Toeplitz sequences, their construction and characterisation

Álvaro Bustos Gajardo
Universidad de Chile

STIC-AmSud – Entropy and Probabilistic Analysis in Algorithms
9 October, 2025

Toeplitz sequences

Definition. *(Toeplitz sequence)*

A Toeplitz sequence over the alphabet $\mathscr{A}$ is an infinte word $\eta \in \mathscr{A}^{\mathbb{N}}$ (or $\mathscr{A}^{\mathbb{Z}}$) such that, for every n, there exists $p_n > 0$ such that $\eta_n = \eta_{n+kp_n}$, for all values of k.

Example: Period-doubling (or Feigenbaum) sequence:

$$0\ 1\ 0\ 0\ 0\ 1\ 0\ 1\ 0\ 1\ 0\ 0\ 0\ 1\ 0\ 0\ 0\ 1\ 0\ 0\ 0\ 1\ 0\ 1\ 0\ 1\ 0\ \ldots$$

Black symbols appear with period 2, **red** symbols appear with period 4, **blue** symbols appear with period 8, and so on.

Periodic sequences are always Toeplitz, but kind of “boring”.

Combinatorics of a Toeplitz sequence

Definition. *(Skeleton)*

The *p*-periodic part of a sequence x is the set of all indices n where we observe the symbols repeating with period p:

$$\mathrm{Per}_p(x) := \{n : x_n = x_{n+kp} \text{ for all values of } k\}.$$

A sequence η is Toeplitz iff $\bigcup_{p\geqslant 1} \mathrm{Per}_p(\eta)$ contains all numbers.

Definition. *(Period sequence)*

An essential period for a Toeplitz sequence η is a p such that $\mathrm{Per}_p(\eta) \neq \mathrm{Per}_q(\eta)$ for all $q < p$. To each such η, we associate the strictly increasing sequence of essential periods $(p_j)_{j\geqslant 1}$, which satisfies $p_j \mid p_{j+1}$ for all j; we also define $q_1 = p_1, q_{j+1} = \frac{p_{j+1}}{p_j}$.

Combinatorics of a Toeplitz sequence

Let $?$ be a symbol not in $\mathscr{A}$. For each number p, define:

$$\mathrm{skel}_p(x)_j = \begin{cases} x_j & \text{if } j \in \mathrm{Per}_p(x), \\ ? & \text{otherwise;} \end{cases} \qquad \mathrm{skel}_{p_n}(x) \xrightarrow{n\to\infty} x \text{ iff } x \text{ is Toeplitz.}$$

We are sequentially "building" x from an "empty" sequence $S_1(x) = ??????????\ldots$: each $\mathrm{skel}_{p_n}(x)$ has some unfilled "gaps" that appear p_n-periodically, and we fill some (hopefully not all) empty cosets of $p_n\mathbb{Z}$ in $\mathrm{skel}_{p_n}(x)$ to create $\mathrm{skel}_{p_{n+1}}(x)$:

a	?	a	?	a	?	a	?	a	?	a	?	a	?	a	?	a	?	a	?	a	?	a	$\ldots$	$p_1 = 2$
a	b	a	b	a	?	a	b	a	b	a	?	a	b	a	b	a	?	a	b	a	b	a	?	$\ldots$	$p_2 = 6$
a	b	a	b	a	c	a	b	a	b	a	?	a	b	a	b	a	c	a	b	a	b	a	?	$\ldots$	$p_3 = 12$
a	b	a	b	a	c	a	b	a	b	a	a	a	b	a	b	a	c	a	b	a	b	a	?	$\ldots$	$p_4 = 24$
a	b	a	b	a	c	a	b	a	b	a	a	a	b	a	b	a	c	a	b	a	b	a	b	$\ldots$	$p_5 = 120$

Toeplitz shifts and dynamics

Definition. *(Toeplitz shift space)*

A Toeplitz shift X_η is the orbit closure $\overline{\mathrm{Orb}_\sigma(\eta)}$ of a Toeplitz sequence η under the shift action $\mathbb{N} \overset{\sigma}{\curvearrowright} \mathcal{A}^{\mathbb{N}}$ (or $\mathbb{Z} \overset{\sigma}{\curvearrowright} \mathcal{A}^{\mathbb{Z}}$) by translations.

