Weak convergence of rescaled discrete objects in combinatorics

Jean-François Marckert (LaBRI - Bordeaux)

— ○ — ○ — ○ — ○ — ○ — ○ — ○ — ○ —

O. What are we talking about? - Pictures

I. Random variables, distributions. Characterization, convergence.

II. Convergence of rescaled paths

- Weak convergence in $C[0,1]$. Definition / Characterization.
- Example: Convergence to the Brownian processes.
- byproducts?!

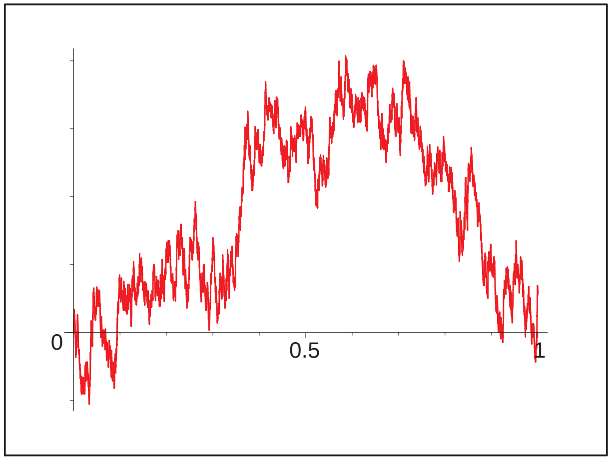
III. Convergence of trees... Convergence to continuum random trees

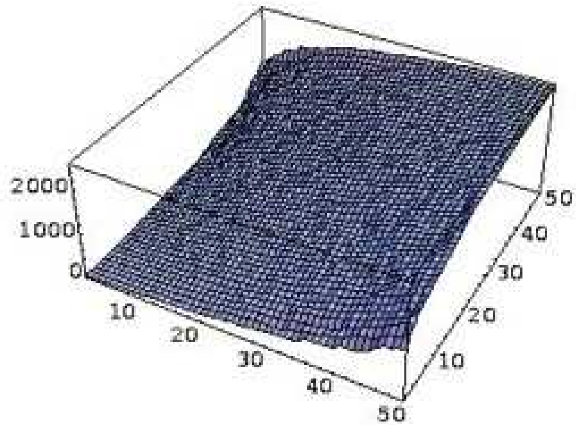
- Convergence of rescaled planar trees
- The Gromov-Hausdorff topology
- Convergences to continuum trees + Examples.

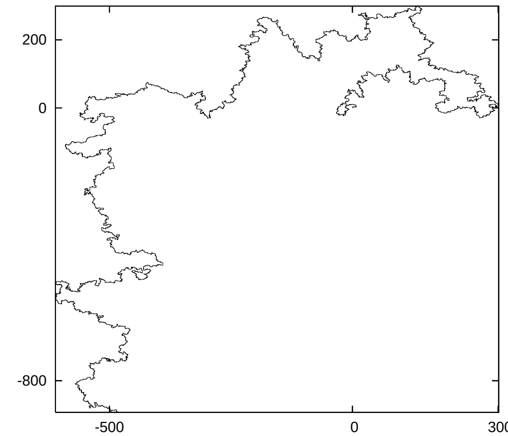
IV. Other examples!

Maresias, AofA 2008.

O. What are we talking about? - Pictures

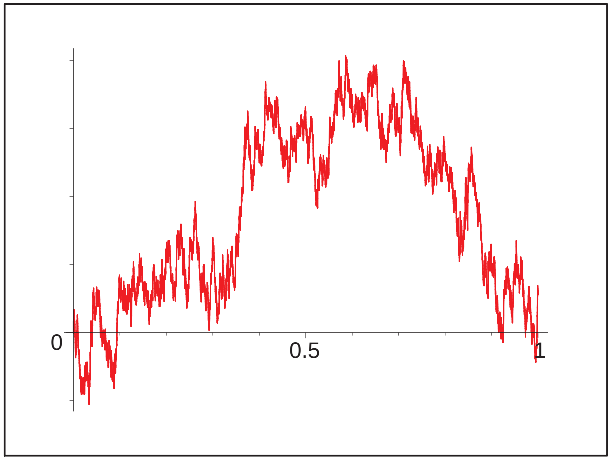


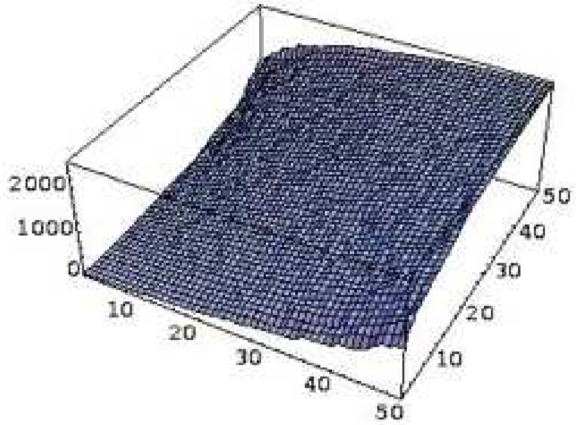


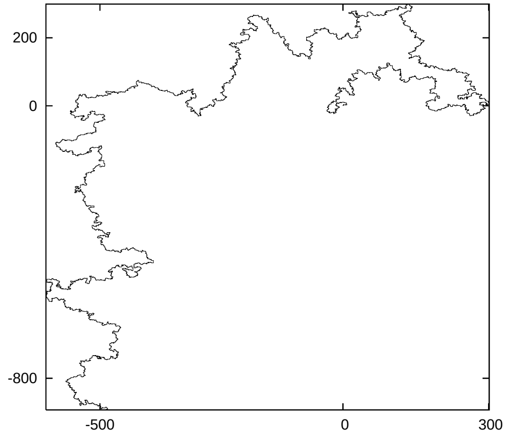


The talk deals with these situations when simulating random combinatorial objects with size $10^3, 10^6, 10^9$ in a window of fixed size, one sees essentially the same picture

O. What are we talking about? - Pictures







Questions :

- What sense can we give to this:

– a sequence of (normalized) combinatorial structures converges?

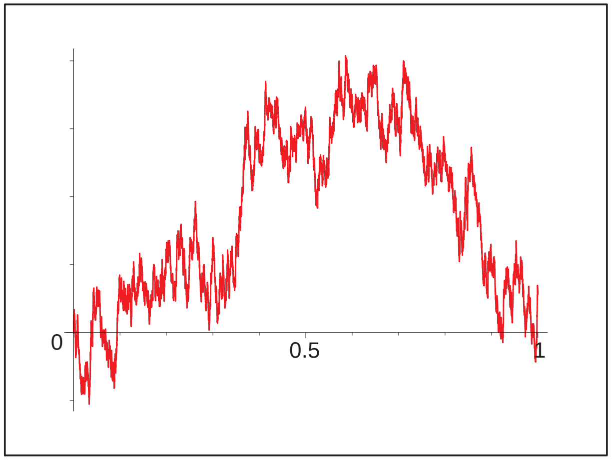
– a sequence of random normalized combinatorial structures converges"?

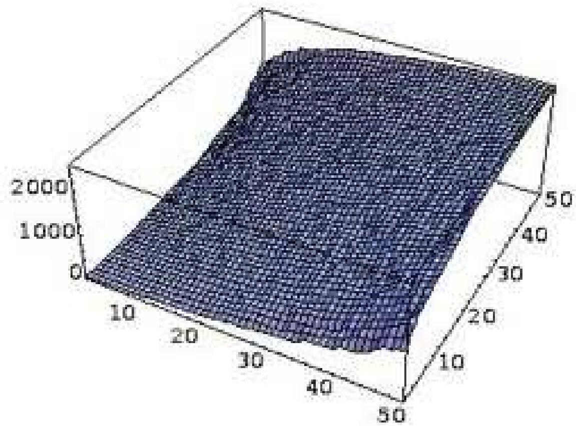
- If we are able to prove such a result...:

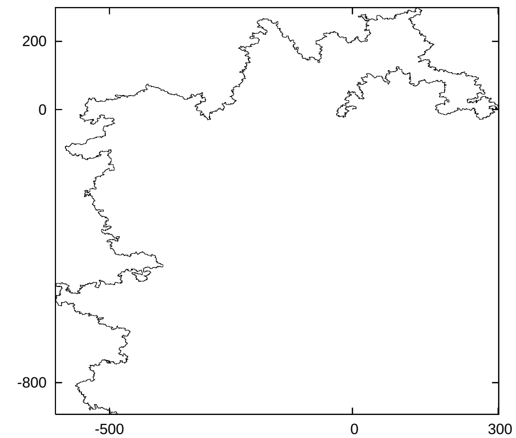
– What can be deduced?

