VII Summer School in Probabilities
and Stochastic Processes (CMM)

Luis Fredes
U-Bordeaux

Heaps of Pieces and Applications to Probability Theory

Contents

January 8, 2025

0 Introduction

The main objective of this course is to convince you of two facts :

- Generating (weighted) functions help to condense the combinatorial information of a class of objects by means of a function. Derivating it gives combinatorial countings, in the case where weights are equal for objects of the same size, and for probabilistic purposes we will be motivated to work with weights that represent probabilities for the objects.
- The quantity $\det(Id - M^{\{r\}})$ is a central object in probability : we will present this quantity as the weighted alternating counting of heap of cycles, an identity coming from "reversibility" that let us change M by a reversibility kernel, as the counting of weighted trees at rooted r and as one invariant measure for M. Finally we will present different examples where combinatorial identities and $\det(Id - M^{\{r\}})$ are key steps for probabilistic purposes.

1 Combinatorics

1.1 Generating function

Definition 1. Consider a class A and a size function $\ell : \mathsf{A} \longrightarrow \mathbb{N}$, the formal generating funcion $\mathcal{A}$ of A is given by

$$\mathcal{A}(x) = \sum_{a \in \mathsf{A}} x^{\ell(a)} = \sum_{n=1}^{\infty} A_n x^n$$

where A_n is the number of elements of size n in A.

This kind functions have a lot of good properties when the class A is a regular language for the size function counting the number of letters in a word or where the objects can be described by an iterative procedure mixing sizes of the objects into play.

Example 1. The class D of Dyck paths are paths only taking steps in $\{-1, 1\}$ starting and ending at 0 and staying positive all along the trajectory. We define $\mathcal{D}(x)$ the generating function of Dyck paths with size funcion the half of the steps. Notice that a path of size n can be decompose as an upstep continued by a Dyck path, then a downstep and finally a dyck path again. If $n = 0$ this decomposition is not valid so we have to treat it separately, but notice that $D_0 = 1$. This let us stablish the following formula for $\mathcal{D}$ the generating function of Dyck paths according to size

$$x^n D_n = \sum_{k=0}^{n-1} x^{1/2} \cdot x^k D_k \cdot x^{1/2} \cdot x^{n-1-k} D_{n-1-k}$$
$$\mathcal{D}(x) = x\mathcal{D}(x)^2 + 1$$

whose solution is given by

$$\mathcal{D}(x) = \frac{1 - \sqrt{1-4x}}{2x} = 1 + x + 2x^2 + 5x^3 + 14x^4 + 42x^5 + 132x^6 + O(x^7)$$

Here we discarded the solution with + since we look for the sign that gives an expansion where the coefficients are natural numbers, i.e. such that D_n counts the number of elements of size n. This generating function converges when $|x| < 1/4$, where D_n the number of Dyck paths of length $2n$ is given by the famous Catalan numbers

$$D_n = \frac{1}{n+1}\binom{2n}{n}$$

For probabilistic purposes, we consider the following generalization of a generating function, called the *Weighted Generating Function*

Definition 2. Consider a class A together with a weight function $W : A \to \mathbb{R}$, the weighted generating function $\mathcal{A}$ of A is given by

$$\mathcal{A}(W) = \sum_{a \in \mathsf{A}} W(a)$$

We will revisit this example

Example 2. Consider an unoriented graph $G = (V, E)$, a rooted tree t of G is a subgraph with a marked vertex such that there exists a unique path for each pair of poins (this is equivalent to being connected and with no cycles). A rooted spanning tree of G is a rooted tree containing all the vertices of G. When a tree t is rooted at r one can associate canonical orientation $\vec{t}$: orient the edges of the tree in the direction that points towards r the root. Consider $\mathsf{ST}(G, r)$ the set of rooted spanning trees of G rooted at r. Consider a weight function $W : \cup_{r \in V(G)} \mathsf{ST}(G, r) \to \mathbb{R}$ for $t \in \mathsf{ST}(G, r)$ as follows

$$W(t, r) = \prod_{\vec{e}=(e_1,e_2)\in\vec{t}} K_{\vec{e}}$$

For $K : E \to \mathbb{R}$ an edge weight function (a matrix of size $|V| \times |V|$). This gives a weighted generating function SG

$$SG(K; r) = \sum_{t \in T(G,r)} W(t, r)$$

Here the variable x has a value depending on each oriented edge, this is K_e.

1.2 Heaps

A heap of pieces is, informally, a collection of pieces, that are placed on a discrete space ($\mathcal{B} \times \mathbb{N}$, where $\mathcal{B}$ is a set of elements, and $\mathbb{N}$ is the height space). The definition uses a reflexive and symmetric relation $\mathcal{R}$ on the set of pieces $\mathcal{B}$. Some pieces are said to be in relation, which implies that they cannot be placed at the same height (if $p\mathcal{R}p'$, then (p, i) and (p', i) cannot belong to the same heap); moreover, a piece p at height i with $i \geq 1$ must be supported by a piece p at height $i - 1$, which is related to it (that is, if (p, i) is in a heap H and $i \geq 1$, then H must contain $(p', i - 1)$ with $p'\mathcal{R}p$).

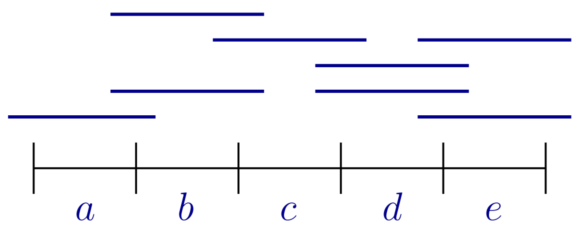

Figure 1: An example of heap of pieces.

Definition 3 (Geometric Viennot [Vie86]). A heaps of pieces H is a finite sets of pairs $\{(p, i) : p \in \mathcal{B}, i \in \mathbb{N}\}$, such that

1. If $(p, i), (p', j) \in H$ and $p\mathcal{R}p'$, then $i \neq j$ (pieces in relation cannot be put at the same height).
2. If $(p, i) \in H$ and $i > 0$, then there exists $(p', i - 1) \in H$ with $p\mathcal{R}p'$ (each piece must be supported).

There are alternative ways to define formally the notion of heap of pieces:

- As a Partially Commutative Monoid (Viennot [Vie86], Krattenthaler [Kra06]).
- And as a set where each Heap can be viewed as a Partially Ordered Set (POSET).

We will stick to the geometric point of view in what follows.

Definition 4. A *trivial heap* of pieces is a heap in which all the pieces are at level 0, i.e. the pieces it contains are not in relation.

Definition 5. A piece (p, i) in H is said to be *maximal* in H, if H does not contain any pair (p', j) such that $p'\mathcal{R}p$ and $j \geq i$, i.e. there are no pieces in relation, above it.

