Extensions of the Pascal triangle to words

Joint work with Julien Leroy and Michel Rigo (ULiège)

Manon Stipulanti (ULiège) FRIA grantee

Most Informal Probability Seminar
Leiden (The Netherlands)
May 28, 2019

Classical Pascal triangle

$\mathrm{P}\colon (m,k) \in \mathbb{N} \times \mathbb{N} \mapsto \binom{m}{k} \in \mathbb{N}$

$\binom{m}{k}$		k								
		0	1	2	3	4	5	6	7	$\cdots$
	0	1	0	0	0	0	0	0	0	
	1	1	1	0	0	0	0	0	0	
	2	1	2	1	0	0	0	0	0	
m	3	1	3	3	1	0	0	0	0	
	4	1	4	6	4	1	0	0	0	
	5	1	5	**10**	**10**	5	1	0	0	
	6	1	6	15	**20**	15	6	1	0	
	7	1	7	21	35	35	21	7	1	
	$\vdots$									$\ddots$

Binomial coefficients:

$$\binom{m}{k} = \frac{m!}{(m-k)!\,k!}$$

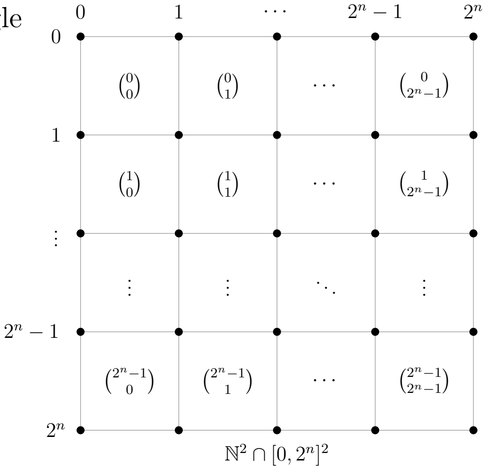
Pascal's rule:

$$\binom{m}{k} = \binom{m-1}{k} + \binom{m-1}{k-1}$$

A specific construction

- Grid: first 2^n rows and columns of the Pascal triangle

$$\left(\binom{m}{k} \bmod 2\right)_{0\le m,k<2^n}$$

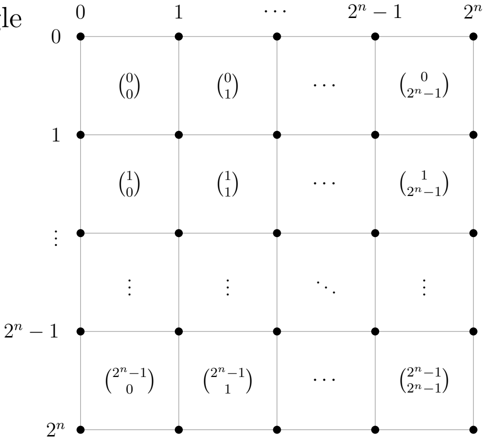


A specific construction

- Grid: first 2^n rows and columns of the Pascal triangle

$$\left(\binom{m}{k} \bmod 2\right)_{0\le m,k<2^n}$$

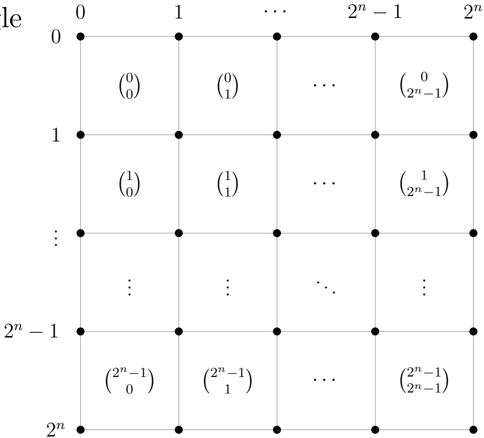
- Color each square in
  - white if $\binom{m}{k} \equiv 0 \bmod 2$
  - black if $\binom{m}{k} \equiv 1 \bmod 2$



$\mathbb{N}^2 \cap [0,2^n]^2$

A specific construction

- Grid: first 2^n rows and columns of the Pascal triangle
$$\left(\binom{m}{k} \bmod 2\right)_{0\le m,k<2^n}$$
- Color each square in
  - white if $\binom{m}{k} \equiv 0 \bmod 2$
  - black if $\binom{m}{k} \equiv 1 \bmod 2$
- Normalize by a homothety of ratio $1/2^n$ (bring into $[0,1]^2$)



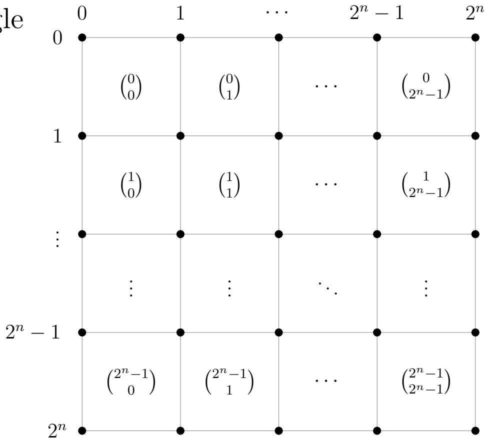
$\mathbb{N}^2 \cap [0,2^n]^2$

- Grid: first 2^n rows and columns of the Pascal triangle

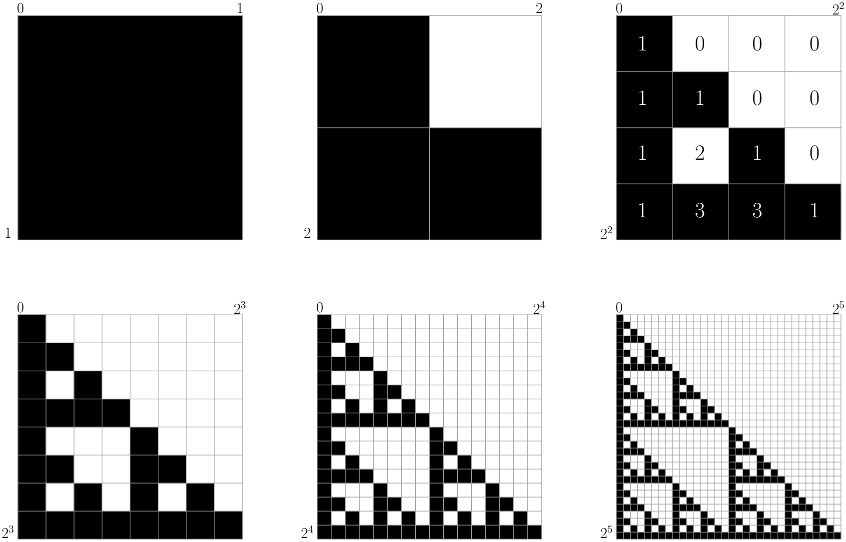
$$\left(\binom{m}{k} \bmod 2\right)_{0\le m,k<2^n}$$

- Color each square in
  - white if $\binom{m}{k} \equiv 0 \bmod 2$
  - black if $\binom{m}{k} \equiv 1 \bmod 2$
- Normalize by a homothety of ratio $1/2^n$ (bring into $[0,1]^2$)