– What cannot be deduced?

O. What are we talking about? - Pictures





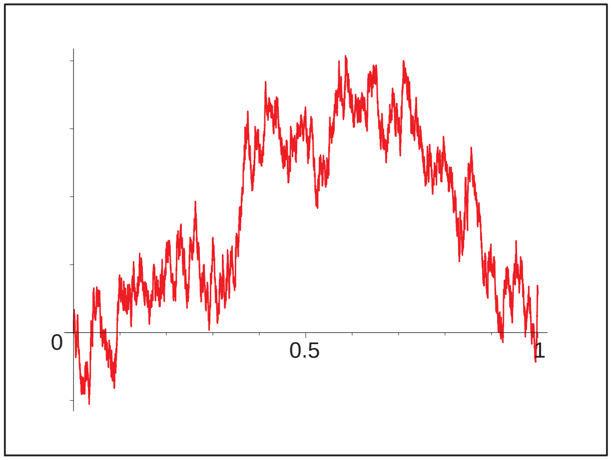


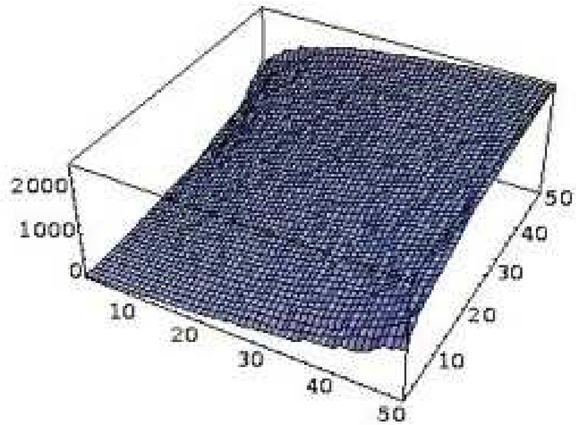
- What sense can we give to this:

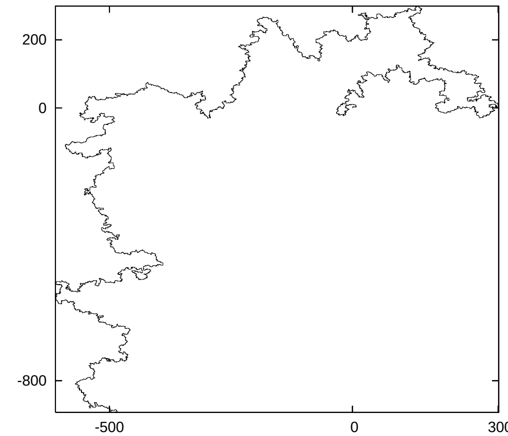
– a sequence of normalized combinatorial structure converges?

answer: this is a question of topology...

O. What are we talking about? - Pictures







- What sense can we give to this:

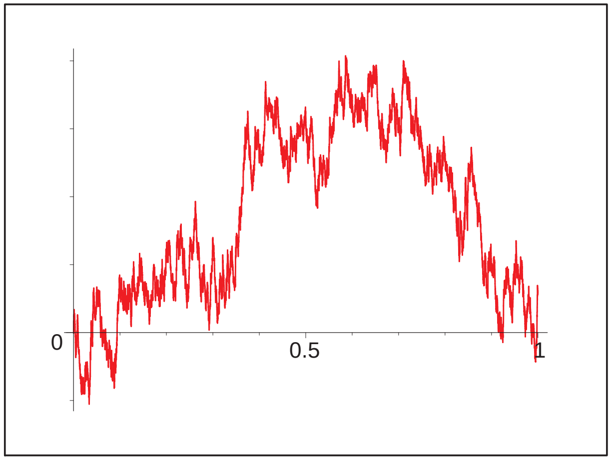
– a sequence of normalized combinatorial structure converges?

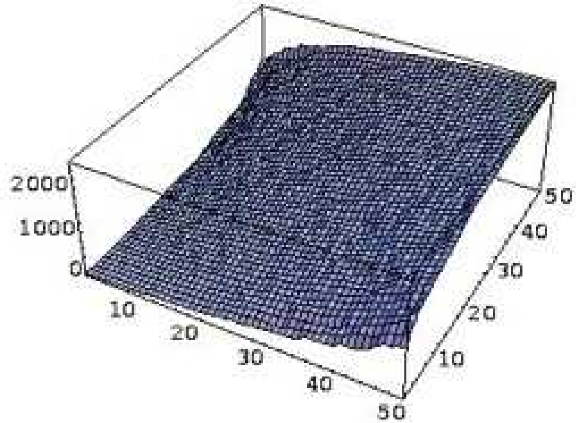
answer: this is a question of topology...

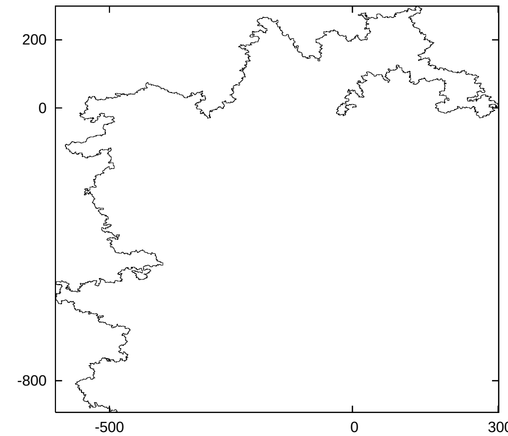
– a sequence of random normalized combinatorial structure converges"?

answer: this is a question of weak convergence associated with the topology.

O. What are we talking about? - Pictures







- What sense can we give to this:

– a sequence of normalized combinatorial structure converges?

answer: this is a question of topology...

– a sequence of random normalized combinatorial structure converges"?

answer: this is a question of weak convergence associated with the topology.

- If we are able to prove such a result...:

What can be deduced?

answer: infinitely many things... but it depends on the topology

What cannot be deduced?

answer: infinitely many things: but it depends on the topology

O. What are we talking about? - Pictures

First - we recall what means convergence in distribution
- in $\mathbb{R}$
- in a Polish space
– Then we treat examples... and see the byproducts

Random variables on $\mathbb{R}$

- A distribution μ on $\mathbb{R}$ is a (positive) measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ with total mass 1.

- a random variable X is a function $X : (\Omega, \mathcal{A}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$, measurable.

- distribution of X: the measure μ,

$$\mu(A) = \mathbb{P}(X \in A) = \mathbb{P}(\{\omega, X(\omega) \in A\}).$$

Characterization of the distributions on $\mathbb{R}$

– the way they integrate some classes of functions

$$f \mapsto \mathbb{E}(f(X)) = \int f(x) d\mu(x),$$

e.g. Continuous bounded functions, Continuous with bounded support

Other characterization: Characteristic function, distribution function $x \mapsto F(x) = \mathbb{P}(X \leq x)$

Convergence in probability

$$X_n \xrightarrow[n]{(proba.)} X \quad \text{if} \quad \forall \varepsilon > 0, \quad \mathbb{P}(|X_n - X| \geq \varepsilon) \underset{n}{\rightarrow} 0.$$

Almost sure convergence

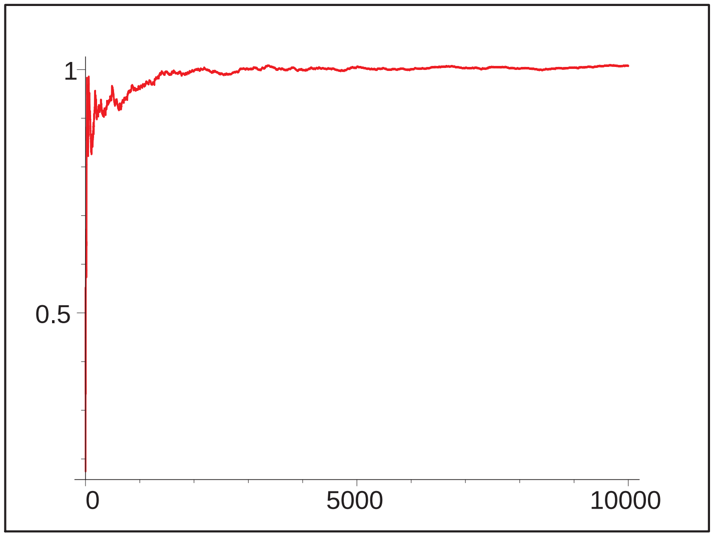
$$X_n \xrightarrow[n]{(as.)} X \quad \text{if} \quad \mathbb{P}(\lim X_n = X) = \mathbb{P}(\{\omega \mid \lim X_n(\omega) = X(\omega)\}) = 1.$$

$X, X_1, X_2, \ldots$ are to be defined on the same probability space Ω:

In these two cases, this is a convergence of RV.

