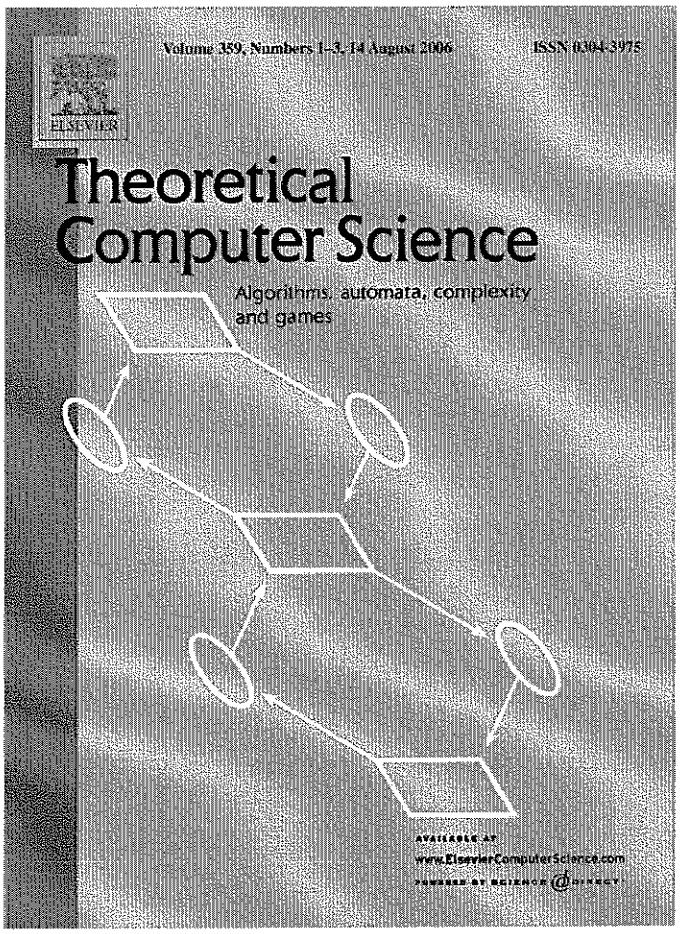
Provided for non-commercial research and educational use only.
Not for reproduction or distribution or commercial use.



This article was originally published in a journal published by Elsevier, and the attached copy is provided by Elsevier for the author's benefit and for the benefit of the author's institution, for non-commercial research and educational use including without limitation use in instruction at your institution, sending it to specific colleagues that you know, and providing a copy to your institution's administrator.

All other uses, reproduction and distribution, including without limitation commercial reprints, selling or licensing copies or access, or posting on open internet sites, your personal or institution's website or repository, are prohibited. For exceptions, permission may be sought for such use through Elsevier's permissions site at:

http://www.elsevier.com/locate/permissionusematerial

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Theoretical Computer Science 359 (2006) 255–281
Theoretical Computer Science
www.elsevier.com/locate/tcs

Averages of automatic sequences

A. Barbé*, F. von Haeseler

Department of Electrical Engineering, K.U. Leuven, Kasteelpark Arenberg 10, B 3001 Leuven, Belgium

Received 4 February 2005; received in revised form 23 March 2006; accepted 8 April 2006

Communicated by G. Ausiello

Abstract

We give some sufficient criteria for the existence of certain averages (mean, correlation functions) of generalized higher-dimensional automatic sequences and show how to calculate these averages. Then follows an exploration of the nature of necessary and sufficient conditions for the existence of averages. Some of these criteria are applied to averages which play a central role in the determination of the correlation function of an automatic sequence.
© 2006 Elsevier B.V. All rights reserved.

MSC: 11B85; 68Q45

Keywords: Automatic sequence; Substitution sequence; Thue–Morse sequence; Paperfolding sequence; Rudin–Shapiro sequence; Correlation function

1. Introduction

The motivation of this work is rooted in the mathematical study of crystals and quasicrystals. Before going into the technical details we present a nonrigorous and nontechnical approach to the problem of quasicrystals, for more information about the physical as well as about the mathematical background we refer to [21,4].

Usually a collection of infinitely many atoms is modeled by a countable set $X \in \mathbb{R}^m$ such that every sphere contains only finitely many points of X. The goal is to find ways to compute the diffraction pattern of this set X. This is known to be the Fourier transform of the autocorrelation function associated with the set X. As it turns out, computing the correlation function involves the computation of certain limits, which unfortunately need not exist. Thus for a given set one is first faced with the task to guarantee the existence of a unique correlation function for a set X (as a unique limit point of a well-chosen averaging sequence).

Due to the insurmountable difficulties to do so in the general situation, one is forced to restrict the choices of the set X. For example, if $X = \mathbb{Z}^m$, then X forms a very regular m-dimensional point lattice (the simplest perfect crystal model) and the correlation function is unique. Moreover, the computation of the correlation function in this case is almost trivial (see, e.g., [3, formula 3.6], for general regular lattices). Since the m-dimensional lattice $\mathbb{Z}^m$ is one of the simplest examples of an automatic set (a special type of automatic sequence, see [6]), it seems natural to consider automatic sets $X \subseteq \mathbb{Z}^m$ as models for quasicrystals or structures with long range aperiodic order and try to find means

* Corresponding author.
E-mail address: andre.barbe@esat.kuleuven.be (A. Barbé).

0304-3975/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.tcs.2006.04.001

of computing the corresponding correlation functions. Since automatic sets are a special case of automatic sequences, there is actually no difference with computing the correlation function of automatic sequences. Automatic sequences have classically been defined as one-dimensional sequences $f : \mathbb{N} \to S$, S being a finite symbol set, where the name automaticity originates from the fact that $f(n)$ can be generated by a finite automaton whose input sequence is given by the digit sequence in the number representation of n in a certain base (see e.g. [2]). Equivalently, these sequences are characterized by the fact that they have a finite number of decimations, i.e. properly defined subsequences, which constitute the so-called kernel of the sequence. Moreover, they are also the image, under a certain projection, of a fixed point of a constant length substitution. Famous prototype examples of automatic sequences with long-range aperiodic order are the Thue–Morse, paperfolding, and Rudin–Shapiro sequences, see e.g. [1], which have in essence a singular continuous, discrete, and absolute continuous spectrum, respectively.

Classical higher-dimensional versions of automatic sequences are obtained by Cartesian product extensions of the one-dimensional case (see, e.g. [2,20] for more details and further references). The m-dimensional automatic sequences that are under consideration here are sequences $f : \mathbb{Z}^m \to \mathbb{C}$ which are obtained by considering similar automata structures as for the classical one-dimensional case, but using number representations of points in $\mathbb{Z}^m$ instead of in $\mathbb{N}$. See Section 2, [13,6,7] for this particular higher-dimensional generalization of the automaticity concept, and [9] for the particular class of binary number systems in $\mathbb{Z}^m$. Also for this type of generalized higher-dimensional automatic sequences, the equivalence with the finiteness of a properly defined decimation kernel, and with the fixed point of a proper analog of a constant length substitution (though more complicated in nature than for the classical case, see Section 2), still holds.

The computation of the correlation function relies on the existence of certain averages. Let $a : \mathbb{Z}^m \to \mathbb{C}$ be a sequence. Its average is given as the arithmetic mean

$$\lim_{R\to\infty} \frac{1}{\mathrm{vol}(B_R(0))} \sum_{x\in B_R(0)\cap\mathbb{Z}^m} a(x),$$

provided that the limit exists, and where $B_R(0)$ is a ball centered at 0 of radius R with respect to a proper metric. The topic of this paper is to establish criteria on automatic sequences that guarantee the existence of such a limit, and to determine it.

Given two sequences $a, b : \mathbb{Z}^m \to \mathbb{C}$ then the correlation at $k \in \mathbb{Z}^m$ is the limit, if it exists,

$$\gamma_{ab}(k) = \lim_{R\to\infty} \frac{1}{\mathrm{vol}(B_R(0))} \sum_{x\in B_R(0)\cap\mathbb{Z}^m} a(x)\overline{b}(x+k),$$

i.e., the average of the product of a with shifted version of $\overline{b}$, where $\overline{b}(x)$ is the complex conjugate of $b(x)$. In [7], it was shown how the correlation function $\gamma_{gh}(k)$, $k \in \mathbb{Z}^m$, of any two kernel sequences g, h of an automatic sequence $f : \mathbb{Z}^m \to \mathbb{C}$ can be calculated solely from knowledge of

$$\gamma_{ab}(0) = \lim_{R\to\infty} \frac{1}{\mathrm{vol}(B_R(0))} \sum_{x\in B_R(0)\cap\mathbb{Z}^m} a(x)\overline{b}(x) \tag{1}$$

for all a, b in the kernel of f. Thus, in general, one needs to calculate $\gamma_{ab}(0)$. This is the average of the sequence $a(x)\overline{b}(x)$, and in Section 5 it will be shown that this sequence is automatic itself. In the last mentioned paper [7], we succeeded to calculate (1) for certain two-dimensional ± 1-valued versions of the Thue–Morse, the paperfolding and the Rudin–Shapiro sequences, using particularities of these sequences. In [8] the correlation function and related spectrum for multidimensional versions of the paperfolding and the Rudin–Shapiro sequences were derived, and in [10] for the multidimensional Thue–Morse sequences (basically also considering ± 1-valued versions, but by referring to results obtained in Example 1 in this paper the correlation function can also be calculated for other-valued versions of the Rudin–Shapiro and Thue–Morse sequences). For correlation functions and spectra for the "classical" higher-dimensional generalizations of the named sequences, see [14,18,19]. This present paper aims at establishing criteria for the existence of correlation functions in a more general setting.

The paper is organized as follows. In the next section we present the necessary technical prerequisites about the generalized type of higher-dimensional automatic sequences under consideration here. In the third section we present some sufficient conditions for the existence of the average, in the sense of arithmetic mean, of an automatic sequence. Taking certain "geometric" considerations into account it is even possible to establish necessary and sufficient conditions

for the existence of the average of one-dimensional and certain higher-dimensional sequences; this is done in Section 4. Finally, Section 5 uses the results of the previous sections in order to establish some sufficient conditions for the existence of a unique correlation function of an automatic sequence.

2. Preliminaries

We briefly recall the essentials concerning automatic sequences, for more information consult [2,13]. The notion of an m-dimensional automatic sequence $f : \mathbb{Z}^m \to \mathbb{C}$ is based on an *expanding* $m \times m$-matrix H, i.e. all eigenvalues have absolute value greater than 1, and an associated *complete residue set* W. This is a set $W \subset \mathbb{Z}^m$ such that for all $u, v \in W$, $u \neq v$, it holds that $u - v \notin H(\mathbb{Z}^m)$, and such that for every $x \in \mathbb{Z}^m$ there exists a $u \in W$ such that $x - u \in H(\mathbb{Z}^m)$. In particular, a complete residue set has cardinality

$$\chi_H = |\det H|.$$

If W is a complete residue set of H, then every $x \in \mathbb{Z}^m$ has a unique representation

$$x = H\kappa(x) + \zeta(x) \tag{2}$$

with $\kappa(x) \in \mathbb{Z}^m$ and $\zeta(x) \in W$. Notice that κ and ζ depend on both H and W. Since H is expanding, it follows that the iterates $\kappa^n(x)$ eventually stay in a ball centered at $0 \in \mathbb{R}^m$ with a radius depending on (H, W), see [13,15]. This shows that the sequence $(\kappa^n(x))_{n\in\mathbb{N}}$ becomes eventually periodic for all $x \in \mathbb{Z}^m$.

A complete residue set W of H is called a *complete digit set* if $0 \in \mathbb{Z}^m$ is the only periodic point of κ. Clearly this requires $0 \in W$. Furthermore, if W is a complete digit set for H, then every $x \in \mathbb{Z}^m$, $x \neq 0$, has a unique finite (H, W)-representation

$$x = H^n v_n + H^{n-1} v_{n-1} + \cdots + Hv_1 + v_0$$

with $v_i \in W$, and $v_n \neq 0$. In this case, (H, W) can be considered as a number system for $\mathbb{Z}^m$, where W provides the set of digits. For more on number systems we refer to [16,17].

In what follows W denotes always a complete residue set of H. Only if necessary, it will be mentioned explicitly that it is a complete digit set.

For $w \in W$, the (H, w)-decimation of a sequence $f : \mathbb{Z}^m \to \mathbb{C}$ is the sequence $\partial_{H,w}(f) : \mathbb{Z}^m \to \mathbb{C}$ defined as

$$\partial_{H,w}(f)(x) = f(Hx + w)$$

and we simply write ∂_w if H and W are clear from the context. Repeated application of decimations to f leads to

$$\partial_{v_n} \circ \partial_{v_{n-1}} \circ \cdots \circ \partial_{v_0}(f)(x) = f(H^{n+1}x + H^n v_n + H^{n-1} v_{n-1} + \cdots + v_0). \tag{3}$$

The (H, W)-*kernel* of f, denoted as $\ker(f) = \ker_{H,W}(f)$ is the set of all sequences obtained by repeated applications of the decimations ∂_w, $w \in W$ to f, together with f.

Definition 2.1. A sequence f is (H, W)-*automatic* if $\ker(f)$ is finite.

According to Theorem 3.2.2 in [13], the finiteness of the kernel is independent from the particular residue set W. It is therefore justified to speak of an H-automatic sequence. However, in what follows, we will almost always fix a residue set and, moreover, some of the presented results depend on the residue set W. Therefore, to emphasize this dependence we will speak of (H, W)-automatic sequence.

Before we continue we introduce some notation which will be useful for all what follows. If K is a finite set then $\mathbb{C}^K$ denotes the $\mathbb{C}$-vector space of functions $h : K \to \mathbb{C}$ with pointwise addition. The elements A of the vector space $\mathbb{C}^{K\times K}$ operate from the left on the elements of $\mathbb{C}^K$ by defining

$$(Ah)(k) = \sum_{l\in K} A(k, l)h(l)$$

for $k \in K$. The product AB of two elements $A, B \in \mathbb{C}^{K\times K}$ is an element of $\mathbb{C}^{K\times K}$ and it is defined as

$$C(k, l) = \sum_{m\in K} A(k, m)B(m, l)$$

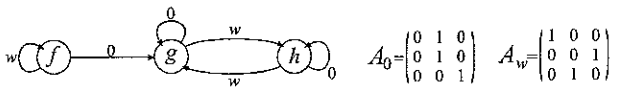


Fig. 1. A kernel-graph and corresponding decimation matrices. Row and column labeling of the matrices in the order f, g, h.

for $k, l \in K$. Note that after introducing a fixed linear order on the elements of the set K, these operations correspond to the common matrix multiplication from the left with a vector and the product of two matrices. In this sense we will use the notion $\mathbb{C}^{\ker(f)}$ as a vector whose components correspond to this fixed order of the elements of $\ker(f)$, and $\mathbb{C}^{\ker(f)\times\ker(f)}$ as a corresponding matrix.

