Igor Kortchemski
Oxford probability workshop 2025
Discrete random structures

I | Bienaymé trees and random walks

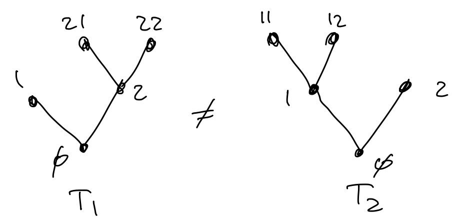
Outline:
1) Coding trees
2) Connection with conditioned random walks
3) The cycle lemma

1) Coding trees

Recall that we work with plane trees, for example:

$T_1 = \{\phi, 1, 2, 21, 22\}$

$T_2 = \{\phi, 1, 2, 11, 12\}$



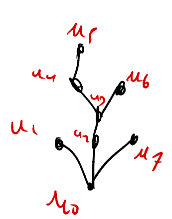
Formally, they can be defined as certain sets of labels (sequences of integers), we skip the formal definition.
Informally, a plane tree can be seen as a genealogical tree where individuals are the vertices.

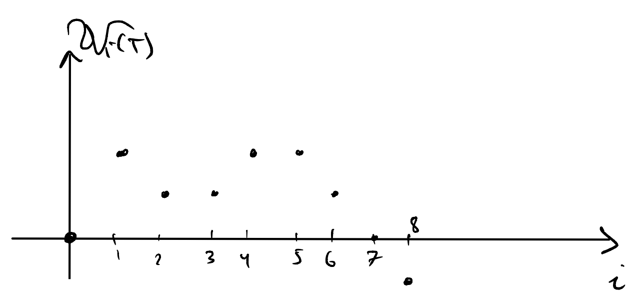
Vertices of a plane tree can be equiped with the depth-first search order (informally, label vertices as soon as possible when doing the "contour" of the tree from left to right)

Definition Let T be a tree with size n, with vertices ordered in depth-first search order: $u_0 < u_1 < \cdots < u_{n-1}$.
The Łukasiewicz path $\mathcal{W}(T) = (\mathcal{W}_0(T), \ldots, \mathcal{W}_n(T))$ is defined by:

- $\mathcal{W}_0(T) = 0$
- $\mathcal{W}_{i+1}(T) = \mathcal{W}_i(T) + k_{u_i}(T) - 1$ for $0 \leq i \leq |T|-1$ with $k_u(T)$ = number of children of u.

Example





Proposition The map $\{\text{trees with } n \text{ vertices}\} \longrightarrow \overline{S}_n^{(1)}$

$$T \longmapsto (k_{u_i}(T) - 1 : 0 \leq i \leq n-1)$$

is a bijection, where $\overline{S}_n^{(1)} = \{ (x_1, \dots, x_n) \in \{-1, 0, 1, \dots\}^n : x_1 + \dots + x_n = -1, \; x_1 + \dots + x_i > -1 \text{ for } 1 \leq i \leq n \}$

This can be readily shown by induction. The complete proof is a bit tedious to write and it is skipped here (the reader should convince him/her that this is true)

2) Connection with conditioned random walks

Let $\mu = (\mu(i) : i \geq 0)$ be a probability distribution on $\mathbb{Z}_+$, called the offspring distribution.
Assume that $\mu(1) \neq 1$ and $\sum_{i=0}^{\infty} i\mu(i) \leq 1$
Let $\mathbb{T}$ be the set of all finite trees ($\mathbb{T}$ is countable)

Recall that for $T \in \mathbb{T}$, $\mathbb{P}_\mu(T) \overset{(*)}{=} \prod_{u \in T} \mu(k_u(T))$ defines a probability measure on $\mathbb{T}$.
A $\mathbb{B}_\mu$ random tree will be a $\mathbb{T}$-valued random variable, with law $\mathbb{P}_\mu$.

To make the connection with the Lukasiewicz path, we introduce the random walk $(W_n)_{n \geq 0}$: Let $(X_i)_{i \geq 1}$ be iid random variables with law $\mathbb{P}(X_1 = k) = \mu(k+1)$ for $k \geq -1$.
Set $W_0 = 0$ and $W_n = X_1 + \dots + X_n$ for $n \geq 1$.
Also set $\zeta_k = \inf\{n \geq 1 : W_n = -k\} \in \mathbb{N} \cup \{+\infty\}$ for $k \geq 1$.

Important remark Observe that $E[X_1] = \sum_{k \geq -1} k\mu(k+1) = \sum_{k=1}^{\infty} k\mu(k) - 1$. In particular, critical trees (for which $\sum_{k=1}^{\infty} k\mu(k) = 1$) play a special role: we have $E[X_1] = 0$ iff μ is critical.

Proposition Let τ be a $\mathbb{B}_\mu$ random tree. Then

$$(\mathcal{W}_0(\tau), \dots, \mathcal{W}_{|\tau|}(\tau)) \overset{\text{law}}{=} (W_0, \dots, W_{\zeta_1})$$

In particular $\zeta < \infty$ a.s.

The proof is straightforward using (★) by computing the probability that the 2 random vectors are equal to $(w_0, \dots, w_n)$.

In the sequel, τ_n denotes a $\mathbb{B}_\mu$ random tree conditioned on having n vertices (we implicitly restrict to values of n such that $\mathbb{P}(|\tau|=n)>0$).

Corollary
- $|\tau| \overset{\text{law}}{=} \zeta$
- $(\mathcal{W}_0(\tau_n), \dots, \mathcal{W}_n(\tau_n)) \overset{\text{law}}{=} (W_0, \dots, W_n)$ under $\mathbb{P}(\cdot \mid \zeta_1 = n)$

The main difficulty is that this conditioning is "non local"! To make it "local" we we are going to use the so-called cycle lemma.

