An extension of the Pascal triangle and the Sierpiński gasket to finite words

Joint work with Julien Leroy and Michel Rigo (ULiège)

Manon Stipulanti (ULiège)

FRIA grantee

LaBRI, Bordeaux (France)

December 15, 2017

Classical Pascal triangle

$\binom{m}{k}$		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	
$\vdots$									$\ddots$

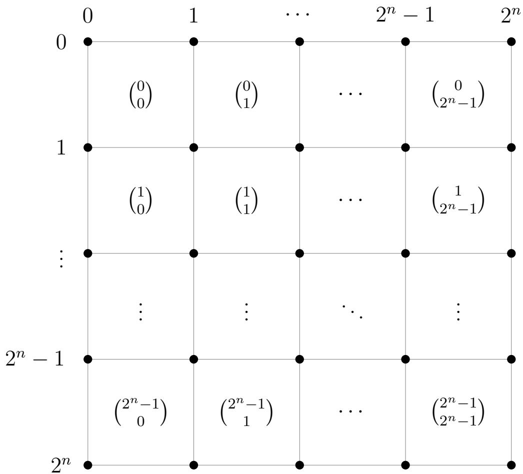
Usual binomial coefficients of integers:

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

Pascal's rule:

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

- Grid: intersection between $\mathbb{N}^2$ and $[0, 2^n] \times [0, 2^n]$



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

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

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

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

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

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

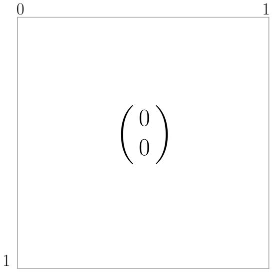
  in
  - white if $\binom{m}{k} \equiv 0 \bmod 2$
  - black if $\binom{m}{k} \equiv 1 \bmod 2$
- Normalize by a homothety of ratio $1/2^n$
  (bring into $[0,1] \times [0,1]$)
  $\rightsquigarrow$ sequence belonging to $[0,1] \times [0,1]$

What happens for $n \in \{0, 1\}$

$n = 0$

What happens for $n \in \{0, 1\}$

$n = 0$



What happens for $n \in \{0, 1\}$

$n = 0$

0 1

$$\binom{0}{0}$$

1

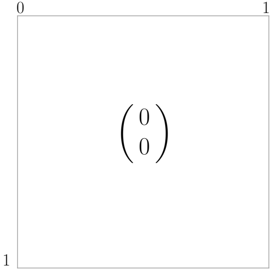
0 1

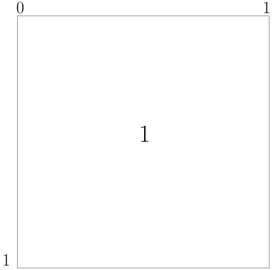
$$1$$

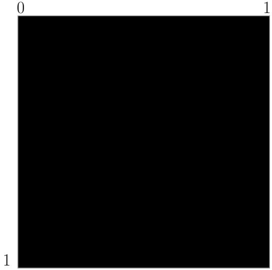
1

What happens for $n \in \{0, 1\}$

$n = 0$

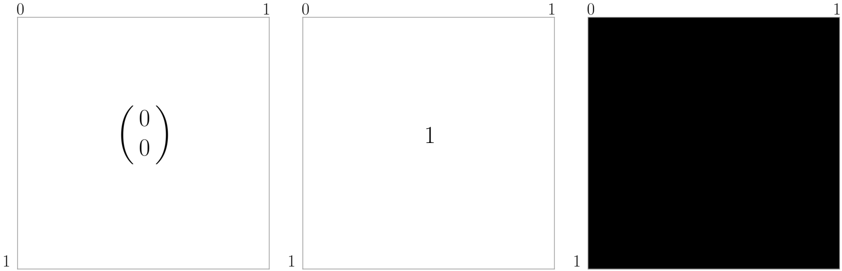






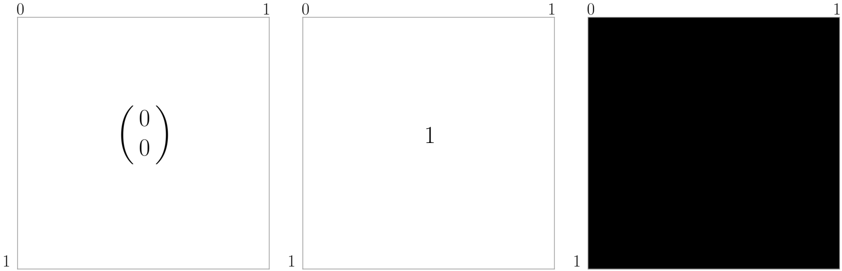
What happens for $n \in \{0, 1\}$

$n = 0$

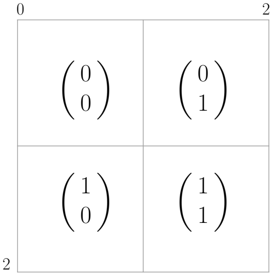


$n = 1$

$n = 0$



$n = 1$



What happens for $n \in \{0, 1\}$

$n = 0$

$\binom{0}{0}$

1

$n = 1$

$\binom{0}{0}$	$\binom{0}{1}$
$\binom{1}{0}$	$\binom{1}{1}$

1	0
1	1

Axis labels: 0, 1 (top) and 1 (left) for $n=0$; 0, 2 (top) and 2 (left) for $n=1$.

$n = 0$

0 1

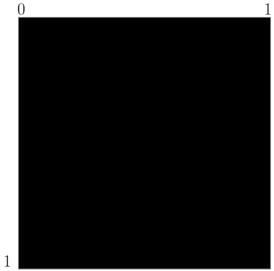
$\binom{0}{0}$

1

0 1

1

1



$n = 1$

0 2

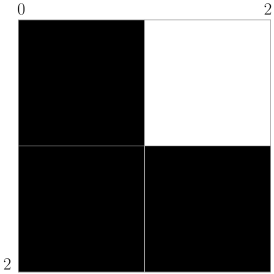
$\binom{0}{0}$	$\binom{0}{1}$
$\binom{1}{0}$	$\binom{1}{1}$