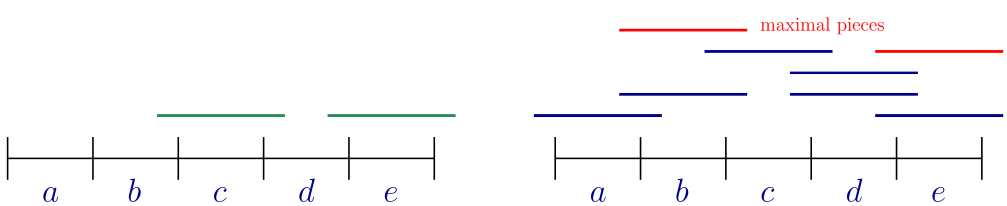

Figure 2: Left: Trivial heap of pieces T. Right: Heap of pieces H with maximal pieces marked.

Theorem 1. *[Prop.5.3 [Vie86]] Let $\mathcal{M}$ be a subset of the pieces $\mathcal{B}$. Let W be a multiplicative weight function on heaps, such that for all heap H its weight $W(H)$ is the product of the elementary weights $W(p)$ of the pairs (p, i) it contains (the weight of a piece is independent of "its place or height" in the heap). Then, the total weight of the heap of pieces having their maximal pieces included in $\mathcal{M}$ is given by*

$$\sum_{\substack{H \text{ heaps in } (\mathcal{B},\mathcal{R}) \\ \text{maximal pieces} \subset \mathcal{M}}} W(H) = \Bigg(\sum_{\substack{T \text{ trivial} \\ \text{heap in } (\mathcal{B},\mathcal{R})}} (-1)^{|T|} W(T) \Bigg)^{-1} \Bigg(\sum_{\substack{T \text{ trivial} \\ \text{heap in } (\mathcal{B} \setminus \mathcal{M},\mathcal{R})}} (-1)^{|T|} W(T) \Bigg),$$

where $|H|$ denotes the number of pieces in the heap H.

Viennot [Vie86, Proposition 5.3] gave this result at the level of combinatorial objects; here, we preferred a projected version, in terms of their weights (which is what we need). (See also Theorem 4.1 [Kra06]).

Proof. The statement is equivalent to

$$\sum_{\substack{H \text{ heaps in } (\mathcal{B},\mathcal{R}) \\ \text{maximal pieces} \subset \mathcal{M}}} \ \sum_{\substack{T \text{ trivial} \\ \text{heap in } (\mathcal{B},\mathcal{R})}} (-1)^{|T|} W(H) \times W(T) = \Big(\sum_{\substack{T \text{ trivial} \\ \text{heap in } (\mathcal{B}\setminus\mathcal{M},\mathcal{R})}} (-1)^{|T|} W(T) \Big) \tag{1}$$

And notice that the term $(-1)^{|T|} W(H) \times W(T) = (-1)^{|T|} W(T \circ H)$, for $T \circ H$ denoting the heap obtained by putting H on top of T and then letting fall the pieces that are not supported.

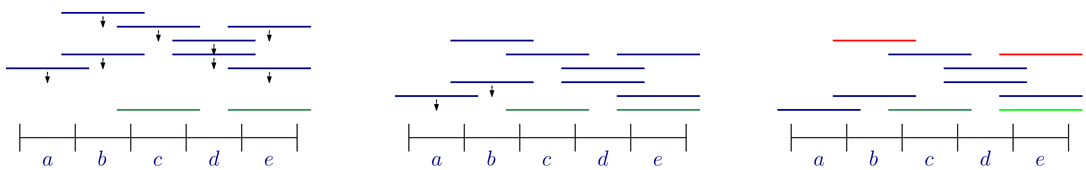

Figure 3: Left and middle: Putting H (dark blue) on top of T (green) at letting the pieces fall. Right: the full transformation $T \circ H$ after letting fall the pieces of H maximal pieces in red and minimal piece (with respect to the total order in $\mathcal{B}$) b supported a maximal piece on $\mathcal{M}$ in green .

The idea of the proof will be to consider pairs of heaps (H, T) and (H', T') such that $H \circ T = H' \circ T'$ and such that $(-1)^{|T|} W(T \circ H) = -(-1)^{|T'|} W(T' \circ H')$. We will obtain this with an involution.

We start by considering a total order of the pieces in $\mathcal{B}$ and we denote by b the piece which is minimal w.r.t. the total order in $\mathcal{B}$ which is contained in $T \circ H$ and at level 1 such that it supports a maximal piece on $\mathcal{M}$. From (H, T) we form a new pair (H', T') as follows :

1. if $b \in T$, we set $T' = T \setminus \{b\}$ and $H' = b \circ H$.

2. if $b \notin T$, then $T' = T \circ b$ and $H' = H \setminus b$.

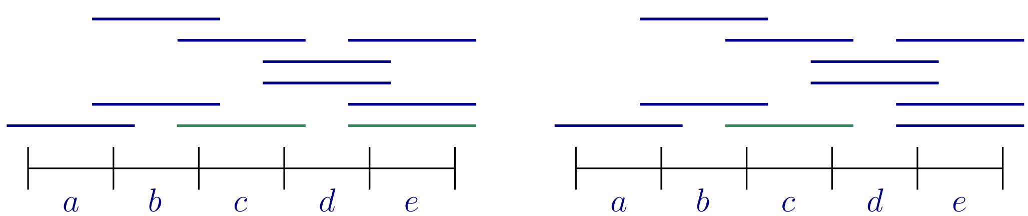

Figure 4: Left: $T \circ H$ with T in green and H in dark blue. Right: $T' \circ H'$ with T' in green and H' in dark blue .

Notice that $T \circ H = T' \circ H'$ and that $(-1)^{|T|} W(T \circ H) = -(-1)^{|T'|} W(T' \circ H')$, moreover this transformation is an involution, meaning that applying it twice gives the identity, i.e. $H'' = H$ and $T'' = T$.

Now it remains to retrieve the elements that survive, since the weight associated to $T \circ H$ cancels out in the left hand side of equation (1) with the weight associated to $T' \circ H'$. To see this notice that our transformation works as soon as $T \circ H$ contains a maximal piece in M, therefore the elements resisting the transformation are such that $T \circ H$ does not have any maximal piece in $\mathcal{M}$, but this forces H to be the empty heap and T to have only pieces belonging to $\mathcal{B} \setminus \mathcal{M}$, since all the pieces in a trivial heap are maximal. This gives the right hand side of equation (1). □

Corollary 1. *Let W be a multiplicative weight function on heaps, such that for all heap H its weight $W(H)$ is the product of the elementary weights $W(p)$ of the pairs (p, i) it contains (the weight of a piece is independent* *of "its place or height" in the heap). Then, the total weight of the heap of pieces is given by*

$$\sum_{H \text{ heaps in } (\mathcal{B},\mathcal{R})} W(H) = \Big(\sum_{\substack{T \text{ trivial} \\ \text{heap in } (\mathcal{B},\mathcal{R})}} (-1)^{|T|} W(T) \Big)^{-1},$$

where $|H|$ denotes the number of pieces in the heap H.