3) The cycle lemma

We first introduce some notation. Fix $k \geq 1$. For $n \geq k$:

Set $S_n^{(k)} = \{ (x_1, \dots, x_n) \in \{-1, 0, 1, \dots\}^n : x_1 + \dots + x_n = -k \}$,

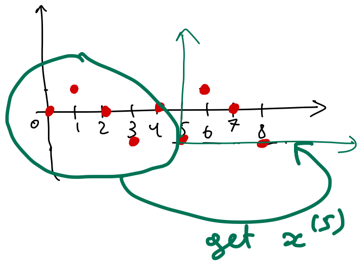
set $\overline{S}_n^{(k)} = \{ (x_1, \dots, x_n) \in \{-1, 0, 1, \dots\}^n : x_1 + \dots + x_n = -1$ and $x_1 + \dots + x_i > -k$ for $1 \leq i \leq n \}$

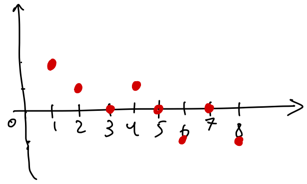
We identify $\mathbb{Z}/n\mathbb{Z}$ with $\{0, 1, \dots, n-1\}$

For $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ and $i \in \mathbb{Z}/n\mathbb{Z}$, we define $x^{(i)} = (x_{i+1}, x_{i+2}, \dots, x_{i+n})$ with addition considered modulo n.

Example $x = (1, -1, -1, 1, -1, 2, -1, -1)$, $x^{(5)} = (2, -1, -1, 1, -1, -1, 1, -1)$.

Walks with jumps given by x and $x^{(5)}$:





Theorem (Cycle lemma) Fix $1 \le k \le n$. For every $x \in S_n^{(k)}$, set $I(x) = \{ i \in \mathbb{Z}/n\mathbb{Z} : x^{(i)} \in \overline{S}_n^{(k)} \}$.
Then $\# I(x) = k$.

The proof is a bit tedious and left to the reader. In the previous example $I(x) = \{5\}$.

Corollary For every $1 \le k \le n$, $\mathbb{P}(\zeta_k = n) = \frac{k}{n} \mathbb{P}(W_n = -k)$

Proof Set $\vec{X}_n = (X_1, \dots, X_n)$ and write

$$\mathbb{P}(\zeta_k = n) = \mathbb{P}(\vec{X}_n \in \overline{S}_n^{(k)}) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{P}(\vec{X}_n^{(i)} \in \overline{S}_n^{(k)}) = \frac{1}{n} \mathbb{E}\Big[\underbrace{\sum_{i=1}^{n} \mathbb{1}_{\vec{X}_n^{(i)} \in \overline{S}_n^{(k)}}}_{= k \, \mathbb{1}_{\vec{X}_n \in S_n^{(k)}}} \Big] = \frac{k}{n} \mathbb{P}(W_n = -k)$$

Application

(1) Set $\mathbb{F}_n^f = \{$forests of f trees having in total n edges$\}$. Then $\# \mathbb{F}_n^f = \frac{f}{2n+f} \binom{2n+f}{n}$

(2) Set H_n be the height of a uniform vertex in a uniform plane tree with n edges. Then for $0 \le h \le n$:

$$\mathbb{P}(H_n = h) = \frac{2h+1}{2n+1} \cdot \frac{\binom{2n+1}{n-h}}{\binom{2n}{n}}$$

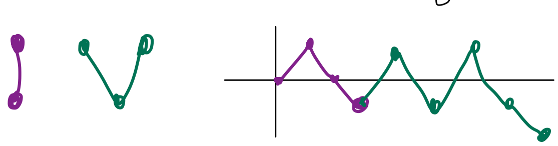
Using Stirling's formula, (2) implies that $\forall x > 0$: $\mathbb{P}(H_n \ge x\sqrt{n}) \xrightarrow[n \to \infty]{} e^{-x^2}$: a typical vertex of a uniform plane tree has height of order $\sqrt{n}$.

Proof (1)

Set $R_n^{(k)} = \{ (x_1, \dots, x_n) \in \{-1,1\}^n : x_1 + \dots + x_n = -k \}$,

set $\overline{R}_n^{(k)} = \{ (x_1, \dots, x_n) \in \{-1,1\}^n : x_1 + \dots + x_n = -1$ and $x_1 + \dots + x_i > -k$ for $1 \le i \le n \}$

observe that $\mathbb{F}_n^f$ is in bijection with $R_{2n+f}^{(f)}$ by using the "contour" encoding, adding -1 after each tree:



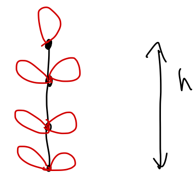
Taking $\mathbb{P}(X_1=\pm 1)=\frac{1}{2}$, when $k \equiv n\ [2]$ we get by the corollary

$$\frac{1}{2^n}\#\bar{R}_n^{(k)} = \mathbb{P}(\zeta_k = n) = \frac{k}{n}\mathbb{P}(W_n=-k) = \frac{k}{n}\times\frac{1}{2^n}\#R_n^{(k)}$$

We conclude that $\#\bar{R}_n^{(k)} = \frac{k}{n}\#R_n^{(k)} = \frac{k}{n}\binom{n}{\frac{n+k}{2}}$

because an element in $R_n^{(k)}$ has $\frac{n+k}{2}$ +1 elements among n elements.

(2) A pair (tree with n edges, distinguished point at height h) is in bijection with $\mathbb{F}_{n-h}^{2h+1}$:



Thus $\mathbb{P}(H_n=h) = \dfrac{\#\mathbb{F}_{n-h}^{2h+1}}{(n+1)\#\mathbb{F}_n^{1}}$, which gives the result.

References for II :

- J.-F. Le Gall, Random trees and applications, Probability surveys 2005
- J. Pitman, Combinatorial stochastic processes, lecture notes in Mathematics 2006 (Chap. 5)

II) Local limits of critical Bienaymé trees

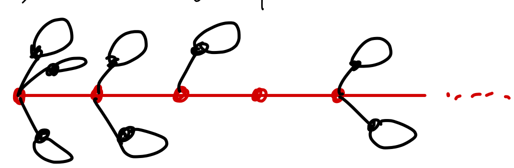
Outline 1) Kesten's tree
2) Local estimates for r.v.

