#A11 INTEGERS 14 (2014)

ABELIAN COMPLEXITY OF FIXED POINT OF MORPHISM $0 \mapsto 012, 1 \mapsto 02, 2 \mapsto 1$

F. Blanchet-Sadri[1]
Department of Computer Science, University of North Carolina, Greensboro, North Carolina
blanchet@uncg.edu

James D. Currie[2]
Department of Mathematics and Statistics, University of Winnipeg, Winnipeg, Manitoba, Canada
j.currie@uwinnipeg.ca

Narad Rampersad[2]
Department of Mathematics and Statistics, University of Winnipeg, Winnipeg, Manitoba, Canada
narad.rampersad@gmail.com

Nathan Fox[1]
Department of Mathematics, Rutgers University, Piscataway, New Jersey
fox@math.rutgers.edu

Received: 6/4/13, Revised: 1/8/14, Accepted: 1/30/14, Published: 2/20/14

Abstract

We study the combinatorics of vtm, a variant of the Thue-Morse word generated by the non-uniform morphism $0 \mapsto 012, 1 \mapsto 02, 2 \mapsto 1$ starting with 0. This infinite ternary sequence appears a lot in the literature and finds applications in several fields such as combinatorics on words; for example, in pattern avoidance it is often used to construct infinite words avoiding given patterns. It has been shown that the factor complexity of vtm, i.e., the number of factors of length n, is $\Theta(n)$; in fact, it is bounded by $\frac{10}{3}n$ for all n, and it reaches that bound precisely when n can be written as 3 times a power of 2. In this paper, we show that the abelian complexity of vtm, i.e., the number of Parikh vectors of length n, is $O(\log n)$ with constant approaching $\frac{3}{4}$ (assuming base 2 logarithm), and it is $\Omega(1)$ with constant 3 (and these are the best possible bounds). We also prove some results regarding factor indices in vtm.

[1]F. Blanchet-Sadri and Nathan Fox's research was supported by the National Science Foundation under Grant No. DMS–1060775.
[2]James D. Currie and Narad Rampersad's research was supported by NSERC Discovery grants.

1. Introduction

The *Thue-Morse word*, which we denote by $\mathtt{tm}$, is defined as the fixed point of the uniform morphism $0 \mapsto 01, 1 \mapsto 10$ that starts at 0:

$$\mathtt{tm} = 0110100110010110100101100110100 1\cdots .$$

In [1], Allouche and Shallit surveyed this well-known infinite binary sequence and discussed some of its applications in various fields such as combinatorics on words, differential geometry, number theory, semigroup and group theory, real analysis, and physics. There are several alternative definitions of this sequence other than the abovementioned one; for instance, the Thue-Morse word is the lexicographically largest overlap-free binary sequence starting with 0.

The *factor complexity* of an infinite sequence w, denoted by ρ_w, counts the number of distinct factors of w, i.e., $\rho_w(n)$ is the number of factors of w of length n. Recent references on factor complexity include [2, Chapter 10] and [4, Chapter 4]. Closed-form formulas for the factor complexity of $\mathtt{tm}$ are known [5, 8, 20, 3]. We recall the recursive definition from [3, Proposition 2.10]: $\rho_{\mathtt{tm}}(0) = 1$, $\rho_{\mathtt{tm}}(1) = 2$, $\rho_{\mathtt{tm}}(2) = 4$, $\rho_{\mathtt{tm}}(3) = 6$, and for $n \geq 2$,

$$\rho_{\mathtt{tm}}(2n+1) = 2\rho_{\mathtt{tm}}(n+1), \qquad \rho_{\mathtt{tm}}(2n) = \rho_{\mathtt{tm}}(n+1) + \rho_{\mathtt{tm}}(n). \tag{1}$$

In [9], a formula was obtained for the factor complexity of the fixed points of binary uniform morphisms.

Rather than counting factors, one may also count Parikh vectors (see, for example, [19]). Let x be a word over an ordered alphabet Σ of size k. The *Parikh vector* of x is the k-vector $\Psi(x)$ defined as follows: the i-th component of $\Psi(x)$ equals the number of occurrences of the i-th letter of Σ in x. The *abelian complexity* of an infinite sequence w, denoted by ρ^{ab}_w, is defined by

$$\rho^{ab}_w(n) = |\,\{\Psi(x) : x \text{ is a factor of } w \text{ of length } n\}\,|.$$

It is easy to see that $\rho^{ab}_{\mathtt{tm}}(n) = 2$ for n odd and $\rho^{ab}_{\mathtt{tm}}(n) = 3$ for $n \neq 0$ even.

Define $\mathtt{pd}$ as the fixed point of the morphism $0 \mapsto 01, 1 \mapsto 00$ beginning with 0:

$$\mathtt{pd} = 0100010101000100010001 01\cdots .$$

The word $\mathtt{pd}$ is also classical and has its own name: the *period-doubling word*. Different ways to construct $\mathtt{pd}$ as well as its connection to the Thue-Morse word for which it is the "derivative" sequence, are well-known. The abelian complexity of $\mathtt{pd}$ has been studied by [13].

In this paper, we discuss a variant of the Thue-Morse word that we denote by $\mathtt{vtm}$. It is the fixed point of the non-uniform morphism $0 \mapsto 012, 1 \mapsto 02, 2 \mapsto 1$ beginning with 0:

$$\texttt{vtm} = 012021012102012021020121\cdots.$$

This infinite ternary sequence also appears quite a lot in the literature and finds applications in different research areas such as square-free walks on labelled graphs [12], infinite words avoiding patterns [17], etc. It has the property of being *square-free*, i.e., it does not contain any factor of the form $x^2 = xx$ for a non-empty word x. Recently, Rao, Rigo, and Salimov [18] showed an even stronger avoidability result, namely that vtm avoids 2-*binomial squares* (see their paper for details).

Here, we discuss in particular the first five following equivalent definitions of vtm and the other five definitions of pd that provide us with the abelian complexity (positions in sequences are labelled starting with 0):

- First, vtm's digits are the number of 0s between consecutive 1s in tm, i.e., the ith symbol of vtm is the number of 0s between the ith 1 and the $(i+1)$st 1 in tm. Similarly, the image of the word whose ith symbol is the number of 1s between the ith 0 and the $(i+1)$st 0 in tm under the map $0 \mapsto 2, 1 \mapsto 1, 2 \mapsto 0$ is the word vtm.

- Second, the image of the fixed point of the morphism $a \mapsto ab$, $b \mapsto ax$, $c \mapsto \beta\alpha$, $d \mapsto dy$, $x \mapsto \gamma\delta$, $y \mapsto dc$, $\alpha \mapsto \gamma c$, $\beta \mapsto d\delta$, $\gamma \mapsto a\alpha$, $\delta \mapsto \beta x$ beginning with a under the map $a \mapsto 012$, $b \mapsto 021$, $c \mapsto 120$, $d \mapsto 210$, $x \mapsto 102$, $y \mapsto 201$, $\alpha \mapsto 101$, $\beta \mapsto 202$, $\gamma \mapsto 020$, $\delta \mapsto 121$ is the word vtm (Theorem 1).

