RANDOM TREES –
AN ANALYTIC APPROACH

Michael Drmota

Institute of Discrete Mathematics and Geometry

TU Wien

A 1040 Wien, Austria

michael.drmota@tuwien.ac.at

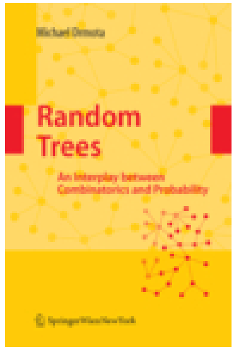
http://www.dmg.tuwien.ac.at/drmota/

Munich Summer School, Discrete Random Systems, Schliersee, Sept. 28–30, 2022

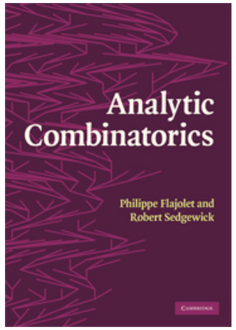
Contents

References

Books



Michael Drmota,
Random Trees, Springer, Wien-New York, 2009.



Philippe Flajolet and Robert Sedgewick,
Analytic Combinatorics, Cambridge University Press, 2009.
(http://algo.inria.fr/flajolet/Publications/books.html)

Asymptotic analysis of random objects

Levels of complexity:

1. Asymptotic enumeration

2. Distribution of (shape) parameters

3. Asymptotic shape (= continuous limiting object)

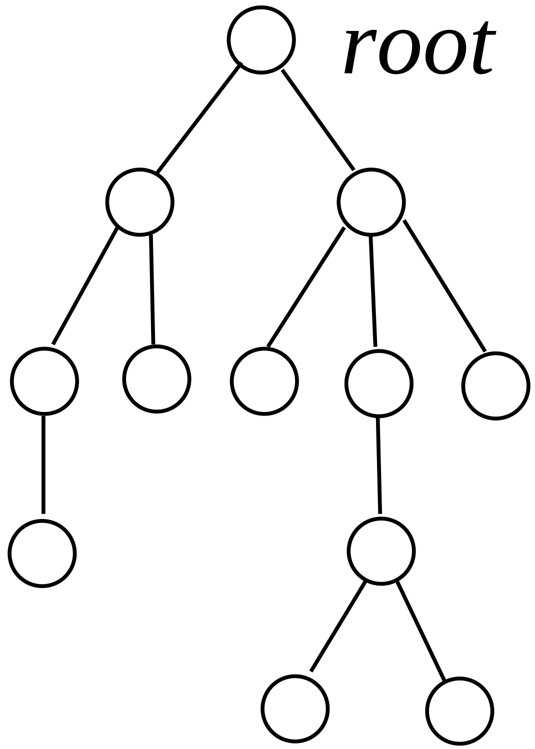
Contents 1

I. COMBINATORIAL RANDOM TREES

- Catalan trees and Cayley trees
- Functional equations and algebraic singularities
- A combinatorial central limit theorem
- The degree distribution of random trees

Random Trees

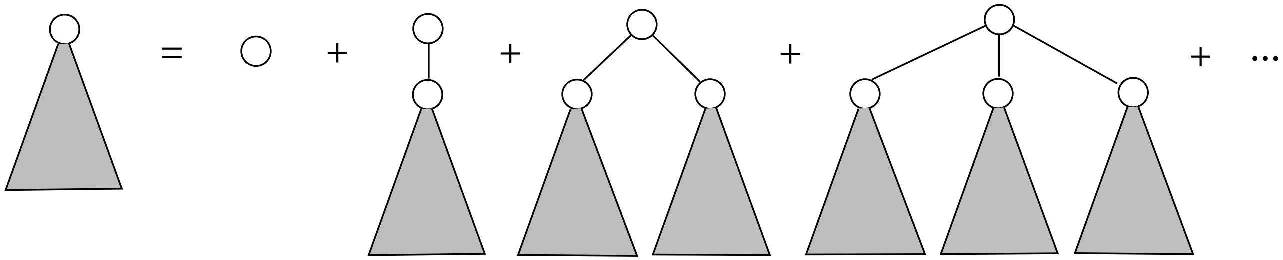
Catalan trees



rooted, ordered (or plane) tree

Random Trees

Catalan trees. g_n = number of Catalan trees of size n; $G(x) = \sum_{n \geq 1} g_n x^n$



$$G(x) = x(1 + G(x) + G(x)^2 + \cdots) = \frac{x}{1 - G(x)}$$

$$G(x) = \frac{1 - \sqrt{1 - 4x}}{2} \quad \Longrightarrow \quad g_n = \frac{1}{n}\binom{2n-2}{n-1} \sim \frac{4^{n-1}}{\sqrt{\pi} \cdot n^{3/2}}$$

(Catalan numbers)

Random Trees

Catalan trees with singularity analysis (to be discussed later)

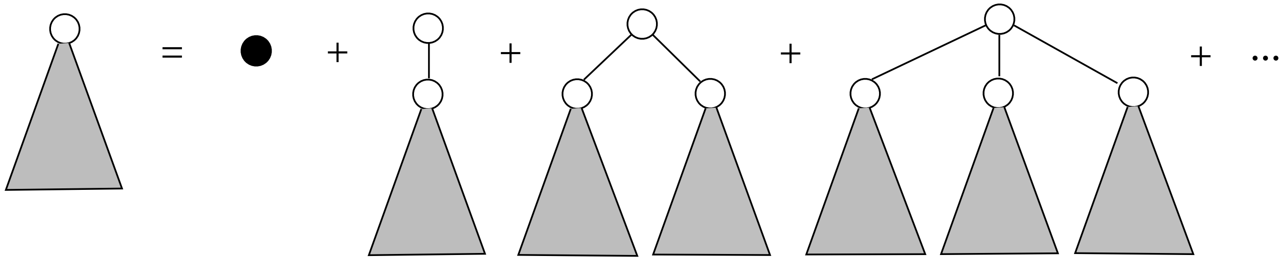
$$G(x) = \frac{1-\sqrt{1-4x}}{2} = \frac{1}{2} - \frac{1}{2}\sqrt{1-4x}$$

$$\Longrightarrow \quad g_n \sim -\frac{1}{2} \cdot \frac{4^n n^{-3/2}}{\Gamma(-\frac{1}{2})} = \frac{4^{n-1}}{\sqrt{\pi} \cdot n^{3/2}}$$

Random Trees

Number of leaves of Catalan trees

$g_{n,k}$ = number of Catalan trees of size n with k leaves.



$$\boxed{G(x,u) = xu + x(G(x,u) + G(x,u)^2 + \cdots = xu + \frac{xG(x,u)}{1 - G(x,u)}}$$

$$\Longrightarrow \quad G(x,u) = \frac{1}{2}\left(1 + (u-1)x - \sqrt{1 - 2(u+1)x + (u-1)^2x^2}\right)$$

$$\Longrightarrow \quad g_{n,k} = \frac{1}{n}\binom{n}{k}\binom{n-1}{k} \sim \frac{4^n}{\pi n^2} \exp\left(-\frac{(k-\frac{n}{2})^2}{\frac{1}{4}n}\right) \quad \text{for } k \approx \frac{n}{2}$$

Random Trees

Number of leaves of Catalan trees

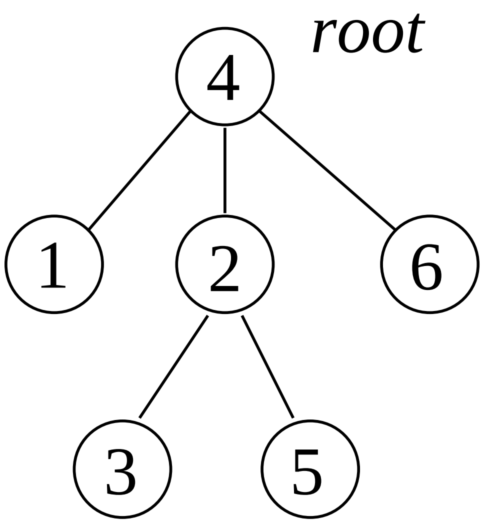
$$\boxed{G(x,u) = g(x,u) - h(x,u)\sqrt{1 - \frac{x}{\rho(u)}}}$$

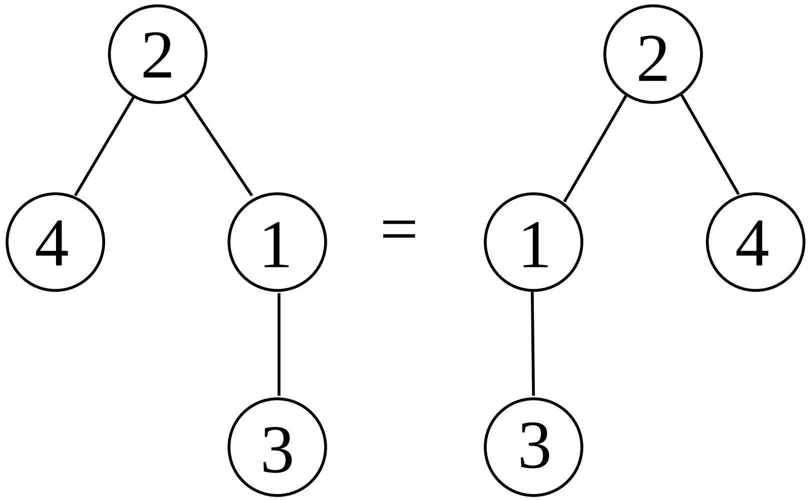
for certain analytic function $g(x,u), h(x,u)$, and $\rho(u)$.

$$\Longrightarrow \quad g_{n,k} = ???$$

Random Trees

Cayley Trees:

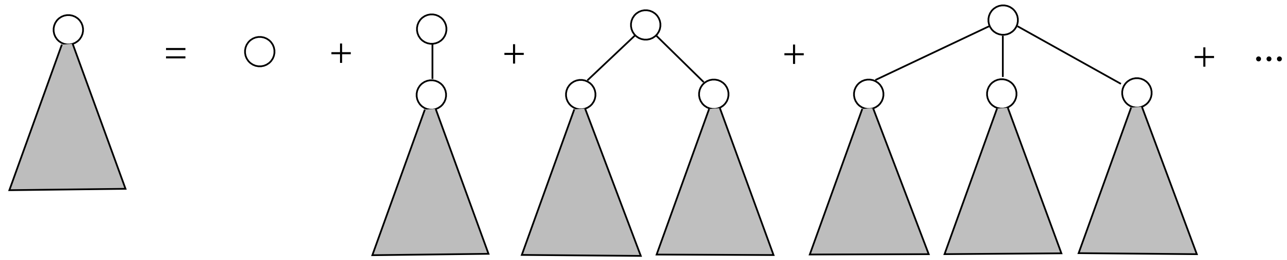




labelled, rooted, unordered (or non-plane) tree

Random Trees

Cayley Trees. r_n =number of Cayley trees of size n; $\boxed{R(x) = \sum_{n \geq 1} r_n \frac{x^n}{n!}}$



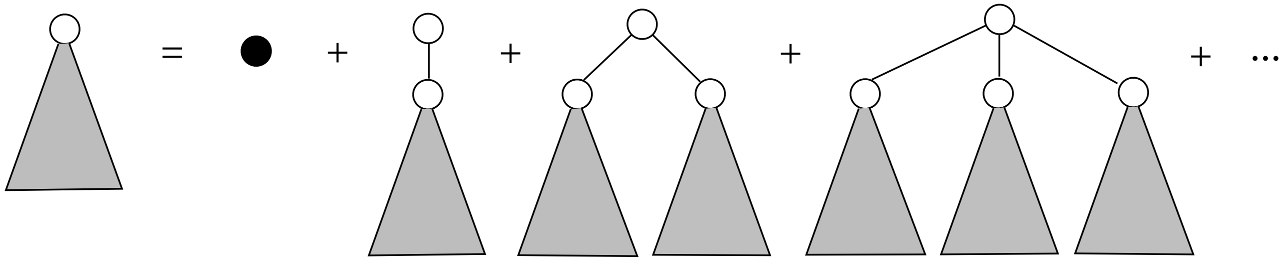
$$\boxed{R(x) = x\left(1 + R(x) + \frac{R(x)^2}{2!} + \frac{R(x)^3}{3!} + \cdots\right) = x\, e^{R(x)}}$$

$\Longrightarrow$ $\boxed{r_n = n^{n-1}}$ ... by Lagrange inversion

Random Trees

Number of leaves of Cayley trees

$r_{n,k} =$ number of Cayley trees of size n with k leaves.



$$R(x,u) = xu + x\left(R(x,u) + \frac{R(x,u)^2}{2!} + \frac{R(x,u)^3}{3!} + \cdots\right) = xe^{R(x,u)} + x(u-1)$$

$$\Longrightarrow \quad R(x,u) = ???$$

Functional equations

Catalan trees: $G(x,u) = xu + xG(x,u)/(1 - G(x,u))$

Cayley trees: $R(x,u) = xe^{R(x,u)} + x(u-1)$

Recursive structure leads to functional equation for gen. func.:

$$\boxed{\boxed{A(x,u) = \Phi(x,u,A(x,u))}}$$

Functional equations

Linear functional equation: $\Phi(x,u,a) = \Phi_0(x,u) + a\Phi_1(x,u)$

$$\Longrightarrow \quad A(x,u) = \frac{\Phi_0(x,u)}{1-\Phi_1(x,u)}$$

Usually these kinds of generating functions are easy to handle, since they are explicit.

Functional equations

Non-linear functional equations: $\Phi_{aa}(x,u,a) \neq 0$.

Suppose that $\boxed{A(x,u) = \Phi(x,u,A(x,u))}$, where $\Phi(x,u,a)$ has a power series expansion at $(0,0,0)$ with non-negative coefficients and $\Phi_{aa}(x,u,a) \neq 0$.

Let $x_0 > 0$, $a_0 > 0$ (inside the region of convergence) satisfy the system of equations:

$$\boxed{a_0 = \Phi(x_0, 1, a_0), \quad 1 = \Phi_a(x_0, 1, a_0)}.$$

Then there exists analytic function $g(x,u), h(x,u)$, and $\rho(u)$ such that locally

$$\boxed{A(x,u) = g(x,u) - h(x,u)\sqrt{1 - \frac{x}{\rho(u)}}}.$$

Functional equations

Idea of the Proof.

Set $\boxed{F(x,u,a) = \Phi(x,u,a) - a}$. Then we have

$$
\begin{aligned}
F(x_0, 1, a_0) &= 0 \\
F_a(x_0, 1, a_0) &= 0 \\
F_x(x_0, 1, a_0) &\neq 0 \\
F_{aa}(x_0, 1, a_0) &\neq 0.
\end{aligned}
$$

Weierstrass preparation theorem implies that there exist analytic functions $H(x,u,a)$, $p(x,u)$, $q(x,u)$ with $H(x_0,1,a_0) \neq 0$, $p(x_0,1) = q(x_0,1) = 0$ and

$$
\boxed{F(x,u,a) = H(x,u,a)\Big((a-a_0)^2 + p(x,u)(a-a_0) + q(x,u)\Big)}.
$$

Functional equations

$$F(x,u,a)=0 \quad \Longleftrightarrow \quad \boxed{(a-a_0)^2+p(x,u)(a-a_0)+q(x,u)=0}.$$

Consequently

$$\begin{aligned} A(x,u) &= a_0 - \frac{p(x,u)}{2} \pm \sqrt{\frac{p(x,u)^2}{4} - q(x,u)} \\ &= \boxed{g(x,u) - h(x,u)\sqrt{1-\frac{x}{\rho(u)}}}, \end{aligned}$$

where we write

$$\frac{p(x,u)^2}{4} - q(x,u) = K(x,u)(x-\rho(u))$$

which is again granted by the Weierstrass preparation theorem and we set

$$g(x,u) = a_0 - \frac{p(x,u)}{2} \quad \text{and} \quad h(x,u) = \sqrt{-K(x,u)\rho(u)}.$$

Random Trees

Catalan Trees $G(x,u) = xu + \frac{xG(x,u)}{1-G(x,u)}$

$$\Longrightarrow \quad G(x,u) = g(x,u) - h(x,u)\sqrt{1 - \frac{x}{\rho(u)}}$$

$$G(x,1) = G(x) = g(x,1) - h(x,1)\sqrt{1 - \frac{x}{\rho(1)}}, \quad \rho(1) = \frac{1}{4}$$

Cayley Trees $T(x,u) = xe^{T(x,u)} + x(u-1)$

$$\Longrightarrow \quad T(x,u) = g(x,u) - h(x,u)\sqrt{1 - \frac{x}{\rho(u)}}$$

$$T(x,1) = T(x) = g(x,1) - h(x,1)\sqrt{1 - \frac{x}{\rho(1)}}, \quad \rho(1) = \frac{1}{e}$$

Algebraic Singularities

Singular expansion

$$
\begin{aligned}
A(x) &= g(x) - h(x)\sqrt{1-\frac{x}{\rho}} \\
&= \Big(g_0 + g_1(x-\rho) + g_2(x-\rho)^2 + \cdots\Big) \\
&\quad + \Big(h_0 + h_1(x-\rho) + h_2(x-\rho)^2 + \cdots\Big)\sqrt{1-\frac{x}{\rho}} \\
&= a_0 + a_1\left(1-\frac{x}{\rho}\right)^{\frac{1}{2}} + a_2\left(1-\frac{x}{\rho}\right)^{\frac{2}{2}} + a_3\left(1-\frac{x}{\rho}\right)^{\frac{3}{2}} + \cdots \\
&= a_0 + a_1\left(1-\frac{x}{\rho}\right)^{\frac{1}{2}} + a_2\left(1-\frac{x}{\rho}\right) + O\left(\left(1-\frac{x}{\rho}\right)^{\frac{3}{2}}\right)
\end{aligned}
$$

Algebraic Singularities

Singular expansion

$$
\begin{aligned}
A(x) &= \boxed{g(x) - h(x)\sqrt{1-\frac{x}{\rho}}} \\
&= \left(g_0 + g_1(x-\rho) + g_2(x-\rho)^2 + \cdots\right) \\
&\quad + \left(h_0 + h_1(x-\rho) + h_2(x-\rho)^2 + \cdots\right)\sqrt{1-\frac{x}{\rho}} \\
&= a_0 + a_1\left(1-\frac{x}{\rho}\right)^{\frac{1}{2}} + a_2\left(1-\frac{x}{\rho}\right)^{\frac{2}{2}} + a_3\left(1-\frac{x}{\rho}\right)^{\frac{3}{2}} + \cdots \\
&= a_0 + a_1\boxed{\left(1-\frac{x}{\rho}\right)^{\frac{1}{2}}} + a_2\left(1-\frac{x}{\rho}\right) + \boxed{O\left(\left(1-\frac{x}{\rho}\right)^{\frac{3}{2}}\right)}
\end{aligned}
$$

Algebraic Singularities

Singularity Analysis

Lemma 1 *Suppose that*

$$\boxed{y(x) = \left(1 - \frac{x}{x_0}\right)^{-\alpha}}.$$

Then

$$\boxed{y_n = (-1)^n \binom{-\alpha}{n} x_0^{-n} = \frac{n^{\alpha-1}}{\Gamma(\alpha)} x_0^{-n} + \mathcal{O}\left(n^{\alpha-2} x_0^{-n}\right).}$$

Remark: This asymptotic expansion is uniform in α if α varies in a compact region of the complex plane.

Algebraic Singularities

Singularity Analysis

Lemma 2 (Flajolet and Odlyzko) *Let*

$$y(x) = \sum_{n \ge 0} y_n \, x^n$$

be analytic in a region

$$\Delta = \{x : |x| < x_0 + \eta, \, |\arg(x - x_0)| > \delta\},$$

$x_0 > 0$, $\eta > 0$, $0 < \delta < \pi/2$.

