\geq 0}$ be a sequence with $x_n \in [0,1)$ for all $n \geq 0$, and

$$D(N, I) = \#\{0 \leq n < N : x_n \in I\} - N\lambda(I)$$

its *discrepancy function* (or *local discrepancy*) on the interval I, where $\lambda(I)$ is the length of I. Then, $(x_n)_{n\geq 0}$ is said to be a *low discrepancy sequence* if $\sup_I D(N, I) = \mathcal{O}(\log N)$, where the supremum is taken over all intervals $I \subseteq [0, 1)$. If $D(N, I)$ is bounded in N, then I is called a *bounded remainder set*. For details on discrepancy, we refer to [KN74, DT97]. References to results on bounded remainder sets can be found in the introduction of [Ste06].

In [BG96, Nin98a, Nin98b], β-adic van der Corput sequences are defined, and it is shown that they are low discrepancy sequences if β is a Pisot number with irreducible β-polynomial. Recall that a *Pisot number* is an algebraic integer greater than 1 with all its conjugates lying in the interior of the unit disk. We refer to Section 3 for the definition of the β-polynomial. In [Mor98, IM04], these results were extended to piecewise linear maps which generalize the

2000 Mathematics Subject Classification: 11K38, 11K31, 11K16, 37B10, 68Q45.
Keywords: Local discrepancy, bounded remainder set, numeration system, automaton, Pisot number, β-transformation, van der Corput sequence, abstract numeration system.
This work was supported by the French Agence Nationale de la Recherche, grant ANR–06-JCJC–0073 "DyCoNum".

β-transformation. The proof in [Nin98a, Nin98b] relies on the fact that cylinder sets of the β-transformation are bounded remainder sets if β is a Pisot number with irreducible β-polynomial. Under the same conditions on β, bounded remainder sets of the form $[0, y)$, $0 \le y \le 1$, were completely characterized in [Ste06]: the β-expansion of y is finite or its tail is the same as that of the expansion of 1.

If β is a Pisot number, then the language of the β-expansions is regular, which means that it is recognized by a finite automaton. Therefore, these β-expansions are special cases of abstract numeration systems as defined in [LR01, LR02], see Section 3. In Section 2, we define van der Corput sequences related to more general abstract numeration systems. Theorem 1 in Section 4 provides a new class of low discrepancy sequences. Finally, Theorem 2 in Section 5 characterizes the bounded remainder sets of the form $[0, y)$ with respect to these abstract van der Corput sequences, generalizing the results in [Ste06].

2. Definitions and first results

Let $(A, \le)$ be a finite and totally ordered alphabet. Denote by A^* the free monoid generated by A for the concatenation product, i.e., the set of finite words with letters in A. The length of a word $w \in A^*$ is denoted by $|w|$. Extend the order on A to A^* by the *shortlex* (or genealogical) *order*, that is to say $v \le w$ if $v = w$ or $v < w$, where $v < w$ means that either $|v| < |w|$ or $|v| = |w|$ and there exist $p, v', w' \in A^*$, $a, b \in A$ such that $v = pav'$, $w = pbw'$ and $a < b$.

According to [LR01], the triple $S = (L, A, \le)$ is an *abstract numeration system* if L is an infinite regular language over A, and the *numerical value* of a word $w \in L$ is defined by

$$\mathrm{val}_S(w) = \#\{v \in L : v < w\}.$$

If $\mathrm{val}_S(w) = n$, then we say that w is the *representation* of n and write

$$\mathrm{rep}_S(n) = w.$$

Denote by A^ω the set of (right) infinite words with letters in A. It is ordered by the *lexicographical order*, that is to say $t \le u$ if $t = u$ or $t < u$, where $t < u$ means that there exist $p \in A^*$, $a, b \in A$, $t', u' \in A^\omega$, such that $t = pat'$, $u = pbu'$ and $a < b$.

Assume that the language L grows exponentially, with

$$\lim_{k\to\infty} \frac{\log \#\{v \in L : |v| \le k\}}{k} = \log \beta > 0.$$

Suppose that $u \in A^\omega$ is the limit of words $w^{(k)} \in L$, i.e., every finite prefix of u is a prefix of $w^{(k)}$ for all but a finite number of k's. Then, the value of u is the real number

$$\mathrm{val}_S^\omega(u) = \lim_{k\to\infty} \frac{\mathrm{val}_S\left(w^{(k)}\right)}{\#\left\{v \in L : \left|v\right| \le \left|w^{(k)}\right|\right\}}, \tag{1}$$

if this limit exists and does not depend on the choice of $w^{(k)}$. Conditions assuring the existence of this value are given in [LR02], see also Lemma 3 and its proof. Let

$$L_\omega = \left\{u \in A^\omega : u = \lim_{k\to\infty} w^{(k)} \text{ for some } w^{(k)} \in L\right\}.$$

Since $\mathrm{val}_S^\omega(u) \in [1/\beta, 1]$, we define the *normalized value*

$$\langle u\rangle = \frac{\beta\,\mathrm{val}_S^\omega(u) - 1}{\beta - 1} \in [0, 1].$$

We extend this definition to finite words $w \in L$ which are prefixes of words in L_ω by setting $\langle w\rangle = \langle u\rangle$, where u is the smallest word in L_ω with prefix w. Since we want to define a sequence without multiple occurrences of the same value, we set

$$L' = \{w \in L : \langle w\rangle \ne \langle v\rangle \text{ for every } v \in L \text{ with } v < w\}.$$

Recall that the *mirror image* of a word $w = w_1w_2\cdots w_k$, $w_j \in A$, is $\widetilde{w} = w_k\cdots w_2w_1$ and that the mirror image of a language L is $\widetilde{L} = \{\widetilde{w} : w \in L\}$. Now, we are ready to define the main object of this paper, abstract van der Corput sequences.

Definition 1 (Abstract van der Corput sequence)**.** Let $S = (L, A, \le)$ be an abstract numeration system, where L is a regular language of exponential growth, every word $w \in L$ is the prefix of some infinite word $u \in L_\omega$, and the limit in (1) exists for every $u \in L_\omega$. Then, the *abstract van der Corput sequence* corresponding to S is given by

$$x_n = \langle w\rangle \quad \text{with} \quad \widetilde{w} = \mathrm{rep}_{\widetilde{S'}}(n),$$

where $\widetilde{S}'$ is the abstract numeration system $\left(\widetilde{L'}, A, \le\right)$.

Thus, the set of values of an abstract van der Corput sequence is

$$\{x_n : n \ge 0\} = \{\langle w\rangle : w \in L\} = \{\langle w\rangle : w \in L'\},$$ and the position of $\langle w\rangle, w \in L'$, in the sequence is determined by the shortlex order on the mirror image of L'. We need a number of further assumptions on the language L in order to get precise formulae for the discrepancy. All these assumptions are satisfied by the β-adic van der Corput sequence when the language of the β-expansions is regular, cf. Section 3, and by Example 1 at the end of this section.

Let $\mathcal{A}_L = (Q, A, \tau, q_0, F)$ be a (complete) deterministic finite automaton recognizing L, with set of states Q, transition function $\tau : Q \times A \to Q$, initial state q_0 and set of final states F. The transition function is extended to words, $\tau : Q \times A^* \to Q$, by setting $\tau(q, \varepsilon) = q$ for the empty word ε and $\tau(q, wa) = \tau(\tau(q, w), a)$. A word $w \in A^*$ is accepted by $\mathcal{A}_L$, and thus in L, if and only if $\tau(q_0, w) \in F$.

Definition 2 (Totally ordered automaton)**.** A deterministic automaton (Q, A, τ, q_0, F) is said to be a *totally ordered automaton* if there exists a total order on the set of states Q such that, for all $q, r \in Q$,

$$q \le r \quad \text{implies} \quad \tau(q, a) \le \tau(r, a) \text{ for every } a \in A.$$

From now on, all automata will be totally ordered automata. Furthermore, the maximal state will be the initial state and every state except the minimal one will be final. (In case $Q = F$, where the automaton recognizes A^*, we add a non-accessible state to Q.) W.l.o.g., the set of states will be $Q = \{0, 1, \dots, d\}$ for some positive integer d, and the order on Q will be the usual order on the integers, hence $q_0 = d$ and $F = \{1, \dots, d\}$. Moreover, we will assume that $\tau(0, a) = 0$ for every $a \in A$, i.e., the state 0 is a sink.

