Balance in billiard words and cohomology

OWCOW – One Worℓd seminar on Combinatorics of Words

Antoine Julien
jt. work with N. Bédaride & V. Berthé

Nord University
Levanger, Norway

October 10^{th}, 2025

Our results

Theorem

Sequences arising as coding of hypercubic billiards are not balanced on all factors. (for cubes of dimension at least 3)

Theorem

Cubic (dimension 3) billiard sequences are not balanced on factors of length 2.

Our results

Theorem

Sequences arising as coding of hypercubic billiards are not balanced on all factors. (for cubes of dimension at least 3)

Theorem

Cubic (dimension 3) billiard sequences are not balanced on factors of length 2.

- To be defined: billiard sequences; subshifts; balance.

Our results

Theorem

Sequences arising as coding of hypercubic billiards are not balanced on all factors. (for cubes of dimension at least 3)

Theorem

Cubic (dimension 3) billiard sequences are not balanced on factors of length 2.

- To be defined: billiard sequences; subshifts; balance.
- Assumptions: irrational direction.

Our results

Theorem

Sequences arising as coding of hypercubic billiards are not balanced on all factors. (for cubes of dimension at least 3)

Theorem

Cubic (dimension 3) billiard sequences are not balanced on factors of length 2.

- To be defined: billiard sequences; subshifts; balance.
- Assumptions: irrational direction.
- A topological surprise: cohomology.

The plan

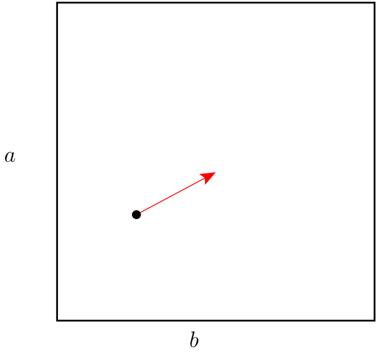
1. Define billiard sequences and subshifts;
2. Define balance;
3. *(show the results again)*
4. Define cohomology for subshifts;
5. "What does cohomology got to do with it?": some ideas of the proof.

Billiard sequences

A well known example: Sturmian sequences.

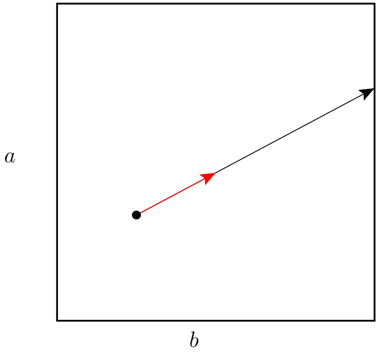
Billiard sequences

A well known example: Sturmian sequences.



Billiard sequences

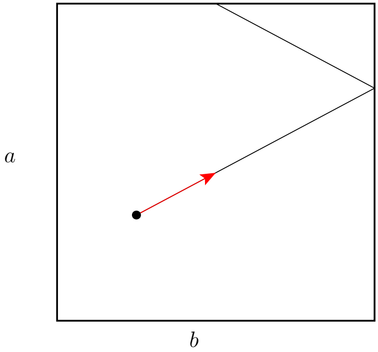
A well known example: Sturmian sequences.



a

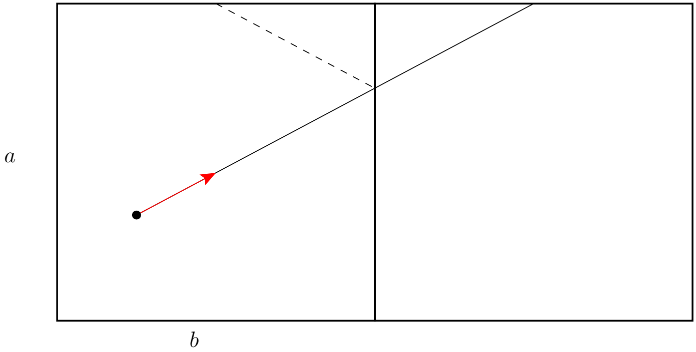
Billiard sequences

A well known example: Sturmian sequences.



