

**DISTRIBUTIONAL ANALYSES  
OF EUCLIDEAN ALGORITHMS**

*Or...*

**EUCLIDEAN ALGORITHMS ARE GAUSSIAN**

*An Instance of a Dynamical Analysis*

Brigitte VALLÉE (CNRS and Université de Caen, France)

Joint work with Viviane BALADI

(CNRS and Université de Paris VI)

## The Euclid Algorithm.

On the input  $(u, v)$ , it computes the **gcd** of  $u$  and  $v$ , together with the **Continued Fraction Expansion** of  $u/v$ .  $v_0 := v$ ;  $v_1 := u$ ;  $v_0 \geq v_1$

$$\left. \begin{array}{l} v_0 = m_1 v_1 + v_2 \quad 0 \leq v_2 < v_1 \\ v_1 = m_2 v_2 + v_3 \quad 0 \leq v_3 < v_2 \\ \dots = \dots + \dots \\ v_{p-2} = m_{p-1} v_{p-1} + v_p \quad 0 \leq v_p < v_{p-1} \\ v_{p-1} = m_p v_p + 0 \quad v_{p+1} = 0 \end{array} \right\}$$

$v_p$  is the **gcd** of  $u$  and  $v$ . ( $m_1, m_2, \dots, m_p$ ) are the **digits**.

$$\text{CFE of } \frac{u}{v}: \quad \frac{u}{v} = \frac{1}{m_1 + \frac{1}{m_2 + \frac{1}{\dots + \frac{1}{m_p}}}},$$

## Variants of the Euclidean Algorithms.

A Euclidean algorithm:= Any algorithm which performs

a **sequence of divisions**  $v = mu + r$ .

Various possible divisions, according to

- the **position of the remainder**  $r$   
(Division By-Default, By Excess, Centered)
- the **parity of the quotient**  $m$   
(Odd divisions, Even divisions)
- A **sequence of  $m$  subtractions** may replace the division with quotient  $m$ .

A division  $v = mu + r$  can be replaced by a **pseudo-division** where powers of 2 are removed from the remainder  $r$ ,  $v = mu + 2^b s$ ,  $s$  odd.

(Binary Algorithm, Hensel divisions)

## Cost of an execution.

Given a step-cost  $c : \mathbf{N}^* \mapsto \mathbf{R}^+$  which depends only on the digit,

$$\text{the total cost } C \text{ is additive } \quad C(u, v) := \sum_{i=1}^p c(m_i)$$

Here, step-cost  $c$  of moderate growth, i.e.,  $c(m) = O(\log m)$

### Main costs of moderate growth.

- if  $c \equiv 1$ , then  $C = P$  is the number of iterations
- if  $c = c_m$  characteristic fn of a given digit  $m$ , then  $C$  is the number of occurrences of  $m$  in the CF.
- if  $c = \ell(m)$ , the binary length of digit  $m$ , then  $C$  is the encoding length of the CF.

**Important Question:** Compare the behaviour of these various

Euclidean algorithms with respect to different costs.

## Previous results on the Average-Case Analysis

Set of possible inputs  $\Omega_N := \{(u, v); \gcd(u, v) = 1, 0 \leq \frac{u}{v} \leq 1, v \leq N\}$ .

First results obtained only for  $C = P$  and for particular algorithms,

Due to Heilbronn, Dixon, Rieger (70), for Standard, Centered Alg.

Heuristic results by Brent (78) for the Binary Alg.

Then a Complete Classification into two classes [Va 1998].

Fast Class = {Standard, Centered, Odd, Binary}  $E_N[P] = A \log N$

Slow Class = {By-Excess, Even, Subtractive}  $E_N[P] = B \log^2 N$

And an analysis of a broad class of costs

[Not only additive costs relative to step-costs of moderate growth],

amongst them: the Bit-Complexity [Akhavi, Va, 2000]

*Instances of a Dynamical Analysis=*

*Analysis of Algorithms + Dynamical Systems*

**Here: Distributional analysis of cost  $C$  on  $\Omega_N$**   
 related to a step-cost  $c$  of  $\mathcal{MG}$  for three Algorithms of the Fast Class

**Main result : The cost  $C$  is asymptotically Gaussian**

First a **CLT Theorem**:

$$\mathbb{P}_N \left[ \frac{C(u, v) - \mu(c) \log N}{\delta(c) \sqrt{\log N}} \leq x \right] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{t^2/2} dt + O\left(\frac{1}{\sqrt{\log N}}\right)$$

