Balance properties of multi-dimensional words

Valérie Berthé* and Robert Tijdeman†

Abstract A word u is called 1-balanced if for any two factors v and w of u of equal length, we have $-1 \leq |v|_i - |w|_i \leq 1$ for each letter i, where $|v|_i$ denotes the number of occurrences of i in the factor v. The aim of this paper is to extend the notion of balance to multi-dimensional words. We first characterize all 1-balanced words on $\mathbb{Z}^n$. In particular we prove they are fully periodic for $n > 1$. We then give a quantitative measure of balancedness for some words on $\mathbb{Z}^2$ with irrational density, including two-dimensional Sturmian words.

1 Introduction

A word u is called 1-balanced if for any two factors (=finite subwords) v and w of u of equal length, we have $-1 \leq |v|_i - |w|_i \leq 1$ for each letter i. Here $|v|_i$ denotes the number of occurrences of i in the factor v. This notion has been widely studied from many points of view, for instance in ergodic theory [6], in number theory [27], in theoretical computer science [7], and in operations research [1]. In the literature a 1-balanced word is usually called balanced. We shall use the term balanced for a weaker property.

1-Balanced words were first studied by Morse and Hedlund in their seminal papers [20], [21] published in 1938 and 1940. They studied 1-balanced words from a two-letter alphabet with as possible domains the integers, the positive integers and a finite interval I. We call such words $\mathbb{Z}$-words, $\mathbb{N}$-words and F-words, respectively. It is now known [12] that every 1-balanced $\mathbb{N}$-word and every 1-balanced F-word is a subword of a 1-balanced $\mathbb{Z}$-word. A classification of all 1-balanced $\mathbb{Z}$-words therefore induces classifications of 1-balanced $\mathbb{N}$-words and F-words. Morse and Hedlund showed that each

*Institut de Mathématiques de Luminy, CNRS-UPR 9016, Case 907, 163 avenue de Luminy, F-13288 Marseille Cedex 9, France; e-mail: berthe@iml.univ-mrs.fr

†Mathematical Institute, Leiden University, Postbus 9512, 2300 RA Leiden, The Netherlands; e-mail: tijdeman@math.leidenuniv.nl

letter of a 1-balanced word has a density (=asymptotic frequency). They further proved that there are three subclasses of 1-balanced $\mathbb{Z}$-words: periodic, irrational and skew. The letters of irrational words have irrational densities, the others have rational densities. For the reader's convenience we give the complete classification in Section 3.1.

1-Balanced words have the remarkable property that they can be characterized in some seemingly independent ways (cf. [18]). For 1-balanced $\mathbb{Z}$-words and 1-balanced $\mathbb{N}$-words with letters 0 and 1 where the letter 1 has density α we have as alternative definitions:
(a) the codings of a rotation of an (open, half-open or closed) interval of length $2\pi\alpha$ on the unit circle (so-called *interval exchange)*
(b) in case α is irrational, the words for which the number of distinct factors of length n is at most $n+1$ for every positive integer n *(low-complexity words)*. 1-Balanced words with rational density have also such low complexity.
Moreover, the periodic and irrational 1-balanced words with density $\alpha > 0$ are precisely
(c) the words for which there is some real constant β such that the symbol at place n either equals $\lfloor (n+1)\alpha + \beta \rfloor - \lfloor n\alpha + \beta \rfloor$ for all n, or $\lceil (n+1)\alpha + \beta \rceil - \lceil n\alpha + \beta \rceil$ for all n, *(Beatty sequences)*
(d) the words for which there exists a real constant β such that the letters 1 occur at the places $\lfloor n\alpha^{-1} + \beta \rfloor$ for all n or at the places $\lceil n\alpha^{-1} + \beta \rceil$ for all n.
The original definition of 1-balanced word is of conceptual interest in queueing theory. Because of (a) the 1-balanced irrational $\mathbb{N}$-words, often called *Sturmian words* or *Sturmian sequences*, play an important role in ergodic theory. Characterization (b) is more relevant for theoretical computer science and characterizations (c) and (d) for number theory.

It is easy to create from each 1-balanced word to a given alphabet a 1-balanced word to a larger alphabet by replacing one letter cyclicly by some other letters (see for instance [14, 26]). By doing so, the two least frequent letters in the new word have the same densities. It is a classical open problem to characterize all 1-balanced words to finite alphabets where the letters have distinct densities. For two-letter words this boils down to determining all 1-balanced words of density $\neq 1/2$. According to Fraenkel's conjecture, for every $m > 2$ there is essentially only one 1-balanced word from m letters with distinct densities. This has been established for $m = 3, 4, 5, 6$ in [1, 27], cf. [14, 28]. Note that on the other hand we can reduce 1-balanced words from m letters to 1-balanced words from 2 letters by identifying all but one letters.

Some authors have generalised the notion of 1-balance to C-balance by requiring $-C \leq |v|_i - |w|_i \leq C$, for all factors v, w with $|v| = |w|$, and for each letter i, where C is some constant. It was believed that Arnoux-Rauzy sequences [2], and more generally Episturmian words to a 3-letter alphabet [8], would be 2-balanced. An example of an Arnoux-Rauzy sequence is constructed in [6] which is balanced, that is, not C-balanced for any $C > 0$. As an application of this result, the authors of [6] deduce that there exist Arnoux-Rauzy sequences which are not natural codings of a rotation on the two-dimensional torus. Similar studies have been made for interval exchanges in [30]. We note that C-balancedness for $C > 1$ without further assumption seems to be uninteresting because of the following observation. The number of 1-balanced words of length n is polynomial in n [17, 19] and being 1-balanced is therefore rare. The number of C-balanced words of length n for $C > 1$ is exponential in n [17, 11] and therefore being C-balanced is relatively common. At the end of Section 3 we shall show a similar phenomenon for higher dimensions.

The aim of this paper is to extend the notion of balance to multi-dimensional words. Instead of intervals we require now that the number of occurrences of each letter in two equal (rectangular) blocks differ by at most 1. In Section 3 we shall characterize all 1-balanced words on $\mathbb{Z}^n$. The following result is a consequence of our theorems. The density of a 1-balanced word to $\{0, 1\}$ denotes here the asymptotic frequency of the letter 1.

Corollary *Let $f : \mathbb{Z}^n \to \{0, 1\}$ be 1-balanced. Then f has a density α. Furthermore when $n = 2$,*

$$\alpha \in \{0, \frac{1}{5}, \frac{1}{3}, \frac{2}{5}, \frac{1}{2}, \frac{3}{5}, \frac{2}{3}, \frac{4}{5}, 1\}.$$

When $n \geq 3$, then $\alpha \in \{0, \frac{1}{3}, \frac{1}{2}, \frac{2}{3}, 1\}$. If $\alpha \neq 0, 1$, then f is fully periodic.

Here fully periodic means that the $\mathbb{Z}$-module of periods has full rank. Note that there are no 1-balanced $\mathbb{Z}$-words with irrational density in the multi-dimensional case. Note that not every 1-balanced $\mathbb{Z}$-word can be extended to a 1-balanced $\mathbb{Z}^2$-word, since the Sturmian words cannot be extended to 1-balanced words on $\mathbb{Z}^2$. In fact, we shall characterize all 1-balanced words on all infinite intervals and it will turn out that there exist irrational words in multi-dimensional intervals if and only if the dimension is 2 and the width of the strip is 2 (cf. Theorem 3). In all other cases the behaviour is as in the corollary (cf. Theorem 4). The 1-balanced words in dimension $2, 3$, and greater than 3 are given in Sections 3.2, 3.3 and 3.4, respectively.

All the examples we provide here have rational density. It is thus natural to ask whether there do exist balanced words on $\mathbb{Z}^2$ to a two-letter alphabet with irrational density. There is a natural candidate to be simultaneously balanced and of irrational density: consider a word on $\mathbb{Z}^2$ defined by shifting from row to row a given Sturmian word on $\mathbb{Z}$ with the periodicity vector $(1, -1)$. We prove in Section 4.1 that such a word cannot be balanced. The proof is based on the estimates of [5] for $\sum_{1 \le j \le N}(\{j\alpha\} - 1/2)$, involving Ostrowski's numeration system (where $\{x\}$ denotes the fractional part of x, that is $x - \lfloor x \rfloor$). This leads us to ask whether the densities of the letters of a balanced word on $\mathbb{Z}^2$ are rational.

We also consider in Section 4.2 balance properties for two-dimensional Sturmian words. These words correspond to the approximation of a plane by a discrete plane [29]. We prove they are not balanced and we provide a quantitative measure of their non-balancedness, based on the expression of [23] for $\sum_{1 \le j \le N}(\{j\alpha + \beta\} - 1/2)$. Consequently these two-dimensional words do not fully generalise all the properties of classical Sturmian words.

2 Definitions and basic results

2.1 Balance properties

Let n be a positive integer. Let $\vec{a}, \vec{b} \in \mathbb{Z}^n, \vec{a} = (a_1, ..., a_n), \vec{b} = (b_1, ..., b_n)$. We write $\vec{a} < \vec{b}$ if $a_i < b_i$ for $i = 1, ..., n$. By a *block*, we mean the set of the vectors $(x_1, \ldots, x_n) \in \mathbb{Z}^n$ satisfying $a_i \le x_i < b_i$ for $i = 1, \ldots, n$. We denote the block by $[\vec{a}, \vec{b})$ or $[\vec{a}, \vec{b} - \vec{1}]$ or $\prod_{j=1}^{n}[a_i, b_j)$ and denote its volume $\prod_{i=1}^{n}(b_i - a_i)$ by $|\vec{a}, \vec{b}|$. Further $|\vec{a}, \vec{b}|_i$ denotes the number of letters (=function values) i in the block $[\vec{a}, \vec{b})$. In the two-dimensional case an m by n block means a block $[\vec{a}, \vec{b})$, where $b_1 - a_1 = m$ and $b_2 - a_2 = n$. We also denote this block by $[a_1, b_1 - 1] \times [a_2, b_2 - 1]$.

We call I an *(n-dimensional) interval* if I is the Carthesian product $I_1 \times I_2 \times \ldots \times I_n$ where I_i is $\mathbb{Z}$ or $\mathbb{N}$ or some finite interval of integers for $i = 1, 2, ..., n$. In the sequel I is such an interval. If I is not a block, it is called an *infinite interval.*

We call $f : I \to \{0, 1\}$ a *1-balanced word* if the numbers of vectors $\vec{x}$ with $f(\vec{x}) = 1$ in any two subblocks (=factors) $[\vec{a}, \vec{a} + \vec{c}), [\vec{b}, \vec{b} + \vec{c})$ of I differ by at most 1.

Let $\mathcal{A}$ be a finite alphabet. Let $C > 0$. We call $f : I \to \mathcal{A}$ *C-balanced* *on the letter* $i \in \mathcal{A}$ if the numbers of vectors $\vec{x}$ with $f(\vec{x}) = i$ in any two subblocks $[\vec{a}, \vec{a} + \vec{c}), [\vec{b}, \vec{b} + \vec{c})$ of I differ by at most C.

A function f is *balanced* if there exists a constant C such that f is C-balanced for every letter from the alphabet $\mathcal{A}$.

2.2 Density

We say that $f : I \to \{0, 1\}$ has *density* α if the quotient of the number of vectors $\vec{x}$ with $f(\vec{x}) = 1$ in $[\vec{a}, \vec{b})$ and the volume $|\vec{a}, \vec{b}|$ tends to α when $[\vec{a}, \vec{b})$ runs in any way through a non-decreasing sequence of blocks with union I. We similarly define the density of the letter $i \in \mathcal{A}$ in the word $f : I \to \mathcal{A}$ to be α if the limit of the quotient of the number of vectors $\vec{x}$ with $f(\vec{x}) = i$ in $[\vec{a}, \vec{b})$ and the volume $|\vec{a}, \vec{b}|$ tends to α when $[\vec{a}, \vec{b})$ runs in any way through a non-decreasing sequence of blocks with union I.

