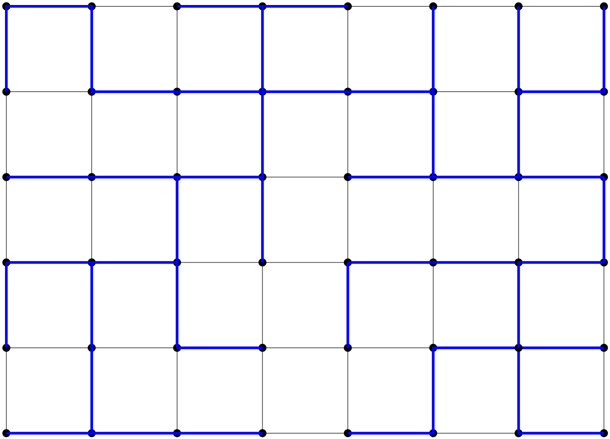
A new proof of the Aldous-Broder theorem

Luis Fredes
(Work with J.F. Marckert)

ERC GeoBrown
Journées MAS, August 2021

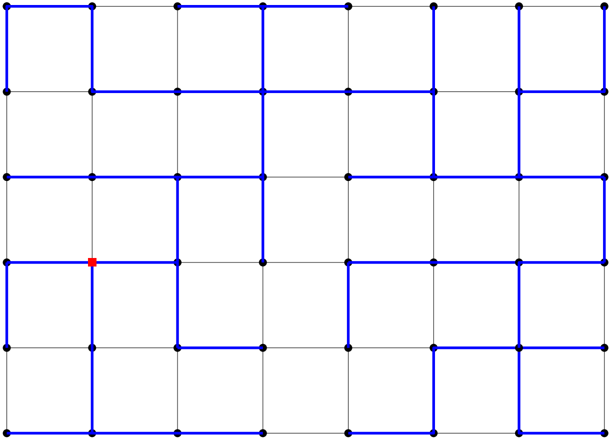
Definition (Spanning tree)

Given a graph G, we say that T is a spanning tree of G if it is a subgraph of G that is a tree containing all the vertices of G.



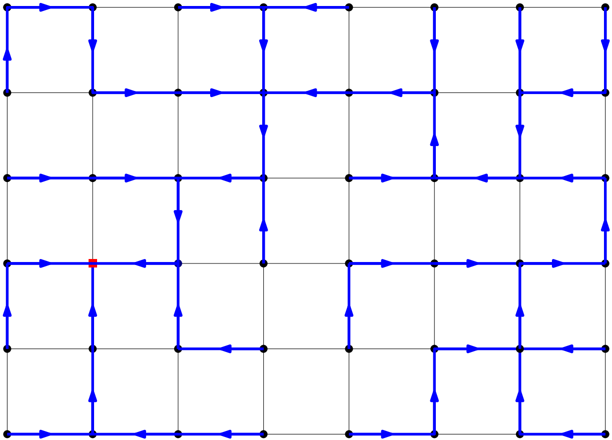
Definition (Spanning tree)

Given a graph G, we say that T is a spanning tree of G if it is a subgraph of G that is a tree containing all the vertices of G.



Definition (Spanning tree)

Given a graph G, we say that T is a spanning tree of G if it is a subgraph of G that is a tree containing all the vertices of G.



I. Combinatorics: Counting (weighted) spanning trees

ST(G)= set of spanning trees of G.

Matrix-tree theorem [Kirchhoff]

$$|\mathrm{ST}(G)| = \det\left(\mathrm{Laplacian}_G^{(r)}\right),$$

where $\mathrm{Laplacian}_G^{(r)}$ is the Laplacian matrix of G deprived of the line and column associated to r.

$$\mathrm{Laplacian}_G(i,j) = \left[deg(u_i)\mathbb{1}_{i=j} - |\{u_i, u_j\} \in E|\right]$$

I. Combinatorics: Counting (weighted) spanning trees

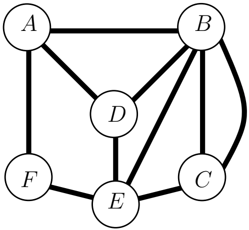
$\mathsf{ST}(G)$= set of spanning trees of G.

Matrix-tree theorem [Kirchhoff]

$$|\mathsf{ST}(G)| = \det\left(\mathsf{Laplacian}_G^{(r)}\right),$$

where $\mathsf{Laplacian}_G^{(r)}$ is the Laplacian matrix of G deprived of the line and column associated to r.

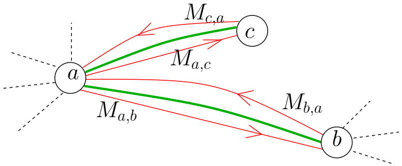
$$\mathsf{Laplacian}_G(i,j) = \left[deg(u_i)\mathbb{1}_{i=j} - |\{u_i, u_j\} \in E|\right]$$



$$\mathsf{Laplacian}_\mathsf{G} = \begin{pmatrix} 3 & -1 & 0 & -1 & 0 & -1 \\ -1 & 5 & -2 & -1 & -1 & 0 \\ 0 & -2 & 3 & 0 & -1 & 0 \\ -1 & -1 & 0 & 3 & -1 & 0 \\ 0 & -1 & -1 & -1 & 4 & -1 \\ -1 & 0 & 0 & 0 & -1 & 2 \end{pmatrix}$$

$$|\mathsf{ST}(G)| = \det\left(\mathsf{Laplacian}_G^{(A)}\right) = 98.$$

ST(G, r):= set of spanning trees of G rooted at r.
M := Markov kernel on G such that $\{u, v\} \in E \implies M_{u,v} > 0$ and $M_{v,u} > 0$.



$W(T, r) := \prod_{\vec{e} \in T} M_{\vec{e}}$ with edges pointing towards the root r.

Weighted Matrix-tree theorem [Kirchhoff]

$$\sum_{T \in \mathrm{ST}(G,r)} W(T, r) = \det\left((I - M)^{(r)}\right),$$

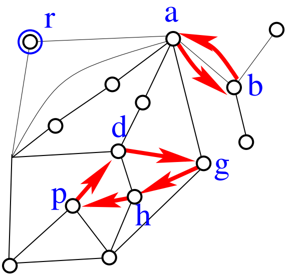
where $(I - M)^{(r)}$ is the matrix $(I - M)$ deprived of the line and column r.

Determinant expansion consequence

$$\det\left((I-M)^{(r)}\right) = \sum_{C\in\mathcal{C}} (-1)^{N(C)} \prod_{c \text{ cycles of } C} \prod_{\vec{e}\in c} M_{\vec{e}},$$

where the sum ranges over
$\mathcal{C}$ = set of collection of disjoint oriented cycles of length ≥ 1 avoiding r.

