Cut Languages in Rational Bases

Jiří Šíma and Petr Savický*

Institute of Computer Science
The Czech Academy of Sciences
P. O. Box 5, 18207 Prague 8, Czech Republic
{sima,savicky}@cs.cas.cz

Abstract. We introduce a so-called cut language which contains the representations of numbers in a rational base that are less than a given threshold. The cut languages can be used to refine the analysis of neural net models between integer and rational weights. We prove a necessary and sufficient condition when a cut language is regular, which is based on the concept of a quasi-periodic power series. We achieve a dichotomy that a cut language is either regular or non-context-free while examples of regular and non-context-free cut languages are presented. We show that any cut language with a rational threshold is context-sensitive.

Keywords: grammars, quasi-periodic power series, cut language

1 Cut Languages

We study so-called cut languages which contain the representations of numbers in a rational base [1, 2, 5–7, 10, 12–15] that are less than a given threshold. Hereafter, let a be a rational number such that $0 < |a| < 1$, which is the inverse of a base (radix) $1/a$ where $|1/a| > 1$, and let $B \subset \mathbb{Q}$ be a finite set of rational digits. We say that $L \subseteq \Sigma^*$ is a *cut language* over a finite alphabet $\Sigma \neq \emptyset$ if there is a bijection $b : \Sigma \longrightarrow B$ and a real threshold c such that

$$L = L_{<c} = \left\{ x_1 \dots x_n \in \Sigma^* \;\middle|\; \sum_{i=0}^{n-1} b(x_{n-i}) a^i < c \right\} . \tag{1}$$

The cut languages can be used to refine the analysis of computational power of neural network models [17, 23]. This analysis is satisfactorily fine-grained in terms of Kolmogorov complexity when changing from rational to arbitrary real weights [4, 18]. In contrast, there is still a gap between integer and rational weights, which results in a jump from regular to recursively enumerable languages in the Chomsky hierarchy. In particular, neural nets with *integer* weights, corresponding to binary-state networks, coincide with finite automata [3, 8, 9, 11, 16, 20, 25]. On the other hand, a neural network that contains *two* analog-state

* Research was done with institutional support RVO: 67985807 and partially supported by the grant of the Czech Science Foundation No. P202/12/G061.

units with *rational* weights, can implement two stacks of pushdown automata, a model equivalent to Turing machines [19]. A natural question arises: what is the computational power of binary-state networks including one extra analog unit with rational weights? Such a model is equivalent to finite automata with a register [21], which accept languages that can be represented by some cut languages combined in a certain way by usual operations (e.g. intersection with a regular language, concatenation, union); see [22] for the exact representation.

In this paper we prove a necessary and sufficient condition when a given cut language is regular (Section 3). For this purpose, we introduce and characterize an a-quasi-periodic number within B whose all representations in basis $1/a$ using the digits from B, are eventually quasi-periodic power series (Section 2). The concept of quasi-periodicity represents a natural generalization of periodicity, allowing for different quasi-repetends even of unbounded length. There are numbers with uncountably many representations, all of which are eventually quasi-periodic, although only countably many of them can be eventually periodic. We achieve a dichotomy that a cut language is either regular or non-context-free. In addition, we present examples of cut languages that are not context-free and we show that any cut language with a rational threshold is context-sensitive (Section 4). Finally, we summarize the results and present some open problems (Section 5).

2 Quasi-Periodic Power Series

In this section, we introduce and analyze a notion of a-quasi-periodic numbers within B which will be employed for characterizing the class of regular cut languages in Section 3. We say that a power series $\sum_{k=0}^{\infty} b_k a^k$ with coefficients $b_k \in B$ for all $k \geq 0$, is *eventually quasi-periodic* with *period sum* P if there is an increasing infinite sequence of its term indices $0 \leq k_1 < k_2 < \cdots$ such that for every $i \geq 1$,

$$\frac{\sum_{k=0}^{m_i-1} b_{k_i+k}\, a^k}{1-a^{m_i}} = P \tag{2}$$

where $m_i = k_{i+1} - k_i > 0$ is the length of *quasi-repetend* $b_{k_i}, \ldots, b_{k_{i+1}-1}$, while k_1 is the length of *preperiodic part* $b_0, \ldots, b_{k_1-1}$. For $k_1 = 0$, we call such a power series *quasi-periodic*. One can calculate the sum of any eventually quasi-periodic power series as

$$\sum_{k=0}^{\infty} b_k a^k = \sum_{k=0}^{k_1-1} b_k a^k + a^{k_1} P \tag{3}$$

since $\sum_{k=k_1}^{\infty} b_k a^k = \sum_{i=1}^{\infty} a^{k_i} \sum_{k=0}^{m_i-1} b_{k_i+k}\, a^k = P \cdot \sum_{i=1}^{\infty} a^{k_i}(1-a^{m_i}) = P \cdot \sum_{i=1}^{\infty}(a^{k_i} - a^{k_{i+1}}) = a^{k_1} P$ is an absolutely convergent series. It follows that the sum (3) does not change if any quasi-repetend is removed from associated sequence $(b_k)_{k=0}^{\infty}$ or if it is inserted in between two other quasi-repetends, which means that the quasi-repetends can be permuted arbitrarily.

Example 1. A quasi-periodic power series can be composed of quasi-repetends having unbounded length. For example, for any rational period sum $P \neq 0$, we define three rational digits as $\beta_1 = (1-a^2)P$, $\beta_2 = a(1-a)P$, and $\beta_3 = 0$, that is, $B = \{\beta_1, \beta_2, \beta_3\}$. Then $\beta_1, \beta_2^n, \beta_3$ where β_2^n means β_2 repeated n times, creates a quasi-repetend of length $n+2$ for every integer $n \geq 0$, because $(\beta_1 + \sum_{k=1}^{n} \beta_2 a^k + \beta_3 a^{n+1})/(1-a^{n+2}) = P$ whereas for any integer r such that $0 \leq r < n$, it holds $(\beta_1 + \sum_{k=1}^{r} \beta_2 a^k)/(1-a^{r+1}) \neq P$.

Furthermore, given a power series $\sum_{k=0}^{\infty} b_k a^k$, we define its *tail sequence* $(d_n)_{n=0}^{\infty}$ as $d_n = \sum_{k=0}^{\infty} b_{n+k}\, a^k$ for every $n \geq 0$. Denote by $D(\sum_{k=0}^{\infty} b_k a^k) = \{d_n \,|\, n \geq 0\}$ the set of tail values.

