Philippe Flajolet

and

Dynamical Combinatorics.

Brigitte Vallée
GREYC (CNRS and Université de Caen)

Conference in the memory of Philippe
Philippe Flajolet and Analytical Combinatorics

Paris, December 2011.

A long joint scientific story, which lasted twenty years ...

A long joint scientific story, which lasted twenty years ...

A long joint scientific story, which lasted twenty years ...

... with four main steps

– Beginnings : analysis of the Gauss Algorithm (1990–1995)

– Developments (1995–1998)

– Foundation of dynamical combinatorics (1998–2006)

– Towards a realistic analysis of algorithms? (2008–2011)

I – Beginnings
Analysis of the Gauss Algorithm

Philippe always asked the “good” question to the “good” person.

I – Beginnings
Analysis of the Gauss Algorithm

Philippe always asked the "good" question to the "good" person.

– I have studied the Gauss Algorithm in my PhD Thesis (1986)
(from the worst–case point of view)

I – Beginnings
Analysis of the Gauss Algorithm

Philippe always asked the “good” question to the “good” person.

– I have studied the Gauss Algorithm in my PhD Thesis (1986)
(from the worst–case point of view)

– In 1990, I decided to work in Analysis of Algorithms.

I – Beginnings
Analysis of the Gauss Algorithm

Philippe always asked the "good" question to the "good" person.

– I have studied the Gauss Algorithm in my PhD Thesis (1986)
(from the worst–case point of view)

– In 1990, I decided to work in Analysis of Algorithms.

Then, Philippe proposed to me

to study the average–case analysis of the Gauss Algorithm

I – Beginnings
Analysis of the Gauss Algorithm

Philippe always asked the "good" question to the "good" person.

– I have studied the Gauss Algorithm in my PhD Thesis (1986)
(from the worst–case point of view)

– In 1990, I decided to work in Analysis of Algorithms.

Then, Philippe proposed to me

to study the average–case analysis of the Gauss Algorithm

Two steps:

– A first paper [Flajolet, V. 1990]: Mean number of iterations

I – Beginnings
Analysis of the Gauss Algorithm

Philippe always asked the “good” question to the “good” person.

– I have studied the Gauss Algorithm in my PhD Thesis (1986)
(from the worst–case point of view)

– In 1990, I decided to work in Analysis of Algorithms.

Then, Philippe proposed to me

to study the average–case analysis of the Gauss Algorithm

Two steps:

– A first paper [Flajolet, V. 1990]: Mean number of iterations

– Then two papers [Daudé, Flajolet, V., 1994, 1997] :
Distribution of the number of iterations

I – Beginnings
Analysis of the Gauss Algorithm

Philippe always asked the "good" question to the "good" person.

– I have studied the Gauss Algorithm in my PhD Thesis (1986)
(from the worst–case point of view)

– In 1990, I decided to work in Analysis of Algorithms.

Then, Philippe proposed to me

to study the average–case analysis of the Gauss Algorithm

Two steps:

– A first paper [Flajolet, V. 1990]: Mean number of iterations

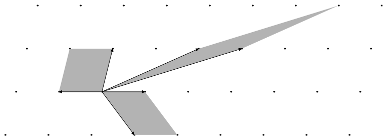
– Then two papers [Daudé, Flajolet, V., 1994, 1997] :
Distribution of the number of iterations

The Gauss algorithm

A lattice of $\mathbb{R}^2$

$=$ a discrete additive subgroup of $\mathbb{R}^2$

$$\mathcal{L} := \{w \in \mathbb{R}^2;\ w = xu + yv,\ x, y \in \mathbb{Z}\}$$



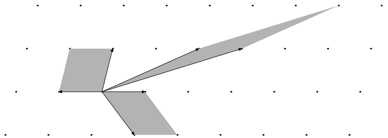
A lattice with three possible bases (u, v)

The Gauss algorithm

A lattice of $\mathbb{R}^2$

$=$ a discrete additive subgroup of $\mathbb{R}^2$

$\mathcal{L} := \{w \in \mathbb{R}^2;\ w = xu + yv,\ x, y \in \mathbb{Z}\}$



A lattice with three possible bases (u, v)

In two dimensions, an algorithm due to Gauss finds,

from any basis of a lattice $\mathcal{L}$, a minimal basis of $\mathcal{L}$

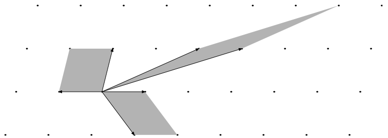
The Gauss algorithm is a main tool for finding short bases in any dimensions.

The Gauss algorithm

A lattice of $\mathbb{R}^2$

$=$ a discrete additive subgroup of $\mathbb{R}^2$

$$\mathcal{L} := \{w \in \mathbb{R}^2;\ w = xu + yv,\ x, y \in \mathbb{Z}\}$$

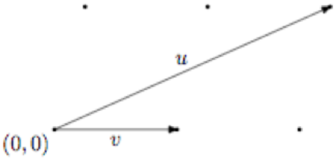


A lattice with three possible bases (u, v)

In two dimensions, an algorithm due to Gauss finds,

from any basis of a lattice $\mathcal{L}$, a minimal basis of $\mathcal{L}$

The Gauss algorithm is a main tool for finding short bases in any dimensions.



The Gauss alg. performs integer translations seen as "vectorial" divisions

$$u = mv + r \qquad \text{with} \qquad m = \left\lfloor \frac{u \cdot v}{|v|^2} \right\rfloor,$$

Here $m = 2$

Our main idea: A projective point of view.

$$(u, v) \in \mathbb{C}^2, \text{ with } u \neq 0 \longrightarrow z := \frac{v}{u} = \frac{(u \cdot v)}{|u|^2} + i\frac{\det(u, v)}{|u|^2}.$$

Up to similarity, the lattice $\mathcal{L}(u, v)$ becomes $\mathcal{L}(1, z) =: L(z)$.

Our main idea: A projective point of view.

$$(u, v) \in \mathbb{C}^2, \text{ with } u \neq 0 \longrightarrow z := \frac{v}{u} = \frac{(u \cdot v)}{|u|^2} + i\frac{\det(u, v)}{|u|^2}.$$

Up to similarity, the lattice $\mathcal{L}(u, v)$ becomes $\mathcal{L}(1, z) =: L(z)$.

The Gauss algorithm is an extension of the Euclid algorithm.

Our main idea: A projective point of view.

$$(u, v) \in \mathbb{C}^2, \text{ with } u \neq 0 \longrightarrow z := \frac{v}{u} = \frac{(u \cdot v)}{|u|^2} + i\frac{\det(u, v)}{|u|^2}.$$

Up to similarity, the lattice $\mathcal{L}(u, v)$ becomes $\mathcal{L}(1, z) =: L(z).$

The Gauss algorithm is an extension of the Euclid algorithm.

Euclid's algorithm	Gauss' algorithm
Division between real numbers $v = mu + r$ with $m = \left\lfloor \frac{u}{v} \right\rfloor$	Division between complex vectors $v = mu + r$ with $m = \left\lfloor \Re\left(\frac{u}{v}\right) \right\rfloor$

Our main idea: A projective point of view.

$$(u, v) \in \mathbb{C}^2, \text{ with } u \neq 0 \longrightarrow z := \frac{v}{u} = \frac{(u \cdot v)}{|u|^2} + i \frac{\det(u, v)}{|u|^2}.$$

Up to similarity, the lattice $\mathcal{L}(u, v)$ becomes $\mathcal{L}(1, z) =: L(z)$.

The Gauss algorithm is an extension of the Euclid algorithm.

