

Pavages par dominos dans le plan (4)

Petite école de combinatoire

Philippe Duchon

LaBRI

20 juin 2022

Ce qu'on a vu...

Pour une région R , simplement connexe :

- On a une relation d'**ordre partiel** naturelle sur les pavages.

Ce qu'on a vu...

Pour une région R , simplement connexe :

- On a une relation d'**ordre partiel** naturelle sur les pavages.
- Il y a un unique pavage minimal, et un unique pavage maximal.

Ce qu'on a vu...

Pour une région R , simplement connexe :

- On a une relation d'**ordre partiel** naturelle sur les pavages.
- Il y a un unique pavage minimal, et un unique pavage maximal.
- La relation de couverture est décrite par les flips (repavage d'un carré ou d'un hexagone)

Ce qu'on a vu...

Pour une région R , simplement connexe :

- On a une relation d'**ordre partiel** naturelle sur les pavages.
- Il y a un unique pavage minimal, et un unique pavage maximal.
- La relation de couverture est décrite par les flips (repavage d'un carré ou d'un hexagone)
- Plus précisément, cet ordre partiel décrit un **treillis distributif**

Ce qu'on a vu...

Pour une région R , simplement connexe :

- On a une relation d'**ordre partiel** naturelle sur les pavages.
- Il y a un unique pavage minimal, et un unique pavage maximal.
- La relation de couverture est décrite par les flips (repavage d'un carré ou d'un hexagone)
- Plus précisément, cet ordre partiel décrit un **treillis distributif**

C'est une situation qui se prête particulièrement bien à la génération aléatoire uniforme par l'algorithme de *couplage depuis le passé* ("coupling from the past" ; CFTP ; Propp-Wilson '96).

Quelques notations

- R : la région (simplement connexe)
- $H(R)$: l'ensemble de toutes les fonctions de hauteur sur R
- f_{\min}, f_{\max} : les fonctions de hauteur minimale et maximale de R
- $m(R)$: le nombre de fonctions de hauteur (ou de pavages) de R
- $a(R)$: le nombre de sommets intérieurs à la région (c'est très approximativement une aire)

Flips et génération quasi uniforme

Connexité par les flips : à partir du moment où on sait qu'on peut passer de n'importe quel pavage à n'importe quel autre en répétant des flips, on a une méthode dont on est assuré qu'elle permet d'obtenir des objets presque uniformes :

- Partir d'un pavage $M_0 = m_0$ quelconque.
- On définit une suite de pavages (M_n) : partant de M_n ,
 - on choisit une position aléatoire pour un flip, et un "sens" (haut/bas)
 - si un flip dans ce sens est possible sur M_n , le résultat est M_{n+1} ; sinon $M_{n+1} = M_n$

Pour peu que la probabilité de flip "vers le haut" soit toujours $1/2$ (et indépendante de la position choisie), M_n est asymptotiquement uniforme : pour tout pavage m ,

$$\lim_n \mathbb{P}(M_n = m) = \frac{1}{m(R)},$$

où $m(R)$ désigne le nombre de pavages de R .

Chaînes de Markov

On est en fait en train de définir une **chaîne de Markov** sur l'ensemble M des pavages : on définit (implicitement) des probabilités de transitions $p_{m,m'}$ (probabilité que le prochain pavage soit m' , si le pavage courant est m).

Chaînes de Markov

On est en fait en train de définir une **chaîne de Markov** sur l'ensemble M des pavages : on définit (implicitement) des probabilités de transitions $p_{m,m'}$ (probabilité que le prochain pavage soit m' , si le pavage courant est m).

La probabilité que les $n + 1$ premiers pavages soient m_0, \dots, m_n est alors donnée par $p_{m_0,m_1} p_{m_1,m_2} \cdots p_{m_{n-1},m_n}$.

