Improved bounds on the number of ternary square-free words

Journal Item

IMPROVED BOUNDS ON THE NUMBER OF TERNARY SQUARE-FREE WORDS

UWE GRIMM

ABSTRACT. Improved upper and lower bounds on the number of square-free ternary words are obtained. The upper bound is based on the enumeration of square-free ternary words up to length 110. The lower bound is derived by constructing generalised Brinkhuis triples. The problem of finding such triples can essentially be reduced to a combinatorial problem, which can efficiently be treated by computer. In particular, it is shown that the number of square-free ternary words of length n grows at least as $65^{n/40}$, replacing the previous best lower bound of $2^{n/17}$.

1. INTRODUCTION

A word w is a string of letters from a certain alphabet Σ, the number of letters of w is called the length of the word. The set of words of length n is $\mathcal{L}(n) = \Sigma^n$, and the union

$$\mathcal{L} = \bigcup_{n\geq 0} \mathcal{L}(n) = \Sigma^{\mathbb{N}_0} \tag{1}$$

is called the language of words in the alphabet Σ. This is a monoid with concatenation of words as operation and the empty word λ, which has zero length, as neutral element [11]. For a word w, we denote by $\bar{w}$ the corresponding reversed word, i.e., the word obtained by reading w from back to front. A palindrome is a word w that is symmetric, $w = \bar{w}$.

Square-free words [1–13] are words w that do not contain a "square" yy of a word y as a subword (factor). In other words, w can only be written in the form $xyyz$, with words x, y and z, if $y = \lambda$ is the empty word. In a two-letter alphabet $\{0, 1\}$, the complete list of square-free words is $\{\lambda, 0, 1, 01, 10, 010, 101\}$. However, in a three-letter alphabet $\Sigma = \{0, 1, 2\}$, square-free words of arbitrary length exist, and the number of square-free words of a given length n grows exponentially with n [4, 3, 7].

We denote the set of square-free words of length n in the alphabet $\Sigma = \{0, 1, 2\}$ by $\mathcal{A}(n) \subset \mathcal{L}(n)$. The language of ternary square-free words is

$$\mathcal{A} = \bigcup_{n\geq 0} \mathcal{A}(n) \subset \Sigma^{\mathbb{N}_0}. \tag{2}$$

We are interested in the number of square-free words of length n

$$a(n) = |\mathcal{A}(n)| \tag{3}$$

and in estimating the growth of $a(n)$ with the length n. For $n = 0, 1, 2, 3$, the sets of ternary square-free words are

$$\mathcal{A}(0) = \{\lambda\}, \tag{4}$$
$$\mathcal{A}(1) = \{0, 1, 2\}, \tag{5}$$
$$\mathcal{A}(2) = \{01, 02, 10, 12, 20, 21\}, \tag{6}$$
$$\mathcal{A}(3) = \{010, 012, 020, 021, 101, 102, 120, 121, 201, 202, 210, 212\}, \tag{7}$$

where λ denotes the empty word. Hence $a(0) = 1$, $a(1) = 3$, $a(2) = 6$, $a(3) = 12$, and so on, see [1] where the values of $a(n)$ for $n \leq 90$ are tabulated. In [15], the sequence is listed as A006156 (formerly M2550).

2. Upper bounds obtained by enumeration

Obviously, a word w of length $m+n$, obtained by concatenation of words w_1 of length m and w_2 of length n, can only be square-free if w_1 and w_2 are square-free. This necessary, but not sufficient, condition implies the inequality

$$a(m+n) \leq a(m)a(n) \tag{8}$$

for all $m, n \geq 0$. By standard arguments, see also [1], this guarantees the existence of the limit

$$s := \lim_{n\to\infty} a(n)^{\frac{1}{n}}, \tag{9}$$

the growth rate or "connective constant" of ternary square-free words [8]. The precise value of s is not known, but lower [4, 3, 7] and upper bounds [1] have been established. It is the purpose of this paper to improve both the lower and the upper bounds.

It is relatively easy to derive reasonable upper bounds from the inequality (8). In fact [1], one can slightly improve on (8) by considering two words w_1 and w_2 of length $m \geq 2$ and $n \geq 2$, such that the last two letters of w_1 are equal to the first two letters of w_2, and we join them to a word w of length $m+n-2$ by having the two words overlap on these two letters. This yields

$$a(m+n-2) \leq \frac{1}{6}a(m)a(n), \tag{10}$$

for all $m, n \geq 2$, because there are precisely $a(n)/6$ square-free letters of length $n \geq 2$ that start with the last two letters of w_1. Taking n fixed, one obtains

$$s^{n-2} = \lim_{m\to\infty} \frac{a(m+n-2)}{a(m)} \leq \frac{a(n)}{6} \tag{11}$$

and hence the upper bound

$$s \leq \left(\frac{a(n)}{6}\right)^{\frac{1}{n-2}} \tag{12}$$

TABLE 1. The number of ternary square-free words $a(n)$ of length n for $91 \leq n \leq 110$.

n	$a(n)$	n	$a(n)$
91	336 655 224 582	101	4 704 369 434 772
92	438 245 025 942	102	6 123 969 129 810
93	570 491 023 872	103	7 971 950 000 520
94	742 643 501 460	104	10 377 579 748 374
95	966 745 068 408	105	13 509 138 183 162
96	1 258 471 821 174	106	17 585 681 474 148
97	1 638 231 187 596	107	22 892 370 891 330
98	2 132 586 986 466	108	29 800 413 809 730
99	2 776 120 525 176	109	38 793 041 799 498
100	3 613 847 436 684	110	50 499 301 907 904

for any $n \geq 3$. This bound can be systematically improved by calculating $a(n)$ for as large values of n as possible. The bound given in [1], from $a(90) = 258\,615\,015\,792$, is

$$s \leq 43\,102\,502\,632^{\frac{1}{88}} = 1.320\,829\,\ldots \tag{13}$$

The results given in table 1 extend the previously known values of $a(n)$ [1] to lengths $n \leq 110$. They were obtained by a simple algorithm, extending square-free words letter by letter and checking that the new letter does not lead to the formation of any square. The value $a(110)$ yields an improved upper bound of

$$s \leq 8\,416\,550\,317\,984^{\frac{1}{108}} = 1.317\,277\,\ldots \tag{14}$$

3. Brinkhuis triples and lower bounds

While the upper bound is already relatively close to the actual value of s, which was estimated in reference [1] to be about 1.301 76 on the basis of the first 90 values, it is much more difficult to obtain any reasonable lower bound for s. In order to derive a lower bound, one has to show that $a(n)$ grows exponentially in n with optimal growth bound. This can be achieved by demonstrating that each square-free word of length n gives rise to sufficiently many different square-free words of some length $m > n$. This was first done by Brinkhuis [4], by constructing what is now known as a Brinkhuis triple or a Brinkhuis triple pair.