Define $\overleftarrow{M}_{x,y} := \rho_y M_{y,x}/\rho_x$, where ρ is the invariant measure associated to M.



$$\text{Weight} = (-1)^2 (M_{a,b}M_{b,a})(M_{d,g}M_{g,h}M_{h,p}M_{p,d})$$

Determinant expansion consequence

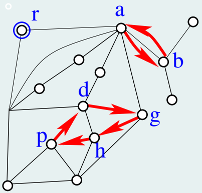
$$\det\left((I-M)^{(r)}\right)=\sum_{C\in\mathcal{C}}(-1)^{N(C)}\prod_{c \text{ cycles of } C}\prod_{\vec{e}\in c}M_{\vec{e}},$$

where the sum ranges over
$\mathcal{C}=$ set of collection of disjoint oriented cycles of length ≥ 1 avoiding r.

Define $\overleftarrow{M}_{x,y} := \rho_y M_{y,x}/\rho_x$, where ρ is the invariant measure associated to M.

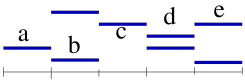
Claim

$$\det\left((I-M)^{(v)}\right)=\det\left((I-\overleftarrow{M})^{(v)}\right)$$



Heaps of pieces

Informally: some "elements" that are stacked.



General Heap

Trivial Heap

General heap: (left) Equivalence class of words describing the history of the stack $= baeddecb = baeddceb = ebaddbce = \ldots.$

Trivial heap: (right) All the pieces on the ground $ae = ea$

Formally: a set of letters $\mathcal{P}$ is given and a binary relation R:

– $x \not R y$ means that x commutes with y (that is $xy = yx$),

– xRy means that x does not commute with y.

Heap of dominos: $\mathcal{P} = \{a, b, c, d, e\}$ aRb, bRc, cRd, dRe.

Heaps: Equivalence classes of words

$w \sim w'$ if they are equal up to a finite number of allowed commutations of consecutive letters.

Heap of pieces

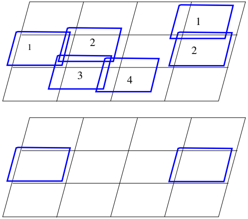


Figure: Heaps of squares. They do not commute if they share a side.

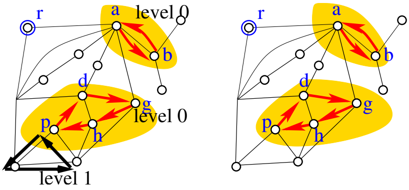


Figure: Heaps of cycles. They do not conmute if they share a vertex.

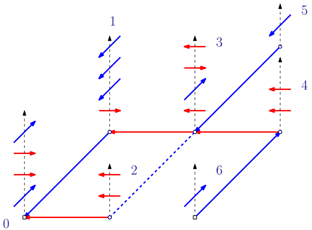


Figure: Heaps of outgoing edges. They do not commute if they start at the same point.

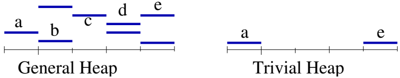


Figure: Heaps of dominoes. They do not commute if they share one extremity.

Heap of pieces

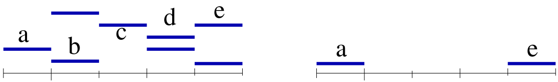
For a heap H

$$Weight(H) = \prod_{e \in H} w(e)$$

where $w : \mathcal{P} \to \mathbb{R}$ (or any formal commutative set)

Inversion lemma

$$\sum_{H \in \text{Heaps}} Weight(H) = \frac{1}{\sum_{H \in \text{TrivialHeaps}} (-1)^{|H|} Weight(H)}$$



General Heap Trivial Heap

Example : Weight x for each piece, $Weight(H) = x^{|H|}$,

$$\sum_{H \in \text{Heaps}} Weight(H) = \frac{1}{1 - 5x + 6x^2 - x^3}$$

In particular for the heaps of cycles with weights given by M one has that

$$\sum_{HC \in \text{Heaps of cycles avoiding } r} W(HC)$$

$$= \frac{1}{\sum_{HC \in \text{Trivial heaps of cycles avoiding } r} (-1)^{|HC|} W(HC)}$$

$$= \det\left((I - M)^{(r)}\right)^{-1}$$

Summary of determinant formulas

1. **Matrix tree theorem (MTT):**

$$\sum_{T \in \mathrm{ST}(G,r)} \prod_{\vec{e} \in T} M_{\vec{e}} = \det\left((I-M)^{(r)}\right),$$

2. **Cycles expansion:**

$$\det\left((I-M)^{(r)}\right) = \sum_{HC \in \text{Trivial heaps of cycles avoiding } r} (-1)^{|H|} \prod_{\vec{e} \in H} M_{\vec{e}},$$

3. **Claim:**

$$\det\left((I-M)^{(r)}\right) = \det\left((I-\overleftarrow{M})^{(r)}\right)$$

Consequences

1. **Heaps of cycles:**

$$\sum_{HC \in \text{Heaps of cycles avoiding } r} W(HC) \overset{\text{(Inv. Lem.)}}{=} \det\left((I-M)^{(r)}\right)^{-1}$$

2. **Markov chain tree theorem:** the invariant measure of M satisfies

$$\rho_v \overset{\text{(Alg)}}{=} \frac{\det((I-M)^{(v)})}{Z} \overset{\text{(MTT)}}{=} \frac{\sum_{T \in \mathrm{ST}(G,v)} \prod_{\vec{e} \in T} M_{\vec{e}}}{Z}$$

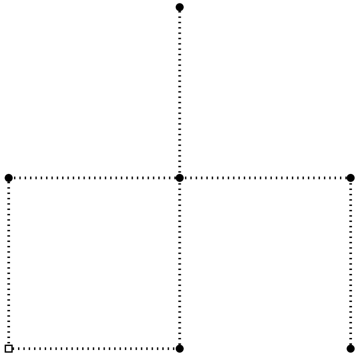
Important

$$Z \times \rho_v \times \left(\sum_{HC \in \text{Heaps of cycles avoinding } v} W(HC) \right) = 1$$

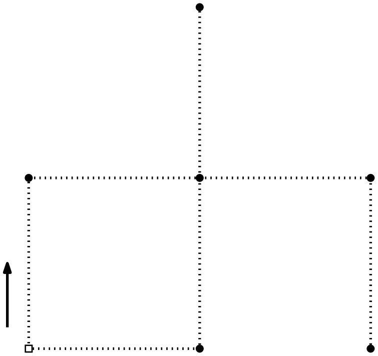
Consider a given graph G.
Algorithms to sample a UST:

1. Aldous-Broder algorithm.
2. Wilson algorithm.
3. Tutte polynomial + Matrix tree theorem.
4. ...

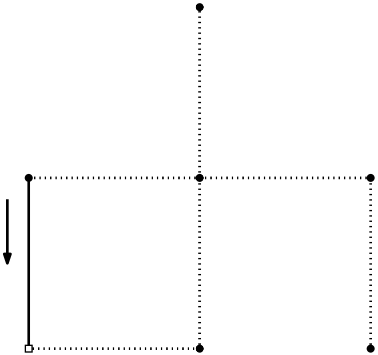
Consider an M-walk W in the invariant regime started at $r \in V$ up to the cover time.
Denote by **FirstEntrance**$(W) = (T, r)$, where r is the starting point of W and T is the spanning tree formed by the first edge used to visit each vertex.



