Finitely valued-generalised polynomials

Jakub Konieczny

Camille Jordan Institute
Claude Bernard University Lyon 1

One World Combinatorics on Words Seminar, 20 XII 2021

Based on joint work with B. Adamczewski & joint work with J. Byszewski.

Motivation: Sturmian sequences

Sturmian words $\mathbf{a} = (a_n)_{n=0}^{\infty} \in \{0,1\}^{\infty}$ have several characterisations:

Sturmian words $\mathbf{a} = (a_n)_{n=0}^{\infty} \in \{0,1\}^{\infty}$ have several characterisations:

Explicit formula: The word $\mathbf{a}$ is Sturmian if and only if it is given by

$$a_n = \lfloor \alpha n + \beta \rfloor - \lfloor \alpha(n-1) + \beta \rfloor, \qquad (n \in \mathbb{N}),$$

for some irrational $\alpha \in (0,1)$ and $\beta \in \mathbb{R}$ (or possibly the same formula with the floor $\lfloor \bullet \rfloor$ replaced by the ceiling $\lceil \bullet \rceil$).

Sturmian words $\mathbf{a} = (a_n)_{n=0}^{\infty} \in \{0,1\}^{\infty}$ have several characterisations:

Explicit formula: The word $\mathbf{a}$ is Sturmian if and only if it is given by

$$a_n = \lfloor \alpha n + \beta \rfloor - \lfloor \alpha(n-1) + \beta \rfloor, \qquad (n \in \mathbb{N}),$$

for some irrational $\alpha \in (0,1)$ and $\beta \in \mathbb{R}$ (or possibly the same formula with the floor $\lfloor \bullet \rfloor$ replaced by the ceiling $\lceil \bullet \rceil$).

Coding of a rotation: Let R_α denote the rotation $x \mapsto x + \alpha$ on the unit circle $\mathbb{R}/\mathbb{Z}$. The word $\mathbf{a}$ is Sturmian if and only if, for some irrational rotation R_α and base point $\beta \in \mathbb{R}/\mathbb{Z}$, a_n is given by

$$a_n = 1 \iff R_\alpha^n(\beta) \in [0, \alpha).$$

Sturmian words $\mathbf{a} = (a_n)_{n=0}^{\infty} \in \{0,1\}^{\infty}$ have several characterisations:

Explicit formula: The word $\mathbf{a}$ is Sturmian if and only if it is given by

$$a_n = \lfloor \alpha n + \beta \rfloor - \lfloor \alpha(n-1) + \beta \rfloor, \qquad (n \in \mathbb{N}),$$

for some irrational $\alpha \in (0,1)$ and $\beta \in \mathbb{R}$ (or possibly the same formula with the floor $\lfloor \bullet \rfloor$ replaced by the ceiling $\lceil \bullet \rceil$).

Coding of a rotation: Let R_α denote the rotation $x \mapsto x + \alpha$ on the unit circle $\mathbb{R}/\mathbb{Z}$. The word $\mathbf{a}$ is Sturmian if and only if, for some irrational rotation R_α and base point $\beta \in \mathbb{R}/\mathbb{Z}$, a_n is given by

$$a_n = 1 \iff R_\alpha^n(\beta) \in [0, \alpha).$$

Subword complexity: The word $\mathbf{a}$ over is Sturmian if and only if it has subword complexity $p_{\mathbf{a}}(N) = N + 1$ ($N \in \mathbb{N}$), i.e. for each N there are precisely $N + 1$ length-N subwords that appear in $\mathbf{a}$.

Sturmian words $\mathbf{a} = (a_n)_{n=0}^{\infty} \in \{0,1\}^{\infty}$ have several characterisations:

Explicit formula: The word $\mathbf{a}$ is Sturmian if and only if it is given by

$$a_n = \lfloor \alpha n + \beta \rfloor - \lfloor \alpha(n-1) + \beta \rfloor, \qquad (n \in \mathbb{N}),$$

for some irrational $\alpha \in (0,1)$ and $\beta \in \mathbb{R}$ (or possibly the same formula with the floor $\lfloor \bullet \rfloor$ replaced by the ceiling $\lceil \bullet \rceil$).

Coding of a rotation: Let R_α denote the rotation $x \mapsto x + \alpha$ on the unit circle $\mathbb{R}/\mathbb{Z}$. The word $\mathbf{a}$ is Sturmian if and only if, for some irrational rotation R_α and base point $\beta \in \mathbb{R}/\mathbb{Z}$, a_n is given by

$$a_n = 1 \iff R_\alpha^n(\beta) \in [0, \alpha).$$

Subword complexity: The word $\mathbf{a}$ over is Sturmian if and only if it has subword complexity $p_{\mathbf{a}}(N) = N + 1$ ($N \in \mathbb{N}$), i.e. for each N there are precisely $N + 1$ length-N subwords that appear in $\mathbf{a}$.

Balance: The word $\mathbf{a}$ is Sturmian if and only if for each $N \in \mathbb{N}$, its length-N subwords are balanced, i.e., for each $u, v \in \{0,1\}^N$ that appear in $\mathbf{a}$, the number of 1s in u and v differs by at most 1.

Overview

We will study a more general class, which we dub *bracket words*.

We will study a more general class, which we dub *bracket words*.

Explicit formula Bracket words are defined as codings of finitely-valued generalised polynomials (expressions built with $[\bullet], +, \times$).

Overview

We will study a more general class, which we dub *bracket words*.

Explicit formula Bracket words are defined as codings of finitely-valued generalised polynomials (expressions built with $[\bullet], +, \times$).

Coding of a rotation: Bracket words are given by codings of semialgebraic partitions of rotations on nilmanifolds.

We will study a more general class, which we dub *bracket words*.

Explicit formula Bracket words are defined as codings of finitely-valued generalised polynomials (expressions built with $[\bullet]$, $+$, $\times$).

Coding of a rotation: Bracket words are given by codings of semialgebraic partitions of rotations on nilmanifolds.

Subword complexity: Bracket words have polynomial subword complexity.

Overview

We will study a more general class, which we dub *bracket words*.

Explicit formula Bracket words are defined as codings of finitely-valued generalised polynomials (expressions built with $[\bullet], +, \times$).

Coding of a rotation: Bracket words are given by codings of semialgebraic partitions of rotations on nilmanifolds.

Subword complexity: Bracket words have polynomial subword complexity.

Balance: All symbols in a bracket word appear with uniform frequency.

We will study a more general class, which we dub *bracket words*.

Explicit formula Bracket words are defined as codings of finitely-valued generalised polynomials (expressions built with $[\bullet], +, \times$).

Coding of a rotation: Bracket words are given by codings of semialgebraic partitions of rotations on nilmanifolds.

Subword complexity: Bracket words have polynomial subword complexity.

Balance: All symbols in a bracket word appear with uniform frequency.

Other notable properties:

Overview

We will study a more general class, which we dub *bracket words*.

Explicit formula Bracket words are defined as codings of finitely-valued generalised polynomials (expressions built with $[\bullet]$, $+$, $\times$).

Coding of a rotation: Bracket words are given by codings of semialgebraic partitions of rotations on nilmanifolds.

Subword complexity: Bracket words have polynomial subword complexity.

Balance: All symbols in a bracket word appear with uniform frequency.

Other notable properties:

- Good closure properties (products, codings, orbit closure).

Overview

We will study a more general class, which we dub *bracket words*.

Explicit formula Bracket words are defined as codings of finitely-valued generalised polynomials (expressions built with $[\bullet], +, \times$).

Coding of a rotation: Bracket words are given by codings of semialgebraic partitions of rotations on nilmanifolds.

Subword complexity: Bracket words have polynomial subword complexity.

Balance: All symbols in a bracket word appear with uniform frequency.

Other notable properties:

- Good closure properties (products, codings, orbit closure).
- Unexpected examples, such at $a_n = \begin{cases} 1 & \text{if } n = \text{Fibonacci number}, \\ 0 & \text{otherwise.} \end{cases}$

Generalised polynomials

Definition (Generalised polynomial sequences)

The generalised polynomial sequences $\mathbb{N} \to \mathbb{R}$ (denoted GP) is the smallest family such that

- all polynomial sequences $\mathbb{Z} \to \mathbb{R}$ belong to GP;
- GP is a ring, i.e., $g, h \in \mathrm{GP} \implies g + h, g \cdot h \in \mathrm{GP}$;
- GP is closed under the floor function, i.e. $g \in \mathrm{GP} \Rightarrow [g] \in \mathrm{GP}$.

(Operations are pointwise: $(g + h)(n) = g(n) + h(n)$, $[g]\,(n) = [g(n)]$, etc.)

Definition (Generalised polynomial sequences)

The generalised polynomial sequences $\mathbb{N} \to \mathbb{R}$ (denoted GP) is the smallest family such that

- all polynomial sequences $\mathbb{Z} \to \mathbb{R}$ belong to GP;
- GP is a ring, i.e., $g, h \in \mathrm{GP} \implies g + h, g \cdot h \in \mathrm{GP}$;
- GP is closed under the floor function, i.e. $g \in \mathrm{GP} \Rightarrow [g] \in \mathrm{GP}$.

(Operations are pointwise: $(g+h)(n) = g(n) + h(n)$, $[g](n) = [g(n)]$, etc.)

Remark:

- Since $\{x\} = x - [x]$, GP is also closed under the fractional part function, i.e. $g \in \mathrm{GP} \Rightarrow \{g\} = g - [g] \in \mathrm{GP}$.
- Since $\|x\|_{\mathbb{R}/\mathbb{Z}} = \min_{a \in \mathbb{Z}} |x - a| = \left(\left\{x + \frac{1}{2}\right\} - \frac{1}{2}\right) \cdot \left(2\left[\left\{x + \frac{1}{2}\right\} - \frac{1}{2}\right] + 1\right)$, GP is also closed under the circle norm, i.e. $g \in \mathrm{GP} \Rightarrow \|g\|_{\mathbb{R}/\mathbb{Z}} \in \mathrm{GP}$.

Definition (Generalised polynomial sequences)

The generalised polynomial sequences $\mathbb{N} \to \mathbb{R}$ (denoted GP) is the smallest family such that

- all polynomial sequences $\mathbb{Z} \to \mathbb{R}$ belong to GP;
- GP is a ring, i.e., $g, h \in \mathrm{GP} \implies g + h, g \cdot h \in \mathrm{GP}$;
- GP is closed under the floor function, i.e. $g \in \mathrm{GP} \Rightarrow [g] \in \mathrm{GP}$.

(Operations are pointwise: $(g + h)(n) = g(n) + h(n)$, $[g](n) = [g(n)]$, etc.)