The relation between the kernel elements of f is visualized by the so-called *kernel-graph*. This is a labeled directed graph where the vertices are the sequences in $\ker(f)$, and with a directed edge labeled $w \in W$ pointing from vertex g to vertex h if and only if $\partial_w(g) = h$. This graph can be described also by the *decimation matrices* $A_w \in \{0, 1\}^{\ker(f)\times\ker(f)}$, $w \in W$, i.e. matrices whose rows and columns are labeled by the kernel elements of f,

$$A_w = (a^w_{g,h}),$$

where $a^w_{g,h} = 1$ if $\partial_w(g) = h$ and $a^w_{g,h} = 0$ otherwise. Note that decimation matrices have precisely a single 1 in each row. As a consequence, products and sums of decimation matrices are non-negative matrices such that every row contains at least one positive entry.

As an example, Fig. 1 displays the kernel-graph and decimation matrices of an $(H, \{0, w\})$-automatic sequence with $\chi_H = 2$.

It turns out that an (H, W)-automatic sequence f can also be obtained as a fixed point of a substitution Σ_f on the set of sequences $F : \mathbb{Z}^m \to \mathbb{C}^{\ker(f)}$, defined as

$$\Sigma_f(F)(Hx + w) = A_w F(x) \tag{4}$$

for all $w \in W$ and all $x \in \mathbb{Z}^m$, where A_w is the corresponding decimation matrix. Then $\mathcal{F} : \mathbb{Z}^m \to \mathbb{C}^{\ker(f)}$ defined by

$$\mathcal{F}(x) = (g(x))_{g\in\ker(f)} \tag{5}$$

is a fixed point of this substitution, i.e., $\Sigma_f(\mathcal{F}) = \mathcal{F}$, see [13, Theorem 2.2.19]. $\mathcal{F}$ is a sequence of vectors over $\mathbb{Z}^m$. The element of $\mathcal{F}$ at $x \in \mathbb{Z}^m$ is the vector whose components correspond to the value of kernel sequences at x, in the given fixed order on $\ker(f)$.

When W is a complete digit set, and only in that case, then, as a consequence of (4) and using the fact that $x = \sum_{j=0}^{n} H^j v_j$ is the unique finite (H, W)-representation of $x \in \mathbb{Z}^m \setminus \{0\}$, one has

$$\mathcal{F}(x) = A_{v_0} A_{v_1} \cdots A_{v_n} \mathcal{F}(0) \tag{6}$$

and that $\mathcal{F}(0) = A_0 \mathcal{F}(0)$. In this case, the kernel-graph of an (H, W)-automatic sequence f can also be interpreted as a finite automaton that generates f, thus justifying the name "automatic sequence". This goes as follows: the representation $x = \sum_{j=1}^{n} H^{j-1} v_j$, $x \neq 0$, defines a path in the kernel-graph: the path starts in f, follows the arrows labeled $v_0, v_1, \ldots, v_n$ and terminates in an element $g \in \ker(f)$. The value of f at x is equal to the value of g at 0, i.e., $f(x) = g(0)$. This is a consequence of (3) with $x = 0$.

As argued in [7], in order to define a proper metric defining the ball $B_R(0)$ over which averages are taken, we assume that the expanding matrix is such that there exists a matrix $P \in \mathbb{R}^{m\times m}$ such that $\Lambda = P^{-1} H P$ is an $m \times m$ block diagonal matrix

$$\Lambda = P^{-1} H P = \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_s, \Lambda_1, \Lambda_2, \ldots, \Lambda_t), \tag{7}$$

where the λ_j, $|\lambda_j| > 1$, $j = 1, \ldots, s$, correspond to the real eigenvalues of H, and the Λ_j, $j = 1, \ldots, t$, are real 2×2-matrices of the form

$$\Lambda_j = \begin{pmatrix} a_j & -b_j \\ b_j & a_j \end{pmatrix}, \tag{8}$$

such that $|\det(\Lambda_j)| = a_j^2 + b_j^2 > 1$ for $j = 1, \ldots, t$. Each Λ_j corresponds to a pair of complex eigenvalues $(a_j \pm b_j \mathrm{i})$ of H.

If the elements of $\mathbb{R}^m$, where $m = s + 2t$, are denoted as $\mathbf{x} = (z_1, \ldots, z_s, x_1, y_1, x_2, y_2, \ldots, x_t, y_t)$, then it is easy to see that

$$\|\mathbf{x}\|_\infty = \max\left\{\max\{|z_i| \mid i = 1, \ldots, s\}, \max\left\{\sqrt{x_i^2 + y_i^2} \;\middle|\; i = 1, \ldots, t\right\}\right\}$$

defines a norm on $\mathbb{R}^m$. Moreover, the map $A : \mathbb{R}^m \to \mathbb{R}^m$ is expanding w.r.t. this norm. Then $H : \mathbb{R}^m \to \mathbb{R}^m$ is expanding w.r.t. the norm defined by

$$\|\mathbf{x}\| = \|P^{-1}\mathbf{x}\|_\infty \tag{9}$$

and one has $\|H(\mathbf{x})\| \geqslant c^- \|\mathbf{x}\|$ for all $\mathbf{x} \in \mathbb{R}^m$ and $c^- = \min\{c_i \mid i = 1, \ldots, s + t\}$, where $\underline{c} = (c_1, \ldots, c_s, c_{s+1}, \ldots, c_{s+t})$ is given as $c_i = |\lambda_i|$ for $i = 1, \ldots, s$ and $c_{s+i} = \sqrt{a_i^2 + b_i^2}$ for $i = 1, \ldots, t$.

Let $\underline{R}$ be an $(s + t)$-vector with positive real entries. Then the cylinder $\mathcal{C}(\underline{R}) \subset \mathbb{R}^m$ is defined as the set

$$\mathcal{C}(\underline{R}) = \{(z_1, \ldots, z_s) \mid |x_i| \leqslant R_i,\ i = 1, \ldots, s\} \times \{(x_1, y_1, \ldots, x_t, y_t) \mid x_j^2 + y_j^2 \leqslant R_{s+j}^2,\ j = 1, \ldots, t\}. \tag{10}$$

Then one easily sees that $A(\mathcal{C}(\underline{R})) = \mathcal{C}(\underline{cR})$, where $\underline{cR} = (c_1 R_1, \ldots, c_{s+t} R_{s+t})$. Note that, due to the expansion assumption on H one has that $c_i > 1$ for all $i = 1, \ldots, s + t$. We also introduce the notion $\underline{c}^l$ as $(c_1^l, \ldots, c_{s+t}^l)$ for $l \in \mathbb{Z}$. Finally we note that $H(P\mathcal{C}(\underline{R})) = P\mathcal{C}(\underline{cR})$. Recall that $\chi_H = |\det(H)|$, and note that

$$\chi_H = |\det(H)| = \prod_{i=1}^{s} c_i \prod_{j=1}^{t} c_{s+j}^2. \tag{11}$$

We conclude these preliminaries with two technical results which will be used several times throughout the text.

Let X be a Banach space, i.e., a real/complex, normed, complete vector space. Furthermore, let $\mathsf{A} : X \to X$ be a contraction, i.e., there exists $0 \leqslant \lambda < 1$ such that

$$\|\mathsf{A}x - \mathsf{A}y\| \leqslant \lambda \|x - y\|$$

for all $x, y \in X$. Note that, due to Banach's fixed point theorem, there exists a unique $x^* \in X$ with $\mathsf{A}x^* = x^*$.

Let $\mathsf{f} : \mathbb{R}^n_{\geqslant 0} \to X$ be a function which is bounded on every bounded subset of $\mathbb{R}^n_{\geqslant 0}$ and which satisfies

$$\mathsf{f}(\underline{R}) = \mathsf{A} \circ \mathsf{f}\left(\frac{\underline{R}}{\underline{c}}\right) + \varepsilon(\underline{R}), \tag{12}$$

for a contraction A and for all $\underline{R} = (R_1, \ldots, R_n)$, $R_i \geqslant 0$, $i = 1, \ldots, n$, and $\underline{c} = (c_1, \ldots, c_n)$ with $c_i > 1$, $i = 1, \ldots, n$, and $\underline{R}/\underline{c} = (R_1/c_1, \ldots, R_n/c_n)$. Furthermore, assume that

$$\lim_{\underline{R} \Rightarrow \infty} \varepsilon(\underline{R}) = 0,$$

where $\underline{R} \Rightarrow \infty$ means that every component of $\underline{R}$ tends to infinity.

Theorem 2.2 (*Barbé and von Haeseler [7]*). *Under the above assumptions*

$$\lim_{\underline{R} \Rightarrow \infty} \mathsf{f}(\underline{R}) = x^* \quad (\textit{where } x^* \textit{ is the unique solution of } \mathsf{A}x = x).$$

We also need the following summation formula [7, Lemma 3.10].

Author's personal copy

Lemma 2.3. *Let $H \in \mathbb{Z}^{m\times m}$ be an expanding integer matrix and let W be a complete digit set of H. Furthermore, let $F : \mathbb{Z}^m \to X$ be a bounded function, i.e., $\|F(x)\| \leqslant M$ for all $x \in \mathbb{Z}^m$. For $\underline{R} \in \mathbb{R}_+^{s+t}$ one has*

$$\sum_{x\in PC(\underline{R})\cap\mathbb{Z}^m} F(x) = \sum_{v\in W} \sum_{x\in PC(\underline{R}/c)\cap\mathbb{Z}^m} F(Hx+v) + o_C(\underline{R}),$$

where

$$\lim_{\underline{R}\Rightarrow\infty} \frac{o_C(\underline{R})}{\mathrm{vol}(PC(\underline{R}))} = 0$$

and $\mathrm{vol}(PC(\underline{R}))$ *denotes the volume of the cylinder* $PC(\underline{R})$.

3. Averages of automatic sequences: sufficient conditions

In this section we state some general sufficient criteria for the existence of the average, in the sense of arithmetic mean, of automatic sequences. For any sequence $F : \mathbb{Z}^m \to X$ with values in a Banach space X, we write

$$m_F(\underline{R}) = \frac{1}{\mathrm{vol}(PC(\underline{R}))} \sum_{x\in\mathbb{Z}^m\cap PC(\underline{R})} F(x), \tag{13}$$

where $\underline{R} = (R_1, \ldots, R_m) \in \mathbb{R}^m$ such that $R_i > 0$ for all $i = 1, \ldots, m$.

μ_F denotes the average of F, i.e.,

$$\mu_F = \lim_{\underline{R}\Rightarrow\infty} m_F(\underline{R}).$$

Let f be an (H, W)-automatic sequence, then the main theme of this section is the relationship between the average of f (existence as well as value) and the averages of all its kernel sequences $g \in \ker(f)$. Or, in other words, between μ_f and $\mu_{\mathcal{F}}$, where $\mathcal{F} : \mathbb{Z}^m \to \mathbb{C}^{\ker(f)}$ is the m-dimensional vector sequence defined by $\mathcal{F}(x) = (g(x))_{g\in\ker(f)}$. In a certain sense that will become clear soon, $\mu_{\mathcal{F}}$ is easier to deal with than μ_f, while the latter is just a component of the first.

A central role in this will be played by the *sum decimation matrix* S_W, given as the sum of the decimation matrices A_w, $w \in W$, defining the (H, W)-automatic sequence f, i.e.,

$$S_W = \sum_{w\in W} A_w.$$

Note that S_W is a non-negative matrix in which $S_W(g, h)$ represents the number of distinct decimation operations ∂_w, $w \in W$, that bring the kernel sequence g into kernel sequence h. Moreover, the sum over every row in S_W equals χ_H. This follows from the fact that each summand A_w is a decimation matrix, i.e., each row contains precisely one 1, and from the fact that there are $|W| = \chi_H$ summands. This implies that $\chi_H (= |\det H|)$ is the leading eigenvalue of S_W (with eigenvector having all its components equal to 1), i.e., every eigenvalue λ of S_W satisfies $|\lambda| \leqslant \chi_H$, see [12]. We remind the reader that a non-negative matrix $B \in \mathbb{R}^{\ker(f)\times\ker(f)}$ is called *primitive* if there exists an $n_0 \in \mathbb{N}$ such that for every pair $g, h \in \ker(f)$ one has $B^{n_0}(g, h) > 0$, see [12].

The first result deals with the case that S_W is a primitive matrix. This result is easily extended to the case of an irreducible matrix S_W. As an application of the obtained results we give a group theoretic criterion for the existence of the limit $\mu_{\mathcal{F}}$. All these results provide a link between the existence of μ_f and $\mu_{\mathcal{F}}$. Finally, we establish criteria for the existence of m_f without reference to the existence of $\mu_{\mathcal{F}}$.

3.1. Criteria for the existence of averages based on the sum decimation matrix S_W

3.1.1. S_W is primitive

The next theorem shows how the average of an automatic sequence is related to the average of a weighted sum of all kernel sequences.

Personal Copy

Theorem 3.1. *Let $f : \mathbb{Z}^m \to \mathbb{C}$ be (H, W)-automatic such that S_W is primitive. Then the following hold:*

(a) *Limit $\mu_{\mathcal{F}} = \lim_{\underline{R}\Rightarrow\infty} m_{\mathcal{F}}(\underline{R})$ exists if and only if there exist constants $\alpha_g \geqslant 0$, $g \in \ker(f)$, such that $\zeta = \sum_{g\in\ker(f)} \alpha_g > 0$ and such that for the sequence $\xi(x) = \sum_{g\in\ker(f)} \alpha_g g(x)$, $x \in \mathbb{Z}^m$, the limit*

$$\mu_\xi = \lim_{\underline{R}\Rightarrow\infty} m_\xi(\underline{R})$$

exists.

(b) *Moreover, all components of the limit $\mu_{\mathcal{F}}$ are equal to μ_ξ/ζ, i.e., all sequences in $\ker(f)$ have the same average.*

Proof.

