RAIRO-Theor. Inf. Appl. **41** (2007) 403–409
DOI: 10.1051/ita:2007018

DEFECT THEOREM IN THE PLANE

Włodzimierz Moczurad[1]

Abstract. We consider the defect theorem in the context of labelled polyominoes, *i.e.*, two-dimensional figures. The classical version of this property states that if a set of n words is not a code then the words can be expressed as a product of at most $n-1$ words, the smaller set being a code. We survey several two-dimensional extensions exhibiting the boundaries where the theorem fails. In particular, we establish the defect property in the case of three dominoes ($n \times 1$ or $1 \times n$ rectangles).

Mathematics Subject Classification. 68Q70, 68R15.

Introduction

The defect theorem is one of the fundamental results of combinatorics on words, providing a kind of dimension property of words (*cf.* Lothaire [10, 11]). Various authors have studied its variants and extensions to other structures, including trees and two-dimensional figures, *e.g.* [7–9, 12, 13]. See Harju and Karhumäki [5] for a comprehensive treatment of the field.

In its classical version, the defect theorem states that if $X \subseteq A^*$ is a finite non-code, *i.e.*, there exists a word in X^* with two different X-factorizations, then there exists a code $Y \subseteq A^*$ such that $X \subseteq Y^*$ and $|Y| < |X|$. More precisely, Y can be taken to be the free hull of X (the smallest free submonoid of A^* containing X).

Whilst it has been shown that the defect property can be extended to trees, the property is not satisfied in many simple cases of plane figures. In Section 2 we give several counterexamples for squares, rectangles and dominoes. We also mention here the fact that sets of two rectangles (as well as squares and dominoes) do possess the defect property. This allows us to exhibit some of the boundary set sizes where the theorem fails.

Keywords and phrases. Defect theorem, codes, polyominoes.

[1] Institute of Computer Science, Jagiellonian University, Nawojki 11, 30-072 Kraków, Poland; `wkm@ii.uj.edu.pl`

In Section 3 we prove the defect theorem for sets of three dominoes ($n \times 1$ or $1 \times n$ rectangles), thus solving a problem stated in [5] and establishing the exact bound for the set size where the property fails.

1. Definitions and notations

Let A be a finite alphabet. We use the usual notation of A^* to denote the free monoid over A, and X^* to denote the submonoid generated by $X \subseteq A^*$. A set of words $X \subseteq A^*$ is a *code,* if X^* is free over X, *i.e.*, every word in X^* has a unique factorization over X.

A *labelled polyomino,* also called a *figure* or a *brick,* is a partial mapping $x : \mathbb{Z}^2 \to A$, where the domain of x is a polyomino (a finite and connected subset of $\mathbb{Z}^2$). Points in the domain of x are referred to as *cells.* The set of all figures over A is denoted by $A^{\bowtie}$.

Given a set of figures $X \subseteq A^{\bowtie}$, the set of all figures tilable with (translated copies of) the elements of X is denoted by $X^{\bowtie}$. Note that we do not allow rotations of figures. $X \subseteq A^{\bowtie}$ is a *code,* if every element of $X^{\bowtie}$ admits exactly one tiling with the elements of X.

The terms *rectangles, squares* and *dominoes* will refer to figures with respective domains. In particular, the domain of a domino is a $1 \times n$ or $n \times 1$ rectangle with $n \geq 1$ (a *vertical domino* and a *horizontal domino,* respectively). Note that since we actually consider all figures up to a translation, horizontal dominoes can be identified with words.

2. In the plane

In this section we give counterexamples for several combinations of figure shapes and set sizes, *i.e.*, there exist non-codes composed of figures of a specified shape that cannot be tiled with fewer figures of the same shape. We also recall a positive result, the defect theorem for two rectangles, squares and dominoes (*cf.* [5, 17]).

More formally, the defect property for figures of class $\mathcal{C} \subseteq \mathbb{Z}^2$ means the following: let $X \subseteq A^{\bowtie}$, with the domains of all $x \in X$ belonging to $\mathcal{C}$, be a finite non-code. Then there exists a code $Y \subseteq A^{\bowtie}$, with the domains of all $y \in Y$ belonging to $\mathcal{C}$, such that $X \subseteq Y^{\bowtie}$ and $|Y| < |X|$.

