# Quasi-Periodic $\beta$-Expansions and Cut Languages

Jiří Šíma[a,*], Petr Savický[a]

[a] *Institute of Computer Science of the Czech Academy of Sciences, P.O. Box 5, 182 07 Prague 8, Czech Republic*

**Abstract**

Motivated by the analysis of neural net models between integer and rational weights, we introduce a so-called cut language over a real digit alphabet, which contains finite $\beta$-expansions (i.e. base-$\beta$ representations) of the numbers less than a given threshold. We say that an infinite $\beta$-expansion is eventually quasi-periodic if its tail sequence formed by the numbers whose representations are obtained by removing leading digits, contains an infinite constant subsequence. We prove that a cut language is regular iff its threshold is a quasi-periodic number whose all $\beta$-expansions are eventually quasi-periodic, by showing that altogether they have a finite number of tail values. For algebraic bases $\beta$, we prove that there is an eventually quasi-periodic $\beta$-expansion with an infinite number of tail values iff there is a conjugate of $\beta$ on the unit circle. For transcendental $\beta$ combined with algebraic digits, a $\beta$-expansion is eventually quasi-periodic iff it has a finite number of tail values. For a Pisot base $\beta$ and digits from the smallest field extension $\mathbb{Q}(\beta)$ over rational numbers including $\beta$, we show that any number from $\mathbb{Q}(\beta)$ is quasi-periodic. In addition, we achieve a dichotomy that a cut language is either regular or non-context-free and we show that any cut language with rational parameters is context-sensitive.

*Keywords:* $\beta$-expansion, quasi-periodicity, Pisot number, cut language, Chomsky hierarchy

## 1. Introduction

Hereafter, let $\beta$ be a real number such that $|\beta| > 1$, which represents a *base (radix)* of non-standard positional numeral system, and let $A \neq \emptyset$ be a finite set of real numbers corresponding to *digits*. We say that a word (string) $a = a_1 \dots a_n \in A^*$ over alphabet $A$ is a finite *base-$\beta$ representation*, or briefly a *$\beta$-expansion* of a real number $x$ if

$$x = (a)_\beta = (a_1 \dots a_n)_\beta = \sum_{k=1}^{n} a_k \beta^{-k} \,. \tag{1}$$

[*] Corresponding author

*Email addresses:* sima@cs.cas.cz (Jiří Šíma), savicky@cs.cas.cz (Petr Savický)

*Preprint submitted to Elsevier* *February 23, 2018*

Note that we use only negative powers of $\beta$ while omitting the radix point at the left of $\beta$-expansions.

We introduce a so-called *cut language* $L_{<c} \subseteq A^*$ over alphabet $A$, which contains all finite $\beta$-expansions of the numbers that are less than a given real threshold $c$, that is,

$$L_{<c} = \{a \in A^* \mid (a)_\beta < c\} = \left\{ a_1 \dots a_n \in A^* \;\middle|\; \sum_{k=1}^{n} a_k \beta^{-k} < c \right\} . \tag{2}$$

In other words, a cut language is composed of finite $\beta$-expansions of a *Dedekind cut* from which its name comes from. One can analogously define the cut language $L_{>c}$ for the numbers greater than $c$, which is used in Paragraph 1.3. Moreover, a cut language can be defined over any finite alphabet $\Gamma \neq \emptyset$ when a mapping $\alpha : \Gamma \longrightarrow A$ is introduced so that each symbol $u \in \Gamma$ represents a digit $\alpha(u) \in A$. In this paper we classify the class of cut languages within the Chomsky hierarchy, which is related to the theory of $\beta$-expansions reviewed in Paragraph 1.1.

*1.1. $\beta$-Expansions: Uniqueness and Periodicity*

We first review the definitions and results concerning $\beta$-expansions related to our study of cut languages which is introduced in Paragraph 1.2 including an outline of the paper. We say that $a = a_1 a_2 a_3 \ldots \in A^\omega$ is an *infinite* base-$\beta$ representation, or briefly a $\beta$-expansion of a real number $x$ if

$$x = (a)_\beta = (a_1 a_2 a_3 \ldots)_\beta = \sum_{k=1}^{\infty} a_k \beta^{-k} . \tag{3}$$

Note that the infinite sum in (3) can be viewed as a power series in variable $\beta^{-1}$ which is convergent due to $|\beta| > 1$. Further denote $A = \{\alpha_1, \ldots, \alpha_p\}$ so that $\alpha_1 < \alpha_2 < \cdots < \alpha_p$. It can be shown [1] for $\beta > 1$ that every real number $x$ from the interval $[\alpha_1/(\beta - 1), \alpha_p/(\beta - 1)]$ has an infinite $\beta$-expansion iff

$$\max_{1<j\leq p} (\alpha_j - \alpha_{j-1}) \leq \frac{\alpha_p - \alpha_1}{\beta - 1} \tag{4}$$

(a slightly more complicated condition can also be derived for $\beta < -1$).

Furthermore, we say that an infinite $\beta$-expansion $a \in A^\omega$ is *eventually periodic* if $a = a_1 a_2 \ldots a_{k_1} (a_{k_1+1} a_{k_1+2} \ldots a_{k_2})^\omega$ where $m = k_2 - k_1 > 0$ is the length of a *repetend* (repeating digits) $a_{k_1+1} a_{k_1+2} \ldots a_{k_2} \in A^m$, whose minimum is called the *period* of $a$, while $k_1$ is the length of *preperiodic part* $a_1 a_2 \ldots a_{k_1} \in A^{k_1}$. For $k_1 = 0$, we call such a $\beta$-expansion (purely) *periodic*. Any eventually periodic $\beta$-expansion can be evaluated as

$$(a_1 a_2 \ldots a_{k_1} (a_{k_1+1} a_{k_1+2} \ldots a_{k_2})^\omega)_\beta = (a_1 \ldots a_{k_1})_\beta + \beta^{-k_1} \varrho \tag{5}$$

where a so-called *periodic point* $\varrho = ((a_{k_1+1} a_{k_1+2} \ldots a_{k_2})^\omega)_\beta \in \mathbb{R}$ satisfies

$$(a_{k_1+1} a_{k_1+2} \ldots a_{k_2})_\beta = \sum_{k=1}^{m} a_{k_1+k} \beta^{-k} = \varrho \left(1 - \beta^{-m}\right) \tag{6}$$

by the sum of geometric series with the common ratio $\beta^{-m}$.

Obviously, $\beta$-expansions are a generalization of the classical *decimal* expansions in integer base $\beta = 10$ with the digits from $A = \{0, 1, \ldots, 9\}$. Another important example are the *binary* expansions in base $\beta = 2$ with the digit alphabet $A = \{0, 1\}$, which are widely used in contemporary computers. The representations in non-integer bases have been systematically studied since late 1950's, starting with the seminal papers due to Rényi [2] and Parry [3]. For simplicity, it is usually assumed in the literature that a real base meets $\beta > 1$ and the standard digits are integers from $A = \{0, 1, \ldots, \lceil\beta\rceil - 1\}$, which ensures that a $\beta$-expansion exists for every real number $x \in \overline{D_\beta}$ where $\overline{D_\beta}$ is the closure of an open interval

$$D_\beta = \left(0\,,\ \frac{\lceil\beta\rceil - 1}{\beta - 1}\right)\,, \tag{7}$$

according to (4).

For an integer base $\beta \in \mathbb{N}$ when $\overline{D_\beta} = [0, 1]$, it is well known that every irrational number $x \in D_\beta \setminus \mathbb{Q}$ has a *unique* infinite $\beta$-expansion, while any rational $x \in \overline{D_\beta} \cap \mathbb{Q}$ has either a *unique* eventually periodic $\beta$-expansion (e.g., the endpoints 0 and 1 of $\overline{D_\beta}$ have the trivial $\beta$-expansions $0^\omega$ and $(\beta-1)^\omega$, respectively), or exactly the *two* distinct eventually periodic $\beta$-expansions, $a_1 a_2 \ldots a_n 0^\omega$ and $a_1 a_2 \ldots a_{n-1}(a_n - 1)(\beta - 1)^\omega$, if there exists a finite $\beta$-expansion $a_1 a_2 \ldots a_n \in A^*$ of $x = (a_1 a_2 \ldots a_n)_\beta$ such that $a_n \neq 0$. An example of such an ambiguity is $3/4 = (75)_{10} = (750^\omega)_{10} = (749^\omega)_{10}$.

For a non-integer base $\beta$, by contrast, *almost every* number $x \in \overline{D_\beta}$ has a continuum of distinct $\beta$-expansions [4]. In particular, for $1 < \beta < 2$ when $A = \{0, 1\}$, *every* number $x \in D_\beta = (0, 1/(\beta - 1))$ has a continuum of distinct $\beta$-expansions if $1 < \beta < \varphi$ where $\varphi = (1 + \sqrt{5})/2 \approx 1.618034$ is the golden ratio [5]. On the other hand, for $\varphi \leq \beta < q$ where $q \approx 1.787232$ is the (transcendental) *Komornik-Loreti constant* (i.e. the unique solution of equation $\sum_{k=1}^{\infty} t_k q^{-k} = 1$ where $(t_k)_{k=1}^{\infty}$ is the *Thue-Morse sequence* in which $t_k \in \{0, 1\}$ is the parity of the number of 1s in the binary representation of $k$), there are *countably* many numbers in $D_\beta$ that have *unique* $\beta$-expansions which are eventually periodic [6], e.g. $0^n(10)^\omega$ or $1^n(01)^\omega$ for every $n \geq 0$. In contrast, there are countably many eventually periodic $\varphi$-expansions of $x = 1$ including only $(10)^n 110^\omega$, $(10)^\omega$, and $(10)^n 01^\omega$ for every $n \geq 0$. Moreover, for $q \leq \beta < 2$, there is a *continuum* of numbers in $D_\beta$ with the unique $\beta$-expansions, whose Hausdorff dimension is 0 for $\beta = q$ and strictly between 0 and 1 for $q < \beta < 2$ (while its Lebesgue measure remains zero). In addition, for $q_2 \leq \beta < 2$ where $q_2 \approx 1.839287$ meets $q_2^3 - q_2^2 - q_2 - 1 = 0$, there is $x \in D_\beta$ which has exactly *two* $\beta$-expansions (a similar result holds for any fixed number of $\beta$-expansions) [7]. Furthermore, for every $m \geq 2$, there exists a bound $\beta_m$ satisfying $\varphi \leq \beta_m < 2$, such that there is $x \in D_\beta$ that has a periodic unique $\beta$-expansion with *period* $m$, provided that $\beta > \beta_m$, while there is no such a number if $\beta \leq \beta_m$ [8].

These results have further been generalized to any non-integer base $\beta > 2$ combined with the standard integer digits from $A = \{0, 1, \ldots, \lceil\beta\rceil - 1\}$, by using generalized golden ratios [9] and generalized Komornik-Loreti constants [10, 11].

For a general alphabet $A$ that may include non-integer digits, which is the case considered in this paper, the only unique $\beta$-expansions are the trivial ones, $\alpha_1^\omega$ and $\alpha_p^\omega$, when $\beta$ is sufficiently close to 1 [12]. The number of unique $\beta$-expansions increases when $\beta$ satisfying (4) increases, while for $\beta$ that meets

$$\beta > 1 + \frac{\alpha_p - \alpha_1}{\min_{1<j\le p}(\alpha_j - \alpha_{j-1})}\,, \tag{8}$$

all the $\beta$-expansions from $A^\omega$ are unique [13]. In particular, for $\beta$ and $A$ satisfying (4), there exist two critical bases $\varphi_A$ and $q_A$ such that $1 < \varphi_A \le q_A$ and the number of unique $\beta$-expansions is finite if $1 < \beta < \varphi_A$, countable if $\varphi_A < \beta < q_A$, and uncountable if $\beta > q_A$. Nevertheless, the determination of these critical bases (and even the existence of $q_A$) for arbitrary $A$ is still not complete even for three digits [12, 13].

The multiple $\beta$-expansions of a number can be ordered lexicographically and its maximal (resp. minimal) $\beta$-expansion is called *greedy* (resp. *lazy*). Obviously, any unique $\beta$-expansion is simultaneously greedy and lazy. Denote by $\mathrm{Per}(\beta)$ the set of numbers whose greedy $\beta$-expansion using the digits from $A$, is eventually periodic, where $A$ will always be clear from the context. For simplicity, assume $\beta > 1$ and $A = \{0, 1, \ldots, \lceil\beta\rceil - 1\}$. For any integer base $\beta \in \mathbb{Z}$, it is well known that $\mathrm{Per}(\beta) = \mathbb{Q} \cap [0, 1]$. For a non-integer base $\beta$, we have $\mathrm{Per}(\beta) \subseteq \mathbb{Q}(\beta) \cap \overline{D_\beta}$ where $\mathbb{Q}(\beta)$ denotes the smallest field extension over $\mathbb{Q}$ including $\beta$, according to (5) and (6). On the other hand, if $\mathbb{Q} \cap [0, 1] \subset \mathrm{Per}(\beta)$, then $\beta$ must be either a Pisot or a Salem number [14] where a *Pisot* (resp. *Salem*) number is a real algebraic integer (a root of some monic polynomial with integer coefficients) greater than 1 such that all its Galois conjugates (other roots of such a unique monic polynomial with minimal degree) are in absolute value less than 1 (resp. less or equal to 1 and at least one equals 1). In particular, for any $\beta \in \mathbb{Q} \setminus \mathbb{Z}$ which cannot be a Pisot nor Salem number by the integral root theorem, there exists a rational number from $\overline{D_\beta} \cap \mathbb{Q}$ whose greedy $\beta$-expansion is not eventually periodic. Nevertheless, it was shown for any Pisot number $\beta$ that $\mathrm{Per}(\beta) = \mathbb{Q}(\beta) \cap \overline{D_\beta}$ [14], while for Salem numbers this is still an open problem [15]. Recently, these results have partially been generalized to negative base $\beta < -1$ and non-integer digits in $A$ [16, 17, 18].

### *1.2. Paper Overview: Quasi-Periodic Numbers and Cut Languages*

For the purpose of classifying the cut languages within the Chomsky hierarchy, we introduce and analyze an interesting concept of a *$\beta$-quasi-periodic number within* $A$ (see Section 4) whose *all* infinite $\beta$-expansions using the digits from $A$, are so-called eventually quasi-periodic. Namely, an *eventually quasi-periodic $\beta$-expansion* (see Section 2) represents a natural generalization of eventually periodic $\beta$-expansion (5), which can be composed of *distinct* quasi-repetends that satisfy (6) for the same periodic point $\varrho$.

The quasi-repetends in a single eventually quasi-periodic $\beta$-expansion can also have different length which can even be unbounded, allowing for e.g. a constant, polynomial, or exponential number of distinct quasi-repetends of a given length. Thus, a $\beta$-quasi-periodic number within $A$ with at least two different quasi-repetends has uncountably many eventually quasi-periodic $\beta$-expansions, while countably many of them are eventually periodic which are generated using individual quasi-repetends as repetends. Moreover, any greedy eventually quasi-periodic $\beta$-expansion of a number can employ only one identical repetend, and hence it must be eventually periodic, although the number itself need not be $\beta$-quasi-periodic within $A$. This implies $\mathrm{QPer}(\beta) \subseteq \mathrm{Per}(\beta)$ where $\mathrm{QPer}(\beta)$ denotes the set of $\beta$-quasi-periodic numbers within $A$, which have an infinite $\beta$-expansion. In general, a number that is not $\beta$-quasi-periodic within $A$ can still have some of its $\beta$-expansions eventually quasi-periodic.

In Section 2, we present examples of $\beta$-expansions that are eventually quasi-periodic, including those with quasi-repetends of unbounded length for the plastic constant $\beta$ which is the minimal Pisot number. For any base $\beta$ that is an algebraic number whose conjugates are in absolute value not 1, we prove that a $\beta$-expansion is eventually quasi-periodic iff its tail sequence formed by the numbers whose representations are obtained by removing leading digits, contains a finite number of values. As it has been pointed to us, this result has already appeared in Akiyama, Thuswaldner, Zaïmi's paper [19, Theorem 3] within a slightly different context restricted to integer digits of bounded absolute value and zero periodic point (cf. Theorem 1). This equivalence is also shown for transcendental $\beta$ when the digits are assumed to be algebraic numbers.

In contrast, for any algebraic base $\beta$ that has a conjugate which is a complex number of absolute value 1, we construct a quasi-periodic $\beta$-expansion with infinitely many distinct tail values in Section 3. Thus, for algebraic bases $\beta$, we obtain the equivalence that for any digit alphabet $A$, every eventually quasi-periodic $\beta$-expansion has a finite number of tail values iff the conjugates of $\beta$ are outside the unit circle. This result yields the opposite implication to [19, Theorem 3] concerning the regularity of language (95). In addition, for both rational and irrational bases $\beta$, we provide examples of a real number that has no eventually quasi-periodic $\beta$-expansion since the tail values of any of its $\beta$-expansions are pairwise distinct. Apart from Section 3, all the examples presented in the paper exploit the binary alphabet $A = \{0, 1\}$ which is the simplest case widely used in the literature. Simpler examples can easily be found for larger alphabets and/or for arbitrary real digits.

