# Probabilistic studies in Number Theory and Word Combinatorics: instances of dynamical analysis

Pablo Rotondo

IRIF, Paris 7 Diderot,

Universidad de la República, Uruguay

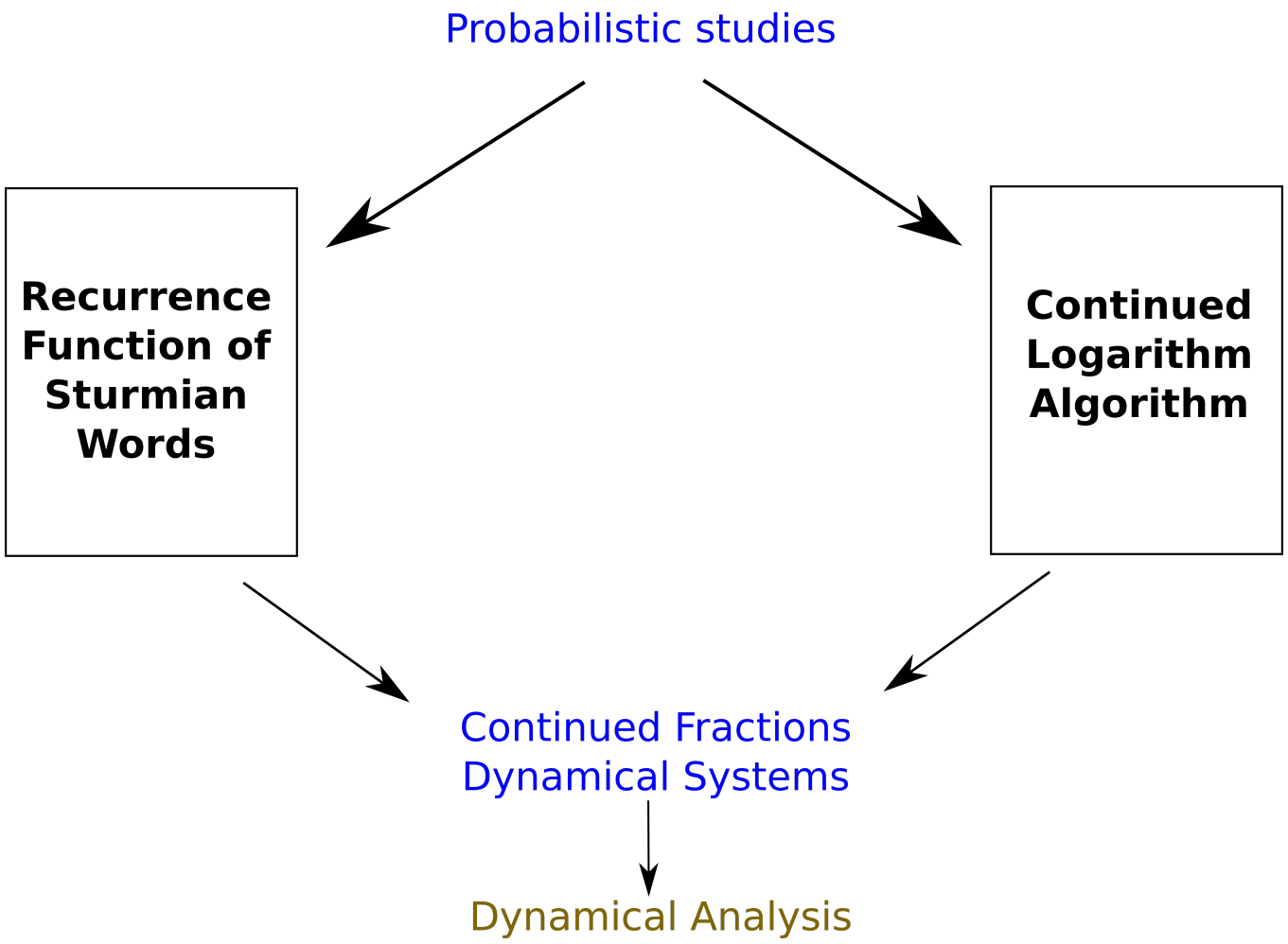
GREYC, associate

Joint works with

Valérie Berthé, Eda Cesaratto, Brigitte Vallée, Alfredo Viola

**Seminaire de Combinatoire**, **IRIF**, May 24, 2018.

# Some key words



# This talk

1. General Introduction: continued fractions and dynamical systems
   - Continued Fractions
   - Euclidean dynamical system
2. The recurrence function of a random Sturmian word
   - Sturmian words and recurrence
   - Our models and results
3. The Continued Logarithm
   - Origins and algorithm
   - The CL dynamical system
   - Extended system and results
   - Conclusions and extensions

# Section

# Continued Fractions

Every irrational number $\alpha \in (0,1)$ has a unique representation

$$\alpha = \cfrac{1}{m_1 + \cfrac{1}{m_2 + \ddots}}$$

where $m_1, m_2, \ldots \geq 1$ are integers called the digits or quotients.

# Continued Fractions

Every irrational number $\alpha \in (0,1)$ has a unique representation

$$\alpha = \cfrac{1}{m_1 + \cfrac{1}{m_2 + \ddots}}$$

where $m_1, m_2, \ldots \geq 1$ are integers called the digits or quotients.

Truncating the expansion at depth $k$ we get a convergent

$$\frac{p_k(\alpha)}{q_k(\alpha)} = \cfrac{1}{m_1 + \cfrac{1}{m_2 + \ddots \frac{1}{m_k}}}.$$

The denominators $q_k(\alpha)$ are called the continuants of $\alpha$.

# Euclidean Algorithm and Continued Fractions

### Property

Given integers $x$ and $y$ with $x \geq y \geq 0$

$$\gcd(x, y) = \gcd(y, x \bmod y)\,.$$

In conjunction with $\gcd(x, 0) = x$, we get the Euclidean Algorithm.

# Euclidean Algorithm and Continued Fractions

### Property

Given integers $x$ and $y$ with $x \geq y \geq 0$

$$\gcd(x, y) = \gcd(y, x \bmod y)\,.$$

In conjunction with $\gcd(x, 0) = x$, we get the Euclidean Algorithm.

This algorithm is equivalent to the continued fraction expansion:

- given the integer division $y = mx + r$,

$$\frac{x}{y} = \frac{1}{m + \frac{r}{x}}\,,$$

  and the process continues with $\frac{r}{x}$.

# Euclidean Algorithm and Continued Fractions

### Property

Given integers $x$ and $y$ with $x \geq y \geq 0$

$$\gcd(x, y) = \gcd(y, x \bmod y)\,.$$

In conjunction with $\gcd(x, 0) = x$, we get the Euclidean Algorithm.

This algorithm is equivalent to the continued fraction expansion:

- given the integer division $y = mx + r$,

$$\frac{x}{y} = \frac{1}{m + \frac{r}{x}}\,,$$

  and the process continues with $\frac{r}{x}$.

# Euclidean dynamical system

To get the digits of the continued fraction expansion observe

$$\alpha = \cfrac{1}{m_1 + \cfrac{1}{m_2 + \ddots}}$$

$$\Longrightarrow m_1 = \left\lfloor \frac{1}{\alpha} \right\rfloor , \qquad \cfrac{1}{m_2 + \cfrac{1}{m_3 + \ddots}} = \left\{ \frac{1}{\alpha} \right\} .$$

The map

$$T \colon (0,1) \to (0,1) , \qquad x \mapsto \left\{ \frac{1}{x} \right\} ,$$

is known as the Gauss map.

# Euclidean dynamical system

To get the digits of the continued fraction expansion observe

$$\alpha = \cfrac{1}{m_1 + \cfrac{1}{m_2 + \ddots}}$$

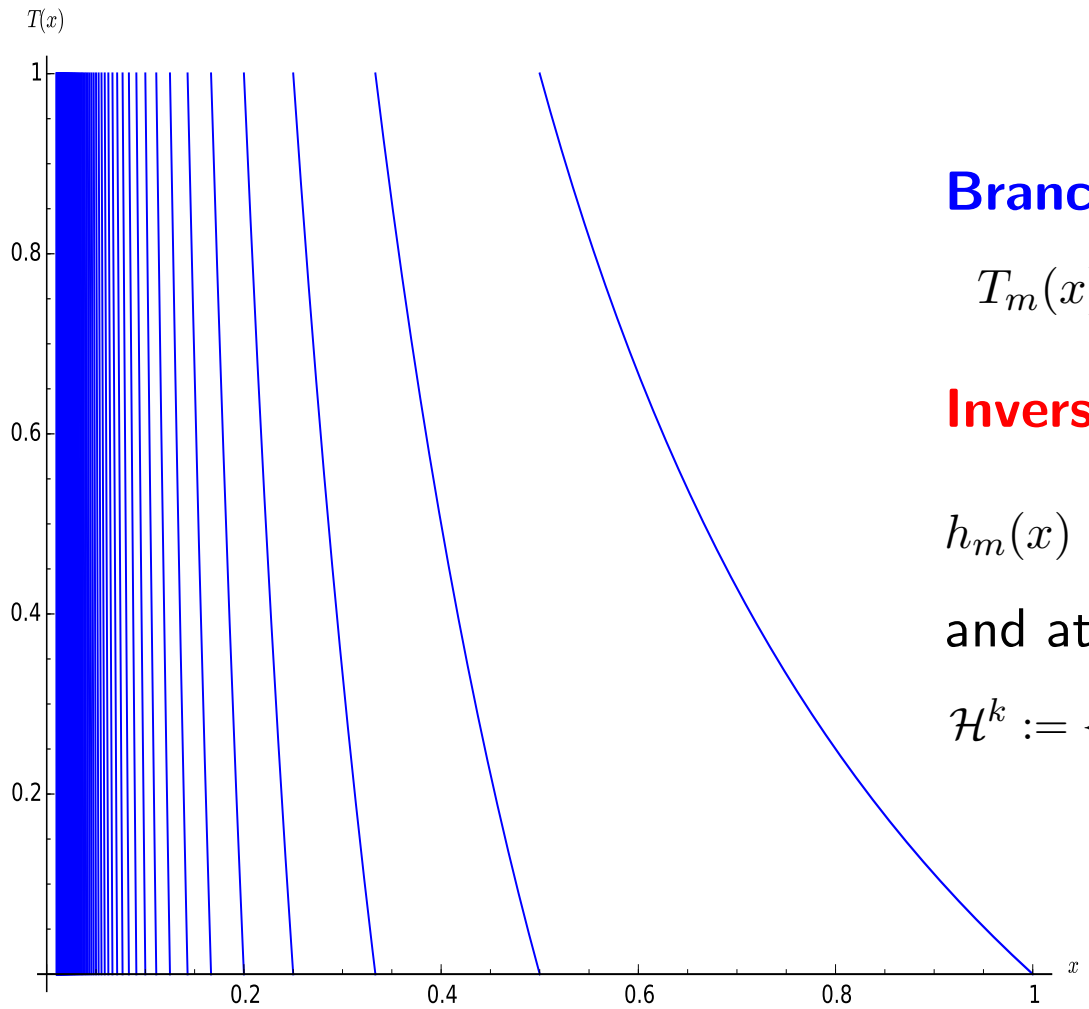
$$\Longrightarrow m_1 = \left\lfloor \frac{1}{\alpha} \right\rfloor , \qquad \cfrac{1}{m_2 + \cfrac{1}{m_3 + \ddots}} = \left\{ \frac{1}{\alpha} \right\} .$$

The map

$$T \colon (0,1) \to (0,1) , \qquad x \mapsto \left\{ \frac{1}{x} \right\} ,$$

is known as the Gauss map.

