

Animaux dirigés, systèmes quadratiques et systèmes de réécriture.

Jean-François Marckert
Résumé par Cécile Mailler et Marie Albenque

Exposé donné le 27/01/2012 au séminaire Philippe Flajolet

Résumé

Cet exposé est consacré à l'énumération des animaux dirigés sur le réseau carré énumérés à la fois selon leur aire et selon leur périmètre. Ce problème lié à la percolation dirigée dans le quart de plan reste complètement ouvert, à la différence de l'énumération selon l'aire uniquement qui est bien compris. La méthode proposée pour résoudre ce problème est dans un premier temps de généraliser la méthode des modèles de gaz et de ramener ensuite ce problème combinatoire en un problème de résolution de systèmes matriciels. Les détails de ces travaux sont présentés dans [8]. Cet exposé est une narration des différentes étapes, qui, d'échecs en réussites, ont permis de transformer ce problème et de mieux l'appréhender.

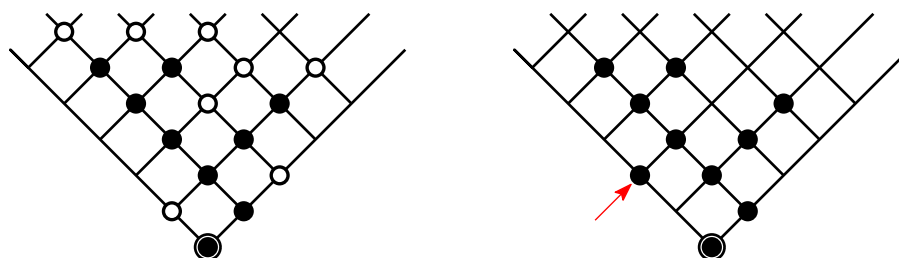
1 Introduction

On se place ici dans le quart de plan orienté, c'est-à-dire sur le graphe dont l'ensemble des sommets est indexé par \mathbb{N}^2 et tel que l'ensemble des arêtes est l'ensemble des arêtes $(i, j) \rightarrow (i+1, j)$ et $(i, j) \rightarrow (i, j+1)$ pour tout $(i, j) \in \mathbb{N}^2$. Dans toute la suite, on représentera le réseau carré tourné de $\pi/4$, de sorte que les arêtes sont naturellement orientées vers le haut, voir Figure 1.

Un **animal dirigé** est un sous-ensemble fini du quart de plan qui contient l'origine et est connexe dans les directions NORD-EST et NORD-OUEST (cf. Figure 1). Autrement dit, étant donné un animal dirigé A , pour tout sommet v de A , il existe un chemin orienté de $(0, 0)$ à v dont tous les sommets n'empruntent que des arêtes $(i, j) \rightarrow (i+1, j)$ ou $(i, j) \rightarrow (i, j+1)$. L'**aire** d'un animal dirigé A est son cardinal, noté $|A|$. L'ensemble des animaux dirigés est noté \mathcal{A} .

Un site x du quart de plan est dit **périmétrique** pour A si $A \cup \{x\}$ est un animal dirigé. On note $\mathcal{P}(A)$ l'ensemble des sites périmétriques de A ; son cardinal $|\mathcal{P}(A)|$ est appelé **périmètre** de A .

L'opérateur s'intéresse à la série génératrice bivariée $G_{A, \mathcal{P}}(x, y) = \sum_{A \text{ fini}} x^{|A|} y^{|\mathcal{P}(A)|}$ qui dénombre les animaux dirigés à la fois selon leur aire et selon leur périmètre. Avant d'étudier cette série, on commence par expliquer un lien naturel avec la percolation dirigée.



(a) Un animal dirigé d'aire 9 et de périmètre 8. Les sommets noirs représentent les cellules de l'animal et les sommets blancs les sites périmétriques

(b) Ce sous-ensemble n'est pas un animal dirigé; il n'existe pas de chemin orienté de l'origine au sommet pointé.

FIGURE 1 – Définition d'un animal dirigé dans le quart de plan.