Lemma 1. *Let $L \subseteq A^*$ be recognized by a totally ordered automaton $\mathcal{A}_L = (Q, A, \tau, d, Q \setminus \{0\})$, with $Q = \{0, 1, \dots, d\}$ and $\tau(0, a) = 0$ for every $a \in A$. Then, $\widetilde{L}$ is recognized by the totally ordered automaton $\mathcal{A}_{\widetilde{L}} = (Q, A, \widetilde{\tau}, 0, Q \setminus \{0\})$, where*

$$\widetilde{\tau}(r, a) = \#\big\{q \in Q : \tau(q, a) + r > d\big\} \quad \textit{for every } r \in Q,\ a \in A. \tag{2}$$

In particular, we have $\widetilde{\tau}(0, a) = 0$ for every $a \in A$.

P r o o f. A deterministic automaton $\mathcal{A}' = (Q', A, \tau', q_0', F')$ recognizing $\widetilde{L}$ is obtained by determinizing the automaton which is given by inverting the transition function τ, see e.g. [Sak09]. This means that

- Q' is a subset of the power set $\mathcal{P}(Q)$,
- $\tau'(q', a) = \{q \in Q : \tau(q, a) \in q'\}$ for every $q' \in Q'$, $a \in A$,
- the set of final states in $\mathcal{A}_L$ is the initial state of $\mathcal{A}'$, i.e., $q_0' = \{1, \dots, d\}$,
- the final states in $\mathcal{A}'$ are those elements of Q' which contain the initial state of $\mathcal{A}_L$, i.e., $F' = \{q' \in Q' : d \in q'\}$.

We show that $Q' \subseteq \{q_0', q_1', \dots, q_d'\}$, where $q_r' = \{r + 1, \dots, d\}$ (q_d' being the empty set). Since $\mathcal{A}_L$ is totally ordered and $\tau(0, a) = 0$, we obtain that

$$\tau'(q_0', a) = \{q \in Q : \tau(q, a) > 0\} = \{r + 1, \dots, d\} = q_r' \quad \text{for some } r \in Q.$$

In the same way, we get, for every $r \in Q$ with $q'_r \in Q'$, that

$$\tau'(q'_r, a) = \{q \in Q : \tau(q, a) > r\} = \{s+1, \dots, d\} = q'_s \quad \text{for some } s \in Q.$$

This shows that $Q' \subseteq \{q'_0, q'_1, \dots, q'_d\}$. It is easy to see that $\mathcal{A}'$ is a totally ordered automaton, with the order on Q' given by $q'_r \le q'_s$ if $q'_r \subseteq q'_s$, i.e., $r \ge s$. We clearly have $F' = \{q'_0, q'_1, \dots, q'_{d-1}\} \cap Q'$. If we extend the set of states to $\{q'_0, q'_1, \dots, q'_d\}$ (with possibly non-accessible states) and label the states by $d-r$ instead of q'_r, we obtain $\mathcal{A}_{\widetilde{L}}$. Therefore, $\mathcal{A}_{\widetilde{L}}$ is a totally ordered automaton, with Q ordered by the usual order on the integers. □

The next lemma provides a fundamental characterization of the words in a language L recognized by a totally ordered automaton $\mathcal{A}_L = (Q, A, \tau, d, Q \setminus \{0\})$ with $Q = \{0, 1, \dots, d\}$ and $\tau(0, a) = 0$ for every $a \in A$.

Lemma 2. *Let $L, \tau, \widetilde{\tau}$ be as in Lemma 1, $w_1 \cdots w_k \in A^*$, $0 \le j \le k$. We have $w_1 \cdots w_k \in L$ if and only if $\tau(d, w_1 \cdots w_j) + \widetilde{\tau}(d, w_k \cdots w_{j+1}) > d$.*

Proof. Let $0 \le j \le k$. With the notation of the proof of Lemma 1, $\widetilde{\tau}(d, w_k \cdots w_{j+1}) = d - r$ can be written as

$$\tau'(q'_0, w_k \cdots w_{j+1}) = q'_r = \{r+1, \dots, d\}$$

. In $\mathcal{A}_L$, this means that $w_{j+1} \cdots w_k$ leads to a final state from the state q, i.e., $\tau(q, w_{j+1} \cdots w_k) > 0$, if and only if $q > r$. Therefore, we have

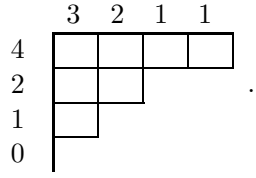
$$\tau(d, w_1 \cdots w_k) = \tau(\tau(d, w_1 \cdots w_j), w_{j+1} \cdots w_k) > 0$$

if and only if $\tau(d, w_1 \cdots w_j) > r$, i.e., $\tau(d, w_1 \cdots w_j) + \widetilde{\tau}(d, w_k \cdots w_{j+1}) > d$. □

Remark 1. If we consider $\tau(d, a) + \cdots + \tau(1, a)$ as a partition of an integer, then $\widetilde{\tau}(d, a) + \cdots + \widetilde{\tau}(1, a)$ is the conjugate partition, since

$$\widetilde{\tau}(d - r, a) = \#\big\{q \in Q : \tau(q, a) > r\big\}.$$

E.g., if $(\tau(d, a), \dots, \tau(1, a)) = (4, 2, 1, 0)$, then $(\widetilde{\tau}(d, a), \dots, \widetilde{\tau}(1, a)) = (3, 2, 1, 1)$, and the corresponding Ferrers diagram is



.

Next, we characterize the values of the abstract van der Corput sequence, under the assumption that the incidence matrix of the co-accessible part of $\mathcal{A}_L$ is primitive. (A state q is co-accessible if $\tau(q, w) \in F$ for some $w \in A^*$.)

Lemma 3. *Let L be as in Lemma 1 and assume that $M_L = (\#\{a \in A : \tau(q,a) = r\})_{1\le q,r\le d}$ is a primitive matrix. Then, the normalized value exists for every $u \in L_\omega$ and is given by*

$$\langle u\rangle = \sum_{j=1}^{\infty} \epsilon_j(u)\beta^{-j} \quad \text{with} \quad \epsilon_j(u) = \sum_{a<u_j} \eta_{\tau(d,u_1\cdots u_{j-1}a)}, \tag{3}$$

where β is the Perron-Frobenius eigenvalue of M_L, $(\eta_1,\dots,\eta_d)^t$ is the corresponding right (column) eigenvector with $\eta_d = 1$, $\eta_0 = 0$, and $u = u_1u_2\cdots$ with $u_j \in A$ for all $j \ge 1$.

P r o o f. The definition and primitivity of the incidence matrix M_L give

$$\#\{v \in A^k : \tau(q,v) > 0\} = (0,\dots,1,\dots,0)M_L^k(1,\dots,1)^t = c\eta_q\beta^k + \mathcal{O}(\rho^k) \tag{4}$$

with constants $c > 0$ and $\rho < \beta$ such that every eigenvalue $\alpha \ne \beta$ of M_L satisfies $|\alpha| < \rho$. Due to the assumptions on $\mathcal{A}_L$, we have $u_1\cdots u_k \in L$ for every $k \ge 1$. Similarly to [LR01, LR02], we split up

$$\begin{aligned}\{v \in L :\, & v < u_1\cdots u_k\}\\ &= \{v \in L :\, |v| < k\} \cup \bigcup_{1\le j\le k}\bigcup_{a<u_j} \left\{u_1\cdots u_{j-1}aw \in L :\, w \in A^{k-j}\right\}.\end{aligned}$$

Since

$$\#\{v \in L :\, |v| \le k\} = \sum_{j=0}^{k}(c\beta^j + \mathcal{O}(\rho^j)) = c\frac{\beta^{k+1}}{\beta-1} + \mathcal{O}(\max(1,\rho)^k),$$

we obtain

$$\frac{\mathrm{val}_S(u_1\cdots u_k)}{\#\{v \in L :\, |v| \le k\}} = \frac{1}{\beta} + \frac{\beta-1}{\beta}\sum_{j=1}^{k}\sum_{a<u_j}\eta_{\tau(d,u_1\cdots u_{j-1}a)}\beta^{-j} + \mathcal{O}\left(\frac{\max(1,\rho)^k}{\beta^k}\right).$$

Therefore, we have

$$\mathrm{val}_S^\omega(u) = \lim_{k\to\infty}\frac{\mathrm{val}_S(u_1\cdots u_k)}{\#\{v \in L :\, |v| \le k\}} = \frac{1}{\beta} + \frac{\beta-1}{\beta}\sum_{j=1}^{\infty}\epsilon_j(u)\beta^{-j},$$

and

$$\langle u\rangle = (\beta\,\mathrm{val}_S^\omega(u) - 1)/(\beta-1)$$

yields (3). □

As a last preparation for the study of the discrepancy of abstract van der Corput sequences, we consider the language L'.