In Section 4, we present examples of $\beta$-quasi-periodic numbers within $A$ for bases $\beta$ that are or are not Salem or Pisot numbers, including those numbers having eventually quasi-periodic $\beta$-expansions with a constant, linear, and exponential number of distinct quasi-repetends in terms of their length. On the other hand, we provide examples of real numbers that are not $\beta$-quasi-periodic within $A$, despite their greedy and/or lazy $\beta$-expansion is eventually periodic. Furthermore, we prove that a real number is $\beta$-quasi-periodic within $A$ iff all its $\beta$-expansions have altogether a finite number of tail values. For any number that is not $\beta$-quasi-periodic within $A$, an infinite $\beta$-expansion exists whose tail sequence contains only pairwise distinct values. For Pisot bases $\beta$ and digits from $\mathbb{Q}(\beta)$, we show that every number in $\mathbb{Q}(\beta)$ is $\beta$-quasi-periodic within $A$, which means $\mathrm{QPer}(\beta) = \mathrm{Per}(\beta)$.

In Section 5, we prove that a cut language is regular iff its threshold is $\beta$-quasi-periodic within $A$. In Section 6, we achieve a dichotomy that a cut language is either regular or non-context-free, depending on whether its threshold is or is not a $\beta$-quasi-periodic number within $A$, respectively. This provides examples of cut languages that are not context-free. We show that any cut language with a rational threshold is context-sensitive when the base $\beta$ and the digits in $A$ are also rationals. Finally, we summarize the results and present some open problems in Section 7. A preliminary version of this paper which considered mainly rational bases, appeared in [20].

### *1.3. Motivations from Neural Networks*

The cut languages can be used to refine the analysis of the computational power of neural network models which depends on the information content of their weight parameters [21, 22]. This analysis is satisfactorily fine-grained in terms of Kolmogorov complexity when changing from rational to arbitrary real weights, which is approved by an infinite proper hierarchy of nonuniform complexity classes between P and P/poly for polynomial-time computations [23, 24]. In contrast, there is still a gap between integer and rational weights, which results in a jump from regular to recursively enumerable languages within the Chomsky hierarchy.

In particular, neural nets with *integer* weights, corresponding to binary-state networks, coincide with finite automata [25, 26, 27, 28, 29, 30, 31]. On the other hand, a neural network that contains a few analog-state units with *rational* weights, can implement two stacks of pushdown automata, a model equivalent to Turing machines [32]. A natural question arises: What is the computational power of binary-state networks including one extra analog unit with rational (or even real) weights? Such a model is equivalent to finite automata with a register [33], which accept the languages $L \subseteq \Sigma^*$ over alphabet $\Sigma \neq \emptyset$ that can roughly be written in the following form [34, 35]:

$$L = h\left(\left(\left(\left(\bigcup_{r=1}^{p} \left(\overline{L_{<c_r}} \cap L_{<c_{r+1}}\right)^{\mathrm{R}} \cdot \Gamma_r\right)^{\mathrm{Pref}} \cap R_0\right)^{*} \cap R\right)\right) \tag{9}$$

($\overline{L_{<c_r}} \cap L_{<c_{r+1}}$ can be replaced with $L_{<0}$ , $L_{>c_r} \cap L_{<c_{r+1}}$ , $L_{>c_r} \cap \overline{L_{>c_{r+1}}}$ , $\overline{L_{<c_r}} \cap \overline{L_{>c_{r+1}}}$ , and $L_{>1}$ , so that the underlying intervals create a disjoint cover of the real line) including usual language operations. In particular, $h : \Gamma^* \longrightarrow \Sigma^*$ is a letter-to-letter morphism where a finite alphabet $\Gamma \neq \emptyset$ is partitioned into $\Gamma_1, \ldots, \Gamma_p$ . Moreover, $R, R_0 \subseteq \Gamma^*$ are regular languages, $S^{\mathrm{Pref}}$ denotes the largest prefix-closed subset of $S \cup \Gamma \cup \{\varepsilon\}$, and $S^{\mathrm{R}}$ denotes the reversal of language $S$. Nevertheless, the core of representation (9) is based on the cut languages $L_{<c_r}, L_{>c_r}$ parametrized by rational thresholds $0 = c_1 \leq c_2 \leq \cdots \leq c_{p+1} = 1$ , and defined over the alphabet $\Gamma$ using a base $\beta$ where a mapping $\alpha : \Gamma \longrightarrow A$ transforms $\Gamma$ into digit alphabet $A$. The digits, the base, and the thresholds are derived from the network weights so that $\beta \in \mathbb{Q}$ and $A \subset \mathbb{Q}$ if the weights of the extra analog-state unit are rationals.

This representation together with the present results on the cut languages show that the computational power of neural networks having integer weights can increase from regular languages to that between context-free and context-sensitive languages, when an extra analog-state unit with rational weights is added, while this does not bring any additional power even for real weights if the thresholds of cut languages in (9) are $\beta$-quasi-periodic within $A$ [34, 35].

## 2. Quasi-Periodic $\beta$-Expansions

In this section, we introduce and analyze a notion of eventually quasi-periodic $\beta$-expansions which is a natural generalization of the eventual periodicity defined by (5) and (6) for non-integer bases. We say that an infinite $\beta$-expansion $a_1a_2a_3\ldots \in A^\omega$ is *eventually quasi-periodic* with a *periodic point* $\varrho \in \mathbb{R}$ if there is an increasing infinite sequence of indices, $0 \le k_1 < k_2 < \cdots$, such that for every $i \ge 1$,

$$(a_{k_i+1}\ldots a_{k_{i+1}})_\beta = \sum_{k=1}^{m_i} a_{k_i+k}\beta^{-k} = \varrho\left(1-\beta^{-m_i}\right) \tag{10}$$

(cf. (6)) where $m_i = k_{i+1}-k_i > 0$ is the length of *quasi-repetend* $a_{k_i+1}\ldots a_{k_{i+1}} \in A^{m_i}$, while $k_1$ is the length of *preperiodic part* $a_1\ldots a_{k_1} \in A^{k_1}$. For $k_1 = 0$, we call such a $\beta$-expansion (purely) *quasi-periodic*. Any eventually quasi-periodic $\beta$-expansion can be evaluated as

$$(a_1a_2a_3\ldots)_\beta = \sum_{k=1}^{\infty} a_k\beta^{-k} = (a_1\ldots a_{k_1})_\beta + \beta^{-k_1}\varrho \tag{11}$$

(cf. (5)) by using (10) since

$$\begin{aligned}\sum_{k=k_1+1}^{\infty} a_k\beta^{-k} &= \sum_{i=1}^{\infty}\beta^{-k_i}\sum_{k=1}^{m_i} a_{k_i+k}\beta^{-k} = \varrho\sum_{i=1}^{\infty}\beta^{-k_i}(1-\beta^{-m_i}) \\ &= \varrho\sum_{i=1}^{\infty}(\beta^{-k_i}-\beta^{-k_{i+1}}) = \beta^{-k_1}\varrho \end{aligned}\tag{12}$$

is an absolutely convergent series. In fact, condition (10) is equivalent to the statement that every quasi-repetend creates a periodic $\beta$-expansion of $\varrho$, that is, for every $i \ge 1$,

$$((a_{k_i+1}\ldots a_{k_{i+1}})^\omega)_\beta = \varrho\,. \tag{13}$$

It follows that the sum (11) does not change if any quasi-repetend is removed from the $\beta$-expansion or if it is inserted in between two other quasi-repetends. More generally, the preperiodic part together with an arbitrary sequence of quasi-repetends satisfying (10) for the same periodic point $\varrho$, yields a $\beta$-expansion of the same number. Clearly, every eventually periodic $\beta$-expansion is eventually quasi-periodic with a sequence of identical quasi-repetends. An eventually periodic $\beta$-expansion can be decomposed into repetends in different ways by extending the preperiodic part and using a cyclic shift of the repetends. Although the periodic points $\rho$ are different in these decompositions, such decompositions are closely related to each other. On the other hand, a general eventually quasi-periodic $\beta$-expansion can be decomposed into quasi-repetends in ways which are completely unrelated, as illustrated in the following example.

**Example 1.** Assume $A = \{0, 1\}$ and let $\beta \approx 1.220744$ be the real root of the polynomial $x^4 - x - 1$, which means

$$\beta^4 - \beta - 1 = 0\,, \tag{14}$$

such that $1 < \beta < 2$. Any infinite word $a \in A^\omega$ generated by the $\omega$-regular expression $00\,(010 + 1000)^\omega$ is an eventually quasi-periodic $\beta$-expansion of the number 1 with the periodic point $\varrho = \beta^2$. In particular, the prefix 00 is the preperiodic part of length $k_1 = 2$ while 010 and 1000 represent two quasi-repetends of length 3 and 4, respectively, satisfying condition (10):

$$\begin{aligned}(010)_\beta &= \beta^{-2} = \beta^2\left(1 - \beta^{-3}\right) && (15)\\ (1000)_\beta &= \beta^{-1} = \beta^2\left(1 - \beta^{-4}\right) && (16)\end{aligned}$$

according to (14). Clearly, formula (11) reads as

$$(a)_\beta = (00)_\beta + \beta^{-2}\beta^2 = 1\,. \tag{17}$$

For instance, $a = 00\,(010\,1000\,010)^\omega = 000\,(1010000\,100)^\omega$ can also be decomposed into the preperiodic part 000 and two quasi-repetends 1010000 and 100 with the periodic point $\varrho = \beta^3$, which are not related to the original quasi-repetends 010 and 1000.

We characterize an eventually quasi-periodic $\beta$-expansion by using a so-called tail sequence. In particular, given an infinite $\beta$-expansion $a = a_1a_2a_3\ldots \in A^\omega$, we define its *tail sequence* $(r_n)_{n=0}^\infty$ as

$$r_n = (a_{n+1}a_{n+2}a_{n+3}\ldots)_\beta = \sum_{k=1}^{\infty} a_{n+k}\beta^{-k}\,, \tag{18}$$

which implies

$$r_{n+1} = \beta r_n - a_{n+1} \quad \text{for every } n \geq 0\,. \tag{19}$$

Denote by

$$R(a) = \{r_n \,|\, n \geq 0\} = \left\{ \sum_{k=1}^{\infty} a_{n+k}\beta^{-k} \;\middle|\; n \geq 0 \right\} \tag{20}$$

the set of its tail values. In addition, we define a directed *transition graph* $G(a) = (R(a), E(a))$ on the vertex set $R(a)$ with the edges from $E(a) = \{(r_n, r_{n+1}) \in (R(a))^2 \,|\, n \geq 0\}$. Each edge $(r_n, r_{n+1}) \in E(a)$ is labeled with the digit $a_{n+1} \in A$, for every $n \geq 0$, so that the $\beta$-expansion $a \in A^\omega$ defines an infinite directed walk in $G(a)$, traversing vertices $r_0, r_1, r_2, \ldots$ which satisfy the recurrence condition (19).

**Lemma 1.** *A $\beta$-expansion $a \in A^\omega$ is eventually quasi-periodic with a periodic point $\varrho$ iff its tail sequence $(r_n)_{n=0}^\infty$ contains a constant infinite subsequence $(r_{k_i})_{i=1}^\infty$ such that*

$$r_{k_i} = \varrho \quad \textit{for every } i \geq 1\,. \tag{21}$$

*Thus, if $R(a)$ is finite, then $a$ is eventually quasi-periodic.*

Proof. Let $a = a_1 a_2 a_3 \ldots \in A^\omega$ be an eventually quasi-periodic $\beta$-expansion with periodic point $\varrho$, which means there is an increasing infinite sequence of indices $0 \leq k_1 < k_2 < \cdots$ such that equation (10) holds for every $i \geq 1$. It follows that

$$\begin{aligned} \beta^{-k_i} r_{k_i} &= \sum_{k=k_i+1}^{\infty} a_k \beta^{-k} = \sum_{j=i}^{\infty} \beta^{-k_j} \sum_{k=1}^{m_j} a_{k_j+k} \beta^{-k} \\ &= \varrho \sum_{j=i}^{\infty} \beta^{-k_j} \left(1 - \beta^{-m_j}\right) = \varrho \sum_{j=i}^{\infty} \left(\beta^{-k_j} - \beta^{-k_{j+1}}\right) = \beta^{-k_i}\, \varrho\,, \end{aligned} \tag{22}$$

which implies (21).

Conversely, assume that $(r_n)_{n=0}^\infty$ contains a constant subsequence $(r_{k_i})_{i=1}^\infty$ that meets (21). We have

$$(a_{k_i+1} \ldots a_{k_{i+1}})_\beta = \sum_{k=1}^{m_i} a_{k_i+k} \beta^{-k} = r_{k_i} - \beta^{-m_i} r_{k_{i+1}} = \varrho \left(1 - \beta^{-m_i}\right) \tag{23}$$

where $m_i = k_{i+1} - k_i > 0$, which implies (10) for every $i \geq 1$.

Finally, assume that $R(a)$ is a finite set, which means there must be a real number $\varrho \in R(a)$ such that $r_{k_i} = \varrho$ for infinitely many indices $0 \leq k_1 < k_2 < \cdots$, that is, $(r_{k_i})_{i=1}^\infty$ creates a constant infinite subsequence of tail sequence $(r_n)_{n=0}^\infty$. Hence, $\beta$-expansion $a$ is eventually quasi-periodic. □

Denote by $P(a)$ the set of all possible periodic points of an eventually quasi-periodic $\beta$-expansion $a = a_1 a_2 a_3 \ldots \in A^\omega$, which meets $P(a) \subseteq R(a)$ by Lemma 1. As illustrated in Example 1, the decomposition of $a$ into a preperiodic part $a_1 \ldots a_{k_1}$ and quasi-repetends $a_{k_i+1} \ldots a_{k_{i+1}}$ for $i \geq 1$, is ambiguous by definition. Nevertheless, any choice of the periodic point $\varrho \in P(a)$ determines the unique quasi-repetends (inducing the unique preperiodic part) that are delimited by *all* its occurrences $r_{k_i} = \varrho$ for $i \geq 1$, in the tail sequence $(r_n)_{n=0}^\infty$ of $a$, according to Lemma 1, which means $r_n \neq \varrho$ whenever $n \notin \{k_i \mid i \geq 1\}$.

**Example 2.** We present an example of an eventually quasi-periodic $\beta$-expansion which is composed of quasi-repetends of unbounded length. Assume $A = \{0, 1\}$ and let $\beta \approx 1.324718$ be the *plastic constant* (i.e. the minimal Pisot number) which is the unique real root of the polynomial $x^3 - x - 1$, that is,

$$\beta^3 - \beta - 1 = 0\,. \tag{24}$$

We define an infinite word $a \in A^\omega$ as

$$a = 0\ (100)\, 0(011)1\ (100)^2\ 0(011)^2 1\ \ldots\ (100)^i\ 0(011)^i 1\ \ldots \tag{25}$$

which proves to be an eventually quasi-periodic $\beta$-expansion of the number 1 by the same argument as in Example 1. In particular, for the periodic point $\varrho = \beta$, the prefix 0 is a preperiodic part of length $k_1 = 1$ while 100 and $0(011)^i 1$, for every $i \geq 1$, represent the quasi-repetends of length 3 and $\ell_i = 3i + 2$, respectively, satisfying condition (10):

$$\begin{aligned}
(100)_\beta &= \beta^{-1} = \beta\left(1 - \beta^{-3}\right) && (26)\\
(0(011)^i 1)_\beta &= \sum_{k=1}^{i} \beta^{-3k} + \sum_{k=1}^{i} \beta^{-3k-1} + \beta^{-\ell_i} \\
&= \left(1 + \beta^{-1}\right) \frac{1 - \beta^{-\ell_i + 2}}{\beta^3 - 1} + \beta^{-\ell_i} \\
&= \frac{(\beta + 1)\left(1 - \beta^{-\ell_i + 2}\right) + \beta^{-\ell_i + 2}}{\beta^2} = \frac{\beta^3 - \beta^{-\ell_i + 3}}{\beta^2} \\
&= \beta\left(1 - \beta^{-\ell_i}\right) && (27)
\end{aligned}$$

according (24).

An alternative way of showing that the $\beta$-expansion $a$ defined in (25) is eventually quasi-periodic, is to generate its tail sequence $(r_n)_{n=0}^\infty$ starting with $r_0 = 1$ and using the recurrence (19) and condition (24). For this purpose, we introduce a sequence of integer polynomials $f_n \in \mathbb{Z}[x]$ as $f_0(x) = 1$ and

$$f_{n+1}(x) = (x f_n(x) - a_{n+1}) \bmod \left(x^3 - x - 1\right) \quad \text{for every } n \geq 0\,, \tag{28}$$

satisfying $r_n = f_n(\beta)$ by induction on $n$, which produces

$$\begin{aligned}
&1,\ \boldsymbol{\beta}\ , \beta^2 - 1, 1,\ \boldsymbol{\beta}\ , \beta^2, \beta + 1, \beta^2 + \beta - 1, \beta^2, \\
&\quad \ldots,\ \left(\boldsymbol{\beta}\ , \beta^2 - 1, 1\right)^i,\ \boldsymbol{\beta}\ , \left(\beta^2, \beta + 1, \beta^2 + \beta - 1\right)^i, \beta^2, \ldots
\end{aligned} \tag{29}$$

where $(\beta, \beta^2 - 1, 1)^i$ denotes the three tail values $\beta, \beta^2 - 1, 1$ repeated $i$ times and $\boldsymbol{\beta}$ in boldface highlights a constant infinite subsequence from Lemma 1. Hence, the set of tail values,

$$R(a) = \left\{\beta^2 - 1, 1, \beta, \beta^2, \beta^2 + \beta - 1, \beta + 1\right\}\,, \tag{30}$$

is finite, which proves $a$ to be eventually quasi-periodic according Lemma 1. In fact, this procedure also generates the corresponding transition graph $G(a) = (R(a), E(a))$ which is depicted in Figure 1. It is clear that each vertex in this graph is traversed infinitely many times following the walk (29), which implies $P(a) = R(a)$. For each periodic point $\varrho \in P(a)$, we obtain the unique decomposition of $a$ into quasi-repetends, e.g. for $\varrho = \beta^2 + \beta - 1$, we have the preperiodic part 0100001 and the quasi-repetends 101 and $11(100)^i 001$, for $i \geq 2$.