If f is C-balanced, then the number of letters 1 in some block $[\vec{a}, \vec{a} + \vec{c})$ is in some interval $[N, N + C]$, where N is an integer independent of $\vec{a}$. We denote the largest integer N with this property by $N_{\vec{c}}$.

Let I be an infinite interval and $f : I \to \{0, 1\}$ 1-balanced. Without loss of generality we may assume that $\mathbb{N} \subseteq I_1$. We shall do so tacitly in the sequel. We first show that $N_{\vec{c}}$ depends only on the volume $|\vec{0}, \vec{c}|$ of the block, subsequently that every 1-balanced function has some density α and finally that $N_{\vec{c}} = \lfloor \alpha \times |\vec{0}, \vec{c}| \rfloor$ or $N_{\vec{c}} = \lceil \alpha \times |\vec{0}, \vec{c}| \rceil - 1$.

Lemma 1 *Let $\vec{c} > \vec{0}, \vec{d} > \vec{0}$ such that for some $\vec{a}$ blocks $[\vec{a}, \vec{a}+\vec{c})$ and $[\vec{a}, \vec{a}+\vec{d})$ fit in I. If f is 1-balanced and $|\vec{0}, \vec{c}| = |\vec{0}, \vec{d}|$, then $N_{\vec{c}} = N_{\vec{d}}$.*

Proof. Suppose $|\vec{0}, \vec{c}| = |\vec{0}, \vec{d}|$. It suffices to prove the theorem with $\vec{d} = (|\vec{0}, \vec{c}|, 1, ..., 1)$. Let $\vec{c} = (c_1, \ldots, c_n), \vec{d} = (d_1, 1, \ldots, 1)$. Let $k \in \mathbb{N}$. Put $\vec{e} = (kc_1d_1, c_2, \ldots, c_n)$. Then the block $[\vec{b}, \vec{b} + \vec{e})$ can be partitioned into blocks of the form $[\vec{a}, \vec{a} + \vec{c})$ and into blocks of the form $[\vec{a}, \vec{a} + \vec{d})$, where $\vec{a}$ varies. Put $m = |\vec{b}, \vec{b} + \vec{e}|/|\vec{0}, \vec{c}|$. As $|\vec{0}, \vec{c}| = |\vec{0}, \vec{d}|$, we have $m = |\vec{b}, \vec{b} + \vec{e}|/|\vec{0}, \vec{d}|$ and $mN_{\vec{c}} \leq N_{\vec{e}} \leq m(N_{\vec{c}} + 1)$ and $mN_{\vec{d}} \leq N_{\vec{e}} \leq m(N_{\vec{d}} + 1)$. If $|[\vec{a}, \vec{a} + \vec{c})|_1$ is independent of $\vec{a}$, then $|[\vec{a}, \vec{a} + \vec{d})|_1$ is independent of $\vec{a}$ by 1-balancedness on $[\vec{b}, \vec{b} + \vec{e})$, whence $N_{\vec{c}} = N_{\vec{d}}$ by the maximality condition in the definition of $N_{\vec{c}}$. If $|[\vec{a}, \vec{a} + \vec{c})|_1$ is non-constant as a function of $\vec{a}$, then its values can be $N_{\vec{c}}$ and $N_{\vec{c}} + 1$. In this case $|[\vec{a}, \vec{a} + \vec{d})|_1$ assumes both the value $N_{\vec{d}}$ and $N_{\vec{d}} + 1$ and we conclude $N_{\vec{c}} = N_{\vec{d}}$ too. ∎

In the sequel we write N_k for $N_{\vec{c}}$ with $k := |\vec{0}, \vec{c}|$. The following lemma shows that N_k/k has some limit α.

Lemma 2 *If f is 1-balanced on I, then f has some density α.*

Proof. By Lemma 1 we may restrict our attention to blocks of the form $(*, 1, ..., 1)$. Let $c_1, d_1 \in \mathbb{Z}_{>0}$ and $\vec{c} = (c_1, 1, \ldots, 1)$, $\vec{d} = (d_1, 1, \ldots, 1) \in \mathbb{Z}^n_{>0}$. Put $\vec{e} = (c_1 d_1, 1, \ldots, 1)$. Then any subblock $[\vec{b}, \vec{b} + \vec{e})$ of I can be partitioned into blocks of the form $[\vec{a}, \vec{a} + \vec{c})$ and into blocks of the form $[\vec{a}, \vec{a} + \vec{d})$. Put $m_1 = |\vec{0}, \vec{e}|/|\vec{0}, \vec{c}|$ and $m_2 = |\vec{0}, \vec{e}|/|\vec{0}, \vec{d}|$. Then

$$m_1 N_{\vec{c}} \leq N_{\vec{e}} \leq m_1 (N_{\vec{c}} + 1)$$

and

$$m_2 N_{\vec{d}} \leq N_{\vec{e}} \leq m_2 (N_{\vec{d}} + 1).$$

Hence $-m_1 \leq m_1 N_{\vec{c}} - m_2 N_{\vec{d}} \leq m_2$. This implies

$$-\frac{1}{|\vec{0}, \vec{c}|} \leq \frac{1}{|\vec{0}, \vec{c}|} N_{\vec{c}} - \frac{1}{|\vec{0}, \vec{d}|} N_{\vec{d}} \leq \frac{1}{|\vec{0}, \vec{d}|}.$$

We infer that $(\frac{N_k}{k})_{k \in \mathbb{N}}$ is a Cauchy sequence and is therefore convergent to some limit, α say. Thus f has density α. ■

Lemma 3 *If f is 1-balanced and has density α, then $N_k = \lfloor k\alpha \rfloor$ or $\lceil k\alpha \rceil - 1$.*

Proof. Since a block of size k contains N_k or $N_k + 1$ ones, and on average $k\alpha$, we have $N_k \leq k\alpha$ and $N_k + 1 \geq k\alpha$. ■

We obtain similarly if f is C-balanced on the letter i.

Lemma 4 *Let $\mathcal{A}$ be a finite alphabet. Let $f : I \to \mathcal{A}$ be C-balanced on the letter i. Let $\vec{c}, \vec{d} \in \mathbb{Z}^n_{>0}$ such that for some $\vec{a}$ blocks $[\vec{a}, \vec{a} + \vec{c})$ and $[\vec{a}, \vec{a} + \vec{d})$ fit in I and that $|\vec{0}, \vec{c}| = |\vec{0}, \vec{d}|$. Then*

$$|N^i_{\vec{c}} - N^i_{\vec{d}}| \leq C - 1,$$

and the letter i has a density, α_i say, satisfying:

$$\forall \vec{N} \in \mathbb{N}^n, \; ||\vec{0}, \vec{N}|_i - \alpha_i |\vec{N}|| \leq C.$$

2.3 Bounded remainder sets

Let f be a word on $\mathbb{Z}^2$ with values in a finite alphabet $\mathcal{A}$. In analogy to what Kesten [15], Rauzy [24] and Ferenczi [10], define for dynamical systems we call a set $A \subseteq \mathcal{A}$ a *bounded remainder set* if there exist two real numbers α and C such that

$$\forall N \in \mathbb{Z}_{>0} \quad |\mathrm{Card}\{n < N;\ u_n \in A\} - \alpha N| \leq C.$$

See also [9] for a connected generalisation of the balance property in the case of a Sturmian dynamical system.

We recall the following classical discrepancy result, which will be of some use in the sequel. The "if" part is due to Ostrowski [22], the "only if" part to Kesten [15].

Theorem 1 (Ostrowski, Kesten) *Consider a rotation R_α of irrational angle α on $\mathbb{R}/\mathbb{Z}$. Let I be an interval in $\mathbb{R}/\mathbb{Z}$. There exists a real numbers C such that*

$$\forall N \in \mathbb{Z}_{>0} \quad |Card\{n < N;\ n\alpha \in I\} - \alpha N| \leq C$$

if and only if its length $|I|$ equals the fractional part of $k\alpha$ for some nonzero integer k.

We extend the definition to multi-dimensional words. Let $f : \mathbb{Z}^n \to \mathcal{A}$ be a word with values in $\mathcal{A}$. A set $A \subseteq \mathcal{A}$ is called a *bounded remainder set for f* if there exist two real numbers α and C such that

$$\forall \vec{N} \in \mathbb{Z}^n_{>0} \quad |\mathrm{Card}\{\vec{x} \in [\vec{0}, \vec{N});\ f(\vec{x}) \in A\} - \alpha|\vec{N}|| \leq C.$$

Note that a set $\{i\}$ with $i \in A$ is a bounded remainder set if and only if there is some C such that f is C-balanced on the letter i. Indeed the "if" part is a direct consequence of Lemma 4. The "only if" part is a direct consequence of the triangle inequality.

3 A complete description of 1-balanced functions

In this section we classify all 1-balanced words $f : I \to \{0, 1\}$ on infinite intervals I. We deal in subsequent subsections with dimension $1, 2, 3$ and > 3. Without further reference we shall use the following observations which classify all 1-balanced functions of densities 0 and 1 and show that we can restrict our attention to density at most $1/2$.

Suppose f is 1-balanced and has density α. Then $0 \leq \alpha \leq 1$. By interchanging all letters 0 and 1 if $\alpha \geq 1/2$, we may restrict our attention to the case $0 \leq \alpha \leq 1/2$. If $\alpha = 0$, then there exist arbitrarily large blocks without 1's so that every arbitrarily large block has at most one 1. Therefore the letter 1 occurs at most once. Thus only the following 1-balanced functions are possible for density 0:

- $f = 0$;
- $f(\vec{x_0}) = 1$ for some $\vec{x_0} \in \mathbb{Z}^n$ and $f(\vec{x}) = 0$ for $\vec{x} \neq \vec{x_0}$.

3.1 The one-dimensional case

Let $I = \mathbb{N}$ or $\mathbb{Z}$. Let $f : I \to \{0, 1\}$ be 1-balanced of density α. If $\alpha = 0$ we consider the complementary word with $\alpha = 1$. Let $S = (s_i)_{i \in I}$ be the sequence of increasing integers for which $f(s_i) = 1$. Below we shall classify all 1-balanced words in terms of the sequence (s_i). The classification is due to Morse and Hedlund [21], but we give a new concise self-contained proof.

Lemma 5 *For $i, j \in I$ with $i < j$ we have*

$$(j - i)\alpha^{-1} - 1 \leq s_j - s_i \leq (j - i)\alpha^{-1} + 1$$

where at most one of the equality signs can occur.

Proof. On average we have $j - i = (s_j - s_i)\alpha$. So there are pairs i, j with $s_j - s_i \geq (j - i)\alpha^{-1}$ and pairs with $s_j - s_i \leq (j - i)\alpha^{-1}$. Since $[s_i + 1, s_j)$ contains exactly $j - i - 2$ ones and $[s_i, s_j + 1)$ contains $j - i$ ones, the 1-balancedness of f implies that the difference $s_j - s_i$ with $j - i$ fixed can attain at most two (consecutive) values. This implies Lemma 5. ∎

Lemma 6 *For every $\beta \in \mathbb{R}$ we have*

$$s_i \geq \lfloor i\alpha^{-1} + \beta \rfloor \text{ for all } i \in I$$

or

$$s_i \leq \lfloor i\alpha^{-1} + \beta \rfloor \text{ for all } i \in I.$$

Proof. Suppose the assertion is false. Then there exists a real number β and integers $i, j \in I$ such that

$$s_i \leq \lfloor i\alpha^{-1} + \beta \rfloor - 1 \text{ and } s_j \geq \lfloor j\alpha^{-1} + \beta \rfloor + 1.$$

Suppose $i < j$. Note that

$$\lfloor j\alpha^{-1} + \beta \rfloor - \lfloor i\alpha^{-1} + \beta \rfloor > (j - i)\alpha^{-1} - 1.$$

Hence $s_j - s_i > (j - i)\alpha^{-1} + 1$. This contradicts Lemma 5. The case $i > j$ is similar. ∎

Theorem 2 (Morse-Hedlund) *Let $I = \mathbb{N}$ or $\mathbb{Z}$. Let $f : I \to \{0, 1\}$ be 1-balanced with density $\alpha > 0$. Let $S = (s_i)_{i \in I}$ be the sequence of increasing integers for which $f(s_i) = 1$. Then one of the following three cases occur, where β is some suitable real number.*
(i) (periodic case) *α is rational and $s_i = \lfloor i\alpha^{-1} + \beta \rfloor$ for $i \in I$.*
(ii) (irrational case) *α is irrational and $s_i = \lfloor i\alpha^{-1} + \beta \rfloor$ for $i \in I$ or $s_i = \lceil i\alpha^{-1} + \beta \rceil$ for $i \in I$.*
(iii) (skew case) *α is rational and there are $\delta \in \{-1, 1\}$ and $i_0 \in I$ such that*

$$s_i = \lfloor i\alpha^{-1} + \beta \rfloor \text{ for } i \in I, i < i_0,$$

$$s_i = \lfloor i\alpha^{-1} + \beta + \delta\epsilon \rfloor \text{ for } i \in I, i \geq i_0,$$

where ϵ^{-1} is the numerator of the rational number α.