Suppose that for some real α

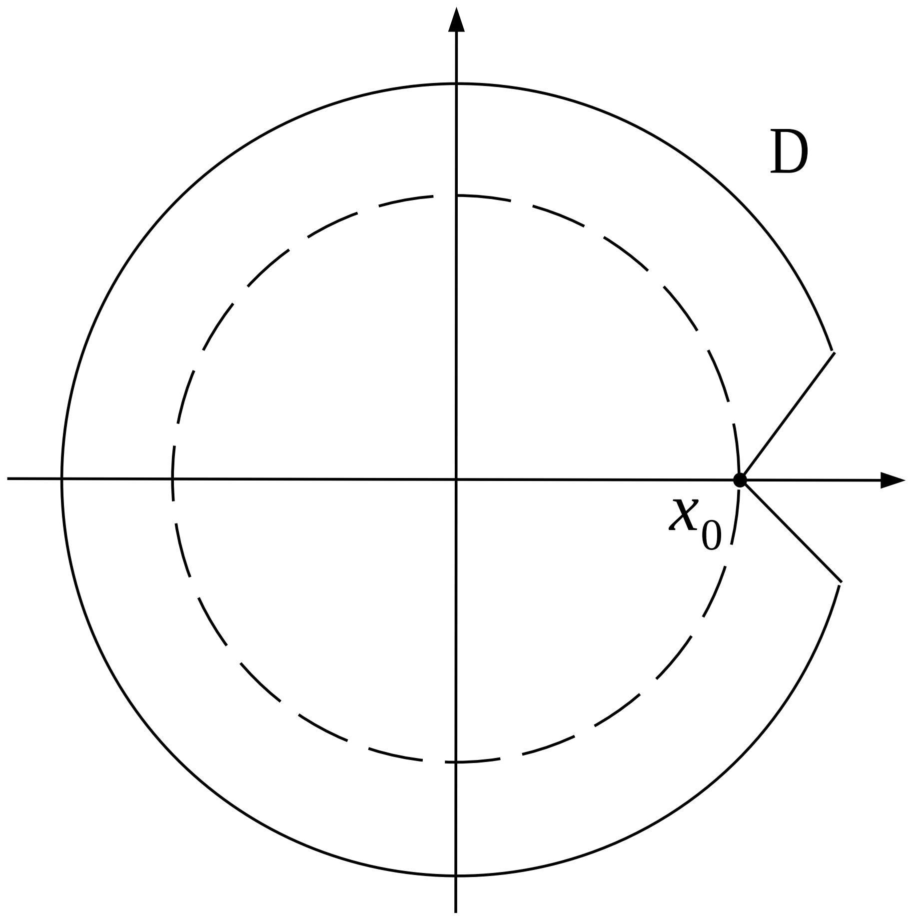
$$\boxed{y(x) = \mathcal{O}\left((1 - x/x_0)^{-\alpha}\right) \qquad (x \in \Delta).}$$

Then

$$\boxed{y_n = \mathcal{O}\left(x_0^{-n} n^{\alpha - 1}\right).}$$

Algebraic Singularities

Δ-region



Algebraic Singularities

Singularity Analysis

Suppose that

$$A(x) = g(x) - h(x)\sqrt{1-\frac{x}{\rho}}$$
$$= a_0 + a_1\left(1-\frac{x}{\rho}\right)^{\frac{1}{2}} + a_2\left(1-\frac{x}{\rho}\right) + O\left(\left(1-\frac{x}{\rho}\right)^{\frac{3}{2}}\right)$$

for $x \in \Delta$ then

$$\boxed{a_n = [x^n]\,A(x) = \frac{h(\rho)}{2\sqrt{\pi}}\rho^{-n}n^{-\frac{3}{2}}\left(1+O\left(\frac{1}{n}\right)\right)}.$$

Algebraic Singularities

Singularity Analysis

Suppose that

$$A(x,u) = g(x,u) - h(x,u)\sqrt{1-\frac{x}{\rho(u)}}$$
$$= a_0(u) + a_1(u)\left(1-\frac{x}{\rho(u)}\right)^{\frac{1}{2}} + a_2(u)\left(1-\frac{x}{\rho(u)}\right) + O\left(\left(1-\frac{x}{\rho(u)}\right)^{\frac{3}{2}}\right)$$

for $x \in \Delta = \Delta(u)$ then

$$\boxed{a_n(u) = [x^n]\,A(x,u) = \frac{h(\rho(u),u)}{2\sqrt{\pi}}\rho(u)^{-n}n^{-\frac{3}{2}}\left(1+O\left(\frac{1}{n}\right)\right)}.$$

Probabilistic Model

a_n ... number of objects of size n

$a_{n,k}$ ... number of objects of size n, where a certain **parameter** has value k

If all objects of size n are considered to be **equally likely** then the parameter can be considered as a random variable X_n with distribution

$$\boxed{\mathbb{P}\{X_n = k\} = \frac{a_{nk}}{a_n}}.$$

Probabilistic Model

Generating functions and the probability generating function

$$A(x,u) = \sum_{n,k} a_{n,k} x^n u^k$$

$$\Longrightarrow \quad \boxed{\mathbb{E}\, u^{X_n}} = \sum_{k \geq 0} \mathbb{P}\{X_n = k\}\, u^k$$

$$= \sum_{k \geq 0} \frac{a_{nk}}{a_n}\, u^k$$

$$= \boxed{\frac{[x^n]\, A(x,u)}{[x^n]\, A(x,1)}} = \frac{a_n(u)}{a_n}$$

Probabilistic Model

Generating functions and the probability generating function

$$A(x,u) = g(x,u) - h(x,u)\sqrt{1 - \frac{x}{\rho(u)}}$$

$$\Longrightarrow \quad \boxed{\mathbb{E}\, u^{X_n}} = \frac{[x^n]\,A(x,u)}{[x^n]\,A(x,1)}$$

$$= \frac{\frac{h(\rho(u),u)}{2\sqrt{\pi}}\rho(u)^{-n} n^{-\frac{3}{2}}\left(1 + O\left(\frac{1}{n}\right)\right)}{\frac{h(\rho(1),1)}{2\sqrt{\pi}}\rho(1)^{-n} n^{-\frac{3}{2}}\left(1 + O\left(\frac{1}{n}\right)\right)}$$

$$= \boxed{\frac{h(\rho(u),u)}{h(\rho(1),1)}\left(\frac{\rho(1)}{\rho(u)}\right)^n \left(1 + O\left(\frac{1}{n}\right)\right)}.$$

Probabilistic Model

Quasi-Power Theorem (Hwang)

Let X_n be a sequence of random variables with the property that

$$\boxed{\mathbb{E}\, u^{X_n} = A(u) \cdot B(u)^{\lambda_n} \cdot \left(1 + O\left(\frac{1}{\phi_n}\right)\right)}$$

holds uniformly in a complex neighborhood of $u = 1$, $\boxed{\lambda_n \to \infty}$ and $\boxed{\phi_n \to \infty}$, and $A(u)$ and $B(u)$ are analytic functions in a neighborhood of $u = 1$ with $A(1) = B(1) = 1$. Set

$$\mu = B'(1) \quad \text{and} \quad \sigma^2 = B''(1) + B'(1) - B'(1)^2.$$

$$\Longrightarrow \quad \boxed{\mathbb{E}\, X_n = \mu\lambda_n + O\left(1 + \lambda_n/\phi_n\right)}, \quad \boxed{\mathbb{V}\, X_n = \sigma^2\lambda_n + O\left(1 + \lambda_n/\phi_n\right)},$$

$$\boxed{\frac{X_n - \mathbb{E}\, X_n}{\sqrt{\mathbb{V}\, X_n}} \xrightarrow{\mathrm{d}} N(0,1)} \quad (\sigma^2 \neq 0).$$

Probabilistic Model

Sums of independent random variables

$X_n = \xi_1 + \xi_2 + \cdots + \xi_n$, where ξ_j are i.i.d.

$$B(u) = \mathbb{E}\, u^{\xi_j}$$

$$\Longrightarrow \quad \begin{aligned} \mathbb{E}\, u^{X_n} &= \mathbb{E}\, u^{\xi_1+\xi_2+\cdots+\xi_n} \\ &= \mathbb{E}\, u^{\xi_1} \cdot \mathbb{E}\, u^{\xi_2} \cdots \mathbb{E}\, u^{\xi_n} \\ &= B(u)^n. \end{aligned}$$

Probabilistic Model

COMBINATORIAL CENTRAL LIMIT THEOREM

Suppose that a sequence of random variables X_n has distribution

$$\mathbb{P}\{X_n = k\} = \frac{a_{nk}}{a_n},$$

where the generating function $A(x,u) = \sum_{n,k} a_{n,k} x^n u^k$ satisfies a functional equation of the form $\boxed{A(x,u) = \Phi(x,u,A(x,u))}$, where $\Phi(x,u,a)$ has a power series expansion at $(0,0,0)$ with non-negative coefficients and $\Phi_{aa}(x,u,a) \neq 0$.

Let $x_0 > 0$, $a_0 > 0$ (inside the region of convergence) satisfy the system of equations:

$$\boxed{a_0 = \Phi(x_0, 1, a_0), \quad 1 = \Phi_a(x_0, 1, a_0)}.$$

Probabilistic Model

COMBINATORIAL CENTRAL LIMIT THEOREM (cont.)

Set

$$\mu = \frac{\Phi_u}{x_0 \Phi_x},$$

$$\sigma^2 = \mu + \mu^2 + \frac{1}{x_0 \Phi_x^3 \Phi_{aa}} \Big(\Phi_x^2 (\Phi_{aa}\Phi_{uu} - \Phi_{au}^2) - 2\Phi_x \Phi_u (\Phi_{aa}\Phi_{xu} - \Phi_{ax}\Phi_{au}) + \Phi_u^2 (\Phi_{aa}\Phi_{xx} - \Phi_{ax}^2) \Big),$$

(where all partial derivatives are evaluated at the point $(x_0, a_0, 1)$)

Then we have

$$\boxed{\mathbb{E}\, X_n = \mu n + O(1)} \quad \text{and} \quad \boxed{\mathbb{V}\mathrm{ar}\, X_n = \sigma^2 n + O(1)}$$

and if $\sigma^2 > 0$ then

$$\boxed{\boxed{\frac{X_n - \mathbb{E}\, X_n}{\sqrt{\mathbb{V}\mathrm{ar}\, X_n}} \to N(0,1)}}.$$

Random Trees

Leaves in Catalan trees

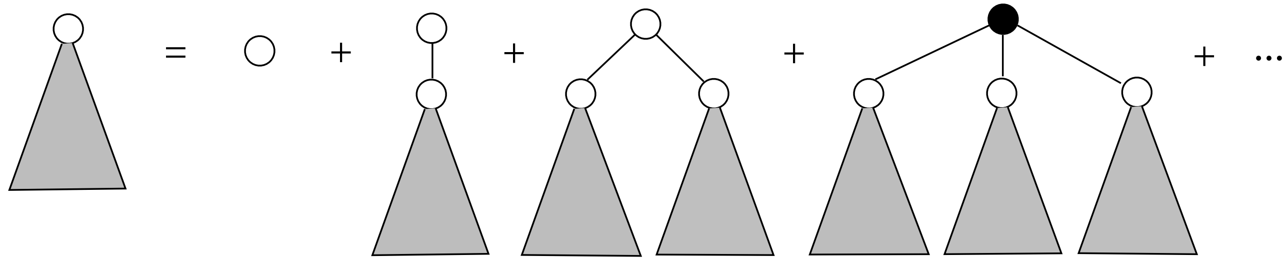
The number of leaves in Catalan trees of size n satisfy a **central limit theorem** with mean $\sim \frac{1}{2}n$ and variance $\sim \frac{1}{8}n$

Leaves in Cayley trees

The number of leaves in Cayley trees of size n satisfy a **central limit theorem** with mean $\sim \frac{1}{e}n$ and variance $\sim \left(\frac{1}{e^2} + \frac{1}{e}\right) n$

Random Trees

Nodes of out-degree d in Catalan trees



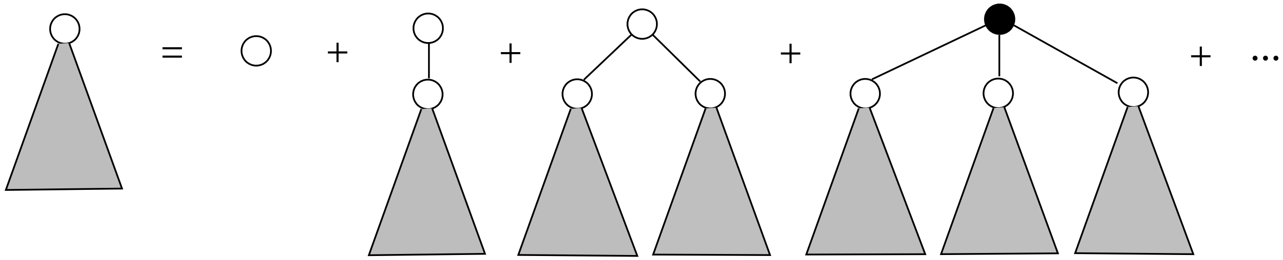
$$G(x,u) = \frac{x}{1-G(x,u)} + x(u-1)G(x,u)^d$$

The number $X_n^{(d)}$ of nodes with out-degree d in Catalan trees of size n satisfy a **central limit theorem** with mean $\sim \mu_d n$ and variance $\sim \sigma_d^2 n$, where

$$\mu_d = \frac{1}{2^{d+1}} \quad \text{and} \quad \sigma_d^2 = \frac{1}{2^{d+1}} + \frac{1}{2^{2(d+1)}} - \frac{(d-1)^2}{2^{2d+3}}.$$

Random Trees

Nodes of out-degree d in Cayley trees



$$R(x,u) = xe^{R(x,u)} + x(u-1)\frac{R(x,u)^d}{d!}$$

The number of nodes with out-degree d in Cayley trees of size n satisfy a **central limit theorem** with mean $\sim \mu_d n$ and variance $\sim \sigma_d^2 n$, where

$$\mu_d = \frac{1}{e\,d!} \quad \text{and} \quad \sigma_d^2 = \frac{1+(d-1)^2}{e^2(d!)^2} + \frac{1}{e\,d!}$$

Random Trees

Degree distribution for Catalan trees

$p_{n,d}$ ... probability that a random node in a random Catalan tree of size n has out-degree d:

$$\mathbb{E}\, X_n^{(d)} = n\, p_{n,d}$$

$$p_d := \lim_{n\to\infty} p_{n,d} = \frac{1}{2^{d+1}} = \mu_d$$

Probability generating function of the out-degree distribution:

$$\boxed{p(w) := \sum_{d\geq 0} p_d w^d = \frac{1}{2-w}}$$

Random Trees

Degree distribution for Cayley trees

$p_{n,d}$ ... probability that a random node in a random Cayley tree of size n has out-degree d:

$$\mathbb{E}\, X_n^{(d)} = n\, p_{n,d}$$

$$p_d := \lim_{n\to\infty} p_{n,d} = \frac{1}{e\, d!} = \mu_d$$

Probability generating function of the out-degree distribution:

$$\boxed{p(w) := \sum_{d\geq 1} p_d w^d = e^{w-1}}$$

Contents 2

I. COMBINATORIAL RANDOM TREES

- Maximum degree
- Unrooted trees

II. PATTERN COUNTS IN RANDOM TREES

- Pattern in trees
- Systems of functional equations

Random Trees

Maximum degree

Δ_n ... maximum out-degree

$X_n^{(>d)} = X_n^{(d+1)} + X_n^{(d+2)} + \cdots$ ... **number of nodes of out-degree** $> d$.

$$\boxed{\Delta_n > d \iff X_n^{(>d)} > 0}$$

Random Trees

First moment method

X ... a discrete random variable on non-negative integers.

$$\Longrightarrow \quad \boxed{\mathbb{P}\{X > 0\} \le \min\{1, \mathbb{E}\, X\}}$$

Proof

$$\mathbb{E}\, X = \sum_{k \ge 0} k\, \mathbb{P}\{X = k\} \ge \sum_{k \ge 1} \mathbb{P}\{X = k\} = \mathbb{P}\{X > 0\}.$$

Random Trees

Second moment method

X is a non-negative random variable with finite second moment.

$$\Longrightarrow \quad \boxed{\mathbb{P}\{X > 0\} \geq \frac{(\mathbb{E}\, X)^2}{\mathbb{E}\,(X^2)}}$$

Proof

$$\mathbb{E}\, X = \mathbb{E}\left(X \cdot \mathbf{1}_{[X>0]}\right) \leq \sqrt{\mathbb{E}\,(X^2)}\sqrt{\mathbb{E}\,(\mathbf{1}^2_{[X>0]})} = \sqrt{\mathbb{E}(X^2)}\sqrt{\mathbb{P}\{X > 0\}}.$$

Random Trees

Tail estimates and expected value

- $\mathbb{P}\{\Delta_n > d\} \le \min\{1, \mathbb{E}\, X_n^{(>d)}\}$

- $\mathbb{P}\{\Delta_n > d\} \ge \dfrac{(\mathbb{E}\, X_n^{(>d)})^2}{\mathbb{E}\,(X_n^{(>d)})^2}$

$$\Longrightarrow \quad \mathbb{P}\{\Delta_n \le d\} \le 1 - \frac{(\mathbb{E}\, X_n^{(>d)})^2}{\mathbb{E}\,(X_n^{(>d)})^2} = \frac{\mathbb{V}\mathrm{ar}\, X_n^{(>d)}}{\mathbb{E}\,(X_n^{(>d)})^2}$$

- $\mathbb{E}\,\Delta_n = \sum_{d\ge 0} \mathbb{P}\{\Delta_n > d\}$

Random Trees

Maximum degree of Catalan trees

$$\mathbb{E}\, X_n^{(>d)} \sim \frac{n}{2^{d+1}}, \qquad \mathbb{V}\mathrm{ar}\,(X_n^{(>d)})^2 \sim n \left(\frac{1}{2^{d+1}} + \frac{1}{2^{2(d+1)}} - \frac{(d-1)^2}{2^{2d+3}} \right)$$

$$\Longrightarrow \quad \mathbb{P}\{\Delta_n > d\} \le \min\left\{1, \frac{n}{2^{d+1}}\right\},$$

$$\mathbb{P}\{\Delta_n \le d\} = 1 - \mathbb{P}\{\Delta_n > d\}$$

$$\le \frac{1}{n} \frac{\frac{1}{2^{d+1}} + \frac{1}{2^{2(d+1)}} - \frac{(d-1)^2}{2^{2d+3}}}{\frac{1}{2^{2(d+1)}}} \sim \frac{2^{d+1}}{n}$$

$\Longrightarrow$ $\boxed{\Delta_n \text{ is concentrated at } \log_2 n + O(1)}$

Random Trees

Maximum degree of Catalan trees (Carr, Goh and Schmutz)

$$\boxed{\mathbb{P}\{\Delta_n \le k\} = \exp\left(-2^{-(k-\log_2 n+1)}\right) + o(1)}$$

$$\boxed{\mathbb{E}\,\Delta_n = \log_2 n + O(1)}$$

Random Trees

Unrooted trees

p_n ... number of different embeddings of **unrooted** trees of size n in the plane, $\boxed{P(x) = \sum_{n\ge 1} p_n x^n}$:

$$\boxed{P(x) = x \sum_{k\ge 0} Z_{\mathfrak{C}_k}(G(x), G(x^2), \ldots, G(x^k)) - \frac{1}{2}G(x)^2 + \frac{1}{2}G(x^2)},$$

where $G(x) = x/(1 - G(x)) = (1 - \sqrt{1-4x})/2$ and

$$Z_{\mathfrak{C}_k}(x_1, x_2, \ldots, x_k) = \frac{1}{k}\sum_{d|k} \varphi(d) x_d^{k/d}$$

is the cycle index of the cyclic group $\mathfrak{C}_k$ of k elements

Random Trees

Unrooted trees

Cancellation of the $\sqrt{1-4x}$-term:

$$G(x) = \frac{1-\sqrt{1-4x}}{2} \quad \Longrightarrow \quad P(x) = a_0 + a_2(1-4x) + \frac{1}{6}(1-4x)^{3/2} + \cdots$$

$$\Longrightarrow \quad \boxed{p_n = \frac{1}{8\sqrt{\pi}} 4^n n^{-5/2} \left(1 + O(n^{-1})\right)}$$