Definition 1. An n-Brinkhuis triple pair is a set $\mathcal{B} = \{\mathcal{B}^{(0)}, \mathcal{B}^{(1)}, \mathcal{B}^{(2)}\}$ of three pairs $\mathcal{B}^{(i)} = \{U^{(i)}, V^{(i)}\} \subset \mathcal{A}(n)$, $i \in \{0,1,2\}$, of pairwise different square-free words such that the set of 96 words of length $3n$

$$\bigcup_{w_1 w_2 w_3 \in \mathcal{A}(3)} \Big\{ W_1 W_2 W_3 \mid W_j \in \mathcal{B}^{(w_j)}, j = 1,2,3 \Big\} \subset \mathcal{A}(3n).$$

In other words, it is required that all $3n$-letter images of the twelve elements of $\mathcal{A}(3)$ under any combination of the eight maps

$$\varrho_{x,y,z} : \begin{cases} 0 \to x \in \mathcal{B}^{(0)} \\ 1 \to y \in \mathcal{B}^{(1)} \\ 2 \to z \in \mathcal{B}^{(2)} \end{cases} \tag{15}$$

are square-free. This property is sufficient to ensure that images of any square-free word in the alphabet Σ under any combination of the eight maps to each of its letters is again square-free. This can be shown as follows.

Consider the six-letter alphabet $\tilde{\Sigma} = \{0, 0', 1, 1', 2, 2'\}$ and a language $\tilde{\mathcal{A}}$ consisting of all words of $\mathcal{A}$ with an arbitrary number of letters replaced by their primed companions. In other words,

$$\tilde{\mathcal{A}} = \bigcup_{n\geq 0} \tilde{\mathcal{A}}(n), \qquad \tilde{\mathcal{A}}(n) = \left\{ w \in \tilde{\Sigma}^n \mid \pi(w) \in \mathcal{A}(n) \right\} \tag{16}$$

where π is the map

$$\pi : \tilde{\Sigma} \to \Sigma, \quad \pi(0) = \pi(0') = 0,\ \pi(1) = \pi(1') = 1,\ \pi(2) = \pi(2') = 2, \tag{17}$$

that projects back to the three-letter alphabet Σ. The map

$$\varrho : \begin{cases} 0 \to U^{(0)},\ 0' \to V^{(0)} \\ 1 \to U^{(1)},\ 1' \to V^{(1)} \\ 2 \to U^{(2)},\ 2' \to V^{(2)} \end{cases} \tag{18}$$

is a uniformly growing morphism from the language $\tilde{\mathcal{A}}$ into the language $\mathcal{L}$. By the condition (1), this morphism is square-free on all three-letter words in $\tilde{\mathcal{A}}$, i.e., the images of elements in $\tilde{\mathcal{A}}(3)$ are square-free. As ϱ is a uniformly growing morphisms, being square-free on $\tilde{\mathcal{A}}(3)$ implies, as proven in [5] and [3], that ϱ is a square-free morphism, i.e., it maps square-free words in $\tilde{\mathcal{A}}$ onto square-free words in $\mathcal{L}$, thus onto words in $\mathcal{A}$.

Lemma 1. The existence of an n-Brinkhuis triple pair implies the lower bound $s \geq 2^{1/(n-1)}$.

Proof. The existence of an n-Brinkhuis triple pair implies the inequality

$$a(mn) \geq 2^m a(m) \tag{19}$$

for any $m > 0$, because each square-free word of length m yields 2^m different square-free words of length mn. This means

$$\left(\frac{a(mn)}{a(m)} \right)^{\frac{1}{m}} \geq 2, \tag{20}$$

for any $m > 0$, and hence

$$s^{n-1} = \lim_{m\to\infty} \left(\frac{a(mn)}{a(m)} \right)^{\frac{1}{m}} \geq 2, \tag{21}$$

establishing the lower bound. □

The first lower bound was derived by Brinkhuis [4], who showed that $s \geq 2^{1/24}$ by constructing a 25-Brinkhuis triple pair consisting entirely of palindromic words. In that case, the conditions on the square-freeness of the images of three-letter words can be simplified to the square-freeness of the images of two-letter words and certain conditions on the "heads" and the "tails" of the words, which are easier to check explicitly. Brandenburg [3] produced a 22-Brinkhuis triple pair, which proves a lower bound of $s \geq 2^{1/21}$. For a long time, this was the best lower bound available, until, quite recently, Ekhad and Zeilberger [7] came up with a 18-Brinkhuis triple pair equivalent to

$$\begin{aligned} U^{(0)} &= 012021020102120210 & V^{(0)} &= 012021201020120210 = \bar{U}^{(0)} \\ U^{(1)} &= 120102101210201021 & V^{(1)} &= 120102012101201021 = \bar{U}^{(1)} \\ U^{(2)} &= 201210212021012102 & V^{(2)} &= 201210120212012102 = \bar{U}^{(2)}, \end{aligned} \tag{22}$$

thus establishing the bound $s \geq 2^{1/17}$. We note that the simpler definition for a Brinkhuis triple pair in [7], which is akin to Brinkhuis' original approach, is in fact incomplete, as it does not rule out a square that overlaps three adjacent words if the words are not palindromic. Nevertheless, the Brinkhuis triple (22) given in [7] is correct, and so is the lower bound $s \geq 2^{1/17} = 1.041\,616\ldots$ derived from it. In fact, it has been claimed (see [17]) that this is the optimal bound that can be obtained in this way, and this is indeed the case, see the discussion below.