Proof. Put $\beta_0 = \inf\{\beta | \lfloor i\alpha^{-1} + \beta \rfloor \geq s_i \text{ for } i \in I\}$. Then $\lfloor i\alpha^{-1} + \beta_0 \rfloor \geq s_i$ for all $i \in I$ and $\lfloor i\alpha^{-1} + \beta \rfloor \leq s_i$ for all $i \in I$ when $\beta < \beta_0$ by Lemma 6. It follows that $s_i = \lfloor i\alpha^{-1} + \beta_0 \rfloor$ if $i\alpha^{-1} + \beta_0 \notin \mathbb{Z}$ and $s_i = \lfloor i\alpha^{-1} + \beta_0 \rfloor$ or $s_i = \lfloor i\alpha^{-1} + \beta_0 \rfloor - 1$ otherwise.
If $\alpha \notin \mathbb{Q}$, then $i\alpha^{-1} + \beta_0 \in \mathbb{Z}$ can happen for only one value of i. Thus, if it is not true that $s_i = \lfloor i\alpha^{-1} + \beta_0 \rfloor$ for all i, then there is an i_0 such that $s_i = \lfloor i\alpha^{-1} + \beta_0 \rfloor$ for $i \neq i_0$ and $s_{i_0} = i\alpha^{-1} + \beta_0 - 1 \in \mathbb{Z}$. Then $s_i = \lceil i\alpha^{-1} + \beta_0 - 1 \rceil$ for $i \in I$. This proves the claim for the irrational case, with $\beta = \beta_0$ in the former case and $\beta = \beta_0 - 1$ in the latter.
If $\alpha \in \mathbb{Q}$ and it is not true that $s_i = \lfloor i\alpha^{-1} + \beta_0 \rfloor$ for $i \in I$, then $i\alpha^{-1} + \beta \in \mathbb{Z}$ for an arithmetic progression of values $i \in I$. If s_i always attains the higher value, we have $s_i = \lfloor i\alpha^{-1} + \beta_0 \rfloor$ for all $i \in I$. If s_i always attains the lower value, we have $s_i = \lfloor i\alpha^{-1} + \beta_0 - \epsilon \rfloor$ for all $i \in I$. This yields periodic cases with $\beta = \beta_0$ and $\beta = \beta_0 - \epsilon$, respectively.
The remaining case is the skew case where s_i assumes sometimes the high integer value and sometimes the low integer value. If a high integer value is followed by a low integer value, then we have the equality sign in the left inequality of Lemma 5. If it jumps from a low integer value to a high integer value, then we have the equality sign in the right inequality of Lemma 5.

Since not both equality signs can occur, there can be only one such a jump. If the jump is from high to low, then we have the former skew case with $\delta = -1$ and $\beta = \beta_0$. If the jump is from low to high at the same transition, then we have the latter skew case with $\delta = 1$ and $\beta = \beta_0 - \epsilon$. ∎

We call a word Sturmian if it corresponds with case (ii).

3.2 The two-dimensional case

The 1-balanced words on two-dimensional infinite intervals split into two classes with totally different behaviour. The first class of intervals has width 2 and is described in Theorem 3 and closely related to the one-dimensional case. The class of intervals with width > 2 is given in Theorem 4 and contains only fully periodic words with an exception when $\alpha = 0, 1$.
We recall that all 1-balanced functions with density $\alpha = 0, 1$ have been classified before and that by interchanging 0's and 1's if necessary we can secure that the density α is at most $1/2$. In particular, in the following theorem we may assume without loss of generality that there are no columns with two 1's.

Theorem 3 *Let I be $I_1 \times [0,1]$ with $I_1 \in \{\mathbb{N}, \mathbb{Z}\}$. Let $f : I \to \{0,1\}$ be 1-balanced of positive density and such that $f(x,0) + f(x,1) \leq 1$ for $x \in I_1$. Then $(F(x))_{x \in I_1}$ with $F(x) := f(x,0) + f(x,1)$ is a 1-balanced word on I_1 (as described in Theorem 2). Moreover, if $F(x) = F(y) = 1, F(z) = 0$ for $x < z < y$ then either $f(x,0) = f(y,1) = 1$ or $f(x,1) = f(y,0) = 1$.*
On the other hand, any function $\{f(x,y) | x \in I_1, y \in \{0,1\}\}$ which corresponds with some 1-balanced function $F : I \to \{0,1\}$ in the above indicated way is 1-balanced.

Proof. By the condition imposed on f, we have $F : I_1 \to \{0,1\}$. Since f is 1-balanced on m by 2 blocks, F has to be 1-balanced on blocks of length m for every m. Thus F is a 1-balanced word as described in Theorem 3.
On the other hand, suppose $F : I_1 \to \{0,1\}$ is a 1-balanced word as described in Theorem 2. Split $(s_i)_{i \in I}$ in two subsequences s_{2i} and s_{2i+1} and define $f : I \to \{0,1\}$ by $f(x,0) = 1$ if and only if $x \in s_{2i}$, $f(x,1) = 1$ if and only if $x \in s_{2i+1}$. Then f is 1-balanced on all m by 2 blocks and $f(x,0)+f(x,1) \leq 1$ for $x \in I_1$. Suppose f is not 1-balanced on some m by 1 block. Then there exist $a, b \in I_1$ such that the number of i's with $s_{2i} \in [a, a+m)$ differs more than one from the number of i's with $s_{2i+1} \in [b, b+m)$. This implies that the number of i's with $s_i \in [a, a+m)$ differs more than one from the number of

i's with $s_i \in [b, b+m)$. The latter statement contradicts the 1-balancedness of F. ■

In the remaining case of infinite two-dimensional intervals I we may assume without loss of generality that I contains $[0,\infty) \times [-1,1]$.

Theorem 4 *Let I be a two-dimensional interval which contains $[0,\infty) \times [-1,1]$. Let $f : I \to \{0,1\}$ be 1-balanced of density α.*
Then $\alpha \in \{0, \frac{1}{5}, \frac{1}{3}, \frac{2}{5}, \frac{1}{2}, \frac{3}{5}, \frac{2}{3}, \frac{4}{5}, 1\}$.
Moreover, by complementation, reflection and translation f can be transformed into g with
if $\alpha = 0$ or 1, then $g(i,j) = 0$ for all i and j or $g(0,0) = 1$ and $g(i,j) = 0$ otherwise;
if $\alpha = \frac{1}{5}$ or $\frac{4}{5}$, then $g(0,0) = 1, g(1,0) = g(2,0) = g(3,0) = g(4,0) = 0$ and g has period lattice with basis $(5,0),\ (2,1)$;
if $\alpha = \frac{1}{3}$ or $\frac{2}{3}$, then $g(0,0) = 1, g(1,0) = g(2,0) = 0$ and g has period lattice with basis $(3,0),\ (1,1)$;
if $\alpha = \frac{2}{5}$ or $\frac{3}{5}$ then $g(0,0) = g(2,0) = 1, g(1,0) = g(3,0) = g(4,0) = 0$ and g has period lattice with basis $(5,0),\ (1,1)$;
if $\alpha = \frac{1}{2}$, then $g(0,0) = 1, g(1,0) = 0$ and g has period lattice with basis $(2,0),\ (1,1)$.
On the other hand, all the above functions are 1-balanced on every interval I.

Note that it is not stated that the domain of g is I. By reflection and translation it may well be $[-3,\infty) \times [-1,1]$ or $(-\infty, 8], [-7, \infty)$.
We give some lemmas about 1-balancedness on finite blocks which will be used in the proof of Theorem 4. They may also be used to derive a complete characterization of all 1-balanced words on all two-dimensional finite intervals (=blocks). Lemma 7 is a variant of Lemma 1.

Lemma 7 *Let $l, m \in \mathbb{N}$ with $l|m$. Let f be 1-balanced on some m by l block. If every m by 1 block contains at least (at most) k letters 1, then every m/l by l block contains at least (at most) k letters 1.*

Proof. Obviously an m by l block contains at least (at most) kl letters 1. Let B denote any m/l by l block. Then the m by l block can be split into l blocks B. If one of the B's contains less (more) than k letters 1, then another B has to contain more (less) than k letters 1. This causes a contradiction with the 1-balancedness of f. ■

Lemma 8 *Let $m \in \mathbb{N}_{>3}$. Let f be 1-balanced on an interval containing the block $[0, m] \times [-1, 1]$. Suppose $f(0,0) = f(m,0) = 1$ and $f(i,0) = 0$ for $0 < i < m$. Then $m = 5$.*

Proof. Since there is an $m-1$ by 1 block with only 0's, every $m-1$ by 1 block has at most one 1. By Lemma 7 every 1 by 3 block has at most one 1. Since there is an $m+1$ block with two 1's, every $m+1$ by 1 block has at least one 1. By Lemma 7 every $\lfloor \frac{m-1}{2} \rfloor$ by 2 block contains at most one 1. Hence $f(0, \pm 1) = f(1, \pm 1) = ... = f(\lfloor \frac{m-3}{2} \rfloor, \pm 1) = 0$. By the same lemma every $\lceil \frac{m+1}{3} \rceil$ by 3 block contains at least one 1. Hence there exists an i with $0 < i \leq \lceil \frac{m+1}{3} \rceil$ and a j with $-1 \leq j \leq 1$ and $f(i,j) = 1$. Thus $\lceil \frac{m+1}{3} \rceil > \lfloor \frac{m-3}{2} \rfloor$ which implies that $m \in \{4, 5, 6, 7, 8, 9, 10, 12\}$.
If $m \in \{7, 9\}$, then we obtain by considering $\frac{m-1}{2}$ by 2 blocks that $f(m, \pm 1) = f(m-1, \pm 1) = ... = f(\frac{m+3}{2}, \pm 1) = 0$. By considering $m+1$ by 1 blocks we discover that $f(\frac{m-1}{2}, 1) = 1$ or $f(\frac{m+1}{2}, 1) = 1$ and also $f(\frac{m-1}{2}, -1) = 1$ or $f(\frac{m+1}{2}, -1) = 1$. By symmetry it is no loss of generality to assume $f(\frac{m-1}{2}, 1) = 1$ whence $f(\frac{m+1}{2}, 1) = 0$, $f(\frac{m-1}{2}, -1) = 0$, $f(\frac{m+1}{2}, -1) = 1$. We find a contradiction by comparing the number of 1's in 2 by 3 blocks.
If $m \in \{8, 10, 12\}$, then we obtain by considering $\frac{m-2}{2}$ by 2 blocks that $f(m, \pm 1) = f(m-1, \pm 1) = ... = f(\frac{m}{2} + 2, \pm 1) = 0$. By considering $m+1$ by 1 blocks we find that $f(\frac{m}{2} - 1, 1) = 1$ or $f(\frac{m}{2}, 1) = 1$ or $f(\frac{m}{2} + 1, 1) = 1$ and similarly that $f(\frac{m}{2} - 1, -1) = 1$ or $f(\frac{m}{2}, -1) = 1$ or $f(\frac{m}{2} + 1, -1) = 1$. Since a 3 by 3 block can contain at most one 1, this excludes that $m = 10$ or $m = 12$. If $m = 8$ we have the pattern

$$\begin{array}{ccccccccc}
0 & 0 & 0 & & & & 0 & 0 & 0 \\
\underline{1} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & & & & 0 & 0 & 0
\end{array}$$

and every 2 by 3 block can have at most one 1. (Here and below the underlined 1 indicates the value at the origin.) By symmetry it is no loss of generality to assume $f(3, 1) = f(5, -1) = 1$. Then $f(4, 1) = f(5, 1) = f(3, -1) = f(4, -1) = 0$. By comparing 4 by 2 blocks we reach a contradiction.
If $m = 6$, then we have by the same arguments as used before

$$\begin{array}{ccccccc}
0 & 0 & & & & 0 & 0 \\
\underline{1} & 0 & 0 & 0 & 0 & 0 & 1 \\
0 & 0 & & & & 0 & 0
\end{array}$$

