Journal de Théorie des Nombres
de Bordeaux **17** (2005), 347–357

Complexity of Hartman sequences

par Christian STEINEDER et Reinhard WINKLER

Résumé. Soit $T : x \mapsto x + g$ une translation ergodique sur un groupe abélien compact C et soit M une partie de C dont la frontière est de measure de Haar nulle. La suite binaire infinie $\mathbf{a} : \mathbb{Z} \mapsto \{0, 1\}$ définie par $\mathbf{a}(k) = 1$ si $T^k(0_C) \in M$ et $\mathbf{a}(k) = 0$ sinon, est dite de Hartman. Notons $P_{\mathbf{a}}(n)$ le nombre de mots binaires de longueur n qui apparaissent dans la suite $\mathbf{a}$ vue comme un mot bi-infini. Cet article étudie la vitesse de croissance de $P_{\mathbf{a}}(n)$. Celle-ci est toujours sous-exponentielle et ce résultat est optimal. Dans le cas où T est une translation ergodique $x \mapsto x+\alpha$ ($\alpha = (\alpha_1, \ldots, \alpha_s)$) sur $\mathbb{T}^s$ et M un parallélotope rectangle pour lequel la longueur du j-ème coté ρ_j n'est pas dans $\alpha_j\mathbb{Z} + \mathbb{Z}$ pour tout $j = 1, \ldots, s$, on obtient $\lim_n P_{\mathbf{a}}(n)/n^s = 2^s \prod_{j=1}^{s} \rho_j^{s-1}$.

Abstract. Let $T : x \mapsto x + g$ be an ergodic translation on the compact group C and $M \subseteq C$ a continuity set, i.e. a subset with topological boundary of Haar measure 0. An infinite binary sequence $\mathbf{a} : \mathbb{Z} \mapsto \{0, 1\}$ defined by $\mathbf{a}(k) = 1$ if $T^k(0_C) \in M$ and $\mathbf{a}(k) = 0$ otherwise, is called a Hartman sequence. This paper studies the growth rate of $P_{\mathbf{a}}(n)$, where $P_{\mathbf{a}}(n)$ denotes the number of binary words of length $n \in \mathbb{N}$ occurring in $\mathbf{a}$. The growth rate is always subexponential and this result is optimal. If T is an ergodic translation $x \mapsto x + \alpha$ ($\alpha = (\alpha_1, \ldots, \alpha_s)$) on $\mathbb{T}^s$ and M is a box with side lengths ρ_j not equal $\alpha_j\mathbb{Z} + \mathbb{Z}$ for all $j = 1, \ldots, s$, we show that $\lim_n P_{\mathbf{a}}(n)/n^s = 2^s \prod_{j=1}^{s} \rho_j^{s-1}$.

Acknowledgement. *The authors thank the Austrian Science Foundation FWF for supporting the research for this paper through projects no. S8312 and no. S8302.*

1. Motivation and Notation

The notion of a Hartman sequence has recently been introduced and studied (cf. [5], [10], [12]). It can be seen as a generalisation of the notion of a Sturmian sequence. Sturmian sequences (and their close relatives, the Beatty sequences) are very interesting objects, as well from the combinatorial, the number theoretic and the dynamical point of view. Let us sketch two approaches.

Consider sequences $\mathbf{a} = (a_k)$ of two symbols, say 0 and 1, where k runs through the set $\mathbb{Z}$ of all integers or $\mathbb{N}$ of all positive integers. Such sequences are also called two resp. one sided infinite binary words. Let $P_{\mathbf{a}}(n)$ be the number of different binary words of length n occurring in $\mathbf{a}$. The mapping $n \mapsto P_{\mathbf{a}}(n)$ is called the complexity function of $\mathbf{a}$. It is easily seen that the complexity function is bounded if and only if $\mathbf{a}$ is (in the one sided case: eventually) periodic. Among all aperiodic sequences Sturmian sequences have minimal complexity, namely $P(n) = n + 1$. This is the combinatorial approach to characterise Sturmian sequences, which has been introduced in [6] and [7].

A different characterisation uses the symbolic coding of irrational rotations. If $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ denotes the circle group (one dimensional torus) and M is a segment of $\mathbb{T}$ of angle $2\pi\alpha$ with irrational $\alpha \in \mathbb{T}$, then the definition $a_k = 1$ if and only if $k\alpha \in M$ defines a Sturmian sequence (cf. for instance [2]).