Considérons le modèle de percolation de Bernoulli classique dans lequel chaque site du quart de plan est conservé avec probabilité p ou effacé avec probabilité $1 - p$. L'*amas* (ou *cluster*) de percolation issu de l'origine $A(p)$ est la composante connexe dirigée dans le graphe ainsi obtenu. On dit que le système *percole* ou qu'il y a *percolation*, s'il existe un cluster infini à partir de v , autrement dit si $A_v(p)$ est infini.

On peut montrer qu'il existe $0 < p_{crit} < 1$, appelé *seuil de percolation* ou *percolation critique*, tel que pour $p < p_{crit}$, il n'y a pas percolation avec probabilité 1 et pour $p > p_{crit}$, il y a percolation avec probabilité strictement positive. Déterminer p_{crit} pour le réseau carré est un problème ouvert. L'observation suivante permet d'établir un lien entre le calcul de p_{crit} et le calcul de la série aire-périmètre. Soit A un animal dirigé fixé, la probabilité que pour une réalisation aléatoire de percolation $A(p)$ coïncide avec A est égale à :

$$P(A(p) = A) = p^{|A|}(1 - p)^{|P(A)|},$$

en effet il faut et il suffit que chaque cellule de l'animal soit conservée et que chaque site périmétrique soit effacé (car $A(p)$ est maximal pour l'inclusion). La probabilité que le cluster de percolation soit fini est donc égal à :

$$P(|A(p)| < \infty) = \sum_{A \in \mathcal{A}} p^{|A|}(1 - p)^{|P(A)|}.$$

Autrement dit, $p \leq p_{crit}$ si et seulement si $G_{\mathcal{A}, \mathcal{P}}(p, 1 - p) = 1$. Connaître la série aire/périmètre permettrait donc probablement de pouvoir calculer la valeur de p_{crit} .

L'orateur souligne deux résultats connus sur cette série aire/périmètre : le premier dû à A. Bacher [1] qui a déterminé la série $\sum_{A \text{ fini}} |\mathcal{P}(A)|x^{|A|}$, et le second, détaillé dans une première partie de l'exposé.

2 Énumération des animaux dirigés selon leur aire

Le série génératrice $G_A(x) = \sum_{A \text{ fini}} x^{|A|} = G_{\mathcal{A}, \mathcal{P}}(x, 1)$ qui énumère les animaux dirigés uniquement selon leur aire est bien connue et a fait l'objet de nombreux travaux. L'orateur cite notamment le résultat suivant

Proposition 2.1. $G_A(x) = \frac{2x}{1 - 3x + (1 - 2x - 3x^2)}.$

Plusieurs démonstrations de cette identité sont connues : des démonstrations utilisant des méthodes combinatoires (Viennot 86, Bétréma et Penaud 93, Gouyou-Beauchamps et Viennot 88, etc) et d'autres fondées sur l'étude de modèles de gaz (Hakim et Nadal 83, Dhar 83, Bousquet-Melou 98, Le Borgne et Marckert 07, etc). Comme c'est cette approche qui sera développée dans la suite pour l'étude de la série aire/périmètre, l'orateur détaille les étapes de l'approche reposant sur les modèles de gaz.

Animaux sur un cylindre. La première idée pour simplifier l'étude est de travailler sur un cylindre de largeur N fixée et non plus sur un cône (cf. Figure 2). Le théorème limite suivant permet de relier les deux modèles : si $G_A^{(N)}(x)$ est la fonction génératrice des animaux dirigés sur le cylindre, où x marque l'aire, alors

$$\lim_{N \rightarrow \infty} G_A^{(N)}(x) \rightarrow G_A(x) \text{ quand } N \rightarrow \infty,$$

au sens des séries formelles. En effet le nombre d'animaux dirigés de taille n fixée est le même sur le cylindre de largeur N et sur le quart de plan dès que $N > n$.