# Gauss map




**Branches**

$$T_m(x) := \frac{1}{x} - m\,, \quad x \in \left(\frac{1}{m+1}, \frac{1}{m+0}\right].$$

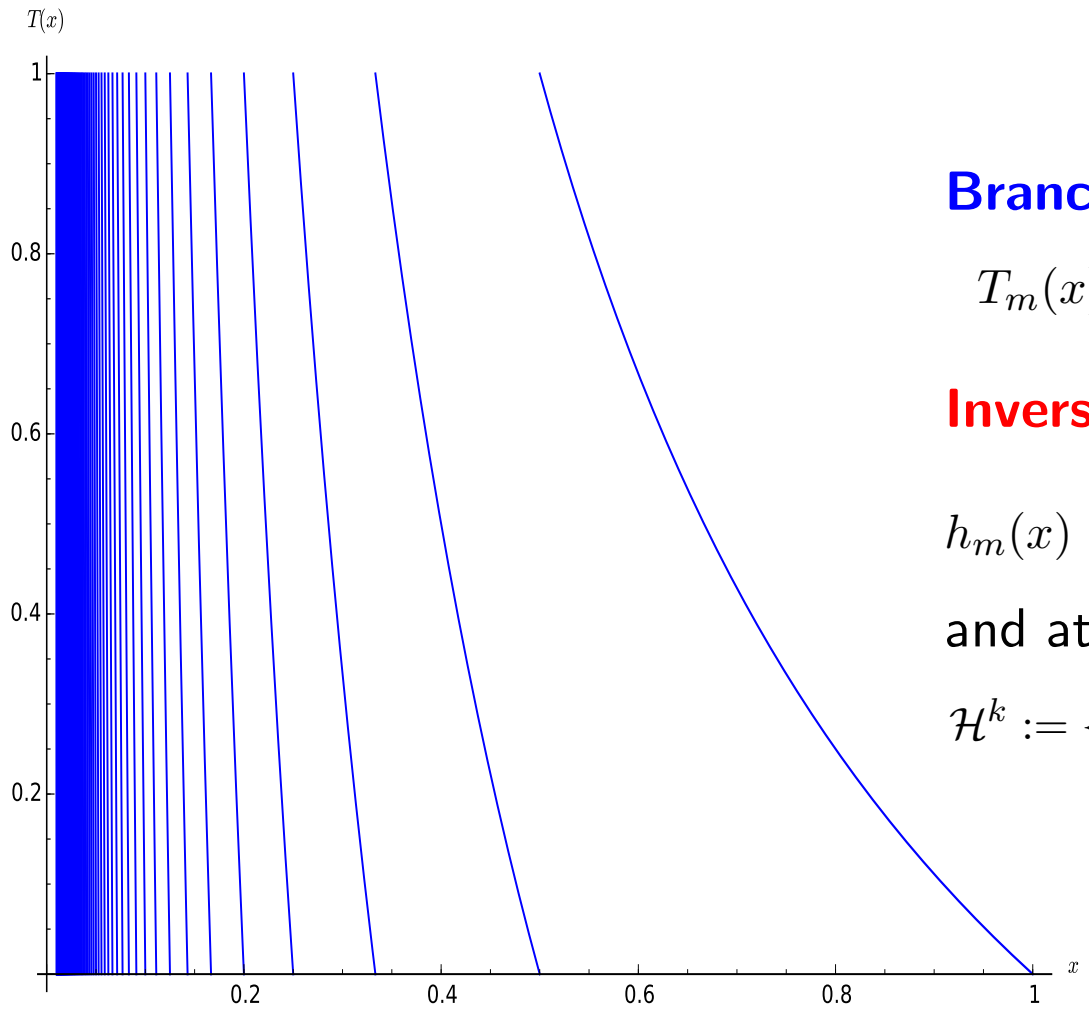
**Inverse branches**

$$h_m(x) := \frac{1}{m+x}\,, \quad \mathcal{H} := \{h_m : m \in \mathbb{N}\}\,,$$

and at depth $k$

$$\mathcal{H}^k := \{h_{m_1} \circ \cdots \circ h_{m_k} : m_1, \ldots, m_k \in \mathbb{N}\}\,.$$

# Gauss map




**Branches**

$$T_m(x) := \frac{1}{x} - m\,, \quad x \in \left(\frac{1}{m+1}, \frac{1}{m+0}\right].$$

**Inverse branches**

$$h_m(x) := \frac{1}{m+x}\,, \quad \mathcal{H} := \{h_m : m \in \mathbb{N}\}\,,$$

and at depth $k$

$$\mathcal{H}^k := \{h_{m_1} \circ \cdots \circ h_{m_k} : m_1, \ldots, m_k \in \mathbb{N}\}\,.$$

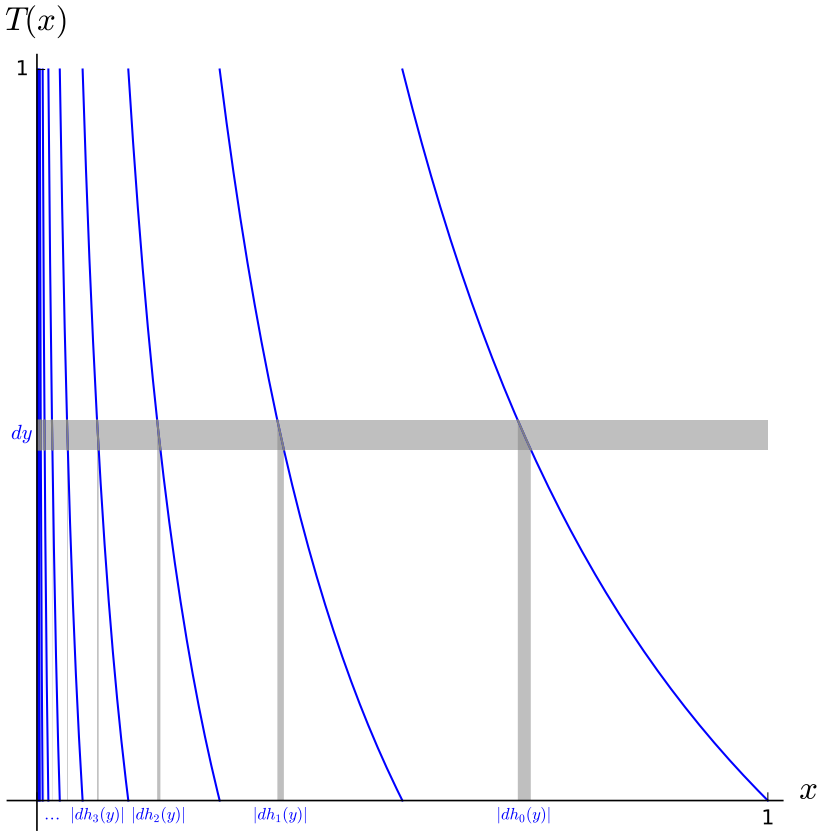
**Property.** Let $h := h_{m_1} \circ \cdots h_{m_k}$, if $h(0) = \frac{p}{q} \Longrightarrow |h'(0)| = \frac{1}{q^2}$ .

# Density transformer

Question: If $g \in \mathcal{C}^0(\mathcal{I})$ were the density of $x \Longrightarrow$ density of $T(x)$?

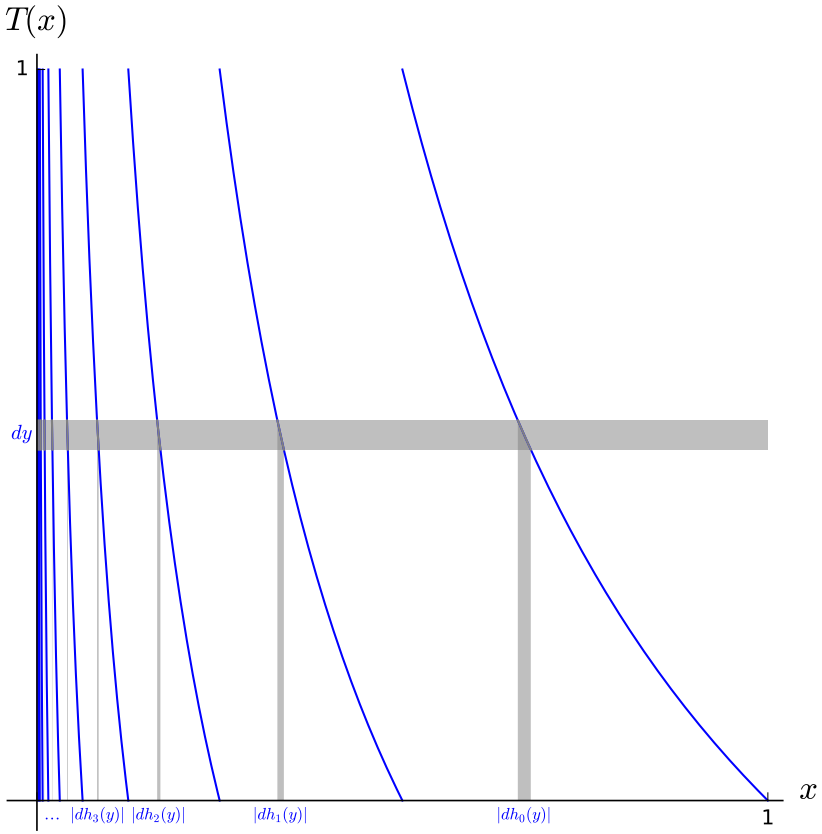
# Density transformer

Question: If $g \in \mathcal{C}^0(\mathcal{I})$ were the density of $x \implies$ density of $T(x)$?



# Density transformer

Question: If $g \in \mathcal{C}^0(\mathcal{I})$ were the density of $x \Longrightarrow$ density of $T(x)$?