In general, *i.e.*, for sets of arbitrary cardinality, the property holds neither for unrestricted shapes ($\mathcal{C} = \mathbb{Z}^2$), nor for $\mathcal{C}$ corresponding to rectangles, squares or dominoes.

Theorem 2.1. *Let $X = \{k, l\} \subseteq A^{\bowtie}$ be a non-code containing two rectangles. Then there exists a common rectangular tiler for k, l, i.e., a rectangle $t \in A^{\bowtie}$ such that $k, l \in \{t\}^{\bowtie}$.*

Proof. Since X is not a code, there exists $y \in X^{\bowtie}$ such that y admits two different tilings with the elements of X.

Take the leftmost cell in the topmost row of y. If both tilings use the same rectangle $x \in \{k, l\}$ to cover this cell, then remove x from y and repeat this procedure until y' is obtained such that the two tilings place different rectangles in the leftmost cell in the topmost row of y'. In other words, we take y to be a minimal figure admitting two tilings over X. Call the tilings T_k and T_l, according to the rectangle being used to cover the cell specified above.

Denote the (horizontal) widths of k and l by $w(k)$ and $w(l)$, and assume *e.g.* $w(k) < w(l)$; the argument becomes trivial when $w(k) = w(l)$. Now this implies that k tiles the top-left corner of l. Moving rightwards along the top edge of l, we observe that when T_k covers the remaining part of l, it cannot place a tile higher than the top edge of l, since we have started in the topmost row. Thus the next tile to the right of k has to be aligned along the top edge. Since we already know that l contains a copy of k in its top-left corner, we have another copy of k "along the way" (although l may have actually been used). Continuing in this way, we observe that the two tilings will eventually arrive at a common right edge when they reach the lowest common multiple of the widths of k and l.

$$
\begin{array}{ll}
T_l\colon & \overbrace{\boxed{k_1 \ \dots\ k_n\ k_1\ \dots\ k_i}}^{l}\ k_1\quad k_2\quad \dots\ k_{n-i} \\
T_k\colon & \underbrace{\boxed{k_1\ \dots\ k_n}}_{k}\ k_1\ \dots\ k_i\ \ k_{i+1}\,k_{i+2}\ \dots\ k_n
\end{array}
$$

Considering the situation to the right of the first copy of l in T_l and denoting the columns of k by $k_1, k_2, ..., k_n$ we obtain $k_1 = k_{i+1}$, $k_2 = k_{i+2}$, ..., $k_{n-i} = k_n$ where $i = w(l) \bmod w(k)$ (*cf.* figure above). If $i = 0$, then $w(k)$ is the width of the common tiler. Otherwise this width is equal to i.

The above argument can be repeated in vertical direction, thus giving a rectangle that tiles both k and l.

Note that the above reasoning is still valid when y is not a rectangle, *e.g.*, the leftmost cell in the topmost row of y is not the leftmost cell of the whole figure. We observe that the width of any part of y protruding to the left has to be a multiple of the width of the common tiler.

Also note that the proof becomes trivial if the effective alphabet of X is just one symbol, since X is never a code then with, *e.g.*, the unit square being the common tiler for k and l. ☐

Corollary 2.2. *Let $X = \{k, l\} \subseteq A^{\bowtie}$ be a non-code containing two squares. Then there exists a common square tiler for k, l, i.e., a square $t \in A^{\bowtie}$ such that $k, l \in \{t\}^{\bowtie}$.*

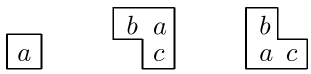
Proof. By the Theorem above, there exists a rectangle t such that $k, l \in \{t\}^{\bowtie}$. Assume that t is of size $p \times q$ and let r be the lowest common multiple of p and q. Both $w(k)$ and $w(l)$ are multiples of r. Hence, a square formed by stacking together

r/p copies of t horizontally and then r/q copies of this compound vertically can be taken as the common square tiler. $\square$

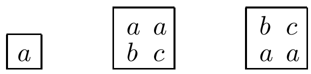
Corollary 2.3. *Let $X = \{k, l\} \subseteq A^{\bowtie}$ be a non-code containing two dominoes. Then there exists a common domino tiler for k, l, i.e., a domino $t \in A^{\bowtie}$ such that $k, l \in \{t\}^{\bowtie}$.*

The following counterexamples disprove the defect property for unrestricted figures, squares, rectangles and dominoes.