Since either $f(2, 1) = 1$ or $f(3, 1) = 1$ or $f(4, 1) = 1$ and similarly either $f(2, -1) = 1$ or $f(3, -1) = 1$ or $f(4, -1) = 1$, we may assume without loss of generality that $f(2,1)=1, f(4,-1)=1$ in view of 1-balancedness of 2 by 3 blocks. It follows that $f(3,1)=f(4,1)=f(2,-1)=f(3,-1)=0$.

$$\begin{array}{ccccccc} 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ \underline{1} & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{array}$$

We now obtain a contradiction by considering 3 by 2 blocks.
If $m=4$, then we have

$$\begin{array}{ccccc} 0 & & & & 0 \\ \underline{1} & 0 & 0 & 0 & 1 \\ 0 & & & & 0 \end{array}$$

There should be at least one 1 in the upper row and at least one 1 in the lower row. Since they cannot be in the same column, we may assume without loss of generality that $f(1,1)=1$. Then every 2 by 2 block contains at least one 1, but every block of size 3 has at most one 1. This yields a contradiction at place (3,1).

$$\begin{array}{ccccc} 0 & 1 & 0 & * & 0 \\ \underline{1} & 0 & 0 & 0 & 1 \\ 0 & & & & 0 \end{array}$$

We conclude that $m=5$. ■

Lemma 9 *Let f be 1-balanced on an interval containing $[0,10]\times[-1,1]$ as a subblock. Let $f(0,0)=f(5,0)=1$ and $f(i,0)=0$ for $0<i<5$. Then f is periodic with period vectors $(5,0)$ and either $(2,1)$ or $(2,-1)$.*

Proof. By Lemma 7 every block of size ≤ 4 contains at most one 1 and every block of size ≥ 6 at least one 1. Therefore $f(0,\pm1)=f(1,\pm1)=f(4,\pm1)=f(5,\pm1)=f(6,\pm1)=0$. By considering 6 by 1 blocks we see that either $f(2,1)=1$ or $f(3,1)=1$ and similarly $f(2,-1)=1$ or $f(3,-1)=1$. We may assume without loss of generality that $f(2,1)=f(3,-1)=1, f(2,-1)=f(3,1)=0$.

$$\begin{array}{cccccc} 0 & 0 & 1 & 0 & 0 & 0 \\ \underline{1} & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{array}$$

Since in between two 1's there have to be at least three 0's, we have $f(10,0)=1$ by Lemma 8. By considering 2 by 2 blocks we obtain $f(9,\pm1)=f(10,\pm1)=$

0. This yields $f(7,1) = f(8,-1) = 1$ or $f(7,-1) = f(8,1) = 1$, but the latter possibility is excluded by considering 5 by 1 blocks. In this way we find $f(a+2, b+1) = f(a+5, b) = f(a,b)$ for every $(a,b) \in I$ with $b \leq 2a$. Since we can also extend in the opposite directions, we obtain a period lattice with basis vectors $(5,0)$ and $(2,1)$. ■

Lemma 10 *Let f be 1-balanced on an interval containing $[0,8] \times [-1,1]$ as a subblock. Let $f(0,0) = f(3,0) = 1$ and $f(1,0) = f(2,0) = f(4,0) = f(5,0) = 0$. Then f is periodic with period lattice basis $(3,0)$ and either $(1,1)$ or $(1,-1)$.*

Proof. We see that every block of size 2 contains at most one 1 and that every block of size 5 has at most two 1's. Hence $f(0, \pm 1) = f(3, \pm 1) = f(4,0) = 0$. Furthermore, every block of size 4 contains at least one 1. Hence either $f(1,1) = 1$ or $f(2,1) = 1$ and either $f(1,-1) = 1$ or $f(2,-1) = 1$. By symmetry we may assume without loss if generality that $f(1,1) = f(2,-1) = 1, f(1,-1) = f(2,1) = 0$.
Suppose $f(4,1) = 0$. Then $f(4,-1) = 0$ by comparing 3 by 1 blocks, $f(5,1) = 1$ by comparing 4 by 1 blocks and $f(5,-1) = 0$ by comparing 1 by 3 blocks. This yields a contradiction when comparing 2 by 2 blocks.
Thus $f(4,1) = 1, f(5,1) = 0$. We see that either $f(4,-1)$ or $f(5,-1)$ equals 1 by comparing 2 by 2 blocks and that in fact $f(4,-1) = 0, f(5,-1) = 1$ by comparing 1 by 3 blocks. Comparisons of 2 by 2 blocks and 3 by 1 blocks give $f(6,0) = 1, f(6, \pm 1) = 0, f(7,0) = 0$. The supposition $f(7,1) = 0$ leads to a contradiction as the supposition $f(4,1) = 0$ did above.
So $f(7,1) = 1, f(8,1) = 0$. Suppose $f(8,0) = 1$. Then $f(8,-1) = 0$ and, by comparing 3 by 1 blocks, $f(7,-1) = 1$. This yields a contradiction when comparing 4 by 2 blocks. Thus $f(8,0) = 0$ and we have shown that $(3,0)$ and $(1,1)$ are period vectors. We can extend in the opposite directions as well and we therefore have a period lattice with basis $(3,0)$ and $(1,1)$. ■

Lemma 11 *Let f be 1-balanced on an interval containing $[0,10] \times [-1,1]$ as a subblock. Let $f(0,0) = f(2,0) = f(5,0) = 1$ and $f(1,0) = f(3,0) = f(4,0) = 0$. Then f is periodic with period lattice basis $(5,0)$ and either $(1,1)$ or $(1,-1)$.*

Proof. From the conditions and Lemma 7 we see that every block of size 2 contains at most one 1 and every block of size 3 has at least one 1. Thus $f(0, \pm 1) = f(2, \pm 1) = f(5, \pm 1) = f(6,0) = 0$. By comparing 3 by 1 blocks we obtain $f(1, \pm 1) = 1$ and either $f(3,1) = 1$ or $f(4,1) = 1$ and either

$f(3,-1) = 1$ or $f(4,-1) = 1$. Without loss of generality we may assume $f(3,1) = f(4,-1) = 1, f(3,-1) = f(4,1) = 0$. Then $f(6,1) = 1$ by comparing 3 by 1 blocks and $f(6,-1) = 1$ by comparing 4 by 2 blocks. Hence $f(7,\pm1) = 0, f(7,0) = 1, f(8,0) = 0$. By comparing 1 by 3 blocks and 3 by 3 blocks we see that either $f(8,1)$ or $f(8,-1)$ equals 1. By comparing 5 by 1 blocks we conclude that $f(8,1) = 1, f(8,-1) = 0, f(9,1) = 0, f(9,-1) = 1, f(9,0) = 0, f(10,0) = 1, f(10,\pm1) = 0$. Thus f has period vectors $(5,0)$and $1,1)$. We can extend in the directions $(-5,0)$ and $(-1,-1)$ as well. ∎

Lemma 12 *Let f be 1-balanced on an interval containing $[0,7] \times [-1,1]$ as a subblock. Let $f(0,0) = f(2,0) = f(4,0) = 1$ and $f(1,0) = f(3,0) = f(5,0) = 0$. Then f is periodic with period lattice basis $(2,0)$ and $(1,1)$.*

Proof. Note that every 5 by 1 block contains at least two 1's. By comparing 1 by 2 blocks and 3 by 1 blocks we deduce that $f(1,\pm1) = f(3,\pm1) = 1$. Since every 5 by 1 block contains at least two 0's, we obtain similarly that $f(2,\pm1) = f(4,\pm1) = 0$.
Suppose $f(5,1) = 0$. Then $f(6,1) = 1$ by comparing 3 by 1 blocks, $f(5,-1) = 1$ by comparing 1 by 3 blocks, $f(6,0) = f(6,-1) = 0$ by comparing blocks of size 2. This yields a contradiction when comparing 3 by 3 blocks.
Thus $f(5,1) = 1$. By symmetry $f(5,-1) = 1$. Suppose $f(6,0) = 0$. Then $f(6,\pm1) = 0$ by considering 2 by 1 blocks and we obtain a contradiction by comparing 3 by 3 blocks. Thus $f(6,0) = 1$. We can show $f(6,\pm1) = 0$ in a similar way as we showed $f(5,\pm1) = 1$, and $f(7,0) = 0$ as we derived $f(6,0) = 1$. We have found that $(2,0)$ and $(1,1)$ are period vectors. We can go into the directions $(-2,0)$ and $(-1,-1)$ as well. ∎

Proof of Theorem 2. We split the proof into three cases.
(a) The case $0 < \alpha < \frac{1}{5}$.
By Lemma 3 every block of size 4 contains at most one 1. According to Lemma 8 applied to $f(x-a, y-b)$ we obtain that $f(a,b) = 1$ implies $f(a+5,b) = 1$. Hence $\alpha \geq \frac{1}{5}$. Thus there are no 1-balanced words in case (a).
(b) The case $\frac{1}{5} \leq \alpha < \frac{1}{3}$.
There should exist a and b with $f(a,b) = 1, f(a+1,b) = f(a+2,b) = f(a+3,b) = 0$, since otherwise $\alpha \geq \frac{1}{3}$. By Lemma 8 it follows that $f(a+4,b) = 0, f(a+5,b) = 1$. By Lemma 9 applied to $f(a+x, b+y)$ we obtain that f is periodic with period vector $(5,0)$ and either $(2,1)$ or $(2,-1)$. By symmetry this yields that in case (b) there only can be the word described

in the theorem with $\alpha = \frac{1}{5}$.
(c) The case $\frac{1}{3} \leq \alpha \leq \frac{1}{2}$.
Since a, b with $f(a, b) = 1, f(a+1, b) = f(a+2, b) = f(a+3, b) = 0, f(a+4, b) = 1$ are excluded by Lemma 8, there can be only one or two zeros between two consecutive 1's in a row.
We distinguish between three subcases:
(c1) in the row $f(x, 0)$ there are always two 0's between consecutive 1's;
(c2) in the row $f(x, 0)$ there is always one 0 between consecutive 1's;
(c3) in the row $f(x, 0)$ there are sometimes two and sometimes one 0 between consecutive 1's.
In case (c1) we apply Lemma 10 to $f(x-a, 0)$ for some a with $f(a, 0) = 1$. In case (c2) we apply Lemma 12 to $f(x-a, 0)$ for some a with $f(a, 0) = 1$. In case (c3) there is a situation with $f(a, 0) = 1, f(a+1, 0) = f(a+4, 0) = 0, f(a+5, 0) = 1$ and either $f(a+2, 0) = 1, f(a+3, 0) = 0$ or $f(a+2, 0) = 0, f(a+3, 0) = 1$. In the former case we apply Lemma 11. A similar argument applies in the latter case.
In each subcase we find that f is fully periodic and has one of the densities and corresponding period lattice bases as stated in the theorem.