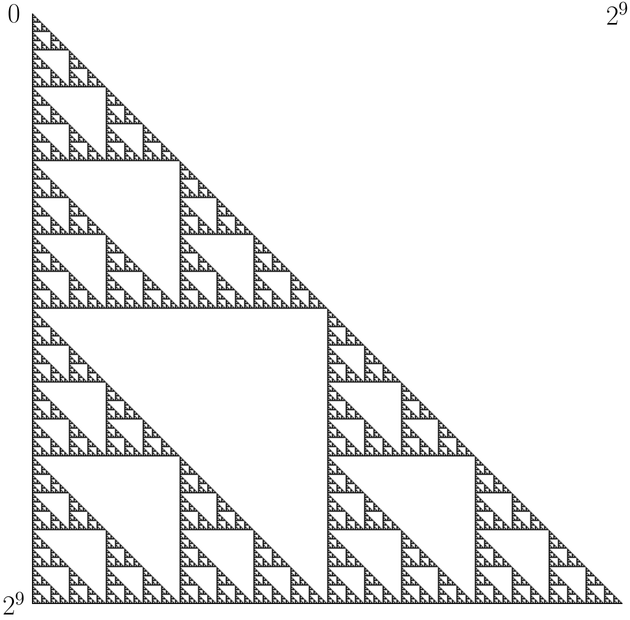
$\rightsquigarrow$ sequence of compact sets belonging to $[0,1]^2$



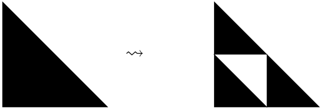
$\mathbb{N}^2 \cap [0, 2^n]^2$


0
1
1
0
2
2
0
2^2
2^2
1 0 0 0
1 1 0 0
1 2 1 0
1 3 3 1
0
2^3
2^3
0
2^4
2^4
0
2^5
2^5

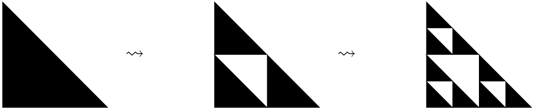
The tenth element of the sequence



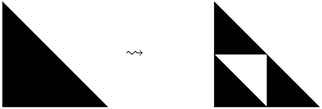
The Sierpiński gasket

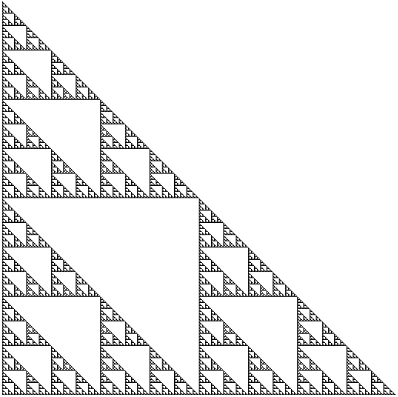


The Sierpiński gasket



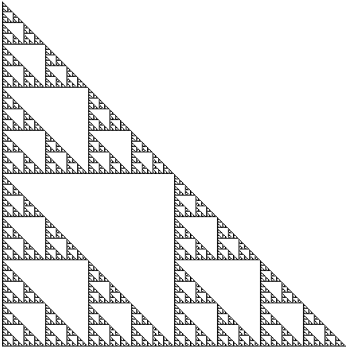
The Sierpiński gasket





Folklore fact

The latter sequence of compact sets converges to the Sierpiński gasket (w.r.t. the Hausdorff distance).



Definitions:

- ϵ-*fattening* of a subset $S \subset \mathbb{R}^2$

$$[S]_\epsilon = \bigcup_{x \in S} B(x, \epsilon)$$

- $(\mathcal{H}(\mathbb{R}^2), d_h)$ complete space of the non-empty compact subsets of $\mathbb{R}^2$ equipped with the *Hausdorff distance* d_h

$$d_h(S, S') = \inf\{\epsilon \in \mathbb{R}_{>0} \mid S \subset [S']_\epsilon \quad \text{and} \quad S' \subset [S]_\epsilon\}$$

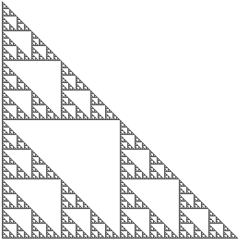
Theorem (von Haeseler, Peitgen, and Skordev, 1992)

Let p be a prime and $s > 0$.
The sequence of compact sets corresponding to

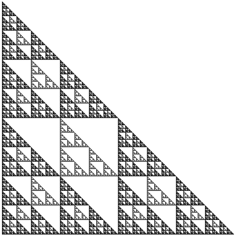
$$\left(\binom{m}{k} \bmod p^s\right)_{0 \leq m,k < p^n}$$

converges when n tends to infinity (w.r.t. the Hausdorff distance).

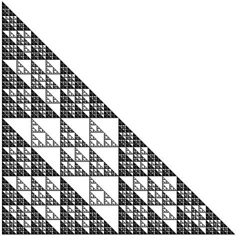
$p = 2,\ s = 1$



$p = 2,\ s = 2$



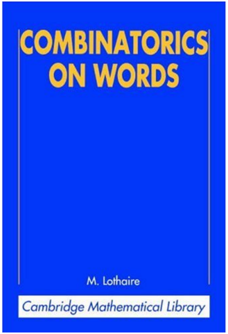
$p = 2,\ s = 3$



Replace integers by finite words.

Combinatorics on words (CoW)

- new area of discrete mathematics (± 1900)
- study sequences of symbols (called letters)
- topics include:
  - ◇ regularities and patterns in words
  - ◇ important types of words (e.g. automatic, regular, de Bruijn, Lyndon, Sturmian)
  - ◇ coding of structures (e.g. paths, trees or curves in the plane)



M. Lothaire, 1983.

Binomial coefficient of finite words

Definition: A *finite word* is a finite sequence of letters belonging to a finite set called *alphabet*.

Example: $101, 101001 \in \{0,1\}^*$

Binomial coefficient of words

Let u, v be two finite words.
The *binomial coefficient* $\binom{u}{v}$ of u and v is the number of times v occurs as a subsequence of u (meaning as a "scattered" subword).

Binomial coefficient of finite words

Definition: A *finite word* is a finite sequence of letters belonging to a finite set called *alphabet*.

Example: $101, 101001 \in \{0,1\}^*$

Binomial coefficient of words

Let u, v be two finite words.
The *binomial coefficient* $\binom{u}{v}$ of u and v is the number of times v occurs as a subsequence of u (meaning as a "scattered" subword).

Example: $u = 101001$ $\qquad v = 101$

Binomial coefficient of finite words

Definition: A *finite word* is a finite sequence of letters belonging to a finite set called *alphabet*.

Example: $101, 101001 \in \{0,1\}^*$

Binomial coefficient of words

Let u, v be two finite words.
The *binomial coefficient* $\binom{u}{v}$ of u and v is the number of times v occurs as a subsequence of u (meaning as a "scattered" subword).

Example: $u = \mathbf{101}001$ $\quad v = 101$ $\quad$ 1 occurrence

Binomial coefficient of finite words

Definition: A *finite word* is a finite sequence of letters belonging to a finite set called *alphabet*.

Example: $101, 101001 \in \{0,1\}^*$

Binomial coefficient of words

Let u, v be two finite words.
The *binomial coefficient* $\binom{u}{v}$ of u and v is the number of times v occurs as a subsequence of u (meaning as a "scattered" subword).

Example: $u = \mathbf{10}100\mathbf{1}$ $\quad v = 101 \quad$ 2 occurrences

Binomial coefficient of finite words

Definition: A *finite word* is a finite sequence of letters belonging to a finite set called *alphabet*.

Example: $101, 101001 \in \{0,1\}^*$

Binomial coefficient of words

Let u, v be two finite words.
The *binomial coefficient* $\binom{u}{v}$ of u and v is the number of times v occurs as a subsequence of u (meaning as a "scattered" subword).

Example: $u = \mathbf{101001}$ $\qquad v = 101 \qquad$ 3 occurrences

Binomial coefficient of finite words

Definition: A *finite word* is a finite sequence of letters belonging to a finite set called *alphabet*.

Example: $101, 101001 \in \{0,1\}^*$

Binomial coefficient of words

Let u, v be two finite words.
The *binomial coefficient* $\binom{u}{v}$ of u and v is the number of times v occurs as a subsequence of u (meaning as a "scattered" subword).

Example: $u = \mathbf{101001}$ $\quad v = 101 \quad$ 4 occurrences

Definition: A *finite word* is a finite sequence of letters belonging to a finite set called *alphabet*.

