CDMTCS
Research
Report
Series

On the Generative Power of Quasiperiods

Ludwig Staiger
Martin-Luther-Universität
Halle-Wittenberg

CDMTCS-548
October 2020

Centre for Discrete Mathematics and
Theoretical Computer Science

On the Generative Power of Quasiperiods

Ludwig Staiger *
Martin-Luther-Universität Halle-Wittenberg
Institut für Informatik
von-Seckendorff-Platz 1, D–06099 Halle (Saale), Germany

Abstract

It is shown that, for every length $l \geq 3$, a quasiperiod of the form $a^n ba^n$ (or $a^n bba^n$ if l is even) generates the largest language Q of words having this word as quasiperiod. As a means of comparison we use the growth of the function which counts the number of words of length l in the language Q.

Moreover, we give the exact ordering of the lengths l with respect to the largest language Q generated by a quasiperiod of length l.

Contents

*email: ludwig.staiger@informatik.uni-halle.de

1 Introduction

Informally, a word q is a quasiperiod of another word w if q is a prefix and a suffix of w and every position of w is covered by q.

In this paper we investigate the languages Q_q of words w having q as quasiperiod. We are interested in the question of which quasiperiods q generate large languages Q_q. Since different quasiperiods may have incomparable w.r.t. set inclusion languages Q_q, we compare the languages Q_q by their functions $s_q : \mathbb{N} \to \mathbb{N}$ which count the number of words of length n in Q_q. As a means of comparison we use their asymptotical growth. It turns out that the languages Q_q are essentially regular star-languages, therefore their function s_q satisfies $s_q(n) \approx \text{const.} \cdot \lambda_q^n$, where the value $\lambda_q \geq 1$ depends on the quasiperiod q.

The aim of this paper is to estimate, for every length $n \geq 3$ the words q which have the largest value λ_q. To this end we consider along with language-theoretical properties of Q_q some combinatorial properties of quasiperiods. Moreover, we need to consider a special class of integer polynomials related to quasiperiods.

The paper is organised as follows. After some preliminaries we deal with combinatorial properties of quasiperiods and the generated languages. The asymptotic growth of Q_q is the subject of Section 4. Then we deal with basic properties of polynomials related to quasiperiods. In these sections we mainly report results of the papers [PS10] and [Sta18]. The following Sections 6 and 7 deal with the proof of the main theorem. Here we derive also the complete ordering of the values $\lambda_n = \max\{\lambda_q : |q| = n\}$.

2 Notation and Preliminaries

We introduce the notation used throughout the paper. By $\mathbb{N} = \{0, 1, 2, \ldots\}$ we denote the set of natural numbers. Let X be a finite alphabet. Usually by

$a,b \in X$ we mean two different letters. X^* is the set (monoid) of words on X, including the *empty word* e.

For $w,v \in X^*$ let $w \cdot v$ be their *concatenation*. This concatenation product extends in an obvious way to subsets $W,L \subseteq X^*$. For a language W let $W^* := \bigcup_{i\in\mathbb{N}} W^i$ be the *submonoid* of X^* generated by W. The smallest subset of a language W which generates W^* is called its *star root* $\sqrt[*]{W}$ [Brz67]. It holds

$$\sqrt[*]{W} = (W \setminus \{e\}) \setminus (W \setminus \{e\})^2 \cdot W^* .$$

Furthermore $|w|$ is the *length* of the word $w \in X^*$, and by $w \sqsubseteq v$ we denote the fact that w is a *prefix* of v.

A word $w \in X^* \setminus \{e\}$ is called *primitive* if $w = v^n$ implies $n = 1$, that is, w is not the power of a shorter word.

As usual a language $L \subseteq X^*$ is called a *code* provided $w_1 \cdots w_l = v_1 \cdots v_k$ for $w_1,\dots,w_l$, $v_1,\dots,v_k \in L$ implies $l = k$ and $w_i = v_i$. A code L is said to be a *suffix code* provided no codeword is a suffix of another codeword.

Finally, we define the language Q_q of words having $q \in X^* \setminus \{e\}$ as quasiperiod.

(0) $e \in Q_q$, and
(1) $w \in Q_q$, if and only if $w \in X^* \cdot q$ and there is a $w' \sqsubset w, w' \in Q_q$, with $w \sqsubseteq w' \cdot q$.

3 Quasiperiodic Words

In this part we consider the finite language P_q ($\mathcal{L}(q)$ in [Mou00]) which is tightly related to Q_q. Most of the results are contained in [Mou00, PS10] and [Sta18].

We set

$$P_q := \{v : e \sqsubset v \sqsubseteq q \sqsubset v \cdot q\} . \tag{1}$$

We have the following property.

$$Q_q \setminus \{e\} = P_q^* \cdot q \subseteq P_q^* \cap q \cdot X^* . \tag{2}$$

3.1 Combinatorial properties of P_q

We investigate basic properties of P_q using simple facts from combinatorics on words (see [BP85, Lot97, Shy01]).

Proposition 1 *$v \in P_q$ if and only if $|v| \le |q|$ and there is a prefix $\bar{v} \sqsubset v$ such that $q = v^k \cdot \bar{v}$ for $k = \lfloor |q|/|v| \rfloor$.*

Corollary 1 *$v \in P_q$ if and only if $|v| \le |q|$ and there is a $k' \in \mathbb{N}$ such that $q \sqsubseteq v^{k'}$.*

Now set $q_0 := \min_{\sqsubseteq} P_q$. Then in view of Proposition 1 and Corollary 1 we have the following canonical representation.

$$q = q_0^k \cdot \bar{q} \text{ where } k = \lfloor |q|/|q_0| \rfloor \text{ and } \bar{q} \sqsubset q_0 . \tag{3}$$

We will refer to q_0 as the *repeated prefix* and to k as the *repetition factor*. If $|q_0| > |q|/2$, that is, if $k = 1$ we will refer to q as *irreducible*.[1]

