

THÈSE
PRÉSENTÉE À
L'UNIVERSITÉ BORDEAUX I
ÉCOLE DOCTORALE DE MATHÉMATIQUES ET
D'INFORMATIQUE
Par **Yvan Le Borgne**
POUR OBTENIR LE GRADE DE
DOCTEUR
SPÉCIALITÉ : INFORMATIQUE

**Variations combinatoires sur des classes d'objets comptées par la
suite de Catalan**

Soutenue le : 13 Décembre 2004

Après avis des rapporteurs :

Dominique Gouyou-Beauchamps Professeur
Volker Strehl Professeur

Devant la commission d'examen composée de :

Mireille Bousquet-Mélou	Directrice de Recherche	Directrice
Robert Cori	Professeur	Président
Dominique Gouyou-Beauchamps	Professeur	Rapporteur
Bruno Salvy	Directeur de Recherche	Examineur
Volker Strehl	Professeur	Rapporteur
Xavier Viennot	Directeur de Recherche	Examineur

Remerciements

Je remercie ma directrice, Mireille Bousquet-Mélou, pour son encadrement bienveillant, le temps qu'elle m'a consacré et ses précieux conseils lors de la rédaction de ce manuscrit. J'ai apprécié l'intérêt constant et rigoureux qu'elle a manifesté pour toutes mes recherches, m'offrant un avis précieux sur mes travaux. Sa présence m'a aussi permis d'oser croiser le fer avec des équations fonctionnelles conséquentes, un combat que jusque là ma morale réprouvait. Enfin, le questionnement accompagnant la lecture de ses travaux fut aussi pour moi une importante source d'inspiration.

Robert Cori m'a fait l'honneur de présider mon jury. Avant cela, il a su rendu agréable mes premiers pas en combinatoire lors de mon stage de DEA. Notre collaboration s'est prolongée pendant ma thèse dans la suite logique de ce travail sur le modèle du tas de sable. Merci.

Dominique Gouyou-Beauchamps et Volker Strehl ont immédiatement accepté d'être les rapporteurs de ma thèse, je les remercie pour leur lecture attentive et leurs commentaires.

Je remercie Xavier Viennot et Bruno Salvy d'avoir participé à mon jury.

J'ai une pensée pour les membres du groupe de travail "Combinatoire énumérative et Génération Aléatoire" qui contribuent à rendre ce rendez-vous hebdomadaire stimulant.

Merci aux membres du laboratoire avec qui j'ai collaboré pour mes enseignements. Je remercie en particulier ceux qui ont allégé mon travail pendant ces derniers mois de rédaction. Merci également à Pierre-André Wacrenier pour nos discussions pédagogiques et Raymond Namyst pour m'avoir fait confiance.

Un grand merci aux (ex-)doctorants bordelais : Afif, Aymeric, Bilel, Bertrand, Bertrand, David, David, Claire, Emmanuelle, Florian, Nicolas, Pascal, Pascal, Coach Pierre, Pierre, Sylvie, Valère, Yon.

Merci à ma famille et en particulier mon père et ma mère.

Merci Sandra.

Cette thèse se situe dans le domaine de la combinatoire énumérative.

Il existe plusieurs techniques d'énumération des animaux dirigés selon l'aire : empilements, bijections avec des chemins, équivalence avec la densité d'un modèle de gaz. On en propose des extensions pour prendre aussi en compte des informations sur le périmètre de site.

En utilisant un formalisme proche des grammaires algébriques et l'ordinateur, on propose un schéma commun à plusieurs bijections autour des mots de Dyck.

La série génératrice des chemins de Dyck énumérés selon leur longueur et les interactions supérieures est un quotient de q -séries hypergéométriques basiques dans lesquelles apparaît un terme algébrique, une forme récente dans le cadre de la combinatoire énumérative. On donne plusieurs preuves de plus en plus combinatoires de ce résultat.

Pour tout graphe G , on donne une preuve bijective du lien entre la distribution de l'activité externe des arbres couvrants, utilisé pour définir le polynôme de Tutte, et la distribution du nombre de grains des configurations récurrentes du modèle du tas de sable.

Mots-Clef : Combinatoire, Énumération, Séries Génératrices, Équations fonctionnelles, Bijections, Graphes.

Table des matières

Introduction	1
1 Autour des animaux dirigés	17
1.1 Problèmes combinatoires liés à la percolation dirigée	18
1.1.1 Formulation combinatoire de la percolation	18
1.1.2 Résultats exacts sur les animaux dirigés	22
1.2 Empilements de pièces	27
1.2.1 Empilements généraux	27
1.2.2 Empilements de dominos et animaux dirigés	30
1.3 Conjectures pour le premier moment sur un ruban borné	34
1.3.1 Une différence entre les réseaux triangulaire et carré?	34
1.3.2 Matrice de transfert pour le périmètre de site	36
1.3.3 Conjecture	40
1.4 Modèles de gaz	42
1.4.1 Une énumération de Dhar utilisant un modèle de gaz	43
1.4.2 Interprétation sur les préfixes de chemins de Motzkin de l'équation algébrique pour le périmètre droit	46
1.4.3 Résolution de la percolation pour une sous-famille d'animaux	49
2 Des bijections pour les chemins de Dyck	55
2.1 Une famille d'applications définies par insertion	56
2.1.1 Les bijections "classiques"	56
2.1.2 Un algorithme pour des applications parentes d'une de ces bijections	59
2.1.3 Étude des équivalences et symétries de ces applications	62
2.1.4 Preuve du caractère bijectif de 12 applications	67

2.2	Exemples de bijections définies par insertion	78
2.2.1	Demi-pyramides strictes et chemins de Motzkin	78
2.2.2	Méthode du noyau et développement de Taylor	80
2.2.3	Les bistatistiques de Haiman et Haglund	94
2.2.4	Codage de bijections entre classes combinatoires comptées par la suite de Catalan	105
2.2.5	Bijections non codables	114
3	Modèle de polymères dirigés, interactions dans les chemins de Dyck	119
3.1	Présentation des modèles	120
3.1.1	Modèles de polymères	120
3.1.2	Interactions sur les chemins de Dyck	125
3.1.3	Traitement des interactions inférieures	125
3.2	Interactions supérieures par équation à la Temperley	128
3.2.1	Mise en équation	129
3.2.2	Résolution de l'équation	132
3.3	Interactions supérieures par équation q -algébrique	136
3.3.1	Changement de modèle pour la mise en équation q -algébrique	137
3.3.2	Mise en équation q -algébrique	139
3.3.3	Résolution de l'équation q -algébrique par linéarisation	141
3.4	Interactions supérieures en une page	147
3.4.1	Résolution utilisant une équation fonctionnelle à la Temperley	148
3.4.2	Une seconde linéarisation	154
3.5	Interprétations combinatoires	155
3.5.1	Interprétation de σ et lien avec les interactions inférieures	156
3.5.2	Chemins à creux sur l'axe et empilements de segments	162
4	Sandpile model and Tutte polynomials	175
4.1	Preliminaries	176
4.2	Decreasing traversals	178
4.3	Decreasing traversals and spanning trees	179
4.4	Decreasing traversals and recurrent configurations	181
4.5	Counting spanning subgraphs	182

Perspectives	185
A Résolution d'équations q-linéaires	189
Bibliographie	204

Table des figures

0.1	Une triangularisation d'un polygone régulier	2
0.2	Énumération exhaustive des triangulations de taille inférieure à 5	2
0.3	Énumération exhaustive des arbres planaires de taille inférieure à 3	3
0.4	Transformation d'une triangulation en arbre planaire	4
0.5	Une décomposition "algébrique" des triangulations	7
0.6	Deux codages par des ordres différents	10
1.1	Une configuration à trois amas	18
1.2	Forme schématique de la probabilité de percolation	19
1.3	Un animal d'aire 38 et de périmètre de site 46	20
1.4	Une configuration empêchant la percolation dirigée	21
1.5	Un animal dirigé d'aire 17 et de périmètre de site dirigé 13	21
1.6	Un animal dirigé sur le réseau triangulaire et ses sites périmétriques	23
1.7	Des demi-animaux sur les réseaux triangulaire et carré	24
1.8	Automate lié aux animaux sur ruban borné	25
1.9	Périmètre droit des animaux dirigés sur réseau carré	26
1.10	Un exemple d'empilement	28
1.11	Demi-animaux sur le réseau triangulaire et demi-pyramides de dominos	31
1.12	Demi-animaux sur le réseau carré et demi-pyramides strictes de dominos	31
1.13	Décomposition des empilements triviaux sur P_6	32
1.14	Lien entre empilements et empilements stricts	33
1.15	Décomposition algébrique des demi-pyramides	33
1.16	Factorisation en deux demi-pyramides	34
1.17	Perturbation du périmètre de site sur le réseau triangulaire	35
1.18	Préservation du périmètre de site sur le réseau carré	35
1.19	Périmètre de site des demi-pyramides strictes	36

1.20	Codage par ligne d'un animal sur le réseau carré	37
1.21	Règles de transition des modèles de gaz	43
1.22	Diagramme espace-temps du modèle de gaz	44
1.23	Décomposition des chemins quelconques	47
1.24	Factorisation des chemins quelconques en préfixes de Motzkin	48
1.25	Un animal à source compacte 4 et sa factorisation	49
1.26	Demi-animal gauche de source compacte et animal	50
1.27	Coloration et occupation d'un animal	50
1.28	Déduction à partir du diagramme aléatoire	51
1.29	Triangle de calcul et densité du modèle de gaz	53
1.30	Stratégie pour le jeu de Nim	53
2.1	Un mot de Dyck et son chemin de Dyck associé	57
2.2	Bijections entre chemins de Dyck, suites quasi-décroissantes et demi-pyramides de dominos	58
2.3	Représentation graphique d'un chemin marqué	59
2.4	Constructions de l'algorithme pour la suite 0, 1, 0, 1, 2, 1	60
2.5	Factorisation d'une suite quasi-décroissante	64
2.6	Les douze applications regroupées en trois classes	67
2.7	Fonction de marquage sur la dernière descente	68
2.8	Un exemple de recherche de l'inverse de $\Phi_{(xB\bar{B}, \bar{x}A\bar{B})}$	69
2.9	Fonction de marquage sur l'axe	70
2.10	Fonction de marquage sur l'axe pour le mode $(AB\bar{B}x, \bar{x}A)$	71
2.11	Un exemple de couverture avec des parapluies	73
2.12	Fonction de marquage des parapluies pour le mode $(BA\bar{B}x, \bar{A}\bar{x})$	73
2.13	Les deux types d'insertion dans les factorisations "avec parapluies"	75
2.14	Fonction de marquage dans les pics avant la première double-descente pour le mode $(xA\bar{A}, \bar{x}B\bar{B})$	77
2.15	Fonction de marquage des préfixes croissants pour le mode $(BA\bar{A}x, \bar{x}\bar{B})$	77
2.16	Demi-pyramides strictes et Chemins de Motzkin : un exemple	79
2.17	Décomposition des demi-pyramides strictes et des chemins de Motzkin	80
2.18	Décomposition des chemins de Dyck "catalysée" par la longueur de la dernière descente	81
2.19	Décomposition des chemins à descente marquée d'ordre k	85

2.20	Définitions des rayons et de la hauteur de rayon	86
2.21	Rayons marqués et hauteur de rayon marqué	87
2.22	L'histoire d'un chemin de Dyck	87
2.23	Un exemple de calcul de l'application $\Phi_{\langle(B\bar{B}x, A\bar{A}\bar{x});(A\bar{A}x, B\bar{B}\bar{x})\rangle}$	89
2.24	Règles pour le marquage associé à la hauteur de rayon	91
2.25	Fonction de marquage pour la hauteur de rayon	91
2.26	Insertion pour les rayons (Cas pair)	92
2.27	Insertion pour les rayons (Cas impair)	93
2.28	Aire d'un chemin de Dyck	94
2.29	Couplage des pas montants définissant la statistique de Haiman	95
2.30	Fonction de marquage la statistique de Haiman <i>dinv</i>	96
2.31	Un exemple de recherche de l'inverse de $\Phi_{\sigma\circ\bar{\sigma}\circ(xB\bar{B}, \bar{x}A\bar{A})}$	98
2.32	Moments des ricochets d'un chemin de Dyck	100
2.33	Ricochets sur les pics d'un chemin de Dyck	101
2.34	Une insertion de pas pour le moment des ricochets	102
2.35	Les derniers pas insérés dans un chemin par <i>f_haglund</i>	103
2.36	p et \bar{p} non consécutifs	104
2.37	Une symétrie naturelle peut induire un paramètre analogue à la dernière descente	105
2.38	Chemins de Dyck et polyominos parallélogrammes	106
2.39	Symétrie selon la diagonale principale des polyominos parallélogrammes	107
2.40	La fonction de marquage <i>f_reflexion</i>	107
2.41	Une insertion selon la fonction de marquage <i>f_reflexion</i>	108
2.42	L'involution de Kreweras sur les chemins de Dyck	110
2.43	Insertion de pas pour une involution de Kreweras	111
2.44	Compatibilité du $[\sigma \circ (xB, \bar{x}A\bar{B}\bar{A})](xA, \bar{x}B\bar{B}\bar{A})$ avec <i>f_reflexion</i>	113
2.45	Un exemple d'application artificielle non codable par mélange de pics	115
2.46	Définition de <i>g_cyclique</i>	115
2.47	Le lemme cyclique et <i>g_cyclique</i>	117
3.1	Un chemin contre-exemple aux polymères dirigés	121
3.2	Exemple de polymère dirigé	122
3.3	Définition des paramètres étudiés sur les polymères dirigés	123
3.4	Récapitulation des paramètres étudiés sur les polymères dirigés	124