Chaînes de Markov

On est en fait en train de définir une **chaîne de Markov** sur l'ensemble M des pavages : on définit (implicitement) des probabilités de transitions $p_{m,m'}$ (probabilité que le prochain pavage soit m' , si le pavage courant est m).

La probabilité que les $n + 1$ premiers pavages soient m_0, \dots, m_n est alors donnée par $p_{m_0,m_1} p_{m_1,m_2} \cdots p_{m_{n-1},m_n}$.

Résultat classique sur les chaînes de Markov à espace d'états fini : si on a à la fois

- pour tous objets m et m' , il existe une séquence (m_i) avec $m_0 = m$, $m_n = m'$, et $p_{m_i,m_{i+1}} > 0$ (connexité ; “irréductible”)
- $p_{m,m} > 0$ pour au moins un m (“apériodique”)

alors il y a convergence vers une loi de probabilité $(\pi(m))_{m \in M}$; de plus, si on identifie une loi telle que $\pi(m)p_{m,m'} = \pi(m')p_{m',m}$ pour tous m et m' , alors c'est elle la loi limite.

(Dans notre cas on veut la loi uniforme, et pour cela on a juste besoin de $p_{m,m'} = p_{m',m}$)

Chaînes de Markov et pavages aléatoires

- Ce que ça nous dit, c'est que pour une région quelconque, il suffit de prendre n "assez grand" et M_n sera "quasiment un pavage uniforme".

Chaînes de Markov et pavages aléatoires

- Ce que ça nous dit, c'est que pour une région quelconque, il suffit de prendre n "assez grand" et M_n sera "quasiment un pavage uniforme".
- **Ce que ça ne dit pas**, c'est comment trouver un n "assez grand" (trouver une estimation de la "vitesse de convergence" est souvent beaucoup plus difficile que de définir une chaîne de Markov facile à simuler)

Chaînes de Markov et pavages aléatoires

- Ce que ça nous dit, c'est que pour une région quelconque, il suffit de prendre n "assez grand" et M_n sera "quasiment un pavage uniforme".
- **Ce que ça ne dit pas**, c'est comment trouver un n "assez grand" (trouver une estimation de la "vitesse de convergence" est souvent beaucoup plus difficile que de définir une chaîne de Markov facile à simuler)
- L'algorithme de CFTP permet de s'affranchir de cette difficulté : au prix d'une légère complexification, on obtient un algorithme qui s'arrête quand il s'arrête, mais dont on est assuré que le résultat est **exactement uniforme**.

Mise à jour d'une fonction de hauteur

Un “pas” de la chaîne de Markov se fait selon l'algorithme suivant : on part d'une fonction de hauteur f

- Choisir un sommet s (intérieur) aléatoire, et un signe aléatoire uniforme $\epsilon = \pm 1$.
- **Si** on obtient une fonction de hauteur f' en posant $f'(s) = f(s) + 4\epsilon$ et $f'(t) = f(t)$ pour tout autre sommet t , c'est la nouvelle fonction de hauteur ; sinon, la fonction de hauteur ne change pas.

Uniformité : si le sommet s est choisi uniformément parmi les $a(R)$ sommets intérieurs, la probabilité de transition $p_{f,f'}$ est définie par $p_{f,f'} = 1/(2a(R))$ si f et f' diffèrent en un unique sommet, 0 sinon (et $p_{f,f}$ “ce qu'il faut” pour que $\sum_{f'} p_{f,f'} = 1$). C'est symétrique en f et f' , donc la limite est uniforme.

Définition d'un couplage

En fait, on peut juste définir $2a(R)$ “fonctions de mise à jour” (de fonctions de hauteur) comme des fonctions $F_{s,\epsilon} : H(R) \rightarrow H(R)$, par : pour toute $f \in H(R)$, si la fonction f' définie par

- $f'(s) = f(s) + 4\epsilon$
- pour tout autre sommet t , $f'(t) = f(t)$

est une fonction de hauteur, alors on pose $F_{s,\epsilon}(f) = f'$; sinon, on pose $F_{s,\epsilon} = f$.