Euclid's algorithm	Gauss' algorithm
Division between real numbers $v = mu + r$ with $m = \left\lfloor \frac{u}{v} \right\rfloor$	Division between complex vectors $v = mu + r$ with $m = \left\lfloor \Re\left(\frac{u}{v}\right) \right\rfloor$
Division + exchange $(v, u) \to (r, v)$ "read" on $x = v/u$ $U(x) = \frac{1}{x} - \left\lfloor \frac{1}{x} \right\rfloor$	Division + exchange $(v, u) \to (r, v)$ "read" on $z = v/u$ $U(z) = \frac{1}{z} - \left\lfloor \Re\left(\frac{1}{z}\right) \right\rfloor$

Our main idea: A projective point of view.

$$(u,v) \in \mathbb{C}^2, \text{ with } u \neq 0 \longrightarrow z := \frac{v}{u} = \frac{(u \cdot v)}{|u|^2} + i\frac{\det(u,v)}{|u|^2}.$$

Up to similarity, the lattice $\mathcal{L}(u,v)$ becomes $\mathcal{L}(1,z) =: L(z)$.

The Gauss algorithm is an extension of the Euclid algorithm.

Euclid's algorithm	Gauss' algorithm
Division between real numbers $v = mu + r$ with $m = \left\lfloor \frac{u}{v} \right\rfloor$	Division between complex vectors $v = mu + r$ with $m = \left\lfloor \Re\left(\frac{u}{v}\right) \right\rfloor$
Division + exchange $(v,u) \to (r,v)$ "read" on $x = v/u$ $U(x) = \frac{1}{x} - \left\lfloor \frac{1}{x} \right\rfloor$	Division + exchange $(v,u) \to (r,v)$ "read" on $z = v/u$ $U(z) = \frac{1}{z} - \left\lfloor \Re\left(\frac{1}{z}\right) \right\rfloor$
Stopping condition: $x = 0$	Stopping condition: $z \in \mathcal{F}$ $\mathcal{F} := \{z; \ \ \|z\| \geq 1, \ \ \|\Re z\| \leq 1/2\}$

$\mathcal{D}$ is the disk of diameter $[0,1]$, and $$U(z) = \frac{1}{z} - \left\lfloor \Re\left(\frac{1}{z}\right) \right\rfloor$$

On an input $z \in \mathcal{D}$, the `CoreGauss` Alg performs $z := U(z)$, until $z \notin \mathcal{D}$.
It uses at each step the set of (inverse) branches of mapping U

$$\mathcal{H} := \{z \mapsto \frac{1}{m+z};\ m \geq 1\},$$ the same set as the Euclid Alg.

The domain $[R \geq k+1]$ is a union of disks : $$[R \geq k+1] = \bigcup_{h \in \mathcal{H}^k} h(\mathcal{D})$$

$\mathcal{D}$ is the disk of diameter $[0,1]$, and $$U(z) = \frac{1}{z} - \left\lfloor \Re\left(\frac{1}{z}\right) \right\rfloor$$

On an input $z \in \mathcal{D}$, the `CoreGauss` Alg performs $z := U(z)$, until $z \notin \mathcal{D}$.

It uses at each step the set of (inverse) branches of mapping U

$$\mathcal{H} := \{z \mapsto \frac{1}{m+z};\ m \geq 1\},$$ the same set as the Euclid Alg.

The domain $[R \geq k+1]$ is a union of disks : $$[R \geq k+1] = \bigcup_{h \in \mathcal{H}^k} h(\mathcal{D})$$

Inside disk $\mathcal{D}$,
the domains $[R = k]$
alternatively
in black and white

$\mathcal{D}$ is the disk of diameter $[0, 1]$, and $\qquad U(z) = \frac{1}{z} - \left\lfloor \Re\left(\frac{1}{z}\right) \right\rfloor$

On an input $z \in \mathcal{D}$, the `CoreGauss` Alg performs $z := U(z)$, until $z \notin \mathcal{D}$.
It uses at each step the set of (inverse) branches of mapping U

$$\mathcal{H} := \{z \mapsto \frac{1}{m+z};\ m \geq 1\},$$ the same set as the Euclid Alg.

The domain $[R \geq k+1]$ is a union of disks : $[R \geq k+1] = \bigcup_{h \in \mathcal{H}^k} h(\mathcal{D})$

Inside disk $\mathcal{D}$,
the domains $[R = k]$
alternatively
in black and white

The disks $h(\mathcal{D})$ are the "fundamental" disks

$$\mathbb{E}[R] = \sum_{h \in \mathcal{H}^\star} [h(0) - h(1)]^2,$$

$\mathcal{D}$ is the disk of diameter $[0,1]$, and $\quad U(z) = \frac{1}{z} - \left\lfloor \Re\left(\frac{1}{z}\right) \right\rfloor$

On an input $z \in \mathcal{D}$, the `CoreGauss` Alg performs $z := U(z)$, until $z \notin \mathcal{D}$.
It uses at each step the set of (inverse) branches of mapping U

$\mathcal{H} := \{z \mapsto \frac{1}{m+z};\ m \geq 1\}$, the same set as the Euclid Alg.

The domain $[R \geq k+1]$ is a union of disks : $[R \geq k+1] = \bigcup_{h \in \mathcal{H}^k} h(\mathcal{D})$

Inside disk $\mathcal{D}$,
the domains $[R = k]$
alternatively
in black and white

The disks $h(\mathcal{D})$ are the "fundamental" disks

$$\mathbb{E}[R] = \sum_{h \in \mathcal{H}^\star} [h(0) - h(1)]^2,$$

A simple characterization of $\mathcal{H}^\star$ leads to

$$\mathbb{E}[R] = \frac{3}{4} + \frac{2}{\zeta(4)} \sum_{d \geq 1} \frac{1}{d^2} \sum_{d < c \leq 2d} \frac{1}{d^2} = 1.351094...$$

An alternative expression involves ....
$\zeta(3)$ and the tetralogarithm Li_4... [FV. FOCS, 1990]

$\mathcal{D}$ is the disk of diameter $[0,1]$, and $\quad U(z) = \frac{1}{z} - \left\lfloor \Re\left(\frac{1}{z}\right) \right\rceil$

On an input $z \in \mathcal{D}$, the `CoreGauss` Alg performs $z := U(z)$, until $z \notin \mathcal{D}$.

It uses at each step the set of (inverse) branches of mapping U

$$\mathcal{H} := \{z \mapsto \frac{1}{m+z};\ m \geq 1\},$$ the same set as the Euclid Alg.

The domain $[R \geq k+1]$ is a union of disks : $[R \geq k+1] = \bigcup_{h \in \mathcal{H}^k} h(\mathcal{D})$

Inside disk $\mathcal{D}$,
the domains $[R = k]$
alternatively
in black and white

The disks $h(\mathcal{D})$ are the "fundamental" disks

$$\mathbb{E}[R] = \sum_{h \in \mathcal{H}^\star} [h(0) - h(1)]^2,$$

A simple characterization of $\mathcal{H}^\star$ leads to

$$\mathbb{E}[R] = \frac{3}{4} + \frac{2}{\zeta(4)} \sum_{d \geq 1} \frac{1}{d^2} \sum_{d < c \leq 2d} \frac{1}{d^2} = 1.351094...$$

An alternative expression involves ....
$\zeta(3)$ and the tetralogarithm Li_4... [FV. FOCS, 1990]

The domains $[R = k]$
alternatively
in black and white

$$\Pr[R \geq k+1] = \sum_{h \in \mathcal{H}^k} |h(0) - h(1)|^2$$

But how to characterize $\mathcal{H}^k$? More difficult! ...

The domains $[R = k]$

alternatively

in black and white

$$\Pr[R \geq k+1] = \sum_{h \in \mathcal{H}^k} |h(0) - h(1)|^2$$

But how to characterize $\mathcal{H}^k$? More difficult! ...

Then ..... Hervé Daudé's PhD

Hensley's paper

Mayer's Chapter in the Green Book

The domains $[R = k]$
alternatively
in black and white

$$\Pr[R \geq k+1] = \sum_{h \in \mathcal{H}^k} |h(0) - h(1)|^2$$

But how to characterize $\mathcal{H}^k$? More difficult! ...