Also a **LIT Theorem**

for cost  $C$  with a lattice step-cost  $c$  [  $\text{Im}(c) \in LN$  with  $L > 0$  ],

$$\mathbb{P}_N [C(u, v) \sim \mu(c) \log N + x \delta(c) \sqrt{\log N}] = \frac{e^{-x^2/2}}{\delta(c) \sqrt{2\pi \log N}} + O\left(\frac{1}{\log N}\right)$$

**Optimal speed of convergence** in both cases (LIT and CLT)

A **major improvement of previous results** due to Hensley (94): our proof is **more natural, our result is more general and more precise.**

## Expressions of constants $\mu(c)$ and $\delta(c)$ as mathematical functions

Central rôle played by the Pressure Fonction  $\Lambda(s, w) := \log \lambda(s, w)$ , where  $\lambda(s, w)$  is the **dominant eigenvalue** of a **weighted transfer operator**  $\mathbf{H}_{s, w}$  associated to the **Euclidean Dynamical System**.

Constants  $\mu(c)$  and  $\delta(c)$  are expressed with the first five partial derivatives of  $(s, w) \longrightarrow \Lambda(s, w)$  at  $(s, w) = (1, 0)$ .

### Five main tools involved in the proofs

- The **dynamical system** and its **weighted transfer operator**  $\mathbf{H}_{s, w}$
- The **Quasi-Powers Theorem** on the moment generating function
- **Perron's formula**
- **Dolgopyat's results**
- An intermediary probabilistic model, called a **smoothed model**

## The Euclidean dynamical System (I).

The trace of the execution of the Euclid Algorithm on  $(v_1, v_0)$  is:

$$(v_1, v_0) \rightarrow (v_2, v_1) \rightarrow (v_3, v_2) \rightarrow \dots \rightarrow (v_{p-1}, v_p) \rightarrow (v_{p+1}, v_p) = (0, v_p)$$

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

The division  $v_{i-1} = m_i v_i + v_{i+1}$  is then written as

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

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

An execution of the Euclidean Algorithm

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



## The Euclidean dynamical System (II).

A dynamical system with a denumerable system of branches  $(T_{[m]})_{m \geq 1}$ ,

$$T_{[m]} : ]\frac{1}{m+1}, \frac{1}{m}[ \longrightarrow ]0, 1[, \quad T_{[m]}(x) := \frac{1}{x} - m$$

The set  $\mathcal{H}$  of the inverse branches of  $T$  is

$$\mathcal{H} := \left\{ h_{[m]} : ]0, 1[ \longrightarrow ]\frac{1}{m+1}, \frac{1}{m}[; \quad h_{[m]}(x) := \frac{1}{m+x} \right\}$$

The set  $\mathcal{H}$  builds **one step** of the CF's.

The set  $\mathcal{H}^n$  is the set of the **inverse branches of  $T^n$** ;

it builds CF's of **depth  $n$** .

The set  $\mathcal{H}^* := \bigcup \mathcal{H}^n$  builds **all the** (finite) CF's.

**The density transformer**  $\mathbf{H}$  expresses the new density  $f_1$  as a function of the old density  $f_0$ , as  $f_1 = \mathbf{H}[f_0]$ . It involves the set  $\mathcal{H}$

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

With a cost  $c : \mathcal{H} \rightarrow \mathbf{R}^+$  defined by  $c(h_{[m]}) := c(m)$ , it extends to the **weighted transfer operator**  $\mathbf{H}_{s,w}$

$$\mathbf{H}_{s,w}[f](x) := \sum_{h \in \mathcal{H}} \exp[wc(h)] \cdot |h'(x)|^s \cdot f \circ h(x)$$

$\left\{ \begin{array}{l} \text{Multiplicative properties of the derivative} \\ \text{Additive properties of the cost} \end{array} \right\} \implies$

$$\mathbf{H}_{s,w}^n[f](x) := \sum_{h \in \mathcal{H}^n} \exp[wc(h)] \cdot |h'(x)|^s \cdot f \circ h(x)$$

The  $n$ -th iterate of  $\mathbf{H}_{s,w}$  generates the CFs of depth  $n$ . The **quasi inverse**  $(I - \mathbf{H}_{s,w})^{-1} = \sum_{n \geq 0} \mathbf{H}_{s,w}^n$  generates **all the finite** CFs.