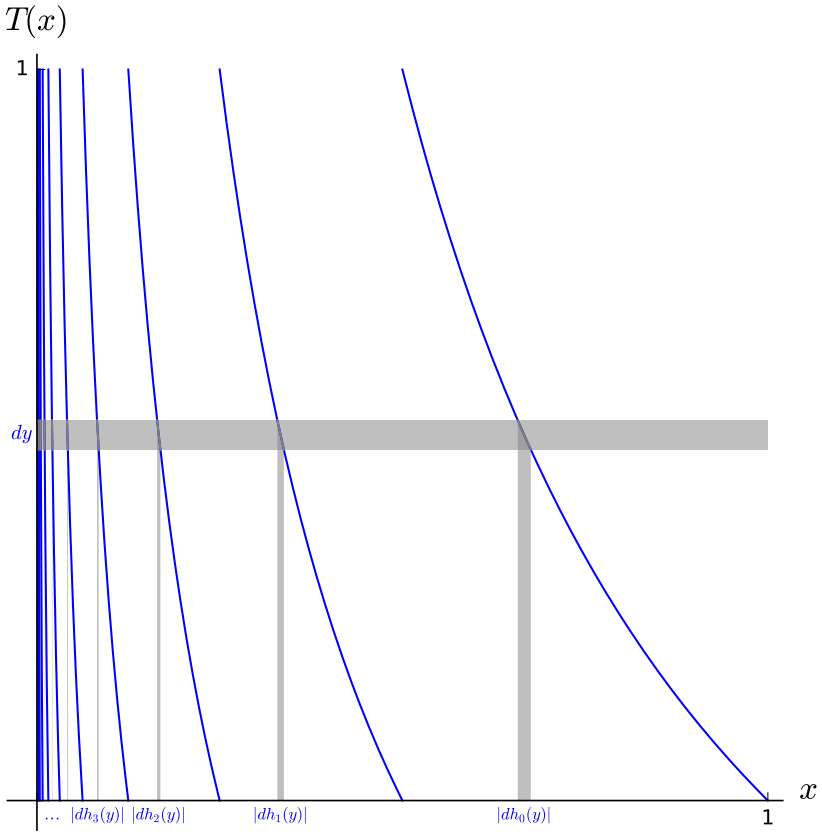

Answer: The density is

$$\mathbf{H}[g](x) = \sum_{h \in \mathcal{H}} \left|h'(x)\right| \, g\left(h(x)\right)$$

$$= \sum_{m \geq 0} \frac{1}{(m+x)^2} \, g\left(\frac{1}{m+x}\right) .$$

# Density transformer

Question: If $g \in \mathcal{C}^0(\mathcal{I})$ were the density of $x \Longrightarrow$ density of $T(x)$?




Answer: The density is

$$\mathbf{H}[g](x) = \sum_{h \in \mathcal{H}} \left|h'(x)\right| \, g\left(h(x)\right)$$

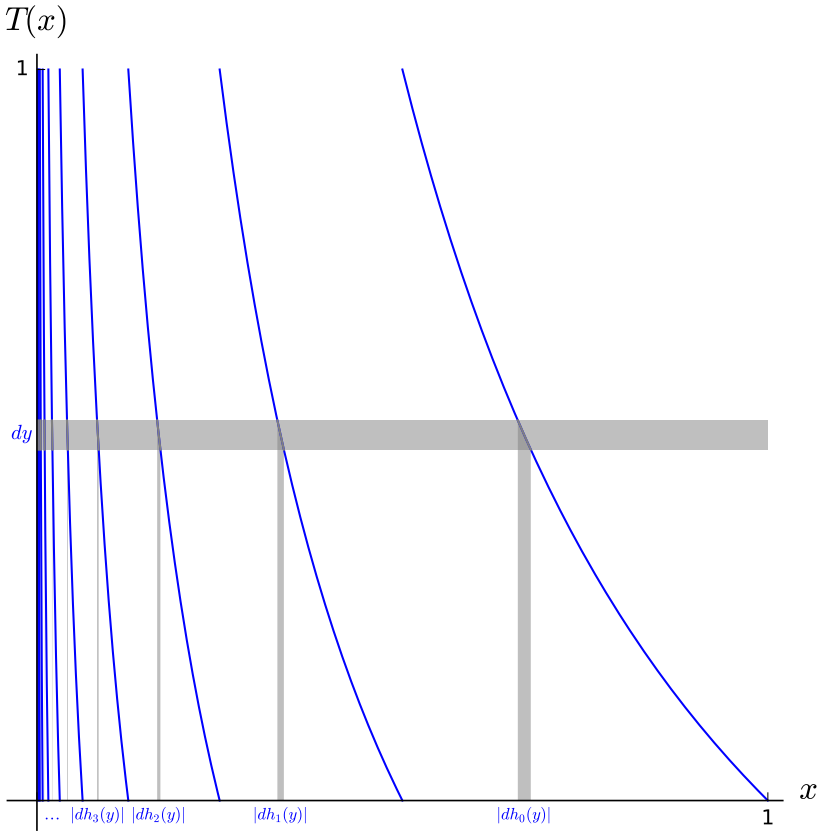
$$= \sum_{m \geq 0} \frac{1}{(m+x)^2} \, g\left(\frac{1}{m+x}\right) .$$

In general $T^k(x)$ has density

$$\mathbf{H}^k[g](x) = \sum_{h \in \mathcal{H}^k} \left|h'(x)\right| \, g\left(h(x)\right) .$$

# Density transformer

Question: If $g \in \mathcal{C}^0(\mathcal{I})$ were the density of $x \Longrightarrow$ density of $T(x)$?




Answer: The density is

$$\mathbf{H}[g](x) = \sum_{h \in \mathcal{H}} \left|h'(x)\right| \, g\left(h(x)\right)$$

$$= \sum_{m \geq 0} \frac{1}{(m+x)^2} \, g\left(\frac{1}{m+x}\right) .$$

In general $T^k(x)$ has density

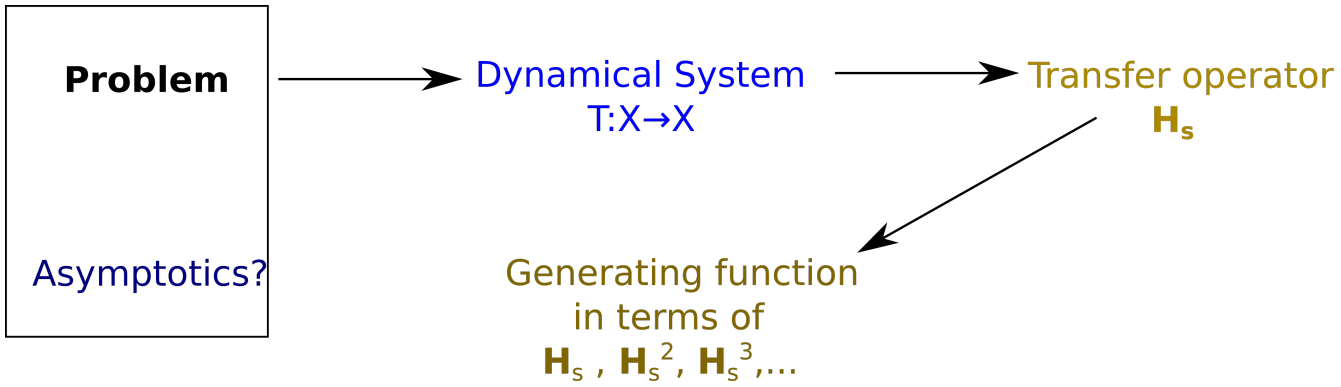
$$\mathbf{H}^k[g](x) = \sum_{h \in \mathcal{H}^k} \left|h'(x)\right| \, g\left(h(x)\right) .$$

$\Longrightarrow$ Transfer operator $\mathbf{H}_s$ extends $\mathbf{H}$, introducing a variable $s$

$$\mathbf{H}_s[g](x) = \sum_{h \in \mathcal{H}} \left|h'(x)\right|^s \, g\left(h(x)\right) .$$

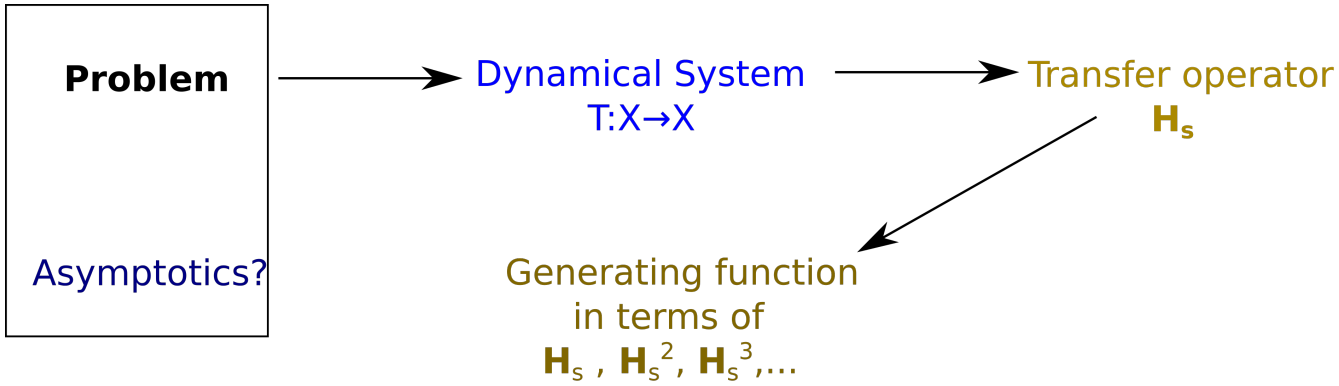
**Principles of dynamical analysis** [Vallée,Flajolet,Baladi,...]:

- Transfer operator $\mathbf{H}_s \Rightarrow$ expressions for generating functions.



## Principles of dynamical analysis [Vallée,Flajolet,Baladi,...]:

- Transfer operator $\mathbf{H}_s \Rightarrow$ expressions for generating functions.

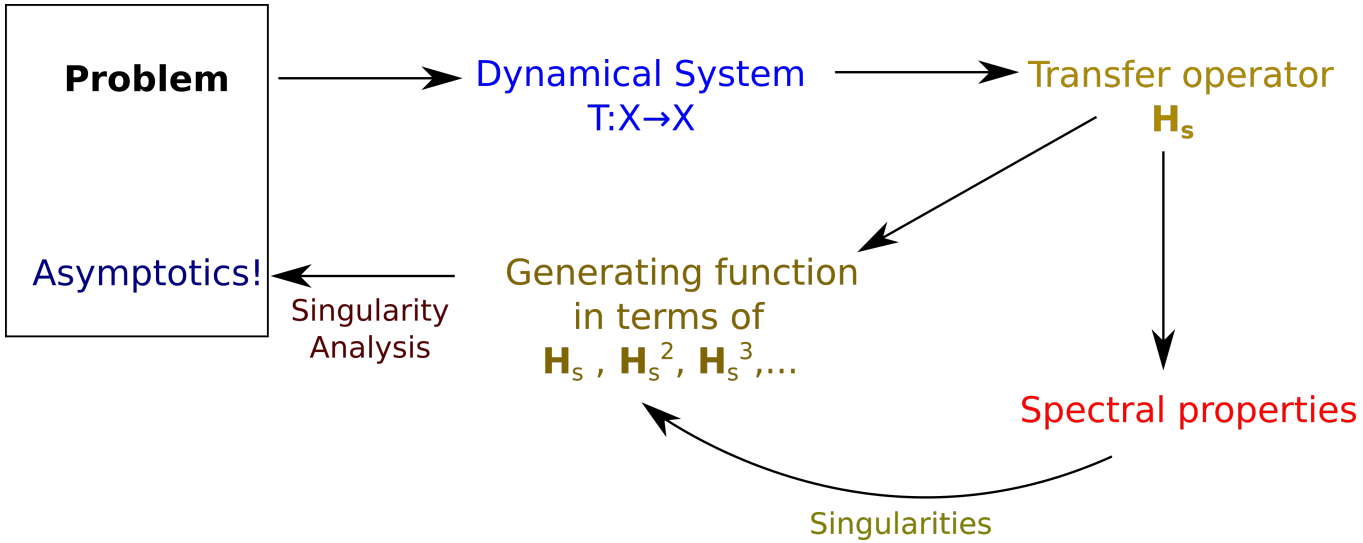



## Note

- $\mathbf{H}_s$ describes all execution of depth $1$.
- $\mathbf{H}_s^2 = \mathbf{H}_s \circ \mathbf{H}_s$ describes all execution of depth $2$.
- $\vdots$
- and $(\mathbf{I} - \mathbf{H}_s)^{-1} = \mathbf{I} + \mathbf{H}_s + \mathbf{H}_s^2 + \ldots$ describes *all* executions.