	Hervé Daudé's PhD
Then	Hensley's paper
	Mayer's Chapter in the Green Book

The domains $[R = k]$

alternatively

in black and white

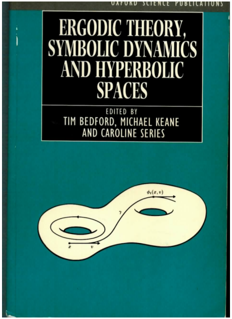
$$\Pr[R \geq k+1] = \sum_{h \in \mathcal{H}^k} |h(0) - h(1)|^2$$

But how to characterize $\mathcal{H}^k$? More difficult! ...

Then ..... Hervé Daudé's PhD

Hensley's paper

Mayer's Chapter in the Green Book

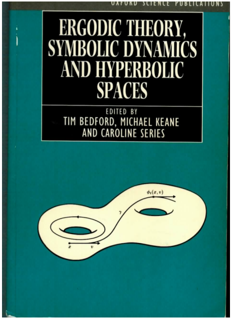


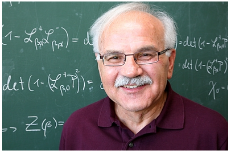
The domains $[R = k]$
alternatively
in black and white

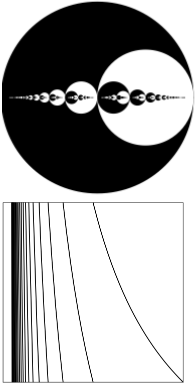
$$\Pr[R \geq k+1] = \sum_{h \in \mathcal{H}^k} |h(0) - h(1)|^2$$

But how to characterize $\mathcal{H}^k$? More difficult! ...

Then ..... Hervé Daudé's PhD
Hensley's paper
Mayer's Chapter in the Green Book

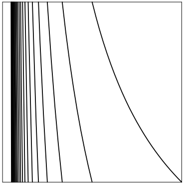






The Euclidean dynamical system

(EDS)



The Euclidean dynamical system

(EDS)

$$\Pr[R \geq k+1] = \sum_{h \in \mathcal{H}^k} |h(0) - h(1)|^2$$

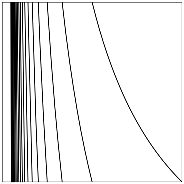
The Ruelle-Mayer operator $\mathbf{G}_s$, defined as

$$\mathbf{G}_s[f](x) := \sum_{h \in \mathcal{H}} |h'(x)|^s \; f \circ h(x)$$

is the transfer operator related to the EDS.

It generates (with its k–th iterate) the subset $\mathcal{H}^k$,

$$\mathbf{G}_s^k[f](x) := \sum_{h \in \mathcal{H}^k} |h'(x)|^s \; f \circ h(x)$$



The Euclidean dynamical system
(EDS)

$$\Pr[R \geq k+1] = \sum_{h \in \mathcal{H}^k} |h(0) - h(1)|^2$$

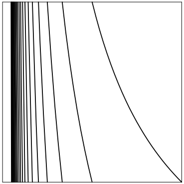
The Ruelle-Mayer operator $\mathbf{G}_s$, defined as

$$\mathbf{G}_s[f](x) := \sum_{h \in \mathcal{H}} |h'(x)|^s \; f \circ h(x)$$

is the transfer operator related to the EDS.
It generates (with its k–th iterate) the subset $\mathcal{H}^k$,

$$\mathbf{G}_s^k[f](x) := \sum_{h \in \mathcal{H}^k} |h'(x)|^s \; f \circ h(x)$$

and $\quad \Pr[R \geq k+1] = \mathbf{G}_2^k \left[\frac{1}{(1+x)^2} \right](0)$



The Euclidean dynamical system (EDS)

$$\Pr[R \geq k+1] = \sum_{h \in \mathcal{H}^k} |h(0) - h(1)|^2$$

The Ruelle-Mayer operator $\mathbf{G}_s$, defined as

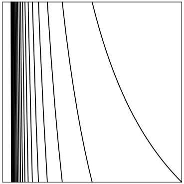
$$\mathbf{G}_s[f](x) := \sum_{h \in \mathcal{H}} |h'(x)|^s \ f \circ h(x)$$

is the transfer operator related to the EDS.
It generates (with its k–th iterate) the subset $\mathcal{H}^k$,

$$\mathbf{G}_s^k[f](x) := \sum_{h \in \mathcal{H}^k} |h'(x)|^s \ f \circ h(x)$$

and $\qquad \Pr[R \geq k+1] = \mathbf{G}_2^k \left[\frac{1}{(1+x)^2} \right](0)$

Spectral properties of $\mathbf{G}_2$ on a convenient functional space prove :



The Euclidean dynamical system

(EDS)

$$\Pr[R \geq k+1] = \sum_{h \in \mathcal{H}^k} |h(0) - h(1)|^2$$

The Ruelle-Mayer operator $\mathbf{G}_s$, defined as

$$\mathbf{G}_s[f](x) := \sum_{h \in \mathcal{H}} |h'(x)|^s \ f \circ h(x)$$

is the transfer operator related to the EDS.

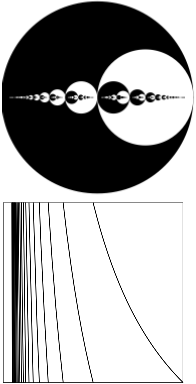
It generates (with its k–th iterate) the subset $\mathcal{H}^k$,

$$\mathbf{G}_s^k[f](x) := \sum_{h \in \mathcal{H}^k} |h'(x)|^s \ f \circ h(x)$$

and $\qquad \Pr[R \geq k+1] = \mathbf{G}_2^k \left[\frac{1}{(1+x)^2} \right] (0)$

Spectral properties of $\mathbf{G}_2$ on a convenient functional space prove :

$\Pr[R \geq k+1] \sim C \ \lambda(2)^k$ with $\lambda(2) :=$ the dominant eigenvalue of $\mathbf{G}_2$



The Euclidean dynamical system (EDS)

$$\Pr[R \geq k+1] = \sum_{h \in \mathcal{H}^k} |h(0) - h(1)|^2$$

The Ruelle-Mayer operator $\mathbf{G}_s$, defined as

$$\mathbf{G}_s[f](x) := \sum_{h \in \mathcal{H}} |h'(x)|^s \; f \circ h(x)$$

is the transfer operator related to the EDS.
It generates (with its k–th iterate) the subset $\mathcal{H}^k$,

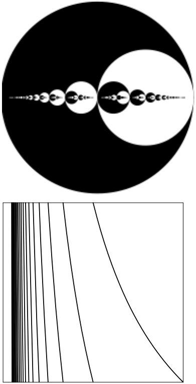
$$\mathbf{G}_s^k[f](x) := \sum_{h \in \mathcal{H}^k} |h'(x)|^s \; f \circ h(x)$$

and $\quad \Pr[R \geq k+1] = \mathbf{G}_2^k \left[\frac{1}{(1+x)^2} \right] (0)$

Spectral properties of $\mathbf{G}_2$ on a convenient functional space prove :

$\Pr[R \geq k+1] \sim C\ \lambda(2)^k$ with $\lambda(2) :=$ the dominant eigenvalue of $\mathbf{G}_2$

We "discovered" $\lambda(2)$ in 1994,



The Euclidean dynamical system (EDS)

$$\Pr[R \geq k+1] = \sum_{h \in \mathcal{H}^k} |h(0) - h(1)|^2$$

The Ruelle-Mayer operator $\mathbf{G}_s$, defined as

$$\mathbf{G}_s[f](x) := \sum_{h \in \mathcal{H}} |h'(x)|^s \; f \circ h(x)$$

is the transfer operator related to the EDS.
It generates (with its k–th iterate) the subset $\mathcal{H}^k$,

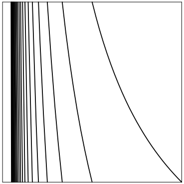
$$\mathbf{G}_s^k[f](x) := \sum_{h \in \mathcal{H}^k} |h'(x)|^s \; f \circ h(x)$$

and $\Pr[R \geq k+1] = \mathbf{G}_2^k \left[\frac{1}{(1+x)^2} \right](0)$

Spectral properties of $\mathbf{G}_2$ on a convenient functional space prove :