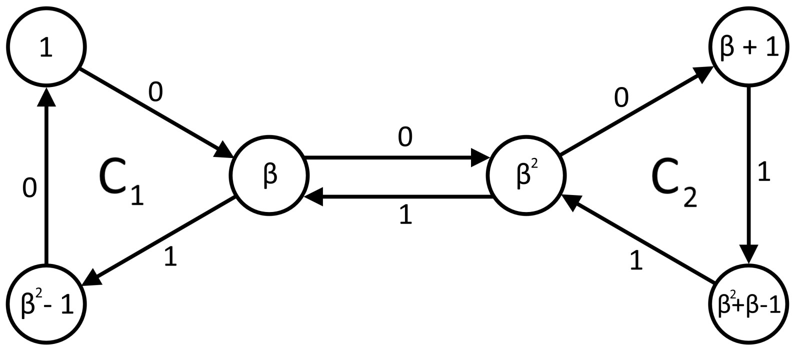


Figure 1: The transition graph $G(a)$ for $a = 0\,(100)\,0(011)1\ldots(100)^i\;0(011)^i1\ldots$ when $A = \{0,1\}$ and $\beta > 1$ is the plastic constant satisfying $\beta^3 - \beta - 1 = 0$.

Graph $G(a)$ contains two directed vertex-disjoint cycles $C_1 = \beta, \beta^2 - 1, 1$ and $C_2 = \beta^2, \beta + 1, \beta^2 + \beta - 1$ of length 3, which are bidirectionally connected by $(\beta, \beta^2) \in E(a)$ and $(\beta^2, \beta) \in E(a)$, respectively. Suppose a periodic point $\varrho \in R(a)$ is taken from the cycle $C_1$ (symmetrically for the cycle $C_2$). For every integer $i \geq 1$, after the cycle $C_1$ is consecutively traversed $i$ times, the walk crosses the edge $(\beta, \beta^2) \in E(a)$ and enters the cycle $C_2$ which is also consecutively traversed $i$ times according to (29). Thus the vertex $\varrho$ in the cycle $C_1$ is again visited first after at least $i$ passes of cycle $C_2$ take place, which ensures the gap of length at least $3i$ between two consecutive occurrences of $\varrho$ in the tail sequence $(r_n)_{n=0}^{\infty}$. It follows that for every choice of periodic point $\varrho \in P(a)$, the length of quasi-repetends of eventually quasi-periodic $\beta$-expansion (25) is unbounded.

**Example 3.** On the other hand, we present below an example of both rational and irrational base $\beta$ such that the tail sequence of each infinite $\beta$-expansion of a number contains only pairwise distinct values, which implies there is no eventually periodic $\beta$-expansion of this number according to Lemma 1. We assume $A = \{0,1\}$.

We first consider rational $\beta = \frac{3}{2} < \varphi$ which ensures there are uncountably many infinite $\beta$-expansions of the number 1 (see Paragraph 1.1). Denote by $a = a_1a_2a_3\ldots \in A^\omega$ any such a $\beta$-expansion whose tail sequence $(r_n)_{n=0}^{\infty}$ thus starts with $r_0 = 1$. We prove by induction on $n$ that for every $n \geq 0$, $r_n = c_n/2^n$ for some odd integer $c_n$. Obviously, $r_0 = 1/2^0$, and let $r_n = c_n/2^n$ holds for an odd integer $c_n$. If $a_{n+1} = 0$, then $r_{n+1} = 3c_n/2^{n+1}$ by (19) where $c_{n+1} = 3c_n$ is an odd integer, while for $a_{n+1} = 1$, we have $r_{n+1} = (3c_n - 2^{n+1})/2^{n+1}$ where $c_{n+1} = 3c_n - 2^{n+1}$ remains odd. Consequently, all the values in the tail sequence of each $\frac{3}{2}$-expansion of 1 are different.

Now consider the case of irrational $\beta = \sqrt{2} \approx 1.414214 < \varphi$ and let $a = a_1a_2a_3\ldots \in A^\omega$ be any of the uncountably many infinite $\beta$-expansions of the number $\beta - 1$. Denote by $\beta_1 = \beta$ and $\beta_2 = -\beta$ the roots of the polynomial $x^2 - 2$. We define a sequence of integer polynomials $f_n \in \mathbb{Z}[x]$ of degree at most 1 as $f_0(x) = x - 1$ and

$$f_{n+1}(x) = (xf_n(x) - a_{n+1}) \bmod \left(x^2 - 2\right) \quad \text{for every } n \geq 0\,. \tag{31}$$

By induction on $n$, the tail sequence $(r_n)_{n=0}^\infty$ of $a$ meets $r_n = f_n(\beta_1)$ according to (19). Similarly, the sequence $(r'_n)_{n=0}^\infty$ which is defined as $r'_n = f_n(\beta_2)$, satisfies

$$r'_{n+1} = \beta_2 r'_n - a_{n+1} \quad \text{for every } n \geq 0\,. \tag{32}$$

Moreover, we define the sequence

$$d_n = \max\left(0, \beta_2 - r'_n, r'_n - 1\right) \quad \text{for every } n \geq 0\,, \tag{33}$$

which is the distance of $r'_n$ from the interval $[\beta_2, 1]$. We prove that

$$d_{n+1} \geq \beta_1 d_n \quad \text{for every } n \geq 0\,. \tag{34}$$

If $r'_n \in [\beta_2, 1]$, then $d_n = 0$ and (34) is trivially met. If $r'_n < \beta_2$, then $d_n = \beta_2 - r'_n$ and $r'_{n+1} > 1$ by (32) since $\beta_2 r'_n > 2$ and $a_{n+1} \in \{0,1\}$, which implies $d_{n+1} = r'_{n+1} - 1$. Thus, (34) reduces to $d_{n+1} = \beta_2 r'_n - a_{n+1} - 1 \geq -\beta_2(\beta_2 - r'_n) = \beta_1 d_n$ which holds due to $\beta_2^2 = 2$. On the other hand, if $r'_n > 1$, then $d_n = r'_n - 1$ and $r'_{n+1} < \beta_2$ since $\beta_2 r'_n < \beta_2$, which implies $d_{n+1} = \beta_2 - r'_{n+1}$. Thus, (34) reduces to $d_{n+1} = \beta_2 - \beta_2 r'_n + a_{n+1} \geq -\beta_2(r'_n - 1) = \beta_1 d_n$ which is met. This completes the proof of inequality (34) which, together with $d_0 = 1$ following from $r'_0 = f_0(\beta_2) = \beta_2 - 1$, ensures that all $d_n$ respectively $r'_n$ are distinct. Hence, all the polynomials $f_n$ are different. Since for each $n \geq 0$, the degree of $f_n$ is less than the degree of the minimal polynomial of $\beta_1$ by (31), we have that also the tail values $r_n = f_n(\beta_1)$ for $n \geq 0$ are pairwise distinct.

For any algebraic base $\beta$ whose conjugates are not on the unit circle, we show in the following Theorem 1 that the sufficient condition from Lemma 1 that a set of tail values $R(a)$ is finite, is also necessary for a $\beta$-expansion $a \in A^\omega$ to be eventually quasi-periodic. In a preliminary version of this paper, we have proven this claim separately for the bases that are *rationals* [20] and *algebraic integers* with conjugates outside the unit circle (including Pisot numbers; cf. Theorem 5), respectively. The ideas of these two proofs can be combined in order to prove the result in Theorem 1, which has in fact appeared already in [19, Theorem 3] within a slightly different context restricted to integer digits of bounded absolute value and zero periodic point, as has been pointed to us by an anonymous referee and others (see Acknowledgments).

**Theorem 1.** *Let $\beta \in \mathbb{A} \cap \mathbb{R}$ be a real algebraic number whose conjugates are in absolute value not* 1*. Then $\beta$-expansion $a \in A^\omega$ is eventually quasi-periodic iff $R(a)$ is finite.*

PROOF. From Lemma 1, we already know even for arbitrary $\beta \in \mathbb{R}$ that if $R(a)$ is finite for some $\beta$-expansion $a \in A^\omega$, then $a$ is eventually quasi-periodic.

Conversely, let $a = a_1 a_2 a_3 \ldots \in A^\omega$ be an eventually quasi-periodic $\beta$-expansion with periodic point $\varrho$. Its tail sequence $(r_n)_{n=0}^\infty$ satisfying the recurrence (19), contains a constant infinite subsequence $(r_{k_i})_{i=1}^\infty$ that meets (21) according to Lemma 1. In the following claim, we first prove that $R(a) = \{r_n \,|\, n \geq 0\}$ is finite for digits $A \subset \mathbb{Q}(\beta)$ from the field extension $\mathbb{Q}(\beta) \subset \mathbb{A}$ for algebraic $\beta \in \mathbb{A} \cap \mathbb{R}$, which will then generalize to arbitrary real digits $A \subset \mathbb{R}$.

**Claim 1.** *Let $A \subset \mathbb{Q}(\beta)$ and assume that $(r_n)_{n=0}^\infty$ is a sequence of real numbers satisfying (19) and (21). Then $R(a) = \{r_n \,|\, n \geq 0\}$ is finite.*

PROOF. Denote by $\mathcal{O}_{\mathbb{Q}(\beta)}$ the ring of algebraic integers contained in $\mathbb{Q}(\beta)$. Reversely, $\mathbb{Q}(\beta)$ is the field of fractions of the integral domain $\mathcal{O}_{\mathbb{Q}(\beta)}$. Hence, there exists $\gamma \in \mathcal{O}_{\mathbb{Q}(\beta)}$ such that $\gamma \neq 0$ and

$$A' = \{\gamma\alpha \mid \alpha \in A\} \quad \subset \quad \mathcal{O}_{\mathbb{Q}(\beta)} \tag{35}$$

$$\varrho' = \gamma\varrho \quad \in \quad \mathcal{O}_{\mathbb{Q}(\beta)} \tag{36}$$

$$r'_0 = \gamma r_0 \quad \in \quad \mathcal{O}_{\mathbb{Q}(\beta)} \tag{37}$$

since we assume $A \subset \mathbb{Q}(\beta)$ which implies $\varrho \in \mathbb{Q}(\beta)$ and $r_0 \in \mathbb{Q}(\beta)$ according to (10) and (11), respectively. It follows that $a' = a'_1 a'_2 a'_3 \ldots \in (A')^\omega$ where $a'_n = \gamma a_n \in A' \subset \mathcal{O}_{\mathbb{Q}(\beta)}$ for every $n \geq 1$, is an eventually quasi-periodic $\beta$-expansion with the periodic point $\varrho' \in \mathcal{O}_{\mathbb{Q}(\beta)}$. By the assumption, its tail sequence $(r'_n)_{n=0}^\infty$ with $r'_n = \gamma r_n \in \mathbb{Q}(\beta)$ for every $n \geq 0$, satisfies the recurrence (19),

$$r'_{n+1} = \beta r'_n - a'_{n+1} \tag{38}$$

and condition (21),

$$r'_{k_i} = \varrho' \quad \text{for every } i \geq 1\,. \tag{39}$$

**Lemma 2.** *If $r'_n \notin \mathcal{O}_{\mathbb{Q}(\beta)}$, then $r'_{n+1} \notin \mathcal{O}_{\mathbb{Q}(\beta)}$.*

PROOF. Assume $r'_n \notin \mathcal{O}_{\mathbb{Q}(\beta)}$ and consider the fractional ideal $(r'_n)$ of $\mathcal{O}_{\mathbb{Q}(\beta)}$ generated by $r'_n \in \mathbb{Q}(\beta)$, which is a non-zero $\mathcal{O}_{\mathbb{Q}(\beta)}$-submodule of $\mathbb{Q}(\beta)$ (i.e. closed under linear combinations with scalars from $\mathcal{O}_{\mathbb{Q}(\beta)}$) such that there exists $d \in \mathcal{O}_{\mathbb{Q}(\beta)} \setminus \{0\}$ satisfying $d \cdot (r'_n) \subseteq \mathcal{O}_{\mathbb{Q}(\beta)}$. Recall that $(r'_n)$ decomposes uniquely up to ordering into the product of integer powers of prime ideals $P \subset \mathcal{O}_{\mathbb{Q}(\beta)}$ since the ring $\mathcal{O}_{\mathbb{Q}(\beta)}$ is a Dedekind domain [36]. For any prime ideal $P \subset \mathcal{O}_{\mathbb{Q}(\beta)}$ in this decomposition, let $v_P : \mathbb{Q}(\beta) \longrightarrow \mathbb{Z} \cup \{\infty\}$ be the discrete valuation on $\mathbb{Q}(\beta)$ over $P$ which defines the discrete valuation subring $R_P = \{x \in \mathbb{Q}(\beta) \,|\, v_P(x) \geq 0\} \supset \mathcal{O}_{\mathbb{Q}(\beta)}$ of $\mathbb{Q}(\beta)$ [36]. Since $r'_n \in \mathbb{Q}(\beta) \setminus \mathcal{O}_{\mathbb{Q}(\beta)}$, $P$ can be chosen so that it has a negative exponent in the decomposition and then the discrete valuation $v = v_P$ over $P$ meets $v\,(r'_n) < 0$. Recall that the discrete valuation $v$ satisfies the axioms: $v(0) = \infty$, $v(x \cdot y) = v(x) + v(y)$, and $v(x + y) \geq \min(v(x), v(y))$ for all $x, y \in \mathbb{Q}(\beta)$.

By induction on $n$, there is a polynomial $g(x) = \sum_{i=0}^n c_i x^i \in \mathcal{O}_{\mathbb{Q}(\beta)}[x]$ with the coefficients $c_n = r'_0 \in \mathcal{O}_{\mathbb{Q}(\beta)}$ and $c_i = -a_{n-i} \in \mathcal{O}_{\mathbb{Q}(\beta)}$ for $i = 0, \ldots, n-1$, such that $r'_n = g(\beta)$ according to (37), (38), and (35). We have $v(\beta) < 0$ since $v(\beta) \geq 0$ leads to a contradiction,

$$v(r'_n) = v(g(\beta)) = v\left(\sum_{i=0}^{n} c_i \beta^i\right) \geq \min_{i=0,\ldots,n} (v(c_i) + i \cdot v(\beta)) \geq 0 \qquad (40)$$

by the properties of $v$ and $v(c_i) \geq 0$ following from $c_i \in \mathcal{O}_{\mathbb{Q}(\beta)}$ for every $i = 0, \ldots, n$. According to (38),

$$\min\left(v\left(r'_{n+1}\right), v\left(a'_{n+1}\right)\right) \leq v\left(r'_{n+1} + a'_{n+1}\right) = v\left(\beta r'_n\right) = v(\beta) + v\left(r'_n\right) < 0 \quad (41)$$

which implies $v\left(r'_{n+1}\right) < 0$ due to $v\left(a'_{n+1}\right) \geq 0$ by (35), and hence, $r'_{n+1} \notin \mathcal{O}_{\mathbb{Q}(\beta)}$, completing the proof of Lemma 2. □

Let $p \in \mathbb{Q}[x]$ be a minimal (monic) polynomial of $\beta$ having degree $m \geq 1$. Let $\beta = \beta_1, \beta_2, \ldots, \beta_m$ be its roots which are pairwise different. We introduce a sequence of polynomials $f_n \in \mathbb{Q}[x]$ for every $n \geq 0$, so that $f_0(\beta) = r'_0 \in \mathbb{Q}(\beta) = \mathbb{Q}[\beta]$ and

$$f_{n+1} = \left(x f_n - g_{a'_{n+1}}\right) \bmod p \qquad (42)$$

where $g_{\alpha'} \in \mathbb{Q}[x]$ is a polynomial corresponding to the digit $\alpha' = g_{\alpha'}(\beta) \in A' \subset \mathbb{Q}[\beta]$. The polynomials $f_n(x) = \sum_{i=0}^{m-1} \phi_{ni} x^i$ have degree at most $m-1$ and satisfy

$$r'_n = f_n(\beta) = \sum_{i=0}^{m-1} \phi_{ni} \beta^i \qquad (43)$$

by induction on $n$, using (38) and $p(\beta) = 0$.

**Lemma 3.** *There exists a natural number $d \in \mathbb{N}$ such that for every $n \geq 0$, the coefficients of polynomial $f_n$ satisfy $d \cdot \phi_{ni} \in \mathbb{Z}$ for $i = 0, \ldots, m-1$.*

Proof. Suppose there exists $n \geq 0$ such that $r'_n \notin \mathcal{O}_{\mathbb{Q}(\beta)}$ and let $i \geq 1$ be the least index in (39) that meets $k_i > n$. By applying Lemma 2 $(k_i - n)$ times we obtain $r_{k_i} \notin \mathcal{O}_{\mathbb{Q}(\beta)}$ which contradicts $r_{k_i} = \varrho' \in \mathcal{O}_{\mathbb{Q}(\beta)}$ according to (36). Therefore, $r'_n \in \mathcal{O}_{\mathbb{Q}(\beta)}$ for all $n \geq 0$.