## Other Euclidean Dynamical Systems.

A **continuous** dynamical system can be associated to each **discrete** division: Replace the rational  $u/v$  by a generic real  $x$  of  $\mathcal{I}$ .

The DS relative to a “true” division is **deterministic**.

The DS relative to a pseudo-division is **random**: The 2-adic valuation  $b$  becomes a random variable  $B$  with  $\mathbb{P}[B = b] = 1/2^b$  for  $b \geq 1$ .

**Key Property : Expansiveness of branches**

$$|T'(x)| \geq \rho > 1 \text{ for all } x \text{ in } \mathcal{I}$$

When **true**, this implies a **chaotic** behaviour for trajectories and good properties for the density transformer when it acts on  $\mathcal{C}^1(\mathcal{I})$ .

The associated algorithms are **Fast** and belong to the **Good Class**

When this condition is **violated at only one fixed point**, this leads to **intermittency phenomena**. The associated algorithms are **Slow**.

**Main Analytical Properties of  $\mathbf{H}_{s,w}$**  for an algorithm of the **Good Class** and a digit-cost  $c$  of **moderate growth**.

$\mathbf{H}_{s,w}$  acts on  $\mathcal{C}^1(\mathcal{I})$ ;

The map  $(s, w) \mapsto \mathbf{H}_{s,w}$  is **analytic** near the reference point  $(1, 0)$

For  $s$  and  $w$  real :

**Property *UDE*** : **Unique dominant eigenvalue**  $\lambda(s, w)$ ,

**Property *SG*** : Existence of a **spectral gap**.

With perturbation theory, these properties remain true when  $(s, w)$  is near  $(1, 0)$ ,  $\lambda(s, w)$  is **analytic w.r.t.  $s$  and  $w$** .

**A spectral decomposition**  $\mathbf{H}_{s,w} = \lambda(s, w) \cdot \mathbf{P}_{s,w} + \mathbf{N}_{s,w}$ .

$\mathbf{P}_{s,w}$  is the projector on the dominant eigensubspace.

$\mathbf{N}_{s,w}$  is the operator relative to the remainder of the spectrum, whose spectral radius  $\rho_{s,w}$  satisfies  $\rho_{s,w} \leq \theta \lambda(s, w)$  with  $\theta < 1$ .

**.....which extends to all  $n \geq 1$ ,  $\mathbf{H}_{s,w}^n = \lambda^n(s, w) \cdot \mathbf{P}_{s,w} + \mathbf{N}_{s,w}^n$ .**

Then, if  $\int_I f(t)dt > 0$ , a Quasi-Power-property

$$\mathbf{H}_{s,w}^n [f] = \lambda^n(s, w) \cdot \mathbf{P}_{s,w} [f] \cdot [1 + \mathcal{O}(\theta^n)]$$

and, a decomposition for the quasi-inverse

$$(I - \mathbf{H}_{s,w})^{-1} = \lambda(s, w) \frac{\mathbf{P}_{s,w}}{1 - \lambda(s, w)} + (I - \mathbf{N}_{s,w})^{-1}$$

Since  $\mathbf{H}_{1,0}$  is a density transformer, one has  $\lambda(1, 0) = 1$ .

“Dominant” (polar) singularities of  $(I - \mathbf{H}_{s,w})^{-1}$  near the point  $(1, 0)$ :  
along a curve  $s = \sigma(w)$  on which the dominant eigenvalue satisfies

$$\lambda(\sigma(w), w) = 1$$

## How to prove an asymptotic gaussian law?

With the moment generating fn  $\mathbb{E}_N[\exp(wC_N)]$  of cost  $C_N := C|\Omega_N$ .

**Quasi-Powers Theorem.** If  $\mathbb{E}_N[\exp(wC_N)]$  is a uniform quasi-power when  $w$  is near 0, then  $C_N$  is asymptotically gaussian on  $\Omega_N$ .