For understanding the definition of a Hartman sequence the second approach is more appropriate. Replace $\mathbb{T}$ by more general compact abelian groups C with normalised Haar measure $\mu = \mu_C$ and replace α by any ergodic group translation. This means that we are interested in the transformation $T : C \mapsto C$, $T : x \mapsto x + g$, where g is a generating element of C, i.e. the orbit $\{kg \ : \ k \in \mathbb{Z}\}$ is dense in C. Thus C is required to be monothetic. C can also be interpreted as a group compactification of $\mathbb{Z}$ since C is the closure of the image of $\mathbb{Z}$ under the dense homomorphic embedding $\iota : \mathbb{Z} \to C$, $\iota(k) = kg$. (Note that for group compactifications one usually does not require ι to be injective. But to avoid trivial case distinctions we will demand that $\iota(\mathbb{Z})$ is infinite.) This approach is particularly appropriate for Theorems 1 and 2. For the classical theory of ergodic group translations we refer for instance to [11].

A set $M \subset C$ is called a (μ_C-)continuity set if $\mu_C(\partial M) = 0$ holds for its topological boundary ∂M. For a continuity set M consider the induced binary sequence $\mathbf{a} = (a_k)_{k=-\infty}^{\infty}$ defined by $a_k = 1$ if $T^k(0_C) \in M$ and $a_k = 0$ otherwise. Such sequences are called Hartman sequences. The set $H \subset \mathbb{Z}$ defined by $k \in H$ if and only if $a_k = 1$ is accordingly called a Hartman set. Thus, by definition, a Hartman set H is the preimage $H = \iota^{-1}(M)$ of a continuity set $M \subseteq C$ where (C, ι) is a group compactification of $\mathbb{Z}$ and we can write $\mathbf{a} = \mathbf{1}_H$. Note that for $C = \mathbb{T}^s$ and $g = (\alpha_1, \ldots, \alpha_s)$, where the family $\{1, \alpha_1, \ldots, \alpha_s\}$ is linearly independent over $\mathbb{Z}$, Hartman sequences are binary coding sequences of Kronecker sequences.

As a consequence of uniform distribution of ergodic group translations (for the theory of uniform distribution of sequences we refer to [9]), Hartman sequences share some nice properties with Sturmian sequences (cf. [5], [10]). In particular each finite subword occurs with a uniform asymptotic density.

More precisely: If the measure of a Hartman set $H = \iota^{-1}(M)$ is defined by $\mu(H) := \mu_C(M)$, then

$$\mu(H) = \mathrm{dens}(H) = \lim_{n\to\infty} \frac{\mathrm{card}(H \cap \{k, k+1, \ldots, k+n-1\})}{n}$$

holds uniformly in $k \in \mathbb{Z}$, cf. [5]. In particular, for a given Hartman set H, this value neither depends on (C, ι) nor on M. In terms of Hartman sequences: If, for $\mathbf{a} = (a_k) = \mathbf{1}_H$, $A_k(n)$ denotes the number of occurrences of 1's in the block $a_k a_{k+1} \ldots a_{k+n-1}$ of length n, there exists a bound $c_M(n) = o(1)$, $n \to \infty$, such that

$$\left| \frac{A_k(n)}{n} - \mu_C(M) \right| \leq c_M(n)$$

for all $k \in \mathbb{N}$. This has been used in [12] to develop a new approach to identify the underlying dynamical system from its symbolic coding.

In this paper we start to investigate the complexity of Hartman sequences. Corresponding to the zero entropy of the underlying dynamical system, the growth rate of the complexity function of any Hartman sequence is subexponential. This upper bound turns out to be best possible. In particular, the complexity of a Hartman sequence might be much higher than, for instance, interval coding sequences for which the complexity essentially is linear (cf. for instance [1] and [3]). In some sense this fact is due to the more general choice of M rather than to the more general choice of the compact group C. Nevertheless, for the case of an M with a very simple geometric structure, namely a box in a finite dimensional compactification $C = \mathbb{T}^s$ the complexity grows polynomially of maximal order s. A more systematic investigation of the role of the geometric properties of M and the further development of the arguments used here is to be the object of future research. So the results in this paper are the following:

- For any Hartman sequence $\mathbf{a}$ and $\lambda > 1$ we have $P_{\mathbf{a}}(n) = o(\lambda^n)$ for $n \to \infty$ **(Theorem 1)**.
- For any sequence P_n of subexponential growth rate and any compactification (C, ι) there is a Hartman sequence $\mathbf{a}$ coming from (C, ι) with $P_n = o(P_{\mathbf{a}}(n))$. This is even true, if $P_{\mathbf{a}}(n)$ counts only binary words which occur in $\mathbf{a}$ with strictly positive density **(Theorem 2)**.
- Assume that M is a box of side lengths ρ_j, $j = 1, \ldots, s$, in $\mathbb{T}^s$. Then the complexity $P_{\mathbf{a}}(n)$ of the induced Hartman sequence $\mathbf{a}$ has the asymptotic growth rate cn^s if $\rho_j \notin \alpha_j\mathbb{Z} + \mathbb{Z}$ for all $j = 1, \ldots, s$. The multiplicative constant c is given by $c = 2^s \prod_{j=1}^{s} \rho_j^{s-1}$ **(Theorem 3)**.

2. A universal upper bound for the complexity of Hartman sequences

Corresponding to the fact that ergodic group translations have entropy 0, one conjectures that the complexity of a Hartman sequence is subexponential. Note that the topological entropy of shift spaces and complexity are very closely related. Nevertheless we cannot apply immediately pertinent theorems in textbooks as [4] or [11] to obtain an upper bound for the complexity of a Hartman sequence in terms of the entropy of the underlying group rotation. Therefore we give a direct proof that the above conjecture is true.

Theorem 1. *For any compactification (C, ι) of $\mathbb{Z}$ and any continuity set $M \subseteq C$ the complexity $P_{\mathbf{a}}(n)$ of the corresponding Hartman sequence $\mathbf{a} = \mathbf{1}_H$ with $H = \iota^{-1}(M)$ satisfies*

$$\lim_{n\to\infty} \frac{\log P_{\mathbf{a}}(n)}{n} = 0.$$

Proof. By Theorem 4 in [12] we may presume that there is a metric d for the topology on C. Let $g = \iota(1) \in C$ denote the generating element of the compactification.

We write M' for the complement $C \setminus M$ and M_δ for the set of all $x \in C$ with $d(x, y) < \delta$ for some $y \in M$. Fix $\varepsilon > 0$. Using the regularity of the Haar measure μ_C and the μ_C-continuity of M, we obtain $\mu_C(R) < \varepsilon$ for $R = (M_{\delta_1} \setminus M) \cup (M'_{\delta_1} \setminus M')$ whenever $\delta_1 > 0$ is sufficiently small. By a standard argument we may assume that R is a continuity set. At least one of the sets M and M', say M, has nonempty interior. This means that there is some open ball B with center x and positive diameter $\delta < \delta_1/2$ with $B \subset M$. For the sake of simpler notation we assume $x = 0$.

Let W_l denote the set of all binary words $a_0 \dots a_{l-1}$ of length l with $a_k = 1$ whenever $kg + B \subseteq M$ and $a_k = 0$ whenever $kg + B \subseteq M'$.

By compactness of C there is some $N_0 \in \mathbb{N}$ such that

$$C \subseteq \bigcup_{n=0}^{N_0-1} (-ng + B),$$

showing that for every $y \in C$ there is some $n \in \{0, 1, \dots, N_0 - 1\}$ with $y + ng \in B$.

Thus any word w of length $N_0 + l$ occurring in $\mathbf{a}$ lies in some of the sets $W_{N_0+l,i}$, $0 \le i \le N_0 - 1$, consisting of all words

$$b_0 b_1 \dots b_{i-1} a_0 \dots a_{l-1} b_i \dots b_{N_0-1}$$

with $a_0 a_1 \dots a_{l-1} \in W_l$ and $b_0 b_1 \dots b_{N_0-1} \in \{0, 1\}^n$. Since $|W_{N_0+l,i}| = 2^{N_0}|W_l|$, this shows $P_{\mathbf{a}}(N_0 + l) \le N_0 2^{N_0} |W_l|$.

Note that each translate $y + B$ is totally contained either in M or in M' whenever $y \notin R$. Thus, by the uniform distribution of $(ng)_n$ in C, the subset $T \subseteq \mathbb{Z}$ of all $k \in \mathbb{Z}$ such that $y = kg \notin R$ has density $\mu_C(C \setminus R) > 1 - \varepsilon$.