Remark:

- Since $\{x\} = x - [x]$, GP is also closed under the fractional part function, i.e. $g \in \mathrm{GP} \Rightarrow \{g\} = g - [g] \in \mathrm{GP}$.
- Since $\|x\|_{\mathbb{R}/\mathbb{Z}} = \min_{a \in \mathbb{Z}} |x - a| = \left(\left\{x + \frac{1}{2}\right\} - \frac{1}{2}\right) \cdot \left(2\left[\left\{x + \frac{1}{2}\right\} - \frac{1}{2}\right] + 1\right)$, GP is also closed under the circle norm, i.e. $g \in \mathrm{GP} \Rightarrow \|g\|_{\mathbb{R}/\mathbb{Z}} \in \mathrm{GP}$.

Related concepts:

- The GP sequences $\mathbb{N} \to \mathbb{R}^d$ are d-tuples $(g_1, \ldots, g_d)$ with $g_i \in \mathrm{GP}$.
- The GP maps $\mathbb{Z}^k \to \mathbb{R}^d$ or $\mathbb{R}^k \to \mathbb{R}^d$ are defined similarly.
- A set $A \subset \mathbb{N}$ is a GP set if 1_A is a GP sequence.

Definition (Bracket word)

A *bracket word* is a coding of a finitely-valued GP sequence.
More precisely, an infinite word $\mathbf{a} = (a_n)_{n=0}^{\infty}$ over a finite alphabet Σ is a *bracket word* if there exists a GP sequence $g\colon \mathbb{N} \to \mathbb{R}$ such that $g(\mathbb{N})$ is finite, and a coding $\pi\colon g(\mathbb{N}) \to \Sigma$ such that

$$a_n = \pi(g(n)), \qquad n \in \mathbb{N}.$$

Definition (Bracket word)

A *bracket word* is a coding of a finitely-valued GP sequence.
More precisely, an infinite word $\mathbf{a} = (a_n)_{n=0}^{\infty}$ over a finite alphabet Σ is a *bracket word* if there exists a GP sequence $g\colon \mathbb{N} \to \mathbb{R}$ such that $g(\mathbb{N})$ is finite, and a coding $\pi\colon g(\mathbb{N}) \to \Sigma$ such that

$$a_n = \pi(g(n)), \qquad n \in \mathbb{N}.$$

Proposal: Let's study bracket words from the point of view of Combinatorics on Words.

Definition (Bracket word)

A *bracket word* is a coding of a finitely-valued GP sequence.
More precisely, an infinite word $\mathbf{a} = (a_n)_{n=0}^{\infty}$ over a finite alphabet Σ is a *bracket word* if there exists a GP sequence $g \colon \mathbb{N} \to \mathbb{R}$ such that $g(\mathbb{N})$ is finite, and a coding $\pi \colon g(\mathbb{N}) \to \Sigma$ such that

$$a_n = \pi(g(n)), \qquad n \in \mathbb{N}.$$

Proposal: Let's study bracket words from the point of view of Combinatorics on Words.

Remark

The following objects are "equivalent":

- bracket words;
- finitely-valued GP sequences;
- GP subsets of $\mathbb{N}$.

For instance, if $E \subset \mathbb{N}$ is a GP subset, then the characteristic word $\mathbf{1}_E = (1_E(n))_{n=0}^{\infty}$ is a bracket word. In this talk, we focus on bracket words.

Lemma

Suppose that $g\colon \mathbb{N} \to \mathbb{R}$ *is a GP sequence and* $I \subset \mathbb{R}$ *is an interval. Put*

$$a_n = 1_I(g(n)) = \begin{cases} 1 & \text{if } g(n) \in I, \\ 0 & \text{otherwise.} \end{cases}$$

Then $\mathbf{a} = (a_n)_{n=0}^{\infty}$ *is a bracket word.*

Remark: The interval I can be proper, infinite or degenerate (e.g. $I = [x, y)$, $I = (x, \infty)$, $I = \{x\}$).

Lemma

Suppose that $g\colon \mathbb{N} \to \mathbb{R}$ *is a GP sequence and* $I \subset \mathbb{R}$ *is an interval. Put*

$$a_n = 1_I(g(n)) = \begin{cases} 1 & \text{if } g(n) \in I, \\ 0 & \text{otherwise.} \end{cases}$$

Then $\mathbf{a} = (a_n)_{n=0}^{\infty}$ *is a bracket word.*

Remark: The interval I can be proper, infinite or degenerate (e.g. $I = [x, y)$, $I = (x, \infty)$, $I = \{x\}$).

Example: To get a Sturmian sequence, let $g(n) = \{\alpha n + \beta\}$ and $I = [0, \alpha)$.

Lemma

Suppose that $g\colon \mathbb{N} \to \mathbb{R}$ is a GP sequence and $I \subset \mathbb{R}$ is an interval. Put

$$a_n = 1_I(g(n)) = \begin{cases} 1 & \text{if } g(n) \in I, \\ 0 & \text{otherwise.} \end{cases}$$

Then $\mathbf{a} = (a_n)_{n=0}^{\infty}$ is a bracket word.

Remark: The interval I can be proper, infinite or degenerate (e.g. $I = [x, y)$, $I = (x, \infty)$, $I = \{x\}$).

Example: To get a Sturmian sequence, let $g(n) = \{\alpha n + \beta\}$ and $I = [0, \alpha)$.

Proof for $I = \{0\}$: For $x \in \mathbb{R}$, we have

$$x = 0 \Leftrightarrow \{x\} = \{\sqrt{2}x\} = 0 \Leftrightarrow \left[1 - \frac{1}{2}\{x\} - \frac{1}{2}\{\sqrt{2}x\}\right] = 1.$$

Hence, $1_I(g(n)) = \left[1 - \frac{1}{2}\{g(n)\} - \frac{1}{2}\{\sqrt{2}g(n)\}\right]$.

Lemma

Suppose that $g\colon \mathbb{N} \to \mathbb{R}$ is a GP sequence and $I \subset \mathbb{R}$ is an interval. Put

$$a_n = 1_I(g(n)) = \begin{cases} 1 & \textit{if } g(n) \in I, \\ 0 & \textit{otherwise.} \end{cases}$$

Then $\mathbf{a} = (a_n)_{n=0}^{\infty}$ is a bracket word.

Remark: The interval I can be proper, infinite or degenerate (e.g. $I = [x, y)$, $I = (x, \infty)$, $I = \{x\}$).

Example: To get a Sturmian sequence, let $g(n) = \{\alpha n + \beta\}$ and $I = [0, \alpha)$.

Proof for $I = \{0\}$: For $x \in \mathbb{R}$, we have

$$x = 0 \Leftrightarrow \{x\} = \{\sqrt{2}x\} = 0 \Leftrightarrow \left[1 - \frac{1}{2}\{x\} - \frac{1}{2}\{\sqrt{2}x\}\right] = 1.$$

Hence, $1_I(g(n)) = \left[1 - \frac{1}{2}\{g(n)\} - \frac{1}{2}\{\sqrt{2}g(n)\}\right]$.

Remark: As long as I is bounded, similar tricks work. When I is unbounded, situation becomes more complicated; the analogue in $\mathbb{Z}$ if false!

Corollary

For a set $E \subset \mathbb{N}$ the following conditions are equivalent:

1. *The set E is GP (i.e., $1_E : \mathbb{N} \to \mathbb{R}$ is a GP sequence);*
2. *The characteristic word $\mathbf{1}_E = (1_E(n))_{n=0}^{\infty} \in \{0,1\}^{\infty}$ is a bracket word;*
3. *There is a GP sequence g with $E = \{n \in \mathbb{N} : g(n) = 0\}$;*
4. *There is a GP sequence h with $E = \{n \in \mathbb{N} : h(n) > 0\}$.*

Corollary

For a set $E \subset \mathbb{N}$ the following conditions are equivalent:

1. *The set E is GP (i.e., $1_E : \mathbb{N} \to \mathbb{R}$ is a GP sequence);*
2. *The characteristic word $\mathbf{1}_E = (1_E(n))_{n=0}^{\infty} \in \{0,1\}^{\infty}$ is a bracket word;*
3. *There is a GP sequence g with $E = \{n \in \mathbb{N} : g(n) = 0\}$;*
4. *There is a GP sequence h with $E = \{n \in \mathbb{N} : h(n) > 0\}$.*

Examples: The following are bracket words:

- $a_n = \begin{cases} 1 & \text{if } \{p(n)\} \in [0, \alpha), \\ 0 & \text{otherwise,} \end{cases}$ where $p(x) \in \mathbb{R}[x]$ and $\alpha \in (0,1)$;
- $a_n = \begin{cases} 1 & \text{if } \left\{\sqrt{2}n\left[\sqrt{3}n\right]\right\} \in [0, \alpha), \\ 0 & \text{otherwise,} \end{cases}$ where $\alpha \in (0,1)$.

Corollary

For a set $E \subset \mathbb{N}$ the following conditions are equivalent:

1. *The set E is GP (i.e., $1_E : \mathbb{N} \to \mathbb{R}$ is a GP sequence);*
2. *The characteristic word $\mathbf{1}_E = (1_E(n))_{n=0}^{\infty} \in \{0,1\}^{\infty}$ is a bracket word;*
3. *There is a GP sequence g with $E = \{n \in \mathbb{N} : g(n) = 0\}$;*
4. *There is a GP sequence h with $E = \{n \in \mathbb{N} : h(n) > 0\}$.*

Examples: The following are bracket words:

- $a_n = \begin{cases} 1 & \text{if } \{p(n)\} \in [0,\alpha), \\ 0 & \text{otherwise}, \end{cases}$ where $p(x) \in \mathbb{R}[x]$ and $\alpha \in (0,1)$;
- $a_n = \begin{cases} 1 & \text{if } \{\sqrt{2}n\,[\sqrt{3}n]\} \in [0,\alpha), \\ 0 & \text{otherwise}, \end{cases}$ where $\alpha \in (0,1)$.

Example

Let F_i denote the i-th Fibonacci number. Then $\{F_i : i \geq 1\}$ is a GP set.

Proof: An integer n is a Fibonacci number if and only if $\|n\varphi\|_{\mathbb{R}/\mathbb{Z}} < 1/2n$, where $\|x\|_{\mathbb{R}/\mathbb{Z}} = \min\{|x-a| : a \in \mathbb{Z}\}$. Take $g(n) = n\|n\varphi\|_{\mathbb{R}/\mathbb{Z}}$, $I = [0, 1/2)$.