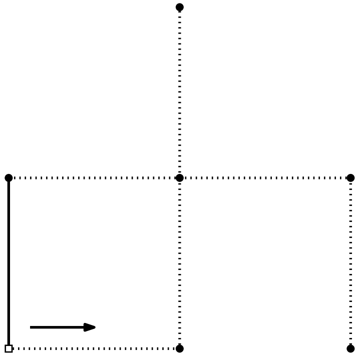
Consider an M-walk W in the invariant regime started at $r \in V$ up to the cover time.
Denote by **FirstEntrance**$(W) = (T, r)$, where r is the starting point of W and T is the spanning tree formed by the first edge used to visit each vertex.



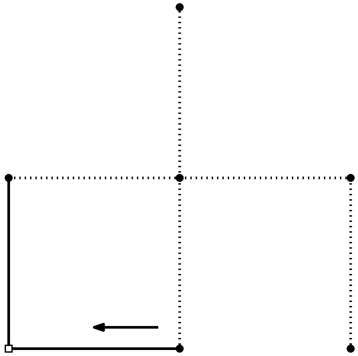
Consider an M-walk W in the invariant regime started at $r \in V$ up to the cover time.
Denote by **FirstEntrance**$(W) = (T, r)$, where r is the starting point of W and T is the spanning tree formed by the first edge used to visit each vertex.



Consider an M-walk W in the invariant regime started at $r \in V$ up to the cover time.
Denote by **FirstEntrance**$(W) = (T, r)$, where r is the starting point of W and T is the spanning tree formed by the first edge used to visit each vertex.

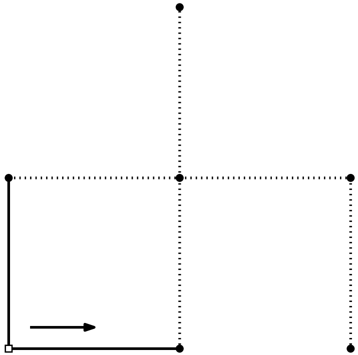


Consider an M-walk W in the invariant regime started at $r \in V$ up to the cover time.
Denote by **FirstEntrance**$(W) = (T, r)$, where r is the starting point of W and T is the spanning tree formed by the first edge used to visit each vertex.

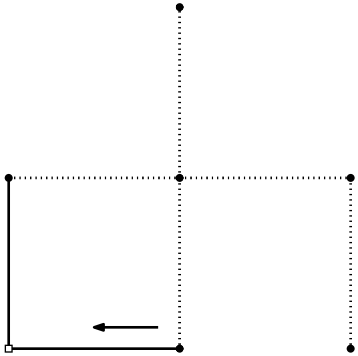


II. UST sampling: Aldous-Broder

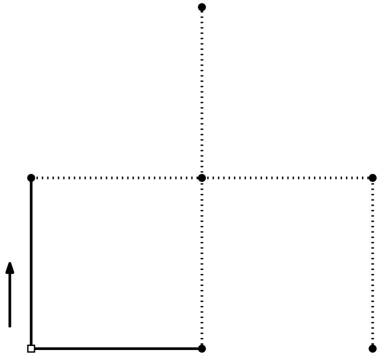
Consider an M-walk W in the invariant regime started at $r \in V$ up to the cover time.
Denote by **FirstEntrance**$(W) = (T, r)$, where r is the starting point of W and T is the spanning tree formed by the first edge used to visit each vertex.



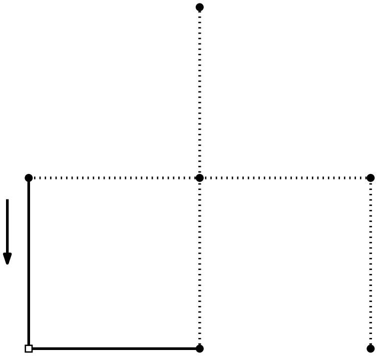
Consider an M-walk W in the invariant regime started at $r \in V$ up to the cover time.
Denote by **FirstEntrance**$(W) = (T, r)$, where r is the starting point of W and T is the spanning tree formed by the first edge used to visit each vertex.



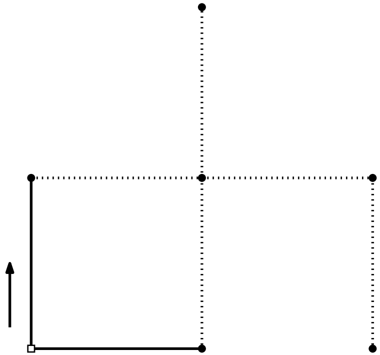
Consider an M-walk W in the invariant regime started at $r \in V$ up to the cover time.
Denote by **FirstEntrance**$(W) = (T, r)$, where r is the starting point of W and T is the spanning tree formed by the first edge used to visit each vertex.



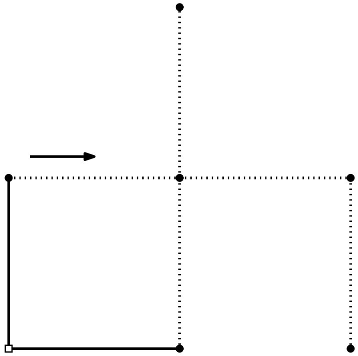
Consider an M-walk W in the invariant regime started at $r \in V$ up to the cover time.
Denote by **FirstEntrance**$(W) = (T, r)$, where r is the starting point of W and T is the spanning tree formed by the first edge used to visit each vertex.



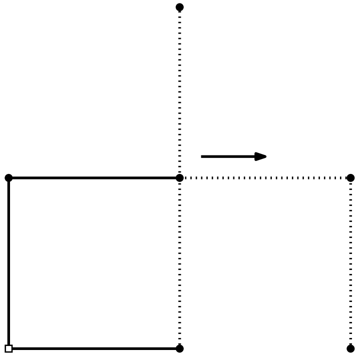
Consider an M-walk W in the invariant regime started at $r \in V$ up to the cover time.
Denote by **FirstEntrance**$(W) = (T, r)$, where r is the starting point of W and T is the spanning tree formed by the first edge used to visit each vertex.