Example: strong law of large number: if Y_i i.i.d. mean m,

$$X_n := \frac{\sum_{i=1}^{n} Y_i}{n} \xrightarrow[n]{(as.)} m$$



Convergence in distribution: DEFINITION:

$$X_n \xrightarrow[n]{(d)} X \ \text{ if } \ \mathbb{E}(f(X_n)) \underset{n}{\to} \mathbb{E}(f(X))$$

for any $f : \mathbb{R} \mapsto \mathbb{R}$ bounded, continuous

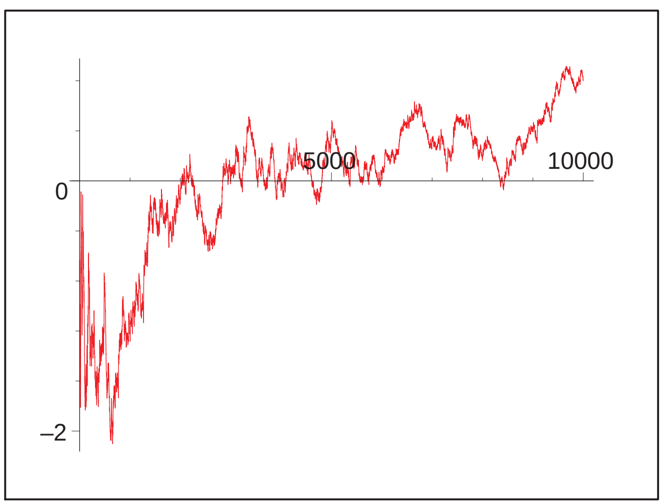
The variables need not to be defined on the same Ω

Other characterizations:

- $F_n(x) = \mathbb{P}(X_n \leq x) \to F(x) = \mathbb{P}(X \leq x)$, for all x where F is continuous
- $\Phi_n(t) = \mathbb{E}(e^{itX_n}) \to \Phi(t) = \mathbb{E}(e^{itX})$, forall t

Example: the central limit theorem: if Y_i i.i.d. mean m, variance $\sigma^2 \in (0, +\infty)$

$$X_n := \frac{\sum_{i=1}^n (Y_i - m)}{\sqrt{n}} \xrightarrow[n]{(d)} \sigma\mathcal{N}(0,1)$$



The sequence (X_n) does not converge!
(Exercice)

Where define (weak) convergence of combinatorial structures?

to define convergence we need a nice topological space:
– to state the convergence.
– this space must contain the (rescaled) discrete objects and all the limits
– this space should give access to weak convergence

Nice topological spaces on which everything works like on $\mathbb{R}$ are Polish spaces.

Where define (weak) convergence of combinatorial structures?

Nice topological spaces on which everything works like on $\mathbb{R}$ are Polish spaces.

Polish space (S, ρ) **:**metric + separable + complete

$\longrightarrow$ open balls, topology, Borelians, may be defined as on $\mathbb{R}$

Examples: – $\mathbb{R}^d$ with the usual distance,

– $(C[0,1], \|.\|_\infty)$, $d(f,g) = \|f-g\|_\infty$

– ... .

Distribution μ on $(S, \mathcal{B}(S))$: measure with total mass 1.

Random variable: $X : (S, \mathcal{B}(S), \mathbb{P}) \longrightarrow (S', \mathcal{B}(S'))$ measurable.

Distribution of X, $\mu(A) = \mathbb{P}(X \in A)$.

Characterization of measures

– The way they integrate continuous bounded functions. $\mathbb{E}(f(X)) = \int f(x) d\mu(x)$.

f **continuous in** x_0 **means:**$\forall \varepsilon > 0, \exists \eta > 0,\ \rho(x, x_0) \le \eta \Rightarrow |f(x) - f(x_0)| \le \varepsilon$.

Random variables on a Polish space

Polish space (S,ρ) **:**metric + separable + complete

Convergence in probab.:

$$\forall\varepsilon>0,\quad \mathbb{P}(\rho(X_n,X)\geq\varepsilon)\underset{n}{\longrightarrow}0.$$

Convergence in distribution

$$\mathbb{E}(f(X_n))\rightarrow\mathbb{E}(f(X)),\qquad \text{for any continuous bounded function } f:S\rightarrow\mathbb{R}$$

$\boxed{\text{Byproduct : if } X_n \xrightarrow[n]{(d)} X \text{ then } f(X_n) \xrightarrow[n]{(d)} f(X) \text{ for any } f:S\rightarrow S' \text{ continuous}}$

Comments

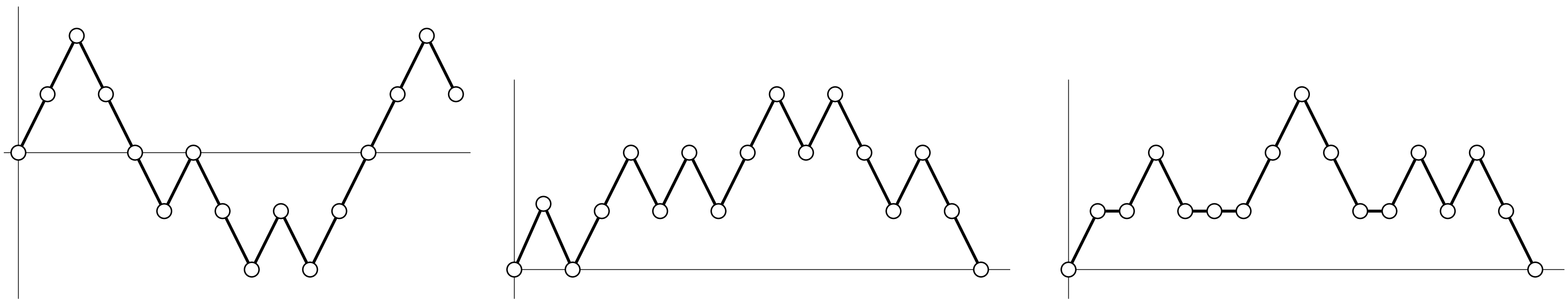
“Are we free to choose the topology we want??”

Yes, but if one takes a ’bad topology’, the convergence will give few informations

II. Convergence of rescaled paths

Paths are fundamental objects in combinatorics.

Walks ± 1, Dyck paths, paths conditioned to stay between some walls, with increments included in $I \subset \mathbb{Z}$.



Convergence of rescaled paths?

In general the only pertinent question is:

does they converge in distribution (after rescaling)?

Here

distribution $=$ distribution on $C[0,1]$ (up to encoding $+$ normalisation).

Here, we choose $C[0,1]$ as Polish space to work in... It is natural, no?

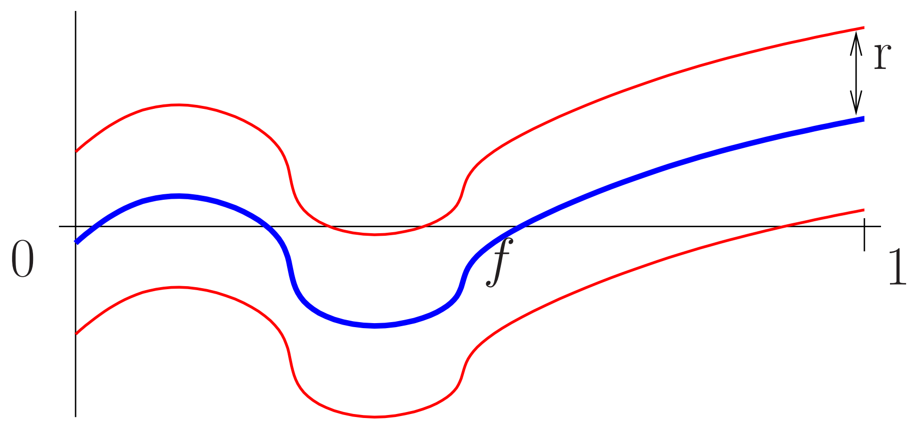
II. Convergence of rescaled paths

How are characterized the distributions on $C[0,1]$?