It follows that $|W_l| \leq 2^{2\varepsilon l}$, hence $P_{\mathbf{a}}(N_0+l) \leq N_0 2^{N_0+2\varepsilon l}$ for l sufficiently large. This yields

$$\log P_{\mathbf{a}}(N_0 + l) \leq \log N_0 + (N_0 + 2\varepsilon l) \log 2$$

and, for $n = N_0 + l$

$$\limsup_{n\to\infty} \frac{\log P_{\mathbf{a}}(n)}{n} \leq \limsup_{l\to\infty} \frac{\log N_0 + (N_0 + 2\varepsilon l) \log 2}{N_0 + l} \leq 2\varepsilon \log 2.$$

Since $\varepsilon > 0$ can be chosen arbitrarily small this proves the theorem. □

3. A Hartman sequence of arbitrarily subexponential complexity

We are going to show that the bound deduced in Theorem 1 is best possible.

Let (C, ι) be any group compactification of $\mathbb{Z}$ and $\phi : \mathbb{N} \mapsto \mathbb{N}$. Suppose $\phi(n) = \varepsilon_n n \leq n$ with $\lim_{n\to\infty} \varepsilon_n = 0$. We have to show that there exists a continuity set $M \subset C$ such that the Hartman sequence $\mathbf{a} := \mathbf{1}_H$ with $H = \iota^{-1}(M)$ fulfills $P_{\mathbf{a}}(n) \geq 2^{\phi(n)}$.

By Theorem 4 in [12] it suffices to prove the assertion for metrisable C and by Theorem 8.3 in [8] there is an invariant metric d for the topology on C. For $c \in C$ we write $\|c\| = d(c, 0)$. For each $n \in \mathbb{N}$ choose a subset $H^{(n)}$ of $\{0, \ldots, n-1\}$ of cardinality $A_n \geq \varepsilon_n n$ and containing 0 such that the diameter d_n of $\iota(H^{(n)})$ is minimal. We claim that $\lim_{n\to\infty} d_n = 0$.

Otherwise we had a sequence $n_1 < n_2 < \ldots$ and a $\delta > 0$ such that $d_{n_k} \geq 2\delta$ for all k. There is some $r \in (0, \delta)$ such that the open ball B with center $0 \in C$ and radius r is a continuity set. By construction, the lower density of the set of all n with $\iota(n) \in B$ is at most ε_{n_k} for all k. By uniform distribution of $\iota(n)$, the lower density is a density and coincides with the Haar measure, hence $\mu(B) \leq \lim_{k\to\infty} \varepsilon_{n_k} = 0$. This contradicts the fact that nonempty open sets have positive measure.

Let now $H_n(0), H_n(1), \ldots, H_n(2^{A_n} - 1)$ be an enumeration of all subsets of $H^{(n)}$. Define recursively $m_n(0) = 0$ and $m_n(i + 1)$ to be the minimal integer $> m_n(i) + n$ such that $\|\iota(m_n(i + 1))\| < d_n$. We put $H_n = \bigcup_{i=0}^{2^{A_n}-1} (m_n(i) + H_n(i))$. Obviously H_n is a finite set of nonnegative integers bounded by, say $h_n \in \mathbb{N}$. Observe furthermore that by construction $\|\iota(h)\| < 2d_n$ for all $h \in H_n$. Define, again recursively, $l_0 = 1$ and l_{n+1} to be the minimal integer $> l_n + h_n$ such that $\|\iota(l_{n+1})\| < d_n$. For the union $H = \bigcup_{n=0}^{\infty} (H_n + l_n)$ this implies $\lim_{n\to\infty, n\in H} \iota(n) = 0$. Thus $M = \iota(H)$ is a countable closed subset of C with the only accumulation point 0, hence a continuity set of measure 0 and $H = \iota^{-1}(M)$ is a Hartman set.

In the corresponding Hartman sequence each H_n induces at least 2^{A_n} different binary words of length n. Thus the complexity function $P(n)$ is bounded below by

$$P(n) \geq 2^{A_n} \geq 2^{\varepsilon_n n} = 2^{\phi(n)}.$$

This construction generates a zero set M. Hence each word in $\mathbf{a}$ containing the letter 1 has asymptotic density 0. It would be nice to get a positive frequency for many words. Let therefore $M = \{0, m_1, m_2, \ldots\}$ be an enumeration of M. There are $\delta_n > 0$ with $\delta_n \to 0$ such that balls B_n, $n \in \mathbb{N}$, with center m_n and radius δ_n are pairwise disjoint continuity sets. Replace M by the union of all B_n, which is again a continuity set. This shows:

Theorem 2. *Let (C, ι) be any group compactification of $\mathbb{Z}$. Assume $\phi(n) \leq n$ and $\phi(n) = o(n)$ for $n \to \infty$. Then there exists a continuity set $M \subset C$ such that its Hartman sequence $\mathbf{a} := \mathbf{1}_{\iota^{-1}(M)}$ fulfills $P_{\mathbf{a}}(n) \geq 2^{\phi(n)}$. Furthermore M can be chosen in such a way that for each $n \in \mathbb{N}$ at least $2^{\phi(n)}$ words of length n occur in $\mathbf{a}$ with strictly positive density.*

4. The case of Kronecker sequences and Boxes

We finally restrict our attention to the special case that $C = \mathbb{T}^s$ is a finite dimensional compactification with generating element $g = (\alpha_1, \ldots, \alpha_s)$ modulo 1, i.e. $\iota : k \mapsto kg = k(\alpha_1, \ldots, \alpha_s)$, where the family $\{1, \alpha_1, \ldots, \alpha_s\}$ is linearly independent over $\mathbb{Z}$ (such $(kg)_k$ are also called Kronecker sequences), and M a box in $\mathbb{T}^s$. To be more precise we use the following notational convention.

As usual, $\{r\} = r - [r]$ and $[r] = \max\{k \in \mathbb{Z} : \ k \leq r\}$ denote the fractional respectively integer part of $r \in \mathbb{R}$. Thus $\mathbb{T}^s = (\mathbb{R}/\mathbb{Z})^s = \kappa(\mathbb{R}^s)$ is considered to be the image of the additive group $\mathbb{R}^s$ under the mapping $\kappa = \kappa_s : (x_1, \ldots, x_s) \mapsto (\{x_1\}, \ldots, \{x_s\})$. Although $\mathbb{T}$ has no order structure it is useful to think about intervals in $\mathbb{T}$ as images of intervals in $\mathbb{R}$ under κ_1, boxes in $\mathbb{T}^s$ as images of boxes in $\mathbb{R}^s$ etc. To avoid too cumbersome notation we therefore write, for instance, $\prod_{j=1}^s [-\rho_j/2, \rho_j/2)$, $\rho_j \in (0,1)$ also for the set $M = \kappa_s(\prod_{j=1}^s [-\rho_j/2, \rho_j/2))$. It is natural to call a set $M = \prod_{j=1}^s [m_j, m_j + \rho_j) \subseteq \mathbb{T}^s$ an s-dimensional box in $\mathbb{T}^s$ with side lengths ρ_j, $j = 1, \ldots, s$. We are especially interested in Hartman sequences $\mathbf{a} = \mathbf{1}_H$, $H = \iota^{-1}(M)$, for this kind of M and call such $\mathbf{a}$ Bohr sequences.

Let us fix a box M of side lengths ρ_j, $j = 1, \ldots, s$, and assume that no ρ_j is in $\alpha_j\mathbb{Z} + \mathbb{Z}$. We are going to determine the asymptotic behaviour of $P_{\mathbf{b}}(n)$ for the Bohr sequence $\mathbf{b} = \mathbf{1}_H$, $H = \iota^{-1}(M)$.

We will use the following notation: For a word $w = a_0 \ldots a_{n-1} \in \{0,1\}^n$ we introduce the set

$$A_w := \{x \in \mathbb{T}^s : (x + ig \in M \Leftrightarrow a_i = 1) \text{ for } i = 0, \ldots, n-1\}$$

and write $w = w(x)$ for $x \in A_w$. Note that, provided $A_w \neq \emptyset$, A_w has inner points. Because of the density of the set $\{ng : n \in \mathbb{N}\}$, the continuity of T and the special form of M this implies

$$P_{\mathbf{b}}(n) = |\{w \in \{0,1\}^n : A_w \neq \emptyset\}|.$$

To compute the number of all nonempty A_w we first consider a half open cube $C_0 := c_0 + [-\sigma/2, \sigma/2)^s \subseteq \mathbb{T}^s$ with center c_0 and side length $\sigma < \rho_j$ for all $j = 1, \ldots, s$. We are going to estimate the local complexity function $P(C_0, n) := |W|$ for $W = W(C_0) := \{w \in \{0,1\}^n : A_w \cap C_0 \neq \emptyset\}$. Note that for k cubes $C_1, \ldots, C_k$ in $\mathbb{T}^s$ with disjoint closures we have

$$P_{\mathbf{a}}(n) \geq \sum_{j=1}^{k} P(C_j, n)$$

for sufficiently large n. As above, $A_w \cap C_0 \neq \emptyset$ implies $\mu_C(A_w \cap C_0) > 0$. So $P(C_0, n)$ is the number of different words $w = b_i \ldots b_{i+n-1}$ of length n in $\mathbf{b}$ with $ig \in C_0$. Define