Lemma 4. *Let L be as in Lemma 3 and assume that $\tau(q,a_0) > 0$ for every $q > 0$, where a_0 denotes the smallest letter of A. Then, L' consists exactly of those words in L which do not end with a_0.*

P r o o f. We clearly have $\langle t\rangle \le \langle u\rangle$ if $t < u$, $t, u \in L_\omega$. The primitivity of M_L implies that $\eta_q > 0$ for all $q > 0$. Therefore, $\langle t\rangle = \langle u\rangle$ if and only if there exists no word $u' \in L_\omega$ with $t < u' < u$. For $v < w$, $v, w \in L$, this means that $\langle v\rangle = \langle w\rangle$ if and only if v is a prefix of w and w is the smallest right extension of v in L of length $|w|$. Since $\tau(q, a_0) > 0$ for every $q > 0$, we have $va_0^k \in L$ for all $k \ge 0$. It follows that $\langle v\rangle = \langle w\rangle$ with $v < w$, $v, w \in L$, if and only if $w = va_0^{|w|-|v|}$. Thus, $w \notin L'$ if and only if w ends with a_0. □

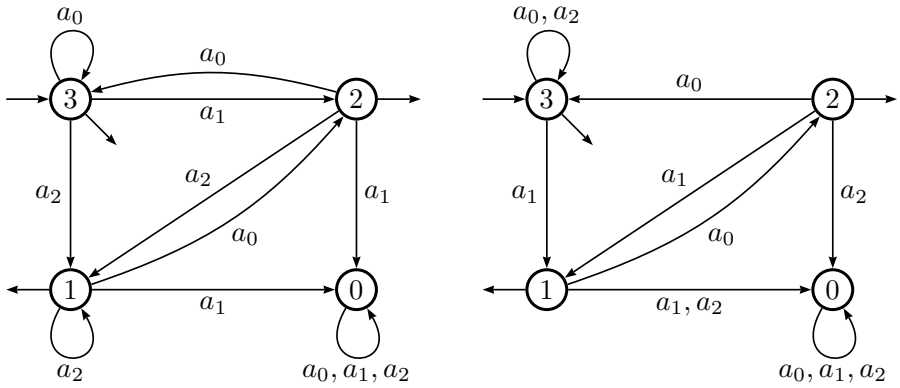


FIGURE 1. The totally ordered automata $\mathcal{A}_L$ (left) and $\mathcal{A}_{\widetilde{L}}$ (right) of Example 1.

EXAMPLE 1. Let $\mathcal{A}_L = (\{0,1,2,3\}, \{a_0, a_1, a_2\}, \tau, 3, \{1,2,3\})$ be the totally ordered automaton in Figure 1 on the left. The first words in L (in the shortlex order) are

$$\varepsilon, a_0, a_1, a_2, a_0a_0, a_0a_1, a_0a_2, a_1a_0, a_1a_2, a_2a_0, a_2a_2,$$
$$a_0a_0a_0, a_0a_0a_1, a_0a_0a_2, a_0a_1a_0, a_0a_1a_2, a_0a_2a_0, a_0a_2a_2,$$
$$a_1a_0a_0, a_1a_0a_1, a_1a_0a_2, a_1a_2a_0, a_1a_2a_2,$$
$$a_2a_0a_0, a_2a_0a_2, a_2a_2a_0, a_2a_2a_2.$$

The transition functions τ, $\widetilde{\tau}$ and the incidence matrices M_L, $M_{\widetilde{L}}$ are given by

τ	a_0	a_1	a_2
0	0	0	0
1	2	0	1
2	3	0	1
3	3	2	1

$$M_L = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix},$$

$$\begin{array}{c|ccc} \widetilde{\tau} & a_0 & a_1 & a_2 \\ \hline 0 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ 2 & 3 & 1 & 0 \\ 3 & 3 & 1 & 3 \end{array}, \quad M_{\widetilde{L}} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 2 \end{pmatrix}.$$

Recall that

$$M_L = (\#\{a \in A : \tau(q,a) = r\})_{1\le q,r\le 3},\ M_{\widetilde{L}} = (\#\{a \in A : \widetilde{\tau}(q,a) = r\})_{1\le q,r\le 3}$$

and that $\widetilde{\tau}$ can be calculated using Remark 1. Thus, $\widetilde{L}$ is recognized by the totally ordered automaton $\mathcal{A}_{\widetilde{L}}$ in Figure 1 on the right. The characteristic polynomial of M_L is x^3-2x^2-x+1, the dominant eigenvalue is $\beta \approx 2.247$, and $(\eta_1,\eta_2,\eta_3)^t = (\beta^2-2\beta, -\beta^2+3\beta-1, 1)^t \approx (0.555, 0.692, 1)^t$ is a right eigenvector of M_L. Since x^3-2x^2-x+1 is irreducible, it must be the characteristic polynomial of $M_{\widetilde{L}}$ as well. The conditions of Lemma 4 are satisfied, hence the first elements of the abstract van der Corput sequence corresponding to L are

$$x_0 = \langle\varepsilon\rangle = 0, \quad x_1 = \langle a_1\rangle = \frac{\eta_3}{\beta}, \quad x_2 = \langle a_2\rangle = \frac{\eta_3+\eta_2}{\beta}, \quad x_3 = \langle a_0a_1\rangle = \frac{\eta_3}{\beta^2},$$

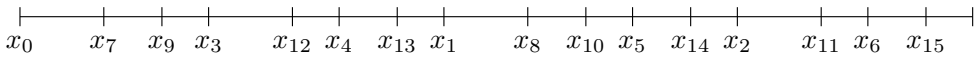
$$x_4 = \langle a_0a_2\rangle = \frac{\eta_3+\eta_2}{\beta^2},\ x_5 = \langle a_1a_2\rangle = \frac{\eta_3}{\beta} + \frac{\eta_3}{\beta^2}, \quad x_6 = \langle a_2a_2\rangle = \frac{\eta_3+\eta_2}{\beta} + \frac{\eta_2}{\beta^2},$$

$$x_7 = \langle a_0a_0a_1\rangle = \frac{\eta_3}{\beta^3}, \quad x_8 = \langle a_1a_0a_1\rangle = \frac{\eta_3}{\beta} + \frac{\eta_3}{\beta^3},\ x_9 = \langle a_0a_0a_2\rangle = \frac{\eta_3+\eta_2}{\beta^3},$$

$$x_{10} = \langle a_1a_0a_2\rangle = \frac{\eta_3}{\beta} + \frac{\eta_3+\eta_2}{\beta^3}, \quad x_{11} = \langle a_2a_0a_2\rangle = \frac{\eta_3+\eta_2}{\beta} + \frac{\eta_3}{\beta^3},$$

$$x_{12} = \langle a_0a_1a_2\rangle = \frac{\eta_3}{\beta^2} + \frac{\eta_3}{\beta^3}, \quad x_{13} = \langle a_0a_2a_2\rangle = \frac{\eta_3+\eta_2}{\beta^2} + \frac{\eta_2}{\beta^3},$$

$$x_{14} = \langle a_1a_2a_2\rangle = \frac{\eta_3}{\beta} + \frac{\eta_3}{\beta^2} + \frac{\eta_2}{\beta^3}, \quad x_{15} = \langle a_2a_2a_2\rangle = \frac{\eta_3+\eta_2}{\beta} + \frac{\eta_2}{\beta^2} + \frac{\eta_2}{\beta^3}.$$



3. β-adic van der Corput sequences

To obtain Ninomiya's β-adic van der Corput sequences, consider a totally ordered automaton $\mathcal{A}_L = (\{0,1,\dots,d\}, A, \tau, d, \{1,\dots,d\})$ on the alphabet $A = \{0,1,\dots,B\}$, with integers $b_q \in A$, $1 \le q \le d$, such that $\tau(q,a) = d$ for all $a < b_q$, and $\tau(q,a) = 0$ if and only if $a > b_q$ or $q = 0$. Assume that M_L is

primitive (which is the only interesting case since $\tau(q,0) \neq d$ implies $b_q = 0$), and let β be its Perron-Frobenius eigenvalue. Then, we have

$$\epsilon_j(u) = \sum_{a<u_j} \eta_{\tau(q,u_1\cdots u_{j-1}a)} = \sum_{a<u_j} \eta_d = u_j \quad \text{for every } u = u_1u_2\cdots \in L_\omega,\ j \geq 1.$$