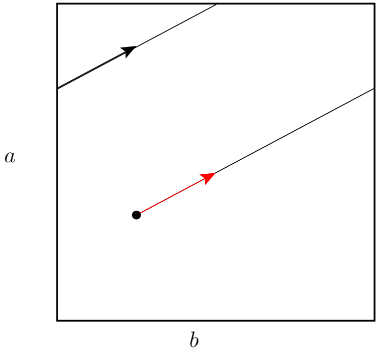
ab

A well known example: Sturmian sequences.



ab

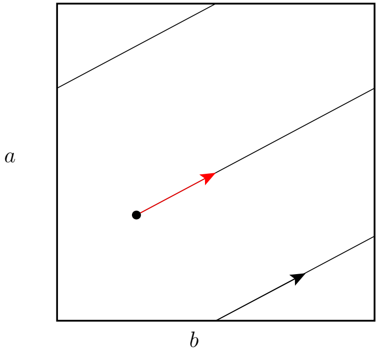
A well known example: Sturmian sequences.



ab

Billiard sequences

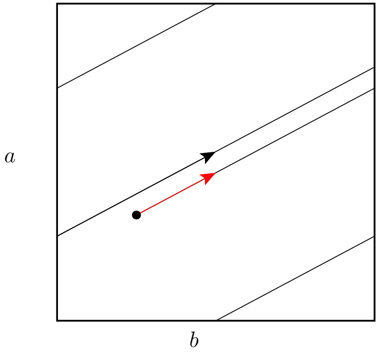
A well known example: Sturmian sequences.



aba

Billiard sequences

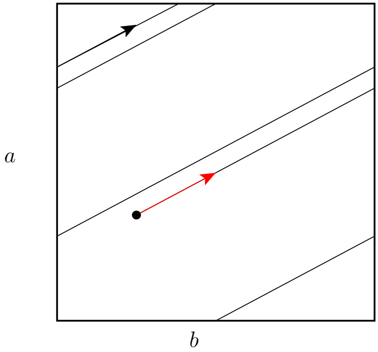
A well known example: Sturmian sequences.



$abaa$

Billiard sequences

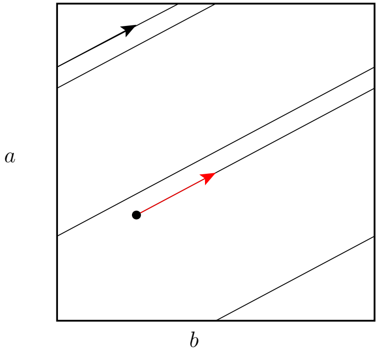
A well known example: Sturmian sequences.



abaab

Billiard sequences

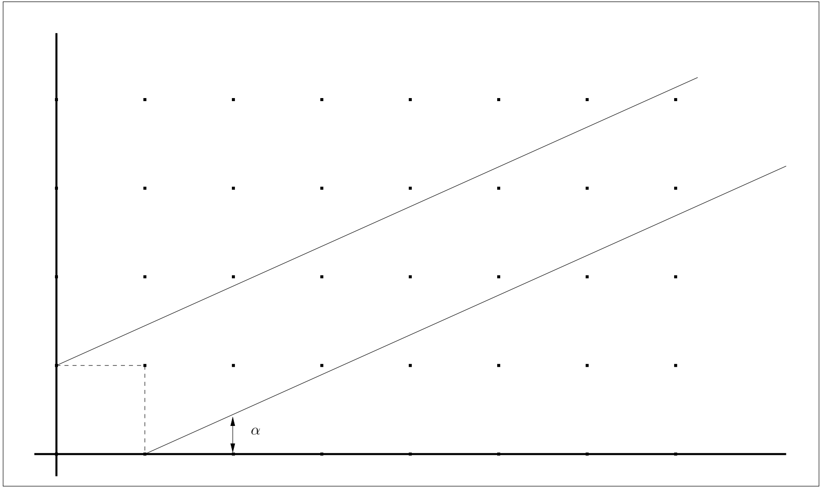
A well known example: Sturmian sequences.