Example 3. For $\mathcal{B} = \{a, b, c, d, e\}$ as in the example (meaning that we use the relation imposed there) with weight function given by $W(H) = x^{|H|}$. Which is the total weight of all possible heap of pieces? From the precedent corollary one has that the total mass is given by the alternating mass of trivial heaps, this gives

$$\sum_{H \text{ heaps in } (\mathcal{B},\mathcal{R})} W(H) = \frac{1}{1 - 5x + 6x^2 - x^3} = 1{+}5x{+}19x^2{+}66x^3{+}221x^4{+}728x^5{+}2380x^6{+}7753x^7{+}25213x^8{+}O(x^9)$$

Since there are 1 heap with no pieces, 5 with one piece, 6 with two pieces and 1 with three pieces.

2 Probabilities

2.1 The quantity $\det\left(\mathrm{Id} - M^{(r)}\right)$ is a central object in probability;

The quantity $\det\left(\mathrm{Id} - M^{(r)}\right)$ is the partition function of the family of weighted spanning trees rooted at r, but not only. It will appear under the following forms in what follows.

2.1.1 Heap of cycles with weights given by a Markov Chain

In a heap of cycles, the pieces are oriented cycles on a given graph G, and two cycles are in relation if they share a vertex. We consider the weight of an oriented cycle C as the transitions following M over the directed edges it contains, i.e.

$$W(\vec{C}) = \prod_{\vec{e} \in C} M_{\vec{e}};$$

and the weight of a heap of cycles is the weight of the collection of cycles it contains. A heap of cycles is then trivial when all the cycles it contains are non-intersecting. Denote by A_G the alternating weight of trivial heaps of cycles

$$A_G = \sum_{C=(C(1),\cdots,C(|C|)) \in \text{Trivial heap of cycles on } G} (-1)^{|C|} \prod_{j=1}^{|C|} \mathsf{Weight}(C(j)),$$

The counting of trivial heap of cycles can be easily be obtained thank to the interpretation using permutations, since a permutation σ over a set of vertices can be seen as a collection of non intersecting cycles C defined as $C = \{(v, \sigma(v)) : v \in V\}$ this interpretation gives for cycles of length bigger than 1 (no self-loops) that

$$(-1)^{|C|} \prod_{j=1}^{|C|} \mathsf{Weight}(C(j)) = \mathrm{sgn}(\sigma) \prod_{v \in V} (-M_{v,\sigma(v)})$$

This follows since the signature of an even cycle is -1 and the signature of an odd cycle is $+1$; meaning that $\mathrm{sgn}(\sigma) = (-1)^{\text{even cycles of } \sigma}$, but when considering $-M_{\vec{e}}$ instead of $M_{\vec{e}}$ gives to each odd cycle a total weight of -1, giving finally that $\mathrm{sgn}(\sigma) = (-1)^{|C|}$. While for cycles of length 1 one should impose a diagonal term equals to $(1 - M_{i,i})$ since it counts : with weight $-M_{i,i}$ when the cycle (loop of vertex i) is present in the collection and 1 when it is not. From this interpretation we deduce that

$$A_G = \det\left(\mathrm{Id} - M\right) = 0,$$

since M has 1 as eigenvalue; hence the set of heaps of cycles on G has total weight $+\infty$. What is of greater interest is the value of $A_{G\setminus S}$, the alternating weight of trivial heaps of cycles avoiding some set of vertices S, which is

$$A_{G\setminus S} = \det\left(\mathrm{Id} - M^{\{S\}}\right), \tag{2}$$

as well as its inverse corresponding to the total weight of heaps of cycles on $G \setminus S$:

$$\sum_{H \in \text{Heap of Cycles on } G\setminus S} \text{Weight}(H) = \det\left(\mathrm{Id} - M^{\{S\}}\right)^{-1}. \tag{3}$$

2.1.2 Matrix tree theorem.

Recall that for a rooted tree (t, r) we set the weight as

$$W(t,r) = \prod_{\vec{e}=(e_1,e_2)\in\vec{t}} K_{\vec{e}}$$

For $K : E \longrightarrow \mathbb{R}$ an edge weight function (a matrix of size $|V| \times |V|$).

Theorem 2 (Matrix Tree Theorem)**.**

$$\sum_{t\in T(G,r)} W(t,r) = \det((D-K)^{\{r\}})$$

Where D is the diagonal matrix with diagonal values equal to the sum of the line of K and where $(D-K)^{\{r\}}$ is the matrix $(D-K)$ without the line and column associated to r.

The Matrix $D - K$ is called a Laplacian matrix since it gives information about the flow of a vertex. More formally

$$((D-K)x)_u = D_{u,u}x_u - \sum_{v\in V} K_{u,v}x_v = \sum_{v\in V} K_{u,v}(x_u - x_v)$$

Imposing this equals to 0 says that the flow in and the flow out compensate.

Proof. For this proof we borrow the argument in Zeilberger [Zei85, Section 4]. Observe that

$$D - K^{\{r\}} = \left[\left(\textstyle\sum_{v'\in V} K_{u,v'}\right) \mathbf{1}_{(u=v)} - K_{u,v}\right]_{(u,v)\in(V\setminus\{r\})^2}. \tag{4}$$

Denote by $\mathcal{B}$ the set of pairs (B, C) such that:
• B is a directed graph on V, where each vertex of $V \setminus \{r\}$ has either 0 or 1 outgoing edge ending in V (including $\{r\}$ this time). Denote by V_B the set of vertices from which there is one outgoing edge.
• C is a collection of directed disjoint cycles on $V_C = (V \setminus \{r\}) \setminus V_B$.