- Third, if we insert a 0 between consecutive 2s, a 2 between consecutive 0s, and a 1 between a 0 and a 2 (in either order) in the fixed point of the morphism $0 \mapsto 02, 2 \mapsto 20$ starting with 0, we obtain vtm.

- Fourth, if we insert a 1 into $(02)^\omega = 020202\cdots$ at each position congruent to $2^{2n-1} - 1$ modulo 2^{2n} for some $n \geq 1$, we obtain vtm (Lemmas 4 and 5).

- Fifth, if we replace every other 0 in pd by a 2, beginning with the second 0, we obtain vtm.

- Sixth, the image of vtm under the map $0 \mapsto 0, 1 \mapsto 1, 2 \mapsto 0$ is pd (Lemma 7).

- Seventh, the image of the fixed point of the morphism $a \mapsto ab$, $b \mapsto ac$, $c \mapsto \beta\alpha$, $\alpha \mapsto \beta c$, $\beta \mapsto a\alpha$ starting with a under the map $a \mapsto 010$, $b \mapsto 001$, $c \mapsto 100$, $\alpha \mapsto 101$, $\beta \mapsto 000$ is pd.

- Eighth, the word with a 1 at each position congruent to $2^{2n-1} - 1$ modulo 2^{2n} for some $n \geq 1$, and with a 0 otherwise, is the word pd (Lemma 5).

- Ninth, let $v_0 = 0$, and define a recursive sequence of words by $v_{2n} = v_{2n-1}0v_{2n-1}$ and $v_{2n+1} = v_{2n}1v_{2n}$. Then, $\lim_{n\to\infty} v_n = \texttt{pd}$ (Proposition 1).

- Tenth, let φ be the morphism $1 \mapsto 131, 3 \mapsto 13331$ and let ϕ be the map $1 \mapsto 01, 3 \mapsto 0001$. Then, $\phi(\varphi^{\omega}(1)) = \texttt{pd}$.

It is a standard exercise, for those who know the technique of (bi)special words or know the relationships between vtm, tm, and pd, to show that the factor complexity of vtm is $\Theta(n)$. In fact, $\rho_{\texttt{vtm}}(n) \leq \frac{10}{3}n$ holds for all n, and $\rho_{\texttt{vtm}}(n) = \frac{10}{3}n$ if and only if n can be written as 3 times a power of 2. This can be found in [7] and also in [10]. The $\Theta(n)$ factor complexity is also a consequence of several theorems, such as [2, Theorem 10.4.12], using primitivity, or [2, Theorem 10.3.1], using the fact that vtm is an automatic word. The technique of (bi)special words has first been described in [6] and has further been developed in particular in [14].

The contents of our paper are as follows: In Section 2, we study combinatorics on vtm. In Section 3, we prove that the abelian complexity of vtm is $O(\log n)$ with constant approaching $\frac{3}{4}$ (assuming base 2 logarithm), and it is $\Omega(1)$ with constant 3 (and these are the best possible bounds) (Corollary 1). Finally in Section 4, we prove two results regarding factor indices in vtm. Let w be an infinite word and let x be a factor of w. Then an integer i is an *index of x in w* if x occurs at position i of w. We prove: (1) If u is a factor of vtm and m is an odd number, then the set of indices of u in vtm contains a representative of every congruence class modulo m. (2) If u is a factor of vtm and m is a positive integer then the set of indices of u modulo m is the same as the set of the indices of $\tilde{u}$, where $\tilde{u}$ is obtained from u by replacing 0s with 2s and vice versa.

2. Combinatorics on vtm

Our first aim is to give a few combinatorial properties of vtm. We consider blocks of three letters in vtm, i.e., all factors of vtm of length three including those that occur at positions not divisible by 3. We start with two lemmas.

Lemma 1. vtm *does not contain factors* 010 *or* 212.

Proof. Recall that vtm is square-free. Assume for a contradiction that it contains 010. Applying the morphism to this yields 01202012, which contains $(20)^2$, a contradiction. Now, assume for a contradiction that it contains 212. Then, it must contain 02120, as anything else flanking 212 would create a square. Applying the morphism to this yields 0121021012, which contains $(210)^2$, a contradiction. □

Lemma 2. *Let $\alpha = 101$, $\beta = 202$, $\gamma = 020$, and $\delta = 121$. Let u and v represent any other length 3 ternary word. Then, the following are the only possible forms of factors of* vtm *containing α, β, γ, or δ: $u\alpha v$, $u\beta v$, $u\gamma v$, $u\delta v$, $u\alpha\beta v$, $u\beta\alpha v$, $u\gamma\delta v$, $u\delta\gamma v$, $u\alpha\beta\alpha v$, $u\beta\alpha\beta v$, $u\gamma\delta\gamma v$, $u\delta\gamma\delta v$.*

Proof. Let us refer to the words α, β, γ, and δ as 3-*palindromes*. Since `vtm` is square-free and by Lemma 1, any occurrence of α must be in the form $2\alpha 2$, any occurrence of β must be in the form $1\beta 1$, any occurrence of γ must be in the form $1\gamma 1$, and any occurrence of δ must be in the form $0\delta 0$. Hence, if a 3-palindrome is going to be followed by or preceded by a 3-palindrome, α must be preceded by (or followed by) β, β by α (Lemma 1), γ by δ (Lemma 1), and δ by γ. There cannot be four such 3-palindromes in a row, as then they would have to alternate between two, which would create a square. □

The next theorem plays a role in Section 3 for analyzing the abelian complexity of `vtm`.

