The EUCLID ALGORITHM is “TOTALLY” GAUSSIAN

Brigitte VALLÉE
GREYC (CNRS and University of Caen)

Journées du GT ALEA, Mars 2012

Study of Local Limit Theorems, with their speed of convergence.

Less studied than Central Limit Theorems,

even in the simplest probabilistic framework.

Here: focus on the case of the output of the Euclid Algorithm.

Study of Local Limit Theorems, with their speed of convergence.

Less studied than Central Limit Theorems,

even in the simplest probabilistic framework.

Here: focus on the case of the output of the Euclid Algorithm.

I – The Euclid Algorithm

II- Distributional results which are already known

Central Limit theorems

Local limit theorems in the particular case of a lattice cost.

Study of Local Limit Theorems, with their speed of convergence.

Less studied than Central Limit Theorems,

even in the simplest probabilistic framework.

Here: focus on the case of the output of the Euclid Algorithm.

I – The Euclid Algorithm

II- Distributional results which are already known

Central Limit theorems

Local limit theorems in the particular case of a lattice cost.

III – Local limit theorems for a non-lattice cost

The easier case of memoryless processes

Study of Local Limit Theorems, with their speed of convergence.

Less studied than Central Limit Theorems,

even in the simplest probabilistic framework.

Here: focus on the case of the output of the Euclid Algorithm.

I – The Euclid Algorithm

II- Distributional results which are already known

Central Limit theorems

Local limit theorems in the particular case of a lattice cost.

III – Local limit theorems for a non-lattice cost

The easier case of memoryless processes

IV – Local limit theorems in the case of a dynamical system

Discrete trajectories versus continuous trajectories.

Return to the Euclid algorithm.

I – The Euclid Algorithm

II- Distributional results which are already known

Central Limit theorems

Local limit theorems in the particular case of a lattice cost.

III – Local limit theorems for a non-lattice cost

The easier case of memoryless processes

IV – Local limit theorems in the case of a dynamical system

Discrete trajectories versus continuous trajectories.

Return to the Euclid algorithm.

The (standard) Euclid Algorithm

On the input (u, v), it computes the gcd of u and v,
together with the Continued Fraction Expansion of u/v.

The (standard) Euclid Algorithm

On the input (u, v), it computes the gcd of u and v,
together with the Continued Fraction Expansion of u/v.

$$u_0 := v;\ u_1 := u; u_0 \geq u_1$$

$$\left\{\begin{array}{ccccccc} u_0 & = & m_1 u_1 & + & u_2 & & 0 < u_2 < u_1 \\ u_1 & = & m_2 u_2 & + & u_3 & & 0 < u_3 < u_2 \\ \dots & = & \dots & + & & & \\ u_{p-2} & = & m_{p-1} u_{p-1} & + & u_p & & 0 < u_p < u_{p-1} \\ u_{p-1} & = & m_p u_p & + & 0 & & u_{p+1} = 0 \end{array}\right\}$$

u_p is the gcd of u and v, the m_i's are the digits. p is the depth.

The (standard) Euclid Algorithm

On the input (u, v), it computes the gcd of u and v, together with the Continued Fraction Expansion of u/v.

$$u_0 := v;\ \ u_1 := u; u_0 \geq u_1$$

$$\left\{\begin{array}{ccccccc} u_0 & = & m_1 u_1 & + & u_2 & & 0 < u_2 < u_1 \\ u_1 & = & m_2 u_2 & + & u_3 & & 0 < u_3 < u_2 \\ \ldots & = & \ldots & + & & & \\ u_{p-2} & = & m_{p-1} u_{p-1} & + & u_p & & 0 < u_p < u_{p-1} \\ u_{p-1} & = & m_p u_p & + & 0 & & u_{p+1} = 0 \end{array}\right\}$$

u_p is the gcd of u and v, the m_i's are the digits. p is the depth.

CFE of $\frac{u}{v}$:

$$\frac{u}{v} = \cfrac{1}{m_1 + \cfrac{1}{m_2 + \cfrac{1}{\ddots + \cfrac{1}{m_p}}}},$$

Three main outputs for the Euclid Algorithm

– the $\gcd(u, v)$ itself

Essential in exact rational computations,

for keeping rational numbers under their irreducible forms

60% of the computation time in some symbolic computations

Three main outputs for the Euclid Algorithm

– the $\gcd(u, v)$ itself

Essential in exact rational computations,

for keeping rational numbers under their irreducible forms

60% of the computation time in some symbolic computations

– the modular inverse $u^{-1} \mod v$, when $\gcd(u, v) = 1$.

Extensively used in cryptography

Three main outputs for the Euclid Algorithm

– the $\gcd(u, v)$ itself

Essential in exact rational computations,

for keeping rational numbers under their irreducible forms

60% of the computation time in some symbolic computations

– the modular inverse $u^{-1} \mod v$, when $\gcd(u, v) = 1$.

Extensively used in cryptography

– the Continued Fraction Expansion CFE (u/v)

Often used directly in computation over rationals.

The main object of interest here.

A basic algorithm ... Perhaps the fifth main operation?

The main costs of interest for the continued fraction expansion

With some "digit-cost" d defined on digits m_i, one defines:

$$\widehat{D}(u, v) := \sum_{i=1}^{p} d(m_i)$$

The main costs of interest for the continued fraction expansion

With some "digit-cost" d defined on digits m_i, one defines:

$$\widehat{D}(u,v) := \sum_{i=1}^{p} d(m_i)$$

Main instances:

if $d = 1$, then $\widehat{D}$:= the number of iterations

if $d = \mathbf{1}_{m_0}$, then $\widehat{D}$:= the number of digits equal to m_0

if $d = \ell$ (the binary length), then $\widehat{D}$:= the length of the CFE

The natural costs d take integer values.

