# Disposability in Square-Free Words 

Tero Harju<br>Department of Mathematics and Statistics, University of Turku, Finland


#### Abstract

We consider words $w$ over the alphabet $\Sigma=\{0,1,2\}$. It is shown that there are irreducibly square-free words of all lengths $n$ except $4,5,7$ and 12 . Such a word is square-free (i.e., it has no repetitions $u u$ as factors), but by removing any one internal letter creates a square in the word.


Keywords: Square-free ternary words, irreducibly square-free, Thue word 2000 MSC: 68R15

## 1. Introduction

Grytczuk et al. [1] showed that there are infinitely many 'extremal' squarefree ternary words where one cannot augment a single new letter anywhere without creating a square; see also Mol and Rampersad [2] for further results. In this article we consider the dual problem of this and show that there are square-free ternary words of all lengths, except $4,5,7$ and 12 , where removing any single interior letter creates a square. Although the problems resemble each other, the results and the proof techniques are quite different.

Let $\Sigma=\{0,1,2\}$ be a fixed ternary alphabet and denote by $\Sigma^{*}$ and $\Sigma^{\omega}$ the sets of all finite and infinite length words over $\Sigma$, respectively. A finite word $u$ is called a factor of a word $w \in \Sigma^{*} \cup \Sigma^{\omega}$ if $w=w_{1} u w_{2}$ for some, possibly empty, words $w_{1}$ and $w_{2}$. Moreover, $w$ is square-free if it does not have a nonempty factor of the form $u u$.

Let $w \in \Sigma^{*}$ be a square-free word with a factorization $w=w_{1} a w_{2}$ where $a \in \Sigma$. We say that the occurrence of the letter $a$ is disposable if $w_{1} w_{2}$ is squarefree. The definition extends naturally to infinite words. An occurrence of a letter $a$ is interior, if $w_{1}$ and $w_{2}$ are both nonempty.

Email address: harju@utu.fi (Tero Harju)

If a square-free word $w \in \Sigma^{*} \cup \Sigma^{\omega}$ does not have disposable occurrences of interior letters then $w$ is said to be irreducibly square-free, i.e., by deleting any interior occurrence of a letter results in a square in the remaining word.

The nonemptiness condition on the prefixes and suffixes is required since all prefixes and suffixes of square-free words are disposable.
Remark 1. The words of length at most two have no internal letters, and therefore we consider the property of being irreducibly square-free only for words of length at least three.
Example 1. Let $\tau: \Sigma^{*} \rightarrow \Sigma^{*}$ be the morphism determined by

$$
\tau(0)=012, \quad \tau(1)=02, \quad \tau(2)=1
$$

The Thue word $\mathbf{t}$ is the fixed point $\mathbf{t}=\tau^{\omega}(0)$ of $\tau$ obtained by iterating $\tau$ on the start word 0 . Then $\mathbf{t}$ is an infinite square-free word; see, e.g., Lothaire [3]:

$$
\mathbf{t}=012021012102012021020121012021012 \cdots
$$

We show that the Thue word is not irreducibly square-free. For this, we first notice that $\mathbf{t}$ avoids 010 and 212 as factors. Also, it avoids 1021, since this word would have to be a factor of $\tau(212)$. Deleting the letter 2 at the third position results in a square-free word $01021012102012 \cdots$. Indeed, a potential square would have to start either from the beginning, but the prefix 010 does not occur in $\mathbf{t}$, or from the second position, but 1021 does not occur in $\mathbf{t}$.

Later checking of irreducibility of (infinite) words is based on the following procedure that depends on a morphism $\alpha: \Sigma^{*} \rightarrow \Sigma^{*}$ for which $|\alpha(a)|>1$ for all letters $a$.

## Procedure I.

1 Check that the morphism $\alpha$ generates an infinite square-free word; say, $\alpha^{\omega}(0)$ or $\alpha(w)$, where $w$ is a given infinite square-free word.

2 For any pair $(a, b)$ of different letters, check that $\alpha(a b)$ is irreducibly square-free. This takes care that the last letter of $\alpha(a)$ and the first letter of $\alpha(b)$ are not disposable in $\alpha(a b)$. This guarantees that these occurrences are not disposable in any $\alpha(w)$ where $w=w_{1} a b w_{2}$ is square-free.
The first item of Procedure I is often taken care of by Crochemore's criterion [4]:
Theorem 2. A morphism $\alpha: \Sigma^{*} \rightarrow \Sigma^{*}$ preserves square-free words if and only if it preserves square-freeness of words of length five.

