C O L L O Q U I U M   M A T H E M A T I C U M

VOL. 80 1999 NO. 2

READING ALONG ARITHMETIC PROGRESSIONS

BY

T. DOWNAROWICZ (WROCŁAW)

Abstract. Given a 0-1 sequence x in which both letters occur with density $1/2$, do there exist arbitrarily long arithmetic progressions along which x reads $010101\ldots$? We answer the above negatively by showing that a certain regular triadic Toeplitz sequence does not have this property. On the other hand, we prove that if x is a generalized binary Morse sequence then each block can be read in x along some arithmetic progression.

This note contains the complete solution of a problem in symbolic dynamics raised by H. Furstenberg during an informal "problem session" held in Hebrew University, Jerusalem in December 1997. One of the variations of the famous van der Waerden's Theorem asserts that every subset of $\mathbb{N}$ of positive density contains arbitrarily long arithmetic progressions (Szemeredi, 1975). Perhaps, if we partition $\mathbb{N}$ into two subsets of equal density, there always exist arithmetic progressions whose elements belong alternately to these sets. The seemingly natural class to look for counterexamples are Morse sequences (for example, it is known that no Morse sequence contains an infinite periodic subsequence). Thus the question can be asked separately for this class. Surprisingly, it turns out that in general the conjecture is false and that for Morse sequences an even stronger statement holds; we prove that if x is a generalized binary Morse sequence then any binary block can be read in x along some arithmetic progression.

This note is fully selfcontained. The definitions of Toeplitz and Morse sequences are included in our constructions. Nevertheless, we direct the interested reader to [F] and [W] for more detailed references.

The triadic Toeplitz sequence. Below we provide an example of a 0-1 sequence with 0 and 1 appearing with density $1/2$ where the longest subsequence with alternate values along an arithmetic progression has 4 elements.

Let $x = (x(i))_{i\in\mathbb{N}}$ be the (one-sided triadic Toeplitz) sequence defined inductively as follows:

1991 *Mathematics Subject Classification*: Primary 54H20, 58F03.

Key words and phrases: Toeplitz sequence, Morse sequence, Szemeredi Theorem.

Step 1.

$$x(i)=\begin{cases}0 & \text{whenever } i=1 \bmod 3,\\ 1 & \text{whenever } i=2 \bmod 3.\end{cases}$$

Positions divisible by 3 will be filled in next steps. After n steps we have filled all positions but those divisible by 3^n, for which we let:

Step $n+1$.

$$x(i)=\begin{cases}0 & \text{whenever } i=1\cdot 3^n \bmod 3^{n+1},\\ 1 & \text{whenever } i=2\cdot 3^n \bmod 3^{n+1}.\end{cases}$$

Eventually all positive positions are filled. For each position i we can determine the inductive step in which this position has been filled:

$$\mathrm{st}(i)=\min\{n : i \text{ is not divisible by } 3^n\}.$$

It is clear that the densities of zeros and ones in x are both $1/2$.

Suppose we have found an arithmetic progression with step j along which x reads 010, i.e., $x(i)=0$, $x(i+j)=1$, and $x(i+2j)=0$, for some $i,j\in\mathbb{N}$. If $\mathrm{st}(i)=\mathrm{st}(i+2j)=n$ then

$$i=i+2j \bmod 3^n(=1\cdot 3^{n-1}),$$

hence $2j$ is a multiple of 3^n and so is j. This implies $x(i+j)=x(i)=0$, which eliminates the case.

Assume that $\mathrm{st}(i)=n<\mathrm{st}(i+2j)$ (for $\mathrm{st}(i)>\mathrm{st}(i+2j)$ we apply a symmetric argument). Now we have

$$i=1\cdot 3^{n-1} \bmod 3^n \quad \text{and} \quad i+2j=0 \bmod 3^n,$$

which implies that $i-2j=i+4j=2\cdot 3^{n-1} \bmod 3^n$, hence $x(i-2j)=x(i+4j)=1$. We have shown that 010 cannot be extended to anything like 0?010 or 010?0, hence the longest alternating sequence along an arithmetic progression is either 0101 or 1010.

Morse sequences. For convenience we change the alphabet to be $\{-1,1\}$ rather than $\{0,1\}$. Also, we change the indexation of one-sided sequences so that they start at coordinate 0 rather than 1. Recall the definition of a generalized binary Morse sequence:

Definition. Let $A=(A(0)A(1)\dots A(a-1))$ and $B=(B(0)B(1)\dots \dots B(b-1))$ be two blocks. By $A\times B$ we denote the block of length ab defined by $A\times B(i)=A(s)B(t)$, where $i=s+at$. Let $(B_n)_{n\in\mathbb{N}}$ be an infinite sequence of blocks satisfying

$$B_n(0)=1 \tag{1}$$

and

$$B_n\neq 1,1,\dots,1 \tag{2}$$

for any $n \in \mathbb{N}$. Let x be the coordinatewise limit of the blocks

$$A_n = B_1 \times \ldots \times B_n$$

(convergence is granted by the condition (1)). If A is not a periodic sequence then we call it a (one-sided) *generalized Morse sequence* determined by (B_n). (The non-periodicity assumption is added because for example if $B_n = (1, -1, 1)$ for each n, then $A = (1, -1, 1, -1, 1, -1, \ldots)$.)