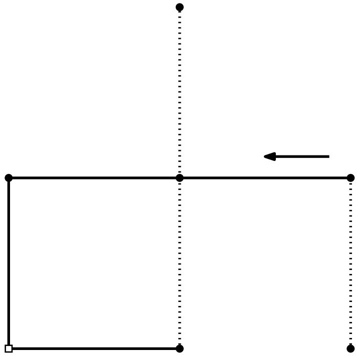
Consider an M-walk W in the invariant regime started at $r \in V$ up to the cover time.
Denote by **FirstEntrance**$(W) = (T, r)$, where r is the starting point of W and T is the spanning tree formed by the first edge used to visit each vertex.



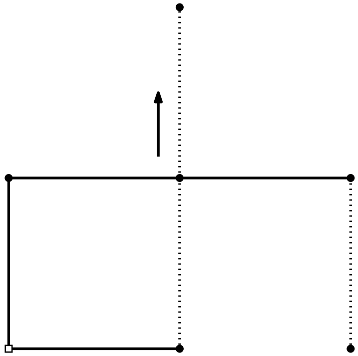
Consider an M-walk W in the invariant regime started at $r \in V$ up to the cover time.
Denote by **FirstEntrance**$(W) = (T, r)$, where r is the starting point of W and T is the spanning tree formed by the first edge used to visit each vertex.



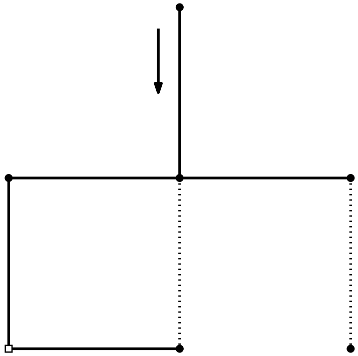
Consider an M-walk W in the invariant regime started at $r \in V$ up to the cover time.
Denote by **FirstEntrance**$(W) = (T, r)$, where r is the starting point of W and T is the spanning tree formed by the first edge used to visit each vertex.



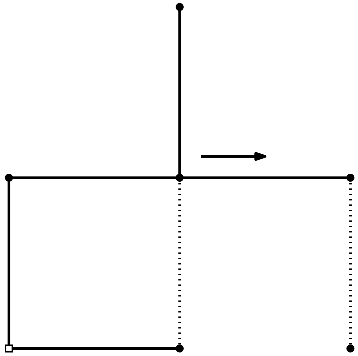
Consider an M-walk W in the invariant regime started at $r \in V$ up to the cover time.
Denote by **FirstEntrance**$(W) = (T, r)$, where r is the starting point of W and T is the spanning tree formed by the first edge used to visit each vertex.



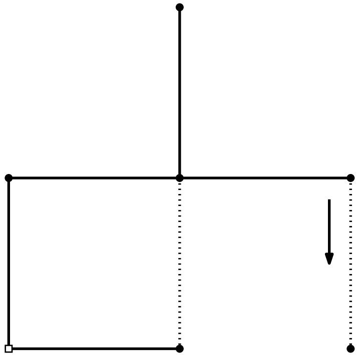
Consider an M-walk W in the invariant regime started at $r \in V$ up to the cover time.
Denote by **FirstEntrance**$(W) = (T, r)$, where r is the starting point of W and T is the spanning tree formed by the first edge used to visit each vertex.



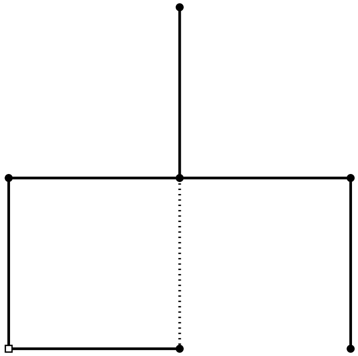
Consider an M-walk W in the invariant regime started at $r \in V$ up to the cover time.
Denote by **FirstEntrance**$(W) = (T, r)$, where r is the starting point of W and T is the spanning tree formed by the first edge used to visit each vertex.



Consider an M-walk W in the invariant regime started at $r \in V$ up to the cover time.
Denote by **FirstEntrance**$(W) = (T, r)$, where r is the starting point of W and T is the spanning tree formed by the first edge used to visit each vertex.



Consider an M-walk W in the invariant regime started at $r \in V$ up to the cover time.
Denote by **FirstEntrance**$(W) = (T, r)$, where r is the starting point of W and T is the spanning tree formed by the first edge used to visit each vertex.



Reversible: $\rho_u M_{u,v} = \rho_v M_{v,u}$

Theorem (Aldous-Broder ('89))

*For M positive and **reversible** Markov kernel with invariant distribution ρ. For any $T \in ST(G)$ one has*

$$\mathbb{P}\left(\textbf{\textit{FirstEntrance}}(W) = (T, r)\right) = \frac{\prod_{\vec{e} \in T} M_{\vec{e}}}{\sum_{w \in V} \det(I - M^{(w)})},$$

Reversible: $\rho_u M_{u,v} = \rho_v M_{v,u}$

Theorem (Aldous-Broder ('89))

*For M positive and **reversible** Markov kernel with invariant distribution ρ. For any $T \in ST(G)$ one has*

$$\mathbb{P}\left(\boldsymbol{FirstEntrance}(W) = (T, r)\right) = \frac{\prod_{\vec{e} \in T} M_{\vec{e}}}{\sum_{w \in V} \det(I - M^{(w)})},$$

Reversible Kernel: $\overleftarrow{M}_{x,y} = \rho_y / \rho_x M_{y,x}$

Reversible: $\rho_u M_{u,v} = \rho_v M_{v,u}$

Theorem (Aldous-Broder ('89))

*For M positive and **reversible** Markov kernel with invariant distribution ρ. For any $T \in ST(G)$ one has*

$$\mathbb{P}\left(\textbf{\textit{FirstEntrance}}(W) = (T, r)\right) = \frac{\prod_{\vec{e} \in T} M_{\vec{e}}}{\sum_{w \in V} \det(I - M^{(w)})},$$

Reversible Kernel: $\overleftarrow{M}_{x,y} = \rho_y / \rho_x M_{y,x}$

Theorem (F.- Marckert ('21+))

For M positive with invariant distribution ρ. For any $T \in ST(G)$ one has

$$\mathbb{P}\left(\textbf{\textit{FirstEntrance}}(W) = (T, r)\right) = \frac{\prod_{\vec{e} \in T} \overleftarrow{M}_{\vec{e}}}{\sum_{w \in V} \det(I - \overleftarrow{M}^{(w)})},$$

Reversible: $\rho_u M_{u,v} = \rho_v M_{v,u}$

Theorem (Aldous-Broder ('89))

*For M positive and **reversible** Markov kernel with invariant distribution ρ. For any $T \in ST(G)$ one has*

$$\mathbb{P}\left(\textbf{\textit{FirstEntrance}}(W) = (T, r)\right) = \frac{\prod_{\vec{e} \in T} M_{\vec{e}}}{\sum_{w \in V} \det(I - M^{(w)})},$$