## Principles of dynamical analysis [Vallée,Flajolet,Baladi,...]:

- Transfer operator $\mathbf{H}_s \Rightarrow$ expressions for generating functions.
- Dominant eigenvalue of $\mathbf{H}_s \Rightarrow$ dominant singularities.




## Note

- $\mathbf{H}_s$ describes all execution of depth $1$.
- $\mathbf{H}_s^2 = \mathbf{H}_s \circ \mathbf{H}_s$ describes all execution of depth $2$.
- $\vdots$
- and $(\mathbf{I} - \mathbf{H}_s)^{-1} = \mathbf{I} + \mathbf{H}_s + \mathbf{H}_s^2 + \ldots$ describes *all* executions.

# Section

# Definition of Sturmian words

## Definition

Complexity function of an infinite word $\boldsymbol{u} \in \mathcal{A}^{\mathbb{N}}$

$$p_{\boldsymbol{u}} \colon \mathbb{N} \to \mathbb{N}\,, \qquad p_{\boldsymbol{u}}(n) = \#\{\text{factors of length } n \text{ in } \boldsymbol{u}\}\,.$$

# Definition of Sturmian words

## Definition

Complexity function of an infinite word $\boldsymbol{u} \in \mathcal{A}^{\mathbb{N}}$

$$p_{\boldsymbol{u}} \colon \mathbb{N} \to \mathbb{N}\,, \qquad p_{\boldsymbol{u}}(n) = \#\{\text{factors of length } n \text{ in } \boldsymbol{u}\}\,.$$

# Definition of Sturmian words

## Definition

Complexity function of an infinite word $\boldsymbol{u} \in \mathcal{A}^{\mathbb{N}}$

$$p_{\boldsymbol{u}} \colon \mathbb{N} \to \mathbb{N}\,, \qquad p_{\boldsymbol{u}}(n) = \#\{\text{factors of length } n \text{ in } \boldsymbol{u}\}\,.$$

## Important property

$\boldsymbol{u} \in \mathcal{A}^{\mathbb{N}}$ is not eventually periodic

$\iff p_{\boldsymbol{u}}(n+1) > p_{\boldsymbol{u}}(n)$ for all $n \in \mathbb{N}$

.

# Definition of Sturmian words

## Definition

Complexity function of an infinite word $\boldsymbol{u} \in \mathcal{A}^{\mathbb{N}}$

$$p_{\boldsymbol{u}} \colon \mathbb{N} \to \mathbb{N}\,, \qquad p_{\boldsymbol{u}}(n) = \#\{\text{factors of length } n \text{ in } \boldsymbol{u}\}\,.$$

## Important property

$$\begin{aligned} &\boldsymbol{u} \in \mathcal{A}^{\mathbb{N}} \text{ is not eventually periodic} \\ &\iff p_{\boldsymbol{u}}(n+1) > p_{\boldsymbol{u}}(n) \text{ for all } n \in \mathbb{N} \\ &\implies p_{\boldsymbol{u}}(n) \geq n+1\,. \end{aligned}$$

Sturmian words are the "simplest" that are not eventually periodic.

# Definition of Sturmian words

## Definition

Complexity function of an infinite word $\boldsymbol{u} \in \mathcal{A}^{\mathbb{N}}$

$$p_{\boldsymbol{u}} \colon \mathbb{N} \to \mathbb{N}\,, \qquad p_{\boldsymbol{u}}(n) = \#\{\text{factors of length } n \text{ in } \boldsymbol{u}\}\,.$$

## Important property

$$\begin{aligned} &\boldsymbol{u} \in \mathcal{A}^{\mathbb{N}} \text{ is not eventually periodic} \\ &\iff p_{\boldsymbol{u}}(n+1) > p_{\boldsymbol{u}}(n) \text{ for all } n \in \mathbb{N} \\ &\implies p_{\boldsymbol{u}}(n) \geq n+1\,. \end{aligned}$$

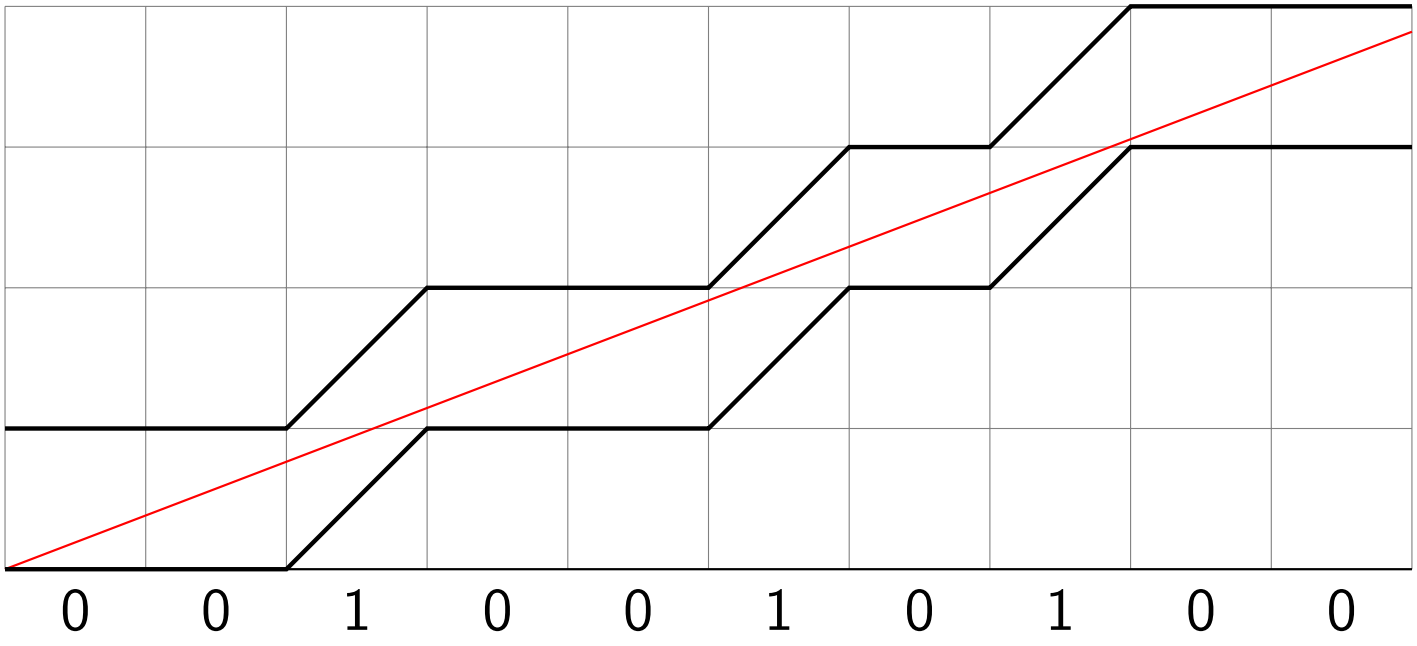
Sturmian words are the "simplest" that are not eventually periodic.

## Definition

$\boldsymbol{u} \in \{0,1\}^{\mathbb{N}}$ is Sturmian $\iff p_{\boldsymbol{u}}(n) = n+1$ for each $n \geq 0$.

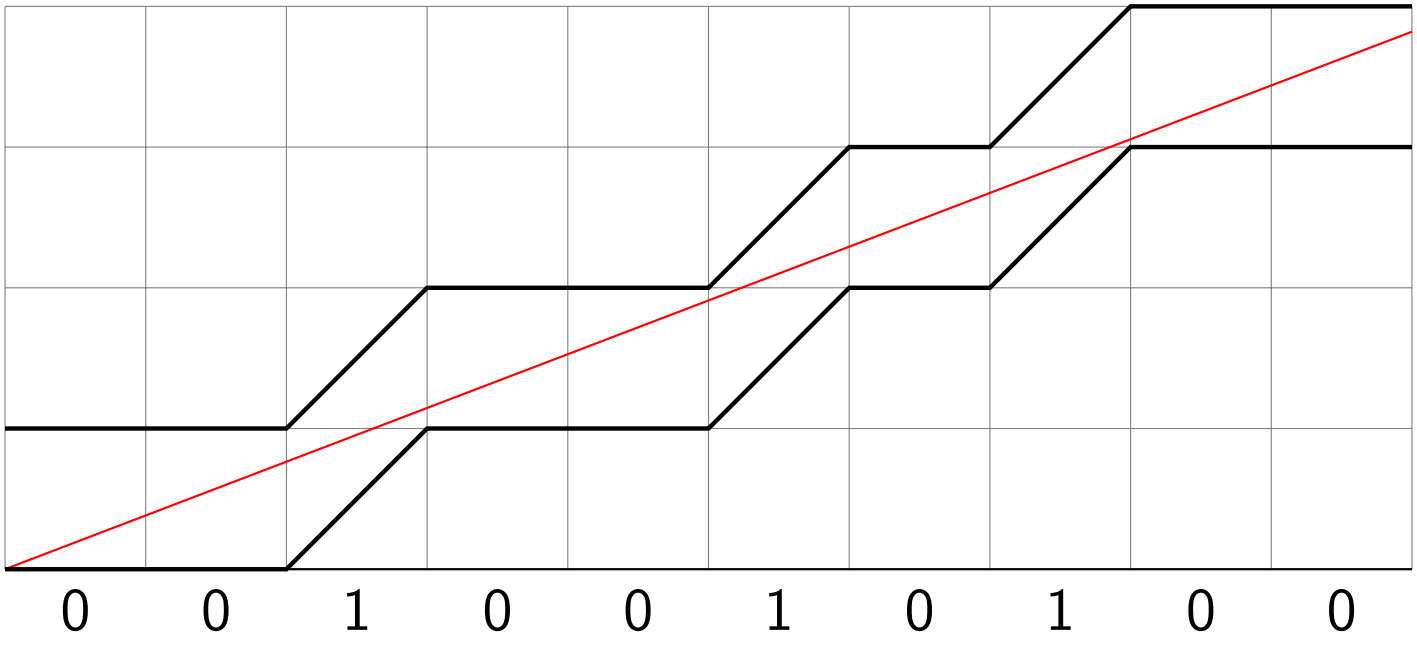
# Sturmian words and digital lines

Sturmian words correspond to discrete codings of lines, from below or above, by horizontal lines and diagonals.