If  $\mathbb{E}_N[\exp(wC_N)] = \exp[\beta_N U(w) + V(w)] \cdot \left[ 1 + O\left(\frac{1}{\kappa_N}\right) \right]$

with a  $O$ -term uniform when  $w$  is near 0,

$U, V$  analytic,  $U''(0) \neq 0$ , and  $\beta_N, \kappa_N \rightarrow \infty$ ,

Then: (i)  $\frac{C_N - U'(0) \cdot \beta_N}{\sqrt{U'''(0)\beta_N}}$  is asymptotically Gaussian,

with a speed of convergence  $O(\kappa_N^{-1} + \beta_N^{-1/2})$

(ii) Precise estimates hold

for the expectation  $\mathbb{E}_N[C_N]$  and the variance  $\mathbb{V}_N[C_N]$

$$\mathbb{E}_N[C_N] = \beta_N U'(0) + V'(0) + \mathcal{O}(\kappa_N^{-1}),$$

$$\mathbb{V}_N[C_N] = \beta_N U''(0) + V''(0) + \mathcal{O}(\kappa_N^{-1}).$$

(iii) and for all moments of order  $k$

$$\mathbb{E}_N[C_N^k] = P_k(\beta_N) + \mathcal{O}\left(\frac{\beta_N^{k-1}}{\kappa_N}\right)$$

with a polynomial  $P_k$  of degree at most  $k$ , with coefficients depending on the derivatives of order at most  $k$  at 0 of  $U$  and  $V$ .

## Distribution of real truncated trajectories. (I) Methods

Endow the interval  $\mathcal{I}$  with density  $f$ , and consider, for any real  $x$ , the cost  $C_n$  relative to the  $n$  first digits

$$C_n(x) := \sum_{i=1}^n c(m_i)$$

Limit distribution of  $C_n$  when  $n \rightarrow \infty$ ?

$$\begin{aligned} \mathbb{E}[\exp(wC_n)] &= \sum_{h \in \mathcal{H}_n} \exp[wc(h)] \cdot \int_{h(\mathcal{I})} f(y) dy, \text{ and, with } y = f(u), \\ &= \int_{\mathcal{I}} \sum_{h \in \mathcal{H}_n} \exp[wc(h)] \cdot |h'(u)| \cdot f \circ h(u) du = \int_{\mathcal{I}} \mathbf{H}_{1,w}^n[f](u) du. \end{aligned}$$

With **UDE + SG**,

$$\mathbb{E}[\exp(wC_n)] = \left( \lambda(1, w)^n \int_{\mathcal{I}} \mathbf{P}_{1,w}[f](u) du \right) (1 + \mathcal{O}(\theta^n)),$$

A uniform quasi power ! with  $U(w) = \Lambda(1, w)$ .



## Distribution of real truncated trajectories. (II) Results

For a triple  $(\mathcal{I}, T, c)$  of  $GMG$  type with non-constant  $c$  and any probability  $\Pr$  on  $\mathcal{I}$  with a  $C^1$  density, there is

**An asymptotic Gaussian law for  $C_n$ :**

$$\mathbb{P} \left[ x \mid \frac{C_n(x) - \hat{\mu}(c)n}{\hat{\delta}(c)\sqrt{n}} \leq Y \right] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^Y e^{-y^2/2} dy + O\left(\frac{1}{\sqrt{n}}\right).$$

with  $\hat{\mu}(c) = \Lambda'_w(1, 0)$ ,  $\hat{\delta}^2(c) = \Lambda''_{w^2}(1, 0)$

[Convexity properties of  $\Lambda$  (w.r.t  $w$ ) prove that  $\Lambda''_{w^2}(1, 0) \neq 0$  for a non-constant cost  $c$ .]

For any  $\theta$  which satisfies  $r_1 < \theta < 1$ , (with  $r_1 =$  the subdominant spectral radius of the density transformer  $\mathbf{H}$ ),

$$\mathbb{E}[C_n] = \hat{\mu}(c) \cdot n + \hat{\eta}(c) + O(\theta^n), \quad \mathbb{V}[C_n] = \hat{\delta}^2(c) \cdot n + \hat{\delta}_1(c) + O(\theta^n),$$

[An easy proof for a quite well-known result.]