2

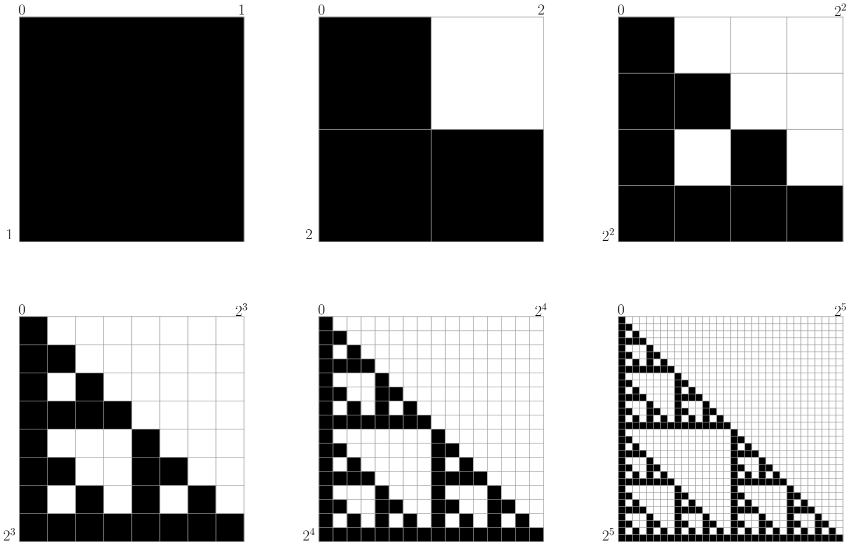
0 2

1	0
1	1

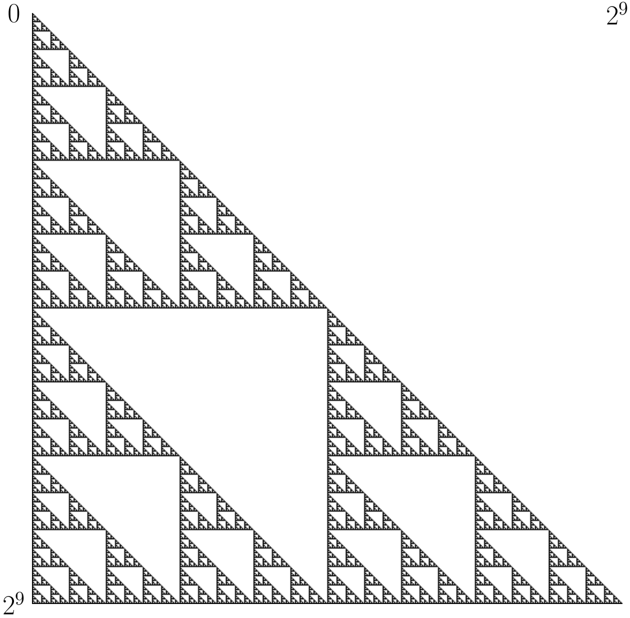
2



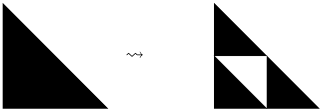
The first six elements of the sequence



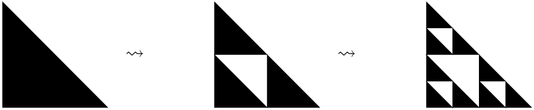
The tenth element of the sequence



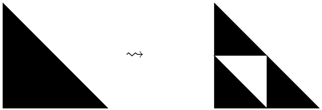
The Sierpiński gasket

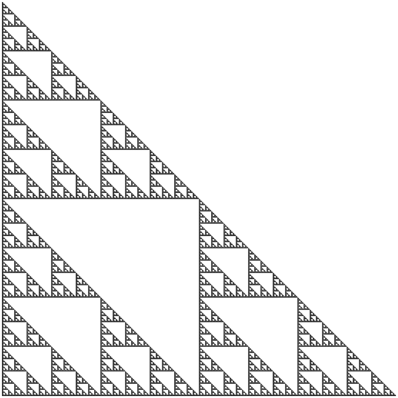


The Sierpiński gasket



The Sierpiński gasket





Folklore fact

The latter sequence converges to the Sierpiński gasket when n tends to infinity (for the Hausdorff distance).

Folklore fact

The latter sequence converges to the Sierpiński gasket when n tends to infinity (for 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

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

Remark
(von Haeseler, Peitgen, Skordev, 1992)

The sequence also converges for other modulos.
For instance, the sequence converges when the Pascal triangle is considered modulo p^s where p is a prime and s is a positive integer.

This talk...

- Part I: Generalized Pascal triangle
- Part II: Counting Subword Occurrences
- Part III: Behavior of the Summatory Function

Part I: Generalized Pascal triangle

New idea

Replace usual binomial coefficients of integers by binomial coefficients of finite words

Binomial coefficient of finite words

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

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

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 (Lothaire, 1997)

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: $u = 101001$ $\quad v = 101$

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 (Lothaire, 1997)

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

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 (Lothaire, 1997)

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: $u = \mathbf{10}100\mathbf{1}$ $\qquad v = 101 \qquad$ 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 (Lothaire, 1997)

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: $u = \mathbf{101001}$ $v = 101$ 3 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 (Lothaire, 1997)

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: $u = \mathbf{101001}$ $v = 101$ 4 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 (Lothaire, 1997)

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: $u = 101001$ $\quad v = 101$ $\quad$ 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 (Lothaire, 1997)

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: $u = 10\mathbf{1001}$ $\quad v = 101$ $\quad$ 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 (Lothaire, 1997)

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

With a one-letter alphabet $\{a\}$

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

Definitions:

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

n		$\text{rep}_2(n)$
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
$\vdots$	$\vdots$	$\vdots$
		$\{\varepsilon\} \cup 1\{0,1\}^*$

Generalized Pascal triangle in base 2

$\binom{u}{v}$		v							
		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	
u	11	1	**2**	0	**1**	0	0	0	0	
	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	
	$\vdots$									$\ddots$

Binomial coefficient of finite words: $\binom{u}{v}$

Rule (not local):

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

Generalized Pascal triangle in base 2

		v								
		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	
u	**11**	**1**	**2**	0	**1**	0	0	0	0	
	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**	
	$\vdots$									$\ddots$

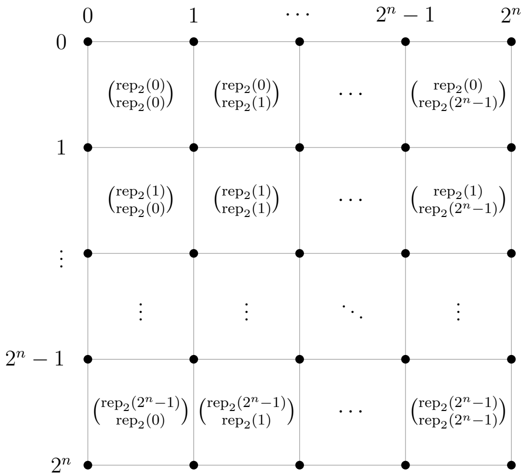
The classical Pascal triangle

Questions:

- After coloring and normalization can we expect the convergence to an analogue of the Sierpiński gasket?
- Could we describe this limit object ?