Goal: show that large critical Bienaymé trees converge "locally".

1) Kesten's tree

Let μ be a critical offspring distribution. Define $\hat{\mu}$, the size-biased distribution by $\hat{\mu}(k) = k\mu(k)$ for $k \geq 1$.
Build $\mathcal{T}_\infty$ as follows:

- there are two types of vertices: normal and special, with the root being special
- normal vertices have children $\sim \mu$, special vertices have children $\sim \hat{\mu}$
- children of normal vertices are normal, a child of a special vertex selected uniformly at random is special, the other are normal
- all choices are $\perp\!\!\!\perp$



Definition For a tree t and $k \geq 0$:

- set $Z_k(t) = \{u \in t : |u| = k\}$ the number of vertices at generation k.
- denote by $[t]_k = \{u \in t : |u| \leq k\}$ the tree obtained by keeping the vertices in the first k generations

The goal of II is to prove the following theorem:

Theorem (Kesten, Janson, Abraham-Delmas)
Assume that μ is critical. Then $\forall k \geq 0$, $[\mathcal{T}_n]_k \xrightarrow[n\to\infty]{(d)} [\mathcal{T}_\infty]_k$ (convergence in distribution in the countable space of finite trees).

Remark On the set $\mathbb{T}_{lc}$ of locally finite plane trees, it is possible to check that

$$d_{loc}(t,t') = \frac{1}{1+\sup\{n\geq 0 : [t]_n = [t']_n\}}$$

defines a distance (called the local distance) so that

$(\mathbb{T}_{lc}, d_{loc})$ is Polish and $t_n \to t$ in $(\mathbb{T}_{lc}, d_{loc})$ iff $\forall k\geq 0$, $[t_n]_k \xrightarrow[n\to\infty]{} [t]_k$

In the sequel, τ denotes a B_μ tree

Proposition For every $k\geq 0$ and tree t with height k:

$$\mathbb{P}([\tau_\infty]_k = t) = Z_k(t)\, \mathbb{P}([\tau]_k = t)$$

Proof Fix $x\in t$ with $|x|=k$. Write

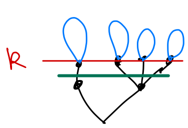
$$\mathbb{P}([\tau_\infty]=t,\ \text{spine of } \tau_\infty \text{ goes through } x) = \prod_{\substack{u\in t\\ |u|\neq k\\ u\notin [\![\emptyset,x]\!]}} \mu(k_u(t)) \cdot \prod_{\substack{u\in [\![\emptyset,x]\!]\\ u\neq x}} \hat{\mu}(k_u(t)) \times \frac{1}{k_u(t)} = \prod_{\substack{u\in t\\ |u|\neq k}} \mu(k_u(t)) = \mathbb{P}([\tau]_k = t)$$

We get the result by summing over all the possible $Z_k(t)$ values of x.

Beginning of the proof We need to show that $\forall k\geq 0$ and $\forall$ tree t with height k:

$$\mathbb{P}([\tau_n]_k = t) \xrightarrow[n\to\infty]{} Z_k(t)\, \mathbb{P}([\tau]_k = t)$$

Write $\mathbb{P}([\tau_n]_k = t) = \mathbb{P}([\tau]_k = t) \times \dfrac{\mathbb{P}(\text{a forest of } Z_k(t) \text{ trees has size } n-\#[t]_{k-1})}{\mathbb{P}(\#\tau = n)}$



$$= \mathbb{P}([\tau]_k = t)\cdot \frac{Z_k(t)}{n-\#[t]_{k-1}} \times n \times \frac{\mathbb{P}(W_{n-\#[t]_{k-1}} = -Z_k(t))}{\mathbb{P}(W_n = -1)}$$

Goal: $\forall a,b\geq 0$, $\dfrac{\mathbb{P}(W_{n-a} = -b)}{\mathbb{P}(W_n = -1)} \xrightarrow[n\to\infty]{} 1$ $(*)$

2) Local estimates for r.v.

Def A real-valued r.v. X is lattice if there exists $b \in \mathbb{R}$ and $h > 0$ s.t $\mathbb{P}(X \in b + h\mathbb{Z}) = 1$. If the greatest such h is called the span of X. If X has span 1 we say that X is aperiodic.

Example If $\mathbb{P}(X=1) = \mathbb{P}(X=-1) = \frac{1}{2}$, X has span 2.

Theorem (Strong ratio limit theorem) Assume that X_1 is $\mathbb{Z}$-valued, aperiodic with $\mathbb{E}[X_1] = 0$. Then $\forall p > 0, \forall q \in \mathbb{Z}, \forall r \in \mathbb{Z}$

$$\frac{\mathbb{P}(W_{n-p} = -q)}{\mathbb{P}(W_n = r)} \xrightarrow[n \to \infty]{} 1$$

This entails the desired local convergence of 1)

Remark: aperiodicity is needed: if $\mathbb{P}(X=1) = \mathbb{P}(X=-1) = \frac{1}{2}$ we have $\mathbb{P}(W_n = r) = 0$ if $n \not\equiv r\ [2]$

Exercise If X is $\mathbb{Z}$-valued, aperiodic with $\mathbb{P}(X>0)>0$, $\mathbb{P}(X<0)>0$, then $\forall r \in \mathbb{Z}$, $\exists m \geq 1$ s.t for every $k \geq m$, $\exists x_1, \dots, x_k \in \mathrm{Support}(X)$ with $r = x_1 + \dots + x_k$. In particular $\mathbb{P}(W_k = r) > 0$

The main ingredient is the fact that $\mathbb{P}(W_n = r)$ cannot decrease too fast:

Proposition If X_1 is aperiodic with $\mathbb{E}[X_1]=0$, for every $r \in \mathbb{Z}$, we have $\mathbb{P}(W_n = r) = e^{-o(n)}$

Proof of the theorem (Neveu)

Step 1 It is enough to show that $\frac{\mathbb{P}(W_{n-1} = r - x)}{\mathbb{P}(W_n = r)} \xrightarrow[n \to \infty]{(*)} 1$ for every $x \in S = \mathrm{Support}(X_1)$ and for every $r \in \mathbb{Z}$.