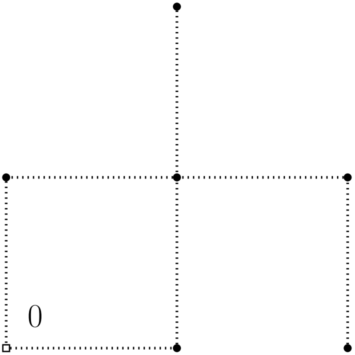
Reversible Kernel: $\overleftarrow{M}_{x,y} = \rho_y / \rho_x M_{y,x}$

Theorem (Hu-Lyons-Tang (20), F.- Marckert ('21+))

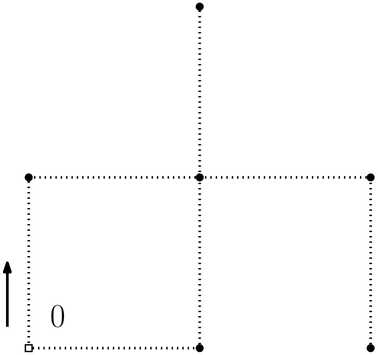
For M positive with invariant distribution ρ. For any $T \in ST(G)$ one has

$$\mathbb{P}\left(\textbf{\textit{FirstEntrance}}(W) = (T, r)\right) = \frac{\prod_{\vec{e} \in T} \overleftarrow{M}_{\vec{e}}}{\sum_{w \in V} \det(I - \overleftarrow{M}^{(w)})},$$