$\Pr[R \geq k+1] \sim C\ \lambda(2)^k$ with $\lambda(2) :=$ the dominant eigenvalue of $\mathbf{G}_2$

We "discovered" $\lambda(2)$ in 1994, its value is $\lambda(2) \sim 0.1994...$



The Euclidean dynamical system (EDS)

$$\Pr[R \geq k+1] = \sum_{h \in \mathcal{H}^k} |h(0) - h(1)|^2$$

The Ruelle-Mayer operator $\mathbf{G}_s$, defined as

$$\mathbf{G}_s[f](x) := \sum_{h \in \mathcal{H}} |h'(x)|^s \ f \circ h(x)$$

is the transfer operator related to the EDS.
It generates (with its k–th iterate) the subset $\mathcal{H}^k$,

$$\mathbf{G}_s^k[f](x) := \sum_{h \in \mathcal{H}^k} |h'(x)|^s \ f \circ h(x)$$

and $\qquad \Pr[R \geq k+1] = \mathbf{G}_2^k \left[\frac{1}{(1+x)^2} \right] (0)$

Spectral properties of $\mathbf{G}_2$ on a convenient functional space prove :

$\Pr[R \geq k+1] \sim C\ \lambda(2)^k$ with $\lambda(2) :=$ the dominant eigenvalue of $\mathbf{G}_2$

We "discovered" $\lambda(2)$ in 1994, its value is $\lambda(2) \sim 0.1994\ldots$

Philippe wished to call it the Vallée constant...

He described 1994 as "the year of the constant $\lambda(2)$"...

II – Developments.

How to use the Ruelle-Mayer operator
in (almost) any Euclidean problem.

– Any problem related to continued fraction expansions
(continuous world, real variable)

– Any problem related to the Euclid Algorithm
(discrete world, rational variable)

Some instances of easy solutions for Euclidean problems – Real case (I)

Continued fraction expansion of x

$$x = \cfrac{1}{m_1 + \cfrac{1}{m_2 + \cfrac{1}{\ddots + \cfrac{1}{m_k + \cfrac{1}{\ddots}}}}}$$

Some instances of easy solutions for Euclidean problems – Real case (I)

Continued fraction expansion of x

$$x = \cfrac{1}{m_1 + \cfrac{1}{m_2 + \cfrac{1}{\ddots + \cfrac{1}{m_k + \cfrac{1}{\ddots}}}}}$$

Iterates $\mathbf{G}_s^k$ of the operator $\mathbf{G}_s$ generate

– the truncated $CFE's$ at depth k,

Some instances of easy solutions for Euclidean problems – Real case (I)

Continued fraction expansion of x

$$x = \cfrac{1}{m_1 + \cfrac{1}{m_2 + \cfrac{1}{\ddots + \cfrac{1}{m_k + \cfrac{1}{\ddots}}}}}$$

Iterates $\mathbf{G}_s^k$ of the operator $\mathbf{G}_s$ generate

– the truncated $CFE's$ at depth k,

– the fundamental intervals

$$h(\mathcal{I}) := [h(0), h(1)] \text{ with } h \in \mathcal{H}^k.$$

Some instances of easy solutions for Euclidean problems – Real case (I)

Continued fraction expansion of x

$$x = \cfrac{1}{m_1 + \cfrac{1}{m_2 + \cfrac{1}{\ddots + \cfrac{1}{m_k + \cfrac{1}{\ddots}}}}}$$

Iterates $\mathbf{G}_s^k$ of the operator $\mathbf{G}_s$ generate

– the truncated $CFE's$ at depth k,

– the fundamental intervals

$$h(\mathcal{I}) := [h(0), h(1)] \text{ with } h \in \mathcal{H}^k.$$

– the best rational approximation of a real x,

namely the rational $h(0) = \dfrac{p_k}{q_k}(x)$

Some instances of easy solutions for Euclidean problems – Real case (I)

Continued fraction expansion of x

$$x = \cfrac{1}{m_1 + \cfrac{1}{m_2 + \cfrac{1}{\ddots + \cfrac{1}{m_k + \cfrac{1}{\ddots}}}}}$$

Iterates $\mathbf{G}_s^k$ of the operator $\mathbf{G}_s$ generate

– the truncated $CFE's$ at depth k,

– the fundamental intervals

$$h(\mathcal{I}) := [h(0), h(1)] \text{ with } h \in \mathcal{H}^k.$$

– the best rational approximation of a real x,

namely the rational $h(0) = \dfrac{p_k}{q_k}(x)$

Distance $\delta_k(x)$ between the real x and its best approximation $p_k/q_k(x)$

$\mathbb{E}[\delta_k]$ expressed with $\mathbf{G}_2^k$ $\Longrightarrow$ $\mathbb{E}[\delta_k] \sim C \;\; \lambda(2)^k$

Some instances of easy solutions for Euclidean problems – Real case (I)

Continued fraction expansion of x

$$x = \cfrac{1}{m_1 + \cfrac{1}{m_2 + \cfrac{1}{\ddots + \cfrac{1}{m_k + \cfrac{1}{\ddots}}}}}$$

Iterates $\mathbf{G}_s^k$ of the operator $\mathbf{G}_s$ generate

– the truncated $CFE's$ at depth k,

– the fundamental intervals

$$h(\mathcal{I}) := [h(0), h(1)] \text{ with } h \in \mathcal{H}^k.$$

– the best rational approximation of a real x, namely the rational $h(0) = \dfrac{p_k}{q_k}(x)$

Distance $\delta_k(x)$ between the real x and its best approximation $p_k/q_k(x)$

$\mathbb{E}[\delta_k]$ expressed with $\mathbf{G}_2^k$ $\qquad \Longrightarrow \qquad$ $\mathbb{E}[\delta_k] \sim C \;\; \lambda(2)^k$

Behaviour of $\log q_k(x)$:

Some instances of easy solutions for Euclidean problems – Real case (I)

Continued fraction expansion of x

$$x = \cfrac{1}{m_1 + \cfrac{1}{m_2 + \cfrac{1}{\ddots + \cfrac{1}{m_k + \cfrac{1}{\ddots}}}}}$$

Iterates $\mathbf{G}_s^k$ of the operator $\mathbf{G}_s$ generate

– the truncated $CFE's$ at depth k,

– the fundamental intervals

$h(\mathcal{I}) := [h(0), h(1)]$ with $h \in \mathcal{H}^k$.

– the best rational approximation of a real x,

namely the rational $h(0) = \frac{p_k}{q_k}(x)$

Distance $\delta_k(x)$ between the real x and its best approximation $p_k/q_k(x)$

$\mathbb{E}[\delta_k]$ expressed with $\mathbf{G}_2^k$ $\Longrightarrow$ $\mathbb{E}[\delta_k] \sim C \quad \lambda(2)^k$

Behaviour of $\log q_k(x)$: $\mathbb{E}[q_k^{2s}]$ expressed with $\mathbf{G}_{1-s}^k$

Then, [spectral dominant properties+ quasi-power theorem]

$\Longrightarrow$ an asymptotic normal law for $\log q_k$

Some instances of easy solutions for Euclidean problems – Real case (I)

Continued fraction expansion of x

$$x = \cfrac{1}{m_1 + \cfrac{1}{m_2 + \cfrac{1}{\ddots + \cfrac{1}{m_k + \cfrac{1}{\ddots}}}}}$$

Iterates $\mathbf{G}_s^k$ of the operator $\mathbf{G}_s$ generate
– the truncated $CFE's$ at depth k,
– the fundamental intervals

$$h(\mathcal{I}) := [h(0), h(1)] \text{ with } h \in \mathcal{H}^k.$$

– the best rational approximation of a real x,

namely the rational $h(0) = \frac{p_k}{q_k}(x)$

Distance $\delta_k(x)$ between the real x and its best approximation $p_k/q_k(x)$

$\mathbb{E}[\delta_k]$ expressed with $\mathbf{G}_2^k$ $\quad\Longrightarrow\quad$ $\mathbb{E}[\delta_k] \sim C \quad \lambda(2)^k$

Behaviour of $\log q_k(x)$: $\mathbb{E}[q_k^{2s}]$ expressed with $\mathbf{G}_{1-s}^k$

Then, [spectral dominant properties+ quasi-power theorem]

$\Longrightarrow$ an asymptotic normal law for $\log q_k$

$$\mathbb{E}[\log q_k] \sim k|\Lambda'(1)|, \quad \mathbb{V}[\log q_k] \sim k\Lambda''(1) \qquad \text{with} \qquad \Lambda(s) := \log \lambda(s),$$

$\Lambda'(1)$ is the entropy, and $\Lambda''(1)$ is the Hensley constant.