Corollary 2 *Every word $v \in \sqrt[*]{P_q}$ is primitive.*

Proof. Assume $v = v_1^l$ for some $v \in \sqrt[*]{P_q}$ and $l > 1$. Then $q \sqsubseteq v^{k'} = v_1^{l \cdot k'}$, and, according to Corollary 1 $v_1 \in P_q$ contradicting $v \in \sqrt[*]{P_q}$. ❏

Proposition 2 *Let $q \in X^*, q \neq e$, $q_0 = \min_{\sqsubseteq} P_q$, $q = q_0^k \cdot \bar{q}$ and $v \in P_q^* \setminus \{e\}$.*

(1) *If $w \sqsubseteq q$ then $v \cdot w \sqsubseteq q$ or $q \sqsubseteq v \cdot w$.*

(2) *If $w \cdot v \sqsubseteq q$ then $w \in \{q_0\}^*$.*

(3) *If $|v| \le |q| - |q_0|$ then $v \in \{q_0\}^*$.*

Corollary 3 *If $q \notin \{q_0\}^*$ then q_0 is not a suffix of q.*

Proof. Let $q = w \cdot q_0$. Then according to Proposition 2.2 $w \in \{q_0\}^*$. ❏

Next we derive a slight improvement of Proposition 2.3. To this end, we use the Theorem of Fine and Wilf.

Theorem 1 ([FW65]) *Let $v, w \in X^*$. Suppose v^m and w^n, for some $m, n \in \mathbb{N}$, have a common prefix of length $|v| + |w| - \gcd(|v|, |w|)$. Then v and w are powers of a common word $u \in X^*$ of length $|u| = \gcd(|v|, |w|)$.*[2]

Proposition 3 *Let $q \in X^*, q \neq e$, $q_0 = \min_{\sqsubseteq} P_q$, $q = q_0^k \cdot \bar{q}$ and $v \in P_q$. If $|v| \le |q| - |q_0| + \gcd(|v|, |q_0|)$ then $v \in \{q_0\}^*$.*

Proof. $q_0, v \in P_q$ imply that q is a common prefix of q_0^{k+1} and $v^{k'}$ for some $k' \in \mathbb{N}$. In view of $|v| \le |q| - |q_0| + \gcd(|v|, |q_0|)$ Theorem 1 implies that q_0 and v are powers of a common word, that is, $v \in \{q_0\}^*$. ❏

---

[1] Superprimitive in the sense of [AFI91, Mou00] quasiperiods are irreducible but not vice versa (see [Sta18, Section 2.3.4]).

[2] Here $\gcd(k, l)$ denotes the greatest common divisor of two numbers $k, l \in \mathbb{N}$.

3.2 The reduced quasiperiod $\hat{q}$

Next we investigate the relation between a quasiperiod $q = q_0^k \cdot \bar{q}$ where $q_0 = \min_{\sqsubseteq} P_q$ and $\bar{q} \sqsubset q_0$ and its *reduced quasiperiod* $\hat{q} := q_0 \cdot \bar{q}$. Since $q \in Q_{\hat{q}}$, we have $Q_{\hat{q}} \supseteq Q_q$.

We continue with a relation between P_q and $P_{\hat{q}}$. It is obvious that $q_0^i \in P_q$ for every $i = 1, \ldots, k$. Then Proposition 3 shows that

$$\sqrt[*]{P_q} \subseteq \{q_0\} \cup \{v' : v' \sqsubseteq q \wedge |v'| > |q| - |q_0| + \gcd(|v'|, |q_0|)\}. \tag{4}$$

Lemma 1 ([Sta18, Lemma 2.2]) *Let $q \in X^*, q \neq e$, $q_0 = \min_{\sqsubseteq} P_q$, $q = q_0^k \cdot \bar{q}$ and $\hat{q} = q_0 \cdot \bar{q}$ the reduced quasiperiod of q. Then*

$$P_q = \{q_0^i : i = 1, \ldots, k-1\} \cup \{q_0^{k-1} \cdot v : v \in P_{\hat{q}}\}.$$

This implies

$$\sqrt[*]{P_q} \subseteq \{q_0\} \cup q_0^{k-1} \cdot (P_{\hat{q}} \setminus \{q_0\}), \text{ and} \tag{5}$$

$$P_{\hat{q}} \subseteq \{v : \hat{q}_0 \sqsubseteq v \sqsubseteq \hat{q}\} \tag{6}$$

Moreover, we have the following.

Lemma 2 *Let $q = q_0^k \cdot \bar{q}$ with $k \geq 2$, $\bar{q} \sqsubset q_0$ and $\hat{q} = q_0 \cdot \bar{q}$.*

If $\hat{q}_0 \neq q_0$ for the repeated prefix of $\hat{q}_0$ then $\bar{q} \sqsubset \hat{q}_0 \sqsubset q_0$ and $|\hat{q}_0| > |\bar{q}| + \gcd(|q_0|, |\hat{q}_0|)$. Moreover, then there is a nonempty suffix $v \neq e$ of q_0 such that $v \sqsubset \hat{q}_0$ and $v \cdot \bar{q} \sqsubset \hat{q}_0^2$.

Proof. We have $\bar{q} \sqsubseteq q_0$ and, since $q_0 \in P_{\hat{q}}$, also $\hat{q}_0 \sqsubseteq q_0$. Moreover, $\hat{q} \sqsubseteq q_0^2$ and $\hat{q} \sqsubseteq \hat{q}_0^{k'}$ for some $k' \in \mathbb{N}$. Since $q_0 \neq \hat{q}_0$ and both prefixes are primitive words, Theorem 1 shows that the common prefix $\hat{q} = q_0 \cdot \bar{q}$ has to satisfy $|\hat{q}| < |q_0| + |\hat{q}_0| - \gcd(|q_0|, |\hat{q}_0|)$, that is, $|\hat{q}_0| > |\bar{q}| + \gcd(|q_0|, |\hat{q}_0|)$. The assertion $\bar{q} \sqsubset \hat{q}_0 \sqsubset q_0$ now follows from a comparison of the lengths of $\bar{q}, \hat{q}_0 \sqsubseteq q_0$.