# Sturmian words and digital lines

Sturmian words correspond to discrete codings of lines, from below or above, by horizontal lines and diagonals.




The *slope* $\alpha$ plays a key role:

the finite factors are determined exclusively by $\alpha$.

# Recurrence of Sturmian words

## Definition (Recurrence function)

Consider an infinite word $\boldsymbol{u}$. Its recurrence function is:

$$R_{\boldsymbol{u}}(n) = \inf \{m \in \mathbb{N} : \text{every factor of length } m \\ \textbf{contains} \text{ all the factors of length } n\}.$$

# Recurrence of Sturmian words

## Definition (Recurrence function)

Consider an infinite word $\boldsymbol{u}$. Its recurrence function is:

$$R_{\boldsymbol{u}}(n) = \inf \{m \in \mathbb{N} : \text{every factor of length } m \text{ } \textbf{contains} \text{ all the factors of length } n\}.$$

- Cost we have to pay to discover the factors if we start from an arbitrary point in $\boldsymbol{u} = u_1 u_2 \ldots$

# Recurrence of Sturmian words

## Definition (Recurrence function)

Consider an infinite word $\boldsymbol{u}$. Its recurrence function is:

$$R_{\boldsymbol{u}}(n) = \inf \{m \in \mathbb{N} : \text{every factor of length } m \text{ } \textbf{contains} \text{ all the factors of length } n\}.$$

- Cost we have to pay to discover the factors if we start from an arbitrary point in $\boldsymbol{u} = u_1 u_2 \ldots$

## Theorem (Morse, Hedlund, 1940)

*The recurrence function is piecewise affine and satisfies*

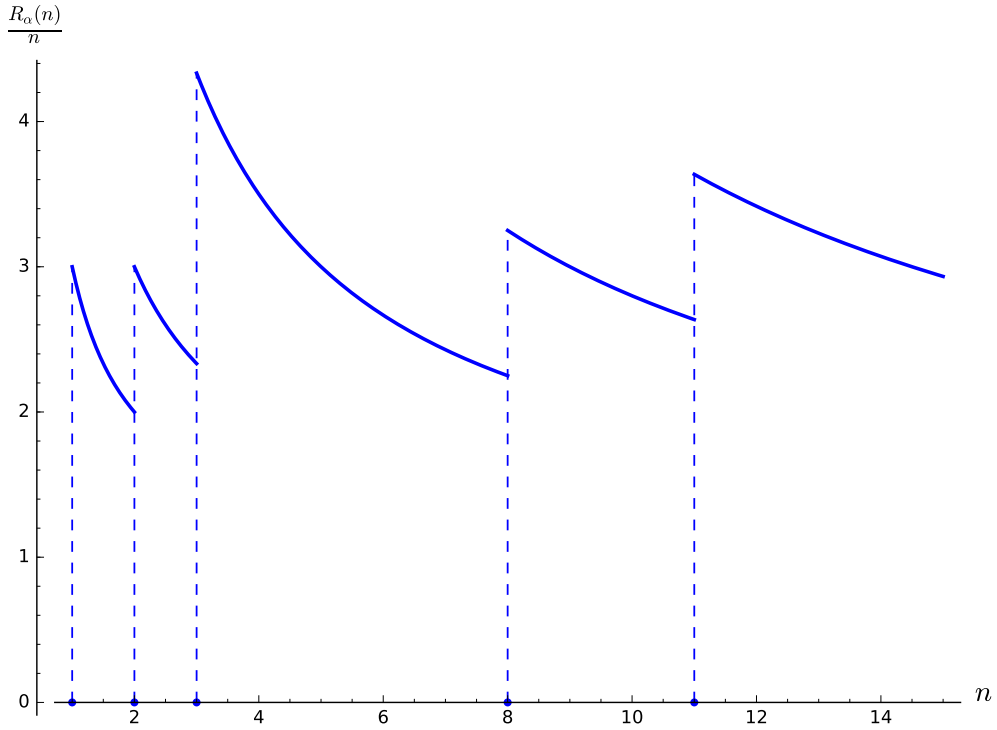
$$R_\alpha(n) = n - 1 + q_{k-1}(\alpha) + q_k(\alpha), \qquad \text{for } q_{k-1}(\alpha) \leq n < q_k(\alpha).$$

# Recurrence quotient and its parameters

$$S(\alpha, n) := \frac{R_\alpha(n) + 1}{n} = 1 + \frac{q_{k-1}(\alpha) + q_k(\alpha)}{n}, \quad q_{k-1}(\alpha) \leq n < q_k(\alpha).$$

# Recurrence quotient and its parameters

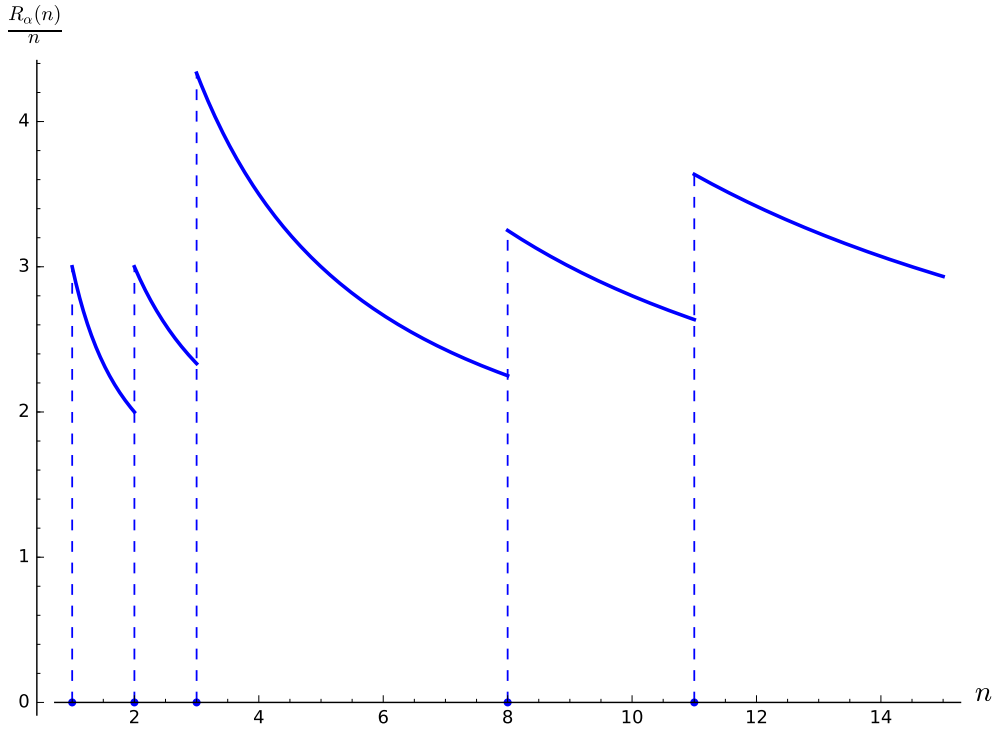
$$S(\alpha, n) := \frac{R_\alpha(n) + 1}{n} = 1 + \frac{q_{k-1}(\alpha) + q_k(\alpha)}{n}, \quad q_{k-1}(\alpha) \leq n < q_k(\alpha).$$




Recurrence quotient $\alpha = e^{-1}$.

# Recurrence quotient and its parameters

$$S(\alpha, n) := \frac{R_\alpha(n) + 1}{n} = 1 + \frac{q_{k-1}(\alpha) + q_k(\alpha)}{n}, \quad q_{k-1}(\alpha) \le n < q_k(\alpha).$$




Recurrence quotient $\alpha = e^{-1}$.

*Size* of $S(\alpha, n)$ dictated by

- the relative position of $n$ within the interval

$$\mu(\alpha, n) = \frac{n - q_{k-1}(\alpha)}{q_k(\alpha) - q_{k-1}(\alpha)},$$

- the quotient between the ends of the interval

$$\rho(\alpha, n) = \frac{q_{k-1}(\alpha)}{q_k(\alpha)}.$$

# Studies of the recurrence function

Classical results concern the worst case scenarios for *fixed* $\alpha$: let $\epsilon > 0$, for almost every $\alpha$

$$\limsup_{n\to\infty} \frac{S(\alpha, n)}{\log n} = \infty\,, \quad \lim_{n\to\infty} \frac{S(\alpha, n)}{(\log n)^{1+\epsilon}} = 0\,.$$

(Morse&Hedlund '40)

# Studies of the recurrence function

Classical results concern the worst case scenarios for *fixed* $\alpha$: let $\epsilon > 0$, for almost every $\alpha$

$$\limsup_{n\to\infty} \frac{S(\alpha,n)}{\log n} = \infty\,, \quad \lim_{n\to\infty} \frac{S(\alpha,n)}{(\log n)^{1+\epsilon}} = 0\,.$$

(Morse&Hedlund '40)

We define two probabilistic models

in both cases $\alpha$ is drawn *uniformly at random*

1) fix the length $n$ $\Rightarrow$ random variables $S_n(\alpha) := S(\alpha, n)$.
   in distribution and expectation as $n \to \infty$.
2) fix index $k$ of interval $[q_{k-1}(\alpha), q_k(\alpha))$ and
   the relative position $\mu$ $\Rightarrow$ sequence $(n_k(\alpha))_k$.

$\Rightarrow$ random variables $S_k(\alpha) := S(\alpha, n_k(\alpha))$ as $k \to \infty$.

# Studies of the recurrence function

Classical results concern the worst case scenarios for *fixed* $\alpha$: let $\epsilon > 0$, for almost every $\alpha$

$$\limsup_{n\to\infty} \frac{S(\alpha,n)}{\log n} = \infty\,, \qquad \lim_{n\to\infty} \frac{S(\alpha,n)}{(\log n)^{1+\epsilon}} = 0\,.$$

(Morse&Hedlund '40)

We define two probabilistic models

in both cases $\alpha$ is drawn *uniformly at random*

1) fix the length $n$ $\Rightarrow$ random variables $S_n(\alpha) := S(\alpha, n)$.
   in distribution and expectation as $n \to \infty$.
2) fix index $k$ of interval $[q_{k-1}(\alpha), q_k(\alpha))$ and
   the relative position $\mu$ $\Rightarrow$ sequence $(n_k(\alpha))_k$.

$\Rightarrow$ random variables $S_k(\alpha) := S(\alpha, n_k(\alpha))$ as $k \to \infty$.

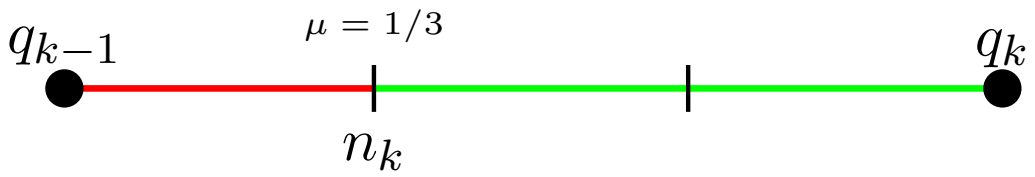



Figure : Sequence of indices $(n_k(\alpha))_k$ for $\mu = 1/3$.