Other Euclidean problems – Real case (II)

Comparing reals via their continued fraction expansion

Other Euclidean problems – Real case (II)

Comparing reals via their continued fraction expansion

Other Euclidean problems – Real case (II)

Comparing reals via their continued fraction expansion

Comparison of two real numbers:

– related to the sign of the determinant,

– and the Hakmem algorithm.

The Hakmem Alg. and the Gauss Alg: the same geometry.

$$\Pr[C \geq k+1] \sim D\lambda(2)^k$$

Other Euclidean problems – Real case (II)

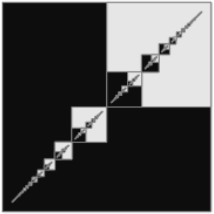
Comparing reals via their continued fraction expansion

Comparison of two real numbers:

– related to the sign of the determinant,

– and the Hakmem algorithm.

The Hakmem Alg. and the Gauss Alg: the same geometry.



$$\Pr[C \geq k+1] \sim D\lambda(2)^k$$

Other Euclidean problems – Real case (II)

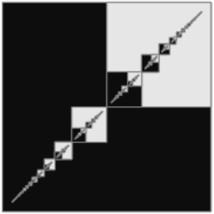
Comparing reals via their continued fraction expansion

Comparison of two real numbers:

– related to the sign of the determinant,

– and the Hakmem algorithm.

The Hakmem Alg. and the Gauss Alg: the same geometry.



$$\Pr[C \geq k+1] \sim D\lambda(2)^k$$

Comparison of n real numbers via the trie
built on their continued fraction expansions
[J. Clément, Master Thesis, 98]

Other Euclidean problems – Real case (II)

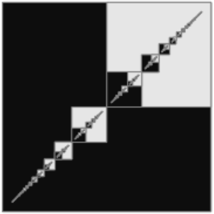
Comparing reals via their continued fraction expansion

Comparison of two real numbers:

– related to the sign of the determinant,

– and the Hakmem algorithm.

The Hakmem Alg. and the Gauss Alg: the same geometry.



$$\Pr[C \geq k+1] \sim D\lambda(2)^k$$

Comparison of n real numbers via the trie
built on their continued fraction expansions
[J. Clément, Master Thesis, 98]

Other Euclidean problems – Real case (II)

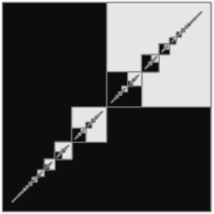
Comparing reals via their continued fraction expansion

Comparison of two real numbers:

– related to the sign of the determinant,

– and the Hakmem algorithm.

The Hakmem Alg. and the Gauss Alg: the same geometry.



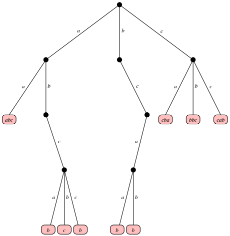
$$\Pr[C \geq k+1] \sim D\lambda(2)^k$$

Comparison of n real numbers via the trie
built on their continued fraction expansions
[J. Clément, Master Thesis, 98]

Typical depth D_n, Height H_n

$$\mathbb{E}[D_n] \sim \frac{1}{|\lambda'(1)|} \log n$$

$$\mathbb{E}[H_n] \sim \frac{1}{|\log \lambda(2)|} \log n$$



Euclidean problems – Rational case. $CFE(u/v) \sim \texttt{Euclid}(u, v)$

Study of the number of iterations P of the Euclid Algorithm.

Euclidean problems – Rational case. $CFE(u/v) \sim \mathtt{Euclid}(u,v)$

Study of the number of iterations P of the Euclid Algorithm.

For any integer pair of $\Omega := \{(u,v); \quad u < v, \quad u, v \text{ coprime}\}$

$$\exists k, \quad \exists h \in \mathcal{H}^k, \quad \frac{u}{v} = h(0), \quad \frac{1}{v^2} = |h'(0)|, \quad P(u,v) = k.$$

Euclidean problems – Rational case. $CFE(u/v) \sim \texttt{Euclid}(u,v)$

Study of the number of iterations P of the Euclid Algorithm.

For any integer pair of $\Omega := \{(u,v); \quad u < v, \quad u,v$ coprime$\}$

$$\exists k, \quad \exists h \in \mathcal{H}^k, \quad \frac{u}{v} = h(0), \quad \frac{1}{v^2} = |h'(0)|, \quad P(u,v) = k.$$

Main idea: we introduce a generating function $S(s)$ of Dirichlet type and we relate it to the Ruelle-Mayer operator $\mathbf{G}_s$:

$$S(s) := \sum_{(u,v)\in\Omega} \frac{P(u,v)}{v^{2s}} = \sum_{k\geq 1} k \sum_{h\in\mathcal{H}^k} |h'(0)|^s = \sum_{k\geq 1} k\mathbf{G}_s^k[1](0) = \mathbf{G}_s \circ (I - \mathbf{G}_s)^{-2}[1]($$

Euclidean problems – Rational case. $CFE(u/v) \sim \mathtt{Euclid}(u,v)$

Study of the number of iterations P of the Euclid Algorithm.

For any integer pair of $\Omega := \{(u,v); \quad u < v, \quad u, v \text{ coprime}\}$

$$\exists k, \quad \exists h \in \mathcal{H}^k, \quad \frac{u}{v} = h(0), \quad \frac{1}{v^2} = |h'(0)|, \quad P(u,v) = k.$$

Main idea: we introduce a generating function $S(s)$ of Dirichlet type and we relate it to the Ruelle-Mayer operator $\mathbf{G}_s$:

$$S(s) := \sum_{(u,v)\in\Omega} \frac{P(u,v)}{v^{2s}} = \sum_{k\geq 1} k \sum_{h\in\mathcal{H}^k} |h'(0)|^s = \sum_{k\geq 1} k\mathbf{G}_s^k[1](0) = \mathbf{G}_s \circ (I - \mathbf{G}_s)^{-2}[1]($$

Two properties of $\mathbf{G}_s$ (on a convenient functional space)

– The map $s \mapsto (I - \mathbf{G}_s)^{-1}$ is analytic on $\Re s \geq 1; s \neq 1$.

– A simple pole at $s = 1$

$\Longrightarrow$ Possible extraction of coefficients with a Tauberian Theorem

$\Longrightarrow$ A short proof of the estimate $\mathbb{E}[P_N] \sim \dfrac{2}{|\lambda'(1)|} \log N = \dfrac{12 \log 2}{\pi^2} \log N$

III– Foundations of Dynamical Combinatorics.

Extensions of previous ideas to more general contexts.

In our previous studies, the Ruelle-Mayer operator $\mathbf{G}_s^k$ generates

III– Foundations of Dynamical Combinatorics.

Extensions of previous ideas to more general contexts.

In our previous studies, the Ruelle-Mayer operator $\mathbf{G}_s^k$ generates

– the fundamental intervals $h(\mathcal{I})$ for $h \in \mathcal{H}^k$

III– Foundations of Dynamical Combinatorics.

Extensions of previous ideas to more general contexts.

In our previous studies, the Ruelle-Mayer operator $\mathbf{G}_s^k$ generates

- the fundamental intervals $h(\mathcal{I})$ for $h \in \mathcal{H}^k$
- the fundamental probabilities p_w, for $|w| = k$

  $p_w :=$ the probability that $CFE(x)$ begins with the sequence w

III– Foundations of Dynamical Combinatorics.