Cette restriction permet d'obtenir une équation de décomposition des animaux dirigés ligne à ligne. Pour cela, on généralise un peu la notion d'animaux dirigés, soit C un ensemble de sommets de la première ligne du cylindre, l'ensemble de sommets A est appelé *animal dirigé de source C* si et seulement s'il contient C et pour tout $x \in A$, il existe $c \in C$ et un chemin orienté de c à a

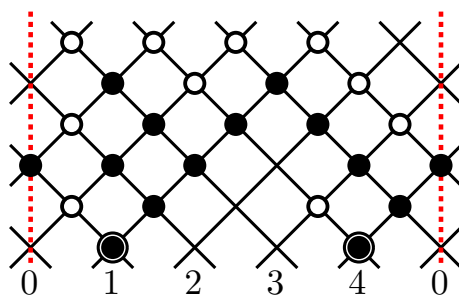


FIGURE 2 – Un animal dirigé sur le cylindre de largeur 5. Les sommets noirs représentent les cellules de l’animal et les sommets entourés ses sources, les sommets blancs désignent les sites périmétriques.

n’empruntant que des sommets de A . On note $G_C(x)$ la génératrice des animaux dirigés de source C . On a alors l’équation de décomposition suivante :

$$G_C^{(N)}(x) = x^{|C|} \sum_{D \subset V(C)} G_D^{(N)}(x),$$

où $V(C)$ désigne le voisinage de C c’est à dire les sommets qui sont atteints, partant d’un point de C par un pas NORD-EST ou un pas NORD-OUEST.

Modèle de gaz et animaux. Un *modèle de gaz* sur un graphe est une collection de variables aléatoires $(X_x)_{x \in V(G)}$ à valeurs dans $\{0, 1\}$ indexée par les sommets du graphe (ici par les sommets du cylindre). La terminologie provient de la physique statistique où cet objet permet de modéliser l’occupation d’un gaz : la valeur du gaz en un site est égale à 1 si et seulement si une particule est présente à cet endroit. Soit x un sommet, on dit que x est *occupé* si $X_x = 1$ et *libre* sinon

Définissons à présent un modèle de gaz sur le cylindre. On commence par définir une chaîne de Markov verticale sur les lignes du cylindre de largeur N grâce aux transitions représentées en Figure 3. Autrement dit, si les deux enfants d’un sommet sont libres alors celui-ci sera occupé avec probabilité p et libre dans tous les autres cas. Cela nous donne une chaîne de Markov irréductible sur un espace à 2^N états qui admet donc une unique loi stationnaire $\pi^{(N)}$.

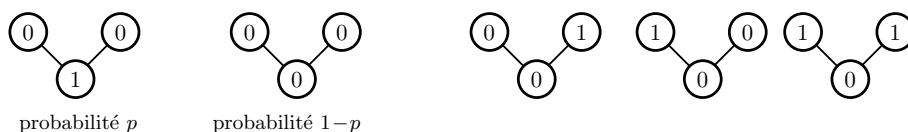


FIGURE 3 – Transitions de la chaîne de Markov du modèle de gaz.

Pour toute source C , notons $F_C^{(N)}(p) = \pi^{(N)}(X_x = 1 \Leftrightarrow x \in C)$, le lien entre ce modèle de gaz et l’énumération d’animaux dirigés est donné par le résultat suivant, dû à Dhar :

Théorème 2.2. *Pour toute source C , pour tout $N \in \mathbb{N}$,*

$$F_C^{(N)}(p) = (-1)^{|C|} G_C^{(N)}(-p).$$

En particulier, pour une source réduite à un sommet, on obtient :

$$\pi^{(N)}(X_x = 1) = -G^{(N)}(-p).$$

Il “suffit” donc de trouver l’expression de la loi stationnaire $\pi^{(N)}$ et de calculer la limite de $\pi^{(N)}(X_x = 1)$ lorsque N tend vers l’infini, pour trouver l’expression de la fonction génératrice de l’aire. Son expression peut-être devinée grâce à l’utilisation d’un logiciel de calcul formel et il suffit ensuite de montrer que la mesure devinée convient bien.