Now, let v be the suffix of q_0 defined by $\hat{q}_0^{k'} \cdot v = q_0 \sqsubset \hat{q}_0^{k'+1}$. Then $v \sqsubset \hat{q}_0$ and $v \cdot \bar{q} \sqsubset (\hat{q}_0)^2$. ❏

4 Asymptotic Growth

In this section we use the fact that $\sqrt[*]{P_q}$ is a suffix code to estimate the exponential growth of the family $(|Q_q \cap X^n|)_{n \in \mathbb{N}}$. In view of the identity $Q_q \setminus \{e\} = P_q^* \cdot q$ we have $|Q_q \cap X^{n+|q|}| = |P_q^* \cap X^n|$. So we may use P_q^* instead of Q_q.

First we mention that $\sqrt[*]{P_q}$ is a suffix code. This generalises Proposition 7 of [Mou00].

Proposition 4 ([PS10, Sta18]) *$\sqrt[*]{P_q}$ is a suffix code.*

In order to derive the announced exponential growth we refer to Corollary 4 of [Sta85] which shows that for every regular language $L \subseteq X^*$ there are constants $c_1, c_2 > 0$ and a $\lambda \geq 1$ such that

$$c_1 \cdot \lambda^n \leq_{\text{i.o.}} |L^* \cap X^n| \leq c_2 \cdot \lambda^n . \tag{7}$$

In the remainder of this section we use, without explicit reference, known results from the theory of formal power series, in particular about generating functions of languages and codes which can be found in the literature, e.g. in [BP85, BR88] or [SS78].

As P_q^* is a regular language the value λ_q for $L = P_q$ in Eq. (7) is $\lambda_q = \limsup_{n\to\infty} \sqrt[n]{|P_q^* \cap X^n|}$ which is the inverse of the convergence radius of the power series $\mathfrak{s}_q^*(t) := \sum_{n\in\mathbb{N}} |P_q^* \cap X^n| \cdot t^n$. The series $\mathfrak{s}_q^*$ is also known as the *structure generating function* of the language P_q^*.

Since $\sqrt[*]{P_q}$ is a code, we have $\mathfrak{s}_q^*(t) = \frac{1}{1-\mathfrak{s}_q(t)}$ where $\mathfrak{s}_q(t) := \sum_{v\in \sqrt[*]{P_q}} t^{|v|}$ is the structure generating function of the finite language $\sqrt[*]{P_q}$. Thus λ_q^{-1} is the smallest root of $1-\mathfrak{s}_q(t)$. Hence λ_q is the largest root of the polynomial $p_q(t) := t^{|q|} - \sum_{v\in \sqrt[*]{P_q}} t^{|q|-|v|}$.

Summarising our observations we obtain the following.

Lemma 3 *Let $q \in X^* \setminus \{e\}$. Then there are constants $c_{q,1}, c_{q,2} > 0$ such that*

$$c_{q,1} \cdot \lambda_q^n \leq_{\text{i.o.}} |P_q^* \cap X^n| \leq c_{q,2} \cdot \lambda_q^n$$

where λ_q is the largest (positive) root of the polynomial $p_q(t)$.

5 Polynomials

Before proceeding to the proof of our main theorem we derive some properties of polynomials of the form $p(t) = t^n - \sum_{i\in M} t^i, M \subseteq \{i : i \in \mathbb{N} \wedge i < n\}$. We are mainly interested in results which are useful for comparing their maximal roots.

The polynomials $p(t) \in \hat{\mathscr{P}} := \{t^n - \sum_{i\in M} t^i : \emptyset \neq M \subseteq \{0,\ldots,n-1\}\}$ have the following easily verified properties.

$$p(0) \leq 0, p(1) \leq 0, p(2) \geq 1 \text{ and } p(t) < 0 \text{ for } 0 < t < 1 . \tag{8}$$

$$\text{If } \varepsilon > 0 \text{ and } p(t') \geq 0 \text{ for some } t' > 0 \text{ then } p((1+\varepsilon)\cdot t') > 0 . \tag{9}$$

Since $p(1) \leq 0$ and $p(2) \geq 1$ for $p(t) \in \hat{\mathscr{P}}$, Eq. (9) shows that once $p(t') \geq 0$, $t' \geq 1$, the polynomial $p(t)$ has no further root in the interval (t', ∞) and $p(t) \in \hat{\mathscr{P}}$ has exactly one root in the interval $[1,2)$. This yields the following fundamental property. If t_0 is the positive root of the polynomial $p(t) \in \hat{\mathscr{P}}$ in $[1,2)$ and $1 \leq t' < 2$ then $p(t') \leq 0$ if and only if $t' \leq t_0$. For the roots of maximal modulus we have the following theorem.

Theorem 2 (Cauchy) *Let $p(t) = \sum_{i=0}^{n} a_i \cdot t^i$ be a complex polynomial. Then every root t' of $p(t)$ satisfies $|t'| \leq t_0$ where t_0 is the maximal root of the polynomial $|a_n| \cdot t^n - \sum_{i=0}^{n-1} |a_i| \cdot t^i$.*

This implies the following property of polynomials $p(t) \in \hat{\mathscr{P}}$.

$$\text{If } p(t) = 0 \text{ then } |t| \leq t_0 . \tag{10}$$

From Property 5 we derive the following criterion to compare the maximal roots of polynomials in $\hat{\mathscr{P}}$.

Criterion 1 *Let $p_1(t), p_2(t) \in \hat{\mathscr{P}}$ have maximal roots t_1 and t_2, respectively. Then $p_2(t_1) > 0$ if and only if $t_1 > t_2$. In particular, $p_2(t_1) > 0$ implies $t_1 > t_2$.*

We conclude this section with a bound on the maximal root of certain polynomials in $\hat{\mathscr{P}}$.