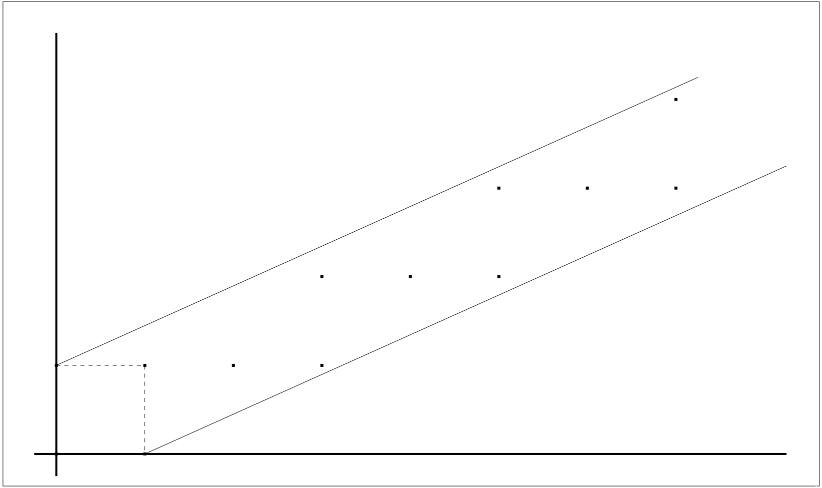
abaab

Interesting when the slope is irrational.