Définition d'un couplage

En fait, on peut juste définir $2a(R)$ “fonctions de mise à jour” (de fonctions de hauteur) comme des fonctions $F_{s,\epsilon} : H(R) \rightarrow H(R)$, par : pour toute $f \in H(R)$, si la fonction f' définie par

- $f'(s) = f(s) + 4\epsilon$
- pour tout autre sommet t , $f'(t) = f(t)$

est une fonction de hauteur, alors on pose $F_{s,\epsilon}(f) = f'$; sinon, on pose $F_{s,\epsilon} = f$.

On peut alors définir un pas de la chaîne de Markov comme le choix d'une fonction de mise à jour aléatoire (qu'on applique à la fonction de hauteur courante).

Définition d'un couplage

En fait, on peut juste définir $2a(R)$ “fonctions de mise à jour” (de fonctions de hauteur) comme des fonctions $F_{s,\epsilon} : H(R) \rightarrow H(R)$, par : pour toute $f \in H(R)$, si la fonction f' définie par

- $f'(s) = f(s) + 4\epsilon$
- pour tout autre sommet t , $f'(t) = f(t)$

est une fonction de hauteur, alors on pose $F_{s,\epsilon}(f) = f'$; sinon, on pose $F_{s,\epsilon} = f$.

On peut alors définir un pas de la chaîne de Markov comme le choix d'une fonction de mise à jour aléatoire (qu'on applique à la fonction de hauteur courante).

Ce faisant, on définit en fait un **couplage** d'autant de copies qu'on veut de la chaîne : si on a k fonctions de hauteurs “courantes” f_1, \dots, f_k , on peut toutes les mettre à jour avec la même fonction F : pour tout j , $f'_j = F(f_j)$.

Couplage avant, couplage arrière

Pour n'importe quel k , on peut définir une suite de k -uplets aléatoires de fonctions de hauteurs de deux manières différentes : on tire une séquence de couples $(s_n, \epsilon_n)_{n \geq 1}$, puis, pour un k -uplet $(f_0^1, f_0^2, \dots, f_0^k)$ de fonctions de départ,

- **couplage avant** : on peut poser

$$f_n^i = F_{s_n, \epsilon_n} \circ F_{s_{n-1}, \epsilon_{n-1}} \circ \dots \circ F_{s_1, \epsilon_1}(f_0^i), \text{ soit } f_n^i = F_{s_n, \epsilon_n}(f_{n-1}^i).$$

Couplage avant, couplage arrière

Pour n'importe quel k , on peut définir une suite de k -uplets aléatoires de fonctions de hauteurs de deux manières différentes : on tire une séquence de couples $(s_n, \epsilon_n)_{n \geq 1}$, puis, pour un k -uplet $(f_0^1, f_0^2, \dots, f_0^k)$ de fonctions de départ,

- **couplage avant** : on peut poser

$$f_n^i = F_{s_n, \epsilon_n} \circ F_{s_{n-1}, \epsilon_{n-1}} \circ \dots \circ F_{s_1, \epsilon_1}(f_0^i), \text{ soit } f_n^i = F_{s_n, \epsilon_n}(f_{n-1}^i).$$

- **couplage arrière** : on pose $g_n^i = F_{s_1, \epsilon_1} \circ F_{s_2, \epsilon_2} \circ \dots \circ F_{s_n, \epsilon_n}(f_0^i)$ (ordre de composition non naturel)

Couplage avant, couplage arrière

Pour n'importe quel k , on peut définir une suite de k -uplets aléatoires de fonctions de hauteurs de deux manières différentes : on tire une séquence de couples $(s_n, \epsilon_n)_{n \geq 1}$, puis, pour un k -uplet $(f_0^1, f_0^2, \dots, f_0^k)$ de fonctions de départ,