Lemma 2. *A power series $\sum_{k=0}^{\infty} b_k a^k$ with $b_k \in B$ for all $k \geq 0$, is eventually quasi-periodic with period sum P iff its tail sequence $(d_n)_{n=0}^{\infty}$ contains a constant infinite subsequence $(d_{k_i})_{i=1}^{\infty}$ such that $d_{k_i} = P$ for every $i \geq 1$.*

Proof. Let $\sum_{k=0}^{\infty} b_k a^k$ be an eventually quasi-periodic power series with period sum P, which means there is an increasing infinite sequence of its term indices $0 \leq k_1 < k_2 < \cdots$ such that equation (2) holds for every $i \geq 1$. It follows that $a^{k_i}\, d_{k_i} = \sum_{k=k_i}^{\infty} b_k\, a^k = \sum_{j=i}^{\infty} a^{k_j} \sum_{k=0}^{m_j-1} b_{k_j+k}\, a^k = P \cdot \sum_{j=i}^{\infty} a^{k_j}(1-a^{m_j}) = P \cdot \sum_{j=i}^{\infty}(a^{k_j} - a^{k_{j+1}}) = a^{k_i} P$, which implies $d_{k_i} = P$ for every $i \geq 1$.

Conversely, assume that $(d_n)_{n=0}^{\infty}$ contains a constant subsequence $(d_{k_i})_{i=1}^{\infty}$ such that $d_{k_i} = P$ for every $i \geq 1$. We have $\sum_{k=0}^{m_i-1} b_{k_i+k}\, a^k = d_{k_i} - a^{m_i} d_{k_{i+1}} = (1-a^{m_i})\, P$ where $m_i = k_{i+1} - k_i > 0$, which implies (2) for every $i \geq 1$. $\square$

Theorem 3. *A power series $\sum_{k=0}^{\infty} b_k a^k$ with $b_k \in B$ for all $k \geq 0$, is eventually quasi-periodic iff the set of its tail values, $D = D(\sum_{k=0}^{\infty} b_k a^k)$, is finite.*

Proof. Assume that D is a finite set, which means there must be a real number $P \in D$ such that $d_{k_i} = P$ for infinitely many indices $0 \leq k_1 < k_2 < \cdots$, that is, $(d_{k_i})_{i=1}^{\infty}$ creates a constant infinite subsequence of tail sequence $(d_n)_{n=0}^{\infty}$. According to Lemma 2, this ensures that $\sum_{k=0}^{\infty} b_k a^k$ is eventually quasi-periodic.

Conversely, let $\sum_{k=0}^{\infty} b_k a^k$ with $b_k \in B$ for all $k \geq 0$, be an eventually quasi-periodic power series with period sum P. Since $a \in \mathbb{Q}$ and $B \subset \mathbb{Q}$ is finite, P is a rational number by (2) and there exists a natural number $\beta > 0$ such that $B' = \{\beta(b-(1-a)P)/a \,|\, b \in B\} \subset \mathbb{Z}$ is a finite set of integers. According to Lemma 2, the tail sequence $(d_n)_{n=0}^{\infty}$ of $\sum_{k=0}^{\infty} b_k a^k$ contains a constant infinite subsequence $(d_{k_i})_{i=1}^{\infty}$ such that $d_{k_i} = P$ for every $i \geq 1$. Assume to the contrary that $D = \{d_n \,|\, n \geq 0\}$ is an infinite set.

We define a modified sequence $(d'_n)_{n=0}^{\infty}$ as $d'_n = \beta(d_{k_1+n} - P)$ for all $n \geq 0$, which satisfies $d'_{k'_i} = 0$ where $k'_i = k_i - k_1$, for every $i \geq 1$, and $D' = \{d'_n \,|\, n \geq 0\}$ is an infinite set. Furthermore, for each $n \geq 0$,

$$\frac{d'_n}{a} - d'_{n+1} = \frac{\beta(d_{k_1+n} - P)}{a} - \beta(d_{k_1+n+1} - P) = \beta\, \frac{b_{k_1+n} - (1-a)P}{a} \in B' \quad (4)$$

is an integer by the definition of B'. In addition, denote $1/a = \alpha/q \in \mathbb{Q}$ where natural number $\alpha > 0$ and integer $q \neq 0$ are coprime.

Lemma 4. *For every $n \geq 0$, there exists an integer δ and a natural number $p \geq 0$ such that $d'_n = \delta/q^p$.*

Proof. We proceed by induction on n. The assertion is obvious for $n = 0$ when $d'_0 = 0$. Assume that $d'_n = \delta/q^p$ for some $\delta \in \mathbb{Z}$ and $p \geq 0$. Then $d'_{n+1} = d'_n/a - b'$ for some integer $b' \in B' \subset \mathbb{Z}$ according to (4), which can be rewritten as $d'_{n+1} = (\alpha/q) \cdot (\delta/q^p) - b' = (\alpha\delta - b'q^{p+1})/q^{p+1} = \delta_1/q^{p+1}$ where $\delta_1 = \alpha\delta - b'q^{p+1} \in \mathbb{Z}$, completing the proof of Lemma 4. $\square$

Lemma 5. *If $d'_{n+1} \in \mathbb{Z}$, then $d'_n \in \mathbb{Z}$.*

Proof. Let $d'_{n+1} \in \mathbb{Z}$. By (4) there is $b' \in B' \subset \mathbb{Z}$ such that $d'_n/a = d'_{n+1} + b' \in \mathbb{Z}$. According to Lemma 4, $d'_n = \delta/q^p$ for some $\delta \in \mathbb{Z}$ and $p \geq 0$, which gives $d'_n/a = \alpha\delta/q^{p+1} \in \mathbb{Z}$. Since α and q are coprime, q^{p+1} must be a factor of δ, which means $\delta = \delta' q^{p+1}$ for some $\delta' \in \mathbb{Z}$, and hence $d'_n = \delta/q^p = \delta' q \in \mathbb{Z}$, completing the proof of Lemma 5. $\square$

We will show for each $n \geq 0$ that $d'_n \in \mathbb{Z}$. Let $i \geq 1$ be the least index such that $k'_i \geq n$ for which we know $d'_{k'_i} = 0 \in \mathbb{Z}$. By applying Lemma 5 $(k'_i - n)$ times we obtain $d'_{k'_i-1}, d'_{k'_i-2}, \ldots, d'_n \in \mathbb{Z}$.

Thus, $D' \subset \mathbb{Z}$ and since D' is infinite, there exists an index $m \geq 0$ such that $|d'_m| \geq (|a| \cdot M)/(1 - |a|) > 0$ where $M = \max_{b' \in B'} |b'|$. Note that $M > 0$ since for $M = 0$, we would have $B = \{(1-a)P\}$ implying $D = \{P\}$ which contradicts that D is infinite. According to (4), $|d'_{m+1}| \geq |d'_m|/|a| - M$ which implies $|d'_{m+1}| - |d'_m| \geq (1/|a| - 1)|d'_m| - M \geq 0$ by the definition of m. Hence, $|d'_{m+1}| \geq |d'_m|$, and by induction we obtain $|d'_n| \geq (|a| \cdot M)/(1 - |a|) > 0$ for every $n \geq m$. On the other hand, we know that there is an index i such that $k'_i \geq m$ for which $d'_{k'_i} = 0$, which is a contradiction completing the proof of Theorem 3. $\square$

We say that a real number c is *a-quasi-periodic within B* if any power series $\sum_{k=0}^{\infty} b_k a^k = c$ with $b_k \in B$ for all $k \geq 0$, is eventually quasi-periodic. Note that c that cannot not be written as a respective power series at all, or can, in addition, be expressed as a finite sum $\sum_{k=0}^{h} b_k a^k = c$ whereas $0 \notin B$, is also considered formally to be a-quasi-periodic. For example, the numbers from the complement of the Cantor set are formally $(1/3)$-quasi-periodic within $\{0, 2\}$.