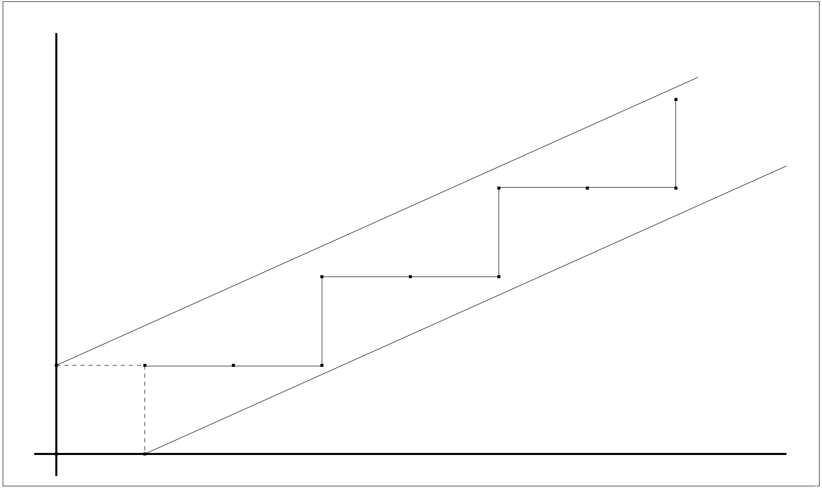
Billiard sequences as cut-and-project



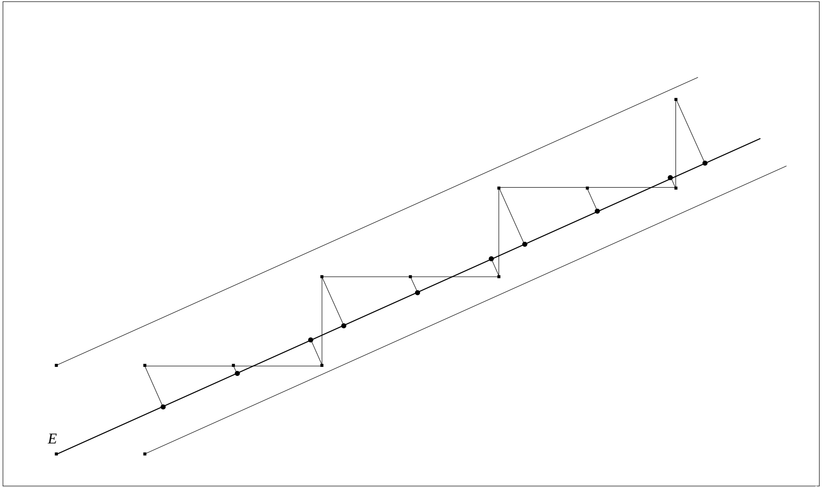
Billiard sequences as cut-and-project



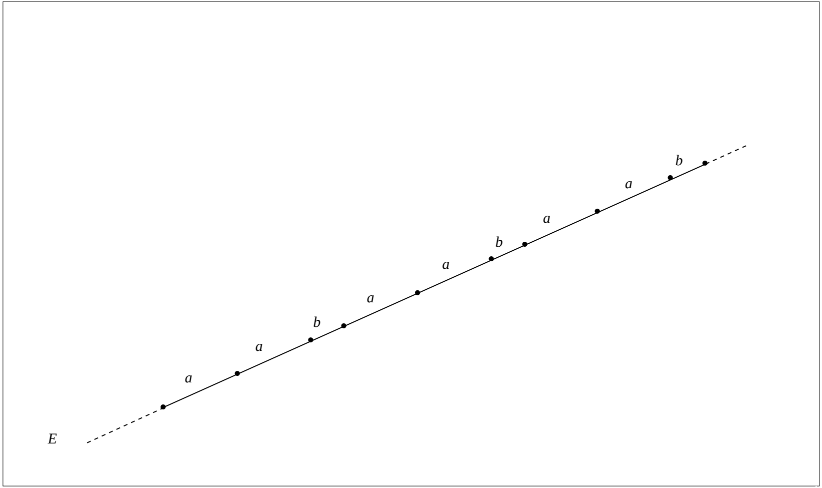
Billiard sequences as cut-and-project



Billiard sequences as cut-and-project



Billiard sequences as cut-and-project



The object of our results

Our results deal with:

- billiard sequences in higher dimension
- still a line trajectory in a $(d+1)$-dimensional cube
- code the trajectory by the d-faces it meets

The object of our results

Our results deal with:

- billiard sequences in higher dimension
- still a line trajectory in a $(d+1)$-dimensional cube
- code the trajectory by the d-faces it meets

How we choose to see it:

- Cut-and-project sequences
- $\mathbb{R}^{d+1} = L_\theta \oplus L_\theta^\perp$
- Which points of $(\mathbb{Z}^{d+1} + m)$ land in a strip $[0;1]^{d+1} + L_\theta$?

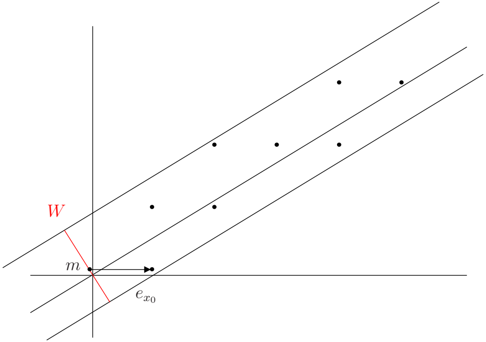
Assume: the trajectory is irrational, meaning $\theta = (1, \theta_1, \ldots, \theta_d)$ are rationally independents.

Why it's useful

The cut-and-project framing can be useful.

Why it's useful

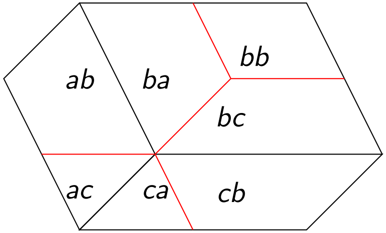
The cut-and-project framing can be useful.



Use $(\mathbb{Z}^2 + m)$ instead of $\mathbb{Z}^2$.
Parameter m: starting point of the billiard trajectory.
Sufficient to take m in a window $W = \pi^{\perp}([0;1]^{d+1})$.

Why it's useful, II

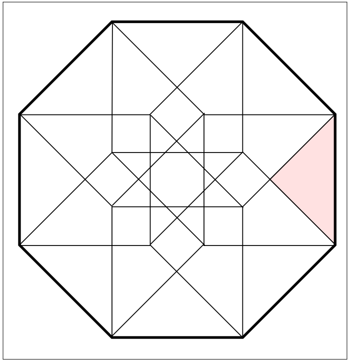
A good grip on the window gives insight on the language of the sequence.
Below, for a cubic billiard.