Same construction

- Grid: intersection between $\mathbb{N}^2$ and $[0, 2^n] \times [0, 2^n]$



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

- Color the grid:
  Color the first 2^n rows and columns of the generalized Pascal triangle

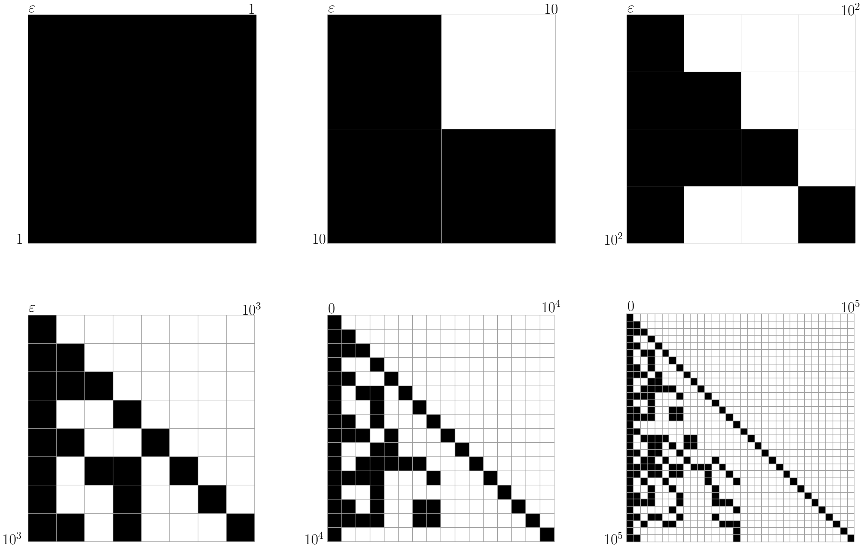
$$\left(\binom{\text{rep}_2(m)}{\text{rep}_2(k)} \bmod 2\right)_{0\leq m,k<2^n}$$

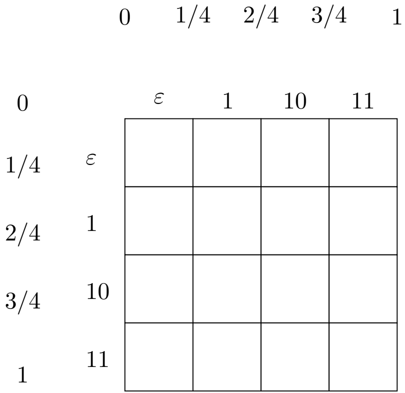
  in
  - white if $\binom{\text{rep}_2(m)}{\text{rep}_2(k)} \equiv 0 \bmod 2$
  - black if $\binom{\text{rep}_2(m)}{\text{rep}_2(k)} \equiv 1 \bmod 2$
- Normalize by a homothety of ratio $1/2^n$
  (bring into $[0,1] \times [0,1]$)
  $\rightsquigarrow$ sequence $(U_n)_{n\geq 0}$ belonging to $[0,1] \times [0,1]$

$$U_n := \frac{1}{2^n} \bigcup_{\substack{u,v\in\{\varepsilon\}\cup 1\{0,1\}^* \\ \text{s.t. } \binom{u}{v}\equiv 1 \bmod 2}} \{(\text{val}_2(v), \text{val}_2(u)) + Q\}$$

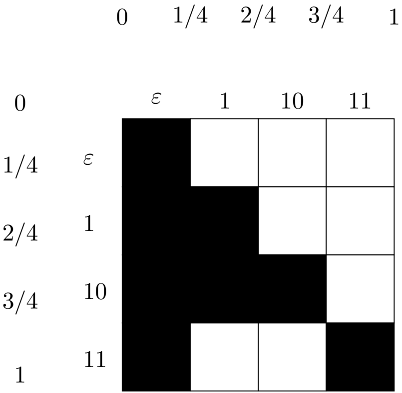
$$Q := [0,1] \times [0,1]$$

The elements $U_0, \ldots, U_5$




0
1/4
2/4
3/4
1
ε
1
10
11
0
1/4
2/4
3/4
1
ε
1
10
11

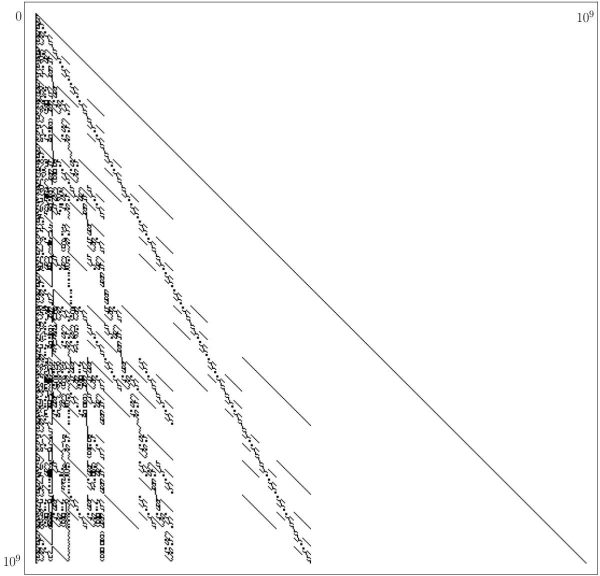
The element U_2



$$\varepsilon \rightsquigarrow 0, \quad 1 \rightsquigarrow 1/4, \quad 10 \rightsquigarrow 2/4 = 1/2, \quad 11 \rightsquigarrow 3/4$$

$$w \in \{\varepsilon\} \cup 1\{0,1\}^* \text{ with } |w| \leq 2 \rightsquigarrow \frac{\mathrm{val}_2(w)}{2^2}$$

The element U_9



Lines of different slopes...