It is interesting to note that, although this minimal-length Brinkhuis triple pair does not consist of palindromes, it is nevertheless invariant under reversion of words, as $V^{(i)} = \bar{U}^{(i)}$. In addition, it also shares the property with Brinkhuis' orginial triple that the words $U^{(1)}$, $V^{(1)}$ and $U^{(2)}$, $V^{(2)}$ which replace the letters 1 and 2, respectively, are obtained from $U^{(0)}$, $V^{(0)}$ by a global permutation τ of the three letters

$$\tau : \begin{cases} 0 \to 1 \\ 1 \to 2 \\ 2 \to 0 \end{cases} \tag{23}$$

i.e.,

$$U^{(2)} = \tau(U^{(1)}) = \tau^2(U^{(0)}), \quad V^{(2)} = \tau(V^{(1)}) = \tau^2(V^{(0)}). \tag{24}$$

Clearly, given any Brinkhuis triple pair, the sets of words obtained by reversion or by applying any permutation of the letters are again Brinkhuis triple pairs, so it may not be too surprising that a Brinkhuis triple pair of minimal length turns out to be invariant under these two operations.

4. Generalised Brinkhuis triples

As we cannot improve on the lower bound by constructing a shorter Brinkhuis triple, we proceed by generalising the notion. The idea is to allow for more than two words that replace each letter. This leads to the following general definition.

Definition 2. An n-Brinkhuis (k_0, k_1, k_2)-triple is a set of $k_0+k_1+k_2$ square-free words $\mathcal{B} = \{\mathcal{B}^{(0)}, \mathcal{B}^{(1)}, \mathcal{B}^{(2)}\}$, $\mathcal{B}^{(i)} = \{w_j^{(i)} \in S(n) \mid 1 \leq j \leq k_i\}$, $k_i \geq 1$, such that, for any square-free word $ii'i''$ of length 3 and any $1 \leq j \leq k_i$, $1 \leq j' \leq k_{i'}$, $1 \leq j'' \leq k_{i''}$, the composed word $w_j^{(i)} w_{j'}^{(i')} w_{j''}^{(i'')}$ of length $3n$ is square-free.

Note that the definition reduces to definition 1 in the case $k_0 = k_1 = k_2 = 2$ of an "ordinary" Brinkhuis triple pair. From the set of square-free words of length 3, we deduce that the number of composed words that enter is $6k_0k_1k_2 + k_0^2(k_1 + k_2) + k_1^2(k_0 + k_2) + k_2^2(k_0 + k_1)$.

Lemma 2. The existence of an n-Brinkhuis (k_0, k_1, k_2)-triple implies the lower bound $s \geq k^{1/(n-1)}$, where $k = \min(k_0, k_1, k_2)$.

Proof. The proof proceeds as in lemma 1 above, with 2 replaced by $k = \min(k_0, k_1, k_2)$. □

As far as the lower bound is concerned, we do not gain anything by considering triples where the number of words k_0, k_1 and k_2 differ from each other. Nevertheless, the generality of definition 2 shall be of use below. In order to derive improved lower bounds, we shall in fact concentrate on a more restricted class of triples.

Definition 3. A special n-Brinkhuis k-triple is an n-Brinkhuis (k, k, k)-triple $\mathcal{B} = \{\mathcal{B}^{(0)}, \mathcal{B}^{(1)}, \mathcal{B}^{(2)}\}$ such that $\mathcal{B}^{(2)} = \tau(\mathcal{B}^{(1)}) = \tau^2(\mathcal{B}^{(0)})$ and $w \in \mathcal{B}^{(0)}$ implies $\bar{w} \in \mathcal{B}^{(0)}$, where τ is the permutation of letters defined in equation (23).

The first condition means that all words in $\mathcal{B}^{(1)}$ and $\mathcal{B}^{(2)}$ can be obtained from the words in $\mathcal{B}^{(0)}$ by the global permutation τ. The second condition implies that the words in $\mathcal{B}^{(0)}$, and hence also in $\mathcal{B}^{(1)}$ and $\mathcal{B}^{(2)}$, are either palindromes, i.e., $w = \bar{w}$, or occur as pairs $(w, \bar{w})$. This means that a special n-Brinkhuis k-triple is characterised by the set of palindromes $w = \bar{w} \in \mathcal{B}^{(0)}$ and by one member of each pairs of non-palindromic words $(w, \bar{w}) \in \mathcal{B}^{(0)}$. If there are k_{p} palindromes and k_{n} pairs in $\mathcal{B}^{(0)}$, then these generate a special Brinkhuis k-triple with $k = k_{\mathrm{p}} + 2k_{\mathrm{n}}$. We shall call $K = (k_{\mathrm{p}}, k_{\mathrm{n}})$ the signature of the special Brinkhuis k-triple, and denote a set of $k_{\mathrm{p}} + k_{\mathrm{n}}$ generating words by $\mathcal{G}$.

In order to obtain the best lower bound possible, we are looking for optimal choices of the length n and the number of words k. There are two possibilities, we may look for the largest k for given length n, or for the smallest length n for a given number k. This is made precise by the following definitions.

Definition 4. An optimal special n-Brinkhuis triple is a special n-Brinkhuis k-triple such that any special n-Brinkhuis l-triple has $l \leq k$.

Definition 5. A minimal-length special Brinkhuis k-triple is a special n-Brinkhuis k-triple such that any special m-Brinkhuis k-triple has $m \geq n$.

If $\mathcal{B}$ is a special n-Brinkhuis k-triple, so is its image $\sigma(\mathcal{B})$ under any permutation $\sigma \in S_3$ of the three letters. Therefore, without loss of generality, we may assume that the first word $w_1^{(0)} \in \mathcal{B}$ starts with the letters 01. This has the following consequences on the other words of the triple.