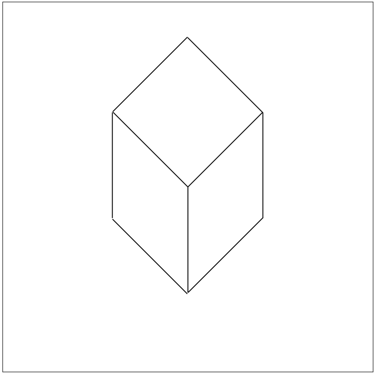


Partition of the window encoding the first two letters of the billiard trajectory.

Window and patterns

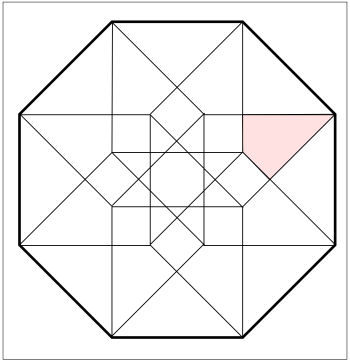
A good grip on the window gives insight on the language of the sequence.
Here illustrated in higher dimension.

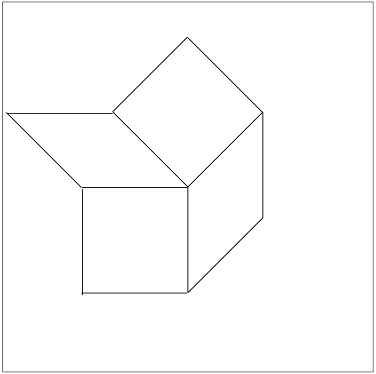




Window and patterns

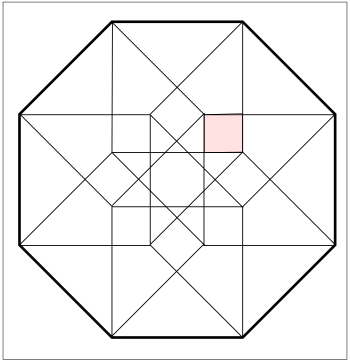
A good grip on the window gives insight on the language of the sequence.
Here illustrated in higher dimension.

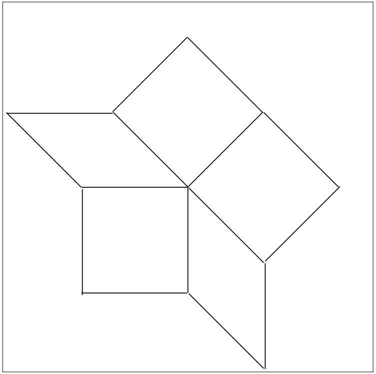




Window and patterns

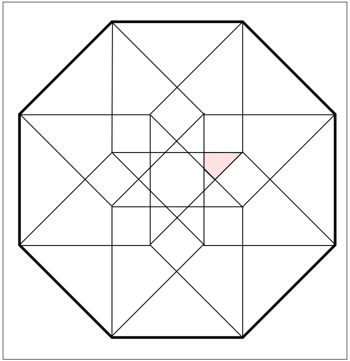
A good grip on the window gives insight on the language of the sequence.
Here illustrated in higher dimension.

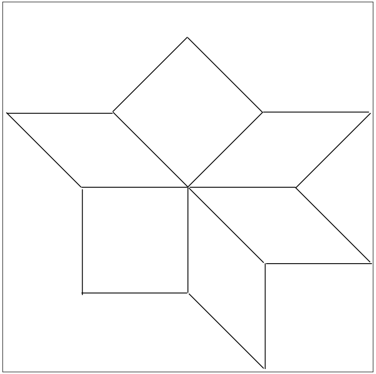




Window and patterns

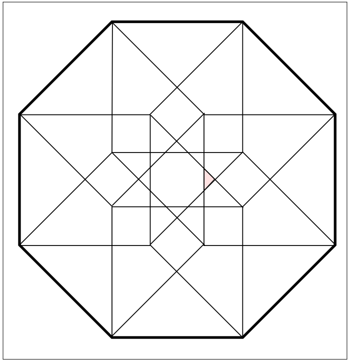
A good grip on the window gives insight on the language of the sequence.
Here illustrated in higher dimension.

