The Avoidability of Cubes under Permutations

Florin Manea, Mike Müller, Dirk Nowotka

Christian-Albrechts-Universität zu Kiel, Institut für Informatik,
D-24098 Kiel, Germany.
{flm,mimu,dn}@informatik.uni-kiel.de

Abstract. In this paper we consider the avoidance of patterns in infinite words. Generalising the traditional problem setting, functional dependencies between pattern variables are allowed here, in particular, patterns involving permutations. One of the remarkable facts is that in this setting the notion of avoidability index (the smallest alphabet size for which a pattern is avoidable) is meaningless since a pattern with permutations that is avoidable in one alphabet can be unavoidable in a larger alphabet. We characterise the (un-)avoidability of all patterns of the form $\pi^i(x)\pi^j(x)\pi^k(x)$, called cubes under permutations here, for all alphabet sizes in both the morphic and antimorphic case.

1 Introduction

The avoidability of patterns in infinite words is an old area of interest with a first systematic study going back to Thue [1,2]. This field includes discoveries and studies by many authors over the last one hundred years; see for example [3] and [4] for surveys. In this article, we are concerned with a generalisation of the theme by considering patterns with functional dependencies between variables, in particular, we investigate permutations. More precisely, we do allow function variables in the pattern that are either morphic or antimorphic extensions of permutations on the alphabet. Consider the following pattern for example:

$$x\,\pi(x)\,x$$

where an instance of the pattern is a word uvu that consists of three parts of equal length, that is, $|u| = |v|$, and v is the image of (the reversal of) u under any permutation on the alphabet. For example, $aab|bba|aab$ ($aab|abb|aab$) is an instance of $x\pi(x)x$ for the morphic (respectively, antimorphic) extension of permutation $a \mapsto b$ and $b \mapsto a$.

Recently, there has been some initial work on avoidance of patterns with involutions which is a special case of the permutation setting considered in this paper (as involutions are permutations of order at most two); see [5,6,7]. The original interest of investigating patterns under involution was motivated by possible applications in biology where the Watson-Crick complement corresponds to an antimorphic involution over four letters. Our considerations here are much more general, however, and the relation to direct applications in microbiology are admittedly scant.

Since these are the very first considerations on this kind of pattern avoidance at all, we restrict ourselves to cube-like patterns. The cube xxx is the most basic and well-investigated pattern that lends itself to nontrivial considerations on patterns with functional dependencies (a square would hardly be interesting in that context). So, we have one variable, occurring three times, and only one function variable, that is, we investigate patterns of the form:

$$\pi^i(x)\,\pi^j(x)\,\pi^k(x)$$

where $i, j, k \geq 0$.

It is worth noting that the notion of avoidability index plays no role in the setting of patterns involving permutations. Contrary to the traditional setting, where once a pattern is avoidable for some alphabet size it remains avoidable in larger alphabets, a pattern with permutations can become unavoidable in a larger alphabet. This is a new and somewhat unexpected phenomenon in the field of pattern avoidance. It does not occur, for example, in the involution setting but requires permutations of higher order.

2 Preliminaries

We define $\Sigma_k = \{0, \ldots, k-1\}$ to be an alphabet with k letters. For words u and w, we say that u is a prefix (resp. suffix) of w, if there exists a word v such that $w = uv$ (resp. $w = vu$). We denote that by $u \leq_p w$ (resp. $u \leq_s w$).

For a word w and an integer i with $1 \leq i \leq |w|$ we denote the i-th letter of w by $w[i]$. We also denote the factor that starts with the i-th letter and ends with the j-th letter in w by $w[i..j]$. If w is a word of length n then w^R, the reversal of w, is defined as the word $w[n]w[n-1]\ldots w[1]$.

If $f : \Sigma_k \to \Sigma_k$ is a permutation, we say that the order of f, denoted $\mathbf{ord}(f)$, is the minimum value $m > 0$ such that f^m is the identity. If $a \in \Sigma_k$ is a letter, the order of a with respect to f, denoted $\mathbf{ord}_f(a)$, is the minimum number m such that $f^m(a) = a$.

A pattern which involves functional dependencies is a term over (word) variables and function variables (where concatenation is an implicit functional constant). For example, $x\pi(y)\pi(\pi(x))y$ is a pattern involving the variables x and y and the function variable π. An instance of a pattern p in Σ_k is the result of substituting every variable by a word in Σ_k^* and every function variable by a function over Σ_k^*. A pattern is avoidable in Σ_k if there is an infinite word over Σ_k that does not contain any instance of the pattern.

In this paper, we consider patterns with anti-/morphic permutations, that is, all function variables are unary and are substituted by anti-/morphic permutations only. A more formal definition of patterns and avoidability under permutations is given in the Appendix, together with several conventions we make for our notations.

The infinite Thue-Morse word t is defined as

$$t = \lim_{n\to\infty} \phi_t^n(0),$$

for $\phi_t : \Sigma_2^* \to \Sigma_2^*$ where $\phi_t(0) = 01$ and $\phi_t(1) = 10$. The word t avoids the patterns xxx (cubes) and $xyxyx$ (overlaps).

Let h be the infinite word defined as

$$h = \lim_{n\to\infty} \phi_h^n(0),$$

where $\phi_h : \Sigma_3^* \to \Sigma_3^*$ is a morphism due to Hall [8], defined by $\phi_h(0) = 012$, $\phi_h(1) = 02$ and $\phi_h(2) = 1$. The infinite word h avoids the pattern xx (squares).

The reader is referred to [9] for further details on the concepts discussed in this paper. Finally, note that some proofs are omitted due to space limitations.

3 The morphic case

In this section, the function variable π is always substituted by a morphic permutation.

We begin this section by showing the avoidability of a series of basic patterns. These results are then be used to show the avoidability of more general patterns. Our first result uses the morphism $\alpha : \Sigma_2^* \to \Sigma_3^*$ that is defined by

$$0 \mapsto 02110, \qquad\qquad 1 \mapsto 02210.$$

Lemma 1. *The infinite word $t_\alpha = \alpha(t)$ avoids the patterns xxx and $x\pi(x)x$ in Σ_m, for all $m \geq 3$. These patterns cannot be simultaneously avoided by words over smaller alphabets.*

The following lemma is the main tool that we use to analyse the avoidability of cubes under morphic permutations. To obtain this result we apply the morphism $\beta : \Sigma_2^* \to \Sigma_4^*$ defined by

$$0 \mapsto 012013213, \qquad\qquad 1 \mapsto 012031023.$$

Lemma 2. *Let $t_\beta = \beta(t)$ for the morphism β defined before and let $i, j \in \mathbb{N}$ and f, g be morphic permutations of Σ_m with $m \geq 4$. The word t_β obtained as such does not contain any factor of the form $uf(u)g(u)$ for any $u \in \Sigma_4^*$ with $|u| \geq 7$. Furthermore, t_β does not contain any factor of the form $uf^i(u)f^j(u)$ with*

$$\left|\left\{u[\ell], f^i(u)[\ell], f^j(u)[\ell]\right\}\right| \leq 2,$$

for all $\ell \leq |u|$ and $|u| \leq 6$.

Proof. We begin with addressing the first claim. One can easily show that t_β contains no cube. For $|u| \in \{7, 8\}$, the length of $uf(u)g(u)$ is 21 or 24 and so it is completely contained in $\beta(v)$ for some factor v of the Thue-Morse word with $|v| = 4$. Thus, it is sufficient to check that there is no factor of the form $uf(u)g(u)$ in the image of the set of factors of length 4 of the Thue-Morse word. We did this using a computer program[1].

[1] Implementations of all programs mentioned in this paper can be found at the webpage `http://www.informatik.uni-kiel.de/zs/taocup`.

For $|u| \geq 9$ we have that at least one of the factors $u, f(u), g(u)$ has 3 occurrences of the letter 1. Indeed, any factor $uf(u)g(u)$ of t_β, having length greater than or equal to 27, has a factor $x\beta(s_1s_2)y$ where $s_1, s_2 \in \{0,1\}$ and $|xy| = 9$. Clearly, x is a suffix of $\beta(s_3)$ and y is a prefix of $\beta(s_4)$ for some letters s_3 and s_4 from $\{0,1\}$. Now, regardless of the way we choose the letters s_1, s_2, s_3 and s_4 from $\{0,1\}$, such that $s_1s_2s_3s_4$ is a factor of t, we obtain that any factor of length 27 of $\beta(s_1s_2s_3s_4)$ contains at least 7 occurrences of the letter 1. By the pigeonhole principle, it follows that at least one of the factors $u, f(u), g(u)$ has 3 occurrences of the letter 1. In fact, this factor contains one of the words $w_1 = 1201321$, $w_2 = 1321301$, $w_3 = 1301201$, or $w_4 = 13012031$. Also, denote $y_1 = 0120310$, $y_2 = 0310230$, and $y_3 = 0230120$. Let us assume first that u contains three occurrences of the letter 1, and assume that $u[i..i+\ell]$, with $\ell \in \{6,7\}$, is the leftmost subfactor of u that contains three 1-letters and begins with 1. But this means that also $f(u)[i..i+\ell]$ and $g(u)[i..i+\ell]$ contain three identical letters. It is rather easy to note that, whenever $w_i \leq_p u[i..i+\ell]$ for $i \in \{2,3,4\}$, then the only possibility is that also $w_i \leq_p f(u)[i..i+\ell]$ and $w_i \leq_p g(u)[i..i+\ell]$; otherwise, f and g would map the same letter in two different ways, a contradiction. However, in that case, f and g would be the identical mappings, which means that t_β would contain a cube, again a contradiction.