Lemma 3. Consider a special n-Brinkhuis k-triple $\mathcal{B}$, with $n > 1$, such that the word $w_1^{(0)} \in \mathcal{B}^{(0)}$ starts with the letters 01. Then $n \geq 7$, and all words in $\mathcal{B}^{(0)}$ start with the three letters 012 and end on 210.

Proof. As $w_1^{(1)} = \tau(w_1^{(0)})$ and $w_1^{(2)} = \tau^2(w_1^{(0)})$, the words $w_1^{(1)}$ and $w_1^{(2)}$ start with letters 12 and 20, respectively. If $n = 2$, then $w_1^{(0)}w_1^{(1)} = 0112$ contains the square 11, so $n \geq 3$. Square-freeness of the composed words $w_j^{(0)}w_1^{(1)}$ and $w_j^{(0)}w_1^{(2)}$, $1 \leq j \leq k$, implies that the words $w_j^{(0)}$ have to end on 210, because $w = 210$ is the only word in $\mathcal{A}(3)$ such that $w12$ and $w20$ are both square-free. This in turn implies that all words in $\mathcal{B}^{(1)}$ and $\mathcal{B}^{(2)}$ end on 021 and 102, respectively. Now, square-freeness of the composed words $w_j^{(1)}w_{j'}^{(0)}$ and $w_j^{(2)}w_{j'}^{(0)}$ implies that the first three letter of $w_j^{(0)}$, for any $1 \leq j \leq k$, have to be $w = 012$, because this is the only word $\mathcal{A}(3)$ in such that $021w$ and $102w$ are both square-free. For $n = 3$ and $n = 4$, no such words exist, and the only possibility for $n = 6$ would be 012210 which is not square-free. For $n = 5$, the square-free word 01210 starts with 012 and ends on 210, but $w_1^{(0)}w_1^{(2)}w_1^{(0)} = 012102010201210$ contains the square of 0201. $\square$

One can even say more about the "heads" and "tails" of the words in a special Brinkhuis triple. There are two possible choices for the forth letter of $w_1^{(0)}$, and both possibilities fix further letters and cannot appear within the same special Brinkhuis triple. Therefore, we can distinguish two different types of special Brinkhuis triples.

Proposition 1. Consider a special n-Brinkhuis k-triple $\mathcal{B}$, with $n > 1$, such that the word $w_1^{(0)} \in \mathcal{B}^{(0)}$ starts with the letters 01. Then $n \geq 13$ and either all words in $\mathcal{B}^{(0)}$ are of the form $012021\ldots120210$, or all words are of the form $012102\ldots201210$.

Proof. From lemma 3, we know that $n \geq 7$ and $w_1^{(0)}$ starts with 012 and ends on 210. There are now two choices for the forth letter. Let us consider the case that $w_1^{(0)}$ starts with 0120. Then $w_1^{(1)}$ starts with 1201. Now, from lemma 3, $w_1^{(2)}$ ends on 102, and square-freeness of $w_1^{(2)}w_j^{(0)}$ implies that $w_j^{(0)}$ starts with 01202, and hence with 012021. Now $w_1^{(1)}$ starts with 120102 and $w_1^{(2)}$ with 201210. From square-freeness of $w_j^{(0)}w_1^{(1)}$ and $w_j^{(0)}w_1^{(2)}$, we can rule out $w_j^{(0)}$ ends on 1210, because both possible extension 101210 and 201210 result in squares. Hence $w_j^{(0)}$ ends on 0210 and, from square-freeness of $w_j^{(0)}w_1^{(w)}$, it has to end on 20210, and thus on 120210.

Consider now the second possibility, i.e., $w_1^{(0)}$ starts with 0121. Necessarily, it then starts with 01210. As $w_1^{(1)}$ ends on 021, square-freeness of $w_1^{(1)} w_j^{(0)}$ implies that $w_j^{(0)}$ starts with 012102. Then $w_1^{(1)}$ starts with 120210 and $w_1^{(2)}$ with 201021. Square-freeness of $w_j^{(0)} w_1^{(1)}$ and $w_j^{(0)} w_1^{(2)}$ rules out an ending 0210 for $w_j^{(0)}$, as the only possible extension 20120 and 120210 both result in squares. Hence $w_j^{(0)}$ ends on 01210, and, from square-freeness of $w_j^{(0)} w_1^{(1)}$, actually has to end on 201210.

Now, in both cases it is obviously impossible to find square-free words of length $n = 8, 9, 10, 12$ that satisfy these conditions. For the first case, the one choice left for $n = 11$ is 01202120210, which contains the square of 1202. In the second case, the only word for $n = 11$ that satisfies the conditions is 01210201210. In this case, $w_1^{(0)} w_1^{(1)} = 0121020121012021012021$ contains the square of 210120. □

The proofs of lemmas 3 and 1 are very explicit, but you may simplify the argument by realising that the conditions at both ends are essentially equivalent, as they follow from reversing the order of letters in combined words. The results restrict the number of words that have to be taken into account when looking for a special n-Brinkhuis k-triple. In what follows, we can restrict ourselves to the case $n \geq 13$. We denote the set of such square-free words by

$$\mathcal{A}_1(n) = \{w \in \mathcal{A}(n) \mid w = 012021 \ldots 120210\} \subset \mathcal{A}(n), \tag{25}$$

$$\mathcal{A}_2(n) = \{w \in \mathcal{A}(n) \mid w = 012102 \ldots 201210\} \subset \mathcal{A}(n), \tag{26}$$

and the number of such words by

$$a_1(n) := |\mathcal{A}_1(n)|, \tag{27}$$

$$a_2(n) := |\mathcal{A}_2(n)|. \tag{28}$$

We denote the number of palindromes by

$$a_{1\mathrm{p}}(n) := |\{w \in \mathcal{A}_1(n) \mid w = \bar{w}\}|, \tag{29}$$

$$a_{2\mathrm{p}}(n) := |\{w \in \mathcal{A}_2(n) \mid w = \bar{w}\}|, \tag{30}$$

and the number of non-palindromic pairs by

$$a_{1\mathrm{n}}(n) := \frac{1}{2}(a_1(n) - a_{1\mathrm{p}}(n)), \tag{31}$$

$$a_{2\mathrm{n}}(n) := \frac{1}{2}(a_2(n) - a_{2\mathrm{p}}(n)). \tag{32}$$

Clearly, there are no palindromic square-free words of even length, and thus $a_{1\mathrm{p}}(2n) = a_{2\mathrm{p}}(2n) = 0$, $a_{1\mathrm{n}}(2n) = a_1(2n)/2$ and $a_{2\mathrm{n}}(2n) = a_2(2n)/2$.