3.5	Paramètres étudiés sur les chemins de Dyck	126
3.6	Factorisation des chemins de Dyck préservant les interactions inférieures .	127
3.7	Factorisation pour les interactions inférieures et les contacts	127
3.8	Difficulté des interactions supérieures	129
3.9	Décomposition à la Temperley pour les chemins de Dyck	130
3.10	Insertion strictement croissante de pic	130
3.11	Insertion décroissante de pic	131
3.12	Une classe de chemins de Dyck pics-équivalents	138
3.13	Valuation d'un petit creux d'un représentant	139
3.14	Décomposition algébrique des chemins à petits creux	140
3.15	Décomposition pour les chemins de Dyck à petits creux	148
3.16	Développement des sommes $\square + \blacksquare$	156
3.17	Chemins à petits creux débutant par un pic double et σ	157
3.18	Un chemin de Dyck et son chemin à petits pics	159
3.19	La bijection liée au lemme cyclique pour les petits creux ou pics	159
3.20	Les cas dégénérés de la bijection précédente	160
3.21	La factorisation selon σ des chemins à petits creux blancs	161
3.22	Chemin à creux sur l'axe et chemin à petits creux	162
3.23	La valuation des segments marqués	164
3.24	Bijection entre des chemins à petits creux débutant par un pic double et les empilements de segments marqués	165
3.25	Décomposition des empilements triviaux de segments marqués	169
3.26	Décomposition d'un empilement trivial débutant par un segment d'au moins 3 éléments	170
3.27	Décomposition des chemins à creux sur l'axe débutant par un pic double puis éventuellement un pic	172

Introduction

Cette thèse se situe dans le domaine de la combinatoire énumérative. On y étudie des ensembles d'objets discrets dont tous les éléments ont une taille finie. L'objectif est d'obtenir une expression permettant d'évaluer efficacement, en temps polynômial, le nombre d'éléments de taille n , une quantité qui croît en général exponentiellement en fonction de n . Ce problème d'énumération, ou un de ses raffinements prenant en compte des paramètres supplémentaires sur les objets, est motivé par le contexte informatique, physique ou mathématique où est apparu l'ensemble d'objets discrets étudié. Dans le cadre de l'analyse d'algorithme [Knu00], [FS95], l'évaluation de la complexité en moyenne d'un algorithme peut passer par l'observation de toutes ses exécutions sur une entrée de taille n . On énumère alors les entrées selon leur taille et le coût de traitement par l'algorithme. En physique statistique [Bax82] [JvR00b], les modèles proposés par les physiciens sont supposés admettre une distribution Boltzmanienne à l'équilibre et leur étude passe par le calcul de la fonction de partition. Cette fonction de partition se définit par une sommation sur tous les états accessibles du système et nécessite donc souvent une énumération. En mathématiques, les questions de probabilité discrètes, comme la percolation [Hug96], conduisent à des probabilités qui sont des sommes sur un ensemble de cas possibles à recenser. En combinatoire algébrique [AB03] ou en théorie des nombres, le calcul de la dimension d'un espace vectoriel correspond souvent à l'énumération des éléments d'une base.

Un ensemble d'objets discrets \mathcal{A} peut donc apparaître au détour de travaux informatiques, mathématiques ou physiques. La première caractérisation de cet ensemble est une définition et il ne possède pas nécessairement plus de structure. Souvent, une notion de taille $h(e)$ pour un élément $e \in \mathcal{A}$ accompagne la définition. La taille $h(e)$ est un entier naturel. Si la taille h est suffisamment discriminante pour qu'il y ait dans \mathcal{A} un nombre fini d'éléments d'une taille donnée, \mathcal{A} est une classe combinatoire et la suite $(a_n)_{n \in \mathbb{N}}$ du nombre d'éléments de taille n est la suite d'énumération de l'ensemble \mathcal{A} . Les premières valeurs de cette suite peuvent s'obtenir à partir de la définition par le recensement des objets de petite taille, grâce à la force brutale d'un ordinateur ou d'un crayon sur quelques feuilles de papier.

Cette démarche peut être illustrée sur l'ensemble des triangulations de polygones réguliers dont un sommet est distingué. Par définition, une triangulation d'un polygone régulier s'obtient en ajoutant des cordes pour découper en deux des faces qui ne sont pas des triangles. La figure 0.1 donne un exemple de triangularisation d'un polygone, la figure la plus

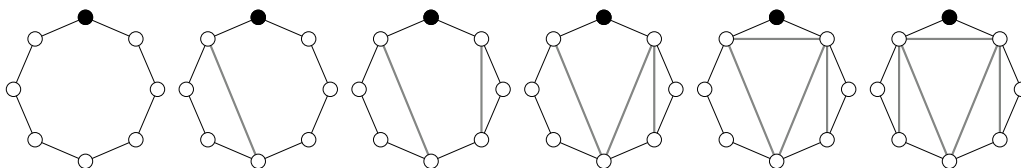


FIG. 0.1 – Une triangularisation d’un polygone régulier

à droite est une triangulation. On définit la taille d’une triangulation par le nombre n de sommets du polygone régulier triangulé. Dans une triangulation de taille n , il y a $n - 3$ cordes à choisir parmi les $n(n-3)/2$ cordes possibles. Il y a au plus $\binom{n(n-3)/2}{n-3}$ triangulations de taille n et donc un nombre fini. La suite d’énumération de ces objets est bien définie. La figure 0.2 contient une énumération exhaustive des triangulations pour les tailles inférieures à 5. Si on ajoute les triangulations de taille 6 et 7, le début de la suite d’énumération de cette ensemble est $0, 0, 0, 1, 2, 5, 14, 42, \dots$

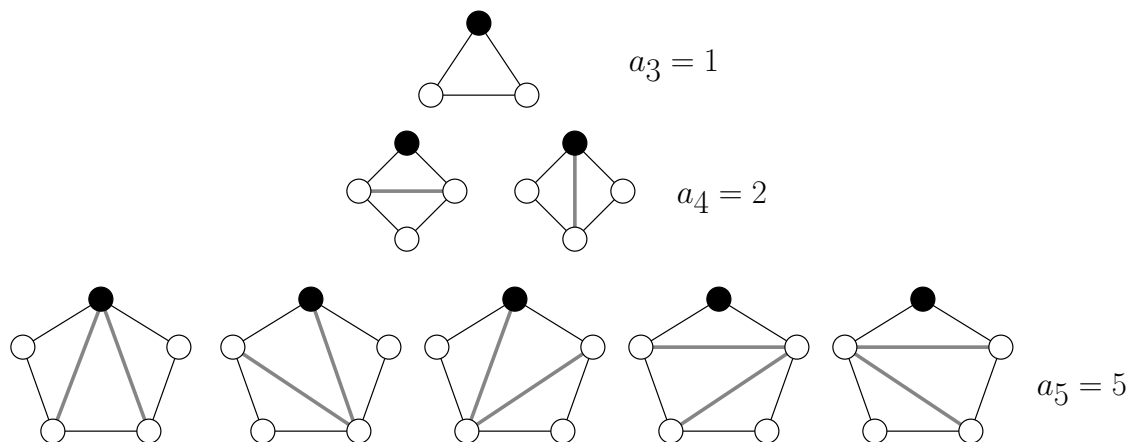


FIG. 0.2 – Énumération exhaustive des triangulations de taille inférieure à 5

Les premiers termes de la suite $(a_n)_{n \in \mathbb{N}}$ forment un “numéro de téléphone” et un annuaire a été construit par Sloane et Plouffe [Slo03]. Une base de données, accessible sur le Web, rassemble “tous” les débuts de suites d’énumérations que des combinatoriciens ont rencontrées. L’existence dans cette base d’un autre ensemble \mathcal{B} , dont la suite d’énumération $(b_n)_{n \in \mathbb{N}}$ coïncide sur les premiers termes avec la suite a_n , suggère l’existence d’une bijection entre les éléments de taille n de \mathcal{A} et \mathcal{B} . La bijection relie les deux ensembles et dans certains cas, un paramètre difficile sur les éléments d’une des classes se traduit par un paramètre plus simple sur l’autre classe.

Dans le cas des polygones réguliers la suite $1, 2, 5, 14, 42$ est identifiée par l’annuaire comme le début de la suite de Catalan $1, 2, 5, 14, 42, 129, \dots$. Cette suite énumère un grand nombre de classes d’objets combinatoires. Stanley recense 66 classes dans [Sta99] et la liste s’allonge de mois en mois sur sa page Web. Parmi elles, il présente les triangulations de polygones réguliers ce qui clôt l’énumération par une référence bibliographique. Les arbres

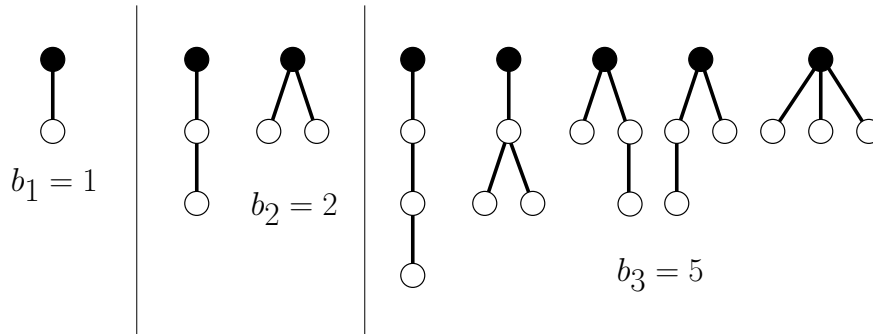


FIG. 0.3 – Énumération exhaustive des arbres planaires de taille inférieure à 3

planaires, dont la taille est le nombre d'arêtes, forment une autre classe combinatoire dont la suite d'énumération est la suite de Catalan. La figure 0.3 liste les arbres planaires de taille inférieure à 3. Il y a autant d'arbres planaires à n arêtes que de triangulations de taille $n+2$ donc il existe une bijection préservant la taille entre ces deux classes. On décrit sommairement une construction bijective d'un arbre planaire à partir d'une triangulation et on l'illustre sur la figure 0.4. On pose une arête en travers de chaque segment du polygone ou de la triangulation pour construire un arbre. On enracine cet arbre comme l'indique la flèche de la figure 0.4 : la racine est un sommet voisin du sommet pointé du polygone. On obtient un arbre planaire en contractant les arêtes entre un père et son fils gauche. Cette bijection permet de prouver que le degré moyen de la racine d'un arbre planaire de n arêtes est $3 - \frac{6}{n+2}$. En effet, sur cette construction on voit que le degré de la racine de l'arbre planaire est aussi le nombre de triangles adjacents au sommet distingué du polygone. Comme l'ensemble des triangulations est invariant par rotation d'angle $2\pi/(n+2)$, ce nombre moyen est aussi celui de n'importe quel sommet d'une triangulation. Or dans une triangulation il y a n triangles donc $3n$ coins, et $n+2$ sommets donc en moyenne un sommet appartient à $3n/(n+2)$ triangles ¹.