Lemma 4 *Let $p(t) = t^n - \sum_{i=0}^{m} t^i, n > m \geq 1$. Then $p(t) > 0$ for $\sqrt[n-m]{m+1} \leq t$ and $p(t) < 0$ for $1 \leq t \leq \sqrt[2n-m]{(m+1)^2}$.*

Proof. The assertion follows from the inequality $t^n - (m+1) \cdot t^m < p(t) < t^n - (m+1) \cdot t^{m/2}$ when $t > 1$. The first part uses the arithmetic-geometric-means inequality $\sum_{i=0}^{m} t^i > (m+1) \cdot \sqrt[m+1]{\prod_{i=0}^{m} t^i} = (m+1) \cdot t^{m/2}$, and the second holds for $t \geq 1$. ❑

The following special case is needed below.

Corollary 4 *If $p(t) = t^n - \sum_{i=0}^{n-3} t^i, n \geq 4$, then $p(t) < 0$ for $1 \leq t \leq \sqrt[n+3]{(n-2)^2}$.*

The subsequent sections are devoted to the proof of our main theorem.

6 Irreducible Quasiperiods

We start with irreducible quasiperiods. As quasiperiods $q, |q| \leq 2$, have trivially $P_q^* = \{q\}^*$, in the subsequent sections. we confine our considerations to quasiperiods q of length $|q| \geq 3$.

6.1 Extremal polynomials

The polynomials $p_q(t)$ of irreducible quasiperiods have non-zero coefficients only for $|q|$ and $i < \frac{|q|}{2}$. Therefore we investigate the set

$$\mathscr{P} := \{t^n - \sum_{i\in M} t^i : n \geq 1 \wedge \emptyset \neq M \subseteq \{j : j \leq \tfrac{n-1}{2}\}\}.$$

Let $p_n(t) := t^n - \sum_{i=0}^{\lfloor \frac{n-1}{2} \rfloor} t^i \in \mathscr{P}$. Let $p(t) \in \mathscr{P}$ a polynomial of degree n. Then $p_n(t) \leq p(t)$ for $t \in [1,2]$, and $p_n(t)$ has the largest positive root among all polynomials of degree n in $\mathscr{P}$.

Proof. This follows from $t^n - \sum_{i=0}^{\lfloor \frac{n-1}{2} \rfloor} t^i \leq p(t)$ for $p(t) \in \mathscr{P}$ when $1 < t \leq 2$ and Criterion 1. ❑

Observe that, for $n \geq 1$,

$$p_{2n+1}(t) = t^{2n+1} - \sum_{i=0}^{n} t^i \text{ and } p_{2n+2}(t) = t^{2n+2} - \sum_{i=0}^{n} t^i .$$

Moreover, $a^n b a^n$ and $a^n w a^n, w \in \{xb, bx\}, x \in X$ are quasiperiods corresponding to the extremal polynomials $p_{2n+1}(t) \in \mathscr{P}$ and $p_{2n+2}(t) \in \mathscr{P}$, respectively.

Let $Q_{\max} := \{a^n b a^n : n \geq 1\} \cup \{a^n w a^n : w \in X \cdot b \cup b \cdot X, n \geq 1\}$.

In what follows we will always assume that the first letter of a quasiperiod q is a. Then $Q_{\max}$ is the set of quasiperiods corresponding to the extremal polynomials.

Lemma 5 $Q_{\max} := \{q : q \in X^* \wedge |q| \geq 3 \wedge p_q(t) = p_{|q|}(t)\}$

Proof. If $q \in Q_{\max}$ then obviously $p_q(t) = p_{|q|}(t)$. Conversely, if $p_q(t) = t^{|q|} - \sum_{v \in \sqrt[*]{P_q}} t^{|q|-|v|} = p_{|q|}(t)$ then $\sqrt[*]{P_q} = \{v : v \sqsubseteq q \wedge |v| > \frac{|q|}{2}\}$. Then, in view of $q \sqsubseteq v \cdot q$, every prefix $w \sqsubseteq q$ of length $|w| < \frac{|q|}{2}$ is also a suffix of q. This is possible only for $q \in Q_{\max}$ or $q \in \{a\}^*$. ❑

In the sequel the positive root of $p_n(t)$ is denoted by λ_n. From Criterion 1 we obtain immediately.

Criterion 2 *Let $t \geq 1$. We have $t < \lambda_n$ if and only if $p_n(t) < 0$.*

Then Property 6.1 implies the following.

Theorem 3 *If $q \in X^*, |q| \geq 3$, is an irreducible quasiperiod then $\lambda_q \leq \lambda_{|q|}$, and $\lambda_q = \lambda_{|q|}$ if and only if $q \in Q_{\max}$.*

6.2 The ordering of the maximal roots λ_n

Before we proceed to the case of reducible quasiperiods we determine the ordering of the maximal roots λ_n. This will not only be interesting for itself but also useful for proving $\lambda_q < \lambda_{|q|}$ when q is reducible (see Eq. (21) below).

The extremal polynomials $p_n(t)$ satisfy the following general relations. [3]

$$t \cdot p_{2n-2}(t) - 1 = p_{2n-1}(t), \tag{11}$$
$$p_{2n}(t) - t^2 \cdot p_{2n-2}(t) = t^n - t - 1, \tag{12}$$
$$t^{n-2} \cdot p_{2n+1}(t) - (t^n + 1) \cdot p_{2n-1}(t) = \sum_{i=0}^{n-3} t^i, \text{ and} \tag{13}$$
$$t^{n-2} \cdot p_{2n+3}(t) - (t^{n+1} + 1) \cdot p_{2n}(t) = -t^n + \sum_{i=0}^{n-3} t^i \tag{14}$$

Lemma 6 *The polynomials* $t^3 - t - 1$ *and* $t^5 - t^2 - t - 1 = (t^2+1)\cdot(t^3 - t - 1)$ *have largest positive roots* $\lambda_3 = \lambda_5$ *among all polynomials in* $\mathscr{P}$, $\lambda_5 > \lambda_4$ *and* $\lambda_{2n-1} > \lambda_{2n+1} > \lambda_{2n}$ *for* $n \geq 3$.