Let $\omega_1, \ldots, \omega_m$ be an integral basis of $\mathcal{O}_{\mathbb{Q}(\beta)}$ [37], which means that for every $n \geq 0$, there are integer coordinates $c_{n1}, \ldots, c_{nm} \in \mathbb{Z}$ of $r'_n$ with respect to this basis,

$$r'_n = \sum_{j=1}^{m} c_{nj} \omega_j \,. \qquad (44)$$

Since $\beta \in \mathbb{A} \cap \mathbb{R}$ is an algebraic number, for each $j \in \{1, \ldots, m\}$, the basis element $\omega_j \in \mathcal{O}_{\mathbb{Q}(\beta)} \subset \mathbb{Q}(\beta) = \mathbb{Q}[\beta]$ can be written as

$$\omega_j = \sum_{i=0}^{m-1} \frac{\mu_{ji}}{d_{ji}} \beta^i \qquad (45)$$

for some integers $\mu_{ji} \in \mathbb{Z}$ and natural numbers $d_{ji} \in \mathbb{N}$. Let $d \in \mathbb{N}$ be the least common multiple of the numbers $d_{ji}$ for every $j = 1, \ldots, m$ and $i = 0, \ldots, m-1$, which implies that the numbers $z_{ji} = d \cdot \frac{\mu_{ji}}{d_{ji}} \in \mathbb{Z}$ are integers. By plugging (45) into (44), we express $r'_n$ as a polynomial in $\beta$,

$$r'_n = \sum_{j=1}^{m} c_{nj} \sum_{i=0}^{m-1} \frac{z_{ji}}{d} \beta^i = \sum_{i=0}^{m-1} \left( \frac{1}{d} \sum_{j=1}^{m} c_{nj} z_{ji} \right) \beta^i . \tag{46}$$

Since the minimal polynomial of $\beta$ has degree $m$, the polynomials in (43) and (46) must coincide, which means

$$d \cdot \phi_{ni} = \sum_{j=1}^{m} c_{nj} z_{ji} \in \mathbb{Z} \tag{47}$$

are integers for every $n \geq 0$ and $i = 0, \ldots, m-1$, completing the proof of Lemma 3. □

For every $n \geq 0$, we define the vector $u_n = (u_{n1}, \ldots, u_{nm})$ as

$$u_n^{\intercal} = V \phi_n^{\intercal} \tag{48}$$

where $^{\intercal}$ denotes the vector transpose, $\phi_n = (\phi_{n0}, \ldots, \phi_{n,m-1})$ is the vector of coefficients of $f_n$, and

$$V = \begin{pmatrix} 1 & \beta_1 & \beta_1^2 & \ldots & \beta_1^{m-1} \\ 1 & \beta_2 & \beta_2^2 & \ldots & \beta_2^{m-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \beta_m & \beta_m^2 & \ldots & \beta_m^{m-1} \end{pmatrix} \tag{49}$$

is the square Vandermonde matrix of order $m$, which is invertible since $\beta_1, \ldots, \beta_m$ are pairwise distinct. Hence,

$$\phi_n^{\intercal} = V^{-1} u_n^{\intercal} \tag{50}$$

which gives the following bound on the maximum norm of $\phi_n$,

$$\|\phi_n\|_\infty \leq \left\|V^{-1}\right\|_\infty \cdot \|u_n\|_\infty . \tag{51}$$

It follows from (48) and (49) that for every $n \geq 0$, $u_n = (f_n(\beta_1), \ldots, f_n(\beta_m))$ which implies

$$u_{n+1,i} = \beta_i u_{ni} - g_{a'_{n+1}}(\beta_i) \quad \text{for } i = 1, \ldots, m \tag{52}$$

according to (42) and $p(\beta_i) = 0$. Note that for $i = 1$, the recurrence (52) coincides with (38). In the following lemmas we will bound $|u_{ni}|$ by using $|\beta_i|$ and $M_i = \max_{\alpha' \in A'} |g_{\alpha'}(\beta_i)|$ for $i = 1, \ldots, m$.

**Lemma 4.** *If $|\beta_i| < 1$, then the sequence $(u_{ni})_{n=0}^{\infty}$ is bounded.*

Proof. Let

$$\mu = \frac{M_i}{1-|\beta_i|}\,. \tag{53}$$

We show that for every $n \geq 0$, if $|u_{ni}| > \mu$, then $|u_{n+1,i}| < |u_{ni}|$. According to (52), we have

$$|u_{n+1,i}| = \left|\beta_i u_{ni} - g_{a'_{n+1}}(\beta_i)\right| \leq |\beta_i| \cdot |u_{ni}| + M_i\,. \tag{54}$$

We multiply inequality (54) by $-1$ and add $|u_{ni}|$, which gives

$$|u_{ni}| - |u_{n+1,i}| \geq (1-|\beta_i|) \cdot |u_{ni}| - M_i > (1-|\beta_i|)\,\mu - M_i = 0\,. \tag{55}$$

On the other hand, if $|u_{ni}| \leq \mu$, then

$$|u_{n+1,i}| \leq |\beta_i|\mu + M_i = \mu\,. \tag{56}$$

Altogether, this implies that for every $n \geq 0$, we have $|u_{ni}| \leq \max(\mu, |u_{0i}|)$ implying the statement of Lemma 4. □

**Lemma 5.** *If $|\beta_i| > 1$, then the sequence $(u_{ni})_{n=0}^{\infty}$ is either bounded or there is an index $n_0 \geq 0$ such that for all $n \geq n_0$, we have $|u_{n+1,i}| > |u_{ni}|$.*

Proof. Let

$$\mu = \frac{M_i}{|\beta_i|-1}\,. \tag{57}$$

We show that for every $n \geq 0$, if $|u_{ni}| > \mu$, then $|u_{n+1,i}| > |u_{ni}|$. According to (52), we have

$$|u_{n+1,i}| = \left|\beta_i u_{ni} - g_{a'_{n+1}}(\beta_i)\right| \geq |\beta_i| \cdot |u_{ni}| - M_i \tag{58}$$

which implies

$$|u_{n+1,i}| - |u_{ni}| \geq (\beta_i - 1) \cdot |u_{ni}| - M_i > (|\beta_i|-1)\,\mu - M_i = 0\,. \tag{59}$$

It follows that either for every $n \geq 0$, we have $|u_{ni}| \leq \mu$, or there is an index $n_0 \geq 0$ such that $|u_{n_0 i}| > \mu$ and then $|u_{n+1,i}| > |u_{ni}|$ for all $n \geq n_0$, which completes the proof of Lemma 5. □

The sequence $(u_{n1})_{n=0}^{\infty}$ is the tail sequence of $\beta$-expansion $a'$ due to $u_{n1} = f_n(\beta) = r'_n$ for every $n \geq 0$, which means $(u_{n1})_{n=0}^{\infty}$ is bounded. For $i \in \{2,\ldots,m\}$ such that $|\beta_i| < 1$, the sequence $(u_{ni})_{n=0}^{\infty}$ is bounded according to Lemma 4. For $i \in \{2,\ldots,m\}$ such that $|\beta_i| > 1$, assume for a contradiction that the sequence $(u_{ni})_{n=0}^{\infty}$ is unbounded. Lemma 5 implies that there is an index $n_0$, such that all the numbers $u_{ni}$ for $n \geq n_0$ are different. Hence, the polynomials $f_n$ for $n \geq n_0$ are distinct because $f_n(\beta_i) = u_{ni}$. By Lemma 1 we know that $(u_{n1})_{n=0}^{\infty}$ contains an infinite constant subsequence and thus, there are two different polynomials $f_{n_1}$ and $f_{n_2}$ of degree at most $m-1$ such that

$f_{n_1}(\beta) - f_{n_2}(\beta) = 0$, which contradicts that the minimal polynomial $p$ of $\beta$ has degree $m$. It follows that $(u_{ni})_{n=0}^{\infty}$ is bounded also for $|\beta_i| > 1$.

Since the sequences $(u_{ni})_{n=0}^{\infty}$ are bounded for all $i \in \{1, \ldots, m\}$, inequality (51) implies that the coefficients of $f_n$ are bounded. Consequently, there is only a finite number of different polynomials $f_n$ according to Lemma 3, and hence, also a finite number of different $r'_n$. This means that $R(a')$ respectively $R(a)$ is finite, completing the proof of Claim 1 for $A \subset \mathbb{Q}(\beta)$. □

Now we generalize the argument to arbitrary real digits $A \subset \mathbb{R}$. Let $\mathcal{U} = span(A \cup \{r_0\}) \subseteq \mathbb{R}$ be the linear span of the set $A \cup \{r_0\}$ in the vector space of real numbers over the field $\mathbb{Q}(\beta)$. Clearly, the dimension of $\mathcal{U}$ is finite due to $A$ is finite, and denote by $B = \{b_1, \ldots, b_d\}$ the basis of $\mathcal{U}$. Since $\mathcal{U}$ is closed under the multiplication by $\beta \in \mathbb{Q}(\beta)$ and $A \subset \mathcal{U}$, we have $r_n \in \mathcal{U}$ for every $n \geq 0$, according to (19), which implies $\varrho \in \mathcal{U}$ by (21). For any $u \in \mathcal{U}$ and $j \in \{1, \ldots, d\}$, let $c_j(u) \in \mathbb{Q}$ be the (unique) $j$-th coordinate of the vector $u \in \mathcal{U}$ with respect to the basis $B$, that is,

$$u = \sum_{j=1}^{d} c_j(u) b_j \,. \tag{60}$$

For every $j = 1, \ldots, d$, denote $A_j = \{c_j(\alpha) \,|\, \alpha \in A\} \subset \mathbb{Q}(\beta)$, $\varrho_j = c_j(\varrho) \in \mathbb{Q}(\beta)$, and $r_{nj} = c_j(r_n) \in \mathbb{Q}(\beta)$. Recurrence (19) reads

$$\beta r_n - r_{n+1} = \sum_{j=1}^{d} (\beta r_{nj} - r_{n+1,j}) b_j \in A \quad \text{for every } n \geq 0 \,, \tag{61}$$

which implies $\beta r_{nj} - r_{n+1,j} \in A_j$. Similarly, condition (21) rewrites to

$$r_{k_i} = \varrho = \sum_{j=1}^{d} c_j(\varrho) b_j = \sum_{j=1}^{d} \varrho_j b_j \quad \text{for every } i \geq 1 \,, \tag{62}$$

which gives $r_{k_i,j} = \varrho_j$ for every $i \geq 1$. Thus for a fixed $j \in \{1, \ldots, d\}$, the sequence $(r_{nj})_{n=0}^{\infty}$ satisfies the assumption of Claim 1 for $A$ and $\varrho$ replaced by $A_j$ and $\varrho_j$, respectively. Therefore, $\{r_{nj} \,|\, n \geq 0\}$ is finite for each $j \in \{1, \ldots, d\}$, which ensures $R(a)$ is finite because

$$r_n = \sum_{j=1}^{d} r_{nj} b_j \,. \tag{63}$$

This completes the proof of Theorem 1. □

Theorem 1 can easily be extended to transcendental bases if the digits are algebraic numbers.

**Theorem 2.** *Let $\beta \in \mathbb{R} \setminus \mathbb{A}$ be a real transcendental base and assume $A \subset \mathbb{A} \cap \mathbb{R}$ is a subset of real algebraic numbers. Then $\beta$-expansion $a \in A^{\omega}$ is eventually quasi-periodic iff $R(a)$ is finite.*

Proof. If $R(a)$ is finite, then $\beta$-expansion $a \in A^\omega$ is eventually quasi-periodic by Lemma 1. Conversely, let $a = a_1a_2a_3\cdots \in A^\omega$ be an eventually quasi-periodic $\beta$-expansion. Clearly, if $a$ is eventually periodic, then $R(a)$ is finite. On the contrary, suppose there are two distinct unique quasi-repetends $a_{k_j+1}\ldots a_{k_{j+1}} \in A^{m_j}$ and $a_{k_{j+1}+1}\ldots a_{k_{j+2}} \in A^{m_{j+1}}$ satisfying condition (10) for $i = j$ and $i = j+1$, respectively, for the same periodic point $\varrho = r_{k_j} = r_{k_{j+1}} \in P(a)$ whereas $r_n \neq \varrho$ for all $n \notin \{k_i \mid i \geq 1\}$. It follows that

$$\left(1-\beta^{-m_{j+1}}\right)\sum_{k=1}^{m_j} a_{k_j+k}\beta^{-k} = \left(1-\beta^{-m_j}\right)\sum_{k=1}^{m_{j+1}} a_{k_{j+1}+k}\beta^{-k} \tag{64}$$

which ensures $\beta \in \mathbb{A}$ because $A \subset \mathbb{A}$ and the field of algebraic numbers $\mathbb{A}$ is algebraically closed. This is a contradiction to our assumption that $\beta$ is transcendental. $\square$

## 3. Quasi-Periodic $\beta$-Expansions with Infinitely Many Tail Values

For algebraic bases $\beta$ whose conjugates are outside the unit circle, Theorem 1 shows that a $\beta$-expansion is eventually quasi-periodic iff the set of its tail values is finite. In this section, we prove for any algebraic base $\beta$ which breaks the assumption of Theorem 1 concerning its conjugates that for some integer digits, there is a quasi-periodic $\beta$-expansion with infinitely many tail values.

**Theorem 3.** *Let $\beta \in \mathbb{A}\cap\mathbb{R}$ be a real algebraic number with a conjugate of absolute value 1. Then there exists $A \subset \mathbb{Z}$ and a quasi-periodic $\beta$-expansion $a \in A^\omega$ of the number 0 such that $R(a)$ is infinite.*

Proof. Let $c_0,\ldots,c_d \in \mathbb{Z}$ where $c_d \neq 0$, be the integer coefficients of a unique (up to its sign) irreducible polynomial

$$p(x) = \sum_{k=0}^{d} c_k x^k \tag{65}$$

of the minimal degree $d$, whose roots include the algebraic base $\beta = \beta_1 \in \mathbb{A}\cap\mathbb{R}$ and its conjugate $\beta_2 \in \mathbb{C}$ from the assumption, which means $|\beta_1| > 1$ and $|\beta_2| = 1$. Note that $\beta_2 \notin \mathbb{R}$ since 1 and $-1$ cannot be the roots of non-linear irreducible $p$. Denote by $\beta_3 \in \mathbb{C}\setminus\mathbb{R}$ the complex conjugate of $\beta_2$ which is also a root of $p$, satisfying $|\beta_3| = |\beta_2| = 1$. Thus, $d \geq 3$ and $\beta_2 = 1/\beta_3$, $\beta_3 = 1/\beta_2$ are the roots of the reciprocal polynomial of $p$, which is the unique minimal polynomial of $\beta_2$ up to sign, that is, either $c_k = c_{d-k}$ for $k = 0\ldots,d$, or $c_k = -c_{d-k}$ for $k = 0\ldots,d$. The latter case would imply that 1 is a root of $p$, and hence $p$ is self-reciprocal. Since any self-reciprocal polynomial of odd degree has $-1$ as a root, we have $d \geq 4$ is even.

For all $m \geq m_0$, where $m_0$ is large enough exceeding the bounds obtained in several places of the proof, we will construct a quasi-repetend $a_1\ldots a_m \in A^m$ of size $m$ over a fixed finite set of integer digits $A \subset \mathbb{Z}$, satisfying (10) as

$$(a_1\ldots a_m)_\beta = 0 \tag{66}$$

for the periodic point $\varrho = r_0 = r_m = 0$, so that there are infinitely many different tail values $r_{m'+d}$ where

$$m' = \left\lfloor \tfrac{m}{2} \right\rfloor \tag{67}$$

and

$$r_n = (a_{n+1} \ldots \ldots a_m)_\beta = \sum_{k=n+1}^{m} a_k \beta^{-k} \qquad \text{for } n = 0, \ldots, m\,, \tag{68}$$

when $m$ goes to infinity. Note that $m_0 > 2d$ in order to have $m' + d < m$ due to (67). By using such quasi-repetends, one can simply define a quasi-periodic $\beta$-expansion $a \in A^\omega$ of the number 0 that is composed of an infinite sequence of these quasi-repetends with increasing $m$, which ensures the set of its tail values $R(a)$ is infinite, including infinite set $\{r_{m'+d} \,|\, m \geq m_0\}$.