So, the only possibility that remains is to have $u[i..i+6] = 1201321$. In this case, we obtain that either $f(u)[i..i+6] = w_1$ or $f(u)[i..i+6]$ is one of the words y_1, y_2, or y_3. When $f(u)[i..i+6] = w_1$ we obtain easily that $|u|$ is divisible by 9, so $g(u)[i..i+6] = w_1$, as well. Again, this shows that f and g are identical, so t_β contains a cube, a contradiction. Now, if $f(u)[i..i+6] = y_1$ we get that the length of u is of the form $9k+8$ for some $k \in \mathbb{N}$. This means that $g(u)[i] = 3$, a contradiction. If $f(u)[i..i+6] = y_2$ we get that the length of u is of the form $9k+2$ for some $k \in \mathbb{N}$. This would mean that $g(u)[i..i+3] = 1023$, again a contradiction. Finally, when $f(u)[i..i+6] = y_3$ we get that the length of u is of the form $9k+5$ for some $k \in \mathbb{N}$ and we get that $g(u)[i] = 2$, which is once more a contradiction. As we have reached a contradiction in every case, we conclude that the assumption we made was false. Similar arguments work for the cases of when $f(u)$ and $g(u)$ contain a factor with three occurrences of the letter 1. Thus, t_β has no factor of the form $uf(u)g(u)$ for any $u \in \Sigma^*$ with $|u| \geq 7$.

To show the second statement, we have that every possible occurrence of such a factor is included in the image under β of a factor of length 4 of t (by the same reasoning as above). Computer calculations show that there are only 12 different factors of the form $ug_1(u)g_2(u)$ for some $u \in \Sigma^*$ with $|u| \leq 6$ and permutations g_1, g_2 such that there is no position $1 \leq \ell \leq |u|$ with $u[l] \neq g_1(u)[\ell] \neq g_2(u)[\ell] \neq u[\ell]$. These factors are: 012|013|213, 013|213|012, 023|012|013, 120|132|130, 130|120|132, 132|130|120, 201|321|301, 213|012|013, 230|120|132, 301|201|321, 321|301|201, 321|301|203, where the vertical lines mark the borders between $u, g_1(u)$ and $g_2(u)$. For every factor we can check that there are no $i, j \in \mathbb{N}$ and no permutation f such that $g_1 = f^i$ and $g_2 = f^j$. For instance, let us assume that there are i, j and f such that 012|013|213 is a factor of the form $uf^i(u)f^j(u)$ (i.e., $u = 012, f^i(u) = 013$ and $f^j(u) = 213$). Since

$u[1] = f^i(u)[1] = f^i(u[1]) = 0$, it follows that $\mathbf{ord}_f(0) \mid i$ and since $f^j(u)[1] = 2$, we conclude that the letter 2 is in the same orbit of f as 0, i.e., $\mathbf{ord}_f(2) = \mathbf{ord}_f(0)$ and $\mathbf{ord}_f(2) \mid i$. This is a contradiction with $u[3] = 2 \neq 3 = f^i(u)[3] = f^i(u[3])$. The analysis of the other factors leads to similar contradictions. □

The next result highlights sets of patterns that cannot be simultaneously avoided.

Lemma 3. *There is no $w \in \Sigma_3^\omega$ that avoids the patterns xxx, $xx\pi(x)$, and $x\pi(x)x$ simultaneously. There is no $w \in \Sigma_3^\omega$ that avoids the patterns xxx, $x\pi(x)\pi(x)$, and $x\pi(x)x$ simultaneously.*

Proof. It can be easily seen (for instance, by checking with a computer program that explores all the possibilities by backtracking) that any word of length at most 9 over Σ_3 contains a word of the form uuu, $uuf(u)$, or $uf(u)u$, for some $u \in \Sigma_3^+$ and some morphic permutation f of Σ_3.

Similarly, any word of length at most 10 over Σ_3 contains a word of the form uuu, $uf(u)f(u)$, or $uf(u)u$, for $u \in \Sigma_3^+$ and morphic permutation f of Σ_3. □

The following result shows the equivalence between the avoidability of several pairs of patterns.

Lemma 4. *A word $w \in \Sigma_m^\omega$ avoids the pattern $xx\pi(x)$ if and only if w avoids the pattern $\pi(x)\pi(x)x$. A word $w \in \Sigma_m^\omega$ avoids the pattern $x\pi(x)\pi(x)$ if and only if w avoids the pattern $\pi(x)xx$. A word $w \in \Sigma_m^\omega$ avoids the pattern $x\pi(x)x$ if and only if w avoids the pattern $\pi(x)x\pi(x)$.*

Proof. If an infinite word w has no factor $uuf(u)$, with $u \in \Sigma_m^+$ and a morphic permutation f of Σ_m, then w does not contain any factor $g(u)g(u)u$, with $u \in \Sigma_m^+$ and a morphic permutation g of Σ_m for which there exists a morphic permutation f of Σ_m such that $g(f(a)) = a$, for all $a \in \Sigma$. This clearly means that w avoids $\pi(x)\pi(x)x$ in Σ_m. The other conclusions follow by the same argument. □

The following two remarks are immediate.

- The pattern $\pi^i(x)\pi^i(x)\pi^i(x)$ is avoidable in Σ_m for $m \geq 2$ by the word t.
- The patterns $\pi^i(x)\pi^i(x)\pi^j(x)$ and $\pi^i(x)\pi^j(x)\pi^j(x)$, $i \neq j$, are avoidable in Σ_m for $m \geq 3$ by the word h.

Another easy case of avoidable patterns is highlighted in the next lemma.

Lemma 5. *The pattern $\pi^i(x)\pi^j(x)\pi^i(x)$, $i \neq j$, is avoidable in Σ_m, for $m \geq 3$.*

Proof. Assume $i < j$. In this case, setting $y = \pi^i(x)$ we get that the pattern $\pi^i(x)\pi^j(x)\pi^i(x)$ is actually $y\pi^{j-i}(y)y$. Avoiding the last pattern is the same as avoiding the pattern $y\pi(y)y$. This pattern is avoidable in alphabets with three or more letters, by Lemma 1. Also, this pattern is clearly unavoidable in Σ_1 and Σ_2.

If $i > j$, we take $y = \pi^j(x)$ and we obtain that $\pi^i(x)\pi^j(x)\pi^i(x)$ is actually $\pi^{i-j}(y)y\pi^{i-j}(y)$, which is avoidable if and only if $\pi(y)y\pi(y)$ is avoidable. This latter pattern is avoidable over alphabets with three or more letters, by Lemmas 1 and 4. The pattern is clearly unavoidable in Σ_1 and Σ_2. □

In the next theorem we present the case of the patterns $x\pi^i(x)\pi^j(x)$, with $i \neq j$. For this we need to define the following values:

$$k_1 = \inf\{t : t \nmid |i-j|, t \nmid i, t \nmid j\} \tag{1}$$
$$k_2 = \inf\{t : t \mid |i-j|, t \nmid i, t \nmid j\} \tag{2}$$
$$k_3 = \inf\{t : t \mid i, t \nmid j\} \tag{3}$$
$$k_4 = \inf\{t : t \nmid i, t \mid j\}. \tag{4}$$

Remember that $\inf \varnothing = +\infty$. However, note that $\{t : t \nmid |i-j|, t \nmid i, t \nmid j\}$ is always not empty, and that $k_1 \geq 3$ (as either $|i-j|$ is even or one of i and j is even, so $k_1 > 2$). Also, as $i \neq j$ at least one of the sets $\{t : t \mid i, t \nmid j\}$ and $\{t : t \nmid i, t \mid j\}$ is also not empty. Further, we define

$$k = \min\{\max\{k_1, k_2\}, \max\{k_1, k_3\}, \max\{k_1, k_4\}\} \tag{5}$$

According to the remarks above, k is always defined (that is $k \neq +\infty$).

Lemma 6. *The pattern $x\pi^i(x)\pi^j(x)$, $i \neq j$, is unavoidable in Σ_m, for $m \geq k$.*

Proof. First, let us note that the fact that $m \geq k_1$ means that for every word $u \in \Sigma_m^*$ there exists a morphism f such that $u \neq f^i(u) \neq f^j(u) \neq u$; indeed, we take f to be a cyclic permutation of Σ_m, which means that the first letters of u, $f^i(u)$ and $f^j(u)$ are pairwise different. Similarly, the fact that $m \geq k_2$ (when $k_2 \neq +\infty$) means that for every word $u \in \Sigma_m^*$ there exist a morphism f such that $u \neq f^i(u) = f^j(u)$. In this case, suppose that a is the first letter of u, and take f a permutation such that $\mathbf{ord}_f(a) = k_2$, and f only changes the letters from the orbit of a (thus, $\mathbf{ord}(f) \mid k_2$). Clearly, the first letters of $f^i(u)$ and $f^j(u)$ are not equal to a, but $f^i(u) = f^j(u)$ as $\mathbf{ord}(f)$ divides $|i-j|$. We get that $u \neq f^i(u) = f^j(u)$, for this choice of f. Finally, one can show by an analogous reasoning that the fact that $m \geq k_3$ (when $k_3 \neq +\infty$) means that for every word $u \in \Sigma_m^*$ there exists a morphism f such that $u = f^i(u) \neq f^j(u)$ and the fact that $m \geq k_4$ (when $k_4 \neq +\infty$) means that for every word $u \in \Sigma_m^*$ there exists a morphism f such that $f^i(u) \neq u = f^j(u)$.