**Rational trajectories: The Dirichlet moment generating function (I).**

**Definition.** Replace the sequence of MGF's  $\mathbb{E}_N[\exp(wC)]$  by a unique Dirichlet moment generating function  $S(s, w)$ ;

Two parameters  $s$  and  $w$ :  $s$  marks the size, and  $w$  marks the cost,

$\Omega$  is the set of all the possible inputs,

$$S(s, w) := \sum_{(u,v) \in \Omega} \frac{1}{\eta^s} \exp[wC(u, v)] = \sum_{n \geq 1} \frac{c_n(w)}{\eta^s}$$

$$\text{with } c_n(w) := \sum_{\substack{(u,v) \in \Omega \\ v=n}} \exp[wC(u, v)]$$

The plain moment generating function  $\mathbb{E}_N[\exp(wC)]$  is expressed with coefficients of  $S(s, w)$

$$\mathbb{E}_N[\exp(wC)] = \frac{\sum_{n \leq N} c_n(w)}{\sum_{n \leq N} c_n(0)}$$

The Dirichlet moment generating function (II).

Link with the transfer operator.

The Euclid Algorithm builds a bijection between  $\Omega$  and  $\mathcal{H}^*$ :

$$(u, v) \mapsto h \quad \text{with} \quad \frac{u}{v} = h(0).$$

$$\text{Then,} \quad \frac{1}{v} = \frac{1}{D[h](0)} = |h'(0)|^{1/2}, \quad C(u, v) = c(h),$$

and the Dirichlet series

$$S(s, w) := \sum_{(u,v) \in \Omega} \frac{1}{v^s} \exp[wC(u, v)] = \sum_{h \in \mathcal{H}^*} |h'(0)|^{s/2} \exp[wc(h)]$$

admits an **alternative expression** with

the quasi inverse  $(I - \mathbf{H}_{s,w})^{-1}$  of the weighted transfer operator  $\mathbf{H}_{s,w}$ ,

$$S(2s, w) = (I - \mathbf{H}_{s,w})^{-1}[\mathbf{1}](0)$$

## Study of the moments $\mathbb{E}_N[C^k]$ . (I) Methods.

Uses the  $k$ -th derivative of  $S(s, w)$  (with respect to  $w$ , at  $w = 0$ )

= A Dirichlet series  $G_k(s)$  which involves  $k$  occurrences of  $(I - \mathbf{H}_s)^{-1}$

**Extraction of coefficients via Tauberian Theorems.**

For a Dirichlet series  $G(s) := \sum_{n \geq 1} a_n n^{-s}$ , with  $a_n \geq 0$

Tauberian Theorem provide estimates for the sums  $\sum_{n \leq N} a_n$   
(but without remainder terms)

Which **properties** of  $\mathbf{H}_s$  are used for applying Tauberian Theorems?

**UDE + SG + Aperiodicity Condition:**

$1 \notin \text{Sp } \mathbf{H}_s$  on the vertical line  $\Re s = 1, s \neq 1$

**Tauberian Theorem.** [Delange] Suppose that a Dirichlet series

$$G(s) := \sum_{n \geq 1} a_n n^{-s}, \text{ with } a_n \geq 0 \text{ converges for } \Re(s) > \sigma > 0.$$

Assume that

- (i)  $G(s)$  is analytic on  $\Re(s) = \sigma, s \neq \sigma$ ,
- (ii) For some  $\gamma > 0$ , when  $s$  is near  $\sigma$ ,

$$G(s) = \frac{A(s)}{(s - \sigma)^{\gamma+1}} + C(s)$$

with  $A, C$  analytic at  $s = \sigma$ , and  $A(\sigma) \neq 0$

Then:

$$\sum_{n \leq N} a_n = \frac{A(\sigma)}{\sigma \Gamma(\gamma + 1)} \cdot N^\sigma \cdot \log^\gamma N \cdot [1 + \epsilon(N)], \quad \epsilon(N) \rightarrow 0.$$

## Study of the moments $\mathbb{E}_N[C^k]$ . (II) Results.

For an algorithm of the Good Class and  $c = 1$ ,  $\mathbb{E}_N[P] \sim \frac{2}{h(S)}$

For an algorithm of the Good Class and a cost  $c$  of moderate growth,

$$\mathbb{E}_N[C^k] \sim (\mathbb{E}_N[C])^k \quad \text{with} \quad \mathbb{E}_N[C] \sim \hat{\mu}(c) \cdot \mathbb{E}_N[P]$$

where  $\hat{\mu}(c)$  = the average of  $c$  along the real trajectories,

= the average of  $c$  with respect to the stationary density.

Two main results.

- Similarity between the behaviour of  $C$  on almost all real trajectories and its average behaviour on rational trajectories.
- The distribution of  $C$  is concentrated around its mean.