Random Trees

Degree distribution of unrooted trees

$X_n^{(d)}$ ... number of nodes of degree d in trees of size n

$$
\begin{aligned}
P(x,u) = x \sum_{k \neq d} & Z_{\mathfrak{C}_k}(G(x,u), G(x^2,u^2), \ldots, G(x^k,u^k)) \\
& + xuZ_{\mathfrak{C}_d}(G(x,u), G(x^2,u^2), \ldots, G(x^d,u^d)) \\
& - \frac{1}{2}G(x,u)^2 + \frac{1}{2}G(x^2,u^2),
\end{aligned}
$$

where

$$G(x,u) = \frac{x}{1 - G(x,u)} + x(u-1)G(x,u)^{d-1}.$$

Random Trees

Degree distribution of unrooted trees

Cancellation of the $\sqrt{1-4x}$-term:

$$G(x,u) = g(x,u) - h(x,u)\sqrt{1-\frac{x}{\rho(u)}}$$

$$\Longrightarrow \quad P(x,u) = a_0(u) + a_2(u)\left(1-\frac{x}{\rho(u)}\right) + a_3(u)\left(1-\frac{x}{\rho(u)}\right)^{\frac{3}{2}} + \cdots$$

$\Longrightarrow X_n^{(d)}$ satisfies a **central limit theorem** with mean $\sim \mu_{d-1}n$ and variance $\sim \sigma_{d-1}^2 n$, where

$$\mu_d = \frac{1}{2^{d+1}} \quad \text{and} \quad \sigma_d^2 = \frac{1}{2^{d+1}} + \frac{1}{2^{2(d+1)}} - \frac{(d-1)^2}{2^{2d+3}}.$$

Random Trees

Degree distribution of unrooted trees

$p_{n,d}$ ... probability that a random node in a tree of size n has degree d:

$$\mathbb{E}\, X_n^{(d)} = n\, p_{n,d}$$

$$p_d = \lim_{n\to\infty} p_{n,d} = \mu_{d-1} = \frac{1}{2^d}$$

Probability generating function of the degree distribution:

$$\boxed{p(w) = \sum_{d\ge 1} p_d w^d = \frac{w}{2-w}}$$

Random Trees

Maximum degree for unrooted trees

Δ_n ... maximum degree of unrooted trees of size n

$$\boxed{\Delta_n \text{ is concentrated at } \log_2 n}$$

$$\boxed{\mathbb{E}\,\Delta_n = \log_2 n + O(1)}$$

Random Trees

Unrooted labelled trees

$t_n = r_n/n = n^{n-2}$ ... number of different **unrooted** labelled trees of size n: $\boxed{T(x) = \sum_{n\ge 1} t_n \frac{x^n}{n!}}$:

$$\boxed{T(x) = xe^{R(x)} - \frac{1}{2}R(x)^2 = R(x) - \frac{1}{2}R(x)^2},$$

where $R(x) = xe^{R(x)}$ (note that $T'(x) = R(x)/x$)

Cancellation of the $\sqrt{1-ex}$-term:

$$R(x) = g(x) - h(x)\sqrt{1-ex} \quad \Longrightarrow \quad T(x) = a_0 + a_2(1-4x) + \frac{1}{6}(1-ex)^{3/2} + \cdots$$

Random Trees

Degree distribution of unrooted labelled trees

$X_n^{(d)}$ ... number of nodes of degree d in trees of size n

$$T(x,u) = xe^{R(x,u)} + x(u-1)\frac{R(x,u)^d}{d!} - \frac{1}{2}R(x,u)^2,$$

where

$$R(x,u) = xe^{R(x,u)} + x(u-1)\frac{R(x,u)^{d-1}}{(d-1)!}.$$

Random Trees

Degree distribution of unrooted labelled trees

Cancellation of the $\sqrt{1-4x}$-term:

$$R(x,u) = g(x,u) - h(x,u)\sqrt{1-\frac{x}{\rho(u)}}$$

$$\Longrightarrow \quad T(x,u) = a_0(u) + a_2(u)\left(1-\frac{x}{\rho(u)}\right) + a_3(u)\left(1-\frac{x}{\rho(u)}\right)^{\frac{3}{2}} + \cdots$$

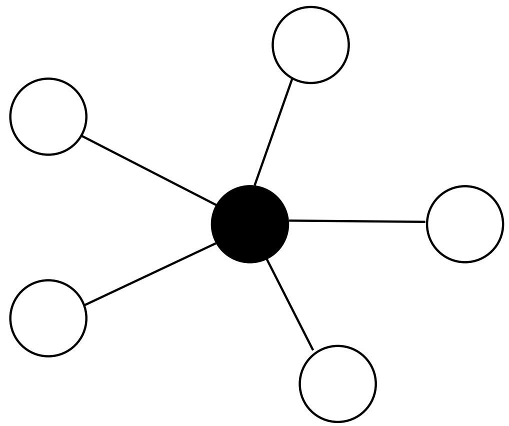
$\Longrightarrow$ $X_n^{(d)}$ satisfies a **central limit theorem** with mean $\sim \mu_{d-1} n$ and variance $\sim \sigma_{d-1}^2 n$, where

$$\mu_d = \frac{1}{e\,d!} \quad \text{and} \quad \sigma_d^2 = \frac{1+(d-1)^2}{e^2(d!)^2} + \frac{1}{e\,d!}$$

(Note again that $\frac{\partial}{\partial x}T(x,u) = R(x,u)/x$)

Random Trees

Star pattern

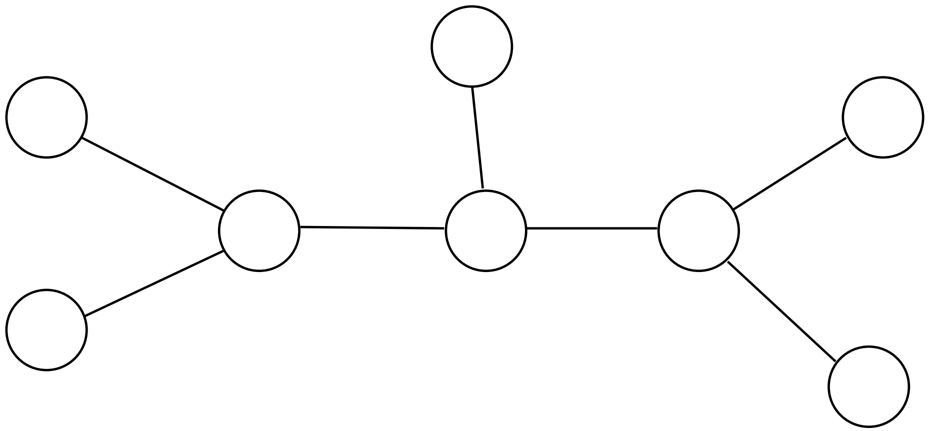


$d = 5$

$$X_n^{(d)} = \text{number of nodes of degree } d \text{ in trees of size } n$$
$$= \text{number of star pattern with } d \text{ rays in trees of size } n$$

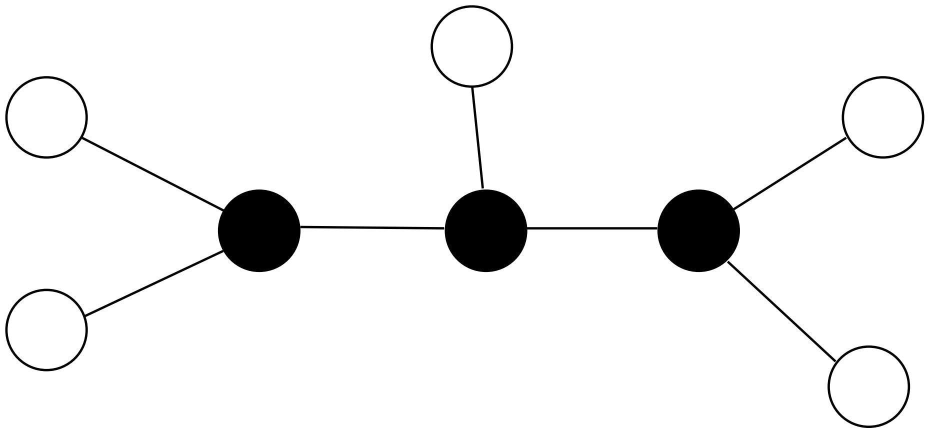
Patterns in Trees

Pattern $\mathcal{M}$

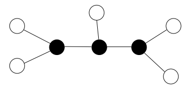


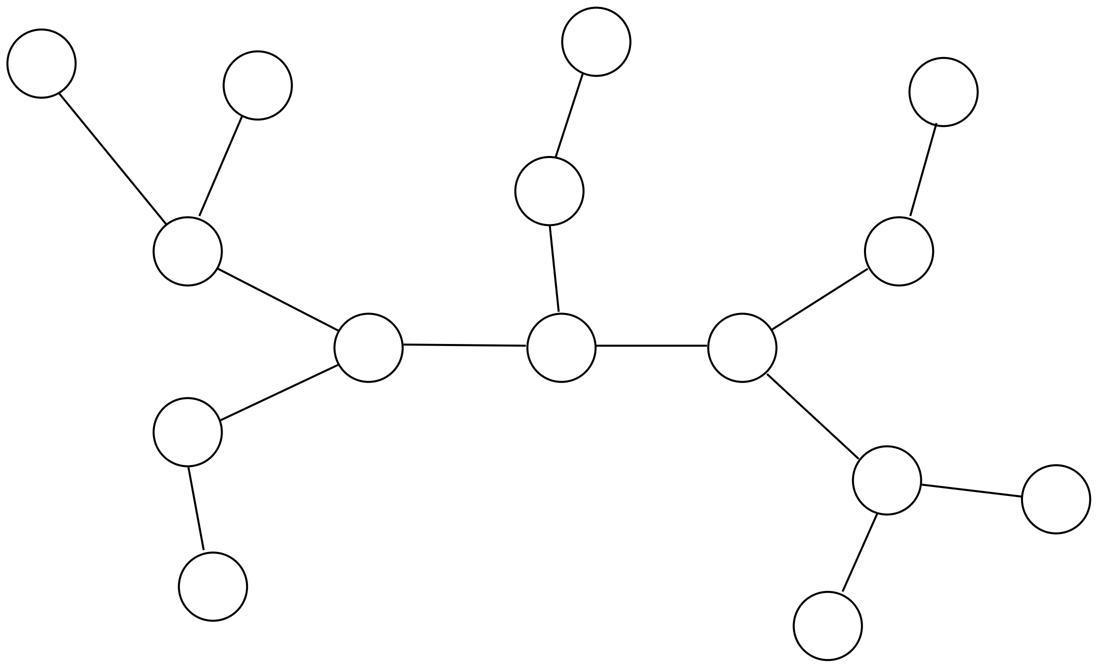
Patterns in Trees

Pattern $\mathcal{M}$

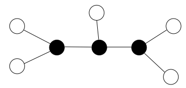


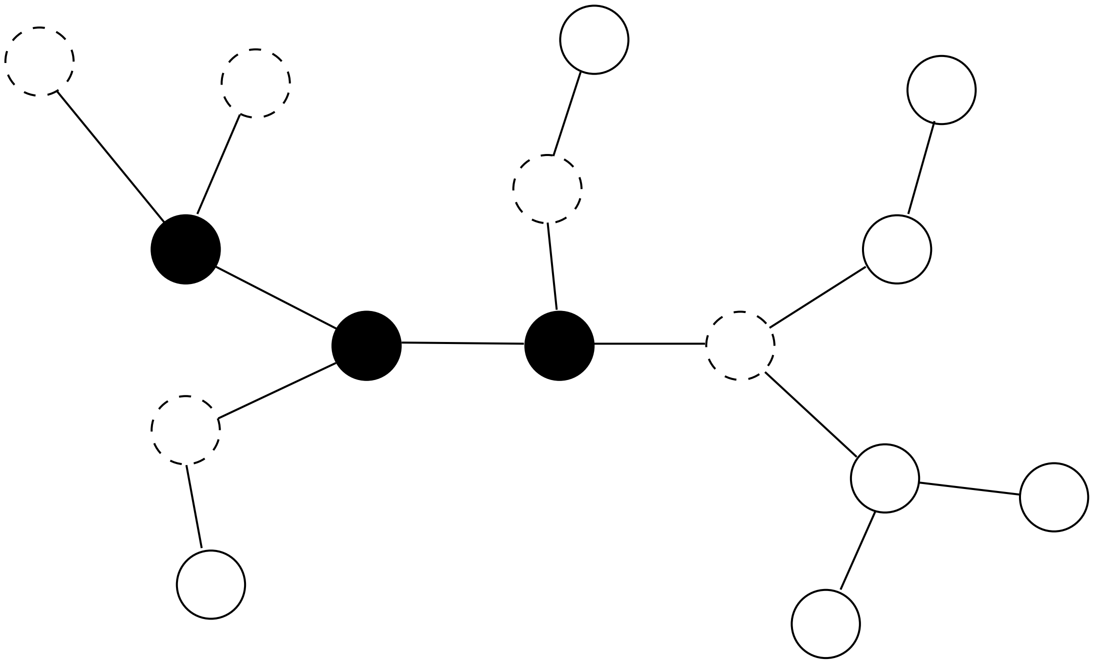
Patterns in Trees

Occurrence of a pattern $\mathcal{M}$ 

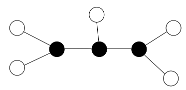


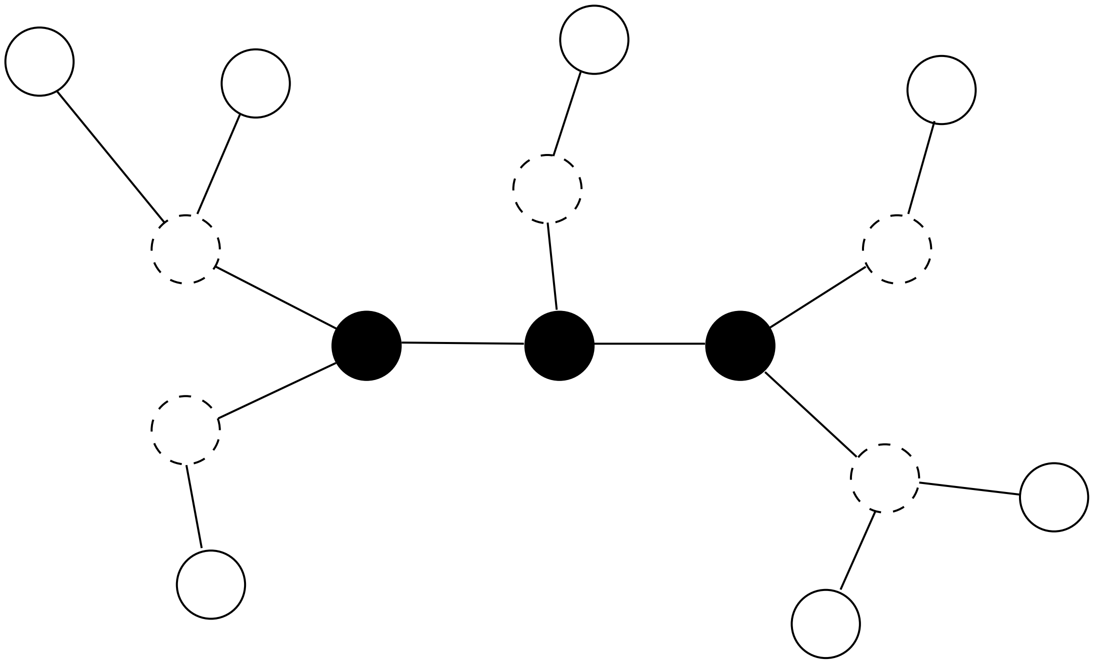
Patterns in Trees

Occurrence of a pattern $\mathcal{M}$ 

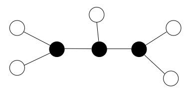


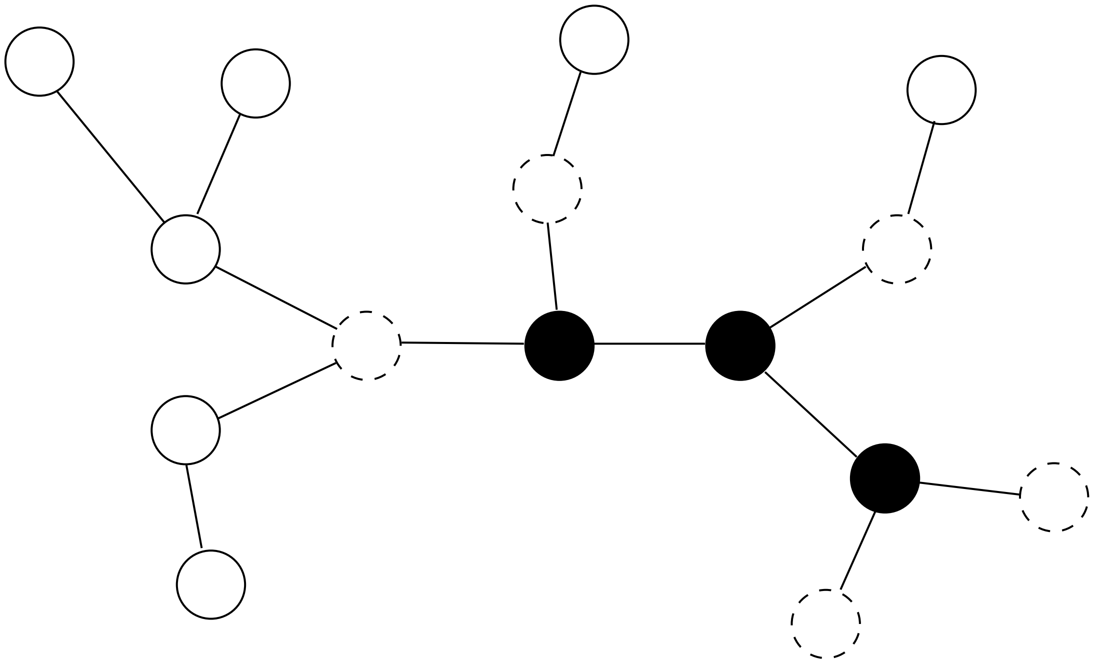
Patterns in Trees

Occurrence of a pattern $\mathcal{M}$ 

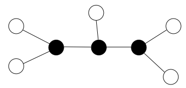


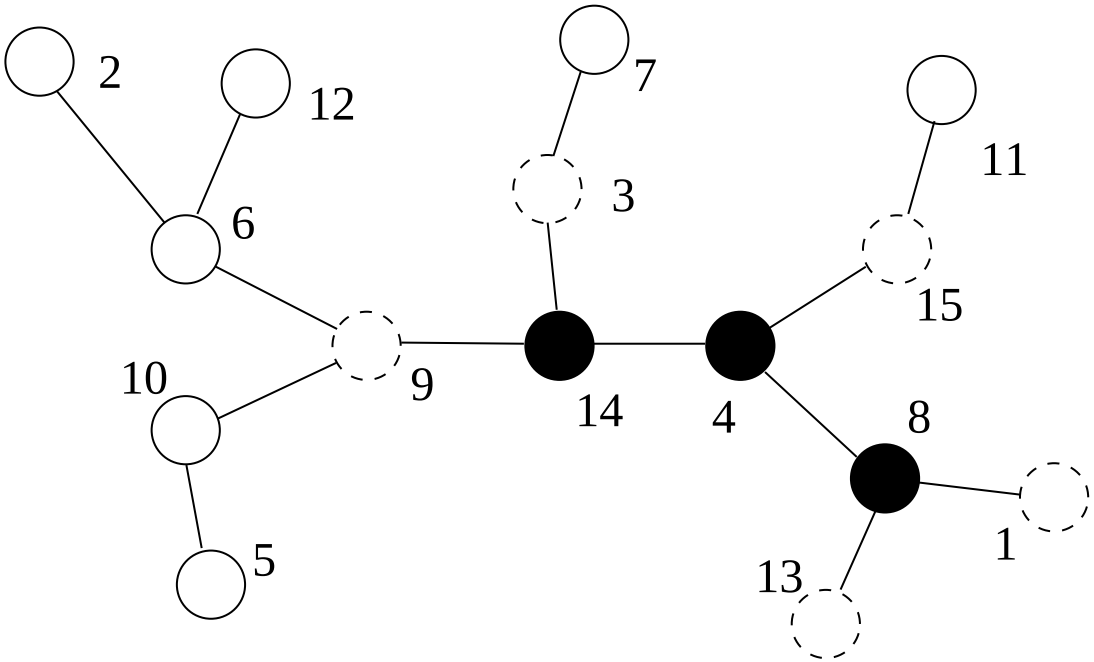
Patterns in Trees

Occurrence of a pattern $\mathcal{M}$ 



Patterns in Trees

Occurrence of a pattern $\mathcal{M}$  in a labelled tree



Patterns in Trees

Cayley's formula

$r_n = n^{n-1}$ ... number of **rooted** labelled trees with n nodes

$t_n = n^{n-2}$ ... number of labelled trees with n nodes

Generating functions

$$R(x) = \sum_{n\ge 1} r_n \frac{x^n}{n!}:$$

$$\boxed{R(x) = xe^{R(x)}}$$

$$T(x) = \sum_{n\ge 1} t_n \frac{x^n}{n!}:$$

$$\boxed{T(x) = R(x) - \frac{1}{2}R(x)^2}$$

(Note that $xT'(x) = R(x)$ so that we also have $T(x) = \int R(x)/x\,dx$.)