Pour une classe combinatoire \mathcal{A} , l'objet de la combinatoire énumérative est de déterminer la suite $(a_n)_{n \in \mathbb{N}}$ pour tout $n \in \mathbb{N}$. Les motivations pour cette question peuvent être directement liée à l'origine du problème. On en a donné un rapide aperçu dans les domaines de l'informatique, de la physique et des mathématiques. Si les motivations pour l'énumération de la classe \mathcal{A} sont variées, la forme de la réponse l'est aussi. On donne trois exemples de réponses acceptables, selon les contextes, à la question de l'énumération de la classe \mathcal{A} :

$$a_n = \frac{1}{2n+1} \binom{2n+1}{n}, \quad (1)$$

$$a_0 = 1 = a_1, a_{n+2} = a_{n+1} + a_n, \quad (2)$$

$$a_n = (n^3 + 2n^2 \log(n)^2) \rho^n + o(n^2 \log(n)^2 \rho^n). \quad (3)$$

¹Cette preuve m'a été communiquée par le bouche-à-oreille bordelais et je n'ai pu en trouver de source officielle

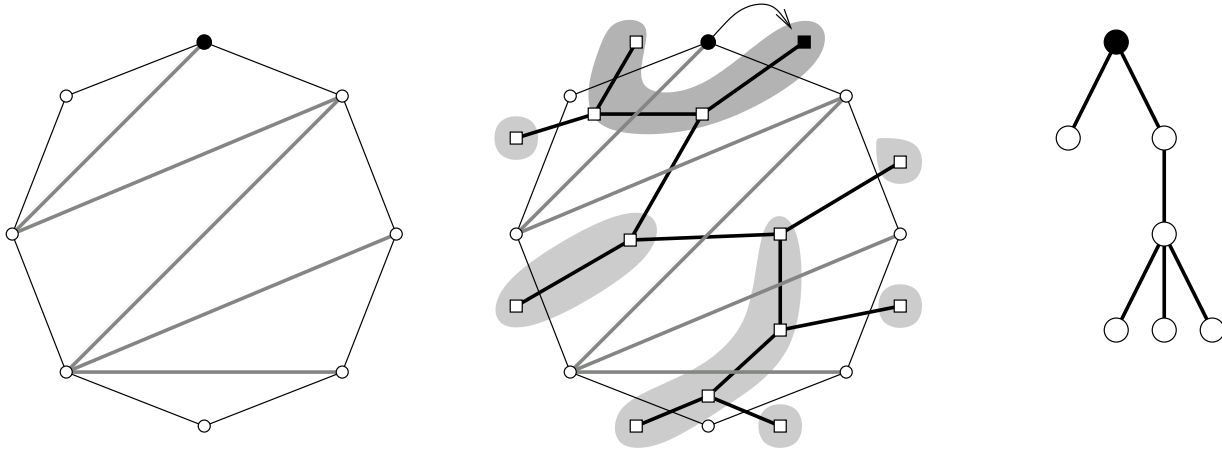


FIG. 0.4 – Transformation d'une triangulation en arbre planaire

Le point commun entre ces trois représentations est de donner une valeur ou une estimation de a_n calculable en un temps polynomial en n bien que souvent a_n dépende exponentiellement de n . Ces représentations de a_n sont donc plus efficaces que l'énumération exhaustive des objets de taille n de la classe \mathcal{A} et montre que quelque chose a été compris. L'équation (1) est une expression close de a_n , l'équation (2) est une récurrence satisfaite par les $(a_n)_{n \in \mathbb{N}}$ et l'équation (3) est une estimation asymptotique de a_n .

Comment obtenir ces résultats d'énumération ? La première approche consiste à trouver une bijection préservant la taille entre les éléments de la classe \mathcal{A} et ceux d'une autre classe \mathcal{B} . On réduit alors l'énumération de la classe \mathcal{A} à celle de la classe \mathcal{B} qui est peut-être déjà connue. La seconde technique consiste à décomposer récursivement les éléments de la classe \mathcal{A} en éléments de cette classe de taille inférieure ou bien d'autres classes que l'on a su énumérer. Ces décompositions se traduisent par des récurrences sur les coefficients $(a_n)_{n \in \mathbb{N}}$ et éventuellement ceux des autres classes. La résolution de ces récurrences mène parfois à une forme close pour les a_n . Une dernière approche ne peut être utilisée que pour déterminer une estimation asymptotique des a_n . Elle consiste à modifier la classe combinatoire étudiée sans perturber cette estimation asymptotique. Cette technique a été utilisée, par exemple, pour l'étude du comportement asymptotique des chemins auto-évitant [MS93].

Séries génératrices

Les séries génératrices sont une représentation des suites d'énumération. La suite d'énumération $(a_n)_{n \in \mathbb{N}}$ peut être codée par différentes séries formelles

$$\sum_{n \geq 0} a_n h_n$$

où $(h_n)_{n \in \mathbb{N}}$ est une échelle, *i.e.* une famille de fonctions linéairement indépendantes. Ces séries formelles sont des séries génératrices de la classe \mathcal{A} . Pour l'échelle $(z^n)_{n \in \mathbb{N}}$, on obtient

la série génératrice ordinaire

$$O(z) = \sum_{n \geq 0} a_n z^n,$$

pour l'échelle $(\frac{z^n}{n!})_{n \in \mathbb{N}}$, la série génératrice exponentielle

$$E(z) = \sum_{n \geq 0} a_n / n! z^n,$$

et pour l'échelle $(1/n^z)_{n \in \mathbb{N}}$, la série de Dirichlet

$$D(z) = \sum_{n \geq 0} \frac{a_n}{n^z}.$$

L'intérêt de ces séries formelles est que pour des suites "simples", on reconnaît le développement de fonctions de l'analyse. Par exemple si pour tout $n \in \mathbb{N}$ il y a dans la classe \mathcal{A} un unique élément de taille n alors $O(z) = \sum_n z^n = \frac{1}{1-z}$, $E(z) = \sum_n z^n / n! = \exp(z)$ et $D(z) = \sum_n 1/n^z = \zeta(z)$. Ces trois fonctions sont une notation plus compacte que la suite $1, 1, 1, 1 \dots$. De plus ces représentations ne se limitent pas à coder les suites d'énumération, il existe des techniques pour extraire de ces séries génératrices les diverses représentations du résultat d'une énumération. Par exemple, les termes a_n sont aussi les coefficients du développement de Taylor en $z = 0$ de la série génératrice exponentielle :

$$a_n = \frac{\partial^n}{\partial z^n} E(z = 0).$$

Les récurrences sur les coefficients a_n peuvent se traduire par des équations fonctionnelles satisfaites par les séries génératrices. Par exemple la récurrence

$$a_0 = 1 = a_1, \quad a_{n+2} = a_{n+1} + a_n$$

implique que la série génératrice ordinaire $O(z) = \sum_{n \geq 0} a_n z^n$ satisfait l'équation

$$O(z) - 1 - z = z(O(z) - 1) + z^2 O(z).$$

Les différents types de séries génératrices se prêtent plus ou moins bien à des récurrences qui sont de natures variées. Par exemple, les séries de Dirichlet s'adaptent bien aux récurrences issues de la théorie des nombres qui utilisent la notion de divisibilité. L'extraction d'une estimation asymptotique comme dans l'équation (3) à partir de la série génératrice peut se faire à l'aide de l'analyse complexe. Flajolet et Sedgewick présentent ces techniques dans [FS04]. Dans ce contexte, la série génératrice formelle est vue comme une fonction analytique en la variable complexe z et la formule de Cauchy est le premier pas vers l'extraction du comportement asymptotique de a_n . La "combinatoire analytique" a permis d'unifier un grand nombre de résultats asymptotiques : le développement est gouverné par le comportement de la série génératrice au voisinage de son cercle de convergence.

Il existe pour les séries génératrices ordinaires un petit dictionnaire traduisant des combinaisons de classes combinatoires en des opérations sur les séries génératrices. L'union disjointe de deux classes combinatoires \mathcal{A} et \mathcal{B} est une classe dont la série génératrice est la somme $A(z) + B(z)$. Le produit cartésien $\mathcal{A} \times \mathcal{B}$ admet comme série génératrice le produit $A(z)B(z)$ à condition que la taille du couple (a, b) soit la somme des tailles de a et b . Si la classe \mathcal{A} ne contient pas d'élément de taille nulle alors la série génératrice de la classe des suites d'éléments de \mathcal{A} est $1/(1 - A(z))$. Les séries génératrices des ensembles ou des cycles d'éléments de \mathcal{A} possèdent également des expressions presque raisonnables en fonction de $A(z)$. Un tel dictionnaire existe aussi pour les séries génératrices exponentielles. Ces dictionnaires sont compilés dans [FS04].

Décompositions combinatoires et nature des séries génératrices

La décomposition des objets d'une classe combinatoire se traduit souvent en récurrences sur les a_n et en équations sur les séries génératrices $A(z)$. Il existe une grande variété de décompositions utilisant plus ou moins d'information sur les objets décomposés. Mais ce sont les formes des équations satisfaites par les séries génératrices qui définissent formellement la nature de ces séries. On présente ici des exemples de décompositions que l'on espère caractéristiques des équations de chaque nature.

On considère la classe \mathcal{R} des mots non-vides sur l'alphabet $\{a, b\}$ qui évitent les facteurs aba et aa . La taille d'un mot est le nombre de lettres qui le compose. On partitionne cette classe en 5 sous-classes, la classe $\mathcal{R}_{<2}$ les mots de longueur strictement inférieure à 2, la classe \mathcal{R}_{aa} , respectivement \mathcal{R}_{ab} , \mathcal{R}_{ba} et \mathcal{R}_{bb} des mots de la classe \mathcal{R} dont les deux premières lettres forment le mot aa , respectivement ab , ba et bb . Pour décomposer chacune de ces sous-classes on considère ses éléments de longueur 1 ou 2 et on décompose les autres mots en isolant leur première lettre. Compte tenu des facteurs évités on obtient les décompositions suivantes :

$$\left\{ \begin{array}{l} \mathcal{R}_{<2} = \{a, b\}, \\ \mathcal{R}_{aa} = \emptyset, \\ \mathcal{R}_{ab} = \{ab\} \cup (\{a\} \times \mathcal{R}_{bb}), \\ \mathcal{R}_{ba} = \{ba\} \cup (\{b\} \times \mathcal{R}_{aa}) \cup (\{b\} \times \mathcal{R}_{ab}), \\ \mathcal{R}_{bb} = \{bb\} \cup (\{b\} \times \mathcal{R}_{ba}) \cup (\{b\} \times \mathcal{R}_{bb}). \end{array} \right.$$

Le produit cartésien des classes sur cette décomposition correspond au produit de concaténation. En terme de séries génératrices, cela donne

$$\left\{ \begin{array}{l} R_{<2}(z) = 2z, \\ R_{aa}(z) = 0, \\ R_{ab}(z) = z^2 + zR_{bb}(z), \\ R_{ba}(z) = z^2 + zR_{aa}(z) + zR_{ab}(z), \\ R_{bb}(z) = z^2 + zR_{ba}(z) + zR_{ab}(z). \end{array} \right.$$

La résolution de ce système linéaire conduit à des solutions sous forme de fractions rationnelles pour les séries génératrices $R_{<2}(z)$, $R_{aa}(z)$, $R_{ab}(z)$, $R_{ba}(z)$ et $R_{bb}(z)$. La série

génératrice $R(z)$ de la classe \mathcal{R} est la somme $R(z)$ de ces séries et on a finalement

$$R(z) = \frac{2z + z^2 + z^3}{1 - z - z^3}. \quad (4)$$

Cette série est *rationnelle* comme toutes les séries $Rat(z)$ qui satisfont une équation non triviale de la forme

$$Q(z)Rat(z) = P(z)$$

où $Q(z)$ et $P(z)$ sont des polynômes en z . Pour écrire cette décomposition, il a suffit de connaître les deux premières lettres de chaque mot de \mathcal{R} . De façon plus générale, une telle décomposition existe pour toutes les classes de mots évitant un nombre fini de facteurs. Si k est la taille du plus long des facteurs éviter, il suffit de discuter sur les $k - 1$ premières lettres des mots pour obtenir un système linéaire admettant donc des solutions rationnelles. Ces séries sont rationnelles car il y a un nombre fini de sous-parties de taille bornée des objets à observer pour les décomposer en une partie finie et un seul reste qui appartient à la classe étudiée.

