



HAL
open science

De la probabilité de creuser un tunnel

Elie de Panafieu, Mohamed Lamine Lamali, Michael Wallner

► **To cite this version:**

Elie de Panafieu, Mohamed Lamine Lamali, Michael Wallner. De la probabilité de creuser un tunnel. ALGOTEL 2019 - 21èmes Rencontres Francophones sur les Aspects Algorithmiques des Télécommunications, Jun 2019, Saint Laurent de la Cabrerisse, France. hal-02123269

HAL Id: hal-02123269

<https://hal.science/hal-02123269>

Submitted on 7 May 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

De la probabilité de creuser un tunnel

Élie de Panafieu¹ et Mohamed Lamine Lamali^{2 †} et Michael Wallner^{2 ‡}

¹Nokia Bell Labs, France

²LaBRI - Univ. de Bordeaux, France

Un tunnel est une portion d'un chemin où un protocole de communication est encapsulé dans un autre. Les tunnels sont de plus en plus présents dans les réseaux. Ils sont utilisés pour la sécurité, l'interopérabilité, les réseaux virtuels, etc. Un chemin contenant un tunnel doit être *faisable*, c'est-à-dire qu'il doit respecter l'ordre des encapsulations et des désencapsulations, parfois imbriquées. Les algorithmes de calcul de chemins avec prise en compte de tunnels ne sont pas les mêmes que dans un réseau classique. Afin de tester ces algorithmes, il serait utile de disposer d'un générateur de topologies aléatoires avec tunnels. Ce générateur devra produire des topologies ayant certaines caractéristiques (par exemple, que la probabilité qu'un chemin soit faisable soit un paramètre p quelconque). Il se trouve que ce problème est lié aux *chemins de Dyck et de Motzkin* en combinatoire. En effet, une encapsulation peut être vue comme un pas montant et une désencapsulation peut être vue comme un pas descendant. La question qui se pose est donc de savoir quelle est la probabilité qu'un chemin de Dyck ou de Motzkin ait certaines propriétés. Dans ce travail, en utilisant des outils de combinatoire analytique, nous donnons la probabilité qu'un chemin soit *faisable* en fonction de la distribution des encapsulations et des désencapsulations sur les sommets.

La version complète de cet article est [dPLW19](<https://arxiv.org/abs/1601.01786>)

Mots-clés : Combinatoire analytique, Encapsulation de protocoles, Tunnels.

1 Introduction

1.1 Problème et motivation

Commençons par une illustration (très) imagée du problème avec des poupées russes. Supposons qu'on ait $n+1$ personnes alignées. Chaque personne i reçoit une poupée de la personne $i-1$, effectue une opération, puis la passe à son tour à la personne $i+1$. Il y a trois sortes de personnes. Celles qui, recevant une poupée, sont uniquement capables de la mettre dans une poupée plus grande. Celles qui sont uniquement capables d'extraire une poupée plus petite de celle reçue. S'ils reçoivent la plus petite, ils la jettent. Enfin, celles qui sont capables de faire les deux. La question qui se pose est la suivante : si on donne la plus petite poupée à la première personne sur la ligne (la personne 1), est-il possible que la personne $n+1$ la reçoive telle quelle ?

Si on remplace les poupées par des paquets, la question est de savoir si, étant donné un chemin avec $n+1$ routeurs, certains étant capables d'encapsuler le paquet, d'autres de désencapsuler le paquet, et enfin d'autres capables de faire les deux, est-ce que le dernier routeur peut recevoir le paquet tel qu'émis par le premier routeur ?

1.2 Analogie avec les chemins de Dyck et de Motzkin

Les chemins de Dyck sont des chemins dans le quart de plan. Démarrant de la position $(0,0)$ (en coordonnées cartésiennes), on dispose des deux pas suivants : $(+1,1)$ (pas montant) et $(-1,1)$ (pas descendant), que nous noterons $\{+1\}$ et $\{-1\}$ par la suite. Les chemins de Motzkin disposent du pas supplémentaire $(0,1)$, noté $\{0\}$ par la suite. La Figure 1a illustre un tel chemin. On dit qu'un

[†]Financé par le projet HÉRA ANR-18-CE25-0002

[‡]Financé par Exzellenzstipendium of the Austrian Federal Ministry of Education, Science and Research and the Erwin Schrödinger Fellowship of the Austrian Science Fund (FWF): J 4162-N35.

chemin qui ne descend jamais sous la ligne d'abscisse 0 et qui se termine au point $(0, n)$ que c'est une *excursion*. Ces chemins ont été extensivement étudiés en combinatoire.