Set $\text{Weight}(B) := \prod_{u\in V_B} K_{u,t(u)}$ where $t(u)$ is the target of the edge starting at u and $\text{Weight}(C) := (-1)^{|C|} \prod_{\vec{c} \text{ cycles of C}} \prod_{\vec{e}\in\vec{c}} K_{\vec{e}}$ the product of the weight of edges along the directed cycles of C and finally define $\text{Weight}(B, C) := \text{Weight}(B)\text{Weight}(C)$.
Claim: $\text{Weight}(\mathcal{B}) := \sum_{(B,C)\in\mathcal{B}} \text{Weight}(B, C) = \det(D - K^{(r)})$.
Proof of the claim : first, expand $\det(D - K^{(r)})$ using Leibniz formula:

$$\det(D-K^{(r)}) = \sum \underbrace{(-1)^{\text{sign}(\sigma)}}_{(-1)^{|\text{even cycles of } \sigma|}} \prod_{i\neq r} (D-K^{(r)})_{i,\sigma(i)},$$

where the sum range on all permutations σ on $V \setminus \{r\}$. Now, consider the set $F(\sigma) = \{i : \sigma(i) = i\}$ of fix points of σ, and rewrite:

$$\prod_{i\neq r}(D-K^{(r)})_{i,\sigma(i)} = \left(\prod_{i\in F(\sigma)}\left(-K_{i,i} + \sum_{j\in V} K_{i,j}\right)\right)\left(\prod_{i\in V\setminus(\{r\}\cup F(\sigma))} -K_{i,\sigma(i)}\right).$$

The second parenthesis can be interpreted as the weight of cycles of σ with lengths at least 2. Now expand the first parenthesis (without simplifying the diagonal term). This first parenthesis can be rewritten as a sum over $A \subset F(\sigma)$ as follows:

$$\prod_{i\in F(\sigma)}\left(-K_{i,i}+\sum_{j\in V}K_{i,j}\right)=\sum_{A\subset F(\sigma)}\left(\prod_{i\in A}(-K_{i,i})\right)\left(\prod_{i\in F(\sigma)\setminus A}\sum_{j}K_{i,j}\right).$$

Each factor $-K_{i,i}$ can be seen to be the weight of a loop over i (that is a cycle of size 1), and by expansion, $\prod_{i\in F(\sigma)\setminus A}\sum_{j\in V}K_{i,j}$ can be interpreted as the sum of weights of directed graph where each vertex of $F(\sigma)\setminus A$ has a single outgoing edge, ending on any vertex of $V\setminus\{A\}$. This ends the argument explaining why $\mathrm{Weight}(\mathcal{B})=\det(D-K^{(r)})$.

Now we return to the proof of the matrix tree theorem.
We show now that $\mathrm{Weight}(\mathcal{B})=\sum_{(t,r)\in\mathbb{T}(G,r)}\prod_{\vec{e}\in\vec{t}}K_{\vec{e}}$. The graphs "$B$" are made of cycles and trees, and C is made of cycles. For any pair (B,C) having (totally) at least one cycle one can define (B',C') as follows: for a total order on the set of oriented cycles, take the greatest cycle c in the union of B and C. Denote by (B',C') the pair obtained by moving c from the component containing it to the other. This map $(B,C)\to(B',C')$ is clearly an involution and satisfies $\mathrm{Weight}(B',C')=-\mathrm{Weight}(B,C)$. Hence, $\mathrm{Weight}(\mathcal{B})$ coincides with the sum of the $\mathrm{Weight}(B,C)$ taken on the set of pairs (B,C) which have no cycles: C is empty, and the graph B has no cycle, and since its number of edges is one less than its number of vertices, it is a spanning tree. □

In the case where K is a transition matrix M, one has that the previous theorem takes the form of

$$\sum_{t\in T(G,r)}W(t,r)=\det((\mathrm{Id}-M)^{\{r\}})$$

2.1.3 Reversibility

Consider M and ρ one invariant measure of it. We denote by $\overleftarrow{M}$ the reversibility Kernel of the transition matrix M, which is defined as

$$\overleftarrow{M}_{i,j}=\frac{\rho_j}{\rho_i}M_{j,i}.$$

By considering the involution that changes the orientation of every cycle one gets

$$\det(\mathrm{Id}-M^{(r)})=\det(\mathrm{Id}-\overleftarrow{M}^{(r)})$$

Remark 1. For this formula, the presence of ρ is an "illusion" : this can be seen from the description over oriented cycles, since for an oriented cycle $\vec{c}=x_0x_1x_2x_3$ one has

$$\overleftarrow{M}_{x_0,x_3}\overleftarrow{M}_{x_3,x_2}\overleftarrow{M}_{x_2,x_1}\overleftarrow{M}_{x_1,x_0}=\underbrace{\frac{\rho_{x_3}}{\rho_{x_0}}\frac{\rho_{x_2}}{\rho_{x_3}}\frac{\rho_{x_2}}{\rho_{x_1}}\frac{\rho_{x_1}}{\rho_{x_0}}}_{=1}M_{x_0,x_1}M_{x_1,x_2}M_{x_2,x_3}M_{x_3,x_0}$$

So that we can always use the constant measure which is always invariant for any transition Matrix, since the sum of rows is equal to 1.

Therefore for a fixed $r\in V$

$$\sum_{(t,r)\in T(G}\prod_{\vec{e}\in\vec{t}}M_e=\sum_{(t,r)\in T(G)}\prod_{\vec{e}\in\vec{t}}\overleftarrow{M}_e.$$

Even though the terms associated to each tree (t,r) may be different, these sums are equal.

2.1.4 Invariant measure

An invariant measure ρ of the Markov kernels M and $\overleftarrow{M}$ is related to these quantities by:

$$\rho(w) = \text{Const.}\det(\text{Id} - M^{(w)}) = \text{Const.}\det(\text{Id} - \overleftarrow{M}^{(w)}) \tag{5}$$

so that, from the matrix tree theorem, this provides a connection between ρ_w and the total mass of spanning trees rooted at w. To prove (5), there are several methods, direct arguments exist, but we will present one for the irreducible case using Aldous-Broder theorem.

A path is a sequence of vertices $w = (w_k, 0 \le k \le m)$, with the property that $\{w_k, w_{k+1}\} \in E$ for every k. This path is said to be covering if it visits all the vertices and if m is the first time with this property. More formally, for every $1 \le k \le |V|$ define

$$\tau_k(w) = \inf\{j, |\{w_0, \cdots, w_j\}| = k\}, \quad 1 \le k \le |V|,$$

the first time the path has visited k different points (we write τ_k instead of $\tau_k(w)$ when it is clear from the context). The path w is then called covering if $\tau_{|V|}(w) = m$.

Denote by $\mathsf{ST}(G)$ the set of spanning trees t of G, and by $\mathsf{ST}^\bullet(G)$ the set of rooted spanning trees (t, r), where r, the root, is a distinguished vertex of V.

Definition 6. For a covering path w, denote by $\mathsf{FE}(w)$ (First entrance tree of w) the rooted spanning tree (t, w_0) whose $|V| - 1$ edges are $(w_{\tau_k}, w_{\tau_k - 1})$ (that is oriented towards the root w_0) for $2 \le k \le \tau_{|V|}$.

In simple words, we start a walk at w_0 and each time a new vertex is discovered we add the edge used to discover it to the current tree (it is connected) pointing towards w_0 and we stop the walk the first time every vertex has been visited.