3 Énumération selon l'aire et le périmètre

Si l'énumération des animaux dirigés selon leur aire est bien comprise, c'est loin d'être le cas dans le cas aire/périmètre. L'orateur présente des idées et des nouveaux résultats sur le sujet : il montre en particulier dans quelle mesure la méthode du modèle de gaz peut-être adaptée au cas aire/périmètre.

Définition du modèle de gaz On se place toujours sur le cylindre de largeur N et on définit un nouveau modèle de gaz (Y_x) toujours en s'appuyant à nouveau sur la donnée des transitions ligne à ligne (cf. Figure 4). Dans ce nouveau modèle 2 types de transition sont possibles : avec probabilité p , on applique le premier type et un sommet va être occupé si et seulement ses deux enfants sont occupés, sinon (donc avec probabilité $1 - p$) un sommet est occupé avec probabilité q indépendamment de tout le reste, autrement dit, en notant c_1 et c_2 les deux enfants d'un sommet x , on a :

$$Y_x = B_p \min(Y_{c_1}, Y_{c_2}) + (1 - B_p)B_q,$$

où B_p et B_q sont deux variables aléatoires indépendantes de loi de Bernoulli de paramètre respectif p et q . À nouveau, cela permet de définir une chaîne de Markov irréductible à états finis qui admet donc une unique probabilité stationnaire \mathbb{P} .

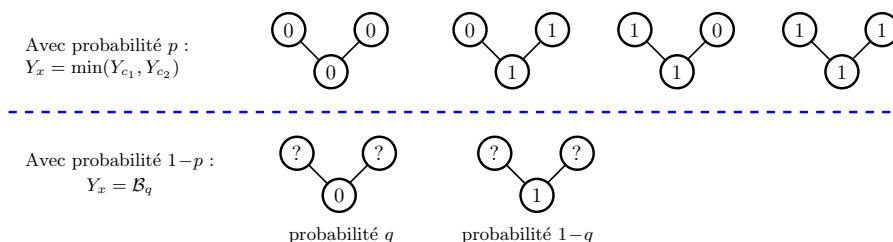


FIGURE 4 – Transitions de la chaîne de Markov du modèle de gaz : cas aire/périmètre.

Le lien entre la série génératrice aire-périmètre et ce modèle de gaz est alors donné par le résultat suivant :

Proposition 3.1. *Pour tout sommet x du graphe, on a :*

$$\mathbb{P}(Y_x = 1) = \sum_{A \text{ de source } x} p^{|A|} (1 - p)^{|\mathcal{P}(A)|} q^{|\mathcal{P}(A)|} = G_{A, \mathcal{P}}(p, (1 - p)q).$$

Ébauche de preuve. Dans ce modèle de gaz, on associe à chaque sommet une variable de Bernoulli de paramètre p et une de paramètre q . On utilise la variable de paramètre p pour définir un modèle de percolation comme décrit dans l'introduction. On peut donc écrire :

$$\mathbb{P}(Y_x = 1) = \sum_{A \text{ de source } x} \mathbb{P}(A(p) = A \text{ et } Y_x = 1) = \sum_{A \text{ de source } x} \mathbb{P}(Y_x = 1 | A(p) = A) \cdot \mathbb{P}(A(p) = A).$$

Il suffit ensuite d'observer que si $A(p)$ est fixé alors Y_x vaut 1 si et seulement si tous les sites paramétriques de $A(p)$ ont une variable de paramètre q égale à 1. On obtient donc :

$$\mathbb{P}(Y_x = 1) = \sum_{A \text{ de source } x} p^{|A|} (1 - p)^{|\mathcal{P}(A)|} q^{|\mathcal{P}(A)|} = G_{A, \mathcal{P}}(p, (1 - p)q).$$

□

Malheureusement, deviner la loi stationnaire de la chaîne de Markov définie ci-dessus s'avère plus difficile que dans le cas de l'aire, et cette loi est aujourd'hui encore inconnue. Plus précisément, pour une petite largeur N de cylindre fixée, on peut calculer la mesure stationnaire, mais les simulations réalisées avec un logiciel de calcul formel laissent à penser qu'il n'existe pas de forme produit qui permettrait de passer ensuite à la limite. En effet contrairement à ce qu'il se passe pour le cas de

l'énumération selon l'aire, il n'existe pas de matrice M de taille 2×2 (qui ne dépend pas de N et une constante de renormalisation Z_N telles que :

$$\mathbb{P}^{(N)}(x_1, \dots, x_N) = Z_N \prod_{i=1}^N M_{x_i, x_{i+1}}. \quad (1)$$

La méthode de l'aire semble donc être en échec sur le problème de l'aire/périmètre : l'orateur montre comment transformer le problème pour utiliser des méthodes algébriques.