- *Only if part* of (a): If $\lim_{\underline{R}\Rightarrow\infty} m_{\mathcal{F}}(\underline{R})$ exists, then it is clear that the constants $\alpha_g \geqslant 0$ as stated, and also μ_ξ, exist.
- *If part* of (a): Let us assume that

$$\lim_{\underline{R}\Rightarrow\infty} m_\xi(\underline{R}) = \mu_\xi$$

exists, with μ_ξ as defined above with proper α_g, ξ. Due to Lemma 2.3, we have

$$\sum_{x\in PC(\underline{R})\cap\mathbb{Z}^m} \mathcal{F}(x) = \sum_{v\in W} \sum_{x\in PC(\underline{R}/c)\cap\mathbb{Z}^m} \mathcal{F}(Hx+v) + o_C(\underline{R}).$$

Since $\mathcal{F}$ satisfies Eq. (4), we have $\mathcal{F}(Hx+v) = A_v\mathcal{F}(x)$, and therefore

$$\sum_{x\in PC(\underline{R})\cap\mathbb{Z}^m} \mathcal{F}(x) = \sum_{v\in W} \sum_{x\in PC(\underline{R}/c)\cap\mathbb{Z}^m} A_v\mathcal{F}(x) + o_C(\underline{R})$$

which becomes with $S_W = \sum_{v\in W} A_v$,

$$\sum_{x\in PC(\underline{R})\cap\mathbb{Z}^m} \mathcal{F}(x) = S_W \left(\sum_{x\in PC(\underline{R}/c)\cap\mathbb{Z}^m} \mathcal{F}(x) \right) + o_C(\underline{R}).$$

Since $\mathrm{vol}(PC(\underline{R})) = \chi_H \,\mathrm{vol}(PC(\underline{R}/c))$, see Eq. (11), one has

$$m_{\mathcal{F}}(\underline{R}) = \frac{1}{\chi_H} S_W m_{\mathcal{F}}(\underline{R}/c) + \varepsilon(\underline{R}), \tag{14}$$

with $\lim_{\underline{R}\Rightarrow\infty} \varepsilon(\underline{R}) = 0$.

Since S_W is primitive, there exists an $n_0 \in \mathbb{N}$ such that $S_W^{n_0}(g,h) \geqslant \max\{\alpha_{g'} \mid g' \in \ker(f)\}$ for all $g, h \in \ker(f)$. Repeated application of Eq. (14) leads to

$$m_{\mathcal{F}}(\underline{R}) = \frac{1}{\chi_H^{n_0}} S_W^{n_0} m_{\mathcal{F}}(\underline{R}/\underline{c}^{n_0}) + \varepsilon_{n_0}(\underline{R}). \tag{15}$$

Let $M \in \mathbb{C}^{\ker(f)\times\ker(f)}$ be defined as the matrix with constant rows

$$M(g,h) = \alpha_h \quad \text{for } g, h \in \ker(f)$$

and let $\delta \in \mathbb{C}^{\ker(f)}$ be defined as the vector with all components equal to 1, i.e., $\delta(g) = 1$ for all $g \in \ker(f)$. We rewrite Eq. (15) as

$$m_{\mathcal{F}}(\underline{R}) = \frac{1}{\chi_H^{n_0}} (S_W^{n_0} - M) m_{\mathcal{F}}(\underline{R}/\underline{c}^{n_0}) + \frac{1}{\chi_H^{n_0}} M m_{\mathcal{F}}(\underline{R}/\underline{c}^{n_0}) + \varepsilon_{n_0}(\underline{R}).$$

Since

$$\frac{1}{\chi_H^{n_0}} M m_{\mathcal{F}}(\underline{R}/\underline{c}^{n_0}) = \frac{1}{\chi_H^{n_0}} m_\xi(\underline{R}/\underline{c}^{n_0})\delta,$$

Author's personal copy

one obtains

$$m_{\mathcal{F}}(\underline{R}) = \frac{1}{\chi_H^{n_0}}(S_W^{n_0} - M)m_{\mathcal{F}}(\underline{R}/\underline{c}^{n_0}) + \frac{1}{\chi_H^{n_0}} m_\xi(\underline{R}/\underline{c}^{n_0})\delta + \varepsilon_{n_0}(\underline{R}).$$

Due to the assumption, one has

$$\lim_{\underline{R}\Rightarrow\infty} \frac{1}{\chi_H^{n_0}} m_\xi(\underline{R}/\underline{c}^{n_0})\delta = \frac{\mu_\xi}{\chi_H^{n_0}}\delta,$$

therefore one has

$$m_{\mathcal{F}}(\underline{R}) = \frac{1}{\chi_H^{n_0}}(S_W^{n_0} - M)m_{\mathcal{F}}(\underline{R}/\underline{c}^{n_0}) + \frac{\mu_\xi}{\chi_H^{n_0}}\delta + \tilde{\varepsilon}_{n_0}(\underline{R}). \tag{16}$$

Since S_W has leading eigenvalue χ_H, it follows that $(1/\chi_H^{n_0})S_W^{n_0}$ has leading eigenvalue 1. Due to the choice of n_0, $(1/\chi_H^{n_0})(S_W^{n_0} - M)$ is a non-negative matrix such that the sum over each row is equal to $1 - \zeta/\chi_H^{n_0}$, which shows that its leading eigenvalue is equal to $1 - \zeta/\chi_H^{n_0} < 1$, implying that $(1/\chi_H^{n_0})(S_W^{n_0} - M)$ is a contraction. As Eq. (16) is of the form (12), with $\mathsf{f}(\underline{R}) = m_{\mathcal{F}}(\underline{R})$ and

$$\mathsf{A} \circ \mathsf{f}\left(\frac{\underline{R}}{\underline{c}}\right) = \frac{1}{\chi_H^{n_0}}(S_W^{n_0} - M)\mathsf{f}(\underline{R}/\underline{c}^{n_0}) + \frac{\mu_\xi}{\chi_H^{n_0}}\delta,$$

Theorem 2.2 applies, showing that $\mu_{\mathcal{F}} = \lim_{\underline{R}\Rightarrow\infty} m_{\mathcal{F}}(\underline{R})$ exists. Moreover, and this proves part (b) of the assertion, Theorem 2.2 also shows that $\mu_{\mathcal{F}}$ is the unique solution of the equation

$$\mu_{\mathcal{F}} = \frac{1}{\chi_H^{n_0}}(S_W^{n_0} - M)\mu_{\mathcal{F}} + \frac{\mu_\xi}{\chi_H^{n_0}}\delta.$$

Since δ (vector with all components equal to 1) is an eigenvector of $S_W^{n_0} - M$ with respect to the leading eigenvalue $\chi_H^{n_0} - \zeta$, it follows that $\mu_{\mathcal{F}} = (\mu_\xi/\zeta)\delta$ is the solution of the above equation. It is a vector with all components equal to μ_ξ/ζ. □

The particular meaning of this theorem is that, if S_W is primitive, then the average of f exists if and only if the averages of all sequences in $\ker(f)$ exist, and that all these averages are equal.

Remark. As we will see in Example 4, item (4), the condition that $\alpha_g \geqslant 0$ in Theorem 3.1 is necessary. Indeed, that example shows an automatic sequence f for which μ_f does not exist although there exist $\alpha_g < 0$ and a sequence $\xi = \sum_{g\in\ker(f)} \alpha_g g$ for which μ_ξ does exist.

Example 1. *Thue–Morse sequence*: Let $(H, \{0, w\})$ be a binary number system for $\mathbb{Z}^m$ [9] such that H is conjugate to a block diagonal matrix, see Eq. (7). Then the decimation matrices

$$A_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad A_w = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

define the sequence $\mathcal{F}(x) = (t(x), \hat{t}(x))$ as $\mathcal{F}(0) = (a, b)^{\mathrm{T}}$, $a, b \in \mathbb{C}$, and for $x \neq 0$ recursively as

$$\mathcal{F}(Hx + v) = A_v\mathcal{F}(x). \tag{17}$$

The first component of $\mathcal{F}$, i.e. the sequence $t(x)$, forms an m-dimensional Thue–Morse sequence with $\ker(t) = \{t, \hat{t}\}$. Due to the form of the decimation matrices, one sees that $\mathcal{F}(x)$ is either equal to $(a, b)^{\mathrm{T}}$ or to $(b, a)^{\mathrm{T}}$. This shows that the sequence $\xi(x) = t(x) + \hat{t}(x)$ is the constant sequence $a + b$, hence

$$\mu_\xi = a + b.$$

Author's personal copy

Since $S_W = A_0 + A_w$ is a positive matrix, it is primitive and it follows from Theorem 3.1 that

$$\mu_t = \mu_{\hat{t}} = \frac{a+b}{2}.$$

Rudin–Shapiro sequence: Let $(H, \{0, w\})$ be a binary number system for $\mathbb{Z}^m$. The decimation matrices

$$A_0 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad A_w = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

define the sequence $\mathcal{F}(x) = (r(x), r_1(x), r_2(x), r_3(x))^{\mathrm{T}}$ using the recursion (4) with $\mathcal{F}(0) = (a, a, b, b)^{\mathrm{T}}$, $a, b \in \mathbb{C}$. Then the first component $(r(x))_{x\in\mathbb{Z}^m}$ defines an m-dimensional Rudin–Shapiro sequence, see [8], with $\ker(r) = \{r, r_1, r_2, r_3\}$. Since $\mathcal{F}(x) \in \{(a, a, b, b)^{\mathrm{T}}, (a, b, a, b)^{\mathrm{T}}, (b, a, b, a)^{\mathrm{T}}, (b, b, a, a)^{\mathrm{T}}\}$, it follows that the sequence $\xi(x) = r(x) + r_1(x) + r_2(x) + r_3(x)$ satisfies $\xi(x) = 2a + 2b$ for all $x \in \mathbb{Z}^m$ and therefore $\mu_\xi = 2a + 2b$. Since the matrix $S_W = A_0 + A_w$ satisfies $S_W^3 > 0$, it is primitive. According to Theorem 3.1 it follows that

$$\mu_h = \frac{a+b}{2}$$

for all $h \in \ker(r)$.

3.1.2. S_W is irreducible

We will briefly describe how Theorem 3.1 can be applied also in the case that the matrix S_W is *irreducible*. This means that for every $g, h \in \ker(f)$, there exists an $n = n(g, h)$ such that $S_W^n(g, h) > 0$, or in terms of decimations: for all $g, h \in \ker(f)$ one has that $g \in \ker(h)$ and $h \in \ker(g)$. Then it is known, see [12], that either S_W is primitive or that S_W can be written as

$$S_W = \begin{pmatrix} 0 & Q_0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & Q_1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & Q_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & \vdots \\ Q_{p-1} & 0 & 0 & 0 & \cdots & 0 \end{pmatrix} \tag{18}$$

with $p \geqslant 2$. If S_W is primitive, then Theorem 3.1 applies. Otherwise one has that S_W^p is a block diagonal matrix $\mathrm{diag}(B_0, \ldots, B_{p-1})$, and each matrix B_i is primitive [12]. In terms of the kernel of f, this translates to: there exist mutually disjoint sets $K_i \subset \ker_{H,W}(f)$, $i = 0, \ldots, p-1$, such that $f \in K_0$, $\ker_{H,W}(f) = K_0 \cup \cdots \cup K_{p-1}$ and such that

$$K_{(i+1) \bmod p} = \{\partial_v(g) \mid g \in K_i,\ v \in W\}.$$

In other words, the decimations of the elements in K_i form the set $K_{(i+1) \bmod p}$. Note further, due to Theorem 3.2.1 in [13], that f is H^p-automatic. Therefore, one has that the (H^p, W_p)-kernel of f, with $W_p = \{v_0 + Hv_1 + H^2v_3 + \cdots + H^{p-1}v_{p-1} \mid v_i \in W,\ i = 0, \ldots, p-1\}$, is K_0 and the associated matrix S_{W_p} is the primitive matrix $B_0 = Q_0Q_1 \ldots Q_{p-1}$. Therefore, if S_W is irreducible but nonprimitive, Theorem 3.1 applies to f regarded as an (H^p, W_p)-automatic sequence.

3.1.3. S_W is not irreducible

We will now deal with the case that S_W is *not irreducible*, i.e., there are vertices in the kernel-graph which cannot be reached from other vertices. Let f be (H, W)-automatic. A subset K of $\ker(f)$ is called *invariant* if

$$\partial_w(K) \subseteq K$$

for all $w \in W$. An invariant set $K \subseteq \ker(f)$ is called minimal if there exists no invariant set K' such that K' is a proper subset of K.

With $A_w^{(K)}$ we denote the restriction of the decimation matrices A_w of f to an invariant set $K \subseteq \ker(f)$. If K is a minimal invariant set, then $\ker(g) = K$ for all $g \in K$, and this yields:

Lemma 3.2. *Let f be (H, W)-automatic. If K is a minimal invariant subset of* $\ker(f)$, *then $S^{(K)} = \sum_{w\in W} A_w^{(K)}$ is irreducible.*

Let f be (H, W)-automatic such that f is not contained in a minimal invariant subset of $\ker(f)$. If $K_i, i = 1, \ldots, l$, denotes the collection of all minimal invariant subsets of $\ker(f)$, then all $S^{(K_i)}$ are irreducible. By similar arguments as given in the case that S_W is irreducible, one can consider f as an (H^q, W_q)-automatic sequence such that the corresponding decimation matrices A_v, $v \in W_q$, restricted to the minimal invariant sets $K_i' \subset \ker_{H^q,W_q}(f)$, $i = 1, \ldots, l'$, are primitive. Bearing this in mind one has:

Theorem 3.3. *Let f be (H, W)-automatic such that* $\ker(f)$ *is not a minimal invariant set. Let $K_i \subset \ker(f)$, $i = 1, \ldots, l$, be the collection of all minimal invariant sets of* $\ker(f)$ *and assume that the matrices $S^{(K_i)}$, $i = 1, \ldots, l$, are primitive.*

If there exist $g_i \in K_i$, $i = 1, \ldots, l$, such that $\mu_{g_i} = \lim_{\underline{R}\Rightarrow\infty} m_{g_i}(\underline{R})$ exist, then $\mu_f = \lim_{\underline{R}\Rightarrow\infty} m_f(\underline{R})$ exists.

Proof. We prove that under these conditions, $\mu_{\mathcal{F}}$ exists, thus implying the existence of μ_f. There exists $K_0 \subset \ker(f)$ such that $f \in K_0$ and $\ker(f) = K_0 \cup K_1 \cup \cdots \cup K_l$. Then S_W is of the form

$$S_W = \begin{pmatrix} Q_0 & * & * & \cdots & * \\ 0 & S^{(K_1)} & 0 & \cdots & 0 \\ \vdots & 0 & \ddots & \ddots & \vdots \\ 0 & \vdots & 0 & S^{(K_{l-1})} & 0 \\ 0 & 0 & \cdots & 0 & S^{(K_l)} \end{pmatrix},$$

where Q_0 is the restriction of S_W on K_0. Since K_0 does not contain any invariant subset of $\ker(f)$, it follows that there exists an n_0 such that the row sums of $Q_0^{n_0}$ are less than $\chi_H^{n_0}$. By Eq. (15) one has

$$m_{\mathcal{F}}(\underline{R}) = \frac{1}{\chi_H^{n_0}} S_W^{n_0} m_{\mathcal{F}}(\underline{R}/\underline{c}^{n_0}) + \varepsilon_{n_0}(\underline{R}). \tag{19}$$

Moreover, note that $S_W^{n_0}$ has the same form as S_W above, but with the n_0th power of the blocks on the diagonal instead.