Example: $101, 101001 \in \{0,1\}^*$

Binomial coefficient of words

Let u, v be two finite words.
The *binomial coefficient* $\binom{u}{v}$ of u and v is the number of times v occurs as a subsequence of u (meaning as a "scattered" subword).

Example: $u = 10\mathbf{1}\mathbf{0}\mathbf{0}\mathbf{1}$ $\quad v = 101 \quad$ 5 occurrences

Definition: A *finite word* is a finite sequence of letters belonging to a finite set called *alphabet*.

Example: $101, 101001 \in \{0,1\}^*$

Binomial coefficient of words

Let u, v be two finite words.
The *binomial coefficient* $\binom{u}{v}$ of u and v is the number of times v occurs as a subsequence of u (meaning as a "scattered" subword).

Example: $u = 10\mathbf{1}0\mathbf{0}\mathbf{1}$ $\quad v = 101 \quad$ 6 occurrences

Definition: A *finite word* is a finite sequence of letters belonging to a finite set called *alphabet*.

Example: $101, 101001 \in \{0,1\}^*$

Binomial coefficient of words

Let u, v be two finite words.
The *binomial coefficient* $\binom{u}{v}$ of u and v is the number of times v occurs as a subsequence of u (meaning as a "scattered" subword).

Example: $u = 101001$ $\qquad$ $v = 101$

$$\Rightarrow \binom{101001}{101} = 6$$

Remark:

Natural generalization of binomial coefficients of integers

If a is a letter,

$$\binom{a^m}{a^k} = \binom{\overbrace{a\cdots a}^{m\text{ times}}}{\underbrace{a\cdots a}_{k\text{ times}}} = \binom{m}{k} \quad \forall\, m,k \in \mathbb{N}\,.$$

Generalized Pascal triangles

Let $(A, <)$ be a totally ordered alphabet.
Let $L \subset A^*$ be an infinite language (set of words) over A.
The words in L are genealogically ordered

$$w_0 <_{\mathrm{gen}} w_1 <_{\mathrm{gen}} w_2 <_{\mathrm{gen}} \cdots .$$

The *generalized Pascal triangle* P_L associated with L is defined by

$$\mathrm{P}_L \colon (m, k) \in \mathbb{N} \times \mathbb{N} \mapsto \binom{w_m}{w_k} \in \mathbb{N} .$$

<u>Questions:</u>

- With a similar construction, can we expect the convergence to an analogue of the Sierpiński gasket?
- In particular, where should we cut to normalize a given generalized Pascal triangle?
- Could we describe this limit object?

Base 2

Definitions:

- $\text{rep}_2(n)$ greedy base-2 representation of $n \in \mathbb{N}_{>0}$ starting with 1
- $\text{rep}_2(0) = \varepsilon$ where ε is the empty word

n	$n = \sum_i c_i 2^i$ with $c_i \in \{0,1\}$	$\text{rep}_2(n)$
0		ε
1	1×2^0	1
2	$1 \times 2^1 + 0 \times 2^0$	10
3	$1 \times 2^1 + 1 \times 2^0$	11
4	$1 \times 2^2 + 0 \times 2^1 + 0 \times 2^0$	100
5	$1 \times 2^2 + 0 \times 2^1 + 1 \times 2^0$	101
6	$1 \times 2^2 + 1 \times 2^1 + 0 \times 2^0$	110
$\vdots$	$\vdots$	$\vdots$
		$L_2 = 1\{0,1\}^* \cup \{\varepsilon\}$

Generalized Pascal triangle P_2 in base 2

	$\binom{\mathrm{rep}_2(m)}{\mathrm{rep}_2(k)}$	$\mathrm{rep}_2(k)$								
		ε	1	10	11	100	101	110	111	$\cdots$
	ε	1	0	0	0	0	0	0	0	
	1	1	1	0	0	0	0	0	0	
	10	1	1	1	0	0	0	0	0	
	11	1	**2**	0	**1**	0	0	0	0	
$\mathrm{rep}_2(m)$	100	1	1	2	0	1	0	0	0	
	101	1	2	1	1	0	1	0	0	
	110	1	2	2	1	0	0	1	0	
	111	1	3	0	**3**	0	0	0	1	
	$\vdots$									$\ddots$

Rule (not local):

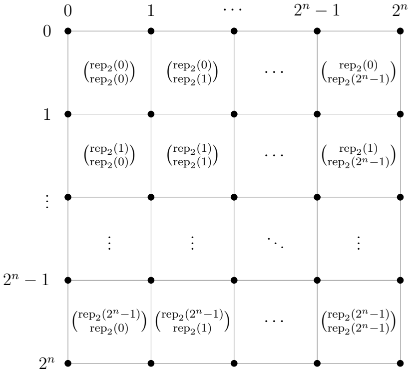
$$\binom{ua}{vb} = \binom{u}{vb} + \delta_{a,b}\binom{u}{v}$$

$\binom{\mathrm{rep}_2(m)}{\mathrm{rep}_2(k)}$		$\mathrm{rep}_2(k)$								
		ε	**1**	10	**11**	100	101	110	**111**	$\cdots$
	ε	**1**	0	0	0	0	0	0	0	
	1	**1**	**1**	0	0	0	0	0	0	
	10	1	1	1	0	0	0	0	0	
	11	**1**	**2**	0	**1**	0	0	0	0	
$\mathrm{rep}_2(m)$	100	1	1	2	0	1	0	0	0	
	101	1	2	1	1	0	1	0	0	
	110	1	2	2	1	0	0	1	0	
	111	**1**	**3**	0	**3**	0	0	0	**1**	
	$\vdots$									$\ddots$

The classical Pascal triangle

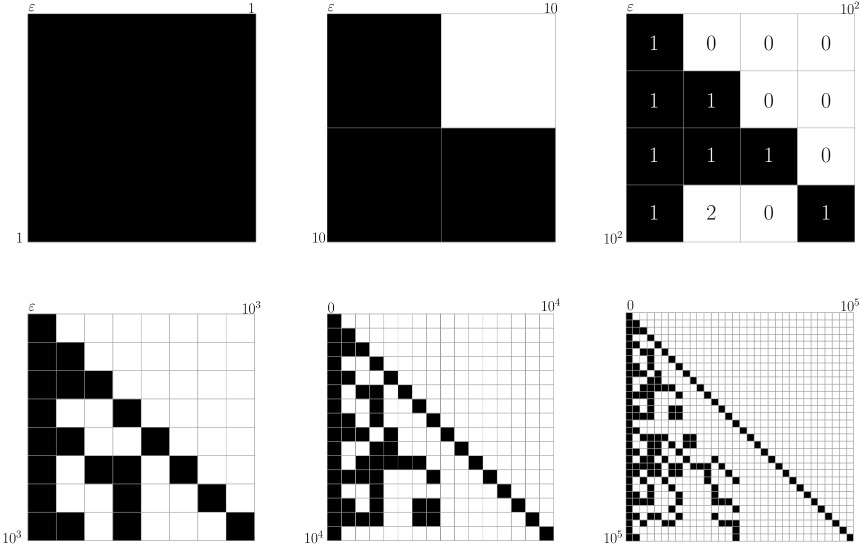
Same construction

- Grid: first 2^n rows and columns of P_2
- Color each square in
  - white if $\binom{\mathrm{rep}_2(m)}{\mathrm{rep}_2(k)} \equiv 0 \bmod 2$
  - black if $\binom{\mathrm{rep}_2(m)}{\mathrm{rep}_2(k)} \equiv 1 \bmod 2$
- Normalize by a homothety of ratio $1/2^n$ (bring into $[0,1]^2$) $\rightsquigarrow$ sequence of compact sets belonging to $[0,1]^2$