Toeplitz shifts are always minimal: every point has a dense orbit. Thus, $|X_\eta| = \infty$ unless η is a periodic point. Beyond this, Toeplitz shifts can vary a lot:

- they can be substitutive and thus have very low word complexity,
- but otherwise they may have any arbitrary entropy (and thus high complexity),
- they can have wildly different spectrums,
- they can have any configuration of invariant measures, and so on…

The maximal equicontinuous factor

A Toeplitz sequence η (and hence, any $x \in X_\eta$) can be thought of as an infinite concatenation of finitely many "chunks" of length p_n (n-supertiles):

$$010001010100010001000101\ldots \begin{array}{l} \nearrow\ 010001010100010001000101\ldots \\ \rightarrow\ 010001010100010001000101\ldots \\ \searrow\ 010001010100010001000101\ldots \end{array}$$

To any $x \in X_\eta$, we associate a list of numbers $\pi(x) = (s_n)_{n \geqslant 1}$, with $0 \leqslant s_n < p_n$, where s_n indicates how "misplaced" the n-supertile closest to the origin is. As every $(n+1)$-supertile is made out of n-supertiles, we must have $s_{n+1} \equiv s_n \pmod{p_n}$.

$$\pi(101010001000100\ldots) = (1, 1, 5, 5, \ldots).$$

If $\pi(x) = \pi(y)$, then $\mathrm{skel}_{p_n}(x) = \mathrm{skel}_{p_n}(y)$ for all n.

The maximal equicontinuous factor

Definition. *(Odometer)*

The $(p_n)_{n\geqslant 1}$-adic group is the topological group:

$$\mathbb{Z}_{(p_n)_{n\geqslant 1}} := \left\{ (s_n)_{n\geqslant 1} \in \prod_{j=1}^{\infty} \mathbb{Z}/p_j\mathbb{Z} : \text{for all } n,\ s_n \equiv s_{n+1} \pmod{p_n} \right\},$$

with the prodiscrete topology. The action $\mathbb{Z} \overset{\omega}{\curvearrowright} \mathbb{Z}_{(p_n)_{n\geqslant 1}}$ given by $(s_n)_{n\geqslant 1} \mapsto (s_n+1)_{n\geqslant 1}$ produces an equicontinuous dynamical system, called a $(p_n)_{n\geqslant 1}$-adic odometer.

The map $\pi : X_\eta \twoheadrightarrow \mathbb{Z}_{(p_n)_{n\geqslant 1}}$ (**MEF**) is a topological factor map, as it is surjective and:

$$\pi(\sigma(x)) = \pi(x) + 1 = \omega(\pi(x)),$$

where we identify $n \in \mathbb{N}$ with the sequence $(n, n, n, n, \ldots)$. Also, $\pi(\eta) = 0$.

Dynamical characterisation of Toeplitz words

The MEF π allows us to recognise which elements of X_η are Toeplitz:

Theorem. *(Williams)*

A point $x \in X_\eta$ is Toeplitz iff $\pi^{-1}[\pi(x)] = \{x\}$.

In terms of the "coset-filling" procedure: if x is not Toeplitz, we will end up with one or more ? symbols ("holes") that are never filled, and thus there is more than one "legal" way to fill them, i.e. there are different points with the same skeletons.

Conversely, we may characterise Toeplitz shifts via the MEF:

Theorem. *(Downarowicz)*

A minimal subshift $X \subset \mathcal{A}^{\mathbb{N}}$ (or $\mathcal{A}^{\mathbb{Z}}$) is Toeplitz iff there exists an odometer $\mathbb{Z} \overset{\omega}{\curvearrowright} \mathbb{Z}_{(p_n)_{n\geqslant 1}}$ and a factor map $\pi : X \twoheadrightarrow \mathbb{Z}_{(p_n)_{n\geqslant 1}}$ for which there exists some $z \in \mathbb{Z}_{(p_n)_{n\geqslant 1}}$ with exactly one preimage.