The main costs of interest for the continued fraction expansion

With some "digit-cost" d defined on digits m_i, one defines:

$$\widehat{D}(u,v) := \sum_{i=1}^{p} d(m_i)$$

Main instances:

- if $d = 1$, then $\widehat{D}$:= the number of iterations
- if $d = \mathbf{1}_{m_0}$, then $\widehat{D}$:= the number of digits equal to m_0
- if $d = \ell$ (the binary length), then $\widehat{D}$:= the length of the CFE

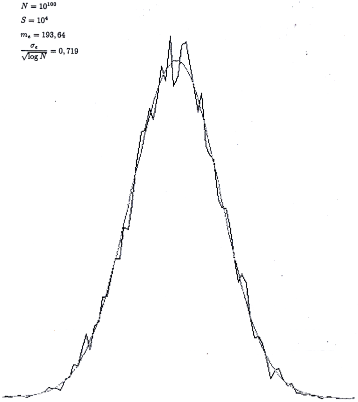
The natural costs d take integer values.

However, it is also interesting to study general digit costs,

They give rise to various observables on the Continued Fraction expansion

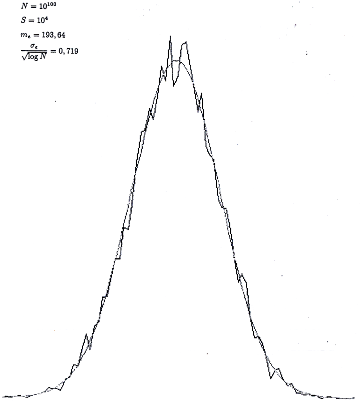
For instance $d(m) = \log m$, ... related to the Khinchine constant.

Main probabilistic questions on the Continued Fraction Expansion
... and its "total" cost $\widehat{D}$



Number of iterations $\widehat{D}$
of the Euclid Algorithm
$d = 1$

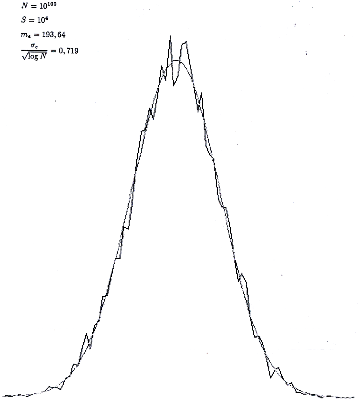
Main probabilistic questions on the Continued Fraction Expansion
... and its "total" cost $\widehat{D}$



Number of iterations $\widehat{D}$
of the Euclid Algorithm
$d = 1$

Analyse in particular, the distribution of D:

Main probabilistic questions on the Continued Fraction Expansion ... and its "total" cost $\widehat{D}$



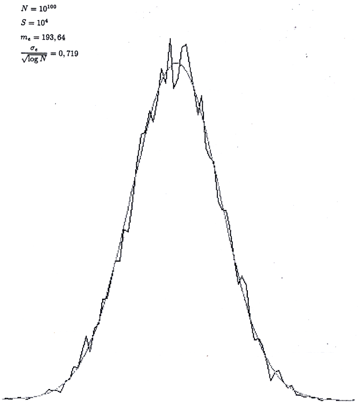
Number of iterations $\widehat{D}$
of the Euclid Algorithm
$d = 1$

Analyse in particular, the distribution of D:

For instance:

A gaussian law for the number of steps?

Main probabilistic questions on the Continued Fraction Expansion
... and its "total" cost $\widehat{D}$



Number of iterations $\widehat{D}$
of the Euclid Algorithm

$d = 1$

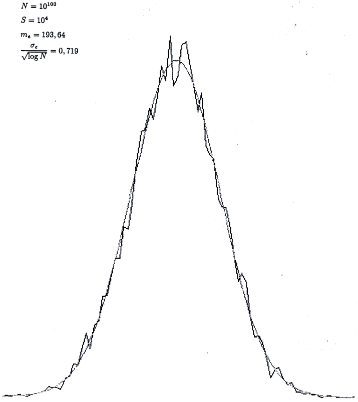
Analyse in particular, the distribution of D:

For instance:

A gaussian law for the number of steps?

Existence of a Central Limit Theorem?

Main probabilistic questions on the Continued Fraction Expansion
... and its "total" cost $\widehat{D}$



Number of iterations $\widehat{D}$
of the Euclid Algorithm
$d = 1$

Analyse in particular, the distribution of D:

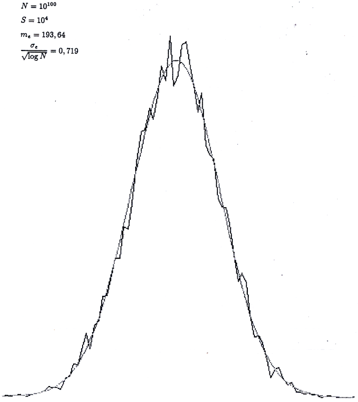
For instance:

A gaussian law for the number of steps?

Existence of a Central Limit Theorem?

Existence of a Local Limit Theorem?

Main probabilistic questions on the Continued Fraction Expansion
... and its "total" cost $\widehat{D}$



Number of iterations $\widehat{D}$
of the Euclid Algorithm
$d = 1$

Analyse in particular, the distribution of D:

For instance:
A gaussian law for the number of steps?

Existence of a Central Limit Theorem?

Existence of a Local Limit Theorem?

Which speed of convergence?

The underlying dynamical system (I).

The trace of the execution of the Euclid Algorithm on (u_1, u_0) is:

$$(u_1, u_0) \to (u_2, u_1) \to (u_3, u_2) \to \ldots \to (u_{p-1}, u_p) \to (u_{p+1}, u_p) = (0, u_p)$$

The underlying dynamical system (I).