Now, for a word $w \in \mathcal{A}_1(n)$ or $w \in \mathcal{A}_2(n)$ to be a member of a special n-Brinkhuis triple, it must at least generate a triple by itself. This motivates the following definition.

Definition 6. A square-free palindrome $w = \bar{w} \in \mathcal{A}(n)$ is called admissible if w generates a special n-Brinkhuis 1-triple. A non-palindromic square-free word w of length n is admissible if w generates a special n-Brinkhuis 2-triple.

The hunt for optimal special n-Brinkhuis triples now proceeds in three steps.

Step 1. The first step consists of selecting all admissible words in $\mathcal{A}_1(n)$ and $\mathcal{A}_2(n)$. Let us denote the number of admissible palindromes in $\mathcal{A}_1(n)$ by $b_{1\mathrm{p}}(n)$ and the number of admissible non-palindromes by $2b_{1\mathrm{n}}(n)$, such that $b_{1\mathrm{n}}$ is the number of admissible pairs $(w, \bar{w})$ of non-palindromic words in $\mathcal{A}_1(n)$. Analogously, we define $b_{2\mathrm{p}}(n)$ and $b_{2\mathrm{n}}(n)$ for admissible words in $\mathcal{A}_2(n)$.

Step 2. The second step consists of finding all triples of admissible words that generate a special n-Brinkhuis triple. Depending on the number of palindromes k_{p} in that triple, which can be $k_{\mathrm{p}} = 0, 1, 2, 3$, these are special Brinkhuis k-triples with $k = 6, 5, 4, 3$, respectively. We denote the number of such admissible triples by $t_1(n)$ and $t_2(n)$. Here, we need to check the conditions of definition 2 for each triple. Using the structure of the special Brinkhuis triple, the number of words that have to be checked is substantially reduced from $12k^3$ to $k(2k^2 + k_{\mathrm{p}})$.

Step 3. The third and final step is purely combinatorial in nature, and does not involve any explicit checking of square-freeness of composed words. The reason is the following. A set $\mathcal{G}$, $|\mathcal{G}| \geq 3$, of words in $\mathcal{A}_1(n)$ or $\mathcal{A}_2(n)$, generates a special n-Brinkhuis triple if and only if all three-elemental subsets of $\mathcal{G}$ generate special n-Brinkhuis triples. This is obvious, because the conditions of definition 2 on three-letter words never involve more than three words simultaneously, so checking the condition for all subsets of three generating words is necessary and sufficient. Thus, the task is to find the largest sets of generating words such that all three-elemental subsets are contained in our list of admissible triples. In order to obtain an optimal special n-Brinkhuis triple, one has to take into account that $k = k_{\mathrm{p}} + 2k_{\mathrm{n}}$, so solutions with maximum number of generators are not necessarily optimal.

Even though this step is purely combinatorial and no further operations on the words are required, it is by far the most expensive part of the algorithm as the length n increases. Therefore, this is the part that limits the maximum length n that we can consider. Using a computer, we found the optimal Brinkhuis triples for $n \leq 41$. The results for generating words from $\mathcal{A}_1(n)$ are given in table 2, those for generating words taken from $\mathcal{A}_2(n)$ are displayed in table 3. We included partial results for $42 \leq n \leq 45$, in order to show how the number of admissible words grows for larger n. Even though we do not know the optimal n-Brinkhuis triples for these cases, it has to be expected that the value of k that can be achieved continues to grow, and it is certainly true for $n = 42$ where $k_{\mathrm{opt}} \geq 72$.

TABLE 2. Results of the algorithm to find optimal special n-Brinkhuis triples with generating words in $\mathcal{A}_1(n)$.

n	a	a_1	$a_{1\mathrm{p}}$	$a_{1\mathrm{n}}$	$b_{1\mathrm{p}}$	$b_{1\mathrm{n}}$	t_1	sign.	k_{opt}
13	342	0	0	0	0	0	0	(0,0)	0
14	456	0	0	0	0	0	0	(0,0)	0
15	618	1	1	0	0	0	0	(0,0)	0
16	798	0	0	0	0	0	0	(0,0)	0
17	1 044	1	1	0	1	0	0	(1,0)	1
18	1 392	4	0	2	0	1	0	(0,1)	2
19	1 830	5	1	2	1	0	0	(0,0)	0
20	2 388	4	0	2	0	0	0	(0,0)	0
21	3 180	1	1	0	0	0	0	(0,0)	0
22	4 146	2	0	1	0	0	0	(0,0)	0
23	5 418	3	1	1	0	1	0	(0,1)	2
24	7 032	4	0	2	0	1	0	(0,1)	2
25	9 198	13	3	5	2	1	1	(2,1)	4
26	11 892	16	0	8	0	1	0	(0,1)	2
27	15 486	18	2	8	2	0	0	(1,0)	1
28	20 220	10	0	5	0	1	0	(0,1)	2
29	26 424	27	3	12	2	3	4	(2,2)	6
30	34 422	52	0	26	0	4	0	(0,2)	4
31	44 862	64	4	30	2	7	8	(1,3)	7
32	58 446	64	0	32	0	6	5	(0,4)	8
33	76 122	60	6	27	3	7	30	(0,6)	12
34	99 276	70	0	35	0	7	13	(0,4)	8
35	129 516	109	9	50	4	13	328	(2,8)	18
36	168 546	174	0	87	0	27	1 304	(0,15)	30
37	219 516	291	9	141	6	27	2 533	(3,14)	31
38	285 750	376	0	188	0	30	973	(0,14)	28
39	372 204	386	12	187	3	35	2 478	(2,15)	32
40	484 446	428	0	214	0	55	10 767	(0,24)	48
41	630 666	593	15	289	4	76	28 971	(3,31)	65
42	821 154	926	0	463	0	114	74 080	?	?
43	1 069 512	1 273	23	625	12	156	229 180	?	?
44	1 392 270	1 518	0	759	0	170	235 539	?	?
45	1 812 876	1 788	26	881	17	191	510 345	?	?