μ : Distribution on $(C[0,1], \|.\|_\infty)$ (measure on the Borelians of $C[0,1]$):
Let $X = (X_t, t \in [0,1])$ a process, with distribution μ.

Intuition: a distribution μ on $C[0,1]$ gives weight to the Borelians of $C[0,1]$.

The balls $B(f,r) = \{g \mid \|f-g\|_\infty < r\}$.



Proposition 1 *The distribution of X is characterized by the finite dimensional distribution FDD:*

$$i.e.\ \textit{the distribution of}\ (X(t_1),\dots,X(t_k)), \quad k \geq 1, t_1 < \cdots < t_k.$$

II. Convergence of rescaled paths

μ_n : Distribution on $(C[0,1], \|.\|_\infty)$ (measure on the Borelians of $C[0,1]$):

Let $X = (X_t, t \in [0,1])$ a process, with distribution μ.

Proposition 2 *The distribution of X is characterized by the finite dimensional distribution FDD:*

$$i.e.\ the\ distribution\ of\ (X(t_1), \dots, X(t_k)), \quad k \geq 1, t_1 < \cdots < t_k.$$

Example:

– your prefered discrete model of random paths, (rescaled to fit in [0,1].

II. Convergence of rescaled paths

How are characterized the convergence in distributions on $C[0,1]$?:

Main difference with $\mathbb{R}$:
FDD characterizes the measure...
But: convergence of FDD does not characterized the convergence of distribution:
If $(X_n(t_1),\dots,X_n(t_k))) \xrightarrow[n]{(d)} (X(t_1),\dots,X(t_k))$ then we are not sure that $X_n \xrightarrow[n]{(d)} X$ in $C[0,1]$.

if $X_n \xrightarrow[n]{(d)} X$ then $X_n(t) \xrightarrow[n]{(d)} X(t)$ (the function $f \to f(t)$) is continuous). Then if $X_n \xrightarrow[n]{(d)} X$ then the FFD of X_n converges to those of X

A tightness argument is needed
(if you are interested... ask me)

Convergence to Brownian processes

A) $X_1, \ldots, X_n$= i.i.d.random variables. $\mathbb{E}(X_1) = 0$, $\mathrm{Var}(X_i) = \sigma^2 \in (0, +\infty)$.

$$S_k = X_1 + \cdots + X_k$$

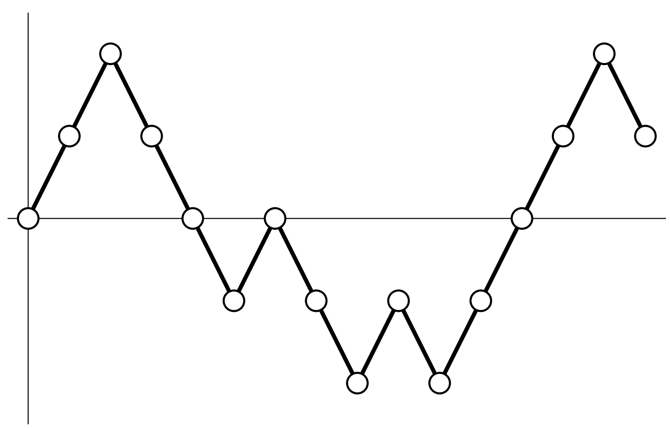
then

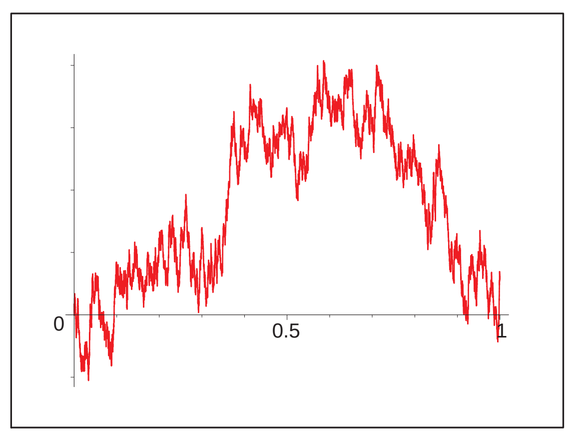
$$(Donsker's\ \ Theorem) \qquad \left(\frac{S_{nt}}{\sqrt{n}}\right)_{t\in[0,1]} \xrightarrow[n]{(d)} (\sigma B_t)_{t\in[0,1]}$$

where $(B_t)_{t\in[0,1]}$ is the **Brownian motion.**

– The BM is a random variable under the limiting distribution: the Wiener measure
The Brownian motion has for FDD: for $0 < t_1 < \cdots < t_k$, $B_{t_1} - B_0, \ldots, B_{t_k} - B_{t_{k-1}}$ are independent, $B_{t_j} - B_{t_{j-1}} \sim \mathcal{N}(0, t_j - t_{j-1})$.

$\left(\frac{S_{nt}}{\sqrt{n}}\right)_{t\in[0,1]}$ **does not converge in probability!**





II. Convergence of rescaled paths

Convergence to Brownian processes

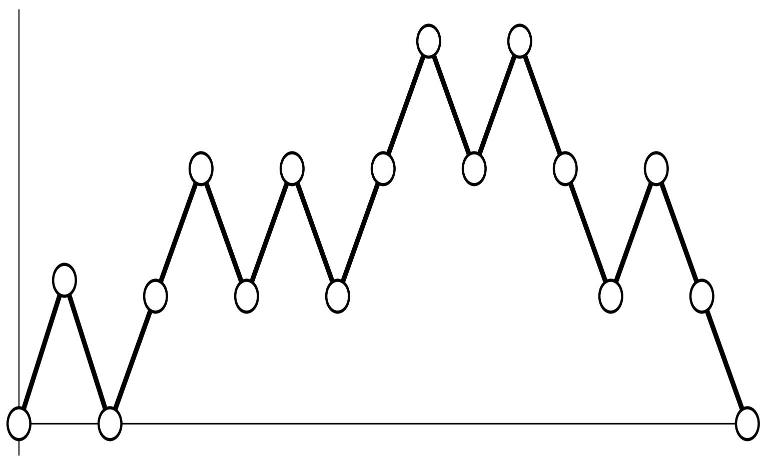
B) $X_1, \ldots, X_n$= i.i.d.random variables. $\mathbb{E}(X_1) = 0$, $\mathrm{Var}(X_i) = \sigma^2 \in (0, +\infty)$, + X_i's lattice support.

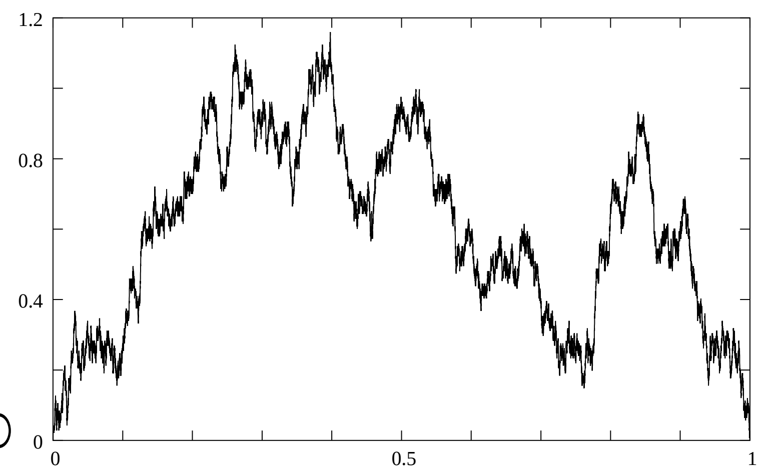
$$S_k = X_1 + \cdots + X_k$$

then

$$(Kaigh's \;\; Theorem) \qquad \left(\frac{S_{nt}}{\sqrt{n}}\right)_{t\in[0,1]} \xrightarrow[n]{(d)} (\sigma \mathbf{e}_t)_{t\in[0,1]}$$

where $(\mathbf{e}_t)_{t\in[0,1]}$ is the **Brownian excursion** .





Similar results capture numerous models of walks appearing in combinatorics

II. Convergence of rescaled paths

Byproducts of $X_n \xrightarrow[n]{(d)} X$ in $C[0,1]$.

1) $\mathbb{E}(f(X_n)) \longrightarrow \mathbb{E}(f(X))$ for any f bounded continuous.