The trace of the execution of the Euclid Algorithm on (u_1, u_0) is:

$$(u_1, u_0) \to (u_2, u_1) \to (u_3, u_2) \to \ldots \to (u_{p-1}, u_p) \to (u_{p+1}, u_p) = (0, u_p)$$

Replace the integer pair (u_i, u_{i-1}) by the rational $x_i := \dfrac{u_i}{u_{i-1}}$.

The division $u_{i-1} = m_i u_i + u_{i+1}$ is then written as

$$x_{i+1} = \frac{1}{x_i} - \left\lfloor \frac{1}{x_i} \right\rfloor \qquad \text{or} \qquad x_{i+1} = T(x_i), \qquad \text{where}$$

$$T : [0,1] \longrightarrow [0,1], \quad T(x) := \frac{1}{x} - \left\lfloor \frac{1}{x} \right\rfloor \quad \text{for} \ \ x \neq 0, \quad T(0) = 0$$

The underlying dynamical system (I).

The trace of the execution of the Euclid Algorithm on (u_1, u_0) is:

$$(u_1, u_0) \to (u_2, u_1) \to (u_3, u_2) \to \ldots \to (u_{p-1}, u_p) \to (u_{p+1}, u_p) = (0, u_p)$$

Replace the integer pair (u_i, u_{i-1}) by the rational $x_i := \frac{u_i}{u_{i-1}}$.

The division $u_{i-1} = m_i u_i + u_{i+1}$ is then written as

$$x_{i+1} = \frac{1}{x_i} - \left\lfloor \frac{1}{x_i} \right\rfloor \qquad \text{or} \qquad x_{i+1} = T(x_i), \qquad \text{where}$$

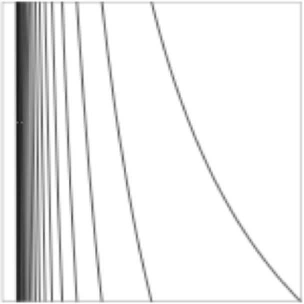
$$T : [0,1] \longrightarrow [0,1], \quad T(x) := \frac{1}{x} - \left\lfloor \frac{1}{x} \right\rfloor \quad \text{for } \; x \neq 0, \quad T(0) = 0$$

An execution of the Euclidean Algorithm $(x, T(x), T^2(x), \ldots, 0)$

= A rational trajectory of the Dynamical System $([0,1], T)$

= a trajectory that reaches 0.

The dynamical system is a continuous extension of the algorithm.



$$T(x) := \frac{1}{x} - \left\lfloor \frac{1}{x} \right\rfloor$$

$$T_{[m]} : \left]\frac{1}{m+1}, \frac{1}{m}\right[\longrightarrow]0,1[,$$

$$T_{[m]}(x) := \frac{1}{x} - m$$

$$h_{[m]} :]0,1[\longrightarrow \left]\frac{1}{m+1}, \frac{1}{m}\right[$$

$$h_{[m]}(x) := \frac{1}{m+x}$$

$$\frac{u}{v} = \cfrac{1}{m_1 + \cfrac{1}{m_2 + \cfrac{1}{\ddots + \cfrac{1}{m_p}}}} = h_{[m_1]} \circ h_{[m_2]} \circ \ldots \circ h_{[m_p]}(0)$$

The discrete algorithm is extended into a continuous process.

Two types of weighted trajectories and two probabilistic models:

The discrete algorithm is extended into a continuous process.

Two types of weighted trajectories and two probabilistic models:

First model : Study of truncated real trajectories "at depth n"

For a random $x \in \mathcal{I}$ $\qquad D_n(x) := \sum_{i=1}^{n} d(m_i(x))$

The discrete algorithm is extended into a continuous process.

Two types of weighted trajectories and two probabilistic models:

First model : Study of truncated real trajectories "at depth n"

For a random $x \in \mathcal{I}$ $\qquad D_n(x) := \sum_{i=1}^{n} d(m_i(x))$

Second model: Study of rational trajectories "of denominator N"

on $\Omega_N := \{x = u/v \in \mathcal{I}, v = N\}$

For a random $x \in \Omega_N$ $\qquad \widehat{D}_N(x) := \sum_{i=1}^{P(x)} d(m_i(x)),$

The discrete algorithm is extended into a continuous process.

Two types of weighted trajectories and two probabilistic models:

First model : Study of truncated real trajectories "at depth n"

For a random $x \in \mathcal{I}$ $$D_n(x) := \sum_{i=1}^{n} d(m_i(x))$$

Second model: Study of rational trajectories "of denominator N"

on $\Omega_N := \{x = u/v \in \mathcal{I}, v = N\}$

For a random $x \in \Omega_N$ $$\widehat{D}_N(x) := \sum_{i=1}^{P(x)} d(m_i(x)),$$

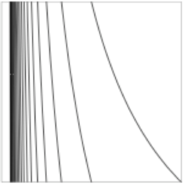
We wish to compare these two "observables" .

Since the discrete data are of zero measure amongst the continuous data,

we need a "transfer from continuous to discrete".

A main tool in both probabilistic models: The transfer operator.

The transfer operator



$\mathcal{H} :=$ the set of the inverse branches of T.

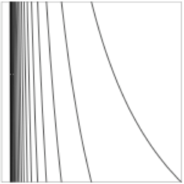
$$T(x) := \frac{1}{x} - \left\lfloor \frac{1}{x} \right\rfloor$$

Density Transformer:

For a density f on $[0,1]$, $\mathbf{H}[f]$ is the density on $[0,1]$ after one iteration of the shift

$$\mathbf{H}[f](x) = \sum_{h \in \mathcal{H}} |h'(x)|\, f \circ h(x) = \sum_{m \in \mathbb{N}} \frac{1}{(m+x)^2} f\left(\frac{1}{m+x}\right).$$

The transfer operator



$\mathcal{H}$:= the set of the inverse branches of T.