# Results

## Model (1): fixed $n$. [RV17]

- Limit distribution for

$$\mu_n, \quad \rho_n, \quad S_n \, .$$

- Convergence of histograms to limit density.
- Conditional expectations

$$\mathbb{E}[S_n | \mu_n \geq \epsilon(n)] \sim \left| \log \epsilon(n) \right| .$$

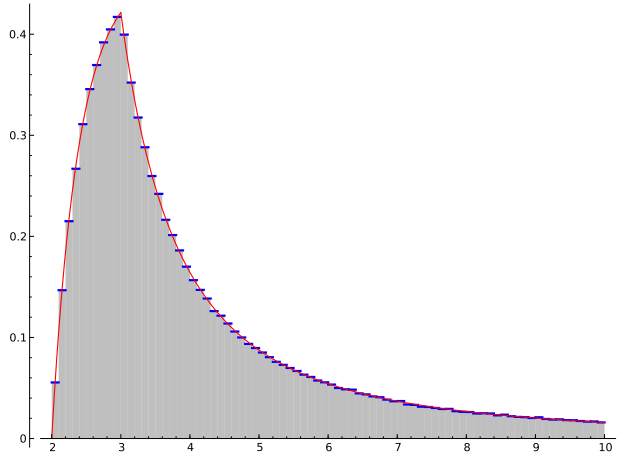



Figure : Limit density of $S_n$.

# Results

## Model (1): fixed $n$. [RV17]

- Limit distribution for
$$\mu_n, \quad \rho_n, \quad S_n\,.$$
- Convergence of histograms to limit density.
- Conditional expectations
$$\mathbb{E}[S_n|\mu_n \geq \epsilon(n)] \sim \left|\log \epsilon(n)\right|.$$

## Model (2): fixed $\mu$. [BCRVV15]

- Limit distribution/expectation of
$$S_k(\alpha) := S(\alpha, n_k)$$
depending on $\mu$.
- Study of $\mathbb{E}[S_k(\alpha)]$ as
$$\mu := \mu_k \to 0\,.$$

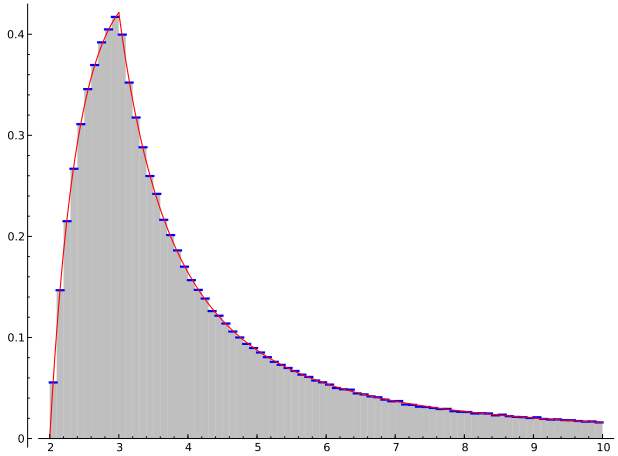



Figure : Limit density of $S_n$.

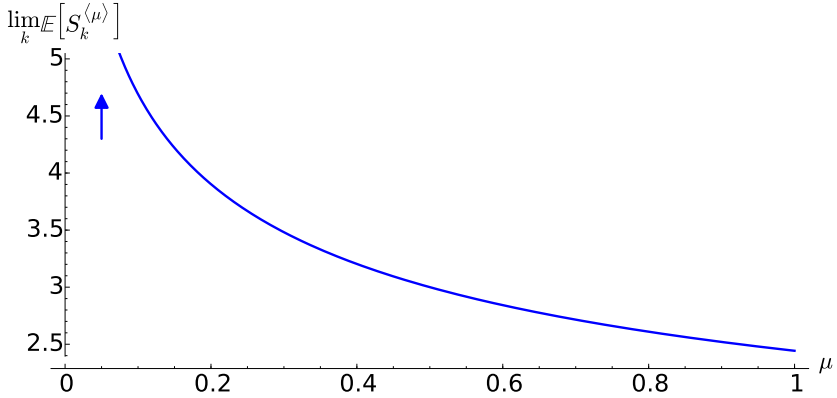



Figure : Limit of $\mathbb{E}[S_k]$ versus $\mu$.

# Elements of the proofs

## Model (1): fixed $n$

- Distribution at $\lambda$ given by *coprime* Riemann sum of step $\frac{1}{n}$ of

$$\omega(x,y) = \frac{2}{y(x+y)}\,,$$

over

$$\Delta_\lambda = \{(x,y) : 0 < x \leq 1 < y,\\ 1 + x + y \leq \lambda\}\,.$$

# Elements of the proofs

## Model (1): fixed $n$

- Distribution at $\lambda$ given by *coprime* Riemann sum of step $\frac{1}{n}$ of

$$\omega(x,y) = \frac{2}{y(x+y)}\,,$$

over

$$\Delta_\lambda = \{(x,y) : 0 < x \leq 1 < y,\ 1 + x + y \leq \lambda\}\,.$$

## Model (2): fixed $\mu$

- Expected values given by Perron-Frobenius operator:

$$\mathbb{E}[S_k] \sim 1 + \mathbf{H}^k[g](0)\,,$$

where

$$g(\rho) = \left(\frac{1+\rho}{\mu+\rho-\mu\rho}\right) / (1+\rho)\,.$$

# Section

# The origins

Introduced by Gosper as a mutation of continued fractions:

- gives rise to a $\gcd$ algorithm akin to Euclid's.
- quotients are powers of two:
  - small information parcel.
  - employs only shifts and substractions.
- appears to be simple and efficient.

# The origins

Introduced by Gosper as a mutation of continued fractions:

- gives rise to a $\gcd$ algorithm akin to Euclid's.
- quotients are powers of two:
  - small information parcel.
  - employs only shifts and substractions.
- appears to be simple and efficient.

More recently:

- Shallit studied its worst-case performance in 2016.
- We consider its average performance!

# Continued Logarithm Algorithm

A sequence of binary "divisions" beginning from $(p, q)$:

$$q = \mathbf{2^a} p + r\,, \qquad 0 \le r < 2^a p\,.$$

# Continued Logarithm Algorithm

A sequence of binary "divisions" beginning from $(p, q)$:

$$q = \mathbf{2^a} p + r\,, \qquad 0 \le r < 2^a p\,.$$

**Note.** $a = \max\{k \ge 0 : 2^k p \le q\}$

# Continued Logarithm Algorithm

A sequence of binary "divisions" beginning from $(p, q)$:

$$q = \mathbf{2^a} p + r\,, \qquad 0 \le r < 2^a p\,.$$

**Note.** $a = \max\{k \ge 0 : 2^k p \le q\}$

Continue with the new pair

$$(p, q) \mapsto (p', q') = (r, 2^a p)\,,$$

until the remainder $r$ equals $0$.

# Continued Logarithm Algorithm

A sequence of binary "divisions" beginning from $(p, q)$:

$$q = \mathbf{2^a} p + r\,, \qquad 0 \le r < 2^a p\,.$$

**Note.** $a = \max\{k \ge 0 : 2^k p \le q\}$

Continue with the new pair

$$(p, q) \mapsto (p', q') = (r, 2^a p)\,,$$

until the remainder $r$ equals $0$.

**Example.** Let us find $\gcd(13, 31)$.

| $a$ | $p$ | $q$ | $r$ | $2^a p$ |
|---|---|---|---|---|
| 1 | 13 | 31 | 5 | 26 |
| 2 | 5 | 26 | 6 | 20 |
| 1 | 6 | 20 | 8 | 12 |
| 0 | 8 | 12 | 4 | 8 |
| 1 | 4 | 8 | 0 | 8 |

# Continued Logarithm Algorithm

A sequence of binary "divisions" beginning from $(p, q)$:

$$q = \mathbf{2^a} p + r\,, \qquad 0 \le r < 2^a p\,.$$

**Note.** $a = \max\{k \ge 0 : 2^k p \le q\}$

Continue with the new pair

$$(p, q) \mapsto (p', q') = (r, 2^a p)\,,$$

until the remainder $r$ equals $0$.

**Example.** Let us find $\gcd(13, 31)$.

| $a$ | $p$ | $q$ | $r$ | $2^a p$ |
|---|---|---|---|---|
| 1 | 13 | 31 | 5 | 26 |
| 2 | 5 | 26 | 6 | 20 |
| 1 | 6 | 20 | 8 | 12 |
| 0 | 8 | 12 | 4 | 8 |
| 1 | 4 | 8 | 0 | 8 |

- Ended with $(0, 8)$, what is the $\gcd$?
  $\Rightarrow$ odd $\gcd \times$ parasitic powers of $2$.

Consider

$$\Omega_N = \{(p,q) \in \mathbb{N} \times \mathbb{N} : p \leq q \leq N\}.$$

Worst-case studied by Shallit (2016): $2\log_2 N + O(1)$ steps.

Consider

$$\Omega_N = \{(p,q) \in \mathbb{N} \times \mathbb{N} : p \leq q \leq N\}\,.$$

Worst-case studied by Shallit (2016): $2\log_2 N + O(1)$ steps.

- Family $(p,q) = (1, 2^n - 1)$ gives the bound asymptotically.

Consider

$$\Omega_N = \{(p,q) \in \mathbb{N} \times \mathbb{N} : p \leq q \leq N\}.$$

Worst-case studied by Shallit (2016): $2\log_2 N + O(1)$ steps.

- Family $(p,q) = (1, 2^n - 1)$ gives the bound asymptotically.

We studied the average number of steps over $\Omega_N$, posed by Shallit.

Consider

$$\Omega_N = \{(p,q) \in \mathbb{N} \times \mathbb{N} : p \le q \le N\}\,.$$

Worst-case studied by Shallit (2016): $2\log_2 N + O(1)$ steps.

◦ Family $(p,q) = (1, 2^n - 1)$ gives the bound asymptotically.

We studied the average number of steps over $\Omega_N$, posed by Shallit.

**Main result (R., Vallée, '18).**

Mean number of **steps** $E_N[K]$ and **shifts** $E_N[S]$ are $\Theta(\log N)$.

More precisely

$$E_N[K] \sim k\,\log N\,, \qquad E_N[S] \sim \tfrac{\log 3 - \log 2}{2\log 2 - \log 3}\,E_N[K]$$

for an *explicit constant* $k \doteq 1.49283\ldots$ given by

$$k = \frac{2}{H}\,, \quad H = \tfrac{1}{\log(4/3)}\left(\tfrac{\pi^2}{6} + 2\sum_j \tfrac{(-1)^j}{2^j j^2} - (\log 2)\tfrac{\log 27}{\log 16}\right)$$

Process depends **only** on $p/q$ rather than $(p, q)$.