Lemma (Closure under codings)

Let $\mathbf{a} = (a_n)_{n=0}^{\infty}$ *be a bracket word over an alphabet* Σ *and let* $\varphi \colon \Sigma \to \Lambda$ *be any map. Then* $\varphi(\mathbf{a}) = (\varphi(a_n))_{n=0}^{\infty}$ *is a bracket words over* Λ*.*

Lemma (Closure under codings)

Let $\mathbf{a} = (a_n)_{n=0}^{\infty}$ *be a bracket word over an alphabet* Σ *and let* $\varphi\colon \Sigma \to \Lambda$ *be any map. Then* $\varphi(\mathbf{a}) = (\varphi(a_n))_{n=0}^{\infty}$ *is a bracket words over* Λ.

Lemma (Closure under direct products)

Let $\mathbf{a} = (a_n)_{n=0}^{\infty}$ *and* $\mathbf{b} = (b_n)_{n=0}^{\infty}$ *be bracket words over alphabets* Σ *and* Λ *respectively. Then* $\mathbf{a} \times \mathbf{b} = ((a_n, b_n))_{n=0}^{\infty}$ *is a bracket word over* $\Sigma \times \Lambda$.

Consequence: Bracket words over a finite ring constitute a ring.

Lemma (Closure under codings)

Let $\mathbf{a} = (a_n)_{n=0}^{\infty}$ *be a bracket word over an alphabet* Σ *and let* $\varphi\colon \Sigma \to \Lambda$ *be any map. Then* $\varphi(\mathbf{a}) = (\varphi(a_n))_{n=0}^{\infty}$ *is a bracket words over* Λ.

Lemma (Closure under direct products)

Let $\mathbf{a} = (a_n)_{n=0}^{\infty}$ *and* $\mathbf{b} = (b_n)_{n=0}^{\infty}$ *be bracket words over alphabets* Σ *and* Λ *respectively. Then* $\mathbf{a} \times \mathbf{b} = ((a_n, b_n))_{n=0}^{\infty}$ *is a bracket word over* $\Sigma \times \Lambda$.

Consequence: Bracket words over a finite ring constitute a ring.

Lemma (Extracting progressions)

Let $\mathbf{a} = (a_n)_{n=0}^{\infty}$ *be a bracket word over an alphabet* Σ *and let* $k \in \mathbb{N}$, $r \in \mathbb{N}_0$ *be any map. Then* $(a_{kn+r})_{n=0}^{\infty}$ *is a bracket word over* Σ.

Lemma (Closure under codings)

Let $\mathbf{a} = (a_n)_{n=0}^{\infty}$ *be a bracket word over an alphabet* Σ *and let* $\varphi\colon \Sigma \to \Lambda$ *be any map. Then* $\varphi(\mathbf{a}) = (\varphi(a_n))_{n=0}^{\infty}$ *is a bracket words over* Λ*.*

Lemma (Closure under direct products)

Let $\mathbf{a} = (a_n)_{n=0}^{\infty}$ *and* $\mathbf{b} = (b_n)_{n=0}^{\infty}$ *be bracket words over alphabets* Σ *and* Λ *respectively. Then* $\mathbf{a} \times \mathbf{b} = ((a_n, b_n))_{n=0}^{\infty}$ *is a bracket word over* $\Sigma \times \Lambda$*.*

Consequence: Bracket words over a finite ring constitute a ring.

Lemma (Extracting progressions)

Let $\mathbf{a} = (a_n)_{n=0}^{\infty}$ *be a bracket word over an alphabet* Σ *and let* $k \in \mathbb{N}$, $r \in \mathbb{N}_0$ *be any map. Then* $(a_{kn+r})_{n=0}^{\infty}$ *is a bracket word over* Σ*.*

Let $\mathbf{a}$ be an infinite word over an alphabet Σ. Then a word $\mathbf{b}$ belongs to the *orbit closure of* $\mathbf{a}$ if each finite prefix $b_0 b_1 \cdots b_{n-1}$ of $\mathbf{b}$ is a subword of $\mathbf{a}$.

Proposition (Orbit closure)

Let $\mathbf{a} = (a_n)_{n=0}^{\infty}$ *be a bracket word over an alphabet* Σ*. Let* $\mathbf{b}$ *belong to the orbit closure of* $\mathbf{a}$*. Then* $\mathbf{b}$ *is a bracket word.*

Definition

Let G be a nilpotent Lie group, and $\Gamma < G$ a *cocompact* discrete subgroup.

1. The space $X = G/\Gamma$ is a *nilmanifold.*
2. For $g \in G$, the map $T_g \colon X \to X$, $x \mapsto gx$ is a *nilrotation.*
3. The dynamical system (X, T_g) is a nilsystem. It has a natural *Haar measure* μ_X which is T_g-invariant.
4. For $F \colon X \to \mathbb{R}$ (Lipschitz) and $x \in X$, $\big(F(T_g^n(x))\big)_{n=0}^{\infty}$ is a *nilsequence.*

Definition

Let G be a nilpotent Lie group, and $\Gamma < G$ a *cocompact* discrete subgroup.

1. The space $X = G/\Gamma$ is a *nilmanifold.*
2. For $g \in G$, the map $T_g \colon X \to X$, $x \mapsto gx$ is a *nilrotation.*
3. The dynamical system (X, T_g) is a nilsystem. It has a natural *Haar measure* μ_X which is T_g-invariant.
4. For $F \colon X \to \mathbb{R}$ (Lipschitz) and $x \in X$, $\big(F(T_g^n(x))\big)_{n=0}^{\infty}$ is a *nilsequence.*

A reassuring example:

- Take $G = \mathbb{R}$, $\Gamma = \mathbb{Z}$. Then $G/\Gamma = \mathbb{T}$, the unit circle.
- The circle comes equipped with $T_\alpha(x) = x + \alpha$.
- The Haar measure is just the Lebesgue measure.
- The additive characters $n \mapsto e^{2\pi i n\alpha}$ are 1-step nilsequences.
- Put $F = 1_{[0,\alpha)}$. Then $(F(T_\alpha^n(x)))_{n=0}^{\infty}$ is Sturmian. (F not continuous!)

Definition

Let G be a nilpotent Lie group, and $\Gamma < G$ a *cocompact* discrete subgroup.

1. The space $X = G/\Gamma$ is a *nilmanifold.*
2. For $g \in G$, the map $T_g \colon X \to X$, $x \mapsto gx$ is a *nilrotation.*
3. The dynamical system (X, T_g) is a nilsystem. It has a natural *Haar measure* μ_X which is T_g-invariant.
4. For $F \colon X \to \mathbb{R}$ (Lipschitz) and $x \in X$, $\big(F(T_g^n(x))\big)_{n=0}^{\infty}$ is a *nilsequence.*

A reassuring example:

- Take $G = \mathbb{R}$, $\Gamma = \mathbb{Z}$. Then $G/\Gamma = \mathbb{T}$, the unit circle.
- The circle comes equipped with $T_\alpha(x) = x + \alpha$.
- The Haar measure is just the Lebesgue measure.
- The additive characters $n \mapsto e^{2\pi i n\alpha}$ are 1-step nilsequences.
- Put $F = 1_{[0,\alpha)}$. Then $(F(T_\alpha^n(x)))_{n=0}^{\infty}$ is Sturmian. (F not continuous!)

Remark: Nilsequences are a central object of interest in higher order Fourier analysis, which we will not discuss here any further.

Theorem (Bergelson, Leibman)

Let $g\colon \mathbb{N} \to \mathbb{R}$ be a bounded GP sequence. Then is exists a nilsystem (X, T), a point $x \in X$ and a piecewise polynomial map $F\colon X \to \mathbb{R}$ such that

$$g(n) = F(T^n(x)), \qquad n \in \mathbb{N}.$$

Theorem (Bergelson, Leibman)

Let $g\colon \mathbb{N} \to \mathbb{R}$ be a bounded GP sequence. Then is exists a nilsystem (X, T), a point $x \in X$ and a piecewise polynomial map $F\colon X \to \mathbb{R}$ such that

$$g(n) = F(T^n(x)), \qquad n \in \mathbb{N}.$$

Definition: A set $S \subset [0,1)^d$ is *semialgebraic* if be defined by a finite number of polynomial equations and inequalities, e.g.

$$S = \left\{(x,y) \in [0,1)^2 \;:\; x^2 + y^2 \leq 1 \text{ and } x + y \neq 1\right\}.$$

A map $p\colon [0,1)^d \to \mathbb{R}$ is *piecewise polynomial* if there is a finite partition $[0,1)^d = \bigcup_i S_i$ with S_i semialgebraic, such that $p|_{S_i}$ are polynomials, e.g.

$$p(x,y) = \begin{cases} 0 & \text{if } x + y = 1, \\ xy + 1 & \text{if } x^2 + y^2 \leq 1 \text{ and } x + y \neq 1, \\ x + 2y & \text{if } x^2 + y^2 > 1. \end{cases}$$

Theorem (Bergelson, Leibman)

Let $g\colon \mathbb{N} \to \mathbb{R}$ be a bounded GP sequence. Then is exists a nilsystem (X, T), a point $x \in X$ and a piecewise polynomial map $F\colon X \to \mathbb{R}$ such that

$$g(n) = F(T^n(x)), \qquad n \in \mathbb{N}.$$

Definition: A set $S \subset [0,1)^d$ is *semialgebraic* if be defined by a finite number of polynomial equations and inequalities, e.g.

$$S = \left\{(x,y) \in [0,1)^2 \,:\, x^2 + y^2 \leq 1 \text{ and } x + y \neq 1\right\}.$$

A map $p\colon [0,1)^d \to \mathbb{R}$ is *piecewise polynomial* if there is a finite partition $[0,1)^d = \bigcup_i S_i$ with S_i semialgebraic, such that $p|_{S_i}$ are polynomials, e.g.

$$p(x,y) = \begin{cases} 0 & \text{if } x + y = 1, \\ xy + 1 & \text{if } x^2 + y^2 \leq 1 \text{ and } x + y \neq 1, \\ x + 2y & \text{if } x^2 + y^2 > 1. \end{cases}$$

A nilmanifold X has a system of *Mal'cev coordinates*, $\tau\colon X \to [0,1)^{\dim X}$. This allows us to speak of *piecewise polynomial maps* $X \to \mathbb{R}$.

$$G = \left\{ \begin{bmatrix} 1 & x & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{bmatrix}, \ x, y, z \in \mathbb{R} \right\}, \qquad \Gamma = \left\{ \begin{bmatrix} 1 & a & c \\ 0 & 1 & b \\ 0 & 0 & 1 \end{bmatrix}, \ a, b, c \in \mathbb{Z} \right\}$$

$$G_2 = [G, G] = \left\{ \begin{bmatrix} 1 & 0 & z \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \ z \in \mathbb{R} \right\}, \qquad G_3 = [G, G_2] = \{e_G\}.$$