Further, we show that if $m \geq \max\{k_1, k_2\}$ (in the case when $k_2 \neq +\infty$) there is no infinite word over Σ_m that avoids $x\pi^i(x)\pi^j(x)$. As $k_1 \geq 3$ it follows that $m \geq 3$. One can quickly check that the longest word that does not contain an instance of this pattern has length six and is 001010 by trying to construct such a word letter by letter. This means that there is no infinite word over Σ_m that avoids this pattern in this case.

By similar arguments, we can show that if $m \geq \max\{k_1, k_3\}$ (in the case when $k_3 \neq +\infty$) there is no infinite word over Σ_m that avoids $x\pi^i(x)\pi^j(x)$. In this case, the longest word that avoids those patterns is 01010.

If $m \geq \max\{k_1, k_4\}$ (in the case when $k_4 \neq +\infty$) we also get that here is no infinite word over Σ_m that avoids $x\pi^i(x)\pi^j(x)$. The construction ends at length six, the longest words without an instance of the pattern are 011001, 011002, 011221, 011223 and 011220.

These last remarks show that the pattern $x\pi^i(x)\pi^j(x)$ is unavoidable by infinite words over Σ_m, for all $m \geq k$. $\square$

The next result represents the main step we take towards characterising the avoidability of cubes under morphic permutations.

Proposition 1. *Given the pattern $x\pi^i(x)\pi^j(x)$ we can determine effectively the values m, such that the pattern is avoidable in Σ_m.*

Proof. Since we already examined the case $m \geq k$ in Lemma 6, it only remains to be seen which is the situation for Σ_m with $m < k$.

The cases for $m = 2$ and $m = 3$ are depicted in Table 1. Note that in the table an entry "✓" (respectively, "×") at the intersection of line (i, Σ_m) and column (j) means that the pattern $xf^i(x)f^j(x)$ is avoidable (respectively, unavoidable) in Σ_m. In building the table we used the fact that the pattern $x\pi^i(x)\pi^j(x)$ is avoidable in Σ_2 if and only if $i \equiv j \equiv 0(mod\ 2)$, and in that case it is avoided by the Thue-Morse word. Also, for Σ_3, when $j \neq 0$, the avoidability of the pattern follows from the fact that an instance of the pattern contains cubes or squares, so it can be avoided by the infinite words t (seen as a word over three letters, that just does not contain one of the letters) or h, respectively. In the case when $j = 0$, we use the word defined in Lemma 2 to show the avoidability of the respective patterns.

We move on to the case $m \geq 4$. In this case, we split the discussion in several further cases, depending on the minimum of k_1, k_2, k_3, and k_4.

Case 1: $k_1 = \min\{k_1, k_2, k_3, k_4\}$. This means that $k > k_1$. If $m < k_1$ it must be the case that $m \mid i$ and $m \mid j$ (since $k_3, k_4 > k_1$). For every letter $a \in \Sigma_m$ and every morphic permutation f of Σ_m, since $\mathbf{ord}_f(a) \leq m$ we get that $\mathbf{ord}_f(a) \mid i$ and $\mathbf{ord}_f(a) \mid j$. So in this case an instance of the pattern $x\pi^i(x)\pi^j(x)$ is actually a cube, which can be avoided by the Thue-Morse word. If $k_1 \leq m < k$, then for every $a \in \Sigma_m$ and morphic permutation f of Σ_m we either have that $\mathbf{ord}_f(a)$ divides both i and j or that $\mathbf{ord}_f(a)$ divides neither i nor j nor $|i - j|$. If we have a letter a occurring in a word u such that the latter holds, it means that we have at least 3 different letters in the the word $uf^i(u)f^j(u)$. If there is no such letter in u, then $uf^i(u)f^j(u)$ is a cube. In both cases, the Thue-Morse word avoids the pattern $x\pi^i(x)\pi^j(x)$.

Case 2: $k_2 = \min\{k_1, k_2, k_3, k_4\}$. In this case, it can easily be seen that $k = k_1$. If $4 \leq m < k_2$ we get for every $a \in \Sigma_m$ and every morphic permutation

		j(mod 6)											
		0		1		2		3		4		5	
i(mod 6)	0	✓	✓	×	✓	✓	✓	×	✓	✓	✓	×	✓
	1	×	✓	×	✓	×	×	×	×	×	×	×	×
	2	✓	✓	×	×	✓	✓	×	×	✓	✓	×	✓
	3	×	✓	×	✓	×	×	×	✓	×	×	×	✓
	4	✓	✓	×	✓	✓	✓	×	×	✓	✓	×	×
	5	×	✓	×	×	×	×	×	×	×	×	×	✓
		Σ_2	Σ_3	Σ_2	Σ_3	Σ_2	Σ_3	Σ_2	Σ_3	Σ_2	Σ_3	Σ_2	Σ_3

Table 1. Avoidability of $x\pi^i(x)\pi^j(x)$ in Σ_2 and Σ_3 for morphic permutations π

f of Σ_m that $\mathbf{ord}_f(a) \mid i$ and $\mathbf{ord}_f(a) \mid j$ (since $k_3, k_4 > k_2$). This means that in this case every instance of the pattern $x\pi^i(x)\pi^j(x)$ is a cube, which can be avoided by the Thue-Morse word. If $k_2 \leq m < k$, we have for each letter $a \in \Sigma_m$ and every morphic permutation f of Σ_m that either $\mathbf{ord}_f(a)$ divides at least one of i and j or $\mathbf{ord}_f(a) \mid |i-j|$. In all cases, this means that for each position l of a word u, we have that at least two of the letters $u[\ell], f^i(u)[\ell]$ and $f^j(u)[\ell]$ are equal, and the word defined in Lemma 2 avoids such patterns.

Case 3: $k_3 = \min\{k_1, k_2, k_3, k_4\}$. As in the previous case we get that $k = k_1$. If $4 \leq m < k_3$ we have that for every letter $a \in \Sigma_m$ and every morphic permutation f it must be the case that $\mathbf{ord}_f(a) \mid i$ and $\mathbf{ord}_f(a) \mid j$. Again, every instance of $x\pi^i(x)\pi^j(x)$ is in fact a cube, and so this pattern is avoided by the Thue-Morse word. If $k_3 \leq m < k = k_1$ we can easily see that for every letter $a \in \Sigma_m$ and every morphic permutation f we have that $\mathbf{ord}_f(a)$ divides i or j or both of them. This means that for every factor of the form $uf^i(u)f^j(u)$ and every position ℓ in u we have that $u[\ell] = f^i(u)[\ell]$ or $u[\ell] = f^j(u)[\ell]$. The word of Lemma 2 avoids such patterns.

Case 4: $k_4 = \min\{k_1, k_2, k_3, k_4\}$. This is symmetric to the previous case, so the pattern $x\pi^i(x)\pi^j(x)$ is avoided by the Thue-Morse word for $4 \leq m < k_4$ and by the word of Lemma 2 for $k_4 \leq m < k$.

Now we can conclude the characterisation of patterns $x\pi^i(x)\pi^j(x)$. Such a pattern is always avoidable in Σ_m for all $4 \leq m < k$. Moreover, it might also be avoidable Σ_2 and Σ_3, or only Σ_3 but not in Σ_2, or neither in Σ_2 nor in Σ_3 (according to Table 1). Therefore, for each pair (i, j) of natural numbers, defining a pattern $x\pi^i(x)\pi^j(x)$, we can effectively compute the values of m such that this pattern is avoidable in Σ_m. ☐

Further we show the following result, as a completion of the previous one.

Proposition 2. *Given the pattern $\pi^i(x)\pi^j(x)x$ we can determine effectively the values m, such that the pattern is avoidable in Σ_m.*

Proof. Let m be a natural number. We want to check whether $\pi^i(x)\pi^j(x)x$ is avoidable in Σ_m or not. Take $M = \max\{i+1, j+1, m\}$. It is not hard to see that $f^{M!}$ equals the identity for all morphic permutations f of the alphabet Σ_m. Let us take $y = \pi^i(x)$. By the fact that the function that can substitute π are permutations, we obtain that $\pi^i(x)\pi^j(x)x$ is avoidable in Σ_m if and only if $y\pi^{M!-i}(y)\pi^{M!-i+j}(y)$ is avoidable in Σ_m. Moreover, note that:

$$\begin{aligned}
\inf\{t : t \nmid j, t \nmid M!-i, t \nmid M!-i+j\} &= \inf\{t : t \nmid |i-j|, t \nmid i, t \nmid j\} \\
\inf\{t : t \mid j, t \nmid M!-i, t \nmid M!-i+j\} &= \inf\{t : t \nmid i, t \mid j\} \\
\inf\{t : t \mid M!-i, t \nmid M!-i+j\} &= \inf\{t : t \mid i, t \nmid j\} \\
\inf\{t : t \nmid M!-i, t \mid M!-i+j\} &= \inf\{t : t \mid |i-j|, t \nmid i, t \nmid j\}
\end{aligned}$$

Therefore, $y\pi^{M!-i}(y)\pi^{M!-i+j}(y)$ is avoidable in Σ_m if and only if $4 \leq m < k$, where k is defined using (5) for i and j. ☐

In the exact same manner we get the following proposition.