Table 3. Results of the algorithm to find optimal special n-Brinkhuis triples with generating words in $\mathcal{A}_2(n)$.

n	a	a_2	$a_{2\text{p}}$	$a_{2\text{n}}$	$b_{2\text{p}}$	$b_{2\text{n}}$	t_2	sign.	k_{opt}
13	342	1	1	0	1	0	0	(1,0)	1
14	456	0	0	0	0	0	0	(0,0)	0
15	618	0	0	0	0	0	0	(0,0)	0
16	798	0	0	0	0	0	0	(0,0)	0
17	1 044	2	0	1	0	0	0	(0,0)	0
18	1 392	2	0	1	0	0	0	(0,0)	0
19	1 830	1	1	0	0	0	0	(0,0)	0
20	2 388	0	0	0	0	0	0	(0,0)	0
21	3 180	1	1	0	0	0	0	(0,0)	0
22	4 146	6	0	3	0	0	0	(0,0)	0
23	5 418	6	2	2	2	1	0	(1,1)	3
24	7 032	10	0	5	0	2	0	(0,1)	2
25	9 198	11	1	5	1	2	1	(1,2)	5
26	11 892	8	0	4	0	1	0	(0,1)	2
27	15 486	8	2	3	1	1	0	(1,1)	3
28	20 220	10	0	5	0	3	0	(0,2)	4
29	26 424	30	4	13	1	3	2	(0,3)	6
30	34 422	40	0	20	0	6	5	(0,4)	8
31	44 862	37	5	16	2	3	2	(1,2)	5
32	58 446	32	0	16	0	4	0	(0,2)	4
33	76 122	49	5	22	2	3	7	(1,3)	7
34	99 276	76	0	38	0	10	39	(0,5)	10
35	129 516	142	6	68	3	20	483	(2,7)	16
36	168 546	188	0	94	0	29	1 602	(0,16)	32
37	219 516	205	9	98	3	32	2 707	(1,13)	27
38	285 750	198	0	99	0	27	1 112	(0,11)	22
39	372 204	231	13	109	6	36	5 117	(2,14)	30
40	484 446	396	0	198	0	56	12 002	(0,19)	38
41	630 666	615	15	300	8	81	54 340	(1,29)	59
42	821 154	820	0	410	0	120	123 610	?	?
43	1 069 512	969	15	477	10	158	332 054	?	?
44	1 392 270	1070	0	535	0	166	362 560	?	?
45	1 812 876	1341	23	659	13	200	792 408	?	?

The optimal Brinkhuis triples are not necessarily unique, and the list also contains a case, $n = 29$, where there exist optimal Brinkhuis triples of both types. In general, several choices exist, which, however, cannot be combined into an even larger triple. A list of optimal n-Brinkhuis triples which at the same time are minimal-length Brinkhuis k_{opt} triples is given below.

Proposition 2. The following sets of words generate optimal and minimal-length Brinkhuis triples:

- $n = 13$, $k_{\mathrm{p}} = 1$, $k_{\mathrm{n}} = 0$, $k = 1$:
$$\mathcal{G}_{13} = \{0121021201210\} \tag{33}$$
- $n = 18$, $k_{\mathrm{p}} = 0$, $k_{\mathrm{n}} = 1$, $k = 2$:
$$\mathcal{G}_{18} = \{012021020102120210\} \tag{34}$$
- $n = 23$, $k_{\mathrm{p}} = 1$, $k_{\mathrm{n}} = 1$, $k = 3$:
$$\begin{aligned}\mathcal{G}_{23} = \{&01210212021012021201210,\\ &01210201021012021201210\}\end{aligned} \tag{35}$$
- $n = 25$, $k_{\mathrm{p}} = 1$, $k_{\mathrm{n}} = 2$, $k = 5$:
$$\begin{aligned}\mathcal{G}_{25} = \{&0121021202102012021201210,\\ &0121020102101201021201210,\\ &0121021201021012021201210\}\end{aligned} \tag{36}$$
- $n = 29$, $k_{\mathrm{p}} = 2$, $k_{\mathrm{n}} = 2$, $k = 6$:
$$\begin{aligned}\mathcal{G}_{29}^{(1)} = \{&01202120102012021020102120210,\\ &01202120121012021012102120210,\\ &01202102012101202120102120210,\\ &01202120102012021012102120210\}\end{aligned} \tag{37}$$
- $n = 29$, $k_{\mathrm{p}} = 0$, $k_{\mathrm{n}} = 3$, $k = 6$:
$$\begin{aligned}\mathcal{G}_{29}^{(2)} = \{&01210201021201020121021201210,\\ &01210201021202101201021201210,\\ &01210201021202102012021201210\}\end{aligned} \tag{38}$$
- $n = 30$, $k_{\mathrm{p}} = 0$, $k_{\mathrm{n}} = 4$, $k = 8$:
$$\begin{aligned}\mathcal{G}_{30} = \{&012102010210120102012021201210,\\ &012102010212012102012021201210,\\ &012102010212021020121021201210,\\ &012102120210120102012021201210\}\end{aligned} \tag{39}$$
- $n = 33$, $k_{\mathrm{p}} = 0$, $k_{\mathrm{n}} = 6$, $k = 12$:
$$\begin{aligned}\mathcal{G}_{33} = \{&012021020121012010212012102120210,\\ &012021020121021201021012102120210,\\ &012021020121021201210120102120210,\\ &012021201020120210121020102120210,\\ &012021201020121012021012102120210,\\ &012021201021012010212012102120210\}\end{aligned} \tag{40}$$