## 2. Irreducibly square-free words of almost all lengths

By a systematic search we find that there are no irreducibly square-free words of lengths $4,5,7$ and 12 . In the following table we have counted the irreducibly square-free words of lengths $3, \ldots, 30$ up to isomorphism (produced by permutations of the letters) and reversal (mirror image) of the words. For instance, 010212010 is the only irreducibly square-free word of length nine up to isomorphism and reversal. It is a palindrome. The table suggests that the irreducibly square-free words are quite rare among the square-free words, e.g., there are (up to isomorphism and reversal) 202 square-free words of length 20 , but only 12 of those are irreducibly square-free. Counting the numbers of (irreducibly) square-free words must take into consideration those words that are palindromes or isomorphic to their reversals.

| length | card |  |  |  |  |  |  |  |  |  |  |  |  |
| ---: | ---: | ---: | ---: | ---: | ---: | ---: | ---: | ---: | ---: | ---: | ---: | ---: | ---: |
| 3 | 1 | 4 | 0 | 5 | 0 | 6 | 1 | 7 | 0 | 8 | 1 | 9 | 1 |
| 10 | 1 | 11 | 3 | 12 | 0 | 13 | 3 | 14 | 4 | 15 | 4 | 16 | 7 |
| 17 | 9 | 18 | 7 | 19 | 12 | 20 | 12 | 21 | 16 | 22 | 18 | 23 | 23 |
| 24 | 24 | 25 | 34 | 26 | 36 | 27 | 48 | 28 | 55 | 29 | 69 | 30 | 78 |

Table 1: The number of irreducibly square-free words of lengths from 3 to 30 up to isomorphism and reversal.

Theorem 3. There exists an infinite irreducibly square-free word.
Proof. Let $\varphi$ be the following uniform palindromic morphism of length 17, i.e., $\varphi(1)=\pi(\varphi(0))$ and $\varphi(2)=\pi^{2}(\varphi(0))$ for the permutation $\pi=\left(\begin{array}{ll}0 & 1\end{array}\right)$ of the letters:

$$
\begin{aligned}
& \varphi(0)=01202120102120210 \\
& \varphi(1)=12010201210201021 \\
& \varphi(2)=20121012021012102
\end{aligned}
$$

By Theorem 2, $\varphi$ preserves square-freeness. It is easy to check that $\varphi(0)$, and so also the isomorphic copies $\varphi$ (1) and $\varphi(2)$, are irreducibly square-free. Finally, Procedure I entails that deleting the 'middle' 17th letters 0 of $\varphi(01)$ and of $\varphi(02)$ gives squares: 11 and 02120212 , respectively. Similarly, deleting the 18th letter of $\varphi(01)$ and of $\varphi(02)$ gives squares: 10201020 and 00 , respectively.

These observations suffice for the proof of the theorem, since now $\varphi(w)$ is irreducibly square-free for all square-free, finite or infinite, words $w$.

Remark 4. The morphism $\varphi$ has an alignment property, i.e., for all letters $a, b, c$ if $\varphi(b c)=u \varphi(a) v$ then $u$ or $v$ is empty, and $a=b$ or $a=c$, respectively.

The morphism $\varphi$ has an infinite fixed point

$$
\Phi=\varphi^{\omega}(0)
$$

that is the limit of the sequence $\varphi(0), \varphi^{2}(0), \ldots$.
Note that the finite prefixes of $\Phi$ are not always irreducibly square-free. For instance, none of the prefixes of $\Phi$ of length $n$ with $19 \leq n \leq 29$ are irreducibly square-free. However, we do have the following result with the help of $\varphi$.

Theorem 5. There are irreducibly square-free words of all lengths $n$ except 4,5,7 and 12 .

Proof. Table 2 gives an example for the cases $n \leq 17$.

| 3 | 010 | 13 | 0102012101202 |
| :---: | :--- | :---: | :--- |
| 6 | 010212 | 14 | 01020120212010 |
| 8 | 01020121 | 15 | 010201210120212 |
| 9 | 010212010 | 16 | 0102012021201020 |
| 10 | 0102012101 | 17 | 01202120102120210 |
| 11 | 01020120212 |  |  |

Table 2: Small irreducibly square-free words. There are no examples for the lengths 4,5,7 and 12.