Vision matricielle. Comme on ne parvient trouver une mesure stationnaire sous forme produit (cf. Equation (1), on va se donner plus de degrés de liberté en recherchant des solutions sous forme matricielle, c'est-à-dire en considérant que les coefficients de la matrice M peuvent être également des matrices. Plus précisément on cherche une mesure sous la forme

$$\pi(x_1, \dots, x_N) = \text{Tr} \left(\prod_{i=1}^N Q^{x_i} \right),$$

où Q^0 et Q^1 sont deux matrices de la forme $Q^0 = V^0 H^0$ et $Q^1 = V^1 H^1$ telles que $V^1 H^0 = V^0 H^1 = 0$.

Chercher des solutions sous cette forme n'est pas restrictif, on peut en effet montrer que pour n'importe quelle mesure de probabilités sur $\{0, 1\}^N$ invariante par rotation (ce qui est le cas ici, car la définition du gaz est invariante par rotation), on peut trouver de telles matrices.

Le résultat principal est le suivant :

Proposition 3.2. *Soit $T_{y,y',x}$ la probabilité d'obtenir x en un site sachant que ses deux voisins (par un pas NORD-EST ou un pas NORD-OUEST) ont pour valeurs respectives y et y' . Supposons qu'il existe des matrices carrées V^0, H^0, V^1 et H^1 telles que :*

$$\begin{cases} V^x H^x &= \sum_{y,y'} H^y V^{y'} T_{y,y',x} \\ V^x H^{1-x} &= 0 \end{cases}$$

alors $\pi(x_1, \dots, x_N) = \text{Tr} \left(\prod_{i=1}^N Q^{x_i} \right)$ pour Q^0 et Q^1 associés à cette solution, est mesure invariante pour le modèle de gaz considéré.

Il "suffirait" donc de résoudre ce système. Malheureusement, on n'y parvient pas. À ce stade, on peut soit se décourager et tout arrêter, soit continuer : bien sûr on continue ! L'orateur introduit une méthode itérative pour contourner ce problème.

Construction d'une suite de lois. Soit μ_0 une loi sur les N -uplets de $\{0, 1\}^N$. On définit par récurrence une suite de lois $(\mu_k)_k$ en appliquant itérativement les transitions de la chaîne de Markov qui définissent le gaz. La suite ainsi définie converge vers la loi stationnaire π . L'idée est de choisir μ_0 de façon à pouvoir calculer μ_k explicitement pour tout $k \geq 0$. Nous allons construire donc construire une suite $(V_k^0, H_k^0, V_k^1, H_k^1)_{k \geq 0}$ telle que $V_0^x H_0^{1-x} = 0$ et pour tout $k \geq 0$:

$$\begin{cases} V_k^x H_k^x &= \sum_{y,y'} H_{k-1}^y V_{k-1}^{y'} T_{y,y',x} \\ V_k^x H_k^{1-x} &= 0. \end{cases}$$

En fait il serait suffisant de trouver des valeurs $(h_{y,x}, v_{y,x})_{x,y \in \{0,1\}}$ telles que

$$h_{y,x} v_{y',x'} = T_{y,y',x} \mathbb{1}_{x=x'}, \quad (2)$$

pour tout x, x', y, y' dans $\{0, 1\}$. En effet dans ce cas, on obtiendrait une solution en construisant une suite de matrices comme suit :

$$\begin{cases} V_k^x &= \sum_y H_{k-1}^y h_{y,x} \\ H_k^x &= \sum_{y'} V_{k-1}^{y'} v_{y',x}, \end{cases}$$

