# A note on the base- $p$ expansions of putative counterexamples to the $p$-adic Littlewood conjecture 

J. Blackman ${ }^{\text {a }}$, S. Kristensen ${ }^{\text {b,* }}$, M.J. Northey ${ }^{\text {c }}$<br>${ }^{\text {a }}$ University of Liverpool, Department of Mathematical Sciences, Liverpool, L69 7ZL, UK<br>${ }^{\mathrm{b}}$ Aarhus University, Department of Mathematics, Ny Munkegade 118, 8000 Aarhus C, Denmark<br>${ }^{\text {c }}$ Department of Mathematical Sciences, Mathematical Sciences and Computer Science Building, Durham<br>University, Upper Mountjoy Campus, Stockton Road, Durham, DH1 3LE, UK

Received 17 November 2023; accepted 29 February 2024


#### Abstract

In this paper, we investigate the base- $p$ expansions of putative counterexamples to the $p$-adic Littlewood conjecture of de Mathan and Teulié. We show that if a counterexample exists, then so does a counterexample whose base- $p$ expansion is uniformly recurrent. Furthermore, we show that if the base- $p$ expansion of $x$ is a morphic word $\tau\left(\varphi^{\omega}(a)\right)$ where $\varphi^{\omega}(a)$ contains a subword of the form $u X u X u$ with $\lim _{n \rightarrow \infty}\left|\varphi^{n}(u)\right|=\infty$, then $x$ satisfies the $p$-adic Littlewood conjecture. In the special case when $p=2$, we show that the conjecture holds for all pure morphic words. © 2024 The Author(s). Published by Elsevier GmbH. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).


MSC: primary 11J04; secondary 11J61
Keywords: Diophantine approximation; p-adic Littlewood conjecture; Combinatorics on words

## 1. Introduction

The $p$-adic Littlewood conjecture ( pLC ) is an open problem in Diophantine approximation, first proposed by de Mathan and Teulié [8] in 2004, which states that for each prime number $p$ and all $x \in \mathbb{R}$ the following equality holds

$$
\begin{equation*}
\liminf _{q \rightarrow \infty} q \cdot|q|_{p} \cdot\|q x\|=0 \tag{1}
\end{equation*}
$$

[^0]Here, $|\cdot|_{p}$ denotes the $p$-adic absolute value, $\|\cdot\|$ denotes the distance to the nearest integer, and $q$ runs over the positive integers. It follows trivially that if a real number $x$ is well-approximable, i.e.,

$$
\liminf _{q \rightarrow \infty} q \cdot\|q x\|=0
$$

then $x$ satisfies pLC , for all primes $p$.
In the paper that introduced this problem, de Mathan and Teulié showed that (1) is equivalent to the condition that for each real number $x$ and all non-negative integers $k$, the partial quotients of $p^{k} x$ are not uniformly bounded from above.

Lemma 1.1 ([8, Lemma 1.3]). For each $k \in \mathbb{Z}_{\geq 0}$, let $\overline{p^{k} x}=\left[a_{0, k} ; a_{1, k}, \ldots\right]$ be the continued fraction expansion of $p^{k} x$. Then condition (1) is equivalent to

$$
\begin{equation*}
\sup \left\{a_{i, k}: i \geq 1, k \geq 0\right\}=+\infty \tag{2}
\end{equation*}
$$

In particular, the p-adic Littlewood conjecture is deeply connected to how the partial quotients of a real number behave under iterative prime multiplication. Note that since

$$
\frac{1}{\sup _{i \geq 1}\left\{a_{i, k}\right\}+2} \leq \inf _{q \geq 1}\left\{q \cdot\left\|q p^{k} x\right\|\right\} \leq \frac{1}{\sup _{i \geq 1}\left\{a_{i, k}\right\}}
$$

for all $k \in \mathbb{Z}_{\geq 0}$ (see [7, Ch. 7]), conditions (1) and (2) are also equivalent to

$$
\begin{equation*}
\inf _{k \geq 0} \inf _{q \geq 1} q \cdot\left\|q p^{k} x\right\|=0 \tag{3}
\end{equation*}
$$

The main results regarding this conjecture can be broadly separated into two categories: (1) results which induce restrictions on the structure of the continued fraction expansions of potential counterexamples to pLC , and (2) results regarding the measure of the set of counterexamples to pLC and related objects. Notable works regarding the continued fraction expansion of putative counterexamples to pLC include that of de Mathan and Teulié [8], which shows that quadratic irrationals satisfy pLC; Bugeaud, Drmota and de Mathan [6], which shows that all real numbers which have arbitrarily many repetitions of a given finite block in their continued fraction expansion satisfy pLC; and Badziahin, Bugeaud, Einsiedler and Kleinbock [4], which shows that the complexity function of the continued fraction expansion of a counterexample to pLC must grow subexponentially, but the continued fraction expansion cannot be recurrent, see Section 2.1 for a definition. In particular, the complexity function cannot grow too quickly or too slowly. The main result regarding the measure of the set of potential counterexamples is that of Einsiedler and Kleinbock [10], which shows that for each prime $p$ the set of real numbers that do not satisfy (1) has Hausdorff dimension 0. In fact, a stronger result was shown: this set is a countable union of sets which have box-counting dimension zero.

In this manuscript, instead of looking at the continued fraction expansions of potential counterexamples to pLC, we will look at the base- $p$ expansions (see Section 2), which for the most part appear to have been largely unexplored. Our main results are presented in Section 2. In Section 2.1, we look at the base- $p$ expansions of potential counterexamples to pLC and put restrictions on the type of repetitive blocks that can occur in these expansions. Furthermore, we show that if any counterexamples to pLC exist, then there
exist counterexamples with uniformly recurrent base-p expansions. In Section 2.2, we utilise the results of Section 2.1 to analyse the 2 -adic Littlewood conjecture. Due to the simpler alphabet, we are able to provide stronger results. In particular, we show that any real number with a pure morphic base-2 expansion satisfies 2LC and that no counterexample to 2LC can have arbitrarily long overlap-free subwords - see Definition 1.2. The proofs of the results of Section 2.1 are contained in Section 3 and the proofs for Section 2.2 are contained in Section 4.

### 1.1. Notation

Let $\mathcal{A}$ be a finite set which we refer to as an alphabet and let $\mathcal{A}^{*}$ be the set of all finite words over $\mathcal{A}$ including the empty word, which we denote as $\epsilon$. The set $\mathcal{A}^{*}$ forms a free monoid over $\mathcal{A}$ generated by concatenation. We denote the set of (right-sided) infinite words of $\mathcal{A}$ as $\mathcal{A}^{\omega}$, and denote the union of this set with $\mathcal{A}^{*}$ as $\mathcal{A}^{\infty}$. Given these notions, we define the length $|\cdot|$ of a word $w \in \mathcal{A}^{\infty}$ to be the number of letters that appear in $w$, where $|\epsilon|=0$ and $|w|=\infty$ if $w \in \mathcal{A}^{\omega}$.

Definition 1.2. A finite word $w \in \mathcal{A}^{*}$ is an $\alpha$-power if it can be written in the form $w=v^{\lfloor\alpha\rfloor} v^{\prime}$ where $\left|v^{\prime}\right| /|v| \geq\{\alpha\}:=\alpha-\lfloor\alpha\rfloor$. A word $w \in \mathcal{A}^{\infty}$ is overlap-free if it contains no subword of the form $u X u X u$, where $u \in \mathcal{A}$ and $X \in \mathcal{A}^{*}$.