$$M_0 := \prod_{j=1}^{s} [-\frac{\rho_j}{2} + \frac{\sigma}{2}, \frac{\rho_j}{2} - \frac{\sigma}{2}), \quad M_1 := \prod_{j=1}^{s} [-\frac{\rho_j}{2} - \frac{\sigma}{2}, \frac{\rho_j}{2} + \frac{\sigma}{2}), \quad \Gamma := M_1 \backslash M_0,$$

and furthermore, for each $j = 1, \ldots, s$,

$$Q_1^{(j)} := \{x = (x_1, \ldots, x_s) \in M_1 : x_j < -\frac{\rho_j}{2} + \frac{\sigma}{2} \text{ or } x_j \geq \frac{\rho_j}{2} - \frac{\sigma}{2}\},$$
$$Q_0^{(j)} := Q_1^{(j)} \setminus \bigcup_{j' \neq j} Q_1^{(j')}.$$

Observe that the sets $Q_0^{(j)}$ (in contrast to the $Q_1^{(j)}$) are pairwise disjoint. For $w = a_0 \ldots a_{n-1}$ in W note that

$$(c_0 + ig \in M_0 \Rightarrow a_i = 1) \text{ and } (c_0 + ig \notin M_1 \Rightarrow a_i = 0).$$

This shows that for $w = a_0 \ldots a_{n-1}$ and $w' = a'_0 \ldots a'_{n-1}$ in W the letters a_i and a'_i can differ only if $c_0 + ig \in \Gamma$. Since $\Gamma = \bigcup_{j=1}^{s} Q_1^{(j)}$, we define, for $j = 1, \ldots, s$ and $l = 0, 1$,

$$I_l^{(j)} := \{i \in \{0, \ldots, n-1\} : c_0 + ig \in Q_l^{(j)}\},$$
$$I_l := \bigcup_{j=1}^{s} I_l^{(j)}.$$

Due to the special geometric situation (C_0 and M are boxes), for $x = (x_1, \ldots, x_s) \in C_0$, $w = w(x) = (a_i(x))_{0 \leq i < n} \in W$, $j \in \{1, \ldots, s\}$, the tuple $(a_i(x))_{i \in I_0^{(j)}}$ depends only on x_j, namely in the following way. Let $X_j = [y_0, y_0 + \sigma)$ be the interval for the j-th coordinate of points in C_0. Then for each $i \in I_0^{(j)}$ there is one point y_i (namely either $\rho_j/2 - ig$ or

$-\rho_j/2 - ig$) such that y_i splits the interval X_j into two subintervals $X_j^{(0)}$ and $X_j^{(1)}$ such that $a_i(x) = 0$ for $x_j \in X_j^{(0)}$ and $a_i(x) = 1$ for $x_j \in X_j^{(1)}$. Since $\rho_j \notin \alpha_j\mathbb{Z} + \mathbb{Z}$, all y_i, $i \in I_0^{(j)}$, are distinct. As a consequence the mapping $x_j \mapsto (a_i(x))_{i\in I_0^{(j)}}$ takes at least $|I_0^{(j)} + 1|$ different values, hence

$$A_j = |\{(a_i(x))_{i\in I_0^{(j)}} \,:\, x \in C_0\}| \geq |I_0^{(j)} + 1|.$$

Since the sets $I_0^{(j)}$, $j = 1, \ldots, s$, are pairwise disjoint and all coordinates j can be treated independently, we conclude

$$P(C_0, n) = |W| \geq \prod_{j=1}^{s} (|I_0^{(j)}| + 1).$$

For $\varepsilon > 0$ we know by uniform distribution of the sequence $(ng)_n$ that

$$|I_0^{(j)}| \geq \mu(Q_0^{(j)})n - \varepsilon n$$

for n sufficiently large. Since $\mu(Q_0^{(j)}) = 2\prod_{j=1,j\neq i}^{s}(\rho_j - \sigma)\sigma$, $j = 1, \ldots, s$, we get

$$|W| \geq n^s \prod_{i=1}^{s} \Big(2 \prod_{j=1,j\neq i}^{s} ((\rho_j - \sigma)\sigma - \varepsilon\Big)$$

for n sufficiently large. Thus we obtain

$$\liminf_{n\to\infty} \frac{P(C_0, n)}{n^s} \geq 2^s \prod_{i=1}^{s} \Big(\prod_{j=1,j\neq i}^{s} (\rho_j - \sigma)\sigma - \varepsilon\Big)$$

for all $\varepsilon > 0$ and therefore

$$\liminf_{n\to\infty} \frac{P(C_0, n)}{n^s} \geq 2^s \prod_{i=1}^{s} \prod_{j=1,j\neq i}^{s} (\rho_j - \sigma)\sigma.$$