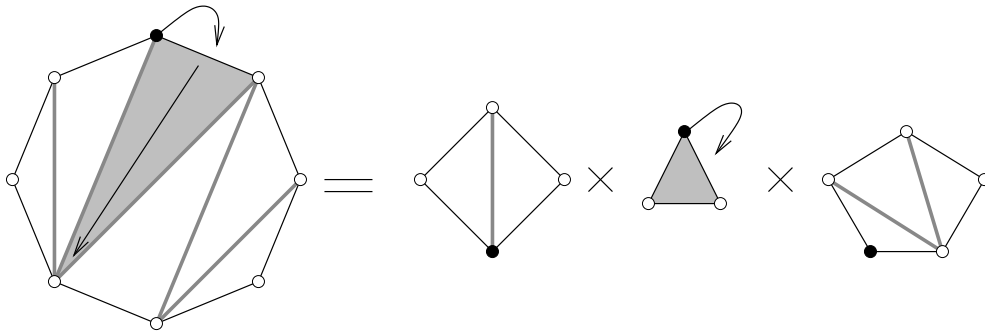


FIG. 0.5 – Une décomposition "algébrique" des triangulations

On note \mathcal{T} la classe combinatoire des triangulations des polygones dont un sommet est pointé, une classe déjà étudiée dans cette introduction. La taille d'une triangulation est maintenant le nombre de triangles. Pour décomposer une triangulation on considère le triangle bordé par le côté du polygone suivant le sommet distingué dans le sens des aiguilles d'une montre. Ce triangle partitionne en deux l'ensemble, éventuellement vide, des autres triangles. Si les ensembles sont non-vides, on reconnaît dans chacun de ces ensembles une triangulation plus petite en distinguant le sommet opposé au côté du triangle distingué. La figure 0.5 illustre un cas non dégénéré de cette décomposition. En terme de classe combinatoire, si on note Δ le triangle et I la triangulation d'un polygone à deux sommets on a

$$\mathcal{T} = (\{I\} \cup \mathcal{T}) \times \{\Delta\} \times (\{I\} \cup \mathcal{T}).$$

La série génératrice des triangulations satisfait donc l'équation

$$T(z) = (1 + T(z)) z (1 + T(z)).$$

Cette série est *algébrique* comme toutes les séries $Alg(z)$ qui satisfont des équations de la forme

$$\sum_{i=0}^k P_i(z) Alg(z)^i = 0$$

où les $P_i(z)$ sont des polynômes en z . En résumé, les décompositions d'éléments en plusieurs éléments plus petits de la même classe conduisent à des équations algébriques.

Un atome dans un élément est une sous-partie de taille 1. Dans la classe \mathcal{T} des triangulations, un triangle est un atome. Dans la classe \mathcal{R} des mots évitant les facteurs aba et aa les atomes sont les lettres. Un élément $a \in \mathcal{A}$ est pointé si un de ses atomes est distingué. Il y a n façons différentes de pointer un élément de taille n donc la série génératrice ordinaire des éléments pointés est

$$\sum_{n \geq 0} n a_n z^n = z \frac{\partial}{\partial z} A(z).$$

Il est possible de pointer plusieurs atomes dans un élément, la série génératrice de tels objets est liée aux dérivées d'ordres supérieurs. Le pointage dans un élément de \mathcal{A} ou bien un de ses sous-éléments dans une décomposition peut permettre de se souvenir de la façon de recombinaison des éléments. Pour illustrer ce propos, on considère la classe \mathcal{P} des permutations. La taille d'une permutation σ sur les entiers $1, 2, \dots, n$ est n . On représente une permutation par la suite de ses images $\sigma(1)\sigma(2)\dots\sigma(n)$. On s'intéresse à la série génératrice *ordinaire* $P(z)$ des permutations, bien qu'elle ne soit pas analytique dans un voisinage de l'origine du plan complexe. On partitionne la classe \mathcal{P} en le singleton contenant la permutation sur l'ensemble vide, la classe $\mathcal{P}_=$ des permutations tels que $\sigma(n) = n$ et la classe $\mathcal{P}_<$ des autres permutations telles que $\sigma(n) < n$. La classe $\mathcal{P}_=$ est en bijection avec la classe des permutations par suppression de n dans la suite des images. On a donc $P_=(z) = zP(z)$. La classe $\mathcal{P}_<$ est en bijection avec la classe des permutations pointées, pour cela on supprime la valeur n de la suite des images et on pointe la valeur suivante dans la suite. On en déduit que $P_<(z) = z \left(z \frac{\partial}{\partial z} P(z) \right)$. La série génératrice ordinaire des permutations satisfait donc l'équation

$$P(z) = 1 + zP(z) + z^2 \frac{\partial}{\partial z} P(z).$$

Cette série est donc *différentiellement finie* comme toutes les séries $D(z)$ satisfaisant une équation de la forme

$$\sum_{i=0}^k P_i(z) \frac{\partial^i}{\partial z^i} D(z) = Q(z)$$

où les $P_i(z)$ et $Q(z)$ sont des polynômes en z . Pour une présentation plus détaillée des séries différentiellement finies on peut se référer à [Sta80].

Les ensembles de séries qui satisfont ces trois types d'équations sont inclus les uns dans les autres : les séries rationnelles sont aussi algébriques et les séries algébriques sont différentiellement finies. De plus ces ensembles sont clos par l'addition et la multiplication.

Le pointage précédent est uniforme sur tous les atomes d'un élément. Si on souhaite ne pointer que dans une sous-partie d'un élément, on introduit un paramètre et donc une variable supplémentaire s , qui compte le nombre de ces atomes "pointables". On décompose maintenant des objets dont on connaît la valeur d'un des paramètres. On obtient une équation dite fonctionnelle de la forme

$$A(z; s) = f(A(z; 1), A(z; s), \frac{\partial}{\partial s} A(z; 1), \dots).$$

Le paramètre supplémentaire ou la taille du sous-élément peuvent aussi être liés à la valuation de l'élément décomposé. Une des formes les plus simples de cette dépendance est d'ajouter un poids q sur chaque site concerné, ce qui peut faire apparaître un terme $A(z; qs)$ ou $A(qz; s)$ dans l'équation. Dans ce contexte on peut définir la notion d'équation q -linéaire, c'est-à-dire de la forme

$$\sum_{i=0}^k P_i(z) A(q^i z) = Q(z)$$

et d'équation q -algébrique comme par exemple celle de la forme

$$\sum_{i=0}^k \left(P_i(z) \prod_{j=0}^i A(q^j z) \right) = Q(z).$$

Lorsque q tend vers 1 on retrouve des équations linéaires ou algébriques. Les solutions de ce type d'équations sont des q -séries, objets pour lesquels il existe une abondante littérature [And86]. La résolution des q -équations linéaires ou algébriques est beaucoup moins systématique que celle des équations linéaires ou algébriques (pour les degrés 2, 3 et 4). Une équation q -linéaire permet certes de calculer de façon efficace les coefficients de la série génératrice mais la recherche d'une formule explicite est plus délicate. Dans [APP98] Abramov, Petkovšek et Paule s'attaquent à la recherche algorithmique de solutions d'une équation q -linéaire quelconque. Ils peuvent conclure en restreignant à la recherche de solutions sous la forme de séries hypergéométriques basiques.

Le principe d'inclusion/exclusion, utilisé par exemple dans les travaux de Gessel et Viennot [GV85], est une autre technique d'énumération qui n'a pas été abordée ici mais qui est fréquente en combinatoire.

Fils directeurs de cette thèse

Le parcours linéaire d'un objet combinatoire

Dans la série génératrice ordinaire $A(z)$ chaque élément de taille n de la classe \mathcal{A} est représenté par un monôme z^n . Deux objets de même taille ont donc le même codage. L'énumération selon la taille peut être raffinée en ajoutant des paramètres qui se traduisent par des variables supplémentaires pour la série génératrice. Par exemple, un objet de taille

n où un premier paramètre vaut k et un second j va être représenté par $z^t u^k v^j$ dans la série génératrice. En général, la valuation d'un élément s'obtient par un produit commutatif des valuations de ses atomes. Il existe souvent un codage de plus bas niveau des objets de taille n en utilisant des mots de n lettres qui sont un produit non-commutatif des valuations des atomes. On construit une bijection entre les objets de la classe \mathcal{A} et les mots en définissant un ordre de parcours des atomes de chaque objet. Ensuite on étudie le langage formé par les mots codant les objets de \mathcal{A} . L'informatique théorique propose un ensemble de formalismes pour décrire ces langages. Les classes combinatoires dont la série génératrice est rationnelle peuvent en général être associées à un langage rationnel, lui-même reconnu par un automate déterministe fini. Les objets combinatoires sont alors en bijection avec un sous-ensemble des chemins dans l'automate et de multiples décompositions peuvent se comprendre sur l'automate. Les classes dont la série génératrice est algébrique sont parfois liées aux langages algébriques, mais ce lien est plus difficile à établir. Parfois une représentation finie de ces langages peut être obtenue sous la forme d'une grammaire algébrique ou d'un automate à pile. Il est parfois possible de court-circuiter le passage par les mots pour obtenir une grammaire algébrique correspondant à une décomposition des objets, c'est l'esprit des grammaires d'objets de Fédou [DFR04]. La méthode ECO, qui consiste à construire les éléments d'une classe en ajoutant un à un ses atomes induit elle aussi un ordre sur les atomes [DFR04]. Les liens entre les grammaires d'objets et la méthode ECO sont étudiés dans [DFR04]. D'autre part il existe des caractérisations ou des exemples de langages qui ne sont pas rationnelles ou algébriques, ce qui permet de comprendre pourquoi un langage n'est pas rationnel ou algébrique.

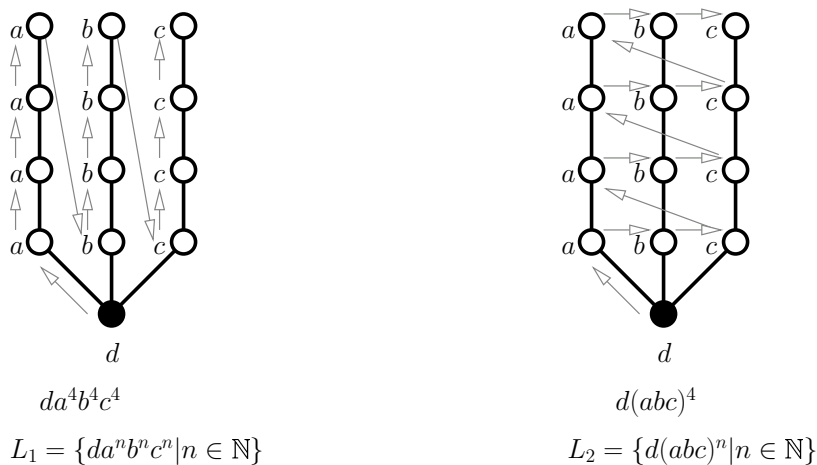


FIG. 0.6 – Deux codages par des ordres différents

Entre la classe combinatoire et le langage, il y a le choix d'un ordre sur les atomes des objets et celui-ci peut déterminer la nature du langage autant que la nature des objets. Par exemple, la classe des arbres formés d'une racine de degré 3 et de trois chemins de longueurs égales, peut se coder par les deux langages $L_1 = \{da^n b^n c^n \mid n \in \mathbb{N}\}$ ou bien $L_2 = \{d(abc)^n \mid n \in \mathbb{N}\}$. Dans les deux cas on value par d la racine et par a , respectivement

b et c , les sommets du premier chemin, respectivement du deuxième et du troisième. Pour le premier langage L_1 on réalise un parcours en profondeur et pour L_2 un parcours en largeur. La figure 0.6 illustre ces deux parcours sur un exemple. Ces deux langages codent bien la même classe mais L_2 est rationnel alors que L_1 n'est pas rationnel (ni même algébrique).

Dans le chapitre 4, on étudie pour tout graphe G , l'égalité des séries génératrices de ses arbres couvrants énumérés selon leur nombre d'arêtes externes actives et des configurations récurrentes énumérées selon leur nombre de grains. Cette égalité était déjà établie grâce la coïncidence de deux équations de décompositions récursives. L'intérêt de la bijection proposée est de définir un ordre de parcours des sommets et des arêtes du graphe qui mette en évidence *simultanément* les arêtes externes d'un arbre couvrant et les grains de la configuration récurrente associée par la bijection. Dans le chapitre 2, les bijections proposées correspondent à différents ordres de parcours des pas d'un mot de Dyck.