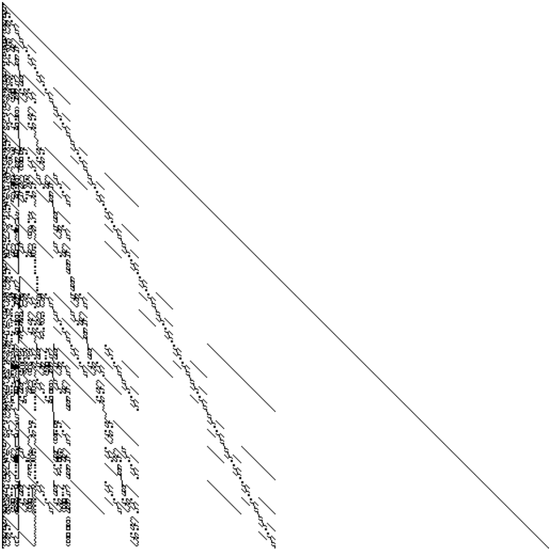


$\mathbb{N}^2 \cap [0, 2^n]^2$

$$U_n = \frac{1}{2^n} \bigcup_{\substack{u,v \in L_2,\ |u|,|v| \le n \\ \binom{u}{v} \equiv 1 \bmod 2}} \mathrm{val}_2(v,u) + [0,1]^2$$


ε
1
1
ε
10
10
ε
10^2
1 0 0 0
1 1 0 0
1 1 1 0
1 2 0 1
10^2
ε
10^3
10^3
0
10^4
10^4
0
10^5
10^5

The element U_9



Lines of different slopes: $1, 2, 4, 8, 16, \ldots$

The $(\star)$ condition

$(\star)$

$$(u,v) \text{ satisfies } (\star) \text{ iff } \begin{cases} u, v \neq \varepsilon \\ \binom{u}{v} \equiv 1 \bmod 2 \\ \binom{u}{v0} = 0 = \binom{u}{v1} \end{cases}$$

Example: $(u,v) = (101, 11)$ satisfies $(\star)$

$$\binom{101}{11} = 1 \qquad \binom{101}{110} = 0 \qquad \binom{101}{111} = 0$$

Lemma: Completion

(u, v) satisfies $(\star) \Rightarrow (u0, v0), (u1, v1)$ satisfy $(\star)$

Proof: Since (u, v) satisfies $(\star)$

$$\binom{u}{v} \equiv 1 \bmod 2, \quad \binom{u}{v0} = 0 = \binom{u}{v1}$$

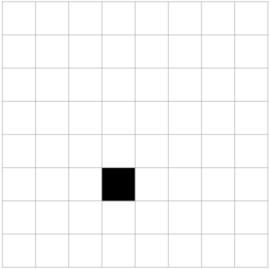
Proof for $(u0, v0)$:

$$\binom{u0}{v0} = \underbrace{\binom{u}{v0}}_{=0 \text{ since } (\star)} + \underbrace{\binom{u}{v}}_{\equiv 1 \bmod 2} \equiv 1 \bmod 2$$

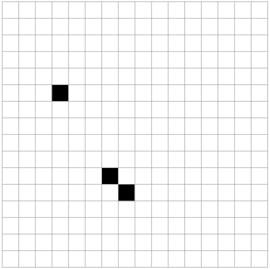
If $\binom{u0}{v00} > 0$ or $\binom{u0}{v01} > 0$, then $v0$ is a subsequence of u.
This contradicts $(\star)$.

Same proof for $(u1, v1)$. $\square$

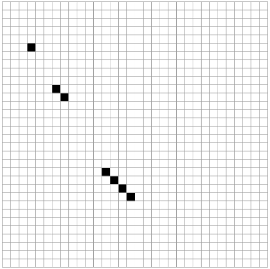
Example: $(u, v) = (101, 11)$ satisfies $(\star) \Rightarrow \binom{u}{v} \equiv 1 \bmod 2$



U_3



U_4

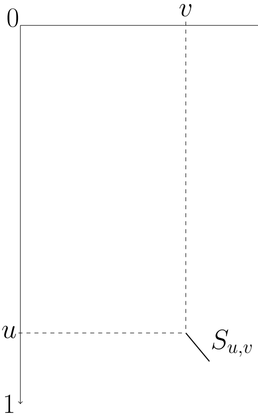


U_5

$\rightsquigarrow$ Creation of a segment of slope 1
Endpoint $(3/8, 5/8) = (\mathrm{val}_2(11)/2^3, \mathrm{val}_2(101)/2^3)$
Length $\sqrt{2} \cdot 2^{-3}$

Segments of slope 1

The $(\star)$ condition describes lines of slope 1 in $[0,1]^2$.



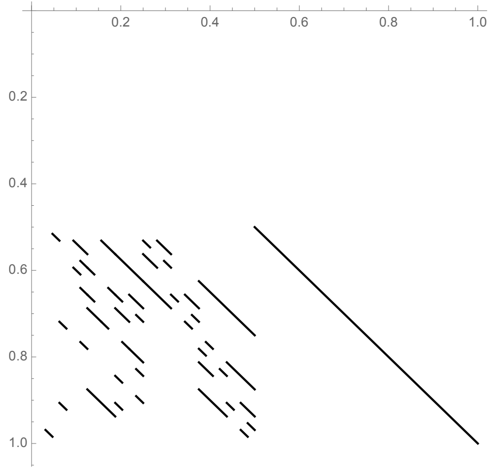
$(u,v) \in L_2 \times L_2$ satisfying $(\star)$
$\rightsquigarrow$ closed segment $S_{u,v}$

- slope 1
- length $\sqrt{2} \cdot 2^{-|u|}$
- origin

$$A_{u,v} = \mathrm{val}_2(v,u)/2^{|u|}$$
$$= (0.0^{|u|-|v|}v, 0.u)$$

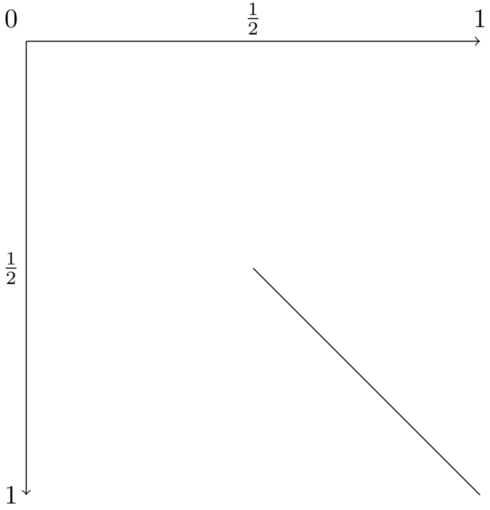
Definition: New compact set containing those lines

$$\mathcal{A}_0 = \overline{\bigcup_{\substack{(u,v)\\ \text{satisfying } (\star)}} S_{u,v}} \subset [0,1]^2$$