Definition. *(Regularity)*

A Toeplitz sequence η is said to be regular if:

$$d := \lim_{n\to\infty} \operatorname{dens}(\operatorname{Per}_{p_n}(\eta)) = \lim_{n\to\infty} \frac{\#\{0 \leqslant k < p_n : \eta_k = \eta_{k+p_n m} \text{ for all } m\}}{p_n} = 1.$$

A regular Toeplitz sequence is one where we fill the "gaps" with symbols quickly.

Example. For the sequence of periods $(3^n)_{n\geqslant 1}$ we have that:

- filling one coset at a time results in an irregular sequence ($d \to \frac{1}{2}$):

$$0\ 1\ 2\ 0\ 3\ 4\ 0\ 5\ 6\ 0\ 1\ 7\ 0\ 8\ 9\ 0\ ?\ ?\ 0\ 1\ ?\ 0\ ?\ ?\ 0\ ?\ 2\ 0\ 1\ \ldots$$

- filling all but one available cosets at a time produces a regular sequence ($d \to 1$):

$$0\ 0\ 1\ 0\ 0\ 1\ 0\ 0\ 2\ 0\ 0\ 1\ 0\ 0\ 1\ 0\ 0\ 2\ 0\ 0\ 1\ 0\ 0\ 1\ 0\ 0\ 3\ 0\ 0\ \ldots$$

Regularity

Regularity has a dynamical interpretation: η is regular iff "almost all" elements of X_η are Toeplitz sequences, and irregular if "almost none" is. Formally:

Theorem. *(Regularity, dynamically)*

A Toeplitz sequence η is regular iff the set $\{z \in \mathbb{Z}_{(p_n)_{n\geqslant 1}} : |\pi^{-1}[z]| = 1\}$ has Haar measure 1. Otherwise, this set has measure 0.

Any shift-invariant measure in X_η is heavily influenced by the Haar measure of the odometer. Thus, in the regular case we get:

Theorem. *(Jacobs–Keane)*

If η is regular, X_η is uniquely ergodic.

In this scenario, X_η has zero entropy, and thus low word complexity.

Ways to construct a Toeplitz sequence

Definition. *(Substitution)*

A length-ℓ substitution is a monoid morphism $\theta\colon \mathcal{A}^* \to \mathcal{A}^*$ such that $|\theta(a)| = \ell$ for all $a \in \mathcal{A}$.

Applying θ to a word $w = w_1 \dots w_k \in \mathcal{A}^*$ replaces each symbol w_j by the corresponding word $\theta(w_j)$ and concatenates the results. By iteration we get a set of words of increasing length, which may "converge" to an infinite word:

$$\theta\colon \begin{array}{ccc} 0 & \mapsto & 0\,1 \\ 1 & \mapsto & 0\,0 \end{array} \quad \rightsquigarrow \quad 0 \mapsto 01 \mapsto 0100 \mapsto 01000101 \mapsto \cdots \mapsto 0100010101000100010\dots$$

If $\theta(a)$ starts with a fixed symbol for all $a \in \mathcal{A}$, this limit word will be Toeplitz.

Moreover, we may use a sequence of many different monoid morphisms (a $\mathcal{S}$-adic sequence) with this property to obtain a wide variety of Toeplitz words.

The Oxtoby construction

This is a "standard" way to construct Toeplitz sequences from a period sequence $(p_n)_{n\geqslant 1}$ with $p_1 \geqslant 3, q_{i+1} = \frac{p_{i+1}}{p_i} \geqslant 3$. Given a non-constant sequence of symbols $(a_j)_{j\geqslant 1}$:

1. we fill the cosets $p_1\mathbb{Z}$ and $p_1\mathbb{Z}-1$ with a_1,
2. we proceed inductively: at step $n+1$, each block $[0, p_{n+1}-1]$ is a concatenation of q_{n+1} n-supertiles with unfilled gaps on the inside of each; we fill the inside of the first and the last supertile with copies of the symbol a_{n+1}.