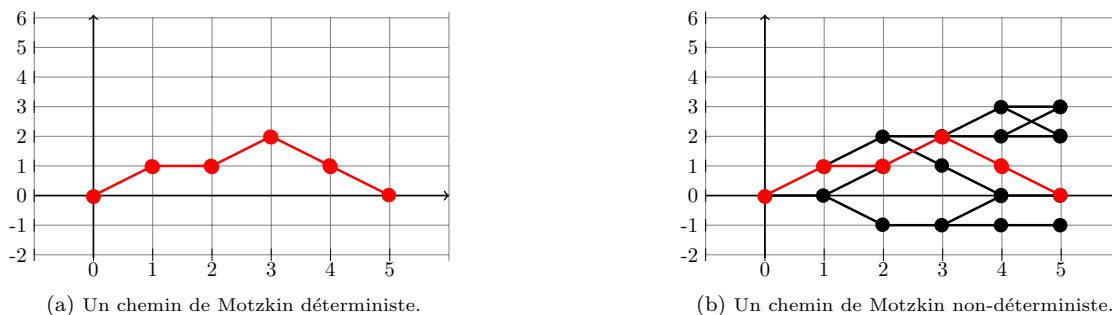


FIGURE 1: Illustration d'un chemin de Motzkin déterministe et non-déterministe.

Néanmoins, on peut qualifier ces chemins de *déterministes* car à chaque étape, on ne dispose que d'un seul choix. Or, la présence de routeurs qui peuvent soit encapsuler, soit désencapsuler, nous oblige à prendre en compte la présence de plusieurs choix possibles. Nous définissons donc les chemins de Dyck et de Motzkin *non-déterministes* (notés N-chemins), qui permettent d'explorer les différents chemins possibles (voir Figure 1b). Ainsi, un N-chemin de Dyck dispose des pas $\{+1, -1, -1, +1\}$ et un N-chemin de Motzkin dispose des pas

$$\{-1, \{0, \{+1, \{-1, 0\}, \{-1, +1\}, \{0, +1\}, \{-1, 0, +1\}\}$$

La question devient : étant donné un ensemble de pas, avec possibilité à un point donné de choisir soit de monter, soit de descendre, soit de rester au même niveau (le cas des routeurs qui peuvent encapsuler, désencapsuler, ou transmettre un paquet tel quel), existe-il une suite de choix permettant d'atteindre le point $(0, n)$ sans jamais passer sous 0? En d'autres termes, est-ce qu'un N-chemin contient une excursion?

2 Combinatoire analytique, séries génératrices et analyse de singularités

Cette section a pour but de présenter la combinatoire analytique, qui fournit de puissants outils pour résoudre ce genre de problèmes. La combinatoire analytique s'intéresse à l'énumération de certaines structures formelles en utilisant des méthodes d'analyse complexe. Soit par exemple C_n le nombre d'arbres binaires § contenant n nœuds internes. On sait que $C_0 = 1$, $C_1 = 1$, $C_2 = 2$, $C_3 = 5$, $C_4 = 14$, $C_5 = 42$, etc.

On peut associer à cette suite une série formelle

$$C(z) = C_0 + C_1z + C_2z^2 + C_3z^3 + \dots = \sum_{i=0}^{\infty} C_i z^i .$$

On peut obtenir une expression close pour la fonction $C(z)$ de manière purement formelle. En effet, un arbre binaire est soit un nœud isolé, soit ayant un fils gauche et un fils droit, eux-mêmes arbres binaires, on obtient formellement

$$C(z) = 1 + zC(z)^2 \quad \text{et donc} \quad C(z) = \frac{1 - \sqrt{1 - 4z}}{2z} .$$

On considère maintenant $C(z)$ comme une fonction sur les nombres complexes. Son comportement près de sa singularité $\rho = 1/4$ dicte l'asymptotique de ses coefficients. Sous certaines conditions

§. Il s'agit des nombres de Catalan.