Theorem 1. *Let ψ be the map $a \mapsto 012$, $b \mapsto 021$, $c \mapsto 120$, $d \mapsto 210$, $x \mapsto 102$, $y \mapsto 201$, $\alpha \mapsto 101$, $\beta \mapsto 202$, $\gamma \mapsto 020$, $\delta \mapsto 121$. Let φ be the morphism $a \mapsto ab$, $b \mapsto ax$, $c \mapsto \beta\alpha$, $d \mapsto dy$, $x \mapsto \gamma\delta$, $y \mapsto dc$, $\alpha \mapsto \gamma c$, $\beta \mapsto d\delta$, $\gamma \mapsto a\alpha$, $\delta \mapsto \beta x$. Then, $\psi\left(\varphi^{\omega}\left(a\right)\right) =$ `vtm`.*

Proof. In the context of `vtm`, we find that $a \mapsto ab$, $b \mapsto ax$, $c \mapsto ba$, $d \mapsto xa$, $x \mapsto \gamma\delta$, $y \mapsto \alpha\beta$, $\alpha \mapsto \gamma c2 = 0y\beta$, $\beta \mapsto \alpha 21 = 10\delta$, $\gamma \mapsto a\alpha 2 = 0\delta a$, and $\delta \mapsto b02 = 02x$. To see this, for example, $\psi(\alpha) = 101 \mapsto 0201202$ by the map $0 \mapsto 012, 1 \mapsto 02, 2 \mapsto 1$, but $0201202 = \psi(\gamma)\psi(c)2 = 0\psi(y)\psi(\beta)$. The mappings for a, b, and x are already what we want, and the mappings for α and γ begin with what we want. We now show that c, d, and y only occur after an odd number of Greek letters and that a, b, and x only occur after an even number of Greek letters. To do this, we show that this map avoids the following types of factors (sufficient by Lemma 2):

- u_1u_2 for $u_1 \in \{a, b, x\}$, $u_2 \in \{c, d, y\}$;
- u_1u_2 for $u_1 \in \{c, d, y\}$, $u_2 \in \{a, b, x\}$;
- $u_1u_2u_3$ for $u_1 \in \{a, b, x\}$, $u_2 \in \{\alpha, \beta, \gamma, \delta, \alpha\beta\alpha, \beta\alpha\beta, \gamma\delta\gamma, \delta\gamma\delta\}$, $u_3 \in \{a, b, x\}$;
- $u_1u_2u_3$ for $u_1 \in \{c, d, y\}$, $u_2 \in \{\alpha, \beta, \gamma, \delta, \alpha\beta\alpha, \beta\alpha\beta, \gamma\delta\gamma, \delta\gamma\delta\}$, $u_3 \in \{c, d, y\}$;
- $u_1u_2u_3$ for $u_1 \in \{a, b, x\}$, $u_2 \in \{\alpha\beta, \beta\alpha, \gamma\delta, \delta\gamma\}$, $u_3 \in \{c, d, y\}$;
- $u_1u_2u_3$ for $u_1 \in \{c, d, y\}$, $u_2 \in \{\alpha\beta, \beta\alpha, \gamma\delta, \delta\gamma\}$, $u_3 \in \{a, b, x\}$.

To cover the first two types of factors, we examine what cannot follow a, b, c, d, x, and y: the word a corresponds to 012, so it cannot be followed by a 12 or a 2. This excludes c, d, and y; the word b corresponds to 021, so it cannot be followed by a 1 or a 2. This excludes c, d, and y (and x). the word c corresponds to 120, so it cannot be followed by a 0, a 10, or a 20. This excludes a, b, and x (and y); the word d corresponds to 210, so it cannot be followed by a 0 or a 10. This excludes a, b, and x; the word x corresponds to 102, so it cannot be followed by a 02, a 12, or a 2. This excludes c, d, and y (and b); the word y corresponds to 201, so it cannot be followed by a 0 or a 1. This excludes a, b, and x (and c).

To cover the next two types of factors, we note that α must appear as $2\alpha 2$, β must appear as $1\beta 1$, γ must appear as $1\gamma 1$, and δ must appear as $0\delta 0$. The middle blocks u_2 here all start and end with the same Greek letter, so the left u_1 and right u_3 sides must end and start with the same ternary letters as each other. For α, the only possibilities are of the form $\{a, x\}\alpha\{d, y\}$; for β, only $y\beta x$; for γ, only $b\gamma c$; and for δ, only $\{c, d\}\delta\{a, b\}$. This excludes all of the given factors.

To cover the last two types of factors, we do a similar analysis and note that the only possibilities are $\{a, x\}\alpha\beta x$, $y\beta\alpha\{d, y\}$, $b\gamma\delta\{a, b\}$, and $\{c, d\}\delta\gamma c$. These avoid the factors in the last two cases.

Now, we examine what happens to a c, a d, or a y. We have established that these three letters occur in blocks beginning with α or γ and ending with β or δ. Applying the morphism to these blocks (including the end Greek letters) yields something from the α or γ followed by a 2 followed by some sequence of ba, xa, and $\alpha\beta$, and terminated by the morphism on the other greek letter. We notice that $2ba = \beta\alpha 2$, $2xa = dy2$, and $2\alpha\beta = dc2$. All of these cause the 2 to propagate through this piece, causing c to map to $\beta\alpha$, d to dy, and y to dc, all as required. Finally, by continuing to propagate the 2, we notice that δ must map to $202x = \beta x$, and β must map to $210\delta = d\delta$, both also as required. □

We also need the following technical lemma.

Lemma 3. *The word consisting of only the even positions in* vtm *is the Thue-Morse word over alphabet* $\{0, 2\}$. *That is, for each* i,

$$\mathtt{vtm}[2i] = 2\mathtt{tm}[i].$$

Furthermore, the odd positions of vtm *satisfy*

$$\mathtt{vtm}[2i+1] = (4 - \mathtt{vtm}[2i] - \mathtt{vtm}[2i+2])/2.$$

Proof. As mentioned in the introduction, an alternative definition of vtm is that its digits are the number of zeroes between consecutive ones in tm. Define ψ by $0 \mapsto 01, 1 \mapsto 10$. As a result of ψ's uniformity, tm can be thought of as a word over alphabet $\{0110, 1001\}$. Let A correspond to 0110, and let B correspond to 1001. The string 1001 is sent to 10010110 by ψ; the string 0110 is sent to 01101001. Hence, the morphism becomes the morphism ψ_{AB} over $\{A, B\}$, where $A \mapsto AB, B \mapsto BA$. Notice that this is the same morphism with A replacing 0 and B replacing 1. The even positions in vtm correspond to the number of zeroes between ones within each A or B. Each A contributes a 0; each B contributes a 2. Hence, the even positions of vtm form the Thue-Morse word over $\{0, 2\}$.

Let a be a symbol in an odd position of vtm. Considering the letters on either side of a, we obtain the following possible length 3 factors of vtm: $0a0$, $0a2$, $2a0$, and $2a2$. Now, $\texttt{tm}$ is obtained from $\texttt{vtm}$ by applying the map $0 \mapsto 011$, $1 \mapsto 01$, and $2 \mapsto 0$. Applying this map to the length 3 factors listed above and comparing the result to factors of $\texttt{tm}$, a simple case analysis shows that the only possible choices of length 3 factor are 020, 012, 210, and 202, which establishes the claim. □

3. Abelian Complexity of vtm

Our goal is to study the abelian complexity of $\texttt{vtm}$. As stated in the following lemma, if we remove all ones from $\texttt{vtm}$, we obtain a periodic sequence.

Lemma 4. *Let* $\tilde{\texttt{vtm}}$ *be* $\texttt{vtm}$ *with all of its ones removed. We have* $\tilde{\texttt{vtm}} = (02)^{\omega}$.