We employ an $m \times m$ integer circulant matrix $C$ whose associated polynomial is $p$. Thus, the first column of $C$ contains its $d+1$ coefficients $c_0, \ldots, c_d$ completed by $m-d-1$ zeros, while the remaining columns of $C$ are each cyclic permutations of the first column with offset equal to the column index, that is,

$$(C)_{i,j} = \begin{cases} c_k = c_{d-k} & \text{if } 0 \leq k = (i-j) \bmod m \leq d \\ 0 & \text{otherwise} \end{cases} \quad \text{for } 1 \leq i, j \leq m\,. \tag{69}$$

It is known that the eigenvectors of circulant matrix $C$ are

$$v(\omega) = (1, \omega^1, \ldots, \omega^{m-1}) \tag{70}$$

for any $m$-th root of unity $\omega \in \mathbb{C}$ satisfying $\omega^m = 1$. The eigenvalue $\lambda(\omega) \in \mathbb{C}$ paired with the eigenvector $v(\omega)$ whose first component is 1, can be determined from the first component of vector $\lambda(\omega) v(\omega)^\intercal = C v(\omega)^\intercal$ as

$$\begin{aligned} \lambda(\omega) &= \sum_{j=1}^{m} (C)_{1,j} (v(\omega))_j = c_0 + \sum_{k=1}^{d} c_k \omega^{m-k} = \sum_{k=0}^{d} c_k \omega^{m-k} \\ &= \omega^{m-d} \sum_{k=0}^{d} c_k \omega^{d-k} = \omega^{-d} \sum_{k=0}^{d} c_k \omega^{k} = \frac{p(\omega)}{\omega^d} \end{aligned} \tag{71}$$

because $p$ is self-reciprocal and $\omega^m = 1$. Hence,

$$|\lambda(\omega)| = |p(\omega)| \tag{72}$$

due to $|\omega| = 1$.

Let $\delta \in \mathbb{C}$ be the $m$-th root of unity having the smallest positive complex argument, that is,

$$\delta = e^{\frac{2\pi \mathrm{i}}{m}}\,. \tag{73}$$

Assume that $\omega$ is the $m$-th root of unity such that in the complex plane, the $m$-th root of unity $\omega \delta^{d/2}$ is the closest possible to $\beta_2$. Then,

$$\left| \omega \delta^i - \beta_2 \right| = O(1/m) \qquad \text{for } i = 0, \ldots, d\,. \tag{74}$$

According to (70)–(72), the absolute value of the eigenvalues $\lambda(\omega\delta^i)$ paired with the eigenvectors $v(\omega\delta^i)$ for $i = 0, \ldots, d$, for the matrix $C$, is bounded as

$$\max_{i=0,\ldots,d} \left|(\lambda(\omega\delta^i)\right| = \max_{i=0,\ldots,d} \left|(p(\omega\delta^i)\right| = O(1/m) \tag{75}$$

since we assume (74) and the derivative of $p$ is bounded in the neighborhood of its root $\beta_2$. In particular, choose $s = \Theta(m)$ so that

$$\left|(\lambda(\omega\delta^i)\right| = \left|(p(\omega\delta^i)\right| \leq 1/s \qquad \text{for } i = 0, \ldots, d\,. \tag{76}$$

In addition, let $\gamma_0, \gamma_1, \ldots, \gamma_d \in \mathbb{C}$ be the complex coefficients of the monic polynomial $h$ of degree $d$ whose roots are $\delta^{m-d}, \delta^{m-d+1}, \ldots, \delta^{m-1} \in \mathbb{C}$, that is,

$$h(z) = \sum_{i=0}^{d} \gamma_i z^i = \prod_{j=1}^{d} \left(z - \delta^{m-j}\right)\,. \tag{77}$$

Now, we can define the quasi-repetend $a_1 \ldots a_m \in \mathbb{Z}^m$ as

$$(a_1, \ldots, a_m)^\intercal = C \cdot u'^\intercal \tag{78}$$

where $C$ is the circulant matrix introduced in (69) and $u' = (u'_1, \ldots, u'_m) \in \mathbb{Z}^m$ is an integer vector of size $m$, such that

$$u' = \mathrm{round}\left(\mathrm{Re}\left(su\right)\right)\,. \tag{79}$$

In particular, the integer vector $u' \in \mathbb{Z}^m$ is created from a complex vector $su \in \mathbb{C}^m$ by rounding the real part of all its components to the nearest integer, where $s$ meets (76) and $u = (u_1, \ldots, u_m) \in \mathbb{C}^m$ is a complex vector of size $m$, which is defined as

$$u = \frac{1}{\omega^{m'}} \sum_{i=0}^{d} \gamma_i v(\omega\delta^i) \tag{80}$$

by using $m'$, $\gamma_i$, and $v(\omega\delta^i)$ from (67), (77), and (70), (74), respectively.

The $j$-th component of vector $u$ introduced in (80), can be expressed as

$$u_j = \frac{1}{\omega^{m'}} \sum_{i=0}^{d} \gamma_i \left(\omega\delta^i\right)^{j-1} = \omega^{j-m'-1} h\left(\delta^{j-1}\right) \qquad \text{for } j = 1, \ldots, m\,, \tag{81}$$

by using (70) and (77), which implies $u_{m-d+1} = u_{m-d+2} = \cdots = u_m = 0$ by (77). Thus,

$$u'_{m-d+1} = u'_{m-d+2} = \cdots = u'_m = 0 \tag{82}$$

from (79). According to (78), (69), and (82), we have

$$a_j = (C \cdot u'^\intercal)_j = \sum_{k=\max(0,j-m+d)}^{\min(d,j-1)} c_k u'_{j-k} \qquad \text{for } j = 1, \ldots, m\,, \tag{83}$$

which produces

$$\sum_{j=1}^{m} a_j x^{-j} = \sum_{j=1}^{m} \left( \sum_{k=\max(0,j-m+d)}^{\min(d,j-1)} c_k u'_{j-k} \right) x^{-j} = p(x) \sum_{i=1}^{m-d} u'_i x^{-i-d} \,. \tag{84}$$

It follows from (84) and $p(\beta) = 0$ that

$$(a_1 \dots a_m)_\beta = \sum_{j=1}^{m} a_j \beta^{-j} = p(\beta) \sum_{i=1}^{m-d} u'_i \beta^{-i-d} = 0 \tag{85}$$

which meets (66) confirming that $a_1 \dots a_m \in \mathbb{Z}^m$ is a quasi-repetend for the periodic point $\varrho = 0$.

In the following lemma, we show that the quasi-repetend introduced in (78) exploits only bounded digits for arbitrarily large $m$, which means there is a finite set of integer digits $A \subset \mathbb{Z}$ such that $a_1 \dots a_m \in A^m$ for any $m \geq m_0$.

**Lemma 6.** *For arbitrarily large $m$, the digits $a_1, \dots, a_m \in \mathbb{Z}$ are bounded.*

Proof. In (79), rounding changes each component of vector $\mathrm{Re}(su)$ by at most $1/2$. According to (69), the entries of matrix $C$ are bounded and each row of $C$ contains at most $d+1$ non-zero entries. Hence, at most $d+1$ rounding errors contribute to each component of $C \cdot \mathrm{Re}(su)^\intercal$. It follows that the difference between $a_j$ in (78) and the $j$-th component of $C \cdot \mathrm{Re}(su)^\intercal$ is bounded for every $j = 1, \dots, m$. Moreover, we have $C \cdot \mathrm{Re}(su)^\intercal = \mathrm{Re}(C \cdot su^\intercal)$ due to $C$ is an integer matrix. Thus, it remains to prove that the components of $\mathrm{Re}(C \cdot su^\intercal)$ are bounded. For this purpose, it suffices to show that $|(C \cdot su^\intercal)_j|$ is bounded for every $j = 1, \dots, m$, since the real part of a complex number is at most its absolute value.

According to (80), we have

$$C \cdot su^\intercal = \frac{s}{\omega^{m'}} \sum_{i=0}^{d} \gamma_i \, C \cdot v(\omega\delta^i)^\intercal = \frac{s}{\omega^{m'}} \sum_{i=0}^{d} \gamma_i \, \lambda(\omega\delta^i) \, v(\omega\delta^i)^\intercal \tag{86}$$

since $\lambda(\omega\delta^i)$ are the eigenvalues paired with the eigenvectors $v(\omega\delta^i)$ for the matrix $C$. We know from (70) that each component of $v(\omega\delta^i)$ has absolute value 1 and hence, the absolute value of each component of $\lambda(\omega\delta^i)\, v(\omega\delta^i)$ is at most $1/s$ by (76). It follows that

$$\left| (C \cdot su^\intercal)_j \right| \leq \frac{s}{|\omega^{m'}|} \sum_{i=0}^{d} \frac{|\gamma_i|}{s} = \sum_{i=0}^{d} |\gamma_i| \qquad \text{for } j = 1, \dots, m \,. \tag{87}$$

For sufficiently large $m \geq m_0$, the roots of the polynomial $h$ defined by (77), are close to 1 in the complex plane, because of (73) whereas $d$ is a constant. Hence, the coefficients $\gamma_0, \dots \gamma_d$ of $h$ are close to the coefficients of the polynomial

$$(z-1)^d = (1-z)^d = \sum_{i=0}^{d} (-1)^{d-i} \binom{d}{i} z^i \tag{88}$$

for even $d$, which extends the bound (87) to

$$\left|(C \cdot su^{\intercal})_j\right| \leq \sum_{i=0}^{d} |\gamma_i| \leq \sum_{i=0}^{d} \left|(-1)^d \binom{d}{i}\right| + o(1) = 2^d + o(1) \tag{89}$$

for every $j = 1, \ldots, m$. This completes the proof of Lemma 6. $\square$

Furthermore, the tail value $r_n$ for $n = 1, \ldots, m$, defined by (68), can be computed by using (19) for $r_0 = 0$ and $p(\beta) = 0$, as

$$r_n = f_n(\beta) \tag{90}$$

where $f_n \in \mathbb{Z}[x]$ is an integer polynomial of degree at most $d-1$ such that

$$f_n(x) = \left(-\sum_{j=1}^{n} a_j x^{n-j}\right) \bmod p\,. \tag{91}$$

**Lemma 7.** *For $1 \leq n \leq m$, the coefficient of the monomial $x^{d-1}$ in $f_n$ is $-c_0 u'_{n-d+1}$.*

Proof. We multiply equation (84) by $-x^n$ and move the terms with the negative powers of $x$ from the left-hand side:

$$g_n(x) = -\sum_{j=1}^{n} a_j x^{n-j} = -\sum_{i=1}^{m-d} u'_i x^{n-i-d} p(x) + \sum_{j=1}^{m-n} a_{n+j} x^{-j} \tag{92}$$

which satisfies $f_n = g_n \bmod p$ according to (91). In order to analyze $f_n$, it is thus sufficient to consider the terms $u'_i x^{n-i-d} p(x)$ on the right-hand side of (92) that contain the non-negative powers of $x$. For $1 \leq i \leq n-d$, the term $u'_i x^{n-i-d} p(x) = \sum_{k=0}^{d} c_k u'_i x^{n-i-k}$ is a polynomial that is a multiple of self-reciprocal $p$. Hence,

$$\begin{aligned} f_n(x) &= \left(-\sum_{i=n-d+1}^{m-d} u'_i x^{n-i-d} p(x) + \sum_{j=1}^{m-n} a_{n+j} x^{-j}\right) \bmod p \\ &= \left(-\sum_{i=n-d+1}^{m-d} \sum_{k=0}^{d} c_k u'_i x^{n-i-k} + \sum_{j=1}^{m-n} a_{n+j} x^{-j}\right) \bmod p \end{aligned} \tag{93}$$

By (92), all the terms with a negative power of $x$ in the right-hand side of (93) cancel each other. Hence, we have

$$f_n(x) = -\sum_{i=n-d+1}^{n} \sum_{k=0}^{n-i} c_k u'_i x^{n-i-k} \tag{94}$$

which is a polynomial of degree at most $d-1$ achieved for $i = n-d+1$ and $k = 0$, corresponding to the coefficient $-c_0 u'_{n-d+1}$. This completes the proof of Lemma 7. $\square$

By Lemma 7 with $n = m' + d$, the coefficient of $x^{d-1}$ in $f_{m'+d}$ is $-c_0 u'_{m'+1}$. For any sufficiently large $m \geq m_0$ and $m'$ defined by (67), $\delta^{m'}$ is arbitrarily close to $-1$ in the complex plane, due to (73). Hence, the complex number $u_{m'+1} = h(\delta^{m'})$ derived from (81), is close to $h(-1)$ which is close to $2^d$ according to (88). It follows from (79) that $u'_{m'+1} \geq s(2^d - o(1)) - 1/2$ is unbounded with increasing $m$ since $s$ grows with $m$. Therefore, there are infinitely many different values of $u'_{m'+1}$, so there are infinitely many different polynomials $f_{m'+d}(x)$, and hence, also infinitely many different tail values $r_{m'+d} = f_{m'+d}(\beta)$ by (90), which completes the proof of Theorem 3. $\square$

Combining Theorem 1 and 3, we obtain the following corollary:

**Corollary 1.** *Let $\beta \in \mathbb{A} \cap \mathbb{R}$ be a real algebraic number. There exists a digit alphabet $A$ such that there is an eventually quasi-periodic $\beta$-expansion $a \in A^\omega$ with infinite $R(a)$ iff $\beta$ has a conjugate of absolute value $1$.*

Theorem 3 also provides the opposite implication to [19, Theorem 3], which means there exists an integer $M > 0$, such that there is no finite automaton recognizing the language

$$\left\{ a_m \ldots a_0 \in \{-M, \ldots, -1, 0, 1, \ldots, M\}^* \;\middle|\; \sum_{k=0}^{m} a_k \beta^k = 0 \right\} \tag{95}$$

iff a conjugate of $\beta$ lies on the unit circle.

## 4. Quasi-Periodic Numbers

In this section, we introduce and study so-called quasi-periodic numbers, which will be employed for characterizing the class of regular cut languages in Section 5. We say that a real number $c$ is *$\beta$-quasi-periodic within $A$* if *every* infinite $\beta$-expansion of $c$ is eventually quasi-periodic. Note that a number $c$ that has no $\beta$-expansion at all, or has, in addition, a finite $\beta$-expansion whereas $0 \notin A$, is also considered formally to be $\beta$-quasi-periodic, since this simplifies the formulation of statements concerning the quasi-periodic numbers. For example, the numbers from the complement of the Cantor set are formally 3-quasi-periodic within $\{0, 2\}$, since they have no 3-expansion for the alphabet $\{0, 2\}$.

We generalize the definition of set $R(a)$, introduced for a single $\beta$-expansion $a \in A^\omega$ in (20), to a real number $c \in \mathbb{R}$ as

$$\mathcal{R}_c = \bigcup_{a \in A^\omega \,:\, (a)_\beta = c} R(a) \tag{96}$$

which contains the tail values of all the infinite $\beta$-expansions of $c$. In addition, we introduce an alternative wider definition of such a set, namely,

$$\mathcal{R}'_c = \{ r_c(a) \,|\, I \leq r_c(a) \leq S \,,\, a \in A^* \} \tag{97}$$

which includes the normalized differences between $c$ and $(a)_\beta$,

$$r_c(a) = \beta^{|a|}(c - (a)_\beta) \tag{98}$$

that are in the admissible interval $[I, S]$, for all finite $\beta$-expansions $a \in A^*$, where $|a|$ stands for length of the string $a$ and

$$I = \inf_{a \in A^*} (a)_\beta\,, \qquad S = \sup_{a \in A^*} (a)_\beta\,. \tag{99}$$

Clearly, $\mathcal{R}_c \subseteq \mathcal{R}'_c$ since the tail values $r_n \in R(a)$ of any $\beta$-expansion $a = a_1a_2a_3\ldots \in A^\omega$ of $c$ can be expressed as $r_n = \beta^n(c-(a_1\ldots a_n)_\beta) = r_c(a_1\ldots a_n)$ for every $n \geq 0$, according to (98). On the other hand, if condition (4) is satisfied, then every real number within the interval $[I, S]$ has an infinite $\beta$-expansion. In particular, for any $a \in A^*$ such that $I \leq r_c(a) \leq S$, which means $r_c(a) \in \mathcal{R}'_c$, there exists $a' \in A^\omega$ that meets $r_c(a) = (a')_\beta$. Hence, $(aa')_\beta = (a)_\beta + \beta^{-|a|}r_c(a) = c$ by (98), which ensures $r_c(a) \in R(aa') \subseteq \mathcal{R}_c$ according to (96). Thus, $\mathcal{R}_c = \mathcal{R}'_c$ provided that condition (4) holds. In general, we will show below that for every $c \in \mathbb{R}$, the two sets, $\mathcal{R}_c$ and $\mathcal{R}'_c$, share the property of being finite.

Furthermore, for every $c \in \mathbb{R}$, we define a directed *transition graph* $\mathcal{G}_c = (\mathcal{R}'_c, \mathcal{E}_c)$ on the vertex set $\mathcal{R}'_c$ with the edges from $\mathcal{E}_c = \{(r_c(a), r_c(a\alpha)) \in (\mathcal{R}'_c)^2 \,|\, a \in A^*\,,\, \alpha \in A\}$. Each edge $(r_c(a), r_c(a\alpha)) \in \mathcal{E}_c$ is labeled with the digit $\alpha \in A$ which satisfies the recurrence condition (19) as

$$r_c(a\alpha) = \beta r_c(a) - \alpha \tag{100}$$

according to (98). We also denote by $\mathcal{P}_c$ the set of possible periodic points of all eventually periodic $\beta$-expansions of $c$, which meets $\mathcal{P}_c \subseteq \mathcal{R}_c \subseteq \mathcal{R}'_c$ by Lemma 1.

We first illustrate these definitions by the following elaborated examples, before formulating the theorems concerning the quasi-periodic numbers at the end of this section.