Patterns in Trees

Theorem

$\mathcal{M}$ ... be a given finite tree.

X_n ... number of occurrences of of $\mathcal{M}$ in a labelled tree of size n

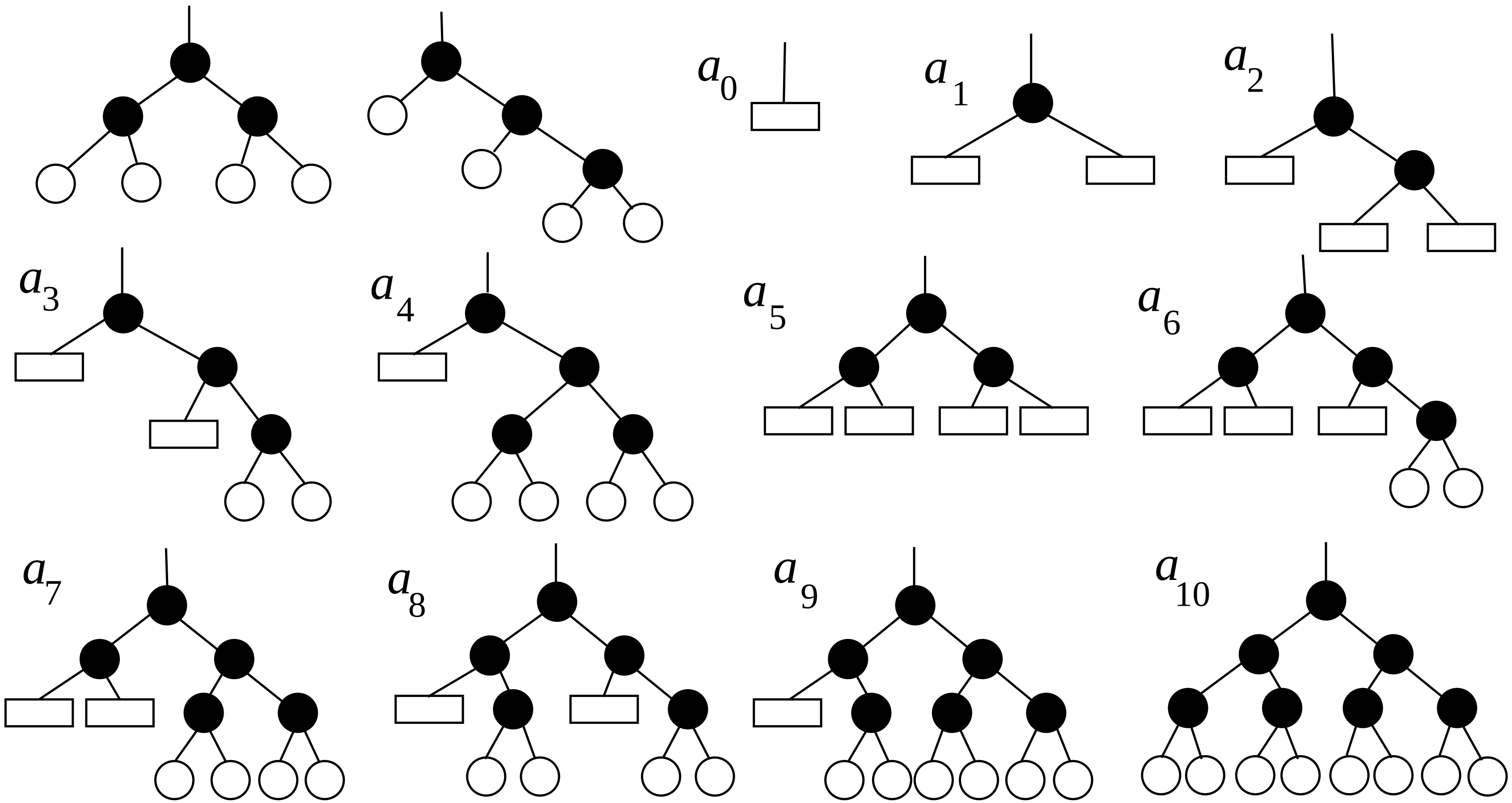
$\Longrightarrow$ X_n satisfies a **central limit theorem** with

$$\mathbb{E}\, X_n \sim \mu n \quad \text{and} \quad \mathbb{V}\, X_n \sim \sigma^2 n.$$

$\mu > 0$ and $\sigma^2 \geq 0$ depend on the pattern $\mathcal{M}$ and can be computed explicitly and algorithmically and can be represented as polynomials (with rational coefficients) in $1/e$.

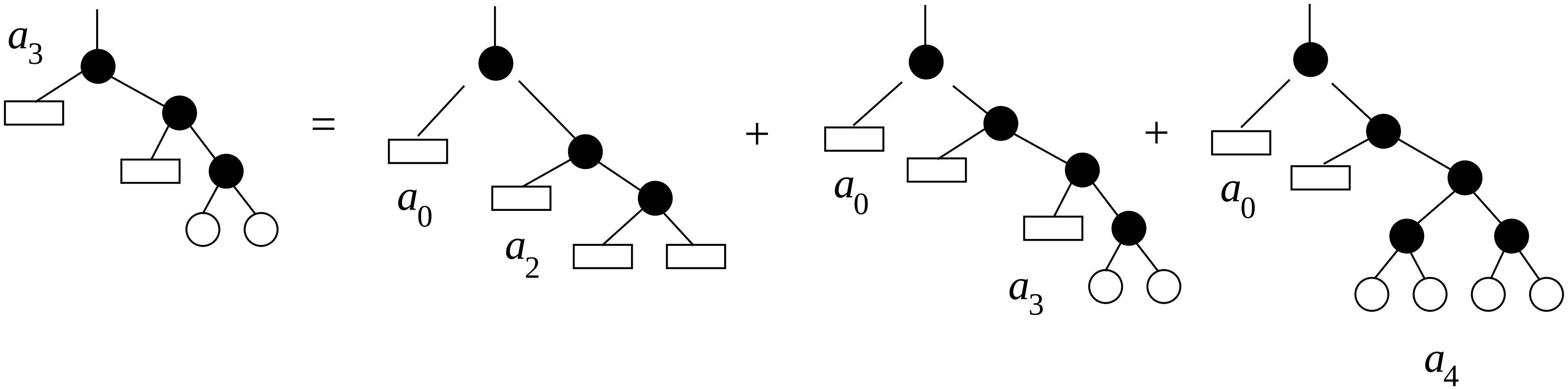
Patterns in Trees

Partition of trees in classes (▭ ... out-degree different from 2)



Patterns in Trees

Recurrences $\boxed{A_3 = xA_0A_2 + xA_0A_3 + xA_0A_4}$

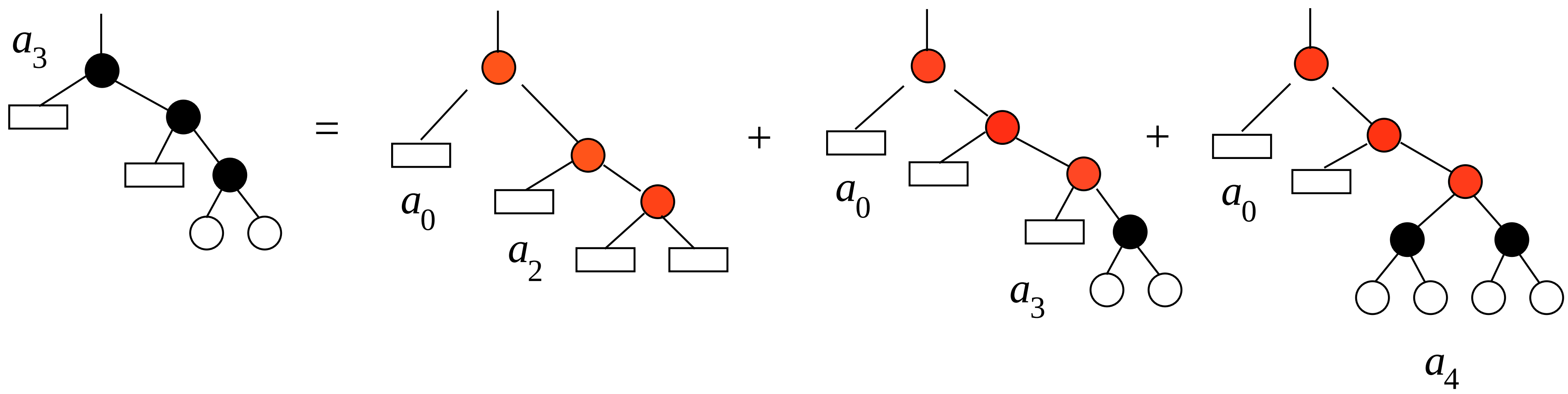


$$\boxed{A_j(x) = \sum_{n,k} a_{j;n} \frac{x^n}{n!}}$$

$a_{j;n}$ ... number of trees of size n in class j

Patterns in Trees

Recurrences $A_3 = xuA_0A_2 + xuA_0A_3 + xuA_0A_4$



$$A_j(x,u) = \sum_{n,k} a_{j;n,k} \frac{x^n}{n!} u^k$$

$a_{j;n,k}$ ... number of trees of size n in class j with k occurrences of $\mathcal{M}$

Patterns in Trees

$$
\begin{aligned}
A_0 &= A_0(x,u) = x + x\sum_{i=0}^{10} A_i + x\sum_{n=3}^{\infty}\frac{1}{n!}\left(\sum_{i=0}^{10} A_i\right)^n,\\
A_1 &= A_1(x,u) = \frac{1}{2}xA_0^2,\\
A_2 &= A_2(x,u) = xA_0A_1,\\
A_3 &= A_3(x,u) = xA_0(A_2+A_3+A_4)u,\\
A_4 &= A_4(x,u) = xA_0(A_5+A_6+A_7+A_8+A_9+A_{10})u^2,\\
A_5 &= A_5(x,u) = \frac{1}{2}xA_1^2u,\\
A_6 &= A_6(x,u) = xA_1(A_2+A_3+A_4)u^2,\\
A_7 &= A_7(x,u) = xA_1(A_5+A_6+A_7+A_8+A_9+A_{10})u^3,\\
A_8 &= A_8(x,u) = \frac{1}{2}x(A_2+A_3+A_4)^2u^3,\\
A_9 &= A_9(x,u) = x(A_2+A_3+A_4)(A_5+A_6+A_7+A_8+A_9+A_{10})u^4,\\
A_{10} &= A_{10}(x,u) = \frac{1}{2}x(A_5+A_6+A_7+A_8+A_9+A_{10})^2u^5.
\end{aligned}
$$

Systems of Functional equations

COMBINATORIAL CENTRAL LIMIT THEOREM II

Suppose that a sequence of random variables X_n has distribution

$$\mathbb{P}[X_n = k] = \frac{a_{nk}}{a_n},$$

where the generating function $A(x,u) = \sum_{n,k} a_{n,k} x^n u^k$ is given by

$$\boxed{A(x,u) = \Psi(x,u,A_1(x,u),\ldots,A_r(x,u))}$$

for an analytic function Ψ and the generating functions

$$A_1(x,u) = \sum_{n,k} a_{1;n,k} u^k x^n, \ \ldots \ , A_r(x,u) = \sum_{n,k} a_{r;n,k} u^k x^n$$

satisfy a **system of non-linear equations**

$$\boxed{A_j(x,u) = \Phi_j(x,u,A_1(x,u),\ldots,A_r(x,u))}, \quad (1 \le j \le r).$$

Systems of Functional equations

COMBINATORIAL CENTRAL LIMIT THEOREM II (cont.)

Suppose that at least one of the functions $\Phi_j(x, u, a_1, \ldots, a_r)$ is non-linear in $a_1, \ldots, a_r$ and they all have a power series expansion at $(0, 0, 0)$ with non-negative coefficients.

Let $x_0 > 0$, $\mathbf{a}_0 = (a_{0,0}, \ldots, a_{r,0}) > 0$ (inside the region of convergence) satisfy the system of equations: $(\mathbf{\Phi} = (\Phi_1, \ldots, \Phi_r))$

$$\boxed{\mathbf{a}_0 = \mathbf{\Phi}(x_0, 1, \mathbf{a}_0), \quad 0 = \det(\mathbb{I} - \mathbf{\Phi}_{\mathbf{a}}(x_0, 1, \mathbf{a}_0)}$$

such that the spectral radius of the Jacobian $\mathbf{\Phi}_{\mathbf{a}}$ equals 1. Suppose further, that the **dependency graph** of the system $\mathbf{a} = \mathbf{\Phi}(x, u, \mathbf{a})$ is **strongly connected** (which means that no subsystem can be solved before the whole system).

Systems of Functional equations

COMBINATORIAL CENTRAL LIMIT THEOREM II (cont.)

Then there exists analytic function $g_j(x,u), h_j(x,u)$, and $\rho(u)$ (that is **independent of** j) such that locally

$$\boxed{A_j(x,u) = g_j(x,u) - h_j(x,u)\sqrt{1-\frac{x}{\rho(u)}}}.$$

and consequently (for some $g(x,u)$, $h(x,u)$)

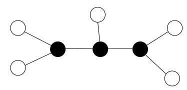
$$\boxed{A(x,u) = g(x,u) - h(x,u)\sqrt{1-\frac{x}{\rho(u)}}}.$$

Consequently the random variable X_n satisfies a **central limit theorem** with

$$\boxed{\mathbb{E}\, X_n \sim n\mu} \quad \text{and} \quad \boxed{\mathbb{V}\mathrm{ar}\, X_n \sim n\sigma^2},$$

where μ and σ^2 can be computed.

Patterns in Trees

Final Result for $\mathcal{M} =$ 

Central limit theorem with

$$\mu = \frac{5}{8e^3} = 0.0311169177\ldots$$

and

$$\sigma^2 = \frac{20e^3 + 72e^2 + 84e - 175}{32e^6} = 0.0764585401\ldots.$$

Contents 3

III. CONTINUOUS LIMITING OBJECTS

- Weak Convergence
- The Depth-First-Search of Rooted Trees
- The Continuum Random Tree
- The Profile of Galton-Watson trees
- Scaling Limit of Series-Parallel Graphs

Asymptotics on Random Discrete Objects

Levels of complexity:

1. **Asymptotic enumeration**

2. **Distribution of (shape) parameters**

3. **Asymptotic shape (= continuous limiting object)**

Weak Convergence

X_n, X ... (real) random variables:

$$\boxed{X_n \xrightarrow{\text{d}} X} \quad :\Longleftrightarrow \quad \boxed{\lim_{n\to\infty} \mathbb{P}\{X_n \le x\} = \mathbb{P}\{X \le x\}}$$

for all points of continuity of $F_X(x) = \mathbb{P}\{X \le x\}$

$$\Longleftrightarrow \quad \boxed{\boxed{\lim_{n\to\infty} \mathbb{E}\, G(X_n) = \mathbb{E}\, G(X)}}$$

for all **bounded** continuous functionals $G : \mathbb{R} \to \mathbb{R}$

$$\Longleftrightarrow \quad \boxed{\lim_{n\to\infty} \mathbb{E}\, e^{itX_n} = \mathbb{E}\, e^{itX}}$$

for all real t
(Levy's criterion)

Weak Convergence

Polish space: (S, d) ... complete, separable, metric space

Examples: $\mathbb{R}$, $\mathbb{R}^k$, $C[0,1]$, $\mathcal{M}_0(X)$ (probability measures on X)

S-**valued random variable**: $\boxed{X : \Omega \to S}$ ... measurable function

$S = \mathbb{R}$: random variable

$S = \mathbb{R}^k$: k-dimensional random vector

$S = C[0,1]$: **stochastic process** $(X(t), 0 \leq t \leq 1)$

$S = \mathcal{M}_0(X)$: random measure

Weak Convergence

Definition

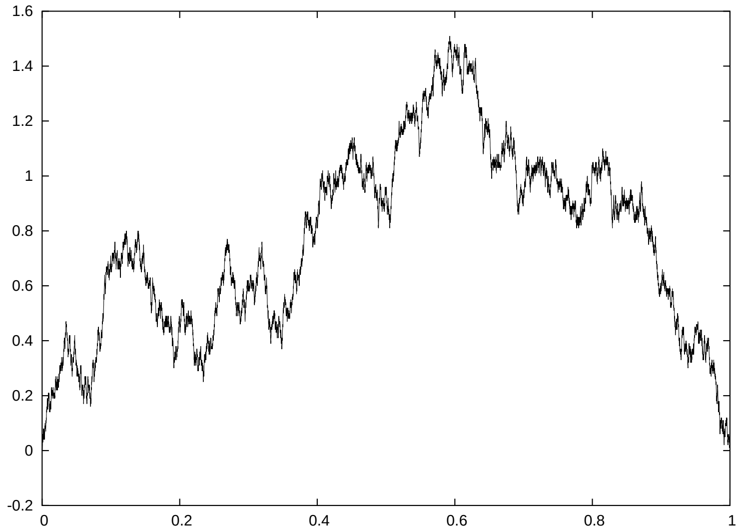
$X_n, X : \Omega \to S$ ... S-valued random variables ((S, d) ... Polish space)

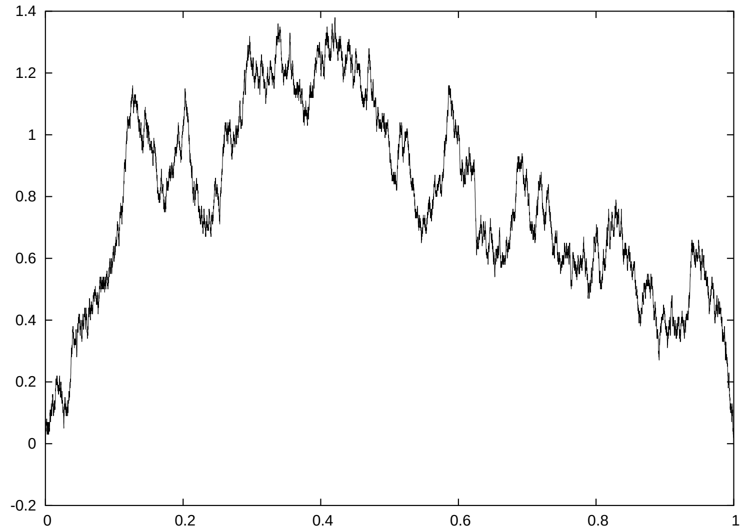
$$\boxed{X_n \xrightarrow{\text{d}} X} \quad :\Longleftrightarrow \quad \boxed{\boxed{\lim_{n\to\infty} \mathbb{E}\, G(X_n) = \mathbb{E}\, G(X)}}$$

for all **bounded** continuous functionals $G : S \to \mathbb{R}$

Weak Convergence

Stochastic process: random function





Weak Convergence

Stochastic process

$X_n : \Omega \to C[0,1]$ sequence of stochastic processes, $X : \Omega \to C[0,1]$

- $X_n \xrightarrow{\text{d}} X \quad \Longrightarrow \quad F(X_n) \xrightarrow{\text{d}} F(X)$ for all continuous $F : S \to S'$.