The $(\star)$ condition

$(\star)$

(u,v) satisfies $(\star)$ 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 $(\star)$

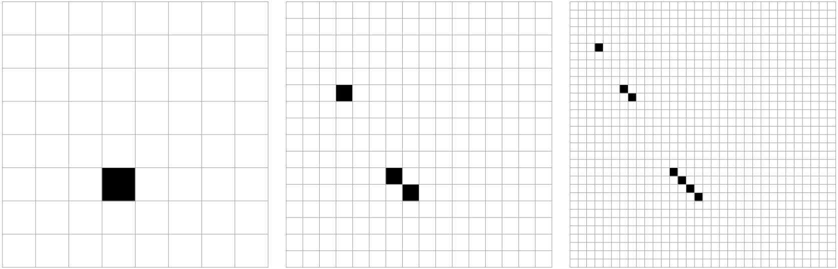
Proof:

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

If $\binom{u0}{v00} > 0$ or $\binom{u0}{v01} > 0$, then $v0$ is a subsequence of u, which contradicts $(\star)$.
Same proof for $(u1, v1)$.

$\square$

Example: $u = 101$, $v = 11$



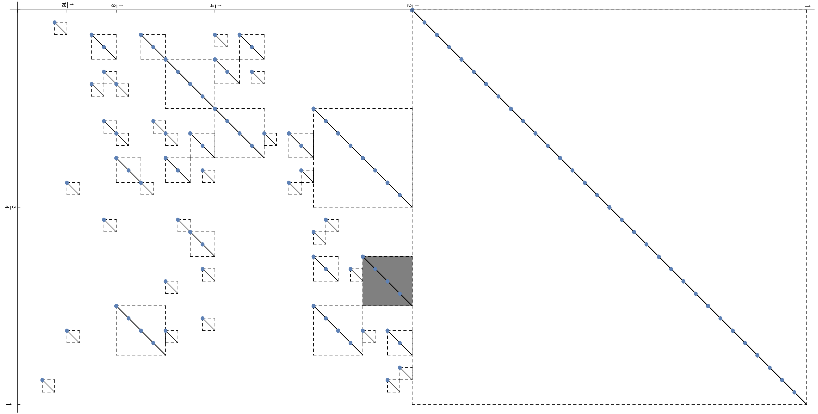
U_3 U_4 U_5

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

$$S_{u,v} \subset [0,1] \times [1/2,1] \quad \text{endpoint } (\mathrm{val}_2(v)/2^{|u|}, \mathrm{val}_2(u)/2^{|u|})$$
$$\text{length } \sqrt{2} \cdot 2^{-|u|}$$

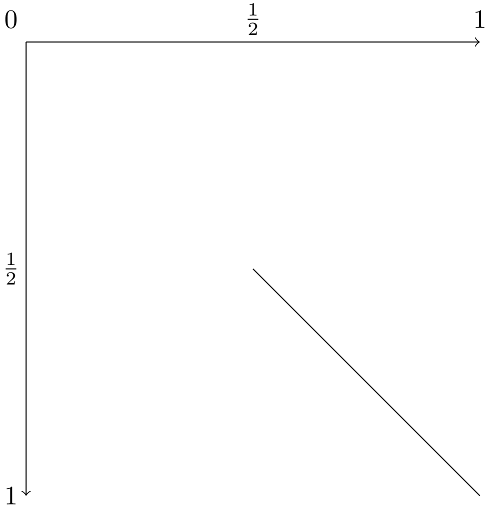
Definition: Set of segments of slope 1

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



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

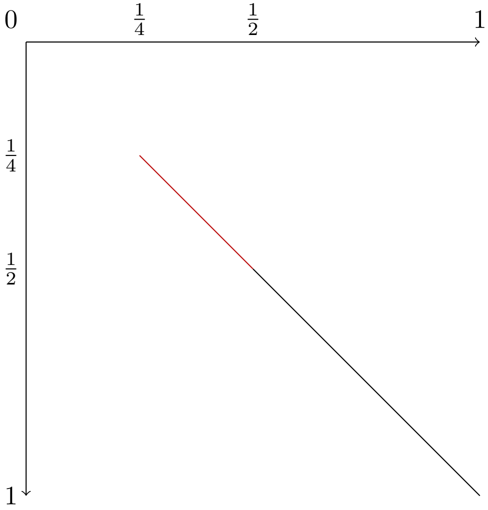
Segment $S_{1,1}$ endpoint $(1/2,1/2)$ length $\sqrt{2}\cdot 2^{-1}$



Modifying the slope

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

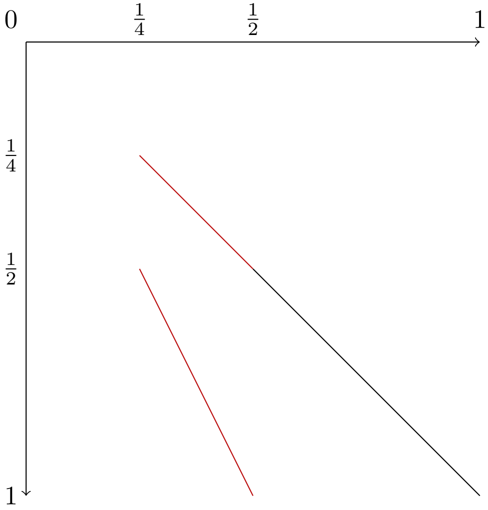
Segment $S_{1,1}$ endpoint $(1/2, 1/2)$ length $\sqrt{2} \cdot 2^{-1}$



Modifying the slope

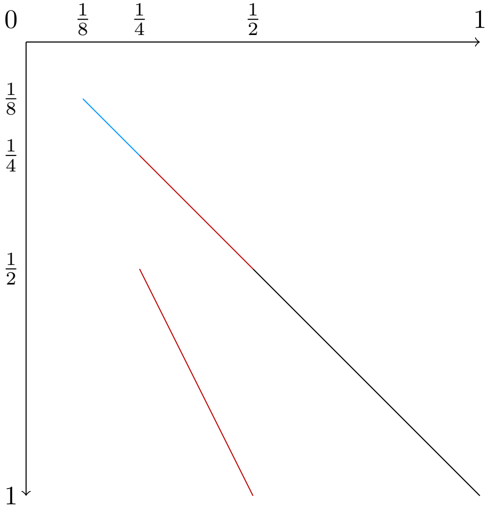
Example: $(1,1)$ satisfies $(\star)$

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



Modifying the slope

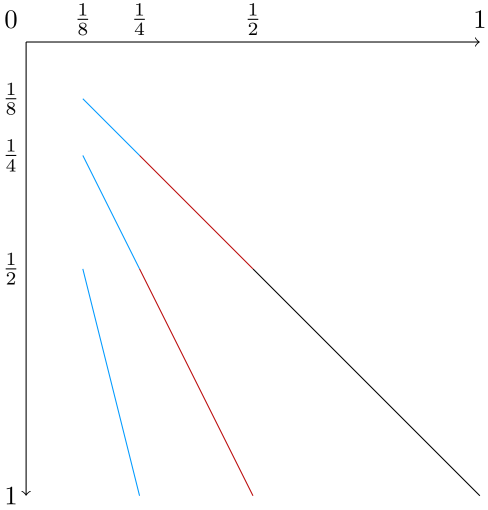
Example: $(1, 1)$ satisfies $(\star)$
Segment $S_{1,1}$ endpoint $(1/2, 1/2)$ length $\sqrt{2} \cdot 2^{-1}$