Modifying the slope

Two maps $c\colon (x,y) \mapsto (x/2, y/2)$ and $h\colon (x,y) \mapsto (x, 2y)$

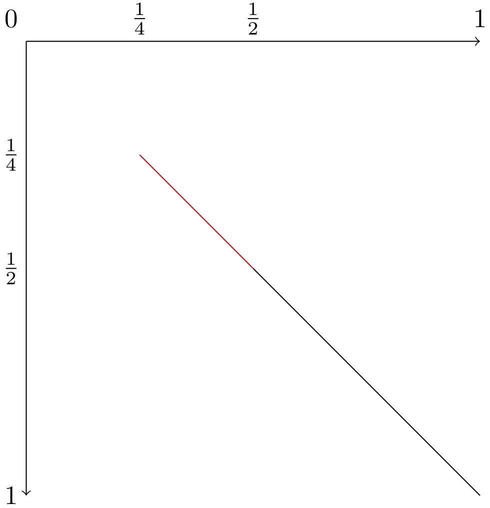


Example: $(1,1)$ satisfies $(\star)$

Segment $S_{1,1}$
endpoint $(1/2, 1/2)$
length $\sqrt{2} \cdot 2^{-1}$

Modifying the slope

Two maps $c\colon (x,y) \mapsto (x/2, y/2)$ and $h\colon (x,y) \mapsto (x, 2y)$

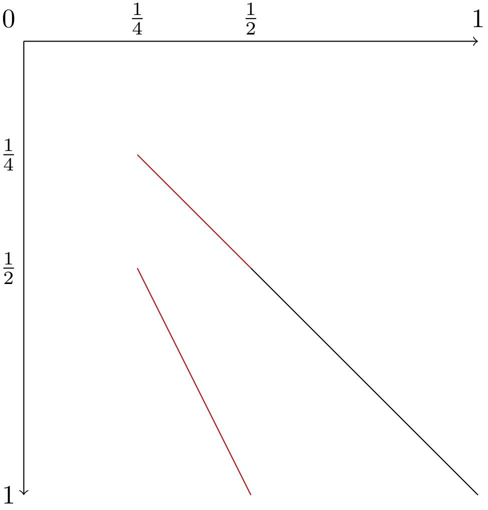


Example: $(1,1)$ satisfies $(\star)$

Segment $S_{1,1}$
endpoint $(1/2, 1/2)$
length $\sqrt{2} \cdot 2^{-1}$

Modifying the slope

Two maps $c\colon (x,y) \mapsto (x/2, y/2)$ and $h\colon (x,y) \mapsto (x, 2y)$



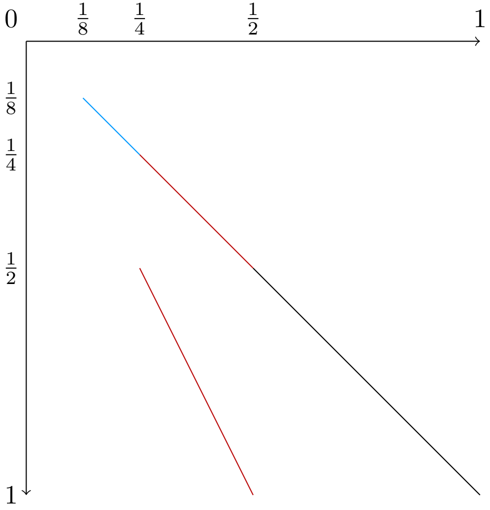
Example: $(1,1)$ satisfies $(\star)$

Segment $S_{1,1}$
endpoint $(1/2, 1/2)$
length $\sqrt{2} \cdot 2^{-1}$

Two maps $c\colon (x,y) \mapsto (x/2, y/2)$ and $h\colon (x,y) \mapsto (x, 2y)$

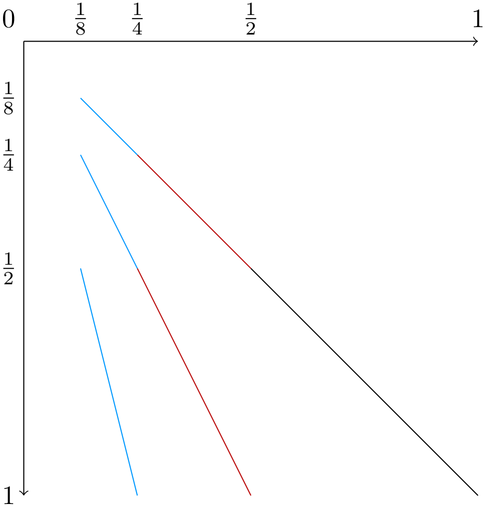
Example: $(1,1)$ satisfies $(\star)$

Segment $S_{1,1}$
endpoint $(1/2, 1/2)$
length $\sqrt{2} \cdot 2^{-1}$



Modifying the slope

Two maps $c\colon (x,y) \mapsto (x/2, y/2)$ and $h\colon (x,y) \mapsto (x, 2y)$

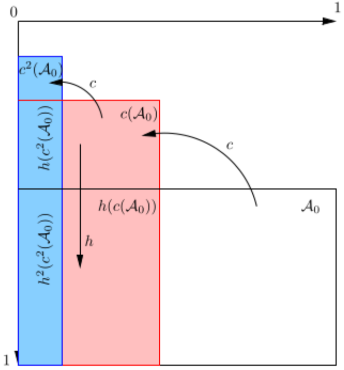


<u>Example</u>: $(1,1)$ satisfies $(\star)$

Segment $S_{1,1}$
endpoint $(1/2, 1/2)$
length $\sqrt{2} \cdot 2^{-1}$

Definition: New compact set containing lines of slopes 1, 2, 2^2, ..., 2^n
$c\colon (x,y) \mapsto (x/2, y/2)$
$h\colon (x,y) \mapsto (x, 2y)$

$$\mathcal{A}_n = \bigcup_{0 \le j \le i \le n} h^j(c^i(\mathcal{A}_0))$$



The compact sets $(\mathcal{A}_n)_{n\ge 0}$ are increasingly nested and their union is bounded.
Thus $(\mathcal{A}_n)_{n\ge 0}$ converges to

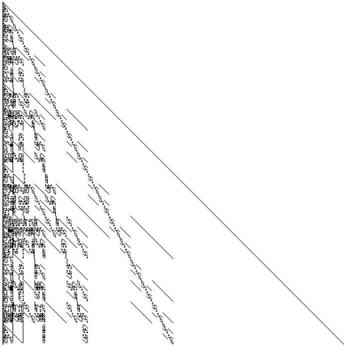
$$\mathcal{L} = \overline{\bigcup_{n\ge 0} \mathcal{A}_n}$$

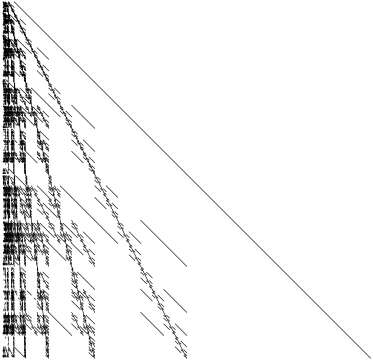
(w.r.t. the Hausdorff distance).

A key result

Theorem (Leroy, Rigo, S., 2016)

The sequence $(U_n)_{n\geq 0}$ of compact sets converges to the compact set $\mathcal{L}$ when n tends to infinity (w.r.t. the Hausdorff distance).





"Simple" characterization of $\mathcal{L}$: $(\star)$ condition

Extension modulo p

Previous result: even and odd coefficients

Theorem (Lucas, 1878)

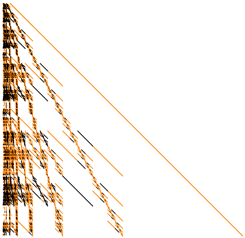
Let p be a prime number.
If $m = m_k p^k + \cdots + m_1 p + m_0$ and $n = n_k p^k + \cdots + n_1 p + n_0$ then

$$\binom{m}{n} \equiv \prod_{i=0}^{k} \binom{m_i}{n_i} \bmod p.$$

Theorem (Leroy, Rigo, S., 2016)

Let p be a prime and $0 < r < p$.
When considering binomial coefficients congruent to $r \pmod p$, the sequence $(U_{n,p,r})_{n \geq 0}$ converges to a well-defined compact set $\mathcal{L}_{p,r}$ (w.r.t. the Hausdorff distance).