Remplacer des calculs par de la combinatoire *ad hoc*

Pour obtenir la série génératrice $R(z)$ des mots évitant les facteurs aba et aa on a utilisé une méthode relativement générale. On a considéré la longueur du plus long facteur évité, ici 3, et on a partitionné les mots de \mathcal{R} selon leurs $3 - 1 = 2$ premières lettres. Pour conclure, un logiciel de calcul formel a résolu un système linéaire de $2 + 2^2 = 6$ équations pour aboutir à l'expression (4). En observant un peu plus les mots de \mathcal{R} on peut éviter la résolution fastidieuse du système. En effet, les mots qui évitent les facteurs aa et aba sont ceux dans lesquels il y a toujours au moins deux b entre deux lettres a . On note $Seq(\mathcal{A})$ les suites d'éléments de la classe générique \mathcal{A} . En discutant sur la position de la première lettre a dans un mot de \mathcal{R} on obtient la décomposition suivante de la classe \mathcal{R} :

$$\begin{aligned} \mathcal{R} &= \{a\} \times Seq(\{b, bba\}) \\ &\cup \{ba\} \times Seq(\{b, bba\}) \\ &\cup \{bba\} \times Seq(\{b, bba\}) \\ &\cup \{b\} \times Seq(\{b, bba\}) \\ &= \{a, ba, bba, b\} \times Seq(\{b, bba\}). \end{aligned}$$

Cette décomposition est moins générale que celle utilisant les deux premières lettres du mot cependant elle est plus proche de la série génératrice puisque le dictionnaire sur les séries ordinaires permet de calculer directement

$$R(z) = (z + z^2 + z^3 + z) \frac{1}{1 - (z + z^3)}.$$

Sur cet exemple, la résolution du système linéaire a été comprise sous la forme d'une factorisation *ad hoc* des mots. Dans cette thèse on aborde parfois des problèmes d'énumération déjà résolus en essayant d'en donner des réponses où certaines étapes de calculs sont remplacées par une décomposition plus précise de la classe étudiée. C'est le fil rouge du chapitre 3, où l'on essaye de comprendre un quotient de séries hypergéométriques basiques dans lesquelles apparaît un terme algébrique. On retrouve cette démarche dans le chapitre 1 dans

lequel on essaye de comprendre des résultats sur les animaux dirigés obtenus par des calculs de densité dans certains modèles de gaz. Dans la partie 2.2.2 on détermine la solution d'un système linéaire à l'aide d'une bijection sur la classe combinatoire interprétant ce système.

Remplacer la combinatoire *ad hoc* par du calcul (sur un processeur)

La combinatoire *ad hoc* permet de mieux comprendre les résultats d'énumération. Mais il y a des situations où l'objectif n'est que de connaître ce résultat. Dans ce cas, une automatisation au prix de calculs supplémentaires est la bienvenue. Par exemple pour les mots évitants des facteurs, il est raisonnable d'envisager un programme qui, prenant en entrée un alphabet et les facteurs évités, calcule la série génératrice des mots évitant ces facteurs. Une première version de cet algorithme pourra construire l'automate sur la partition des mots selon leurs premières lettres et résoudre le système linéaire, éventuellement volumineux, qui s'en déduit. On peut espérer une seconde version plus efficace où une analyse préalable de l'ensemble des facteurs évités permette de trouver des décompositions court-circuitant la résolution d'un système linéaire. Cet exemple illustre la démarche, pour ce problème particulier il existe des résultats généraux, l'exercice 14 du chapitre 4 de [Sta86] propose un panorama sur la question.

Les programmes de la bibliothèque *AlgoLib* pour le logiciel de calcul formel *Maple* sont un exemple concret de l'implémentation de calculs combinatoires. Ces programmes manipulent principalement des séries génératrices. Par exemple, une série rationnelle pour laquelle on s'est donné une borne sur les degrés du numérateur et du dénominateur est caractérisée par le début de son développement et cette recherche peut être automatisée. Dans le chapitre 2, on s'intéresse au codage des bijections sur les chemins de Dyck. On s'intéresse à un sous-ensemble de bijections que l'on espère reconnaissable à partir de leur comportement sur les chemins de petite taille. À ce jour, on a pu prouver le caractère bijectif de certaines applications à partir de leur codage et montrer qu'un certain nombre de bijections de la littérature appartient à ce sous-ensemble. L'espoir est de pouvoir automatiser partiellement le calcul d'un codage d'une telle bijection.

Plan de cette thèse

Elle est composée de quatre chapitres relativement indépendants.

Chapitre 1 : Autour des animaux dirigés

Le problème de la percolation dirigée sur le réseau carré conduit à la question de l'énumération des animaux dirigés selon l'aire et le périmètre de site [BM96]. En combinatoire, les structures arborescentes se factorisent souvent agréablement pour donner des équations algébriques sur les séries génératrices. A priori, les animaux dirigés sont des objets qui

ne sont pas arborescents mais le sont presque. D'ailleurs, l'énumération selon l'aire uniquement utilise une bijection avec des empilements de dominos [BP93] qui met en valeur une structure arborescente cachée grâce à des glissements de dominos qui préservent l'aire mais pas le périmètre de site. Je présente une forme rationnelle d'une conjecture numérique de Conway [Con96] sur l'algébricité de la série génératrice selon l'aire des animaux pointés dans leur périmètre. Les modèles de gaz, introduits par Dhar [Dha83], [Dha82], sont liés à une autre technique pour énumérer les animaux dirigés selon l'aire. Grâce à eux, Bousquet-Mélou [BM98] a obtenu une équation algébrique pour la série génératrice selon l'aire et un sous-ensemble du périmètre : le périmètre droit. Je propose une interprétation combinatoire de cette équation en terme de préfixes de chemins de Motzkin. Enfin, dans l'espoir de mieux la comprendre combinatoirement, je reformule la correspondance entre la densité d'un modèle de gaz et la série génératrice des animaux dirigés selon l'aire en observant le diagramme espace/temps du modèle de gaz. Cette formulation fait apparaître une sous-classe des animaux dirigés : ceux pour lesquels il existe une stratégie gagnante pour le second joueur d'un jeu de Nim disputé sur l'animal.

Chapitre 2 : Des bijections sur les mots de Dyck

Les mots de Dyck sont les mots sur l'alphabet $\{x, \bar{x}\}$ qui contiennent autant de lettres x que \bar{x} et dont aucun préfixe ne contient strictement plus de \bar{x} que de x . La suite d'énumération de cette classe est la suite de Catalan ainsi que pour plus d'une soixantaine d'autres classes recensées par Stanley comme par exemple les suites quasi-décroissantes, définies par $s_1 = 0$ et $s_{n+1} \in \{0, 1, \dots, s_n + 1\}$. Un chemin de Dyck est aussi un mélange de pics, des mots $x\bar{x}$. À l'aide d'un algorithme paramétré par des règles de réécritures inspirées des grammaires algébriques, je défini des appariements des lettres x et \bar{x} et des ordres de lecture de ces couples. Chacun de ces ordres de parcours correspond à une application entre les suites quasi-décroissantes et les chemins de Dyck. Une étude systématique à l'aide de l'ordinateur m'a permis d'isoler 12 applications potentiellement bijectives différentes parmi les 210 règles de réécritures possibles. Je prouve que ces applications sont des bijections. J'en déduis 5 paramètres non-triviaux sur les mots de Dyck, qui ont la même distribution que la longueur de la dernière descente. Certains de ces paramètres étaient déjà connu mais ma démarche unifie leur origine. Je m'intéresse ensuite aux bijections de la littérature qui peuvent se représenter par l'algorithme ou certaines de ces variantes. La correspondance entre les demi-pyramides de dominos et les chemins de Motzkin proposée par Shapiro [Sha], correspond à une symétrie simple sur les règles de réécritures. Pour l'énumération des chemins de Dyck, une variation de la méthode du noyau [BMP00] utilisant le développement de Taylor selon la variable catalytique peut se comprendre bijectivement. Je présente cette bijection et je la code par une généralisation périodique de l'algorithme. L'aire d'un chemin de Dyck est distribuée comme la somme des valeurs des suites quasi-décroissantes. Plus généralement, un analogue de l'aire correspond à chacune des bijections codée par l'algorithme. Des variations de l'algorithme permettent de retrouver des paramètres introduits par Haiman et Haglund [Hag03] dans un contexte de combinatoire algébrique. Le codage

des bijections selon l'algorithme permet parfois d'identifier des égalités entre certaines d'entre elles. Je montre qu'une involution de Kreweras [Lal93] sur les chemins de Dyck est liée à la symétrie par rapport à la première diagonale des polyominos parallélogrammes par une conjugaison par une bijection classique entre ces deux classes combinatoires. Je conclu ce chapitre en présentant une bijection sur les mots de Dyck qui ne peut pas être codée par l'algorithme et ses variantes.

Chapitre 3 : Modèle de polymères dirigés

En énumérant des chemins partiellement dirigés selon leur périmètre de site, Bousquet-Mélou et Rechnitzer [BMR03] ont obtenu une expression ayant la forme d'un quotient de deux séries hypergéométriques basiques semblables dans lesquelles apparaît un terme algébrique. Cette forme de série génératrice n'avait jusque là pas été observée pour un problème combinatoire. J'ai épuré ce problème en travaillant sur les chemins de Dyck énumérés selon leur longueur et le nombre d'*interactions supérieures*. En reprenant les calculs de Bousquet-Mélou et Rechnitzer, qui combinent une mise en équation à la Temperley, la méthode du noyau et une itération, j'ai obtenu la formule close

$$D(t, u) = - \frac{\sum_{n \geq 0} \left(\frac{(q-t)\sigma}{1-q} \right)^n \frac{q^{\binom{n+2}{2}-1}}{(q)_n (qt\sigma^2)_n} t}{\sum_{n \geq 0} \left(\frac{(q-t)\sigma}{1-q} \right)^n \frac{q^{\binom{n+2}{2}-1}}{(q)_n (qt\sigma^2)_n} \frac{q(1-tq^n\sigma)}{(1-q^n\sigma)(1-q^{n+1}\sigma)}} \quad (5)$$

avec $q = ut$ et σ la série entière solution du "noyau", c'est-à-dire ici solution de

$$1 - \left(1 + \frac{q-t}{1-q} \right) \sigma + t\sigma^2 = 0. \quad (6)$$

Une première observation combinatoire m'a permis d'écrire une équation q -algébrique pour ce problème. La résolution de cette équation passe par une linéarisation pour obtenir une équation q -linéaire. En combinant des travaux de Brak et Prellberg [BP95], pour la linéarisation, et Abramov, Paule et Petkovšek [APP98], pour la résolution de l'équation q -linéaire, j'ai retrouvé le résultat précédent. De plus, je propose une nouvelle forme générale de linéarisation qui, sur l'équation traitée, conduit à une solution plus directe.

Ensuite je propose une interprétation combinatoire de la série algébrique σ . Je mets cette interprétation en bijection avec une autre statistique sur les mots de Dyck, définie par Denise et Simion [DS95]. D'autre part, je propose un modèle d'empilement de segments pour énumérer un sous-ensemble des mots de Dyck déjà traité par Janse Van Rensburg [JvR00a] à l'aide d'une équation q -algébrique. Cette technique conduit à un quotient de séries hypergéométriques sans terme algébrique σ . Il reste à unifier ces deux interprétations combinatoires pour voir dans la formule (5) une instance du lemme d'inversion pour un modèle d'empilement.

Chapitre 4 : Tas de sable et polynôme de Tutte

Le modèle du tas de sable est un système dynamique simulant la diffusion de grains de sable posés sur les sommets d'un graphe. Les configurations récurrentes de ce modèle sont celles qui peuvent apparaître infiniment souvent lorsque l'on ajoute périodiquement des grains de sable. Le polynôme de Tutte d'un graphe se définit en considérant la distribution de deux statistiques : les activités externe et interne, sur l'ensemble des arbres couvrants de ce graphe. Depuis des travaux de Dhar et Majumdar [DM92], on sait qu'il y a autant de configurations récurrentes que d'arbres couvrants. Merino-Lopez a montré [ML97], en donnant deux décompositions récursives, que la distribution de l'activité externe est liée à celle du nombre de grains dans les configurations récurrentes. En collaboration avec Robert Cori [CLB03], j'ai construit une bijection explicite entre les arbres couvrants et les configurations récurrentes qui lie l'activité externe d'un arbre couvrant aux nombres de grains de la configuration associée.

Chapitre 1

Autour des animaux dirigés

Les animaux dirigés sont des objets qui apparaissent naturellement dans une interprétation combinatoire d'un problème de probabilité discrète : la percolation de site dirigée sur un réseau carré ou triangulaire. On présente cette interprétation dans la partie 1.1 en s'inspirant de [BM96] et [Hug96].

Les deux paramètres à énumérer simultanément dans ce contexte sont l'aire et le périmètre de site. L'énumération selon l'aire uniquement est bien comprise à l'aide d'une structure arborescente sous-jacente des animaux dirigés qui apparaît à travers le formalisme des empilements de pièces [BP93], [Vie86], [BM96] présenté dans la partie 1.2.

La structure arborescente apparaissant grâce à des déplacements de pièces ne préservant pas le périmètre de site, il reste encore une difficulté avant la résolution du problème initial de percolation. Il existe cependant une conjecture de Conway [Con96] sur l'algébricité de la série génératrice des animaux dirigés sur le réseau carré. On donne dans la partie 1.3 un raffinement de cette conjecture qui fait intervenir des polynômes apparaissant dans les formules d'énumération des empilements.