It remains to prove that the found words f are 1-balanced indeed. We can restrict ourselves to $\mathbb{Z} \times \mathbb{Z}$, since every restriction of a 1-balanced word is 1-balanced. We give the proof for one case. The other cases are similar or simpler. We prove that the word in case $\alpha = \frac{2}{5}$ is 1-balanced. It suffices to prove that for every block $[\vec{a}, \vec{b})$ the number of function values 1 is either $\lfloor \frac{2}{5} |\vec{a}, \vec{b}| \rfloor$ or $\lceil \frac{2}{5} |\vec{a}, \vec{b}| \rceil$. On using that every 5 by 1 block contains exactly two 1's we may assume that $\vec{b} = \vec{a} + \vec{c}$ with $\vec{c} = (c_1, c_2)$ satisfying $0 \leq c_1 < 5$, $0 \leq c_2 < 5$. If $c_1 c_2 = 0$, we are finished. By considering, the complement with respect to the 5 by 5 block if necessary, we may assume $1 \leq c_1 \leq 2$. By considering the complement with respect to the c_1 by 5 block if necessary, we may further restrict our attention to $1 \leq c_2 \leq 2$. Check that every block of size 2 contains at most one 1 and that every 2 by 2 block has one or two letters 1. This implies that f is 1-balanced. ■

3.3 The three-dimensional case

In Theorem 5 we characterize all 1-balanced words on infinite intervals $I = I_1 \times I_2 \times I_3$. Without loss of generality we may assume that $I_3 \subseteq I_2 \subseteq I_1$. It will turn out that there are essentially only four such 1-balanced words.

Theorem 5 *Let $I = I_1 \times I_2 \times I_3$ where $I_1 \in \{\mathbb{N}, \mathbb{Z}\}$, $I_3 \subseteq I_2 \subseteq I_1$ and $I_2, I_3 \in \{\mathbb{N}, \mathbb{Z}, B\}$ where B is the set of one-dimensional blocks containing*

at least two integers. Let $f : I \to \{0,1\}$ be 1-balanced of density α. Then $\alpha \in \{0, \frac{1}{3}, \frac{1}{2}, \frac{2}{3}, 1\}$.
Moreover, by complementation, reflection and translation f can be transformed into g with
if $\alpha = 0$ or 1, then $g(x,y,z) = 0$ for all i and j
or $g(0,0,0) = 1$ and $g(x,y,z) = 0$ otherwise;
if $\alpha = \frac{1}{3}$ or $\frac{2}{3}$, then $g(x,y,z) = 1$ if $x+y+z$ is divisible by 3 and $g(x,y,z) = 0$ otherwise;
if $\alpha = \frac{1}{2}$, then $g(x,y,z) = 1$ if $x+y+z$ is even and $g(x,y,z) = 0$ otherwise.
On the other hand, the above functions are 1-balanced on every interval I.

Proof. If $[-1,1] \in I_2$, then we can apply Theorem 4 to $f(x,y,i)$ for every fixed i. We shall show that there are no 1-balanced words with densities $1/5$, $2/5$, $3/5$ or $4/5$. By complementation it suffices only to consider $1/5$ and $2/5$. Suppose f is 1-balanced with density $\alpha = 1/5$. Then we may assume without loss of generality that

$$f(x,y,0) = \begin{cases} 1 & \text{if } x - 2y \text{ is divisible by } 5 \\ 0 & \text{otherwise.} \end{cases}$$

$$\begin{array}{ccccccccc} 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ \underline{1} & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \end{array}$$

In particular, $f(0,0,0) = f(2,1,0) = f(3,-1,0) = 1$. Since every block of size 4 contains at most one 1 by Lemma 1 (cf. Lemma 7), we have $f(0,0,1) = f(1,0,1) = f(2,0,1) = f(3,0,1) = f(4,0,1) = 0$. However, we have seen in the previous section that no 1-balanced word $f(x,y,0)$ of density $1/5$ can contain a 5 by 1 block without 1's.
Suppose f is 1-balanced with density $2/5$. Then we may assume without loss of generality that

$$f(x,y,0) = \begin{cases} 1 & \text{if } x - y \text{ or } x - y - 2 \text{ is divisible by } 5 \\ 0 & \text{otherwise.} \end{cases}$$

$$\begin{array}{ccccccccc} 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ \underline{1} & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \end{array}$$

In particular, $f(0,0,0) = f(1,1,0) = f(1,-1,0) = f(2,0,0) = f(3,1,0) = f(4,-1,0) = f(5,0,0) = 1$. We know that every block of size 2 contains at most one 1 and every block of size 3 contains at least one 1. Hence $f(0,0,1)=f(1,1,1)=f(2,0,1)=f(3,1,1)=f(5,0,1)=0$ and $f(1,0,1)=f(2,1,1)=1$. Thus we know the following values for $f(x,y,1)$:

$$\begin{array}{ccccccccc}
 & 0 & 1 & 0 & * & * & 0 & 1 & 0 \\
0 & \underline{1} & 0 & * & * & 0 & 1 & 0 & \\
 & 0 & * & * & 0 & 1 & & &
\end{array}$$

Furthermore either $f(3,0,1)=1$, $f(3,-1,1)=f(4,0,1)=0$ or $f(4,0,1)=1$, $f(3,0,1)=f(4,1,1)=0$, . The 2 by 2 by 2 block $[(1,0,0),(2,1,1)]$ contains four 1's, but in the former case the 2 by 2 by 2 block $[(3,-1,0),(4,0,1)]$ has at most two 1's and in the latter case the 2 by 2 by 2 block $[(3,0,0),(4,1,1)]$ has at most two 1's. These contradictions prove the claim.
Subsequently we prove that for the other densities in Theorem 4 the two-dimensional words can only extend to the 1-balanced three-dimensional words given in Theorem 5. Again we assume $\alpha \leq 1/2$.
If $\alpha = 0$, the statement follows from the considerations at the beginning of Section 3.
If $\alpha = \frac{1}{3}$, then according to Theorem 4 every block of size 3 contains exactly one 1. By symmetry this implies periodicity as given in Theorem 5.
If $\alpha = \frac{1}{2}$, then according to Theorem 4 every block of size 2 contains exactly one 1. This implies periodicity as given in Theorem 5.

The only remaining cases are equivalent to the cases with $I_2 = I_3 = [0,1]$. We consider this case more closely. Recall that $f(x,y,0), f(x,y,1), f(x,0,z)$ and $f(x,1,z)$ are 1-balanced. Let m be such that $f(0,0,0)=f(m,0,0)=1, f(i,0,0)=0$ for $0<i<m$. We may assume $m>1$.
If m is even, then we have $f(\frac{m}{2},1,0)=f(\frac{m}{2},0,1)=1$ by 1-balancedness on $\frac{m}{2}\pm1$ by 2 by 1 and $\frac{m}{2}\pm1$ by 1 by 2 blocks. If, moreover, $m>2$ then we obtain by considering 3 by 1 by 1 blocks that $f(\frac{m}{2}\pm1,1,0)=f(\frac{m}{2}\pm1,0,1)=0$ and $f(\frac{m}{2}-1,1,1)=0$ or $f(\frac{m}{2}+1,1,1)=0$. This gives a contradiction for 1 by 2 by 2 blocks. Thus $m=2$, if m is even.
If m is odd, then we have either $f(\frac{m-1}{2},1,0)=1$ or $f(\frac{m+1}{2},1,0)=1$, and similarly either $f(\frac{m-1}{2},0,1)=1$ or $f(\frac{m+1}{2},0,1)=1$. By symmetry we may assume $f(\frac{m-1}{2},1,0)=1$. If, moreover, $m>4$, then every 4 by 1 by 1 block has at most one 1. It follows by considering 1 by 2 by 2 blocks that $f(\frac{m-1}{2},0,1)=f(\frac{m+1}{2},1,0)=0$, $f(\frac{m+1}{2},0,1)=1$. By comparing $\frac{m-1}{2}$ by 2 by 1 blocks and $\frac{m-1}{2}$ by 1 by 2 blocks, we see that $f(i,1,1)=0$ for $0<i<m$. By comparing $m+1$ by 1 by 1 blocks we infer $f(0,1,1)=1$ or $f(m,1,1)=1$. We obtain a contradiction when considering 1 by 2 by 2 blocks. Thus $m=3$ when m is odd.

If every 2 by 1 by 1 block contains exactly one 1, then we have period vectors $(2,0,0),(1,-1,0),(1,0,-1)$ and we obtain the 1-balanced word in Theorem 5 with $\alpha = 1/2$.
If every 3 by 1 by 1 block contains exactly one 1, then by reflection we can reduce to period vectors $(3,0,0),(1,-1,0),(1,0,-1)$ and we obtain the 1-balanced word in Theorem 5 with $\alpha = 1/3$.
Otherwise we may assume that after translation and possibly reflection we have the situation $f(0,0,0) = f(2,0,0) = 1, f(1,0,0) = f(3,0,0) = f(4,0,0) = 0$. Then every block of size 2 has at most one 1, whence $f(0,1,0) = f(0,0,1) = f(2,1,0) = f(2,0,1) = 0$. Since every block of size 3 has at least one 1, we obtain $f(1,0,1) = f(1,1,0) = 1, f(1,1,1) = 0, f(5,0,0) = 1, f(5,1,0) = f(5,0,1) = 0$.

$$\begin{array}{lcccccc}
f(x,1,1) & & 0 & & & & \\
f(x,1,0) & 0 & 1 & 0 & & & 0 \\
 & & & & & & \\
f(x,0,1) & 0 & 1 & 0 & & & 0 \\
f(x,0,0) & \underline{1} & 0 & 1 & 0 & 0 & 1
\end{array}$$

Suppose first $f(3,1,0) = 1$. Then $f(3,1,1) = f(4,1,0) = 0, f(2,1,1) = 1$. By comparing 2 by 2 by 2 blocks we see that $f(3,0,1) = f(4,1,1) = 1, f(4,0,1) = f(5,1,1) = 0$.

$$\begin{array}{lcccccc}
f(x,1,1) & & 0 & 1 & 0 & 1 & 0 \\
f(x,1,0) & 0 & 1 & 0 & 1 & 0 & 0 \\
 & & & & & & \\
f(x,0,1) & 0 & 1 & 0 & 1 & 0 & 0 \\
f(x,0,0) & \underline{1} & 0 & 1 & 0 & 0 & 1
\end{array}$$

By considering 2 by 2 by 2 blocks we derive a contradiction. Thus $f(3,1,0) = 0, f(4,1,0) = 1, f(4,1,1) = 0$.
The assumption $f(2,1,1) = 1$ implies $f(3,1,1) = 0$ and yields a contradiction by considering 2 by 2 by 2 blocks.

$$\begin{array}{lcccccc}
f(x,1,1) & & 0 & 1 & 0 & 0 & \\
f(x,1,0) & 0 & 1 & 0 & 0 & 1 & 0 \\
 & & & & & & \\
f(x,0,1) & 0 & 1 & 0 & & & 0 \\
f(x,0,0) & \underline{1} & 0 & 1 & 0 & 0 & 1
\end{array}$$

Thus $f(2,1,1) = 0, f(0,1,1) = f(3,1,1) = 1, f(3,0,1) = 0$ and we obtain another contradiction by considering 2 by 2 by 2 blocks.

$$\begin{array}{llllllll}
f(x,1,1) & 1 & 0 & 0 & 1 & & \\
f(x,1,0) & 0 & 1 & 0 & 0 & & 0 \\
 & & & & & & \\
f(x,0,1) & 0 & 1 & 0 & 0 & & 0 \\
f(x,0,0) & \underline{1} & 0 & 1 & 0 & 0 & 1
\end{array}$$

We conclude that there are no other 1-balanced words than described in Theorem 5.

The proof that the words in Theorem 5 are 1-balanced is similar to the corresponding part of Theorem 4. ■

3.4 The case of dimension greater than three

Let n be a positive integer with $n > 3$. Suppose $f : \mathbb{Z}^n \to \{0,1\}$ is 1-balanced. Then $f(x,y,z,c_4,\dots,c_n)$ is 1-balanced for every $(c_4,\dots,c_n) \in \mathbb{Z}^{n-3}$. It follows that essentially only the natural extensions of the four words from the previous section are 1-balanced.