Proof. Obvious from the images of the morphism defining $\texttt{vtm}$. □

The following lemma tells us precisely which positions in $\texttt{vtm}$ are ones.

Lemma 5. $\texttt{vtm}\,[i] = 1$ *if and only if* $i \equiv 2^{2n-1} - 1 \left(\bmod 2^{2n}\right)$ *for some* $n \geq 1$.

Proof. By Lemma 3 we may restrict our attention to odd i, so write $i = 2j + 1$. By this same lemma, we have

$$\texttt{vtm}[i] = \texttt{vtm}[2j + 1] = (4 - \texttt{vtm}[2j] - \texttt{vtm}[2j + 2])/2.$$

This implies that $\texttt{vtm}[i] = 1$ if and only if $2 = \texttt{vtm}[2j] + \texttt{vtm}[2j + 2]$. Now,

$$\texttt{vtm}[2j] + \texttt{vtm}[2j + 2] = 2\texttt{tm}[j] + 2\texttt{tm}[j + 1],$$

so $2 = \texttt{vtm}[2j] + \texttt{vtm}[2j + 2]$ if and only if $1 = \texttt{tm}[j] + \texttt{tm}[j + 1]$. Thus, $\texttt{vtm}[i] = 1$ if and only if $\texttt{tm}[j] \neq \texttt{tm}[j + 1]$.

Recall that $\texttt{tm}[j]$ is the sum modulo 2 of the digits of the binary representation of j. It is easy to see that $\texttt{tm}[j] \neq \texttt{tm}[j + 1]$ if and only if the binary representation of j either equals 1^{2k} for some k or ends with 01^{2k} for some k. Since $i = 2j + 1$, this occurs if and only if the binary representation of i either equals 1^{2k+1} or ends with 01^{2k+1}. The result now follows. □

There are various definitions of $\texttt{pd}$ in the literature. The one we are choosing is the following.

Definition 1. *Let* $\texttt{pd}$ *be the result of changing all twos in* $\texttt{vtm}$ *to zeroes.*

Lemma 5 also tells us precisely which positions in $\texttt{pd}$ are ones. We exploit this fact in the following proposition.

Proposition 1. *Let* $v_0 = 0$*, and define a recursive sequence of words by* $v_{2n} = v_{2n-1} 0 v_{2n-1}$ *and* $v_{2n+1} = v_{2n} 1 v_{2n}$*. Then,* $\lim_{n\to\infty} v_n = \texttt{pd}$.

Proof. We show that the positions in pd which are ones are also ones in each of the v_ns. First, note that $|v_n| = 2^{n+1} - 1$. Next, notice that v_0 has its ones in the proper positions (trivially). Now, assume that v_n has its ones in the proper positions. We must consider two cases.

If n is even, then $v_{n+1} = v_n 1 v_n$. All of the symbols in v_n recur in the same positions modulo 2^{n+1}, so this part satisfies Lemma 5. The inserted one is at position $2^{n+1} - 1$ modulo 2^{n+2}, which, since $n + 1$ is odd, must be a 1. Therefore, v_{n+1} has its ones in the necessary positions, as required. On the other hand, if n is odd, then $v_{n+1} = v_n 0 v_n$. All of the symbols in v_n recur in the same positions modulo 2^{n+1}, so this part satisfies Lemma 5. The inserted zero is at position $2^{n+1} - 1$ modulo 2^{n+2}, which, since $n + 1$ is even, must not be a 1 (so it must be a 0). Therefore, v_{n+1} has its ones in the necessary positions. □

The following two lemmas are useful for our purposes.

Lemma 6. *Let u be a factor of* vtm*. Define $\tilde{u}$ to be the result of replacing all zeroes in u with twos and vice versa (while preserving its ones). Then $\tilde{u}$ is also a factor of* vtm*.*

Proof. Let ψ be the map $0 \mapsto 0, 1 \mapsto 1, 2 \mapsto 0$. Using the definition from Proposition 1, choose a positive integer m such that $\psi(u)$ is a factor of v_{2m}. Then, $v_{2m+1} = v_{2m} 1 v_{2m}$, so $\psi(u)$ occurs as a factor of the second v_{2m} at the same position within it as it occurs in the first v_{2m}. Every occurrence of $\psi(u)$ must be derived from either u or $\tilde{u}$. By the alternating property of 0 and 2 (Lemma 4) and the fact that the two copies of v_{2m} are joined by a 1, one of the occurrences of $\psi(u)$ must be derived from u and the other from $\tilde{u}$. Hence, $\tilde{u}$ occurs as a factor of vtm. □

Lemma 7. *Define the morphism ζ by $0 \mapsto 01, 1 \mapsto 00$. Then,* $\mathtt{pd} = \zeta^{\omega}(0)$.

Proof. The morphism ϕ defined by $a \mapsto ab$, $b \mapsto ac$, $c \mapsto \beta\alpha$, $\alpha \mapsto \beta c$, $\beta \mapsto a\alpha$ is obtained from the one in Theorem 1 by identifying a with d, b with y, c with x, α with δ, and β with γ. Notice that this essentially identifies 0 with 2 in vtm. Hence, if we consider the morphism ψ defined by $a \mapsto 010$, $b \mapsto 001$, $c \mapsto 100$, $\alpha \mapsto 101$, $\beta \mapsto 000$ and consider $\psi(\phi^{\omega}(a))$, we obtain pd. We now show what happens to *pairs* of $\{a, b, c, \alpha, \beta\}$ under ϕ. We only consider the pairs that occur in positions congruent to 0 modulo 2. To find these, we begin with ab and then consider additional ones as they occur on the right side of the paired morphism: $\phi(ab) = abac$, $\phi(ac) = ab\beta\alpha$, $\phi(\beta\alpha) = a\alpha\beta c$, $\phi(a\alpha) = ab\beta c$, and $\phi(\beta c) = a\alpha\beta\alpha$. We now examine this mapping with both sides expanded under ψ: $010001 \mapsto 010001010100$, $010100 \mapsto 010001000101$, $000101 \mapsto 010101000100$, $010101 \mapsto 010001000100$, $000100 \mapsto 010101000101$. Notice that the right side of each of these is ζ applied to the left side. This implies that $\zeta^{\omega}(0) = \mathtt{pd}$. □

We now prove some combinatorial properties of $f_m(n)$ (resp., $f_M(n)$), defined as the minimal (resp., maximal) number of ones in a factor of $\texttt{vtm}$ of length n. These properties will be used to derive the abelian complexity of $\texttt{vtm}$.