Extensions of previous ideas to more general contexts.

In our previous studies, the Ruelle-Mayer operator $\mathbf{G}_s^k$ generates

- the fundamental intervals $h(\mathcal{I})$ for $h \in \mathcal{H}^k$
- the fundamental probabilities p_w, for $|w| = k$

  $p_w :=$ the probability that $CFE(x)$ begins with the sequence w
- the rationals u/v of depth k via their denominators v

  as the branches are LFT's : $u/v = h(0) \Longrightarrow 1/v^2 = |h'(0)|$

III– Foundations of Dynamical Combinatorics.

Extensions of previous ideas to more general contexts.

In our previous studies, the Ruelle-Mayer operator $\mathbf{G}_s^k$ generates

- the fundamental intervals $h(\mathcal{I})$ for $h \in \mathcal{H}^k$
- the fundamental probabilities p_w, for $|w| = k$

  $p_w :=$ the probability that $CFE(x)$ begins with the sequence w
- the rationals u/v of depth k via their denominators v

  as the branches are LFT's : $u/v = h(0) \Longrightarrow 1/v^2 = |h'(0)|$

The role of $\mathbf{G}_s$ is central

because it is the transfer operator of the underlying dynamical system.

III– Foundations of Dynamical Combinatorics.

Extensions of previous ideas to more general contexts.

In our previous studies, the Ruelle-Mayer operator $\mathbf{G}_s^k$ generates

- the fundamental intervals $h(\mathcal{I})$ for $h \in \mathcal{H}^k$
- the fundamental probabilities p_w, for $|w| = k$

  $p_w :=$ the probability that $CFE(x)$ begins with the sequence w
- the rationals u/v of depth k via their denominators v

  as the branches are LFT's : $u/v = h(0) \Longrightarrow 1/v^2 = |h'(0)|$

The role of $\mathbf{G}_s$ is central

because it is the transfer operator of the underlying dynamical system.

We wished to extend this use of transfer operators to other contexts

- Information theory: any dynamical system

  viewed as a processus to produce symbols

III– Foundations of Dynamical Combinatorics.

Extensions of previous ideas to more general contexts.

In our previous studies, the Ruelle-Mayer operator $\mathbf{G}_s^k$ generates

– the fundamental intervals $h(\mathcal{I})$ for $h \in \mathcal{H}^k$

– the fundamental probabilities p_w, for $|w| = k$

$p_w :=$ the probability that $CFE(x)$ begins with the sequence w

– the rationals u/v of depth k via their denominators v

as the branches are LFT's : $u/v = h(0) \Longrightarrow 1/v^2 = |h'(0)|$

The role of $\mathbf{G}_s$ is central

because it is the transfer operator of the underlying dynamical system.

We wished to extend this use of transfer operators to other contexts

– Information theory: any dynamical system

viewed as a processus to produce symbols

– Euclidean context: any dynamical system underlying a Euclid alg

whose branches are LFT's.

Very often, these extensions began with Philippe, and

– I mostly continued with other collaborators.

– He had a strong influence on my works

– I had very regular discussions with him.

Very often, these extensions began with Philippe, and

– I mostly continued with other collaborators.

– He had a strong influence on my works

– I had very regular discussions with him.

A typical conversation (by phone, usually at the end of the afternoon)

Very often, these extensions began with Philippe, and

- I mostly continued with other collaborators.
- He had a strong influence on my works
- I had very regular discussions with him.

A typical conversation (by phone, usually at the end of the afternoon)

Very often, these extensions began with Philippe, and

– I mostly continued with other collaborators.

– He had a strong influence on my works

– I had very regular discussions with him.

A typical conversation (by phone, usually at the end of the afternoon)

P: On which subject have you worked today?

B: [...] Do you think this question interesting?

The answer was often "yes", but not always "yes"...

Very often, these extensions began with Philippe, and

– I mostly continued with other collaborators.

– He had a strong influence on my works

– I had very regular discussions with him.

A typical conversation (by phone, usually at the end of the afternoon)

P: On which subject have you worked today?

B: [...] Do you think this question interesting?

The answer was often "yes", but not always "yes"...

B: I obtain this formula [*the formula is spelled*].

Do you think it is possible to deal with it?

P: The x, do you put it on the numerator or in the denominator?

B: [....]

Very often, these extensions began with Philippe, and

- I mostly continued with other collaborators.
- He had a strong influence on my works
- I had very regular discussions with him.

A typical conversation (by phone, usually at the end of the afternoon)

P: On which subject have you worked today?

B: [...] Do you think this question interesting?

The answer was often "yes", but not always "yes"...

B: I obtain this formula [*the formula is spelled*].
Do you think it is possible to deal with it?

P: The x, do you put it on the numerator or in the denominator?

B: [....]

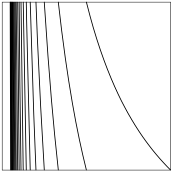
P: I don't know for this formula... But if you have another one ...
[*the formula is spelled*], it would be easier...

B: The i, is it in the exponent?

Dynamical Combinatorics = Extension of Analytical Combinatorics
(Modified)transfer operators viewed as Generating Operators

In Period II, we studied the Euclid dynamical system,
with its transfer operator

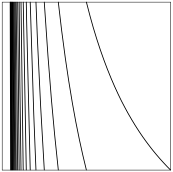
$$\mathbf{G}_s[f](x) = \sum_{h \in \mathcal{H}} |h'(x)|^s \, f \circ h(x)$$



Dynamical Combinatorics = Extension of Analytical Combinatorics

(Modified)transfer operators viewed as Generating Operators

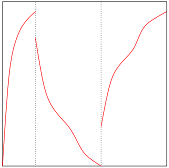
In Period II, we studied the Euclid dynamical system, with its transfer operator



$$\mathbf{G}_s[f](x) = \sum_{h \in \mathcal{H}} |h'(x)|^s \, f \circ h(x)$$

In Period III, extension to any weighted dynamical system with its set $\mathcal{H}$ of branches weighted by some cost c.

It is useful to consider two points x and y,



$$\mathbf{H}_{s,w}[F](x,y) := \sum_{h \in \mathcal{H}} w^{c(h)} \left| \frac{h(x) - h(y)}{x - y} \right|^s F(h(x), h(y))$$

s marks the input size, and w marks the cost c

The secant $\left| \frac{h(x) - h(y)}{x - y} \right|$ replaces the tangent $|h'(x)|$

Dynamical Combinatorics = extension of Analytical Combinatorics.

(Modified) transfer operator $\mathbf{H}_{s,w}$ viewed as a generating operator.

$$\mathbf{H}_{s,w}[F](x,y) := \sum_{h\in\mathcal{H}} w^{c(h)} \left|\frac{h(x)-h(y)}{x-y}\right|^s F(h(x),h(y))$$

Dynamical Combinatorics = extension of Analytical Combinatorics.

(Modified) transfer operator $\mathbf{H}_{s,w}$ viewed as a generating operator.

$$\mathbf{H}_{s,w}[F](x,y) := \sum_{h \in \mathcal{H}} w^{c(h)} \left| \frac{h(x) - h(y)}{x - y} \right|^s F(h(x), h(y))$$

Very often, generating functions involve the quasi-inverse $(I - \mathbf{H}_{s,w})^{-1}$

Dynamical Combinatorics = extension of Analytical Combinatorics.