Each element of G/Γ has a represenation $\begin{bmatrix} 1 & x & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{bmatrix} \Gamma$ with $x, y, z \in [0, 1)$.

$$\begin{bmatrix} 1 & \alpha & \gamma \\ 0 & 1 & \beta \\ 0 & 0 & 1 \end{bmatrix}^n \Gamma = \begin{bmatrix} 1 & n\alpha & n\gamma + \binom{n}{2}\alpha\beta \\ 0 & 1 & n\beta \\ 0 & 0 & 1 \end{bmatrix} \Gamma = \begin{bmatrix} 1 & \{n\alpha\} & \{f(n)\} \\ 0 & 1 & \{n\beta\} \\ 0 & 0 & 1 \end{bmatrix} \Gamma,$$

where $f(n) = n\gamma + \binom{n}{2}\alpha\beta - [n\alpha]\, n\beta.$

Subword complexity

Definition

Let $\mathbf{a} = (a_n)_{n=0}^{\infty} \in \Sigma^{\infty}$. The *subword complexity* $p_{\mathbf{a}}$ is given by

$$p_{\mathbf{a}}(N) = \#\left\{ w \in \Sigma^N \ : \ w \text{ is a subword of } \mathbf{a} \right\}.$$

Definition

Let $\mathbf{a} = (a_n)_{n=0}^{\infty} \in \Sigma^{\infty}$. The *subword complexity* $p_{\mathbf{a}}$ is given by

$$p_{\mathbf{a}}(N) = \# \left\{ w \in \Sigma^N \ : \ w \text{ is a subword of } \mathbf{a} \right\}.$$

- If $\mathbf{a}$ is eventually periodic then $p_{\mathbf{a}}(N)$ is bounded.
- If $p_{\mathbf{a}}(N) \leq N$ for some N then $\mathbf{a}$ is eventually periodic.
- The sequence $\mathbf{a}$ is Sturmian if and only if $p_{\mathbf{a}}(N) = N + 1$ for all N.
- If a is automatic then $p_{\mathbf{a}}(N) = O(N)$.
- If a is morphic then $p_{\mathbf{a}}(N) = O(N^2)$.

Definition

Let $\mathbf{a} = (a_n)_{n=0}^{\infty} \in \Sigma^{\infty}$. The *subword complexity* $p_{\mathbf{a}}$ is given by

$$p_{\mathbf{a}}(N) = \# \left\{ w \in \Sigma^N \ : \ w \text{ is a subword of } \mathbf{a} \right\}.$$

- If $\mathbf{a}$ is eventually periodic then $p_{\mathbf{a}}(N)$ is bounded.
- If $p_{\mathbf{a}}(N) \leq N$ for some N then $\mathbf{a}$ is eventually periodic.
- The sequence $\mathbf{a}$ is Sturmian if and only if $p_{\mathbf{a}}(N) = N + 1$ for all N.
- If a is automatic then $p_{\mathbf{a}}(N) = O(N)$.
- If a is morphic then $p_{\mathbf{a}}(N) = O(N^2)$.

Theorem (Adamczewski, K.)

If $\mathbf{a}$ is a bracket word then $p_{\mathbf{a}}(N) = O(N^C)$ for a constant $C > 0$.

Question: Given a bracket word $\mathbf{a}$, what is the best possible value of C?

Proof ideas

- As an example, consider $a_n = \left[2\left\{n^2\sqrt{2}\right\}\right] = \begin{cases} 0 & \text{if } \left\{n^2\sqrt{2}\right\} \in [0, \frac{1}{2}), \\ 1 & \text{if } \left\{n^2\sqrt{2}\right\} \in [\frac{1}{2}, 1). \end{cases}$

Proof ideas

- As an example, consider $a_n = \left[2\left\{n^2\sqrt{2}\right\}\right] = \begin{cases} 0 & \text{if } \left\{n^2\sqrt{2}\right\} \in [0, \frac{1}{2}), \\ 1 & \text{if } \left\{n^2\sqrt{2}\right\} \in [\frac{1}{2}, 1). \end{cases}$
- Note that $a_{n+m} = \left[2\left\{n^2\sqrt{2} + n\alpha_m + \beta_m\right\}\right]$ where $\alpha_m = \left\{2m\sqrt{2}\right\} \in [0, 1)$ and $\beta_m = \left\{m^2\sqrt{2}\right\} \in [0, 1)$.

Proof ideas

- As an example, consider $a_n = \left[2\left\{n^2\sqrt{2}\right\}\right] = \begin{cases} 0 & \text{if } \left\{n^2\sqrt{2}\right\} \in [0,\frac{1}{2}), \\ 1 & \text{if } \left\{n^2\sqrt{2}\right\} \in [\frac{1}{2},1). \end{cases}$.
- Note that $a_{n+m} = \left[2\left\{n^2\sqrt{2} + n\alpha_m + \beta_m\right\}\right]$ where $\alpha_m = \left\{2m\sqrt{2}\right\} \in [0,1)$ and $\beta_m = \left\{m^2\sqrt{2}\right\} \in [0,1)$.
- Let $N \in \mathbb{N}$. Task: count length-N subwords $\mathbf{a}$.
  Instead, we will count length-N prefixes of words $\mathbf{a}^{\alpha,\beta}$ $(\alpha, \beta \in [0,1))$ given by $a_n^{\alpha,\beta} = \left[2\left\{n^2\sqrt{2} + n\alpha + \beta\right\}\right]$.

Proof ideas

- As an example, consider $a_n = \left[2\left\{n^2\sqrt{2}\right\}\right] = \begin{cases} 0 & \text{if } \left\{n^2\sqrt{2}\right\} \in [0, \frac{1}{2}), \\ 1 & \text{if } \left\{n^2\sqrt{2}\right\} \in [\frac{1}{2}, 1). \end{cases}$.
- Note that $a_{n+m} = \left[2\left\{n^2\sqrt{2} + n\alpha_m + \beta_m\right\}\right]$ where $\alpha_m = \left\{2m\sqrt{2}\right\} \in [0,1)$ and $\beta_m = \left\{m^2\sqrt{2}\right\} \in [0,1)$.
- Let $N \in \mathbb{N}$. Task: count length-N subwords $\mathbf{a}$.
  Instead, we will count length-N prefixes of words $\mathbf{a}^{\alpha,\beta}$ $(\alpha, \beta \in [0,1))$ given by $a_n^{\alpha,\beta} = \left[2\left\{n^2\sqrt{2} + n\alpha + \beta\right\}\right]$.
- Approximate $\alpha \simeq \alpha^*$, $\beta \simeq \beta^*$ by rationals, error $\leq N^{-10}$ (N^{20} choices).
  Note that $n^2\sqrt{2} + n\alpha + \beta = n^2\sqrt{2} + n\alpha^* + \beta^* + O(N^{-9})$ $(n < N)$.

Proof ideas

- As an example, consider $a_n = \left[2\left\{n^2\sqrt{2}\right\}\right] = \begin{cases} 0 & \text{if } \left\{n^2\sqrt{2}\right\} \in [0, \frac{1}{2}), \\ 1 & \text{if } \left\{n^2\sqrt{2}\right\} \in [\frac{1}{2}, 1). \end{cases}$
- Note that $a_{n+m} = \left[2\left\{n^2\sqrt{2} + n\alpha_m + \beta_m\right\}\right]$ where $\alpha_m = \left\{2m\sqrt{2}\right\} \in [0,1)$ and $\beta_m = \left\{m^2\sqrt{2}\right\} \in [0,1)$.
- Let $N \in \mathbb{N}$. Task: count length-N subwords $\mathbf{a}$.
  Instead, we will count length-N prefixes of words $\mathbf{a}^{\alpha,\beta}$ $(\alpha, \beta \in [0,1))$ given by $a_n^{\alpha,\beta} = \left[2\left\{n^2\sqrt{2} + n\alpha + \beta\right\}\right]$.
- Approximate $\alpha \simeq \alpha^*$, $\beta \simeq \beta^*$ by rationals, error $\leq N^{-10}$ $(N^{20}$ choices).
  Note that $n^2\sqrt{2} + n\alpha + \beta = n^2\sqrt{2} + n\alpha^* + \beta^* + O(N^{-9})$ $(n < N)$.
- There are two types of points $n < N$:
  The Good: $\|2n^2\sqrt{2} + 2n\alpha^* + 2\beta^*\|_{\mathbb{R}/\mathbb{Z}} > N^{-8}$, so we can compute
  $a_n^{\alpha,\beta} = \left[2\left\{n^2\sqrt{2} + n\alpha^* + \beta^*\right\}\right]$. If so, we are done.
  The Bad: $\|2n^2\sqrt{2} + 2n\alpha^* + 2\beta^*\|_{\mathbb{R}/\mathbb{Z}} < N^{-8}$. Let
  $\delta_n := 2n^2\sqrt{2} + 2n\alpha + 2\beta - \left\lfloor 2n^2\sqrt{2} + 2n\alpha^* + 2\beta^* \right\rceil = O(N^{-8})$.

Proof ideas

- As an example, consider $a_n = \left[2\left\{n^2\sqrt{2}\right\}\right] = \begin{cases} 0 & \text{if } \left\{n^2\sqrt{2}\right\} \in [0, \frac{1}{2}), \\ 1 & \text{if } \left\{n^2\sqrt{2}\right\} \in [\frac{1}{2}, 1). \end{cases}$
- Note that $a_{n+m} = \left[2\left\{n^2\sqrt{2} + n\alpha_m + \beta_m\right\}\right]$ where $\alpha_m = \left\{2m\sqrt{2}\right\} \in [0,1)$ and $\beta_m = \left\{m^2\sqrt{2}\right\} \in [0,1)$.
- Let $N \in \mathbb{N}$. Task: count length-N subwords $\mathbf{a}$.
  Instead, we will count length-N prefixes of words $\mathbf{a}^{\alpha,\beta}$ ($\alpha, \beta \in [0,1)$) given by $a_n^{\alpha,\beta} = \left[2\left\{n^2\sqrt{2} + n\alpha + \beta\right\}\right]$.
- Approximate $\alpha \simeq \alpha^*$, $\beta \simeq \beta^*$ by rationals, error $\leq N^{-10}$ (N^{20} choices).
  Note that $n^2\sqrt{2} + n\alpha + \beta = n^2\sqrt{2} + n\alpha^* + \beta^* + O(N^{-9})$ ($n < N$).
- There are two types of points $n < N$:
  The Good: $\|2n^2\sqrt{2} + 2n\alpha^* + 2\beta^*\|_{\mathbb{R}/\mathbb{Z}} > N^{-8}$, so we can compute
  $a_n^{\alpha,\beta} = \left[2\left\{n^2\sqrt{2} + n\alpha^* + \beta^*\right\}\right]$. If so, we are done.
  The Bad: $\|2n^2\sqrt{2} + 2n\alpha^* + 2\beta^*\|_{\mathbb{R}/\mathbb{Z}} < N^{-8}$. Let
  $\delta_n := 2n^2\sqrt{2} + 2n\alpha + 2\beta - \left\lfloor 2n^2\sqrt{2} + 2n\alpha^* + 2\beta^* \right\rceil = O(N^{-8})$.
- If we know $\operatorname{sgn}(\delta_n)$ and α^* and β^*, then we can recover $a_n^{\alpha,\beta}$. E.g.:
  - If $\left\{n^2\sqrt{2} + n\alpha^* + \beta^*\right\} \approx 0$ and $\delta_n > 0$ then $a_n^{\alpha,\beta} = 0$.