Lemma 8. *The following are true:*

1. *For all integers ℓ satisfying $f_m(n) \leq \ell \leq f_M(n)$, there exists a factor u_ℓ of $\texttt{vtm}$ satisfying $|u_\ell| = n$ such that u_ℓ contains exactly ℓ ones.*
2. $f_m(n+1) - f_m(n) \in \{0,1\}$.
3. $f_M(n+1) - f_M(n) \in \{0,1\}$.
4. $f_m(2n) = n - f_M(n)$.
5. $f_M(2n) = n - f_m(n)$.
6. $f_m(4n) = n + f_m(n)$.
7. $f_M(4n) = n + f_M(n)$.
8. $f_m(4n-1) = f_m(4n) - 1 = n + f_m(n) - 1$.
9. $f_M(4n-1) = f_M(4n) = n + f_M(n)$.
10. $f_m(4n+1) = f_m(4n) = n + f_m(n)$.
11. $f_M(4n+1) = f_m(4n) + 1 = n + f_M(n) + 1$.

Proof. Let ψ be the map $0 \mapsto 0, 1 \mapsto 1, 2 \mapsto 0$. Also, let $\texttt{pd} = \psi(\texttt{vtm})$ and let morphism ζ be defined as in Lemma 7.

For Statement 1, let $u_\ell = \texttt{vtm}[i] \cdots \texttt{vtm}[i+n-1]$ be a factor of $\texttt{vtm}$ of length n containing exactly ℓ ones (and, without loss of generality, assume $i > 0$, which can be done because every prefix of $\texttt{vtm}$ recurs later). Let $u'_\ell = \texttt{vtm}[i-1] \cdots \texttt{vtm}[i+n-2]$ and $u''_\ell = \texttt{vtm}[i+1] \cdots \texttt{vtm}[i+n]$. Since u'_ℓ and u''_ℓ each share $n-1$ symbols with u_ℓ, the number of ones in each of them can differ from ℓ by at most 1. Hence, given positions i_m and i_M such that $\texttt{vtm}[i_m] \cdots \texttt{vtm}[i_m+n-1]$ contains $f_m(n)$ ones and $\texttt{vtm}[i_M] \cdots \texttt{vtm}[i_M+n-1]$ contains $f_M(n)$ ones, the factors of length n beginning at positions between i_m and i_M contain all numbers of ones between $f_m(n)$ and $f_M(n)$.

For Statement 2, $f_m(n+1) \geq f_m(n)$ because every factor of length $n+1$ contains a factor of length n. Also, $f_m(n+1) - f_m(n) < 2$ because appending one symbol to a length n factor cannot make the number of ones grow by more than 1.

Statement 3 follows from similar reasoning to that used in Statement 2.

For Statement 4, first, let u be a factor of $\texttt{vtm}$ of length n containing exactly ℓ ones. Then, $\psi(u)$ is a factor of $\texttt{pd}$ of length n containing exactly ℓ ones. Then,

$\zeta(\psi(u))$ is a factor of pd of length $2n$ containing exactly $n-\ell$ ones. Every factor of pd of length $2n$ beginning at an even position must come from ζ of some length n factor of pd, so taking $\ell = f_M(n)$ yields a factor of length $2n$ with the minimal number of ones that can begin at an even position. Now, assume for a contradiction that there is a factor of pd of length $2n$ with fewer ones than $n - f_M(n)$. Let i be the position of this factor in pd. We know that i is odd and that ones only occur in odd positions in pd, so the last digit of $\texttt{pd}[i]\cdots\texttt{pd}[i+2n-1]$ must be a 0. Also, $\texttt{pd}[i-1]\cdots\texttt{pd}[i+2n-2]$ must begin with a 0. So, these factors contain the same number of ones, so $\texttt{pd}[i-1]\cdots\texttt{pd}[i+2n-2]$ contains fewer than $n-f_M(n)$ ones and begins at an even position, a contradiction.

Statement 5 follows from similar reasoning to that used in Statement 4. Statement 6 follows from composing Statement 4 with itself, while Statement 7 follows from composing Statement 5 with itself.

The second equality of Statement 8 follows from Statement 6. We now prove the equality $f_m(4n-1) = n + f_m(n) - 1$. Consider the morphism ζ^2 defined by $0 \mapsto 0100, 1 \mapsto 0101$. Let u be a factor of vtm of length n containing exactly ℓ ones. Then, $\psi(u)$ is a factor of pd of length n containing exactly ℓ ones. Then, $\zeta^2(\psi(u))$ is a factor of pd of length $4n$ containing exactly $n+\ell$ ones. Also, $\zeta^2(\psi(u))$ has a 1 in position 1, so the factor of pd of length $4n-1$ beginning two positions later contains exactly $n+\ell-1$ ones (since its last symbol must be a 0). This quantity is minimized by $\ell = f_m(n)$. Hence, the minimal number of ones in a factor of pd of length $4n-1$ beginning at a position congruent to 2 modulo 4 is $n + f_m(n) - 1$. By Statement 2, this must be the overall minimum.

The second equality of Statement 9 follows from Statement 7. We now prove the first. Let u be a factor of vtm of length n containing exactly ℓ ones. Then, $\psi(u)$ is a factor of pd of length n containing exactly ℓ ones. Then, $\zeta^2(\psi(u))$ is a factor of pd of length $4n$ containing exactly $n+\ell$ ones. Also, $\zeta^2(\psi(u))$ begins with a 0, so the factor of pd of length $4n-1$ beginning at the next position contains exactly $n+\ell$ ones. This quantity is maximized by $\ell = f_M(n)$. Hence, the maximal number of ones in a factor of pd of length $4n-1$ beginning at a position congruent to 1 modulo 4 is $n + f_M(n)$. By 3, this must be the overall maximum, as required.

The second equality of Statement 10 follows from Statement 6. We now prove the first. Let u be a factor of vtm of length n containing exactly ℓ ones. Then, $\psi(u)$ is a factor of pd of length n containing exactly ℓ ones. Then, $\zeta^2(\psi(u))$ is a factor of pd of length $4n$ containing exactly $n+\ell$ ones. Also, $\zeta^2(\psi(u))$ the factor of pd of length $4n+1$ beginning at that same position contains exactly $n+\ell$ ones (as the new final position is an even position of pd, and, hence, a 0). This quantity is minimized by $\ell = f_m(n)$. Hence, the minimal number of ones in a factor of pd of length $4n+1$ beginning at a position congruent to 0 modulo 4 is $n + f_m(n)$. By Statement 2, this must be the overall minimum, as required.

The second equality of Statement 11 follows from Statement 7. We now prove the first. Let u be a factor of $\texttt{vtm}$ of length n containing exactly ℓ ones. Then, $\psi(u)$ is a factor of $\texttt{pd}$ of length n containing exactly ℓ ones. Then, $\zeta^2(\psi(u))$ is a factor of $\texttt{pd}$ of length $4n$ containing exactly $n+\ell$ ones. Also, $\zeta^2(\psi(u))$ begins with 01, so the factor of $\texttt{pd}$ of length $4n+1$ beginning at the next position contains exactly $n+\ell+1$ ones (as the two positions after $\zeta^2(\psi(u))$ in $\texttt{pd}$ must be 01). This quantity is maximized by $\ell = f_M(n)$. Hence, the maximal number of ones in a factor of $\texttt{pd}$ of length $4n+1$ beginning at a position congruent to 1 modulo 4 is $n + f_M(n) + 1$. By Statement 3, this must be the overall maximum, as required. □

Properties 2, 3, and 4 of the following proposition allow for complete calculation of $\rho^{ab}_{\texttt{pd}}$ in logarithmic time. Similar relations were independently obtained by Karhumäki, Saarela, and Zamboni [13].