Proof. From Eq. (11) we have $p_{2n+1}(\lambda_{2n}) = -1 < 0$ and, therefore, $\lambda_{2n} < \lambda_{2n+1}$ when $n \geq 1$.

Similarly, Eq. (13) yields $p_{2n+1}(\lambda_{2n-1}) = \lambda_{2n-1}^{-(n-2)} \cdot \sum_{i=0}^{n-3} \lambda_{2n-1}^i > 0$ which implies $\lambda_{2n+1} < \lambda_{2n-1}$ for $n \geq 3$ and $\lambda_3 = \lambda_5$ when $n = 2$. ❑

So far we have ordered the 'odd' roots: $\lambda_3 = \lambda_5 > \lambda_7 > \lambda_9 > \cdots$. Next we are going to investigate the ordering of the 'even' roots λ_{2n}, $n \geq 2$.

To this end we derive the following bounds.

Lemma 7 (1) $\sqrt[3n+1]{n^2} \leq \lambda_{2n} \leq \sqrt[n+1]{n}$ *and* $\sqrt[3n-1]{n^2} \leq \lambda_{2n-1} \leq \sqrt[n]{n}$ *for* $n \geq 2$.

(2) *Let* $n \geq 5$. *Then* $\lambda_{2n} \geq \sqrt[n-1]{2}$.

Proof. 1. follows from Lemma 4.

2. We calculate $p_{2n}(\sqrt[n-1]{2}) = 4 \cdot \sqrt[n-1]{4} - \sum_{i=0}^{n-1} \sqrt[n-1]{2^i} \leq 4 \cdot \sqrt[4]{4} - (2 + (n-1)) = 4 \cdot \sqrt{2} - (n+1) < 0$ if $n \geq 5$ and the assertion follows with Property 5. ❑

Remark 1 The lower bound of Lemma 7.2 does not exceed the lower bound in Lemma 7.1. However, the latter is more convenient for the purposes of the subsequent Lemma 8.

Lemma 8 *If* $n \geq 5$ *then* $\lambda_{2n-2} > \lambda_{2n}$ *and* $\lambda_{2n} > \lambda_{2n+3}$.

[3] By convention, $\sum_{i=k}^{m} a_i = 0$ if $k > m$.

Proof. If $t \geq \sqrt[n-1]{2}$ then $t^n - t - 1 \geq t - 1 > 0$. Consequently, Eq. (12) implies $p_{2n}(\lambda_{2n-2}) > 0$ whence $\lambda_{2n} < \lambda_{2n-2}$.

Eq. (14), Corollary 4 and the inequality $\lambda_{2n} \leq \sqrt[n+1]{n} \leq \sqrt[n+3]{(n-2)^2}$ when $n \geq 5$ imply $\lambda_{2n} \cdot p_{2n+3}(\lambda_{2n}) = -(\lambda_{2n}^n - \sum_{i=0}^{n-3} \lambda_{2n}^i) > 0$ whence $\lambda_{2n} > \lambda_{2n+3}$ for $n \geq 5$. ❑

Since $p_8(\sqrt[3]{2}) > 0$, the proof of Lemma 8 cannot be applied to lower values of n. Thus it remains to establish the order of the λ_i for $i \leq 13$. To this end, we consider some special identities and use Criterion 2 and Lemma 8.

$$p_{12}(t) - (t^8 + t^5 + t^4 + t^2 + t) \cdot p_4(t) = t^2 - 1 \text{ and} \tag{15}$$
$$p_{13}(t) - t \cdot (t^8 + t^5 + t^4 + t^2 + t) \cdot p_4(t) = t^3 - t - 1 = p_3(t). \tag{16}$$

Lemma 9 $\lambda_8 > \lambda_{10} > \lambda_{13} > \lambda_4 > \lambda_{12}$

Proof. Lemma 8 shows $\lambda_8 > \lambda_{10} > \lambda_{13}$. Eq. (15) yields $p_{12}(\lambda_4) = \lambda_4^2 - 1 > 0$ whence $\lambda_4 > \lambda_{12}$, and Eq. (16) yields $p_{13}(\lambda_4) = p_3(\lambda_4) < 0$, that is $\lambda_{13} > \lambda_4$. This shows our assertion. ❑

For the remaining part we consider the identities

$$t^2 \cdot p_{11}(t) - (t^5 + 1) \cdot p_8(t) = -t^4 + t + 1 = -p_4(t) \tag{17}$$
$$p_{11}(t) - (t^5 + 1) \cdot p_6(t) = t^3 \cdot p_4(t) \text{ and} \tag{18}$$
$$t \cdot p_9(t) - (t^4 + 1) \cdot p_6(t) = -t^3 + 1. \tag{19}$$

Lemma 10 $\lambda_9 > \lambda_6 > \lambda_{11} > \lambda_8$

Proof. We use Eqs. (17) to (19). Then $p_{11}(\lambda_8) = -p_4(\lambda_8) < 0$ implies $\lambda_{11} > \lambda_8$, $p_{11}(\lambda_6) = \lambda_6^3 \cdot p_4(\lambda_6) > 0$ implies $\lambda_6 > \lambda_{11}$, and, finally, $\lambda_6 \cdot p_9(\lambda_6) = -\lambda_6^3 + 1$ implies $\lambda_9 > \lambda_6$. ❑

Now Lemmata 6, 8, 9 and 10 yield the complete ordering of the values λ_n.

Theorem 4 *Let $\lambda_n, n \geq 3$, be the maximal root of the polynomial $p_n(t)$. Then the overall ordering of the values λ_n starts with*

$$\lambda_3 = \lambda_5 > \lambda_7 > \lambda_9 > \lambda_6 > \lambda_{11} > \lambda_8 > \lambda_{10} > \lambda_{13} > \lambda_4 > \lambda_{12}$$

and continues as follows $\lambda_{2n+1} > \lambda_{2n} > \lambda_{2n+3}, n \geq 7$.

From Lemma 7.1 we obtain immediately.

Corollary 5 *Let $M \subseteq \mathbb{N} \setminus \{0, 1, 2\}$ be infinite. Then* $\inf\{\lambda_i : i \in M\} = 1$.