Le calcul de la densité de certains modèles de gaz est la première méthode [Dha82], [Dha83] qui a permis d'énumérer les animaux dirigés. On présente cette approche dans la partie 1.4. En la perfectionnant, Bousquet-Mélou [BM98] a obtenu la série génératrice des animaux du réseau carré selon l'aire et un sous-ensemble du périmètre de site : le périmètre droit. Cependant on ne dispose pas d'interprétation simple de l'équation algébrique bivariée satisfaite par cette série. Dans cette même partie on présente une interprétation combinatoire de cette équation en terme de chemins : les préfixes de Motzkin. Enfin on présente une autre interprétation du lien entre la densité d'un modèle de gaz et la série génératrice des animaux dirigés ; elle est basée sur l'observation du diagramme espace/temps du modèle de gaz. Cela fait apparaître une sous-classe des animaux dirigés pour laquelle on sait calculer la série génératrice liée à la percolation dirigée.

1.1 Problèmes combinatoires liés à la percolation dirigée

1.1.1 Formulation combinatoire de la percolation

La percolation est un problème probabiliste élémentaire qui est aussi étudié en physique statistique et en combinatoire [Hug96],[SA92], [BM96]. Ce problème peut se définir pour un grand nombre de graphes infinis, on le formule ici pour le réseau carré. On se place sur la grille \mathbb{Z}^2 . Les sommets de cette grille peuvent être *ouverts* avec une probabilité p ou *fermés* avec une probabilité $q = 1 - p$, indépendamment les uns des autres. Le réel $p \in [0, 1]$ est un paramètre de ce problème. Un *amas* est un ensemble connexe maximal de sommets ouverts. Par *connexe*, on entend que deux sommets quelconques de l'amas peuvent être reliés l'un à l'autre par un chemin de \mathbb{Z}^2 , formé de pas nord, sud, est et ouest et dont les sommets appartiennent à l'amas. De plus l'ensemble est *maximal* au sens de l'inclusion. Par exemple, sur la figure 1.1, on a une configuration avec trois amas.

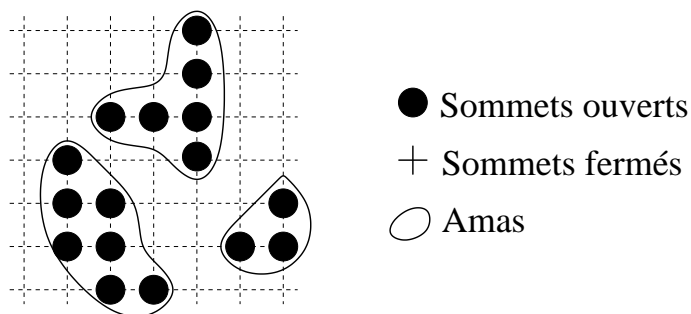


FIG. 1.1 – Une configuration à trois amas

On dit qu'il y a percolation si l'origine $(0, 0)$ du réseau appartient à un amas de cardinal infini. La probabilité de percolation $P(p)$ est une fonction de p , croissante, telle que $P(0) = 0$ et $P(1) = 1$. Aussi simple que puisse paraître ce modèle, on ne sait pas calculer cette fonction, et même sa continuité n'est pas très facile à établir. On s'attend à ce que son graphe ressemble à celui de la figure 1.2, et en particulier, à ce qu'il n'y ait pas percolation pour de petites valeurs de p .

La quantité

$$p_c = \inf\{p | P(p) > 0\}$$

est appelée *seuil de percolation* et est estimée autour de 0.59. À défaut de déterminer complètement la probabilité de percolation $P(p)$, on s'intéresse à la valeur de p_c , et également à un *exposant critique* régissant le comportement de $P(p)$ au voisinage de p_c . On conjecture en effet que $P(p) \sim (p - p_c)^\beta$, à une constante multiplicative près, lorsque $p \rightarrow p_c^+$. La valeur conjecturale de l'exposant β est $5/36$.

Il existe d'autres façons de définir un seuil, par exemple en considérant la probabilité qu'il existe un amas infini, plutôt que la probabilité de percolation (qui est par définition

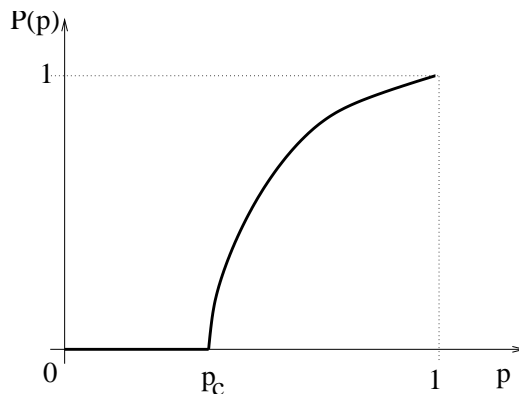


FIG. 1.2 – Forme schématique de la probabilité de percolation

inférieure). On s'attend à ce que ces deux définitions de p_c coïncident. On s'intéresse aussi à d'autres quantités, comme à la taille moyenne des amas finis, et au comportement de cette taille au voisinage de p_c . Enfin, on peut considérer d'autres réseaux, planaires ou non, ou étudier la percolation *de lien* plutôt que *de site*, dans laquelle les arêtes du réseau, plutôt que ses sommets, ont une nature aléatoire déterminée par la probabilité p .

Le problème probabiliste de percolation est un cas particulier d'un problème de combinatoire : l'énumération des animaux selon l'aire et le périmètre de site. Un *animal* (de site) sur un réseau carré est un ensemble fini et connexe de sommets contenant l'origine $(0, 0)$. Son *aire* est son cardinal. Un sommet non contenu dans l'animal, mais adjacent à l'un de ses points, est un *site périmétrique* ; le nombre de sites périmétriques de l'animal est son *périmètre de site*. Par convention, l'ensemble vide est un animal dont l'origine est le seul site périmétrique. La figure 1.3 contient un exemple d'animal. Si $a(n, k)$ est le nombre d'animaux d'aire n et de périmètre de site k , on définit la série génératrice

$$A(t, r) = \sum_{n, k \geq 0} a(n, k) t^n r^k$$

correspondant à l'énumération des animaux selon l'aire et le périmètre de site.

Le lien entre la percolation et l'énumération des animaux s'obtient en considérant la probabilité de non-percolation. Dans ce cas, l'origine appartient à un amas fini par exemple d'aire n et de périmètre de site k . La probabilité de cet amas est celle que les n sites de l'amas soient ouverts, un événement de probabilité p^n , et que les k sites voisins de l'amas soient fermés, un événement de probabilité $(1 - p)^k$. Ces deux probabilités étant indépendantes, la probabilité de l'amas est $p^n(1 - p)^k$. Les occurrences de chaque amas étant disjointes, on obtient la probabilité de non-percolation en sommant sur tous les amas finis :

$$1 - P(p) = \sum_{n, k \geq 0} a_{n, k} p^n (1 - p)^k = A(p, 1 - p).$$

Le calcul de $P(p)$ se réduit donc au calcul de $A(t, r)$ qui est un problème *a priori* plus difficile. La série $A(t, r)$ contient des informations supplémentaires sur la question de la

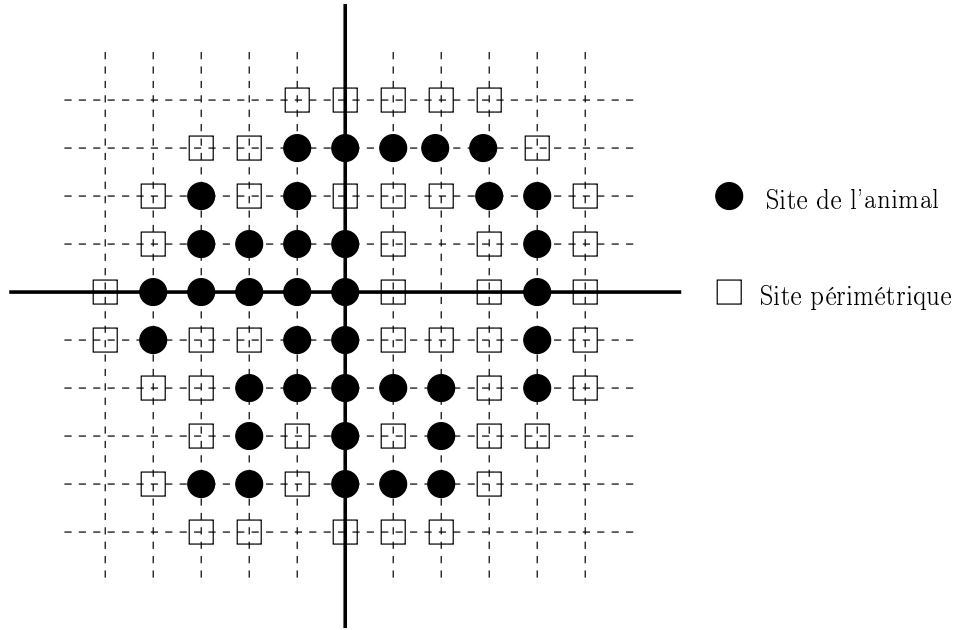


FIG. 1.3 – Un animal d'aire 38 et de périmètre de site 46

percolation. Par exemple, en cas de non-percolation, la taille moyenne des amas se calcule par la formule

$$\sum_{k,n \leq 0} na_{n,k} p^n (1-p)^k = p \frac{\partial}{\partial t} A(p, 1-p).$$

Cependant, le calcul exact de $A(t, r)$ et même celui de $A(t, 1)$ paraissent aujourd'hui totalement hors de portée. Il est cependant permis d'être plus optimiste si on se restreint à l'étude d'une sous-classe d'animaux, dit *animaux dirigés*, naturellement reliée aux modèles de percolation dirigée. On reprend la définition du problème de percolation de site sur réseau carré pour en donner la version dirigée. Les sommets sont toujours ouverts avec la probabilité p , mais la définition de la percolation est modifiée : on dit qu'il y a percolation (dirigée) s'il existe un chemin infini du réseau, issu de l'origine et *seulement formé de pas nord et est*, ne visitant que des sommets ouverts. Ainsi, le phénomène de percolation ne dépend que de l'état du premier quadrant. Dans la configuration de la figure 1.4, il n'y a pas percolation (dirigée) mais rien n'exclut qu'il y ait percolation isotrope.

Les questions étudiées sont les mêmes que dans le cas de la percolation non dirigée. Les modèles dirigés sont naturellement liés à l'énumération des animaux dirigés. Un *animal dirigé* est un animal tel qu'en tout sommet se termine un chemin issu de l'origine et formé uniquement de pas nord et est. L'aire d'un animal dirigé est à nouveau le cardinal de l'ensemble. Par contre, il y a une version dirigée du périmètre de site : les sites périmétriques d'un animal dirigé sont uniquement ceux qui sont accessibles depuis l'animal par un pas nord ou est. Le *périmètre de site dirigé* est le nombre de sites périmétriques (dirigés). Par convention, l'ensemble vide est un animal dirigé et l'origine est son unique site périmétrique

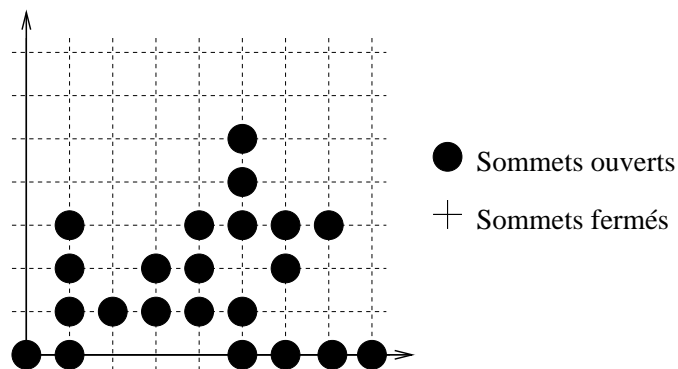


FIG. 1.4 – Une configuration empêchant la percolation dirigée

dirigé. La figure 1.5 contient un animal dirigé. De nouveau, la probabilité de non-percolation s'exprime en fonction de la série génératrice double :

$$D(t, r) = \sum_{n, k \geq 0} d(n, k) t^n r^k$$

où $d(n, k)$ est le nombre d'animaux dirigés d'aire n et de périmètre de site dirigé k .