Let $t_1t_2\cdots$ be the maximal sequence in L_ω, i.e., $t_j = b_{\tau(d,t_1\cdots t_{j-1})}$ for all $j \geq 1$. Since $\mathcal{A}_L$ is a totally ordered automaton, $\tau(d,t_1\cdots t_{j-1}) \leq \tau(d,t_1\cdots t_{k-1})$ implies $t_j < t_k$ or $t_j = t_k$, $\tau(d,t_1\cdots t_j) \leq \tau(d,t_1\cdots t_k)$, thus $t_jt_{j+1}\cdots \leq t_kt_{k+1}\cdots$, in particular $t_jt_{j+1}\cdots \leq t_1t_2\cdots$. By the maximality of $t_1t_2\cdots$, we have $\sum_{j=1}^{\infty} t_j\beta^{-j} = 1$. The sequence $t_1t_2\cdots$ is called the *infinite expansion of* 1 in base β or *quasi-greedy expansion* of 1, cf. [Par60]. By the special structure of $\mathcal{A}_L$ and the primitivity of M_L, we have $\{\tau(d,t_1\cdots t_j) : 0 \leq j < d\} = \{1,\dots,d\}$, hence $\tau(d,t_1\cdots t_d) = \tau(d,t_1\cdots t_m)$ for some $m < d$. Therefore, $t_1t_2\cdots$ is eventually periodic with preperiod length m and period length $d-m$, which implies

$$1 - \sum_{j=1}^{d} t_j\beta^{-j} = \left(1 - \sum_{j=1}^{m} t_j\beta^{-j}\right)\beta^{m-d},$$

thus β is a root of the polynomial

$$(x^d - t_1x^{d-1} - \cdots - t_dx^0) - (x^m - t_1x^{m-1} - \cdots - t_mx^0). \tag{5}$$

If $\mathcal{A}_L$ is the minimal deterministic automaton recognizing L or, equivalently, m and $d-m$ are the minimal preperiod and period lengths of $t_1t_2\cdots$, then (5) is called *β-polynomial*.

If $u = u_1u_2\cdots \in L_\omega$, then we have either $u = t_1t_2\cdots$ or $u_1\cdots u_{k-1} = t_1\cdots t_{k-1}$, $u_k < t_k$ for some $k \geq 1$. Since $\tau(d,u_1\cdots u_k) = d$ in the latter case, $u \in L_\omega$ is equivalent with $u_ju_{j+1}\cdots \leq t_1t_2\cdots$ for all $j \geq 1$, cf. [Par60]. Therefore, u is either the greedy or the quasi-greedy β-expansion of $\langle u\rangle$. Since $w00\cdots$ is a greedy β-expansion for every $w \in L$, the abstract van der Corput sequence given by $S = (L,A,\leq)$ is exactly the β-adic van der Corput sequence defined in [Nin98a]. Consequently, we call $\mathcal{A}_L$ a *β-automaton*.

Conversely, let $t_1t_2\cdots$ be the infinite expansion of 1 in base $\beta > 1$, i.e., the unique sequence of integers satisfying $\sum_{j=1}^{\infty} t_j\beta^{-j} = 1$ and $00\cdots < t_jt_{j+1}\cdots \leq t_1t_2\cdots$ for all $j \geq 1$, cf. [Par60]. Assume that $t_1t_2\cdots$ is eventually periodic. In particular, this holds when β is a Pisot number, see [Ber77, Sch80]. Let d be the sum of the minimal preperiod and period lengths, and $q_j = \#\{1 \leq k \leq d : t_kt_{k+1}\cdots \leq t_{j+1}t_{j+2}\cdots\}$ for $j \geq 0$. Then, the β-automaton is given by $b_{q_j} = t_{j+1}$ and $\tau(q_j,t_{j+1}) = q_{j+1}$, see Example 2.

Note that β-adic van der Corput sequences were first considered in [BG96] for some cases in which $\widetilde{L} = L$. In this case, the definition of the sequence is

simpler beacause the β-expansion of x_n is the mirror image of the expansion of n in a numeration system with respect to the linear recurrence corresponding to the β-polynomial. In our notation, $\widetilde{L} = L$ is equivalent with $\widetilde{\tau} = \tau$, see also [BY00, Kwo09] for a different characterization.

Example 2. Let β be the real root of $x^3 - 4x^2 - 2$. Then, $t_1 t_2 \cdots = 401401\cdots$, hence we have $d = 3$, $q_0 = 3$, $q_1 = 1$, $q_2 = 2$, $q_3 = 3$, thus $b_3 = 4$, $b_1 = 0$, $b_2 = 1$, and

τ	0	1	2	3	4
0	0	0	0	0	0
1	2	0	0	0	0
2	3	3	0	0	0
3	3	3	3	3	1

$$M_L = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 1 & 0 & 4 \end{pmatrix},$$

$\widetilde{\tau}$	0	1	2	3	4
0	0	0	0	0	0
1	2	2	1	1	0
2	3	2	1	1	0
3	3	2	1	1	1

$$M_{\widetilde{L}} = \begin{pmatrix} 2 & 2 & 0 \\ 2 & 1 & 1 \\ 3 & 1 & 1 \end{pmatrix}.$$

This example is also given in Section 2.3 in [Ste06], with different notation. The substitution τ in [Ste06] plays the role of the transition function $\widetilde{\tau}$ in this paper.

4. Discrepancy function

For a given abstract van der Corput sequence $(x_n)_{n\ge0}$ and $y \in [0,1]$, we study now the behavior of the function $D(N,[0,y)) = \#\{0 \le n < N : x_n < y\} - Ny$ as $N \to \infty$. Since $D(N,[y,z)) = D(N,[0,z)) - D(N,[0,y))$, this determines the discrepancy of $(x_n)_{n\ge0}$.

If L satisfies the conditions of Lemma 3, then every $y \in [0,1]$ is the numerical value of some $u \in L_\omega$, see [LR02]. We have the following lemma, where $a_0^\omega = a_0 a_0 \cdots$.

Lemma 5. *Let $(x_n)_{n\ge0}$ be an abstract van der Corput sequence with L as in Lemma 4. For $y = \langle u \rangle$ with $u = u_1 u_2 \cdots \in L_\omega$, $N \ge 0$ with $\mathrm{rep}_{\widetilde{S}'}(N) = w_\ell \cdots w_1$, we have*

$$\#\{0 \le n < N : x_n < y\} = \sum_{j=1}^{\ell} \sum_{a<u_j} \sum_{k=j+1}^{\ell} \sum_{b<w_k} \#L^{k-j-1}_{\tau(d,u_1\cdots u_{j-1}a),\widetilde{\tau}(d,w_\ell\cdots w_{k+1}b)} + C(N,u),$$

where

$$L^m_{q,r} = \{v \in A^m : \tau(q,v) + r > d\} = \{v \in A^m : q + \widetilde{\tau}(r,\widetilde{v}) > d\}$$

and

$$C(N,u) = \sum_{k=1}^{\ell} \#\{b < w_k : bw_{k+1}\cdots w_\ell a_0^\omega < u_k u_{k+1}\cdots,\ u_1\cdots u_{k-1}bw_{k+1}\cdots w_\ell \in L\}.$$

Since $\ell = \mathcal{O}(\log N)$, we have $C(N,u) = \mathcal{O}(\log N)$.

P r o o f. We have to count the number of words $v \in L'$ with $\mathrm{val}_{\widetilde{S'}}(\widetilde{v}) < N$, i.e., $\widetilde{v} < w_\ell \cdots w_1$, and $\langle v\rangle < y$. As in the proof of Lemma 4, we have $va_0^{\ell-|v|} \in L$, thus $\langle v\rangle = \langle va_0^{\ell-|v|}\rangle$. Since $w_\ell > a_0$, $\widetilde{v} < w_\ell \cdots w_1$ holds if and only if $a_0^{\ell-|v|}\widetilde{v} < w_\ell \cdots w_1$. Therefore, we can count the number of words $v \in L \cap A^\ell$ with $\widetilde{v} < w_\ell \cdots w_1$ and $\langle v\rangle < y$, instead of those in L'.