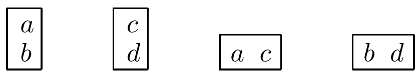
Example 2.4. The following non-code containing three figures cannot be tiled with two figures.



Example 2.5. The following non-code containing three squares cannot be tiled with two squares. Note, however, that it can be tiled with two rectangles. It remains open whether the defect property holds in the more general setting of three rectangles.



Example 2.6. The following non-code containing four rectangles/dominoes cannot be tiled with three rectangles.



Notice that if the defect property – within a certain class of figures – fails at a certain set size, it also fails for bigger sets. If an n-set cannot be tiled by fewer than n figures, an $(n+1)$-set that cannot be tiled by fewer than $n+1$ figures may be constructed by adding a figure labelled with a new alphabet symbol.

3. Defect theorem for sets of three dominoes

We now prove the defect theorem for sets of three dominoes. The proof is directly combinatorial, resorting to the defect theorem for words in its basic cases.

Theorem 3.1. *Let $X \subseteq A^{\bowtie}$ be a set of three dominoes. If X is not a code then there exists a code $Y \subseteq A^{\bowtie}$ such that $|Y| < 3$ and Y tiles the dominoes of X.*

Proof. If $|A| < 3$ then $Y = A$ may be taken, since unit square(s) labelled with the letter(s) of A tile X. Thus assume that the size of the alphabet A is at least 3, with letters $a, b, c \in A$ each appearing at least once in X.

If X contains three horizontal dominoes or three vertical dominoes, use the defect theorem for words to obtain the proper Y immediately. Thus assume that X contains two horizontal dominoes and one vertical domino. The case of two vertical and one horizontal domino can be dealt with similarly.

If one of 2-subsets of X is not a code, use the defect theorem in its two-rectangle version (see Sect. 2) to obtain a single tiler for the subset. The tiler and the remaining domino can now be taken as Y (this procedure can be repeated should Y fail to be a code). Thus assume that all 2-subsets of X are codes.

Since X is not a code, there exists a minimal (in the sense of domain inclusion) figure F with two different X-factorizations. We will refer to the two factorizations as "layer 1" and "layer 2". We will consider rows of the figure defined as maximal horizontal words contained within its domain.

Notice that an intersection with the vertical domino has to appear in every row, in at least one of the layers. Otherwise, one of the rows would have two different factorizations with just the horizontal dominoes, contradicting the assumption that all 2-subsets of X are codes.

Now consider the following cases.

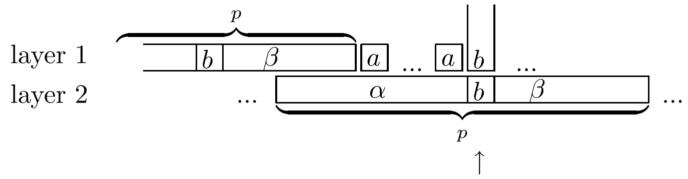
Case 1. The vertical domino is labelled with one letter, *e.g.*, it is of the form $(a^k)^T$.

Consider any row of the figure F. Since it intersects with the vertical domino, it is tiled by the horizontal dominoes and the unit square a. By the defect theorem for words, the horizontal dominoes (and a) are tiled by two words, one of which has to be a. Call the other one p. Thus, $Y = \{p, a\}$ tiles all three dominoes.

Case 2. The vertical domino is of the form $(a^k b ...)^T$ with $k > 0$.

Consider the topmost, leftmost row of the figure F. Similarly as in case 1, there exists a word $p \in A^*$ such that $Y = \{p, a\}$ tiles the horizontal dominoes (and a). Now consider the $(k+1)$-th row of the figure F which intersects with the (first from the left) vertical domino on b. Assume that layer 1 contains the b of the vertical domino (this is the point marked with an arrow in the figure below). Layer 2 covers it with a b belonging to one of the horizontal dominoes, and hence to p. Notice that this is the first b of the vertical domino and thus cannot be covered in layer 2 by another copy of the vertical domino. This implies that p can be written as $\alpha b \beta$ with $\alpha, \beta \in A^*$.

