Journal of Integer Sequences, Vol. 6 (2003),
Article 03.3.2

New Lower Bound On The Number of Ternary Square-Free Words

Xinyu Sun
Department of Mathematics
Temple University
Philadelphia, PA 19122
USA
xysun@math.temple.edu

Abstract

A new lower bound on the number of n-letter ternary square-free words is presented: $110^{n/42}$, which improves the previous best result of $65^{n/40}$.

1. Introduction

A word w is a finite sequence of letters from a certain alphabet Σ. The length of a word is the number of letters of the word. Binary words are the words from a two-letter alphabet $\{0, 1\}$, whereas ternary words are from a three-letter alphabet $\{0, 1, 2\}$. A word is square-free if it does not contain two identical consecutive subwords (factors), i.e., w cannot be written as $axxb$ where a, b, x are words with x non-empty.

It is easy to see that there are only finitely many binary square-free words. However, there are infinitely many ternary square-free words. The fact was proved by utilizing what is now called the Prouhet-Thue-Morse sequence (see [10]). Brinkhuis[3], Brandenburg[2] (also in [1]), Zeilberger[5] and Grimm[8] showed that the numbers of such words of length n are greater than $2^{n/24}$, $2^{n/21}$, $2^{n/17}$, and $65^{n/40}$ respectively. Details on words and related topics can be found in [6] and [11].

While the best available upper bound has been very close to the estimate as described later, the available lower bounds still have much room for improvement. Finding better lower bounds has posed as a algorithmic challenge, as well as a theoretic one. As explained later, the complexity of the algorithm used here is likely (very) exponential.

2. Brinkhuis Triples

We denote $a(n)$ to be the number of ternary square-free words of length n. It is easy to see that

$$a(m+n) \leq a(m)a(n) \tag{2.1}$$

for all $m,\ n \geq 0$, which implies (see in [1]) the existence of the limit

$$s := \lim_{n\to\infty} a(n)^{1/n}, \tag{2.2}$$

which is also called the growth rate or "connective constant" of ternary square-free words.

It is widely believed that the available upper bounds are very close to the actual value of s. In fact, it has been estimated by Noonan and Zeilberger [7] that $s \approx 1.302$ using the Zinn-Justin method, and they have also proved that $s \leq 1.30201064$ by implementing the Golden-Jackson method.

Definition 1. An n-Brinkhuis k-triple is three sets of words $\mathcal{B} = \{\mathcal{B}^0, \mathcal{B}^1, \mathcal{B}^2\}$, $\mathcal{B}^i = \{w_j^i | 1 \leq j \leq k\}$, where w_j^i are square-free words of length n, such that for any square-free word $i_1 i_2 i_3$, $0 \leq i_1, i_2, i_3 \leq 2$, and any $1 \leq j_1, j_2, j_3 \leq k$, the word $w_{j_1}^{i_1} w_{j_2}^{i_2} w_{j_3}^{i_3}$ of length $3n$ is also square-free.

Based on an n-Brinkhuis k-triple, we can define the following set of uniformly growing morphisms:

$$\rho = \begin{cases} 0 \to w_{j_0}^0,\ 1 \leq j_0 \leq k; \\ 1 \to w_{j_1}^1,\ 1 \leq j_1 \leq k; \\ 2 \to w_{j_2}^2,\ 1 \leq j_2 \leq k. \end{cases} \tag{2.3}$$

As proven in [2], [4] and [9], ρ are square-free morphisms, i.e., they map each square-free word of length m onto k^m different images of square-free words of length nm.

Therefore, the existence of an n-Brinkhuis k-triple indicates that

$$\frac{a(mn)}{a(m)} \geq k^m \tag{2.4}$$

for any $m \geq 1$, which implies

$$s^{n-1} = \lim_{n\to\infty} \left(\frac{a(mn)}{a(m)}\right)^{1/m} \geq k, \tag{2.5}$$

and thus yields the lower bound of $s \geq k^{1/(n-1)}$.