- **couplage avant** : on peut poser

$$f_n^i = F_{s_n, \epsilon_n} \circ F_{s_{n-1}, \epsilon_{n-1}} \circ \dots \circ F_{s_1, \epsilon_1}(f_0^i), \text{ soit } f_n^i = F_{s_n, \epsilon_n}(f_{n-1}^i).$$

- **couplage arrière** : on pose $g_n^i = F_{s_1, \epsilon_1} \circ F_{s_2, \epsilon_2} \circ \dots \circ F_{s_n, \epsilon_n}(f_0^i)$ (ordre de composition non naturel)

Le couplage avant décrit une chaîne de Markov, pas le couplage arrière ; **mais** si on s'intéresse seulement au "temps n ", les k -uplets (f_n^1, \dots, f_n^k) et (g_n^1, \dots, g_n^k) ont la même loi.

Coalescence et arrêt

- Dans le couplage avant et arrière, on peut (théoriquement du moins) prendre k égal au nombre de couplages, et prendre tous les couplages existants comme points de départ.

Coalescence et arrêt

- Dans le couplage avant et arrière, on peut (théoriquement du moins) prendre k égal au nombre de couplages, et prendre tous les couplages existants comme points de départ.
- On peut alors définir le “temps” aléatoire N (N') comme le premier n tel que tous les f_n^i (tous les g_n^i) sont égaux (“temps de coalescence”).

Coalescence et arrêt

- Dans le couplage avant et arrière, on peut (théoriquement du moins) prendre k égal au nombre de couplages, et prendre tous les couplages existants comme points de départ.
- On peut alors définir le “temps” aléatoire N (N') comme le premier n tel que tous les f_n^i (tous les g_n^i) sont égaux (“temps de coalescence”).
- On peut alors se demander si la fonction de hauteur $f_N^?$ (ou $g_{N'}^?$) ne serait pas uniforme. . .

Coalescence et arrêt

- Dans le couplage avant et arrière, on peut (théoriquement du moins) prendre k égal au nombre de couplages, et prendre tous les couplages existants comme points de départ.
- On peut alors définir le “temps” aléatoire N (N') comme le premier n tel que tous les f_n^i (tous les g_n^i) sont égaux (“temps de coalescence”).
- On peut alors se demander si la fonction de hauteur $f_N^?$ (ou $g_{N'}^?$) ne serait pas uniforme. . .
- Double résultat un peu surprenant : $f_N^?$ n'est en général pas uniforme ; en revanche, $g_{N'}^?$ l'est.

Coalescence et arrêt

- Dans le couplage avant et arrière, on peut (théoriquement du moins) prendre k égal au nombre de couplages, et prendre tous les couplages existants comme points de départ.
- On peut alors définir le “temps” aléatoire N (N') comme le premier n tel que tous les f_n^i (tous les g_n^i) sont égaux (“temps de coalescence”).
- On peut alors se demander si la fonction de hauteur $f_N^?$ (ou $g_{N'}^?$) ne serait pas uniforme. . .
- Double résultat un peu surprenant : $f_N^?$ n'est en général pas uniforme ; en revanche, $g_{N'}^?$ l'est.
- **Note** : Dans les deux cas on compose des fonctions (aléatoires) de l'espace d'états dans lui-même ; le temps de coalescence, c'est le premier temps où la fonction composée est une fonction constante.

Non uniformité du couplage avant

On prend un ensemble d'états $S = \{1, 2, \dots, k\}$ ($k > 2$) et deux fonctions de mise à jour, $F^+(i) = \min(k, i + 1)$ et $F^-(i) = \max(1, i - 1)$.

- La chaîne de Markov sous-jacente, c'est la marche aléatoire sur S (on monte ou on descend avec probabilité $1/2$, avec interdiction de sortir de l'intervalle S).