Note that a word contains an overlap if and only if it contains a subword that is a $(2+\delta)$-power for some $\delta>0$.

### 1.1.1. Morphic words

An important class of words are the morphic words. As a special case, these include all automatic words, i.e., words which can be generated by a finite automaton with output. Let $\varphi: \mathcal{A} \rightarrow \mathcal{A}^{*}$ be a morphism. If there is some natural number $j \geq 1$ such that $\varphi^{j}(a)=\epsilon$, for $a \in \mathcal{A}$, then $a$ is said to be mortal. The set of mortal letters is denoted by $M_{\varphi}$. A morphism $\varphi$ is prolongable on the letter $a \in \mathcal{A}$, if $\varphi(a)=a x$ and $x \notin M_{\varphi}{ }^{*}$. If a morphism is prolongable on $a$, then the words $a, \varphi(a), \varphi^{2}(a), \ldots$ converge to an infinite word $\varphi^{\omega}(a)$ of the form

$$
\begin{equation*}
\varphi^{\omega}(a)=a x \cdot \varphi(x) \cdot \varphi^{2}(x) \cdot \ldots \tag{4}
\end{equation*}
$$

Any word that can be formed in this way is referred to as a pure morphic word. If there is a coding $\tau: \mathcal{A} \rightarrow \mathcal{B}$ - i.e., a morphism that maps letter to letter - such that $w=\tau\left(\varphi^{\omega}(a)\right)$, then $w$ is referred to as a morphic word. A morphism $\varphi: \mathcal{A} \rightarrow \mathcal{A}^{*}$ is $k$-uniform if $|\varphi(a)|=k$ for all $a \in \mathcal{A}$ and is expanding if $|\varphi(a)| \geq 2$ for all $a \in \mathcal{A}$. A morphism $\varphi$ is primitive if there exists some exponent $n \geq 1$ such that for every $a, b \in \mathcal{A}$, the letter $b$ appears in the word $\varphi^{n}(a)$ at least once.

Example 1.3. The Thue-Morse word $M$ is the overlap-free, infinite word that is the limit $\mu^{\omega}(0)$ of the morphism $\mu:\{0,1\} \rightarrow\{0,1\}^{*}$ with $\mu(0):=01$ and $\mu(1):=10$. The first few letters are

$$
M=0110100110010110 \cdots
$$

The complement of the Thue-Morse word $\tilde{M}$ is the word given by $\mu^{\omega}(1)$.

## 2. Main results

For every $x \in[0,1]$ and every natural number $n \geq 2$, we can rewrite $x$ in the following form

$$
x=\sum_{i=1}^{\infty} a_{i} n^{-i}
$$

where $a_{i} \in\{0,1, \ldots, n-1\}$ for all $i \in \mathbb{N}$. Unless the number $x$ is a rational number with denominator $n^{k}$ for some $k \geq 1$, this series expansion is unique. Since pLC is clearly satisfied for rational numbers, we will disregard this case and only consider real numbers that correspond to a unique sequence of digits. The word formed by taking the coefficients of this power series is called the base-n expansion of $x$. We denote this word as $w(x, n)$, i.e., $w(x, n):=a_{1} a_{2} \cdots$. Conversely, given a word $w \in\{0,1, \ldots, n-1\}^{\omega}$, we will denote the real number whose base- $n$ expansion coincides with $w$ as $w_{n}$. If $\{n x\}$ is the fractional part of $n x$, i.e., $\{n x\}:=n x-\lfloor n x\rfloor$, then the corresponding base- $n$ expansion is $T\left(a_{1} a_{2} a_{3} \cdots\right):=a_{2} a_{3} \cdots$. In particular, up to taking the number modulo 1 , the shift map $T$ induces multiplication by $n$. More generally, the base- $n$ expansion of $\left\{n^{k} x\right\}$ corresponds to the word $T^{k}\left(a_{1} a_{2} a_{3} \cdots\right)=a_{k+1} a_{k+2} \cdots$.

Due to this structure, the base-n expansion is very well-equipped for producing information regarding the limiting behaviour of a real point under repeated multiplication by $n$. Whilst the rational approximations coming from the base- $n$ (or base- $p$ ) expansion are typically worse than the rational approximations coming from the continued fraction expansion, in a number of cases this approximation is still good enough to induce restrictions on the potential counterexamples of pLC. On the other hand, whilst the continued fraction expansion gives a very good rational approximation of a real number, the integer multiplication of continued fractions is far more complicated - see [12,14].

For our purposes, it will also be useful to deal with base-n representations of integers. For any integer $a \geq 0$, we can uniquely write $a$ as

$$
\sum_{i=1}^{m} a_{i} n^{m-i}
$$

with $a_{i} \in\{0,1, \ldots, n-1\}$ and $a_{m} \neq 0$ (unless $m=1$ ). The word $v(a, n)$ formed by taking the coefficients of this sum is the base-n representation of $a$. Given a finite, non-empty word $v$, let $v_{n}^{+}$denote the integer whose base- $n$ representation coincides with $v$.

### 2.1. The p-adic Littlewood conjecture

For a finite word $w$ on some alphabet $\mathcal{A}$ and a $\delta \in(0,1)$, we will denote the prefix of the word $w$ of length $\lfloor\delta \cdot|w|\rfloor$ as $w^{\delta}$. Note that by construction, $w w w^{\delta}$ is an $\alpha$ power for all $\alpha \leq 2+(\lfloor\delta|w|\rfloor /|w|)$. The following theorem shows that if the base- $p$ expansion of a real number $x$ has a sequence of subwords of the form $w_{j} w_{j} w_{j}^{\delta_{j}}$ with $\lim _{j \rightarrow \infty}\left|w_{j}^{\delta_{j}}\right|=\lim _{j \rightarrow \infty}\left\lfloor\delta_{j} \cdot\left|w_{j}\right|\right\rfloor=\infty$, then $x$ satisfies pLC.

Theorem 2.1. Let $w=\left(a_{n}\right)_{n=1}^{\infty}$ be an infinite word on the alphabet $\{0,1, \ldots, p-1\}$ satisfying the property that there is a sequence $\left(w_{j}\right)_{j=1}^{\infty}$ of finite words and a sequence of
positive real numbers $\left(\delta_{j}\right)_{j=1}^{\infty}$ which are less than 1 such that the word $w_{j} w_{j} w_{j}^{\delta_{j}}$ occurs as a subword in $w$ and $\lim _{j \rightarrow \infty}\left|w_{j}^{\delta_{j}}\right|=\infty$. Then $w_{p}=\sum_{n=1}^{\infty} a_{n} p^{-n}$ satisfies the p-adic Littlewood conjecture.

Taking $\left(\delta_{j}\right)_{j=1}^{\infty}$ to be a constant sequence leads to the following corollary.
Corollary 2.2. Assume $x$ is a counterexample to $p L C$ and let $w(x, p)$ be the corresponding base-p expansion. For each fixed $\alpha>2$, the length of the $\alpha$-powers appearing in $w(x, p)$ are bounded.