7 Reducible Quasiperiods

Reducible quasiperiods q have a repeated prefix $q_0 = \min_{\sqsubseteq} P_q$ with $|q_0| \leq |q|/2$ and a repetition factor $k \geq 2$ such that $q = q_0^k \cdot \bar{q}$ where $\bar{q} \sqsubset q_0$. Moreover $|\bar{q}| < |q_0| \leq |q|/2$. Observe that q_0 is primitive.

We shall consider three cases depending on the relation between the lengths $n = |q|$, $\ell = |q_0|$, the length of the suffix $|\bar{q}| < |q_0|$ and the repetition factor $k \geq 2$.

7.1 The case $|\bar{q}| + |q_0| \leq 2$

The case $|\bar{q}| + |q_0| \leq 2$ is the simplest one. Here, in view of $\bar{q} \sqsubset q_0$ we have necessarily $\bar{q} = e$ and $q \in a^* \cup \{ab\}^*$, $a, b \in X, a \neq b$ and, therefore, $\lambda_q = 1$ for $q \in a^* \cup \{ab\}^*$.

The remaining cases are divided according to the additional requirement $|q| - 2|q_0| \geq 3$ and its complementary one $|q| - 2|q_0| \leq 2$.

7.2 The case $|q| - 2|q_0| \geq 3 \wedge |\bar{q}| + |q_0| \geq 3$

Under the additional requirements $|\bar{q}| + |q_0| \geq 3$ and $|\bar{q}| < |q_0|$ this condition is equivalent to the fact that $|\bar{q}| \geq 3$ or the repetition factor $k \geq 3$. Moreover, then $|q| = 7$ (where $q = (ab)^3 a$) or $|q| \geq 9$.

From Eq. (4) we have

$$\sqrt[*]{P_q} \subseteq \{q_0\} \cup \{v : v \sqsubseteq q \wedge |v| > |q| - |q_0|\} \tag{20}$$

This implies that for $|q_0| \leq |q|/2$ the polynomials $p_q(t)$ have non-zero coefficients only for $|q| = n$, $|q| - |q_0| = n - \ell$ and $i < |q_0|$., that is, are of the form $p_q(t) = t^n - t^{n-\ell} - \sum_{i \in M_q} t^i$ where $M_q \subseteq \{i : i < \ell\}$.[4] Therefore, in the sequel we consider the positive roots of polynomials in

$$\mathscr{P}_{\text{red}} := \{t^n - t^{n-\ell} - \sum_{i \in M} t^i : n \geq 1 \wedge \ell \leq \frac{n}{2} \wedge M \subseteq \{i : i < \ell\}\}$$

Let $p_{n,\ell}(t) := t^n - t^{n-\ell} - \sum_{i=0}^{\ell-1} t^i \in \mathscr{P}_{\text{red}}$ and $\lambda_{n,\ell}$ be its maximal root. Similar to Property 6.1, Criterion 2 and Theorem 3 we have the following. Let $n \geq 3, \ell \leq \frac{n}{2}$ and $p(t) \in \mathscr{P}_{\text{red}}$. Then $p(t) \geq p_{n,\ell}(t)$ for $t \in [1,2]$, and $p_{n,\ell}(t)$ has the largest positive root among all polynomials of degree n and parameter ℓ in $\mathscr{P}_{\text{red}}$.

[4] Eq. (4) shows that even $M_q \subseteq \{i : i < \ell - 1\}$. For the Eq. (21) below this stronger version is not needed.

Lemma 11 *If $q, |q| = n$, is a quasiperiod with $|q_0| = \ell \le n/2$ then $p_q(t) \ge p_{n,\ell}(t)$ for $t \ge 1$, in particular, $\lambda_q \le \lambda_{n,\ell}$.*

We have the following relation between the polynomials $p_n(t)$ and $p_{n,\ell}(t)$.

$$p_n(t) - t^\ell \cdot p_{n-2\ell}(t) = p_{n,\ell}(t), \text{ for } n - 2\ell \ge 3 \tag{21}$$

This yields

Corollary 6 *Let $n - 2 \cdot \ell \ge 3$. If $\lambda_n < \lambda_{n-2\ell}$ then $\lambda_{n,\ell} < \lambda_n$.*

Proof. If $\lambda_n < \lambda_{n-2\ell}$ then $p_{n-2\ell}(\lambda_n) < p_{n-2\ell}(\lambda_{n-2\ell}) = 0$. Thus $p_{n,\ell}(\lambda_n) = \lambda_n^\ell \cdot p_{n-2\ell}(\lambda_n) > 0$, that is, $\lambda_n > \lambda_{n,\ell}$. ❑

Next we show the relation $\lambda_q < \lambda_{|q|}$ for all quasiperiods q having $|q_0| \le |q|/2$ and $|q_0| + |\bar{q}| \ge 3$.

Lemma 12 *Let $|q| - 2|q_0| \ge 3$ and $|q_0| + |\bar{q}| \ge 3$. Then $\lambda_q < \lambda_{|q|}$.*

Proof. Above we have shown that $|q| - 2|q_0| \ge 3$ and $|q_0| + |\bar{q}| \ge 3$ imply $|q| \ge 7$ or $|q| \ge 10$ according to whether $|q|$ is odd or even.

The ordering of Theorem 4 and Corollary 6 show $\lambda_n > \lambda_{n,\ell}$ for all odd values $n \ge 7$ and for all even values $n \ge 12$.

It remains to consider the exceptional case when $|q| = 10$. Here $|q| - 2|q_0| \ge 3$ and $|q_0| + |\bar{q}| \ge 3$ imply $|q_0| = 3$. Then Eq. (4) shows $\sqrt[*]{P_q} = \{q_0, q\}$ whence $p_q(t) = t^{10} - t^7 - 1 = p_{10}(t) - t^2 \cdot p_5(t)$.