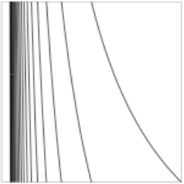
$$T(x) := \frac{1}{x} - \left\lfloor \frac{1}{x} \right\rfloor$$

Density Transformer:

For a density f on $[0,1]$, $\mathbf{H}[f]$ is the density on $[0,1]$ after one iteration of the shift

$$\mathbf{H}[f](x) = \sum_{h \in \mathcal{H}} |h'(x)| \, f \circ h(x) = \sum_{m \in \mathbb{N}} \frac{1}{(m+x)^2} f\left(\frac{1}{m+x}\right).$$

The transfer operator



$\mathcal{H} :=$ the set of the inverse branches of T.

$$T(x) := \frac{1}{x} - \left\lfloor \frac{1}{x} \right\rfloor$$

Density Transformer:

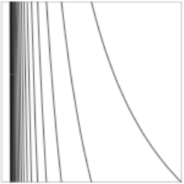
For a density f on $[0,1]$, $\mathbf{H}[f]$ is the density on $[0,1]$ after one iteration of the shift

$$\mathbf{H}[f](x) = \sum_{h \in \mathcal{H}} |h'(x)| \, f \circ h(x) = \sum_{m \in \mathbb{N}} \frac{1}{(m+x)^2} f\left(\frac{1}{m+x}\right).$$

Weighted transfer operator relative to a digit-cost d

$$\mathbf{H}_{s,w}[f](x) = \sum_{h \in \mathcal{H}} |h'(x)|^s \, e^{wd(h)} \, f \circ h(x).$$

The transfer operator



$\mathcal{H} :=$ the set of the inverse branches of T.

$$T(x) := \frac{1}{x} - \left\lfloor \frac{1}{x} \right\rfloor$$

Density Transformer:

For a density f on $[0,1]$, $\mathbf{H}[f]$ is the density on $[0,1]$ after one iteration of the shift

$$\mathbf{H}[f](x) = \sum_{h\in\mathcal{H}} |h'(x)|\, f\circ h(x) = \sum_{m\in\mathbb{N}} \frac{1}{(m+x)^2} f(\frac{1}{m+x}).$$

Weighted transfer operator relative to a digit-cost d

$$\mathbf{H}_{s,w}[f](x) = \sum_{h\in\mathcal{H}} |h'(x)|^s\, e^{wd(h)}\, f\circ h(x).$$

The k-th iterate satisfies, with d extended in an additive way

$$\mathbf{H}^k_{s,w}[f](x) = \sum_{h\in\mathcal{H}^k} |h'(x)|^s\; e^{wd(h)} f\circ h(x)$$

II- Distributional results for weighted trajectories

Transfer operator and distributional study of weighted trajectories

In distributional studies, the main tools are the characteristic functions

$$\mathbb{E}[\exp(wD_n)], \qquad \mathbb{E}_N[\exp(w\widehat{D})]$$

Transfer operator and distributional study of weighted trajectories

In distributional studies, the main tools are the characteristic functions

$$\mathbb{E}[\exp(wD_n)], \qquad \mathbb{E}_N[\exp(w\widehat{D})]$$

Real case: $$\mathbb{E}[\exp(wD_n)] = \int_{\mathcal{I}} \mathbf{H}_{1,w}^n[1](t)dt$$

Transfer operator and distributional study of weighted trajectories

In distributional studies, the main tools are the characteristic functions

$$\mathbb{E}[\exp(wD_n)], \qquad \mathbb{E}_N[\exp(w\widehat{D})]$$

Real case: $\quad \mathbb{E}[\exp(wD_n)] = \int_{\mathcal{I}} \mathbf{H}_{1,w}^n[1](t)dt$

Rational case : $\quad \mathbb{E}_N[\exp(w\widehat{D})]$ related to $\quad [N^{-s}](I - \mathbf{H}_{s,w})^{-1}[1](0)$

due to the relation between

Dirichlet generating functions and quasi-inverses of the transfer operator,

$$S_d(s,w) := \sum_{(u,v)\in\Omega} \frac{1}{v^{2s}} \exp[w\widehat{D}(u,v)] = (I - \mathbf{H}_{s,w})^{-1}[1](0)$$

Distributional results for the continued fraction expansion

Already known results [Baladi-V (2003)]

In both cases, Real trajectories or Rational trajectories,

For a cost d of moderate growth $d(m) = O(\log m)$,

(a) Central Limit Theorems hold for $D_n, \widehat{D}_N$

(b) Moreover, for a lattice cost, Local Limit Theorems hold for $D_n, \widehat{D}_N$

$$\exists d_0, L \in \mathbb{R}, \quad \text{with } L > 0, \text{ such that} \quad \forall m \quad \frac{d(m) - d_0}{L} \in \mathbb{Z}$$

(c) With optimal speed of convergence

$$O\left(\frac{1}{\sqrt{n}}\right), \quad O\left(\frac{1}{\sqrt{\log N}}\right)$$

Distributional results for the continued fraction expansion

They deal with the characteristic functions $\mathbb{E}[\exp(wD_n)], \mathbb{E}_N[\exp(w\widehat{D})]$
and thus with the transfer operator $\mathbf{H}_{s,w}$

Different cases of study for parameters s and w

Distributional results for the continued fraction expansion

They deal with the characteristic functions $\mathbb{E}[\exp(wD_n)], \mathbb{E}_N[\exp(w\widehat{D})]$ and thus with the transfer operator $\mathbf{H}_{s,w}$