Theorem 2.1 can be generalised as follows.
Theorem 2.3. Let $w=\left(a_{n}\right)_{n=1}^{\infty}$ be an infinite word on the alphabet $\{0,1, \ldots, p-1\}$ that contains a sequence $\left(w_{j}\right)_{j=1}^{\infty}$ of finite words with $m_{j}=\left|w_{j}\right|$ and a sequence of positive real numbers $\left(\delta_{j}\right)_{j=1}^{\infty}$ such that the word $w_{j} w_{j}^{\delta_{j}}$ occurs as a subword in w. Furthermore, let $\left(\ell_{j}\right)_{j=1}^{\infty}$ be the sequence of natural numbers satisfying

$$
\begin{equation*}
p^{\ell_{j}-1} \leq \frac{p^{m_{j}}-1}{\operatorname{gcd}\left(p^{m_{j}}-1,\left(w_{j}\right)_{p}^{+}\right)} \leq p^{\ell_{j}} \tag{5}
\end{equation*}
$$

If $\lim _{j \rightarrow \infty} m_{j}+\left\lfloor m_{j} \delta_{j}\right\rfloor-2 \ell_{j}=\infty$, then $w_{p}$ satisfies $p L C$.
In the above theorem, the three most useful cases are:

- when $\operatorname{gcd}\left(p^{m_{j}}-1,\left(w_{j}\right)_{p}^{+}\right)=1, \ell_{j}=m_{j}$, and $\lim _{j \rightarrow \infty}\left\lfloor m_{j} \delta_{j}\right\rfloor-m_{j}=\infty$ (Theorem 2.1),
- when $m_{j}=2 n_{j}$ with $n_{j} \in \mathbb{N}, \operatorname{gcd}\left(p^{m_{j}}-1,\left(w_{j}\right)_{p}^{+}\right)=p^{n_{j}}-1, \ell_{j}=n_{j}+1$ and $\lim _{j \rightarrow \infty}\left\lfloor m_{j} \delta_{j}\right\rfloor=\infty$, and
- when $\lim _{j \rightarrow \infty} \delta_{j}=\infty$.

As an example of how the second of the above bullet points can be used, given a word $w=b_{1} b_{2} \cdots b_{n}$ in $\{0,1, \ldots, p-1\}^{*}$, the integer $(w \bar{w})_{p}^{+}$will always be divisible by $p^{n}-1$ where $\bar{b}=p-1-b$ for letter each $b$ in the alphabet $\{0,1, \ldots, p-1\}$. This follows since

$$
\sum_{i=1}^{n} p^{n-i} \cdot\left[p^{n} b_{i}+p-1-b_{i}\right]=\left(p^{n}-1\right)+\sum_{i=1}^{n}\left(p^{n}-1\right) p^{n-i} b_{i}
$$

Thus, we obtain the following corollary.
Corollary 2.4. Let $w=\left(a_{n}\right)_{n=1}^{\infty}$ be an infinite word on the alphabet $\{0,1, \ldots, p-1\}$ satisfying the property that there is a sequence $\left(w_{j}\right)_{j=1}^{\infty}$ of finite words and a sequence of positive real numbers $\left(\delta_{j}\right)_{j=1}^{\infty}$ such that the word $w_{j} \overline{w_{j}} w_{j}^{\delta_{j}}$ occurs as a subword in $w$ and $\lim _{j \rightarrow \infty}\left|w_{j}^{\delta_{j}}\right|=\infty$. Then $w_{p}$ satisfies the p-adic Littlewood conjecture.

Another property that can be deduced is that if a word $w$ contains a sequence of increasing prefixes of another word $v$ and $v_{p}$ satisfies pLC , then so does $w_{p}$.

Proposition 2.5. Let $w, v \in\{0,1, \ldots, p-1\}^{\omega}$ and assume that there exists a sequence of prefixes $\left(v_{k}\right)_{k=1}^{\infty}$ of $v$ such that $\left|v_{k}\right| \rightarrow \infty$ and $v_{k}$ appears as a subword of $w$ for all $k$. If $v_{p}$ satisfies $p L C$, then so does $w_{p}$.

An infinite word $w=\left(a_{n}\right)_{n=1}^{\infty}$ is said to be recurrent if any finite subword $v$ of $w$ occurs infinitely often in $w$. It is said to be uniformly recurrent if for every finite subword $v$ of $w$, there exists a constant $N_{v}$ such that $v$ appears in every subword of $w$ of length $N_{v}$. Using an idea similar to the work of Badziahin [3] on "limit words" of continued fraction expansions, we can look at the topological closure of the set of base- $p$ expansions of the counterexamples to pLC under the action of the shift map. This allows us to deduce that if this set is non-empty, then it contains an element with a uniformly recurrent base- $p$ expansion.

Theorem 2.6. If there is a counterexample to pLC, there is a counterexample with a uniformly recurrent base-p expansion.

Remark 2.7. It is worth noting that none of the above statements rely on $p$ being prime other than to link to the $p$-adic Littlewood conjecture. In particular, we can replace $p$ with a composite number $n$ to obtain analogous results on the " $n$-adic Littlewood conjecture".

The proof of Theorems 2.1 and 2.3 can be found in Section 3.1. The proof of Proposition 2.5 and Theorem 2.6 is in Section 3.2.

### 2.1.1. Results on morphic words

Let $w=\varphi^{\omega}(a)$ be a pure morphic word. If the prefix $\varphi^{k}(a)$ contains overlap of the form $u X u X u$ for some $k \in \mathbb{N}$, then $\varphi^{n}(u) \varphi^{n}(X) \varphi^{n}(u) \varphi^{n}(X) \varphi^{n}(u)$ is a subword of $\varphi^{k+n}(a)$ for all $n \in \mathbb{N}$. Under the assumption that $u$ is not mortal for $\varphi$, infinitely many instances of overlap occur. Furthermore, if $\lim _{n \rightarrow \infty}\left|\varphi^{n}(u)\right|=\infty$, the word satisfies the conditions of Theorem 2.1. This leads to the following proposition.

Proposition 2.8. Let $w=\varphi^{\omega}(a) \in \mathcal{A}^{\omega}$ be a pure morphic word containing a subword uХиХи such that $\lim _{n \rightarrow \infty}\left|\varphi^{n}(u)\right|=\infty$. For any non-erasing morphism $g: \mathcal{A} \rightarrow$ $\{0,1, \ldots, p-1\}$, the real number $g(w)_{p}$ satisfies the $p$-adic Littlewood conjecture.

Remark 2.9. Here we should note that the condition $\lim _{n \rightarrow \infty}\left|\varphi^{n}(u)\right|=\infty$ is instantly satisfied for morphisms which are expanding, including (powers of) primitive morphisms and $k$-uniform morphisms for $k \geq 2$. Furthermore, due to a result of Durand [9], all uniformly recurrent morphic words are primitive morphic. Therefore, if $x$ is a counterexample to pLC with a morphic, uniformly recurrent base- $p$ expansion of the form $\tau\left(\varphi^{\omega}(a)\right)$, then the underlying pure morphic word $\varphi^{\omega}(a)$ must be overlap-free.

Similar to the previous argument, if a morphism $\varphi$ is prolongable on the letters $a, b \in \mathcal{A}^{*}$ and $b$ appears in the word $\varphi^{\omega}(a)$ at least once, then every prefix of $\varphi^{\omega}(b)$ appears in $\varphi^{\omega}(a)$. Proposition 2.5 then directly implies the following corollary.