From $\lambda_5 > \lambda_{10}$ and $p_{10}(\lambda_{10}) = 0$ we have $p_q(\lambda_{10}) = -\lambda_{10}^2 \cdot p_5(\lambda_{10}) > 0$, that is, $\lambda_q < \lambda_{10}$. ❑

Remark 2 *In the exceptional case when $n = 10$ and $\ell = 3$ we have indeed $\lambda_{10,3} > \lambda_{10}$. This follows from $p_{10}(t) - p_{10,3}(t) = t^3 \cdot p_4(t)$ and $\lambda_4 < \lambda_{10}$.*

7.3 The case $|q| - 2|q_0| \le 2 \wedge |q_0| + |\bar{q}| \ge 3$

This amounts to $|q| = 2 \cdot |q_0| + |\bar{q}|$ where $|\bar{q}| \in \{0, 1, 2\}$.

Here we have to go into more detail and to take into consideration also the reduced quasiperiod $\hat{q} = q_0 \cdot \bar{q}$ of q and its repeated prefix $\hat{q}_0 = \min_{\sqsubseteq} P_{\hat{q}}$. Observe that both repeated prefixes $q_0, \hat{q}_0$ are primitive.

Taking into consideration the repeated prefix $\hat{q}_0$, for $q = q_0^k \cdot \bar{q}, k \ge 2$, we have from Eqs. (5) and (6)

$$p_q(t) \in \{t^{|q|} - t^{|q|-|q_0|} - \textstyle\sum_{i \in M} t^i : M \subseteq \{0, \ldots, |\hat{q}| - |\hat{q}_0|\}\}.$$

Observe that $|\hat{q}| - |\hat{q}_0| = |q_0| - (|\hat{q}_0| - |\bar{q}|) < |q_0|$.

Let $\mathscr{P}'_{\text{red}} := \{t^n - t^\ell - \sum_{i\in M} t^i : n > \ell > j \wedge M \subseteq \{0,\dots,\ell-j\}\}$ and $p_{n,\ell,j}(t) = t^n - t^\ell - \sum_{i=0}^{\ell-j} t^i$. Then similar to Property 7.2 and Lemma 11 we have Let $n,\ell \geq 3, \ell \leq \frac{n}{2}, \ell > j$, and $p(t) \in \mathscr{P}'_{\text{red}}$. Then $p(t) \geq p_{n,\ell,j}(t)$ for $t \in [1,2]$, and $p_{n,\ell,j}(t)$ has the largest positive root among all polynomials of degree n and parameters ℓ and j in $\mathscr{P}'_{\text{red}}$.

Lemma 13 *If $q, |q| = n$, is a quasiperiod with $|q_0| = \ell \leq n/2$ and $|\hat{q}_0| - |\bar{q}| \geq j$ then $p_q(t) \geq p_{n,\ell,j}(t)$ for $t \geq 1$, in particular, $\lambda_q \leq \lambda_{n,\ell,j}$.*

We consider the cases $|\bar{q}| \in \{0,1,2\}$ separately.

7.3.1 The case $q = q_0^2 \wedge |\bar{q}| = 0$

In view of Section 7.1 we may consider only the case when $|q_0| \geq 3$. Here we have the following relation between $p_{2\ell}(t)$ and $p_{2\ell,\ell,3}(t)$.

$$p_{2\ell}(t) - p_{2\ell,\ell,3}(t) = t^{\ell-2}(t^2 - t - 1) \tag{22}$$

Lemma 14 *If $q = q_0^2$ and $|q_0| = \ell \geq 3$ then $\lambda_q < \lambda_{|q|}$.*

Proof. First we suppose $|\hat{q}_0| \geq 3$. Then $|\hat{q}_0| - |\bar{q}| \geq 3$ and Property 7.3 and Lemma 13 yield $p_q(t) \geq p_{2\ell,\ell,3}(t)$ for $t \in [1,2]$. Now Eq. (22) shows $p_q(\lambda_{2\ell}) \geq p_{2\ell,\ell,3}(\lambda_{2\ell}) = -\lambda_{2\ell}^{\ell-2}(\lambda_{2\ell}^2 - \lambda_{2\ell} - 1)$. Since $t^2 - t - 1 < 0$ and $p_q(t) \geq p_{2\ell,\ell,3}(t)$ for $1 \leq t \leq \lambda_3 = \max\{\lambda_n : n \in \mathbb{N}\}$ and $\lambda_{2\ell} < \lambda_3$, it follows $p_q(\lambda_{2\ell}) > 0$, that is $\lambda_q < \lambda_{2\ell}$.

It remains to consider $1 \leq |\hat{q}_0| \leq 2$. If $\hat{q}_0 \in a^*$ then $q_0 = a^\ell$ which is not primitive. Thus $\hat{q}_0 = ab$ and, since q_0 is primitive, $q_0 = (ab)^m a$, $m \geq 1$, and $q = q_0^2 = (ab)^m a \cdot (ab)^m a$. We obtain $\sqrt[*]{P_q} = \{(ab)^m a \cdot (ab)^i : i = 0,\dots,m\}$ and, consequently, $p_q(t) = t^{4m+2} - \sum_{i=0}^m t^{2i+1}$. From $p_q(t) = t^{4m+2} - \sum_{i=0}^m t^{2i+1} \geq p_{4m+2}(t) - t^{2m-2}(t^3 - t^2 - 1)$ and $t^3 - t^2 - 1 < 0$ for $1 < t \leq \lambda_3$, in the same way as above, we obtain $p_q(\lambda_{4m+2}) > 0$. ❏

7.3.2 The case $q = q_0^2 \cdot \bar{q} \wedge |\bar{q}| = 1$

Here we have the following relation between $p_{2\ell+1}(t)$ and $p_{2\ell+1,\ell,2}(t)$.

$$p_{2\ell+1}(t) - p_{2\ell+1,\ell,2}(t) = t^{\ell-1}(t^2 - t - 1) \tag{23}$$

Lemma 15 *If* $q = q_0^2 \cdot a, a \in X$, *then* $\lambda_q < \lambda_{|q|}$.