(Modified) transfer operator $\mathbf{H}_{s,w}$ viewed as a generating operator.

$$\mathbf{H}_{s,w}[F](x,y) := \sum_{h \in \mathcal{H}} w^{c(h)} \left| \frac{h(x)-h(y)}{x-y} \right|^s F(h(x),h(y))$$

Very often, generating functions involve the quasi-inverse $(I - \mathbf{H}_{s,w})^{-1}$

Three main steps

Dynamical Combinatorics = extension of Analytical Combinatorics.
(Modified) transfer operator $\mathbf{H}_{s,w}$ viewed as a generating operator.

$$\mathbf{H}_{s,w}[F](x,y) := \sum_{h \in \mathcal{H}} w^{c(h)} \left| \frac{h(x)-h(y)}{x-y} \right|^{s} F(h(x),h(y))$$

Very often, generating functions involve the quasi-inverse $(I - \mathbf{H}_{s,w})^{-1}$

Three main steps

- Determine
  - the good weighted dynamical system
  - and the related transfer operator adapted to the problem

Dynamical Combinatorics = extension of Analytical Combinatorics.
(Modified) transfer operator $\mathbf{H}_{s,w}$ viewed as a generating operator.

$$\mathbf{H}_{s,w}[F](x,y) := \sum_{h\in\mathcal{H}} w^{c(h)} \left|\frac{h(x)-h(y)}{x-y}\right|^s F(h(x),h(y))$$

Very often, generating functions involve the quasi-inverse $(I-\mathbf{H}_{s,w})^{-1}$

Three main steps

– Determine

– the good weighted dynamical system

– and the related transfer operator adapted to the problem

– Translate the geometric properties of the weighted dynamical system into spectral properties of $\mathbf{H}_{s,w}$, then into analytical properties of $(I-\mathbf{H}_{s,w})^{-1}$

(on a convenient functional space)

[a functional analysis step]

Dynamical Combinatorics = extension of Analytical Combinatorics.
(Modified) transfer operator $\mathbf{H}_{s,w}$ viewed as a generating operator.

$$\mathbf{H}_{s,w}[F](x,y) := \sum_{h \in \mathcal{H}} w^{c(h)} \left| \frac{h(x)-h(y)}{x-y} \right|^s F(h(x),h(y))$$

Very often, generating functions involve the quasi-inverse $(I-\mathbf{H}_{s,w})^{-1}$

Three main steps

– Determine
  – the good weighted dynamical system
  – and the related transfer operator adapted to the problem

– Translate the geometric properties of the weighted dynamical system into spectral properties of $\mathbf{H}_{s,w}$, then into analytical properties of $(I-\mathbf{H}_{s,w})^{-1}$
  (on a convenient functional space)
  [a functional analysis step]

– Translate these analytical properties into asymptotics of coefficients
  [a "classical" step in Analytical Combinatorics]

Euclidean dynamics: study of (almost) any algorithm of Euclid type

The branches are always LFT's, but the dynamics may differ a lot.

Extensions of the methods of Period II in four directions

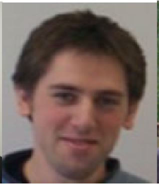
– to other types of Euclid Alg.
  the Binary Alg. – the Lyapounov tortoise and the dyadic hare.

– to other costs (for instance: bit complexity of Euclid Alg).

– to distributional analyses
  with bivariate transfer operators

– to a precise comparison between Euclid and Gauss Alg.



In Information Theory,
the main object is :

[V. 2001]

In Information Theory, the main object is :

[V. 2001]

the Dirichlet series of the source $\Lambda(s) := \sum_{w \in \Sigma^\star} p_w^s$

$p_w :=$ the probability that a word begins with prefix w

In Information Theory, the main object is : [V. 2001]

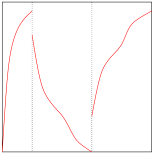
the Dirichlet series of the source $\Lambda(s) := \sum_{w \in \Sigma^\star} p_w^s$

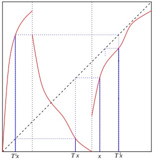
$p_w :=$ the probability that a word begins with prefix w

A general dynamical source

$\Lambda(s)$ closely related to $(I - \mathbf{H}_s)^{-1}$

where $\mathbf{H}_s$ is the (secant) transfer operator of the dynamical system.





$M(x) = (c, b, a, c \ldots)$

In Information Theory, the main object is :

[V. 2001]

the Dirichlet series of the source $\Lambda(s) := \sum_{w \in \Sigma^\star} p_w^s$

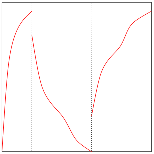
$p_w :=$ the probability that a word begins with prefix w

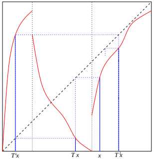
A general dynamical source

$\Lambda(s)$ closely related to $(I - \mathbf{H}_s)^{-1}$

where $\mathbf{H}_s$ is the (secant) transfer operator

of the dynamical system.

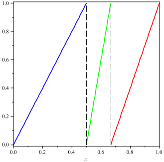




$M(x) = (c, b, a, c \ldots)$

A memoryless source, with probabilities (p_i)

$$\Lambda(s) = \frac{1}{1 - \lambda(s)} \qquad \text{with} \qquad \lambda(s) = \sum_{i=1}^{r} p_i^s$$



In Information Theory, the main object is : [V. 2001]

the Dirichlet series of the source $\Lambda(s) := \sum_{w \in \Sigma^\star} p_w^s$

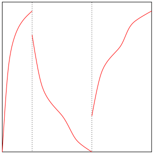
$p_w :=$ the probability that a word begins with prefix w

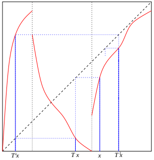
A general dynamical source

$\Lambda(s)$ closely related to $(I - \mathbf{H}_s)^{-1}$

where $\mathbf{H}_s$ is the (secant) transfer operator

of the dynamical system.

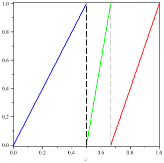




$M(x) = (c, b, a, c \ldots)$

A memoryless source, with probabilities (p_i)

$$\Lambda(s) = \frac{1}{1 - \lambda(s)} \qquad \text{with} \qquad \lambda(s) = \sum_{i=1}^{r} p_i^s$$



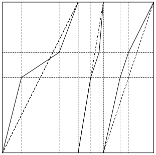
A Markov chain, defined by

– the vector R of initial probabilities (r_i)

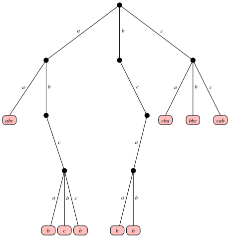
– and the transition matrix $P := (p_{i,j})$

$$\Lambda(s) = 1 + {}^t\mathbf{1}(I - P(s))^{-1} R(s)$$

with $\quad P(s) = (p_{i,j}^s), \quad R(s) = (r_i^s).$



Analysis of general tries



– All the possible node structures: array – list – or BST

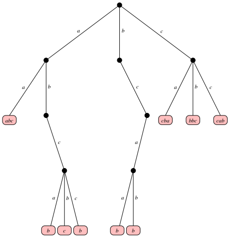
– Words produced by a general (good) dynamical source.

– Possible infinite alphabet

Here, focus on the height and typical depth [2001]:
Both means are of order $\log n$ [FV with J. Clément]

$$\mathbb{E}[D_n] \sim \frac{1}{|\lambda'(1)|} \log n \qquad \mathbb{E}[H_n] \sim \frac{1}{|\log \lambda(2)|} \log n$$

Analysis of general tries



– All the possible node structures: array – list – or BST

– Words produced by a general (good) dynamical source.

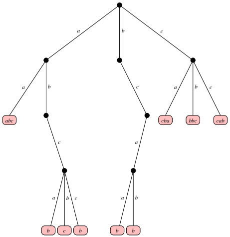
– Possible infinite alphabet

Here, focus on the height and typical depth [2001]:
Both means are of order $\log n$ [FV with J. Clément]

$$\mathbb{E}[D_n] \sim \frac{1}{|\lambda'(1)|} \log n \qquad \mathbb{E}[H_n] \sim \frac{1}{|\log \lambda(2)|} \log n$$

What about the remainder term for the typical depth?
It depends on the shape of a "tameness" region $\mathcal{R} \subset \{\Re s \leq 1, s \neq 1\}$
where $\Lambda(s)$ is analytic and of polynomial growth.