Example: $\mathcal{L}_{3,1} \cup \mathcal{L}_{3,2}$

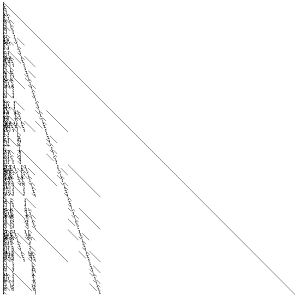


Extension to any integer base

Everything still holds for binomial coefficients $\equiv r \bmod p$ with

- integer base $b \geq 2$
- language L_b of greedy base-b representations of integers
- p a prime
- $r \in \{1, \ldots, p-1\}$

Example: base 3, $\equiv 1 \bmod 2$



Definitions:

- Fibonacci numbers $(F(n))_{n\geq 0}$
  $F(0) = 1,\ F(1) = 2,\ F(n+2) = F(n+1) + F(n) \quad \forall\, n \geq 0$
  $1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144, 233, 377, 610, 987, 1597, \ldots$
- $\mathrm{rep}_F(n)$ greedy Fibonacci representation of $n \in \mathbb{N}_{>0}$ starting with 1
- $\mathrm{rep}_F(0) = \varepsilon$ where ε is the empty word

n	$n = \sum_i c_i F(i)$ with $c_i \in \{0,1\}$	$\mathrm{rep}_F(n)$
0		ε
1	$1 \times F(0)$	1
2	$1 \times F(1) + 0 \times F(0)$	10
3	$1 \times F(2) + 0 \times F(1) + 0 \times F(0)$	100
4	$1 \times F(2) + 0 \times F(1) + 1 \times F(0)$	101
5	$1 \times F(3) + 0 \times F(2) + 0 \times F(1) + 0 \times F(0)$	1000
6	$1 \times F(3) + 0 \times F(2) + 0 \times F(1) + 1 \times F(0)$	1001
$\vdots$	$\vdots$	$\vdots$
		$L_F = 1\{0, 01\}^* \cup \{\varepsilon\}$

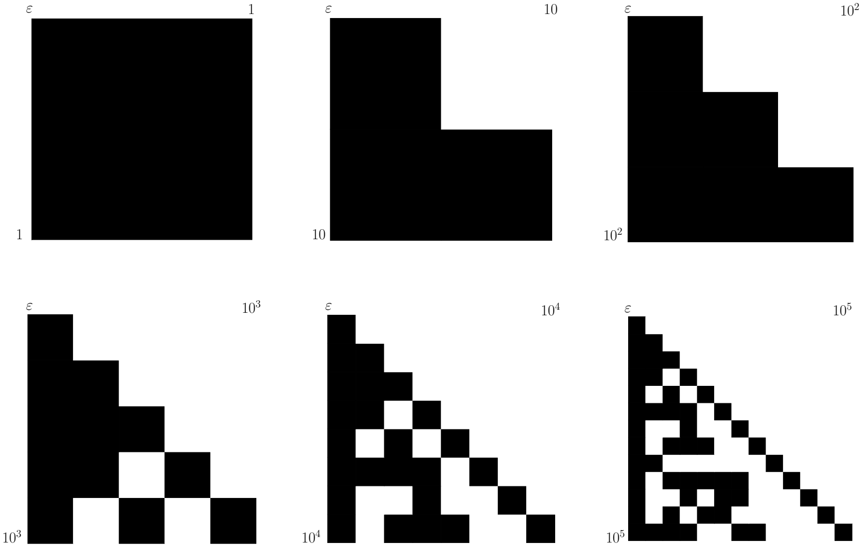
Generalized Pascal triangle P_F in Fibonacci base

$\binom{\mathrm{rep}_F(m)}{\mathrm{rep}_F(k)}$		$\mathrm{rep}_F(k)$								
		ε	1	10	100	101	1000	1001	1010	$\dots$
	ε	1	0	0	0	0	0	0	0	
	1	1	1	0	0	0	0	0	0	
	10	1	1	1	0	0	0	0	0	
	100	1	1	2	1	0	0	0	0	
$\mathrm{rep}_F(m)$	101	1	2	1	0	1	0	0	0	
	1000	1	1	3	3	0	1	0	0	
	1001	1	2	2	1	2	0	1	0	
	1010	1	2	3	1	1	0	0	1	
	$\vdots$									$\ddots$

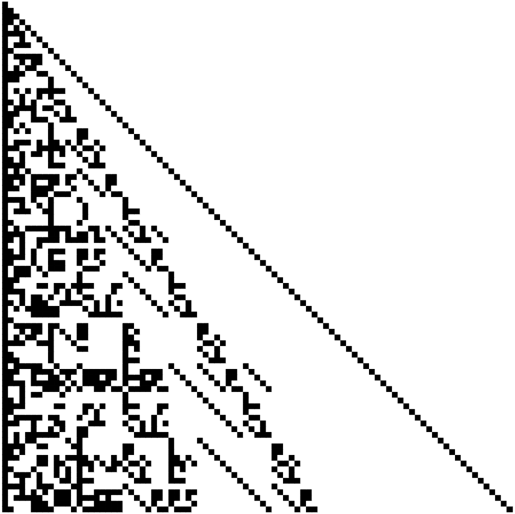
Rule (not local):

$$\binom{ua}{vb} = \binom{u}{vb} + \delta_{a,b}\binom{u}{v}$$

The first six elements of the sequence $(U'_n)_{n\geq 0}$

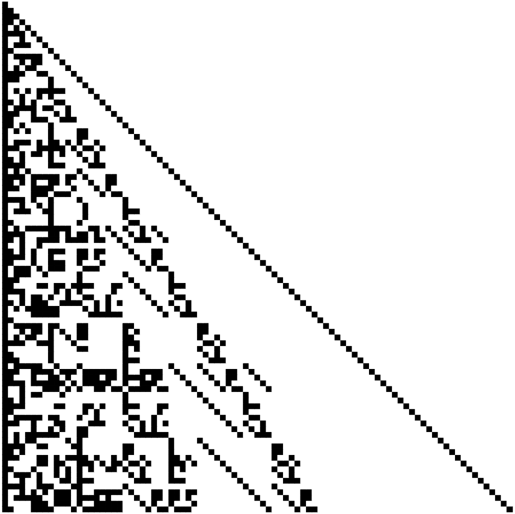


The tenth element



Lines of different slopes:

The tenth element



Lines of different slopes: φ^n, $n \geq 0$, with $\varphi = \frac{1+\sqrt{5}}{2}$ Golden Ratio

The $(\star')$ condition

Recall

- (u, v) satisfies $(\star)$ iff $\binom{u}{v} \equiv 1 \bmod 2$, and $\binom{u}{v0} = 0 = \binom{u}{v1}$.
- (u, v) satisfies $(\star) \Rightarrow (u0, v0), (u1, v1)$ satisfy $(\star)$

Problem: we cannot always add a letter 1 as a suffix in L_F.

Solution: $p(u, v) \in \mathbb{N}$ s.t. $u0^{p(u,v)}w, v0^{p(u,v)}w \in L_F$ for all $w \in 0^* L_F$

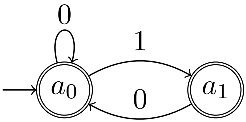
$(\star')$

$$(u, v) \text{ satisfies } (\star') \text{ iff } u = v = \varepsilon \text{ or } \begin{cases} u, v \neq \varepsilon \\ \binom{u0^{p(u,v)}}{v0^{p(u,v)}} \equiv 1 \pmod 2 \\ \binom{u0^{p(u,v)}}{v0^{p(u,v)}a} = 0 \quad \forall\, a \in \{0, 1\}. \end{cases}$$

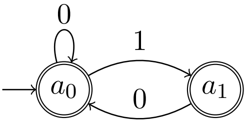
<u>Definition</u>: A deterministic finite automaton (DFA) over an alphabet A is given by a 5-tuple $\mathcal{A} = (Q, q_0, F, A, \delta)$ where

- Q is a finite set of states
- $q_0 \in Q$ is the initial state (incoming arrow)
- $F \subset Q$ is the set of final states (concentric circles)
- $\delta \colon Q \times A \mapsto Q$ is the transition function

<u>Example</u>:



$A = \{0, 1\}$
$Q = \{a_0, a_1\}$
$q_0 = a_0$
$F = \{a_0, a_1\}$
$\delta(a_0, 0) = a_0,\ \delta(a_0, 1) = a_1,\ \delta(a_1, 0) = a_0$



Reading words in the automaton:

100101	$a_0 \xrightarrow{1} a_1 \xrightarrow{0} a_0 \xrightarrow{0} a_0 \xrightarrow{1} a_1 \xrightarrow{0} a_0 \xrightarrow{1} a_1$	✓	accepted
1011	$a_0 \xrightarrow{1} a_1 \xrightarrow{0} a_0 \xrightarrow{1} a_1 \xrightarrow{1} ???$	✗	not accepted
000101	$a_0 \xrightarrow{0} a_0 \xrightarrow{0} a_0 \xrightarrow{0} a_0 \xrightarrow{1} a_1 \xrightarrow{0} a_0 \xrightarrow{1} a_1$	✓	accepted

Accepted language (set of words): 0^*L_F with $L_F = 1\{0,01\}^* \cup \{\varepsilon\}$

Definition: $p(u,v)$ is the smallest p s.t. $\delta(a_0, u0^p) = a_0 = \delta(a_0, v0^p)$

Then $u0^{p(u,v)}w, v0^{p(u,v)}w \in L_F$ for all $w \in 0^*L_F$.