Example 6. Example 1 can be extended to provide a nontrivial instance of an a-quasi-periodic number that has infinitely many different quasi-periodic representations composed of quasi-repetends of arbitrary length (greater than 1). This includes ordinarily periodic representations composed of one of these quasi-repetends and uncountably many non-periodic ones. Let $a \in \mathbb{Q}$ meet $0 < a < \frac{1}{2}$. We show that any positive rational number c is a-quasi-periodic within B where $B = \{\beta_1, \beta_2, \beta_3\}$ is defined in Example 1 so that $P = c$. Obviously, $\beta_1 > \beta_2 > \beta_3 = 0$. Assume that $c = \sum_{k=0}^{\infty} b_k a^k$ for some sequence $(b_k)_{k=0}^{\infty}$ where $b_k \in B$ for all $k \geq 0$. Observe first that it must be $b_0 = \beta_1$ since otherwise $c = \sum_{k=0}^{\infty} b_k a^k \leq \beta_2 + \sum_{k=1}^{\infty} \beta_1 a^k = a(1-a)c + (1-a^2)c \cdot a/(1-a) = 2ac < c$ due to $a < \frac{1}{2}$. Moreover, for any $n \geq 0$ such that $b_k = \beta_2$ for every $k = 1, \ldots, n$, it holds $b_{n+1} \neq \beta_1$ since otherwise $c = \sum_{k=0}^{\infty} b_k a^k \geq \beta_1 + \sum_{k=1}^{n} \beta_2 a^k + \beta_1 a^{n+1} =$

$(1-a^2)c+a(1-a)c\cdot a(1-a^n)/(1-a)+(1-a^2)c\cdot a^{n+1} = c-a^{n+1}(a^2+a-1)c > c$ due to $a^2+a-1<0$ for $0<a<\frac{1}{2}$.

First consider the case when there is $r \geq 1$ such that $b_k = \beta_2$ for all $k \geq r$. Then $b_0,\ldots,b_{r-1}$ is a preperiodic part and $b_k = \beta_2$ for $k \geq r$ represents a repetend of length $m_k = 1$, which proves $\sum_{k=0}^{\infty} b_k a^k$ to be eventually quasi-periodic. Further assume there is no such r, and thus $b_k = \beta_2$ for every $k = 1,\ldots,n_1$ and $b_{n_1+1} = \beta_3$, for some $n_1 \geq 0$. It follows that series $\sum_{k=0}^{\infty} b_k a^k = c$ starts with a quasi-repetend $\beta_1, \beta_2^{n_1}, \beta_3$ of length n_1+2 (cf. Example 1) which can be omitted as $\sum_{k=0}^{\infty} b_{n_1+2+k} a^k = (c - \sum_{k=0}^{n_1+1} b_k a^k)/a^{n_1+2} = c$ due to $\sum_{k=0}^{n_1+1} b_k a^k = c(1-a^{n_1+2})$ by (2), and the argument can be repeated for its tail $\sum_{k=0}^{\infty} b_{n_1+2+k} a^k = c$ to reveal the next quasi-repetend $\beta_1, \beta_2^{n_2}, \beta_3$ for some $n_2 \geq 0$ etc. Hence, $\sum_{k=0}^{\infty} b_k a^k$ is quasi-periodic, which completes the proof that c is a-quasi-periodic within B.

Example 7. On the other hand, we present an example of an eventually quasi-periodic series $\sum_{k=0}^{\infty} b_k a^k = c$ with $b_k \in B$ for all $k \geq 0$, such that c is not a-quasi-periodic within B. Let $a = \frac{2}{3}$, $B = \{0,1\}$, and define an eventually quasi-periodic series $\sum_{k=0}^{\infty} b_k a^k$ with a preperiodic part $b_0 = b_1 = 0$ and a repetend $b_{2+3k} = 0$, $b_{3+3k} = b_{4+3k} = 1$ for every $k \geq 0$, which sums to $c = ((\frac{2}{3})^3 + (\frac{2}{3})^4)\cdot \sum_{k=0}^{\infty} (\frac{2}{3})^{3k} = \frac{40}{57}$.

Furthermore, we employ a greedy approach to generate a series $\sum_{k=0}^{\infty} b'_k a^k = c$ with $b'_k \in \{0,1\}$ for all $k \geq 0$, which is not eventually quasi-periodic. In particular, find minimal $k_1 \geq 0$ such that $a^{k_1} < c$ which gives $b'_0 = \cdots = b'_{k_1-1} = 0$, $b'_{k_1} = 1$, and remainder $c_1 = c/a^{k_1} - 1$. For $n > 1$, let $b'_0,\ldots,b'_{k_{n-1}}$ be 0s except for $b'_{k_1} = b'_{k_2} = \cdots = b'_{k_{n-1}} = 1$. Then find minimal $k_n > k_{n-1}$ such that $a^{k_n-k_{n-1}} < c_{n-1}$ which produces $b'_{k_{n-1}+1} = \cdots = b'_{k_n-1} = 0$, $b'_{k_n} = 1$, and remainder $c_n = c_{n-1}/a^{k_n-k_{n-1}} - 1$. It follows that $c_n = \sum_{k=0}^{\infty} b'_{k_n+k} a^k - 1 = (c - \sum_{i=1}^{n} a^{k_i})/a^{k_n}$ for $n \geq 1$. By plugging $a = \frac{2}{3}$ and $c = \frac{40}{57}$ into this formula, for which $k_1 = 1$ and $k_2 = 9$, we obtain

$$c_n = \frac{20}{19}\left(\frac{3}{2}\right)^{k_n-1} - \sum_{i=1}^{n}\left(\frac{3}{2}\right)^{k_n-k_i} = \frac{3^{k_n-1} - 19\cdot 2\cdot \sum_{i=2}^{n} 2^{k_i-2}\cdot 3^{k_n-k_i}}{19\cdot 2^{k_n-1}} \tag{5}$$

which is an irreducible fraction since both 19 and 2 are not factors of 3^{k_n-1}. Hence, for any natural n_1, n_2 such that $0 < n_1 < n_2$ we know $c_{n_1} \neq c_{n_2}$. It follows that the tail sequence $(d'_n)_{n=0}^{\infty}$ of $\sum_{k=0}^{\infty} b'_k a^k = c$ contains infinitely many different values $d'_{k_n} = c_n + 1$ for $n \geq 1$, which implies that $\sum_{k=0}^{\infty} b'_k a^k$ is not an eventually quasi-periodic series, according to Theorem 3.