0 1 0 0 2 0 0 1 0 0 1 0 0 3 0 0 1 0 0 1 0 0 2 0 0 1 0 ...

Theorem. *(Williams)*

An Oxtoby sequence η is regular iff $\sum_{i=2}^{\infty} \frac{1}{q_i}$ diverges. If it is irregular, the subshift X_η is not uniquely ergodic; more precisely, the set of ergodic invariant measures is in a 1-1 correspondence with the alphabet $\mathcal{A}$.

The Williams construction

This is a modified version of the Oxtoby construction where we fill the “gaps” with all possible legal words from a shift space Y, in such a way so that, from an element $x \in X_\eta$ that is not Toeplitz, we can “extract” a unique point $y \in Y$. We will not go into detail, but we note that:

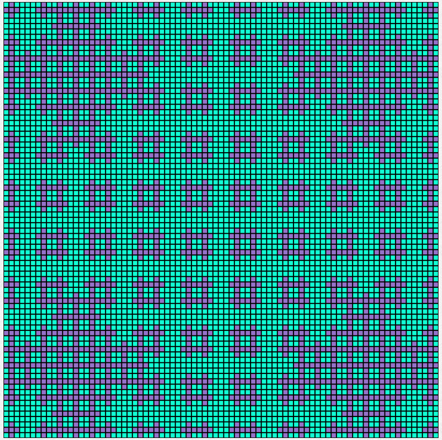
- the Toeplitz sequence obtained from this construction will be irregular iff the sum $\sum_{k=1}^{\infty} \frac{\#\mathscr{L}_k(x)}{q_{k+1}}$ diverges. This will also ensure that, for any non-Toeplitz $x \in X_\eta$, the set of “gaps” outside its skeleton will be unbounded above and below.

- Due to the way in which we insert the symbols in η, if we concatenate all symbols outside of a non-Toeplitz $x \in X_\eta$, we get a point $y \in Y$.

- In the irregular case almost all $x \in X_\eta$ are non-Toeplitz. We can use this to get a measurable isomorphism between X_η and a skew product over $\mathbb{Z}_{(p_n)_{n\geqslant 1}} \times Y$. Hence, X_η “inherits” all ergodic measures from Y and has entropy $(1-d)h_{\mathrm{top}}(Y)$.

A glimpse of the general group case

Beyond (bi-)infinite Toeplitz words, we may define Toeplitz configurations $x \in \mathcal{A}^{\mathbb{Z}^d}$, where for each $\vec{n} \in \mathbb{Z}^d$ there exist d linearly independent periods $\vec{p}_{(\vec{n},1)}, \ldots, \vec{p}_{(\vec{n},d)}$ such that $x_{\vec{n}} = x_{\vec{n}+a_1\vec{p}_{(\vec{n},1)}+\cdots+a_d\vec{p}_{(\vec{n},d)}}$, for all $a_1, \ldots, a_d \in \mathbb{Z}$:



Similar generalisations exist for residually finite groups, where we ask ourselves the same questions, together with new ones coming from geometry or group theory.

Some questions one may ask

- We have several parameters that are involved in the construction of a Toeplitz sequence:
  - the sequence of periods $(p_n)_{n\geqslant 1}$ (or equivalently, the sequence $(q_n)_{n\geqslant 1}$),
  - which cosets of $p_n\mathbb{Z}$ are filled at each step,
  - the actual list of symbols we use to fill the cosets.

  If we know these parameters only partially, how much of η can we reconstruct?
- If a number has as a Toeplitz sequence as its base β expansion, what can we say about it? (e.g. is it transcendent?)
- How do Toeplitz sequences relate to other sequences commonly studied in combinatorics and information theory (e.g. automatic or Sturmian sequences)?

Merci!