An infinity of byproducts (as much as bounded continuous functions)
$g \mapsto f(g) = \min(1, \int_0^1 g(t)dt)$
$g \mapsto \min(\max g, 1)$

$$\mathbb{E}\left(\int \sin(X_n(t))dt\right) \underset{n}{\to} \mathbb{E}\left(\int \sin(X(t))dt\right)$$

II. Convergence of rescaled paths

Byproducts of $X_n \xrightarrow[n]{(d)} X$ in $C[0,1]$

2) $f(X_n) \xrightarrow[n]{(d)} f(X)$ for any $f : C[0,1] \to S'$ continuous.

An infinity of byproducts (as much as continuous functions onto some Polish space)
$g \mapsto \max(g)$,
$g \mapsto \int_{1/2}^{2/3} g^{13}(t)dt$,
$g \mapsto g(\pi/14)$

$$\left(\max X_n, \int_{1/2}^{2/3} X_n^{13}(t)dt, X_n(\pi/14)\right) \xrightarrow[n]{(d)} \left(\max X, \int_{1/2}^{2/3} X^{13}(t)dt, X(\pi/14)\right)$$

Examples of non-continuous important functions :
$g \mapsto \min \operatorname{argmax}(g)$ (the first place where the max is reached),

$$g \mapsto 1/g(1/3), \quad g \mapsto \int 1/g(s)ds.$$

II. Convergence of rescaled paths

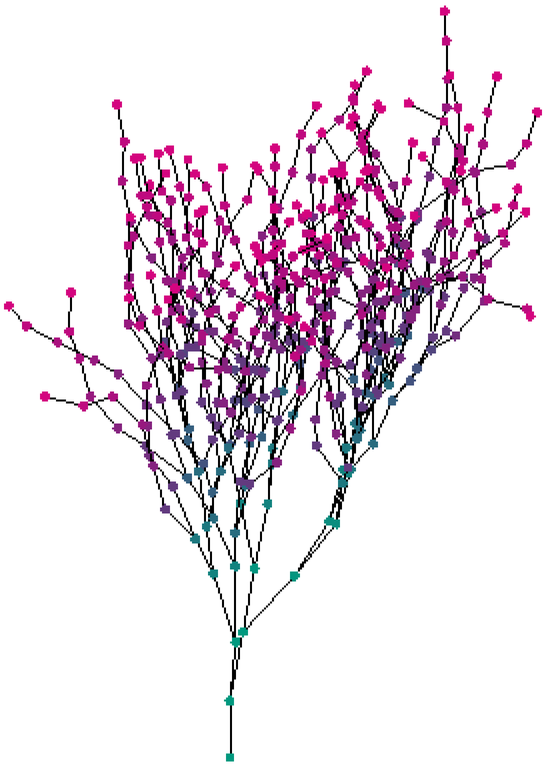
"Reduction of information" at the limit:

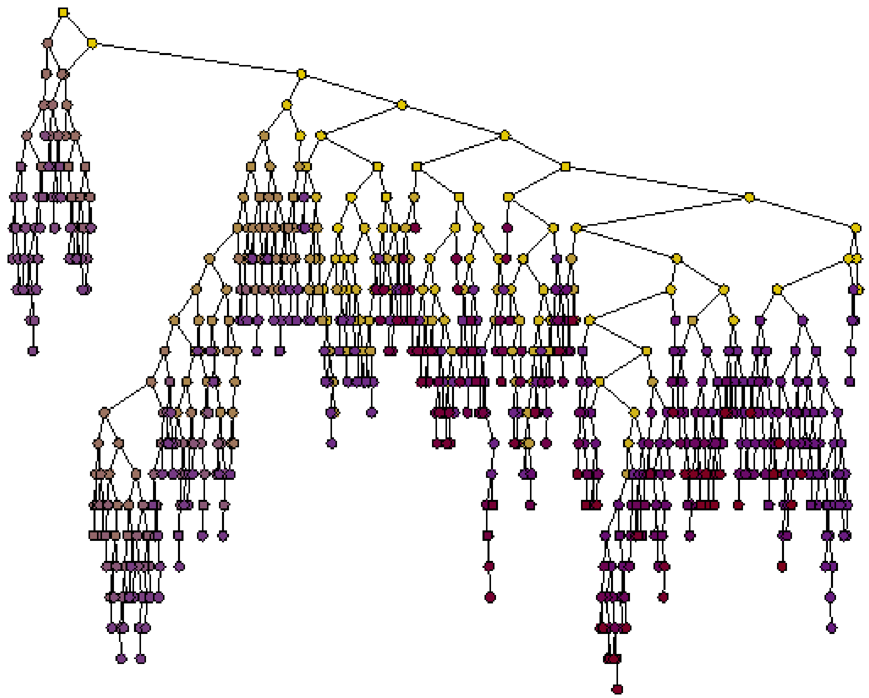
If X_n is a rescaled random discrete object, knowing $X_n \xrightarrow[n]{(d)} X$ in $C[0,1]$ says nothing about any phenomenon which is **not a the right scale.**

Example: Almost surely the Brownian motion reaches is maximum once, traverses the origin an infinite number of times...
This is not the case in the discrete case

III. Convergence of trees... Convergence to continuum random trees

Question: do trees have a limit shape? How can we describe it?
(I am not talking about the profile...)



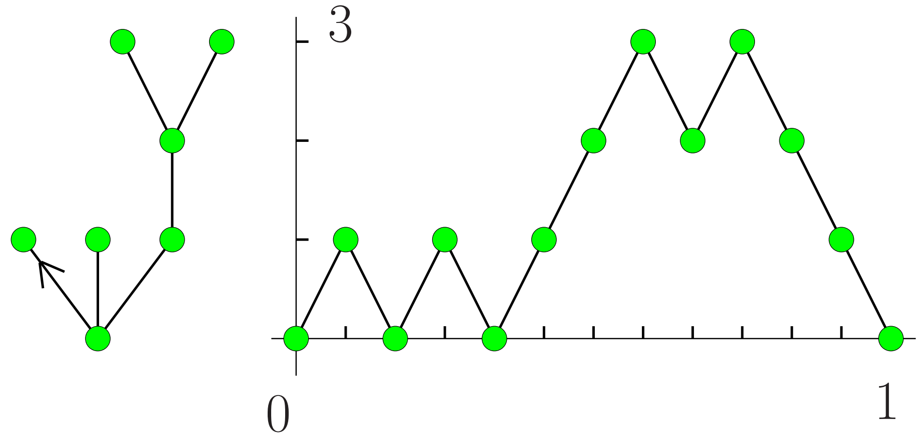


(Luc's trees)

Again...: To prove that rescaled trees converge we search a Polish space containing discrete trees and their limits (continuous trees).

Example of model of random trees: uniform rooted planar tree with n nodes

Trees as element of a Polish space: embedding in $C[0,1]$.



The contour process $(C(k), k = 0, \ldots, 2(n-1))$.

The normalized contour process $\left(\dfrac{C(2(n-1)t)}{\sqrt{n}}\right)_{t\in[0,1]}$.

- This is not the historical path followed by Aldous.
- There exists also some notion of convergence for trees, without normalizations

Notion of real tree

Let $C^+[0,1] = \{f \in C[0,1], f \geq 0, f(0) = f(1) = 0\}$.

With any function $f \in C^+[0,1]$, we associate a tree $A(f)$:

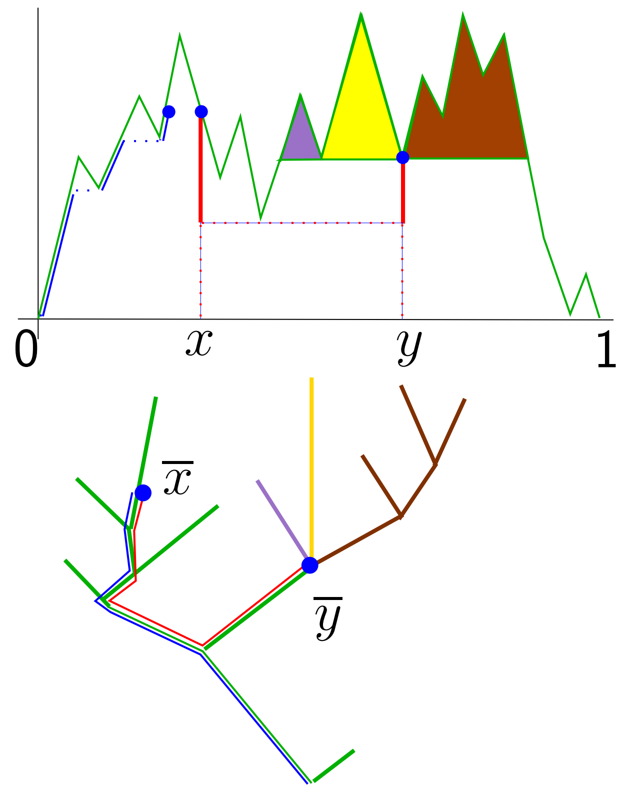
$$A(f) \ := [0,1]/ \underset{f}{\sim} \text{ where}$$

$$x \underset{f}{\sim} y \Longleftrightarrow f(x) = f(y) = \check{f}(x,y) \ := \min_{u \in [x \wedge y, x \vee y]} f(u)$$

$\star$ $A(f)$ equipped with the distance

$$d_f(\overline{x}, \overline{y}) = f(x) + f(y) - 2\check{f}(x,y)$$

is a compact metric space, loop free, connected: it is a tree!