**Example 4.** We present an example of a $\beta$-quasi-periodic number $c$ within $A = \{0, 1\}$. Let $\beta \approx 1.722084$ be the real root of the polynomial $x^4 - x^3 - x^2 - x + 1$, satisfying

$$\beta^4 - \beta^3 - \beta^2 - \beta + 1 = 0\,, \tag{101}$$

such that $\beta > 1$, which is a Salem number. We define a real number

$$c = \frac{1}{9}\left(4\beta^3 - 2\beta^2 - 2\beta - 5\right) \approx 0.672505 \tag{102}$$

whose all $\beta$-expansions are generated by the $\omega$-regular expression

$$(100010 + 011(011101)^*100)^\omega + (100010 + 011(011101)^*100)^*011(011101)^\omega\,, \tag{103}$$

and prove to be eventually quasi-periodic, which ensures that $c$ is $\beta$-quasi-periodic within $A$. This can be shown by constructing a corresponding transition graph $\mathcal{G}_c = (\mathcal{R}'_c, \mathcal{E}_c)$ which is depicted in Figure 2. Note that $\mathcal{R}'_c = \mathcal{R}_c$ because

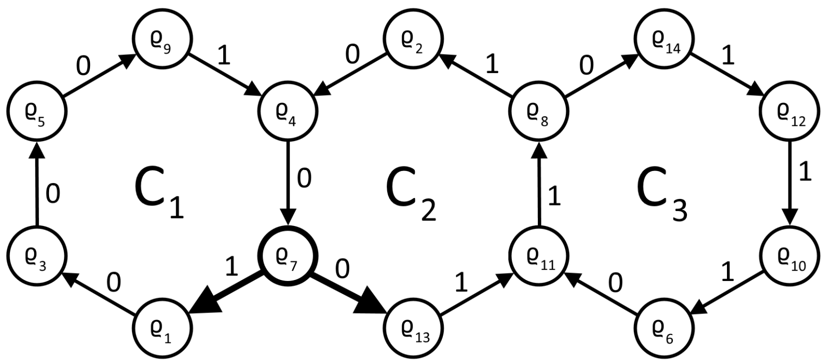


Figure 2: The transition graph $\mathcal{G}_c$ for $c = \frac{1}{9}\left(4\beta^3 - 2\beta^2 - 2\beta - 5\right)$ when $A = \{0, 1\}$ and $\beta > 1$ is the Salem number satisfying $\beta^4 - \beta^3 - \beta^2 - \beta + 1 = 0$.

condition (4) is met, which also guarantees that exactly the real numbers in the interval $[I, S]$ where $I = 0$ and $S = 1/(\beta - 1) \approx 1.384881$, including $c$, have $\beta$-expansions.

In particular, we start with the vertex $\varrho_7 = r_c(\varepsilon) = c \in \mathcal{R}'_c$ which is connected to the two vertices $\varrho_{13} = r_c(0) \in \mathcal{R}'_c$ and $\varrho_1 = r_c(1) \in \mathcal{R}'_c$ via the directed edges $(\varrho_7, \varrho_{13}) \in \mathcal{E}_c$ and $(\varrho_7, \varrho_1) \in \mathcal{E}_c$, labeled with 0 and 1, respectively, as indicated in boldface in Figure 2, whereas the recurrence (100) and condition (101) produce

$$\begin{aligned} \varrho_{13} = r_c(0) = \beta c &= \tfrac{1}{9}\left(4\beta^4 - 2\beta^3 - 2\beta^2 - 5\beta\right) \\ &= \tfrac{1}{9}\left(2\beta^3 + 2\beta^2 - \beta - 4\right) \approx 1.158110 \end{aligned} \tag{104}$$

by using

$$\left(4x^4 - 2x^3 - 2x^2 - 5x\right) \bmod (x^4 - x^3 - x^2 - x + 1) = \left(2x^3 + 2x^2 - x - 4\right), \tag{105}$$

and

$$\begin{aligned} \varrho_1 = r_c(1) = \beta c - 1 &= \tfrac{1}{9}\left(4\beta^4 - 2\beta^3 - 2\beta^2 - 5\beta - 9\right) \\ &= \tfrac{1}{9}\left(2\beta^3 + 2\beta^2 - \beta - 13\right) \approx 0.158110\,, \end{aligned} \tag{106}$$

falling in the interval $[I, S]$. Further, we expand the vertex $\varrho_{13} = r_c(0) \in \mathcal{R}'_c$ through the directed edge $(\varrho_{13}, \varrho_{11}) \in \mathcal{E}_c$ labeled with 1, leading to vertex

$$\varrho_{11} = r_c(01) = \beta r_c(0) - 1 = \tfrac{1}{9}(4\beta^3 + \beta^2 - 2\beta - 11) \approx 0.994363\,, \tag{107}$$

while there is no edge with the label 0 outgoing from $\varrho_{13}$ since $r_c(00) \notin \mathcal{R}'_c$ due to $r_c(00)$, for which $S < r_c(00) = \beta r_c(0) \approx 1.994363$, has no $\beta$-expansion, etc. This procedure of applying the recurrence condition (100) eventually converges to the finite transition graph $\mathcal{G}_c$ presented in Figure 2, containing 14 vertices,

$\mathcal{R}'_c = \{\varrho_1, \varrho_2, \ldots, \varrho_{14}\}$, where

$$\begin{aligned}
\varrho_1 &= \tfrac{1}{9}\left(2\beta^3 + 2\beta^2 - \beta - 13\right) & \varrho_8 &= \tfrac{1}{9}\left(5\beta^3 + 2\beta^2 - 7\beta - 13\right) \\
\varrho_2 &= \tfrac{1}{9}\left(7\beta^3 - 2\beta^2 - 8\beta - 14\right) & \varrho_9 &= \tfrac{1}{9}\left(-2\beta^3 + 7\beta^2 + \beta - 5\right) \\
\varrho_3 &= \tfrac{1}{9}\left(4\beta^3 + \beta^2 - 11\beta - 2\right) & \varrho_{10} &= \tfrac{1}{9}\left(4\beta^3 + 7\beta^2 - 11\beta - 14\right) \\
\varrho_4 &= \tfrac{1}{9}\left(5\beta^3 - \beta^2 - 7\beta - 7\right) & \varrho_{11} &= \tfrac{1}{9}\left(4\beta^3 + \beta^2 - 2\beta - 11\right) \\
\varrho_5 &= \tfrac{1}{9}\left(5\beta^3 - 7\beta^2 + 2\beta - 4\right) & \varrho_{12} &= \tfrac{1}{9}\left(5\beta^3 - \beta^2 + 2\beta - 16\right) \\
\varrho_6 &= \tfrac{1}{9}\left(11\beta^3 - 7\beta^2 - 10\beta - 13\right) & \varrho_{13} &= \tfrac{1}{9}\left(2\beta^3 + 2\beta^2 - \beta - 4\right) \\
\varrho_7 &= \tfrac{1}{9}\left(4\beta^3 - 2\beta^2 - 2\beta - 5\right) & \varrho_{14} &= \tfrac{1}{9}\left(7\beta^3 - 2\beta^2 - 8\beta - 5\right).
\end{aligned} \tag{108}$$

It appears that $\mathcal{G}_c$ is composed of three directed cycles of length 6, $C_1 = \varrho_7, \varrho_1, \varrho_3, \varrho_5, \varrho_9, \varrho_4$, $C_2 = \varrho_7, \varrho_{13}, \varrho_{11}, \varrho_8, \varrho_2, \varrho_4$, and $C_3 = \varrho_8, \varrho_{14}, \varrho_{12}, \varrho_{10}, \varrho_6, \varrho_{11}$.

One can check that the sequences of edge labels on the infinite directed walks through graph $\mathcal{G}_c$, starting at vertex $\varrho_7 = c \in \mathcal{R}'_c$, which correspond to eventually quasi-periodic $\beta$-expansions of $c$, are characterized exactly by the $\omega$-regular expression (103). In particular, the first summand in (103) describes the walks that traverse the cycles $C_1$ or $C_2$ with the edge labels 100010 or 011 000, respectively, for infinitely many times, while the walk through the cycle $C_2$ can be interrupted in the middle by several passes through the cycle $C_3$, which inserts the edge labels $(011101)^*$ in $011(011101)^*000$. The second summand in (103) denotes the same walks restricted to finitely many iterations of cycles $C_1$ and $C_2$ possibly interrupted by several passes through $C_3$, which eventually end up in the infinite loop through the cycle $C_3$, producing the edge labels $(011101)^\omega$.

Moreover, for each vertex $\varrho$ from (108), there exists such a walk whose vertices create a corresponding tail sequence, traversing $\varrho$ infinitely many times, which implies $\mathcal{P}_c = \mathcal{R}'_c$ by Lemma 1. For instance, for the periodic point $\varrho_{12} \in \mathcal{P}_c$, the eventually quasi-periodic $\beta$-expansion

$$a = 01101(110101\,110110001101)^\omega \tag{109}$$

of $c$, corresponding to the infinite walk

$$\begin{aligned}
&\varrho_7, \varrho_{13}, \varrho_{11}, \varrho_8, \varrho_{14}, \\
&\quad(\boldsymbol{\varrho_{12}}, \varrho_{10}, \varrho_6, \varrho_{11}, \varrho_8, \varrho_{14}, \boldsymbol{\varrho_{12}}, \varrho_{10}, \varrho_6, \varrho_{11}, \varrho_8, \varrho_2, \varrho_4, \varrho_7, \varrho_{13}, \varrho_{11}, \varrho_8, \varrho_{14})^\omega
\end{aligned} \tag{110}$$

which is in fact the tail sequence of $a$, is composed of the preperiodic part 01101 and the two quasi-repetends 110101 and 110110001101. In general, there are uncountably many eventually quasi-periodic $\beta$-expansions of $c$ with each periodic point $\varrho \in \mathcal{P}_c$, which are composed of quasi-repetends whose length is divisible by 6.

In addition, $c$ is also an example of a $\beta$-quasi-periodic number within $A$ such that the number of quasi-repetends of length $6n$, which occur in its eventually quasi-periodic $\beta$-expansions, can be constant, linear, or exponential in terms of $n$, depending on the choice of periodic point $\varrho \in \mathcal{P}_c$. For the periodic points that the cycle $C_2$ shares with another cycle in $\mathcal{G}_c$, namely $C_2 \cap (C_1 \cup C_3) = \{\varrho_4, \varrho_7, \varrho_8, \varrho_{11}\} \subset \mathcal{R}'_c$, there is only a *constant* number of quasi-repetends of length $6n$. For instance, consider the periodic point $\varrho_8 \in \mathcal{P}_c$, associated with a preperiodic part of the form $(100010)^k 011$, for $k \geq 0$, corresponding to the directed walk $C_1^k, \varrho_7, \varrho_{13}, \varrho_{11}, \varrho_8$ in $\mathcal{G}_c$ from $\varrho_7$ to $\varrho_8$, which first traverses the cycle $C_1$ for $k$ times. For $n \geq 2$, we have only one quasi-repetend of length $6n$, namely $100(100010)^{n-1}011$, corresponding to the directed closed walk $\varrho_8, \varrho_2, \varrho_4, C_1^{n-1}, \varrho_7, \varrho_{13}, \varrho_{11}, \varrho_8$, which first traverses the path from $\varrho_8$ to $\varrho_7$, then passes through the cycle $C_1$ for $n-1$ times, and finally comes back from $\varrho_7$ to $\varrho_8$. Note that for $n = 1$, we have two quasi-repetends of length 6, 100011 and 011101, corresponding to the cycle $C_2$ and $C_3$, respectively.

For the periodic points from cycle $C_2$ that are not shared by another cycle in $\mathcal{G}_c$, namely $C_2 \setminus (C_1 \cup C_3) = \{\varrho_2, \varrho_{13}\} \subset \mathcal{R}'_c$, there is a *linear* number of quasi-repetends of length $6n$. For example, consider the periodic point $\varrho_{13} \in \mathcal{P}_c$, associated with a preperiodic part of the form $(100010)^k 0$, for $k \geq 0$, corresponding to the directed walk $C_1^k, \varrho_7, \varrho_{13}$ in $\mathcal{G}_c$. For $n \geq 1$, we can divide $n-1$ into two nonnegative integer summands in $n$ ways as $n - 1 = n_1 + n_3$ where $n_1, n_3 \geq 0$, and thus, we have $n$ quasi-repetends of length $6n$, namely $11(011101)^{n_3}100(100010)^{n_1}0$, corresponding to the directed closed walk $\varrho_{13}, \varrho_{11}, C_3^{n_3}, \varrho_8, \varrho_2, \varrho_4, C_1^{n_1}, \varrho_7, \varrho_{13}$.

Finally, for the remaining periodic points in $\mathcal{G}_c$ outside the cycle $C_2$, namely $\mathcal{R}'_c \setminus C_2 = \{\varrho_1, \varrho_3, \varrho_5, \varrho_6, \varrho_9, \varrho_{10}, \varrho_{12}, \varrho_{14}\}$, there is an *exponential* number of quasi-repetends of length $6n$. For instance, the quasi-repetends for the periodic point $\varrho_1 \in \mathcal{P}_c$, are characterized by the regular expression

$$00010(011(011101)^*100)^*1 \tag{111}$$

which corresponds to the directed closed walks in $\mathcal{G}_c$ that first traverse the path from $\varrho_1$ to $\varrho_7$ and then either move directly back to $\varrho_1$ (completing the cycle $C_1$) or, before that, follow the cycle $C_2$ several times. In addition, the cycle $C_2$ includes vertex $\varrho_8$, at which the walk through $C_2$ can be interrupted by several passes of cycle $C_3$. Clearly, there is only one quasi-repetend 000101 of length 6, according to (111). For a quasi-repetend of length $6n$ where $n \geq 2$, if the cycle $C_2$ is traversed $k$ times which means $1 \leq k \leq n-1$, then the cycle $C_3$ is traversed $n-k-1$ times altogether. The number of ways of distributing the $n-k-1$ passes of $C_3$ among the $k$ passes of $C_2$ is $\binom{n-2}{k-1}$. Hence, for $n \geq 2$, the number of quasi-repetends of length $6n$ is

$$\sum_{k=1}^{n-1} \binom{n-2}{k-1} = 2^{n-2} \,. \tag{112}$$

Another example of a $\beta$-quasi-periodic number within $A = \{0, 1\}$ has been presented in Example 2 (in contrast to Example 1 where 1 is not $\beta$-quasi-periodic) for the plastic constant $\beta$ satisfying (24), as it can be shown by generating a corresponding transition graph $\mathcal{G}_c$ for $c = 1$. In addition, there are also $\beta$-quasi-periodic numbers within $A = \{0, 1\}$ for the base $\beta$ that is neither Pisot nor Salem number. In particular, let $\beta \approx 1.685137$ be the unique real root of the polynomial $x^5 - x^4 - x^2 - x - 1$, whose some Galois conjugates are in absolute value greater than 1, which means $\beta$ is neither Pisot nor Salem number. In this case, all the $\beta$-expansions of the real number $c = \frac{1}{3}(-\beta^4 + 3\beta^3 - 2\beta - 1) \approx 0.640563$, which are generated by the $\omega$-regular expression $(10000 + 01(01111)^*10)^\omega + (10000 + 01(01111)^*10)^*01(01111)^\omega$, are eventually quasi-periodic. This can again be verified by constructing the transition graph $\mathcal{G}_c$ where $\mathcal{R}'_c = \mathcal{R}_c = \mathcal{P}_c$ contains now 10 vertices.

**Example 5.** On the other hand, we present examples of real numbers $c$ that are not $\beta$-quasi-periodic within $A = \{0, 1\}$, although their greedy and/or lazy $\beta$-expansion is eventually periodic. Let $\beta = \sqrt{2} \approx 1.414214$ (cf. Example 3) which satisfies condition (4). Hence, the real numbers in the interval $[I, S]$ where $I = 0$ and $S = \beta + 1 \approx 2.414214$, have $\beta$-expansions.

For instance, consider $c = \frac{1}{2}(\beta + 1) \approx 1.207107$ whose eventually periodic greedy $\beta$-expansion is $a = 110^\omega$ by the following reason. The second element of its tail sequence is

$$r_2 = r_c(11) = \beta^2 \left(\tfrac{1}{2}(\beta + 1) - \beta^{-1} - \beta^{-2}\right) = 0 = I \qquad (113)$$

according to (98) and $\beta^2 = 2$. This coincides with the infimum (99), which allows only for one continuation $0^\omega$. Similarly, $a' = 001^\omega$ is the eventually periodic lazy $\beta$-expansion of $c$ due to $r'_2 = r_c(00) = \beta + 1 = S$ is the supremum, for which the only possible continuation is $1^\omega$. Nevertheless, $c$ is not $\beta$-quasi-periodic within $A$ since the $\beta$-expansions of $c$ with the prefix 0111 are not eventually quasi-periodic because the tail value $r_c(0111) = \beta - 1$ has no eventually periodic $\beta$-expansion according to Example 3.