Theorem 3. *[Aldous–Broder reversible case [FM23] and [HLT21]] Let W be a Markov chain with irreducible kernel M and invariant distribution ρ. Then, for any $(t, r) \in ST^\bullet(G)$*

$$\mathbb{P}\left[FE\left((W_0, \cdots, W_{\tau_{|V|}}\right) = (t, r) \mid W_0 = r\right] = \text{Const.}\left[\prod_{\vec{e} \in \vec{t}} \overleftarrow{M}_{\vec{e}}\right] / \rho(r). \tag{6}$$

This theorem was originally proved by Aldous [Ald90] and Broder [Bro89] independently for the reversible case.

Summing over all possible trees one obtains that

$$\rho(w) = \text{Const.}\left[\sum_{t \in \mathsf{ST}(G,w)} \prod_{\vec{e} \in \vec{t}} \overleftarrow{M}_e\right] = \text{Const.}\det(\text{Id} - \overleftarrow{M}^{(w)}) = \text{Const.}\det(\text{Id} - M^{(w)})$$

3 Results

3.1 Kemeny's constant

Consider a Markov chain Y following M an irreducible transition matrix. The following result motivates a work in progress

Theorem 4 (Kemeny's Theorem)**.** *Let M be an irreducible transition matrix in finite state space. The following quantity is constant*

$$K = K_i = \mathbb{E}_i(\tau_X^+) = \sum_j \mathbb{E}_i(\tau_j^+)\pi_j, \forall i$$

where $\tau_j^+ = \inf\{k > 0 : Y_k = j\}$ and the r.v. X is distributed according to π the invariant measure of M.

This idea came from the fact that usually when working with generating functions

$$A(x)=\sum_{k\in\mathbb{N}} a_k x^k \implies \frac{d}{dx}A(x)\Big|_{x=1}=\sum_{k\in\mathbb{N}} k a_k x^{k-1}\Big|_{x=1}=\sum_{k\in\mathbb{N}} k a_k$$

and this can be seen as the expectation of the length when taking the derivative evaluated at $x=1$ for a_k probabilities associated to objects of size k.

We know that $\pi_j = C\det(I-M^{\{j\}})$, where C is the renormalisation constant; therefore, Kemeny's constant can be obtained as :

$$\begin{aligned}
K_i &= \sum_j \mathbb{E}_i(\tau_j^+)\pi_j \\
&= \sum_j \mathbb{E}_i(\tau_j^+)\det(I-M^{\{j\}}) \\
&= C\sum_j \sum_{\ell\neq j} \sum_{k=0}^{\infty}(k+1)\left(\tilde{M}^{(j)}\right)^k_{i,\ell} M_{\ell,j}\det(I-M^{\{j\}}) \\
&= C\frac{d}{dx}\left(\sum_j\sum_\ell\sum_{k=0}^{\infty}\left(x\tilde{M}^{(j)}\right)^k_{i,\ell} xM_{\ell,j}\det(I-M^{\{j\}})\right)\Bigg|_{x=1} \\
&= C\frac{d}{dx}\left(\sum_j\sum_\ell\left(I-x\tilde{M}^j\right)^{-1} xM_{\ell,j}\det(I-M^{\{j\}})\right)\Bigg|_{x=1}
\end{aligned}$$

where $\tilde{M}^{(j)}$ is the matrix with zeroes on the lines and column associated to j and $\left(\tilde{M}^{(j)}\right)^0 = I$. This is in order to consider the possibility of reaching the state j after one step when $i=j$.

Define

$$\begin{aligned}
G_{i,j}(x) &= \sum_\ell\sum_{k=0}^{\infty}\left(xM^{(j)}\right)^k_{i,\ell} xM_{\ell,j} = \sum_{\ell\neq j}(\mathrm{Id}-xM^{(j)})^{-1}_{i,\ell}xM_{\ell,j} \\
\pi_j(x) &= C\rho_j(x) = C\det(I-xM^{\{j\}}) \\
\mathcal{K}_i(x) &= \sum_j G_{i,j}(x)\pi_j(x)
\end{aligned}$$

Where in the last line to say that $\pi_j(1)$ is the invariant probability and that $\rho(x)$ does not take into account the renormalisation. Kemeny's constant rewrites as

$$K = C\sum_j G'_{i,j}(1)\rho_j(1)$$

Theorem 5 (F. and Marckert '24+)**.** *The following generating series does not depend on i*

$$\mathcal{K}_i(x)=\mathcal{K}(x)=\frac{x}{1-x}\det(I-xM)$$

We call this series the Kemeny's generating series.

This theorem does not directly imply Kemeny's Theorem. We show that Kemeny's theorem is a corollary of Theorem 5 in some simple steps. The derivative of Kemeny's series, which is independent of i, gives

$$\frac{d}{dx}\mathcal{K}(x)=\sum_j G'_{i,j}(x)\rho_j(x)+\sum_j G_{i,j}(x)\rho_j(x)'$$

when evaluating at $x=1$ one obtains

$$\frac{d}{dx}\mathcal{K}(x)\Big|_{x=1} = \underbrace{\sum_j G'_{i,j}(1)\rho_j(1)}_{\frac{K}{C}} + \sum_j \underbrace{G_{i,j}(1)}_{=1}\rho_j(x)'\Big|_{x=1}$$

where $G_{i,j}(1) = 1$ he second summand does not depend on i, since it mesures the probability of eventually arriving at j from i; this probability is equal to one by positive recurrence. We conclude by noticing that the left hand side does not depend on i by Theorem 5 and that the rightmost term of the right hand side is independent of i, implying Kemeny's theorem.

Proof of Theorem 5. Notice that

$$\det(I - xM^{(j)}) = \sum_{C\text{trivial heap of cycles}} (-1)^C \prod_{c\in C} W(c)$$

can be seen as the alternating generating function of trivial heaps of cycles, where each cycle/piece has weight according $W(c) = \prod_{\vec{e}\in c} xM_{\vec{e}} = x^{|c|} \prod_{\vec{e}\in c} M_{\vec{e}}$, where $|c|$ is the length of the cycle.

We start by noticing that from the definition of $\mathcal{K}_i(x)$ it consists on the sum over j of pairs formed by :

- A weighted path P starting at i and reaching for the first time j at the last step, with weight equals to xM for the transitions (from $G_{i,j}(x)$).
- A trivial heap of cycles C' avoiding j counted by the alternating series, i.e. each edge adds a -1 (from $\pi_j(x) = \det(\mathrm{Id} - xM^{\{j\}})$).

We interpret this as triplet :

- A self avoinding walk S from i to j weighted according to xM.
- A heaps of cycles C with maximal piece supported on S.
- A trivial heap of cycles C' avoiding j counted by the alternating i.e. each edge adds a -1 (from $\pi_j(x) = \det(\mathrm{Id} - xM^{\{j\}})$).

We define as $SCC_{i,j}(x)$ the generating function of the triplets (S, C, C').