Theorem 8. *A real number c is a-quasi-periodic within B iff the tail sequences of all the power series satisfying $\sum_{k=0}^{\infty} b_k a^k = c$ with $b_k \in B$ for all $k \geq 0$, contain altogether only finitely many values, that is,*

$$\mathcal{D} = \bigcup_{\substack{\sum_{k=0}^{\infty} b_k a^k = c \\ for\ all\ k\geq 0,\, b_k\in B}} D\left(\sum_{k=0}^{\infty} b_k a^k\right) \tag{6}$$

is a finite set. In addition, if c is not a-quasi-periodic within B, then there exists a power series $\sum_{k=0}^{\infty} b_k a^k = c$ with $b_k \in B$ for all $k \geq 0$, whose tail sequence contains pair-wise different values.

Proof. Let $\mathcal{D}$ be a finite set. Then the tail sequence of any power series $\sum_{k=0}^{\infty} b_k a^k = c$ with $b_k \in B$ for all $k \geq 0$, contains only finitely many values and thus includes a constant infinite subsequence. According to Lemma 2, this implies that any $\sum_{k=0}^{\infty} b_k a^k = c$ is eventually quasi-periodic, and hence, c is a-quasi-periodic within B.

Conversely, assume that $\mathcal{D}$ is infinite. Consider a directed tree $T = (V, E)$ with vertex set $V \subseteq B^*$ such that $b_0 \cdots b_{n-1} \in V$ if its tail meets $t(b_0 \cdots b_{n-1}) = (c - \sum_{k=0}^{n-1} b_k a^k)/a^n \in \mathcal{D}$, which includes the empty string ε as a root satisfying $t(\varepsilon) = c$. Define a set of directed edges as

$$E = \{(b_0 \cdots b_{n-1}, b_0 \cdots b_{n-1} b_n) \mid b_0 \cdots b_{n-1}, b_0 \cdots b_{n-1} b_n \in V\} \,, \tag{7}$$

which guarantees the outdegree of T is bounded by $|B|$. Let $T' = (V', E')$ be a subtree of T with a maximal vertex subset $V' \subseteq V$ so that $\varepsilon \in V'$ and $t(v_1) \neq t(v_2)$ for any two different vertices $v_1, v_2 \in V'$.

We show that for any $d \in \mathcal{D}$ there is $v \in V'$ such that $t(v) = d$. On the contrary, suppose $b_0 \cdots b_{n-1} \in V \setminus V'$ is a vertex with minimal n, satisfying $t(v) \neq t(b_0 \cdots b_{n-1}) = d \in \mathcal{D}$ for every $v \in V'$. Clearly, $b_0 \cdots b_{n-2} \in V \setminus V'$ since otherwise vertex $b_0 \cdots b_{n-1}$ could be included into V' which contradicts the maximality of V'. By the minimality of n, we know there is $b'_0 \cdots b'_{m-1} \in V'$ such that $t(b'_0 \cdots b'_{m-1}) = t(b_0 \cdots b_{n-2})$. Thus, we have $t(b'_0 \cdots b'_{m-1} b_{n-1}) = d$ and the maximality of V' implies $b'_0 \cdots b'_{m-1} b_{n-1} \in V'$, which is in contradiction with the definition of $b_0 \cdots b_{n-1}$.

It follows that $\{t(v) \mid v \in V'\} = \mathcal{D}$ implying T' is infinite. According to König's lemma, there exists an infinite directed path in T' corresponding to a power series $\sum_{k=0}^{\infty} b_k a^k = c$ whose tail sequence contains pair-wise different values. By Lemma 2, this series is not eventually quasi-periodic and hence, c is not a-quasi-periodic within B. □

3 Regular Cut Languages

In this section we formulate a necessary and sufficient condition for a cut language $L_{<c}$ to be regular (Theorem 11), which is based on a-quasi-periodic thresholds c within B. The following Lemma 9 provides a technical characterization of the regular cut languages, which is proven by Myhill-Nerode theorem, while subsequent Lemma 10 separates the cases when threshold c is represented by a finite sum or when c has no representation in base $1/a$ using the digits from B.

Lemma 9. *Let Σ be a finite alphabet, $b : \Sigma \longrightarrow B$ be a bijection, and c be a real number. Then the cut language $L_{<c} = \{x_1 \cdots x_n \in \Sigma^* \mid \sum_{i=0}^{n-1} b(x_{n-i}) a^i < c\}$ is regular iff the set*

$$C = \left\{ c(b_0, \ldots, b_{\kappa-1}) \;\middle|\; I_\kappa \leq c - \sum_{k=0}^{\kappa-1} b_k a^k \leq S_\kappa \,;\, b_0, \ldots, b_{\kappa-1} \in B \,;\, \kappa \geq 0 \right\} \tag{8}$$

is finite, where

$$I_\kappa = \inf_{\substack{b_\kappa,\ldots,b_{h-1}\in B\\ h\geq\kappa}} \sum_{k=\kappa}^{h-1} b_k a^k\,, \qquad S_\kappa = \sup_{\substack{b_\kappa,\ldots,b_{h-1}\in B\\ h\geq\kappa}} \sum_{k=\kappa}^{h-1} b_k a^k\,, \tag{9}$$

$$c(b_0,\ldots,b_{\kappa-1}) = \begin{cases} \inf C(b_0,\ldots,b_{\kappa-1}) & \text{if } a^\kappa > 0\\ \sup C(b_0,\ldots,b_{\kappa-1}) & \text{if } a^\kappa < 0\,, \end{cases} \tag{10}$$

$$C(b_0,\ldots,b_{\kappa-1}) = \left\{ \sum_{k=0}^{h-\kappa-1} b_{\kappa+k}a^k \,\middle|\, \sum_{k=0}^{h-1} b_k a^k \geq c\,;\, b_\kappa,\ldots,b_{h-1}\in B\,;\, h\geq\kappa \right\}. \tag{11}$$