In addition, $c = \frac{1}{3}$ is an example of the number with the periodic greedy $\beta$-expansion $(0001)^\omega$ due to

$$r_c(1) < r_c(01) < r_c(001) < 0 = I \leq r_c(0001) = \tfrac{1}{3} = c \leq S = \beta + 1\,, \qquad (114)$$

whose lazy $\beta$-expansion $a = a_1 a_2 a_3 \ldots \in A^\omega$ is not eventually periodic. Recall that any lazy $\beta$-expansion is eventually quasi-periodic iff it is eventually periodic. First observe that $a$ starts with the prefix $0^5 1$ since

$$I \leq r_c(0) < r_c(00) < \cdots < r_c\left(0^5\right) < r_c\left(0^5 1\right) \leq S < r_c\left(0^6\right)\,. \qquad (115)$$

We prove that the continuation of $a$ is not periodic by the same argument as in Example 3. Denote $\beta_1 = \beta$ and $\beta_2 = -\beta$. Let $f_n \in \mathbb{Q}[x]$, for $n \geq 0$, be the sequence of rational polynomials of degree at most 1, which is defined recursively by (31) starting with $f_0(x) = c = \frac{1}{3}$, and satisfies $r_n = f_n(\beta_1)$, where $(r_n)_{n=0}^\infty$ is the tail sequence of $a$. Consider the sequence $r'_n = f_n(\beta_2)$ which coincides with $r_n = r'_n$ for even $n$. By the argument of Example 3, it follows from $r'_4 = r_4 = r_c(0000) = \frac{4}{3} \notin [\beta_2, 1]$ that the tail values $r_n$ for $n \geq 4$ are pairwise distinct. Hence, the lazy $\beta$-expansion $a$ is not eventually periodic according to Lemma 1.

The characterization of a single eventually quasi-periodic $\beta$-expansion $a \in A^\omega$ in Theorems 1 and 2 which employs the finiteness of $R(a)$, can be generalized to any $\beta$-quasi-periodic number $c$ within $A$ for arbitrary real bases and digits by using the finiteness of $\mathcal{R}'_c$ respectively $\mathcal{R}_c$.

**Theorem 4.** *The following four conditions are equivalent:*

(i) $\mathcal{R}'_c$ *is finite.*
(ii) $\mathcal{R}_c$ *is finite.*
(iii) $c$ *is a* $\beta$*-quasi-periodic number within* $A$*.*
(iv) *Every infinite* $\beta$*-expansion of c has at least two tails of the same value.*

Proof. Let $\mathcal{R}'_c$ be a finite set which ensures that $\mathcal{R}_c$ is also finite due to $\mathcal{R}_c \subseteq \mathcal{R}'_c$. Hence, $R(a)$ is finite for every $\beta$-expansion $a \in A^\omega$ of $c$ which is thus eventually quasi-periodic according to Lemma 1. It follows that $c$ is $\beta$-quasi-periodic within $A$, whose every infinite $\beta$-expansion has at least two tails of the same value by Lemma 1. We have (i) $\implies$ (ii) $\implies$ (iii) $\implies$ (iv).

In the rest of the proof we show (iv) $\implies$ (i) by contrapositive. Assume that $\mathcal{R}'_c$ is infinite. Consider a directed tree $T = (V, E)$ with vertex set

$$V = \{v \in A^* \,|\, r_c(v) \in \mathcal{R}'_c\} \tag{116}$$

where $r_c(v)$ is defined by (98), which includes the empty string $\varepsilon$ as a root satisfying $r_c(\varepsilon) = c$. The set of directed edges $E$ is defined as

$$E = \left\{(u, v) \in V^2 \,|\, (\exists\, \alpha \in A)\, v = u\alpha\right\} \tag{117}$$

which guarantees the outdegree of $T$ is bounded by $|A|$. Clearly, the length $|v|$ of string $v \in A^*$ determines the level of vertex $v \in V$ in $T$.

Let $T' = (V', E')$ be a directed rooted tree that is a subgraph of $T$ with a maximal vertex subset $V' \subseteq V$ so that $\varepsilon \in V'$ and for every two vertices $v_1, v_2 \in V'$,

$$v_1 \neq v_2 \implies r_c(v_1) \neq r_c(v_2)\,. \tag{118}$$

The tree $T'$ can be constructed inductively level by level as follows. We start at zero level and include $\varepsilon$ into $V'$. In a general step when $T'$ has been constructed up to some level $\ell \geq 0$, we connect only those vertices $v \in V$ at level $|v| = \ell + 1$ to $T'$ via edges $(u, v) \in E$ from vertices $u \in V'$ at level $|u| = \ell$, that preserve condition (118). In particular, for each value

$$r \in \{r_c(v) \,|\, (u, v) \in E, u \in V' \,\&\, |u| = \ell\} \setminus \{r_c(v') \,|\, v' \in V' \,\&\, |v'| \leq \ell\}\,, \tag{119}$$

we include one representative vertex $v \in V$ with $r_c(v) = r$ at level $|v| = \ell + 1$ into $V'$.

We show that for any $r \in \mathcal{R}'_c$ there is $v \in V'$ such that $r_c(v) = r$. On the contrary, suppose $a_1 \dots a_n \in V \setminus V'$ is a vertex at minimal level $n$, satisfying

$$r_c(v) \neq r_c(a_1 \dots a_n) = r \in \mathcal{R}'_c \quad \text{for every } v \in V'\,. \tag{120}$$

Clearly, $a_1 \dots a_{n-1} \in V \setminus V'$ since otherwise vertex $a_1 \dots a_n$ could be included into $V'$ according to (119). By the minimality of $n$, we know there is $a'_1 \dots a'_m \in V'$ such that $r_c(a'_1 \dots a'_m) = r_c(a_1 \dots a_{n-1})$. Thus, we have $r_c(a'_1 \dots a'_m a_n) = r$ and $a'_1 \dots a'_m a_n \in V'$ by (119), which is in contradiction with (120).

It follows that $\{r_c(v) \,|\, v \in V'\} = \mathcal{R}'_c$ implying $T'$ is infinite. According to König's lemma, there exists an infinite directed path in $T'$ corresponding to a $\beta$-expansion of $c$ whose tail sequence contains pairwise distinct values. $\square$

For Pisot bases and digits from $\mathbb{Q}(\beta)$, the following theorem proves $\mathrm{QPer}(\beta) = \mathrm{Per}(\beta) \subseteq \mathbb{Q}(\beta) \subseteq \overline{\mathrm{QPer}}(\beta)$ where $\overline{\mathrm{QPer}}(\beta)$ denotes the set of $\beta$-quasi-periodic numbers within $A$, including those with no infinite $\beta$-expansion. Recall that $\mathrm{QPer}(\beta)$ contains only the quasi-periodic numbers that have an infinite $\beta$-expansion (see Paragraph 1.2).

**Theorem 5.** *Let $\beta$ be a Pisot number and assume $A \subset \mathbb{Q}(\beta)$. Then any number $c \in \mathbb{Q}(\beta)$ is $\beta$-quasi-periodic within $A$.*

Proof. If there is no infinite $\beta$-expansion of $c$, then $c$ is $\beta$-quasi-periodic within $A$ by definition. Thus, let $a \in A^\omega$ be any $\beta$-expansion of $c$ and let $(r_n)_{n=0}^\infty$ be its tail sequence. The base $\beta$ satisfies the assumption of Theorem 1 due to $\beta$ is a Pisot number which is even an *algebraic integer*. Thus, the proof further proceeds in exactly the same way as the proof of Claim 1 simplified to algebraic integer $\beta$ (e.g. avoiding Lemma 2 and 3). In particular, the digits in $A \subset \mathbb{Q}(\beta) = \mathbb{Q}[\beta]$ and the number $c = r_0 \in \mathbb{Q}(\beta) = \mathbb{Q}[\beta]$ are now multiplied by a suitable natural number $\gamma > 0$ so that $A' = \{\gamma\alpha \,|\, \alpha \in A\} \subset \mathbb{Z}[\beta]$ and $c' = r'_0 = \gamma c \in \mathbb{Z}[\beta]$, respectively, which provides a $\beta$-expansion of $c' = (a')_\beta$ and its tail sequence $(r'_n)_{n=0}^\infty$ satisfying (38). In addition, a sequence of integer polynomials $f_n \in \mathbb{Z}[x]$ which meets (43), is introduced by (42) where $p \in \mathbb{Z}[x]$ is now an integer monic polynomial of algebraic integer $\beta$. The integer coefficients of $f_n$ are bounded by (51) according to Lemma 4 since the Galois conjugates of Pisot $\beta$ lie inside the unit circle. Consequently, there is only a finite number of different polynomials $f_n$, and hence, also a finite number of $r'_n$. This means that $R(a)$ is finite, which implies that $a$ is eventually quasi-periodic according to Lemma 1. It follows that $c$ is $\beta$-quasi-periodic within $A$. $\square$

## 5. Regular Cut Languages

In this section we prove a necessary and sufficient condition for a cut language $L_{<c}$ to be regular by using Myhill-Nerode theorem.

**Theorem 6.** *A cut language $L_{<c} \subseteq A^*$ with base $\beta$ is regular iff $c$ is $\beta$-quasi-periodic within $A$.*

Proof. According to Theorem 4, $c$ is $\beta$-quasi-periodic within $A$ iff $\mathcal{R}'_c$ is finite. First assume that $\mathcal{R}'_c$ is finite. We introduce an equivalence relation $\sim$ on $A^*$ so that for any $u, v \in A^*$, we define $u \sim v$ iff $\beta^{|uv|} > 0$ and either $r_c(u) = r_c(v) \in \mathcal{R}'_c$ or $\max(r_c(u), r_c(v)) < I$ or $\min(r_c(u), r_c(v)) > S$. Obviously, we have only finitely many equivalence classes due to $\mathcal{R}'_c$ is finite. In order to prove that language $L_{<c}$ is regular we employ Myhill-Nerode theorem by showing that for any $u, v \in A^*$, if $u \sim v$, then for every $w \in A^*$, $uw \in L_{<c}$ iff $vw \in L_{<c}$.

Thus, let $u, v \in A^*$ meet $u \sim v$, and $w \in A^*$. We have $uw \in L_{<c}$ iff $(u)_\beta + \beta^{-|u|}(w)_\beta = (uw)_\beta < c$ iff $\beta^{-|u|}(w)_\beta < c - (u)_\beta = \beta^{-|u|} r_c(u)$. This, multiplied by $\beta^{|uv|} > 0$, reduces to

$$\beta^{|v|}(w)_\beta < \beta^{|v|} r_c(u)\,. \tag{121}$$

Similarly, we have $vw \in L_{<c}$ iff

$$\beta^{|u|}(w)_\beta < \beta^{|u|} r_c(v)\,. \tag{122}$$

Moreover, observe that $\beta^{|u|} > 0$ iff $\beta^{|v|} > 0$ due to $\beta^{|uv|} > 0$. Hence, (121) and (122) are equivalent, if $r_c(u) = r_c(v) \in \mathcal{R}'_c$ or $\max(r_c(u), r_c(v)) < I \leq (w)_\beta$ or $\min(r_c(u), r_c(v)) > S \geq (w)_\beta$. Hence, $uw \in L_{<c}$ iff $vw \in L_{<c}$ and this implies that $L_{<c}$ is regular.

Conversely, let $L_{<c}$ be a regular language. According to Myhill-Nerode theorem, there is an equivalence relation $\sim$ on $A^*$ with finitely many equivalence classes such that for any $u, v \in A^*$, if $u \sim v$, then for every $w \in A^*$, $uw \in L_{<c}$ iff $vw \in L_{<c}$. Assume to the contrary that $c$ is not $\beta$-quasi-periodic within $A$, which means there is an infinite $\beta$-expansion $a = a_1 a_2 a_3 \ldots \in A^\omega$ of $c$ whose tail sequence contains pairwise distinct values, according to Theorem 4. It follows that there exist two prefixes $u, v \in A^*$ of $a$, having even length, such that $u \sim v$ and $(u')_\beta = r_c(u) > r_c(v) = (v')_\beta$ where $a = uu' = vv'$ for some $u', v' \in A^\omega$, which implies

$$(uv')_\beta = (u)_\beta + \beta^{-|u|}(v')_\beta < (u)_\beta + \beta^{-|u|}(u')_\beta = (uu')_\beta = c = (vv')_\beta \tag{123}$$

due to $\beta^{-|u|} > 0$. By Lemma 1, we know the $\beta$-expansion $a$ is not eventually periodic, and hence, there is an increasing infinite subsequence of indices, $|v| < k_1 < k_2 < k_3 < \cdots$, satisfying $a_{k_j}\beta^{-k_j} < \max_{\alpha \in A} \alpha\beta^{-k_j}$ for every $j \geq 1$. Denote by $w_j \in A^\omega$ a modified infinite word $v'$ in which the $(k_j - |v|)$-th position is replaced by $\arg\max_{\alpha \in A} \alpha\beta^{-k_j}$, which implies $(vw_j)_\beta > (vv')_\beta = c$ for every $j \geq 1$ and the difference $(w_j)_\beta - (v')_\beta$ is arbitrarily small, if $j$ tends to infinity. Since $(uv')_\beta < (uu')_\beta$ by (123), we achieve $(uw_{j_0})_\beta < (uu')_\beta = c$ for a sufficiently large $j_0 \geq 1$. Thus, $(uw_{j_0})_\beta < c < (vw_{j_0})_\beta$, and hence we have $(uw)_\beta < c < (vw)_\beta$ for a sufficiently long prefix $w \in A^*$ of $w_{j_0}$, which implies $uw \in L_{<c}$ and $vw \notin L_{<c}$, contradicting $u \sim v$. $\square$

**Example 6.** Obviously, not every regular language is a cut language. This can be illustrated by any regular language $L \subset A^*$ where $\{\alpha_1, \alpha_2\} \subseteq A$ such that $L \cap \{\alpha_1, \alpha_2\}^2 = \{\alpha_1\alpha_2, \alpha_2\alpha_1\}$. For $A = \{0, 1\}$, this corresponds to the XOR characteristic function on two-bit words which is often used as a counterexample in neural networks. In particular, assume to the contrary that $L = L_{<c}$ is a cut language for some threshold $c \in \mathbb{R}$. Hence, $\alpha_1\alpha_2, \alpha_2\alpha_1 \in L_{<c}$ implies $(\alpha_1\alpha_2)_\beta < c$ and $(\alpha_2\alpha_1)_\beta < c$, which sums to

$$\alpha_1\beta^{-1} + \alpha_2\beta^{-2} + \alpha_2\beta^{-1} + \alpha_1\beta^{-2} < 2c\,. \tag{124}$$

On the other hand, we know $\alpha_2\alpha_2 \notin L_{<c}$, and thus $(\alpha_2\alpha_2)_\beta \geq c$ which can be subtracted from inequality (124), producing

$$\alpha_1\beta^{-1} + \alpha_1\beta^{-2} < c\,, \tag{125}$$

but this is in contradiction with $\alpha_1\alpha_1 \notin L_{<c}$.

## 6. Non-Context-Free Cut Languages

In this section we show that a cut language $L_{<c}$ is not context-free if its threshold $c$ is not $\beta$-quasi-periodic within $A$. According to Theorem 6, this means that the cut language $L_{<c}$ is context-free iff it is regular. For this purpose, we use a pumping technique introduced in the following lemma. In particular, we say that an infinite word $a \in A^\omega$ is *approximable* in a language $L \subseteq A^*$, if for every finite prefix $u \in A^*$ of $a$, there is $x \in A^*$ such that $ux \in L$.

**Lemma 8.** *Let $a \in A^\omega$ be approximable in a context-free language $L \subseteq A^*$. Then there is a decomposition $a = uvw$ where $u, v \in A^*$ and $w \in A^\omega$, such that $|v| > 0$ is even and for every integer $i \geq 0$, the word $uv^iw$ is approximable in $L$.*

Proof. Consider a context-free grammar $G$ for $L$ in Greibach normal form such that for every nonterminal $N$ of $G$, there is a derivation of a terminal word from $N$. Since $a$ is approximable in $L = L(G)$, there is a left derivation from the start symbol, $S \Rightarrow \ldots \Rightarrow u_n\nu_n$ for every $n$, such that $u_n \in A^n$ is the prefix of $a$ of length $n$, and $\nu_n$ 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 $u\nu$ such that $u$ is a prefix of $a$, and the edges outcoming from $u\nu$ correspond to an application of one production rule to the left-most nonterminal in $\nu$. The degree of each vertex is bounded by the number of production rules. According to König's lemma, there is an infinite left derivation $S \Rightarrow \ldots \Rightarrow u_n\nu_n \Rightarrow \ldots$ such that for every $n$, $u_n$ is the prefix of $a$ of length $n$, and $\nu_n$ is a non-empty sequence of nonterminal symbols.