Corollary 2.10. Let $w=\varphi^{\omega}(a)$ be a pure morphic word over $\mathcal{A}$ and let $\mathcal{B}$ be a subalphabet of $\mathcal{A}$ such that $\varphi: \mathcal{B} \rightarrow \mathcal{B}^{*}$. Furthermore, assume that $\varphi^{\omega}(a)$ contains a letter
$b \in \mathcal{B}$ such that $\varphi$ is prolongable over $b$ and let $\tau: \mathcal{A} \rightarrow\{0,1, \ldots, p-1\}$ be a coding. If $\tau\left(\varphi^{\omega}(b)\right)_{p}$ satisfies $p L C$, then so does $\tau\left(\varphi^{\omega}(a)\right)_{p}$.

### 2.2. Applications to the 2-adic Littlewood conjecture

In the case of the 2-adic Littlewood conjecture, all pure morphic words satisfy at least one of three properties (P1)-(P3) - see Lemma 4.2. Combining this result with Theorem 2.1 and other results in the literature leads to the following theorem.

Theorem 2.11. Let $x \in[0,1]$ and assume that the corresponding base 2 expansion $w(x, 2)$ is a pure morphic word. Then $x$ satisfies $2 L C$.

This theorem can be extended to a class of results regarding pLC by applying Corollary 2.10.

Corollary 2.12. Let $w=\varphi^{\omega}(a)$ be a pure morphic word over $\mathcal{A}$ and let $\mathcal{B}$ be a subalphabet of $\mathcal{A}$ such that $\varphi: \mathcal{B} \rightarrow \mathcal{B}^{*}$ and $|\mathcal{B}|=2$. Furthermore, assume that $\varphi^{\omega}(a)$ contains a letter $b \in \mathcal{B}$ such that $\varphi$ is prolongable over $b$. Then $\tau(w)_{p}$ satisfies $p L C$ for any coding $\tau: \mathcal{A} \rightarrow\{0,1, \ldots, p-1\}$.

Finally, as a result contrasting with Corollary 2.2, we show that the lengths of the overlap-free subwords of the base-2 expansion of a counterexample to 2 LC are bounded.

Theorem 2.13. Assume that $x$ is a counterexample to $2 L C$ and let $w(x, 2)$ be the corresponding base- 2 expansion. Then the length of the overlap-free subwords in $w(x, 2)$ are bounded.

The proofs of Theorem 2.11 and Corollary 2.12 can be found in Section 4.1 and the proofs of Theorem 2.13 can be found in Section 4.2.

## 3. The $\boldsymbol{p}$-adic Littlewood conjecture

### 3.1. Proof of Theorems 2.1 and 2.3

To prove Theorems 2.1 and 2.3, we will show that the conditions of these theorems imply (3). To this end, we will produce sequences $\left(q_{j}\right)_{j=1}^{\infty}$ and $\left(k_{j}\right)_{j=1}^{\infty}$ of natural numbers such that

$$
\begin{equation*}
\lim _{j \rightarrow \infty} q_{j} \cdot\left\|q_{j} p^{k_{j}} x\right\|=0 \tag{6}
\end{equation*}
$$

Proof of Theorem 2.1. For each $j \in \mathbb{N}$, let $k_{j}$ be the length of the prefix of $\left(a_{n}\right)_{n=1}^{\infty}$ up to the first occurrence of the subword $w_{j} w_{j} w_{j}^{\delta_{j}}$. Set

$$
x^{\prime}:=\left\{p^{k_{j}} x\right\}=\left\{p^{k_{j}} \sum_{n=1}^{\infty} a_{n} p^{-n}\right\}=\sum_{n=1}^{\infty} a_{k_{j}+n} p^{-n} .
$$

Then, the base- $p$ expansion of $x^{\prime}$ begins with the subword $w_{j} w_{j} w_{j}^{\delta_{j}}$.

Now, for each $j$, we denote $w_{j}$ as $b_{1}^{(j)} b_{2}^{(j)} \cdots b_{m_{j}}^{(j)}$ where $m_{j}=\left|w_{j}\right|$, and define a sequence of rational numbers

$$
\frac{r_{j}}{q_{j}}:=\sum_{h=0}^{\infty} \sum_{i=1}^{m_{j}} \frac{b_{i}^{(j)}}{p^{i+h m_{j}}}=\sum_{h=0}^{\infty} \frac{1}{p^{h m_{j}}} \sum_{i=1}^{m_{j}} \frac{b_{i}^{(j)}}{p^{i}} .
$$

These are the rational numbers (in reduced form) whose base- $p$ expansion is obtained by extending the word $w_{j}$ periodically. We will show that the sequence of denominators $\left(q_{j}\right)_{j=1}^{\infty}$ can be used in (6).

The numbers $r_{j} / q_{j}$ approximate $x^{\prime}$ rather well. Indeed,

$$
\left|x^{\prime}-\frac{r_{j}}{q_{j}}\right|=\left|\sum_{i=\left\lfloor\delta_{j} m_{j}\right\rfloor+1}^{m_{j}} \frac{c_{2, i}^{(j)}}{p^{i+2 m_{j}}}+\sum_{h=3}^{\infty} \frac{1}{p^{h m_{j}}} \sum_{i=1}^{m_{j}} \frac{c_{h, i}^{(j)}}{p^{i}}\right|<\frac{1}{p^{2 m_{j}+\left\lfloor\delta_{j} m_{j}\right\rfloor}},
$$

where $c_{h, i}^{(j)}=\left(a_{k_{j}+h m_{j}+i}-b_{i}^{(j)}\right)$. On the other hand,

$$
\frac{r_{j}}{q_{j}}=\left(\sum_{h=0}^{\infty} \frac{1}{p^{h m_{j}}}\right)\left(\sum_{i=1}^{m_{j}} \frac{b_{i}^{(j)}}{p^{i}}\right)=\frac{p^{m_{j}}}{p^{m_{j}}-1} \sum_{i=1}^{m_{j}} \frac{b_{i}^{(j)}}{p^{i}}=\frac{r_{j}^{\prime}}{p^{m_{j}}-1},
$$

where $r_{j}^{\prime} \in \mathbb{Z}$. Consequently, $q_{j} \leq p^{m_{j}}-1<p^{m_{j}}$ and therefore,

$$
q_{j} \cdot\left\|q_{j} p^{k_{j}} x\right\| \leq q_{j}^{2} \cdot\left|x^{\prime}-\frac{r_{j}}{q_{j}}\right|<\frac{1}{p^{\left\lfloor\delta_{j} m_{j}\right\rfloor}} .
$$

Since $\delta_{j} \cdot m_{j}$ is assumed to tend to infinity with $j$, the theorem follows.
The above proof illustrates a very useful technique for using combinatorial properties of base- $p$ expansions to show that real numbers satisfy pLC. The proof of Theorem 2.3 serves as generalisation of the above method.

Proof of Theorem 2.3. For each $j \in \mathbb{N}$, let $k_{j}$ be the length of the prefix of $\left(a_{n}\right)_{n=1}^{\infty}$ up to the first occurrence of the subword $w_{j} w_{j}^{\delta_{j}}$ and set

$$
x^{\prime}:=\left\{p^{k_{j}} x\right\}
$$

Then, the base- $p$ expansion of $x^{\prime}$ begins with the subword $w_{j} w_{j}^{\delta_{j}}$. Let $n_{j}=\left\lfloor\delta_{j}\right\rfloor$.
For each $j \in \mathbb{N}$, we denote $w_{j}$ as $b_{1}^{(j)} b_{2}^{(j)} \cdots b_{m_{j}}^{(j)}$ and set

$$
\begin{equation*}
\frac{r_{j}}{q_{j}}:=\sum_{h=0}^{\infty} \sum_{i=1}^{m_{j}} \frac{b_{i}^{(j)}}{p^{i+h m_{j}}}=\left(\sum_{h=0}^{\infty} \frac{1}{p^{(h+1) m_{j}}}\right)\left(\sum_{i=1}^{m_{j}} p^{m_{j}-i} b_{i}^{(j)}\right) \tag{7}
\end{equation*}
$$

These are the rational numbers (in reduced form) whose base- $p$ expansion is obtained by extending the word $w_{j}$ periodically.