Let $\mathcal{F} = (\mathcal{F}_0, \ldots, \mathcal{F}_l)$, where $\mathcal{F}_i$ denotes the restriction of $\mathcal{F}$ on K_i, $i = 0, \ldots, l$. By Theorem 3.1 and due to the assumption $\lim_{\underline{R}\Rightarrow\infty} m_{g_i}(\underline{R}) = \mu_{g_i}$ for $i = 1, \ldots, l$ it follows for $i = 1, \ldots, l$ that $\lim_{\underline{R}\Rightarrow\infty} m_{\mathcal{F}_i}(\underline{R}) = \mu_{g_i}\delta_i$, where δ_i is the $|K_i|$-vector with all components 1.

Using this fact, together with Eq. (19) and the form of $S_W^{n_0}$, one obtains an equation for $m_{\mathcal{F}_0}(\underline{R})$, namely,

$$m_{\mathcal{F}_0}(\underline{R}) = \frac{1}{\chi_H^{n_0}} Q_0^{n_0} m_{\mathcal{F}_0}(\underline{R}/\underline{c}^{n_0}) + \xi + \tilde{\varepsilon}(\underline{R}),$$

where $\xi \in \mathbb{C}^{K_0}$ is determined by $\delta_i\mu_{g_i}$ and $S_W^{n_0}$-matrices, and also $\lim_{\underline{R}\Rightarrow\infty} \tilde{\varepsilon}(\underline{R}) = 0$. Due to the choice of n_0 one has that $Q_0^{n_0}$ is a non-negative matrix with row sums less than $\chi_H^{n_0}$, i.e., the eigenvalues of $Q_0^{n_0}$ have modulus less than $\chi_H^{n_0}$. In other words, the matrix $(1/\chi_H^{n_0})Q_0^{n_0}$ is a contraction. By this, Theorem 2.2 implies the existence of the limit $\lim_{\underline{R}\Rightarrow\infty} m_{\mathcal{F}_0}(\underline{R})$ which is the unique solution of

$$\mu_{\mathcal{F}_0} = \frac{1}{\chi_H^{n_0}} Q_0^{n_0} \mu_{\mathcal{F}_0} + \xi. \qquad \square \tag{20}$$

Author's personal copy

Example 2.

Paperfolding sequence: Let $(H, W = \{0, w\})$ be a binary number system for $\mathbb{Z}^m$ like in the preceding examples. The decimation matrices

$$A_0 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad A_w = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

define the sequence $\mathcal{F}(x) = (p(x), g(x), \alpha(x), \beta(x))^{\mathrm{T}}$ recursively by using

$$\mathcal{F}(Hx + v) = A_v \mathcal{F}(x)$$

for $v \in W$, and starting from $\mathcal{F}(0) = (a, a, a, b)^{\mathrm{T}}$, $a, b \in \mathbb{C}$. Then the first component $(p(x))_{x\in\mathbb{Z}^m}$ is a so-called m-dimensional paperfolding sequence, with $\ker(p) = \{p, g, \alpha, \beta\}$. For later reference, note that one finds quickly that premultiplication of $\mathcal{F}(0)$ with all possible products of A_0, A_w (see (6)) shows that $\mathcal{F}(x)$ takes a value in $\{(a, a, a, b)^{\mathrm{T}}, (a, b, a, b)^{\mathrm{T}}, (b, a, a, b)^{\mathrm{T}}, (b, b, a, b)^{\mathrm{T}}\}$.

Since the decimation matrices imply that $\partial_v(\alpha) = \alpha$ and $\partial_v(\beta) = \beta$ for $v \in W$, it follows that $\alpha(x) = a$ and $\beta(x) = b$ for all $x \in \mathbb{Z}^m$. Moreover, $\{\alpha\}, \{\beta\}$ is the collection of the minimal sets of $\ker(p)$. Since α and β are constant sequences, one has $\mu_\alpha = a$ and $\mu_\beta = b$. By Theorem 3.3 the limit $\mu_\mathcal{F}$ exists. With

$$S_W = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix} \quad \text{it follows that } Q_0 = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}.$$

As Q_0 has eigenvalues 1 and 0, and $\chi_H = 2$, it is clear that setting $n_0 = 1$ in the previous theorem makes $Q_0^{n_0}/2^{n_0} = Q_0/2$ already contracting. As a consequence, (20) becomes

$$\mu_{\mathcal{F}_0} = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}\mu_{\mathcal{F}_0} + \begin{pmatrix} 0 \\ \dfrac{a+b}{2} \end{pmatrix} \quad \text{yielding } \mu_{\mathcal{F}_0} = \begin{pmatrix} \dfrac{a+b}{2} \\ \dfrac{a+b}{2} \end{pmatrix}.$$

It is plain that the above example in connection with Theorem 3.3 readily generalizes to: if f is an automatic sequence such that all minimal invariant components are formed by constant sequences, then μ_f exists.

3.1.4. Yet another S_W-based criterion for the existence of the average $\mu_\mathcal{F}$

The next theorem provides a sufficient condition for the existence of the limit $\mu_\mathcal{F}$ without referring to the primitivity or irreducibility of S_W.

Theorem 3.4. *Let $f : \mathbb{Z}^m \to \mathbb{C}$ be (H, W)-automatic, and let W be a residue set of H such that*

$$\lim_{n\to\infty} \frac{1}{\chi_H^n} S_W^n \mathcal{F}(x) = \eta(x) \quad \textit{for all } x \in \mathbb{Z}^m.$$

If the limit $\mu_\eta = \lim_{\underline{R}\Rightarrow\infty} m_\eta(\underline{R})$ exists, then

$$\mu_\mathcal{F} = \lim_{\underline{R}\Rightarrow\infty} m_\mathcal{F}(\underline{R}) = \mu_\eta.$$

Proof. Since the set $\{\mathcal{F}(x) \mid x \in \mathbb{Z}^m\}$ is a finite set, we may assume uniform convergence, i.e., for every $\varepsilon > 0$ there exists n_0 such that

$$\frac{S_W^n}{\chi_H^n} \mathcal{F}(x) = \eta(x) + \varepsilon_n(x) \tag{21}$$

with $\|\varepsilon_n(x)\| < \varepsilon$ for all $n \geqslant n_0$ and for all $x \in \mathbb{Z}^m$.

For $n \geqslant 1$ we set $W_n = W + HW + \cdots + H^{n-1}W$, then W_n is a residue set of H^n. Furthermore, set

$$v_n = \max\{\|P^{-1}v\|_\infty \mid v \in W_n\},$$

see Eq. (7) for P.

Let $n \in \mathbb{N}$ and let $\underline{R} = (R_1, \ldots, R_{s+t}) > v_n$, i.e., all $R_i > v_n, i = 1, \ldots, s+t$. Let $v \in \mathbb{R}$ then $(R_1+v, \ldots, R_{s+t}+v)$ is written as $\underline{R} + v$.

(a) Then one has that

$$x \in PC\left(\frac{\underline{R} - v_n}{\underline{c}^n}\right) \quad \text{implies } H^n x + v \in PC(\underline{R}) \text{ for all } v \in W_n. \tag{22}$$

Indeed, if $x = Py = P(y_1, \ldots, y_{s+t})^{\mathrm{T}}$, then $x \in PC((\underline{R} - v_n)/\underline{c}^n)$ is equivalent to $|y_i| < (R_i - v_n)/c_i^n$.

Let $v \in W_n$, since $HP = P\Lambda$, it follows that $H^n x + v = H^n Py + v = P\Lambda^n y + v = P(\Lambda^n y + P^{-1}v) = P(z_1, \ldots, z_{s+t})^{\mathrm{T}}$, one gets $|z_i| \leqslant c_i^n |y_i| + v_n < R_i$ for all $i = 1, \ldots, s+t$. This means that $H^n x + v \in PC(\underline{R})$ and proves assertion (22).

(b) For $n \in \mathbb{N}$ let $\underline{R} > v_n$, then due to (22) one has

$$PC(\underline{R}) \cap \mathbb{Z}^m = \left(\bigcup_{\substack{x \in PC((\underline{R}-v_n)/\underline{c}^n) \\ \text{and } x \in \mathbb{Z}^m}} H^n x + W_n \right) \cup C(n, \underline{R}),$$

where $C(n, \underline{R})$ are all points in $PC(\underline{R}) \cap \mathbb{Z}^m$ which are not in the first sets. As far as the cardinality of these sets is concerned, we note that

$$|PC(\underline{R}) \cap \mathbb{Z}^m| = \chi_H^n \left| PC\left(\frac{\underline{R} - v_n}{\underline{c}^n}\right) \cap \mathbb{Z}^m \right| + |C(n, \underline{R})|. \tag{23}$$

This partition of $PC(\underline{R}) \cap \mathbb{Z}^m$ induces also a split of the sum $\sum_{x \in PC(\underline{R}) \cap \mathbb{Z}^m} \mathcal{F}(x)$, i.e., one has a sum which runs over the first set and a sum which runs over $C(n, \underline{R})$. Now observe that the sum over the first set can be split further, namely, for $x \in PC((\underline{R} - v_n)/\underline{c}^n) \cap \mathbb{Z}^m$ one has the partial sum

$$\sum_{x' \in H^n x + W_n} \mathcal{F}(x'),$$

which can be written as

$$\sum_{x' \in H^n x + W_n} \mathcal{F}(x') = \sum_{v_0, \ldots, v_{n-1} \in W} \mathcal{F}(H^n x + H^{n-1} v_{n-1} + \cdots + v_0).$$

Using the property (4) of $\mathcal{F}$ this sum becomes

$$\sum_{v_0, \ldots, v_{n-1} \in W} A_{v_0} \ldots A_{v_{n-1}} \mathcal{F}(x),$$

which is the same as $S_W^n \mathcal{F}(x)$. This shows that

$$\sum_{x \in PC(\underline{R}) \cap \mathbb{Z}^m} \mathcal{F}(x) = \sum_{\substack{x \in PC((\underline{R}-v_n)/\underline{c}^n) \\ \text{and } x \in \mathbb{Z}^m}} S_W^n \mathcal{F}(x) + \sum_{x \in C(n, \underline{R})} \mathcal{F}(x). \tag{24}$$

(c) We are now prepared to establish that $\lim_{\underline{R} \Rightarrow \infty} m_{\mathcal{F}}(\underline{R}) = \mu_\eta$.

Let $\varepsilon > 0$, then, due to the observed uniformity, there exists an $n_0 \in \mathbb{N}$ such that Eq. (21) holds.

Then Eq. (24) can be written as

$$\sum_{x \in PC(\underline{R}) \cap \mathbb{Z}^m} \mathcal{F}(x) = \sum_{x \in PC((\underline{R}-v_{n_0})/\underline{c}^{n_0})} (\chi_H^{n_0} \eta(x) + \chi_H^{n_0} \varepsilon(x)) + \sum_{x \in C(n_0, \underline{R})} \mathcal{F}(x).$$

Author's personal copy

Division by $\mathrm{vol}(PC(\underline{R}))$ gives

$$m_{\mathcal{F}}(\underline{R}) = \frac{\chi_H^{n_0}}{\mathrm{vol}(PC(\underline{R}))} \sum_{x \in PC((\underline{R}-v_{n_0})/\underline{c}^{n_0})} \eta(x) + \frac{\chi_H^{n_0}}{\mathrm{vol}(PC(\underline{R}))} \sum_{x \in PC((\underline{R}-v_{n_0})/\underline{c}^{n_0})} \varepsilon(x) + \frac{1}{\mathrm{vol}(PC(\underline{R}))} \sum_{x \in C(n_0, \underline{R})} \mathcal{F}(x).$$

Note that the first summand converges to μ_η as $\underline{R}$ goes to infinity. It remains to show that the sum of the last two terms, denoted $\mathcal{R}(\underline{R}, n_0)$, vanishes when $\underline{R} \Rightarrow \infty$.

Using the fact that $\|\varepsilon(x)\| < \varepsilon$ for all $x \in \mathbb{Z}^m$ and that the sequence $\|\mathcal{F}(x)\|$, $x \in \mathbb{Z}^m$, is bounded, i.e., $\|\mathcal{F}(x)\| \leqslant M$ for all $x \in \mathbb{Z}^m$, one obtains

$$\|\mathcal{R}(\underline{R}, n_0)\| \leqslant \frac{\chi_H^{n_0}\varepsilon}{\mathrm{vol}(PC(\underline{R}))} |PC((\underline{R} - v_{n_0})/\underline{c}^{n_0}) \cap \mathbb{Z}^m| + \frac{M}{\mathrm{vol}(PC(\underline{R}))} |C(n_0, \underline{R})|.$$

Now Corollary 3.9 in [7] states that for all $\underline{R} > 0$,

$$|PC(\underline{R}) \cap \mathbb{Z}^m| = \mathrm{vol}(PC(\underline{R})) + o_C(\underline{R}),$$

where $o_C(\underline{R})$ satisfies

$$\lim_{\underline{R} \Rightarrow \infty} \frac{o_C(\underline{R})}{\mathrm{vol}(PC(\underline{R}))} = 0.$$

This together with (23) allows us to write

$$\|\mathcal{R}(\underline{R}, n_0)\| \leqslant \frac{\varepsilon\chi_H^{n_0}\mathrm{vol}(PC((\underline{R} - v_{n_0})/\underline{c}^{n_0}))}{\mathrm{vol}(PC(\underline{R}))} + \frac{M(\mathrm{vol}(PC(\underline{R})) - \chi_H^{n_0}\mathrm{vol}(PC((\underline{R} - v_{n_0})/\underline{c}^{n_0})))}{\mathrm{vol}(PC(\underline{R}))} + o_C(\underline{R}).$$

Noting that

$$\mathrm{vol}(PC(\underline{R})) = \rho_m \prod_{i=1}^{s+t} R_i^{\gamma_i},$$

where ρ_m is a positive constant independent from $\underline{R}$ and $\gamma_i = 1$ for $i = 1, \ldots, s$ and $\gamma_i = 2$ for $i = s+1, \ldots, s+t$ and noting that (see Eq. (11))

$$\chi_H = \prod_{i=1}^{s+t} c_i^{\gamma_i}$$

the above inequality transforms into

$$\|\mathcal{R}(\underline{R}, n_0)\| \leqslant \varepsilon + M\left(1 - \prod_{i=1}^{s+t}\left(1 - \frac{v_{n_0}}{R_i}\right)^{\gamma_i}\right) + o_C(\underline{R}).$$

This shows that there exists an $R_0 > 0$ such that

$$M\left(1 - \prod_{i=1}^{s+t}\left(1 - \frac{v_{n_0}}{R_i}\right)^{\gamma_i}\right) < \varepsilon$$

and $|o_C(\underline{R})| < \varepsilon$ for all $\underline{R} > R_0$. Hence we have shown that

$$\|\mathcal{R}(\underline{R}, n_0)\| \leqslant 3\varepsilon$$

for all $\underline{R} = (R_1, \ldots, R_m)$ with $R_i > \max\{R_0, v_{n_0}\}$, $i = 1, \ldots, m$. This completes the proof. □

Remark. If it happens that the sequence of matrices $((1/\chi_H^n)S_W^n)_{n\in\mathbb{N}}$ has itself a limit S_∞, then the above theorem states that the existence of μ_η with $\eta(x) = S_\infty\mathcal{F}(x)$ guarantees that $\mu_{\mathcal{F}} = \mu_\eta$. The Thue–Morse, Rudin–Shapiro and

Author's personal copy

paperfolding sequences presented in the previous examples all satisfy the conditions mentioned in this remark, and hence the conditions of Theorem 3.4. Indeed, that S_∞ exists for the Thue–Morse and Rudin–Shapiro sequences follows from the fact that S_W is primitive (see e.g. [12]). For the paperfolding sequence, S_∞ can be determined by exploiting the simple structure of S_W which allows to find an explicit expression for each element of $(S_W/\chi_H)^n$. The results are

$$S_\infty = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}, \quad S_\infty = \begin{pmatrix} \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{pmatrix}, \quad S_\infty = \begin{pmatrix} 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

for the Thue–Morse, Rudin–Shapiro and paperfolding cases, respectively. Moreover, taking into account the possible values for $\mathcal{F}(x)$ as determined in the previous examples shows that $(S_\infty \mathcal{F}(x))_{x\in\mathbb{Z}^m}$ is in each of these cases a constant sequence, and thus has a limit. As will be demonstrated in Example 5, the existence of equal limits $\lim_{n\to\infty}(1/\chi_H^n)S_W^n\mathcal{F}(x)$ for all x is not a necessary condition for the existence of the limit of $m_{\mathcal{F}}(\underline{R})$.