Theorem 6 *Let $I = I_1 \times ... \times I_n$ where $I_1 \in \{\mathbb{N},\mathbb{Z}\}$, $I_n \subseteq I_{n-1} \subseteq ... \subseteq I_1$ and $I_2,...,I_n \in \{\mathbb{N},\mathbb{Z},B\}$ where B is the set of one-dimensional blocks containing at least two integers. Put $\vec{x} = \{x_1,...,x_n\}$. Let $f : I \to \{0,1\}$ be 1-balanced of density α. Then $\alpha \in \{0, \frac{1}{3}, \frac{1}{2}, \frac{2}{3}, 1\}$.*
Moreover, by complementation, reflection and translation f can be transformed into g with
if $\alpha = 0$ or 1, then $g(\vec{x}) = \vec{0}$ for all $\vec{x}$
or $g(\vec{0}) = 1$ and $g(\vec{x}) = 0$ otherwise;
if $\alpha = \frac{1}{3}$ or $\frac{2}{3}$, then $g(\vec{x}) = 1$ if $x_1 + ... + x_n$ is divisible by 3 and $g(\vec{x}) = 0$ otherwise;
if $\alpha = \frac{1}{2}$, then $g(\vec{x}) = 1$ if $x_1 + ... + x_n$ is even and $g(\vec{x}) = 0$ otherwise.
On the other hand, the above functions are 1-balanced on every interval I.

Proof. Again we may assume without loss of generality that $\alpha \leq 1/2$.
If $\alpha = 0$, the statement follows from the considerations at the beginning of Section 3.
If $\alpha = \frac{1}{3}$, then according to Theorem 5 every block of size 3 contains exactly one 1. By symmetry this implies periodicity as given in Theorem 6.

If $\alpha = \frac{1}{2}$, then according to Theorem 5 every block of size 2 contains exactly one 1. This implies periodicity as given in Theorem 6. ■

Remark

It follows from the above results that in $\mathbb{Z}^n (n > 1)$ there is essentially only one 1-balanced word with density 1/2. The situation changes drastically for 4-balanced words. Partition $\mathbb{Z}^2$ in 2 by 2 blocks and fill every block with one of these patterns:

$$\begin{matrix} 1 & 0 \\ 0 & 1 \end{matrix} \qquad\qquad \begin{matrix} 0 & 1 \\ 1 & 0 \end{matrix}$$

Obviously the resulting function $f : \mathbb{Z}^2 \to \{0, 1\}$ is 4-balanced. Hence there are uncountably many 4-balanced words with density 1/2. Similarly for every rational number p/q with p, q positive integers, $0 < p < q$, there exist uncountably many q^2-balanced $\mathbb{Z}^2$-words with density p/q and uncountably many q^n-balanced $\mathbb{Z}^n$-words with density p/q for any $n > 1$. The situation is unclear to us for irrational density. We present some results on such words in the next section.

4 Imbalances and irrational density

The purpose of this section is to consider the unbalancedness properties of some particular words on $\mathbb{Z}^2$ of irrational density. We have seen that 1-balancedness implies that the density is rational. It is then natural to ask whether a word on $\mathbb{Z}^2$ of irrational density can be balanced. A good candidate for a word to be simultaneously balanced and of irrational density, could be a word on $\mathbb{Z}^2$ built as "regularly" as possible via a Sturmian word on $\mathbb{Z}$. More precisely, consider the following example. Let $\alpha \notin \mathbb{Q}$. Let $f : \mathbb{Z}^2 \to \{0, 1\}$ defined by

$$\forall (m, n) \in \mathbb{Z}^2, \; (f(m, n) = 0 \Longleftrightarrow (m + n)\alpha \in [0, 1 - \alpha) \text{ modulo } 1).$$

This word is periodic and corresponds to a Sturmian word shifted from row to row. We first prove in Section 4.1 that this word cannot be balanced. We then consider in Section 4.2 the case of two-dimensional Sturmian words. Both types of two-dimensional words have the same quantitative behaviour with respect to imbalances. Hence it still remains an open problem whether balance implies rational density.

4.1 An application of Ostrowski's numeration system

Theorem 7 *Let $\alpha \notin \mathbb{Q}$. Let $f : \mathbb{Z}^2 \to \{0, 1\}$ defined by*

$$\forall (m, n) \in \mathbb{Z}^2,\ (f(m, n) = 0 \Longleftrightarrow (m + n)\alpha \in [0, 1 - \alpha) \text{ modulo } 1).$$

Then f is not balanced.

Proof. Suppose f is balanced over 0 and 1. Hence from Lemma 5, there exists C' such that

$$\forall \vec{N} = (N_1, N_2) \in \mathbb{Z}^2_{>0},\ ||\vec{0}, \vec{N}|_1 - N_1 N_2 \alpha| \leq C'. \tag{1}$$

Let us evaluate $|\vec{0}, \vec{N}|_1$, for $\vec{N} = (N_1, N_2) \in \mathbb{Z}^2_{>0}$. Note that

$$f(m, n) = 1 \text{ if and only if } \lfloor (m + n + 1)\alpha \rfloor = \lfloor (m + n)\alpha \rfloor + 1.$$

Hence we have

$$|\vec{0}, \vec{N}|_1 = \sum_{n=0}^{N_2 - 1} (\lfloor (N_1 + n)\alpha \rfloor - \lfloor n\alpha \rfloor) =$$

$$= \sum_{j=N_1}^{N_1 + N_2 - 1} \lfloor j\alpha \rfloor - \sum_{j=0}^{N_2 - 1} \lfloor j\alpha \rfloor,$$

and thus

$$|\vec{0}, \vec{N}|_1 - N_1 N_2 \alpha = \sum_{j=0}^{N_2 - 1} \{j\alpha\} - \sum_{j=N_1}^{N_1 + N_2 - 1} \{j\alpha\}.$$

Let

$$c_\alpha(N) := \sum_{j=1}^{N} (\{j\alpha\} - 1/2).$$

We have

$$|\vec{0}, \vec{N}|_1 - N_1 N_2 \alpha = c_\alpha(N_1 - 1) + c_\alpha(N_2 - 1) - c_\alpha(N_1 + N_2 - 1) - 1/2. \tag{2}$$

There is an abundant literature devoted to the study of $c_\alpha(N)$ and connected discrepancy results, involving Ostrowski's numeration system [22]. Let us follow the notation and use the estimates in [5].

Let $\alpha = [0; a_1, a_2, \ldots]$ be the continued fraction expansion of α with partial quotients a_n and convergents p_n/q_n. For $n \geq 1$, let $\delta_n := |q_n \alpha - p_n|$. Let us expand $N (\geq 1)$ in the numeration scale $(q_n)_{n \in \mathbb{N}}$. Hence there is a unique expansion of N of the form:

1. $N = \sum_{i=1}^{t} z_i q_{i-1}$;

2. $z_t > 0$ and $0 \leq z_i \leq a_i$, for $2 \leq i \leq t$; furthermore, $0 \leq z_1 \leq a_1 - 1$;

3. if $2 \leq i \leq t$ and $z_i = a_i$, then $z_{i-1} = 0$.

For $1 \leq j \leq t$, let

$$m_j = \sum_{i=1}^{j} z_i q_{i-1}, \ m_0 = 0.$$

We have [5]

$$c_\alpha(N) = \sum_{1 \leq j \leq t} (-1)^j z_j \left(\frac{1}{2} - \delta_{j-1}\left(m_{j-1} + \frac{z_j q_{j-1}}{2} + \frac{1}{2}\right)\right). \tag{3}$$

Let us distinguish two cases according to the partial quotients in the continued fraction expansion of α. Suppose that there are infinitely many partial quotients greater than or equal to 2. We can suppose without restriction that infinitely many of the indices of coefficients greater than or equal to 2 are even. Define (j_k) as an infinite strictly increasing sequence satisfying

$$\forall k \in \mathbb{N}_{>0}, \quad a_{2j_k} \geq 2 \text{ and } j_{k+1} - j_k \geq 2.$$

Let $L \in \mathbb{N}_{>0}$. Let

$$N_1 = \sum_{1 \leq k \leq L} \lceil \frac{a_{2j_k}}{2} \rceil q_{2j_k - 1},$$

$$N_2 = \sum_{1 \leq k \leq L} \lfloor \frac{a_{2j_k}}{2} \rfloor q_{2j_k - 1} + q_{2j_k - 2},$$

$$N_3 = N_1 + N_2.$$

We have

$$N_3 = \sum_{1 \leq k \leq L} q_{2j_k}.$$

The three expressions are the Ostrowski's expressions for N_1, N_2, N_3, respectively.

For $i \in \{1, 2, 3\}$, put $N_i = \sum z_k^{(i)} q_{k-1}$ and $m_j^{(i)} = \sum_{1 \leq k \leq j} z_k^{(i)} q_{k-1}$.

- Consider $c_\alpha(N_3)$. We have $z_j^{(3)} \neq 0$ if and only if there exists $1 \leq k \leq L$ such that $j = 2j_k + 1$, and then $z_{2j_k+1}^{(3)} = 1$. We have from Equation (3)

$$c_\alpha(N_3) = -\sum_{1\leq k\leq L}\left(\frac{1}{2} - \delta_{2j_k}\left(m_{2j_k}^{(3)} + \frac{z_{2j_k+1}^{(3)} q_{2j_k}}{2} + \frac{1}{2}\right)\right).$$

Let $1 \leq k \leq L$. We have

$$m_{2j_k}^{(3)} \leq \sum_{1\leq i\leq k-1} q_{2j_i} \leq \frac{q_{2j_{k-1}+1} - 1}{2},$$

since $a_{2j_i} \geq 2$, for every i.

Hence

$$\begin{aligned} m_{2j_k}^{(3)} + \frac{z_{2j_k+1}^{(3)} q_{2j_k}}{2} + \frac{1}{2} &\leq \frac{q_{2j_{k-1}+1} - 1 + q_{2j_k} + 1}{2} \\ &\leq \frac{1}{4} q_{2j_k+1}, \end{aligned}$$

since $j_k - j_{k-1} \geq 2$. Consequently

$$\frac{1}{2} - \delta_{2j_k}\left(m_{2j_k}^{(3)} + \frac{z_{2j_k+1}^{(3)} q_{2j_k}}{2} + \frac{1}{2}\right) \geq \frac{1}{4},$$

since $0 < \delta_{i-1} q_i < 1$, for every i. We thus get

$$-c_\alpha(N_3) \geq L/4.$$

- Consider $c_\alpha(N_1)$. We have $z_j^{(1)} \neq 0$ if and only if there exists $1 \leq k \leq L$ such that $j = 2j_k$, and if so $z_{2j_k} = \lceil \frac{a_{2j_k}}{2} \rceil$.

Let $1 \leq k \leq L$. We have

$$m_{2j_k-1}^{(1)} \leq \sum_{i\leq k-1} \lceil \frac{a_{2j_i}}{2} \rceil q_{2j_i-1} \leq q_{2j_{k-1}}.$$

Hence

$$m_{2j_k-1}^{(1)} + \frac{z_{2j_k}^{(1)} q_{2j_k-1}}{2} + \frac{1}{2} \leq \frac{2q_{2j_{k-1}} + q_{2j_k-1}\lceil a_{2j_k}/2\rceil}{2}.$$

Let us prove that

$$m^{(1)}_{2j_k-1} + \frac{z^{(1)}_{2j_k} q_{2j_k-1}}{2} + \frac{1}{2} \leq \frac{3}{8} q_{2j_k}.$$

Indeed, suppose a_{2j_k} odd. Write $a_{2j_k} = 2p+1$, with $p \geq 1$. Since $p \geq 1$ and $j_k - j_{k-1} \geq 2$, we have

$$(6p+3)q_{2j_k-1} + 3q_{2j_k-2} \geq (4p+4)q_{2j_k-1} + q_{2j_k-1} + 3q_{2j_k-2} \geq$$
$$(4p+4)q_{2j_k-1} + q_{2j_k-2} + q_{2j_k-3} + 3q_{2j_k-3} + 3q_{2j_k-4} \geq$$
$$(4p+4)q_{2j_k-1} + 8q_{2j_k-4}.$$

Consequently

$$2q_{2j_{k-1}} + q_{2j_k-1}(p+1) \leq \frac{3}{4}[(2p+1)q_{2j_k-1} + q_{2j_k-2}] = \frac{3}{4} q_{2j_k}.$$

Suppose a_{2j_k} even. Write $a_{2j_k} = 2p$, with $p \geq 1$. We similarly have

$$2q_{2j_{k-1}} + pq_{2j_k-1} \leq \frac{3}{4}[2pq_{2j_k-1} + q_{2j_k-2}] = \frac{3}{4} q_{2j_k}.$$

We thus get from (3)

$$c_\alpha(N_1) \geq \frac{1}{8} \sum_{1 \leq k \leq L} \lceil \frac{a_{2j_k}}{2} \rceil.$$

- Consider now $c_\alpha(N_2)$. We have $z^{(2)}_j \neq 0$ if and only if there exists $1 \leq k \leq L$ such that $j = 2j_k - 1$ or $j = 2j_k$, and then $z_{2j_k} = \lfloor \frac{a_{2j_k}}{2} \rfloor$, $z_{2j_k-1} = 1$.