Indeed, assume that this holds and let us show $\frac{\mathbb{P}(W_{n-p} = -q)}{\mathbb{P}(W_n = r)} \xrightarrow[n \to \infty]{} 1$.

By aperiodicity $\exists m \geq 1$ s.t $\forall k \geq m$ $r+q$, $r \in S + \dots + S$ (S k times)

By $(*)$, we get $\dfrac{\mathbb{P}(W_{n-(m+p)}=-q)}{\mathbb{P}(W_n=r)} \xrightarrow[n\to\infty]{} 1$ and $\dfrac{\mathbb{P}(W_{n-m}=r)}{\mathbb{P}(W_n=r)} \xrightarrow[n\to\infty]{} 1$

By dividing $\dfrac{\mathbb{P}(W_{n-m-p}=-q)}{\mathbb{P}(W_{n-m}=r)} \xrightarrow[n\to\infty]{} 1$, which is what we want.

Step 2 To show $(*)$, set $N_n = \#\{1 \le i \le n : X_i = x\}$ and $\alpha = \mathbb{P}(X_1 = x)$

$$\mathbb{E}\left[\frac{N_n}{n} \,\Big|\, W_n = r\right] \underset{\text{symmetry}}{=} \mathbb{P}(X_1 = x \mid W_n = r) \underset{\text{Markov}}{=} \alpha \frac{\mathbb{P}(W_{n-1} = r - x)}{\mathbb{P}(W_n = r)}$$

We show that $\mathbb{E}\left[\frac{N_n}{n} \,\Big|\, W_n = r\right] \xrightarrow[n\to\infty]{} 1$ which implies the desired result.

Fix $\varepsilon > 0$. Since $N_n \sim \mathrm{Bin}(n, \alpha)$, $\exists c_\varepsilon > 0$ s.t $\mathbb{P}\left(\left|\frac{N_n}{n} - \alpha\right| \ge \varepsilon\right) \le e^{-c_\varepsilon n}$ $\forall n \ge 1$.

Then
$$\left|\mathbb{E}\left[\frac{N_n}{n} \,\Big|\, W_n = c\right] - \alpha\right| \le \varepsilon + \mathbb{P}\left(\left|\frac{N_n}{n} - \alpha\right| \ge \varepsilon \,\Big|\, W_n = r\right)$$
$$\le \varepsilon + \frac{\mathbb{P}\left(\left|\frac{N_n}{n} - \alpha\right| \ge \varepsilon\right)}{\mathbb{P}(W_n = r)}$$
$$\le \varepsilon + \frac{e^{-c_\varepsilon n}}{e^{-o(n)}} \le 2\varepsilon \quad \text{for } n \text{ large}$$

We give two different proofs of the proposition.

ⓐ Proof 1: properties of 1-D RW

Proof of the proposition

By aperiodicity, $\exists m$ s.t $\mathbb{P}(W_m = r) > 0$. Since $\mathbb{P}(W_n = r) \ge \mathbb{P}(W_{n-m} = 0)\,\mathbb{P}(W_m = r)$, it is enough to show that $\mathbb{P}(W_n = 0)^{1/n} \xrightarrow[n\to\infty]{} 1$

Step 1 Since $\mathbb{P}(W_{a+b} = 0) \ge \mathbb{P}(W_a = 0)\,\mathbb{P}(W_b = 0)$, the sequence $u_n = -\ln(\mathbb{P}(W_n = 0))$ satisfies $u_{a+b} \le u_a + u_b$ for $a, b \ge 1$. By the Fekete subadditivity lemma, $\frac{u_n}{n} \xrightarrow[n\to\infty]{} \ell \in \mathbb{R} \cup \{-\infty\}$, which shows that $\exists c \in [0,1]$ s.t. $\mathbb{P}(W_n = 0)^{1/n} \xrightarrow[n\to\infty]{} c$

Step 2 Argue by contradiction and assume that $c < 1$.

Then $A = \sum_{n=0}^{\infty} \mathbb{P}(W_n = 0) < \infty$.

This implies transience of $(W_n)_{n\geq 0}$. But $\mathbb{E}[W_n]=0$ implies recurrence, contradiction.
To see this, we can proceed as follows.

Fix $\varepsilon>0$.

<u>Claim</u>: for n large enough $\mathbb{E}\left[\sum_{i=0}^{\infty} \mathbb{1}_{|W_i|\leq \varepsilon n}\right] \geq \frac{n}{2}$.

<u>Proof</u> By the SLN, a.s. $\exists N$ s.t $n\geq N \Rightarrow |W_n|\leq \varepsilon n$. Then for $n\geq N$

$$\sum_{i=0}^{\infty} \mathbb{1}_{|W_i|\leq \varepsilon n} \geq \sum_{i=N}^{n} \mathbb{1}_{|W_i|\leq \varepsilon n} = n-N$$ because for $N\leq i\leq n$, $|W_i|\leq \varepsilon i\leq \varepsilon n$.

Thus by Fatou:

$$1 \leq \mathbb{E}\left[\liminf_{n\to\infty} \frac{1}{n}\sum_{i=0}^{\infty} \mathbb{1}_{|W_i|\leq \varepsilon n}\right] \leq \liminf_{n\to\infty} \frac{1}{n}\, \mathbb{E}\left[\sum_{i=0}^{\infty} \mathbb{1}_{|W_i|\leq \varepsilon n}\right]$$

Now, for every $k\in\mathbb{Z}$, setting $\tau_k=\inf\{n\geq 0 : W_n=k\}$:

$$\mathbb{E}\left[\sum_{i=0}^{\infty} \mathbb{1}_{W_i=k}\right] \leq \mathbb{P}(\tau_k<\infty)\, \mathbb{E}\left[\sum_{i=0}^{\infty} \mathbb{1}_{W_i=0}\right] \leq A \quad \text{(Strong Markov property)}$$

By applying this for $k\in[-\varepsilon n, \varepsilon n]\cap\mathbb{Z}$ we get $\mathbb{E}\left[\sum_{i=0}^{\infty} \mathbb{1}_{|S_i|\leq \varepsilon n}\right] \leq (2\varepsilon n+2)C$, which contradicts the claim.