The space $\mathcal{A}$ is equipped with the distance:

$$d(A(f), A(g)) = \|f - g\|_\infty.$$

It is then a Polish space

Convergence of rescaled tree in the space of real trees

Theorem [Aldous: Convergence to the rescaled contour process].

$$\left(\frac{C(2(n-1)t)}{\sqrt{n}}\right)_{t\in[0,1]} \xrightarrow[n]{(d)} \frac{2}{\sigma}\,(\mathsf{e}_t)_{t\in[0,1]}$$

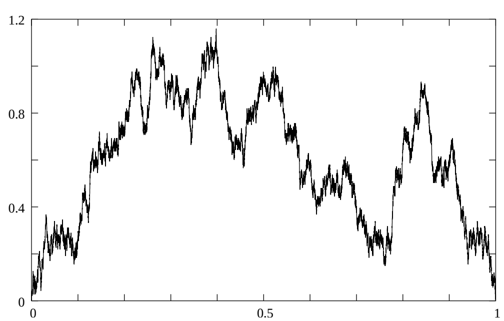
RW: M & Mokkadem, Duquesne.

Result of Aldous valid for critical GW tree conditioned by the size, including Binary tree with n nodes, ...

Theorem [Aldous: convergence of rescaled tree to the Continuum random tree]

$$A\left(\frac{C(2(n-1).)}{\sqrt{n}}\right) \xrightarrow[n]{(d)} A(2\mathsf{e}),$$

in the space of real trees.



This is a convergence (in distribution) of the whole macroscopic structure

Convergence of rescaled tree in the space of real trees

Byproducts of this convergence

- Nice explanation of all phenomenon in $\sqrt{n}$.
- Convergence of the height:

$$H_n/\sqrt{n} \xrightarrow[n]{(d)} \frac{2}{\sigma} \max \mathsf{e}$$

(Found before by Flajolet & Odlyzko (1982) + CV moments)

- Convergence of the total path length PL_n:

$$n^{-3/2} \int_0^1 \frac{C_{2(n-1)t}}{\sqrt{n}} dt \xrightarrow[n]{(d)} \int_0^1 \frac{2}{\sigma} \mathsf{e}(t)\, dt$$

- Convergence of the height of a random node, - convergence of the matrix of the distances $d(U_i, U_j)/\sqrt{n}$ of 12000 random nodes, - joint convergences...

RW: Flajolet, Aldous, Drmota, Gittenberger, Panholzer, Prodinger, Janson, Chassaing, M, ...

But : It does not explain (in general) the phenomenon at a different scale: the continuum random tree is a tree having only binary branching points, degree(root)=1...
One does not see the details of the discrete model on the CRT

Another topology à la mode : the Gromov-Hausdorff topology

The GH topology aims to prove the **convergence of trees or other combinatorials objects** seen **as metric spaces**.

A tree is a metric space, isn't it ?
A connected graph is a metric space, isn't it?
A triangulation is a metric space, isn't it?

The GH distance is a distance on the set of compact metric spaces K.

With this distance, (K, d_{GH}) is a Polish space!!

Another topology à la mode : the Gromov-Hausdorff topology

The idea: up to "isometric relabeling", try to fit as well as possible the two spaces

Hausdorff distance in (E, d_E) = distance between the compact sets of E:

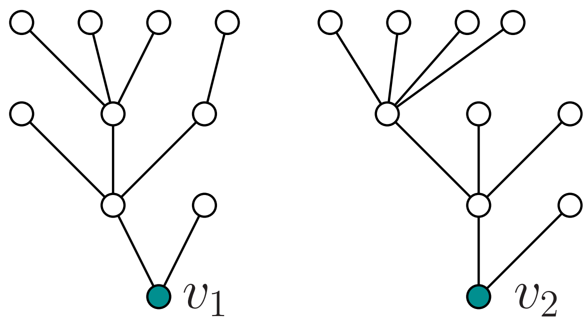
$$d_{Haus(E)}(K_1, K_2) = \inf\{r \mid K_1 \subset K_2^r, K_2 \subset K_1^r\},$$

where $K^r = \cup_{x \in K} B(x, r)$.

Gromov-Hausdorff distance between two compact metric spaces (E_1, d_1) and (E_2, d_2):

$$d_{GH}(E_1, E_2) = \inf d_{Haus(E)}(\phi_1(E_1), \phi_2(E_2))$$

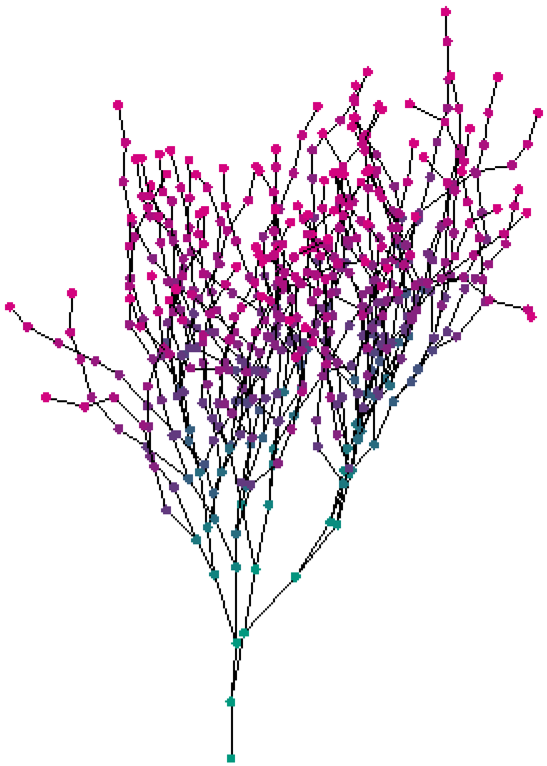
where the infimum is taken on all metric spaces E and all isometric embeddings ϕ_1 and ϕ_2 from (E_1, d_1) and (E_2, d_2) in (E, d_E).

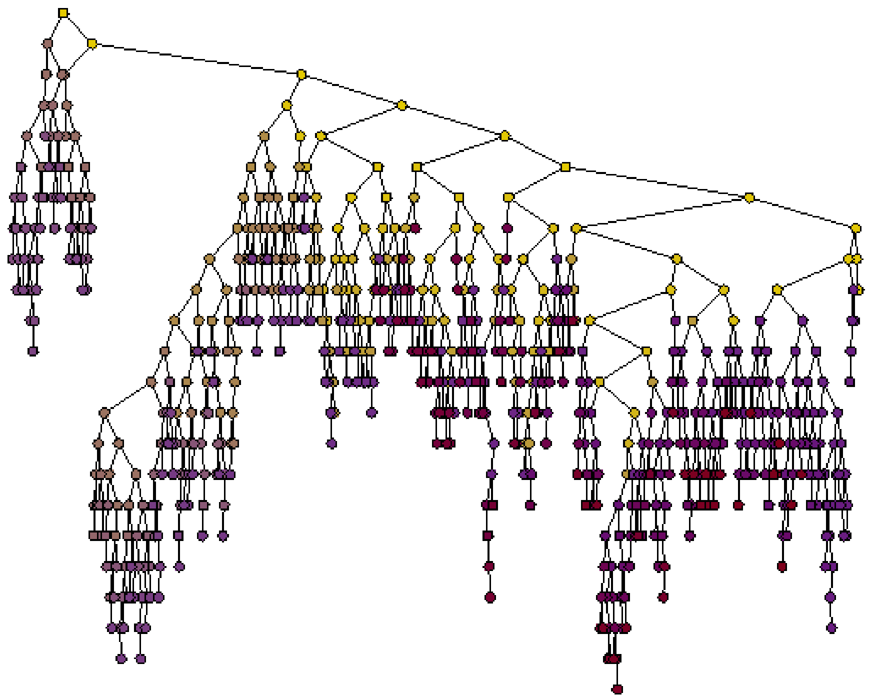


Exercise for everybody (but Bruno):
what if the GH-distance between these trees?