  Let $1 \leq k \leq L$. We have

$$m^{(2)}_{2j_k-1} \leq \sum_{i \leq k-1} (\lfloor \frac{a_{2j_i}}{2} \rfloor q_{2j_i-1} + q_{2j_i-2}) + q_{2j_k-2} \leq \frac{q_{2j_{k-1}} - 1}{2} + q_{2j_k-2}.$$

  Hence

$$m^{(2)}_{2j_k-1} + \frac{z^{(2)}_{2j_k} q_{2j_k-1}}{2} + \frac{1}{2} \leq \frac{q_{2j_{k-1}} + q_{2j_k-1}\lfloor a_{2j_k}/2 \rfloor}{2} + q_{2j_k-2}$$
$$\leq \frac{1}{4} q_{2j_k} + q_{2j_k-2}.$$

We thus get

$$\frac{1}{2}-\delta_{2j_k-1}(m^{(2)}_{2j_k-1}+\frac{z^{(2)}_{2j_k}q_{2j_k-1}}{2}+\frac{1}{2})\geq -\delta_{2j_k-1}q_{2j_k-2}+\frac{1}{4}.$$

We have furthermore for the coefficients of $z^{(2)}_{2j_k-1}$:

$$-(\frac{1}{2}-\delta_{2j_k-2}(m^{(2)}_{2j_k-2}+\frac{z^{(2)}_{2j_k-1}q_{2j_k-2}}{2}+\frac{1}{2}))\geq -\frac{1}{2}+\frac{1}{2}\delta_{2j_k-2}q_{2j_k-2}.$$

¿From $\frac{1}{2}\delta_{2j_k-2}-\delta_{2j_k-1}\geq 0$, we get

$$c_\alpha(N_2)\geq -L/4.$$

Hence

$$c_\alpha(N_1)+c_\alpha(N_2)-c_\alpha(N_3)\geq \sum_{1\leq k\leq L}\frac{1}{8}\lceil\frac{a_{2j_k}}{2}\rceil.$$

Hence $c_\alpha(N_1)+c_\alpha(N_2)-c_\alpha(N_3)$ can be made arbitrarily large by taking L large enough.

It remains to consider the case where the partial quotients in the continued fraction expansion of α are eventually equal to 1. Let n_0 be such that $a_n=1$, for $n\geq n_0$. The denominators and numerators of the convergents satisfy the following linear recurrence relations:

$$\forall n\geq n_0,\quad p_{n+2}=p_{n+1}+p_n,$$

$$\forall n\geq n_0,\quad q_{n+2}=q_{n+1}+q_n.$$

Let $\tau=\frac{\sqrt{5}+1}{2}$. Hence there exists A,B,A',B' such that for $n\geq n_0$

$$q_n=A\tau^n+B(-\tau)^{-n},\ p_n=A'\tau^n+B'(-\tau)^{-n}. \tag{4}$$

Similarly

$$\forall n\geq n_0,\quad \delta_{n+2}=-\delta_{n+1}+\delta_n.$$

Furthermore, as δ_n tends to 0 and $\tau>1$, then there exists C such that

$$\delta_n=C\tau^{-n}.$$

Note that $AC<1/\tau$, since $\delta_n q_{n+1}<1$ for every n.

Let k_0 be such that $4k_0 - 2 \geq n_0$ and let $L \geq k_0$. Let

$$N_1 = \sum_{k_0 \leq k \leq L} q_{4k-1},$$

$$N_2 = \sum_{k_0 \leq k \leq L} q_{4k-2},$$

$$N_3 = N_1 + N_2 = \sum_{k_0 \leq k \leq L} q_{4k}.$$

The three quantities are again Ostrowski's expansions.

- Consider $c_\alpha(N_3)$. By using the expression (4) of q_n, it is easily seen that there exists a constant D (which does not depend on L) such that for every $k_0 \leq k \leq L$, one has:

  $$m_{4k}^{(3)} = \sum_{k_0 \leq i \leq k-1} q_{4i} \leq \frac{q_{4k}}{\tau^4 - 1} + D.$$

  Hence

  $$m_{4k}^{(3)} + \frac{z_{4k+1}^{(3)} q_{4k}}{2} + \frac{1}{2} \leq q_{4k}\left(\frac{1}{\tau^4 - 1} + \frac{1}{2}\right) + \left(D + \frac{1}{2}\right).$$

  Consequently

  $$\delta_{4k}\left(m_{4k}^{(3)} + \frac{z_{4k+1}^{(3)} q_{4k}}{2} + \frac{1}{2}\right) \leq AC\left(\frac{1}{\tau^4 - 1} + \frac{1}{2}\right) + C\left(D + \frac{1}{2}\right)\tau^{-4k}.$$

  Let $E = 1/2 - AC(\frac{1}{\tau^4-1} + \frac{1}{2})$. As $AC < 1/\tau$, then $E > 0$. Let $F = C(D + 1/2)$. We have

  $$-c_\alpha(N_3) \geq \sum_{k_0 \leq k \leq L} (E + F\tau^{-4k}).$$

- A similar computation for N_1 and N_2 leads to $c_\alpha(N_1) + c_\alpha(N_2)$ is bounded below (in L). Indeed there exist similarly D', D'', F', F'' (which do not depend on L) such that for every $k_0 \leq k \leq L$, one has:

  $$m_{4k-1}^{(1)} = \sum_{k_0 \leq i \leq k-1} q_{4i-1} \leq \frac{q_{4k-1}}{\tau^4 - 1} + D',$$

  $$c_\alpha(N_1) \geq \sum_{k_0 \leq k \leq L} (E + F'\tau^{-4k+1}),$$

$$m^{(2)}_{4k-2} = \sum_{k_0 \le i \le k-1} q_{4i-2} \ge \frac{q_{4k-2}}{\tau^4 - 1} + D'',$$

$$c_\alpha(N_2) \ge \sum_{k_0 \le k \le L} (-E + F''\tau^{-4k+2}),$$

hence $c_\alpha(N_1) + c_\alpha(N_2)$ is bounded from below.

In both cases (partial quotients eventually equal to 1 or not), we obtain that $|c_\alpha(N_1) + c_\alpha(N_2) - c_\alpha(N_3)|$ can be made arbitrarily large for a suitable choice of (N_1, N_2), and for L large enough. From (2), the same applies to $||\vec{0}, \vec{N}|_1 - N_1 N_2 \alpha|$, which ends the proof by contradicting (1). ■

Indeed we can deduce from the estimates in [5] an upper-bound for the growth order of $||\vec{P}, N \vec{+} P|_1 - N_1 N_2 \alpha|$, with $\vec{N} = (N_1, N_2) \in \mathbb{Z}^2_{>0}$ and $\vec{P} = (P_1, P_2) \in \mathbb{Z}^2$. Let us define a *balance function* $B(\vec{N})$ associated to a word f on $\mathbb{Z}^2$, for $\vec{N} = (N_1, N_2) \in \mathbb{Z}^2_{>0}$, as follows:

$$B(\vec{N}) := \max_{\vec{P}, \vec{P'} \in \mathbb{Z}^2} ||\vec{P}, \vec{N} + \vec{P}|_1 - |\vec{P'}, \vec{N} + \vec{P'}|_1|.$$

For a word taking its values in an alphabet $\mathcal{A}$ of size at least two, one defines similarly

$$B(\vec{N}) := \max_{a \in \mathcal{A}} \max_{\vec{P}, \vec{P'} \in \mathbb{Z}^2} ||\vec{P}, \vec{N} + \vec{P}|_a - |\vec{P'}, \vec{N} + \vec{P'}|_a|.$$

This function provides a quantitative measure of how far a word is from being balanced.

Theorem 8 *Let $\alpha \notin \mathbb{Q}$. Let $f : \mathbb{Z}^2 \to \{0, 1\}$ defined by*

$$\forall (m, n) \in \mathbb{Z}^2, \ (f(m, n) = 0 \iff (m + n)\alpha \in [0, 1 - \alpha) \ \textit{modulo } 1).$$

Let $\vec{N} = (N_1, N_2) \in \mathbb{Z}^2_{>0}$ and $\vec{P} = (P_1, P_2) \in \mathbb{Z}^2$. We have

$$||\vec{P}, \vec{N} + \vec{P}|_1 - N_1 N_2 \alpha| = o(\sup\{N_1, N_2\}),$$

$$B(\vec{N}) = o(\sup\{N_1, N_2\}).$$

Furthermore, if α has bounded partial quotients

$$||\vec{P}, \vec{N} + \vec{P}|_1 - N_1 N_2 \alpha| = O(\log(\sup\{N_1, N_2\})),$$

$$B(\vec{N}) = O(\log(\sup\{N_1, N_2\})).$$

Proof. Let us evaluate $|\vec{P}, P \vec{+} N|_1$, for $\vec{N} = (N_1, N_2) \in \mathbb{Z}^2_{>0}$ and $\vec{P} = (P_1, P_2) \in \mathbb{Z}^2$. We have

$$|\vec{P}, \vec{P} + \vec{N}|_1 = \sum_{n=P_2}^{P_2+N_2-1} \lfloor (N_1 + n + P_1)\alpha \rfloor - \sum_{n=P_2}^{P_2+N_2-1} \lfloor (n + P_1)\alpha \rfloor =$$

$$= \sum_{j=N_1+P_1+P_2}^{N_1+N_2+P_1+P_2-1} \lfloor j\alpha \rfloor - \sum_{j=P_1+P_2}^{P_1+P_2+N_2-1} \lfloor j\alpha \rfloor.$$

Hence

$$|\vec{P}, \vec{N} + \vec{P}|_1 - N_1 N_2 \alpha = \sum_{j=P_1+P_2}^{P_1+P_2+N_2-1} \{j\alpha\} - \sum_{j=N_1+P_1+P_2}^{P_1+P_2+N_1+N_2-1} \{j\alpha\} =$$

$$= c_\alpha(P_1 + P_2 + N_2 - 1) - c_\alpha(P_1 + P_2 - 1) +$$
$$+ c_\alpha(P_1 + P_2 + N_1 - 1) - c_\alpha(P_1 + P_2 + N_1 + N_2 - 1),$$

with the previous notation: $c_\alpha(N) = \sum_{j=1}^{N}(\{j\alpha\} - 1/2)$. It remains to apply the estimates of [5] (where more refined estimates can be found too), by noticing that it is sufficient to consider a finite number of values for $\vec{P}$:

1. For every α, $c_\alpha(N) = o(N)$ (this result was originally due to Sierpinski [25]).

2. If there exists A such that: for all i, $(1/t)\sum_{1 \le j \le t} a_j \le A$, then $c_\alpha(N) = O(\log N)$, and more precisely $|c_\alpha(N)| < \frac{3}{2} A \log N$ (cf. [16, 13, 22]).