Proof ideas

- As an example, consider $a_n = \left[2\left\{n^2\sqrt{2}\right\}\right] = \begin{cases} 0 & \text{if } \left\{n^2\sqrt{2}\right\} \in [0, \frac{1}{2}), \\ 1 & \text{if } \left\{n^2\sqrt{2}\right\} \in [\frac{1}{2}, 1). \end{cases}$.
- Note that $a_{n+m} = \left[2\left\{n^2\sqrt{2} + n\alpha_m + \beta_m\right\}\right]$ where $\alpha_m = \left\{2m\sqrt{2}\right\} \in [0,1)$ and $\beta_m = \left\{m^2\sqrt{2}\right\} \in [0,1)$.
- Let $N \in \mathbb{N}$. Task: count length-N subwords $\mathbf{a}$.
  Instead, we will count length-N prefixes of words $\mathbf{a}^{\alpha,\beta}$ $(\alpha, \beta \in [0,1))$ given by $a_n^{\alpha,\beta} = \left[2\left\{n^2\sqrt{2} + n\alpha + \beta\right\}\right]$.
- Approximate $\alpha \simeq \alpha^*$, $\beta \simeq \beta^*$ by rationals, error $\leq N^{-10}$ $(N^{20}$ choices).
  Note that $n^2\sqrt{2} + n\alpha + \beta = n^2\sqrt{2} + n\alpha^* + \beta^* + O(N^{-9})$ $(n < N)$.
- There are two types of points $n < N$:
  The Good: $\|2n^2\sqrt{2} + 2n\alpha^* + 2\beta^*\|_{\mathbb{R}/\mathbb{Z}} > N^{-8}$, so we can compute
  $a_n^{\alpha,\beta} = \left[2\left\{n^2\sqrt{2} + n\alpha^* + \beta^*\right\}\right]$. If so, we are done.
  The Bad: $\|2n^2\sqrt{2} + 2n\alpha^* + 2\beta^*\|_{\mathbb{R}/\mathbb{Z}} < N^{-8}$. Let
  $\delta_n := 2n^2\sqrt{2} + 2n\alpha + 2\beta - \left\lfloor 2n^2\sqrt{2} + 2n\alpha^* + 2\beta^* \right\rceil = O(N^{-8})$.
- If we know $\mathrm{sgn}(\delta_n)$ and α^* and β^*, then we can recover $a_n^{\alpha,\beta}$. E.g.:
  - If $\left\{n^2\sqrt{2} + n\alpha^* + \beta^*\right\} \approx 0$ and $\delta_n > 0$ then $a_n^{\alpha,\beta} = 0$.
- For each $n < N$, the set $\{(\alpha, \beta) : \delta_n = 0\}$ is a line.
  Key fact: N lines divide the plane into $O(N^2)$ regions.

Uniform frequency

For an infinite word over Σ and $i \in \Sigma$, let

$$\mathrm{freq}^*(\mathbf{a}, i) = \limsup_{N\to\infty} \sup_{M\in\mathbb{N}} \frac{1}{N} \# \{M \leq n < M + N \, : \, a_n = i\}$$

$$\mathrm{freq}_*(\mathbf{a}, i) = \liminf_{N\to\infty} \inf_{M\in\mathbb{N}} \frac{1}{N} \# \{M \leq n < M + N \, : \, a_n = i\} \, .$$

Uniform frequency

For an infinite word over Σ and $i \in \Sigma$, let

$$\mathrm{freq}^*(\mathbf{a}, i) = \limsup_{N\to\infty} \sup_{M\in\mathbb{N}} \frac{1}{N} \# \{M \le n < M+N \,:\, a_n = i\}$$

$$\mathrm{freq}_*(\mathbf{a}, i) = \liminf_{N\to\infty} \inf_{M\in\mathbb{N}} \frac{1}{N} \# \{M \le n < M+N \,:\, a_n = i\}.$$

Proposition

Let $\mathbf{a}$ *be a bracket word over an alphabet* Σ *and let* $i \in \Sigma$. *Then*

$$\mathrm{freq}^*(\mathbf{a}, i) = \mathrm{freq}_*(\mathbf{a}, i) := \mathrm{freq}(\mathbf{a}, i).$$

Equivalently: If $E \subset \mathbb{N}$ is a GP set, then $d^*(E) = d_*(E)$.

Uniform frequency

For an infinite word over Σ and $i \in \Sigma$, let

$$\mathrm{freq}^*(\mathbf{a}, i) = \limsup_{N\to\infty} \sup_{M\in\mathbb{N}} \frac{1}{N} \# \{M \le n < M+N \,:\, a_n = i\}$$
$$\mathrm{freq}_*(\mathbf{a}, i) = \liminf_{N\to\infty} \inf_{M\in\mathbb{N}} \frac{1}{N} \# \{M \le n < M+N \,:\, a_n = i\}.$$

Proposition

Let $\mathbf{a}$ *be a bracket word over an alphabet* Σ *and let* $i \in \Sigma$*. Then*

$$\mathrm{freq}^*(\mathbf{a}, i) = \mathrm{freq}_*(\mathbf{a}, i) := \mathrm{freq}(\mathbf{a}, i).$$

Equivalently: If $E \subset \mathbb{N}$ is a GP set, then $d^*(E) = d_*(E)$.

Corollary: If $\mathrm{freq}(\mathbf{a}, i) > 0$ then $\{n \in \mathbb{N} \,:\, a_n = i\}$ is syndetic (there exists N such that for each $M \in \mathbb{N}$ there exists $M \le n < M+N$ with $a_n = i$).

For an infinite word over Σ and $i \in \Sigma$, let

$$\mathrm{freq}^*(\mathbf{a}, i) = \limsup_{N\to\infty} \sup_{M\in\mathbb{N}} \frac{1}{N} \# \{M \le n < M+N \,:\, a_n = i\}$$

$$\mathrm{freq}_*(\mathbf{a}, i) = \liminf_{N\to\infty} \inf_{M\in\mathbb{N}} \frac{1}{N} \# \{M \le n < M+N \,:\, a_n = i\}.$$

Proposition

Let $\mathbf{a}$ *be a bracket word over an alphabet* Σ *and let* $i \in \Sigma$*. Then*

$$\mathrm{freq}^*(\mathbf{a}, i) = \mathrm{freq}_*(\mathbf{a}, i) := \mathrm{freq}(\mathbf{a}, i).$$

Equivalently: If $E \subset \mathbb{N}$ is a GP set, then $d^*(E) = d_*(E)$.

Corollary: If $\mathrm{freq}(\mathbf{a}, i) > 0$ then $\{n \in \mathbb{N} \,:\, a_n = i\}$ is syndetic (there exists N such that for each $M \in \mathbb{N}$ there exists $M \le n < M+N$ with $a_n = i$).

Remark: The same results remain valid, *mutatis mutandis*, if $i \in \Sigma$ is replaced with a word $w \in \Sigma^\ell$. (E.g. because $(a_n a_{n+1} \cdots a_{n+\ell-1})_{n=0}^\infty$ is a bracket word over Σ^ℓ.)

Uniform frequency

For an infinite word over Σ and $i \in \Sigma$, let

$$\mathrm{freq}^*(\mathbf{a}, i) = \limsup_{N\to\infty} \sup_{M\in\mathbb{N}} \frac{1}{N} \# \{M \le n < M+N \,:\, a_n = i\}$$

$$\mathrm{freq}_*(\mathbf{a}, i) = \liminf_{N\to\infty} \inf_{M\in\mathbb{N}} \frac{1}{N} \# \{M \le n < M+N \,:\, a_n = i\}.$$

Proposition

Let $\mathbf{a}$ be a bracket word over an alphabet Σ and let $i \in \Sigma$. Then

$$\mathrm{freq}^*(\mathbf{a}, i) = \mathrm{freq}_*(\mathbf{a}, i) := \mathrm{freq}(\mathbf{a}, i).$$

Equivalently: If $E \subset \mathbb{N}$ is a GP set, then $d^*(E) = d_*(E)$.

Corollary: If $\mathrm{freq}(\mathbf{a}, i) > 0$ then $\{n \in \mathbb{N} \,:\, a_n = i\}$ is syndetic (there exists N such that for each $M \in \mathbb{N}$ there exists $M \le n < M+N$ with $a_n = i$).

Remark: The same results remain valid, *mutatis mutandis*, if $i \in \Sigma$ is replaced with a word $w \in \Sigma^\ell$. (E.g. because $(a_n a_{n+1} \cdots a_{n+\ell-1})_{n=0}^\infty$ is a bracket word over Σ^ℓ.)

Proof: Apply Bergelson–Leibman; ergodic nilsystems are uniquely ergodic.

Quantitative frequency

Negative results: Estimates from previous slide cannot be improved. For any function $f\colon \mathbb{N} \to [0,1)$ with $f(N) \to \infty$ as $N \to \infty$, there exists a bracket word $\mathbf{a}$ over $\{0,1\}$ such that

$$f(N) \leq \frac{\#\{n < N \ : \ a_n = 1\}}{N} \to 0 \text{ as } N \to \infty.$$

We can take $a_n = \begin{cases} 1 & \text{if } \|\alpha n\|_{\mathbb{R}/\mathbb{Z}} \|\beta n\|_{\mathbb{R}/\mathbb{Z}} < 1/n, \\ 0 & \text{otherwise,} \end{cases}$ where $\alpha, \beta \in \mathbb{R} \setminus \mathbb{Q}$ are sufficiently well-approximable by rationals.