For $n \geq 18$, we rely on the morphism $\varphi$ in order to have solutions for the lengths $n \equiv p(\bmod 17)$ for $p=0,1, \ldots, 16$.

Claim A. Let $w$ be a nonempty suffix of $\varphi$ (1) or $\varphi$ (2) of length $|w|<17$. Then the word $w \Phi$ is square-free (but not necessarily irreducibly square-free).

The word $w$ is a suffix of exactly one of the words $\varphi(a), a \in \Sigma$. Suppose there is a square in $w \Phi$ and assume that $w$ is of minimal length with this property. Then $w \Phi$ has a prefix $u u$ for $u=w \varphi(x) z$ for some words $x$ and $z$ with $|z|<17$ (when $|x|$ is chosen to be maximal). Hence $u u=w \varphi(x) z w \varphi(x) z$, and so $z w=\varphi(a)$ for $a \in \Sigma$. Therefore $z$ is nonempty. By the alignment property,
$u u$ must be followed in $w \Phi$ by the word $w$. This delivers a square in $\Phi$, namely $\varphi(x) z w \varphi(x) z w=\varphi(x a x a)$; a contradiction since $\Phi$ is square-free. This proves Claim A.

Clearly, there are irreducibly square-free words of lengths $n \equiv 0(\bmod 17)$, since we can take a prefix of $\Phi$ of length $n / 17$ and apply $\varphi$ to it. Next we extend $\Phi$ to the left by considering words of the form $u w$, where $w$ is a prefix of $\Phi$.
Claim B. The words $121 \Phi$ and $0102 \Phi$ are square-free.
First, the words $121 \varphi(0)$ and $0102 \varphi(0)$ are not factors of $\Phi$, since $\varphi$ has the alignment property and the given words are not suffixes of any $\varphi(a), a \in \Sigma$. Therefore, if $121 \Phi$ contains a square, then the square must be a prefix $21 v$ of $21 \Phi$ for some $v$ (and $1 \Phi$ is square-free by Claim A).

Assume that $21 v=21 u 21 u$ where $v=u 21 u$ is a prefix of $\Phi$. Now, $|v|>$ $|\varphi(a)|=17$, since 21 is not followed by the first letter of $u$ in any $\varphi(a)$, i.e., $u 21=\varphi(z) \varphi(1)$ for some $z$, since only $\varphi(1)$ ends in 21 . We have then that $\varphi(1)=y 21$ and $u=\varphi(z) y$. This means that the square $21 v=21 \varphi(z 1 z) y$ is necessarily continued by the rest of $\varphi(1)$, i.e., by 21 , giving a prefix $v 21=$ $u 21 u 21$ of $\Phi$; a contradiction, since $\Phi$ is square-free.

In the case of $0102 \Phi$, Claim A guarantees that $102 \Phi$ is square-free. For the full prefix 0102, the claim follows since the prefix 01020120 of $0102 \Phi$ does not occur in $\Phi$. This proves Claim B.

The special words $w_{i}$ of Table 3 are chosen such that
(iii) $\left|w_{i}\right|=i$,
(iv) $w_{i} \Phi$ is square-free. By Claim A, this follows for $i=1,2,4,5$ and 10 . By Claim B, the claim follows for the other cases.
(v) $w_{i} \varphi(0)$ is irreducibly square-free (by a simple computer check).

The words $w_{i}$, themselves, are not (and, indeed, cannot be) all irreducibly square-free, but they are square-free.

Finally, let $n=17 k+i$. By Table 2, we can assume that $n \geq 18$. We then choose $w_{i}$ from Table 3, and pick a prefix $\varphi(v)$ of $\Phi$ of length $17 k$. This creates an irreducibly square-free word $w_{i} \varphi(v)$ of length $n$.

## 3. Problems on Longer Words to Dispose

The property of being irreducibly square-free can be generalized to longer factors than just letters. Let $w \in \Sigma^{*}$ be a square-free word with a factorization

$$
\begin{array}{ll}
w_{1}=1 & w_{9}=121020121 \\
w_{2}=02 & w_{10}=2021012102 \\
w_{3}=121 & w_{11}=10121020121 \\
w_{4}=2102 & w_{12}=101202120121 \\
w_{5}=12102 & w_{13}=0210121020121 \\
w_{6}=020121 & w_{14}=01021201020121 \\
w_{7}=2120102 & w_{15}=010201202120121 \\
w_{8}=01020121 & w_{16}=0201021201020121
\end{array}
$$

Table 3: The special words with $w_{i} \equiv i(\bmod 17)$. The words $w_{i}$ with $i>1$ that end in 121 or 0102 as called for by Claim B.
$w=w_{1} v w_{2}$ such that both $w_{1}$ and $w_{2}$ are nonempty. We say that the (occurrence of the) factor $v$ is disposable if also $w_{1} w_{2}$ is square-free. If a finite or infinite square-free word $w$ does not have disposable factors of length $k$ then $w$ is called $k$-irreducibly square-free.