The inequality $\langle v\rangle < y$ is equivalent with $va_0^\omega < u$. Thus, we have to count the v in L of the form

$$v = u_1\cdots u_{j-1}av_{j+1}\cdots v_{k-1}bw_{k+1}\cdots w_\ell \ \text{ with } \ a < u_j, b < w_k, 1 \le j < k \le \ell,$$

or

$$v = u_1\cdots u_{k-1}bw_{k+1}\cdots w_\ell \ \text{ with } \ b < w_k, bw_{k+1}\cdots w_\ell a_0^\omega < u_k u_{k+1}\cdots, 1 \le k \le \ell.$$

Lemma 2 yields that $u_1\cdots u_{j-1}av_{j+1}\cdots v_{k-1}bw_{k+1}\cdots w_\ell \in L$ if and only if $v_{j+1}\cdots v_{k-1} \in L^{k-j-1}_{\tau(d,u_1\cdots u_{j-1}a),\widetilde{\tau}(d,w_\ell\cdots w_{k+1}b)}$, which provides the main part of the formula. The other words give $C(N,u)$. We have $\ell = \mathcal{O}(\log N)$ since L grows exponentially and the same holds for L' by Lemma 4. Since A is finite, we obtain

$$C(N,u) = \mathcal{O}(\log N). \qquad \square$$

Similarly to [Nin98a, Nin98b, Ste06], we assume now that the characteristic polynomial of M_L is irreducible. Let $\beta_2, \dots, \beta_d$ be the conjugates of the Perron-Frobenius eigenvalue $\beta_1 = \beta$. For any $z \in \mathbb{Q}(\beta)$, denote by $z^{(i)} \in \mathbb{Q}(\beta_i)$ the image of z by the isomorphism mapping β to β_i. Similarly to (4), we have some constants $\theta_1, \dots, \theta_d \in \mathbb{Q}(\beta)$ such that

$$\#L^k_{q,r} = (0,\dots,1,\dots,0)M_L^k(0,\dots,0,1,\dots,1)^t = \sum_{i=1}^{d} \eta_q^{(i)}\theta_r^{(i)}\beta_i^k \quad \text{for } 1 \le q,r \le d.$$

Note that $(\theta_1,\dots,\theta_d)^t$ is a right eigenvector of $M_{\widetilde{L}}$. Set $\theta_0 = 0$ and define

$$\gamma_k(N) = \sum_{b<w_k} \theta_{\widetilde{\tau}(d,w_\ell\cdots w_{k+1}b)} \quad \text{for } N = \mathrm{val}_{\widetilde{S'}}(w_\ell\cdots w_1),\ 1 \le k \le \ell, \tag{6}$$

similarly to $\epsilon_j(u)$. Then,

$$\sum_{a<u_j}\sum_{b<w_k} \#L^{k-j-1}_{\tau(d,u_1\cdots u_{j-1}a),\widetilde{\tau}(d,w_\ell\cdots w_{k+1}b)} = \sum_{i=1}^{d} \epsilon_j^{(i)}(u)\gamma_k^{(i)}(N)\beta_i^{k-j-1},$$

$$N = \sum_{k=1}^{\ell} \sum_{b<w_k} \#L^{k-1}_{d,\widetilde{\tau}(d,w_\ell\cdots w_{k+1}b)} = \sum_{k=1}^{\ell} \sum_{b<w_k} \sum_{i=1}^{d} \eta_d^{(i)} \theta^{(i)}_{\widetilde{\tau}(d,w_\ell\cdots w_{k+1}b)} \beta_i^{k-1}$$
$$= \sum_{i=1}^{d} \sum_{k=1}^{\ell} \gamma_k^{(i)}(N) \beta_i^{k-1}.$$

For $y = \langle u \rangle$, we have thus

$$\begin{aligned}
&D(N,[0,y)) \\
&= \sum_{j=1}^{\ell} \sum_{a<u_j} \sum_{k=j+1}^{\ell} \sum_{b<w_k} \#L^{k-j-1}_{\tau(d,u_1\cdots u_{j-1}a),\widetilde{\tau}(d,w_\ell\cdots w_{k+1}b)} + C(N,u) - N \sum_{j=1}^{\infty} \epsilon_j(u) \beta^{-j} \\
&= \sum_{j=1}^{\infty} \sum_{i=1}^{d} \left(\sum_{k=j+1}^{\ell} \epsilon_j^{(i)}(u) \gamma_k^{(i)}(N) \beta_i^{k-j-1} - \sum_{k=1}^{\ell} \gamma_k^{(i)}(N) \beta_i^{k-1} \epsilon_j(u) \beta^{-j} \right) + C(N,u) \\
&= \sum_{j=1}^{\infty} \left(\sum_{k=j+1}^{\ell} \sum_{i=2}^{d} \gamma_k^{(i)}(N) \beta_i^{k-1} \left(\epsilon_j^{(i)}(u) \beta_i^{-j} - \epsilon_j(u) \beta^{-j} \right) \right. \\
&\qquad \left. - \sum_{k=1}^{\min(j,\ell)} \sum_{i=1}^{d} \gamma_k^{(i)}(N) \beta_i^{k-1} \epsilon_j(u) \beta^{-j} \right) + C(N,u). \quad (7)
\end{aligned}$$

Since the series converges absolutely, we can change the order of summation, and get

$$\begin{aligned}
D(N,[0,y)) = \sum_{k=1}^{\ell} &\left(\sum_{j=1}^{k-1} \sum_{i=2}^{d} \gamma_k^{(i)}(N) \beta_i^{k-1} \left(\epsilon_j^{(i)}(u) \beta_i^{-j} - \epsilon_j(u) \beta^{-j} \right) \right. \\
&\left. - \sum_{j=k}^{\infty} \sum_{i=1}^{d} \gamma_k^{(i)}(N) \beta_i^{k-1} \epsilon_j(u) \beta^{-j} \right) + C(N,u). \quad (8)
\end{aligned}$$

The following theorem states that $D(N,[0,y)) = \mathcal{O}(\log N)$ if $|\beta_i| < 1$ for $2 \le i \le d$, i.e., if β is a Pisot number. The conditions on M_L are subsumed in the following definition.

Definition 3 (Pisot automaton)**.** A deterministic automaton is said to be a *Pisot automaton* if the incidence matrix of its restriction to the states which are both accessible and co-accessible has one simple eigenvalue $\beta > 1$, and all other eigenvalues satisfy $|\alpha| < 1$.

Theorem 1. *Let $S = (L, A, \le)$ be an abstract numeration system where L is recognized by a totally ordered Pisot automaton $\mathcal{A}_L = (\{0,1,\dots,d\}, A, \tau, d, \{1,\dots,d\})$ with $\tau(0,a) = 0$ for every $a \in A$, $\tau(q,a_0) > 0$ for the minimal letter $a_0 \in A$ and every $q > 0$. Then, the corresponding abstract van der Corput sequence is a low discrepancy sequence.*

P r o o f. If $\mathcal{A}_L$ is a Pisot automaton, then the characteristic polynomial of M_L is irreducible, thus M_L is primitive. Therefore, the conditions of Lemma 5 are satisfied, and $D(N,[0,y))$ is given by (8). Since $\gamma_k(N)$ and $\epsilon_j(u)$ take only finitely many different values and $\ell = \mathcal{O}(\log N)$, we obtain

$$D(N,[0,y)) = \sum_{k=1}^{\ell} \mathcal{O}(1) + \mathcal{O}(\log N) = \mathcal{O}(\log N),$$

where the constants implied by the $\mathcal{O}$-symbols do not depend on y. With

$$D(N,[y,z)) = D(N,[0,z)) - D(N,[0,y)),$$

we get $\sup_I D(N,I) = \mathcal{O}(\log N)$, and the theorem is proved. □

As Section 3 shows, the β-adic van der Corput sequences defined by Pisot numbers β with irreducible β-polynomial considered in [Nin98a, Nin98b] are special cases of the abstract van der Corput sequences in Theorem 1. By Lemma 2, $\widetilde{S} = (\widetilde{L}, A, \leq)$ is another, usually different, abstract numeration system satisfying the conditions of Theorem 1 whenever $\widetilde{\tau}(q,a_0) > 0$ for every $q > 0$, in particular when $\mathcal{A}_L$ is a β-automaton. The abstract van der Corput sequence considered in Example 1 is also a new low discrepancy sequence.

5. Bounded remainder sets

Under the conditions of Theorem 1, (7) gives

$$D(N,[0,y)) = \sum_{j=1}^{\infty} \left(\sum_{k=j+1}^{\ell} \sum_{i=2}^{d} \gamma_k^{(i)}(N) \epsilon_j^{(i)}(u) \beta_i^{k-j-1} - \sum_{k=1}^{\min(j,\ell)} \gamma_k(N) \epsilon_j(u) \beta^{k-j-1} \right) + C(N,u) + \mathcal{O}(1) \tag{9}$$

for $y = \langle u \rangle$. If there exists some $m \geq 0$ such that $u_j = a_0$ for all $j > m$, then $\epsilon_j(u) = 0$ for all $j > m$ and $C(N,u)$ is bounded. It follows that the interval $[0, \langle v \rangle) = [0, \langle v a_0^{\omega} \rangle)$ is a bounded remainder set for every finite word $v \in L$.

For any β-adic van der Corput sequence where β is a Pisot number with irreducible β-polynomial, the bounded remainder sets $[0,y)$ have been characterized in [Ste06] by the fact that the tail of the β-expansion of y is either 0^{ω} or a suffix of $t_1 t_2 \cdots$. With our notation and the β-automaton defined in Section 3, this means that $\sum_{j=m+1}^{\infty} u_j \beta^{m-j} = \eta_q$ for some $m \geq 0$, $q \in Q$. In the more general case, we have the following partial characterization. Note that the proof is simpler than the one in [Ste06].