We shall use the following fact concerning generalized Morse sequences:

LEMMA 1. *Let x be a generalized Morse sequence. There are no infinite arithmetic progressions along which x is constant.*

Proof. Suppose that, on the contrary, $(x(s+it))_{i\in\mathbb{N}}$ is constant for some $s, t \in \mathbb{N}$. For fixed $k \in \mathbb{N}$ define $B_1' = B_1 \times \ldots \times B_k$ and $B_n' = B_{n+k-1}$ for $n > 1$. Then the sequence of blocks (B_n') determines the same Morse sequence as (B_n). Thus, we can assume B_1 to be as long as we want, say $b_1 > t$.

Consider the blocks $x[(n-1)b_1, nb_1 - 1]$. What we see for each n is either B_1 or its negation. Since x has a constant subsequence along a progression with step tb_1, two such blocks with indices n and $n+t$ have the same entry at at least one place. This implies that these blocks are either both B_1 or both $-B_1$. We have shown that x is periodic with period tb_1, which contradicts the definition of Morse sequences. ∎

REMARK. The above argument (unlike the next one) works for Morse sequences taking values in any finite abelian group.

We are in a position to prove the statement announced in the introduction:

THEOREM 1. *Let x be a generalized binary Morse sequence and let C be an arbitrary binary block. Then there exists an arithmetic progression along which x reads C.*

Proof. We will prove the statement inductively on the length c of the block C.

STEP 1. There are only two blocks of length 1 and both appear in any non-periodic sequence.

STEP $c+1$. Suppose the theorem holds for all blocks of length c. Let $C\sigma$ be an arbitrary block of length $c+1$, where C has length c and $\sigma \in \{-1, 1\}$. Let x be a generalized Morse sequence. By the inductive assumption, C can be read in x along some arithmetic progression with some step t, say

$$x(s+it) = C(i) \quad \text{for } 0 \le i \le c-1.$$

If $x(s+ct)=\sigma$ then we are finished. Thus we only need to consider the case

$$x(s+ct) = -\sigma.$$

Let n be sufficiently large, so that A_n coincides with x on $[0, s+ct]$.

Consider the Morse sequence y defined by the blocks $B_{n+1}, B_{n+2}, \ldots$ (It is easy to see that y is a Morse sequence, i.e., it is not periodic.) Again by the inductive assumption, the block $1, 1, \ldots, 1$ of length c can be read along an arithmetic progression in y. Continuing this progression to the right we will eventually arrive at a position where -1 occurs in y (see Lemma 1). Thus we have an arithmetic progression of length $c+1$ where $1, 1, \ldots, 1, -1$ occurs in y, say

$$\begin{aligned} y(u+iw) &= 1 \quad \text{for } 0 \le i \le c-1, \text{ and} \\ y(u+cw) &= -1. \end{aligned}$$

Let m be sufficiently large, so that $B_{n,m} = B_{n+1} \times \ldots \times B_{n+m}$ coincides with y on $[0, u+cw]$. Note that the positions $s+it+a(u+iw)$ (a denotes the length of A_n) form an arithmetic progression (with step $t+aw$). Compute the values of $A_{n+m} = A_n \times B_{n,m}$ along this progression:

$$\begin{aligned} A_{n+m}(s+it+a(u+iw)) &= A_n(s+it)B_{n,m}(u+iw) \\ &= \begin{cases} C(i) \cdot 1 & \text{for } 0 \le i \le c-1, \\ (-\sigma) \cdot (-1) & \text{for } i = c. \end{cases} \end{aligned}$$

We have found an arithmetic progression where $C\sigma$ can be read in A_{n+m} and hence also in x. ∎

REFERENCES

[F] H. Furstenberg, *Recurrence in Ergodic Theory and Combinatorial Number Theory*, Princeton Univ. Press, Princeton, N.J., 1981.

[W] S. Williams, *Toeplitz minimal flows which are not uniquely ergodic*, Z. Wahrsch. Verw. Gebiete 67 (1984), 95–107.

Institute of Mathematics
Wrocław University of Technology
Wybrzeże Wyspiańskiego 27
50-370 Wrocław, Poland
E-mail: downar@im.pwr.wroc.pl

Received 10 December 1998;
revised 22 January 1999