Different cases of study for parameters s and w

For parameter s

- Real trajectories: $s = 1$
- Rational trajectories $s = 1 + it$, with $t \in \mathbb{R}$

Distributional results for the continued fraction expansion

They deal with the characteristic functions $\mathbb{E}[\exp(wD_n)], \mathbb{E}_N[\exp(w\widehat{D})]$
and thus with the transfer operator $\mathbf{H}_{s,w}$

Different cases of study for parameters s and w

For parameter s

- Real trajectories: $s = 1$
- Rational trajectories $s = 1 + it$, with $t \in \mathbb{R}$

For parameter w:

- Central Limit Theorems:

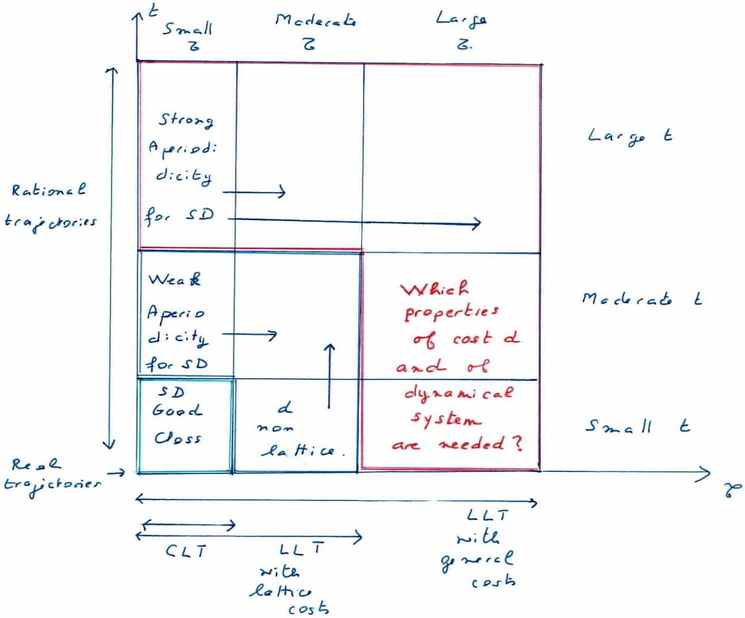
  $w \sim 0$
- Local Limit Theorems for a lattice cost :

  $w = i\tau$ with $\tau \in K$ compact $\subset \mathbb{R}$
- Local Limit Theorems for a non lattice cost :

  $w = i\tau$ with $\tau \in \mathbb{R}$

Properties of the dynamical system and cost needed in distributional studies

for dealing with the operator $\mathbf{H}_{1+it,i\tau}$ in each each domain (t,τ).



III- Local limit theorems with speed of convergence in simpler cases

Memoryless case.

Let (X_i) be a i.i.d sequence with values in $\mathbb{N}$, and $p_m := \Pr[X_i = m]$.
A cost $d : \mathbb{N} \to \mathbb{R}^+$, Some technical conditions:

$$\sigma_0 := \inf\{\sigma;\ \sum_{i=1}^{\infty} p_m^{\sigma} < \infty\} < 1, \qquad d(m) = O(|\log p_m|)$$

The mean $\mu[d]$ and the standard deviation $\sigma[d]$ exist. We assume $\sigma[d] \neq 0$.

Let (X_i) be a i.i.d sequence with values in $\mathbb{N}$, and $p_m := \Pr[X_i = m]$.
A cost $d : \mathbb{N} \to \mathbb{R}^+$, Some technical conditions:

$$\sigma_0 := \inf\{\sigma;\ \sum_{i=1}^{\infty} p_m^\sigma < \infty\} < 1, \qquad d(m) = O(|\log p_m|)$$

The mean $\mu[d]$ and the standard deviation $\sigma[d]$ exist. We assume $\sigma[d] \neq 0$.

Main subject of interest: $\qquad D_n := \sum_{i=1}^{n} d(X_i) \qquad (n \to \infty).$

Let (X_i) be a i.i.d sequence with values in $\mathbb{N}$, and $p_m := \Pr[X_i = m]$.
A cost $d : \mathbb{N} \to \mathbb{R}^+$, Some technical conditions:

$$\sigma_0 := \inf\{\sigma;\ \sum_{i=1}^{\infty} p_m^{\sigma} < \infty\} < 1, \qquad d(m) = O(|\log p_m|)$$

The mean $\mu[d]$ and the standard deviation $\sigma[d]$ exist. We assume $\sigma[d] \neq 0$.

Main subject of interest: $\qquad D_n := \sum_{i=1}^{n} d(X_i) \qquad (n \to \infty).$

There is a Central Limit Theorem (CLT) for D_n
with a speed of convergence of order $O(1/\sqrt{n})$,

$$\Pr\left[\frac{D_n - n\mu[d]}{\sigma[d]\sqrt{n}} \leq y\right] - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{y} e^{-t^2/2} dt = O\left(\frac{1}{\sqrt{n}}\right).$$

A Local Limit Theorem (LLT)

- deals with $Q(x, n) := \mu[d]n + \delta[d]x\sqrt{n}$,
- evaluates the probability that $D_n - Q(x, n)$ belongs to some $J \subset \mathbb{R}$,
- compares it to $(|J|/\sqrt{2\pi n})\, e^{-x^2/2}$.

A Local Limit Theorem (LLT) proves that

$$\sqrt{n}\ \Pr[D_n - \mathbf{Q}(x, n) \in J] - |J| \frac{e^{-x^2/2}}{\delta(d)\sqrt{2\pi}} \to 0 \qquad (n \to \infty).$$

A Local Limit Theorem (LLT)

- deals with $Q(x,n) := \mu[d]n + \delta[d]x\sqrt{n}$,
- evaluates the probability that $D_n - Q(x,n)$ belongs to some $J \subset \mathbb{R}$,
- compares it to $(|J|/\sqrt{2\pi n})\, e^{-x^2/2}$.