Given the permutation $\tau = (0,\ 1,\ 2)$, we can have

Definition 2. A quasi-special n-Brinkhuis k-triple is an n-Brinkhuis k-triple such that $\mathcal{B}^1 = \tau(\mathcal{B}^0)$, $\mathcal{B}^2 = \tau(\mathcal{B}^1)$.

Definition 3. A special n-Brinkhuis k-triple is a quasi-special n-Brinkhuis k-triple such that $w \in \mathcal{B}^0$ implies $\overline{w} \in \mathcal{B}^0$, where $\overline{w}$ is the reversion of w.

Grimm[8] was able to construct a special 41-Brinkhuis 65-triple, hence proved $s \geq 65^{1/40}$.

3. Main Results

Definition 4. A word w is admissible if $(w, \tau(w), \tau^2(w))$ is a quasi-special Brinkhuis 1-triple by itself.

Definition 5. An optimal quasi-special (special) n-Brinkhuis k-triple is a quasi-special (special) n-Brinkhuis k-triple such that any quasi-special (special) n-Brinkhuis l-triple has $l \leq k$.

To find the optimal quasi-special n-Brinkhuis triples, we only need to find the set of all admissible words of length n, and its largest subset in which any three words w_1, w_2, w_3 can form a quasi-special n-Brinkhuis 3-triple, i.e., $\{\{w_1, w_2, w_3\}, \{\tau(w_1), \tau(w_2), \tau(w_3)\}, \{\tau^2(w_1), \tau^2(w_2), \tau^2(w_3)\}\}$ is a quasi-special n-Brinkhuis 3-triple. A Maple package was written to calculate such words and sets. The results are listed below.

Proposition 3.1. *Special n-Brinkhuis triples yield the best possible results for each $13 \leq n \leq 20$, and quasi-special Brinkhuis triples do not yield better results than special n-Brinkhuis triples for each $13 \leq n \leq 39$, except 37.*

n	b_1	k_1	b_2	k_2
13	1	1	1	1
14	0	0	0	0
15	0	0	0	0
16	0	0	0	0
17	1	1	1	1
18	1	2	1	2
19	1	1	1	1
20	0	0	0	0
21	0	0	0	0
22	0	0	0	0
23	1	3	1	3
24	5	2	3	2
25	1	5	1	5
26	2	2	2	2
27	1	3	1	3
28	4	4	2	4
29	2	6	2	6
30	1	8	1	8
31	4	7	2	7
32	1	8	1	8
33	1	12	1	12
34	33	10	5	10
35	2	18	2	18
36	1	32	1	32
37	66	32	24	31
38	9	28	3	28
39	1	32	1	32
40			2	48
41			8	65
42			4	76
43			2	110

In the table above, n is the length of the words; b_1 and k_1 are the numbers of all available optimal quasi-special Brinkhuis triples and the numbers of elements in the triples; b_2 and k_2 are those of the special Brinkhuis triples. Notice the numbers of the triples and their sizes do not always grow as n does, and occasionally there are extraordinary amount of the triples for certain word lengths, i.e., 34 and 37.

Although there are often more choices for the regular and quasi-special Brinkhuis triples than the special Brinkhuis triples as listed above, none of them can be combined to form larger triples. And the exception of $n = 37$ has hardly any significance because the results are superceded by the 36-Brinkhuis 32-triples already. These results strongly suggest that the special Brinkhuis triples will generally yield the best results regardless of n.

It is reasonable to believe that there exist n-Brinkhuis triples that are not quasi-special when $n > 20$, or quasi-special n-Brinkhuis triples that are not special when $n > 39$. However, as explained in the proof of the following proposition, it is vary hard to find such triples due to the complexity.