Exercice for Bruno: 2

Gromov Hausdorff distance between





(Luc's trees)

Another topology à la mode : the Gromov-Hausdorff topology

The GH-topology is a quite weak topology, no?

Since normalized planar trees converges to the continuum random tree for the topology of $C[0,1]$...

Theorem Normalized Galton-Watson trees converge to the Continuum random tree for the Gromov-Hausdorff topology.

III. Convergence of trees... Convergence to continuum random trees

Convergence of rooted non-planar binary trees for the GH topology

Non planar binary trees $U(z) = \sum \#\mathcal{U}_n$ with $\mathcal{U}_n$= binary tree with n leaves.

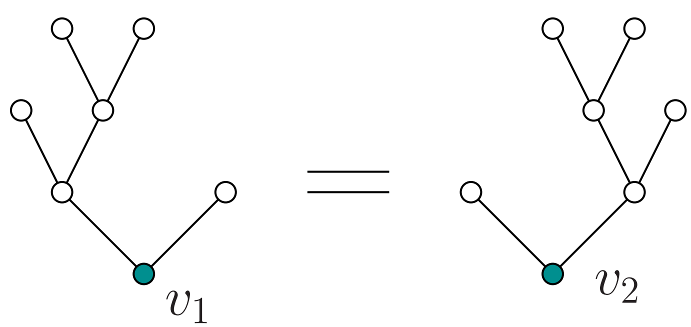
$$U(z) = z + \frac{U(z^2) + U(z)^2}{2}.$$

Let ρ radius of U and

$$\mathbf{c} := \sqrt{2\rho + 2\rho^2 U'(\rho^2)}.$$

Theorem (Work in progress: M & Miermont). Under the uniform distr. on $\mathcal{U}_n$, the metric space $\left(\mathcal{T}_n, \frac{1}{\mathbf{c}\sqrt{n}} d_{\mathcal{T}_n}\right)$ converge in distribution to $(\mathcal{T}_{2\mathrm{e}}, d_{2\mathrm{e}})$ the CRT (encoded by $2\mathrm{e}$) for the GH topology.

Related work: Otter, Drmota, Gittenberger, Broutin & Flajolet



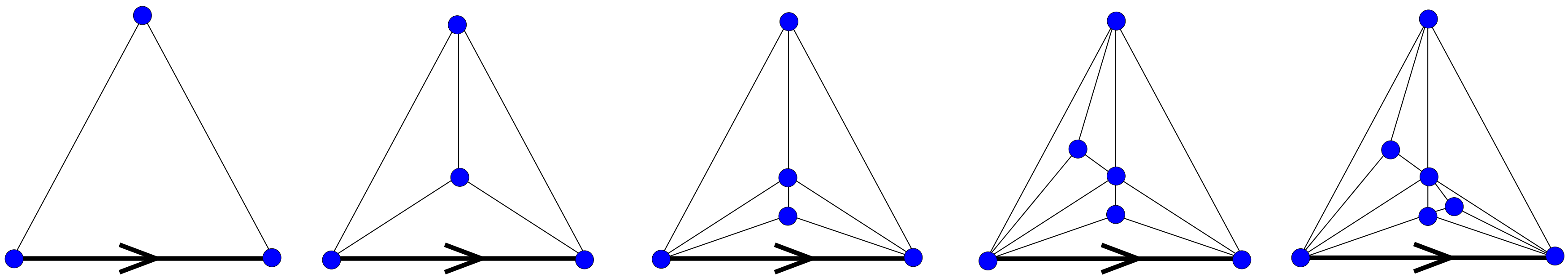
A non-planar-binary tree is a leaf or a multiset of two non-planar-binary trees

Proof absolutely different from the planar case

III. Convergence of trees... Convergence to continuum random trees

Convergence to the CRT for objects that are not trees:

Model of uniform stacked triangulations



M_n= uniform stack-triangulation with $2n$ faces seen as a metric space;

D_{M_n}= graph-distance in M_n

Theorem (Albenque & M)

$$\left(m_n, \frac{D_{m_n}}{\sqrt{6n}/11}\right) \xrightarrow[n]{(d)} (\mathcal{T}_{2\mathrm{e}}, d_{2\mathrm{e}}),$$

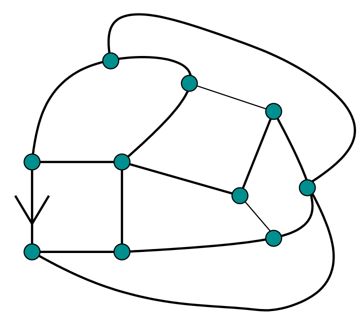
for the Gromov-Hausdorff topology on compact metric spaces.

Related works: Bodini, Darasse, Soria

III. Convergence of trees... Convergence to continuum random trees

The topology of Gromov-Hausdorff

THE important question on maps (say **uniform** triangulations, quadrangulations..)



Seen as metric spaces, do they converges in distribution ?

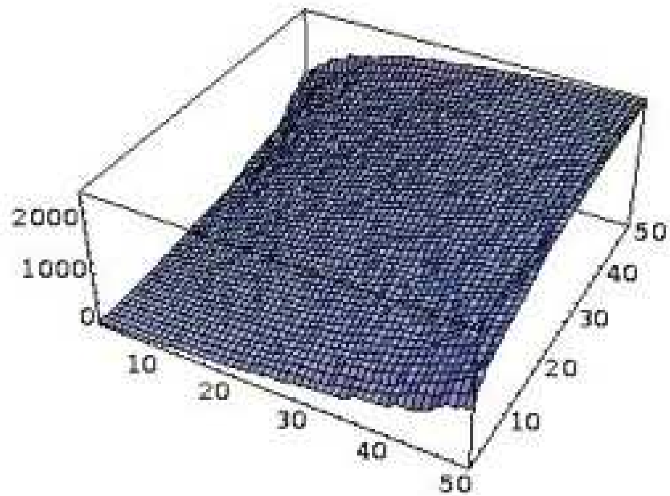
What is known:subsequence converges in distribution to some random metric on the sphere (Le Gall, Miermont) for GH.

Problem:to show uniqueness of the limit

RW: Chassaing-Schaeffer, M-Mokkadem, Miermont, Le Gall...

IV. Other examples!

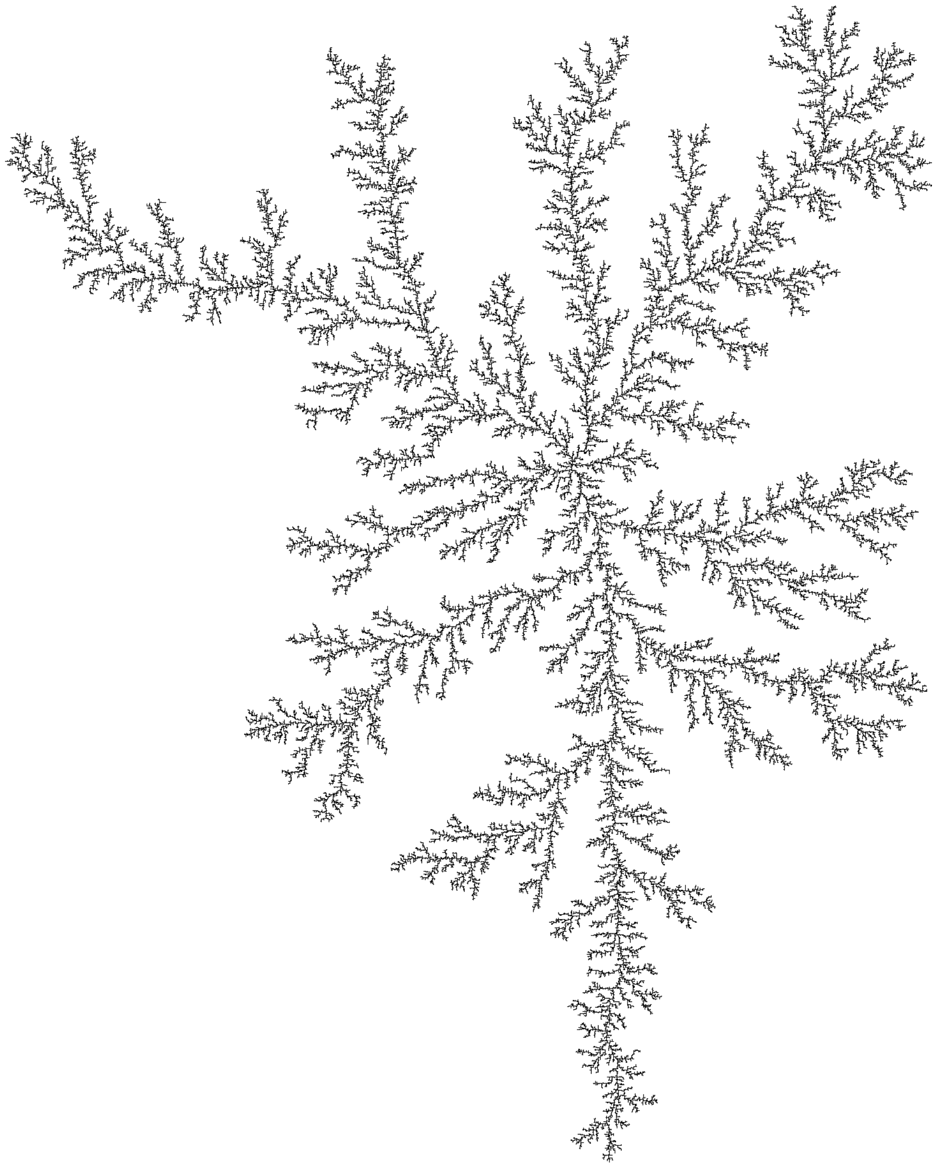
convergence of rescaled combinatorial structures to deterministic limit
Limit shape of a uniform square Young-tableau: Pittel-Romik