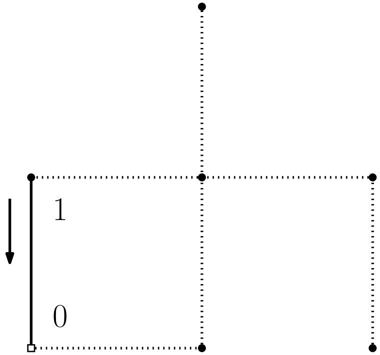
Labelled extension



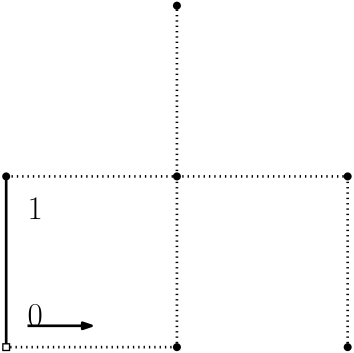
Labelled extension



Labelled extension

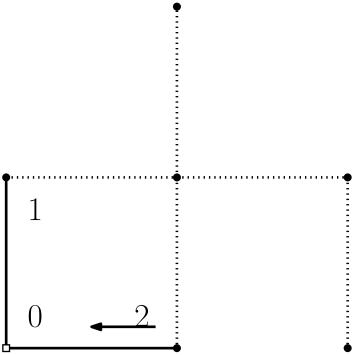


Labelled extension



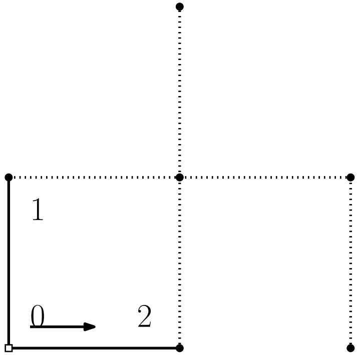
III. New proof

Labelled extension

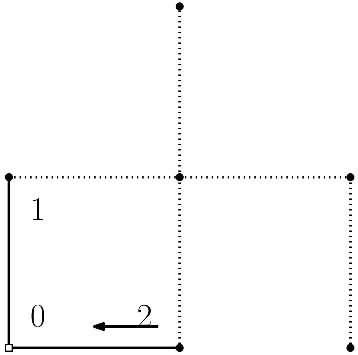


III. New proof

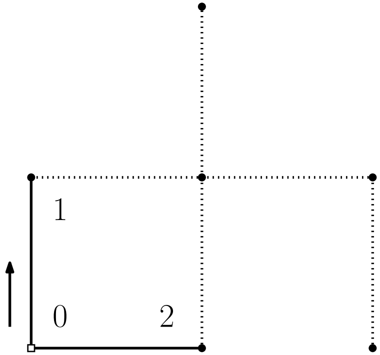
Labelled extension



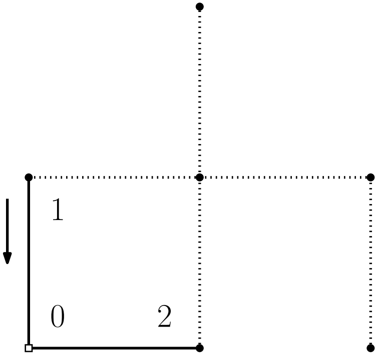
Labelled extension



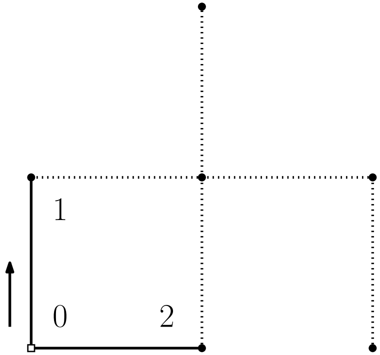
Labelled extension



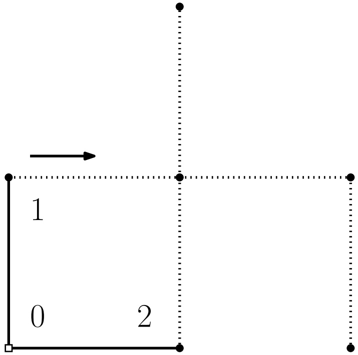
Labelled extension



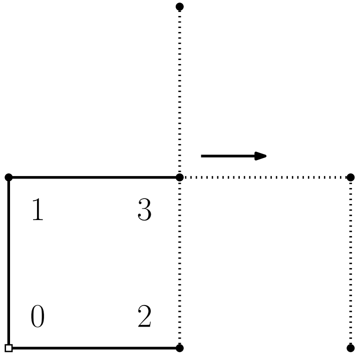
Labelled extension



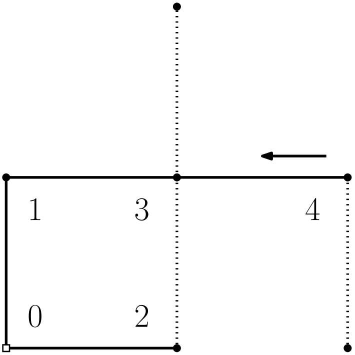
Labelled extension



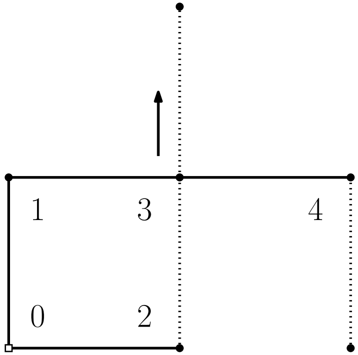
Labelled extension



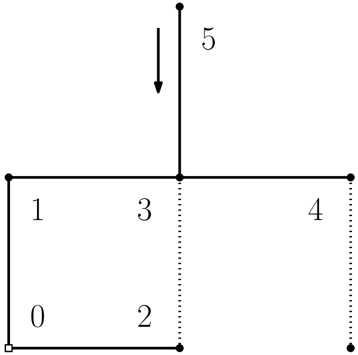
Labelled extension



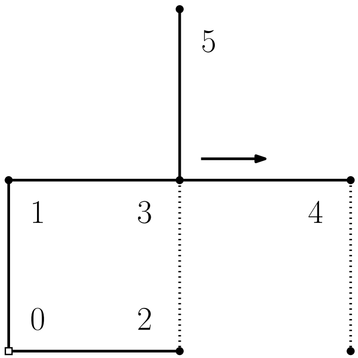
Labelled extension



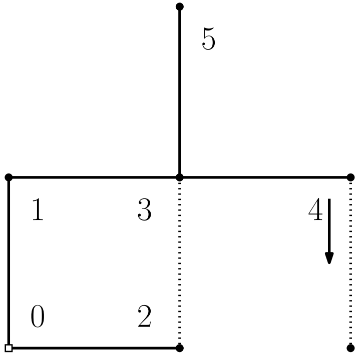
Labelled extension



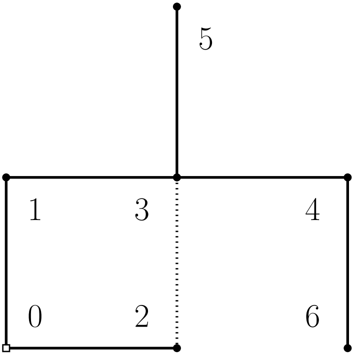
Labelled extension



Labelled extension



Labelled extension



Denote by **Pionner**$(W) = ($**FirstEntrance**$(W), L)$ where L is the labeling.
$H_D(a,b) =$ probability starting from a that a walk following M escapes D at b.
$\overleftarrow{H}_D(a,b) =$ probability starting from a that a walk following $\overleftarrow{M}$ escapes D at b.

$$\begin{aligned}
&\mathbb{P}(\mathbf{Pionner}(W) = ((t,r),\ell)) \\
&= \mathbb{1}_{\ell_{\mathbf{0}}=r}\rho_{\ell_{\mathbf{0}}} \prod_{i=0}^{n-2} \left[H_{\{\ell_{\le i}\}}(\ell_i, a(\ell_{i+1})) M_{a(\ell_{i+\mathbf{1}}),\ell_{i+\mathbf{1}}} \right] \\
&= \left(\mathbb{1}_{\ell_{\mathbf{0}}=r}\rho_{\ell_{n-\mathbf{1}}} \prod_{i=0}^{n-2} \left[\overleftarrow{H}_{\{\ell_{\le i}\}}(a(\ell_{i+1}), \ell_i) \right] Z \right) \frac{\prod_{\vec{e}\in t} \overleftarrow{M}_{\vec{e}}}{Z}
\end{aligned}$$

Can we prove using combinatorics that

$$\sum_{\ell} \mathbb{1}_{\ell_0=r} Z \rho_{\ell_{n-1}} \prod_{i=0}^{n-2} \left[\overleftarrow{H}_{\{\ell_{\leq i}\}}(a(\ell_{i+1}), \ell_i) \right] = 1?$$

(the sum ranges over all decreasing labellings of the tree)

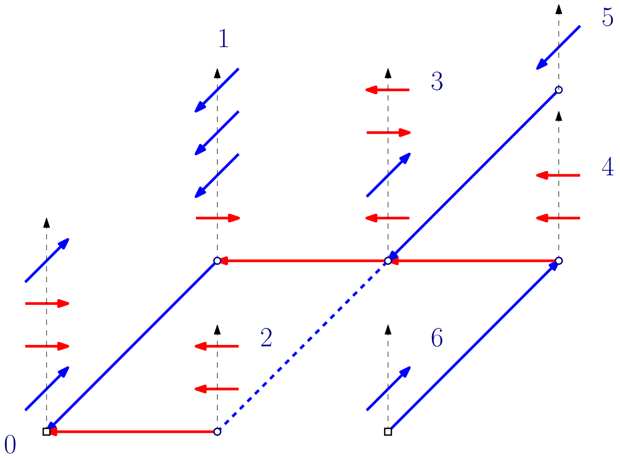


Figure: Path seen backward as a heap of outgoing edges

III. New proof

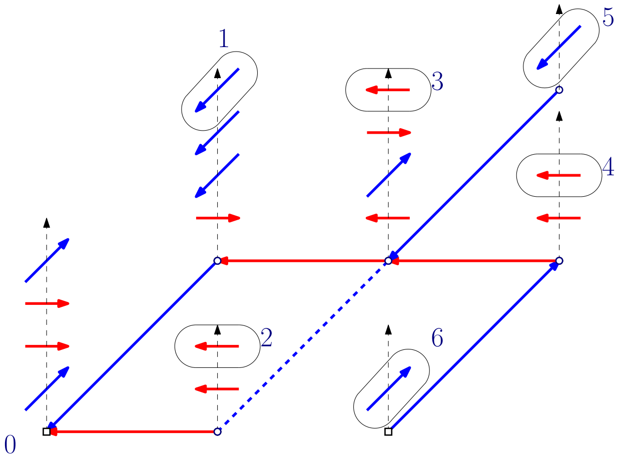


Figure: The tree edges are always on top of the piles.

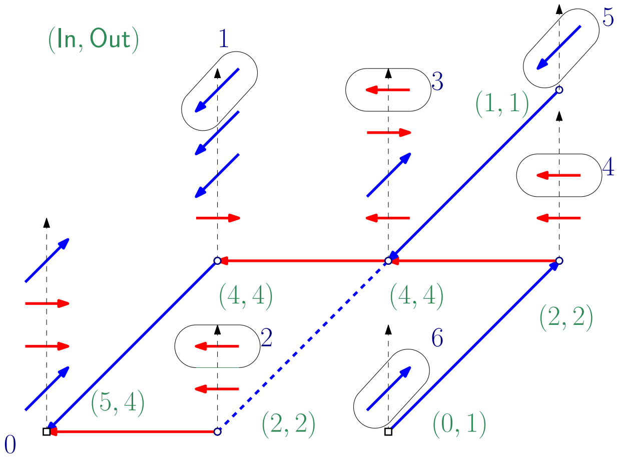


Figure: Count the incoming and outgoing edges

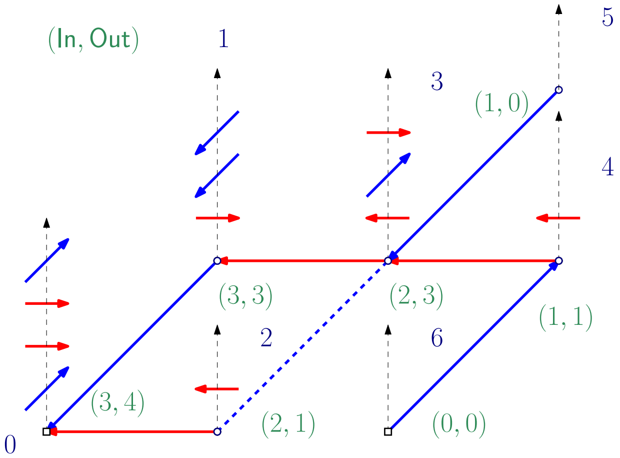


Figure: Pop-out the tree edges to construct H^{-t} (update (In,Out))