Now if c appears in α and not in β (or vice versa), there is no way to match the c's in the two layers. Thus c appears in both α and β. Denote the number of occurrences of c in α and β as n_α and n_β, respectively.



Only p and a are used to the left of $\uparrow$. Let n_1 and n_2 denote the number of times p is used in the respective layers. Counting the occurrences of c to the left of $\uparrow$ we obtain $n_1(n_\alpha + n_\beta) = n_2(n_\alpha + n_\beta) + n_\alpha$. Taking $m = n_1 - n_2$ we get $mn_\beta = (1-m)n_\alpha$ which clearly has no integer solution for m. Thus case 2 is impossible if c appears in p.

Finally, note that in all cases Y is a code. This concludes the proof of the defect theorem for three dominoes. □

Notice that this proof tells us something more about 3-domino non-codes: they are either trivial or they are essentially word non-codes.

4. Final remarks

Status of the defect property for various combinations of figure shapes and set sizes is summarized in the table below. This includes open problems, marked with question marks.

Figures/set size	2	3	≥ 4
Squares	+	−	−
Dominoes	+	+	−
Rectangles	+	?	−
Unrestricted	?	−	−

References

[1] P. Aigrain, and D. Beauquier, Polyomino tilings, cellular automata and codicity. *Theoret. Comput. Sci.* **147** (1995) 165–180.

[2] D. Beauquier, and M. Nivat, A codicity undecidable problem in the plane. *Theoret. Comput. Sci.* **303** (2003) 417–430.

[3] J. Berstel, and D. Perrin, *Theory of Codes.* Academic Press (1985).

[4] V. Bruyère, Cumulative defect. *Theoret. Comput. Sci.* **292** (2003) 97–109.

[5] T. Harju, and J. Karhumäki, Many aspects of the defect effect. *Theoret. Comput. Sci.* **324** (2004) 35–54.

[6] J. Karhumäki, Some open problems in combinatorics of words and related areas. *TUCS Technical Report* **359** (2000).

[7] J. Karhumäki, and S. Mantaci, Defect Theorems for Trees. *Fund. Inform.* **38** (1999) 119–133.

[8] J. Karhumäki, and J. Maňuch, Multiple factorizations of words and defect effect. *Theoret. Comput. Sci.* **273** (2002) 81–97.

[9] J. Karhumäki, J. Maňuch, and W. Plandowski, A defect theorem for bi-infinite words. *Theoret. Comput. Sci.* **292** (2003) 237–243.

[10] M. Lothaire, *Combinatorics on Words.* Cambridge University Press (1997).

[11] M. Lothaire, *Algebraic Combinatorics on Words.* Cambridge University Press (2002).

[12] S. Mantaci, and A. Restivo: Codes and equations on trees. *Theoret. Comput. Sci.* **255** (2001) 483–509.

[13] J. Maňuch, Defect Effect of Bi-infinite Words in the Two-element Case. *Discrete Math. Theor. Comput. Sci.* **4** (2001) 273–290.

[14] W. Moczurad, Algebraic and algorithmic properties of brick codes. *Ph.D. Thesis*, Jagiellonian University, Poland (2000).

[15] W. Moczurad, Brick codes: families, properties, relations. *Intern. J. Comput. Math.* **74** (2000) 133–150.
[16] M. Moczurad, and W. Moczurad, Decidability of simple brick codes, in *Mathematics and Computer Science III (Algorithms, Trees, Combinatorics and Probabilities), Trends in Mathematics.* Birkhäuser (2004).
[17] M. Moczurad, and W. Moczurad, Some open problems in decidability of brick (labelled polyomino) codes, in *Cocoon 2004 Proceedings. Lect. Notes Comput. Sci.* **3106** (2004) 72–81.

Communicated by J. Karhumäki.
Received January 16, 2006. Accepted December 7, 2006.