- Map $p/q \mapsto p'/q'$ can be extended to $\mathcal{I} = (0, 1)$

$$T \colon \mathcal{I} \to \mathcal{I}\,, \quad T(x) = \frac{1}{2^a x} - 1\,,$$

where $a = \lfloor \log_2(1/x) \rfloor$.

- Iteration gives a special continued fraction

$$\frac{p}{q} = \frac{1}{2^a \left(1 + \frac{p'}{q'}\right)}\,.$$

Process depends **only** on $p/q$ rather than $(p, q)$.

- Map $p/q \mapsto p'/q'$ can be extended to $\mathcal{I} = (0, 1)$

$$T\colon \mathcal{I} \to \mathcal{I}\,, \quad T(x) = \frac{1}{2^a x} - 1\,,$$

where $a = \lfloor \log_2(1/x) \rfloor$.

- Iteration gives a special continued fraction

$$\frac{p}{q} = \frac{1}{2^a \left(1 + \frac{p'}{q'}\right)}\,.$$

- For Euclid's algorithm, we get the Gauss map

$$T\colon \mathcal{I} \to \mathcal{I}\,, \quad T(x) = \frac{1}{x} - m\,,$$

where $m = \lfloor 1/x \rfloor$.

- Iteration gives classical continued fractions

$$\frac{p}{q} = \frac{1}{m + \frac{p'}{q'}}\,.$$

Process depends **only** on $p/q$ rather than $(p, q)$.

- Map $p/q \mapsto p'/q'$ can be extended to $\mathcal{I} = (0, 1)$

$$T\colon \mathcal{I} \to \mathcal{I}\,, \quad T(x) = \frac{1}{2^a x} - 1\,,$$

where $a = \lfloor \log_2(1/x) \rfloor$.

- Iteration gives a special continued fraction

$$\frac{p}{q} = \frac{1}{2^a \left(1 + \frac{p'}{q'}\right)}\,.$$

- For Euclid's algorithm, we get the Gauss map

$$T\colon \mathcal{I} \to \mathcal{I}\,, \quad T(x) = \frac{1}{x} - m\,,$$

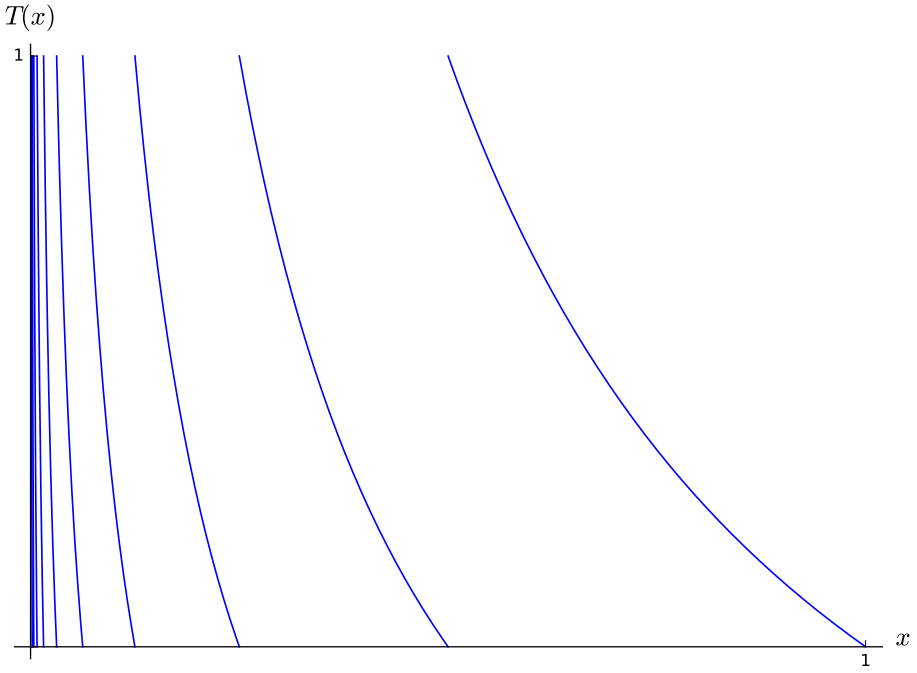
where $m = \lfloor 1/x \rfloor$.

- Iteration gives classical continued fractions

$$\frac{p}{q} = \frac{1}{m + \frac{p'}{q'}}\,.$$

The continued fraction expansion ends (is finite) when we get $0$.

# Dynamical system $(\mathcal{I}, T)$




The map $T\colon \mathcal{I} \to \mathcal{I}$

**Branches**

For $x \in \mathcal{I}_a := [2^{-a-1}, 2^{-a}]$

$$x \mapsto T_a(x) := \frac{2^{-a}}{x} - 1\,.$$

where $a(x) := \lfloor \log_2(1/x) \rfloor\,.$

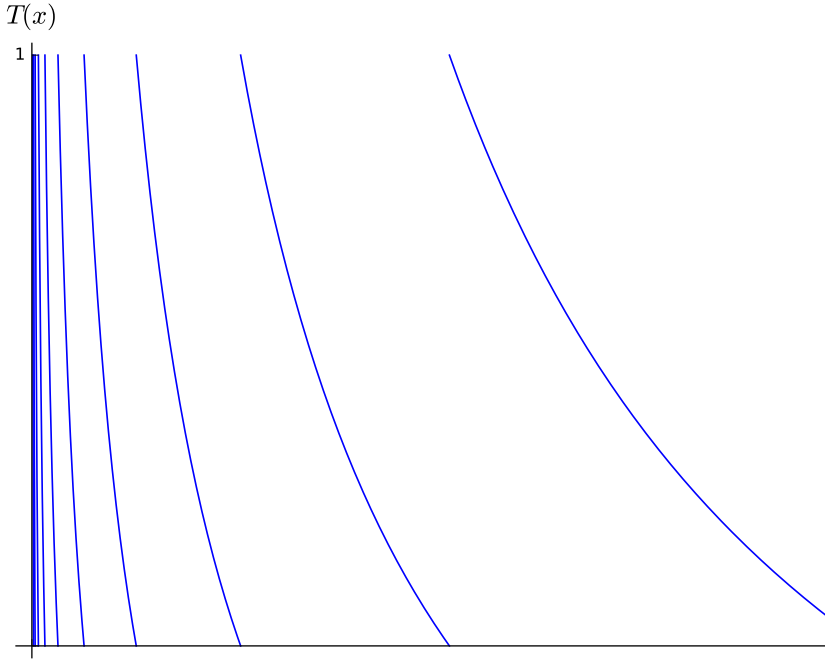
**Inverse branches**

$$h_a(x) := \frac{2^{-a}}{1+x}\,, \quad \mathcal{H} := \{h_a : a \in \mathbb{N}\}\,,$$

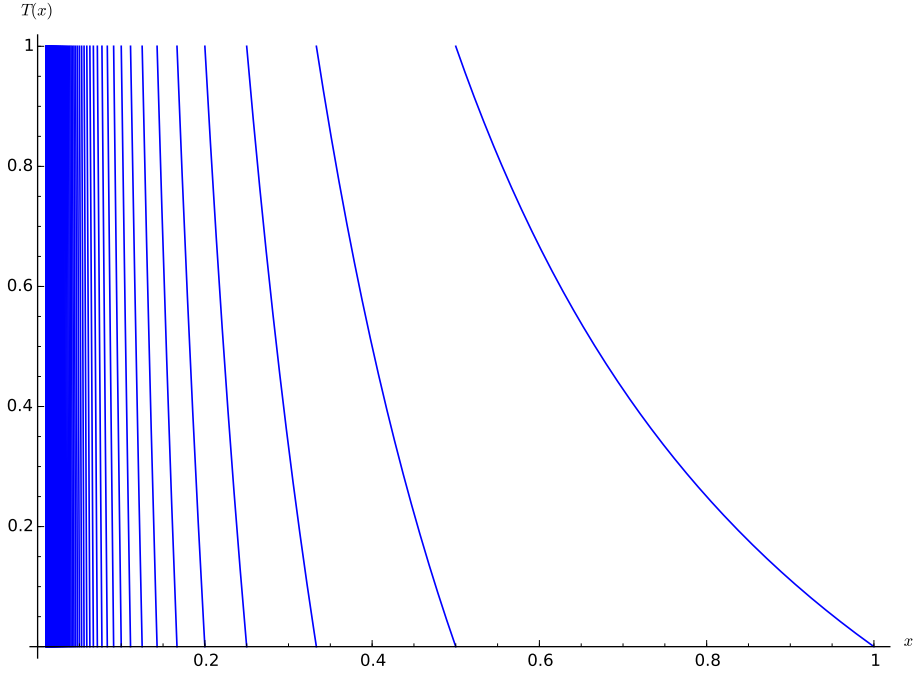
and at depth $k$

$$\mathcal{H}^k := \{h_{a_1} \circ \cdots \circ h_{a_k} : a_1, \ldots, a_k \in \mathbb{N}\}\,.$$

# Dynamical system $(\mathcal{I}, T)$




The map $T\colon \mathcal{I} \to \mathcal{I}$




The map for Euclid's algorithm.

# Reduced denominators and inverse branches

Observe for $f(x) = \frac{ax+b}{cx+d}$ :

$$f'(x) = \frac{\det f}{(cx+d)^2}.$$

Euclidean algorithm:

- Homographies

$$h_m(x) = \frac{1}{m+x},$$

with $\det h_m = -1$.

- For $h = h_{m_1} \circ \ddots \circ h_{m_k}$

$$h(0) = \frac{p}{q} \Rightarrow |h'(0)| = \frac{1}{q^2},$$

$p/q$ reduced.

# Reduced denominators and inverse branches

Observe for $f(x) = \frac{ax+b}{cx+d}$ :

$$f'(x) = \frac{\det f}{(cx+d)^2}.$$

Euclidean algorithm:

- Homographies

$$h_m(x) = \frac{1}{m+x},$$

with $\det h_m = -1$.

- For $h = h_{m_1} \circ \ddots \circ h_{m_k}$

$$h(0) = \frac{p}{q} \Rightarrow |h'(0)| = \frac{1}{q^2},$$

$p/q$ reduced.

CL algorithm:

- Homographies

$$h_a(x) = \frac{1}{2^a(1+x)},$$

with $\det h_a = -2^a$.

- For $h = h_{m_1} \circ \ddots \circ h_{m_k}$

$$h(0) = \frac{p}{q} \Rightarrow |h'(0)| \text{ vs. } \frac{1}{q^2}?$$

$p/q$ reduced.

# Reduced denominators and inverse branches

Observe for $f(x) = \frac{ax+b}{cx+d}$ :

$$f'(x) = \frac{\det f}{(cx+d)^2}.$$

Euclidean algorithm:

- Homographies

$$h_m(x) = \frac{1}{m+x},$$

with $\det h_m = -1$.

- For $h = h_{m_1} \circ \ddots \circ h_{m_k}$

$$h(0) = \frac{p}{q} \Rightarrow |h'(0)| = \frac{1}{q^2},$$

$p/q$ reduced.