Proof. First we suppose $|\hat{q}_0| - |\bar{q}| \geq 2$. Then $\ell = |q_0| \geq |\hat{q}_0| \geq 3$, and Property 7.3 and Eq. (23) yield $p_q(\lambda_{2\ell+1}) \geq p_{2\ell+1,\ell,2}(\lambda_{2\ell+1}) = p_{2\ell+1}(\lambda_{2\ell+1}) - \lambda_{2\ell+1}^{\ell-1}(\lambda_{2\ell+1}^2 - \lambda_{2\ell+1} - 1)$.

Since $t^2 - t - 1 < 0$ and $p_q(t) \geq p_{2\ell+1,\ell,2}(t)$ for $1 < t \leq \lambda_3$ and $\lambda_{2\ell+1} < \lambda_3$, it follows $p_q(\lambda_{2\ell+1}) > 0$, that is $\lambda_q < \lambda_{2\ell+1}$.

It remains to consider $|\hat{q}_0| = 2$. Then Lemma 2 implies $\hat{q}_0 = q_0$ whence $q = ababa$. Now, one easily verifies $\lambda_{ababa} < \lambda_5 = \lambda_3$ ❑

7.3.3 The case $q = q_0^2 \cdot \bar{q} \wedge |\bar{q}| = 2$

Here we have the following relation between $p_{2\ell+2}(t)$ and $p_{2\ell+2,\ell,2}(t)$.

$$p_{2\ell+2}(t) - p_{2\ell+2,\ell,2}(t) = t^{\ell-1}(t^3 - t - 1) \tag{24}$$

Lemma 16 *If* $q = q_0^2 \cdot \bar{q}$ *with* $|\bar{q}| = 2$ *then* $\lambda_q < \lambda_{|q|}$.

Proof. First we suppose $|\hat{q}_0| \geq 4$. Then Property 7.3 and Eq. (24) yield $p_{2\ell+2}(\lambda_{2\ell+2}) - p_q(\lambda_{2\ell+2}) \leq p_{2\ell+2}(\lambda_{2\ell+2}) - p_{2\ell+2,\ell,2}(\lambda_{2\ell+2}) = \lambda_{2\ell+2}^{\ell-1}(\lambda_{2\ell+2}^3 - \lambda_{2\ell+2} - 1)$.

Since $t^3 - t - 1 < 0$ and $p_q(t) \geq p_{2\ell+2,\ell,2}(t)$ for $1 < t \leq \max\{\lambda_{2n} : n \in \mathbb{N}\} < \lambda_3$ and $\lambda_{2\ell+2} < \lambda_3$, it follows $p_q(\lambda_{2\ell+2}) > 0$, that is, $\lambda_q < \lambda_{2\ell+2}$.

It remains to consider $|\hat{q}_0| = 3$. Again, Lemma 2 implies $\hat{q}_0 = q_0$. Then $|q_0| = 3$ and $|q| = 8$, and Eq. (4) yields $\sqrt[*]{P_q} \subseteq \{q_0, v, q\}$ where $v \sqsubset q$ and $|v| = |q| - 1 = 7$ whence $p_q(t) \geq t^8 - t^5 - t - 1 = p_8(t) - t^2 \cdot p_3(t)$ for $1 \leq t \leq \lambda_3$.

This shows $p_q(\lambda_8) \geq -\lambda_8^2 \cdot p_3(\lambda_8) > 0$, that is, $\lambda_q < \lambda_8$. ❑

Our main theorem then follows from Theorem 3 and the results of Section 7.

Theorem 5 *If* $q \in X^*, |q| \geq 3$, *is a quasiperiod then* $\lambda_q \leq \lambda_{|q|}$, *and* $\lambda_q = \lambda_{|q|}$ *if and only if* $q \in Q_{\max}$.

References

[AFI91] Alberto Apostolico, Martin Farach, and Costas S. Iliopoulos. Optimal superprimitivity testing for strings. *Inform. Process. Lett.*, 39(1):17–20, 1991.

[BP85] Jean Berstel and Dominique Perrin. *Theory of Codes*. Academic Press, Orlando, FL, 1985.

[BR88] Jean Berstel and Christophe Reutenauer. *Rational Series and Their Languages*, volume 12 of *EATCS Monographs on Theoretical Computer Science*. Springer-Verlag, Berlin, 1988.

[Brz67] Janusz A. Brzozowski. Roots of star events. *J. ACM*, 14(3):466–477, 1967.

[FW65] Nathan Jacob Fine and Herbert Saul Wilf. Uniqueness theorems for periodic functions. *Proc. Am. Math. Soc.*, 16:109–114, 1965.

[Lot97] M. Lothaire. *Combinatorics on Words*. Cambridge University Press, Cambridge, 2nd edition, 1997.

[Mou00] Laurent Mouchard. Normal forms of quasiperiodic strings. *Theoret. Comput. Sci.*, 249(2):313–324, 2000.

[PS10] Ronny Polley and Ludwig Staiger. The maximal subword complexity of quasiperiodic infinite words. In Ian McQuillan and Giovanni Pighizzini, editors, *Proceedings 12th DCFS*, volume 31 of *Electr. Proc. Theor. Comput. Sci. (EPTCS)*, pages 169–176. Open Publishing Association, 2010.

[Shy01] Huei-Jan Shyr. *Free Monoids and Languages*. Hon Min Book Company, Taichung, third edition, 2001.

[SS78] Arto Salomaa and Matti Soittola. *Automata-Theoretic Aspects of Formal Power Series*. Springer-Verlag, New York, 1978.

[Sta85] Ludwig Staiger. The entropy of finite-state ω-languages. *Problems Control Inform. Theory/Problemy Upravlen. Teor. Inform.*, 14(5):383–392, 1985.

[Sta18] Ludwig Staiger. Quasiperiods of infinite words. In Alexandra Bellow, Cristian S. Calude, and Tudor Zamfirescu, editors, *Mathematics Almost Everywhere. In Memory of Solomon Marcus*, pages 17–36. World Scientific, Hackensack, NJ, 2018.