- $X_n \xrightarrow{\text{d}} X \quad \Longrightarrow \quad X_n(t_0) \xrightarrow{\text{d}} X(t_0)$ for all fixed $t_0 \in [0,1]$.

- $X_n \xrightarrow{\text{d}} X \quad \Longrightarrow \quad (X_n(t_1), \ldots, X_n(t_k)) \xrightarrow{\text{d}} (X(t_1), \ldots, X(t_k))$
  for all $k \geq 1$ and all fixed $t_1, \ldots, t_k \in [0,1]$.

The converse statement is not necessarily true, one needs **tightness**.

Weak Convergence

Stochastic process

$X_n : \Omega \to C[0,1]$ sequence of stochastic processes, $X : \Omega \to C[0,1]$

1. $(X_n(t_1), \dots, X_n(t_k)) \xrightarrow{\text{d}} (X(t_1), \dots, X(t_k))$
   for all $k \geq 1$ and all fixed $t_1, \dots, t_k \in [0,1]$

2. $\mathbb{E}\left(|X_n(0)|^\beta\right) \leq C$
   for some constant $C > 0$ and an exponent $\beta > 0$

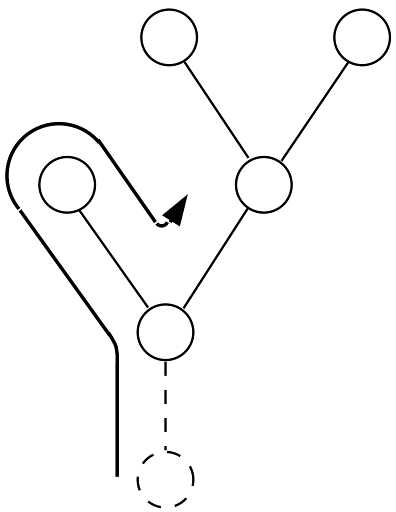
3. $\mathbb{E}\left(|X_n(t) - X_n(s)|^\beta\right) \leq C|t-s|^\alpha$ for all $s, t \in [0,1]$
   for some constant $C > 0$ and exponents $\alpha > 1$ and $\beta > 0$.

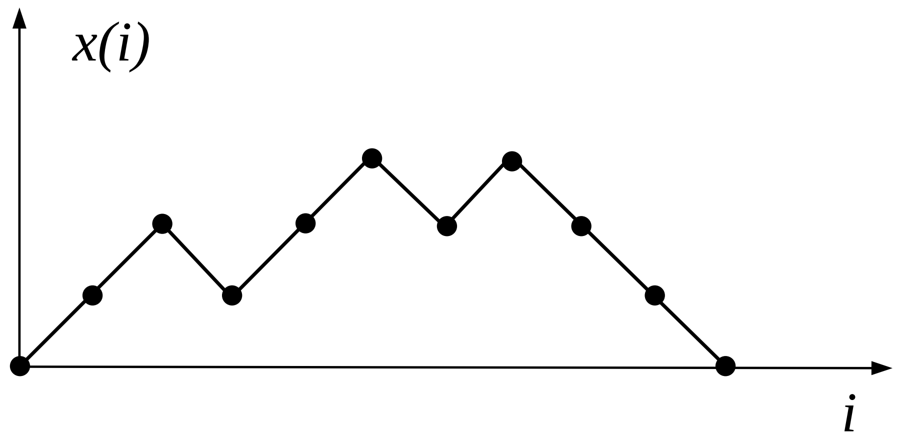
Then

$$\boxed{(X_n(t), 0 \leq t \leq 1) \xrightarrow{\text{d}} (X(t), 0 \leq t \leq 1)}.$$

Depth-First-Search

Rooted trees and discrete excursions





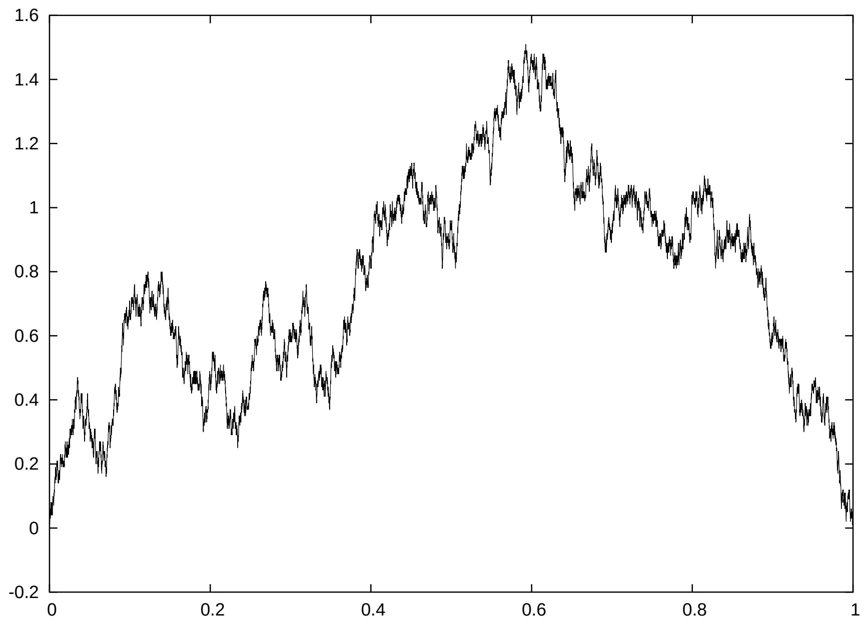
Bijection between

Catalan trees $\longleftrightarrow$ Dyck paths

random trees of size n $\longleftrightarrow$ **random** Dyck paths of length $2n$

Depth-First-Search

Brownian excursion $(e(t), 0 \leq t \leq 1)$



Rescaled Brownian motion between 2 zeros.

Random function in $C[0,1]$.

Depth-First-Search

Kaigh's Theorem

$(X_n(t), 0 \le t \le 2n)$ ... random Dyck path of length $2n$.

$$\Longrightarrow \quad \left(\frac{1}{\sqrt{2n}} X_n(2nt), 0 \le t \le 1\right) \xrightarrow{\text{d}} (2e(t), 0 \le t \le 1).$$

Remark. This theorem also holds for more general random walks with independent increments conditioned to be an excursion.

Real Trees

T ... tree, $\mathcal{T}$ ... embedding of T into the plane $\mathbb{R}^2$

$\Longrightarrow$ $\mathcal{T}$ is a metric space (and a **real tree** in the following sense):

Definition

A metric space $(\mathcal{T}, d)$ is a **real tree** if the following two properties hold for every $x, y \in \mathcal{T}$.

1. There is a unique isometric map $h_{x,y} : [0, d(x,y)] \to \mathcal{T}$ such that $h_{x,y}(0) = x$ and $h_{x,y}(d(x,y)) = y$.

2. If q is a continuous injective map from $[0,1]$ into $\mathcal{T}$ with $q(0) = x$ and $q(1) = y$ then
$$q([0,1]) = h_{x,y}([0, d(x,y)]).$$

A rooted real tree $(\mathcal{T}, d)$ is a real tree with a distinguished vertex $r = r(\mathcal{T})$ called the root.

Real Trees

Two real trees $(\mathcal{T}_1, d_1)$, $(\mathcal{T}_2, d_2)$ are **equivalent** if there is a root-preserving isometry that maps $\mathcal{T}_1$ onto $\mathcal{T}_2$.

$\mathbb{T}$ ... set of all equivalence classes of rooted compact real trees.

Gromov-Hausdorff Distance $d_{\mathsf{GH}}(\mathcal{T}_1, \mathcal{T}_2)$ of two real trees $\mathcal{T}_1, \mathcal{T}_2$ is the infimum of the Hausdorff distance of all isometric embeddings of $\mathcal{T}_1, \mathcal{T}_2$ into the same metric space.

Hausdorff distance: $\delta_{\mathsf{Haus}}(X, Y) = \max\left\{\sup_{x\in X}\inf_{y\in Y} d(x,y), \sup_{y\in Y}\inf_{x\in X} d(x,y)\right\}$

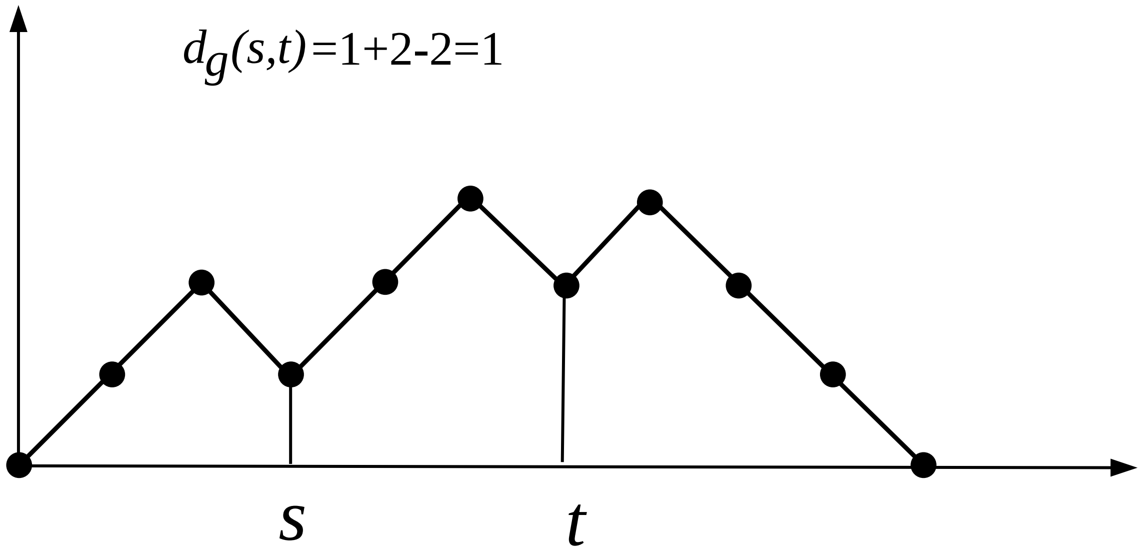
Theorem

The metric space $(\mathbb{T}, d_{\mathsf{GH}})$ is a Polish space.

Real Trees

$g : [0,1] \to [0,\infty)$ ... continuous, ≥ 0, $g(0) = g(1) = 0$

$$d_g(s,t) = g(s) + g(t) - 2 \inf_{\min\{s,t\} \leq u \leq \max\{s,t\}} g(u)$$

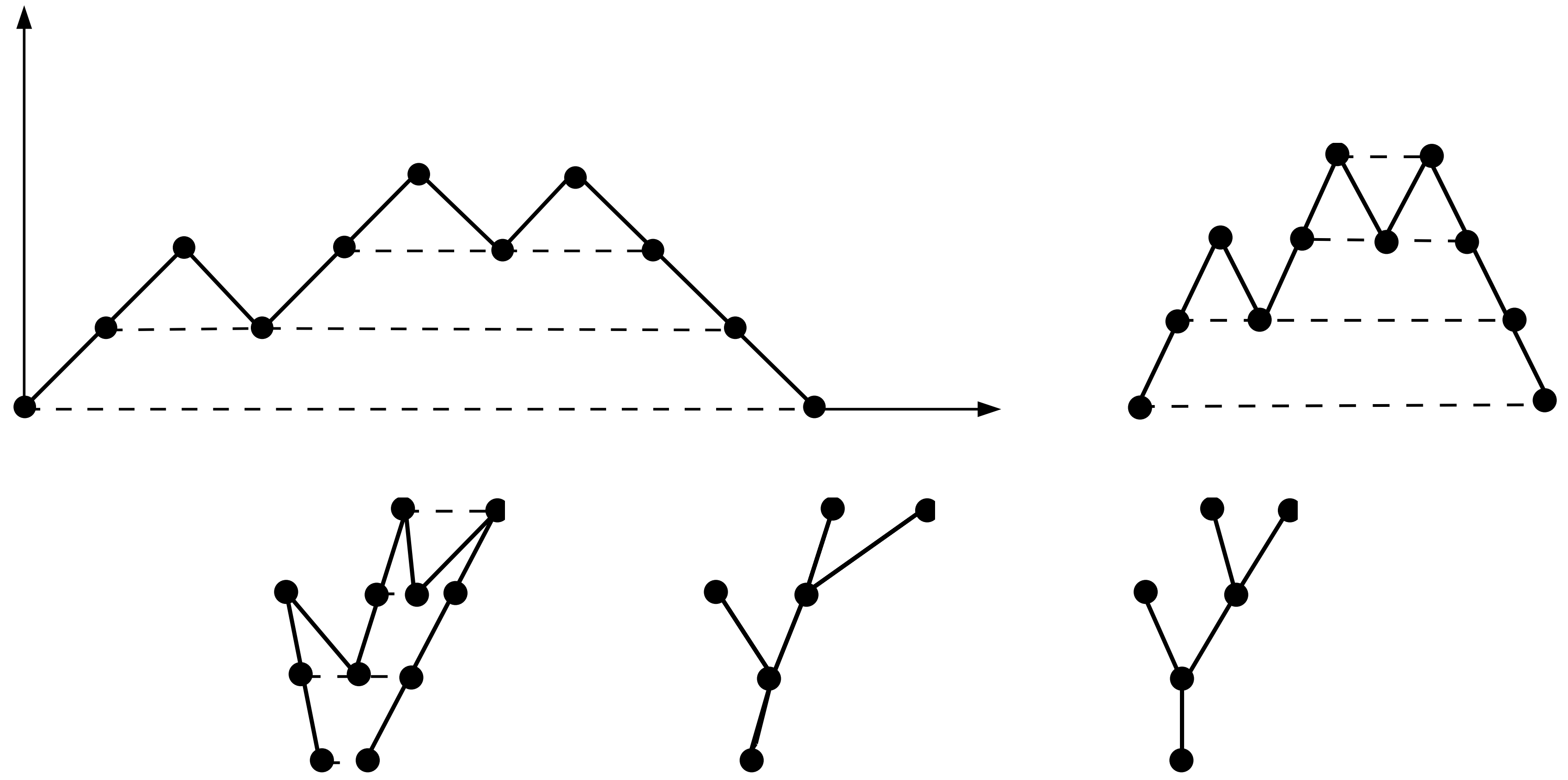


$$s \sim t \iff d_g(s,t) = 0 \qquad \mathcal{T}_g = [0,1]/\sim$$

$$\Longrightarrow \quad (\mathcal{T}_g, d_g) \quad \text{is a compact real tree.}$$

Real Trees

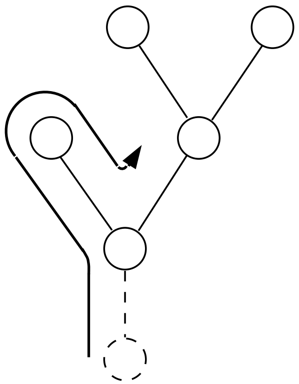
Construction of a real tree $\mathcal{T}_g$



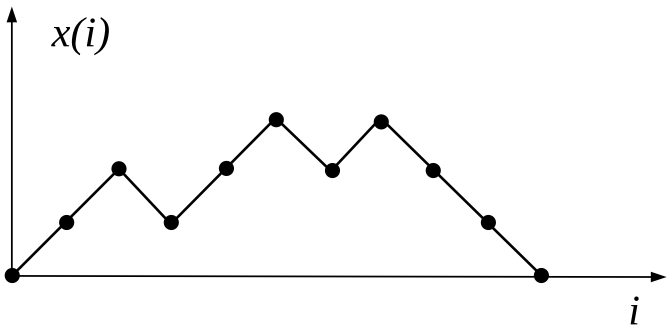
The mapping $C[0,1] \to \mathbb{T}$, $g \mapsto \mathcal{T}_g$ is **continuous**.

Real Trees

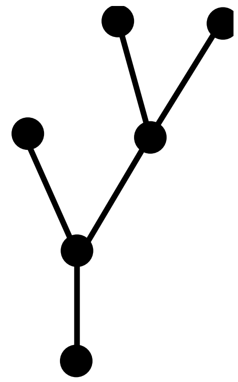
Catalan trees as real trees



T_n



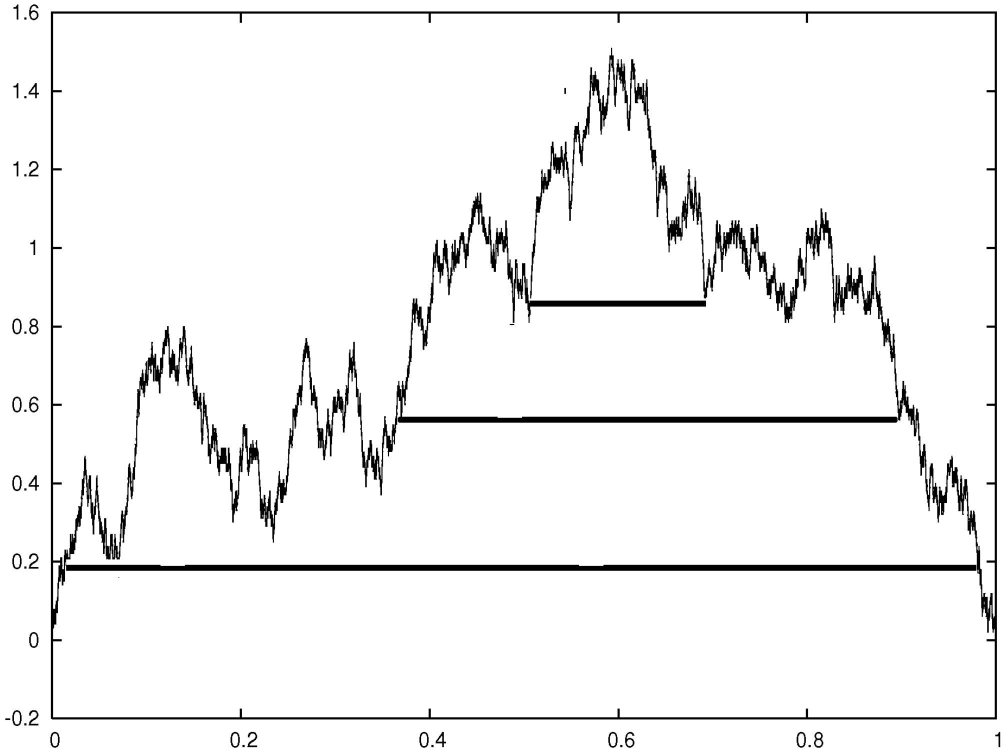
$X_n = X_{T_n}$



$\mathcal{T}_{X_n}$

Real Trees

Continuum random tree $\mathcal{T}_{2e}$ (with Brownian excursion $e(t)$)



Real Trees

Theorem

$(X_n(t), 0 \le t \le 2n)$ ... random Dyck paths of length $2n$
or the depth-first-search process of Catalan trees of size n.

$$\Longrightarrow \quad \boxed{\frac{1}{\sqrt{2n}} \mathcal{T}_{X_n} \xrightarrow{\text{d}} \mathcal{T}_{2e}}$$

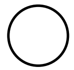
In other words...