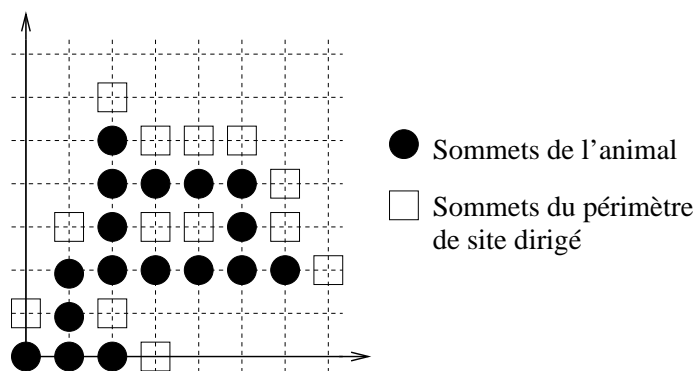


FIG. 1.5 – Un animal dirigé d'aire 17 et de périmètre de site dirigé 13

Dans ce cas, la probabilité critique p_c a été estimée numériquement à 0,705 (*cf.* [Hug96] page 186). Du point de vue de l'énumération exacte, on dispose de résultats et conjectures qui vont faire l'objet de la prochaine partie.

1.1.2 Résultats exacts sur les animaux dirigés

La série $A(t, 1)$ des animaux semble inaccessible mais dans le cas des animaux dirigés on connaît la série génératrice

$$D(t, 1) = \frac{1}{2} \left(\sqrt{\frac{1+t}{1-3t}} + 1 \right).$$

Ce résultat a d'abord été conjecturé par Dhar, Phani et Barma [DPB82] en observant les premiers termes de la série génératrice, des termes calculés grâce à l'ordinateur. Ensuite Dhar [Dha82] a mis en relation la série génératrice $-D(-t; 1)$ avec la densité d'un modèle de gaz de Baxter [Bax82] : les particules dures sur le réseau carré. On ne sait dire des choses sur ce modèle que pour certaines valeurs des paramètres apparaissant dans sa définition. En utilisant des travaux de Verhagen¹, Dhar en déduit l'expression de $D(t, 1)$. Dhar [Dha83] a redémontré ce résultat en utilisant un autre modèle de gaz, employé lors de simulations dans des algorithmes de type Monte Carlo. À nouveau, il établit une égalité entre $-D(-t, 1)$ et la densité du modèle de gaz considéré. Ce modèle de gaz est de dimension 1 en espace et se définit par une transition probabiliste entre l'état à la date t et celui à la date $t + 1$. La distribution des états invariante par cette transition est boltzmanienne. Ces deux éléments permettent de calculer plus simplement la densité. Gouyou-Beauchamps et Viennot [GBV88] ont prouvé ce résultat grâce une bijection "compliquée" avec les préfixes de chemins de Motzkin. Bétréma et Penaud [BP93] ont donné une preuve plus simple grâce à une interprétation des animaux dirigés sous forme d'empilement de pièces. Ce résultat semble donc bien compris et la forme de la solution est remarquablement simple.

On présente maintenant des variantes et/ou des sous-classes des animaux dirigés sur réseau carré. Ces variantes sont motivées par des raisons d'ordre combinatoire que l'on présentera après les définitions.

La première variante consiste à remplacer le réseau carré par le réseau triangulaire. Dans ce réseau dirigé, un site (i, j) peut accéder à ses trois sites voisins par un pas nord pour $(i, j + 1)$, nord-est pour $(i + 1, j + 1)$ et est pour $(i + 1, j)$. Un animal dirigé du réseau triangulaire est tel que pour tout sommet s de l'animal il existe un chemin issu de l'origine, formé de pas nord, nord-est ou est qui se termine en s . L'aire est toujours le cardinal de l'animal. Le périmètre de site est formé des sites accessibles depuis l'animal par des pas nord, nord-est et est. La figure 1.6 donne un exemple d'animal dirigé sur le réseau triangulaire.

La série génératrice bivariée pour ce modèle est

$$E(t, r) = \sum_{n, k} e(n, k) t^n r^k$$

où $e(n, k)$ est le nombre d'animaux dirigés d'aire n et de périmètre de site k . La série

¹Que je n'ai pas réussi à trouver

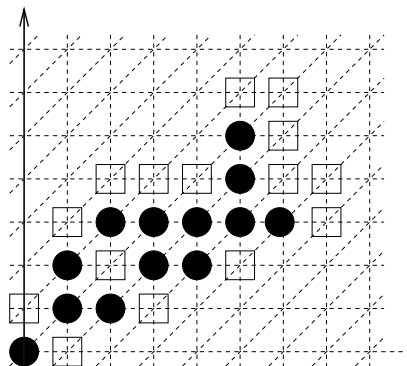


FIG. 1.6 – Un animal dirigé sur le réseau triangulaire et ses sites périmétriques

génératrice selon l'aire est également algébrique et est plus simple que $D(t, 1)$:

$$E(t, 1) = \frac{1}{2} \left(1 + \frac{1}{\sqrt{1-4t}} \right) = 1 + \frac{1}{2} \sum_{n \geq 1} \binom{2n}{n} t^n.$$

Il existe une correspondance simple, exposée plus loin, entre les animaux dirigés sur réseau carré et triangulaire qui explique l'égalité

$$E(t, 1) = D\left(\frac{t}{1-t}, 1\right).$$

La simplicité de la formule d'énumération des animaux selon l'aire sur ces réseaux contraste avec l'énumération des animaux sur le réseau hexagonal qui n'est pas connue [ONGJ01].

Pour définir une sous-classe des animaux dirigés sur le réseau triangulaire ou carré on contraint les animaux à ne contenir aucun site au dessus (strictement) de la diagonale principale définie par les sites de la forme (i, i) . De tels animaux sont nommés des *demi-animaux*. La figure 1.7 contient un exemple de demi-animal. Les sites périmétriques peuvent être au dessus de la diagonale. On note $E_{1/2}(t, r)$ la série génératrice pour les demi-animaux sur réseau triangulaire et $D_{1/2}(t, r)$ celle des demi-animaux sur réseau carré. Ces deux séries génératrices selon l'aire de ces demi-animaux sont particulièrement fréquentes en combinatoire : les coefficients de

$$E_{1/2}(t, 1) = \frac{1 - \sqrt{1-4t}}{2t} = \sum_{n \geq 0} \frac{1}{2n+1} \binom{2n+1}{n} t^n$$

forment la suite des nombres de Catalan et les coefficients de

$$D_{1/2}(t, 1) = \frac{1 - t - 2t^2 - \sqrt{1 - 2t - 3t^2}}{2t^3}$$

sont, à un décalage d'indice près, la suite des nombres de Motzkin. Ces deux suites de nombres comptent aussi des chemins dans le plan. Les nombres de Catalan comptent

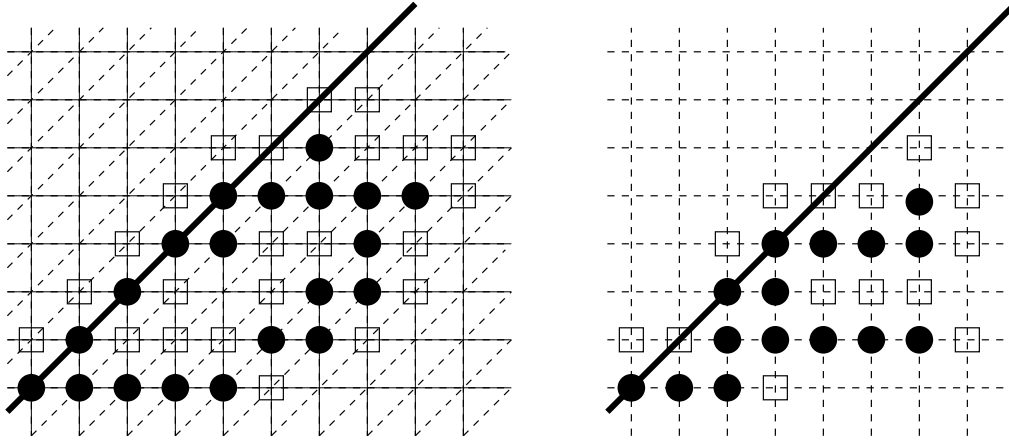


FIG. 1.7 – Des demi-animaux sur les réseaux triangulaire et carré

les chemins de Dyck. Bétréma et Penaud [BP93] ont donné une bijection simple entre ces chemins et les demi-animaux dirigés du réseau triangulaire. Les nombres de Motzkin comptent les chemins de Motzkin, et Shapiro [Sha] a construit une bijection avec les demi-animaux dirigés du réseau carré.

Jusqu'à présent les séries génératrices connues sont toutes algébriques sans être rationnelles. Les séries génératrices rationnelles étant souvent liées à des objets plus "simples" à décomposer, il est intéressant d'obtenir des sous-classes des animaux dont la série génératrice est rationnelle. Une façon naturelle de définir de telles sous-classes est de contraindre les sites des animaux à être compris dans un ruban de largeur borné du réseau défini par deux diagonales. La *largeur* d'un animal a est définie par la quantité

$$1 + \max\{((i_r - j_r) - (i_l - j_l)), (i_l, j_l) \in a, (i_r, j_r) \in a\}.$$

C'est, à une constante près, la distance entre les deux diagonales tangentes à l'animal des deux côtés. Les demi-animaux de largeur au plus n sont les animaux inclus dans le ruban borné défini par les deux diagonales contenant les sites (i, i) et $(i, i - n - 1)$. La série génératrice de ces objets est rationnelle. Sur la figure 1.8, on illustre le raisonnement suivant. Les animaux sur un ruban borné sont en bijection avec des chemins dans un automate dont le nombre d'états est fini. Pour s'en convaincre sur le réseau carré, on peut observer les états de cet automate en considérant l'occupation par l'animal des sites des contre-diagonales $(i, k - i)$. Il y a environ $2^{\lfloor n/2 \rfloor}$ configurations possibles d'une contre-diagonale. Les configurations possibles de la contre-diagonale $(i, k + 1 - i)$ ne dépendent que de la configuration de la contre-diagonale $(i, k - i)$. La série génératrice des chemins dans un automate est rationnelle et donc celle des animaux sur ruban borné l'est aussi. Pour les demi-animaux sur le réseau triangulaire les fractions rationnelles intervenant dans les séries génératrices s'expriment à l'aide des polynômes de Fibonacci. Ces polynômes $(T_n(t))_{n \geq 0}$ sont

$$T_{-1}(t) = 1, \quad T_0(t) = 1$$

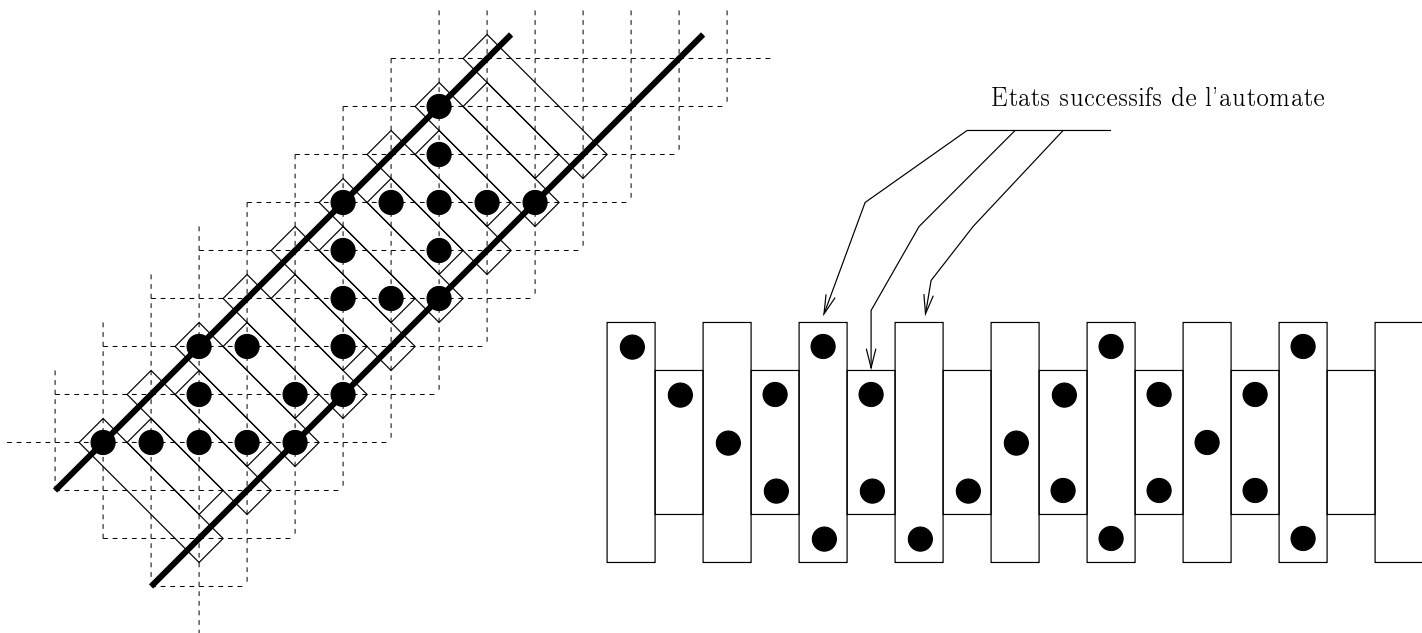


FIG. 1.8 – Automate lié aux animaux sur ruban borné

et pour tout $n \geq 1$,

$$T_{n+1}(t) = T_n(t) - tT_{n-1}(t).$$

En $t = -1$, on retrouve la suite de Fibonacci. La série génératrice selon l'aire des demi-animaux sur réseau triangulaire de largeur au plus n est