As a consequence of uniform distribution we know that $d(x, x') > \delta$ implies $w(x) \neq w(x')$ if the words are sufficiently long. Thus $W(C_0)$ and $W(C_0')$ are disjoint whenever two cubes C_0 and C_0' are separated by a strictly positive distance δ. Fix now $k \in \mathbb{N}$ and consider the disjoint cubes $C_1, \ldots, C_{k^s}$ with centers $c_i = (m_i/k)$, $m_i \in \{0, \ldots, k-1\}$ and side length $\sigma = 1/k - \delta$, $0 < \delta < \frac{1}{k}$. We get

$$\liminf_{n\to\infty} \frac{P_{\mathbf{b}}(n)}{n^s} \geq \sum_{i=1}^{k^s} \liminf_{n\to\infty} \frac{P(C_i, n)}{n^s} \geq k^s 2^s (\frac{1}{k} - \delta)^s \prod_{i=1}^{s} \prod_{j=1,j\neq i}^{s} (\rho_j - \frac{1}{k} + \delta).$$

Since this holds for all $\delta > 0$ we can consider the limit $\delta \to 0$ to get

$$\liminf_{n\to\infty} \frac{P_{\mathbf{b}}(n)}{n^s} \geq k^s 2^s \frac{1}{k^s} \prod_{i=1}^{s} \prod_{j=1,j\neq i}^{s} (\rho_j - \frac{1}{k}) = 2^s \prod_{j=1}^{s} (\rho_j - \frac{1}{k})^{s-1}.$$

For $k \to \infty$ this finally shows the lower bound

$$\liminf_{n\to\infty} \frac{P_{\mathbf{b}}(n)}{n^s} \geq 2^s \prod_{j=1}^{s} \rho_j^{s-1}.$$

To obtain an upper bound for the complexity we consider instead of A_j as defined above the numbers

$$B_j = |\{(a_i(x))_{i\in I_1^{(j)}} \,:\, x \in C_0\}| \leq |I_1^{(j)} + 1|.$$

Note that the sets $I_1^{(j)}$, $j = 1, \ldots, s$, are (in contrast to the sets $I_0^{(j)}$) not disjoint. This implies that $a_i(x)$ possibly depends on more than one component of x. Comparison with the argument for the lower bound shows that the relevant mapping $x \mapsto (a_i(x))_{i\in I_1^{(j)}}$, $x \in C_0$, can only take one additional value, namely the zero word $a_i(x) = 0$ for all $i \in I_1^{(j)}$. Thus arguments similar (in fact even simpler) to those above show that $|B_j| \leq |I_1^{(j)}| + 2$ and finally

$$\limsup_{n\to\infty} \frac{P_{\mathbf{b}}(n)}{n^s} \leq 2^s \prod_{j=1}^{s} \rho_j^{s-1}.$$

Since the same argument applies if M is not centered at 0 we have proved:

Theorem 3. *Consider an ergodic translation $T : x \mapsto x + g$ on $\mathbb{T}^s$ with $g = (\alpha_1, \ldots, \alpha_s)$. Assume $\rho_j \in (0,1) \setminus (\alpha_j\mathbb{Z} + \mathbb{Z})$ for all $j = 1, \ldots, s$. For $m_j \in [0,1)$, $j = 1, \ldots, s$, let $M = \prod_{j=1}^{s}[m_j, m_j + \rho_j)$ denote an s-dimensional box of side lengths ρ_j, and $\mathbf{b}$ the corresponding Bohr sequence. Then the complexity function of $\mathbf{b}$ satisfies*

$$\lim_{n\to\infty} \frac{P_{\mathbf{b}}(n)}{n^s} = 2^s \prod_{j=1}^{s} \rho_j^{s-1}.$$

We add the following remarks.