Similar technique

- Completion lemma with the $(\star')$ condition
- Creation of segments of slope 1
- New compact set $\mathcal{A}'_0$ containing those lines

$$\mathcal{A}'_0 = \overline{\bigcup_{\substack{(u,v) \\ \text{satisfying } (\star')}} S_{u,v}} \subset [0,1]^2$$

- Modification of the slopes with $c\colon (x,y) \mapsto (x/\varphi, y/\varphi)$ and $h\colon (x,y) \mapsto (x, \varphi y)$
- New compact set $\mathcal{A}'_n$ containing lines of slopes $1, \varphi, \varphi^2, \ldots, \varphi^n$

$$\mathcal{A}'_n = \bigcup_{0 \le j \le i \le n} h^j(c^i(\mathcal{A}'_0))$$

- $(\mathcal{A}'_n)_{n\ge 0}$ converges to

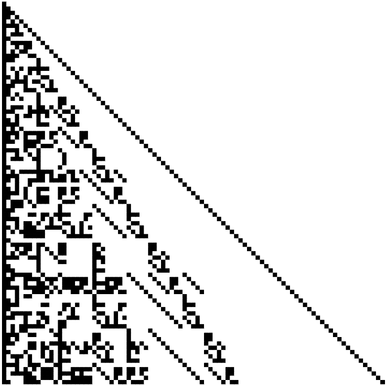
$$\mathcal{L}' = \overline{\bigcup_{n\ge 0} \mathcal{A}'_n}$$

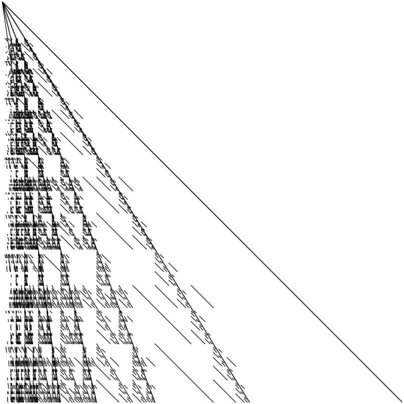
(w.r.t. the Hausdorff distance).

Similar convergence result

Theorem (S., 2018)

The sequence $(U'_n)_{n\geq 0}$ of compact sets converges to the compact set $\mathcal{L}'$ when n tends to infinity (w.r.t. the Hausdorff distance).





"Simple" characterization of $\mathcal{L}'$: $(\star')$ condition

Linear numeration system

Definition

A *numeration system* is a sequence $U = (U(n))_{n\geq 0}$ of integers s.t.

- U increasing
- $U(0) = 1$
- $\sup_{n\geq 0} \frac{U(n+1)}{U(n)}$ bounded by a constant $\rightsquigarrow$ finite alphabet.

A numeration system U is *linear* if $\exists k \geq 1, \exists a_0, \ldots, a_{k-1} \in \mathbb{Z}$ s.t.

$$U(n+k) = a_{k-1}\, U(n+k-1) + \cdots + a_0\, U(n) \quad \forall\, n \geq 0.$$

Greedy representation in $(U(n))_{n\geq 0}$:

$$n = \sum_{i=0}^{\ell} c_i U(i) \qquad \text{with} \qquad \sum_{i=0}^{j-1} c_i U(i) < U(j)$$

$$\text{rep}_U(n) = c_\ell \cdots c_0 \in \underbrace{L_U = \text{rep}_U(\mathbb{N})}_{\text{numeration language}}$$

Example: integer base $(b^n)_{n\geq 0}$ with $b \in \mathbb{N}_{>1}$

Fibonacci numeration system $(F(n))_{n\geq 0}$

Parry numbers

$\beta \in \mathbb{R}_{>1}$ $\qquad A_\beta = \{0, 1, \ldots, \lceil \beta \rceil - 1\}$

$$x \in [0,1] \rightsquigarrow x = \sum_{j=1}^{+\infty} c_j \beta^{-j}, \quad c_j \in A_\beta$$

Greedy way: $c_j \beta^{-j} + c_{j+1}\beta^{-j-1} + \cdots < \beta^{-(j-1)}$
β-expansion of x: $d_\beta(x) = c_1 c_2 c_3 \cdots$

Definition

$\beta \in \mathbb{R}_{>1}$ is a *Parry number* if $d_\beta(1)$ is ultimately periodic.

Example: $b \in \mathbb{N}_{>1}$: $d_b(1) = (b-1)^\omega$
Golden ratio φ: $d_\varphi(1) = 110^\omega$

Parry numeration system

Parry number $\beta \in \mathbb{R}_{>1} \rightsquigarrow$ linear numeration system $(U_\beta(n))_{n\geq 0}$

- $d_\beta(1) = t_1 \cdots t_m 0^\omega$

$$\begin{array}{lcll} U_\beta(0) & = & 1 & \\ U_\beta(i) & = & t_1 U_\beta(i-1) + \cdots + t_i U_\beta(0) + 1 & \forall\, 1 \leq i \leq m-1 \\ U_\beta(n) & = & t_1 U_\beta(n-1) + \cdots + t_m U_\beta(n-m) & \forall\, n \geq m \end{array}$$

- $d_\beta(1) = t_1 \cdots t_m (t_{m+1} \cdots t_{m+k})^\omega$

$$\begin{array}{lcll} U_\beta(0) & = & 1 & \\ U_\beta(i) & = & t_1 U_\beta(i-1) + \cdots + t_i U_\beta(0) + 1 & \forall\, 1 \leq i \leq m+k-1 \\ U_\beta(n) & = & t_1 U_\beta(n-1) + \cdots + t_{m+k} U_\beta(n-m-k) & \forall\, n \geq m+k \\ & & +U_\beta(n-k) & \\ & & -t_1 U_\beta(n-k-1) - \cdots - t_m U_\beta(n-m-k) & \end{array}$$

Examples:

$b \in \mathbb{N}_{>1} \rightsquigarrow (b^n)_{n\geq 0}$ base b

Golden ratio $\varphi \rightsquigarrow (F(n))_{n\geq 0}$ Fibonacci numeration system