Proposition 2. *The following hold for all positive integers n:*

1. $\rho^{ab}_{\texttt{pd}}(n) = f_M(n) - f_m(n) + 1$.
2. $\rho^{ab}_{\texttt{pd}}(2n) = \rho^{ab}_{\texttt{pd}}(n)$.
3. $\rho^{ab}_{\texttt{pd}}(4n-1) = \rho^{ab}_{\texttt{pd}}(n) + 1$.
4. $\rho^{ab}_{\texttt{pd}}(4n+1) = \rho^{ab}_{\texttt{pd}}(n) + 1$.

Proof. Note that Statement 1 (resp., 2, 3, 4) follows from Lemma 8(1) (resp., Lemma 8(1,4,5), (1,8,9), (1,10,11)), as the number of ones determines the Parikh vector (in the binary case) and all intermediate numbers of ones are possible. □

This proposition also implies that the abelian complexity function of $\texttt{pd}$ is 2-regular; see [15] for definitions and for a similar result concerning the paperfolding word.

The following theorem helps us derive the abelian complexity of $\texttt{vtm}$, stated as a corollary.

Theorem 2. *The following hold for all positive integers n:*

1. $\rho^{ab}_{\texttt{vtm}}(n) = \frac{3}{2}(f_M(n) - f_m(n) + 1)$ *if* $f_m(n) + f_M(n)$ *is odd.*
2. $\rho^{ab}_{\texttt{vtm}}(n) = \frac{3}{2}(f_M(n) - f_m(n) + 2)$ *if* $f_m(n) + f_M(n)$ *is even and* $n + f_m(n)$ *is odd.*
3. $\rho^{ab}_{\texttt{vtm}}(n) = \frac{3}{2}(f_M(n) - f_m(n))$ *if* $f_m(n) + f_M(n)$ *is even and* $n + f_m(n)$ *is even.*
4. $\rho^{ab}_{\texttt{vtm}}(2n) = \rho^{ab}_{\texttt{vtm}}(n)$.

Proof. We first prove Statements 1, 2 and 3. By Lemma 8(1), the number of ones in factors of length n ranges over all values from $f_m(n)$ to $f_M(n)$. By Lemma 4, the number of zeroes and twos can differ by at most 1. By Lemma 6, when the number of zeroes and twos differ by 1 for a given number of ones, both permissible Parikh vectors occur. Hence, each value ℓ for the number of ones in a factor of length n such that $n-\ell$ is odd contributes two Parikh vectors, and each value ℓ for the number of ones in a factor of length n such that $n-\ell$ is even contributes one Parikh vector.

In case 1, there are an even number of possibilities for ℓ, half of which leave $n-\ell$ even and half of which leave $n-\ell$ odd. Hence, the first formula holds. In case 2, there are an odd number of possibilities for ℓ, one more of which leave $n-\ell$ odd than leave $n-\ell$ even. Hence, as we must account for one additional Parikh vector, the second formula holds. In case 3, there are an odd number of possibilities for ℓ, one more of which leave $n-\ell$ even than leave $n-\ell$ odd. Hence, as we must account for one fewer Parikh vector, the third formula holds.

We now prove Statement 4. Since $f_m(2n) = n - f_M(n)$ and $f_M(2n) = n - f_m(n)$, we obtain the equation $f_m(2n) + f_M(2n) = 2n - (f_m(n) + f_M(n))$. Next, $f_M(2n) - f_m(2n) = n - f_m(n) - (n - f_M(n)) = f_M(n) - f_m(n)$. Hence, if we can show that we always remain in the same case of 1, 2, or 3 when doubling n, we have proved the desired result.

When $f_m(n)+f_M(n)$ is odd, we begin and remain in case 1 since $f_m(2n)+f_M(2n)$ is odd, as required. When $f_m(n)$ is even, $f_M(n)$ is even, and n is even, we begin in case 3. Also, $2n + f_m(2n) = 2n + n - f_M(n) = 3n - f_M(n)$ is even, so we remain in case 3, as required. When $f_m(n)$ is even, $f_M(n)$ is even, and n is odd, we begin in case 2. Also, $2n+f_m(2n) = 2n+n-f_M(n) = 3n-f_M(n)$ is odd, so we remain in case 2, as required. When $f_m(n)$ is odd, $f_M(n)$ is odd, and n is even, we begin in case 2. Also, $2n+f_m(2n) = 2n+n-f_M(n) = 3n-f_M(n)$ is odd, so we remain in case 2, as required. Finally when $f_m(n)$ is odd, $f_M(n)$ is odd, and n odd, we begin in case 3. Also, $2n+f_m(2n) = 2n+n-f_M(n) = 3n-f_M(n)$ is even, so we remain in case 3, as required. □

Corollary 1. *The abelian complexity of* vtm *is* $O(\log n)$ *with constant approaching* $\frac{3}{4}$ *(assuming base* 2 *logarithm), and it is* $\Omega(1)$ *with constant* 3.

Proof. Theorem 2(1,2,3) along with Proposition 2(1) imply that the inequality $\left|\rho^{ab}_{\texttt{vtm}}(n) - \frac{3}{2}\rho^{ab}_{\texttt{pd}}(n)\right| \le \frac{3}{2}$ holds. Hence, we prove an upper bound for $\rho^{ab}_{\texttt{pd}}(n)$ and then we multiply it by $\frac{3}{2}$ to obtain an upper bound for $\rho^{ab}_{\texttt{vtm}}(n)$.

Let m be an integer greater than 1. By Proposition 2, the first time that $\rho^{ab}_{\texttt{pd}}(n) = m$ is a_m in the sequence $a_2 = 1$, and $a_n = 4a_{n-1} - 1$ for $n \ge 3$. The solution to this recurrence is $a_{n+2} = \frac{2\cdot 4^n+1}{3}$. So $\frac{3}{2}m = \rho^{ab}_{\texttt{vtm}}(4^m+1) = \rho^{ab}_{\texttt{vtm}}\left(2^{2m}+1\right)$. Taking logs (and ignoring additive constants and renaming m to n) yields that the largest values taken by $\rho^{ab}_{\mathtt{vtm}}(n)$ grow asymptotically like $\frac{1}{2}\log n$. Multiplying by $\frac{3}{2}$ yields the big-O bound of $\log n$ with constant approaching $\frac{3}{4}$ for $\rho^{ab}_{\mathtt{vtm}}(n)$, as required.