Negative results: Estimates from previous slide cannot be improved. For any function $f\colon \mathbb{N} \to [0,1)$ with $f(N) \to \infty$ as $N \to \infty$, there exists a bracket word $\mathbf{a}$ over $\{0,1\}$ such that

$$f(N) \leq \frac{\#\{n < N \ : \ a_n = 1\}}{N} \to 0 \text{ as } N \to \infty.$$

We can take $a_n = \begin{cases} 1 & \text{if } \|\alpha n\|_{\mathbb{R}/\mathbb{Z}} \|\beta n\|_{\mathbb{R}/\mathbb{Z}} < 1/n, \\ 0 & \text{otherwise,} \end{cases}$ where $\alpha, \beta \in \mathbb{R} \setminus \mathbb{Q}$ are sufficiently well-approximable by rationals.

Positive results: Quantitative results on equidistribution of orbits on nilmanifolds (e.g. Green–Tao) translate into estimates on frequencies of symbols in bracket words, but we must either

- make additional assumptions about Diophantine properties of the coefficients in the definition of $\mathbf{a}$; or
- deal with the possibility that $\mathbf{a}$ has different behaviour on different long arithmetic progressions.

IP sets

Finite sums. For a sequence $(n_i)_{i=1}^{\infty}$, $n_i \in \mathbb{N}$, define:

$$\mathrm{FS}\left((n_i)_{i=1}^{\infty}\right) = \Big\{ \sum_{i \in I} n_i \ : \ I \subset \mathbb{N}, \ \text{finite}, \ I \neq \emptyset \Big\}.$$

IP sets

Finite sums. For a sequence $(n_i)_{i=1}^{\infty}$, $n_i \in \mathbb{N}$, define:

$$\mathrm{FS}\left((n_i)_{i=1}^{\infty}\right) = \Big\{ \sum_{i \in I} n_i \; : \; I \subset \mathbb{N}, \text{ finite}, \; I \neq \emptyset \Big\}.$$

- A set $A \subset \mathbb{N}$ is IP if $A \supset \mathrm{FS}\left((n_i)_{i=1}^{\infty}\right)$ for some sequence $(n_i)_{i=1}^{\infty}$.
- A set $B \subset \mathbb{N}$ is IP* if $B \cap \mathrm{FS}\left((n_i)_{i=1}^{\infty}\right) \neq \emptyset$ for each sequence $(n_i)_{i=1}^{\infty}$.

IP sets

Finite sums. For a sequence $(n_i)_{i=1}^{\infty}$, $n_i \in \mathbb{N}$, define:

$$\mathrm{FS}\left((n_i)_{i=1}^{\infty}\right) = \left\{ \sum_{i \in I} n_i \ : \ I \subset \mathbb{N},\ \text{finite},\ I \neq \emptyset \right\}.$$

- A set $A \subset \mathbb{N}$ is IP if $A \supset \mathrm{FS}\left((n_i)_{i=1}^{\infty}\right)$ for some sequence $(n_i)_{i=1}^{\infty}$.
- A set $B \subset \mathbb{N}$ is IP* if $B \cap \mathrm{FS}\left((n_i)_{i=1}^{\infty}\right) \neq \emptyset$ for each sequence $(n_i)_{i=1}^{\infty}$.

Fact

Any IP* *set is syndetic (i.e. intersects any sufficiently long interval).*

IP sets

Finite sums. For a sequence $(n_i)_{i=1}^{\infty}$, $n_i \in \mathbb{N}$, define:

$$\mathrm{FS}\left((n_i)_{i=1}^{\infty}\right) = \left\{ \sum_{i \in I} n_i \; : \; I \subset \mathbb{N}, \text{ finite}, \; I \neq \emptyset \right\}.$$

- A set $A \subset \mathbb{N}$ is IP if $A \supset \mathrm{FS}\left((n_i)_{i=1}^{\infty}\right)$ for some sequence $(n_i)_{i=1}^{\infty}$.
- A set $B \subset \mathbb{N}$ is IP* if $B \cap \mathrm{FS}\left((n_i)_{i=1}^{\infty}\right) \neq \emptyset$ for each sequence $(n_i)_{i=1}^{\infty}$.

Fact

Any IP* *set is syndetic (i.e. intersects any sufficiently long interval).*

Theorem (Hindman)

- *If A is an* IP *set, $A = A_1 \cup A_2 \cup \cdots \cup A_r$ then $\exists j :$ A_j is* IP.
- *If $B_1, B_2 \ldots, B_r$ are* IP* *sets then $B = B_1 \cap B_2 \cap \cdots \cap B_r$ is* IP*.

Corollary (Bergelson, Leibman)

- *Let* $g\colon \mathbb{N} \to \mathbb{R}$ *be a bounded GP sequence. Then for almost all* $n \in \mathbb{N}$*, for any* $\delta > 0$*, the set* $R = \{m \in \mathbb{N} : |g(n+m) - g(n)| < \delta\}$ *is* IP^*.
- *If* $\mathbf{a}$ *is a bracket word then for almost all* $n \in \mathbb{N}$ *the set* $\{m \in \mathbb{N} : a_{n+m} = a_n\}$ *is* IP^*.
- *For any GP set* $E \subset \mathbb{N}$ *with* $d(E) > 0$*, the set* $E - n$ *is* IP^* *for some* n.

Corollary (Bergelson, Leibman)

- *Let* $g\colon \mathbb{N} \to \mathbb{R}$ *be a bounded GP sequence. Then for almost all* $n \in \mathbb{N}$, *for any* $\delta > 0$, *the set* $R = \{m \in \mathbb{N} : |g(n+m) - g(n)| < \delta\}$ *is* IP*.
- *If* $\mathbf{a}$ *is a bracket word then for almost all* $n \in \mathbb{N}$ *the set* $\{m \in \mathbb{N} : a_{n+m} = a_n\}$ *is* IP*.
- *For any GP set* $E \subset \mathbb{N}$ *with* $d(E) > 0$, *the set* $E - n$ *is* IP* *for some* n.

Stronger versions: Let $(n_i)_{i=1}^{\infty}$ be a sequence, and let R be as above.

- IP$_r$ recurrence: We can find $\sum_{i\in I} n_i \in R$ with $I \subset [r]$ for some r that depends on g, δ and n [Bergelson, Leibman];
- SG$_d$ recurrence: We can find $\sum_{i\in I} n_i \in R$ where I has gaps bounded by d for some d that depends on g and δ [K.];
- VIP recurrence: [Bergelson, Håland Knutson, McCutcheon];

One bracket word can have many distinct representations. For instance:

$$\begin{aligned} 1_{\{0\}}(n) &= \left[1 - \left\{\sqrt{2}n\right\}\right] = \left[1 - \left\{\sqrt{3}n\right\}\right] \\ &= \left[[\sqrt{2}n]2\sqrt{2}n\right] - \left[\sqrt{2}n\right]^2 - 2n^2 + 1, \qquad n \in \mathbb{N}. \end{aligned}$$

Representations and decidability

One bracket word can have many distinct representations. For instance:

$$
\begin{aligned}
1_{\{0\}}(n) &= \left[1 - \left\{\sqrt{2}n\right\}\right] = \left[1 - \left\{\sqrt{3}n\right\}\right] \\
&= \left[[\sqrt{2}n]2\sqrt{2}n\right] - \left[\sqrt{2}n\right]^2 - 2n^2 + 1, \qquad n \in \mathbb{N}.
\end{aligned}
$$

Question

Can we determine if two bracket words (or GP sequences) are equal?

One bracket word can have many distinct representations. For instance:

$$1_{\{0\}}(n) = \left[1 - \left\{\sqrt{2}n\right\}\right] = \left[1 - \left\{\sqrt{3}n\right\}\right]$$
$$= \left[[\sqrt{2}n]2\sqrt{2}n\right] - \left[\sqrt{2}n\right]^2 - 2n^2 + 1, \qquad n \in \mathbb{N}.$$

Question

Can we determine if two bracket words (or GP sequences) are equal?

For polynomial sequences, there are also unexpected identities, like

$$n^4 + 4 = (n^2 + 2n + 2)(n^2 - 2n + 2).$$

This is not a problem, since polynomials have a "canonical" representation

$$p(n) = \sum_{i=0}^{d} n^i \alpha_i, \qquad (d \in \mathbb{N}_0,\ \alpha_i \in \mathbb{R}).$$

Question

Is there an analogous statement for bracket words (or GP sequences)?

Canonical representation

Let us call a GP sequence $g\colon \mathbb{N} \to \mathbb{R}$ a *generalised monomial* if it can be expressed using only polynomials, fractional part, and multiplication, e.g.

$$p_1(n)\{p_2(n)\},\ p_1(n)\{p_2(n)\}\{p_3(n)\},\ p_1(n)\{p_2(n)\}\{p_3(n)\{p_4(n)\}\},$$

where $p_i\colon \mathbb{N} \to \mathbb{R}$ are (ordinary) polynomial sequences.

Canonical representation

Let us call a GP sequence $g\colon \mathbb{N} \to \mathbb{R}$ a *generalised monomial* if it can be expressed using only polynomials, fractional part, and multiplication, e.g.

$$p_1(n)\left\{p_2(n)\right\},\; p_1(n)\left\{p_2(n)\right\}\left\{p_3(n)\right\},\; p_1(n)\left\{p_2(n)\right\}\left\{p_3(n)\left\{p_4(n)\right\}\right\},$$

where $p_i\colon \mathbb{N} \to \mathbb{R}$ are (ordinary) polynomial sequences.

Theorem (Leibman)

Each bounded GP sequence g has a "canonical representation"

$$g(n) = F\left(h_0(n), h_1(n), h_2(n), \ldots, h_d(n)\right), \qquad n \in \mathbb{N},$$

where F is a piecewise polynomial function, h_0 is periodic and $h_1, h_2, \ldots, h_d\colon \mathbb{N} \to [0,1)$ are jointly equidistributed generalised monomials.

Remark: The representation is explicitly computable.

Let us call a GP sequence $g\colon \mathbb{N} \to \mathbb{R}$ a *generalised monomial* if it can be expressed using only polynomials, fractional part, and multiplication, e.g.

$$p_1(n)\{p_2(n)\}, \; p_1(n)\{p_2(n)\}\{p_3(n)\}, \; p_1(n)\{p_2(n)\}\{p_3(n)\{p_4(n)\}\},$$

where $p_i\colon \mathbb{N} \to \mathbb{R}$ are (ordinary) polynomial sequences.