Proposition 1. Let $(x_n)_{n\ge 0}$ be an abstract van der Corput sequence satisfying the conditions of Theorem 1, $u \in L_\omega$. If there exists some $m \ge 0$, $q \in Q$, such that

$$\epsilon_{m+1}(u)\epsilon_{m+2}(u)\cdots = \epsilon_{q,1}(t_q)\epsilon_{q,2}(t_q)\cdots,$$

where $t_q = t_{q,1}t_{q,2}\cdots$ is the maximal sequence in A^ω with $\tau(q, t_{q,1}\cdots t_{q,j}) > 0$ for all $j \ge 1$ if $q > 0$, $t_0 = a_0^\omega$, $\epsilon_{q,j}(t_q) = \sum_{a<t_{q,j}} \eta_{\tau(q,t_{q,1}\cdots t_{q,j-1}a)}$, then $D(N, [0, \langle u\rangle))$ is bounded.

P r o o f. If $q = 0$, then $\epsilon_j(u) = 0$ for every $j > m$, and we have already seen that the boundedness of $D(N, [0, \langle u\rangle))$ follows from (9).

For $q > 0$, note that

$$\epsilon_{q,2}(t_q)\epsilon_{q,3}(t_q)\cdots = \epsilon_{q',1}(t_{q'})\epsilon_{q',2}(t_{q'})\cdots,$$

where $q' = \tau(q, t_{q,1})$, and that M_L is primitive. Therefore, we can assume that m is large enough such that M_L^m has only positive entries. Hence, there exists some $v \in L \cap A^m$ such that $\tau(d, v) = q$. Then, we have $vt_q \in L_\omega$ and

$$\epsilon_{m+1}(vt_q)\epsilon_{m+2}(vt_q)\cdots = \epsilon_{q,1}(t_q)\epsilon_{q,2}(t_q)\cdots,$$

thus $D\big(N, [0, \langle u\rangle)\big) = D\big(N, [0, \langle vt_q\rangle)\big) + \mathcal{O}(1)$ by (9).

It only remains to show that $D\big(N, [0, \langle vt_q\rangle)\big) = \mathcal{O}(1)$. If $vt_q = t_d$, then

$$\langle vt_q\rangle = 1 \quad \text{and} \quad D\big(N, [0, \langle vt_q\rangle)\big) = 0.$$

Otherwise, the successor v' of v in L has length m, and $\langle vt_q\rangle = \langle v'\rangle$ since no sequence $u' \in L_\omega$ satisfies $vt_q < u' < v'a_0^\omega$, hence

$$D\big(N, [0, \langle u\rangle)\big) = D\big(N, [0, \langle vt_q\rangle)\big) + \mathcal{O}(1) = D\big(N, [0, \langle v'\rangle\big) + \mathcal{O}(1) = \mathcal{O}(1). \quad \square$$

Clearly $[y, z)$ is a bounded remainder set if both $[0, y)$ and $[0, z)$ have this property. For the converse, the following holds.

Proposition 2. Let $(x_n)_{n\ge 0}$ be an abstract van der Corput sequence satisfying the conditions of Theorem 1. If $D(N, I)$ is bounded, then $\lambda(I) \in \mathbb{Q}(\beta)$.

P r o o f. The proof is very similar to that of Theorem 1 in [Ste06]. Therefore, we only give the main steps. Define a substitution $q \mapsto \widetilde{\tau}(q, a_{q,0})\cdots\widetilde{\tau}(q, a_{q,m_q})$, with $\{a_{q,0}, a_{q,1}, \ldots, a_{q,m_q}\} = \{a \in A : \widetilde{\tau}(q, a) > 0\}$ and $a_{q,0} < a_{q,1} < \cdots < a_{q,m_q}$, $1 \le q \le d$, which plays the role of the substitution τ in [Ste06]. Since $\widetilde{\tau}(d, a_0) = d$, we have $d \mapsto dw$ for some $w \in A^*$. Then, a continuous successor function on L_ω satisfying $\mathrm{rep}_{\widetilde{S}'}(n)a_0^\omega \mapsto \mathrm{rep}_{\widetilde{S}'}(n+1)a_0^\omega$ is topologically conjugate to the successor function S on $\mathcal{D}$ defined in [Ste06], see also [BR07]. If $D(N, I)$ is bounded, then $\lambda(I)$ is an eigenvalue of the dynamical system $(\mathcal{D}, S)$, see Theorem 5.1 in [Sha78] and [Ste06]. By Proposition 5 in [FMN96], these eigenvalues are in $\mathbb{Q}(\beta)$. $\square$

If $y \in \mathbb{Q}(\beta)$, then $y = \langle u \rangle$ for some eventually periodic sequence u, see [RS05]. Let p be the period length of $\epsilon_1(u)\epsilon_2(u)\cdots$ and m the preperiod length. From (8), we get that

$$D(N,[0,y)) = \sum_{k=1}^{\ell}\left(\sum_{i=2}^{d}\gamma_k^{(i)}(N)\sum_{j=1}^{k-1}\epsilon_j^{(i)}(u)\beta_i^{k-j-1} - \gamma_k(N)\sum_{j=k}^{\infty}\epsilon_j(u)\beta^{k-j-1}\right) + C(N,u) + \mathcal{O}(1).$$

Set

$$y_k = \sum_{j=k}^{\infty}\epsilon_j(u)\beta^{k-j-1}.$$

For $k > m$, we have

$$y_k = \frac{\epsilon_k(u)\beta^{p-1} + \cdots + \epsilon_{k+p-1}(u)}{\beta^p - 1},$$

$$\sum_{j=1}^{k-1}\epsilon_j^{(i)}(u)\beta_i^{k-j-1} = \frac{\epsilon_{k-p}^{(i)}(u)\beta_i^{p-1} + \cdots + \epsilon_{k-1}^{(i)}(u)}{1-\beta_i^p} + \mathcal{O}(\beta_i^k) = -y_k^{(i)} + \mathcal{O}(\beta_i^k)$$

for $2 \le i \le d$, which gives

$$D(N,[0,y)) = C(N,u) - \sum_{k=1}^{\ell}\sum_{i=1}^{d}\gamma_k^{(i)}(N)y_k^{(i)} + \mathcal{O}(1).$$

Recall the definition of $\gamma_k(N)$ in (6), and set

$$\zeta_r(z) = \sum_{i=1}^{d}\theta_r^{(i)}z^{(i)} \quad \text{for } r \in Q,\ z \in \mathbb{Q}(\beta). \tag{10}$$

Then, we have

$$D(N,[0,y)) = C(N,u) - \sum_{k=1}^{\ell}\sum_{b<w_k}\zeta_{\widetilde{\tau}(d,w_\ell\cdots w_{k+1}b)}(y_k) + \mathcal{O}(1). \tag{11}$$

Note that

$$\zeta_r(\eta_q) = \#L_{q,r}^0 = \begin{cases} 1 & \text{if } q+r > d, \\ 0 & \text{else.} \end{cases} \tag{12}$$

The main result of this section is the following theorem.

Theorem 2. *Let $(x_n)_{n\ge 0}$ be an abstract van der Corput sequence defined by an abstract numeration system $S = (L, A, \le)$, where L is recognized by a totally ordered Pisot automaton $\mathcal{A}_L = (\{0,1,\ldots,d\}, A, \tau, d, \{1,\ldots,d\})$ with $\tau(0,a) = 0$ for every $a \in A$, $\tau(q,a_0) > q$ for the minimal letter $a_0 \in A$ and every $q \in \{1,\ldots,d-1\}$. Then, $D(N,[0,\langle u\rangle))$, $u \in L_\omega$, is bounded in N if and only if $\langle u\rangle \in \mathbb{Q}(\beta)$ and there exists some $m \ge 0$ such that*

$$\zeta_{\widetilde{\tau}(d,\widetilde{v})}(y_k) = \begin{cases} 1 & \text{if } va_0^\omega < u_k u_{k+1}\cdots \text{ and } u_1\cdots u_{k-1}v \in L, \\ 0 & \text{else,} \end{cases} \tag{13}$$

for every $v \in A^*$, $k > m$, *where* $y_k = \sum_{j=k}^{\infty} \epsilon_j(u)\beta^{k-j-1}$, $\widetilde{\tau}$ *is as in* (2), ζ *as in* (10).