The next corollary provides a useful extension of Theorem 3.4 which will be used in the next section. A map $G : \mathbb{Z}^m \to \mathbb{Z}^m$ is called *H-adapted* if G is bijective and if there exists a $C > 0$ such that

$$PC(\underline{R} - C) \subseteq G(PC(\underline{R})) \subseteq PC(\underline{R} + C)$$

holds for all $\underline{R} > 0$. If G is H-adapted, then one sees that for every bounded $F : \mathbb{Z}^m \to X$ (where X is any Banach space)

$$\sum_{x\in PC(\underline{R})} F(x) - \sum_{x\in PC(\underline{R})} F(G(x)) = o_C(\underline{R}) \tag{25}$$

with

$$\lim_{\underline{R}\Rightarrow\infty} \frac{o_C(\underline{R})}{\mathrm{vol}(PC(\underline{R}))} = 0.$$

The extension of Theorem 3.4 reads as:

Corollary 3.5. *Let* $\mathbf{G}$ *be a finite set of H-adapted maps. If* $f : \mathbb{Z}^m \to \mathbb{C}$ *is* (H, W)*-automatic and* W *a residue set of* H *such that*

$$\lim_{n\to\infty} \frac{1}{|\mathbf{G}|\chi_H^n} S_W^n \left(\sum_{G\in\mathbf{G}} \mathcal{F}(G(x)) \right) = \eta(x)$$

for all $x \in \mathbb{Z}^m$ *and such that* μ_η *exists, then*

$$\lim_{\underline{R}\Rightarrow\infty} m_{\mathcal{F}}(\underline{R}) = \mu_\eta.$$

Proof (*Sketch*). One certainly has a uniform convergence of

$$\lim_{n\to\infty} \frac{1}{|\mathbf{G}|\chi_H^n} S_W^n \left(\sum_{G\in\mathbf{G}} \mathcal{F}(G(x)) \right) = \eta(x).$$

The fundamental Eq. (24) reads

$$\sum_{x\in PC(\underline{R})} \mathcal{F}(x) = \sum_{x\in PC((\underline{R}-v_n)/\underline{c}^n)} S_W^n \mathcal{F}(x) + \sum_{x\in C(n,\underline{R})} \mathcal{F}(x).$$

Since Eq. (25) applies to every $G \in \mathbf{G}$ a summation of these $|\mathbf{G}|$ equalities and a subsequent normalization leads to

$$\sum_{x\in PC((\underline{R}-v_n)/\underline{c}^n)} S_W^n \mathcal{F}(x) = \frac{1}{|\mathbf{G}|} \sum_{x\in PC((\underline{R}-v_n)/\underline{c}^n)} S_W^n \left(\sum_{G\in\mathbf{G}} \mathcal{F}(g(x)) \right) + o_C(\underline{R}).$$

Author's personal copy

Using this identity in the fundamental equation and following similar lines of reasoning as in the proof of Theorem 3.4 yields the desired result. □

3.1.5. About μ_f without referring to $\mu_{\mathcal{F}}$

This section states results that ensure the existence of μ_f rather than $\mu_{\mathcal{F}}$. We remind first that $S_W^n(g,h)$ is, after a proper linear ordering of the kernel elements, the entry of the nth power of S_W in row g at column h.

Corollary 3.6. *Let $f : Z^m \to \mathbb{C}$ be (H, W)-automatic. If*

$$\lim_{n\to\infty} \frac{1}{\chi_H^n} \sum_{g\in\ker(f)} S_W^n(f,g)g(x) = \eta(x)$$

exists for all $x \in \mathbb{Z}^m$ and if

$$\lim_{\underline{R}\Rightarrow\infty} m_\eta(\underline{R}) = \mu_\eta,$$

then

$$\lim_{\underline{R}\Rightarrow\infty} m_f(\underline{R}) = \mu_\eta.$$

Proof. First observe that

$$\sum_{g\in\ker(f)} S_W^n(f,g)g(x)$$

is the scalar product of the vector $(S_W^n(f,g))_{g\in\ker(f)}$ with the vector $\mathcal{F}(x)$. Since $\mathcal{F}(x)$ has only finitely many values, the convergence of

$$\lim_{n\to\infty} \frac{1}{\chi_H^n} \sum_{g\in\ker(f)} S_W^n(f,g)g(x) = \eta(x)$$

is uniform in $x \in \mathbb{Z}^m$. Now the proof is very similar to the proof of Theorem 3.4, using the same notation. In particular, Eq. (24) becomes

$$\sum_{x\in PC(\underline{R})} f(x) = \sum_{x\in PC((\underline{R}-v_n)/\underline{c}^n)} \left(\sum_{g\in\ker(f)} S_W^n(f,g)g(x) \right) + \sum_{x\in C(n\underline{R})} f(x).$$

Let $\varepsilon > 0$, then there exists an n_0 such that

$$\sum_{g\in\ker(f)} S_W^n(f,g)g(x) = \chi_H^{n_0}\eta(x) + \chi_H^{n_0}\varepsilon(x)$$

with $\|\varepsilon(x)\| < \varepsilon$ for all $x \in \mathbb{Z}^m$. Using this and dividing by $\mathrm{vol}(PC(\underline{R}))$ one obtains

$$\frac{1}{\mathrm{vol}(PC(\underline{R}))} \sum_{x\in PC(\underline{R})} f(x) = \frac{\chi_H^{n_0}}{\mathrm{vol}(PC(\underline{R}))} \sum_{x\in PC((\underline{R}-v_{n_0})/\underline{c}^{n_0}))} \eta(x) + \frac{\chi_H^{n_0}}{\mathrm{vol}(PC(\underline{R}))} \sum_{x\in PC((\underline{R}-v_{n_0})/\underline{c}^{n_0}))} \varepsilon(x) + \frac{\chi_H^{n_0}}{\mathrm{vol}(PC(\underline{R}))} \sum_{x\in C(n_0,\underline{R})} f(x). \tag{26}$$

The first sum on the right-hand side converges to μ_η as $\underline{R} \Rightarrow \infty$, due to the assumptions on η. By arguments similar to those of the proof of Theorem 3.4 and the fact that $\|\varepsilon(x)\| < \varepsilon$ for all $x \in \mathbb{Z}^m$, there exists an $R_0 > 0$ such that

$$\left| \frac{\chi_H^{n_0}}{\mathrm{vol}(PC(\underline{R}))} \sum_{x\in PC((\underline{R}-v_{n_0})/\underline{c}^{n_0}))} \varepsilon(x) + \frac{\chi_H^{n_0}}{\mathrm{vol}(PC(\underline{R}))} \sum_{x\in C(n_0,\underline{R})} f(x) \right| < 2\varepsilon$$

for all $\underline{R} > R_0$. □

Author personal copy

As a consequence of Eq. (26) we also have:

Corollary 3.7. *Let $f : \mathbb{Z}^m \to \mathbb{C}$ be (H, W)-automatic. If there exists an $n_0 \in \mathbb{N}$ such that the sequence*

$$\xi(x) = \frac{1}{\chi_H^{n_0}} \sum_{g \in \ker(f)} S_W^{n_0}(f, g)g(x)$$

has the limit μ_ξ, then

$$\lim_{\underline{R} \Rightarrow \infty} m_f(\underline{R}) = \mu_\xi.$$

Proof. Consider Eq. (26) for $\varepsilon(x) = 0$ and replace $\eta(x)$ by $\xi(x)$, taking the limit $\underline{R} \Rightarrow \infty$ proves the corollary. □

Example 3. Let p be the (H, W)-paperfolding sequence from the previous example, then one computes that

$$\lim_{n \to \infty} \frac{1}{2^n} \sum_{g \in \ker(p)} S_{p,W}^n(p, g)g(x) = \frac{\alpha(x)}{2} + \frac{\beta(x)}{2} = \frac{a+b}{2}$$

is a constant sequence. Due to Corollary 3.6, $\mu_p = (a+b)/2$, confirming the earlier result. Note further that Corollary 3.7 is not immediately applicable.

Let t be the Thue–Morse sequence, then one has

$$\frac{1}{2} \sum_{g \in \ker(t)} S_{t,W}(t, g)g(x) = \frac{1}{2}(t(x) + \hat{t}(x)) = \frac{a+b}{2},$$

implying $\mu_t = (a+b)/2$, due to Corollary 3.7.

3.2. A group-theoretic criterion for the existence of averages

If $f : \mathbb{Z}^m \to \mathbb{C}$ is (H, W)-automatic with decimation matrices A_w, $w \in W$, then the set of all possible products of decimation matrices, i.e.,

$$\Gamma_{f,W} = \{A_{v_0} \dots A_{v_n} \mid v_0, \dots, v_n \in W,\ n \in \mathbb{N}\}$$

forms a semigroup under matrix multiplication. As already observed in [6], properties of this semigroup are related to certain geometric properties of automatic sets. Note that the elements of $\Gamma_{f,W}$ are elements of $\{0, 1\}^{\ker(f) \times \ker(f)}$ and have exactly one 1 in each row, i.e., they have the same characteristics as a decimation matrix. The elements in $\Gamma_{f,W}$ also have an interpretation for the kernel-graph of f. If

$$\gamma = A_{w_0} \dots A_{w_n} \in \Gamma_{f,W},$$

then $\gamma(g, h) = 1$ if and only if the path $w_0, w_1, \dots, w_n$ which starts in g terminates in h. In terms of decimations this means

$$\partial_{w_n} \circ \cdots \circ \partial_{w_0}(g) = h.$$

Thus there is a natural relation between products of the decimation matrices and the composition of the decimation maps.

As far as the limit $\mu_{\mathcal{F}}$ is concerned one has:

Theorem 3.8. *Let $f : \mathbb{Z}^m \to \mathbb{C}$ be (H, W)-automatic with W a complete digit set. If $\Gamma_{f,W}$ is a group, then all components of $\mu_{\mathcal{F}}$ are equal to*

$$\mu_{\mathcal{F}} = \frac{1}{|\ker(f)|} \sum_{g \in \ker(f)} g(0).$$

Author Personal Copy

Proof. Since $\Gamma_{f,W}$ is a group, it follows that the decimation matrices are permutation matrices, i.e., each row and column contains precisely one 1. This is equivalent to the fact that the maps $\partial_w : \ker(f) \to \ker(f)$, $w \in W$, are bijective.

Consider the sequence ξ defined as

$$\xi(x) = \sum_{g \in \ker(f)} g(x)$$

for $x \in \mathbb{Z}^m$. Since each $\partial_w : \ker(f) \to \ker(f)$ is bijective it follows that for all $w \in W$

$$\partial_w(\xi) = \xi.$$

Due to the fact that W is a complete digit set it follows that ξ is the constant sequence with value

$$\mu = \sum_{g \in \ker(f)} g(0).$$

As a next step we show that S_W is irreducible, i.e., for every pair $g, h \in \ker(f)$ there exist $n \in \mathbb{N}$ and $w_0, \ldots, w_n \in W$ such that $\partial_{w_0} \circ \cdots \circ \partial_{w_n}(g) = h$. If $g \in \ker(f)$, $f \neq g$, then there exist $n \in \mathbb{N}$ and $w_0, \ldots, w_n \in W$ such that $\partial_{w_n} \circ \cdots \circ \partial_{w_0}(f) = g$. It remains to establish the existence of a path back to f. The decimation $\Delta = \partial_{w_n} \circ \cdots \circ \partial_{w_0} : \ker(f) \to \ker(f)$ corresponds to the matrix product

$$\gamma = A_{w_0} \ldots A_{w_n} \in \Gamma_{f,W}.$$

Since $f \neq g$, it follows that γ is not the identity matrix, I, and since $\Gamma_{f,W}$ is a finite group there exists $M \geqslant 2$ such that $\gamma^M = I$. This means that Δ^M is the identity on $\ker(f)$ and therefore $f = \Delta^M(f) = \Delta^{M-1}(\Delta(f)) = \Delta^{M-1}(g)$. This shows that S_W is irreducible. Two cases are possible.

Case S_W is primitive: Then the assertion follows from Theorem 3.1, taking into account that the sequence $\xi(x)$ is the constant sequence with value μ.

Case S_W is not primitive: Then S_W can be written as in Eq. (18) with an appropriate partition $\ker(f) = K_0 \cup K_1 \cup \cdots \cup K_{p-1}$ with $f \in K_0$. Since $\Gamma_{f,W}$ is a group, the maps $\partial_w : K_i \to K_{(i+1) \bmod p}$ with $w \in W$ are all bijective. This implies that all K_i have the same cardinality ρ, yielding $|\ker(f)| = p\rho$, and that the matrices Q_i in Eq. (18) are of size $\rho \times \rho$.

If the sequences ξ_i, $i = 0, \ldots, p-1$, are defined as

$$\xi_i(x) = \sum_{g \in K_i} g(x)$$

for $x \in \mathbb{Z}^m$, then they are identical and constant with value μ/p. Indeed, one has $\partial_v(\xi_i) = \xi_{(i+1) \bmod p}$ for all $v \in W$, in particular one has $\xi_i(0) = \xi_{(i+1) \bmod p}(0)$ for all $i = 0, \ldots, p-1$. This shows that each ξ_i is the constant sequence with value μ/p.