Theorem (Leibman)

Each bounded GP sequence g has a "canonical representation"

$$g(n) = F\left(h_0(n), h_1(n), h_2(n), \dots, h_d(n)\right), \qquad n \in \mathbb{N},$$

where F is a piecewise polynomial function, h_0 is periodic and $h_1, h_2, \dots, h_d\colon \mathbb{N} \to [0,1)$ are jointly equidistributed generalised monomials.

Remark: The representation is explicitly computable.

Corollary

Given (representations of) two bracket words $\mathbf{a}, \mathbf{b}$ over the same alphabet Σ, there is a procedure to determine if $\mathbf{a}$ and $\mathbf{b}$ are equal almost everywhere, i.e. if $\#\{n \leq N : a_n \neq b_n\}/N \to 0$ as $N \to \infty$.

Canonical representation

Example (Leibman)

Let $g(n) = \left\{\sqrt{2}n\left[\sqrt{3}n\right]\right\}$. Then:

$$g(n) = \left\{\sqrt{6}n^2 - \sqrt{2}n\left\{\sqrt{3}n\right\}\right\}$$
$$= \left\{\left\{\sqrt{6}n^2\right\} - \left\{\sqrt{2}n\left\{\sqrt{3}n\right\}\right\}\right\} = F\left(\left\{\sqrt{6}n^2\right\}, \left\{\sqrt{2}n\left\{\sqrt{3}n\right\}\right\}\right),$$

where F is the piecewise polynomial function given by

$$F(x,y) = \{x-y\} = \begin{cases} x-y & \text{if } x \geq y, \\ x-y+1 & \text{if } x < y. \end{cases}$$

Example (Leibman)

Let $g(n) = \left\{\sqrt{2}n\left[\sqrt{3}n\right]\right\}$. Then:

$$g(n) = \left\{\sqrt{6}n^2 - \sqrt{2}n\left\{\sqrt{3}n\right\}\right\}$$
$$= \left\{\left\{\sqrt{6}n^2\right\} - \left\{\sqrt{2}n\left\{\sqrt{3}n\right\}\right\}\right\} = F\left(\left\{\sqrt{6}n^2\right\}, \left\{\sqrt{2}n\left\{\sqrt{3}n\right\}\right\}\right),$$

where F is the piecewise polynomial function given by

$$F(x,y) = \{x-y\} = \begin{cases} x-y & \text{if } x \geq y, \\ x-y+1 & \text{if } x < y. \end{cases}$$

Construction: Start by taking all polynomials p_i that appear in the representation of g, and iteratively construct expressions involving p_i, $\{\bullet\}$ and $\times$. It is possible to express g using these generalised monomials.

1. We have to be careful which basic sequences to include, e.g. if we add $p(n)\{q(n)\}$ then we cannot add $q(n)\{p(n)\}$ because of identities like
$$xy - x\{y\} - y\{x\} - \{x\}\{y\} \equiv [x][y] = 0 \bmod 1.$$
2. We might need to replace p_i with p_i/M $(M \in \mathbb{N})$.
3. We might need to pass to an arithmetic progression.

Equality is undecidable

Theorem (Adamczewski, K.)

There is no algorithm which, given a representation of a GP sequence g *with algebraic coefficients, determines if* $g(n) = 0$ *for all* $n \in \mathbb{N}$.

Corollary

There is no algorithm which, given representations of two bracket words $\mathbf{a}, \mathbf{b}$, *each involving only algebraic coefficients, determines if* $\mathbf{a} = \mathbf{b}$.

Equality is undecidable

Theorem (Adamczewski, K.)

There is no algorithm which, given a representation of a GP sequence g with algebraic coefficients, determines if $g(n) = 0$ for all $n \in \mathbb{N}$.

Corollary

There is no algorithm which, given representations of two bracket words $\mathbf{a}, \mathbf{b}$, *each involving only algebraic coefficients, determines if* $\mathbf{a} = \mathbf{b}$.

Sketch of proof:
Key fact: There exists a surjective GP map $\mathbb{N} \to \mathbb{N}^2$, for instance:

$$n \mapsto \left(\left[n \cdot \{\sqrt{2}n\}^{10} \right], \left[n \cdot \{\sqrt{3}n\}^{10} \right] \right).$$

Iterating, we can construct a surjective GP map $\iota_d \colon \mathbb{N} \to \mathbb{Z}^d$ for each $d \in \mathbb{N}$. If we could recognise if a GP sequence is identically zero, then we could also recognise solvable polynomial equations in $\mathbb{Z}$:

$$(\exists x_1, \dots, x_d \in \mathbb{Z})\ p(x_1, \dots, x_d) = 0 \Leftrightarrow (\exists n \in \mathbb{N})\ 1_{\{0\}} \left(p \circ \iota_d(n)\right) \neq 0$$

But it is well-known that this is impossible (cf. Hilbert's 10th problem).

Let $\beta > 1$ be an algebraic integer with minimal polynomial

$$p(X) = X^d - \sum_{j=0}^{d-1} a_j X_j.$$

Let $E = \{n_i : i \geq 0\}$, where n_i satisfy recurrence:

$$n_{i+d} = \sum_{j=0}^{d-1} a_j n_{i+j}$$

Let $\beta > 1$ be an algebraic integer with minimal polynomial

$$p(X) = X^d - \sum_{j=0}^{d-1} a_j X_j.$$

Let $E = \{n_i : i \geq 0\}$, where n_i satisfy recurrence:

$$n_{i+d} = \sum_{j=0}^{d-1} a_j n_{i+j}$$

Theorem

The set E is GP if β is any of the following:

Let $\beta > 1$ be an algebraic integer with minimal polynomial

$$p(X) = X^d - \sum_{j=0}^{d-1} a_j X_j.$$

Let $E = \{n_i : i \geq 0\}$, where n_i satisfy recurrence:

$$n_{i+d} = \sum_{j=0}^{d-1} a_j n_{i+j}$$

Theorem

The set E is GP if β is any of the following:

- *Quadratic Pisot unit ($d = 2$ and $a_0 = \pm 1$);*

Let $\beta > 1$ be an algebraic integer with minimal polynomial

$$p(X) = X^d - \sum_{j=0}^{d-1} a_j X_j.$$

Let $E = \{n_i : i \geq 0\}$, where n_i satisfy recurrence:

$$n_{i+d} = \sum_{j=0}^{d-1} a_j n_{i+j}$$

Theorem

The set E is GP if β is any of the following:

- *Quadratic Pisot unit ($d = 2$ and $a_0 = \pm 1$);*
- *Not totally real cubic Pisot unit ($d = 3$, $a_0 = 1$, p has two complex roots);*

Let $\beta > 1$ be an algebraic integer with minimal polynomial

$$p(X) = X^d - \sum_{j=0}^{d-1} a_j X_j.$$

Let $E = \{n_i : i \geq 0\}$, where n_i satisfy recurrence:

$$n_{i+d} = \sum_{j=0}^{d-1} a_j n_{i+j}$$

Theorem

The set E is GP if β is any of the following:

- *Quadratic Pisot unit ($d = 2$ and $a_0 = \pm 1$);*
- *Not totally real cubic Pisot unit ($d = 3$, $a_0 = 1$, p has two complex roots);*
- *Salem unit ($d \geq 4$ even, $a_0 = 1$, $p(1/\beta) = 0$, other roots on the unit circle).*

Let $\beta > 1$ be an algebraic integer with minimal polynomial

$$p(X) = X^d - \sum_{j=0}^{d-1} a_j X_j .$$

Let $E = \{n_i : i \geq 0\}$, where n_i satisfy recurrence:

$$n_{i+d} = \sum_{j=0}^{d-1} a_j n_{i+j}$$

Theorem

The set E is GP if β is any of the following:

- *Quadratic Pisot unit ($d = 2$ and $a_0 = \pm 1$);*
- *Not totally real cubic Pisot unit ($d = 3$, $a_0 = 1$, p has two complex roots);*
- *Salem unit ($d \geq 4$ even, $a_0 = 1$, $p(1/\beta) = 0$, other roots on the unit circle).*

Conjecture: These are the only cases where E is GP.

Extremely sparse sequences

Proposition (Byszewski, K.)

Suppose that $(n_i)_{i=0}^{\infty}$ *is a sequence with* $\liminf_{i\to\infty} \frac{\log n_{i+1}}{\log n_i} > 1$.

Then, $E = \{n_i : i \geq 0\}$ *is a GP set.* *(E.g.* $n_i = 2^{2^i}$*.)*

Proposition (Byszewski, K.)

Suppose that $(n_i)_{i=0}^{\infty}$ *is a sequence with* $\liminf_{i\to\infty} \frac{\log n_{i+1}}{\log n_i} > 1$.

Then, $E = \{n_i : i \geq 0\}$ *is a GP set.* *(E.g.* $n_i = 2^{2^i}$*.)*

Proof strategy: Hope to find $\alpha \in \mathbb{R}$ and C so that $n \in E$ if and only if

$$\|\alpha n\|_{\mathbb{R}/\mathbb{Z}} < 1/n^C \qquad (\dagger)$$

By inductive construction, one can show that there exists a Cantor set of α such that $(\dagger)$ holds for all $n \in E$.

Proposition (Byszewski, K.)

Suppose that $(n_i)_{i=0}^{\infty}$ *is a sequence with* $\liminf_{i\to\infty} \frac{\log n_{i+1}}{\log n_i} > 1$.

Then, $E = \{n_i : i \geq 0\}$ *is a GP set.* (*E.g.* $n_i = 2^{2^i}$.)

Proof strategy: Hope to find $\alpha \in \mathbb{R}$ and C so that $n \in E$ if and only if

$$\|\alpha n\|_{\mathbb{R}/\mathbb{Z}} < 1/n^C \qquad (\dagger)$$

By inductive construction, one can show that there exists a Cantor set of α such that $(\dagger)$ holds for all $n \in E$.

Problem: $(\dagger)$ might hold for some $n \in \mathbb{N} \setminus E$. First, reduce to the case where $D < \log n_{i+1}/\log n_i < 2D$ for a constant D. We strengthen $(\dagger)$ to

$$1/2n^C < \|\alpha n\|_{\mathbb{R}/\mathbb{Z}} < 1/n^C \qquad (\ddagger)$$

Under suitable conditions on C and D (e.g. $C = 5$, $D = 6$), we use continued fractions to check that no spurious n satisfy $(\ddagger)$.

Negative results: Criteria

How does one prove that a set (or a sequence) is *not* GP?

Negative results: Criteria

How does one prove that a set (or a sequence) is *not* GP?

- (Boring answer) Use properties from previous slides.