## Distribution study: Extraction of coefficients via the Perron Formula:

The Perron Formula of order two,

$$\text{For } F(s) := \sum_{n \geq 1} \frac{a_n}{n^s}, \quad \sum_{n \leq N} \sum_{q \leq n} a_q = \frac{1}{2i\pi} \int_{D-i\infty}^{D+i\infty} F(s) \frac{N^{s+1}}{s(s+1)} ds$$

is a first step for estimating  $E_N[\exp(wC)]$ .. **uniformly in  $w$ .**

Perron's formula relates the MGF  $E_N[\exp(wC)]$  to

$$\begin{aligned} & \frac{1}{2i\pi} \int_{D-i\infty}^{D+i\infty} S(2s, w) \frac{N^{2s+1}}{s(2s+1)} ds \\ &= \frac{1}{2i\pi} \int_{D-i\infty}^{D+i\infty} (I - \mathbf{H}_{s,w})^{-1} [1](0) \frac{N^{2s+1}}{s(2s+1)} ds \end{aligned}$$

What can be expected on  $S(2s, w)$  (closely related to  $(I - \mathbf{H}_{s,w})^{-1}$ ) for dealing with the Perron Formula?

## Property $US$

**Property  $US$ .** There exists a strip  $S := \{s; |\Re(s) - 1| < \alpha\}$  such that, uniformly w.r.t.  $w$  when  $w$  is near 0,

(i) [Strong aperiodicity]  $S(2s, w)$  has a unique pole inside  $S$ ; it is located at  $s = \sigma(w)$

(ii) [Uniform estimates] On the left line  $\Re s = 1 - \alpha$ :

$$S(2s, w) = O(|Ss|^\beta) \text{ with } \beta < 1$$

**Remark.** Property  $US$  is not always true; For instance, Property (i) is false for Dynamical Systems with affine branches.

**Three main facts.**

- (1) There exists a Condition, the **Condition  $UNI$** , that expresses that the dynamical system is quite different from a piecewise affine map.
- (2) The **Condition  $UNI$**  is sufficient to imply the **Property  $US$** .
- (3) The **Condition  $UNI$**  is true in our Euclidean context.



## Condition *UNI*.

With a “distance”  $\Delta$  between two inverse branches  $h$  and  $k$

$$\Delta_{h,h} := \inf_{x \in \mathcal{I}} \Psi'_{h,k}(x), \quad \text{with} \quad \Psi_{h,k}(x) := \log \frac{|h'(x)|}{|k'(x)|},$$

Condition *UNI* says: The inverse branches of **same depth** are **not too often too close** w.r.t.  $\Delta$ .

Condition *UNI* is **never true** for D S with **affine branches** ( $\Delta \equiv 0$ ), but the *UNI* Condition is **true** in our **Euclidean** context. (Item 3)

Dolgopyat (98) proves the Item 2 but **only for**

- Dynamical Systems with a **finite** number of branches
- **Plain** transfer operators (not weighted)

We adapt his arguments to **generalize** this result to our framework and prove (2).

**Coming back to the proof of the asymptotic gaussian law .**

**Step 1.** Introduce the smoothed model  $\bar{\Omega}_N$ . Choose first  $N$ ; then draw uniformly  $q$  between  $N$  and  $\lfloor N - N^{1-\gamma} \rfloor$ ; finally draw uniformly  $(u, v) \in \Omega_q$

The Perron Formula with the *US* Property entail a **uniform quasi-power behaviour** for the MGF of the smoothed version of cost  $C$ ,

$$\mathbb{E}_N[\exp(wC)] = (1 + O(N^{-\gamma})) \exp(2[\sigma(w) - \sigma(0)] \log N + A(w))$$

with a  $O$ -term uniform in  $w$ .

**Step 2.** The Quasi-Power Theorem proves: the Cost  $C_N$  follows asymptotically a **Gaussian Law in the smoothed model**.

**Step 3.** The two distributions of  $C$  [on  $\Omega_N$  and on  $\bar{\Omega}_N$ ] are  $O(N^{-\gamma})$ -close and, finally, the Cost  $C$  follows asymptotically a **Gaussian Law in the plain model**.