Modifying the slope

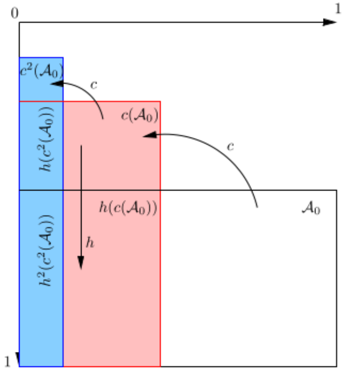
Example: $(1,1)$ satisfies $(\star)$

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



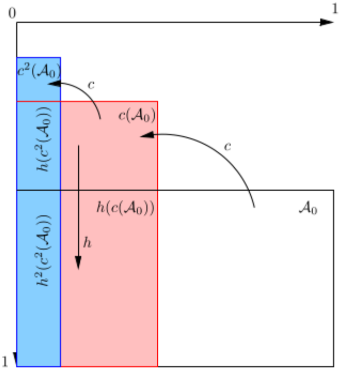
Definition: Set of segments of different slopes
$c : (x, y) \mapsto (x/2, y/2)$ (homothety of center $(0, 0)$, ratio $1/2$)
$h : (x, y) \mapsto (x, 2y)$

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



Definition: Set of segments of different slopes
$c : (x, y) \mapsto (x/2, y/2)$ (homothety of center $(0, 0)$, ratio $1/2$)
$h : (x, y) \mapsto (x, 2y)$

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



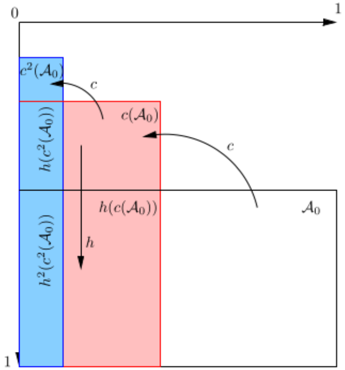
Lemma: $(\mathcal{A}_n)_{n \ge 0}$ is a Cauchy sequence

Definition: Set of segments of different slopes

$c : (x, y) \mapsto (x/2, y/2)$ (homothety of center $(0,0)$, ratio $1/2$)

$h : (x, y) \mapsto (x, 2y)$

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



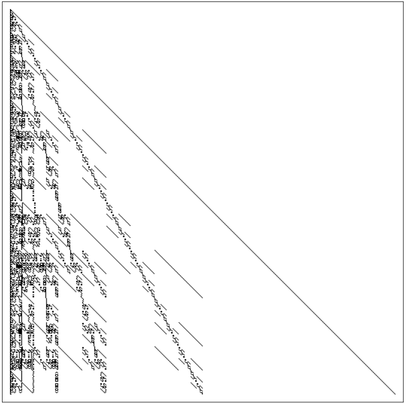
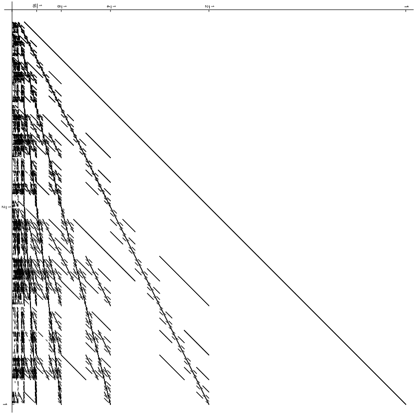
Lemma: $(\mathcal{A}_n)_{n \ge 0}$ is a Cauchy sequence

Definition: Limit object $\mathcal{L}$

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 (for the Hausdorff distance).

"Simple" characterization of $\mathcal{L}$: topological closure of a union of segments described through a "simple" combinatorial property

Extension modulo p

Simplicity: coloring the cells of the grids regarding their parity

Extension using Lucas' theorem

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

- base-2 expansions of integers
- p a prime
- $r \in \{1, \ldots, p-1\}$

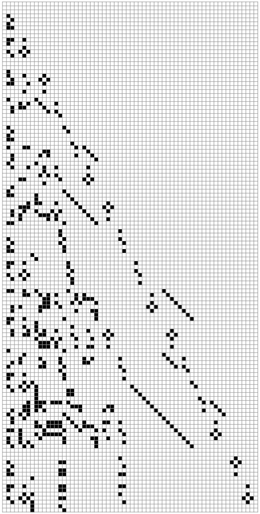
Theorem (Lucas, 1878)

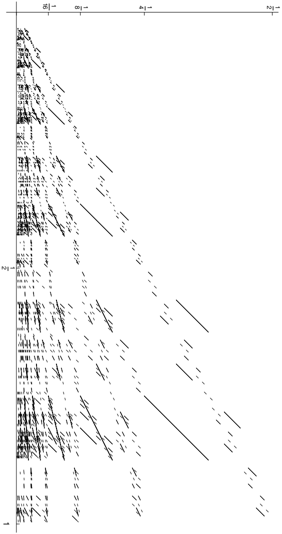
Let p be a prime number.
If $m = m_k p^k + \cdots + m_1 p + m_0$ and $n = n_k p^k + \cdots + n_1 p + n_0$, then

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

Example with $p = 3$, $r = 2$

Left: binomial coefficients $\equiv 2 \bmod 3$
Right: estimate of the corresponding limit object





Part II: Counting Subword Occurrences

Generalized Pascal triangle in base 2

	$\binom{u}{v}$	ε	1	10	11	100	101	110	111	n	$S_2(n)$
	ε	1	0	0	0	0	0	0	0	0	**1**
	1	1	1	0	0	0	0	0	0	1	**2**
	10	1	1	1	0	0	0	0	0	2	**3**
u	11	1	2	0	1	0	0	0	0	3	**3**
	100	1	1	2	0	1	0	0	0	4	**4**
	101	1	2	1	1	0	1	0	0	5	**5**
	110	1	2	2	1	0	0	1	0	6	**5**
	111	1	3	0	3	0	0	0	1	7	**4**