De la probabilité de creuser un tunnel

techniques ici vérifiées, le *théorème de transfert* [FS09] affirme que pour une fonction donnée $F(z)$, si $C(z) \underset{z \rightarrow \rho}{\sim} F(z)$, alors $C_n \underset{n \rightarrow \infty}{\sim} F_n$. On obtient ici

$$C(z) \underset{z \rightarrow \rho}{\sim} 2 - 2\sqrt{1-4z} \quad \text{et donc} \quad C_n \underset{n \rightarrow \infty}{\sim} \frac{4^n}{n^{3/2}\sqrt{\pi}}.$$

Ce que vérifient bien les nombres de Catalan. Cette technique, appelée *analyse de singularité*, s'étend à toute série génératrice algébrique et permet également d'extraire des lois limites de paramètres des objets étudiés.

3 Résultats sur le chemins de Dyck et de Motzkin

Nous appliquons cette technique au problème de l'énumération des N-chemins de Dyck et de Motzkin qui contiennent des excursions. À partir de la même méthode que précédemment, nous construisons les fonctions génératrices des N-chemins contenant des excursions. Les preuves des différents théorèmes peuvent être trouvées dans [dPLW19]. Plus précisément, nous donnons la fonction génératrice de la probabilité qu'un N-chemin contienne une excursion en fonction de la distribution de probabilité sur les pas.

Soit p_{-1} (resp. p_1 et $p_{-1,1}$) la probabilité qu'un nœud dispose du pas $\{-1\}$ (resp. $\{+1\}$ et $\{-1, +1\}$). Le théorème 1 donne la fonction génératrice de la probabilité qu'un N-chemin de Dyck contienne une excursion.

Théorème 1. *La fonction génératrice de la probabilité qu'un N-chemin de Dyck contienne une excursion est*

$$D(z) = \frac{X(z)}{1 - X(z)^2} \frac{1 - X(z)Y(z)}{(p_{-1} + p_{-1,1})z} \quad (1)$$

où

$$X(z) = \frac{1 - \sqrt{1 - 4p_1(p_{-1} + p_{-1,1})z^2}}{2p_1z} \quad \text{et} \quad Y(z) = \frac{1 - \sqrt{1 - 4p_{-1}(p_1 + p_{-1,1})z^2}}{2(p_1 + p_{-1,1})z}.$$

Évidemment, un N-chemin de Dyck ne peut contenir une excursion que s'il est de longueur paire. En analysant les singularités de la fonction (1), on obtient l'asymptotique des probabilités.

Théorème 2. *La probabilité pour un N-chemin de Dyck aléatoire de longueur $2n$ quand $n \rightarrow \infty$ de contenir une excursion est asymptotique à :*

$$\begin{aligned} & - \frac{(1-2p_1)(1-2p_{-1})}{(1-p_1)(1-p_{-1})} + O\left(\frac{(4p_{-1}(1-p_{-1}))^n}{n^{3/2}}\right) \quad \text{si } 0 < p_1 \leq p_{-1} < \frac{1}{2}, \\ & - \frac{1-2p_1}{(1-p_1)\sqrt{\pi n}} + O\left(\frac{1}{n^{3/2}}\right) \quad \text{si } 0 < p_1 < \frac{1}{2} \text{ et } p_{-1} = \frac{1}{2}, \\ & - \frac{1}{\sqrt{\pi n^3}} + O\left(\frac{1}{n^{5/2}}\right) \quad \text{si } p_1 = p_{-1} = \frac{1}{2}, \\ & - O\left(\frac{(4p_{-1}(1-p_{-1}))^n}{n^{3/2}}\right) \quad \text{si } 0 < p_1 < \frac{1}{2} < p_{-1} < 1 \text{ et } p_{-1} + p_1 \leq 1. \end{aligned}$$

Par exemple, si $p_{-1} = p_1 = p_{-1,1} = 1/3$, alors cette probabilité tend vers $1/4$.

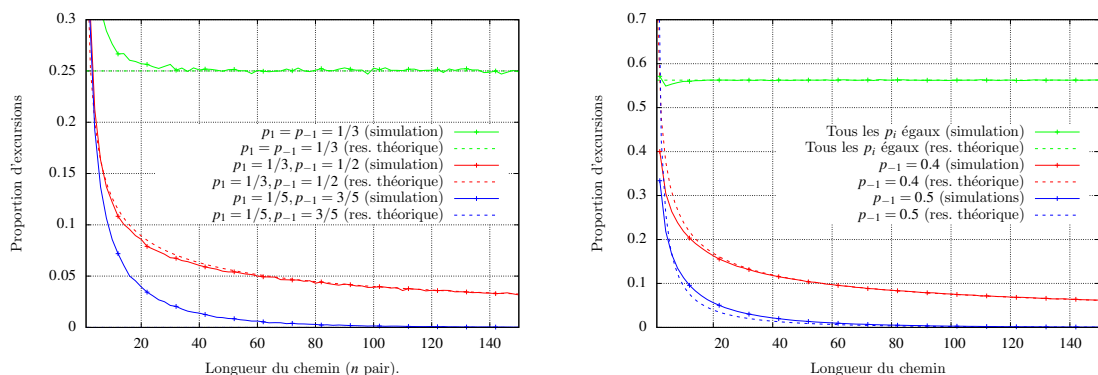
La fonction génératrice de la probabilité qu'un N-chemin de Motzkin contienne une excursion est malheureusement trop volumineuse pour être donnée ici. Néanmoins l'analyse de singularités donne le résultat suivant