Non uniformité du couplage avant

On prend un ensemble d'états $S = \{1, 2, \dots, k\}$ ($k > 2$) et deux fonctions de mise à jour, $F^+(i) = \min(k, i + 1)$ et $F^-(i) = \max(1, i - 1)$.

- La chaîne de Markov sous-jacente, c'est la marche aléatoire sur S (on monte ou on descend avec probabilité $1/2$, avec interdiction de sortir de l'intervalle S).
- La loi limite est bien uniforme sur S (condition d'équilibre satisfaite).

Non uniformité du couplage avant

On prend un ensemble d'états $S = \{1, 2, \dots, k\}$ ($k > 2$) et deux fonctions de mise à jour, $F^+(i) = \min(k, i + 1)$ et $F^-(i) = \max(1, i - 1)$.

- La chaîne de Markov sous-jacente, c'est la marche aléatoire sur S (on monte ou on descend avec probabilité $1/2$, avec interdiction de sortir de l'intervalle S).
- La loi limite est bien uniforme sur S (condition d'équilibre satisfaite).
- **Mais** les seules valeurs possibles au temps de coalescence sont 1 et k , donc la valeur obtenue au temps de coalescence par le couplage avant n'est pas uniforme sur S .

Uniformité du couplage arrière

- Observation critique : si $F_1 \circ F_2 \cdots \circ F_n$ est une fonction constante, alors $F_1 \circ F_2 \cdots \circ F_{n+1}$ est aussi une fonction constante, **avec la même valeur**.

Uniformité du couplage arrière

- Observation critique : si $F_1 \circ F_2 \cdots \circ F_n$ est une fonction constante, alors $F_1 \circ F_2 \cdots \circ F_{n+1}$ est aussi une fonction constante, **avec la même valeur**.
- Donc (informellement), qu'on regarde la valeur du couplage arrière au temps de coalescence ou plus tard, le résultat est le même ; mais en faisant tendre le "plus tard" vers l'infini, on obtient un résultat uniforme ; donc le résultat du couplage arrière au temps de coalescence est uniforme.

Uniformité du couplage arrière

- Observation critique : si $F_1 \circ F_2 \cdots \circ F_n$ est une fonction constante, alors $F_1 \circ F_2 \cdots \circ F_{n+1}$ est aussi une fonction constante, **avec la même valeur**.
- Donc (informellement), qu'on regarde la valeur du couplage arrière au temps de coalescence ou plus tard, le résultat est le même ; mais en faisant tendre le "plus tard" vers l'infini, on obtient un résultat uniforme ; donc le résultat du couplage arrière au temps de coalescence est uniforme.
- Pour une preuve plus propre, il faut écrire des familles croissantes d'événements et faire explicitement appel au fait que le temps de coalescence est fini avec probabilité 1.

Oui mais...

- On a un algorithme de génération aléatoire totalement idiot : pour le mettre en oeuvre concrètement, il faut composer des fonctions de l'ensemble d'états dans lui-même, et détecter si la composée est constante ; *a priori* ça demande de calculer les images de tous les états (les pavages), donc si on est partis pour calculer tous les pavages, autant le faire une seule fois et en choisir un de manière uniforme.

Oui mais...

- On a un algorithme de génération aléatoire totalement idiot : pour le mettre en oeuvre concrètement, il faut composer des fonctions de l'ensemble d'états dans lui-même, et détecter si la composée est constante ; *a priori* ça demande de calculer les images de tous les états (les pavages), donc si on est partis pour calculer tous les pavages, autant le faire une seule fois et en choisir un de manière uniforme.
- C'est là qu'on va utiliser l'ordre sur les pavages/fonctions de hauteur...

Couplage arrière monotone, version réaliste

- **Remarque clé** : les fonctions de mise à jour (flips) sont **croissantes** sur l'ensemble des pavages :

$$\text{si } f \leq f', \text{ alors } F(f) \leq F(f')$$

pour n'importe quelle fonction de mise à jour F .