Since S_W^p is a diagonal block matrix with primitive matrices on the diagonal, there exists an n_0 such that $S_W^{pn_0}$ has positive block matrices on the diagonal. Let M be the $(\ker(f) \times \ker(f))$-diagonal block matrix, defined by

$$M(g,h) = \begin{cases} 1 & \text{if } g, h \in K_i \text{ for an } i = 0, \ldots, p-1, \\ 0 & \text{otherwise,} \end{cases}$$

i.e., $M(g,h) = 1$ if and only if $S_W^{pn_0}(g,h) \geqslant 1$. This yields, according to (15),

$$m_{\mathcal{F}}(\underline{R}) = \frac{1}{\chi_H^{pn_0}} (S_W^{pn_0} - M) m_{\mathcal{F}}(\underline{R}/\underline{c}^{pn_0}) + \frac{1}{\chi_H^{pn_0}} M m_{\mathcal{F}}(\underline{R}/\underline{c}^{pn_0}) + \varepsilon_{pn_0}(\underline{R}).$$

Now, by its very construction, one has that

$$\frac{1}{\chi_H^{pn_0}} M m_{\mathcal{F}}(\underline{R}/\underline{c}^{pn_0}) = \frac{1}{\chi_H^{pn_0}} m_v(\underline{R}/\underline{c}^{pn_0}),$$

where v is the constant sequence with value $(\mu/p)\delta$ (recall that δ is the vector with $|\ker(f)|$ components all equal to 1). One therefore has

$$\lim_{\underline{R}\Rightarrow\infty} \frac{1}{\chi_H^{pn_0}} m_v(\underline{R}/\underline{c}^{pn_0}) = \frac{\mu}{p\chi_H^{pn_0}}\delta.$$

Since $S_W^{pn_0} - M$ is a non-negative matrix with all row sums equal to $\chi_H^{pn_0} - \rho < \chi_H^{pn_0}$, it follows that $m_{\mathcal{F}}(\underline{R})$ has a limit and that the limit is the solution of the equation

$$x = \frac{1}{\chi_H^{pn_0}}(S_W^{pn_0} - M)x + \frac{\mu}{p\chi_H^{pn_0}}\delta,$$

which is $(\mu/|\ker(f)|)\delta$, since δ is an eigenvector of $S_W^{pn_0} - M$ with eigenvalue $\chi_H^{pn_0} - \rho$, and since $|\ker(f)| = p\rho$. □

Remarks.

(1) For the m-dimensional Thue–Morse sequence t introduced in Example 1, the semigroup $\Gamma_{t,W}$ is indeed a group. This provides another reason why the limit exists.

(2) The m-dimensional Rudin–Shapiro sequence r introduced in Example 1 shows that a limit $\mu_{\mathcal{F}}$ exists even though $\Gamma_{f,W}$ is not a group.

(3) The fact that W is a complete digit set is essential for the above theorem. This is illustrated by the following one-dimensional example with $H = 3$. Then $W = \{-4, 0, 1\}$ is a complete residue set of H, which is not a complete digit set, since $2 = H \cdot 2 + (-4)$, i.e., $\kappa_{W,H}(2) = 2$. Define the decimation matrices $A_{-4} = A_0 = A_1 = 1$. Then it can be shown that there exists a nonultimately periodic sequence $f : \mathbb{Z} \to \mathbb{C}$ such that

$$\partial_w(f) = f \tag{27}$$

for all $w \in \{-4, 0, 1\}$. Moreover, $f(0) = a$ and $f(2) = b$, with $a, b \in \mathbb{C}$ completely determine all solutions of (27), see [11,5] for more details. The most general form of this sequence $f(k)$ for $k = 0, \ldots, 69$ is

$$\begin{array}{l} a\ a\ b\ a\ a\ a\ b\ b\ a\ a\ a\ a\ a\ a\ b\ a\ a\ b\ b\ b\ a\ b\ b\ a\ a\ a\ a\ a\ a\ a\ a\ a\ a\ a \\ \quad a\ a\ a\ b\ a\ a\ a\ b\ b\ a\ a\ a\ b\ a\ a\ b\ b\ b\ b\ b\ b\ a\ b\ b\ b\ a\ a\ b\ b\ b\ a\ b\ b\ b\ a\ a\ldots. \end{array}$$

Both a and b are recurring in $(f(k))_{k>0}$. For $k < 0$, $f(k) = a$.

Now consider $V = \{-1, 0, 1\}$ which is a complete digit set for $H = 3$. Using the fact that $f(3x + w) = f(x)$ for $w \in \{-4, 0, 1\}$ one can compute the kernel of f w.r.t. the complete digit set V. This produces the three kernel elements f, $g(x) = f(x + 1)$ and $h(x) = f(x + 2)$ (we consider this order f, g, h to be the order of the kernel elements) and the decimation matrices

$$A'_0 = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad A'_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad A'_{-1} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

Then $\mathcal{F} = (f, g, h)^{\mathrm{T}}$ is given by $\mathcal{F}(0) = (f(0), g(0), h(0)) = (a, a, b)$, $a, b \in \mathbb{C}$ (a fixed point of A'_0), and $\mathcal{F}(3x + v) = A'_v\mathcal{F}(x)$ for $v \in V$.

Moreover, using the theory of Mahler equations for automatic sequences, it can be shown that every sequence f which satisfies Eq. (27) is H-automatic and therefore possesses a related vector sequence $\mathcal{F}$, see [13, Chapter 5.2].

Note that $\Gamma_{f,V}$ is not a group and note further that $S_V = A'_{-1} + A'_0 + A'_1$ is primitive, as S_V^2 is positive. The recursive definition of $\mathcal{F}$ leads to $\mathcal{F}(x) = (a, a, a)^{\mathrm{T}}$ for $x < 0$, $\mathcal{F}(3) = (a, a, a)^{\mathrm{T}}$, and $\mathcal{F}(17) = (b, b, b)^{\mathrm{T}}$. This induces

$$\lim_{n\to\infty} \frac{1}{3^n} S_V^n \mathcal{F}(0) = \frac{7a + b}{8}\delta,$$

$$\lim_{n\to\infty} \frac{1}{3^n} S_V^n (\mathcal{F}(-3) + \mathcal{F}(3)) = a\delta,$$

$$\lim_{n\to\infty} \frac{1}{3^n} S_V^n (\mathcal{F}(-17) + \mathcal{F}(17)) = (a + b)\delta,$$

where $\delta = (1, 1, 1)^{\mathrm{T}}$. According to Theorem 4.2 (see next section) these three limits are equal if the limit $\lim_{R\to\infty} m_{\mathcal{F}}(R)$ exists. This is only possible if $a = b = 0$. Thus if $a, b \in \mathbb{C}$ are different, then $\mu_{\mathcal{F}}$ does not exist. Since S_V is primitive, Theorem 3.1 implies that μ_f does not exist either. Thus, this example shows that f is 3-automatic, $\Gamma_{f,W}$ is a group, W is not a complete digit set and μ_f does not exist. As a consequence, $\Gamma_{f,W}$ being a group only implies that μ_f exists in case W is a complete digit set.

4. Necessary and sufficient conditions for the existence of averages for some restricted classes of automatic sequences

So far the existence of the average of an H-automatic sequence was linked to algebraic properties (in the broadest sense) of the sum decimation matrix S_W without explicit reference to particularities of the expanding matrix H and the residue set. This shows that the results of Section 3 are "universal". In this section, we establish several necessary and sufficient conditions for the existence of $\mu_{\mathcal{F}}$, and hence of μ_f, for several special cases in which suitable "geometric" properties of H and W play a crucial role. I.e., we put restrictions on the expanding matrices H as well as on the complete digit sets and *also* on the dimension of $\mathbb{Z}^m$.

4.1. One-dimensional k-automatic sequences over $\mathbb{N}$

We begin with the "classical case" of k-automatic sequences $f : \mathbb{N} \to \mathbb{C}$, for which H is the (single element) matrix representing the map $H : \mathbb{N} \to \mathbb{N}$ with $H(x) = kx$, $k \geqslant 2$, and complete digit set $W = \{0, 1, \ldots, k-1\}$ (the only complete digit set for $\mathbb{N}$). Note that in this case, $\chi_H = k$.

Theorem 4.1. *Let $f : \mathbb{N} \to \mathbb{C}$ be k-automatic with complete digit set $W = \{0, \ldots, k-1\}$. Then*

$$\mu_{\mathcal{F}} = \lim_{R\to\infty} \frac{1}{R} \sum_{0\leqslant x<R} \mathcal{F}(x) = \mu \tag{28}$$

if and only if

$$\lim_{n\to\infty} \frac{1}{\chi_H^n} S_W^n \mathcal{F}(x) = \mu \tag{29}$$

for all $x \in \mathbb{N}$.

Proof.

If part: If (29) holds for all $x \in \mathbb{N}$, then the proof of Theorem 3.4 shows, by replacing the cylinders by $[0, R[$, $R > 0$, that

$$\lim_{R\to\infty} \frac{1}{R} \sum_{0\leqslant x<R} \mathcal{F}(x) = \mu. \tag{30}$$

Only if part: Because of the specific form of the complete digit set W for $\mathbb{N}$, it holds that

$$\frac{1}{k^n} \sum_{0\leqslant x<k^n} \mathcal{F}(x) = \frac{1}{k^n} S_W^n \mathcal{F}(0).$$

If (29) holds, the limit as $n \to \infty$ on the left-hand side of this equality is equal to μ, and the assertion is proved for $x = 0$. The rest of the proof is by induction on x. Suppose that

$$\lim_{n\to\infty} \frac{1}{k^n} S_W^n \mathcal{F}(x) = \mu$$

is true for $x = 0, \ldots, N-1$. Now, again as a consequence of the specific form of the complete digit set W, one has for $n \in \mathbb{N}$ that

$$\frac{1}{(N+1)k^n} \sum_{0\leqslant x<(N+1)k^n} \mathcal{F}(x) = \frac{1}{(N+1)k^n} S_W^n (\mathcal{F}(0) + \cdots + \mathcal{F}(N-1) + \mathcal{F}(N)).$$

Due to the assumption (30), the limit for the left-hand side of this equality for $n \to \infty$ exists and is equal to μ. This shows that also the limit of the right-hand side exists, and due to the induction hypothesis, the limit form of the previous equation becomes

$$\mu = \frac{N}{N+1}\mu + \lim_{n\to\infty} \frac{1}{(N+1)k^n} S_W^n \mathcal{F}(N).$$

This yields $\mu = \lim_{n\to\infty}(1/\chi_H^n) S_W^n \mathcal{F}(N)$. By induction, this holds for all N, as it holds for $N = 0$. This completes the proof. □

Example 4.

(1) The "classical" two-valued one-dimensional Thue–Morse, Rudin–Shapiro, and paperfolding sequences over $\mathbb{N}$, which have the decimation matrices given in the previous examples for the complete digit set $W = \{0, w\} = \{0, 1\}$, provide examples of 2-automatic sequences on $\mathbb{N}$ for which the limit $\mu_{\mathcal{F}}$ can be obtained via the above theorem, since $\mu_{\mathcal{F}} = \lim_{n\to\infty} 2^{-n} S_W^n \mathcal{F}(x) = \mu$ for all $x \in \mathbb{N}$.

(2) Let $H : \mathbb{N} \to \mathbb{N}$ be defined as $H(x) = 2x$ and the corresponding complete digit (for $\mathbb{N}$) set $W = \{0, 1\}$. The decimation matrices

$$A_0 = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad A_1 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \text{ yielding } S_W = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{pmatrix}$$

define the sequence $\mathcal{F}(x) = (f(x), g(x), h(x))^{\mathrm{T}}$, $x \in \mathbb{N}$, recursively from $\mathcal{F}(0) = (a, a, b)$, $a, b \in \mathbb{C}$ and by setting

$$\mathcal{F}(2x + v) = A_v \mathcal{F}(x)$$

for $x \in \mathbb{N} \setminus \{0\}$ and $v \in W$. The first component f of $\mathcal{F}$ is a 2-automatic sequence with $\ker(f) = \{f, g, h\}$ and decimation matrices A_0 and A_1. Then one computes

$$\mathcal{F}(2) = A_0 A_1 \mathcal{F}_0 = (a, a, a)^{\mathrm{T}} \quad \text{and} \quad \mathcal{F}(26) = A_0 A_1 A_0 A_1 A_1 \mathcal{F}(0) = (b, b, b)^{\mathrm{T}}$$

which gives, after observing that

$$\frac{1}{2^n} S_W^n = \begin{pmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \end{pmatrix} \quad \text{for } n \geqslant 2,$$

that

$$\lim_{n\to\infty} \frac{1}{2^n} S_W^n \mathcal{F}(2) = (a, a, a)^{\mathrm{T}} \quad \text{and} \quad \lim_{n\to\infty} \frac{1}{2^n} S_W^n \mathcal{F}(26) = (b, b, b)^{\mathrm{T}}.$$

Therefore, $\mu_{\mathcal{F}}$ exists if and only if $a = b$, but then all the sequences f, g, and h equal the constant sequence with value a.

Note further, that S_W is a primitive matrix. Therefore, due to Theorem 3.1, the average μ_f of f does not exist, as long as $a \neq b$. Moreover, as a consequence of the same Theorem 3.1, neither do the averages of g and h exist.

(3) Consider the sequence f just mentioned under point (2), with $a \neq b$, and kernel elements f, g, h. Define the sequence $F : \mathbb{N} \to \mathbb{C}$ as

$$f(0), -f(0), f(1), -f(1), f(2), -f(2), f(3), -f(3), \ldots,$$

i.e.,

$$F(x) = \begin{cases} f(x/2) & \text{if } x \text{ is even,} \\ -f((x-1)/2) & \text{if } x \text{ is odd.} \end{cases}$$

Author's personal copy

The sequence F is also $(H, W) = (2, \{0, 1\})$-automatic, and has

$$\ker(F) = \{F, f, g, h, -f, -g, -h\},$$

yielding the corresponding $\mathcal{F} = (F, f, g, h, -f, -g, -h)^{\mathrm{T}}$. As a consequence of what was mentioned under point (2), $\mu_{\mathcal{F}}$ does not exist. However, it is clear from the sequence F itself that

$$\lim_{N\to\infty} \frac{1}{N} \sum_{0\leqslant x<N} F(x)$$

does exist and is equal to zero. This illustrates that it is not necessary for $\mu_{\mathcal{F}}$ to exist in order for μ_f to exist. The conclusion on the existence of μ_F can also be drawn from applying Corollary 3.7. If the rows and columns of the decimation matrices of F and of S_W are labeled by the kernel elements in the order given above, then one computes that for all $n \geqslant 3$,

$$\begin{aligned}\xi(x) &= \frac{1}{2^n} \sum_{a\in\ker(F)} S_W^n(F, a)a(x) \\ &= \frac{2f(x) + g(x) + h(x)}{8} - \frac{2f + g(x) + h(x)}{8} = 0\end{aligned}$$

and thus $\lim_{R\to\infty} m_F(R) = \mu_\xi = 0$.