## Asymptotic Gaussian Law: the Central Limit Theorem

$$\mathbb{P}_N \left[ \frac{C(u, v) - \mu(c) \log N}{\delta(c) \sqrt{\log N}} \leq x \right] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-t^2/2} dt + O\left(\frac{1}{\sqrt{\log N}}\right)$$

The constants  $\mu(c)$  and  $\delta(c)$  are expressed with the first and second derivatives of the function  $w \mapsto \sigma(w)$  defined by  $\Lambda(\sigma(w), w) = 0$ .

$$\mu(c) = 2\sigma'(0) = \frac{-2\Lambda'_w(1, 0)}{\Lambda'(1)},$$

With  $L(w) := \Lambda(1 + \sigma'(0)w, w)$ ,

$$\delta^2(c) = 2\sigma''(0) = \frac{2}{|\Lambda'(1)|} L''(0).$$

The **strict positivity** of  $L''(0)$  is related to the **UNI Property**.

**Computation of the constants.** Example of the Standard Case.

**Mean constants** [related to the **first derivatives** of  $\Lambda(s, w)$ ] admit alternative expressions which involve the **stationary density**  $f_1$ .

The entropy of the system  $h(\mathcal{S}) = \Lambda'_s(1, 0) = \int_{\mathcal{I}} \log |T'(x)| \cdot f_1(x) dx$

The constants  $\hat{\mu}(c) = \Lambda'_w(1, 0) = \sum_{h \in \mathcal{H}} c(h) \cdot \int_{h(\mathcal{I})} f_1(t) dt$

Since  $f_1(x) = \frac{1}{\log 2} \frac{1}{1+x}$ , the entropy  $h(\mathcal{S}) = \frac{\pi^2}{6 \log 2}$ ,

$\hat{\mu}(c_m) = \frac{1}{\log 2} \log \left( 1 + \frac{1}{m(m+2)} \right)$  [ $c_m =$  characteristic fn of digit  $m$ ]

$\hat{\mu}(\ell) = \frac{1}{\log 2} \log \prod_{k=0}^{\infty} \left( 1 + \frac{1}{2^k} \right)$  [ $\ell =$  the binary length of the digit]

Not such explicit expressions for **variance constants**  $\delta(c)$ ; however, they are **polynomial-time computable** [Lhote].

## Asymptotic Gaussian Law: the Local Limit Theorem for lattice costs

It is sufficient to consider **integer** costs in the **smoothed** model.

$$\bar{\mathbb{E}}_N[e^{i\tau C}] = \sum_{\ell \geq 0} \bar{\mathbb{P}}_N[C = \ell] e^{i\tau \ell} \implies \bar{\mathbb{P}}_N[C = \ell] = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-i\tau \ell} \cdot \bar{\mathbb{E}}_N[e^{i\tau C}] d\tau.$$

LLT Study  $\implies \ell$  near  $q_x(n) := \lfloor \mu(c)n + \delta(c)x\sqrt{n} \rfloor$ , with  $n := \log N$ .

$$I_n := 2\pi \sqrt{\log N} \cdot \bar{\mathbb{P}}_N[C = q_x(\log N)] = \sqrt{n} \int_{-\pi}^{+\pi} \exp[-i\tau q_x(n)] \cdot \bar{\mathbb{E}}_N[e^{i\tau C}] d\tau.$$

Decompose  $[-\pi, +\pi]$  into  $[-\nu, \nu]$  and its complement, so that

$$I_n = I_n^{(0)} + I_n^{(1)}.$$

For  $I_n^{(0)}$ , with the **saddle-point method**,  $I_n^{(0)} = \sqrt{2\pi} \frac{e^{-x^2/2}}{\delta(c)} + O\left(\frac{1}{\sqrt{n}}\right)$ .

For  $I_n^{(1)}$ , with the **UMI Property for lattice costs**,

$$|\bar{\mathbb{E}}_N[\exp(i\tau C)]| \leq QN^{-\gamma_0}, \forall |\tau| \in [\nu, \pi] \quad I_n^{(1)} = O(e^{-n^\gamma}).$$

## Conclusion

An instance of a **Dynamical Analysis**,

Only previously used for Average–Case Analyses,

Here used for a **Distributional Analysis**.

## Open problems

Study of **other algorithms**,

**Fast** ones (for instance the Binary Algorithm?) or **Slow** ones?

Study of **other costs**,

with **non moderate** growth?

**non additive** [for instance the bit–complexity?]