D) Proof 2: Local limit theorem + cutoff argument

The second proof actually shows the following stronger version

<u>Proposition</u> Let X_1 be a $\mathbb{Z}$-valued r.v. Write $X_1=X_1^+-X_1^-$ with $X_1^+, X_1^-\geq 0$. Assume that X_1 is aperiodic and that $\mathbb{E}[X_1^+]=\mathbb{E}[X_1^-]\in\mathbb{R}_+\cup\{+\infty\}$. Then for every $r\in\mathbb{Z}$, we have $\mathbb{P}(W_n=r)=e^{-o(n)}$

We rely on the following result.

Theorem (local limit theorem)

Let $(X_i)_{i\geq 1}$ be iid aperiodic $\mathbb{Z}$-valued r.v. Assume that $\sigma^2 = \mathrm{Var}(X_1) \in (0,\infty)$. Set $a = \mathbb{E}[X_1]$ and $W_n = X_1 + \dots + X_n$. Then

$$\sup_{k\in\mathbb{Z}} \left| \sqrt{n}\, \mathbb{P}(W_n = k) - \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{k-an}{\sigma\sqrt{n}}\right)^2} \right| \xrightarrow[n\to\infty]{} 0$$

Alternatively, $\mathbb{P}(W_n = k) = \frac{1}{\sqrt{2\pi\sigma^2 n}} e^{-\frac{1}{2}\left(\frac{k-an}{\sigma\sqrt{n}}\right)^2} + \frac{\varepsilon(k,n)}{\sqrt{n}}$ with $\sup_{k\in\mathbb{Z}} |\varepsilon(k,n)| \xrightarrow[n\to\infty]{} 0$

The standard proof (see e.g. Ibragimov & Linnik Theorem 4.2.1) is based on characteristic functions by writing $\mathbb{E}[e^{itW_n}] = \sum_{j\in\mathbb{Z}} e^{itj}\, \mathbb{P}(W_n = j)$, so $\mathbb{P}(W_n = k) = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-itk}\, \mathbb{E}[e^{itW_n}]\, dt$

$$= \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-itk}\, \phi(t)^n\, dt$$

with $\phi(t) = \mathbb{E}[e^{itX_1}]$

Observe that this gives the proposition for aperiodic X_1 and finite variance μ.

Proof of the proposition (Chung–Erdős, Kakutani).

Idea: cutoff argument

If $\mathrm{Var}(X_1) < \infty$, $\mathbb{P}(W_n = c) \underset{n\to\infty}{\sim} \frac{A}{\sqrt{n}}$ by the local limit theorem.

We can thus assume $\mathrm{Var}(X_1) = \infty$. To simplify notation set $p_k = \mathbb{P}(X_1 = k)$ for $k \in \mathbb{Z}$ and $S^+ = \mathrm{Support}(X_1) \cap \{1,2,3,\dots\}$, $S^- = \mathrm{Support}(X_1) \cap \{-1,-2,-3,\dots\}$ and $X = X_1$.

Let $M_0 > 0$ be such that $\mathbb{P}(|X| > M_0) \leq \varepsilon$ and $X\, \mathbb{1}_{|X|\leq M_0}$ is aperiodic.

Case 1: $\# S^+ = \infty$ and $\# S^- = \infty$

Fix $\varepsilon \in (0, \frac{1}{2})$.

Let $M \geq M_0$ be such that $\mathbb{E}[X\, \mathbb{1}_{-M\leq X\leq M_0}] < 0$.

Let $N \geq M_0$ be the smallest integer such that $c = \mathbb{E}[X\, \mathbb{1}_{-M\leq X\leq N}] \geq 0$

Observe that $c < N p_N$ $(*)$ because $c - N p_N = \mathbb{E}[X\, \mathbb{1}_{-M\leq X\leq N-1}] < 0$

Set $b = \mathbb{P}(-M \leq X \leq N) - \frac{c}{N} \geq 1 - \varepsilon - p_N \geq 1 - 2\varepsilon$ by $(*)$.

Set $p'_k = \begin{cases} p_k & \text{for } -M \leq k < N \\ p_N - \frac{c}{N} & \text{for } k = N \quad (\geq 0 \text{ by } (*)) \\ 0 & \text{otherwise} \end{cases}$

Observe that $\sum_{k \in \mathbb{Z}} p'_k = b$ and $\sum_{k \in \mathbb{Z}} k p'_k = 0$.

Case 2 $\# S^+ < \infty$ or $\# S^- < \infty$. Without loss of generality (consider $-X$ instead of X) assume $\# S^+ = \infty$ and $\# S^- < \infty$. Write $-M = \inf S^-$. Fix $\varepsilon > 0$ with $\varepsilon < \min(-\mathbb{E}[X \mathbb{1}_{-M \leq X \leq M_0}], M p_{-M})$

let $N \geq M_0$ be the smallest integer such that $c = \mathbb{E}[X \mathbb{1}_{-M \leq X \leq N}] + \varepsilon \geq 0$

Observe that $c \leq N p_N$ $(**)$ because $c - N p_N = \mathbb{E}[X \mathbb{1}_{-M \leq X \leq N-1}] + \varepsilon < 0$

Set $b = \mathbb{P}(-M \leq X \leq N) - \frac{c}{N} - \frac{\varepsilon}{M} \geq 1 - \varepsilon - p_N - \varepsilon \geq 1 - 3\varepsilon$ by $(**)$

Set $p'_k = \begin{cases} p_{-M} - \frac{\varepsilon}{M} & \text{for } k = -M \\ p_k & \text{for } -M \leq k < N \\ p_N - \frac{c}{N} & \text{for } k = N \quad (\geq 0 \text{ by } (**)) \\ 0 & \text{otherwise} \end{cases}$

Observe that $\sum_{k \in \mathbb{Z}} p'_k = b$ and $\sum_{k \in \mathbb{Z}} k p'_k = 0$.