Proof. Let $C=\{c_1,\ldots,c_p\}$ in (8) be a finite set such that $c_1 < c_2 < \cdots < c_p$. We introduce an equivalence relation $\sim$ on Σ^* as follows. For any $x,y\in\Sigma^*$ of length $n_x=|x|$ and $n_y=|y|$, respectively, we define $x\sim y$ iff both $z_x=\sum_{i=0}^{n_x-1} b(x_{n_x-i})a^i$ and $z_y=\sum_{i=0}^{n_y-1} b(y_{n_x-i})a^i$ belong either to one of the $p+1$ open intervals $(-\infty,c_1),(c_1,c_2),\ldots,(c_{p-1},c_p),(c_p,\infty)$, or to one of the p singletons $\{c_1\},\{c_2\},\ldots,\{c_p\}$. Obviously, we have $2p+1$ equivalence classes. In order to prove that language $L_{<c}$ is regular we employ Myhill-Nerode theorem by showing that for any $x,y\in\Sigma^*$, if $x\sim y$, then for every $w\in\Sigma^*$, $xw\in L_{<c}$ iff $yw\in L_{<c}$. Thus, consider $x,y\in\Sigma^*$ such that $x\sim y$, and on the contrary, suppose there is $w\in\Sigma^*$ of length $\kappa=|w|$ with $z_w=\sum_{i=0}^{\kappa-1} b(w_{\kappa-i})a^i$, such that $xw\in L_{<c}$ and $yw\notin L_{<c}$. This means $z_w+I_\kappa \leq z_w+a^\kappa z_x < c \leq z_w+a^\kappa z_y \leq z_w+S_\kappa$ by (9), implying $I_\kappa < c-z_w \leq S_\kappa$ which ensures $c_j=c(b(w_\kappa),\ldots,b(w_1))\in C$ for some $j\in\{1,\ldots,p\}$, according to (8). It follows from (10) and (11) that $z_w+a^\kappa z_x < c \leq z_w+a^\kappa c_j \leq z_w+a^\kappa z_y$ which gives $a^\kappa z_x < a^\kappa c_j \leq a^\kappa z_y$ contradicting $x\sim y$.

Conversely, let $L_{<c}$ be a regular languages. According to Myhill-Nerode theorem, there is an equivalence relation $\sim$ on Σ^* with a finite number p of equivalence classes such that for any $x,y\in\Sigma^*$, if $x\sim y$, then for every $w\in\Sigma^*$, $xw\in L_{<c}$ iff $yw\in L_{<c}$. Assume to the contrary that C in (8) is infinite. Choose $c_0,c_1,\ldots,c_{2p+2}\in C$ so that $c_0<c_1<\cdots<c_{2p+2}$, and for each $j\in\{0,\ldots,2p+2\}$, let $c_j=c(b_{j0},\ldots,b_{j,\kappa_j-1})$ for some $b_{j0},\ldots,b_{j,\kappa_j-1}\in B$ and $\kappa_j\geq 0$, according to (8). Definition (10) and (11) ensures that for each odd $j\in\{1,3,\ldots,2p+1\}$, there exists $h_j\geq\kappa_j$ and $b_{j,\kappa_j},\ldots,b_{j,h_j-1}\in B$ such that $c'_j=\sum_{k=0}^{h_j-\kappa_j-1} b_{j\kappa_j+k}a^k$ is sufficiently close to c_j so that $c_{j-1}<c'_j<c_{j+1}$. Since there are only p equivalence classes, there must be two odd indices $j_x,j_y\in\{1,3,\ldots,2p+1\}$, say $j_x<j_y$, determining $x,y\in\Sigma^*$ of length $n_x=|x|=h_{j_x}-\kappa_{j_x}$ and $n_y=|y|=h_{j_y}-\kappa_{j_y}$, respectively, by $b(x_{n_x-i})=b_{j_x,\kappa_{j_x}+i}$ for $i=0,\ldots,n_x-1$ and $b(y_{n_y-i})=b_{j_y,\kappa_{j_y}+i}$ for $i=0,\ldots,n_y-1$, such that $x\sim y$. Thus, $c'_{j_x}=\sum_{i=0}^{n_x-1} b(x_{n_x-i})a^i$ and $c'_{j_y}=\sum_{i=0}^{n_y-1} b(y_{n_y-i})a^i$. For $a^\kappa>0$, choose $w\in\Sigma^*$ of length $\kappa=|w|=\kappa_{j_x+1}$ so that $c_{j_x+1}=c(b(w_\kappa),\ldots,b(w_1))$, and denote $z_w=\sum_{i=0}^{\kappa-1} b(w_{\kappa-i})a^i$. We know $c'_{j_x}<c_{j_x+1}<c'_{j_y}$. It follows that $z_w+a^\kappa c'_{j_x} < c \leq z_w+a^\kappa c_{j_x+1} < z_w+a^\kappa c'_{j_y}$ since $z_w+a^\kappa c'_{j_x}\geq c$ would contradict that c_{j_x+1} is the infimum according to (10) and (11). Hence, $xw\in L_{<c}$ and

$yw \notin L_{<c}$, which gives the contradiction. Similarly for $a^\kappa < 0$, choose $w \in \Sigma^*$ so that $c_{j_y-1} = c(b(w_\kappa), \ldots, b(w_1))$, which gives $z_w + a^\kappa c'_{j_y} < c \leq z_w + a^\kappa c_{j_y-1} < z_w + a^\kappa c'_{j_x}$, leading to the contradiction $xw \notin L_{<c}$ and $yw \in L_{<c}$. $\square$

Lemma 10. *Assume the notation as in Lemma 9. Then the two subsets of C, $C_1 = \{c(b_0, \ldots, b_{\kappa-1}) \in C \mid \sum_{k=0}^{\kappa-1} b_k a^k + a^\kappa c(b_0, \ldots, b_{\kappa-1}) > c\}$ and $C_2 = \{c(b_0, \ldots, b_{\kappa-1}) \in C \mid (\exists\, b_\kappa, \ldots, b_{h-1} \in B\,,\, h \geq \kappa) \sum_{k=0}^{h-1} b_k a^k = c\ \&\ (\forall\, b \in B)\ c(b_0, \ldots, b_{h-1}, b) \in C_1\}$ are finite.*

Proof. We define a directed rooted tree $T = (V, E)$ with vertex set $V = \{b_0 \cdots b_{k-1} \in B^* \mid (\exists\, b_k, \ldots, b_{\kappa-1} \in B)\, c(b_0, \ldots, b_{k-1}, b_k \ldots, b_{\kappa-1}) \in C_1\}$, including an empty string as a root, and a set of directed edges (7). Clearly, T covers all the directed paths starting at the root and leading to $b_0 \cdots b_{\kappa-1} \in V$ such that $c(b_0, \ldots, b_{\kappa-1}) \in C_1$. This also guarantees that T includes all $b_0 \cdots b_{\kappa-1} \in V$ such that $c(b_0, \ldots, b_{\kappa-1}) \in C_2$, by the definition of C_2. For each vertex $b_0 \cdots b_{k-1} \in V$ we define a closed interval $I(b_0, \ldots, b_{k-1}) = [\sum_{i=0}^{k-1} b_i a^i + I_k\,,\, \sum_{i=0}^{k-1} b_i a^i + S_k]$ by using (9). Obviously, $I(b_0, \ldots, b_{k-1}, b_k) \subset I(b_0, \ldots, b_{k-1})$ for any edge $(b_0 \cdots b_{k-1}, b_0 \cdots b_{k-1} b_k) \in E$. Hence, $c \in I(b_0, \ldots, b_{k-1})$ for every vertex $b_0 \cdots b_{k-1} \in V$ since $b_0 \cdots b_{k-1} \cdots b_{\kappa-1} \in V$ such that $c(b_0, \ldots, b_{\kappa-1}) \in C_1$ satisfies $c \in I(b_0, \ldots, b_{\kappa-1}) \subset I(b_0, \ldots, b_{k-1})$ according to (8).