As in the proof of Theorem 2.1, this sequence of rational numbers $\frac{r_{j}}{q_{j}}$ approximates $x^{\prime}$ very well,

$$
\left|x^{\prime}-\frac{r_{j}}{q_{j}}\right|=\left|\sum_{i=\left\lfloor\left(\delta_{j}-n_{j}\right) m_{j}\right\rfloor+1}^{m_{j}} \frac{c_{n_{j}, i}^{(j)}}{p^{i+n_{j} m_{j}}}+\sum_{h=n_{j}+1}^{\infty} \frac{1}{p^{h m_{j}}} \sum_{i=1}^{m_{j}} \frac{c_{h, i}^{(j)}}{p^{i}}\right|<\frac{1}{p^{m_{j}+\left\lfloor\delta_{j} m_{j}\right\rfloor}},
$$

where $c_{h, i}^{(j)}=\left(a_{k_{j}+2 h m_{j}+i}-b_{i}^{(j)}\right)$.

Let $d_{j}=\operatorname{gcd}\left(p^{m_{j}}-1,\left(w_{j}\right)_{p}^{+}\right)$. Then there exists some $a \in \mathbb{N}$ such that

$$
a d_{j}=\sum_{i=1}^{m_{j}} p^{m_{j}-i} b_{i}^{(j)} .
$$

Combining this with (7) shows

$$
\frac{r_{j}}{q_{j}}=a \cdot \frac{d_{j}}{p^{m_{j}}-1} .
$$

From (5), it follows that $q_{j} \leq\left(p^{m_{j}}-1\right) / d_{j} \leq p^{\ell_{j}}$ and therefore

$$
q_{j}^{2} \cdot\left|x^{\prime}-\frac{r_{j}}{q_{j}}\right|<\frac{p^{2 \ell_{j}}}{p^{m_{j}+\left\lfloor\delta_{j} m_{j}\right\rfloor}}=\frac{1}{p^{m_{j}+\left\lfloor\delta_{j} m_{j}\right\rfloor-2 \ell_{j}}}
$$

Since it was assumed that $\lim _{j \rightarrow \infty} m_{j}+\left\lfloor\delta_{j} m_{j}\right\rfloor-2 \ell_{j}=\infty$, this completes the proof.

### 3.2. Proof of Proposition 2.5 and Theorem 2.6

Given an infinite word $w \in \mathcal{A}^{\omega}$, we define the set of suffixes $\mathcal{S}(w)$ of $w \in \mathcal{A}^{\omega}$ to be

$$
\mathcal{S}(w):=\left\{T^{k}(w): k \in \mathbb{Z}_{\geq 0}\right\}
$$

We can turn $\mathcal{A}^{\omega}$ into a metric space by defining a metric $d(x, y)=2^{-|u|}$, where $u$ is the largest common prefix of $x$ and $y$ and $d(x, x)=0$. From this, we can take the topological closure of the set of suffixes $\overline{\mathcal{S}(w)}$. A word $v \in \mathcal{A}^{\omega}$ is an element of $\overline{\mathcal{S}(w)}$ if and only if every prefix of $v$ appears in $w$. Analogously, for any $x \in[0,1]$, we can define the set

$$
T_{p}(x):=\left\{\left\{p^{n} x\right\}: n \in \mathbb{N} \cup\{0\}\right\}
$$

Assuming $x$ is not a rational number with denominator equal to $p^{k}$ for some natural number $k \geq 1$, the sets $T_{p}(x)$ and $\mathcal{S}(w(x, p))$ are in bijection, where each real number corresponds to its base-p expansion. Likewise, the topological closures $\overline{\mathcal{T}_{p}(x)}$ (using the Euclidean metric) and $\overline{\mathcal{S}(w(x, p))}$ are also in bijection. This comes from the observation that there is a subsequence $\left\{p^{k_{j}} x\right\}$ that limits to $y$ if and only if the base- $p$ expansions of $\left\{p^{k_{j}} x\right\}$ limit to the base- $p$ expansion of $y$. Using the notions above, the proof of Proposition 2.5 essentially comes down to showing that if any accumulation point of $T_{p}(x)$ satisfies pLC , then $x$ satisfies pLC . The contrapositive of this statement is shown in the next lemma.

Lemma 3.1. Let $x$ be a counterexample to $p L C$ and assume that there exists some $\varepsilon>0$ such that $m_{p}(x) \geq \varepsilon$. Then $m_{p}(y) \geq \varepsilon$ for all $y \in \overline{T_{p}(x)}$.

Proof. We trivially have that $m_{p}(x) \geq \varepsilon$ implies that $m_{p}\left(\left\{p^{n} x\right\}\right) \geq \varepsilon$ for every $n \in \mathbb{N} \cup\{0\}$, since

$$
\liminf _{q \rightarrow \infty} q \cdot|q|_{p} \cdot\left\|q p^{n} x\right\|=\liminf _{q \rightarrow \infty} p^{n} q \cdot\left|p^{n} q\right|_{p} \cdot\left\|p^{n} q x\right\| \geq \liminf _{q \rightarrow \infty} q \cdot|q|_{p} \cdot\|q x\| .
$$

Assume that $y$ is a limit point of $T_{p}(x)$ such that $m_{p}(y)<\varepsilon$. Then there exists some $\varepsilon^{\prime} \in \mathbb{R}$ satisfying $0<\varepsilon^{\prime}<\varepsilon<1$ and some sequence $\left(q_{n}\right)_{n=1}^{\infty}$ such that

$$
\lim _{n \rightarrow \infty} q_{n} \cdot\left|q_{n}\right|_{p} \cdot\left\|q_{n} y\right\|=\varepsilon^{\prime}
$$

For every $n \in \mathbb{N}$, let $\delta_{n}=2^{-1} q_{n}^{-2}\left(\varepsilon-\varepsilon^{\prime}\right)$ and let $k_{n}$ be the smallest natural number such that $\left\{p^{k_{n}} x\right\}=y+\Delta_{n}$ with $\left|\Delta_{n}\right|<\delta_{n}$. The existence of each $k_{n}$ follows from the fact that $y$ is an accumulation point. This implies

$$
\begin{aligned}
p^{k_{n}} q_{n} \cdot\left|p^{k_{n}} q_{n}\right|_{p} \cdot\left\|p^{k_{n}} q_{n} x\right\| & =q_{n} \cdot\left|q_{n}\right|_{p} \cdot\left\|q_{n} p^{k_{n}} x\right\| \\
& =q_{n} \cdot\left|q_{n}\right|_{p} \cdot\left\|q_{n}\left(y+\Delta_{n}\right)\right\| \\
& \leq q_{n} \cdot\left|q_{n}\right|_{p} \cdot\left\|q_{n} y\right\|+q_{n} \cdot\left|q_{n}\right|_{p} \cdot\left\|q_{n} \Delta_{n}\right\| \\
& \leq q_{n} \cdot\left|q_{n}\right|_{p} \cdot\left\|q_{n} y\right\|+q_{n} \cdot\left|q_{n} \Delta_{n}\right| .
\end{aligned}
$$

Then

$$
\lim _{n \rightarrow \infty} p^{k_{n}} q_{n} \cdot\left|p^{k_{n}} q_{n}\right|_{p} \cdot\left\|p^{k_{n}} q_{n} x\right\| \leq \varepsilon^{\prime}+2^{-1}\left(\varepsilon-\varepsilon^{\prime}\right)<\varepsilon
$$

Therefore, $m_{p}(x)<\varepsilon$ which is a contradiction.
We will now use Lemma 3.1 to deduce that should a counterexample $x$ to pLC exist, there exists an element $y$ of $\overline{T_{p}(x)}$ with a uniformly recurrent base- $p$ expansion that is a counterexample to pLC. This is sufficient for Theorem 2.6.