Théorème 3. *La probabilité qu'un N-chemin de Motzkin de longueur n contienne une excursion si tous les pas sont équiprobables (i.e., égaux à $1/7$) est asymptotique à*

$$\frac{9}{16} - \frac{\gamma}{\sqrt{\pi n^3}} \left(\frac{6}{7}\right)^n + O\left(\frac{(6/7)^n}{n^{5/2}}\right),$$

où $\gamma \approx 0.6183$ est la solution positive de l'équation $1024\gamma^4 - 8019\gamma^2 + 2916 = 0$.

Ainsi, pour des pas équiprobables, environ $56,25\%$ des N-chemins de Motzkin contiennent des excursions. On peut obtenir cette probabilité pour des pas non-équiprobables (comme on le verra dans la Section 4), mais cela nécessite un calcul particulier pour chaque cas.



(a) Proportion de N-chemins de Dyck contenant une excursion.

(b) Proportion de N-chemins de Motzkin contenant une excursion.

FIGURE 2: Comparaison des prédictions théoriques et des résultats de simulations (sur 10^5 runs) de la proportion de N-chemins de Dyck et de Motzkin contenant une excursion.

4 Simulations

Nous avons effectué des simulations pour comparer les résultats théoriques et empiriques. La Figure 2a montre la probabilité qu'un N-chemin de Dyck contienne une excursion. Selon le Théorème 2, si les trois pas sont équiprobables, la probabilité converge vers $1/4$. Si $p_1 = 1/3$ et $p_{-1} = 1/2$, la probabilité converge vers $1/\sqrt{2\pi n}$. Et si $p_1 = 1/5$ et $p_{-1} = 3/5$, la probabilité décroît exponentiellement. On peut voir sur la Figure 2a que les prédictions théoriques sont parfaitement vérifiées.

La Figure 2b montre l'évolution de la probabilité qu'un N-chemin de Motzkin contienne une excursion. Le calcul des probabilités est plus complexe. D'après le Théorème 3, la probabilité tend vers $9/16$ si les pas sont équiprobables. Des calculs plus complexes nous ont permis de déduire que cette probabilité est asymptotique à $3/7\sqrt{10/\pi n}$ si $p_{-1} = 2/5$ et que les autres pas sont équiprobables. Pour le cas $p_{-1} = 1/2$, la formule est trop complexe pour être donnée ici, mais elle décroît exponentiellement. La Figure 2b montre encore que les simulations confirment parfaitement les résultats théoriques.

5 Conclusion et questions ouvertes

Dans cet article, nous avons donné la probabilité qu'un chemin puisse contenir un tunnel en fonction de la distribution de probabilités des capacités d'encapsulations et de désencapsulations des routeurs. Nous avons utilisés des outils de combinatoire analytique qui nous ont donné des résultats très précis confirmés par les simulations. Néanmoins, plusieurs questions restent ouvertes. Dans ce travail, nous avons supposé qu'un routeur peut encapsuler n'importe quel protocole dans un autre. Or, si on différencie les protocoles, le problème est bien plus complexe car la suite des opérations n'est plus commutative (encapsuler dans IPv4 puis IPv6, n'est pas la même chose que de faire l'inverse.) Il serait également intéressant de généraliser le problème à un graphe quelconque au lieu d'un chemin.

Remerciements : les auteurs remercient Mireille Bousquet-Mélou pour ses commentaires très utiles et pertinents.

Références

- [dPLW19] É. de Panafieu, M. L. Lamali, and M. Wallner. Combinatorics of nondeterministic walks of the dyck and motzkin type. In *Analytic Algorithmics and Combinatorics, ANALCO*, 2019.
- [FS09] Philippe Flajolet and Robert Sedgewick. *Analytic Combinatorics*. Cambridge University Press, 2009.