On the contrary, suppose that tree T whose outdegree is bounded by $|B|$, is infinite. According to König's lemma, there exists an infinite directed path corresponding to an infinite sequence $(b_k^*)_{k=0}^\infty$ with $b_k^* \in B$ for all $k \geq 0$, which contains infinitely many vertices $b_0^* \cdots b_{\kappa-1}^* \in V$ such that $c(b_0^*, \ldots, b_{\kappa-1}^*) \in C_1$. On the other hand, interval $I(b_0^*, \ldots, b_{k-1}^*)$ is a nonempty compact set satisfying $c \in I(b_0^*, \ldots, b_{k-1}^*) \supset I(b_0^*, \ldots, b_k^*)$ for every $k \geq 1$, which yields $c \in \bigcap_{k\geq 0} I(b_0^*, \ldots, b_{k-1}^*) \neq \emptyset$ by Cantor's intersection theorem. Hence, $\sum_{k=0}^\infty b_k^* a^k = c$ which implies $\sum_{k=0}^{\kappa-1} b_k^* a^k + a^\kappa c(b_0^*, \ldots, b_{\kappa-1}^*) = c$ for any $b_0^* \cdots b_{\kappa-1}^* \in V$ such that $c(b_0^*, \ldots, b_{\kappa-1}^*) \in C_1$, according to (10) and (11), which contradicts the definition of C_1. It follows that T is finite which implies that C_1, C_2 are finite. $\square$

Theorem 11. *A cut language $L_{<c}$ is regular iff c is a-quasi-periodic within B.*

Proof. According to Lemma 9, language $L_{<c}$ is regular iff set C is finite which is equivalent to the condition that $C \setminus (C_1 \cup C_2)$ is finite, by Lemma 10. It follows from (8)–(11) that for any $b_0, \ldots, b_{\kappa-1} \in B$ and $\kappa \geq 0$, $c(b_0, \ldots, b_{\kappa-1}) \in C \setminus (C_1 \cup C_2)$ iff there exists sequence $(b_k)_{k=\kappa}^\infty$ with $b_k \in B$ for all $k \geq 0$, such that $\sum_{k=0}^{\kappa-1} b_k a^k + a^\kappa c(b_0, \ldots, b_{\kappa-1}) = c$ ($c(b_0, \ldots, b_{\kappa-1}) \notin C_1$) and $\sum_{k=0}^\infty b_k a^k = c$ ($c(b_0, \ldots, b_{\kappa-1}) \notin C_2$), which yields $c(b_0, \ldots, b_{\kappa-1}) = \sum_{k=0}^\infty b_{\kappa+k} a^k$. It follows that $C \setminus (C_1 \cup C_2) = \mathcal{D}$ by the definition of $\mathcal{D}$, which is finite iff c is a-quasi-periodic within B, according to Theorem 8. $\square$

4 Non-Context-Free Cut Languages

In this section we show in Theorem 13 that a cut language $L_{<c}$ is not context-free if threshold c is not a-quasi-periodic within B, which is proven by a pumping

technique introduced in Lemma 12. According to Theorem 11, we thus achieve a dichotomy that, a cut language is either regular or non-context-free. We present explicit instances of rational numbers with no eventually quasi-periodic representations in Example 14. On the other hand, the cut languages with rational thresholds are shown to be context-sensitive in Theorem 15.

We say that an infinite word $x \in \Sigma^\omega$ is *approximable* in a language $L \subseteq \Sigma^*$, if for every finite prefix $u \in \Sigma^*$ of x, there is $y \in \Sigma^*$ such that $uy \in L$.

Lemma 12. *Let $x \in \Sigma^\omega$ be approximable in a context-free language $L \subseteq \Sigma^*$. Then there is a decomposition $x = uvw$ where $u, v \in \Sigma^*$ and $w \in \Sigma^\omega$, such that $|v| > 0$ is even and for every integer $i \geq 0$, word $uv^i w$ is approximable in L.*

Proof. Consider a context-free grammar G for L in Greibach normal form such that for every nonterminal A of G, there is a derivation of a terminal word from A. Since x is approximable in $L = L(G)$, there is a left derivation $S \Rightarrow \ldots \Rightarrow u_n\alpha_n$ for every n, such that $u_n \in \Sigma^*$ is the prefix of x of length n, and α_n is a sequence of nonterminal symbols. These derivations form an infinite directed rooted tree with the root S, whose vertices are the left sentential forms $u\alpha$ such that u is a prefix of x, and the edges outcoming from $u\alpha$ correspond to an application of one production rule to the left-most nonterminal in α. The degree of each vertex is bounded by the number of production rules. According to König's lemma, there is an infinite left derivation $S \Rightarrow \ldots \Rightarrow u_n\alpha_n \Rightarrow \ldots$ such that for every n, u_n is the prefix of x of length n, and α_n is a non-empty sequence of nonterminal symbols.

Let us call an occurrence of a nonterminal in α_n *temporary*, if it is substituted by a production rule of G in some of the following steps, and *stable* otherwise. We prove that for every n, there is $m \geq n$ such that α_m contains exactly one temporary nonterminal. We know the left-most nonterminal A_1 in $\alpha_n = A_1 \ldots A_i \ldots A_k$ is temporary, and let A_i be the right-most temporary nonterminal in α_n. If $i = 1$, then choose $m = n$. For $i \geq 2$, there is an index m, such that all the temporary nonterminals $A_1, \ldots, A_{i-1}$ in α_n are transformed into terminal words in u_m. If m is the smallest such index, then A_i is the first and the only temporary nonterminal of α_m. It follows that there is an infinite number of indices n such that α_n contains exactly one temporary nonterminal.