Scaled Catalan trees (interpreted as "real trees") converge weakly to the continuum random tree.

Galton-Watson Trees

Galton-Watson branching process

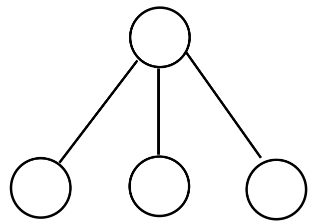
ξ ... offspring distribution, $\varphi_k = \mathbb{P}\{\xi = k\}$, $\varphi_0 > 0$



Galton-Watson Trees

Galton-Watson branching process

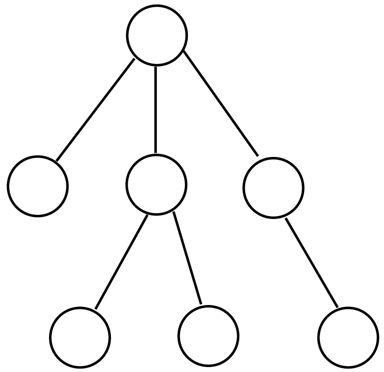
ξ ... offspring distribution, $\varphi_k = \mathbb{P}\{\xi = k\}$, $\varphi_0 > 0$



Galton-Watson Trees

Galton-Watson branching process

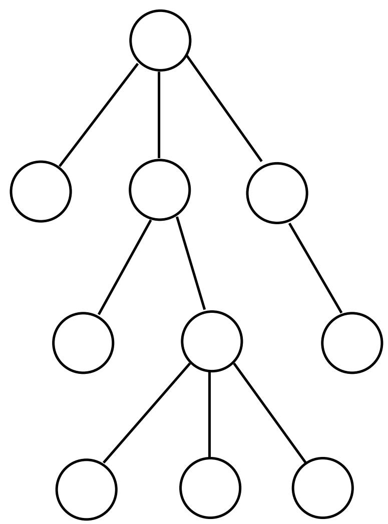
ξ ... offspring distribution, $\varphi_k = \mathbb{P}\{\xi = k\}$, $\varphi_0 > 0$



Galton-Watson Trees

Galton-Watson branching process

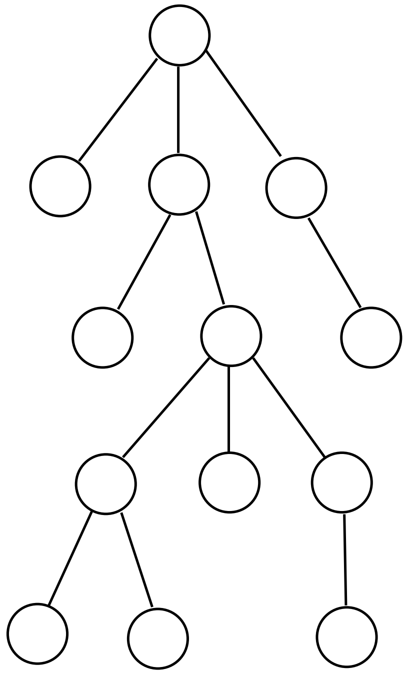
ξ ... offspring distribution, $\varphi_k = \mathbb{P}\{\xi = k\}$, $\varphi_0 > 0$



Galton-Watson Trees

Galton-Watson branching process

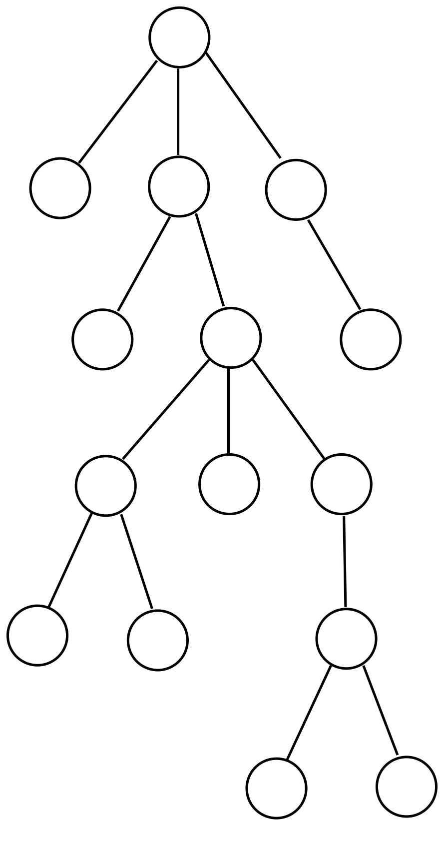
ξ ... offspring distribution, $\varphi_k = \mathbb{P}\{\xi = k\}$, $\varphi_0 > 0$



Galton-Watson Trees

Galton-Watson branching process

ξ … offspring distribution, $\varphi_k = \mathbb{P}\{\xi = k\}$, $\varphi_0 > 0$



Galton-Watson Trees

Galton-Watson branching process. $(Z_k)_{k\ge 0}$

$Z_0 = 1$, and for $k \ge 1$

$$\boxed{Z_k = \sum_{j=1}^{Z_{k-1}} \xi_j^{(k)}},$$

where the $(\xi_j^{(k)})_{k,j}$ are iid random variables distributed as ξ.

Z_k ... number of nodes in k-th generation

$Z = Z_0 + Z_1 + Z_2 + \cdots$ ... total progeny

Galton-Watson Trees

Generating functions

$$y_n = \mathbb{P}\{Z = n\}, \qquad y(x) = \sum_{n \geq 1} y_n x^n$$

$$\Phi(w) = \mathbb{E}\, w^{\xi} = \sum_{k \geq 0} \varphi_k w^k$$

$$\Longrightarrow \quad \boxed{y(x) = x\,\Phi(y(x))}$$

Conditioned Galton-Watson tree

GW-branching process conditioned on the total progeny $Z = n$.

Galton-Watson Trees

Example. $\mathbb{P}\{\xi = k\} = 2^{-k-1}$, $\Phi(w) = 1/(2-w)$

$\Longrightarrow$ all trees of size n have the same probability

$\Longrightarrow$ conditioned GW-tree of size n is the same model as the **Catalan tree model** (with the uniform distribution on trees of size n)

Example. $\Phi(w) = \frac{1}{2}(1+w)^2$: **binary trees** with n internal nodes.

Example. $\Phi(w) = \frac{1}{3}(1+w+w^2)$: **Motzkin trees**

Example. $\Phi(w) = e^{w-1}$: **Cayley trees**

Galton-Watson Trees

General assumption: $\boxed{\mathbb{E}\,\xi = 1}$, $\boxed{0 < \mathbb{V}\mathrm{ar}\,\xi = \sigma^2 < \infty}$

Theorem (Aldous)

$X_n(t)$ ... depth-first-search of conditioned GW-trees of size n

$$\Longrightarrow \quad \boxed{\left(\frac{\sigma}{2\sqrt{n}} X_n(2nt), 0 \le t \le 1\right) \xrightarrow{\mathrm{d}} (e(t), 0 \le t \le 1)}.$$

Corollary

$$\boxed{\frac{\sigma}{\sqrt{n}}\,\mathcal{T}_{X_n} \xrightarrow{\mathrm{d}} \mathcal{T}_{2e}}$$

Galton-Watson Trees

Corollary H_n ... height of conditioned GW-trees of size n:

$$\Longrightarrow \quad \boxed{\frac{1}{\sqrt{n}} H_n \xrightarrow{\mathrm{d}} \frac{2}{\sigma} \max_{0 \le t \le 1} e(t)}$$

Remark. Distribution function of $\max_{0 \le t \le 1} e(t)$:

$$\mathbb{P}\{\max_{0 \le t \le 1} e(t) \le x\} = 1 - 2 \sum_{k=1}^{\infty} (4x^2k^2 - 1) e^{-2x^2k^2}$$

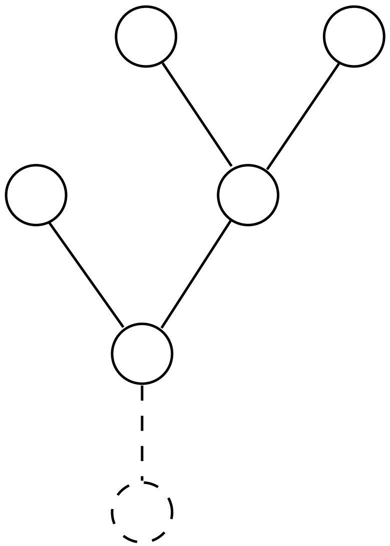
Galton-Watson Trees

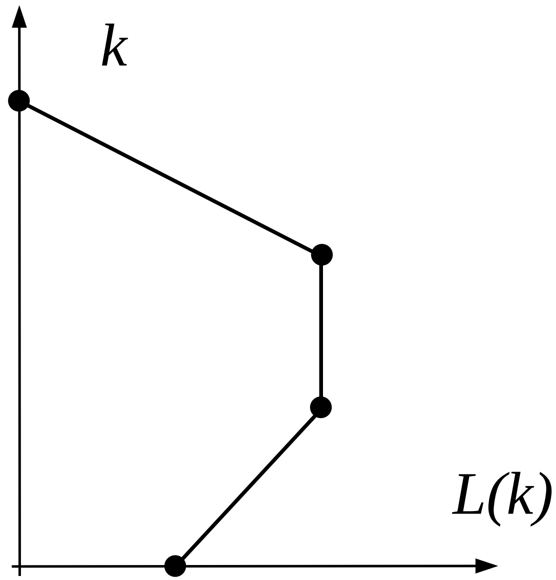
Profile

$L_T(k)$ ... number of nodes at distance k from the root

$(L_T(k))_{k \geq 0}$ ... profile of T

$(L_T(s), s \geq 0)$ ... linearly interpolated profile of T





Galton-Watson Trees

Value distribution

$$\mu_T = \frac{1}{|T|} \sum_{k \geq 0} L_T(k)\, \delta_k$$

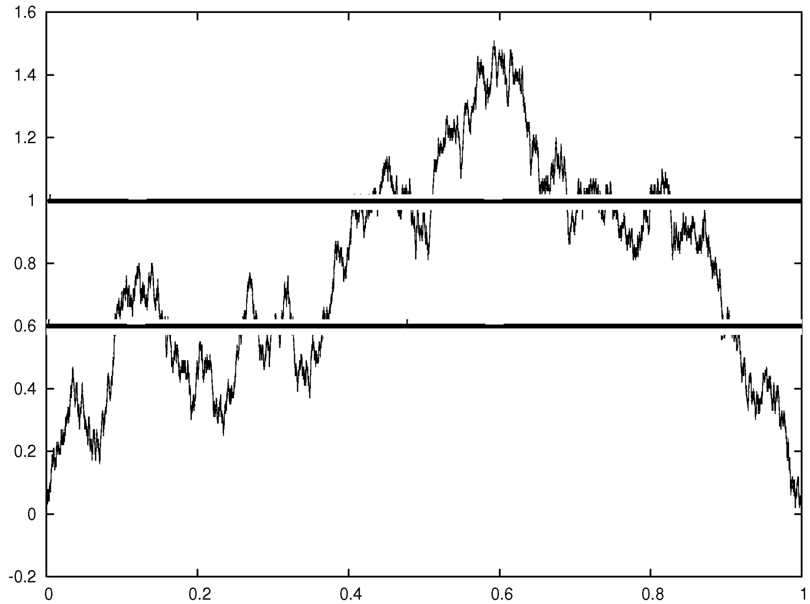
δ_x ... δ-distribution concentrated at x

Galton-Watson Trees

Occupation measure: random measure on $\mathbb{R}$

$$\boxed{\mu(A) = \int_0^1 \mathbf{1}_A(e(t)\,dt}$$

measure how long $e(t)$ stays in set A



Galton-Watson Trees

Theorem (Aldous)

$(L_n(k), k \geq 0)$ ... random profile of conditioned GW-trees of size n

$$\Longrightarrow \quad \boxed{\frac{1}{n}\sum_{k\geq 0} L_n(k)\,\delta_{(\sigma/2)k/\sqrt{n}} \xrightarrow{\text{d}} \mu}$$

Galton-Watson Trees

Local time of the Brownian excursion: random density of μ

$$l(s) = \lim_{\varepsilon\to 0} \frac{1}{\varepsilon} \int_0^1 \mathbf{1}_{[s,s+\varepsilon]}(e(t))\, dt$$

Theorem (D.+Gittenberger)

$(L_n(s), s \ge 0)$ ... random profile of conditioned GW-trees of size n

$$\Longrightarrow \quad \left(\frac{1}{\sqrt{n}} L_n(s\sqrt{n}),\ s \ge 0 \right) \xrightarrow{\mathrm{d}} \left(\frac{\sigma}{2} l\left(\frac{\sigma}{2} s\right),\ s \ge 0 \right)$$

Proof with asymptotics on generating functions (very involved)!!!

Galton-Watson Trees

Width

$$W = \max_{k \geq 0} L(k) = \max_{t \geq 0} L(t),$$

maximal number of nodes in a level.

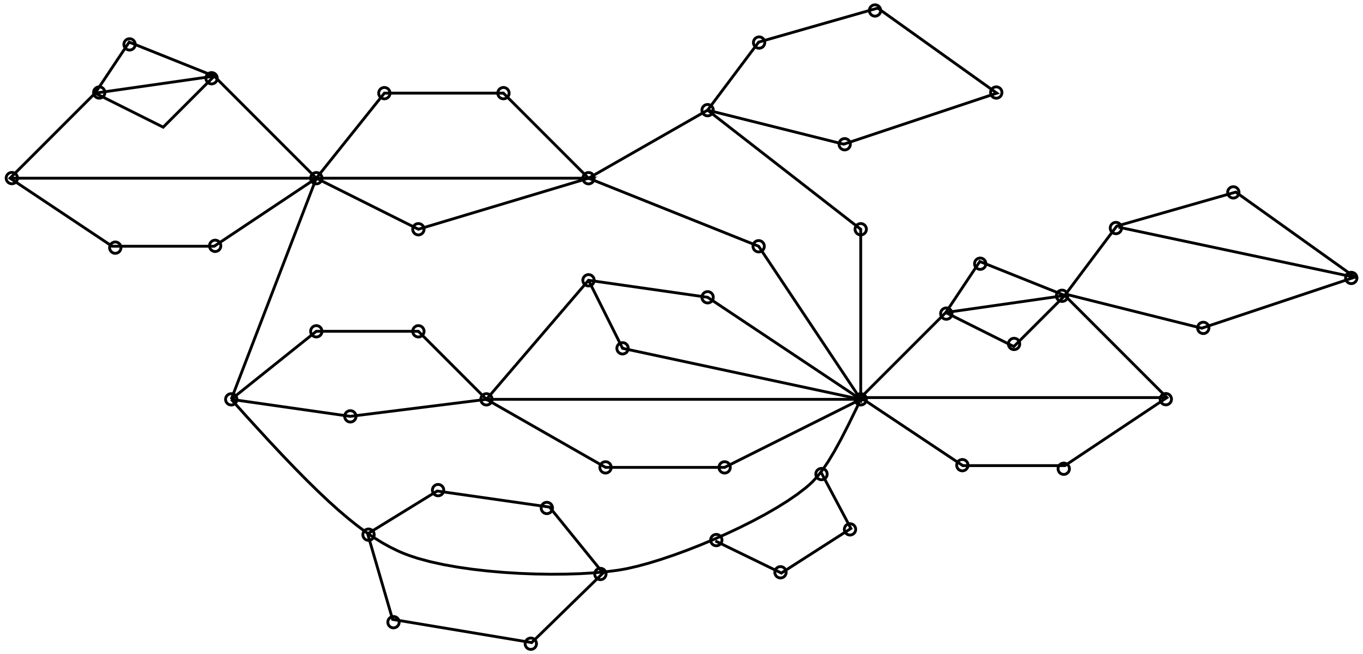
Corollary

$$\boxed{\frac{1}{\sqrt{n}} W_n \xrightarrow{\text{d}} \frac{\sigma}{2} \sup_{0 \leq t \leq 1} l(t)}$$

Remark. $\sup_{t \geq 0} l(t) = 2 \sup_{0 \leq t \leq 1} e(t)$ (in distribution)

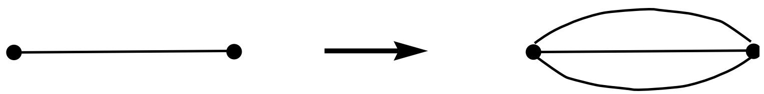
Series-Parallel Graphs

Connected Series-Parallel Graphs



Series-parallel extension of a tree (or no K_4 as a minor)

Series-extension: 

Parallel-extension: 

Scaling Limit of Series Parallel Graphs

A typcial series-parallel graph of size n has $\approx c_1 n$ 2-connected components that form a **tree**

The 2-connected components do not scale in distribution, their expected size is finite and they behave *almost) independent and identically distributed.*

So, series-parallel graphs look **tree-like***.*

Scaling Limit of Series Parallel Graphs

Theorem *(Panagiotou, Stufler, and Weller)*

C_n ... *connected, vertex labelled series-parallel graphs with* n *vertices*

$$\boxed{\frac{c}{\sqrt{n}} C_n \xrightarrow{\mathrm{d}} \mathcal{T}_{2e}}$$

for some constant $c > 0$.