We found in an analysis term by term of $SCC_{i,j}(x)$, that the only terms resisting in this generating function are the ones consisting of (S, C'') where S is as before and C'' is an alternating heap of cycles not supported on S. We denote as $H_{i,j}(x)$ the generating function of the tuples (S, C''). We have that

$$SCC_{i,j}(x) = H_{i,j}(x).$$

(This follows from an argument similar to the one used in the Matrix tree theorem)

We noticed that

$$\frac{\det(I - xM)H_{i,j}(x)}{\det(I - xM)} = \det(I - xM)I_{i,j}(x),$$

where $I_{i,j}(x)$ is the generating function of all the paths from i to j according to the number of steps.

(This follows from an argument as in the theorem of the generating function of heaps with maximal pieces in terms of trivial heaps)

From this result we obtain that

$$\sum_j \det(I - xM)I_{i,j}(x) = \det(I - xM) \sum_j I_{i,j}(x) = \det(I - xM)A_i(x)$$

where $A_i(x)$ is the generating function of all the paths starting at i of length at least one. We remarked that $A_i(x) = \sum_{k\geq 1} a_k x^k$ has coefficients (a_k) that are all equal to 1 : this follows since M^k is also a transition matrix and this is deduced from the discrete Chapman-Kolmogorov equations. From this we conclude that $A_i(x) = x/(1-x)$ and the result follows. □

We are currently working on new ways to compute Kemeny's constant, together with other combinatorial quantities that do not depend on the initial point.

Another important fact is that if we denote by $\lambda_1 = 1$ and λ_i the i-th biggest eigenvalue of M one gets

$$\mathcal{K}(x) = x \prod_{i\neq 1} (1 - x\lambda_i)$$

This follows since $\det(I - xM) = (1-x)\prod_{i\neq 1}(1 - x\lambda_i)$.

3.2 Upper directed random walks on trees

Let T be a finite or infinite tree rooted at $\varnothing$. We denote by $p(u)$ the parent of u in T when $u \neq \varnothing$ and by T_u the tree of descendants of u rooted at u (see Figure 5) and the set of children of u in T is denoted by $c_T(u)$.

Definition 7. Let T be a finite or infinite tree . A transition matrix $\mathsf{U} = \left(\mathsf{U}_{u,v}\right)_{u,v\in T}$ is said to be almost upper-directed , "AUD " for short, if

$$\forall u, v \in T, \quad \mathsf{U}_{u,v} > 0 \implies (v \in \{p(u)\} \cup T_u)$$

that is, if either v is the parent $p(u)$ of u or a descendant of u. Since $u \in T_u$, $\mathsf{U}_{u,u}$ is allowed to be positive.

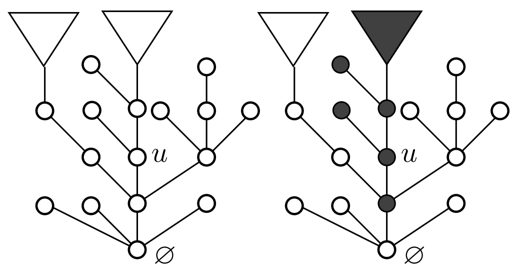

Figure 5: The node $\varnothing$ is the root, and u a node. Triangles represent finite or infinite subtrees, and disks, nodes. On the second picture, only the transitions from u toward dark nodes are possible for almost upper-directed transitions matrices.

For a transition matrix U, a node a, and a set of nodes B, we let $\mathsf{U}_{a,B} = \sum_{b\in B} \mathsf{U}_{a,b}$. The tree T_a is the subtree of T rooted at a. For a node u in T and $v \in [\![\varnothing, p(u)]\!]$, the successor of v in the direction of u is denoted by (v, u) (this is the unique child of v on the path $[\![\varnothing, u]\!]$).

Theorem 6 ([FM24]). *Consider* U *an irreducible* AUD *transition matrix of a finite or infinite tree* T. *For any node* u *of* T, *consider the weighted graph* $({}^uG, {}^u\mathsf{U})$ *with set of nodes* ${}^uV := [\![\varnothing, u]\!]$, *and in which the weight of*

$${}^u\mathsf{U}_{a,b} = (\mathsf{U}_{a,T_b} - \mathsf{U}_{a,T_{(b,u)}})\boldsymbol{1}_{b\in[\![p(a),p(u)]\!]} + \boldsymbol{1}_{b=u}\mathsf{U}_{a,T_u}$$

is defined for all $a \in [\![\varnothing, p(u)]\!]$ *and* $b \in [\![\varnothing, u]\!]$. *Set*

$$\pi(u) = \pi(\varnothing)\frac{\mathsf{Weight}([\![\varnothing, u]\!], {}^u\mathsf{U}, u)}{\displaystyle\prod_{v\in]\!]\varnothing,u]\!]} \mathsf{U}_{v,p(v)}} = \pi(\varnothing)\frac{\det((\mathrm{Id} - {}^u\mathsf{U})^{(u)})}{\displaystyle\prod_{v\in]\!]\varnothing,u]\!]} \mathsf{U}_{v,p(v)}}. \tag{7}$$

1. *The measure* π *is an invariant measure of* U *(with positive coordinates).*

2. *An* AUD *transition matrix is positive recurrent if and only if* $\sum \pi(u) < +\infty$.

Proof. ▪ Finite tree case. Assume first that U is an AUD irreducible transition matrix on a finite tree T. In this finite state case, there is uniqueness of the invariant measure and positive recurrence.

It remains to prove that (7) gives the invariant measure of U. We will use the matrix tree theorem and (5), and then define for $u \in T$, the set $\mathsf{ST}(T, E_{\mathsf{U}}, u)$ of spanning trees t, with root u, on the graph with vertex set $V = T$, and set of edge sets corresponding to (almost upper directed) possible transitions $E_{\mathsf{U}} := \{(i, j) \in T, \mathsf{U}_{i,j} > 0\}$. The edges of a tree are oriented towards its root: thus, for a spanning tree rooted at u, there is an edge going out from each vertex, except from u.

⋆ An illustration in Figure 6 could help the reader to understand more quickly the argument.