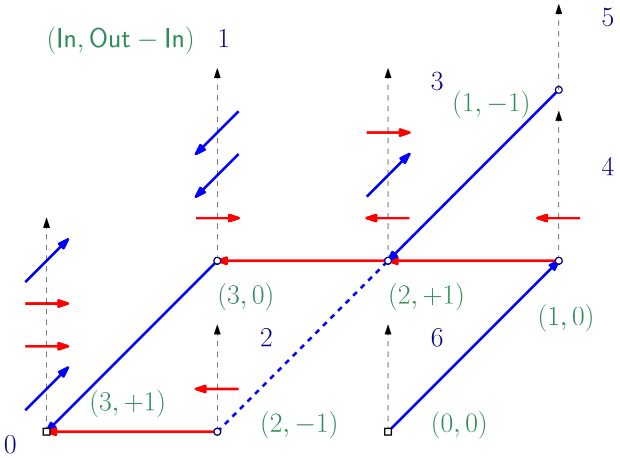


Figure: Convenient to keep an eye on (In,Out-In)

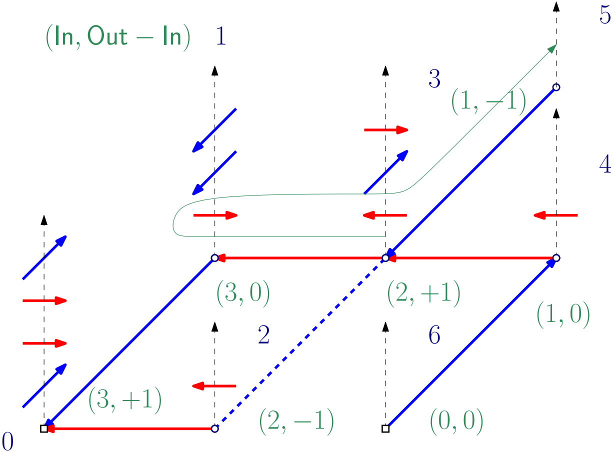


Figure: Play golf!

III. New proof

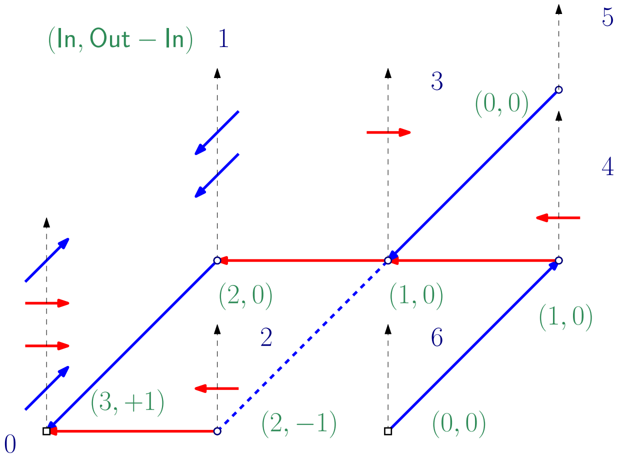


Figure: Supress the path and update (In,Out-In)

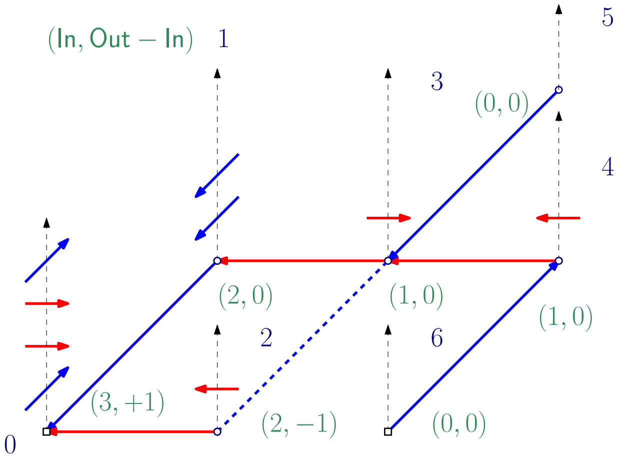


Figure: Let the pieces fall

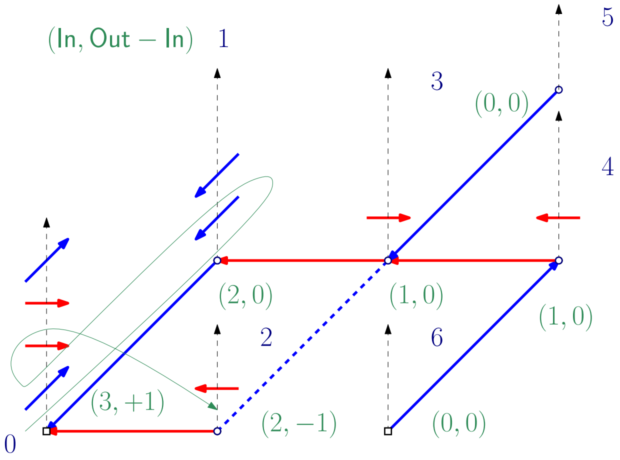


Figure: Continue playing golf with next emitting vertex.