∎

4.2 Imbalances in two-dimensional Sturmian words

A higer-dimensional generalisation of Sturmian words can be defined either on a two-letter alphabet or on a three-letter alphabet [29, 3].

Definition 1 *Let α, β, ρ be real numbers, with $1, \alpha, \beta$ rationally independent, and $0 < \alpha + \beta < 1$. We define the two-dimensional Sturmian word over the three-letter alphabet $\{1, 2, 3\}$ (with parameters α, β, ρ) as the function $f : \mathbb{Z}^2 \to \{1, 2, 3\}$, with*

$$\forall (m, n) \in \mathbb{Z}^2,\ (f(m, n) = i \Longleftrightarrow m\alpha + n\beta + \rho \in I_i \textit{ modulo } 1),$$

where

$$I_3 = [0, \alpha),\ I_2 = [\alpha, \alpha + \beta),\ I_1 = [\alpha + \beta, 1).$$

We similarly define the two-dimensional Sturmian word over the two-letter alphabet $\{0, 1\}$ *(with parameters* α, β, ρ*) as the function* $f : \mathbb{Z}^2 \to \{0, 1\}$*, with*

$$\forall (m, n) \in \mathbb{Z}^2,\ (f(m, n) = i \Longleftrightarrow m\alpha + n\beta + \rho \in I_i \textit{ modulo } 1),$$

where

$$I_0 = [0, 1 - \alpha),\ I_1 = [1 - \alpha, 1).$$

Note that these words are non-periodic (i.e., there is no non-zero vector of periodicity with integer coefficients) and uniformly recurrent. For further properties, see [29, 3, 4].

Recall that (classic) Sturmian words code the approximation of a line by a discrete line made of horizontal and vertical segments with integer vertices. Two-dimensional Sturmian words over a three-letter alphabet code discrete planes as follows. Consider the set of all unit cubes, with vertices at integer lattice points, which intersect a given plane. The discrete plane approximating this plane is the (upper or lower) surface of the union of these unit cubes. The discrete plane thus consists of three kinds of square faces, orientated according to the three coordinate planes. After projection, we obtain a tiling of the plane by three kinds of diamonds, being the projections of the square faces. This tiling is associated in a natural way with a $\mathbb{Z}^2$-lattice. We thus can code this tiling over a two dimensional sequence defined on a three-letter alphabet. Now one defines two-dimensional Sturmian sequences over a two-letter alphabet via a letter-to-letter projection p. With the notation in Definition 1, the projection p satisfies: $p(3) = 1$, $p(2) = p(1) = 0$, with the addition of $1 - \alpha$ to the parameter γ. Such sequences have many interesting combinatorial properties which allow us to consider them as a two-dimensional generalisation of Sturmian words. In particular, they are characterized among uniformly recurrent sequences by their rectangle complexity function $P(m, n) = mn + n$ [4]. Here the rectangle complexity function $P(m, n)$ counts the number of m by n rectangular factors.

Note that the words studied in the previous section correspond to the case $\alpha = \beta$.

Theorem 9 *Two-dimensional Sturmian words are not balanced. Let* $B(\vec{N})$*, for* $\vec{N} = (N_1, N_2) \in \mathbb{Z}^2_{>0}$*, be the balance function of a two-dimensional*

Sturmian sequence of parameters α, β, ρ (defined either on a three-letter or on a two-letter alphabet). We have

$$B(\vec{N}) = o(\sup\{N_1, N_2\}).$$

Furthermore, if α or β has bounded partial quotients, then

$$B(\vec{N}) = O(\log(\sup\{N_1, N_2\})).$$

Proof. Let f be a two-dimensional Sturmian word on the three-letter alphabet $\{1, 2, 3\}$, with parameters α, β, ρ. Suppose that f is C-balanced on each letter. Fix an index $j \in \mathbb{Z}$. This implies in particular that the one-dimensional word (in row) $f_j : \mathbb{Z} \to \{1, 2, 3\}$ defined by: $\forall m \in \mathbb{Z}$, $f_j(m) = f(m, j)$, is also C-balanced on the letter 2. We have

$$\forall m \in \mathbb{Z},\ (f_j(m) = 2 \Longleftrightarrow m\alpha + (j\beta + \rho) \in [\alpha, \alpha + \beta) \text{ modulo } 1).$$

However as $1, \alpha, \beta$ are rationally independent, we get that $\{2\}$ is not a bounded remainder set for f_j, from Theorem 1, hence the contradiction. Note that we similarly prove that the word f is not balanced neither on the letter 1 nor on the letter 3 (by considering words in columns). The same reasoning applies to two-letter Sturmian words.

Consider now a two-dimensional Sturmian word f on the two-letter alphabet $\{0, 1\}$:

$$\forall (m, n),\ f(m, n) = 1 \Longleftrightarrow m\alpha + n\beta + \gamma \in [1 - \alpha, 1).$$

Let us evaluate $|\vec{P}, \vec{P} + \vec{N}|_1$, for $\vec{N} = (N_1, N_2) \in \mathbb{Z}^2_{>0}$ and $\vec{P} = (P_1, P_2) \in \mathbb{Z}^2$. We have

$$|\vec{P}, \vec{P} + \vec{N}|_1 = \sum_{j=P_2}^{P_2+N_2-1} \lfloor (N_1 + P_1)\alpha + j\beta + \gamma \rfloor - \sum_{j=P_2}^{P_2+N_2-1} \lfloor P_1\alpha + j\beta + \gamma \rfloor.$$

Hence

$$|\vec{P}, \vec{N} + \vec{P}|_1 - N_1 N_2 \alpha = \sum_{j=P_2}^{P_2+N_2-1} \{j\beta + P_1\alpha + \gamma\} - \sum_{j=P_2}^{P_2+N_2-1} \{j\beta + (N_1 + P_1)\alpha + \gamma\}.$$

Let us introduce now, following the notation of [23], the non-homogeneous extension of $c_\alpha(N)$:

$$C_N(\alpha, \gamma) = \sum_{j=1}^{N} (\{j\alpha + \gamma\} - 1/2).$$

We thus have

$$|\vec{P}, \vec{N}+\vec{P}|_1 - N_1 N_2 \alpha = C_{P_2+N_2-1}(\beta, P_1\alpha+\gamma) - C_{P_2-1}(\beta, P_1\alpha+\gamma) -$$

$$C_{P_2+N_2-1}(\beta, (N_1+P_1)\alpha+\gamma) + C_{P_2-1}(\beta, (N_1+P_1)\alpha+\gamma).$$

We can get a similar expression involving quantities of the kind $C_N(\alpha, \gamma')$ by adding in columns instead of rows.

It then remains to apply the results of [23]: let $N = \sum_{i=1}^{t} z_i q_{i-1}$ be the Ostrowski expansion of N; we have

$$|C_N(\alpha, \gamma)| \leq \frac{3}{2} \sum_{i=1}^{t} z_i,$$

which implies (see for instance [5], Fact 1)

$$C_N(\alpha, \gamma) = o(N).$$

Let (a_i) be the partial quotients in the continued fraction expansion of α. If $(1/t) \sum_{1 \leq j \leq t} a_j \leq A$ for all t, then $C_N(\alpha, \gamma) < \frac{1}{3}(A+24)\log 3N$.

The case of a two-dimensional word on a three-letter alphabet can be handled exactly in the same way. One gets similar expressions for each of $|\vec{P}, \vec{N}+\vec{P}|_a$, with $a = 2, 3$, which is enough to conclude to the required estimates. ■

References

[1] E. ALTMAN, B. GAUJAL, A. HORDIJK *balanced sequences and optimal routing*, to appear in J. ACM.

[2] P. ARNOUX, G. RAUZY *Représentation géométrique de suites de complexité* $2n+1$, Bull. Soc. math. France **119** (1991), 199–215.

[3] V. BERTHÉ, L. VUILLON *A two-dimensional generalization of Sturmian sequences : tilings and rotations*, to appear in Discr. Math.

[4] V. BERTHÉ, L. VUILLON *Suites doubles de faible complexité*, to appear in J. Théor. Nombres Bordeaux.

[5] T. C. BROWN, P. J.-S. SHIUE *Sums of fractional parts of integer multiples of an irrational*, J. Number Theory **50** (1995), 181–192.

[6] J. CASSAIGNE, S. FERENCZI, L. Q. ZAMBONI *Imbalances in Arnoux-Rauzy sequences*, preprint.

[7] S. DULUCQ, D. GOUYOU-BEAUCHAMP *Sur les facteurs des suites de Sturm*, Theoret. Comput. Sci. **71** (1990), 381–400.

[8] X. DROUBAY, J. JUSTIN, G. PIRILLO *Episturmian words and some constructions of de Luca and Rauzy*, Theoret. Comput. Sci., to appear.

[9] I. FAGNOT, L. VUILLON *Generalized balances in Sturmian words*, preprint 2000.

[10] S. FERENCZI *Bounded remainder sets*, Acta Arith. **61** (1992), 319–326.

[11] A. HEINIS *On low-complexity* $\mathbf{Z}$*-words and their factors*, to appear.

[12] A. HEINIS, R.TIJDEMAN *Extending balanced and stiff words*, Report 97-15, Math. Inst. Leiden Univ., The Netherlands, 7 pp.

[13] G. H. HARDY, J. E. LITTLEWOOD *Some problems of Diophantine approximations: The lattice-points of a right-angled triangle*, Proc. London Math. Soc. **20** (1922), 15–36.

[14] P. HUBERT *Well balanced sequences*, to appear in Theoret. Comput. Sci.

[15] H. KESTEN *On a conjecture of Erdős and Szűsz related to uniform distribution mod 1*, Acta Arith. **12** (1966), 193–212.

[16] M. LERCH *Question 1547*, Intermédiaire Math. **11** (1904), 145–146.

[17] E. P. LIPATOV *A classification of binary collections and properties of homogeneity classes* (Russian), Problemy Kibernet. **39** (1982), 67–84.

[18] M. LOTHAIRE *Algebraic Combinatoric on Words, Chapter 2: Sturmian words*, by J. Berstel, P. Séébold, to appear.

[19] F. MIGNOSI *On the number of factors of Sturmian words*, Theoret. Comput. Sci. **82** (1991), 71-84.

[20] M. MORSE, G. A. HEDLUND *Symbolic dynamics*, Amer. J. Math. **60** (1938), 815–866.

[21] M. MORSE, G. A. HEDLUND *Symbolic dynamics II: Sturmian trajectories*, Amer. J. Math. **62** (1940), 1–42.

[22] A. OSTROWSKI *Bemerkungen zur Theorie der Diophantischen Approximationen I, II*, Abh. Math. Sem. Hamburg I (1922), 77–98 et 250–251.

[23] C. PINNER *On sums of fractional parts* $\{n\alpha + \gamma\}$, J. Number Theory **65** (1997), 48–73.

[24] G. RAUZY *Ensembles à restes bornés*, Séminaire de Théorie des Nombres de Bordeaux, 1983-1984, exposé 24. *On the discrepancy of* $n\alpha$, Acta Arith. **44** (1984), 241–279.

[25] W. SIERPINSKI *Pewne twierdzenie tyczace sie liczb niewymiernych. Un théorème sur les nombres irrationnels*, Bull. Internat. Acad. Polon. Sci. Lett. Cl. Sci. Math. Naturelles Sér. A (Cracovie) (1909), 725–727.

[26] R. TIJDEMAN *On complementary triples of Sturmian bisequences*, Indag. Math. N. S. **7** (1996), 419–424.

[27] R. TIJDEMAN *Fraenkel's conjecture for six sequences*, Discrete Math., to appear.

[28] R. TIJDEMAN *Exact covers of balanced sequences and Fraenkel's conjecture*, Proc. Number Theory Conference Graz (1998), De Gruyter.

[29] L. VUILLON *Combinatoire des motifs d'une suite sturmienne bidimensionnelle,* Theoret. Comput. Sci. **209** (1998), 261–285.

[30] A. ZORICH *Deviation for interval exchange transformations*, Ergod. Th. & Dynam. Sys. **17** (1997), 1477-1499.