APERIODIC ORDER – LECTURE 2 SUMMARY

1. Sturmian sequences (see Chapter 6 in [2])

We will consider sequences in a finite alphabet \( \mathcal{A} \). Denote by \( \mathcal{A}^n \) the set of “words” of length \( n \) in the alphabet \( \mathcal{A} \). Given an infinite sequence \( u \in \mathcal{A}^\mathbb{N} \), let \( \mathcal{L}_n(u) \) be the set of words of length \( n \) which occur in \( u \). The cardinality
\[
p_u(n) = \#\mathcal{L}_n(u)
\]
is called the complexity of a sequence \( u \). Recall the following

**Proposition 1.1.** If there exists \( n \) such that \( p_u(n) \leq n \), then \( u \) is eventually periodic (in which case \( p_u(n) \leq C \)).

Thus, \( p_u(n) = n + 1 \) is the minimal possible complexity of a non-periodic sequence.

**Definition 1.2.** A sequence \( u \in \{0, 1\}^\mathbb{N} \) is called Sturmian if \( p_u(n) = n + 1 \) for all \( n \).

**Lemma 1.3.** Every Sturmian sequence \( u \) is recurrent, which means that every word that occurs in \( u \) appears infinitely often.

**Example 1.4.** (Exercise) The Fibonacci sequence obtained from the substitution \( \zeta : 0 \rightarrow 01, \ 1 \rightarrow 0 \) by
\[
u = \lim_{n \to \infty} \zeta^n(0) = 01001010\ldots
\]
is Sturmian.

1.1. Rotation sequences. Let \( \alpha, \beta \in (0, 1), \ \alpha \notin \mathbb{Q} \). Consider the sequence
\[
u_n = \lfloor (n+1)\alpha + \beta \rfloor - \lfloor n\alpha + \beta \rfloor
\]
or
\[
u_n = \lceil (n+1)\alpha + \beta \rceil - \lceil n\alpha + \beta \rceil
\]
It is called the rotation sequence with rotation angle \( \alpha \) and initial point \( \beta \). Why?

In fact, consider the rotation \( R_\alpha : x \mapsto x + \alpha \pmod{1} \) on \( T = \mathbb{R}/\mathbb{Z} \sim [0, 1) \). Let \( I_0 = [0, 1 - \alpha) \) and \( I_1 = [1 - \alpha, 1) \), and consider the function \( \nu(x) = 0 \), if \( x \in I_0 \) and \( \nu(x) = 1 \) if \( x \in I_1 \).
Lemma 1.5. The sequence $u_n$ in (1.1) can be obtained as

$$ u_n = \nu(R^*_\alpha(\beta)), $$

where $R^*_\alpha$ denotes the $n$-th iterate of $R_\alpha$. The sequence $u_n$ in (1.2) can be obtained in a similar way, if we take $I_0 = (0, 1 - \alpha]$ and $I_1 = (1 - \alpha, 1]$.

Proposition 1.6. Every rotation sequence is Sturmian (the converse is also true, but we will not prove it).

Proof idea. Consider the partition of $[0, 1)$ by the points of the “reverse orbit” of 0 of length $n + 1$, namely, $0, -\alpha, -2\alpha, \ldots, -n\alpha$ (mod 1). We claim that the coding sequence of length $n$ is exactly determined by where we are with respect to this partition. Therefore, the number of “words” of length $n$ is $n + 1$. □

Remark 1.7. The Fibonacci sequence is the rotation sequence with $\alpha = \tau^{-2} = 1 - \tau^{-1} = 0.381\ldots$, where $\tau = \frac{1 + \sqrt{5}}{2}$ is the golden ratio, and $\beta = \tau^{-2}$ (alternatively, we start at $\beta = 0$, but drop the first symbol 0).

1.2. Digression: sequence of intermediate growth. Let $c$ be a sequence of maximal complexity, i.e. $p_c(n) = 2^n$ for all $n$, say, the so-called “Champernowne word”:

$$ c = 1.10.11.100.101.110.111\ldots $$

Then let

$$ q_k = c_{k-1} c_{k-2} 00 c_{k-3} 0000 \ldots c_0 0^{2k-2} $$

Note that $|q_k| = k^2$. Then let $u = q_1 q_2 q_3 \ldots$ be the concatenation of all words $q_k$. Observe that $p_u(k^2) \geq 2^k$. Cassaigne [1] proved that

$$ 2^{\lceil \sqrt{n} \rceil} \leq p_u(n) \leq n^{2\lceil \sqrt{n} \rceil} $$

for all $n$, hence $\log p_u(n) \sim \sqrt{n}$. The sequence $u$ is not recurrent, but it is possible to modify this construction to get $u$ recurrent and even uniformly recurrent, see [1].