Proposition 3.2. Let $x$ be a counterexample of $p L C$. Then $\overline{T_{p}(x)}$ contains a counterexample of pLC with a uniformly recurrent base-p expansion.

Proof. By construction $\overline{T_{p}(x)}$ is closed and bounded. Therefore, $\overline{T_{p}(x)}$ is compact and invariant under multiplication by $p$. The corresponding set of base- $p$ expansions is given by $\overline{S(w(x, p))}$ and is also compact and invariant under the shift map $T$. At least one minimal, invariant, compact subset $R$ of $\overline{S(w(x, p))}$ exists, and by [13, Theorem 1.5.9], this is a set comprised of numbers with uniformly recurrent base-p expansions. By Lemma 3.1, all elements in $R$ are counterexamples to pLC.

## 4. The 2-adic Littlewood conjecture

### 4.1. Proof of Theorem 2.11

In order to prove Theorem 2.11, it will be useful to first introduce a number of auxiliary results. The first result is that of Seébold [15], which shows that the only pure morphic words over $\{0,1\}$ which are overlap-free are the Thue-Morse word $M$ and its complement $\tilde{M}$.

Theorem 4.1 ([15]). $M$ and $\tilde{M}$ are the only pure morphic overlap-free words in $\{0,1\}^{\omega}$.
Using this theorem, we can give the following characterisation of all binary pure morphic words.

Lemma 4.2. Let $w$ be a pure morphic word in $\{0,1\}^{\omega}$, where $\varphi$ is the underlying morphism. Then (at least) one of the following statements holds:
(P1) $w$ is $M$ or $\tilde{M}$.
(P2) There is a non-trivial subword $v$ of $w$, such that $v^{n}$ is a subword of $w$ for all $n \in \mathbb{N}$.
(P3) $w$ contains overlap of the form $a X a X a$ where $a \in \mathcal{A}$ and $X \in \mathcal{A}^{*}$ and $\lim _{n \rightarrow \infty}\left|\varphi^{n}(a)\right|=\infty$.

Proof. In the case that $w$ is overlap-free, Theorem 4.1 shows that $w$ is the Thue-Morse word $M$ or its complement $\tilde{M}$ (P1).

Assume that $w$ is not overlap-free and that $w=\varphi^{\omega}(0)$ - the case that $w=\varphi^{\omega}(1)$ follows by symmetry. Then, for some words $u, v \in\{0,1\}^{*}$, we have

$$
\varphi(0)=0 u \quad \text { and } \quad \varphi(1)=v .
$$

Since $\varphi$ is prolongable over 0 , the word $u$ is not the empty word. In particular, $|\varphi(0)| \geq 2$. Note that if $u$ consists only 0 's, i.e., $u=0^{n}$ for $n \in \mathbb{N}$, then

$$
w=\varphi^{\omega}(0)=0^{\omega},
$$

where $x^{\omega}$ is the periodic word $x x x \cdots$. Thus, $w$ satisfies (P2).

## Case I: $v$ is the empty word $\epsilon$.

Since $\varphi(1)=\epsilon$, applying the morphism to $\varphi^{k}(0)$ will ignore any 1 's in this sequence. In other words, if $i_{k}$ is the number of 0 's that appear in $\varphi^{k}(0)$, then

$$
\varphi^{k+1}(0)=(\varphi(0))^{i_{k}}
$$

Therefore, $i_{k+1}=i_{k} \cdot i_{1}=i_{1}^{k+1}$. Since $\varphi$ is prolongable, $u$ contains the letter 0 at least once, and so $i_{1} \geq 2$. Since $\lim _{k \rightarrow \infty} i_{k}=\infty$,

$$
w=\varphi^{\omega}(0)=(\varphi(0))^{\omega} .
$$

In this case, $w$ satisfies ( $\mathbf{P 2 ) .}$
Case II: $v=1^{n}$.
As discussed above, we can assume that $u$ contains the letter 1 at least once. If $\varphi(1)=1^{n}$ for some $n \geq 2$, then $\varphi^{k}(1)=1^{n^{k}}$. Since $\varphi(0)$ contains the letter 1 , the word $\varphi^{k+1}(0)$ contains the subword $\varphi^{k}(1)$ for all $k \in \mathbb{N}$. Therefore, $w=\varphi^{\omega}(0)$ satisfies (P2).

Let $v=1$. Note that if $u$ does not contain the letter 0 , i.e., $u=1^{k}$, then

$$
\varphi^{2}(0)=\varphi(0) \varphi\left(1^{k}\right)=01^{2 k}, \quad \varphi^{3}(0)=\varphi(0) \varphi\left(1^{2 k}\right)=01^{3 k}, \text { and } \varphi^{m}\left(01^{k}\right)=01^{m k}
$$

In this case, $\varphi^{\omega}(0)=01^{\infty}$ and (P2) is satisfied. Furthermore, if $u=u^{\prime} 01^{k}$ with $k \in \mathbb{N}$, we note that for all $m \in \mathbb{N}$

$$
\varphi\left(01^{m}\right)=0 u^{\prime} 01^{k} 1^{m}=0 u^{\prime} 01^{(k+m)} .
$$

In particular, for all $n \in \mathbb{N}$ the word $\varphi^{n}(u)$ ends in the term $1^{(n+1) k}$. Therefore, $w$ satisfies (P2), and so we have now reduced our considerations to the cases where $u$ ends in the letter 0 .

If $u=1^{k} 0$ with $k \in \mathbb{N}$, then

$$
\varphi^{2}(0)=\varphi(0) 1^{k} \varphi(0)=\left(01^{k} 0\right) 1^{k}\left(01^{k} 0\right)
$$

Since this word contains the subword $01^{k} 01^{k} 0$ and $\varphi$ is prolongable on 0 , the length of $\varphi^{n}(0)$ tends to infinity and (P3) is satisfied.