Proposition 3. *Given the pattern $\pi^i(x)x\pi^j(x)$ we can determine effectively the values m, such that the pattern is avoidable in Σ_m.* □

We can now summarise the results of this section in the following theorem:

Theorem 1. *Given the pattern $\pi^i(x)\pi^j(x)\pi^k(x)$ where π is substituted by morphic permutations, we can determine effectively the values m such that the pattern is avoidable in Σ_m.*

Proof. Let us assume that i is the minimum between i, j, and k. Let us take $y = \pi^i(x)$. The pattern becomes $y\pi^\ell(y)\pi^t(y)$, and we can identify all the alphabets where this pattern is avoidable by Proposition 1.

If j is the minimum between i, j, and k we use Proposition 3 to identify all the alphabets where this pattern is avoidable. Finally, if k is the minimum between i, j, and k we use Proposition 2 to identify all the alphabets where this pattern is avoidable. □

4 The antimorphic case

In this section, the function variable π is always replaced by an antimorphic permutation. Most of the proofs of this section can be found in the Appendix. They follow mainly the same lines as the proofs presented in Section 3; however, in some cases, they are more technically involved.

As in the morphic case, we first establish a series of results regarding basic patterns. To begin with, we introduce the morphism $\gamma : \Sigma_2^* \to \Sigma_3^*$ defined by

$$0 \mapsto 0011022, \qquad 1 \mapsto 1100122.$$

Lemma 7. *The word $t_\gamma = \gamma(t)$ avoids the pattern $x\pi(x)x$ in Σ_m, for $m \geq 3$.*

The following lemma shows the avoidability of a particular type of patterns where the function variable is a morphism; this result becomes useful in the sequel. For this, we define the morphism $\delta : \Sigma_3^* \to \Sigma_4^*$ by

$$0 \mapsto 012031, \qquad 1 \mapsto 032132, \qquad 2 \mapsto 032102130132.$$

Lemma 8. *The word $h_\delta = \delta(h)$ contains no factor uu and $uf(u)u^R$ where $u \in \Sigma_m^+$ and f is a morphic permutation of Σ_m, for all $m \geq 4$.*

The previous lemma has a corollary that is important in the context of avoidability of cubes under antimorphic permutations.

Corollary 1. *There exists an infinite word that avoids the patterns xx and $x\pi(x)x^R$ in Σ_m, for all $m \geq 4$.*

Proof. By the previous proof we obtain that there exist infinitely many finite words that contain no factors uu and $uf(u)u^R$ for $u \in \Sigma_m^+$ and morphic permutations f over alphabets Σ_m with $m \geq 4$. By reversing these words, we obtain that there exist infinitely many finite words over Σ_m that contain neither squares nor factors $uf(u)u^R$ for $u \in \Sigma_m^*$ and antimorphic permutations f on Σ_m, with $m \geq 4$. Therefore, there exists an infinite word that contains no such factors, and the statement of the corollary holds. □

As in the case of the morphic permutations, we first study the avoidability of the pattern $x\pi^i(x)\pi^j(x)$. However, a finer analysis must be performed here.

In the next lemma we look at case when the exponent i is even and j is odd. For this purpose let the morphism $\zeta : \Sigma_2^* \to \Sigma_5^*$ be defined by

$$0 \mapsto 012034, \qquad\qquad 1 \mapsto 120324.$$

Lemma 9. *Let $t_\zeta = \zeta(t)$ for the morphism ζ defined above. Also, let $i \in \mathbb{N}$ be even and $j \in \mathbb{N}$ be odd, and f and g be morphic and, respectively, antimorphic permutations of Σ_m, with $m \geq 5$. The word t_ζ obtained as such does not contain any factor of the form $uf(u)g(u)$ for $u \in \Sigma_5^*$ with $|u| \geq 6$. Furthermore, t_ζ does not contain any factor of the form $uf^i(u)f^j(u)$ such that*

$$\left|\left\{u[\ell], f^i(u)[\ell], f^j(u)^R[\ell]\right\}\right| \leq 2,$$

for all $\ell \leq |u|$ and $|u| \leq 5$.

In the case when the exponent i is odd and j is even, we examine the morphism $\eta : \Sigma_2^* \to \Sigma_5^*$ defined by

$$0 \mapsto 01234012431024301234012431023410243012431 0234,$$
$$1 \mapsto 0123401243102430123410234012430123410243 10234.$$

Note that this morphism is equivalent to $\theta \circ \beta$, where β is the morphism defined in Lemma 2 and $\theta : \Sigma_4^* \to \Sigma_5^*$ is defined by

$$0 \mapsto 01234, \qquad\qquad 1 \mapsto 01243,$$
$$2 \mapsto 10243, \qquad\qquad 3 \mapsto 10234.$$

Lemma 10. *Let $t_\eta = \eta(t)$ for the morphism η defined above. Also, let $i \in \mathbb{N}$ be odd and $j \in \mathbb{N}$ be even and f and g be antimorphic and, respectively, morphic permutations of Σ_m, with $m \geq 5$ The word t_η obtained as such does not contain any factor of the form $uf(u)g(u)$ for $u \in \Sigma_5^*$ with $|u| \geq 11$. Furthermore, t_η does not contain any factor of the form $uf^i(u)f^j(u)$ such that*

$$\left|\left\{u[\ell], f^i(u)^R[\ell], f^j(u)[\ell]\right\}\right| \leq 2,$$

for all $\ell \leq |u|$ and $|u| \leq 10$.

We now move further to the main results regarding the avoidability of cubes under antimorphic permutations.

It is not hard to see that the results on the avoidability of the patterns $\pi^i(x)\pi^i(x)\pi^i(x)$ with $i \in \mathbb{N}$ and $\pi^i(x)\pi^i(x)\pi^j(x)$ with $i, j \in \mathbb{N}$ for morphic permutations also hold in the case of antimorphic permutations. An equivalent of Lemma 5 also holds in the antimorphic case.

Lemma 11. *The pattern $\pi^i(x)\pi^j(x)\pi^i(x)$, $i \neq j$, is avoidable in Σ_m for $m \geq 3$.*

We now look at patterns of the form $x\pi^i(x)\pi^j(x)$ with $i \neq j$ and antimorphic f. Let k_1, k_2, k_3, k_4 and k be defined as in (1) to (5).

Lemma 12. *The pattern $x\pi^i(x)\pi^j(x)$, $i \neq j$, is unavoidable in Σ_m for $m \geq k$.*

Proposition 4. *Given the pattern $x\pi^i(x)\pi^j(x)$, we can determine effectively the values m, such that the pattern is avoidable in Σ_m.*

Proof. The cases when $m = 2$ and $m = 3$ are exactly like those depicted in Table 1 for the morphic case.

The case when $m = 4$ is based on the remark that it is sufficient to know how to decide the avoidability of the pattern $x\pi^i(x)\pi^j(x)$ for $i, j < 12$. Indeed, it is not hard to see that if i and j are arbitrary natural numbers, then $x\pi^i(x)\pi^j(x)$ is avoidable in Σ_4 if and only if $x\pi^{i'}(x)\pi^{j'}(x)$ is avoidable, for i' (resp. j') being the remainder of i (resp. j) divided by 12. With this in mind, one can analyse every pair (i, j) with $1 \leq i, j \leq 12$, and decide in each case the avoidability of the pattern $x\pi^i(x)\pi^j(x)$. The pattern is clearly unavoidable whenever the value k computed for i and j in (5) is less than or equal to 4. When $i = 0$ the pattern $x\pi^i(x)\pi^j(x)$ is avoided by the word h as any instance of the pattern contains squares, and when $j = 0$ the pattern is avoided by the word from Lemma 7. Also, in the case when i and j are both even we can decide the avoidability of the pattern using the results obtained for morphisms in the previous sections, as, in this case, f can be seen as a morphism instead of an antimorphism. Moreover, when $i = j$ we can avoid the pattern $x\pi^i(x)\pi^i(x)$ by the word h that contains no squares. The same word h avoids the pattern in the cases when $(i, j) \in \{(4, 1), (9, 1), (8, 5), (9, 5), (3, 7), (4, 7), (3, 11), (8, 11)\}$. To complete the picture, we note that a word avoids the pattern $x\pi(x^R)x^R$ if and only if it avoids the pattern $x\pi'(x)x^R$ where π' is mapped to a morphic permutation. Therefore, by Lemma 8 we obtain that the pattern $x\pi^i(x)\pi^j(x)$ is avoided by the infinite word h_δ for $(i, j) \in \{(4, 3), (8, 3), (4, 9), (8, 9)\}$ and by Corollary 1 we obtain that it is avoidable for $(i, j) \in \{(7, 3), (11, 3), (1, 9), (5, 9)\}$.