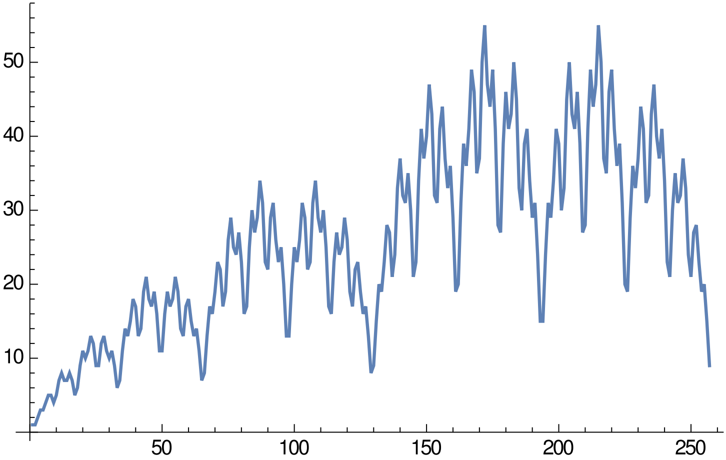
(Columns ε through 111 are values of v.)

Definition:

$$S_2(n) = \#\left\{m \in \mathbb{N} \mid \binom{\mathrm{rep}_2(n)}{\mathrm{rep}_2(m)} > 0\right\} \quad \forall n \geq 0$$

$$= \#\text{ (scattered) subwords in } \{\varepsilon\} \cup 1\{0,1\}^* \text{ of } \mathrm{rep}_2(n)$$

The sequence $(S_2(n))_{n\geq 0}$ in the interval $[0, 256]$



Palindromic structure $\rightsquigarrow$ regularity

- 2-*kernel* of $s = (s(n))_{n\geq 0}$

$$\begin{aligned}\mathcal{K}_2(s) = &\{(s(n))_{n\geq 0}, (s(2n))_{n\geq 0}, (s(2n+1))_{n\geq 0}, (s(4n))_{n\geq 0},\\ &(s(4n+1))_{n\geq 0}, (s(4n+2))_{n\geq 0}, \ldots\}\\ = &\{(s(2^i n + j))_{n\geq 0} \mid i \geq 0 \text{ and } 0 \leq j < 2^i\}\end{aligned}$$

Regularity in the sense of Allouche and Shallit (1992)

- 2-*kernel* of $s = (s(n))_{n\geq 0}$

$$\begin{aligned}\mathcal{K}_2(s) =& \{(s(n))_{n\geq 0}, (s(2n))_{n\geq 0}, (s(2n+1))_{n\geq 0}, (s(4n))_{n\geq 0},\\ & (s(4n+1))_{n\geq 0}, (s(4n+2))_{n\geq 0}, \ldots\}\\ =& \{(s(2^i n + j))_{n\geq 0} \mid i \geq 0 \text{ and } 0 \leq j < 2^i\}\end{aligned}$$

- 2-*regular* if there exist

$$(t_1(n))_{n\geq 0}, \ldots, (t_\ell(n))_{n\geq 0}$$

s.t. each $(t(n))_{n\geq 0} \in \mathcal{K}_2(s)$ is a $\mathbb{Z}$-linear combination of the t_j's

The case of $(S_2(n))_{n\geq 0}$

Theorem (Leroy, Rigo, S., 2017)

The sequence $(S_2(n))_{n\geq 0}$ satisfies, for all $n \geq 0$,

$$\begin{aligned} S_2(2n+1) &= 3\,S_2(n) - S_2(2n) \\ S_2(4n) &= -S_2(n) + 2\,S_2(2n) \\ S_2(4n+2) &= 4\,S_2(n) - S_2(2n). \end{aligned}$$

Proof using a special tree structure...

Corollary (Leroy, Rigo, S., 2017)

$(S_2(n))_{n\geq 0}$ is 2-regular.

Proof: Generators $(S_2(n))_{n\geq 0}$ and $(S_2(2n))_{n\geq 0}$

$S_2(n) = S_2(n)$
$S_2(2n) = S_2(2n)$
$S_2(2n+1) = 3S_2(n) - S_2(2n)$
$S_2(4n) = -S_2(n) + 2S_2(2n)$

$$\begin{aligned} S_2(4n+1) &= S_2(2(2n)+1) = 3S_2(2n) - S_2(4n) \\ &= 3S_2(2n) - (-S_2(n) + 2S_2(2n)) = S_2(n) + S_2(2n) \end{aligned}$$

$$S_2(4n+2) = 4S_2(n) - 2S_2(2n)$$

$$\begin{aligned} S_2(4n+3) &= S_2(2(2n+1)+1) = 3S_2(2n+1) - S_2(4n+2) \\ &= 3 \cdot 3S_2(n) - 3S_2(2n) - (4S_2(n) - 2S_2(2n)) \\ &= 5S_2(n) - S_2(2n) \end{aligned}$$

etc.

□

Matrix representation to compute $(S_2(n))_{n\geq 0}$ easily

$$V(n) = \begin{pmatrix} S_2(n) \\ S_2(2n) \end{pmatrix}$$

$$V(2n) = \begin{pmatrix} S_2(2n) \\ S_2(4n) \end{pmatrix} = \underbrace{\begin{pmatrix} 0 & 1 \\ -1 & 2 \end{pmatrix}}_{:=\mu(0)} \begin{pmatrix} S_2(n) \\ S_2(2n) \end{pmatrix}$$

$$V(2n+1) = \begin{pmatrix} S_2(2n+1) \\ S_2(4n+2) \end{pmatrix} = \underbrace{\begin{pmatrix} 3 & -1 \\ 4 & -1 \end{pmatrix}}_{:=\mu(1)} \begin{pmatrix} S_2(n) \\ S_2(2n) \end{pmatrix}$$

If $\text{rep}_2(m) = m_\ell \cdots m_0$ with $m_i \in \{0,1\}$, then

$$S_2(m) = \begin{pmatrix} 1 & 0 \end{pmatrix} \mu(m_0) \cdots \mu(m_\ell) \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

A special tree structure: the tree of subwords

Definition: w a word in $\{0,1\}^*$

The *tree of subwords of* w is the tree $\mathcal{T}(w)$

- root ε
- if u and ua are subwords of w with $a \in \{0,1\}$, then ua is a child of u