Finally, let $(Y_i)_{i \geq 1}$ be iid with $\mathbb{P}(Y_1 = k) = \frac{p'_k}{b}$ for $k \in \mathbb{Z}$ and set $T_n = Y_1 + \cdots + Y_n$. Then

$$\mathbb{P}(T_n = r) = \sum_{\substack{x_1, \ldots, x_n \in \mathbb{Z} \\ -M \leq x_i \leq N \\ x_1 + \cdots + x_n = r}} \frac{p'_1}{b} \times \cdots \times \frac{p'_n}{b} \leq \frac{1}{b^n} \sum_{\substack{x_1, \ldots, x_n \in \mathbb{Z} \\ -M \leq x_i \leq N \\ x_1 + \cdots + x_n = r}} p_1 \cdots p_n \quad \text{because } p'_i \leq p_i$$

$$= \frac{1}{b^n} \mathbb{P}(S_n = r, -M \leq X_i \leq N \text{ for } 1 \leq i \leq n) \leq \frac{1}{b^n} \mathbb{P}(S_n = r)$$

We conclude that $\mathbb{P}(W_n = r) \geq b^n \, \mathbb{P}(T_n = r) \geq (1 - 3\varepsilon)^n \frac{A}{\sqrt{n}}$ (LLT for T_n)

This shows that $\liminf_{n \to \infty} \mathbb{P}(W_n = r)^{\frac{1}{n}} \geq 1 - 3\varepsilon$, which implies the result.

References for VI:

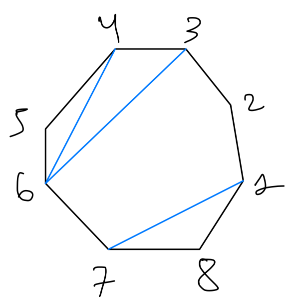
- K.L. Chung, P. Erdős, Probability limit theorems assuming only the first moment, Memoirs AMS, 1951
- J. Neveu, Sur le théorème ergodique de Chung-Erdos, C.R. Acad. Sci. Paris 257, 1963
- N. Curien, Random walks and graphs, lecture notes (Chapter 3)

III Random dissections

Outline
1) The dual tree
2) Minami's mapping
3) Local limit and application

A dissection of a regular polygon P_n with n vertices $\{1, 2, \dots, n\}$ is a collection of non-crossing diagonals

Example:



For a dissection D, let:
- $V(D)$ be the number of diagonals adjacent to vertex 1 (1 in the example)
- $F(D)$ be the size of the face containing vertices 1 and 2 (5 in the example)

Let D_n be a dissection of P_n chosen uniformly at random
What is the asymptotic behavior of $V(D_n), F(D_n)$ as $n \to \infty$?

The goal of this section is to prove the following result:

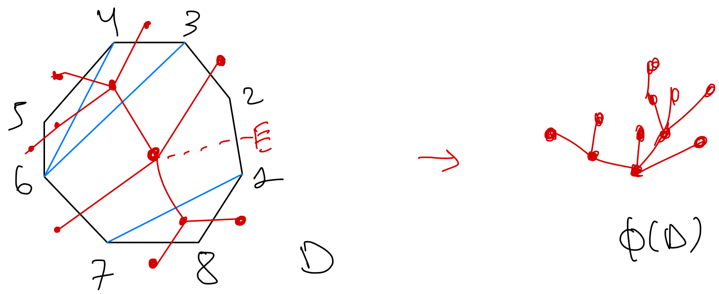
Theorem [Curien, K.]
We have
- $\mathbb{P}(F_n = k) \xrightarrow[n\to\infty]{} (k-1)\left(\frac{2-\sqrt{2}}{2}\right)^{k-2}$ for $k \geq 3$
- $\mathbb{P}(V_n = k) \xrightarrow[n\to\infty]{} (k+1)(2-\sqrt{2})^2(\sqrt{2}-1)^k$ for $k \geq 0$

1) The dual tree

$1 - \frac{1}{\sqrt{2}}$

$\frac{2-\sqrt{2}}{2}$

One naturally associates a dual tree $\phi(D)$ with a dissection D



D $\qquad$ $\phi(D)$

Observations

- Dissections of P_{n+1} are in bijection with $\Pi_n = \{$trees with n leaves and no vertex with 1 child$\}$
- $F(D)$ = number of children of the root of $\phi(D)$
- $V(D)$ = length of the leftmost path from the root of $\phi(D) - 1$.

Lemma Set $\mu(0) = 2-\sqrt{2}$, $\mu(1) = 0$, $\mu(k) = \left(\frac{2-\sqrt{2}}{2}\right)^{k-1}$ for $k \geq 2$, which is a critical offspring distribution. Let τ_n be a $\mathcal{B}_\mu$ tree conditioned on having n leaves.
Then τ_n is uniform on Π_n.