Further, the discussion is split in four cases. If both i and j are even, we can decide the avoidability of the pattern just as in the case of morphisms (as the instance of π can be seen, in fact, as a morphism). If both i and j are odd, we compute the value k defined in (5) and define $M = \max\{k, j + 1, i + 1\}$. Now, $x\pi^i(x)\pi^j(x)$ is avoidable in Σ_m if and only if $(x\pi^i(x)\pi^j(x))^R = \pi^j(x^R)\pi^i(x^R)x^R$ is avoidable in Σ_m. The last condition is equivalent to the avoidability of the pattern $\pi^j(y)f^i(y)y$ in Σ_m. Taking $z = \pi^j(y)$, we obtain that $\pi^j(y)\pi^i(y)y$ is avoidable in Σ_m if and only if $z\pi^{M!-j+i}(z)\pi^{M!-j}(z)$ is avoidable in Σ_m. Now we only have to notice that $M! - j + i$ is even and $M! - j$ is odd, as $M!$ is always even. Therefore, the case when i and j are odd can be reduced to the case when i is even and j is odd.

So there remain only two cases to be analysed: the case when i is even and j is odd as well as the case when i is odd and j is even. In this cases the proofs follows similar to the morphic case. $\square$

As in the case of morphic permutations we can easily derive the following two results.

Proposition 5. *Given the pattern $\pi^i(x)\pi^j(x)x$, we can determine effectively the values m such that the pattern is avoidable in Σ_m.* □

Proposition 6. *Given the pattern $\pi^i(x)x\pi^j(x)$, we can determine effectively the values m such that the pattern is avoidable in Σ_m.* □

Finally, as consequence of the last three propositions, we state the main result of this section in the following theorem:

Theorem 2. *Given the pattern $\pi^i(x)\pi^j(x)\pi^k(x)$ where π is substituted by antimorphic permutations, we can determine effectively the values m such that the pattern is avoidable in Σ_m.* □

5 Conclusions

In this paper, we have extended the concept of avoidability of patterns to avoidability of patterns with permutations. We have characterised for all m whether a cube, that is, a pattern of the form $\pi^i(x)\pi^j(x)\pi^k(x)$, is avoidable in Σ_m for all $i, j, k \geq 0$. We have given these characterisations for both the morphic and antimorphic case.

The next natural question is of course concerning the avoidance of longer patterns. Note that a first step towards answering that question follows from Lemma 2 (morphic case) and 9 (antimorphic case). They each give a word over four letters or five letters, respectively, that avoids sequences of permutations of length 3 or more for all factors of length 7 or more.

References

1. Thue, A.: Über unendliche Zeichenreihen. Norske Vid. Skrifter I. Mat.-Nat. Kl., Christiania **7** (1906) 1–22
2. Thue, A.: Über die gegenseitige Lage gleicher Teile gewisser Zeichenreihen. Norske Vid. Skrifter I. Mat.-Nat. Kl., Christiania **1** (1912) 1–67
3. Cassaigne, J.: Unavoidable Patterns. In: Algebraic Combinatorics on Words. Cambridge University Press, Cambridge, UK (2002) 111–134
4. Currie, J.: Pattern avoidance: themes and variations. Theoret. Comput. Sci. **339**(1) (2005) 7–18
5. Bischoff, B., Nowotka, D.: Pattern avoidability with involution. In: Words 2011, Prague. Volume 63 of Electron. Proc. in Theor. Comput. Sci. (2011) 65–70
6. Chiniforooshan, E., Kari, L., Xu, Z.: Pseudopower avoidance. Fundamenta Informaticae **114** (2012) 1–18
7. Currie, J.: Pattern avoidance with involution. CoRR **abs/1105.2849** (2011)
8. Hall, M.: Generators and relations in groups The Burnside problem. In: Lectures on Modern Mathematics. Volume 2. Wiley, New York (1964) 42–92
9. Lothaire, M.: Combinatorics on Words. Cambridge University Press (1997)

Appendix

Preliminaries

We first mention several notations that we use throughout the paper. Word variables are usually denoted by x, y, or z, while words are denoted by u, v, w, s, t, etc. On the other hand, we denote the function variables with π or π' and the instances of such variables (i.e., functions on Σ_k^*) with letters like f, g, f', g', f_1, g_1, etc.

Let us give an algebraic definition of the general notion of pattern which involves functional dependencies and pattern avoidability in this setting.

In order to define the notions of pattern and pattern avoidability we make use of some basic universal algebra notions. A signature is a pair (S, Σ) with S a finite set, whose elements are called sorts, and

$$\Sigma = \{\sigma : s_1 \times \cdots \times s_n \to s \quad : \quad n \geq 0, s_1, \ldots, s_n, s \in S\}$$

a set of function symbols; for a function symbol $\sigma : s_1 \times \cdots \times s_n \to s$ we say that n is its arity. In this paper, we assume that $S = \{s\}$ is a singleton and that all the functions have their arity greater than or equal to 1. An (S, Σ)-algebra is a structure $A = (A_s, A_\Sigma)$ where A_s is a set and

$$A_\Sigma = \{A_\sigma : A_s^n \to A_s \quad : \quad \sigma : s^n \to s \in \Sigma\}.$$

For a set of variables X we define the term algebra generated by X, denoted $T_\Sigma(X)$, as follows:

- $T_0 = X$;
- $T_{i+1} = \{\sigma(\alpha_1, \ldots, \alpha_n) \quad : \quad \sigma : s^n \to s \in \Sigma, \alpha_1, \ldots, \alpha_n \in T_i\} \cup T_i$, for $i \geq 0$;
- $(T_\Sigma(X))_s = \cup_{i \geq 0} T_i$;
- $(T_\Sigma(X))_\sigma(\alpha_1, \ldots, \alpha_n) = \sigma(\alpha_1, \ldots, \alpha_n)$, for $\sigma : s^n \to s \in \Sigma, \alpha_1, \ldots, \alpha_n \in (T_\Sigma(X))_s$.

A morphism ϕ of (S, Σ)-algebras $A = (A_s, A_\Sigma)$ and $B = (B_s, B_\Sigma)$ is a function $\phi : A_s \to B_s$ such that $\phi(A_\sigma(\alpha_1, \ldots, \alpha_n)) = B_\sigma(\phi(\alpha_1), \ldots, \phi(\alpha_n))$. Given an algebra $A = (A_s, A_\Sigma)$ and a function $f : X \to A_s$, there exists a unique morphism F from $T_\Sigma(X)$ to A such that $F(x) = f(x)$, for all $x \in X$.

Given a set of variables X and a set of function letters $\mathcal{F}$ (all having arity greater than or equal to one, and working over a single sort s), a pattern p is an element of $T_\mathcal{F}(X)$, the term algebra generated by X over the signature $(\{s\}, \mathcal{F})$. A pattern p is said to be *avoidable in the alphabet Σ_k with functions from the family* $(\mathcal{F}_n)_{n \geq 1}$ (where the functions of $\mathcal{F}_n$ are defined over $(\Sigma_k^*)^n$ with values in Σ_k^*, for $n \geq 1$) if there is an infinite word w over Σ_k that has no factor u verifying the following properties:

- there exists an algebra $A^u = (\Sigma_k^*, A^u_\mathcal{F})$ whose functions of arity n are contained in $\mathcal{F}_n$ for $n \geq 1$,

– u is the image of p through the unique morphism ϕ of $(\{s\}, \mathcal{F})$-algebras from $T_{\mathcal{F}}(X)$ to A^u.

In this paper, we only analyse the avoidability of a very restricted class of patterns. First, we only deal with the case when X is a singleton and $\mathcal{F}$ contains exactly one function letter for each arity greater than or equal to 1. For $n \geq 2$, we always take $\mathcal{F}_n = \{c_n\}$, where c_n is the n-ary catenation of words. Moreover, for an alphabet Σ_k we take the family $\mathcal{F}_1$ as either the family of morphic permutations over Σ_k or the family of antimorphic permutations over the same alphabet. Finally, almost all our results concern patterns with only three occurrences of the variable x; such patterns are, clearly, of the form $\pi^i(x)\pi^j(x)\pi^k(x)$.

For a pattern $\pi^i(x)\pi^j(x)\pi^k(x)$, we say that a word $f^i(u)f^j(u)f^k(u)$, where $u \in \Sigma_k^*$ and $f : \Sigma_k^* \to \Sigma_k^*$ is a anti-/morphism that permutes Σ_k, is an instance of the given pattern.

The morphic case

Proof of Lemma 1:
We start by noting that t_α does not contain any factor of length less or equal to 15 which is an instance of the pattern $x\pi(x)x$. This can be shown by checking whether such a word is a factor of $\alpha(v)$, for all the factors v of length 4 of the Thue-Morse word; a simple computer program shows that indeed there are no such words.

It remains to show that any factor of t_α, having more than 15 letters, is not an instance of the pattern $x\pi(x)x$. For this, assume that w is such a factor (i.e., w appears in t_α and $|w| > 15$).