For the lower bound, by Proposition 2(2,3,4), $\rho^{ab}_{\mathtt{pd}}(n) = 2$ for n a power of 2, so $\rho^{ab}_{\mathtt{vtm}}(n) = 3$ for n a power of 2 (that is, infinitely often). The value 3 is minimal. Therefore, 3 is the best possible lower bound. □

Note that the above corollary gives the best possible bounds.

4. Factor Indices in vtm

We prove two results regarding factor indices in vtm. The first one states that if u is a factor of vtm and i, m are positive integers, then there is an occurrence of u in vtm beginning at a position congruent to i modulo $(2m + 1)$ (Section 4.1). The second one refers to Lemma 6 and states that if u is a factor of vtm and $\tilde{u}$ is the result of replacing all zeroes in u with twos and vice versa (while preserving its ones), then u occurs in vtm beginning at a position congruent to i modulo m if and only if $\tilde{u}$ occurs in vtm beginning at a position congruent to i modulo m (Section 4.2).

We begin with some preliminaries. If pu is a prefix of word v, we say that u *appears in* v *with index* $|p|$. More formally, if $v = puw$, we refer to the triple $\langle p, u, w\rangle$ as an *occurrence of* u *in* v *of index* $|p|$.

The ith letter of tm (starting the count with 0) is obtained as the modulo 2 sum of the binary digits of i; thus the binary representation of 5 is 101, so that the 5th letter of tm is $1 + 0 + 1 = 0 \pmod 2$.

The words tm and vtm are related by the morphism $h : \{0, 1, 2\}^* \to \{0, 1\}^*$ given by

$$h(0) = 011; \qquad h(1) = 01; \qquad h(2) = 0,$$

in that $h(\mathtt{vtm}) = \mathtt{tm}$. The word tm can be uniquely parsed in terms of 011, 01, and 0, so by "desubstitution" one can thus alternatively define vtm as $h^{-1}(\mathtt{tm})$.

Recall from Lemma 3 that for each i,

$$\mathtt{vtm}_{2i} = 2\mathtt{tm}_i,$$

while

$$\mathtt{vtm}_{2i+1} = (4 - \mathtt{vtm}_{2i} - \mathtt{vtm}_{2i+2})/2.$$

This means that the even index letters of vtm totally determine the odd index letters.

4.1. First Result on Indices

We begin with the following observation.

Observation 1. *If u is a prefix of* vtm*, then $h(u)$ is a prefix of* tm*, and $|h(u)|_0 = |u|$.*

Lemma 9. *If $u2$ is a prefix of* vtm*, then $h(u)00$ is a prefix of* tm*, and $|h(u)| = 2|u| + 1$.*

Proof. Suppose that $u2a$ is a prefix of vtm where $a \in \{0, 1, 2\}$. Then $h(u2a)$ is a prefix of tm by the observation. However, $h(u2a) = h(u)001^{2-a}$, which has $h(u)00$ as a prefix. Since $\texttt{tm} \in \{01, 10\}^*$, the factor 00 of tm only ever appears in tm with odd index. We therefore deduce that $h(u)0 \in \{01, 10\}^*$ whence $|h(u)0|_0 = |h(u)0|/2$. Therefore, $|h(u)| = |h(u)0| - 1 = 2|h(u)0|_0 - 1 = 2|u2| - 1 = 2|u| + 1$. □

Lemma 10. *If $u0$ is a prefix of* vtm*, then $|h(u)| = 2|u|$.*

Proof. Suppose that $u0$ is a prefix of tm. Then $h(u0) = h(u)011$ is a prefix of tm by the observation. The factor 11 of tm only ever appears with odd index. We therefore deduce that $h(u) \in \{01, 10\}^*$ whence $|u| = |h(u)|_0 = |h(u)|/2$. □

Lemma 11. *Let u be a factor of* vtm*, $u \neq 1$. Let m be an odd number. Let S be the set of indices at which u appears in* vtm*. Let T be the set of indices at which $h(u)0$ appears in* tm*. Then S contains a representative of every congruence class modulo m if and only if T contains a representative of every congruence class modulo m.*

Proof. Suppose the first letter of u is 2. By Lemma 9, u will occur in vtm with index i if and only if $h(u)0$ appears in tm with index $2i + 1$. Since 2 is relatively prime to m, the map $\phi : \mathbb{Z}_m \to \mathbb{Z}_m$ given by $i \mapsto 2i+1$ is a bijection, and the result follows. A similar proof applies if the first letter of u is 0.

Consider the case when the first letter of u is a 1. If u commences 12, it follows from Lemma 9 that u occurs in vtm with index i if and only if $h(u)0$ occurs in tm with index $2i + 1$; an index of $h(1^{-1}u)0$ in tm is $2(i + 1) + 1$, giving $01h(1^{-1}u)0 = h(u)0$ with index $2i + 1$. Similarly, if u commences 10, then u occurs in vtm with index i if and only if $h(u)0$ occurs in tm with index $2i$. Thus if u commences 12 or 10, the proof of the previous paragraph is adapted to establish our result. □

Lemma 12. *Let u be a factor of* tm*, i an integer and m an odd integer. There exists an occurrence of u in* tm *whose index is congruent to i modulo m.*

Proof. First we show that 0 occurs with index $i \bmod m$. In fact, this is a consequence of a deep result of Gelfond [11, Théorème I], but for completeness we give a simple proof of this weaker claim. Since 2 is relatively prime to m, choose positive integer e such that $2^e \equiv 1 (\mathrm{mod}\ m)$. Construct the sequence of integers

$$1, 2^e + 1, 2^e(2^e + 1) + 1, 2^e(2^e(2^e + 1) + 1) + 1, \ldots$$

where each integer is obtained from the previous by multiplying by 2^e and adding 1. Modulo m then, each element of the sequence is one greater than the previous.

We continue multiplying by 2^e and adding 1 until we get a number n congruent to $i \pmod m$. The binary representation of n will contain i 1s. If i is even, we are done: $\mathtt{tm}_n = 0$. If i is odd, multiply by 2^e and add 1 an additional m times to get a new number n'. The binary representation of n' now has an even number $(i+m)$ of 1s and is still congruent to $i \pmod m$. However, $\mathtt{tm}_{n'} = 0$, as desired.