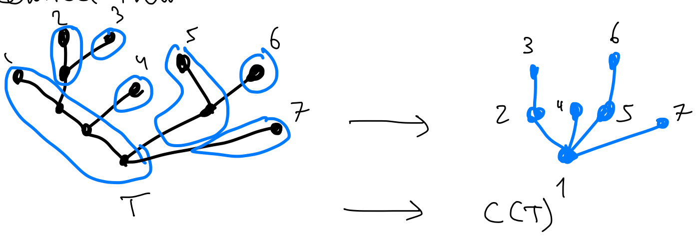
Proof We show the result for μ of the form $\mu(0) = a, \mu(1) = 0, \mu(k) = b^{k-1}$ $k \geq 2$ (this implies the lemma with $b = \frac{2-\sqrt{2}}{2}$).
Let t be a tree with n leaves and τ a $\mathcal{B}_\mu$ tree. It is enough to show that $\mathbb{P}(\tau = t)$ only depends on n. Write

$$\mathbb{P}(\tau = t) = \prod_{\substack{u \in t \\ k_u(t) = 0}} a \times \prod_{\substack{u \in t \\ k_u(t) \geq 2}} b^{k_u(t)-1} = a^n \times b^{\sum_{\substack{u \in t \\ k_u(t) \geq 2}} (k_u(t)-1)} = a^n b^{n-1}.$$

2) Minami's mapping

Minami has found a very nice mapping which turns Bienaymé trees with n leaves to Bienaymé trees with n vertices.

Construction decompose a tree into "left-most twigs" according to increasing order of the leaves, then contract them



Proposition

If τ is a B_μ tree, then $C(\tau)$ is a $B_{\tilde{\mu}}$ tree for some offspring distribution $\tilde{\mu}$.
If μ is critical, then $\tilde{\mu}$ is critical.

Proof: • the branching property of $C(\tau)$ comes from the fact that the trees attached to the left-most twig of τ are iid B_μ trees.

• To find $\tilde{\mu}$ we compute the law of $k_\emptyset(C(\tau))$, the number of children of the root of $C(\tau)$.

Let N be defined by $\mathbb{P}(N=n) = (1-\mu_0)^{n-1}\mu(0)$ for $n \geq 1$, so that N is the length of the left-most twig of τ. Then $k_\emptyset(C(\tau)) \overset{(d)}{=} \sum_{k=1}^{N-1} (Y_k - 1)$ with $(Y_i)_{i\geq 1}$ iid distributed as μ conditioned on being $\neq 0$, i.e. $\mathbb{P}(Y_1 = i) = \frac{\mu_i}{1-\mu_0}$ for $i \geq 1$.

• Now assume μ critical, so $\mathbb{E}[Y_1] = \frac{1}{1-\mu_0}$

Thus by Wald's identity, the mean of $\tilde{\mu}$ is $\mathbb{E}[N-1] \times \mathbb{E}[Y_1 - 1] = \left(\frac{1}{\mu_0} - 1\right)\left(\frac{1}{1-\mu_0} - 1\right) = 1$.

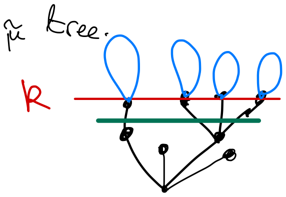
3) Local limit and application

Recall that $\mu(0) = 2-\sqrt{2}$, $\mu(1) = 0$, $\mu(k) = \left(\frac{2-\sqrt{2}}{2}\right)^{k-1}$ for $k \geq 3$. Let τ^n be a B_μ tree conditioned on having n leaves. Recall that that the dual tree of a uniform dissection of P_{n+1} has the same law as τ^n.

Proposition $\forall k \geq 0$, $[\mathcal{T}^n]_k \xrightarrow{(d)} [\mathcal{T}_\infty]_k$, where $\mathcal{T}_\infty$ is Kesten's tree associated with μ.

This implies that
$$\begin{cases} F(\mathcal{D}_n) \xrightarrow{(d)} \#\text{ children of the root of } \mathcal{T}_\infty \\ V(\mathcal{D}_n) \xrightarrow{(d)} \text{length of the left-most path starting from the path of } \mathcal{T}_\infty - 1 \end{cases}$$
which using the explicit form of μ gives the theorem stated at the beginning of the section.

Proof Denote by $\Lambda(T)$ the number of leaves of T. Let $\tilde{\mathcal{T}}$ be a $B_{\tilde{\mu}}$ tree.



Fix $k \geq 0$ and t a tree of height k. Write

$\mathbb{P}([\mathcal{T}^n]_k = t)$

$$= \mathbb{P}([\mathcal{T}]_k = t) \cdot \frac{1}{\mathbb{P}(\#\tilde{\mathcal{T}} = n)} \mathbb{P}(\text{a forest of } k \ B_{\tilde{\mu}} \text{ trees has } n - \Lambda([t]_{k-1}) \text{ vertices})$$

$$= \mathbb{P}([\mathcal{T}]_k = t) \frac{1}{\mathbb{P}(\Lambda(\mathcal{T}) = n)} \cdot \mathbb{P}(\text{a forest of } k \ B_{\mu} \text{ trees has } n - \Lambda([t]_{k-1}) \text{ leaves})$$

$$= \mathbb{P}([\mathcal{T}]_k = t) \cdot \frac{Z_k(t)}{n - \Lambda([t]_{k-1})} \times n \times \frac{\mathbb{P}(\tilde{W}_{n - \#[t]_{k-1}} = -Z_k(t))}{\mathbb{P}(\tilde{W}_n = -1)}$$

with $\tilde{W}_n = \tilde{X}_1 + \dots + \tilde{X}_n$ with $(\tilde{X}_i)_{i \geq 1}$ iid $\mathbb{P}(\tilde{X}_1 = k) = \tilde{\mu}(k+1)$ for $k \geq -1$.

Since $\tilde{\mu}$ is critical, $\mathbb{E}[\tilde{X}_1] = 0$, so $\dfrac{\mathbb{P}(\tilde{W}_{n - \#[t]_{k-1}} = -Z_k(t))}{\mathbb{P}(\tilde{W}_n = -1)} \xrightarrow[n \to \infty]{} 1$

We conclude that

$$\mathbb{P}([\mathcal{T}^n]_k = t) \xrightarrow[n \to \infty]{} Z_k(t) \, \mathbb{P}([\mathcal{T}]_k = t) = \mathbb{P}([\mathcal{T}_\infty]_k = t).$$

References for IV:

- N. Curien, I. Kortchemski, Random non-crossing plane configurations: a conditioned Galton-Watson tree approach, Random Struct. Alg 45(2), 2014.
- N. Minami, On the number of vertices with a given degree in a Galton-Watson tree, Adv. Appl. Prob. 37(1), 2005
- R. Abraham, J.-F. Delmas, An introduction to Galton-Watson trees and their local limits, Lecture Notes.