Let us call an occurrence of a nonterminal in $\nu_n$ *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 $\nu_m$ contains exactly one temporary nonterminal. We know the left-most nonterminal $N_1$ in $\nu_n = N_1 \ldots N_i \ldots N_k$ is temporary, and let $N_i$ be the right-most temporary nonterminal in $\nu_n$. If $i = 1$, then choose $m = n$. For $i \geq 2$, there is an index $m$, such that all the temporary nonterminals $N_1, \ldots, N_{i-1}$ in $\nu_n$ are transformed into terminal words in $u_m$. If $m$ is the smallest such index, then $N_i$ is the first and the only temporary nonterminal of $\nu_m$. It follows that there is an infinite number of indices $n$ such that $\nu_n$ 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 $u_{m_1}\nu_{m_1} = u'N\mu'_1$, $u_{m_2}\nu_{m_2} = u'v_1N\mu'_2\mu'_1$, $u_{m_3}\nu_{m_3} = u'v_1v_2N\mu'_3\mu'_2\mu'_1$ for some nonterminal $N$, where $u', v_1$, $v_2 \in A^*$, $|v_1| > 0$, $|v_2| > 0$, and $\mu'_1, \mu'_2, \mu'_3$ consist of stable nonterminals in all $\nu_{m_1}, \nu_{m_2}, \nu_{m_3}$. If $|v_1|$ is even, then define $n_1 = m_1$, $n_2 = m_2$, $u = u'$, $v = v_1$,

$\mu_1 = \mu'_1$, and $\mu_2 = \mu'_2$, otherwise, if $|v_2|$ is even, then $n_1 = m_2$, $n_2 = m_3$, $u = u'v_1$, $v = v_2$, $\mu_1 = \mu'_2\mu'_1$, and $\mu_2 = \mu'_3$. On the other hand, if $|v_1|$ and $|v_2|$ are both odd, then $|v_1 v_2|$ is even and define $n_1 = m_1$, $n_2 = m_3$, $u = u'$, $v = v_1 v_2$, $\mu_1 = \mu'_1$, and $\mu_2 = \mu'_3\mu'_2$.

Thus, there are two words $u, v \in A^*$ such that $u_{n_1}\nu_{n_1} = uN\mu_1$, $u_{n_2}\nu_{n_2} = uvN\mu_2\mu_1$, and $|v| > 0$ is even, where $N \stackrel{*}{\Rightarrow} vN\mu_2$. For every $m \geq n_2$, we have $u_m\nu_m = uv\xi_m\mu_2\mu_1$ where $\xi_m$ is such that $N \stackrel{*}{\Rightarrow} \xi_m$. Hence, an infinite word $w \in A^\omega$ is produced from $N$, such that $a = uvw$. Clearly, every finite prefix of $w$ is the terminal part of $\xi_m$ 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 $u_{n_1}\nu_{n_1} = uN\mu_1$. Then, the derivation $N \stackrel{*}{\Rightarrow} vN\mu_2$ is used $i$ times. Finally, the derivations $N \stackrel{*}{\Rightarrow} \xi_m$ are used in an infinite sequence for all $m > n_2$. Altogether, we obtain

$$S \stackrel{*}{\Rightarrow} uN\mu_1 \stackrel{*}{\Rightarrow} uv^iN\mu_2^i\mu_1 \Rightarrow \ldots \Rightarrow uv^i\xi_m\mu_2^i\mu_1 \Rightarrow \ldots \quad \text{for all } m > n_2\,. \tag{126}$$

We show that for every $i \geq 0$, the infinite sequence $uv^iw$ is approximable in $L$. For any prefix $x \in A^*$ of $uv^iw$, we employ the derivation (126) until $x$ 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 $x$. $\square$

**Theorem 7.** *If $c$ is not $\beta$-quasi-periodic within $A$, then the cut language $L_{<c} \subseteq A^*$ with base $\beta$ is not context-free.*

Proof. On the contrary assume that $c$ is not $\beta$-quasi-periodic within $A$ and $L_{<c}$ is a context-free language. Theorem 4 provides an infinite $\beta$-expansion $a = a_1a_2a_3\ldots \in A^\omega$ of $c$ whose tail sequence contains pairwise distinct values. Suppose for a contradiction that $a$ is not approximable in $L_{<c}$. This means there is a prefix $u \in A^*$ of $a$ such that for every $x \in A^*$ it holds $ux \notin L_{<c}$, that is, $(ux)_\beta \geq c = (a)_\beta$. On the other hand, we know $(a)_\beta = \lim_{n\to\infty}(ux_n)_\beta$ where for every $n$, $x_n \in A^*$ is a string of length $n = |x_n|$ such that $ux_n$ is a prefix of $a$, which implies $(a)_\beta = \inf_{x\in A^*}(ux)_\beta$. For $\beta > 0$, this ensures $a_k = \min A$ for every $k > |u|$, whereas for $\beta < 0$, it must be $a_{2k} = \min A$ and $a_{2k+1} = \max A$ for every $k > |u|/2$. Hence, $a$ is periodic, which contradicts the fact that the tail values of $a$ are pairwise different. It follows that $a$ is approximable in $L_{<c}$.

Let $a = uvw$ where $|v| > 0$ is even, be a decomposition guaranteed by Lemma 8. In particular, $uw$ and $uvvw$ are also approximable in $L_{<c}$. We know the tail values $r_c(u) = (vw)_\beta$ and $r_c(uv) = (w)_\beta$ are different. If $\beta^{-|u|}(w)_\beta > \beta^{-|u|}(vw)_\beta$, then define $y = uw$ which meets

$$(y)_\beta = (u)_\beta + \beta^{-|u|}(w)_\beta > (u)_\beta + \beta^{-|u|}(vw)_\beta = (uvw)_\beta = (a)_\beta = c\,. \tag{127}$$

On the other hand, if $\beta^{-|u|}(vw)_\beta > \beta^{-|u|}(w)_\beta$, then define $y = uvvw$ which satisfies

$$(y)_\beta = (uv)_\beta + \beta^{-|uv|}(vw)_\beta > (uv)_\beta + \beta^{-|uv|}(w)_\beta = (uvw)_\beta = (a)_\beta = c\,, \tag{128}$$

due to $\beta^{-|v|} > 0$. Thus, we have $y \in A^\omega$ which is approximable in $L_{<c}$ and $(y)_\beta > c$. This means that for every integer $n \geq 0$, there is $y_n \in L_{<c}$ implying $(y_n)_\beta < c$, such that $y$ and $y_n$ share the same prefix of length at least $n$. Hence,

$$|(y)_\beta - (y_n)_\beta| \leq \frac{\alpha}{|\beta|^n(|\beta| - 1)} \tag{129}$$

where $\alpha = \max\{|\alpha_1 - \alpha_2| \,;\, \alpha_1 \in A,\, \alpha_2 \in A \cup \{0\}\}$. It follows that $(y_n)_\beta$ converges to $(y)_\beta$ as $n$ tends to infinity, which contradicts $(y_n)_\beta < c < (y)_\beta$. □

Theorem 7 represents another proof of the necessary condition from Theorem 6 for a cut language to be regular. According to Theorem 6, we thus achieve a dichotomy that a cut language is either regular or non-context-free.

**Corollary 2.** *Any cut language $L_{<c} \subseteq A^*$ with base $\beta$ is either regular if $c$ is $\beta$-quasi-periodic within $A$, or non-context-free otherwise.*

Examples 3 and 5 provide explicit instances of a number $c$ whose all or some $\beta$-expansions are not eventually quasi-periodic for both rational and irrational $\beta$ and/or $c$, which correspond to examples of non-context-free cut languages $L_{<c}$.

On the other hand, the cut languages with rational thresholds are shown to be context-sensitive for any rational base and digits.

**Theorem 8.** *Let $\beta \in \mathbb{Q}$ be a rational base and $A \subset \mathbb{Q}$ be a set of rational digits. 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 $s_n = \sum_{k=1}^{n} a_k \beta^{-k}$ step by step when reading an input word $a_1 \ldots a_n \in A^*$ from right to left. In particular, $M$ starts with $s_0 = 0$ which updates to $s_k = (s_{k-1} + a_{n-k+1})/\beta$ every time after $M$ reads the next input symbol $a_{n-k+1} \in A$, for $k = 1, \ldots, n$. As the numbers $\beta, a_1, \ldots, a_n, c \in \mathbb{Q}$ can be represented within constant space and the length of the numerator and the denominator in the fractions representing $s_k$ increases at most by a constant in each step, $M$ needs only linear space in terms of input length $n$, for computing $s_n$ and testing whether $s_n < c$. □

## 7. Conclusion

In this paper we have defined the class of cut languages using a positional numeral system with a base $\beta$ and a digit alphabet $A$, which are motivated by the analysis of the computational power of neural net models with the weight parameters between integer and rational numbers. We have classified the cut languages within the Chomsky hierarchy. In particular, we have shown a dichotomy that a cut language $L_{<c}$ is either regular or non-context-free, depending on whether its threshold parameter $c$ is or is not a $\beta$-quasi-periodic number within $A$. For rational parameters $\beta$, $c$, and $A$, any cut language has proven to be context-sensitive.

Furthermore, we have introduced the concept of a $\beta$-quasi-periodic number within $A$, whose all $\beta$-expansions are eventually quasi-periodic, which has been illustrated by detailed examples. The definition of an eventually quasi-periodic $\beta$-expansion which has an infinite subsequence of tails sharing the same value, naturally generalizes the notion of eventually periodic $\beta$-expansions. For any base $\beta$ that is an algebraic number whose conjugates in absolute value differ from 1, or for transcendental $\beta$ combined with algebraic digits, we have shown that a $\beta$-expansion is eventually quasi-periodic iff it has only finitely many tail values.

On the other hand, for any algebraic $\beta$ with a conjugate of absolute value 1, we have constructed a quasi-periodic $\beta$-expansion having infinitely many distinct tail values. For algebraic bases $\beta$, we thus obtain the equivalence that for any digit alphabet $A$, every eventually quasi-periodic $\beta$-expansion has a finite number of tail values iff the conjugates of $\beta$ do not lie on the unit circle. In addition, we have proven that a number is $\beta$-quasi-periodic within $A$ iff all its $\beta$-expansions altogether have a finite number of tail values iff every its $\beta$-expansion has at least two tails of the same value.

Assume $A \subset \mathbb{Q}(\beta)$ and recall $\mathrm{QPer}(\beta) \subseteq \mathrm{Per}(\beta) \subseteq \mathbb{Q}(\beta)$ (see Paragraph 1.2). One can analogously define a set of so-called strongly (resp. weakly) periodic numbers, $\mathrm{SPer}(\beta)$ (resp. $\mathrm{WPer}(\beta)$) containing the real numbers for which an infinite $\beta$-expansion exists and every (resp. at least one) such $\beta$-expansion is eventually periodic. Obviously,

$$\mathrm{SPer}(\beta) \subseteq \mathrm{QPer}(\beta) \subseteq \mathrm{Per}(\beta) \subseteq \mathrm{WPer}(\beta) \subseteq \mathbb{Q}(\beta)\,. \tag{130}$$

For any Pisot base $\beta$, we have shown that $\mathrm{QPer}(\beta) = \mathrm{Per}(\beta) = \mathrm{WPer}(\beta)$. The presented examples including Pisot or non-Pisot bases satisfy the strict inclusion $\mathrm{SPer}(\beta) \subset \mathrm{QPer}(\beta)$ but it is an open question for which bases $\beta$ this is satisfied. For example, we conjecture that for $A = \{0, 1\}$ and any non-Pisot $\beta$ less than the golden ratio $\varphi$, it holds $\mathrm{SPer}(\beta) = \mathrm{QPer}(\beta) = \emptyset$. Another interesting issue for further research would be to generalize these results to arbitrary real digits.


## Acknowledgments

We would like to thank Professor E. Pelantová, Professor Z. Masáková, and Dr. Tomáš Vávra for their encouraging interest in our work, important feedback to our paper, and pointers to relevant results in the literature. The presentation of this paper also benefited from valuable suggestions of an anonymous reviewer. This research was done with institutional support RVO:67985807 and partially supported by the grant of the Czech Science Foundation No. GBP202/12/G061.


## References


[1] M. Pedicini, Greedy expansions and sets with deleted digits, Theoretical Computer Science 332 (1-3) (2005) 313–336.

[2] A. Rényi, Representations for real numbers and their ergodic properties, Acta Mathematica Academiae Scientiarum Hungaricae 8 (3-4) (1957) 477–493.

[3] W. Parry, On the $\beta$-expansions of real numbers, Acta Mathematica Hungarica 11 (3) (1960) 401–416.

[4] N. Sidorov, Almost every number has a continuum of $\beta$-expansions, The American Mathematical Monthly 110 (9) (2003) 838–842.

[5] P. Erdös, I. Joó, V. Komornik, Characterization of the unique expansions $1 = \sum_{i=1}^{\infty} q^{-n_i}$ and related problems, Bulletin de la Société Mathématique de France 118 (3) (1990) 377–390.

[6] P. Glendinning, N. Sidorov, Unique representations of real numbers in non-integer bases, Mathematical Research Letters 8 (4) (2001) 535–543.

[7] N. Sidorov, Expansions in non-integer bases: Lower, middle and top orders, Journal of Number Theory 129 (4) (2009) 741–754.

[8] J.-P. Allouche, M. Clarke, N. Sidorov, Periodic unique beta-expansions: The Sharkovskiĭ ordering, Ergodic Theory and Dynamical Systems 29 (4) (2009) 1055–1074.

[9] S. Baker, Generalized golden ratios over integer alphabet, Integers 14 (2014) A15.

[10] M. de Vries, V. Komornik, Unique expansions of real numbers, Advances in Mathematics 221 (2) (2009) 390–427.

[11] D. Kong, W. Li, F. M. Dekking, Intersections of homogeneous cantor sets and beta-expansions, Nonlinearity 23 (11) (2010) 2815–2834.

[12] V. Komornik, A. C. Lai, M. Pedicini, Generalized golden ratios of ternary alphabets, Journal of the European Mathematical Society 13 (4) (2011) 1113–1146.

[13] V. Komornik, M. Pedicini, Critical bases for ternary alphabets, Acta Mathematica Hungarica 152 (1) (2017) 25–57.

[14] K. Schmidt, On periodic expansions of Pisot numbers and Salem numbers, Bulletin of the London Mathematical Society 12 (4) (1980) 269–278.

[15] K. G. Hare, Beta-expansions of Pisot and Salem numbers, in: Proceedings of the Waterloo Workshop in Computer Algebra 2006: Latest Advances in Symbolic Algorithms, World Scientific, 2007, pp. 67–84.

[16] S. Baker, Z. Masáková, E. Pelantová, T. Vávra, On periodic representations in non-Pisot bases, Monatshefte für Mathematik 184 (1) (2017) 1–19.

[17] D. Dombek, Z. Masáková, E. Pelantová, Number representation using generalized $(-\beta)$-transformation, Theoretical Computer Science 412 (48) (2011) 6653–6665.

[18] C. Frougny, A. C. Lai, Negative bases and automata, Discrete Mathematics & Theoretical Computer Science 13 (1) (2011) 75–94.

[19] S. Akiyama, J. M. Thuswaldner, T. Zaïmi, Comments on the height reducing property II, Indagationes Mathematicae 26 (1) (2015) 28–39.

[20] J. Šíma, P. Savický, Cut languages in rational bases, in: Proceedings of the LATA 2017 Eleventh International Conference on Language and Automata Theory and Applications, Vol. 10168 of LNCS, 2017, pp. 311–322.

[21] H. T. Siegelmann, Neural Networks and Analog Computation: Beyond the Turing Limit, Birkhäuser, Boston, 1999.

[22] J. Šíma, P. Orponen, General-purpose computation with neural networks: A survey of complexity theoretic results, Neural Computation 15 (12) (2003) 2727–2778.

[23] J. L. Balcázar, R. Gavaldà, H. T. Siegelmann, Computational power of neural networks: A characterization in terms of Kolmogorov complexity, IEEE Transactions on Information Theory 43 (4) (1997) 1175–1183.

[24] H. T. Siegelmann, E. D. Sontag, Analog computation via neural networks, Theoretical Computer Science 131 (2) (1994) 331–360.

[25] N. Alon, A. K. Dewdney, T. J. Ott, Efficient simulation of finite automata by neural nets, Journal of the ACM 38 (2) (1991) 495–514.

[26] B. G. Horne, D. R. Hush, Bounds on the complexity of recurrent neural network implementations of finite state machines, Neural Networks 9 (2) (1996) 243–252.

[27] P. Indyk, Optimal simulation of automata by neural nets, in: Proceedings of the STACS 1995 Twelfth Annual Symposium on Theoretical Aspects of Computer Science, Vol. 900 of LNCS, 1995, pp. 337–348.

[28] M. Minsky, Computations: Finite and Infinite Machines, Prentice-Hall, Englewood Cliffs, 1967.

[29] H. T. Siegelmann, Recurrent neural networks and finite automata, Journal of Computational Intelligence 12 (4) (1996) 567–574.

[30] J. Šíma, Energy complexity of recurrent neural networks, Neural Computation 26 (5) (2014) 953–973.

[31] J. Šíma, J. Wiedermann, Theory of neuromata, Journal of the ACM 45 (1) (1998) 155–178.

[32] H. T. Siegelmann, E. D. Sontag, On the computational power of neural nets, Journal of Computer System Science 50 (1) (1995) 132–150.

[33] J. Šíma, The power of extra analog neuron, in: Proceedings of the TPNC 2014 Third International Conference on the Theory and Practice of Natural Computing, Vol. 8890 of LNCS, 2014, pp. 243–254.

[34] J. Šíma, Neural networks between integer and rational weights, in: Proceedings of the IJCNN 2017 Thirties International Joint Conference on Neural Networks, IEEE, 2017, pp. 154–161.

[35] J. Šíma, The computational power of neural networks and representations of numbers in non-integer bases (invited talk), Mendel Softcomputing Journal: Proceedings of the MENDEL 2017 Twenty-Third International Conference on Soft Computing 23 (1) (2017) 103–110.

[36] M. F. Atiyah, I. G. Macdonald, Introduction To Commutative Algebra, Addison-Wesley Series in Mathematics, Addison-Wesley Publishing Co., Reading, MA, 1969.

[37] D. A. Marcus, Number Fields, Springer-Verlag, New York, 1977.