How does one prove that a set (or a sequence) is *not* GP?

- (Boring answer) Use properties from previous slides.

Theorem (Byszewski, K.)

If $E \subset \mathbb{N}$, $d(E) = 0$ and E contains an IP set, then E is not GP.

Example: The set $\left\{\sum_{n \in I} 2^{2^n} : I \subset \mathbb{N}\right\}$ is not GP. The set of integers whose base-10 expansions do not contain 7 is not GP.

How does one prove that a set (or a sequence) is *not* GP?

- (Boring answer) Use properties from previous slides.

Theorem (Byszewski, K.)

If $E \subset \mathbb{N}$, $d(E) = 0$ and E contains an IP set, then E is not GP.

Example: The set $\left\{\sum_{n\in I} 2^{2^n} : I \subset \mathbb{N}\right\}$ is not GP. The set of integers whose base-10 expansions do not contain 7 is not GP.

Theorem (K.)

Fix $k \in \mathbb{N}$. If $E \subset \mathbb{N}$, $d(E) = 0$ and $E = E/k := \{n \in \mathbb{N} : kn \in E\}$ then E is not GP.

Example: $\{k^n : n \geq 0\}$ is not GP.

Negative results: Examples

Negative results: Examples

- The squares $S = \{n^2 : n \in \mathbb{N}\}$ are not a GP set. (Proof: $S/4 = S$.)

Negative results: Examples

- The squares $S = \{n^2 : n \in \mathbb{N}\}$ are not a GP set. (Proof: $S/4 = S$.)
- The primes are not a GP set. (Proof: use Gowers uniformity of μ.)

- The squares $S = \{n^2 : n \in \mathbb{N}\}$ are not a GP set. (Proof: $S/4 = S$.)

- The primes are not a GP set. (Proof: use Gowers uniformity of μ.)

- Let $q \geq 3$. Then $(\varphi(n) \bmod q)_{n=0}^{\infty}$ is not bracket word.
  (Here, $\varphi(n) = \#(\mathbb{Z}/n\mathbb{Z})^{\times}$ is the totient function.)
  Proof: The set $E = \{n : \varphi(n) \not\equiv 0 \bmod q\}$ has $d(E) = 0$ and $E/p = E$ for any $p \in \mathcal{P}$ with $p > q$, $q \nmid p - 1$.

- The squares $S = \{n^2 : n \in \mathbb{N}\}$ are not a GP set. (Proof: $S/4 = S$.)
- The primes are not a GP set. (Proof: use Gowers uniformity of μ.)
- Let $q \geq 3$. Then $(\varphi(n) \bmod q)_{n=0}^{\infty}$ is not bracket word.
  (Here, $\varphi(n) = \#(\mathbb{Z}/n\mathbb{Z})^{\times}$ is the totient function.)
  Proof: The set $E = \{n : \varphi(n) \not\equiv 0 \bmod q\}$ has $d(E) = 0$ and $E/p = E$ for any $p \in \mathcal{P}$ with $p > q$, $q \nmid p - 1$.
- Recall that a word $\mathbf{a} = (a_n)_{n=0}^{\infty}$ is k-automatic if a_n can be computed by a finite automaton, taking base-k expansion of n as input.

Theorem (Byszewski, K.)

If $\mathbf{a}$ *is* k*-automatic and not eventually periodic, then* $\mathbf{a}$ *is not a bracket word.*

The End

THANK YOU FOR YOUR ATTENTION!

Example

Let $\beta > 1$, $\alpha, \bar{\alpha}$ be the roots of $X^3 - X^2 - X - 1$. Then the set $\{\lfloor \beta^i \rceil : i \geq 0\}$ is GP. (NB $\lfloor \beta^{i+3} \rceil = \lfloor \beta^{i+2} \rceil + \lfloor \beta^{i+1} \rceil + \lfloor \beta^i \rceil$; $\lfloor x \rceil = \lfloor x + \frac{1}{2} \rfloor$.)

Bonus: Recursive sequences

Example

Let $\beta > 1$, $\alpha, \bar{\alpha}$ be the roots of $X^3 - X^2 - X - 1$. Then the set $\{\lfloor \beta^i \rceil \,:\, i \geq 0\}$ is GP. (NB $\lfloor \beta^{i+3} \rceil = \lfloor \beta^{i+2} \rceil + \lfloor \beta^{i+1} \rceil + \lfloor \beta^i \rceil$; $\lfloor x \rceil = \lfloor x + \frac{1}{2} \rfloor$.)

Sketch of proof: Note that $|\alpha| < 1$. If we guess that $n = \lfloor \beta^i \rceil$ then

$$\begin{aligned} n &= \beta^i + \alpha^i + \bar{\alpha}^i = \mathrm{Tr}(\beta^i) \\ \lfloor \beta n \rceil &= \beta^{i+1} + \alpha^{i+1} + \bar{\alpha}^{i+1} \\ \lfloor \beta^2 n \rceil &= \beta^{i+2} + \alpha^{i+2} + \bar{\alpha}^{i+2}. \end{aligned}$$

Example

Let $\beta > 1$, $\alpha, \bar{\alpha}$ be the roots of $X^3 - X^2 - X - 1$. Then the set $\{\lfloor \beta^i \rceil : i \geq 0\}$ is GP. (NB $\lfloor \beta^{i+3} \rceil = \lfloor \beta^{i+2} \rceil + \lfloor \beta^{i+1} \rceil + \lfloor \beta^i \rceil$; $\lfloor x \rceil = \lfloor x + \frac{1}{2} \rfloor$.)

Sketch of proof: Note that $|\alpha| < 1$. If we guess that $n = \lfloor \beta^i \rceil$ then

$$\begin{aligned} n &= g + h + \bar{h} \\ \lfloor \beta n \rceil &= \beta g + \alpha h + \bar{\alpha} \bar{h} \\ \lfloor \beta^2 n \rceil &= \beta^2 g + \alpha^2 h + \bar{\alpha}^2 \bar{h}. \end{aligned}$$

with $g = \beta^i$, $h = \alpha^i$. This defines GP sequences $g(n) \in \mathbb{Q}(\beta)$, $h(n) \in \mathbb{Q}(\alpha)$ *for all* n. New goal: $n = \lfloor \beta^i \rceil \iff g(n) = \beta^i$.

Example

Let $\beta > 1$, $\alpha, \bar{\alpha}$ be the roots of $X^3 - X^2 - X - 1$. Then the set $\{\lfloor \beta^i \rceil : i \geq 0\}$ is GP. (NB $\lfloor \beta^{i+3} \rceil = \lfloor \beta^{i+2} \rceil + \lfloor \beta^{i+1} \rceil + \lfloor \beta^i \rceil$; $\lfloor x \rceil = \lfloor x + \frac{1}{2} \rfloor$.)

Sketch of proof: Note that $|\alpha| < 1$. If we guess that $n = \lfloor \beta^i \rceil$ then

$$n = g + h + \bar{h}$$
$$\lfloor \beta n \rceil = \beta g + \alpha h + \bar{\alpha}\bar{h}$$
$$\lfloor \beta^2 n \rceil = \beta^2 g + \alpha^2 h + \bar{\alpha}^2 \bar{h}.$$

with $g = \beta^i$, $h = \alpha^i$. This defines GP sequences $g(n) \in \mathbb{Q}(\beta)$, $h(n) \in \mathbb{Q}(\alpha)$ *for all* n. New goal: $n = \lfloor \beta^i \rceil \iff g(n) = \beta^i$.

Key fact: The group of units of $\mathbb{Q}(\beta)$ has rank 1, β is a generator.

Example

Let $\beta > 1$, $\alpha, \bar{\alpha}$ be the roots of $X^3 - X^2 - X - 1$. Then the set $\{\lfloor \beta^i \rceil : i \geq 0\}$ is GP. (NB $\lfloor \beta^{i+3} \rceil = \lfloor \beta^{i+2} \rceil + \lfloor \beta^{i+1} \rceil + \lfloor \beta^i \rceil$; $\lfloor x \rceil = \lfloor x + \frac{1}{2} \rfloor$.)

Sketch of proof: Note that $|\alpha| < 1$. If we guess that $n = \lfloor \beta^i \rceil$ then

$$\begin{aligned} n &= g + h + \bar{h} \\ \lfloor \beta n \rceil &= \beta g + \alpha h + \bar{\alpha} \bar{h} \\ \lfloor \beta^2 n \rceil &= \beta^2 g + \alpha^2 h + \bar{\alpha}^2 \bar{h}. \end{aligned}$$

with $g = \beta^i$, $h = \alpha^i$. This defines GP sequences $g(n) \in \mathbb{Q}(\beta)$, $h(n) \in \mathbb{Q}(\alpha)$ *for all* n. New goal: $n = \lfloor \beta^i \rceil \iff g(n) = \beta^i$.

Key fact: The group of units of $\mathbb{Q}(\beta)$ has rank 1, β is a generator.

$$g(n) = \beta^i \iff g(n) \text{ is a unit in } \mathbb{Q}(\beta).$$

Bonus: Recursive sequences

Example

Let $\beta > 1$, $\alpha, \bar{\alpha}$ be the roots of $X^3 - X^2 - X - 1$. Then the set $\{\lfloor \beta^i \rceil : i \geq 0\}$ is GP. (NB $\lfloor \beta^{i+3} \rceil = \lfloor \beta^{i+2} \rceil + \lfloor \beta^{i+1} \rceil + \lfloor \beta^i \rceil$; $\lfloor x \rceil = \lfloor x + \frac{1}{2} \rfloor$.)

Sketch of proof: Note that $|\alpha| < 1$. If we guess that $n = \lfloor \beta^i \rceil$ then

$$\begin{aligned} n &= g + h + \bar{h} \\ \lfloor \beta n \rceil &= \beta g + \alpha h + \bar{\alpha}\bar{h} \\ \lfloor \beta^2 n \rceil &= \beta^2 g + \alpha^2 h + \bar{\alpha}^2 \bar{h}. \end{aligned}$$

with $g = \beta^i$, $h = \alpha^i$. This defines GP sequences $g(n) \in \mathbb{Q}(\beta)$, $h(n) \in \mathbb{Q}(\alpha)$ *for all* n. New goal: $n = \lfloor \beta^i \rceil \iff g(n) = \beta^i$.

Key fact: The group of units of $\mathbb{Q}(\beta)$ has rank 1, β is a generator.

$$g(n) = \beta^i \iff \begin{cases} g(n) \text{ is an algebraic integer} \\ \mathrm{N}\big(g(n)\big) = 1 \iff g(n)h(n)\bar{h}(n) = 1. \end{cases}$$