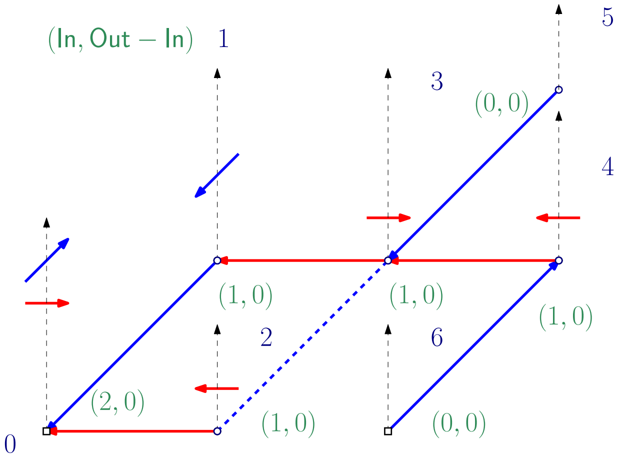


Figure: Supress the path and update (In,Out-In)

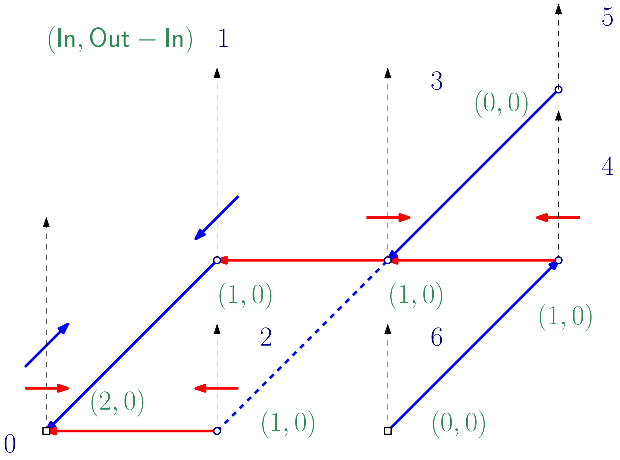


Figure: Let the pieces fall

III. New proof

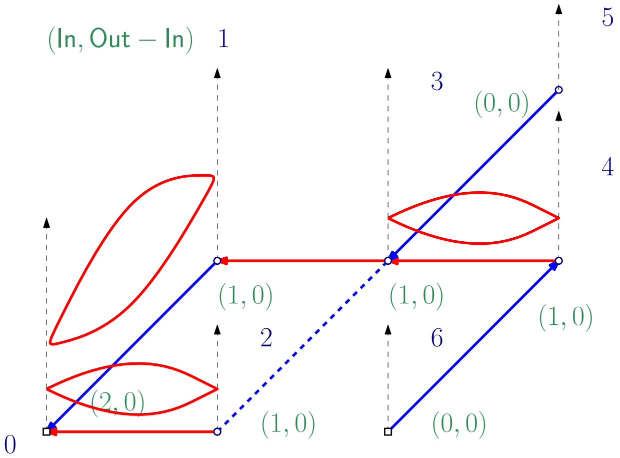


Figure: heap of cycles

III. New proof

The heap of outgoing edges H^{-t} is a heap only on $V \setminus \ell_{n-1}$ and $H^{-t} = \text{Golf} \times HC$. Recall we fix a treatment order to fix an ordering of the starting points in the Golf game.

$$\sum_{\ell} \mathbb{1}_{\ell_0 = r} Z \rho_{\ell_{n-1}} \times \prod_{i=0}^{n-2} \left[\overleftarrow{H}_{\ell_{\leq i}}(a(\ell_{i+1}), \ell_i) \right]$$

III. New proof

The heap of outgoing edges H^{-t} is a heap only on $V \setminus \ell_{n-1}$ and $H^{-t} = \text{Golf} \times HC$. Recall we fix a treatment order to fix an ordering of the starting points in the Golf game.

$$\sum_{\ell} \mathbb{1}_{\ell_0 = r} Z\rho_{\ell_{n-1}} \times \prod_{i=0}^{n-2} \left[\overleftarrow{H}_{\ell_{\leq i}}(a(\ell_{i+1}), \ell_i) \right]$$

$$= \sum_{H^{-t} \text{ valid}} Z\rho_{\ell_{n-1}} \times W(H^{-t})$$

The heap of outgoing edges H^{-t} is a heap only on $V \setminus \ell_{n-1}$ and $H^{-t} = \text{Golf} \times HC$. Recall we fix a treatment order to fix an ordering of the starting points in the Golf game.

$$\sum_{\ell} \mathbb{1}_{\ell_0 = r} Z\rho_{\ell_{n-1}} \times \prod_{i=0}^{n-2} \left[\overleftarrow{H}_{\ell_{\leq i}}(a(\ell_{i+1}), \ell_i) \right]$$

$$= \sum_{H^{-t} \text{ valid}} Z\rho_{\ell_{n-1}} \times W(H^{-t})$$

$$= \sum_{(Golf, HC) \text{ valid}} Z\rho_{\ell_{n-1}} \times W(Golf) \times W(HC)$$

III. New proof

The heap of outgoing edges H^{-t} is a heap only on $V \setminus \ell_{n-1}$ and $H^{-t} = \text{Golf} \times HC$. Recall we fix a treatment order to fix an ordering of the starting points in the Golf game.

$$
\begin{aligned}
&\sum_{\ell} \mathbb{1}_{\ell_{\mathbf{0}}=r} Z\rho_{\ell_{n-\mathbf{1}}} \times \prod_{i=0}^{n-2} \left[\overleftarrow{H}_{\ell_{\leq i}}(a(\ell_{i+1}), \ell_i) \right] \\
&= \sum_{H^{-t} \text{ valid}} Z\rho_{\ell_{n-\mathbf{1}}} \times W(H^{-t}) \\
&= \sum_{(Golf, HC) \text{ valid}} Z\rho_{\ell_{n-\mathbf{1}}} \times W(Golf) \times W(HC) \\
&= \underbrace{\sum_{Golf \text{ valid}} W(Golf)}_{=1} \times \underbrace{Z\rho_{\ell_{n-\mathbf{1}}} \times \left(\sum_{\substack{HC \in \textbf{ heaps of cycles} \\ \text{avoiding } \ell_{n-\mathbf{1}}}} W(HC) \right)}_{\overset{\textbf{Important}}{=} 1}
\end{aligned}
$$

The first by a probabilistic algorithm.

Motivation: Odded Schramm question

Conformally invariant scaling limits: an overview and a collection of problems

Oded Schramm

2.5. Lattice trees. We now present an example of a discrete model where we suspect that perhaps conformal invariance might hold. However, we do not presently have a candidate for the scaling limit.

Fix $n \in \mathbb{N}_+$, and consider the collection of all trees contained in the grid G that contain the origin and have n vertices. Select a tree T from this measure, uniformly at random.

Problem 2.8. What is the growth rate of the expected diameter of such a tree? If we rescale the tree so that the expected (or median) diameter is 1, is there a limit for the law of the tree as $n \to \infty$? What are its geometric and topological properties? Can the limit be determined?

It would be good to be able to produce some pictures. However, we presently do not know how to sample from this measure.

Problem 2.9. Produce an efficient algorithm which samples lattice trees approximately uniformly, or prove that such an algorithm does not exist.

Figure: Schramm ICM 2006.

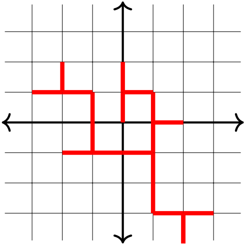


Figure: Subtree of size 20 containing the origin on $\mathbb{Z}^2$.

Corollary (F.-Marckert ('21+))

If W is a SRW stopped when $m < |V|$ vertices has been discovered, then the tree ***FirstEntrance***(W) *is not uniform in the set of subtrees of G of size m.*

THANKS!