2. Substitutions

Let $A$ be a finite alphabet of cardinality $m \geq 2$, usually $A = \{0, \ldots, m-1\}$ or $A = \{1, \ldots, m\}$. Denote by $A^+$ the set of all nonempty “words” using the “letters” from $A$, and let $A^* = A^+ \cup \{\varepsilon\}$, where $\varepsilon$ is the empty word. We will also write $A^\mathbb{N}$ to denote the set of (one-sided) sequences: $u = u_0 u_1 u_2 \ldots \in A^\mathbb{N}$ whenever $u_j \in A$. (Our convention is that $0 \in \mathbb{N}$.)
**Definition 2.1.** A substitution on \( A \) is a map \( \zeta: A \to A^+ \). It extends to a map on \( A^\mathbb{N} \) by concatenation:

\[
\zeta(WW') = \zeta(W)\zeta(W'), \quad \zeta(u_0u_1u_2\ldots) = \zeta(u_0)\zeta(u_1)\zeta(u_2)\ldots
\]

More formally, we can define a substitution \( \zeta \) as a morphism of a free semigroup with the set of generators \( A \).

A fixed point for \( \zeta \) is a sequence \( u \in A^\mathbb{N} \) such that \( \zeta(u) = u \). A periodic point is \( u \) such that \( \zeta^k(u) = u \) for some \( k \). (Note that this is periodicity with respect to the substitution, and not the usual periodicity of a sequence!)

Suppose that \( \zeta(a) \) starts with \( a \) and \( |\zeta(a)| \geq 2 \). Then we have a well-defined limit

\[
u = \lim_{n \to \infty} \zeta^n(a) \in A^\mathbb{N},
\]

which is a fixed point for \( \zeta \).

**Exercise.** Suppose that \( \lim_{n \to \infty} |\zeta^n(a)| = \infty \). Show that there is a periodic point for \( \zeta \).

**Example 2.2.**

(i) Fibonacci substitution: \( \zeta_1(0) = 01, \, \zeta_1(1) = 0 \). The fixed point is \( u = 01001010 \ldots \)

(ii) Morse (or Thue-Morse) substitution: \( \zeta_2(0) = 01, \, \zeta_2(1) = 10 \). The fixed point is \( u = 01101001 \ldots \) It is easy to see by induction that if we denote \( U_n = \zeta^n_2(0) \), then \( U_{n+1} = U_n\overline{U}_n \), where the “overline bar” is the operation of exchanging \( 0 \leftrightarrow 1 \), e.g. \( \overline{01} = 10 \).

(iii) Cantor substitution: \( \zeta_3(0) = 000, \, \zeta_3(1) = 101 \). Here we have two fixed points for \( \zeta \): a boring one \( 000 \ldots = 0^\infty \) and a more interesting one

\[
u = 101000101000000000101000101 \ldots
\]

It is called a “Cantor substitution” (although G. Cantor had nothing to do with it), because if we view 1’s as little line segments, its construction resembles the construction of the middle-third Cantor set (to be rigorous, one should “renormalize” appropriately).

**Definition 2.3.** For a substitution \( \zeta \) on an alphabet of \( m \) symbols the substitution matrix is defined by

\[
S_{\zeta}(i,j) = \text{number of } i \text{'s in } \zeta(j).
\]

In our examples above we have

\[
S_{\zeta_1} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}, \quad S_{\zeta_2} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \quad S_{\zeta_3} = \begin{pmatrix} 3 & 1 \\ 0 & 2 \end{pmatrix}.
\]
Definition 2.4. The canonical homomorphism is the map $\ell : A^* \to \mathbb{N}^m$ defined by $\ell(W) = (\ell_i(W))_{i=1}^m$, where $\ell_i(W)$ is the number of symbols $i$ is the word $W$.

We consider $\ell(W)$ as a column vector; it is sometimes called the “population vector” of $W$. Note that it is indeed a homomorphism from the free semigroup $A^*$ to the free Abelian semigroup $\mathbb{N}^m$, since we have $\ell(WW') = \ell(W) + \ell(W')$. Basically, we replace the non-commuting generators — elements of $A$, by commuting generators. Note that we can express the substitution matrix in terms of its columns as follows:

$S_\zeta = [\ell(\zeta(1)), \ldots, \ell(\zeta(m))]$.

We obtain the following important identities:

$\ell(S_\zeta W) = S_\zeta(\ell(W))$, \hspace{1cm} $S_\zeta^k = (S_\zeta)^k$.

The transformation $W \mapsto S_\zeta W$ is often called the Abelianization of the substitution.

Definition 2.5. A substitution $\zeta$ is called primitive if there exists $k \in \mathbb{N}$ such that for all $a \in A$, the word $\zeta^k(a)$ contains all symbols $b \in A$. In view of the above, this is equivalent to the condition that $S_\zeta^k$ has all entries strictly positive. (Such matrices are also called primitive.)

Note that in the example above, the substitutions $\zeta_1$ and $\zeta_2$ are primitive, whereas $\zeta_3$ is not primitive.

References