(4) Item (2) above in this series of examples allows also to demonstrate the necessity of the condition $\alpha_g \geqslant 0$ in Theorem 3.1. To this end consider the sequence $\xi(x) = f(x) - g(x)$, $x \in \mathbb{N}$. Then one easily computes that $\partial_0(\xi)(x) = 0$ for all $x \in \mathbb{N}$, i.e., the constant sequence $\mathbf{0}$. Moreover, it is not hard to demonstrate that $\mathbf{0}$ is the only minimal invariant set in $\ker(\xi) = \{\xi = f - g, \mathbf{0}, g - h, h - f, -\xi, -(g - h), -(h - f)\}$. Due to Theorem 3.3 one immediately concludes that $\mu_\xi = 0$. However, this does not imply that μ_f exists.

Remark. We rephrase the equality of the limits in Theorem 4.1 a little.

Due to Eq. (6), one has that for every $x \in \mathbb{N}$ there exists a $\gamma = \gamma(x) \in \Gamma_{f,W}$ such that $\mathcal{F}(x) = \gamma\mathcal{F}(0)$. Then condition (29) reads as

$$\lim_{n\to\infty} \frac{1}{k^n} S_W^n \mathcal{F}(0) = \lim_{n\to\infty} \frac{1}{k^n} S_W^n \gamma\mathcal{F}(0) = \mu$$

for all $\gamma \in \Gamma_{f,W}$. In case that $S_\infty = \lim_{n\to\infty}(1/k^n)S_W^n$ exists, e.g., if S_W is primitive, then the convergence condition reads as

$$S_\infty\mathcal{F}(0) = S_\infty\gamma\mathcal{F}(0) = \mu$$

for all $\gamma \in \Gamma_{f,W}$. This is the same as saying $(S_\infty - S_\infty\gamma)\mathcal{F}(0) = 0$ for all $\gamma \in \Gamma_{f,W}$. Together with the fact that $A_0\mathcal{F}(0) = \mathcal{F}(0)$, this is equivalent to the fact that $\mathcal{F}(0)$ belongs to an intersection of certain kernels of matrices, namely,

$$\mathcal{F}(0) \in \ker(I - A_0) \bigcap \bigcap_{\gamma\in\Gamma_{f,W}} \ker(S_\infty - S_\infty\gamma).$$

This shows at least on an intuitive level that the existence of $\mu_{\mathcal{F}}$ in nontrivial cases is not "very likely".

4.2. *One-dimensional k-automatic sequences over $\mathbb{Z}$*

We will now consider some sequences over $\mathbb{Z}$. Let $k \in \mathbb{Z}$ be an odd number and $|k| > 1$. Let $H : \mathbb{Z} \to \mathbb{Z}$ be defined as $H(x) = kx$. Then

$$W = \left\{-\frac{|k| - 1}{2}, \ldots, -1, 0, 1, \ldots, \frac{|k| - 1}{2}\right\} \tag{31}$$

is a complete digit set for H.

Theorem 4.2. *Let $f : \mathbb{Z} \to \mathbb{C}$ be (k, W)-automatic with k odd, $|k| > 1$, and W the complete digit given in* (31). *Then*

$$\mu_{\mathcal{F}} = \lim_{R\to\infty} \frac{1}{2R} \sum_{|x|<R} \mathcal{F}(x) = \mu \tag{32}$$

if and only if

$$\lim_{n\to\infty} \frac{1}{2|k|^n} S_W^n(\mathcal{F}(x) + \mathcal{F}(-x)) = \mu \tag{33}$$

for all $x \in \mathbb{Z}$.

Proof.

(a) *If part*: Suppose that (33) holds for all $x \in \mathbb{Z}$. Since $x \mapsto x$ and $x \mapsto -x$ are H-adapted, Corollary 3.5 implies the existence of $\mu_{\mathcal{F}} = \mu$.

(b) *Only if part*: Assume that (32) holds, i.e., that $\mu_{\mathcal{F}} = \mu$. Due to the specific form of the complete digit set, one has that

$$\frac{1}{|k|^n} \sum_{|x|<|k|^n/2} \mathcal{F}(x) = \frac{1}{|k|^n} S_W^n \mathcal{F}(0)$$

for all $n \in \mathbb{N}$. Taking the limits in this equation gives

$$\mu = \lim_{n\to\infty} \frac{1}{2|k|^n} S_W^n(\mathcal{F}(0) + \mathcal{F}(0)).$$

This shows that (33) holds for $x = 0$. Taking into account that

$$\frac{1}{3|k|^n} \sum_{|x|<3k^n/2} \mathcal{F}(x) = \frac{1}{3|k|^n} S_W^n(\mathcal{F}(-1) + \mathcal{F}(0) + \mathcal{F}(1)),$$

one obtains, by taking limits again, that

$$\mu = \lim_{n\to\infty} \frac{1}{2|k|^n} S_W^n(\mathcal{F}(-1) + \mathcal{F}(1)),$$

i.e., that (33) holds for $x = 1$. The rest of the proof follows by induction, similar to the proof of Theorem 4.1. □

Example 5. This example shows that there exists a 3-automatic sequence $f \in \mathbb{C}^{\mathbb{Z}}$ such that the limits

$$\lim_{n\to\infty} \frac{1}{3^n} S_W^n \mathcal{F}(x)$$

are different for different x, while the limit

$$\lim_{n\to\infty} \frac{1}{3^n} S_W^n(\mathcal{F}(x) + \mathcal{F}(-x))$$

is independent from $x \in \mathbb{Z}$. Define f with the four kernel-elements $\{f, g, h, k\}$, based on the following decimation matrices associated with the complete digit set $W = \{-1, 0, 1\}$ for the expanding map $H(x) = 3x$, $x \in \mathbb{Z}$:

$$A_0 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad A_{-1} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}, \quad A_1 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}.$$

This gives

$$S_W = \begin{pmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 2 & 1 & 0 \\ 2 & 0 & 0 & 1 \end{pmatrix}, \quad S_\infty = \lim_{n\to\infty} \frac{1}{3^n} S_W^n = \frac{1}{4} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}.$$

Author's personal copy

Note that $\Gamma_{f,W}$, the set of products of decimation matrices, does not form a group, and hence Theorem 3.8 cannot be applied.

The corresponding $\mathcal{F}$-sequence can be obtained for all $x \in \mathbb{Z}$ from

$$\mathcal{F}(x) = A_{v_0} A_{v_1} \dots A_{v_n} \mathcal{F}(0) \tag{34}$$

with $\mathcal{F}(0) = (a, b, c, d)^{\mathrm{T}}$, where $\sum_{j=0}^{n} 3^j v_j$ is the $(3, W)$-representation of x. For illustration, the corresponding $\mathcal{F}$-sequence in $[-13, 13]$ is given by

$$\begin{array}{l} c\ b\ c\ b\ c\ a\ c\ a\ c\ b\ c\ a\ c\ \mathbf{a}\ d\ b\ d\ a\ d\ b\ d\ b\ d\ a\ d\ a\ d \\ c\ b\ c\ b\ c\ a\ c\ a\ c\ b\ c\ a\ c\ \mathbf{b}\ d\ b\ d\ a\ d\ b\ d\ b\ d\ a\ d\ a\ d \\ b\ c\ b\ c\ b\ c\ a\ c\ a\ c\ b\ c\ b\ \mathbf{c}\ b\ d\ b\ d\ b\ d\ b\ d\ b\ d\ a\ d\ a \\ b\ c\ b\ c\ a\ c\ a\ c\ a\ c\ a\ c\ a\ \mathbf{d}\ a\ d\ a\ d\ b\ d\ b\ d\ a\ d\ a\ d\ a, \end{array}$$

the boldface values being at position 0. This already shows that $S_\infty \mathcal{F}(x)$ differs for different x.

Now, it can be shown from (34), using the fact that if the $(3, W)$-representation of $x = \sum_{j=0}^{n} 3^j v_j$, then the $(3, W)$-representation of $-x = \sum_{j=0}^{n} 3^j (-v_j)$, that for all $x \in \mathbb{Z}$

$$\frac{1}{2} S_\infty (\mathcal{F}(x) + \mathcal{F}(-x)) = \frac{a+b+c+d}{4} \delta,$$

where $\delta = (1, 1, 1, 1)^{\mathrm{T}}$. Hence f and all its $(3, W)$-kernel sequences have average $(a + b + c + d)/4$.

In the two previous theorems the necessary and sufficient conditions are based on the geometric fact that the cylinders (in the above situations intervals of the form $[0, R[$ or $]-R, R[$) almost look like the set W_n, where n is chosen properly. Note that W_n is nothing else than $\kappa^{-n}(\{0\})$, see Eq. (2).

If k is an even positive number, then $W = \{-1, 0, 1, \dots, k-2\}$ provides a complete digit set for $H(x) = kx, x \in \mathbb{Z}$. However, the sets W_n do not properly approximate intervals of the form $]-R, R[$. Let $W = \{0, 1, \dots, k-1\}$, then W is a complete residue for H. However, since $-1 = k(-1) + k - 1$, it follows that -1 has no finite (k, W)-representation, and thus W is not a complete digit set for H. But, since $W_n = \{0, \dots, k^n - 1\}$ one obtains that

$$\kappa^{-n}(\{-1, 0\}) = \{-k^n, -k^n + 1, \dots, -1, 0, 1, \dots, k^n - 1\},$$

which properly approximates an interval $]-R, R[$.

Theorem 4.3. *Let $f : \mathbb{Z} \to \mathbb{C}$ be (k, W)-automatic with k even and positive with complete residue set $W = \{0, \dots, k-1\}$. Then*

$$\mu_{\mathcal{F}} = \lim_{R \to \infty} \frac{1}{2R} \sum_{|x| < R} \mathcal{F}(x) = \mu$$

if and only if

$$\lim_{n \to \infty} \frac{1}{2k^n} S_W^n (\mathcal{F}(x) + \mathcal{F}(-x-1)) = \mu$$

for all $x \in \mathbb{Z}$.

Proof. (a) *If part*: As the maps $x \mapsto x$ and $x \mapsto -x - 1$ are H-adapted, the existence of

$$\lim_{n \to \infty} \frac{1}{2k^n} S_W^n (\mathcal{F}(x) + \mathcal{F}(-x-1)) = \mu$$

for all $x \in \mathbb{Z}$ implies, by Corollary 3.5, the existence of $\mu_{\mathcal{F}}$.

(b) *Only if part*: Now suppose that

$$\lim_{R \to \infty} \frac{1}{2R} \sum_{|x| < R} \mathcal{F}(x) = \mu.$$

By the specific form of the residue set, one has

$$\frac{1}{2k^n}\sum_{|x|<k^n}\mathcal{F}(x)=\frac{1}{2k^n}S_W^n(\mathcal{F}(-1)+\mathcal{F}(0))-\frac{1}{2k^n}\mathcal{F}(-k^n)$$

for all $n\in\mathbb{N}$. Since the limit of the left side exists and since $\mathcal{F}$ is bounded it follows that

$$\lim_{n\to\infty}\frac{1}{2k^n}S_W^n(\mathcal{F}(-1)+\mathcal{F}(0))=\mu.$$

Now the assertion follows by induction, taking into account that for $x>0$ $\kappa^{-n}(\{-x-1,\dots,x\})=\{k^n(-x-1),\dots,0,\dots,k^nx-1\}$. □

4.3. Higher-dimensional (H, W)-automatic sequences with H diagonal

The developments thus far in this section corroborate the expectation that in higher dimensions the existence of a necessary and sufficient condition for the existence of $\mu_{\mathcal{F}}$ depends on "geometric" properties of H and the associated residue set (digit set) W.

The following theorem provides an example of a necessary and sufficient condition for the existence of $\mu_{\mathcal{F}}$ in two dimensions. It is obvious how to generalize this theorem to higher dimensions and other expanding matrices of diagonal form.

Theorem 4.4. *Let* $H=\left(\begin{smallmatrix}3&0\\0&3\end{smallmatrix}\right)$. *Let* $f:\mathbb{Z}^2\to\mathbb{C}$ *be* (H, W)*-automatic with complete digit set* $W=\{-1,0,1\}^2$. *Then*

$$\mu_{\mathcal{F}}=\lim_{\underline{R}\Rightarrow\infty}m_{\mathcal{F}}(\underline{R})=\mu$$

if and only if

$$\lim_{n\to\infty}\frac{1}{9^n}S_W^n\left(\frac{\mathcal{F}((x,y)^{\mathrm{T}})+\mathcal{F}((x,-y)^{\mathrm{T}})+\mathcal{F}((-x,y)^{\mathrm{T}})+\mathcal{F}((-x,-y)^{\mathrm{T}})}{4}\right)=\mu$$

for all $(x, y)\in\mathbb{Z}^2$.

Proof. (a) *If part*: is again based on the H-adaptedness of the maps $(x,y)^{\mathrm{T}}\mapsto(x,y)^{\mathrm{T}}$, $(x,y)^{\mathrm{T}}\mapsto(x,-y)^{\mathrm{T}}$, $(x,y)^{\mathrm{T}}\mapsto(-x,y)^{\mathrm{T}}$, $(x,y)^{\mathrm{T}}\mapsto(-x,-y)^{\mathrm{T}}$, and application of Corollary 3.5.

(b) *Only if part*: Now suppose that

$$\lim_{\underline{R}\Rightarrow\infty}m_{\mathcal{F}}(\underline{R})=\mu.$$

One has, with $\chi_H=9$ in this case, and because of the specific form of the complete digit set, that

$$\frac{1}{9^n}\sum_{|x|,|y|\leqslant(3^n-1)/2}\mathcal{F}((x,y)^{\mathrm{T}})=\frac{1}{9^n}S_W^n\mathcal{F}((0,0)^{\mathrm{T}})$$

for all $n\in\mathbb{N}$. Taking the limit proves the assertion for $(x, y)=(0, 0)$. Now consider

$$\frac{1}{3\times9^n}\sum_{\substack{|x|\leqslant3^n+(3^n-1)/2\\|y|\leqslant(3^n-1)/2}}\mathcal{F}((x,y)^{\mathrm{T}})=\frac{1}{3\times9^n}S_W^n(\mathcal{F}((-1,0)^{\mathrm{T}})+\mathcal{F}((0,0)^{\mathrm{T}})+\mathcal{F}((1,0)^{\mathrm{T}})).$$

Taking the already known limits one gets

$$\lim_{n\to\infty}\frac{1}{2\times9^n}S_W^n(\mathcal{F}((-1,0)^{\mathrm{T}})+\mathcal{F}((1,0)^{\mathrm{T}}))=\mu,$$

which proves the assertion for $(x, y)=(1, 0)$ and $(-1, 0)$. In a similar way, by interchanging the role of the first and second coordinates one obtains the result for $(x, y)=(0, 1)$ and $(0, -1)$.