Malheureusement l'Équation (2) n'a pas de solution dans \mathbb{R} . On cherche donc une solution sous forme matricielle, c'est-à-dire des vecteurs lignes et colonnes sous la forme :

$$h_{x,1} = [\star, 0, \star, 0] \quad h_{x,0} = [0, \star, 0, \star]$$

$$v_{y,1} = \begin{bmatrix} \star \\ 0 \\ \star \\ 0 \end{bmatrix} \quad v_{y,0} = \begin{bmatrix} 0 \\ \star \\ 0 \\ \star \end{bmatrix}$$

et tels que (2) soit vérifiée. Dans ce cas on obtient une solution en définissant une suite de matrices par :

$$\begin{cases} V_k^x = \sum_y H_{k-1}^y \otimes h_{y,x} \\ H_k^x = \sum_{y'} H_{k-1}^{y'} \otimes v_{y',x} \end{cases}$$

où \otimes représente le produit tensoriel de matrice. Le produit tensoriel de deux matrices $A \otimes B$ est la matrice par blocs

$$A \otimes B = \begin{bmatrix} a_{11}B & a_{21}B & \cdots \\ a_{12}B & a_{22}B & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix}.$$

Notons que $(A_1 \otimes B_1)(A_2 \otimes B_2) = (A_1 A_2) \otimes (B_1 B_2)$, de plus, $\text{Tr}(A \otimes B) = \text{Tr}(A)\text{Tr}(B)$. Introduire de tels vecteurs lignes et colonnes pour se donner plus de degrés de liberté et pour ainsi trouver une solution est une méthode qui apparaît dans d'autres domaines notamment dans l'étude de processus d'exclusion (PASEP par exemple).

Ce nouveau système peut alors être résolu. Il ne reste plus qu'à choisir une mesure initiale – par exemple $V^1 = H^1 = [1]$ et $V^0 = H^0 = [0]$ – et à faire converger la suite des Q_k^x point par point quand k tend vers l'infini vers une matrice (de dimension infinie) Q_∞^x . On a alors le système suivant

$$Q_\infty^x = \sum_y Q_\infty^y \otimes q_{y,x}.$$

Malheureusement, ce système est difficile à résoudre, mais le problème est maintenant énoncé sous une forme algébrique. Cette méthode peut par ailleurs être généralisée à d'autres questions sur les animaux dirigés.

Références

- [1] A. Bacher. *Chemins et animaux : applications de la théorie des empilements de pièces*. PhD thesis, Univ. Bordeaux 1, 2011.
- [2] J. Bétréma and J.-G. Penaud. Modèles avec particules dures, animaux dirigés, et séries en variables partiellement commutatives. [arXiv:math.CO/0106210](https://arxiv.org/abs/math/0106210), 1993.
- [3] M. Bousquet-Mélou. New enumerative results on two-dimensional directed animals. *Discrete Math.*, 180 :73–106, 1998.
- [4] D. Dhar. Exact solution of a directed-site animals-enumeration problem in three dimensions. *Phys. Rev. Lett.*, 51(10) :853–856, 1983.
- [5] D. Gouyou-Beauchamps and G. Viennot. Equivalence of the two-dimensional directed animal problem to a one-dimensional path problem. *Adv. in Appl. Math.*, 9(3) :334–357, 1988.
- [6] V. Hakim and J.P. Nadal. Exact results for 2d directed animals on a strip of finite width. *Journal of Physics A : Mathematical and General*, 16 :L213, 1983.
- [7] Y. Le Borgne and J.-F. Marckert. Directed animals and gas models revisited. *Electronic journal of Combinatorics*, 14, 2007.
- [8] J.-F. Marckert. Directed animals, quadratic and rewriting systems. *Arxiv preprint arXiv :1112.0910*, 2011.
- [9] G. Viennot. Heaps of pieces. I : Basic definitions and combinatorial lemmas. *Lecture Notes in Math.*, 1234 :321–350, 1986.