A Local Limit Theorem (LLT) proves that

$$\sqrt{n}\ \Pr[D_n - \mathbf{Q}(x,n) \in J] - |J|\frac{e^{-x^2/2}}{\delta(d)\sqrt{2\pi}} \to 0 \qquad (n \to \infty).$$

What about the speed of convergence?

It depends on arithmetical properties of cost d. Two main cases:

the lattice case, and the non–lattice case.

A cost d is lattice if

$$\exists d_0, L \in \mathbb{R}, \quad \text{with } L > 0, \text{ such that } \quad \forall m \quad \frac{d(m) - d_0}{L} \in \mathbb{Z}$$

The smallest possible $L > 0$ is called the span of the lattice cost.

If $d_0 = 0$, the cost is called “plain lattice”.

In the lattice case, the optimal speed, of order $O(1/\sqrt{n})$ is attained.
More precisely, for a plain lattice cost of span 1, one has

$$\sqrt{n}\,\Pr[D_n = P(x,n)] = \sqrt{2\pi}\frac{e^{-x^2/2}}{\delta(d)} + O\left(\frac{1}{\sqrt{n}}\right) \qquad P(x,n) := \lfloor Q(x,n) \rfloor.$$

In this case, the characteristic function ϕ is periodic

$$\phi(\tau) := \int_{\mathbb{R}} \exp[i\tau x]\, dP_d(x) = \sum_{m \geq 1} p_m \exp[i\tau d(m)],$$

In the lattice case, the optimal speed, of order $O(1/\sqrt{n})$ is attained.
More precisely, for a plain lattice cost of span 1, one has

$$\sqrt{n}\,\Pr[D_n = P(x,n)] = \sqrt{2\pi}\frac{e^{-x^2/2}}{\delta(d)} + O\left(\frac{1}{\sqrt{n}}\right) \qquad P(x,n) := \lfloor Q(x,n) \rfloor.$$

In this case, the characteristic function ϕ is periodic

$$\phi(\tau) := \int_{\mathbb{R}} \exp[i\tau x]\, dP_d(x) = \sum_{m\geq 1} p_m \exp[i\tau d(m)],$$

In the non–lattice case, the speed in the LLT depends

– on the behaviour of the characteristic function ϕ of cost d, when $\tau \to \infty$

– on arithmetic properties of cost d

which measures the "difference" between the cost d and a lattice cost.

In the lattice case, the optimal speed, of order $O(1/\sqrt{n})$ is attained.
More precisely, for a plain lattice cost of span 1, one has

$$\sqrt{n}\,\Pr[D_n = P(x,n)] = \sqrt{2\pi}\frac{e^{-x^2/2}}{\delta(d)} + O\left(\frac{1}{\sqrt{n}}\right) \qquad P(x,n) := \lfloor Q(x,n) \rfloor.$$

In this case, the characteristic function ϕ is periodic

$$\phi(\tau) := \int_{\mathbb{R}} \exp[i\tau x]\, dP_d(x) = \sum_{m \geq 1} p_m \exp[i\tau d(m)],$$

In the non–lattice case, the speed in the LLT depends

– on the behaviour of the characteristic function ϕ of cost d, when $\tau \to \infty$

– on arithmetic properties of cost d

which measures the "difference" between the cost d and a lattice cost.

Important fact: There is a relation between these two properties.

Proposition (classical and easy). The conditions are equivalent :

(i) The cost d is lattice

(ii) There exists $\tau_0 \neq 0$ for which ϕ_d satisfies $|\phi_d(\tau_0)| = 1$.

Moreover, Condition (ii) entails Condition (iii)

(iii) For any $h, k, \ell \in \mathbb{N}$, the ratio $\dfrac{d(h) - d(k)}{d(h) - d(\ell)}$ is rational.

Reinforcements of negations of Conditions (ii) or (iii).

A cost d is of characteristic exponent χ if

$$\exists K, \tau_0 > 0, \qquad |\phi_d(\tau)| \leq 1 - \frac{K}{|\tau|^\chi} \qquad \text{for } |\tau| \geq \tau_0.$$

A cost d is of diophantine exponent μ if

$$\exists (h, k, \ell) \in \mathbb{N}^3, \quad \text{such that the ratio} \quad \frac{d(h) - d(k)}{d(h) - d(\ell)} \quad \text{is } \mathtt{Diop}\ (\mu)$$

A number x is diophantine of exponent μ if

$$\exists C > 0, \quad \forall (p, q) \in \mathbb{N}^2, \quad \text{one has:} \qquad \left| x - \frac{p}{q} \right| > \frac{C}{q^{2+\mu}}$$

First result (Breuillard)

The cost d is of characteristic exponent χ $\Longrightarrow$ a Local Limit Theorem for D_n with speed $n^{1/\chi}$

For any ϵ with $\epsilon < 1/\chi$, for any compact interval $J \subset \mathbb{R}$,

there exists M_J, so that $\forall x \in \mathbb{R}, \forall n \geq 1$, one has:

$$\left| \sqrt{n}\ \Pr[D_n(u) - \mathbf{Q}(x,n) \in J] - |J| \frac{e^{-x^2/2}}{\delta(d)\sqrt{2\pi}} \right| \leq \frac{M_J}{n^\epsilon}$$

Second result (Breuillard)

The cost d is of diophantine exponent μ, $\Longrightarrow$ d of characteristic exponent χ for any $\chi > 2(\mu+1)$.

Conclusion:

The cost d is of diophantine exponent μ, $\Longrightarrow$ a Local Limit Theorem for D_n with speed $n^{1/2(\mu+1)}$.