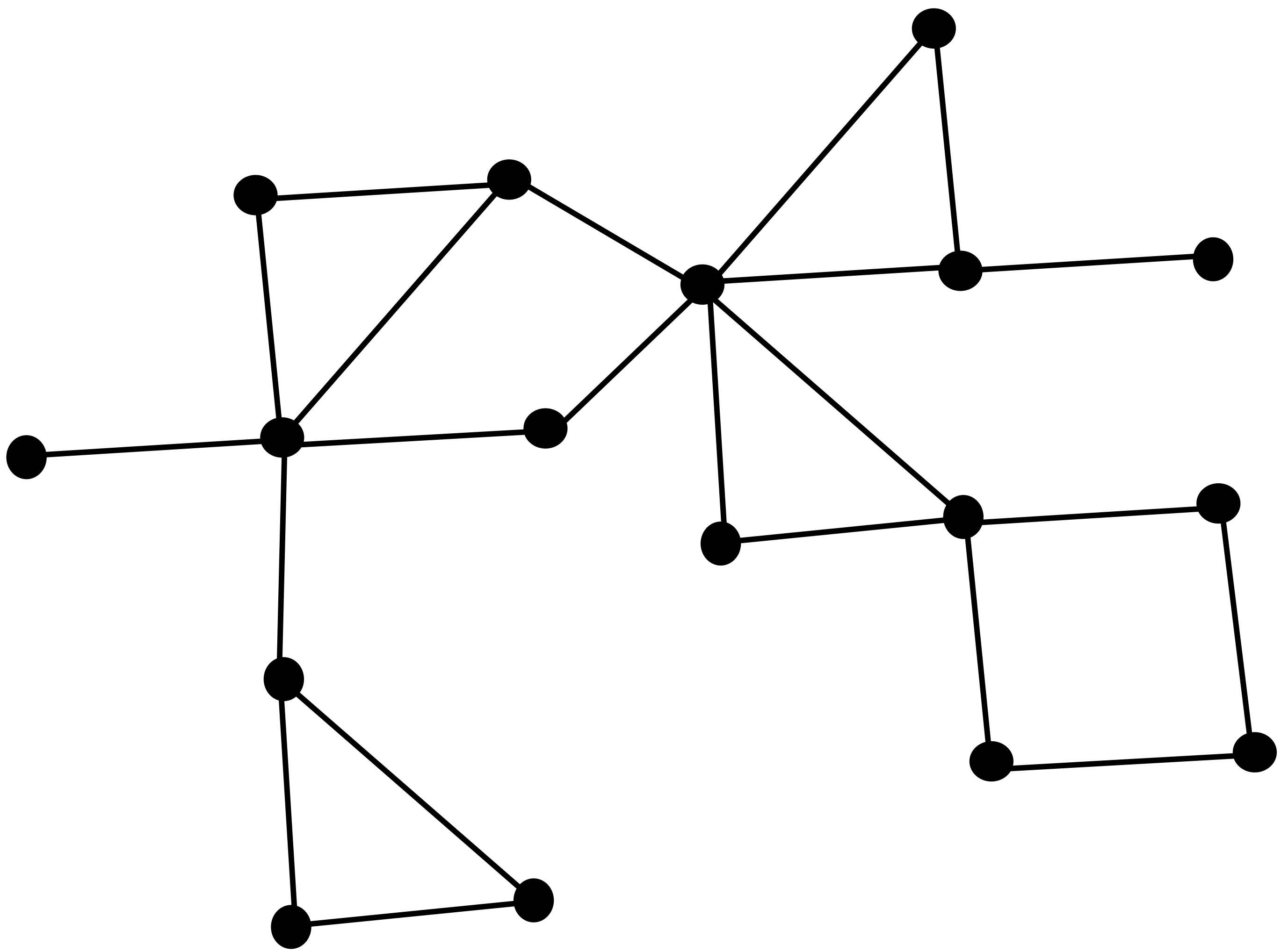
Remark. *The same result holds for so-called* **subcricital graph classes** *like cacti-graphs, outerplanar graphs etc. In all these graph classes the diameter is of oder* $\sqrt{n}$.

Contents 4

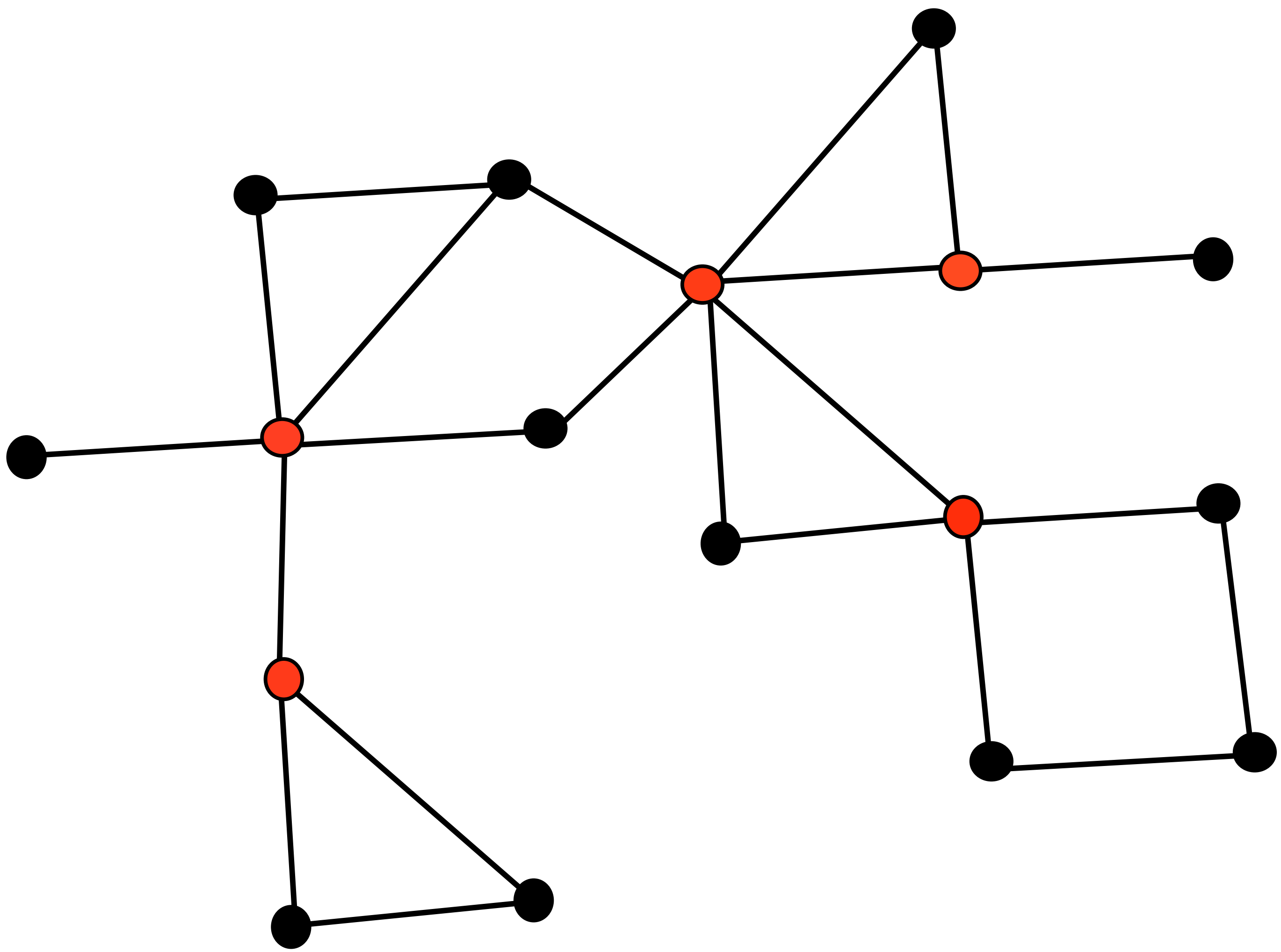
IV. SUBGRAPH COUNTS IN SERIES PARALLEL GRAPHS

- *Sub-critical graph classes*
- *Asymptotic counting of sub-critical graph classes*
- *Series parallel graphs are sub-critical*
- *Subgraph counting*
- *A combinatorial CLT for infinite systems*

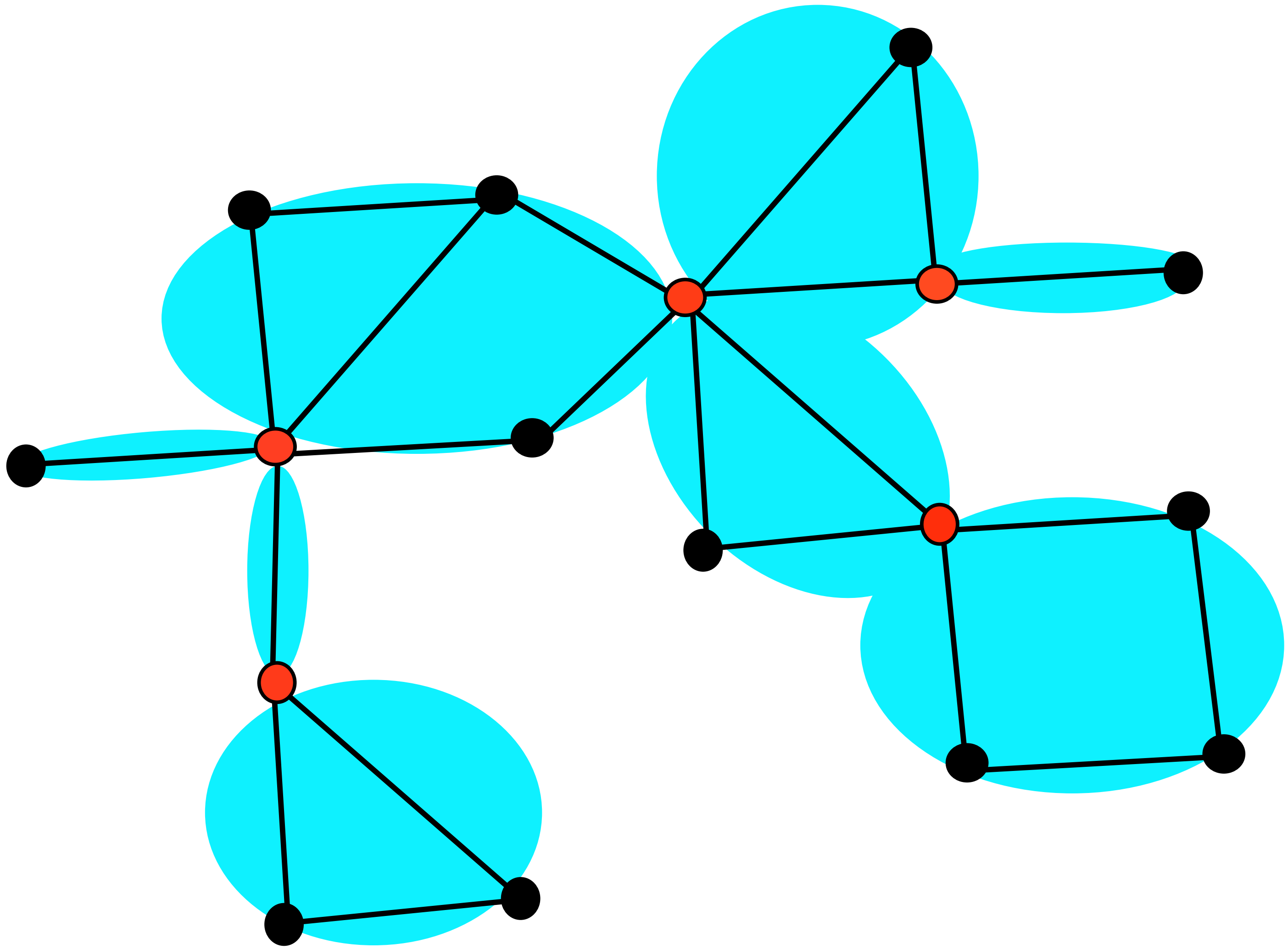
Block-Decomposition



Block-Decomposition



Block-Decomposition



Block-Decomposition

block: *2-connected component (= maximal 2-connected subgraph)*

Block-stable graph class $\mathcal{G}$: *$\mathcal{G}$ contains the one-edge graph and $G \in \mathcal{G}$ if and only if all blocks of G are contained in $\mathcal{G}$.*

Equivalently, the 2-connected graphs of $\mathcal{G}$ and the one-edge graph generate all graphs of $\mathcal{G}$.

Examples: *Planar graphs, series-parallel graphs, minor-closed graph classes etc.*

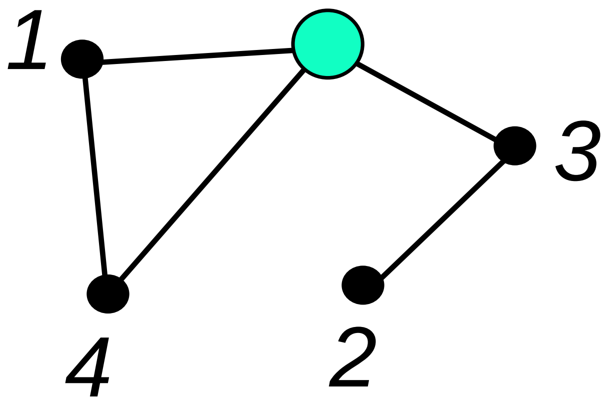
$B(x)$ ... *GF for 2-connected graphs in $\mathcal{G}$*

$C(x)$ ... *GF for connected graphs in $\mathcal{G}$*

[We will consider here only connected graphs]

Generating Functions for Block-Decomposition

Vertex-rooted graphs: *one vertext (the* **root***) is distinguished (and usually discounted, that is, it gets no label)*

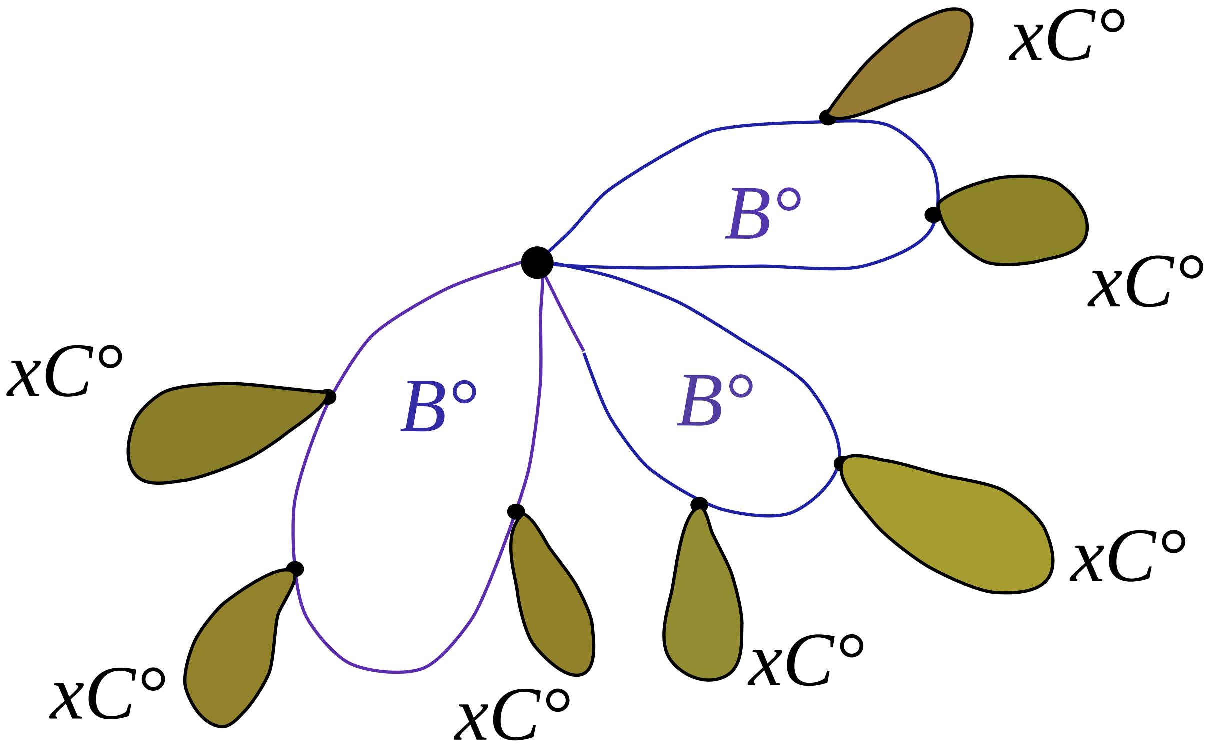


Generating function: *(in den* **labelled** *case)*

$$\boxed{G^{\bullet}(x) = G'(x)}$$

Generating Functions for Block-Decomposition

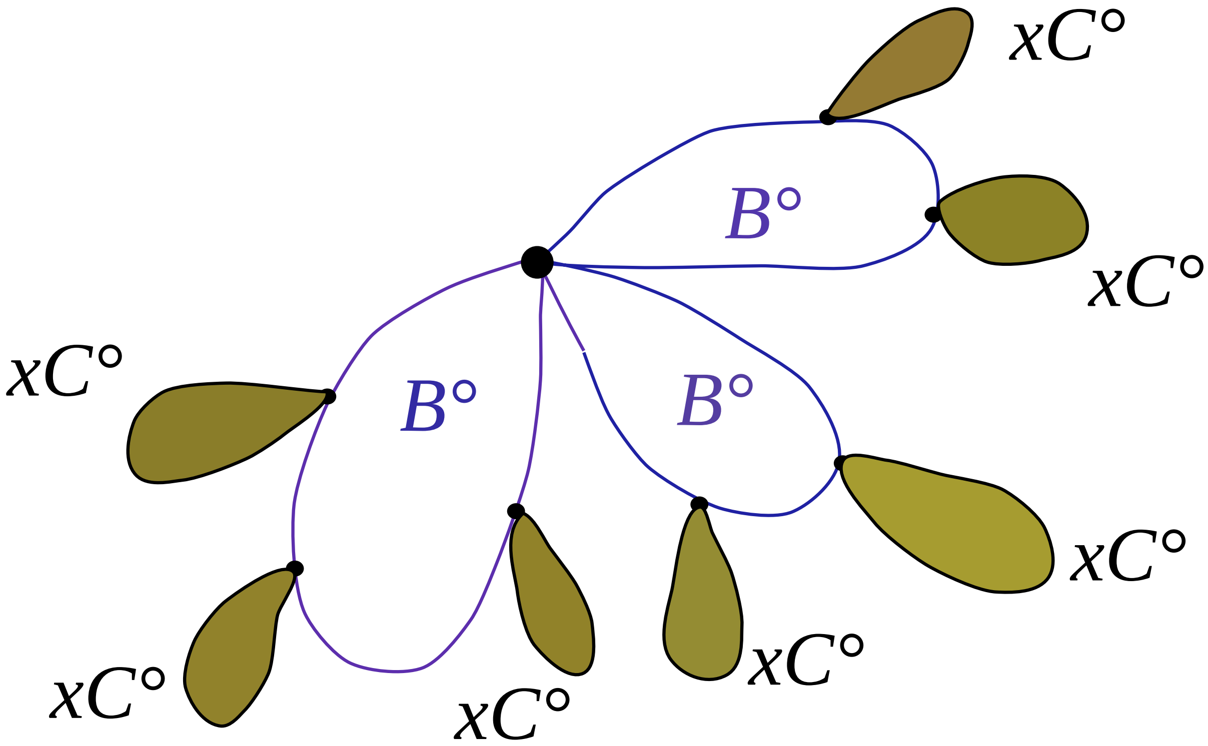
(in the **labelled** *case)*



$$C^{\bullet}(x) = e^{B^{\bullet}(xC^{\bullet}(x))}$$

Generating Functions for Block-Decomposition

(in the **labelled** *case)*



$$\frac{\partial C(x,y)}{\partial x} = \exp\left(\frac{\partial B}{\partial x}\left(x\frac{\partial C(x,y)}{\partial x}, y\right)\right)$$

Labelled Trees

Rooted Trees:

$$B^\bullet(x) = x$$

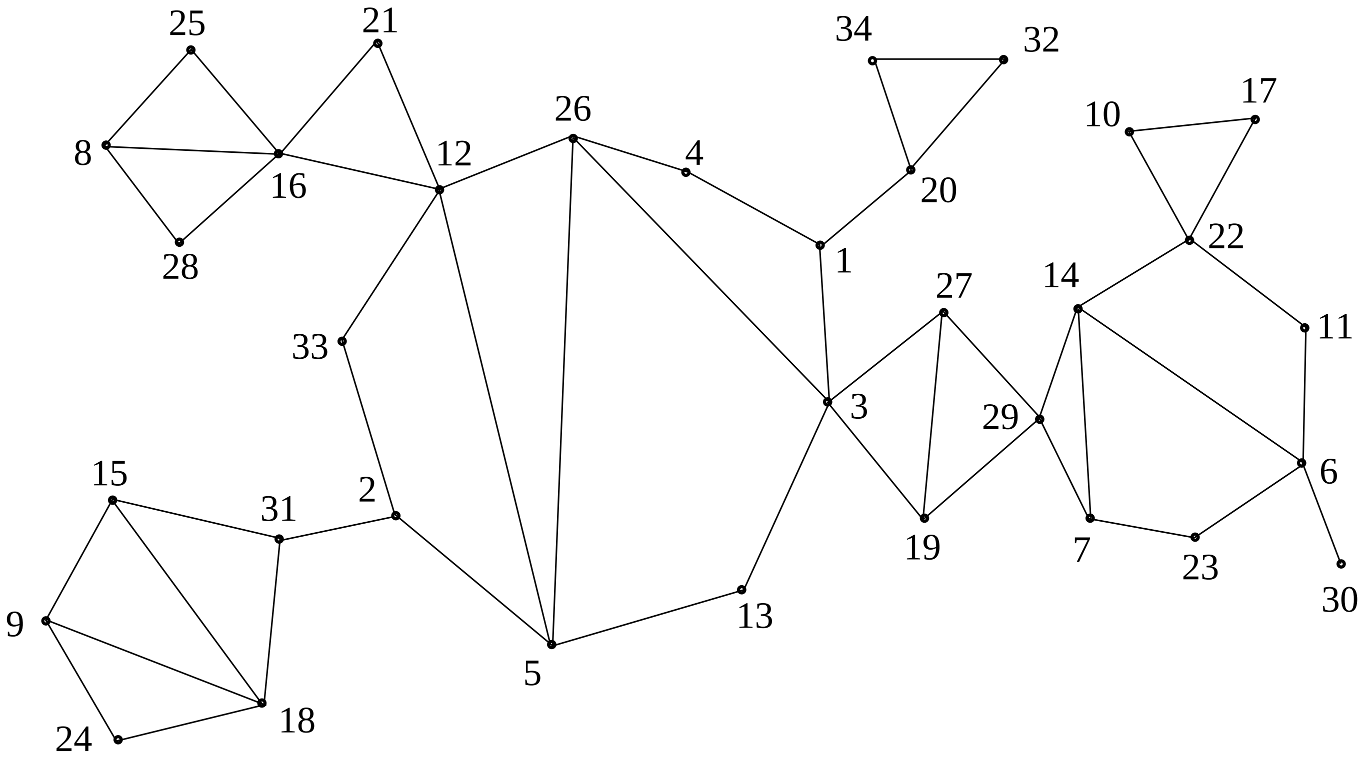
$R(x) = xC^\bullet(x)$ ... *generating function of* **rooted, labelled trees**

$$\boxed{C^\bullet(x) = e^{B^\bullet(xC^\bullet(x))}} \Longrightarrow \boxed{R(x) = xe^{R(x)}}$$

Remark: $T(x)$ ... *GF for unrooted labelled trees:*

$$T(x)' = \frac{1}{x}R(x) \quad \Longrightarrow \quad T(x) = R(x) - \frac{1}{2}R(x)^2$$

Outerplanar Graphs



All vertices are on the infinite face.