First, we show that w cannot be a cube. For the sake of a contradiction, assume that w is a cube and $w = u^3$. If u begins with 0 it follows easily that $|u|$ is divisible by 5; similarly, if u begins with 1 or with 2 we reach the same conclusion, namely $|u|$ is divisible by 5. So, in all cases $|u|$ is divisible by 5. It follows that $w = (s\alpha(v)s')^3$, where $s, s' \in \{0, 1, 2\}^*$, $|s| + |s'| = 5$ and $v \in \{0, 1\}^*$. If $|s| = 5$ (respectively, $|s'| = 5$) we obtain that $u = \alpha(av)$ (respectively, $u = \alpha(va)$), for some $a \in \{0, 1\}$; but this means that the Thue-Morse word contains a cube, a contradiction. Let us now assume that $3 \leq |s| < 5$. It follows that $s's = \alpha(a)$, for some letter $a \in \{0, 1\}$, and that s' is placed exactly before w in t_α. This means, once more, that t_α contains a factor $(s's\alpha(v))^3 = (\alpha(av))^3$. Thus, the Thue-Morse word contains a cube, a contradiction. Finally, we look at the case when $0 < |s| < 3$. It follows that $3 \leq |s'| < 5$. Thus, $s's = \alpha(a)$, for some letter $a \in \{0, 1\}$, and s is placed exactly after w in t_α. This means that t_α contains the factor $(\alpha(v)s's)^3 = (\alpha(va))^3$; it follows that the Thue-Morse word contains a cube, a contradiction. In all the cases we reached contradictions, so t_α contains no cube.

Further, note that 00 occurs always in t_α on positions equal to 4 modulo 5, 11 occurs always on positions equal to 2 modulo 5, and 22 occurs always on positions equal to 1 modulo 5. Moreover, each factor of t_α, longer than 5, contains at least a factor aa with $a \in \{0, 1, 2\}$.

Let us first analyse the case of $m = 3$: i.e., x can be mapped to words over Σ_3 and f is a permutation on the alphabet Σ_3.

Assume now that t_α contains a factor w that can be expressed as $uf(u)u$ for some $u \in \Sigma_3^*$ and a morphism f such that f permutes Σ_3. Let aa be the rightmost occurrence of one of the factors $\{00, 11, 22\}$ in u. On the positions that correspond to the letters aa we also have a square bb in $f(u)$, with $b \in \Sigma_3$. We have $uf(u)u = (saas')(f(s)bbf(s'))(saas')$, for some words s and s' such that $u = saas'$ and $|s| + |s'| + 2 > 5$. It is not hard to see that the length of $aas'f(u)s$ should be divisible by 5 (as the positions of t_α where a pair of letters aa occurs are equal modulo 5). Therefore, $|u| + |f(u)| = 2|u|$ is divisible by 5. This means that $|u|$ is divisible by 5. So aa and bb occur on positions that are equal modulo 5. Therefore, $a = b$. As u contains no other occurrences of one of the factors $\{00, 11, 22\}$ to the right of aa it follows that u has at most two other letters to the right of aa (in the case when $a = 1$ the factor u can have at most one letter to the right of aa). This means that u contains at least one letter exactly before aa, which is different from a; the same letter occurs before $bb = aa$ in $f(u)$. This means that f maps two of the three letters of the alphabet to themselves. Therefore, f is the identity on Σ_3. It follows that $f(u) = u$ and t_α contains a cube uuu. This is a contradiction.

Consequently, the assumption we made is false and t_α avoids the pattern $x\pi(x)x$ on Σ_3.

If $m > 3$, it follows that t_α may contain an instance of $x\pi(x)x$ that is not a cube if and only if π is mapped to a permutation of Σ_m whose restriction to Σ_3 is also a permutation (because t_α contains no other symbols than those of Σ_3). So, basically, the other letters of the alphabet over which the function substituting π is defined are not important, and we can follow the same reasoning as in the case when $m = 3$. □

The antimorphic case

Proof of Lemma 7:
We can easily check that t_γ contains no factor of the form $uf(u)u$ for some antimorphic permutation f with $|u| < 4$. Looking at γ, we see that every factor of t_γ of length at least 4 contains a square of letters ss with $s \in \Sigma_3$. Assuming that t_γ contains a factor of the form $uf(u)u$ for some u and an antimorphic permutation of the alphabet f, we now look at the last of those letter-squares that occurs in u:

Case 1: $ss = 00$. Looking at the occurrences of 00 in t_γ, we see that u either ends with 00, 001 or 0012 then. If u ends with 00, it is followed by 1102 or 122 in t_γ. Considering that $|u| > 4$, u ends with 2200 in the first case. It follows that $f(2200) = f(0)f(0)f(2)f(2) = 1102$, a contradiction since 2 is mapped to both 0 and 2. In the other case we get $f(00) = 12$, which is impossible as well. If u ends with 001, it is followed by 102 or 22 in t_γ. Again, there is no permutation f

such that $f(001) = 102$ or $f(001) = f(1)f(0)f(0) = 22f(0)$ (since the mapping is not injective in the second case). If u ends with 0012, which is always followed by 200 or 211 in t_γ, we also get $f(1) = f(2)$ in both cases, a contradiction as mentioned before.

Case 2: $ss = 11$. Then u ends with 11, 110 or 1102. If u ends with 11, it actually ends with 0011 or 2211, because 11 is always preceded by 00 or 22 in t_γ. If u ends with 0011, which is always followed by 022 in t_γ, it means that $f(011) = f(1)f(1)f(0) = 022$, a contradiction. If it ends with 2211, which is followed by 0012, we get $f(2211) = f(1)f(1)f(2)f(2) = 0012$, which also means mapping a letter onto two different images. In the case when u ends with 110, we observe that 110 is always followed by 22 or 012 in t_γ. In both cases, there is no permutation f such that $f(110) = f(0)f(1)f(1)$ starts with 22 or equals 012, contradicting the assumption. If u ends with 1102, which is followed by 200 or 211, we get from $f(102) = f(2)f(0)f(1)$ that $f(0) = f(1)$, again a contradiction.

Case 3: $ss = 22$. Then u ends with 22, 220 or 221. If u ends with 22, the suffix of length 4 must be 1022 or 0122, because 22 is always preceded by 01 or 10 in t_γ. Both possibilities are followed by 0011 or 1100 in t_γ, so f can not be a permutation in this case, since we have $f(2) = f(1)$. If u ends with 220, this means that it actually ends with 0220 or 1220, and it is followed by 0110 in t_γ. We conclude that u has to end with 0220, otherwise we would get that $f(0) = f(1)$, so f could be no permutation. According to the previous reasoning, we have the following situation in t_γ at the border between u and $f(u)$: 00110220|011022 where | marks this border. Since the factors 10220|01102|10220 and 110220|011022|110220 don't occur in t_γ, we must have $|u| > 6$ and we got the following factor in t_γ: 00110220|01102200 which is obviously followed by 11022 because of the form of γ. This means we have $(0011022)^3$ in $t_\gamma = \gamma(t)$ and since γ is a block code, it follows that 000 is a factor of t, a contradiction since the Thue-Morse avoids cubes. Similar reasoning in the case when u ends with 221 leads us to the conclusion, that there is no word u and no antimorphic permutation of the alphabet f such that $uf(u)u$ is a factor of t_γ. □

Proof of Lemma 8:
We begin with a series of simple remarks on the structure of the word h_δ. First, it is rather plain that in every factor of length 20 of h_δ we have at least one occurrence of the factor 032; indeed, a factor of length 20 of h_δ contains either $\delta(2)$ or the prefix of length 3 of one of $\delta(1)$ or $\delta(2)$. Second, as h contains no factor 121, we obtain in a similar fashion that in every factor of length 26 of h_δ we have at least one occurrence of the factor 012. Further, note that any factor of length less than or equal to 60 that appears in h_δ is contained in the image of a factor of length 10 of h (as every 4 consecutive letters of h contain at least one occurrence of the letter 2). Finally, every factor of length 10 of h is contained in its prefix $\phi_h^8(0)$, so every factor of length at most 60 is contained in the prefix $\delta(\phi_h^8(0))$. Indeed, this holds if we note that all the factors of length 3 that may appear in h appear in its prefix $\phi_h^4(0)$ (these factors are all the possible factors of length 3 over Σ_3 except the ones that contain squares, 010 or 212). Then, all

the factors of length 4 of h are contained in $\phi_h^5(0)$, all the factors of length 5 of h are contained in $\phi_h^6(0)$, all the factors of length 8 of h are contained in $\phi_h^7(0)$, and finally, all the factors of length 10 of h_δ are contained in $\phi_h^8(0)$.

We are now able to prove the lemma. We can check by a computer program that the prefix $\delta(\phi_h^8(0))$ contains neither a factor uu with $|u| < 26$ nor a factor $uf(u)u^R$ with $|u| < 20$ and f an arbitrary morphic permutation of Σ_m with $m \geq 4$.

Further, we show that h_δ contains no occurrences of a factor $uf(u)u^R$ with $|u| \geq 20$ and f an arbitrary morphic permutation of Σ_m with $m \geq 4$. In fact, we show that h_δ contains no occurrences of a factor uvu^R with $|u| \geq 20$ and $|v| = |u|$. By the previous remarks, u contains 032. Thus, u^R should contain 230. But one can easily see that 230 does not appear in h_δ, and this concludes the proof.