IV- Local limit theorems with speed of convergence

Trajectories of dynamical systems.

And now if the X_i are generated by a dynamical system?

For instance the digits of the continued fraction expansion

(they are no longer independent)

Case of real trajectories

Definition: d is of characteristic exponent χ (wrt to the DS), if,

$$||\mathbf{H}_{1,i\tau}^{n(\tau)}|| \leq 1 - \frac{1}{|\tau|^{\chi}}, \qquad \text{for any } \tau \text{ with } |\tau| \geq \tau_0 \qquad n(\tau) := \Theta(\log|\tau]).$$

Two properties:

The cost d is of characteristic exponent χ wrt to the DS

$\Longrightarrow$ a Local Limit Theorem for D_n with speed $n^{1/\chi}$

The cost d is of diophantine exponent μ,

$\Longrightarrow$ d of characteristic exponent χ for any χ with $\chi > K(\mu + 1)$.

K depends on the DS.

A good generalization of the memoryless case.

Case of rational trajectories.

Definition: d is of uniform characteristic exponent χ

$$||\mathbf{H}_{1+it,i\tau}^{n(\tau)}|| \leq 1 - \frac{1}{|\tau|^{\chi}}, \qquad \text{for any } (t,\tau) \text{ with } |t| \leq a \text{ and } |\tau| \geq \tau_0.$$

NOW: (Baladi-Hachemi)

The cost d is of uniform characteristic exponent χ

$\Longrightarrow$ a Local Limit Theorem for $\hat{D}_N$ with speed $(\log N)^{1/\chi}$

Case of rational trajectories.

Definition: d is of uniform characteristic exponent χ

$$||\mathbf{H}_{1+it,i\tau}^{n(\tau)}|| \le 1 - \frac{1}{|\tau|^{\chi}}, \qquad \text{for any } (t,\tau) \text{ with } |t| \le a \text{ and } |\tau| \ge \tau_0.$$

NOW: (Baladi-Hachemi)

The cost d is of uniform characteristic exponent χ

$\Longrightarrow$ a Local Limit Theorem for $\hat{D}_N$ with speed $(\log N)^{1/\chi}$

HOWEVER (Baladi-Hachemi)

The property : "The cost d is of diophantine exponent μ",

is A PRIORI NOT sufficient to entail

"d is of uniform characteristic exponent χ for any χ with $\chi > K(\mu+1)$".

Case of rational trajectories.

Definition: d is of uniform characteristic exponent χ

$$||\mathbf{H}_{1+it,i\tau}^{n(\tau)}|| \le 1 - \frac{1}{|\tau|^{\chi}}, \qquad \text{for any } (t,\tau) \text{ with } |t| \le a \text{ and } |\tau| \ge \tau_0.$$

NOW: (Baladi-Hachemi)

The cost d is of uniform characteristic exponent χ

$\Longrightarrow$ a Local Limit Theorem for $\hat{D}_N$ with speed $(\log N)^{1/\chi}$

HOWEVER (Baladi-Hachemi)

The property : "The cost d is of diophantine exponent μ",

is A PRIORI NOT sufficient to entail

"d is of uniform characteristic exponent χ for any χ with $\chi > K(\mu+1)$".

Baladi and Hachemi proposed an intertwined diophantine condition

involving the branches of the dynamical system AND the cost d

Our result:

A set of two conditions NOT intertwined

- The diophantine condition (D) on the cost d
- A (new) condition (C) on the branches of the DS

  a "diophantine" version of the aperiodicity condition on the DS.

Our result:

A set of two conditions NOT intertwined

- The diophantine condition (D) on the cost d
- A (new) condition (C) on the branches of the DS
  a "diophantine" version of the aperiodicity condition on the DS.

The Aperiodicity Condition says :

"The branches of the system do not have all the same shape".

If $h^\star$ is the fixed point of branch h,

This implies that the cost $c(h) := \log |h'(h^\star)|$ is strongly non additive,

and then very often $\Gamma(h,k) := c(h \circ k) - c(h) - c(k) \neq 0$

Our result:

A set of two conditions NOT intertwined

- The diophantine condition (D) on the cost d
- A (new) condition (C) on the branches of the DS

  a "diophantine" version of the aperiodicity condition on the DS.

The Aperiodicity Condition says :

"The branches of the system do not have all the same shape".

If $h^\star$ is the fixed point of branch h,

This implies that the cost $c(h) := \log|h'(h^\star)|$ is strongly non additive,

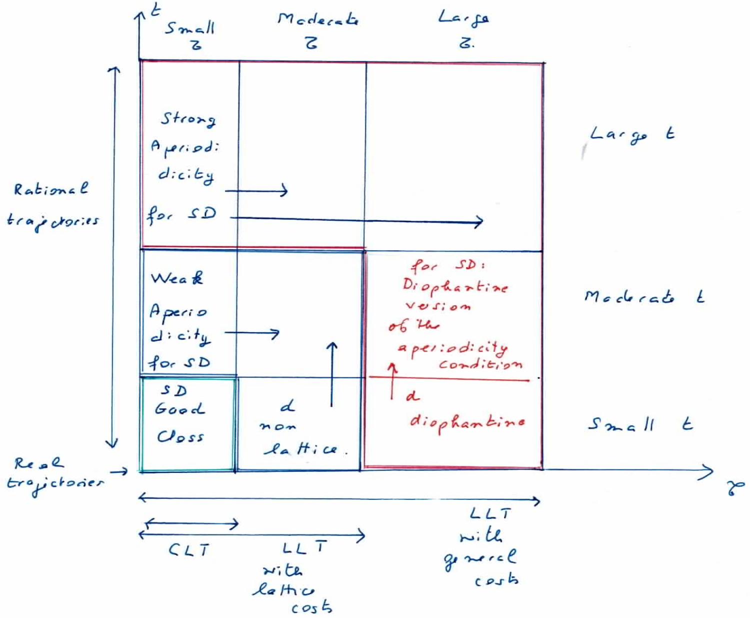
and then very often $\Gamma(h,k) := c(h \circ k) - c(h) - c(k) \neq 0$