Finally, assume that $u=u^{\prime} 01^{k} 0$ with $k \in \mathbb{Z}_{\geq 0}$. The word

$$
\varphi\left(01^{k} 0\right)=0 u^{\prime} 01^{k} 01^{k} 0 u^{\prime} 01^{k} 0
$$

contains $01^{k} 01^{k} 0$ as a subword. Since $\varphi$ is prolongable on 0 , the length of $\varphi^{n}(0)$ tends to infinity and (P3) is satisfied.
Case III: $v$ contains 0 .
Again, we can freely assume that $u$ contains the letter 1 . Since $v$ contains the letter 0 , the morphism $\varphi$ is primitive: if $v$ contains both 0 and 1 , it follows by definition; if $v$ only contains the letter 0 , then $\varphi^{2}(1)$ will contain $\varphi(0)$ which contains both the letters 0 and 1 . Since $\varphi$ is primitive, $\lim _{n \rightarrow \infty}\left|\varphi^{n}(0)\right|=\lim _{n \rightarrow \infty}\left|\varphi^{n}(1)\right|=\infty$. Under the assumption that these words contain overlap (otherwise ( $\mathbf{P 1}$ ) applies), it follows that ( $\mathbf{P 3}$ ) is satisfied.

The final result needed to prove Theorem 2.11 is due to Badziahin and Zorin, which shows that the real number that has the Thue-Morse word (or its complement) as its base- $n$ expansion is well-approximable provided that $n$ is not divisible by 15 . Note that if $n$ is divisible by 15 the result is unknown, as opposed to being false.

Theorem 4.3 ([5]). Let $M_{n}$ be the real number whose base-n expansion is the ThueMorse word. If $n$ is not divisible by 15 , then $M_{n}$ is well-approximable.

Combining together Section 3.2, Lemma 4.2, and Theorem 4.3 provides the proof for Theorem 2.11.

Proof of Theorem 2.11. From Theorem 4.3, $M_{2}$ is well-approximable and therefore, satisfies 2LC. In this case, $\widetilde{M}_{2}$ is given by $1-M_{2}$. Since $M_{2}$ is well-approximable, so is $\widetilde{M}_{2}$. Therefore, the real numbers whose base-2 expansion satisfy (P1) satisfy 2LC. For words satisfying (P2), we note that for any periodic word $v$, i.e., $v=X^{\omega}$, the real number $v_{2}$ is rational and therefore, well-approximable. Applying Proposition 2.5 shows that the real numbers whose base-2 expansions satisfy (P2), also satisfy 2 LC . Finally, Proposition 2.8 implies that for any base-2 expansion which satisfies (P3), the corresponding real number satisfies 2LC.

### 4.1.1. Proof of Corollary 2.12

From Corollary 2.10, we can extend Theorem 2.11 to Corollary 2.12, by showing that for any morphism $\psi:\{a, b\} \rightarrow\{a, b\}^{*}$ which is prolongable on $a$ with $a, b \in \mathcal{A}$ and any coding $\tau: \mathcal{A} \rightarrow\{0,1, \ldots, p-1\}$, the real number $\tau\left(\psi^{\omega}(a)\right)_{p}$ satisfies pLC. Note that since $a$ and $b$ are arbitrary letters, we can consider them to be letters in $\{0,1, \ldots, p-1\}$ and forget the coding. By the same argument, the word $\psi^{\omega}(a)$ can be rewritten as a coding of a pure morphic word $w$ over the alphabet $\{0,1\}$, i.e., $\psi^{\omega}(a)=\sigma(w)$ where $\sigma(0)=a$ and $\sigma(1)=b$. If $w$ satisfies (P2) or (P3), then $\psi^{\omega}(a)_{p}$ satisfies pLC using
the same arguments as in the proof of Theorem 2.11. When $\psi^{\omega}(a)$ is a coding of the Thue-Morse word or its complement, the situation is a bit more complicated.

Let $T M(a, b)$ be the coding of $M$, where 0 is mapped to $a$ and 1 is mapped to $b$. In order to complete the proof of Corollary 2.12 , we will show that $T M(a, b)_{p}$ is well-approximable for all primes $p$ and all $a, b \in\{0,1, \ldots, p-1\}$.

Proposition 4.4. Let $a, b \in\{0,1, \ldots, n-1\}$. If $n$ is not divisible by 15 , then $T M(a, b)_{n}$ is well-approximable.

Proof. We start this proof by noting that if a real number $x$ is well-approximable, then adding a rational number $p / q$ or multiplying by a rational constant will preserve this property. In particular, $T M(0,1)_{n}$ is well-approximable if and only if $r \cdot T M(0,1)_{n}$ is well-approximable for all $r \in \mathbb{Q}$. If we restrict $r$ to $\{0,1, \ldots, n-1\}$, then the base- $n$ expansion of $T M(0, r)_{n}$ is

$$
r \cdot T M(0,1)_{n}=r \cdot \sum_{i=1}^{\infty} \frac{\sigma(i)}{n^{i}}=\sum_{i=1}^{\infty} \frac{r \cdot \sigma(i)}{n^{i}}=T M(0, r)_{n},
$$

where $\sigma(i)$ returns the $i$ th letter in the Thue-Morse word.
Similarly, $T M(0, n-1)_{n}$ is well-approximable if and only if $T M(n-1,0)_{n}$ is well-approximable. This follows from the following observation:

$$
\begin{aligned}
1-T M(0, n-1)_{n} & =1-\sum_{i=1}^{\infty} \frac{(n-1) \cdot \sigma(i)}{n^{i}} \\
& =\sum_{i=1}^{\infty} \frac{(n-1) \cdot(1-\sigma(i))}{n^{i}}=T M(n-1,0)_{n} .
\end{aligned}
$$

Multiplying by $k /(n-1)$ shows that the number $T M(k, 0)_{n}$ is well-approximable for all $k \in\{0,1, \ldots, n-1\}$ if and only if $T M(n-1,0)_{n}$ is well-approximable.

Furthermore, we note that for $\ell \in\{0,1, \ldots, n-1\}$, the real number whose base$n$ expansion is an infinite string of $\ell$ 's corresponds to the rational number $\ell /(n-1)$. Therefore, if $\ell \leq n-1-k$, then

$$
T M(0, k)_{n}+\frac{\ell}{n-1}=\sum_{i=1}^{\infty} \frac{k \cdot \sigma(i)}{n^{i}}+\sum_{i=1}^{\infty} \frac{\ell}{n^{i}}=\sum_{i=1}^{\infty} \frac{k \cdot \sigma(i)+\ell}{n^{i}}=T M(\ell, \ell+k)_{n}
$$

Likewise, $T M(k, 0)_{n}+\ell /(n-1)=T M(k+\ell, \ell)_{n}$. This, combined with the previous arguments, shows that for all $a, b \in\{0,1, \ldots, n-1\}$ the real number $T M(a, b)_{n}$ is well-approximable if and only if $\operatorname{TM}(0,1)_{n}$ is well-approximable. Applying Theorem 4.3 completes the proof.

### 4.2. Proof of Theorem 2.13

Let $\mu$ be the Thue-Morse morphism. In order to prove Theorem 2.13, we will use Section 3.2, Theorem 4.3, and the following two lemmas.

Lemma 4.5. For every overlap-free word $x \in\{0,1\}^{*}$, there exist words $u, v, y \in\{0,1\}^{*}$ with $|u|,|v| \leq 2$ and $x=u \mu(y) v$.

Lemma 4.6. Let $y \in\{0,1\}^{*}$. Then $y$ is overlap-free if and only if $\mu(y)$ is overlap-free.
For Lemma 4.5, see [11, Theorem 6.4] or [1, Lemma 3]. For Lemma 4.6, see [2, Lemma 1.7.4].

Proof of Theorem 2.13. In order to prove this result, we will show that every overlapfree base-2 expansion of length $K$ contains a prefix of $M$ or $\widetilde{M}$ of length $p(K)$, where $\lim _{K \rightarrow \infty} p(K)=\infty$. The result then follows from Proposition 2.5 and Theorem 4.3.