Exercise If X is $\mathbb{Z}$-valued, aperiodic with $\mathbb{P}(X>0)>0$, $\mathbb{P}(X<0)>0$, then $\forall r\in\mathbb{Z}$, $\exists m\geq 1$ s.t for every $k\geq m$, $\exists x_1,\dots,x_k\in \text{Support}(X)$ with $r=x_1+\dots+x_k$. In particular $\mathbb{P}(W_k=r)>0$

This is a variant of the Frobenius coin problem.

Solution of the exercise (argument due to Cheung, Multinomial ration [Paul Erdős solves a problem], Methods and Applications of analysis 5(2), 1998)

We will use the following lemma

Lemma (Frobenius Coin problem)

Let $(a_i)_{i\in I}$ be a finite or countable collection of integers with $\gcd(a_i : i\in I)=1$. Then $\exists M>0$ s.t for every $n\geq M$ there exist ≥ 0 integers $(x_i)_{i\in I}$ with $n=\sum_{i\in I} x_i a_i$.

Proof By Bezout, there exist integers $(u_i)_{i\in I}$ such that $1=\sum_{i\in I} u_i a_i$.

Set $c=\sum_{i\in I}|u_i|a_i$. We check that $M=c^2$ works.

Take $n\geq c^2$ and write $n=qc+r$ with $0\leq r<c$. Since $n\geq c^2$ we have $q\geq c$.

Then $n=q\sum_{i\in I}|u_i|a_i + r\sum_{i\in I}u_i a_i = \sum_{i\in I}\underbrace{(q|u_i|+r u_i)}_{x_i} a_i$

We have $x_i\geq 0$ because $q\geq r$.

$\sim$

Back to the exercise Set $\{a_i : i\in I\}=\text{Support}(X_1)$. By aperiodicity, $\gcd\{a_i-a_j : i\in I, j\in I\}=1$

For $r\in\mathbb{Z}$ let $P(r)$ be the property:

$P(r)$: " $\exists M(r)>0 : \forall n\geq M(r)$ $\exists \geq 0$ integers $(x_i)_{i\in I}$ s.t. $\sum_{i\in I}x_i=n$ and $\sum_{i\in I}x_i a_i=r$ "

We show that $P(r)$ is true for every $r\in\mathbb{Z}$.

Step 1: $P(0)$ is true.

Set $E=\{(k,\ell)\in I^2 : a_k>0>a_\ell\}\neq\emptyset$

We have $\gcd\{a_k-a_\ell : (k,\ell)\in E\}=1$ since any a_k-a_ℓ, $(k,\ell)\notin E$ can be written as a difference of two elements of that set.

By the lemma $\exists M$ s.t $n\geq N$ implies $\exists\geq 0$ integers $(m_{k\ell})_{(k,\ell)\in E}$ s.t.

$$n=\sum_{(k,\ell)\in E} m_{k\ell}(a_k-a_\ell) = \sum_{(k,\ell)\in E} m_{k,\ell}(a_k+|a_\ell|)$$

Then for $i \in I$ define x_i as follows

- if $a_i < 0$, $\quad x_i = \sum_{(k,i) \in E} m_{ki}\, a_k$
- if $a_i > 0$, $\quad x_i = \sum_{(i,\ell) \in E} m_{i\ell}\, |a_\ell|$

Then $\sum_{i \in I} x_i = n$ and $\sum_{i \in I} x_i a_i = \sum_{(k,\ell) \in E} m_{k,\ell}\, a_k a_\ell + \sum_{(k,\ell) \in E} m_{k,\ell}\, |a_\ell|\, a_k$

$$= \sum_{(k,\ell) \in E} m_{k,\ell}\, a_k \underbrace{(a_\ell + |a_\ell|)}_{=0} = 0$$

This shows $P(0)$

Step 2: by Step 1, it is enough to show $\forall r \in \mathbb{Z}$:

$Q(r)$: " $\exists \geq 0$ integers $(x_i)_{i \in I}$ s.t $r = \sum_{i \in I} x_i a_i$.

We show $Q(1)$. This will imply $Q(r)$ $\forall r \geq 1$. $Q(-1)$ is established in the same way, implying $Q(r)$ $\forall r \leq -1$. This will complete the solution.

To show $Q(1)$, set $d_- = \gcd(a_i : i \in I,\ a_i < 0)$ and $d_+ = \gcd(a_i : i \in I,\ a_i > 0)$

Since $\gcd\{a_i : i \in I\} \mid \gcd\{a_i - a_j : i,j \in I\}$, we have $\gcd\{a_i : i \in I\}$, implying $\gcd(d_+, d_-) = 1$.

By the lemma (applied twice), there exists $M > 0$ s.t $\forall n \geq M$,

$(*)$ $\exists \geq 0$ integers $(x_i)_{i \in I}$ s.t $-nd_- = \sum_{\substack{i \in I \\ a_i < 0}} x_i a_i$ and $nd_+ = \sum_{\substack{i \in I \\ a_i > 0}} x_i a_i$

Since $\gcd(d_+, d_-) = 1$, by Euclid $\exists\ a, b \in \mathbb{Z}$ with $1 = ad_- + bd_+$

Then $\forall s \geq 0$, $1 = (-s\, d_+ + a)\, d_- + (s\, d_- + b)\, d_+$

Choose s large enough so that $-sd_+ + a \leq -M$ and $sd_- + b \geq M$

By $(*)$ we can write $(-sd_+ + a)\, d_- = \sum_{\substack{i \in I \\ a_i < 0}} x_i a_i$ and $(sd_- + b)\, d_+ = \sum_{\substack{i \in I \\ a_i > 0}} x_i a_i$ with $x_i \geq 0$.

Then $1 = \sum_{i \in I} x_i a_i$