- $n = 35$, $k_{\mathrm{p}} = 2$, $k_{\mathrm{n}} = 8$, $k = 18$:

$$
\begin{aligned}
\mathcal{G}_{35} = \{&01202120102012102120121020102120210,\\
&01202120102101210201210120102120210,\\
&01202102010212010201202120102120210,\\
&01202102010212010201210120102120210,\\
&01202102012101201020121020102120210,\\
&01202102012101202120121020102120210,\\
&01202102012102120210121020102120210,\\
&01202120102012101201021012102120210,\\
&01202120102012102010210120102120210,\\
&01202120102120210201021012102120210\}
\end{aligned} \tag{41}
$$

- $n = 36$, $k_{\mathrm{p}} = 0$, $k_{\mathrm{n}} = 16$, $k = 32$:

$$
\begin{aligned}
\mathcal{G}_{36} = \{&012102010210120212010210121021201210,\\
&012102010210120212012101201021201210,\\
&012102010210121021201020121021201210,\\
&012102010210121021201021012021201210,\\
&012102010210121021202101201021201210,\\
&012102010212012101201020121021201210,\\
&012102010212012101201021012021201210,\\
&012102010212012102120210121021201210,\\
&012102010212021012010210121021201210,\\
&012102010212021020102101201021201210,\\
&012102120102101202120102012021201210,\\
&012102120102101210201021012021201210,\\
&012102120102101210212021012021201210,\\
&012102120121012010212021012021201210,\\
&012102120121012021201021012021201210,\\
&012102120121020102120102012021201210\}
\end{aligned} \tag{42}
$$

- $n = 40$, $k_{\mathrm{p}} = 0$, $k_{\mathrm{n}} = 24$, $k = 48$:

$$
\begin{aligned}
\mathcal{G}_{40} = \{&0120210201210120102120210121020102120210,\\
&0120210201210120212010210121020102120210,\\
&0120210201210212010210120212012102120210,\\
&0120210201210212012101201021012102120210,\\
&0120210201210212012101202120121020120210,\\
&0120210201210212012102010210120102120210,\\
&0120210201210212012102010212012102120210,\\
&0120210201210212012102012021012102120210,\\
&0120210201210212012102012021020102120210,\\
&0120210201210212021020120212012102120210,\\
&0120212010201202101210201021012102120210,\\
&0120212010201210120102012021012102120210,\\
&0120212010201210120102101202120102120210,\\
&0120212010201210120102120121020102120210,\\
&0120212010201210120210201021012102120210,\\
&0120212010201210120210201202120102120210,\\
&0120212010201210120212010210120102120210,\\
&0120212010201210212010201202120102120210,\\
&0120212010201210212012101202120102120210,
\end{aligned}
$$

$$
\begin{aligned}
&012021201021012010201210120212010212020210,\\
&012021201021202101202120102101210212020210,\\
&012021201021202101210201021201210212020210,\\
&012021201021202101210212010201210212020210,\\
&012021201021202102010212010201210212020210\}
\end{aligned}
\tag{43}
$$

- $n = 41$, $k_{\mathrm{p}} = 3$, $k_{\mathrm{n}} = 31$, $k = 65$:

$$
\begin{aligned}
\mathcal{G}_{41} = \{&01202102012102120210201202120121020120210,\\
&01202120121012010201210201021012102120210,\\
&01202120121021201021012010212012102120210,\\
&01202102012101201021202101202120102120210,\\
&01202102012101202120102012021012102120210,\\
&01202102012101202120102012021020102120210,\\
&01202102012102120102012101202120102120210,\\
&01202102012102120121012010212012102120210,\\
&01202102012102120121020120212012102120210,\\
&01202120102012021012010210121020102120210,\\
&01202120102012021012102010210120102120210,\\
&01202120102012021012102010212012102120210,\\
&01202120102012021012102012021020102120210,\\
&01202120102012021012102120102012102120210,\\
&01202120102012021012102120121020102120210,\\
&01202120102012021020102120102012102120210,\\
&01202120102012021020102120121020102120210,\\
&01202120102012101201020120212012102120210,\\
&01202120102012101202102010210120102120210,\\
&01202120102012101202102010212012102120210,\\
&01202120102012101202102012021012102120210,\\
&01202120102012102120102012021012102120210,\\
&01202120102012102120102101202120102120210,\\
&01202120102012102120121012021012102120210,\\
&01202120102012102120210201021012102120210,\\
&01202120102012102120210201202120102120210,\\
&01202120102101201020121012021012102120210,\\
&01202120102101201021202101202120102120210,\\
&01202120102120210121021201021012102120210,\\
&01202120102120210201202120102012102120210,\\
&01202120121012010201202120102012102120210,\\
&01202120121012021012102010212012102120210,\\
&01202120121012021012102120102012102120210,\\
&01202120121012021020102120102012102120210\}
\end{aligned}
\tag{44}
$$

Proof. The proof that these are indeed special Brinkhuis triples consist of checking the conditions of definition 2 explicitly. This has to be done by computer, as the number of symmetry-inequivalent composed words of length $3n$ that have to be checked for square-freeness is $k(2k^2 + k_{\mathrm{p}})$, which gives 549 445 words of length 123 for $\mathcal{G}_{41}$. A Mathematica [16] program `brinkhuistriples.m` that performs these checks accompanies this paper. This check is independent of the construction algorithm used to find the optimal triples. In order to show that these triples are indeed optimal, one has to go through the algorithm outlined above. This has been done, giving the results of tables 2 and 3. □