Analysis of general tries



- All the possible node structures: array – list – or BST
- Words produced by a general (good) dynamical source.
- Possible infinite alphabet

Here, focus on the height and typical depth [2001]:
Both means are of order $\log n$ [FV with J. Clément]

$$\mathbb{E}[D_n] \sim \frac{1}{|\lambda'(1)|} \log n \qquad \mathbb{E}[H_n] \sim \frac{1}{|\log \lambda(2)|} \log n$$

What about the remainder term for the typical depth?
It depends on the shape of a "tameness" region $\mathcal{R} \subset \{\Re s \le 1, s \neq 1\}$
where $\Lambda(s)$ is analytic and of polynomial growth.

Later, we study the tameness of a memoryless source,
We relate the shape of an optimal tameness region
with diophantine properties of ratios $\log p_i / \log p_j$.

[FV with M. Roux, 2010]

IV– Towards a "realistic" analysis of algorithms?

– Concept of a general source, completely defined by the set

$p_w :=$ the probability that the word begins with w.

$\Longrightarrow$ Parametrization by the unit interval $\mathcal{I}$

There exists a map $M : \mathcal{I} \to \Sigma^{\mathbb{N}}$, continuous, increasing, s.t each emitted word is (a.e uniquely) written as $M(u)$ with $u \in \mathcal{I}$.

– Analysis of the basic algorithms (sorting and searching alg.) when

- the inputs are random words of a general source
- the cost of a comparison between words
  $=$ the number of symbols needed in a lexicographic comparison.

a b a b b b...
a b a a b a...

coincidence=3; #comparisons=4.

Distribution of the coincidence C

$=$ the length of the longest common prefix

$$\Pr[C \geq k+1] = \sum_{w \in \Sigma^k} p_w^2$$

Realistic analysis of an algorithm using comparisons on keys.

Coincidence γ of the source $\mathcal{S}$

$\gamma(u,t)$ is the coincidence between $M(u)$ and $M(t)$

Density ϕ of the algorithm $\mathcal{A}$ (using comparisons):

$\phi(u,t)$:= the "mean" number of key comparisons between u and t when they are given to Alg. $\mathcal{A}$ after being inserted in a random sequence of $\mathcal{I}^\star$.

We wished to build a general dictionary (and began to do it)

Source $\mathcal{S}$, its coincidence γ and Alg $\mathcal{A}$, its density ϕ	$\Longrightarrow$	Mixed Dirichet series $\varpi(s)$ depends both on the source and the alg.	$\Longrightarrow$	The number S_n of symbol comp. performed by Alg $\mathcal{A}$ on n words

$\Longrightarrow$: $$\mathbb{E}[S_n] = \sum_{k=2}^{n} (-1)^k \binom{n}{k} \varpi(k)$$

A general dictionary

Source $\mathcal{S}$, its coincidence γ and Alg $\mathcal{A}$, its density ϕ	$\Longrightarrow$	Mixed Dirichet series $\varpi(s)$ depends both on the source and the alg.	$\Longrightarrow$	The number S_n of symbol comp. performed by Alg $\mathcal{A}$ on n words

$$\mathbb{E}[S_n] = \sum_{k=2}^{n} (-1)^k \binom{n}{k} \varpi(k)$$

We [$=$ FV with Julien] have already determined the mixed series $\varpi(s)$

– for Tries [2001],

– for Binary Search Trees (or QuickSort) [with J.Fill, 2009]

$$\varpi_T(s) = s\Lambda(s), \qquad \varpi_B(s) = \frac{2\Lambda(s)}{s(s-1)}, \qquad \text{with} \quad \Lambda(s) = \sum_{w \in \Sigma^\star} p_w^s$$

A general dictionary

Source $\mathcal{S}$, its coincidence γ and Alg $\mathcal{A}$, its density ϕ	$\Longrightarrow$	Mixed Dirichet series $\varpi(s)$ depends both on the source and the alg.	$\Longrightarrow$	The number S_n of symbol comp. performed by Alg $\mathcal{A}$ on n words

And now? Determine series $\varpi(s)$ in other contexts!

A general dictionary

Source $\mathcal{S}$, its coincidence γ and Alg $\mathcal{A}$, its density ϕ	$\Longrightarrow$	Mixed Dirichet series $\varpi(s)$ depends both on the source and the alg.	$\Longrightarrow$	The number S_n of symbol comp. performed by Alg $\mathcal{A}$ on n words

And now? Determine series $\varpi(s)$ in other contexts!

This autumn, we have begun to study Insertion Sort and Selection Sort

[with Julien and Thu Hien Nguyen Thi, our PhD student]

A general dictionary

Source $\mathcal{S}$, its coincidence γ and Alg $\mathcal{A}$, its density ϕ	$\Longrightarrow$	Mixed Dirichet series $\varpi(s)$ depends both on the source and the alg.	$\Longrightarrow$	The number S_n of symbol comp. performed by Alg $\mathcal{A}$ on n words

And now? Determine series $\varpi(s)$ in other contexts!

This autumn, we have begun to study Insertion Sort and Selection Sort [with Julien and Thu Hien Nguyen Thi, our PhD student]

Philippe strongly wished to study digital search trees in this framework.

A general dictionary

Source $\mathcal{S}$, its coincidence γ and Alg $\mathcal{A}$, its density ϕ	$\Longrightarrow$	Mixed Dirichet series $\varpi(s)$ depends both on the source and the alg.	$\Longrightarrow$	The number S_n of symbol comp. performed by Alg $\mathcal{A}$ on n words

And now? Determine series $\varpi(s)$ in other contexts!

This autumn, we have begun to study Insertion Sort and Selection Sort
[with Julien and Thu Hien Nguyen Thi, our PhD student]

Philippe strongly wished to study digital search trees in this framework.
– We [FV] began to work on this subject on December 2010.

A general dictionary

Source $\mathcal{S}$, its coincidence γ and Alg $\mathcal{A}$, its density ϕ	$\Longrightarrow$	Mixed Dirichet series $\varpi(s)$ depends both on the source and the alg.	$\Longrightarrow$	The number S_n of symbol comp. performed by Alg $\mathcal{A}$ on n words

And now? Determine series $\varpi(s)$ in other contexts!

This autumn, we have begun to study Insertion Sort and Selection Sort
[with Julien and Thu Hien Nguyen Thi, our PhD student]

Philippe strongly wished to study digital search trees in this framework.
– We [FV] began to work on this subject on December 2010.
– in February 2011, we have planned a co-direction of a PhD thesis.

A general dictionary

Source $\mathcal{S}$, its coincidence γ and Alg $\mathcal{A}$, its density ϕ	$\Longrightarrow$	Mixed Dirichet series $\varpi(s)$ depends both on the source and the alg.	$\Longrightarrow$	The number S_n of symbol comp. performed by Alg $\mathcal{A}$ on n words

And now? Determine series $\varpi(s)$ in other contexts!

This autumn, we have begun to study Insertion Sort and Selection Sort
[with Julien and Thu Hien Nguyen Thi, our PhD student]

Philippe strongly wished to study digital search trees in this framework.
– We [FV] began to work on this subject on December 2010.
– in February 2011, we have planned a co-direction of a PhD thesis.
– The student Kanal Hun will arrive from Cambodgia in January 2012.

A general dictionary

Source $\mathcal{S}$, its coincidence γ and Alg $\mathcal{A}$, its density ϕ	$\Longrightarrow$	Mixed Dirichet series $\varpi(s)$ depends both on the source and the alg.	$\Longrightarrow$	The number S_n of symbol comp. performed by Alg $\mathcal{A}$ on n words

And now? Determine series $\varpi(s)$ in other contexts!

This autumn, we have begun to study Insertion Sort and Selection Sort
[with Julien and Thu Hien Nguyen Thi, our PhD student]

Philippe strongly wished to study digital search trees in this framework.
– We [FV] began to work on this subject on December 2010.
– in February 2011, we have planned a co-direction of a PhD thesis.
– The student Kanal Hun will arrive from Cambodgia in January 2012.

Philippe, we will continue this work.
This would have been your strong wish.