Proposition 3.2. *The following 43-Brinkhuis 110-triple exists, and thus shows* $s \geq 110^{1/42} = 1.118419\ldots > 65^{1/40} = 1.109999\ldots$*:*

$$
\begin{array}{ll}
\{ & 0120212012102120102012102010212012102120210, \\
 & 0120212010210120102012102010210120102120210, \\
 & 0120212010201210212021020120212012102120210, \\
 & 0120212012102120210201202120121020102120210, \\
 & 0120210201210120210121020120212012102120210, \\
 & 0120212012102120210201210120210121020120210, \\
 & 0120210201210120102120210120212012102120210, \\
 & 0120212012102120210120212010210121020120210, \\
 & 0120210201210120212010210120212012102120210, \\
 & 0120212012102120210120102120210121020120210, \\
 & 0120212010201210212010210120212012102120210, \\
 & 0120212012102120210120102120121020102120210, \\
 & 0120212010210121021201210120212012102120210, \\
 & 0120212012102120210121021201210120102120210, \\
 & 0120212012101201021201210120212012102120210, \\
 & 0120212012102120210121021201021012102120210, \\
 & 0120212010210121021202102010212012102120210, \\
 & 0120212012102120102012021201210120102120210, \\
 & 0120212012101201021202102010212012102120210, \\
 & 0120212012102120102012021201021012102120210, \\
 & 0120212010210120102012102010212012102120210, \\
 & 0120212012102120102012102010210120102120210, \\
 & 0120210201210120102012102010212012102120210, \\
 & 0120212012102120102012102010210121020120210, \\
 & 0120210201210120212012102010212012102120210,
\end{array}
$$

012021201210212010201210212021012102012020210,
012021201020121021201210201021201210212020210,
012021201210212010201210212012102010212020210,
012021020121021202101210201021201210212020210,
012021201210212010201210120212012102012020210,
012021201210120102101210201021201210212020210,
012021201210212010201210120102101210212020210,
012021201020121020102101201021201210212020210,
012021201210212010210120102012102010212020210,
012021201021202102012101201021201210212020210,
012021201210212010210121020120212010212020210,
012021201021012102012101201021201210212020210,
012021201210212010210121020121012010212020210,
012021201210120102120210201202101210212020210,
012021201210120210201202120102101210212020210,
012021020121012021201210201202101210212020210,
012021201210120210201210212021012102012020210,
012021201020121021201210201202101210212020210,
012021201210120210201210212012102010212020210,
012021201021012010212012101202101210212020210,
012021201210120210121021201021012010212020210,
012021020121012010212012101202101210212020210,
012021201210120210121021201021012102012020210,
012021201021012010212021020102101210212020210,
012021201210120102012021201021012010212020210,
012021020121012010212021020102101210212020210,
012021201210120102012021201021012102012020210,
012021201021012102120121020102101210212020210,
012021201210120102012102120121012010212020210,
012021201021012010201202120102101210212020210,
012021201210120102120210201021012010212020210,
012021020121012010201202120102101210212020210,
012021201210120102120210201021012102012020210,
012021020121021202101202120102101210212020210,
012021201210120102120210120212012102012020210,
012021020121012021012102120102101210212020210,
012021201210120102120121012021012102012020210,
012021201021012010201210120102101210212020210,
012021201210120102101210201021012010212020210,
012021020121012021201210120102101210212020210,
012021201210120102101210212021012102012020210,
012021201020121021201210120102101210212020210,