We move on to the more complicated task of showing that h_δ contains no squares uu with $|u| \geq 26$. Clearly, if h_δ would contain such a factor, then u would contain 012. But, if 012 appears in the first u on position i, then it appears also on position i in the second u. As 012 appears in h_δ only on positions $6t+1$ for some $t > 0$ it follows that the length of u is divisible by 6. Clearly, $u = x\delta(z)y$ for some $x, y \in \Sigma_4^*$ and $z \in \Sigma_3$ such that y (resp. x) has no prefix (resp. suffix) that is the image of a non-empty word through δ. This means that $yx = \delta(s)$ for some $s \in \Sigma_3$. If $|y| \geq 5$ then it follows that uu is followed by x as well (as y would uniquely determine the letter $s \in \Sigma_3$ such that $\delta(s) = yx$), so this means that h_δ contains a square $\delta(zs)^2$; this is a contradiction with the fact that h contains no squares. If $|x| \geq 4$ a similar conclusion would follow; more precisely, h_δ would contain the square $\delta(sz)^2$, again a contradiction. So $|y| \leq 4$ and $|x| \leq 3$. This means that $s \in \{0, 1\}$. If $yx = \delta(0)$ the conclusion follows just as the above. So the only case left to be analysed is when $yx = \delta(1)$. The only case that is different from the above is when uu is contained in $\phi(2z1z2)$. As h contains no squares, we get that z starts and ends with an 0, so h contains 010, again a contradiction. Therefore, we conclude that h_δ contains no squares.

Finally, we showed that h_δ has no factors uu and $uf(u)u^R$ for morphic permutations f on Σ_m for all $m \geq 4$. □

Proof of Lemma 9: We start by proving the first claim of the lemma. If $|u| = 6$, the length of such a factor is 18 and so it is completely contained in the images of factors of length 4 of the Thue-Morse word under ζ. We can verify that there is no such factor in this set by simple computer calculations. If $|u| > 6$, we can show an even stronger statement, namely that there is no factor $uf(u)$ in t_ζ where f is an antimorphic permutation. To see this, we make an extensive case analysis on the suffix of length 7 of u. There are 22 different factors of length 7 in t_ζ. We show two cases explicitly, the others use the same arguments. For example, if $0120340 \leq_s u$, this factor is always followed by 12034 in t_ζ. If this had the form $uf(u)$ for a word u and an antimorphic permutation f, we would get $f(0) = 1$ and $f(0) = 3$, a contradiction. In most other cases we get that one letter would be mapped onto two different images as well. A case where we need some different reasoning is when $0324120 \leq_s u$. We can easily see that this is always followed by 324 in t_ζ, which itself is followed by either 012 or 120. In the case when it is followed by 012, we get a contradiction of the same type as before, since 2 would then be mapped to both 2 and 1. So if this has the form $uf(u)$ for some word u and antimorphic permutation f, we must have 0324.120324.120 as a factor in t_ζ. But in order to get this as a factor of t_ζ, we must have the factor 111 in t, a contradiction. For the proof of the second statement, we only have to look at occurrences in the images of factors of length 4 of t, again because of the length constraint. A short computation shows that there are only two different factors of the form $ug_1(u)g_2(u)$ with $|u| \leq 5$ and g_1 and g_2 antimorphic permutations such that there is no position $1 \leq \ell \leq |u|$ with $u[\ell] \neq g_1(u)[\ell] \neq g_2(u)[\ell] \neq u[\ell]$ in those images: 240|120|341 and 340|120|341. For both factors, we can quickly check that there is no permutation f and no $i, j \in \mathbb{N}$ such that $g_1 = f^i$ and $g_2 = f^j$. If we assume that 240|120|341 has the form $uf^i(u)f^j(u)$, for some u and f as in the statement, we obtain $u = 240$, $f^i(u) = 120$ and $f^j(u) = 341$. By looking at the second letter of each block we get $f^i(4) = 2$ and $f^j(4) = 4$, that is $\mathbf{ord}_f(4) \mid j$ and 2 is in the same orbit of f as 4, so $\mathbf{ord}_f(2) = \mathbf{ord}_f(4)$. But we also get $f^j(240) = f^j(0)f^j(4)f^j(2) = 341$, so $f^j(2) = 1$, contradicting the fact that $\mathbf{ord}_f(2) \mid j$. The same reasoning applies to the factor 340|120|341. □

Proof of Lemma 10:
We first show that $t_\beta = \beta(t)$ does not contain any factor of the form $uf'(u)g'(u)$ for any $u \in \Sigma^*$ with $|u| \geq 7$, where f' is an antimorphism and g' is a morphism: As in Lemma 2, we check the cases when u is short by computer computations. In fact, we check that t_β has no factor of the form $uf'(u)g'(u)$ with $7 \leq |u| \leq 11$. It suffices to check the images of all factors of length 5 of the Thue-Morse word under β. In the case when $|u| \geq 12$ and f' is antimorphic, we can even prove a stronger result. In fact, we show that t_β does not contain any factor of the form $uf'(u)$ for an antimorphic permutation f', when $|u| \geq 12$. Let us assume, for the sake of a contradiction, that t_β has such a factor. Looking at $\beta(0)$ and $\beta(1)$, we see that every factor u of length at least 12 contains an occurrence of a substring s of length 4 that contains 4 different letters (this is already true for $|u| \geq 7$). In the following, we look at the last occurrence of such a factor in u and perform an exhaustive case analysis on its possible values and positions in u. Note that the vertical line marks the border between u and $f'(u)$, while the dot marks the border between the images of two letters under β.

1. $s = 0132$: Note that the factor 0132 is always followed by 130120 in t_β. If $0132 \leq_s u$, we got the following situation in t_β: 0132|1301. This means that $f'(0) = f'(2) = 1$, which contradicts the fact that f' is a permutation. If $01321 \leq_s u$, the situation at the border is 01321|3012. This would mean that 1 is mapped to both 2 and 3 by f', a contradiction. If $013212 \leq_s u$, we have 2013213|0120 (remember that we assumed $|u| \geq 11$ and by the definition of β, 013213 is always preceded by a 2 in t_β). The prefix 0120 of $f'(u)$ is followed by either 13 or 31 in t_β. In the former case, we get that f' maps 2 to

both 2 and 3 and in the latter case f' maps 1 to both 1 and 3, a contradiction in both cases.

2. $s = 2130$: If $2130 \leq_s u$, we would have $32130|1201$ or $32130|12031$ and either $f'(2) = f'(0) = 1$ or $f'(3) = f'(0) = 1$. So let $21301 \leq_s u$. But then we have either $21301|20132$ or $21301|20310$ and either $f'(2) = f'(1) = 2$ or $f'(2) = f'(0) = 0$, a contradiction in both cases.
3. The cases $s \in \{3012, 2013, 2031, 3102, 1023\}$ follow the same reasoning as above and lead to similar contradictions.
4. $s = 1203$. In this case we see that $1203 \leq_s u$, otherwise s would not be the last factor of that shape. From the definition of β, we see that the situation at the border between u and $f'(u)$ must be $301203|102301$ from which it follows that $f(1) = 3$ and $f(2) = 2$. This means that we get a contradiction as above if 301203 is preceded by a 1, since 102301 is always followed by 2 in t_β. So we have $102301213|1023012$ and since f' is completely determined by this, we get that $102301203 \leq_p f'(u)$. So we have the following factor occurring in t_β: $1023.01203|1023.01203$ and since β is a block code and therefore uniquely decodable, we conclude that 111 is a factor of t, a contradiction.
5. If $s = 2301$, we can derive that there is a cube in the Thue-Morse word as above, because we get the following situation: $3.012031023.01|2031023.01203$ (this is where we use the fact that $|u| \geq 12$).

Further, we look at the word $t_\eta = \theta(t_\beta)$. Let us first assume that $|u| > 30$. If $uf(u)g(u)$ appears in t_η then u contains at least six occurrences of the letter 2, and each two consecutive such occurrences have exactly four letters between them. This means that in $g(u)$ each two consecutive occurrences of $g(2)$ have exactly four letters between them as well, and $g(2)$ occurs at least six times in $g(u)$. This only leaves the possibility that $g(2) = 2$. From this, we also obtain easily that $|u|$ is divisible by 5. A similar argument shows that $f(2) = 2$. Consider now the last occurrence of 2 in u. This letter is mapped to the first 2 of $f(u)$, and there are exactly 4 letters between these two consecutive occurrence of the letter 2 in t_η. This means that after the last occurrence of the letter 2 in u there are exactly two more letters in this word and there are exactly two letters in $f(u)$ before the first occurrence of 2. As $5 \mid |u|$ we get that there exists a factor vsw in t_β, with $v, s, w \in \Sigma_4^*$, such that $u = \theta(v)$, $f(u) = \theta(s)$, and $g(u) = \theta(w)$. As f is an antimorphic permutation, there exists an antimorphic permutation f' such that $s = f'(v)$, and as g is a morphic permutation, there exists a morphic permutation g' such that $w = g'(v)$. As $|u| > 30$, we get that $|v| \geq 7$. Thus, t_β would contain a factor $vf'(v)g(v)$ with f antimorphic permutation, g morphic permutation, and $|v| \geq 7$, a contradiction.

For $11 \leq |u| \leq 30$ we can check with the computer that the conclusion holds. Indeed, this means that we must check all the factors of length at most 90 and see whether they are of the form $vf(v)g(v)$ or not. But any factor of t_η of length at most 90 is a factor of $\eta(w)$ where w is a factor of length 3 of t, so our check can be done quite fast. This shows that the first statement of the lemma holds.