A special tree structure: the tree of subwords

Definition: w a word in $\{0,1\}^*$

The *tree of subwords of* w is the tree $\mathcal{T}(w)$

- root ε
- if u and ua are subwords of w with $a \in \{0,1\}$, then ua is a child of u

Example: $w = 1001$

length	subwords
0	ε

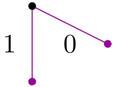
Definition: w a word in $\{0,1\}^*$

The *tree of subwords of* w is the tree $\mathcal{T}(w)$

- root ε
- if u and ua are subwords of w with $a \in \{0,1\}$, then ua is a child of u

Example: $w = 1001$

length	subwords				
0	ε				
1	0		1		



A special tree structure: the tree of subwords

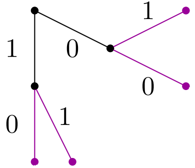
Definition: w a word in $\{0,1\}^*$

The *tree of subwords of* w is the tree $\mathcal{T}(w)$

- root ε
- if u and ua are subwords of w with $a \in \{0,1\}$, then ua is a child of u

Example: $w = 1001$

<table>
<tr><td>length</td><td colspan="5">subwords</td></tr>
<tr><td>0</td><td colspan="5">ε</td></tr>
<tr><td>1</td><td colspan="2">0</td><td colspan="3">1</td></tr>
<tr><td>2</td><td>00</td><td>01</td><td colspan="2">10</td><td>11</td></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td></td></tr>
</table>



A special tree structure: the tree of subwords

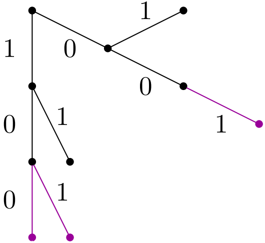
Definition: w a word in $\{0,1\}^*$

The *tree of subwords of* w is the tree $\mathcal{T}(w)$

- root ε
- if u and ua are subwords of w with $a \in \{0,1\}$, then ua is a child of u

Example: $w = 1001$

length	subwords				
0	ε				
1	0		1		
2	00	01	10		11
3	001		100	101	



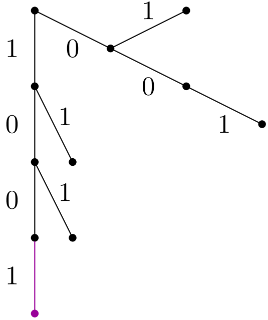
Definition: w a word in $\{0,1\}^*$

The *tree of subwords of* w is the tree $\mathcal{T}(w)$

- root ε
- if u and ua are subwords of w with $a \in \{0,1\}$, then ua is a child of u

Example: $w = 1001$

length	subwords				
0	ε				
1	0		1		
2	00	01	10		11
3	001		100	101	
4			1001		



$\mathcal{T}(1001)$

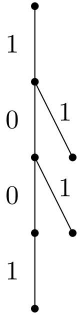
Convention in base 2: no leading 0

⇝ delete the right part of the tree $\mathcal{T}(w)$

⇝ new tree $\mathcal{T}_2(w)$

Example: $w = 1001$

length	subwords
0	ε
1	1
2	10, 11
3	100, 101
4	1001



$\mathcal{T}_2(1001)$

Usefulness:

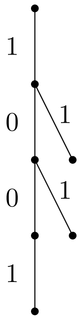
#nodes on level n of $\mathcal{T}_2(w)$ =#subwords of length n of w in $1\{0,1\}^* \cup \{\varepsilon\}$

Usefulness:

$$\#\text{nodes on level } n \text{ of } \mathcal{T}_2(w) = \#\text{subwords of length } n \text{ of } w \text{ in } 1\{0,1\}^* \cup \{\varepsilon\}$$

Example: $w = 1001$

	level	#nodes	subwords
	0	1	ε
	1	1	1
	2	2	10, 11
	3	2	100, 101
	4	1	1001
Total		7	



$\mathcal{T}_2(1001)$

Link with S_2: $\mathrm{val}_2(1001) = 9$ $\qquad S_2(9) = 7$

Palindromic structure (Leroy, Rigo, S., 2017)

For all $\ell > 1$ and all $0 \leq r < 2^{\ell-1}$,

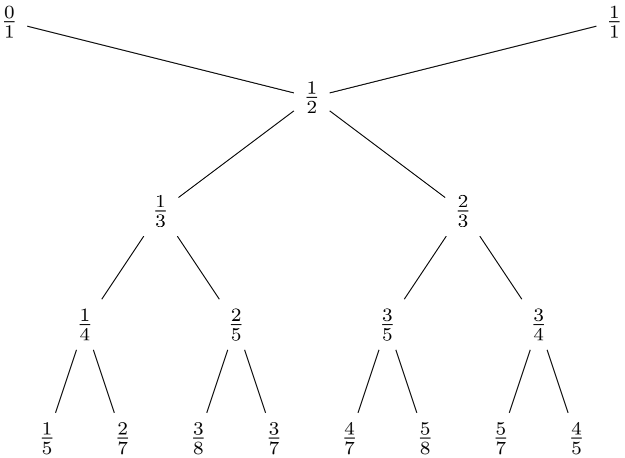
$$S_2(2^\ell + r) = S_2(2^{\ell+1} - r - 1).$$

Example: $\ell = 3, r = 1$

$2^3+1 = 9 \rightsquigarrow \text{rep}_2(9) = 1001 \qquad 2^4-1-1 = 14 \rightsquigarrow \text{rep}_2(14) = 1110$

$$\mathcal{T}_2(1001) \cong \mathcal{T}_2(1110) \Rightarrow \# \text{ nodes of } \mathcal{T}_2(1001) = \# \text{ nodes of } \mathcal{T}_2(1110)$$
$$\Rightarrow S_2(9) = S_2(14)$$

Definition: The *Farey tree* is a tree that contains every (reduced) positive rational less than 1 exactly once.



$S_2(n)$
1
2
3
3
4
5
5
4
5
7
8
7
7
8
7
5
⋮

Part III: Behavior of the Summatory Function

Example: $s(n)$ number of 1 in $\mathrm{rep}_2(n)$

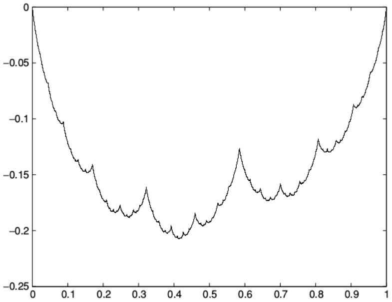
$s(2n) = s(n) \quad s(2n+1) = s(n)+1$

s is 2-regular

Summatory function A:

$$A(0) := 0$$

$$A(n) := \sum_{j=0}^{n-1} s(j) \quad \forall n \geq 1$$



Theorem (Delange, 1975)

$$\frac{A(n)}{n} = \frac{1}{2}\log_2(n) + \mathcal{G}(\log_2(n)) \qquad (1)$$

where $\mathcal{G}$ continuous, nowhere differentiable, periodic of period 1.