Since there are only finitely many nonterminals in G, there exist three indices m_1, m_2, m_3 such that $m_1 < m_2 < m_3$ and $u_{m_1}\alpha_{m_1} = u_1 A \beta_1'$, $u_{m_2}\alpha_{m_2} = u_1 v_1 A \beta_2' \beta_1'$, $u_{m_3}\alpha_{m_3} = u_1 v_1 v_2 A \beta_3' \beta_2' \beta_1'$ for some nonterminal A, where $u_1, v_1, v_2 \in \Sigma^*$, $|v_1| > 0$, $|v_2| > 0$, and $\beta_1', \beta_2', \beta_3'$ consist of stable nonterminals in all $\alpha_{m_1}, \alpha_{m_2}, \alpha_{m_3}$. If $|v_1|$ is even, then define $n_1 = m_1$, $n_2 = m_2$, $u = u_1$, $v = v_1$, $\beta_1 = \beta_1'$, and $\beta_2 = \beta_2'$, otherwise, if $|v_2|$ is even, then $n_1 = m_2$, $n_2 = m_3$, $u = u_1 v_1$, $v = v_2$, $\beta_1 = \beta_2' \beta_1'$, and $\beta_2 = \beta_3'$. On the other hand, if $|v_1|$ and $|v_2|$ are both odd, then $|v_1 v_2|$ is even and define $n_1 = m_1$, $n_2 = m_3$, $u = u_1$, $v = v_1 v_2$, $\beta_1 = \beta_1'$, and $\beta_2 = \beta_3' \beta_2'$. Thus, there are two words $u, v \in \Sigma^*$ such that $u_{n_1}\alpha_{n_1} = uA\beta_1$, $u_{n_2}\alpha_{n_2} = uvA\beta_2\beta_1$, and $|v| > 0$ is even, where $A \stackrel{*}{\Rightarrow} vA\beta_2$. For every $m \geq n_2$, we have $u_m\alpha_m = uv\gamma_m\beta_2\beta_1$ where γ_m is such that $A \stackrel{*}{\Rightarrow} \gamma_m$. Hence, an infinite word $w \in \Sigma^\omega$ is produced from A, such that $x = uvw$. Clearly, every finite prefix of w is the terminal part of γ_m for some $m \geq n_2$.

For every $i \geq 0$, we can construct an infinite left derivation whose sentential forms contain arbitrarily long prefixes of the sequence uv^iw by combining the above derivations similarly as in the proof of the pumping lemma. The derivation starts as the original derivation until $u_{n_1}\alpha_{n_1} = uA\beta_1$. Then, the derivation $A \stackrel{*}{\Rightarrow} vA\beta_2$ is used i times. Finally, the derivations $A \stackrel{*}{\Rightarrow} \gamma_m$ are used in an infinite sequence for all $m > n_2$. Altogether, we obtain

$$S \stackrel{*}{\Rightarrow} uA\beta_1 \stackrel{*}{\Rightarrow} uv^iA\beta_2^i\beta_1 \Rightarrow \ldots \Rightarrow uv^i\gamma_m\beta_2^i\beta_1 \Rightarrow \ldots \quad \text{for all } m > n_2\,. \tag{12}$$

We show that for every $i \geq 0$, the infinite sequence uv^iw is approximable in L. For any prefix $u' \in \Sigma^*$ of uv^iw, we employ the derivation (12) until u' is derived. Then, we include any finite derivation of a terminal word from each of the remaining nonterminals. We obtain a word in $L = L(G)$ with prefix u'. □

Theorem 13. *Assume that Σ is a finite alphabet and $b : \Sigma \longrightarrow B$ is a bijection. If c is not a-quasi-periodic within B (see Examples 7 and 14 for instances of such $c \in \mathbb{Q}$), then the cut language $L_{<c}$ over Σ is not context-free.*

Proof. For any string $x = x_1 \ldots x_n \in \Sigma^*$ of length $n = |x|$, denote $z_x = \sum_{k=0}^{n-1} b(x_{k+1})a^k$, whereas $z_x = \sum_{k=0}^{\infty} b(x_{k+1})a^k$ for an infinite word $x \in \Sigma^\omega$. Assume for a contradiction that $L_{<c}$ is a context-free language, and hence the same holds for its reversal $L = L_{<c}^R = \{x \in \Sigma^* \mid z_x < c\}$. Since c is not eventually a-quasi-periodic within B, Theorem 8 provides an infinite word $x \in \Sigma^\omega$ such that the tail sequence of a power series $z_x = \sum_{k=0}^{\infty} b(x_{k+1})a^k = c$ is composed of pair-wise different values.

On the contrary, suppose that x is not approximable in L. This means there is a prefix $u \in \Sigma^*$ of x such that for every $y \in \Sigma^*$ it holds $uy \notin L$, that is, $z_{uy} \geq c = z_x$. On the other hand, we know $z_x = \lim_{n\to\infty} z_{uy_n}$ where for every n, $y_n \in \Sigma^*$ is a string of length $n = |y_n|$ such that uy_n is a prefix of x, which implies $z_x = \inf_{y\in\Sigma^*} z_{uy}$. For $a > 0$, this ensures $b(x_k) = \min B$ for every $k > |u|$, whereas for $a < 0$, it must be $b(x_{2k}) = \max B$ and $b(x_{2k+1}) = \min B$ for every $k > |u|/2$, which contradicts the fact that the tail values of series z_x are pair-wise different.

It follows that x is approximable in L. Let $x = uvw$ where $|v| > 0$ is even, be a decomposition guaranteed by Lemma 12. In particular, uw and $uvvw$ are also approximable in L. We know the tails z_w and z_{vw} are different. If $a^{|u|}z_w > a^{|u|}z_{vw}$, then define $y = uw$ which meets $z_y = z_{uw} = z_u + a^{|u|}z_w > z_u + a^{|u|}z_{vw} = z_{uvw} = z_x = c$. On the other hand, if $a^{|u|}z_{vw} > a^{|u|}z_w$, then define $y = uvvw$ which satisfies $z_y = z_{uvvw} = z_{uv} + a^{|uv|}z_{vw} > z_{uv} + a^{|uv|}z_w = z_{uvw} = z_x = c$, due to $a^{|v|} > 0$. Thus, we have $y \in \Sigma^\omega$ which is approximable in L and $z_y > c$. This means that for every integer $n \geq 0$, there is $y_n \in L$ implying $z_{y_n} < c$, such that y and y_n share the same prefix of length at least n. Hence, $|z_y - z_{y_n}| \leq \beta a^n/(1-a)$ where $\beta = \max\{|b_1 - b_2|\,;\, b_1 \in B,\, b_2 \in B \cup \{0\}\}$. It follows that z_{y_n} converges to z_y as n tends to infinity, which contradicts $z_{y_n} < c < z_y$. □

Example 14. We generalize Example 7 to provide instances of rational numbers c such that any power series $\sum_{k=0}^{\infty} b'_k a^k = c$ with $b'_k \in B$ for all $k \geq 0$, is not eventually quasi-periodic. Let $B = \{0,1\}$ and $a = \alpha_1/\alpha_2$, $c = \gamma_1/\gamma_2 \in \mathbb{Q}$ be irreducible fractions where $\alpha_1, \gamma_1 \in \mathbb{Z}$ and $\alpha_2, \gamma_2 \in \mathbb{N}$, such that $\alpha_1\gamma_2$ and $\alpha_2\gamma_1$ are coprime. Denote by $0 < k_1 < k_2 < \cdots$ all the indices of a (not necessarily greedy) representation of $c = \sum_{k=0}^{\infty} b'_k a^k$ such that $b'_{k_i} = 1$ for $i \geq 1$. Then formula (5) can be rewritten as

$$c_n = \frac{\gamma_1 \alpha_2^{k_n} - \gamma_2 \alpha_1 \sum_{i=1}^{n} \alpha_1^{k_i - 1} \alpha_2^{k_n - k_i}}{\gamma_2 \alpha_1^{k_n}} \tag{13}$$

which is still an irreducible fraction.