Extension to Parry numeration systems

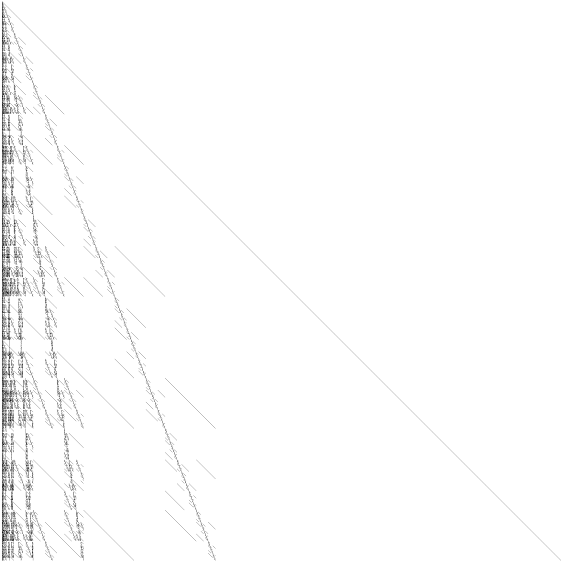
- Parry number $\beta \in \mathbb{R}_{>1}$
- Parry numeration system $(U_\beta(n))_{n\geq 0}$
- Numeration language L_{U_β}
- Generalized Pascal triangle P_β in $(U_\beta(n))_{n\geq 0}$ indexed by words of L_{U_β}
- Sequence of compact sets extracted from P_β (first $U_\beta(n)$ rows and columns of P_β)
- Convergence to a limit object (same technique)
  - Lines of different slopes: β^n, $n \geq 0$
  - $(\star')$ condition and description of segments of slope 1
  - Two maps $c\colon (x,y) \mapsto (x/\beta, y/\beta)$ and $h\colon (x,y) \mapsto (x, \beta y)$
  - Sequence of sets $\mathcal{A}_n^\beta$ containing lines of slopes 1, β, β^2, $\ldots$, β^n
  - $\mathcal{A}_n^\beta$ converges to

$$\mathcal{L}^\beta = \overline{\bigcup_{n\geq 0} \mathcal{A}_n^\beta}$$

- Works modulo any prime number

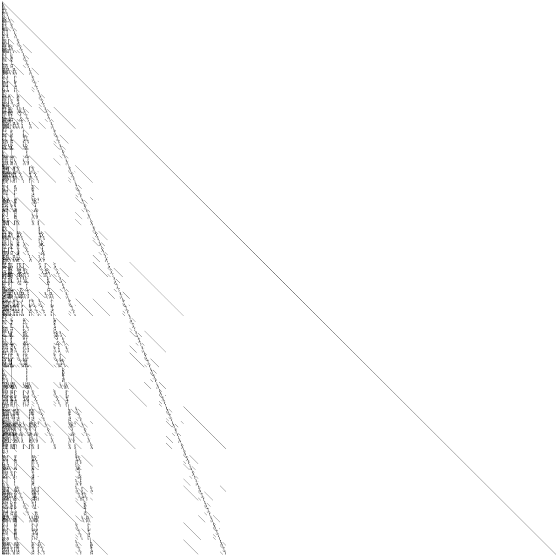
Example 1

φ^2



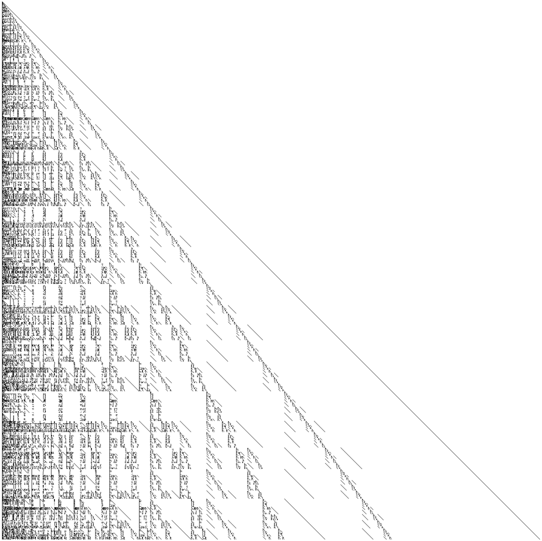
Example 2

$\beta_1 \approx 2.47098$ dominant root of $P(X) = X^4 - 2X^3 - X^2 - 1$



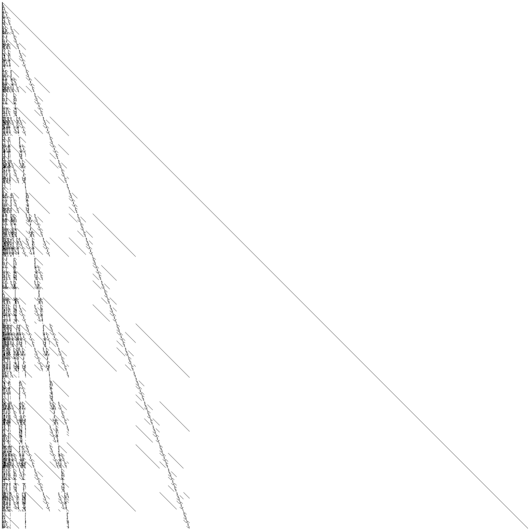
Example 3

$\beta_2 \approx 1.38028$ dominant root of $P(X) = X^4 - X^3 - 1$



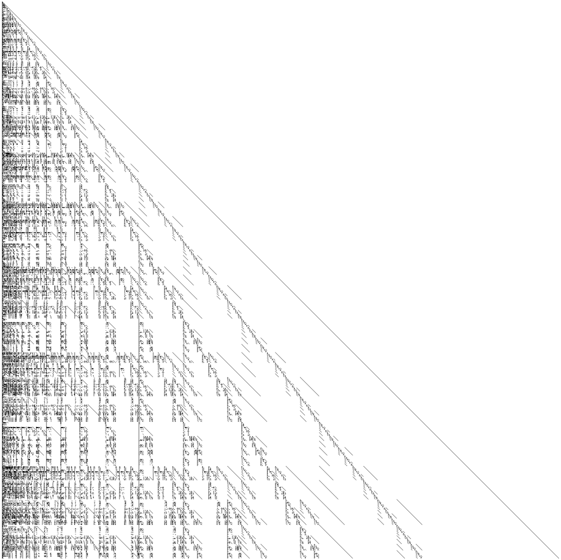
Example 4

$\beta_3 \approx 2.80399$ dominant root of $P(X) = X^4 - 2X^3 - 2X^2 - 2$



Example 5

$\beta_4 \approx 1.32472$ dominant root of polynomial $P(X) = X^5 - X^4 - 1$



In this talk:

Numeration system	Generalized Pascal triangle	Convergence mod p
Base 2	✓	✓
Integer base	✓	✓
Fibonacci	✓	✓
Parry	✓	✓

Other related work

- Regularity of the sequence counting subword occurrences: result for any integer base b and the Fibonacci numeration system
- Behavior of the summatory function: result for any integer base b (exact behavior) and the Fibonacci numeration system (asymptotics)

References

- J. Berstel, D. Perrin, The origins of combinatorics on words, *European J. Combin.* **28** (2007), 996–1022.
- F. von Haeseler, H.-O. Peitgen, G. Skordev, Pascal's triangle, dynamical systems and attractors, *Ergod. Th. & Dynam. Sys.* **12** (1992), 479–486.
- J. Leroy, M. Rigo, M. Stipulanti, Generalized Pascal triangle for binomial coefficients of words, *Adv. in Appl. Math.* **80** (2016), 24–47.
- J. Leroy, M. Rigo, M. Stipulanti, Counting the number of non-zero coefficients in rows of generalized Pascal triangles, *Discrete Math.* **340** (2017), 862–881.
- J. Leroy, M. Rigo, M. Stipulanti, Behavior of digital sequences through exotic numeration systems, *Electron. J. Combin.* **24** (2017), no. 1, Paper 1.44, 36 pp.
- J. Leroy, M. Rigo, M. Stipulanti, Counting Subword Occurrences in Base-b Expansions, *Integers* **18A** (2018), Paper No. A13, 32 pp.
- M. Lothaire, *Combinatorics On Words*, Cambridge Mathematical Library, Cambridge University Press, Cambridge, 1997. Corrected reprint of the 1983 original.
- É. Lucas, Théorie des fonctions numériques simplement périodiques, *Amer. J. Math.* **1** (1878), 197–240.
- M. Stipulanti, Convergence of Pascal-Like Triangles in Parry–Bertrand Numeration Systems, *Theoret. Comput. Sci.* **758** (2019), 42–60.