Choose a positive integer k such that u is a factor of v, the prefix of $\mathtt{tm}$ of length 2^k. Since 0 occurs with indices all values $i \pmod m$, v occurs at all positions $i2^k \pmod m$. However, as i runs through all residues modulo m, so does $i2^k$. Thus we can find v with index congruent to any $i \pmod m$ and the same is true for u. □

Theorem 3. *Let u be a factor of* $\mathtt{vtm}$ *and m an odd number. The set of indices of u in* $\mathtt{vtm}$ *contains a representative of every congruence class modulo m.*

Proof. By Lemma 11 and Lemma 12, the result is true when $u \neq 1$. However, then the set of indices of 01 in $\mathtt{vtm}$ takes on all values modulo m, implying that the factor $u = 1$ of 01 does also. □

4.2. Second Result on Indices

The operation of replacing each 0 in a factor u of $\mathtt{vtm}$ by 2 and vice versa we call *2-complementation*, and the result $\tilde{u}$ is called the *2-complement* of u.

Let pu be a prefix of $\mathtt{vtm}$ such that $|p|$ is even and $|pu|$ is odd. Fix an integer m, and write $m = 2^s r$ where r is odd. Now choose $k \geq s$ so that $2^k > |pu|$. Write $\mathtt{vtm} = v_0 v_1 v_2 \cdots$ where each v_i has length 2^k. The even index letters of v_0 are obtained by multiplying each letter of the length 2^{k-1} prefix of $\mathtt{tm}$ by 2, and these even index letters determine the odd index letters. If we write $\mathtt{tm} = u_0 u_1 u_2 \ldots$ where each u_i has length 2^{k-1}, it is well-known that each u_i is either u_0 or the binary complement of u_0. It follows that the even index letters of each v_i are either the same as those in v_0, or the 2-complement.

Note that the rule

$$\mathtt{vtm}_{2i+1} = (4 - \mathtt{vtm}_{2i} - \mathtt{vtm}_{2i+2})/2$$

commutes with 2-complementation:

$\mathtt{vtm}_{2i}$	$\mathtt{vtm}_{2i+2}$	$\mathtt{vtm}_{2i+1}$
0	0	2
0	2	1
2	0	1
2	2	0

This implies that each v_i is either v_0 or its 2-complement; if $v_i = v_0$, then u appears in v_i with index $|p|$; otherwise, $\tilde{u}$ appears in v_i with index $|p|$. Consider the sequence of words $\{v_{ri}\}_{i=0}^{\infty}$. Each of these words contains either u or $\tilde{u}$ at index

$|p|$. These occurrences of u or $\tilde{u}$ *in* vtm occur at indices differing by $r2^k$, which is a multiple of m. If, in fact, none of v_{ri} contains $\tilde{u}$ at index $|p|$, then the v_{ri} are all equal to v_0. This implies that $\mathtt{tm}_{ir2^k} = 0$ for all i. Given the characterization of tm in terms of binary representations, we have $\mathtt{tm}_{ir} = 0$ for all i. This is a contradiction—for instance, it contradicts the result of Gelfond mentioned in the proof of Lemma 12. (For a simple, direct proof that $\mathtt{tm}_{ir}$ cannot equal 0 for all i, see [16].) We have therefore established the following result.

Theorem 4. *If u is a factor of* vtm *and m is a positive integer then the set of indices of u modulo m is the same as the set of the indices of $\tilde{u}$, where $\tilde{u}$ is obtained from u by replacing 0s with 2s and vice versa.*

Acknowledgement We would like to thank J. Paykin and C. Stevens for valuable comments and suggestions. We also thank the referees of preliminary versions of this paper for their very valuable comments and suggestions.

References

[1] J.-P. Allouche and J. Shallit. The ubiquitous Prouhet-Thue-Morse sequence. In C. Ding, T. Helleseth, and H. Niederreiter, editors, *SETA 1998, Sequences and their Applications*, pages 1–16. Springer, 1998.

[2] J.-P. Allouche and J. Shallit. *Automatic Sequences: Theory, Applications, Generalizations*. Cambridge University Press, 2003.

[3] J. Berstel, A. Lauve, C. Reutenauer, and F. Saliola. *Combinatorics on Words: Christoffel Words and Repetitions in Words*. American Mathematical Society, 2008.

[4] V. Berthé and M. Rigo, editors. *Combinatorics, Automata and Number Theory*. Number 135 in Encyclopedia of Mathematics and its Applications. Cambridge University Press, 2010.

[5] S. Brlek. Enumeration of factors in the Thue-Morse word. *Discrete Appl. Math.*, 24:83–96, 1989.

[6] J. Cassaigne. Complexité et facteurs spéciaux. *Bull. Belg. Math. Soc.*, 4:67–88, 1997.

[7] A. de Luca and S. Varricchio. On the factors of the Thue-Morse word on three symbols. *Inform. Process. Lett.*, 27:281–285, 1988.

[8] A. de Luca and S. Varricchio. Some combinatorial properties of the Thue-Morse sequence and a problem in semigroups. *Theoret. Comput. Sci.*, 63:333–348, 1989.

[9] A. E. Frid. The subword complexity of fixed points of binary uniform morphisms. In *FCT 1997, 11th International Symposium on Fundamentals of Computation Theory*, volume 1279 of *Lecture Notes in Computer Science*, pages 179–187, Berlin, Heidelberg, 1997. Springer-Verlag.

[10] A. E. Frid and S. V. Avgustinovich. On bispecial words and subword complexity of D0L sequences. In C. Ding, T. Helleseth, and H. Niederreiter, editors, *SETA 1998, Sequences and their Applications*, pages 191–204. Springer, 1999.

[11] A. O. Gelfond. Sur les nombres qui ont des propriétés additives et multiplicatives données. *Acta Arith.*, 13:259–265, 1968.

[12] T. Harju. Square-free walks on labelled graphs. *CoRR*, abs/1106.4106, 2011.

[13] J. Karhumäki, A. Saarela, and L. Q. Zamboni. Variations of the Morse–Hedlund Theorem for k-abelian equivalence. `http://arxiv.org/abs/1302.3783`, 2013.

[14] K. Klouda. Bispecial factors in circular non-pushy D0L languages. *Theoret. Comput. Sci.*, 445:63–74, 2012.

[15] B. Madill and N. Rampersad. The abelian complexity of the paperfolding word. *Discrete Math.*, 313:831–838, 2013.

[16] J. F. Morgenbesser, J. Shallit, and T. Stoll. Thue–Morse at multiples of an integer. *J. Number Theory*, 131:1498–1512, 2011.

[17] P. Ochem. Binary words avoiding the pattern *aabbcabba*. *RAIRO Theor. Inform. Appl.*, 44:151–158, 2010.

[18] M. Rao, M. Rigo, and P. Salimov. Avoiding 2-binomial squares and cubes. `http://arxiv.org/abs/1310.4743`, 2013.

[19] G. Richomme, K. Saari, and L. Q. Zamboni. Abelian complexity in minimal subshifts. *J. Lond. Math. Soc.*, 83:79–95, 2011.

[20] J. Tromp and J. Shallit. Subword complexity of a generalized Thue-Morse word. *Inform. Process. Lett.*, 54:313–316, 1995.