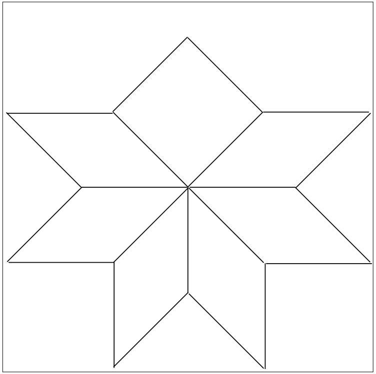




Window and patterns

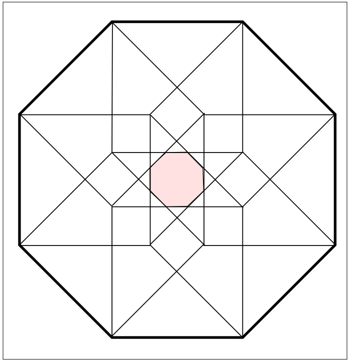
A good grip on the window gives insight on the language of the sequence.
Here illustrated in higher dimension.

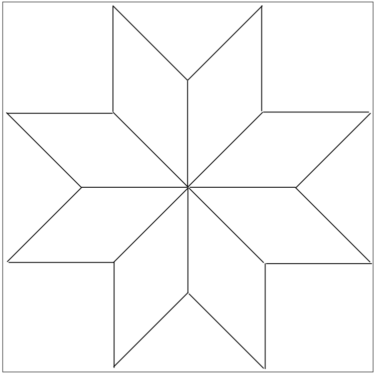




Window and patterns

A good grip on the window gives insight on the language of the sequence.
Here illustrated in higher dimension.





Frequency and balance

Given a word $\cdots x_{-1}x_0x_1\cdots \in \mathcal{A}^{\mathbb{Z}}$,
The frequency of the letter $a \in \mathcal{A}$ is

$$\mu_a = \frac{1}{n} \cdot \lim_{n\to+\infty} \Big(\text{number of } a \text{ in } x_0\cdots x_{n-1}\Big)$$

if it exists. For billiard sequences, it exists.

Frequency and balance

Given a word $\cdots x_{-1}x_0x_1\cdots \in \mathcal{A}^{\mathbb{Z}}$,
The frequency of the letter $a \in \mathcal{A}$ is

$$\mu_a = \frac{1}{n} \cdot \lim_{n\to+\infty} \Big(\text{number of } a \text{ in } x_0\cdots x_{n-1}\Big)$$

if it exists. For billiard sequences, it exists.
How fast does it converge? If

$$\big|(\text{number of } a \text{ in } x_0 \ldots x_{n-1}) - n\cdot\mu_a\big| < C$$

independently of n, we say that the word is balanced on a.

Frequency and balance

Given a word $\cdots x_{-1}x_0x_1\cdots \in \mathcal{A}^{\mathbb{Z}}$,
The frequency of the letter $a \in \mathcal{A}$ is

$$\mu_a = \frac{1}{n} \cdot \lim_{n\to+\infty} \left(\text{number of } a \text{ in } x_0 \cdots x_{n-1}\right)$$

if it exists. For billiard sequences, it exists.
How fast does it converge? If

$$\left|(\text{number of } a \text{ in } x_0 \ldots x_{n-1}) - n \cdot \mu_a\right| < C$$

independently of n, we say that the word is balanced on a.
Can define balance for any letter or finite factor.

Our results (again!)

Theorem

Sequences arising as coding of hypercubic billiards are not balanced on all factors. (for cubes of dimension at least $d+1=3$)

Theorem

Cubic (dimension 3) billiard sequences are not balanced on factors of length 2.

Our results (again!)

Theorem

Sequences arising as coding of hypercubic billiards are not balanced on all factors. (for cubes of dimension at least $d+1=3$)

Theorem

Cubic (dimension 3) billiard sequences are not balanced on factors of length 2.