Theorem 15. *Every cut language $L_{<c}$ with threshold $c \in \mathbb{Q}$ is context-sensitive.*

Proof. A corresponding (deterministic) linear bounded automaton M that accepts a given cut language $L_{<c} = L(M)$, evaluates (and stores) the sum $s_n = \sum_{i=0}^{n-1} b(x_{n-i})a^i$ step by step when reading an input word $x_1 \ldots x_n \in \Sigma^*$ from left to right. In particular, M starts with $s_0 = 0$ which updates to $s_i = as_{i-1} + b(x_i)$ every time after M reads the next input symbol $x_i \in \Sigma$, for $i = 1, \ldots, n$. As the numbers $a, b(x_1), \ldots, b(x_n), c \in \mathbb{Q}$ can be represented within constant space, M needs only linear space in terms of input length n, for computing s_n and testing whether $s_n < c$. □

5 Conclusion

In this paper we have introduced the cut languages in rational bases and classified them within the Chomsky hierarchy, among others, by using the quasi-periodic power series. A natural direction for future research is to generalize the results to arbitrary real bases.

We have already strengthened Theorem 8 whose proof is now based on Lemma 2 which does not require rational bases as opposed to stronger Theorem 3 that was used for the proof in a preliminary version [24]. As a consequence of this improvement, the characterization of regular cut languages in Theorem 11 remains valid for arbitrary real bases. For example, for the only real root $a \approx 0.6823278$ of algebraic equation $a^3 + a - 1 = 0$, which is the inverse of a Pisot number, the number $c = 1$ (similarly for $c = 1/a$) is a-quasi-periodic within $B = \{0,1\}$ and has uncountably many different quasi-periodic representations (including the non-periodic ones) whose tail values form $\mathcal{D} = \{0, a, 1, 1/a, 1+a, a/(1-a), (1+a)/a, 1/(1-a)\}$ (cf. Theorem 8). It is an open question of whether the inverse of the minimal Pisot number (i.e. the inverse of the plastic constant), $a \approx 0.7548777$ which is the unique real solution of the cubic equation $a^3 + a^2 - 1 = 0$, is the greatest such a.

Nevertheless, the generalization of Theorem 3 to arbitrary real bases is still an open problem which can be formulated elementarily as follows. Let a be a real number such that $0 < |a| < 1$, and $(d_n)_{n=0}^{\infty}$ be a sequence of real numbers, containing a constant infinite subsequence (cf. Lemma 2), such that $B = \{d_n - ad_{n+1} \,|\, n \geq 0\}$ is finite. Is $D = \{d_n \,|\, n \geq 0\}$ a finite set?

References

1. Adamczewski, B., Frougny, C., Siegel, A., Steiner, W.: Rational numbers with purely periodic β-expansion. Bulletin of The London Mathematical Society 42(3), 538–552 (2010)
2. Allouche, J.P., Clarke, M., Sidorov, N.: Periodic unique beta-expansions: The Sharkovskiĭ ordering. Ergodic Theory and Dynamical Systems 29(4), 1055–1074 (2009)
3. Alon, N., Dewdney, A.K., Ott, T.J.: Efficient simulation of finite automata by neural nets. Journal of the ACM 38(2), 495–514 (1991)
4. Balcázar, J.L., Gavaldà, R., Siegelmann, H.T.: Computational power of neural networks: A characterization in terms of Kolmogorov complexity. IEEE Transactions on Information Theory 43(4), 1175–1183 (1997)
5. Chunarom, D., Laohakosol, V.: Expansions of real numbers in non-integer bases. Journal of the Korean Mathematical Society 47(4), 861–877 (2010)
6. Glendinning, P., Sidorov, N.: Unique representations of real numbers in non-integer bases. Mathematical Research Letters 8(4), 535–543 (2001)
7. Hare, K.G.: Beta-expansions of Pisot and Salem numbers. In: Proceedings of the Waterloo Workshop in Computer Algebra 2006: Latest Advances in Symbolic Algorithms. pp. 67–84. World Scientific (2007)
8. Horne, B.G., Hush, D.R.: Bounds on the complexity of recurrent neural network implementations of finite state machines. Neural Networks 9(2), 243–252 (1996)
9. Indyk, P.: Optimal simulation of automata by neural nets. In: Proceedings of the STACS 1995 Twelfth Annual Symposium on Theoretical Aspects of Computer Science. LNCS, vol. 900, pp. 337–348 (1995)
10. Komornik, V., Loreti, P.: Subexpansions, superexpansions and uniqueness properties in non-integer bases. Periodica Mathematica Hungarica 44(2), 197–218 (2002)
11. Minsky, M.: Computations: Finite and Infinite Machines. Prentice-Hall, Englewood Cliffs (1967)
12. Parry, W.: On the β-expansions of real numbers. Acta Mathematica Hungarica 11(3), 401–416 (1960)
13. Rényi, A.: Representations for real numbers and their ergodic properties. Acta Mathematica Academiae Scientiarum Hungaricae 8(3-4), 477–493 (1957)
14. Schmidt, K.: On periodic expansions of Pisot numbers and Salem numbers. Bulletin of the London Mathematical Society 12(4), 269–278 (1980)
15. Sidorov, N.: Expansions in non-integer bases: Lower, middle and top orders. Journal of Number Theory 129(4), 741–754 (2009)
16. Siegelmann, H.T.: Recurrent neural networks and finite automata. Journal of Computational Intelligence 12(4), 567–574 (1996)
17. Siegelmann, H.T.: Neural Networks and Analog Computation: Beyond the Turing Limit. Birkhäuser, Boston (1999)
18. Siegelmann, H.T., Sontag, E.D.: Analog computation via neural networks. Theoretical Computer Science 131(2), 331–360 (1994)
19. Siegelmann, H.T., Sontag, E.D.: On the computational power of neural nets. Journal of Computer System Science 50(1), 132–150 (1995)
20. Šíma, J.: Energy complexity of recurrent neural networks. Neural Computation 26(5), 953–973 (2014)
21. Šíma, J.: The power of extra analog neuron. In: Proceedings of the TPNC 2014 Third International Conference on the Theory and Practice of Natural Computing. LNCS, vol. 8890, pp. 243–254 (2014)

22. Šíma, J.: Neural networks between integer and rational weights. Tech. Rep. V-1237, Institute of Computer Science, The Czech Academy of Sciences, Prague (2016)
23. Šíma, J., Orponen, P.: General-purpose computation with neural networks: A survey of complexity theoretic results. Neural Computation 15(12), 2727–2778 (2003)
24. Šíma, J., Savický, P.: Cut languages in rational bases. Tech. Rep. V-1236, Institute of Computer Science, The Czech Academy of Sciences, Prague (2016)
25. Šíma, J., Wiedermann, J.: Theory of neuromata. Journal of the ACM 45(1), 155–178 (1998)