P r o o f. We show first that the conditions on $u \in L_\omega$ are sufficient for the boundedness of $D(N, [0, \langle u \rangle))$. Note that the assumption $\tau(q, a_0) > q$ for $0 < q < d$ is not used here, but only the weaker assumption $\tau(q, a_0) > 0$ of Lemma 4. For $N = \mathrm{val}_{\widetilde{S'}}(w_\ell \cdots w_1)$, we have

$$\sum_{k=1}^{\ell} \sum_{b<w_k} \zeta_{\widetilde{\tau}(d, w_\ell \cdots w_{k+1} b)}(y_k) = \sum_{k=m+1}^{\ell} \#\left\{ b < w_k : \begin{array}{l} bw_{k+1} \cdots w_\ell a_0^\omega < u_k u_{k+1} \cdots, \\ u_1 \cdots u_{k-1} b w_{k+1} \cdots w_\ell \in L \end{array} \right\} + \mathcal{O}(1) = C(N, u) + \mathcal{O}(1),$$

thus by (11)

$$D(N, [0, \langle u \rangle)) = \mathcal{O}(1).$$

Now, suppose that $D(N, [0, \langle u \rangle)) = \mathcal{O}(1)$ for $u \in L_\omega$. By Proposition 2, we have $\langle u \rangle \in \mathbb{Q}(\beta)$, thus u is eventually periodic by [RS05]. Let $m \ge 0$, $p \ge 1$ be such that

$$u = u_1 \cdots u_m (u_{m+1} \cdots u_{m+p})^\omega \quad \text{and} \quad \tau(d, u_1 \cdots u_m) = \tau(d, u_1 \cdots u_{m+p}).$$

Consider $v \in A^*$ and $k > m$. If $v \notin L$, then (13) holds since

$$\zeta_{\widetilde{\tau}(d,v)}(y_k) = \zeta_0(y_k) = 0.$$

If $v = v' a_0$, then (13) holds for v if and only if it holds for v'. Therefore, we can assume $v \in L'$. If v is not the empty word, then we define integers

$$N_h = \mathrm{val}_{\widetilde{S'}}\left((\widetilde{v} a_0^g)^h \widetilde{v} a_0^{k-1} \right) \quad \text{for } h \ge 0, \quad \text{with} \quad g \ge \max(d-1, p)$$

such that $g + |v|$ is a multiple of p. Set $v = v_k \cdots v_\ell$, with $k \le \ell$.

We show that

$$C(N_h, u) = (h+1) C(N_0, u) \quad \text{for all } h \ge 0. \tag{14}$$

If $u_{m+1} \cdots u_{m+p} = a_0^p$, then $C(N_h, u) = 0$ and (14) holds. If $u_{m+1} \cdots u_{m+p} > a_0^p$, then $g \ge p$ implies that $b v_{j+1} \cdots v_\ell a_0^\omega < u_j u_{j+1} \cdots$, $k \le j \le \ell$, if and only if $b v_{j+1} \cdots v_\ell (a_0^g v)^h a_0^\omega < u_j u_{j+1} \cdots$. Since $\mathcal{A}_L$ is totally ordered, the assumption $\tau(q, a_0) > q$ for $0 < q < d$ implies $\tau(d, a_0) = d$ and $\tau(q, a_0^j) = d$ for all $q > 0$, $j \ge d-1$, in particular $\tau(q, a_0^g) = d$. Therefore, we also have

$$u_1 \cdots u_{j-1} b v_{j+1} \cdots v_\ell \in L \ \text{ if and only if } \ u_1 \cdots u_{j-1} b v_{j+1} \cdots v_\ell (a_0^g v)^h \in L.$$

Since $g + |v|$ is a multiple of the period length of u, (14) follows.

Since $\tau(q, a_0^g) = d$ for all $q > 0$, we also have $\widetilde{\tau}(q, a_0^g) = d$ for all $q > 0$, and thus

$$\zeta_{\widetilde{\tau}(d, (\widetilde{v} a_0^g)^h v_\ell \cdots v_{j+1} b)}(y_j) = \zeta_{\widetilde{\tau}(d, v_\ell \cdots v_{j+1} b)}(y_j).$$

By (11), we get

$$D(N_h,[0,\langle u\rangle)) = (h+1)\Big(C(N_0,u) - \sum_{j=k}^{\ell}\sum_{b<v_j}\zeta_{\widetilde{\tau}(d,v_\ell\cdots v_{j+1}b)}(y_j)\Big) + \mathcal{O}(1).$$

Therefore, $D(N,[0,\langle u\rangle)) = \mathcal{O}(1)$ implies that

$$\sum_{j=k}^{\ell}\sum_{b<v_j}\zeta_{\widetilde{\tau}(d,v_\ell\cdots v_{j+1}b)}(y_j) = C(N_0,u) \quad \text{for } N_0 = \mathrm{val}_{\widetilde{S'}}(v_\ell\cdots v_k a_0^{k-1}),\ k>m. \tag{15}$$

We will show that (13) holds, by considering (15) for the integer N_0' defined by the successor of v in $\widetilde{L'}$. Let first v be the empty word, and $N_0' = \mathrm{val}_{\widetilde{S'}}(v_k' a_0^{k-1})$ with v_k' such that $\widetilde{\tau}(d,b) = 0$ for all $b\in A$ with $a_0 < b < v_k'$. Then, (15) gives

$$\zeta_d(y_k) = \zeta_{\widetilde{\tau}(d,a_0)}(y_k) = C(N_0',u) = \begin{cases} 1 & \text{if } a_0^\omega < u_k u_{k+1}\cdots, \\ 0 & \text{else,}\end{cases}$$

thus (13) holds if v is the empty word. Assume next that there exists some

$$v_k' > v_k \ \text{ with } \ \widetilde{\tau}(d, v_\ell\cdots v_{k+1}v_k') > 0.$$

Choose v_k' such that $\widetilde{\tau}(d,v_\ell\cdots v_{k+1}b) = 0$ for all $b\in A$ with $v_k < b < v_k'$, and set $N_0' = \mathrm{val}_{\widetilde{S'}}(v_\ell\cdots v_{k+1}v_k' a_0^{k-1})$. Then, (13) holds because of

$$\begin{aligned}\zeta_{\widetilde{\tau}(d,v_\ell\cdots v_k)}(y_k) &= \sum_{b<v_k'}\zeta_{\widetilde{\tau}(d,v_\ell\cdots v_{k+1}b)}(y_k) - \sum_{b<v_k}\zeta_{\widetilde{\tau}(d,v_\ell\cdots v_{k+1}b)}(y_k)\\ &= C(N_0',u) - C(N_0,u)\\ &= \begin{cases}1 & \text{if } v_k\cdots v_\ell a_0^\omega < u_k u_{k+1}\cdots,\ u_1\cdots u_{k-1}v_k\cdots v_\ell\in L,\\ 0 & \text{else.}\end{cases}\end{aligned}$$

Now, we proceed by induction. We know that (13) holds if v is the empty word. Assume that (13) holds for all words v of length $\ell-k$, and consider

$$\begin{aligned}\zeta_{\widetilde{\tau}(d,v_\ell\cdots v_{k+1})}(y_{k+1}) &= \sum_{i=1}^{d}\theta^{(i)}_{\widetilde{\tau}(d,v_\ell\cdots v_{k+1})}\big(\beta y_k - \epsilon_k(u)\big)^{(i)} \qquad (16)\\ &= \sum_{i=1}^{d}\beta_i\theta^{(i)}_{\widetilde{\tau}(d,v_\ell\cdots v_{k+1})}y_k^{(i)} - \sum_{i=1}^{d}\theta^{(i)}_{\widetilde{\tau}(d,v_\ell\cdots v_{k+1})}\sum_{b<u_k}\eta^{(i)}_{\tau(d,u_1\cdots u_{k-1}b)}.\end{aligned}$$

Let v_k be the maximal letter with $\widetilde{\tau}(d,v_\ell\cdots v_{k+1}v_k) > 0$. Since $(\theta_1,\dots,\theta_d)^t$ is an eigenvector of $M_{\widetilde{L}}$ and we already know that $\zeta_{\widetilde{\tau}(d,v_\ell\cdots v_{k+1}b)}(y_k)$, $b<v_k$, is

given by (13), we have

$$\sum_{i=1}^{d} \beta_i \theta^{(i)}_{\widetilde{\tau}(d,v_\ell\cdots v_{k+1})} y_k^{(i)} = \sum_{i=1}^{d} \sum_{b\in A} \theta^{(i)}_{\widetilde{\tau}(d,v_\ell\cdots v_{k+1}b)} y_k^{(i)}$$
$$= \sum_{b\in A} \zeta_{\widetilde{\tau}(d,v_\ell\cdots v_{k+1}b)}(y_k) = \zeta_{\widetilde{\tau}(d,v_\ell\cdots v_k)}(y_k) +$$
$$\#\{b < v_k : bv_{k+1}\cdots v_\ell a_0^\omega < u_k u_{k+1}\cdots,\ u_1\cdots u_{k-1} b v_{k+1}\cdots v_\ell \in L\}.$$