Couplage arrière monotone, version réaliste

- **Remarque clé** : les fonctions de mise à jour (flips) sont **croissantes** sur l'ensemble des pavages :

$$\text{si } f \leq f', \text{ alors } F(f) \leq F(f')$$

pour n'importe quelle fonction de mise à jour F .

- **De plus**, la composée de fonctions croissantes est croissante.

Couplage arrière monotone, version réaliste

- **Remarque clé** : les fonctions de mise à jour (flips) sont **croissantes** sur l'ensemble des pavages :

$$\text{si } f \leq f', \text{ alors } F(f) \leq F(f')$$

pour n'importe quelle fonction de mise à jour F .

- **De plus**, la composée de fonctions croissantes est croissante.
- **Donc** pour n'importe quelle composée F de fonctions de mise à jour, on a pour toute fonction de hauteur f ,
 $F(f_{\min}) \leq F(f) \leq F(f_{\max})$, et F est constante si et seulement si $F(f_{\min}) = F(f_{\max})$.

Couplage arrière monotone, version réaliste

- **Remarque clé** : les fonctions de mise à jour (flips) sont **croissantes** sur l'ensemble des pavages :

$$\text{si } f \leq f', \text{ alors } F(f) \leq F(f')$$

pour n'importe quelle fonction de mise à jour F .

- **De plus**, la composée de fonctions croissantes est croissante.
- **Donc** pour n'importe quelle composée F de fonctions de mise à jour, on a pour toute fonction de hauteur f ,
 $F(f_{\min}) \leq F(f) \leq F(f_{\max})$, et F est constante si et seulement si $F(f_{\min}) = F(f_{\max})$.
- Par conséquent on peut simuler le couplage arrière en ne calculant que les images des fonctions de hauteur minimale et maximale (par contre il faut bien composer les fonctions dans l'ordre non naturel).

Le CFTP en pratique

- On n'est pas du tout obligé de choisir les sites des flips uniformément, ni même de prendre des fonctions de mise à jour qui ne font que des flips simples ; la seule chose qui compte c'est l'équilibre (pour que la limite soit uniforme) et qu'on puisse obtenir des fonctions de mise à jour monotones.

Le CFTP en pratique

- On n'est pas du tout obligé de choisir les sites des flips uniformément, ni même de prendre des fonctions de mise à jour qui ne font que des flips simples ; la seule chose qui compte c'est l'équilibre (pour que la limite soit uniforme) et qu'on puisse obtenir des fonctions de mise à jour monotones.
- La recherche du temps exact de coalescence conduit à un nombre de calculs qui est de l'ordre de N^2 : composer une fonction, puis 2, puis 3, etc. jusqu'à N . Mais comme on peut aller au-delà du temps de coalescence sans que le résultat ne change, on peut utiliser une stratégie de doublement : composer 1 fonction, puis 2, puis 4, etc. Au pire on va jusqu'à $2N$ et on fait de l'ordre de $4N$ calculs seulement.

Le CFTP en pratique

- On n'est pas du tout obligé de choisir les sites des flips uniformément, ni même de prendre des fonctions de mise à jour qui ne font que des flips simples ; la seule chose qui compte c'est l'équilibre (pour que la limite soit uniforme) et qu'on puisse obtenir des fonctions de mise à jour monotones.
- La recherche du temps exact de coalescence conduit à un nombre de calculs qui est de l'ordre de N^2 : composer une fonction, puis 2, puis 3, etc. jusqu'à N . Mais comme on peut aller au-delà du temps de coalescence sans que le résultat ne change, on peut utiliser une stratégie de doublement : composer 1 fonction, puis 2, puis 4, etc. Au pire on va jusqu'à $2N$ et on fait de l'ordre de $4N$ calculs seulement.
- Il existe même une version "read-once", qui permet d'appliquer l'algorithme de CFTP sans stocker les résultats de tirages aléatoires (source à lecture unique)