Consider $[\![\varnothing, u]\!]$ the ancestral line of u in T. We claim that $\mathsf{ForcedEdges}(u)$, the set of edges $(v, p(v))$ for all $v \in (T \setminus [\![\varnothing, u]\!])$, is a subset of the edge set of any (τ, u) in $\mathsf{SP}(T, E_{\mathsf{U}}, u)$. Indeed since the edges

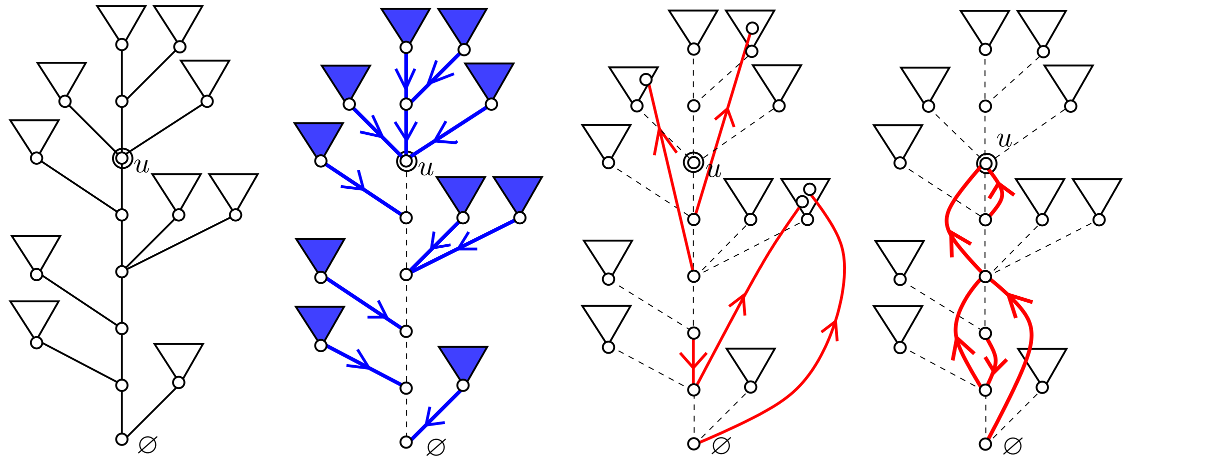

Figure 6: Each triangle represents a finite or infinite subtree. On the second picture, the blue edges (including those, not drawn, in the subtrees) are all oriented downwards: these edges are in all spanning trees rooted at u: the starting point of each oriented edge is any node which is not in $[\![\varnothing, u]\!]$. The set of blue edges forms connected components, that are subtrees, each of them rooted a different vertices of $[\![\varnothing, u]\!]$ (this is the forest $F_u = (f_v, v \in [\![\varnothing, u]\!])$). On the third picture, an example of what could be the set of edges going out from $[\![\varnothing, u]\!]$ of a spanning tree rooted at u: either directed to the parent, or to a node among their descendants. On the last picture, the "spanning tree condition" on $[\![\varnothing, u]\!]$. If we redirect each red edge toward the root of the blue component containing its second extremities, this forms a (red) spanning tree of $[\![\varnothing, u]\!]$: this condition is necessary and sufficient for the blue edges and red edges to form, together, a spanning tree of the global tree, rooted at u.

of (τ, u) are oriented toward u, each edge $(v, p(v)) \in \mathsf{ForcedEdges}(u)$ is needed to get out of T_v. Now, color the edges of $\mathsf{ForcedEdges}(u)$ in blue, together with the vertices they contain. Each blue connected component f is a tree, and contains a unique element v of $[\![\varnothing, u]\!]$, which is the root of f. Thus, each node $v \in [\![\varnothing, u]\!]$ is the root of a blue tree f_v, where, f_v may be reduced to its root. Each node w of T belongs to a single tree $f_{R(w)}$ of the forest $F_u := (f_v, v \in [\![\varnothing, u]\!])$. In other words, $R(w)$ is the identity of the root of this tree, a node of $[\![\varnothing, u]\!]$.

Let us now characterize the other edges of a spanning tree $(\tau, u) \in \mathsf{SP}(T, E_{\mathsf{U}}, u)$, those starting from the nodes $[\![\varnothing, u[\![$.

Lemma 7. *Consider a set of edges, with one outgoing edge from each the elements of* $[\![\varnothing, u[\![$ *:*

$$\mathsf{AddEdges} := ((v, x(v)), v \in [\![\varnothing, u[\![)$$

where $x(v) \in T$ *for all* v*. The set* $\mathsf{AddEdges}$ *union with* $\mathsf{ForcedEdges}(u)$ *is the set of edges of an element* (τ, u) *of* $\mathsf{SP}(T, E_{\mathsf{U}}, u)$ *if the tuple* $[(v, R(x(v))), v \in [\![\varnothing, u[\![]$ *is a spanning tree* (τ', u) *of* $[\![\varnothing, u]\!]$*.*

Remark 2. The "if" in the lemma becomes an "if and only if" if all the edges weight U_{v_1,v_2} are positive between each node v_1 and its descendant v_2 (by hypothesis $\mathsf{U}_{v,p(v)}$ are always positive).

Proof. Given that the blue edges are already present in the forest, we need to understand the structure of the other edges of (τ, u), i.e., those whose starting point in $[\![\varnothing, u[\![$. From each node $v \in [\![\varnothing, u[\![$, there is a single outgoing edge $(v, x(v))$ and $x(v)$ must belong to an element $f_{R(x(v))}$ of the $(f_w, w \in [\![\varnothing, u]\!])$ (since these trees form a partition of T). The coefficient $\mathsf{U}_{v,x(v)}$ must be positive, so $x(v)$ is either a descendant of v, or its parent; from $x(v)$, following the outgoing blue edges forms a path leading to $R(x(v))$, but these edges are not the subject of this discussion, since they are already fixed (they are blue). Intuitively, since there is a fixed path from all the nodes of $f_{R(x(v))}$ toward $R(x(v))$, one can contract any tree f_w and consider it as a vertex $\bar{w}$ (which can be identified with w), locally, in the proof. Notice that if several extreme points $x(v_1)$ and $x(v_2)$ belong to the same tree f_w, then, since in f_w the blue paths are directed to the root w, these paths together do not form a cycle (the contraction does not destroy any cycle). The contraction is coherent with this point of view (the edges are then directed to $R(\bar{x}(v))$).

Once the contractions have been done, the new edges $(\bar{v}, R(\bar{x(v)}))$ must form a spanning tree (τ', u) of $(\bar{v}, \bar{v} \in [\![\emptyset, u]\!])$ (because if it is not the case, the union of AddEdges union with ForcedEdges(u) would be disconnected). □

Extract from a spanning tree (τ, u), the set of edges

$$\mathsf{SetRedEdges}(\tau, u) = \{(v, x(v)) \in E(\tau, u), v \in [\![\emptyset, u[\![\}$$

going out from $[\![\emptyset, u[\![$. Following Lemma 7 and remark 2, we denote by

$$\mathsf{SetOfSetRedEdges} = \{[(v, x(v)), v \in [\![\emptyset, u[\![] \text{ s.t. } [(v, R(x(v))), v \in [\![\emptyset, u[\![] \text{ is a spanning tree of } [\![\emptyset, u]\!]\}.$$

As said in remark 2, in general the set of all possible SetRedEdges(τ, u) is only contained in SetOfSetRedEdges, because SetOfSetRedEdges contains elements, that contain "red edges" with null weight. However, since we are now dealing with weights, adding elements with null weights amounts to adding negligible sets, and this is what we will do. Set

$$\mathsf{RedWeight}(u) = \sum_{\substack{[(v,x(v)),v\in[\![\emptyset,u[\![]\\ \in \mathsf{SetOfSetRedEdges}}} \prod_{v\in[\![\emptyset,u[\![} \mathsf{U}_{v,x(v)},$$

$$\mathsf{BlueWeight}(u) = \prod_{v\in T\setminus[\![\emptyset,u]\!]} \mathsf{U}_{v,p(v)}$$

this latter being the total weight of ForcedEdges(u). We have that

$$W(u) := \mathsf{RedWeight}(u) \times \mathsf{BlueWeight}(u)$$

is the total weight of all spanning trees rooted at u and we need to prove that $W(u) = \mathsf{Cst}.\pi(u)$ given in Equation 7.