Author's personal copy

Finally consider

$$\frac{1}{9\times 9^n}\sum_{|x|,|y|\leqslant 3^n+(3^n-1)/2}\mathcal{F}((x,y)^{\mathrm{T}})=\frac{1}{3\times 9^n}S_W^n\left(\sum_{|x|,|y|<2}\mathcal{F}((x,y)^{\mathrm{T}})\right).$$

Taking the already known limits into account, one obtains

$$\lim_{n\to\infty}\frac{1}{9^n}S_W^n\left(\frac{\mathcal{F}((1,1)^{\mathrm{T}})+\mathcal{F}((1,-1)^{\mathrm{T}})+\mathcal{F}((-1,1)^{\mathrm{T}})+\mathcal{F}((-1,-1)^{\mathrm{T}})}{4}\right)=\mu.$$

By induction, the assertion follows for all $(x,y)\in\mathbb{Z}^2$. □

5. Correlations

As already mentioned in the Introduction, the main motivation for considering averages of automatic sequences was the determination of the correlation function of an automatic sequence f, as given by

$$\gamma_f(k)=\lim_{R\to\infty}\frac{1}{\mathrm{vol}(B_R(0))}\sum_{x\in B_R(0)\cap\mathbb{Z}^m}f(x)\overline{f}(x+k)$$

for all $k\in\mathbb{Z}^m$, where $\overline{f}$ denotes the complex conjugate. According to the results in [7], establishing the existence and calculating the correlation function of f boils down to the problem of determining the correlation function at $k=0$, i.e., the averages

$$\mu_{g\overline{h}}=\lim_{R\to\infty}\frac{1}{\mathrm{vol}(PC(R))}\sum_{x\in PC(R)\cap\mathbb{Z}^m}g(x)\overline{h}(x)$$

for all $g,h\in\ker(f)$. Note that, due to the automaticity of f, the sequences $g\overline{h}$ with $g,h\in\ker(f)$ are automatic, too. Having read the previous sections it comes as no surprise that it is difficult, if not impossible, to establish any general criteria for the existence of these limits. In this section we will present two results on the existence of the correlation function of f which are based on Theorems 3.8 and 3.3. Before that, we recapitulate some notions developed in [7], Section 2.

Let $f:\mathbb{Z}^m\to\mathbb{C}$ be (H,W)-automatic and $\mathcal{F}:\mathbb{Z}^m\to\mathbb{C}^{\ker(f)}$ the usually associated vector sequence defined by $\mathcal{F}(x)=(g(x))_{g\in\ker(f)}$. Consider the sequence $\mathcal{F}\otimes\mathcal{F}:\mathbb{Z}^m\to\mathbb{C}^{\ker(f)\times\ker(f)}$, where $\mathcal{F}\otimes\mathcal{F}(x)$ is defined as the Kronecker (tensor) product

$$(\mathcal{F}\otimes\mathcal{F})(x)=\mathcal{F}(x)\otimes\mathcal{F}(x).$$

Recall that, if $A=(a_{i,j})$ and B are matrices (vectors) with complex entries, then $A\otimes B$ is the matrix (vector) obtained by replacing each element $a_{i,j}$ in A by the matrix $a_{i,j}\times B$. I.e., $\mathcal{F}\otimes\mathcal{F}$ is the sequence of vectors with components labeled (g,h) and the sequence corresponding to (g,h) is given as $gh(x)=g(x)h(x)$, $g,h\in\ker(f)$. Note further that, although it might happen that two components of $\mathcal{F}\otimes\mathcal{F}$ may be the same sequence, these components will be treated as different sequences.

For $w\in W$ and $g,h\in\ker(f)$ one computes

$$\partial_w(gh)(x)=(gh)(Hx+w)=g(Hx+w)h(Hx+w)=\partial_w(g)(x)\partial_w(h)(x),$$

which is again a component of $\mathcal{F}\otimes\mathcal{F}$. Thus we may speak of the decimation matrices $\mathbf{A}_w$, $w\in W$ of $\mathcal{F}\otimes\mathcal{F}$. Due to Lemma 2.2 in [7], these matrices are given by

$$\mathbf{A}_w=A_w\otimes A_w. \tag{35}$$

In other words $(\mathcal{F}\otimes\mathcal{F})(w+Hx)=\mathbf{A}_w(\mathcal{F}\otimes\mathcal{F})(x)=(A_w\otimes A_w)(\mathcal{F}(x)\otimes\mathcal{F}(x))$. Yet another interpretation of the sequence $\mathcal{F}\otimes\mathcal{F}$ is helpful. Consider the set $\mathbf{K}=\ker(f)\times\ker(f)$ as a subset of all maps (sequences) from $\mathbb{Z}^m$ to $\mathbb{C}^2$, i.e., $(g,h)(x)=(g(x),h(x))$ for $(g,h)\in K$. Defining the decimations as $\partial_w(g,h)=(\partial_w(g),\partial_w(h))$ the above

considerations show that $\mathbf{A}_w$ are the decimation matrices of the decimation invariant set $\mathbf{K}$. With these preparations we are able to state the first result on the existence of $\mu_{g\overline{h}}$ for all $g, h \in \ker(f)$ and, by the results in [7], of the correlation function of f.

We remind the reader that $\Gamma_{f,W}$ denotes the semigroup of all products of decimation matrices associated with an (H, W)-automatic sequence f.

Theorem 5.1. *Let $f : \mathbb{Z}^m \to \mathbb{C}$ be (H, W)-automatic and let W be a complete digit set. If $\Gamma_{f,W}$ is a group, then $\mu_{g\overline{h}}$ exists for all $g, h \in \ker(f)$.*

Proof. We shall show that for every pair $(g, h) \in \ker(f) \times \ker(f)$ the sequence $g\overline{h}$ defined as $(g\overline{h})(x) = g(x)\overline{h}(x)$ satisfies the condition of Theorem 3.8, i.e., $\Gamma_{g\overline{h},W}$ is a group.

As $\Gamma_{f,W}$ is a group, it follows that the decimation matrices A_w, $w \in W$, are permutation matrices, i.e., they have a single 1 in each row and each column. As a consequence, also the matrices $\mathbf{A}_w$ are permutation matrices. In the interpretation of $\mathcal{F} \otimes \mathcal{F}$ as $\mathbf{K} = \ker(f) \times \ker(f)$ this is equivalent to $\partial_w : \mathbf{K} \to \mathbf{K}$ is bijective. Moreover, if $\ker((g, h)) \subset \mathbf{K}$ denotes the kernel of the pair (g, h), then the restrictions $\partial_w : \ker((g, h)) \to \ker((g, h))$ are bijective, too.

Now consider the map $\rho : \mathbb{C} \times \mathbb{C} \to \mathbb{C}$ defined as $\rho(u, v) = u\overline{v}$. For $g, h \in \ker(f)$ one has $(g\overline{h})(x) = \rho(g(x), h(x))$, or $g\overline{h} = \rho(g, h)$ for short. Now note that one has the obvious identity

$$\partial_w(g\overline{h}) = \rho(\partial_w(g), \partial_w(h)),$$

or, equivalently,

$$(\partial \circ \rho)(g, h) = (\rho \circ \partial_w)(g, h) \tag{36}$$

for all $g, h \in \ker(f)$. In particular, $\ker(g\overline{h}) = \rho(\ker((g, h)))$. In order to prove that $\Gamma_{g\overline{h},W}$ is a group it is sufficient to prove that the maps $\partial_w : \ker(g\overline{h}) \to \ker(g\overline{h})$, $w \in W$, are bijective. Since $\ker(g\overline{h})$ is a finite set, ∂_w is bijective if $\partial_w(\ker(g\overline{h})) = \ker(g\overline{h})$. Since $\rho : \ker((g, h)) \to \ker(g\overline{h})$ is surjective, one has $\partial_w(\ker(g\overline{h})) = \partial_w(\rho(\ker((g, h))))$, using Eq. (36) this gives $\partial_w(\rho(\ker((g, h)))) = \rho(\partial_w(\ker((g, h))))$. Since $\partial_w : \ker((g, h)) \to \ker((g, h))$ is bijective, one has $\rho(\partial_w(\ker((g, h)))) = \rho(\ker((g, h)))$ which is, due to the surjectivity of ρ, equal to $\ker(g\overline{h})$. Thus, we have shown that $\partial_w(\ker(g, \overline{h})) = \ker(g\overline{h})$ which proves the bijectivity of $\partial_w : \ker(g\overline{h}) \to \ker(g\overline{h})$. Therefore, $\Gamma_{g\overline{h},W}$ is a group and since W is a complete digit set, Theorem 3.8 shows that $\mu_{g\overline{h}}$ exists. □

The second result on the existence of a correlation function is related to Theorem 3.3.

Theorem 5.2. *Let $f : \mathbb{Z}^m \to \mathbb{C}$ be an (H, W)-automatic sequence. If every minimal invariant component K_i, $i = 1, \ldots, s$, of $\ker(f)$ corresponds to a constant sequence, i.e., $K_i = \{g_i\}$ and $g_i \in \ker(f)$, $i = 1, \ldots, s$, then $\mu_{g\overline{h}}$ exists for every $g, h \in \ker(f)$.*

Proof. For every $g \in \ker(f)$ there exists $\Delta_g = \partial_{w_n} \circ \cdots \circ \partial_{w_0}$ such that $\Delta_g(g) \in K_1 \cup \cdots \cup K_s$, which means that $\Delta_g(g)$ is a constant sequence. For $g, h \in \ker(f)$ consider the sequence $g\overline{h}$. Then one gets $\Delta_g(g\overline{h}) = \Delta_g(g)\overline{\Delta_g(h)}$ and $\Delta_g(g)$ is a constant sequence. Since $\Delta_g(h) \in \ker(f)$, there exists $\Delta_{\Delta_g(h)}$ such that $\Delta_{\Delta_g(h)}(\Delta_g(h))$ is a constant sequence. This implies

$$\Delta_{\Delta_g(h)}(\Delta_g(g\overline{h})) = \Delta_{\Delta_g(h)}(\Delta_g(g)\overline{\Delta_g(h)})\Delta_g(g)\Delta_{\Delta_g(h)}(\overline{\Delta_g(h)})$$

is also a constant sequence. This also shows that the minimal invariant subsets of $\ker(g\overline{h})$ correspond to constant sequences. Application of Theorem 3.3 completes the proof. □

The existence of the correlation functions of higher-dimensional analogs of the Thue–Morse, the paperfolding, and the Rudin–Shapiro sequences has been established in [7]. The full determination of the correlation function of the last two sequences was done in [8]. The results there were based on particularities of these sequences. As illustrated, the conditions presented in this paper also apply to them, but they can also be used for other automatic sequences described

in terms of decimation matrices determined by the underlying automaton, and based on proper (H, W)-number systems in $\mathbb{Z}^m$. But this in itself does not guarantee that the conditions that have been discussed here are easily satisfied.

Acknowledgments

This research was supported by the Concerted Action Project GOA-Mefisto of the Flemish Community, by the FWO (Fund for Scientific Research Flanders) project G.0121.03, and by the Belgian Program on Interuniversity Attraction Poles of the Belgian Prime Minister's Office for Science, Technology and Culture (IAP V-22).

References

[1] J.-P. Allouche, M. Mendès France, Automatic sequences, in: F. Axel, D. Gratias (Eds.), Beyond Quasicrystals, Springer, Berlin, 1995 (Les Editions de Physique, Les Ulis).
[2] J.-P. Allouche, J. Shallit, Automatic Sequences. Theory, Applications, Generalizations, Cambridge University Press, Cambridge, 2003.
[3] M. Baake, A guide to mathematical quasicrystals, in: J.-B. Suck, M. Schreiber, P. Haussler (Eds.), Quasicrystals, Springer, Berlin, 2002, pp. 17–48, See also (http://front.math.ucdavis.edu/math-ph/9901014).
[4] M. Baake, R.V. Moody (Eds.), Directions in Mathematical Quasicrystals, American Mathematical Society, Providence, RI, 2000.
[5] A. Barbé, F. von Haeseler, Symmetries of decimation invariant sequences and digit sets, Theoret. Comput. Sci. 289 (2002) 105–136.
[6] A. Barbé, F. von Haeseler, Automatic sets and Delone sets, J. Phys. A 37 (2004) 4017–4038.
[7] A. Barbé, F. von Haeseler, Correlation functions for higher-dimensional automatic sequences, J. Phys. A 37 (2004) 10879–10898.
[8] A. Barbé, F. von Haeseler, Correlation and spectral properties of higher-dimensional paperfolding and Rudin–Shapiro sequences, J. Phys. A 38 (2005) 2599–2622.
[9] A. Barbé, F. von Haeseler, Binary number systems for $\mathbb{Z}^k$, J. Number Theory 117 (2006) 14–30.
[10] A. Barbé, F. von Haeseler, Correlation and spectral properties of multidimensional Thue–Morse sequences, Internat. J. Bifurc. Chaos Appl. Sci. Eng., to appear.
[11] A. Barbé, G. Skordev, Decimation invariant sequences and their automaticity, Theoret. Comput. Sci. 259 (2000) 379–403.
[12] A. Berman, R.J. Plemmons, Nonnegative Matrices in the Mathematical Sciences, Academic Press, New York, 1979.
[13] F. von Haeseler, Automatic Sequences, W. De Gruyter, Berlin, 2003.
[14] M. Höffe, M. Baake, Surprises in diffuse scattering, Z. Kristallogr. 215 (2000) 441–444.
[15] I. Katai, Generalized number systems in Euclidean space, Math. Comput. Modelling 38 (7–9) (2003) 883–892.
[16] M. Lothaire, Algebraic Combinatorics on Words, Cambridge University Press, Cambridge, 2002.
[17] D.W. Matula, Basic digit sets for radix representation, J. Assoc. Comput. Mach. 29 (1982) 1131–1143.
[18] N. Priebe Frank, Substitution sequences in $\mathbb{Z}^d$ with non-simple Lebesgue component in the spectrum, Ergodic Theory Dynamical Systems 23 (2003) 519–532.
[19] N. Priebe Frank, Multidimensional constant-length substitution sequences, Topology Appl. 152 (1–2) (2005) 44–69.
[20] M. Queffélec, Substitution Dynamical Systems—Spectral Analysis, Lecture Notes in Mathematics, Vol. 1294, Springer, Berlin, 1987.
[21] M. Senechal, Quasicrystals and Geometry, Cambridge University Press, Cambridge, 1995.

Author's personal copy