The triple for $n = 18$ is equivalent to the triple (22) of [7]. The corresponding lower bounds on s are

$$\begin{aligned}
n = 13, k = 1: &\quad s \geq 1^{1/12} = 1 \\
n = 18, k = 2: &\quad s \geq 2^{1/17} > 1.041616 \\
n = 23, k = 3: &\quad s \geq 3^{1/22} > 1.051204 \\
n = 25, k = 5: &\quad s \geq 5^{1/24} > 1.069359 \\
n = 29, k = 6: &\quad s \geq 6^{1/28} > 1.066083 \\
n = 30, k = 8: &\quad s \geq 8^{1/29} > 1.074338 \\
n = 33, k = 12: &\quad s \geq 12^{1/32} > 1.080747 \\
n = 35, k = 18: &\quad s \geq 18^{1/34} > 1.088728 \\
n = 36, k = 32: &\quad s \geq 32^{1/35} > 1.104089 \\
n = 40, k = 48: &\quad s \geq 48^{1/39} > 1.104355 \\
n = 41, k = 65: &\quad s \geq 65^{1/40} > 1.109999
\end{aligned} \tag{45}$$

Apparently, the largest value of n considered here yields the best lower bound. This suggests that the bound can be systematically improved by considering special Brinkhuis triples for longer words.

What about the restriction to spectial Brinkhuis triples? In general, it is not clear what the answer is, but for the Brinkhuis triple pair of [7] it can easily be checked by computer that one cannot find a shorter triple by lifting these restriction. In fact, this follows from the following stronger result which is easier to check.

Lemma 4. An n-Brinkhuis $(2, 1, 1)$-triple requires $n > 17$.

Proof. This can be checked by computer. The number of square-free words of length $n = 17$ is 1044. However, we do not need to check all 1044^4 possibilities. Without loss of generality, we may restrict one of the four words to start with the letters 01, leaving only $1044/6 = 174$ choices for this word. Furthermore, the two words in $\mathcal{B}^{(0)}$ may be interchanged, as well as the other two words; so it is sufficient to consider one order of words in both cases. No n-Brinkhuis $(2, 1, 1)$-triple was found for $n \leq 17$. □

5. Concluding remarks

By enumerating square-free ternary words up to length 110 and by constructing generalised Brinkhuis triples, we improved both upper and lower bounds for the number of ternary square-free words. The resulting bounds for the exponential growth rate s (9) are

$$1.109999 < 65^{1/40} \leq s \leq 8\,416\,550\,317\,984^{\frac{1}{108}} < 1.317278. \tag{46}$$

The main difficulty in improving the lower bound further is caused by the combinatorial step in the algorithm to find optimal special Brinkhuis triples.

The data in tables 2 and 3 suggest that generators from the set $\mathcal{A}_1(n)$ (25) are more likely to provide optimal n-Brinkhuis triples for large n than generators from the set $\mathcal{A}_2(n)$ (26). It would be interesting to know whether, in principle, the lower bound obtained in this way eventually converges to the actual value of s.

Acknowledgment

The author would like to thank Jean-Paul Allouche for useful discussions during a workshop at Oberwolfach in May 2001. The author gratefully acknowledges comments from Shalosh B. Ekhad and Doron Zeilberger, who pointed out an error in a previous attempt to improve the lower bound.

References

[1] M. Baake, V. Elser and U. Grimm, The entropy of square-free words, *Mathl. Comput. Modelling* **26** (1997) 13–26; math-ph/9809010.

[2] D. R. Bean, A. Ehrenfeucht and G. F. McNulty, Avoidable patterns in strings of symbols, *Pacific J. Math.* **85** (1979) 261–294.

[3] F.-J. Brandenburg, Uniformly growing k^{th} power-free homomorphisms, *Theor. Comp. Sci.* **23** (1983) 69–82.

[4] J. Brinkhuis, Non-repetitive sequences on three symbols, *Quart. J. Math. Oxford* **34** (1983) 145–149.

[5] M. Crochemore, Sharp characterizations of squarefree morphisms, *Theor. Comp. Sci.* **18** (1982) 221–226.

[6] M. Crochemore, Tests sur les morphismes faiblement sans carré, in: *Combinatorics on Words*, edited by L. J. Cummings, Academic Press, Toronto (1983), 63–89.

[7] S.B. Ekhad and D. Zeilberger, There are more than $2^{n/17}$ n-letter ternary square-free words. *Journal of Integer Sequences* **1** (1998) 98.1.9; math.CO/9809135.

[8] S. Finch, Pattern-free word constants, MathSoft Constants web site, URL: `http://www.mathsoft.com/asolve/constant/words/words.html`.

[9] Y. Kobayashi, Repetition-free words, *Theor. Comp. Sci.* **44** (1986) 175–197.

[10] M. Leconte, k-th power-free codes, in: *Automata on Infinite Words*, edited by M. Nivat and D. Perrin, *Automata on Infinite Words*, Lecture Notes in Computer Science 192, Springer, Berlin (1985), 172–187.

[11] M. Lothaire, *Combinatorics on Words*, Cambridge University Press, 1983.

[12] P. A. B. Pleasants, Nonrepetitive sequences, *Proc. Cambr. Philos. Soc.* **68** (1970) 267–274.

[13] P. Séébold, Overlap-free sequences, in: *Automata on Infinite Words*, edited by M. Nivat and D. Perrin, Lecture Notes in Computer Science 192, Springer, Berlin (1985), 207–215.

[14] R. O. Shelton, On the Structure and Extendability of Square-Free Words, in: *Combinatorics on Words*, edited by L. J. Cummings, Academic Press, Toronto (1983), 101–118.

[15] N. J. A. Sloane and S. Plouffe, *The Encyclopedia of Integer Sequences*, Academic Press, San Diego, 1995; see also the On-Line Encyclopedia of Integer Sequences, URL: `http://www.research.att.com/~njas/sequences/`.

[16] S. Wolfram, *Mathematica book* (4th edition), Cambridge University Press, Cambridge (1999).

[17] D. Zeilberger,
URL: http://www.math.temple.edu/~zeilberg/mamarim/mamarimhtml/jan.html.

Applied Mathematics Department, Faculty of Mathematics and Computing, The Open University, Walton Hall, Milton Keynes MK7 6AA, U.K.
E-mail address: u.g.grimm@open.ac.uk
URL: http://mcs.open.ac.uk/ugg2