In order to complete the proof in the finite case, we need three additional ingredients:

(a) Set $\mathsf{Factor}(u) = \prod_{v\in]\!]\emptyset,u]\!]} \mathsf{U}_{v,p(v)}$, the denominator in equation 7. Since $\{(v, p(v)) :]\!]\emptyset, u]\!]\}$ are the edges we need to add to ForcedEdges(u) to get all the edges of T oriented toward $\emptyset$, we have

$$\mathsf{Factor}(u) \times \mathsf{BlueWeight}(u) = \prod_{v\in T\setminus\{\emptyset\}} \mathsf{U}_{v,p(v)} = \mathsf{Cst}.$$

Hence since we work up to a multiplicative factor, we will use $1/\mathsf{Factor}(u)$ instead of BlueWeight(u), since the first remains under control when the tree is infinite.

(b) Let us compute RedWeight(u). From the "contraction" point of view, an edge from $a \in [\![\emptyset, u[\![$ to some node $\bar{b}$ represents the total weight of edges going from v to f_b: the tree f_b itself, equals to $T_b \setminus T_{s(b,u)}$ (the subtree T_b deprived from $T_{s(b,u)}$). The weight of all edges starting at a and with second extremity in f_b is then ${}^u\mathsf{U}_{a,b}$. By the matrix tree theorem we then have

$$\mathsf{RedWeight}(u) = \det(\mathsf{Laplacian}({}^u\mathsf{U})^{(u)})$$

(c) It remains to justify that we can choose the ${}^u\mathsf{U}_{a,a}$ as we wish: this is standard; in a weighted graphs, the weight of the spanning trees is independent of the weights of the loops, and the user is then free to change them according to his personal motivations... This is what we have done by imposing the values ${}^u\mathsf{U}_{a,a}$ such that the total weights $\sum_v {}^u\mathsf{U}_{a,v} = 1$ for all a. From this choice we find that

$$\det(\mathsf{Laplacian}({}^u\mathsf{U})^{(u)}) = \det((\mathrm{Id} - {}^u\mathsf{U})^{(u)})$$

This concludes the proof for the finite case.

▪ When T is an infinite (local finite) tree, write the balance equation at u for an invariant measure ρ (normalized so that $\rho(\emptyset) = 1$):

$$\rho(u) = \sum_{v\in[\![\emptyset,u]\!]} \rho(v)\mathsf{U}_{v,u} + \sum_{c\in c_T(u)} \rho(c)\mathsf{U}_{c,u}. \tag{8}$$

We want to prove that the measure $\rho = \pi$ given in (7) solves this equation. Thus, if we write the invariance equations for all nodes u below a given level h, then all these equations together involve only ρ and transitions up to level $h+1$. Let us cut down this tree!

Consider a transition matrix $\mathsf{U}^{(h+2)}$ on say, the finite tree $T^{(h+2)}$ of height $h+2$ coinciding with T up to this level, and define $\mathsf{U}^{(h+2)}_{u,v}$ as to be $\mathsf{U}_{u,v}$ if $|v| < h+2$, and $\mathsf{U}^{(h+2)}_{u,v} = \mathsf{U}_{u,T_v}$ if $|v| = h+2$. Then we get that $\pi^{(h+2)}$ the invariant measure of $\mathsf{U}^{(h+2)}$ on the finite tree $T^{(h+2)}$ satisfies, for all u such that $|u| \leq h+1$ the equation

$$\pi^{(h+2)}\mathsf{U}^{(h+2)} = \pi^{(h+2)}$$

and by uniqueness $\pi^{(h+2)}$, it is the h-invariant measure of $\mathsf{U}^{(h+2)}$ given in equation 7. For u such that $|u| \leq h+1$, the equation

$$\pi^{(h+2)}_u = \sum_v \pi^{(h+2)}_v \mathsf{U}^{(h+2)}_{v,u} \tag{9}$$

can be represented by the formula of π instead, on $T^{(h+2)}$; indeed, a quick inspection shows that since the weights involved in the computation of $\pi(u)$ using U and those for the computation of $\pi^{(h+2)}_u$ using $\mathsf{U}^{(h+2)}$ are the same (these are the transitions on the path on $[\![\varnothing, u]\!]$ and the weight toward the subtree f_v rooted on $[\![\varnothing, u]\!]$ thanks to the fact that we took the total $\mathsf{U}^{(h+2)}_{u,v} = \mathsf{U}_{u,T_v}$ if $|v| = h+2$).

From this, we deduce that $\pi_u = \sum_v \pi_v \mathsf{U}_{v,u}$ for all u such that $|u| \leq h$, and since h can be taken as large as wanted, for all $u \in T$. □

References

[Ald90] David J Aldous. The random walk construction of uniform spanning trees and uniform labelled trees. SIAM Journal on Discrete Mathematics, 3(4):450–465, 1990.

[Bro89] Andrei Z Broder. Generating random spanning trees. In FOCS, volume 89, pages 442–447, 1989.

[FM23] Luis Fredes and Jean-François Marckert. A combinatorial proof of aldous–broder theorem for general markov chains. Random Structures & Algorithms, 62(2):430–449, 2023.

[FM24] Luis Fredes and Jean-François Marckert. Markov chains on trees: almost lower and upper directed cases. arXiv preprint arXiv:2411.07158, 2024.

[HLT21] Yiping Hu, Russell Lyons, and Pengfei Tang. A reverse aldous–broder algorithm. 2021.

[Kra06] Christian Krattenthaler. The theory of heaps and the Cartier–Foata monoid, 2006.

[Vie86] Gérard Xavier Viennot. Heaps of pieces, i: Basic definitions and combinatorial lemmas. In Combinatoire énumérative, pages 321–350. Springer, 1986.

[Zei85] Doron Zeilberger. A combinatorial approach to matrix algebra. Discrete Mathematics, 56(1):61–72, 1985.