$$
\begin{array}{l}
012021201210120102101210212012102010212020210,\\
012021020121012021201210201202102010212020210,\\
012021201020120210201210212021012102012020210,\\
012021201020121021201210201202102010212020210,\\
012021201020120210201210212012102010212020210,\\
012021201021012010212012101202102010212020210,\\
012021201020120210121021201021012010212020210,\\
012021020121012010212012101202102010212020210,\\
012021201020120210121021201021012102012020210,\\
012021020121021201210120102012102010212020210,\\
012021201020121020102101210212012102012020210,\\
012021020121021201210120212012102010212020210,\\
012021201020121021202101210212012102012020210,\\
012021201021012010201210212012102010212020210,\\
012021201020121021201210201021012010212020210,\\
012021020121021202101210212012102010212020210,\\
012021201020121021201210120212012102012020210,\\
012021020121021202102010212012102010212020210,\\
012021201020121021201020120212012102012020210,\\
012021201020121012010212021012102010212020210,\\
012021201020121012021201021012102010212020210,\\
012021020121021202102012021012102010212020210,\\
012021201020121012021020120212012102012020210,\\
012021201021012010212021020120212010212020210,\\
012021201021202102012021201021012010212020210,\\
012021020121012010212021020120212010212020210,\\
012021201021202102012021201021012102012020210,\\
012021201021012102120121020120212010212020210,\\
012021201021202102012102120121012010212020210,\\
012021201021012102120210201021012010212020210,\\
012021201021012010201202120121012010212020210,\\
012021020121012010201210201021012010212020210,\\
012021201021012010201210201021012102012020210,\\
012021020121012021201210201021012010212020210,\\
012021201021012010201210212021012102012020210,\\
012021020121021202101210201021012010212020210,\\
012021201021012010201210120212012102012020210,\\
012021020121012010201202120121012010212020210,\\
012021201021012102120210201021012102012020210,\\
012021020121021202101202120121012010212020210,\\
012021201021012102120210120212012102012020210,\\
012021020121012021012102120121012010212020210,\\
012021201021012102120121012021012102012020210 \quad \}
\end{array}
$$

Proof: Each admissible word is of length at least 13 and of the form either $012021 \cdots 120210$ or $012102 \cdots 201210$ as proved by Grimm [8]. So we first find all the square-free words of length $n-12$, attach the two pairs of prefixes and suffixes to these words, then determine if the results are square-free and admissible words, and label them from 1 to m, where m is the total number of such words. The next step is to find all quasi-special (special) Brinkhuis 3-triples and replace the words with the labels we just assigned to them. Thus each triple correspond to a unique ordered list of three different integers, and we have created a set of lists of integers S. Note that if the square-free words of length $n-12$ are known, the rest of the process above only take polynomial time. Now the problem is reduced to find the largest subset T of $\{1,\ldots,m\}$ so that the list of any three elements of T is an element of S. Such a question is obviously NP, because the certificate will be the solution itself, and the time required to verify the certificate will be $O(\binom{n}{3})$, thus polynomial. Fortunately, we are not obliged to tell how long it takes to get the certificate.

We now create a graph G so that each element in S is a vertex of G, and any two vertices are connected if and only if any combination of three different numbers from the two lists can form a quasi-special (special) Brinkhuis 3-triple. For example, if $[1,2,3]$ and $[1,2,4]$ are vertices of the graph, they can be connected if and only if $[1,3,4]$ and $[2,3,4]$ are vertices of the graph too. And in this case, the four vertices will form a complete graph. Now we have reduced the problem into finding the largest complete subgraph of a graph, which is known to be NP-complete, in polynomial time. Although what we did does not imply the original problem to be NP-complete, it does shed some light on how to solve the problem: we will use the backtracking method to find the largest Brinkhuis triple.

We say a number i is compatible with a list of numbers $i_1,\ldots,i_n$ if any three words chosen from the corresponding words $w_i, w_{i_1},\ldots,w_{i_n}$ can form a quasi-special (special) Brinkhuis 3-triple.