Using this equation,

$$\sum_{i=1}^{d} \theta^{(i)}_{\widetilde{\tau}(d,v_\ell\cdots v_{k+1})} \sum_{b<u_k} \eta^{(i)}_{\tau(d,u_1\cdots u_{k-1}b)} = \sum_{b<u_k} \#L^0_{\tau(d,u_1\cdots u_{k-1}b),\widetilde{\tau}(d,v_\ell\cdots v_{k+1})}$$
$$= \#\{b < u_k : u_1\cdots u_{k-1} b v_{k+1}\cdots v_\ell \in L\}$$

and the induction hypothesis, (16) yields

$$\zeta_{\widetilde{\tau}(d,v_\ell\cdots v_k)}(y_k)$$
$$= \zeta_{\widetilde{\tau}(d,v_\ell\cdots v_{k+1})}(y_{k+1}) + \#\{b < u_k : u_1\cdots u_{k-1} b v_{k+1}\cdots v_\ell \in L\}$$
$$- \#\{b < v_k : bv_{k+1}\cdots v_\ell a_0^\omega < u_k u_{k+1}\cdots,\ u_1\cdots u_{k-1} b v_{k+1}\cdots v_\ell \in L\}$$
$$= \begin{cases} 1 & \text{if } v_{k+1}\cdots v_\ell a_0^\omega < u_{k+1}u_{k+2}\cdots \text{ and } u_1\cdots u_k v_{k+1}\cdots v_\ell \in L, \\ 0 & \text{else,} \end{cases}$$
$$+ \begin{cases} 1 & \text{if } v_k < u_k \text{ and } u_1\cdots u_{k-1} v_k\cdots v_\ell \in L, \\ -1 & \text{if } v_k > u_k,\ v_{k+1}\cdots v_\ell a_0^\omega < u_{k+1}u_{k+2}\cdots,\ u_1\cdots u_k v_{k+1}\cdots v_\ell \in L, \\ 0 & \text{else,} \end{cases}$$
$$= \begin{cases} 1 & \text{if } v_k\cdots v_\ell a_0^\omega < u_k u_{k+1}\cdots \text{ and } u_1\cdots u_{k-1} v_k\cdots v_\ell \in L, \\ 0 & \text{else,} \end{cases}$$

where we have used that $u_1\cdots u_{k-1} b v_{k+1}\cdots v_\ell$ can be in L only if $b \le v_k$. Therefore, (13) holds for all words v of length $\ell - k + 1$, and Theorem 2 is proved. □

In the case of β-adic van der Corput sequences, all bounded remainder sets satisfy the conditions of Proposition 1, see [Ste06]. The following example shows that this is probably not true in the more general setting, i.e., that there might be sequences u satisfying the conditions of Theorem 1, but not those of Proposition 1. However, $\mathcal{A}_L$ is not a Pisot automaton in this example, and we have not found an example where $\mathcal{A}_L$ satisfies the conditions of Theorem 1.

Example 3. Let $\mathcal{A}_L = (\{0,1,2,3,4\}, \{a_0,a_1,a_2,a_3\}, \tau, 4, \{1,2,3,4\})$ be the totally ordered automaton with $\tau, \widetilde{\tau}$ given by the transition tables

τ	a_0	a_1	a_2	a_3
0	0	0	0	0
1	3	2	1	0
2	3	2	2	1
3	4	4	3	1
4	4	4	4	2

,

$\widetilde{\tau}$	a_0	a_1	a_2	a_3
0	0	0	0	0
1	2	2	1	0
2	4	2	2	0
3	4	4	3	1
4	4	4	4	3

.

If $u = a_3a_0a_2^\omega$, then $\tau(4, u_1 \cdots u_{k-1}) = 3$ and thus $\epsilon_k(u) = 2\eta_4$ for all $k \geq 3$. We obtain $y_k = \eta_3 - \eta_2 + \eta_1$, which implies $\zeta_4(y_k) = \zeta_2(y_k) = 1$, $\zeta_3(y_k) = \zeta_1(y_k) = 0$ by (12). It can be easily verified that (13) holds for all $v \in A^*$, $k \geq 3$, but the conditions on u of Proposition 1 are not satisfied.

We conclude by the remark that the boundedness of $D(N, I)$ is not invariant under translation of the interval, i.e., $D(N, [y, z))$ can be unbounded if $[0, z - y)$ is a bounded remainder set and vice versa, see [Ste06].

Acknowledgement. I am grateful to Philippe Nadeau for indicating the links with partitions to me, and to the referee for many suggestions improving the quality of the paper.

REFERENCES

[Ber77] BERTRAND, A.: *Développements en base de Pisot et répartition modulo* 1, C. R. Acad. Sci. Paris Sér. A-B **285** (1977), no. 6, A419–A421.

[BG96] BARAT, G. – GRABNER, P. J.: *Distribution properties of G-additive functions*, J. Number Theory **60** (1996), no. 1, 103–123.

[BR07] BERTHÉ, V. – RIGO, M.: *Odometers on regular languages*, Theory Comput. Syst. **40** (2007), no. 1, 1–31.

[BY00] BROWN, G. – YIN, Q.: *β-transformation, natural extension and invariant measure*, Ergodic Theory Dynam. Systems **20** (2000), no. 5, 1271–1285.

[DT97] DRMOTA, M. – TICHY, R. F.: *Sequences, Discrepancies and Applications*, in: Lecture Notes in Math. **1651**, Springer-Verlag, Berlin, 1997.

[FMN96] FERENCZI, S. – MAUDUIT, C. – NOGUEIRA, A.: *Substitution dynamical systems: algebraic characterization of eigenvalues*, Ann. Sci. École Norm. Sup. (4) **29** (1996), no. 4, 519–533.

[IM04] ICHIKAWA, Y. – MORI, M.: *Discrepancy of van der Corput sequences generated by piecewise linear transformations*, Monte Carlo Methods Appl.**10** (2004), no. 2, 107–116.

[KN74] KUIPERS, L. – NIEDERREITER, H.: *Uniform Distribution of Sequences*, Wiley-Interscience [John Wiley & Sons], New York, 1974 (Reprint edition in Dover Publications, Inc. Mineola, New York, 2006).

[Kwo09] KWON, D. Y.: *The natural extensions of β-transformations which generalize baker's transformations*, Nonlinearity **22** (2009), no. 2, 301–310.

[LR01] LECOMTE, P. B. A. – RIGO, M.: *Numeration systems on a regular language*, Theory Comput. Syst. **34** (2001), no. 1, 27–44.

[LR02] LECOMTE, P. – RIGO, M.: *On the representation of real numbers using regular languages*, Theory Comput. Syst. **35** (2002), no. 1, 13–38.

[Mor98] MORI, M.: *Low discrepancy sequences generated by piecewise linear maps*, Monte Carlo Methods Appl. **4** (1998), no. 2, 141–162.

[Nin98a] NINOMIYA, S.: *Constructing a new class of low-discrepancy sequences by using the β-adic transformation*, Math. Comput. Simulation **47** (1998), no. 2–5, 403–418.

[Nin98b] NINOMIYA, S.: *On the discrepancy of the β-adic van der Corput sequence*, J. Math. Sci. Univ. Tokyo **5** (1998), no. 2, 345–366.

[Par60] PARRY, W.: *On the β-expansions of real numbers*, Acta Math. Acad. Sci. Hungar. **11** (1960), 401–416.

[RS05] RIGO, M. – STEINER, V.: *Abstract β-expansions and ultimately periodic representations*, J. Théor. Nombres Bordeaux **17** (2005), no. 1, 283–299.

[Sak09] SAKAROVITCH, J.: *Elements of Automata Theory*, Cambridge University Press, Cambridge, 2009.

[Sch80] SCHMIDT, V.: *On periodic expansions of Pisot numbers and Salem numbers*, Bull. London Math. Soc. **12** (1980), no. 4, 269–278.

[Sha78] SHAPIRO, L.: *Regularities of distribution*, Studies in Probability and Ergodic Theory, Adv. in Math. Suppl. Stud. **2**, Academic Press, New York, 1978, pp. 135–154.

[Ste06] STEINER, W.: *Regularities of the distribution of β-adic van der Corput sequences*, Monatsh. Math. **149** (2006), no. 1, 67–81.

Received October 12, 2008
Accepted December 11, 2009

Wolfgang Steiner
LIAFA, CNRS,
Université Paris Diderot–
–Paris 7, case 7014
FR–75205 Paris Cedex 13
FRANCE
E-mail: steiner@liafa.jussieu.fr