Theorem (Allouche, Shallit, 2003)

Under some hypotheses, the summatory function of every b-regular sequence has a behavior analogous to (1).

$\rightsquigarrow$ Replacing s by S_2: same behavior as (1) but does not satisfy the hypotheses of the theorem

Definition: $A_2(0) := 0$

$$A_2(n) := \sum_{j=0}^{n-1} S_2(j) \quad \forall n \geq 1$$

First few values:

$0, \mathbf{1}, \mathbf{3}, 6, \mathbf{9}, 13, 18, 23, \mathbf{27}, 32, 39, 47, 54, 61, 69, 76, \mathbf{81}, 87, 96, 107, \ldots$

Using several numeration systems

Definition: $A_2(0) := 0$

$$A_2(n) := \sum_{j=0}^{n-1} S_2(j) \quad \forall n \geq 1$$

First few values:

$0, \mathbf{1}, \mathbf{3}, 6, \mathbf{9}, 13, 18, 23, \mathbf{27}, 32, 39, 47, 54, 61, 69, 76, \mathbf{81}, 87, 96, 107, \ldots$

Lemma (Leroy, Rigo, S., 2017)

For all $n \geq 0$, $A_2(2^n) = 3^n$.

Lemma (Leroy, Rigo, S., 2017)

Let $\ell \geq 1$.

- If $0 \leq r \leq 2^{\ell-1}$, then

$$A_2(2^\ell + r) = 2 \cdot 3^{\ell-1} + A_2(2^{\ell-1} + r) + A_2(r).$$

- If $2^{\ell-1} < r < 2^\ell$, then

$$A_2(2^\ell+r) = 4\cdot 3^\ell - 2\cdot 3^{\ell-1} - A_2(2^{\ell-1}+r') - A_2(r') \quad \text{where } r' = 2^\ell - r.$$

Lemma (Leroy, Rigo, S., 2017)

Let $\ell \geq 1$.

- If $0 \leq r \leq 2^{\ell-1}$, then

$$A_2(2^\ell + r) = 2 \cdot 3^{\ell-1} + A_2(2^{\ell-1} + r) + A_2(r).$$

- If $2^{\ell-1} < r < 2^\ell$, then

$$A_2(2^\ell + r) = 4 \cdot 3^\ell - 2 \cdot 3^{\ell-1} - A_2(2^{\ell-1} + r') - A_2(r') \quad \text{where } r' = 2^\ell - r.$$

$\rightsquigarrow$ 3-*decomposition*: particular decomposition of $A_2(n)$ based on powers of 3

Lemma (Leroy, Rigo, S., 2017)

Let $\ell \geq 1$.

- If $0 \leq r \leq 2^{\ell-1}$, then

$$A_2(2^\ell + r) = 2 \cdot 3^{\ell-1} + A_2(2^{\ell-1} + r) + A_2(r).$$

- If $2^{\ell-1} < r < 2^\ell$, then

$$A_2(2^\ell+r) = 4\cdot 3^\ell - 2\cdot 3^{\ell-1} - A_2(2^{\ell-1}+r') - A_2(r') \quad \text{where } r' = 2^\ell - r.$$

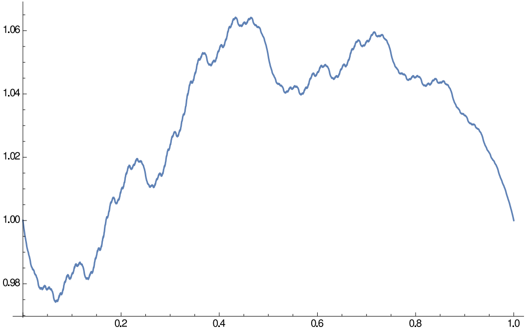
$\rightsquigarrow$ 3-*decomposition*: particular decomposition of $A_2(n)$ based on powers of 3

$\rightsquigarrow$ two numeration systems: base 2 and base 3

Theorem (Leroy, Rigo, S., 2017)

There exists a continuous and periodic function $\mathcal{H}_2$ of period 1 such that, for all $n \geq 1$,

$$A_2(n) = 3^{\log_2(n)}\ \mathcal{H}_2(\log_2(n)).$$



Conclusion

In this talk:

- Generalization of the Pascal triangle in base 2 modulo a prime number
- 2-regularity of the sequence $(S_2(n))_{n\geq 0}$ counting subword occurrences
- Behavior of the summatory function $(A_2(n))_{n\geq 0}$ of the sequence $(S_2(n))_{n\geq 0}$

Other work

Done:

- Generalization of the Pascal triangle modulo a prime number: extension to any Pisot–Bertrand numeration system
- Regularity of the sequence counting subword occurrences: extension to any base b and the Fibonacci numeration system
- Behavior of the summatory function: extension to any base b (exact behavior) and the Fibonacci numeration system (asymptotics)

What's next?
Pisot–Bertrand numeration systems,
Apply the methods for sequences not related to Pascal triangles,
etc.

Conus textile or Cloth of gold cone

Color pattern of its shell $\leftrightsquigarrow$ Sierpiński gasket

References

- J.-P. Allouche, J. Shallit, The ring of k-regular sequences, *Theoret. Comput. Sci.* 98(2) (1992), 163–197.
- J.-P. Allouche, J. Shallit, *Automatic sequences. Theory, applications, generalizations*, Cambridge University Press, Cambridge, 2003.
- H. Delange, Sur la fonction sommatoire de la fonction "somme des chiffres", *Enseignement Math.* 21 (1975), 31–47.
- P. Dumas, *Diviser pour régner : algèbre et analyse*, Journées Aléa 2016 (CIRM).
- 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, *submitted in April 2017.*
- M. Lothaire, *Combinatorics On Words*, Cambridge Mathematical Library, Cambridge University Press, Cambridge, 1997.
- É. Lucas, Théorie des fonctions numériques simplement périodiques, *Amer. J. Math.* 1 (1878) 197–240.
- M. Stipulanti, Convergence of Pascal-Like Triangles in Pisot–Bertrand Numeration Systems, *in progress.*