source: Dan Romik's page

Convergence for the topology of uniform convergence (functions $[0,1]^2 \mapsto [0,1]$).
Same idea: limit for Ferrer diagram (Pittel)
Limit shape for plane partitions in a box (Cohn, Larsen, Propp); random generation (Bodini, Fusy, Pivoteau)

Unknown limits



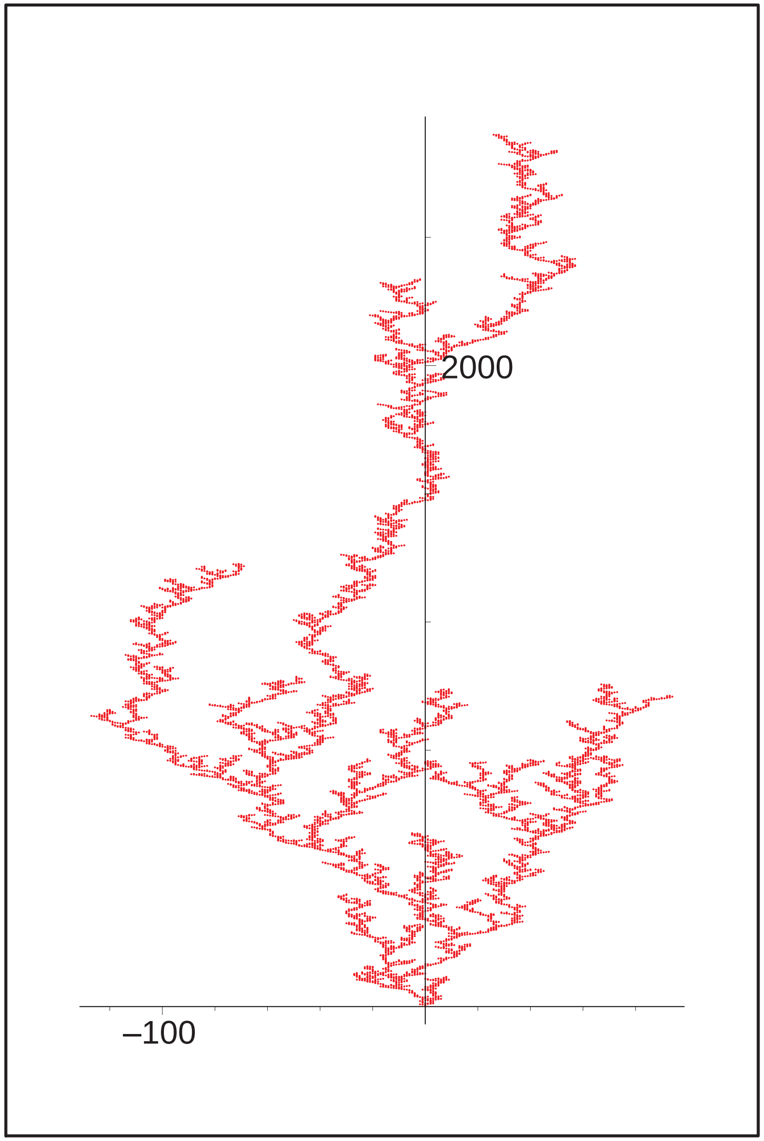
DLA: diffusion limited aggregation
source:Vincent Beffara's page

Other model: internal DLA; the limit is the circle (CV in proba), Bramson, Griffeath, Lawler

Unknown limits

DLA-directed: diffusion limited aggregation

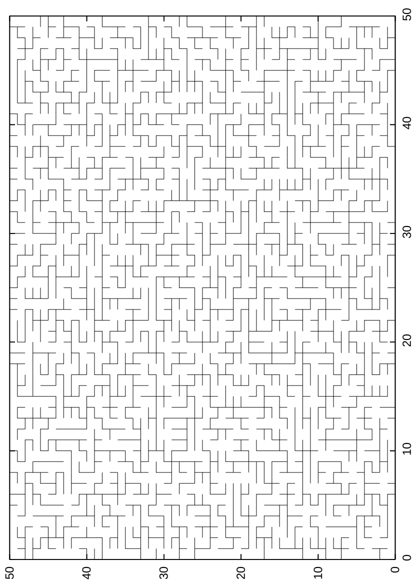
Unknown limits

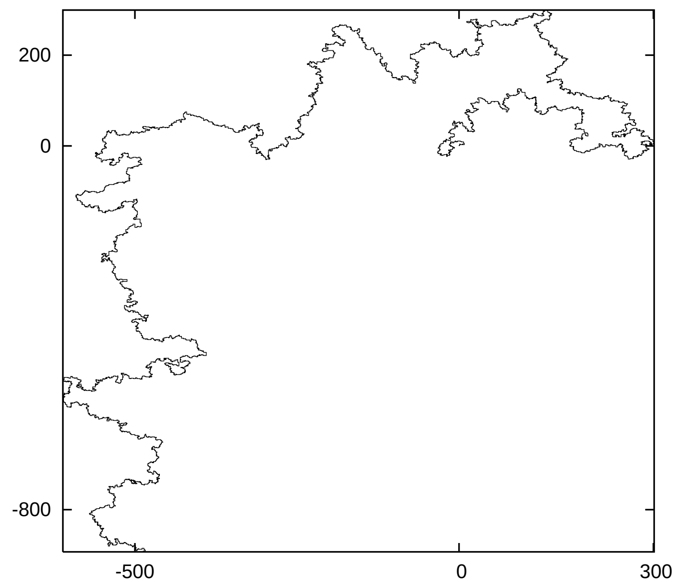


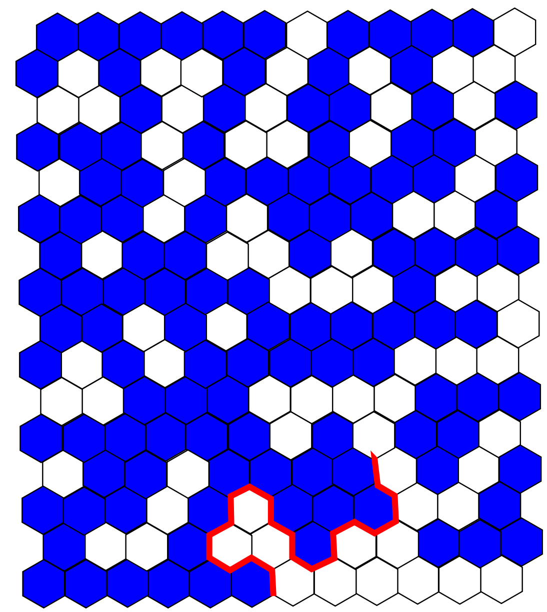
Directed animal

More or less known limits

SLE related process:limit of loop erased random walk, self avoiding random walks, contour process of percolation cluster, uniform spanning tree,...
Works of Lawler, Schramm, Werner







Convergence for the Hausdorff topology to conformally invariant distribution

Other models

Voter models, Ising models, First passage percolation, Richardson's growth model,...

That's all...
Thanks