- Given the direction of the billiard trajectory, the specific word is irrelevant. (words are repetitive / subshifts are minimal).

Cohomology

Our results rely on linking balance property and cohomology.

Cohomology

Our results rely on linking balance property and cohomology.

1. What is cohomology in this context?
2. What is the link?
3. How does it help?

Cohomology for a graph

Consider a cell-complex (in dimension 1: a graph).

- Vertices
- Oriented edges
- A boundary function: if e goes from v_0 to v_1, then $\partial e = v_1 - v_0$.

A cocycle is a function defined on edges with values (for example) in $\mathbb{R}$.

Cohomology for a graph

Consider a cell-complex (in dimension 1: a graph).

- Vertices
- Oriented edges
- A boundary function: if e goes from v_0 to v_1, then $\partial e = v_1 - v_0$.

A cocycle is a function defined on edges with values (for example) in $\mathbb{R}$.

- A cocycle φ pairs with paths $(e_1, \ldots, e_n)$: $\varphi(e_1, \ldots, e_n) = \sum_i \varphi(e_i)$.
- A cocycle is a coboundary if it derives from a potential: $\varphi(e) = b(e^+) - b(e^-)$ for a function b defined on vertices.

Silly example

Consider the following graph decomposition of a line.

$$\cdots \qquad \bullet \longrightarrow \bullet \longrightarrow \bullet \longrightarrow \bullet \longrightarrow \bullet \longrightarrow \bullet \qquad \cdots$$

Silly example

Consider the following graph decomposition of a line.

$$\cdots \qquad \bullet \longrightarrow \bullet \longrightarrow \bullet \longrightarrow \bullet \longrightarrow \bullet \longrightarrow \bullet \qquad \cdots$$

All cocycles derive from a potential.
The cohomology group $H^1 = \{\text{cocycles}\}/\{\text{coboundaries}\}$ is trivial.

Less silly example

Consider the following graph decomposition of a line associated with an infinite word x.

$$\cdots \qquad \bullet \xrightarrow[x_{-1}]{a} \bullet \xrightarrow[x_0]{b} \bullet \xrightarrow[x_1]{a} \bullet \xrightarrow[x_2]{a} \bullet \xrightarrow[x_3]{b} \bullet \qquad \cdots$$

Less silly example

Consider the following graph decomposition of a line associated with an infinite word x.

$$\cdots \quad \bullet \xrightarrow[x_{-1}]{\text{a}} \bullet \xrightarrow[x_0]{\text{b}} \bullet \xrightarrow[x_1]{\text{a}} \bullet \xrightarrow[x_2]{\text{a}} \bullet \xrightarrow[x_3]{\text{b}} \bullet \quad \cdots$$

Additional constraint:

- Not all cocycles are accepted.
- $\varphi(x_i)$ should only depend on a finite number of past and future coordinates.
- Acceptable coboundaries (potential functions on vertices) are defined similarly.

Example: $\varphi(x_i) = 1$ if the letter is a, 0 otherwise.

Less silly example

Consider the following graph decomposition of a line associated with an infinite word x.

$$\cdots \quad \bullet \xrightarrow[x_{-1}]{a} \bullet \xrightarrow[x_0]{b} \bullet \xrightarrow[x_1]{a} \bullet \xrightarrow[x_2]{a} \bullet \xrightarrow[x_3]{b} \bullet \quad \cdots$$

Additional constraint:

- Not all cocycles are accepted.
- $\varphi(x_i)$ should only depend on a finite number of past and future coordinates.
- Acceptable coboundaries (potential functions on vertices) are defined similarly.

Example: $\varphi(x_i) = 1$ if the letter is a, 0 otherwise.

The result: pattern-equivariant cohomology (or Cech cohomology of the subshift).
Also called strong cohomology.

An old example of using cohomology

$$\cdots \qquad \bullet \xrightarrow[x_{-1}]{a} \bullet \xrightarrow[x_0]{b} \bullet \xrightarrow[x_1]{a} \bullet \xrightarrow[x_2]{a} \bullet \xrightarrow[x_3]{b} \bullet \qquad \cdots$$

Given such a cocycle with positive values, it can be used to define the length of the "tiles".