Assuming all the numbers in the vertices are ordered increasingly, we try to construct the largest quasi-special (special) Brinkhuis triples recursively: We start with the pair of numbers, a_1 and a_2, who has the largest set of compatible numbers of all pairs of numbers in $\{1,\ldots,m\}$. After we have a list $a_1,\ldots,a_{n-1}$ such that every three numbers in the list can form a quasi-special (special) Brinkhuis 3-triple, we try to find a_n as the number such that a_n is compatible with $a_1,\ldots,a_{n-1}$, and $a_1,\ldots,a_n$ has the largest possible set of compatible numbers. If there is a tie, we choose the smallest possible number. Once we cannot add another number to the current list of $a_1,\ldots,a_n$, we have found a "locally optimal" Brinkhuis n-triple. We then backtrack to a_{n-1} and search for the next best choice of a_n. When all such choices are analyzed, we backtrack to a_{n-2}. We repeat the process until we backtrack to a_1 and a_2, when we try the pair of numbers who has the next largest set of compatible numbers. We will continue until all the possibilities are considered. Of course, we can always break out of the search if the size of the list of numbers found plus the number of compatible numbers available is less than the best known size of the triples at the time.

The complexity of searching the largest complete subgraph of n vertices is equivalent to searching the largest independent set of vertices of the complement of the graph, whose *average* rate of growth is subexponential, i.e., $O(n^{\log n})$. However, the exact amount of labor required for a specific kind of graphs can be very exponential. Theoretically, we can take advantage of the special structure of the graphs to increased the performance: if vertices $[1,2,3]$ and $[4,5,6]$ are connected, there is automatically a complete subgraph of 20 vertices, namely any combinations of three numbers from 1 to 6. But such an approach will use recursive programming, which would have required exponential space and thus is impractical. Unless we can find other methods to find the lower bound, using Brinkhuis triples cannot provide must better results, even with more powerful (multi-processor) computers. Unfortunately, this is the best method known yet, if not the only one.

The Maple package and the results on optimal Brinkhuis triples are all available at http://www.math.temple.edu/~xysun/ternarysf/ternary_square_free.htm.

4. Acknowledgement

Many thanks to Doron Zeilberger for his encouragements, to Uwe Grimm for pointing out an error in the first attempt, and to Li Zhang for his useful discussions and comments during the development of the programs. Also thanks to the referees who provided suggestions that made this paper more informative and readable.

References

1. M. Baake, V. Elaser and U. Grimm, The entropy of square-free words, *Math. Comput. Modelling* **26** (1997), 13–26.
2. F.-J. Brandenburg, Uniformly growing k^{th} power-free homomorphisms, *Theoret. Comput. Sci.* **23** (1983), 69–82.
3. J. Brinkhuis, Nonrepetitive sequences on three symbols, *Quart. J. Math. Oxford* **34** (1983), 145–149.
4. M. Crochemore, Sharp characterizations of squarefree morphisms, *Theoret. Comput. Sci.* **18** (1982), 221–226.
5. S. B. Ekhad and D. Zeilberger, There are more than $2^{n/17}$ n-letter ternary square-free words, *J. Integer Seq.* **1** (1998), Article 98.1.9.
6. S. Finch, Pattern-free word constants, http://pauillac.inria.fr/algo/bsolve/constant/words/words.html
7. John Noonan and Doron Zeilberger, The Goulden-Jackson cluster method: extensions, applications and implementations, *J. Differ. Equations Appl.* **5** (1999), 355–377.
8. Uwe Grimm, Improved bounds on the number of ternary square-free words, *J. Integer Seq.* **4** (2001), Article 01.2.7.
9. Michel Leconte, A characterization of power-free morphisms, *Theoret. Comput. Sci.* **38** (1985), 117–122.
10. M. Lothaire, *Combinatorics on Words*, Addison-Wesley, 1983.
11. Wolfram Research, Squarefree word, http://mathworld.wolfram.com/SquarefreeWord.html

---

2000 *Mathematics Subject Classification*: Primary 05A20; Secondary 68R15.
Keywords: square-free, ternary, word, Brinkhuis triple

(Concerned with sequence A006156.)

Received November 21, 2002; revised version received August 1, 2003; Published in *Journal of Integer Sequences*, August 18, 2003.

Return to Journal of Integer Sequences home page.