As in the proof of Lemma 2, the second statement can be easily checked by computer, as well. $\square$

Proof of Lemma 11:
If $i < j$, we take $y = \pi^i(x)$ and the pattern $\pi^i(x)\pi^j(x)\pi^i(x)$ becomes $y\pi^{j-i}(y)y$. The latter pattern is avoidable if and only if $y\pi(y)y$ is avoidable, where the instances of π are morphisms if $j - i$ is even and antimorphisms if $j - i$ is odd. In both cases the pattern is avoidable by Lemma 1 or Lemma 7, respectively. If $i > j$. we take $y = \pi^j(x)$ and we have that $\pi^i(x)\pi^j(x)\pi^i(x)$ becomes $\pi^{i-j}(y)y\pi^{i-j}(y)$. This pattern is avoidable if and only if $\pi(y)y\pi(y)$ is avoidable, where the instances of π are morphisms if $i - j$ is even and antimorphisms if $i - j$ is odd. In both cases the pattern is avoidable by Lemma 4 and Lemma 1 or 7 respectively. □

Proof of Lemma 12:
Clearly, the remarks we made at the beginning of Lemma 6 are valid in the antimorphic case as well. We now distinguish different cases depending on the parity of i and j. If both i and j are even, then for every antimorphic permutation f there exists a morphic permutation f' such that $f^i(u) = f'^i(u)$ and $f^j(u) = f'^j(u)$ for all $u \in \Sigma_m^+$ and we can apply Lemma 6. So let us assume first that i is odd and j is even.

By trying to construct an infinite word over Σ_m that avoids the pattern if $m \geq \max\{k_1, k_2\}$, we quickly notice that such a word can not exist. In fact, the longest word without an occurrence of such a pattern is 001010101, which is of length nine. If $m \geq \max\{k_1, k_3\}$ the construction stops even earlier: In this case the longest prefix that avoids the pattern is of length five: 01010.

If $m \geq \max\{k_1, k_4\}$, we can not get a word of length larger than six without having an instance of the pattern. One of those longest words is 011002.

So in all cases we have seen that the pattern $x\pi^i(x)\pi^j(x)$ is unavoidable in Σ_m with $m \geq k$ if i is odd and j is even. The cases when i is even while j is odd and when both i and j are odd are similar and lead to the same results; therefore, the analysis of these is left to the reader. □

Proof of Proposition 4:
Since we already examined the case $m \geq k$ in Lemma 12, it only remains to be seen which is the situation for alphabets with less than k letters.

The cases when $m = 2$ and $m = 3$ are exactly like those depicted in Table 1 for the morphic case. As in the case of morphic permutations, the pattern $x\pi^i(x)\pi^j(x)$ is avoidable in Σ_2 if and only if $i \equiv j \equiv 0(mod\ 2)$, and in that case it is avoided by the Thue-Morse word. In the case of Σ_3, if $j \neq 0$, an instance of the pattern would contain either squares or cubes, so it would be avoided by h or, respectively, t. If $j = 0$, we use the word defined in Lemma 7 to obtain the avoidability of the pattern.

The analysis of the case when $m = 4$ is more involved. First, note that it is sufficient to know how to decide the avoidability of the pattern $x\pi^i(x)\pi^j(x)$ for $i, j < 12$. Indeed, it is not hard to see that if i and j are arbitrary natural numbers, then $x\pi^i(x)\pi^j(x)$ is avoidable in Σ_4 if and only if $x\pi^{i'}(x)\pi^{j'}(x)$ is avoidable, for i' (resp. j') being the remainder of i (resp. j) divided by 12.

Consequently, we only analyse the cases when $i, j < 12$. The pattern is clearly unavoidable whenever the value k computed for i and j in 5 is less than or equal to 4. When $i = 0$ the pattern $x\pi^i(x)\pi^j(x)$ is avoided by the word h as any instance of the pattern contains squares, and when $j = 0$ the pattern is avoided by the word from Lemma 7. Also, in the case when i and j are both even we can decide the avoidability of the pattern using the results obtained for morphisms in the previous sections, as, in this case, f can be seen as a morphism instead of an antimorphism. Moreover, when $i = j$ we can avoid the pattern $x\pi^i(x)\pi^i(x)$ by the word h that contains no squares. The same word h avoids the pattern in the cases when $(i, j) \in \{(4, 1), (9, 1), (8, 5), (9, 5), (3, 7), (4, 7), (3, 11), (8, 11)\}$. To complete the picture, we note that a word avoids the pattern $x\pi(x^R)x^R$ if and only if it avoids the pattern $x\pi'(x)x^R$ where π' is mapped to a morphic permutation. Therefore, by Lemma 8 we obtain that the pattern $x\pi^i(x)\pi^j(x)$ is avoided by the infinite word h_δ for $(i, j) \in \{(4, 3), (8, 3), (4, 9), (8, 9)\}$ and by Corollary 1 we obtain that it is avoidable for $(i, j) \in \{(7, 3), (11, 3), (1, 9), (5, 9)\}$.

Further, the discussion is split in four cases. If both i and j are even, we can decide the avoidability of the pattern just as in the case of morphisms (as the instance of π can be seen, in fact, as a morphism). If both i and j are odd, we compute the value k defined in (5) and define $M = \max\{k, j + 1, i + 1\}$. Now, $x\pi^i(x)\pi^j(x)$ is avoidable in Σ_m if and only if $(x\pi^i(x)\pi^j(x))^R = \pi^j(x^R)\pi^i(x^R)x^R$ is avoidable in Σ_m. The last condition is equivalent to the avoidability of the pattern $\pi^j(y)f^i(y)y$ in Σ_m. Taking $z = \pi^j(y)$, we obtain that $\pi^j(y)\pi^i(y)y$ is avoidable in Σ_m if and only if $z\pi^{M!-j+i}(z)\pi^{M!-j}(z)$ is avoidable in Σ_m. Now we only have to notice that $M! - j + i$ is even and $M! - j$ is odd, as $M!$ is always even. Therefore, the case when i and j are odd can be reduced to the case when i is even and j is odd.

So there remain only two cases to be analysed: the case when i is even and j is odd as well as the case when i is odd and j is even. As in the morphic case, we look at the minimum of k_1, k_2, k_3 and k_4:

Case 1: $k_1 = \min\{k_1, k_2, k_3, k_4\}$. This means that $k > k_1$ and for $m < k_1$ we get that m divides both i and j. For every letter $a \in \Sigma_m$ and antimorphic permutation f of Σ_m, since $\mathbf{ord}_f(a) \leq m$, we get that $\mathbf{ord}_f(a)$ divides both i and j. Thus, every instance of $x\pi^i(x)\pi^j(x)$ is in fact an instance of xx^Rx when i is odd and j is even or an instance of xxx^R when i is even and j is odd. Those patterns are avoided by the word t_γ of Lemma 7 or the word h respectively. If $k_1 \leq m < k$, then for every $a \in \Sigma_m$ and antimorphic permutation f of Σ_m we either have that $\mathbf{ord}_f(a)$ divides both i and j or it divides neither i nor j nor $|i - j|$. If there is no letter that fulfils the latter case we get that the pattern is actually xxx^R (resp. xx^Rx) if i is odd (resp. even) and j is even (resp. odd) and we can avoid it by the word h (resp. the word t_γ from Lemma 7). Otherwise we get that there have to be at least 3 different letters in an instance of this pattern and this is obviously avoided by the Thue-Morse word.

Case 2: $k_2 = \min\{k_1, k_2, k_3, k_4\}$. In this case we can see that $k = k_1$. If $4 \leq m < k_2$, then for every $a \in \Sigma_m$ and antimorphic permutation f of Σ_m we get that $\mathbf{ord}_f(a)$ divides both i and j (since $k_3, k_4 > k_2$) and we get again that the pattern becomes xx^Rx (resp. xxx^R) when i is odd (resp. even) and j is even (resp. odd), which is avoided by the word h (resp. the word t_γ from Lemma 7). If $k_2 \leq m < k$ we get for each letter $a \in \Sigma_m$ and antimorphic permutation f of Σ_m that $\mathbf{ord}_f(a)$ divides at least one of i, j and $|i - j|$. Thus, for a factor $uf^i(u)f^j(u)$, at every position of $\ell \leq |u|$ there are at most 2 different letters appearing in $u, f^i(u), f^j(u)^R$ if i is even and j is odd (resp. in $u, f^i(u)^R, f^j(u)$ if i is odd and j is even). Such factors are avoided by the words of Lemma 9 and Lemma 10.

Case 3: $k_3 = \min\{k_1, k_2, k_3, k_4\}$. Again we get $k = k_1$. If $4 \leq m < k_3$ we get that $\mathbf{ord}_f(a)$ divides both i and j for every letter $a \in \Sigma_m$ and antimorphic permutation f of Σ_m. So every instance of $x\pi^i(x)\pi^j(x)$ is in fact an instance of xxx^R (resp. xx^Rx) if i is even (resp. odd) and j is odd (resp. even) and thus avoided by the word h (resp. the word t_γ from Lemma 7). If $k_3 \leq m < k$, we observe that $\mathbf{ord}_f(a)$ divides at least one of i and j for every letter $a \in \Sigma_m$ and antimorphic permutation f of Σ_m. Again we get that for a factor $uf^i(u)f^j(u)$ at every position of $\ell \leq |u|$ there are at most two different letters appearing in $u, f^i(u), f^j(u)^R$ if i is even and j is odd (resp. in $u, f^i(u)^R, f^j(u)$ if i is odd and j is even) and such factors do not appear in the words of Lemma 9 and Lemma 10.

As in the morphic case, the situation when $k_4 = \min\{k_1, k_2, k_3, k_4\}$ is symmetric to the previous case and therefore the same results hold. $\square$