Outerplanar Graphs

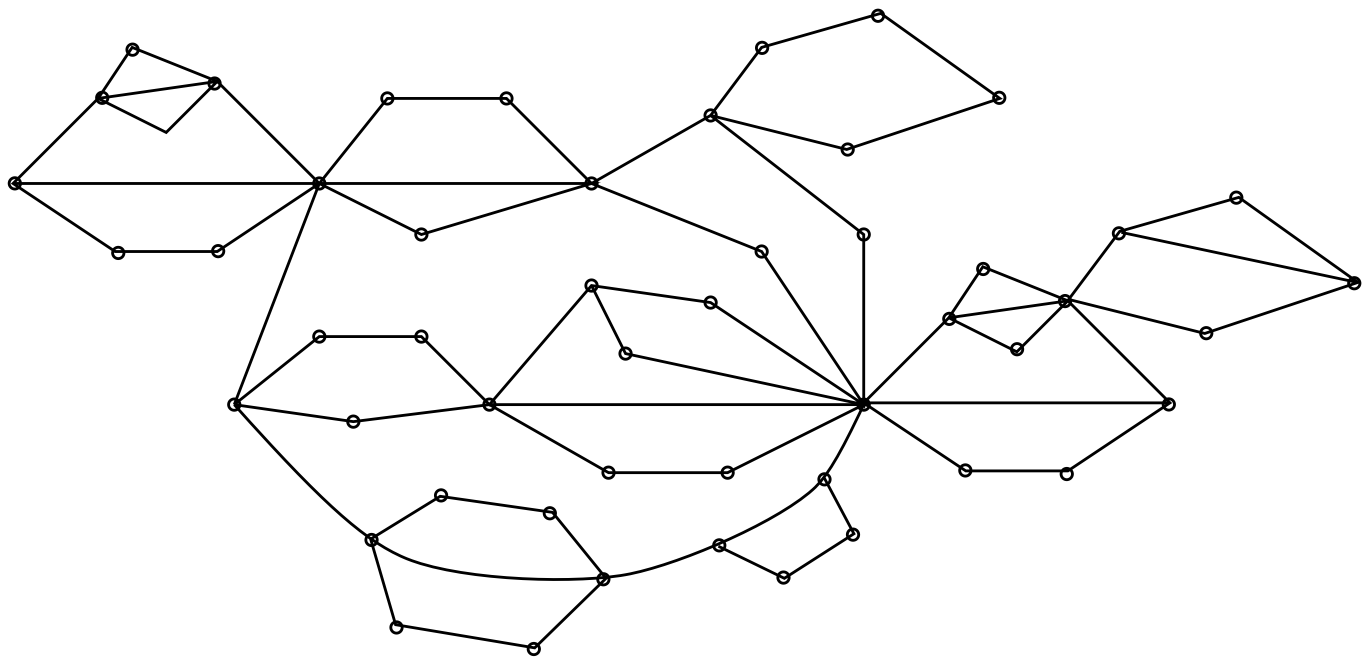
Generating functions

$$\boxed{C^\bullet(x) = e^{B^\bullet(xC^\bullet(x))}},$$

$$B^\bullet(x) = \frac{1 + 5x - \sqrt{1 - 6x + x^2}}{8}.$$

2-connected outerplanar graphs = dissections of the n-gon

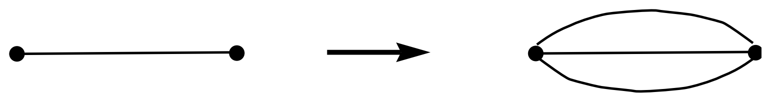
Series-Parallel Graphs



Series-parallel extension of a tree (if we restict to connected graphs)

Series-extension:

Parallel-extension:



Series-Parallel Graphs

Equivalent Definitions

- *Ex*(K_4)
- *tree-width* ≤ 2
- *nested ear decomposition (if connected)*

Series-Parallel Graphs

Generating functions

$$\boxed{\frac{\partial C(x,y)}{\partial x} = \exp\left(\frac{\partial B}{\partial x}\left(x\frac{\partial C(x,y)}{\partial x}, y\right)\right)},$$

$$\frac{\partial B(x,y)}{\partial y} = \frac{x^2}{2}\,e^{S(x,y)},$$

$$S(x,y) = \frac{x(P(x,y)+y)^2}{1-x(P(x,y)+y)},$$

$$P(x,y) = (e^{S(x,y)} - 1 - S(x,y)) + y(e^{S(x,y)} - 1).$$

Sub-critical Graphs

Repetition: Functional equations

Suppose that $\boxed{A(x) = \Phi(x, A(x))}$, *where* $\Phi(x, a)$ *has a power series expansion at* $(0, 0)$ *with non-negative coefficients and* $\Phi_{aa}(x, a) \neq 0$.

Let $x_0 > 0$, $a_0 > 0$ (**inside the region of convergence of** Φ) *satisfy the system of equations:*

$$\boxed{a_0 = \Phi(x_0, a_0), \quad 1 = \Phi_a(x_0, a_0)}.$$

Then there exists analytic function $g(x), h(x)$ *such that locally*

$$\boxed{A(x) = g(x) - h(x)\sqrt{1 - \frac{x}{x_0}}}.$$

Remark. *If there is no* x_0, a_0 *inside the region of convergence of* Φ *then the singular behaviour of* Φ *determines the singular behaviour of* $A(x)$ *!!!*

Sub-critical Graphs

$A(x) = xC^\bullet(x)$, $\Phi(x,a) = xe^{B^\bullet(a)}$, $\boxed{xC^\bullet(x) = xe^{B^\bullet(xC^\bullet(x))}}$

$$\Longrightarrow \quad \boxed{A(x) = \Phi(x, A(x))}$$

A block-stable graph class is called **sub-critical** *if the system (note that $B^\bullet(x) = B'(x)$)*

$$a_0 = x_0 e^{B'(a_0)}, \quad 1 = x_0 e^{B'(a_0)} B''(a_0)$$

has positive solutions x_0, a_0 inside the region of convergence of $\Phi(x,a) = xe^{B^\bullet(a)}$. In particular we get a **squareroot singularity** *for $C^\bullet(x)$.*

This means that "a_0 **is smaller than the radius of convergence** *η* **of** *$B^\bullet$".*

Eliminating x_0 leads to $\boxed{a_0 B''(a_0) = 1}$ *or that*

$$\boxed{\eta B''(\eta) > 1}$$

where η is the radius of convergence of $B(x)$.

Sub-critical Graphs

- **Trees** *are sub-critical*
- **Outerplanar graphs** *are sub-critical*
- **Series-parallel graphs** *are sub-critical*

Sub-critical Graphs

Lemma. *Suppose that* $B(x)$ *has radius of convergence* $\eta \in (0, \infty]$.

$$\lim_{x \to \eta} B''(x) = \infty \quad \Longrightarrow \quad \textbf{sub-critical}.$$

Corollary *If* $B^\bullet(x) = B'(x)$ *is entire or has a squareroot singularity:*

$$B^\bullet(x) = g(x) - h(x)\sqrt{1 - \frac{x}{\eta}},$$

then we are in the **sub-critical** *case.*

This applies for **outerplanar** *and* **series-parallel** *graphs.*

Sub-critical Graphs

What does "sub-critical" mean?

In a sub-critical graph class the **average size of the 2-connected components is bounded***.*

$\Longrightarrow$ *This leads to a* **tree like structure***.*

$\Longrightarrow$ *The* **law of large numbers** *should apply so that we can expect* **universal behaviors** *that are independent of the the precise structure of 2-connected components.*

Sub-critical Graphs

Universal properties

- **Asymptotic enumeration**:

Labelled case:

$$\boxed{c_n \sim c\, n^{-5/2} \rho^{-n} n!}$$

Unlabelled case:

$$\boxed{c_n \sim c\, n^{-5/2} \rho^{-n}}$$

($c > 0$, ρ ... radius of convergence of $C(z)$)

[D.+Fusy+Kang+Kraus+Rue 2011]

Sub-critical Graphs

- **Asymptotic enumeration**:

$$C^\bullet(x) = e^{B^\bullet(xC^\bullet(x))}$$

$$\longrightarrow \quad xC^\bullet(x) = xC'(x) = g(x) - h(x)\sqrt{1-\frac{x}{\rho}}$$

$$\longrightarrow \quad [x^n]xC'(x) = \frac{n\,c_n}{n!} \sim c\,n^{-3/2}\rho^{-n}$$

$$\longrightarrow \quad \boxed{c_n \sim c\,n^{-5/2}\rho^{-n}n!}.$$

Additive Parameters in Subcritical Graph Classes

Theorem 1 *[D.+Fusy+Kang+Kraus+Rue]*

X_n ... *number of* **edges** / *number of* **blocks** / *number of* **cut-vertices** / *number of* **vertices of degree** k

$$\Longrightarrow \quad \boxed{\frac{X_n - \mu n}{\sqrt{n}} \to N(0, \sigma^2)}$$

with $\mu > 0$ *and* $\sigma^2 \ge 0$.

Remark. *There is an easy to check "combinatorial condition" that ensures* $\sigma^2 > 0$.

Additive Parameters in Subcritical Graph Classes

Proof Methods:
Refined versions of the functional equation $C^\bullet(x) = e^{B^\bullet(xC^\bullet(x))}$,
*+ singularity analysis (***always squareroot singularity***)*

E.g: number of edges:

$$C^\bullet(x,y) = e^{B^\bullet(xC^\bullet(x,y),y)}$$

or number of 2-connected components:

$$C^\bullet(x,y) = e^{yB^\bullet(xC^\bullet(x,y))}$$

$$\longrightarrow \quad C^\bullet(x,y) = g(x,y) - h(x,y)\sqrt{1 - \frac{x}{\rho(y)}}$$

$$\longrightarrow \quad [x^n]C^\bullet(x,y) \sim c(y)\rho(y)^{-n}n^{-3/2}$$

+ application of Quasi-Power-Theorem (by Hwang).

Graph Limits

$\mathcal{T}_e$ ... **continuum random tree** *(CRT)*

Theorem *[Panagiotou+Stufler+Weller]*

$\mathcal{C}$ ... *sub-critical graph class of connected graphs*

$$\Longrightarrow \quad \boxed{\frac{c}{\sqrt{n}}\mathcal{C}_n \to \mathcal{T}_e}$$

with respect to the Gromov-Hausdorff metric, where $c > 0$ *is a constant.*

Corollary. *The diameter* D_n *as well as a typical distance in a sub-critical graph is or order* $\sqrt{n}$*.*

Subgraph Counting

Theorem *[D.+Ramos+Rue]*

$\mathcal{G}$ ... *sub-critial graph class,* $\boxed{H \in \mathcal{G}}$ *fixed.*
$X_n^{(H)}$ ... *number of occurences of* H *as a subgraph in graphs of size* n

$$\Longrightarrow \quad \boxed{\frac{X_n^{(H)} - \mu n}{\sqrt{n}} \to N(0, \sigma^2)}$$

with $\mu > 0$ *and* $\sigma^2 \geq 0$.

Remark. *The proof is easy if* H *is 2-connected* = **additive parameter!!!**

Subgraph Counting

$\boxed{H = P_2}$ ... *path of length 2*

$B_j^\bullet(w_1, w_2, w_3, \ldots; u)$ .... *generating function of blocks in* $\mathcal{G}$, *where the root has degree* j, *where* w_i *counts the number of non-root vertices of degree* i, *and where* u *counts the number of occurrences of* $H = P_2$.

$C_j^\bullet(x, u)$ ... *generating function of connected rooted graphs in* $\mathcal{G}$, *where the root vertex has degree* j, *where* x *counts the number of (all) vertices and* u *the number of occurrences of* $H = P_2$.

Subgraph Counting

System of infinite number of equations

$$C_j^\bullet(x,u) = \sum_{s\ge 0} \frac{1}{s!} \sum_{j_1+\cdots+j_s=j} u^{\sum_{i_1<i_2} j_{i_1} j_{i_2}}$$
$$\times \prod_{i=1}^{s} B_{j_i}^\bullet \left(x \sum_{\ell_1\ge 0} u^{\ell_1} C_{\ell_1}^\bullet(x,u), x \sum_{\ell_2\ge 0} u^{2\ell_2} C_{\ell_2}^\bullet(x,u), \ldots; u \right),$$
$$(j \ge 0)$$

$$C_j^\bullet(x,1) = \sum_{s\ge 0} \frac{1}{s!} \sum_{j_1+\cdots+j_s=j} \prod_{i=1}^{s} B_{j_i}^\bullet \left(xC^\bullet(x), xC^\bullet(x), \ldots; 1 \right)$$

$$C^\bullet(x) = \sum_{\ell\ge 0} C_\ell^\bullet(x,1)$$

Subgraph Counting

System of infinite number of equations

Suppose that $\boxed{\mathbf{A}(z) = (A_j(z))_{j\ge 0} = \mathbf{\Phi}(z, \mathbf{A}(z))}$ *is a* **positive**, **non-linear**, **infinite** *and* **strongly connected** *system such that the* **Jacobian** $\mathbf{\Phi}_{\mathbf{a}}(z, \mathbf{a})$ *is* **compact** *for* $z > 0$ *and* $\mathbf{a} > \mathbf{0}$.

Let $z_0 > 0$, $\mathbf{a}_0 = (a_{j,0})_{j\ge 0}$ *(inside the region of convergence) satisfy the system of equations:*

$$\boxed{\mathbf{a}_0 = \mathbf{\Phi}(z_0, \mathbf{a}_0), \quad r(\mathbf{\Phi}_{\mathbf{a}}(z_0, \mathbf{a}_0)) = 1},$$

where $r(\cdot)$ *denotes the spectral radius.*

Then there exists analytic function $g_j(z), h_j(z) \ne 0$ *such that locally*

$$\boxed{A_j(z) = g_j(z) - h_j(z)\sqrt{1 - \frac{z}{z_0}}}.$$

with $g_j(z_0) = a_{j,0}$ *and* $h_j(z_0) > 0$.

Infinite Systems of Functional Equations

COMBINATORIAL CENTRAL LIMIT THEOREM III

Suppose that $\boxed{\mathbf{A}(z,u) = (A_j(z,u))_{j\ge 0} = \mathbf{\Phi}(z,u,\mathbf{A}(z,u))}$ *is a* **positive**, **non-linear**, **infinite** *and* **strongly connected** *system such that the* **Jacobian $\mathbf{\Phi}_{\mathbf{a}}(z,1,\mathbf{a})$ is compact** *for* $z > 0$ *and* $\mathbf{a} > \mathbf{0}$.

Let $z_0 > 0$, $\mathbf{a}_0 = (a_{j,0})_{j\ge 0}$ *(inside the region of convergence) satisfy the system of equations:*

$$\boxed{\mathbf{a}_0 = \mathbf{\Phi}(z_0,1,\mathbf{a}_0), \quad r(\mathbf{\Phi}_{\mathbf{a}}(z_0,1,\mathbf{a}_0)) = 1},$$

where $r(\cdot)$ *denotes the spectral radius.*

Then there exists analytic function $g_j(z,u), h_j(z,u) \ne 0$ *and* $\rho(u)$ *such that locally*

$$\boxed{A_j(z,u) = g_j(z,u) - h_j(z,u)\sqrt{1-\frac{z}{\rho(u)}}}.$$

with $g_j(z_0,1) = a_{j,0}$, $h_j(z_0,1) > 0$, *and* $\rho(1) = z_0$.

Infinite Systems of Functional Equations

COMBINATORIAL CENTRAL LIMIT THEOREM III (cont.)

Suppose that $A(z,u) = \Psi(z,u,(A_j(z,u))_{j\ge 0})$*, where* Ψ *is analytic with non-negative coefficients.*

$$\Longrightarrow \qquad A(z,u) = g(z,u) - h(z,u)\sqrt{1-\frac{z}{\rho(u)}}$$

$$\longrightarrow \qquad [z^n]A(z,u) \sim C(u)\rho(u)^{-n}n^{-3/2}$$

Consider the random variable X_n *giben by*

$$\mathbb{P}\{X_n = k\} = \frac{a_{nk}}{a_n},$$

where $a_{n,k} = [z^n u^k]\, A(z,u)$ *and* $a_n = [z^n]\, A(z,1)$*. Then* X_n *satisfies a* **central limit theorem** *with* $\mathbb{E}\, X_n \sim \mu n$ *and* $\mathbb{V}rmar\, X_n \sim \sigma^2 n$*.*

Subgraph Counting

Special case of infinite system

$$A_j = \Phi_j(z, u, A_0, A_1, \ldots), \qquad j \geq 0,$$

with

$$\boxed{\Phi_j(z, \mathbf{1}, A_0, A_1, \ldots) = \tilde{\Phi}_j(z, A_0 + A_1 + \cdots)},$$

so that $\boxed{A = A_0 + A_1 + \cdots}$ *satisfies*

$$\boxed{A = \tilde{\Phi}(z, A)},$$

where

$$\tilde{\Phi}(z, A) = \sum_{j \geq 0} \tilde{\Phi}_j(z, A) = \sum_{j \geq 0} \Phi(z, 1, A_0, A_1, \ldots)$$

$$\Longrightarrow \quad \frac{\partial \Phi_j}{\partial a_i}(z, 1, \mathbf{a}) \quad \textit{does not depend on } i$$

$$\Longrightarrow \quad \boxed{\mathbf{\Phi_a}(z, 1, \mathbf{a}) \ \ \textit{is compact}}$$

Thank You!