CL algorithm:

- Homographies

$$h_a(x) = \frac{1}{2^a(1+x)},$$

with $\det h_a = -2^a$.

- For $h = h_{m_1} \circ \ddots \circ h_{m_k}$

$$h(0) = \frac{p}{q} \Rightarrow |h'(0)| \text{ vs. } \frac{1}{q^2}?$$

$p/q$ reduced.

**Problem:** Denominator retrieved is engorged by powers of two.

# Recording the dyadic behaviour

Dyadic behaviour is related to divisibility

# Recording the dyadic behaviour

Dyadic behaviour is related to divisibility

$\Longrightarrow$ ... but we employ analysis!

# Recording the dyadic behaviour

Dyadic behaviour is related to divisibility

$\Longrightarrow$ ... but we employ analysis!

**Response:** Dyadic numbers $\mathbb{Q}_2$ !

Dyadic topology = Divisibility by 2 constraints ,

using the dyadic norm $|\cdot|_2$.

# Recording the dyadic behaviour

Dyadic behaviour is related to divisibility

$\Longrightarrow$ ... but we employ analysis!

**Response:** Dyadic numbers $\mathbb{Q}_2$ !

Dyadic topology = Divisibility by 2 constraints ,

using the dyadic norm $|\cdot|_2$.

- Introduce dyadic component

  $\Rightarrow$ mixed dynamical system $(x, y) \in \mathcal{I} \times \mathbb{Q}_2$

# Recording the dyadic behaviour

Dyadic behaviour is related to divisibility

$\Longrightarrow$ ... but we employ analysis!

**Response:** Dyadic numbers $\mathbb{Q}_2$ !

Dyadic topology = Divisibility by 2 constraints ,

using the dyadic norm $|\cdot|_2$.

- Introduce dyadic component

  $\Rightarrow$ mixed dynamical system $(x, y) \in \mathcal{I} \times \mathbb{Q}_2$
- Incorporate $\mathbb{Q}_2$ into the Transfer Operator?

# Recording the dyadic behaviour

Dyadic behaviour is related to divisibility

$\Longrightarrow$ ... but we employ analysis!

**Response:** Dyadic numbers $\mathbb{Q}_2$ !

Dyadic topology = Divisibility by 2 constraints ,

using the dyadic norm $|\cdot|_2$.

- Introduce dyadic component

  $\Rightarrow$ mixed dynamical system $(x, y) \in \mathcal{I} \times \mathbb{Q}_2$
- Incorporate $\mathbb{Q}_2$ into the Transfer Operator?

**Idea works!**

# The extended dynamical system

⊛ Introduce $\underline{\mathcal{I}} := \mathcal{I} \times \mathbb{Q}_2$ and $\underline{T} \colon \underline{\mathcal{I}} \to \underline{\mathcal{I}}$ as follows

$$\underline{T}(x, y) = (T_a(x), T_a(y)) \,,$$

for $x \in \mathcal{I}_a = [2^{-a-1}, 2^{-a}]$. This gives inverse branches

$$\underline{h}_a(x, y) = (h_a(x), h_a(y)) \;, \qquad (x, y) \in \underline{\mathcal{I}} \,.$$

# The extended dynamical system

⊛ Introduce $\underline{\mathcal{I}} := \mathcal{I} \times \mathbb{Q}_2$ and $\underline{T} \colon \underline{\mathcal{I}} \to \underline{\mathcal{I}}$ as follows

$$\underline{T}(x, y) = (T_a(x), T_a(y)) \,,$$

for $x \in \mathcal{I}_a = [2^{-a-1}, 2^{-a}]$. This gives inverse branches

$$\underline{h}_a(x, y) = (h_a(x), h_a(y)) \;, \qquad (x, y) \in \underline{\mathcal{I}} \,.$$

Evolution is led by the real component, which determines $a$.

# The extended dynamical system

⊛ Introduce $\underline{\mathcal{I}} := \mathcal{I} \times \mathbb{Q}_2$ and $\underline{T} \colon \underline{\mathcal{I}} \to \underline{\mathcal{I}}$ as follows

$$\underline{T}(x, y) = (T_a(x), T_a(y)) \,,$$

for $x \in \mathcal{I}_a = [2^{-a-1}, 2^{-a}]$. This gives inverse branches

$$\underline{h}_a(x, y) = (h_a(x), h_a(y)) \;, \qquad (x, y) \in \underline{\mathcal{I}} \,.$$

Evolution is led by the real component, which determines $a$.

⊛ For Transfer operator $\Rightarrow$ need change of variables formula!

# The extended dynamical system

⊛ Introduce $\underline{\mathcal{I}} := \mathcal{I} \times \mathbb{Q}_2$ and $\underline{T} \colon \underline{\mathcal{I}} \to \underline{\mathcal{I}}$ as follows

$$\underline{T}(x, y) = (T_a(x), T_a(y)) \,,$$

for $x \in \mathcal{I}_a = [2^{-a-1}, 2^{-a}]$. This gives inverse branches

$$\underline{h}_a(x, y) = (h_a(x), h_a(y)) \;, \qquad (x, y) \in \underline{\mathcal{I}} \,.$$

Evolution is led by the real component, which determines $a$.

⊛ For Transfer operator $\Rightarrow$ need change of variables formula!

Haar (translation invariant) measure $\nu$ on $\mathbb{Q}_2$ has one!

# Functional space $\mathcal{F}$ for the extended operator $\underline{\mathbf{H}}_s$

Real component directs the dynamical system:

- *sections* $F_y$ fixing $y \in \mathbb{Q}_2$ asked to be $C^1(\mathcal{I})$.
- the dyadic component follows, demanding only integrability.

# Functional space $\mathcal{F}$ for the extended operator $\underline{\mathbf{H}}_s$

Real component directs the dynamical system:

- *sections* $F_y$ fixing $y \in \mathbb{Q}_2$ asked to be $C^1(\mathcal{I})$.
- the dyadic component follows, demanding only integrability.

Ensuing space $\mathcal{F}$ makes $\underline{\mathbf{H}}_s$

- have a dominant eigenvalue and spectral gap
  relying strongly on the real component.

# Functional space $\mathcal{F}$ for the extended operator $\underline{\mathbf{H}}_s$

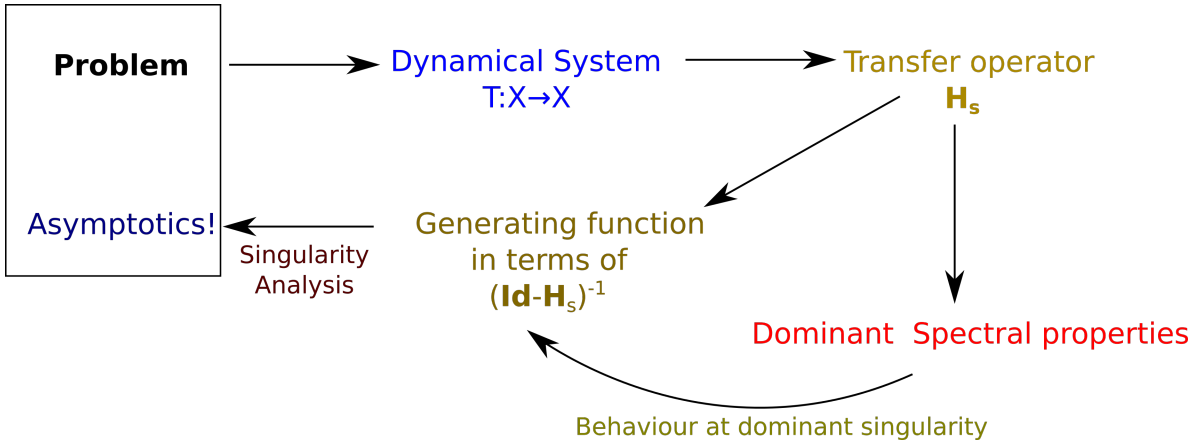
Real component directs the dynamical system:

- *sections* $F_y$ fixing $y \in \mathbb{Q}_2$ asked to be $C^1(\mathcal{I})$.
- the dyadic component follows, demanding only integrability.

Ensuing space $\mathcal{F}$ makes $\underline{\mathbf{H}}_s$

- have a dominant eigenvalue and spectral gap
  relying strongly on the real component.

**We can finish the dynamical analysis!**



# Conclusion and further questions

⊛ We have studied the average number of shifts and substractions for the CL algorithm.

⊛ Study makes an interesting use of the dyadics in the framework of dynamical analysis.

# Conclusion and further questions

⊛ We have studied the average number of shifts and substractions for the CL algorithm.

⊛ Study makes an interesting use of the dyadics in the framework of dynamical analysis.

Questions:

1. **Conjecture:** The sucessive pairs $(p_i, q_i)$ given by the algorithm satisfy

$$\log_2 \gcd(p_i, q_i) \sim i/2\,.$$

Back to $(13, 31)$

| $i$ | $a_i$ | $p_i$ | $q_i$ | $\gcd(p_i, q_i)$ |
|---|---|---|---|---|
| 0 | 1 | 13 | 31 | $2^0$ |
| 1 | 2 | 5 | 26 | $2^0$ |
| 2 | 1 | 6 | 20 | $2^1$ |
| 3 | 0 | 8 | 12 | $2^2$ |
| 4 | 1 | 4 | 8 | $2^2$ |

# Conclusion and further questions

⊛ We have studied the average number of shifts and substractions for the CL algorithm.

⊛ Study makes an interesting use of the dyadics in the framework of dynamical analysis.

Questions:

1. **Conjecture:** The sucessive pairs $(p_i, q_i)$ given by the algorithm satisfy

$$\log_2 \gcd(p_i, q_i) \sim i/2\,.$$

   Back to $(13, 31)$

| $i$ | $a_i$ | $p_i$ | $q_i$ | $\gcd(p_i, q_i)$ |
|---|---|---|---|---|
| 0 | 1 | 13 | 31 | $2^0$ |
| 1 | 2 | 5 | 26 | $2^0$ |
| 2 | 1 | 6 | 20 | $2^1$ |
| 3 | 0 | 8 | 12 | $2^2$ |
| 4 | 1 | 4 | 8 | $2^2$ |

2. **Comparison to other binary algorithms:** binary GCD, LSB.

# Conclusion and further questions

⊛ We have studied the average number of shifts and substractions for the CL algorithm.

⊛ Study makes an interesting use of the dyadics in the framework of dynamical analysis.

Questions:

1. **Conjecture:** The sucessive pairs $(p_i, q_i)$ given by the algorithm satisfy

$$\log_2 \gcd(p_i, q_i) \sim i/2\,.$$

   Back to $(13, 31)$

| $i$ | $a_i$ | $p_i$ | $q_i$ | $\gcd(p_i, q_i)$ |
|---|---|---|---|---|
| 0 | 1 | 13 | 31 | $2^0$ |
| 1 | 2 | 5 | 26 | $2^0$ |
| 2 | 1 | 6 | 20 | $2^1$ |
| 3 | 0 | 8 | 12 | $2^2$ |
| 4 | 1 | 4 | 8 | $2^2$ |

2. **Comparison to other binary algorithms:** binary GCD, LSB.