Let $x$ be an overlap-free word of length $K$. By Lemma 4.5, there exist words $u_{1}, v_{1}, y_{1} \in\{0,1\}^{*}$ with $\left|u_{1}\right|,\left|v_{1}\right| \leq 2$ and $x=u_{1} \mu\left(y_{1}\right) v_{1}$. Using this construction, we can conclude that $\left|\mu\left(y_{1}\right)\right|=K-|u|-|v| \geq K-4$. Furthermore, since $\mu$ is 2-uniform, i.e., $|\mu(0)|=|\mu(1)|=2$, the length of $y_{1}$ is equal to $\left|\mu\left(y_{1}\right)\right| / 2$. Provided that $K-4 \geq 1$, we also have that $y_{1}$ is not the empty word.

Since $x$ is overlap-free, it follows that $\mu\left(y_{1}\right)$ is overlap-free. Additionally, Lemma 4.6 implies that $y_{1}$ is overlap-free. As a result, there exist $u_{2}, v_{2}, y_{2} \in\{0,1\}^{*}$ with $\left|u_{2}\right|,\left|v_{2}\right| \leq$ 2 such that $y_{1}=u_{2} \mu\left(y_{2}\right) v_{2}$. Then $x$ can be rewritten as

$$
x=u_{1} \mu\left(y_{1}\right) v_{1}=u_{1} \mu\left(u_{2}\right) \mu^{2}\left(y_{2}\right) \mu\left(v_{2}\right) v_{1} .
$$

The length of $y_{2}$ is bounded as follows:

$$
\frac{\left|y_{1}\right|-4}{2} \leq\left|y_{2}\right| \leq \frac{\left|y_{1}\right|}{2} .
$$

More generally, for any $k \in \mathbb{N}$, the subword $y_{k}$ can be rewritten as

$$
y_{k}=u_{k+1} \mu\left(y_{k+1}\right) v_{k+1},
$$

where $u_{k+1}, v_{k+1}, y_{k+1} \in\{0,1\}^{*},\left|u_{k+1}\right|,\left|v_{k+1}\right| \leq 2$ and

$$
\frac{\left|y_{k}\right|-4}{2} \leq\left|y_{k+1}\right| \leq \frac{\left|y_{k}\right|}{2} .
$$

Note that $u_{k+1}, v_{k+1}$ and $y_{k+1}$ can all be the empty word.
Using this substitution, $x$ can be rewritten in terms of $y_{k}$ as

$$
x=u_{1} \mu\left(u_{2}\right) \cdots \mu^{k-1}\left(u_{k}\right) \mu^{k}\left(y_{k}\right) \mu^{k-1}\left(v_{k}\right) \cdots \mu\left(v_{2}\right) v_{1},
$$

where the length of $y_{k}$ is bounded below:

$$
\begin{equation*}
\left|y_{k}\right| \geq \frac{K-4 \cdot\left(2^{k}-1\right)}{2^{k}} \tag{8}
\end{equation*}
$$

From (8), the word $y_{k}$ is non-empty provided that $K-4 \cdot\left(2^{k}-1\right)>0$. By rearranging, the largest value of $k$ that guarantees that $y_{k}$ is non-empty is $k=\left\lfloor\log _{2}(K+4)\right\rfloor-2$. For such a value of $k$, let $a$ be any subword of $y_{k}$ of length 1 . Then $\mu^{k}(a)$ is a prefix of either $M$ or $\widetilde{M}$. Since $\mu$ is 2-uniform and $|a|=1$, it follows that the length of this prefix is

$$
\left|\mu^{k}(a)\right| \geq 2^{k}
$$

$$
\begin{aligned}
& \geq 2^{\left(\log _{2}(K+4)-3\right)} \\
& =\frac{K+4}{8}
\end{aligned}
$$

Since $\lim _{K \rightarrow \infty}(K+4) / 8=\infty$, the result follows.

## CRediT authorship contribution statement

J. Blackman: Investigation, Writing - original draft, Writing - review \& editing. S.

Kristensen: Investigation, Writing - original draft, Writing - review \& editing. M.J. Northey: Investigation, Writing - original draft, Writing - review \& editing.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Data availability

No data was used for the research described in the article.

## Acknowledgement

We thank the referee for a thorough reading and helpful comments.

## Funding

Dr Blackman was supported by the Engineering and Physical Sciences Research Council (EPSRC), UK [grant no. EP/W006863/1] awarded through the University of Liverpool. Prof. Kristensen's research was supported by the Independent Research Fund Denmark (Grant ref. 1026-00081B) and Aarhus University Research Foundation, Denmark (Grant ref. AUFF-E-2021-9-20). Dr Northey was supported by the Engineering and Physical Sciences Research Council, UK [grant no. EP/RF060349] awarded through Durham University.

## References

[1] J.-P. Allouche, J. Currie, J. Shallit, Extremal infinite overlap-free binary words, Electron. J. Combin. 5 (R27) (1998).
[2] J.-P. Allouche, J. Shallit, Automatic Sequences: Theory, Applications, Generalisations, Cambridge University Press, 2003.
[3] D. Badziahin, On continued fraction expansion of potential counterexamples to p-adic Littlewood conjecture, 2015.
[4] D. Badziahin, Y. Bugeaud, M. Einsiedler, D. Kleinbock, On the complexity of a putative counterexample to the $p$-adic Littlewood conjecture, Compos. Math. 151 (9) (2015) 1647-1662.
[5] D. Badziahin, E. Zorin, Thue-Morse constant is not badly approximable, Int. Math. Res. Not. IMRN 2015 (19) (2015) 9618-9637.
[6] Y. Bugeaud, M. Drmota, B. de Mathan, On a mixed Littlewood conjecture in diophantine approximation, Acta Arith. 128 (2007) 107-124.
[7] E.B. Burger, Exploring the Number Jungle: A Journey Into Diophantine Analysis, American Mathematical Society, 2000.
[8] B. de Mathan, O. Teulié, Problèmes diophantiens simultanés, Monatsh. Math. 143 (15) (2004) 229-245.
[9] F. Durand, Decidability of uniform recurrence of morphic sequences, Internat. J. Found. Comput. Sci. (2013).
[10] M. Einsiedler, D. Kleinbock, Measure rigidity and p-adic Littlewood-type problems, Compos. Math. 143 (3) (2007) 689-702.
[11] Y. Kobayashi, Repetition-free words, Theoret. Comput. Sci. 44 (1986) 175-197.
[12] P. Liardet, P. Stambul, Algebraic computations with continued fractions, J. Number Theory 73 (1) (1998) 92-121.
[13] M. Lothaire, Algebraic Combinatorics on Words, Cambridge University Press, 2002.
[14] G. Raney, On continued fractions and finite automata, Math. Ann. 206 (4) (1973) 265-283.
[15] P. Séébold, Sequences generated by infinitely iterated morphisms, Discrete Appl. Math. 11 (3) (1985) 255-264.


[^0]:    * Corresponding author.

    E-mail addresses: John.Blackman@liverpool.ac.uk (J. Blackman), sik@math.au.dk (S. Kristensen), m.j.northey @durham.ac.uk (M.J. Northey).
    https://doi.org/10.1016/j.exmath.2024.125548
    0723-0869/© 2024 The Author(s). Published by Elsevier GmbH. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