Simplest example: one length for *a*, one length for *b*.

Can be used to define a suspension of the subshift and study flow equivalence.

See:

- Parry & Sullivan, *A topological invariant of flows on 1-dimensional spaces*, 1975.
- Parry & Tuncel, *Classification problems in ergodic theory*, 1982.

What does it have to do with balance?

How does cohomology help studying balance properties?

What does it have to do with balance?

How does cohomology help studying balance properties?

$$\cdots \qquad \bullet \xrightarrow[1]{\color{red}{a}} \bullet \xrightarrow[0]{\color{red}{b}} \bullet \xrightarrow[0]{\color{red}{a}} \bullet \xrightarrow[1]{\color{red}{a}} \bullet \xrightarrow[0]{\color{red}{b}} \bullet \qquad \cdots$$

Letter-counting functions are cocycles. So are pattern-counting functions. Above, the counting function φ_{ab}.

What does it have to do with balance?

How does cohomology help studying balance properties?

$$\cdots \qquad \bullet \xrightarrow[1]{a} \bullet \xrightarrow[0]{b} \bullet \xrightarrow[0]{a} \bullet \xrightarrow[1]{a} \bullet \xrightarrow[0]{b} \bullet \qquad \cdots$$

Letter-counting functions are cocycles. So are pattern-counting functions. Above, the counting function φ_{ab}.
The word x is balanced for the factor ab if:

$$|\varphi_{ab}(x_0 x_1 \cdots x_{n-1}) - \mu_{ab}\mathbb{1}(x_0 \cdots x_{n-1})|$$

is bounded uniformly in n.
($\mathbb{1}$ is the cocycle which is 1 on all letters).

Balance and cohomology

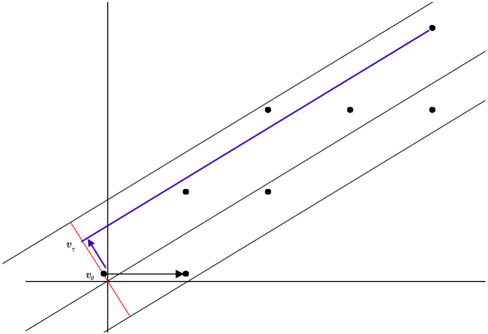
A word x is balanced on the factor w if there exists C such that

$$|\varphi_w(x_0 x_1 \cdots x_{n-1}) - \mu_w \mathbb{1}(x_0 \cdots x_{n-1})| < C$$

Gottshalk-Hedlund's theorem (assuming minimality): this is equivalent to $\varphi_w - \mu_w \mathbb{1}$ being a weak coboundary.

(Meaning: this cocycle derives from a potential; the potential function does not only depend on finitely many past and future coordinates, but it can be approximated by such functions)

Balance on letters



Sturmian sequences are balanced on letters: it means

$$\varphi_a(x_0 x_1 \cdots x_{n-1}) - \mu_a \mathbb{1}(x_0 \cdots x_{n-1})$$

is a difference of potential $b(v_0) - b(v_n)$.

Sturmian, continued

The dimension of H^1 for Sturmian sequences is 2: generated by the letter-counting function $\mathbb{1}$ and the a-discrepancy function $\varphi_a - \mu_a \mathbb{1}$.
Any other factor-discrepancy function is a combination of these two (up to a coboundary).
By a frequency argument, one shows that it is a multiple of $\varphi_a - \mu_a \mathbb{1}$ (up to a coboundary).
Therefore, Sturmian sequences are balanced on all factors.

In higher dimensions

For higher-dimensional billiards,

1. The cohomology group H^1 is generated by the factor-counting functions φ_w;
2. The dimension of the cohomology group H^1 is infinite (Forest–Hunton–Kellendonk);
   (there can be a lot of independent discrepancy functions)
3. The subspace of H^1 generated by weak coboundaries is of finite dimension (Kellendonk–Sadun);
4. Not all discrepancy functions can be bounded: billiard sequences can't be balanced on all factors.