Example 2. We show that $\tau^{2 n}(0)$ is not 2 -irreducibly square-free, for all $n \geq 2$. Indeed,

$$
\begin{aligned}
& \tau^{2}(0)=012021 \\
& \tau^{2}(1)=0121 \\
& \tau^{2}(2)=02
\end{aligned}
$$

Now, for $n \geq 2$, the word $\tau^{2 n}(0)$ has the suffix 121 since $\tau^{2}(1)$ has this suffix and $\tau^{2}(0)$ ends with the letter 1 . But a 2 -irreducibly square-free word cannot be of the form $w 121$, since by removing the pair 12 , we obtain a (square-free) prefix of $w 1$ of $w 121$.

However, the limit $\mathbf{t}=\tau^{\omega}(0)$ is 2 -irreducibly square-free. To see this, we consider the 5-th powers of the morphism $\tau$ :

$$
\begin{aligned}
& \tau^{5}(0)=012021012102012021020121012021012102012101202102 \\
& \tau^{5}(1)=01202101210201202102012101202102 \\
& \tau^{5}(2)=0120210121020121
\end{aligned}
$$

where the lengths of the images are 48,32 and 16, respectively. These images have a common prefix $p=012021$ (and even longer ones). A computer check shows that the words $\tau^{5}(a) p$ are 2 -irreducibly square-free for $a=1,2$.

Moreover, deleting an internal occurrence of a pair $c d$ from $\tau^{5}(0) p$ results in a square-free word only for $c d=20$ and $c d=02$ that lie inside $p$. This proves that the infinite word $\mathbf{t}$ is 2 -irreducibly square-free.

These considerations raise many problems.
Problem 1. Given $k \geq 1$, does there exist an infinite ternary word that is $k$ irreducibly square-free?

Theorem 6. Every infinite ternary square-free word $w$ does have an infinite number of integers $k$ for which $w$ is not $k$-irreducibly square-free.

Proof. We need only to consider repetitions of the first letter of $w$, say $w=$ auaw $w_{0}$. Deleting the factor ua gives $a w_{0}$, a (square-free) suffix of $w$.

We have seen that Problem 1 has a positive solution for $k=1$ and $k=2$. For small values of $k$ a solution may be found using square-free morphisms. E.g., the morphism

$$
\begin{aligned}
& \alpha_{3}(0)=0121012 \\
& \alpha_{3}(1)=01020120212 \\
& \alpha_{3}(2)=0102101210212
\end{aligned}
$$

generates a 3 -irreducibly square-free word $\alpha^{\omega}(0)$. This follows from the fact that $\alpha_{3}(a b)$ is 3-irreducibly square-free for all different letters $a$ and $b$.

Problem 2. Does there exist, for every $k$, a bound $N(k)$ such that there exist $k$-irreducibly square-free words of all lengths $n \geq N(k)$ ?

Finally, we state a problem of the opposite nature:
Problem 3. Does there exist an infinite square-free word $w$ such that $w$ is $k$ irreducibly square-free for no $k \geq 1$ ?

Acknowledgements. I would like to thank the referees of this journal for their comments that improved the presentation of this paper.

## References

[1] J. Grytczuk, H. Kordulewski, A. Niewiadomski, Extremal square-free words, ArXiv: 1920.06226 (2019).
[2] L. Mol, N. Rampersad, Lengths of extremal square-free ternary words, Arxiv: 2001.11763 (2020).
[3] M. Lothaire, Combinatorics on words, Cambridge Mathematical Library, Cambridge University Press, Cambridge, 1997. doi:10.1017/ CB09780511566097.
[4] M. Crochemore, Sharp characterizations of squarefree morphisms, Theoret. Comput. Sci. 18 (2) (1982) 221-226. doi:10.1016/0304-3975(82) 90023-8.