Complexity and volume versus surface: Let $V(M)$ denote the volume of a box M in $\mathbb{T}^s$ and $v_j(M) = \prod_{i=1, i\neq j}^{s} \rho_i$ the $(s-1)$-dimensional measures (surfaces) of the facets of M. Then we can express the formula in Theorem 3 in terms of $V(M)$ as well as in terms of the $v_j(M)$:

$$2^s \prod_{j=1}^{s} \rho_j^{s-1} = 2^s V(M)^{s-1} = 2^s \prod_{j=1}^{s} v_j(M)$$

Consider first $M' := M_0 \cup M_1$, where the M_i are disjoint translates of M. The same arguments as in the proof of Theorem 3 show that M' induces a

Hartman sequence $\mathbf{a}'$ of complexity

$$\lim_{n\to\infty} \frac{P_{\mathbf{a}'}(n)}{n^s} = 2^s \prod_{j=1}^{s} (2v_j(M)).$$

Comparison with the value $2^s \prod_{j=1}^{s} v_j(M)$ for M indicates that the complexity is related to the surface rather than to the volume.

On the other hand we can consider an automorphism A of $\mathbb{T}^s$ (i.e. $A \in \mathrm{SL}(s,\mathbb{Z})$). The topological generator $A(g)$ and the parallelepiped $A(M)$ induce the same Hartman sequence as g and M. Here the volume $V(M) = V(A(M))$ is invariant while the product of the surface measures may change.

The interplay between the geometry of more general sets M and the complexity of the corresponding Hartman sequences might be an interesting topic of more systematic future investigations.

Dropping linear independence: If the assumption $\rho_j \notin \alpha_j\mathbb{Z} + \mathbb{Z}$ fails then a modification of the proof of Theorem 3 based on a careful investigation of the resulting cancellation effects yields a corresponding formula.

Complexity determines dimension: Given $P_{\mathbf{b}}(n)$ and the information that M is a box (of some unknown dimension s and unknown side lengths ρ_j), Theorem 3 tells us how s can be determined.

References

[1] P. Alessandri, V. Berthé, *Three Distance Theorems and Combinatorics on Words*. Enseig. Math. **44** (1998), 103–132.

[2] P. Arnoux, V. Berthé, S. Ferenci, S. Ito, C. Mauduit, M. Mori, J. Peyrière, A. Siegel, J.- I. Tamura, Z.- Y. Wen, *Substitutions in Dynamics, Arithmetics and Combinatorics*. Lecture Notes in Mathematics **1794**, Springer Verlag Berlin, 2002.

[3] V. Berthé, *Sequences of low complexity: Automatic and Sturmian sequences*. Topics in Symbolic Dynamics and Applications. Lond. Math. Soc, Lecture Notes **279**, 1–28, Cambridge University Press, 2000.

[4] M. Denker, Ch. Grillenberger, K. Sigmund, *Ergodic Theory on Compact Spaces*. Lecture Notes in Mathematics **527**, Springer, Heidelberg, 1976.

[5] S. Frisch, M. Pašteka, R. F. Tichy, R. Winkler, *Finitely additive measures on groups and rings*. Rend. Circolo Mat. di Palermo **48** Series II (1999), 323–340.

[6] G. A. Hedlund, M. Morse, *Symbolic Dynamics I*. Amer. J. Math. **60** (1938), 815–866.

[7] G. A. Hedlund, M. Morse, *Symbolic Dynamics II*. Amer. J. Math. **62** (1940), 1–42.

[8] E. Hewitt, K. A. Ross, *Abstract Harmonic Analysis I*. Springer, Berlin—Göttingen—Heidelberg, 1963.

[9] L. Kuipers, H. Niederreiter, *Uniform distribution of sequences*. Wiley, New York, 1974.

[10] J. Schmeling, E. Szabó, R. Winkler, *Hartman and Beatty bisequences*. Algebraic Number Theory and Diophantine Analysis, 405–421, Walter de Gruyter, Berlin, 2000.

[11] P. Walters, *An Introduction to Ergodic Theory*. Grad. Texts in Math. Springer Verlag New York, 2000.

[12] R. Winkler, *Ergodic Group Rotations, Hartman Sets and Kronecker Sequences*. Monatsh. Math. **135** (2002), 333–343.

Christian Steineder
Technische Universität Wien
Institut für Diskrete Mathematik und Geometrie
Wiedner Hauptstraße 8-10
1040 Vienne, Autriche
E-mail : christian.steineder@tuwien.ac.at

Reinhard Winkler
Technische Universität Wien
Institut für Diskrete Mathematik und Geometrie
Wiedner Hauptstraße 8-10
1040 Vienne, Autriche
E-mail : reinhard.winkler@tuwien.ac.at