Our condition (C). There exist three branches h, k, ℓ for which

$$\Gamma(h,k) \neq 0, \quad \Gamma(h,\ell) \neq 0, \quad \text{and} \quad \frac{\Gamma(h,k)}{\Gamma(h,\ell)} \quad \text{is diophantine.}$$

Properties of the dynamical system and cost needed in distributional studies

for dealing with the operator $\mathbf{H}_{1+it,i\tau}$ in each each domain (t,τ).



Return to the Euclid dynamical system.

In this case, the condition (C) is always satisfied.

Let $c(h) := \log |h'(h^\star)|$. There exist three branches h, k, ℓ for which $$\Gamma(h,k) \neq 0, \quad \Gamma(h,\ell) \neq 0, \quad \text{and} \quad \frac{\Gamma(h,k)}{\Gamma(h,\ell)} \quad \text{is diophantine.}$$

Why?

– The fixed point $h^\star$ of h and $|h'(h^\star)|$ are algebraic numbers.

Return to the Euclid dynamical system.

In this case, the condition (C) is always satisfied.

Let $c(h) := \log |h'(h^\star)|$. There exist three branches h, k, ℓ for which $$\Gamma(h,k) \neq 0, \quad \Gamma(h,\ell) \neq 0, \quad \text{and} \quad \frac{\Gamma(h,k)}{\Gamma(h,\ell)} \quad \text{is diophantine.}$$

Why?

- The fixed point $h^\star$ of h and $|h'(h^\star)|$ are algebraic numbers.
- Then $\Gamma(h,k)$ equals the logarithm of an algebraic number $\alpha(h,k)$

Return to the Euclid dynamical system.

In this case, the condition (C) is always satisfied.

Let $c(h) := \log|h'(h^\star)|$. There exist three branches h, k, ℓ for which

$$\Gamma(h,k) \neq 0, \quad \Gamma(h,\ell) \neq 0, \quad \text{and} \quad \frac{\Gamma(h,k)}{\Gamma(h,\ell)} \quad \text{is diophantine.}$$

Why?

- The fixed point $h^\star$ of h and $|h'(h^\star)|$ are algebraic numbers.
- Then $\Gamma(h,k)$ equals the logarithm of an algebraic number $\alpha(h,k)$
- There exist h, k, ℓ such that

  $\alpha(h,k)$ and $\alpha(h,\ell)$ be algebraically independent.

Return to the Euclid dynamical system.

In this case, the condition (C) is always satisfied.

Let $c(h) := \log |h'(h^\star)|$. There exist three branches h, k, ℓ for which

$$\Gamma(h,k) \neq 0, \quad \Gamma(h,\ell) \neq 0, \quad \text{and} \quad \frac{\Gamma(h,k)}{\Gamma(h,\ell)} \quad \text{is diophantine.}$$

Why?

– The fixed point $h^\star$ of h and $|h'(h^\star)|$ are algebraic numbers.

– Then $\Gamma(h,k)$ equals the logarithm of an algebraic number $\alpha(h,k)$

– There exist h, k, ℓ such that

$\alpha(h,k)$ and $\alpha(h,\ell)$ be algebraically independent.

– Baker's theorem proves that the ratio $\Gamma(h,k)/\Gamma(h,\ell)$ is diophantine.

The final result,

for the total costs of a continued fraction relative to some cost d.

$$\widehat{D}_N(x) := \sum_{i=1}^{P(x)} d(m_i(x)) \qquad \text{on} \quad \Omega_N := \{x = p/q; \quad q \le N\}$$

$$D_n(x) := \sum_{i=1}^{n} d(m_i(x)) \qquad \text{on} \quad \mathcal{I}$$

The final result,

for the total costs of a continued fraction relative to some cost d.

$$\widehat{D}_N(x) := \sum_{i=1}^{P(x)} d(m_i(x)) \qquad \text{on} \quad \Omega_N := \{x = p/q; \quad q \le N\}$$

$$D_n(x) := \sum_{i=1}^{n} d(m_i(x)) \qquad \text{on} \quad \mathcal{I}$$

If the non lattice cost d is

- of moderate growth $[d(m) = O(\log m)]$
- of diophantine exponent (μ, θ),

there is a Local Limit Theorem for costs $\widehat{D}_N$, D_n

with a speed of convergence $\quad O\left(\frac{1}{\log^\epsilon N}\right) \quad$ or $\quad O\left(\frac{1}{n^\epsilon}\right)$

The final result,

for the total costs of a continued fraction relative to some cost d.

$$\widehat{D}_N(x) := \sum_{i=1}^{P(x)} d(m_i(x)) \qquad \text{on} \quad \Omega_N := \{x = p/q; \quad q \le N\}$$

$$D_n(x) := \sum_{i=1}^{n} d(m_i(x)) \qquad \text{on} \quad \mathcal{I}$$

If the non lattice cost d is

- of moderate growth $[d(m) = O(\log m)]$
- of diophantine exponent (μ, θ),

there is a Local Limit Theorem for costs $\widehat{D}_N$, D_n

with a speed of convergence $\quad O\left(\frac{1}{\log^\epsilon N}\right) \quad$ or $\quad O\left(\frac{1}{n^\epsilon}\right)$

with $\quad \epsilon > \frac{1}{2(\mu+1)(2+\theta/\theta_0)}$.