$$E_{1/2}^{(n)}(t, 1) = \frac{T_{n-1}(t)}{T_n(t)}.$$

Si les séries selon l'aire de ces divers objets sont bien comprises, le périmètre de site n'a pas pu être traité. Il y a tout de même deux approches qui ont donné des résultats. En généralisant le travail de Dhar utilisant les modèles de gaz, Bousquet-Mélou a pu énumérer les animaux dirigés du réseau carré selon l'aire et le périmètre droit, un ensemble inclus dans le périmètre. Pour un animal a , un site périmétrique (i, j) est un *site périmétrique droit* si $(i - 1, j) \in a$. La figure 1.9 donne un exemple d'animal et de son périmètre droit. On peut noter au passage que pour mettre en valeur la latéralité on a fait tourner le réseau d'un huitième de tour. La série génératrice des animaux dirigés non-vides sur le réseau carré selon l'aire et le périmètre droit est

$$D_{droit}(t, u) = \sum_{n, k} d_{droit}(n, k) t^n u^k$$

où $d_{droit}(n, k)$ est le nombre d'animaux dirigés d'aire n et de périmètre droit k . Cette série est algébrique

$$D_{droit}(t, u) = \frac{1}{2} \left(\left(1 - \frac{4t}{(1+t)(1+t-ut)} \right)^{-1/2} - 1 \right).$$

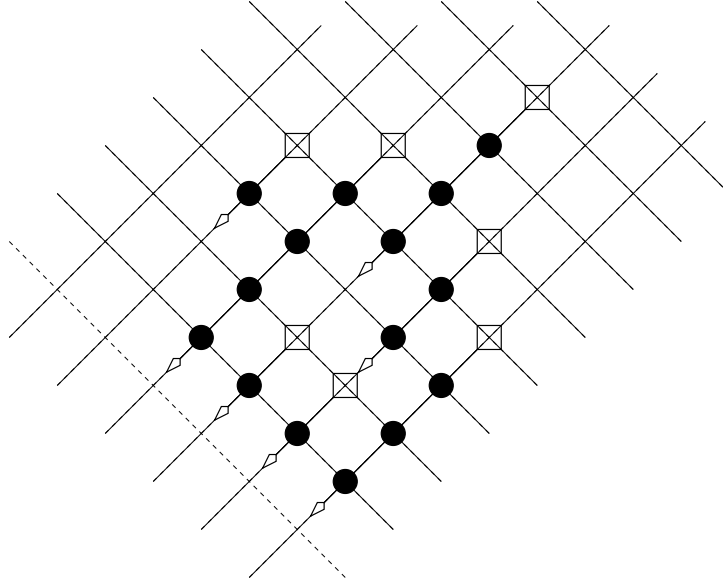


FIG. 1.9 – Périmètre droit des animaux dirigés sur réseau carré

De façon équivalente, $D_{\text{droit}}(t, u)$ est la seule série entière solution de l'équation algébrique

$$(1 - t(2 + u) + t^2(1 - u)) (D_{\text{droit}}(t, u) + D_{\text{droit}}(t, u)^2) = t.$$

Il y a un analogue de cette formule pour le réseau triangulaire au contraire de celle traitant du périmètre de site. La technique employée par Bousquet-Mélou n'utilise pas directement le périmètre droit, elle emploie le nombre de sites (i, j) de l'animal a tels que $(i - 1, j) \notin a$, c'est-à-dire de sites uniquement portés sur la droite. Sur le réseau carré, il existe une bijection entre ces sites et les sites périmétriques droits qui est illustrée par les flèches sur la figure 1.9.

La seconde piste vers le périmètre de site ne mène pas à une solution exacte mais à la conjecture d'une solution partielle exacte. L'objectif est de déterminer la série $D(t, r)$. Une façon d'aborder cette question est de considérer le développement de Taylor en $r = 1$ de cette série :

$$D(t, r) = \sum_{i \geq 0} \frac{(r - 1)^i}{i!} \frac{\partial^i}{\partial r^i} D(t, 1).$$

La connaissance de $D(t, r)$ équivaut à la connaissance de tous les moments $\left(\frac{\partial^i}{\partial r^i} D(t, 1) \right)_{i \in \mathbb{N}}$. Les animaux dirigés dont i sites périmétriques distincts sont étiquetés par les entiers de 1 à i forment une classe combinatoire dont la série génératrice est

$$\sum_{n, k} k(k - 1) \dots (k - (i - 1)) d(n, k) t^n = \frac{\partial^i}{\partial r^i} D(t, 1).$$

Pour $i = 0$, on obtient la série $D(t, 1)$. L'étape suivante serait d'obtenir une expression de $\frac{\partial}{\partial r} D(t, 1)$. En énumérant à l'aide de l'ordinateur les animaux dirigés d'aire inférieure à n ,

Conway [Con96] a déterminé le début du développement de cette série. En observant ce développement, il a conjecturé que ce premier moment satisfait l'équation algébrique

$$t(t+1) \left(\frac{\partial}{\partial r} D(t, 1) \right)^2 + (t^2+t+1) \frac{\partial}{\partial r} D(t, 1) + \frac{t(2-6t-5t^2+12t^3+13t^4+12t^5+9t^6)}{(t+1)^2(3t-1)^3} = 0. \quad (7)$$

La résolution de cette équation conduit à l'expression

$$\frac{\partial}{\partial r} D(t, 1) = \frac{9t^5 + 12t^4 + 7t^3 - t^2 - 4t + 1 - (1 - 3t + 2t^2 + t^3 - 3t^4) \sqrt{(t+1)(1-3t)}}{2t(t+1)^2(3t-1)^2} \quad (8)$$

La suite de ce chapitre discute successivement des deux principales techniques utilisées pour énumérer exactement les animaux dirigés : la bijection de Bétréma et Penaud vers les empilements de pièces de Viennot puis les modèles de gaz de Dhar et Bousquet-Mélou. Pour chacune de ces approches, on présente les travaux réalisés au cours de cette thèse pour essayer d'aller plus loin que la série $D(t, 1)$.

1.2 Empilements de pièces

1.2.1 Empilements généraux

Un *modèle d'empilement* se définit par un ensemble P et une relation binaire $C \subseteq P \times P$ réflexive et symétrique. L'ensemble P est *l'ensemble des types de pièces*. La *concurrence des types de pièces* est codée par la relation C : p_1 et p_2 sont en concurrence si $p_1 C p_2$.

Par exemple, soit $G = (V, E)$ un graphe fini non-orienté dont V est l'ensemble des sommets et E l'ensemble des arêtes. On peut définir un modèle d'empilement sur ce graphe en considérant E comme l'ensemble des types de pièces. Deux arêtes (uv) et (wx) sont en concurrence, $(uv)C(wx)$, si elles ont au moins un sommet en commun.

Un empilement de pièces est le résultat de la chute ordonnée d'une suite de pièces appartenant à P . Lorsque deux pièces de la suite sont en concurrence l'une d'elle se place au-dessus de l'autre dans l'empilement. Par exemple sur la figure 1.10, on considère le graphe $G(V, E)$ où $V = \{0, 1, 2, 3, 4\}$, $E = \{(01), (12), (23), (34)\}$ et l'empilement résultant de la chute dans cet ordre des pièces (01) , (23) , (12) , (23) et (34) . Un empilement peut être le résultat de la chute de différentes suites de pièces. Sur l'exemple la suite (23) , (01) , (12) , (23) et (34) donne le même empilement. Pour représenter les empilements on veut utiliser un codage pour lequel il n'existe qu'une suite par empilement. Une telle forme canonique a été proposée par Cartier et Foata [CF69]. Cela motive l'introduction du niveau d'une pièce dans un empilement. Une *pièce* dans un empilement est un couple (p, h) où $p \in P$ et $h \in \mathbb{N}$. Le type de la pièce est p , une arête sur l'exemple, et h est le *niveau* de cette pièce. Un *empilement* est un ensemble de pièces tel qu'aucune pièce de même niveau ne soit en concurrence et pour toute pièce (p, h) , $h > 0$, il existe une pièce $(p', h - 1)$ telle

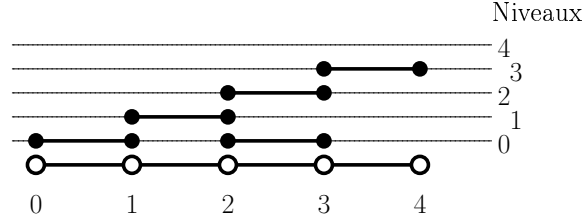


FIG. 1.10 – Un exemple d'empilement

que p et p' soient en concurrence. Une pièce (p, h) est *maximale* dans un empilement e s'il n'existe pas dans e de pièce (p', h') en concurrence tout en étant au-dessus ($h' > h$). Une pièce est *minimale* dans un empilement si son niveau est nul. Sur la figure 1.10 seule la pièce $((34), 3)$ est maximale et les deux pièces minimales sont $((01), 0)$ et $((23), 0)$.

On note \mathcal{H} (h pour heaps), l'ensemble des empilements. Soit \mathbb{K} un corps, $\mathbb{K}[t]$ est l'ensemble des polynômes en t à coefficients dans \mathbb{K} et $t\mathbb{K}[t]$ dénote l'ensemble des polynômes de la forme $tR(t)$ avec $R(t) \in \mathbb{K}[t]$. Une *valuation* v est une application de P dans $t\mathbb{K}[t]$. Si P est fini, on définit la série génératrice $H_v(t)$ des empilements valués par v par

$$H_v(t) = \sum_{e \in \mathcal{H}} \prod_{(p_i, h_i) \in e} v(p_i) = \sum_{e \in \mathcal{H}} v(e)$$

où pour un empilement e le produit porte sur l'ensemble de ses pièces. Par exemple, pour la valuation $v(p) = t$, on trouve la série génératrice usuelle des empilements énumérés selon leur nombre de pièces.

Un empilement est *trivial* si ses pièces ne sont pas en concurrence ; elles sont donc toutes de hauteur nulle. On note \mathcal{T} l'ensemble des empilements triviaux. Pour la valuation v , la série génératrice *alternée* des empilements triviaux est

$$T_{-v}(t) = \sum_{f \in \mathcal{T}} \prod_{(p_i, h_i) \in f} (-v(p_i)) = \sum_{f \in \mathcal{T}} (-1)^{|f|} v(f).$$

Viennot [Vie86] a montré que le calcul de $H_v(t)$ se réduit au calcul de $T_{-v}(t)$.

Il a démontré le lemme suivant qui permet d'énumérer des empilements dont les pièces minimales ou maximales sont de certains types. On présente ici l'étude des empilements dont toutes les pièces minimales sont de types appartenant à un certain sous-ensemble B de P . Pour compter tous les empilements il suffit de choisir l'ensemble des types de pièces pour B .

Proposition 1.1 (Lemme d'inversion) *Soient (P, C) un modèle d'empilement, P fini, v une valuation et $B \subseteq P$. Soit \mathcal{H}_B l'ensemble des empilements dont les pièces minimales sont incluses dans B , $\mathcal{T}_{P \setminus B}$ l'ensemble des empilements triviaux dont toutes les pièces sont d'un type appartenant à $P \setminus B$. La série génératrice des empilements valués de \mathcal{H}_B est une*

fraction rationnelle dont le dénominateur est la série alternée des empilements triviaux de \mathcal{T} et le numérateur est la série alternée des empilements triviaux de $\mathcal{T}_{P \setminus B}$:

$$H_{B,v}(t) = \frac{T_{P \setminus B, -v}(t)}{T_{-v}(t)}.$$

Preuve. On veut montrer que

$$H_{B,v}(t)T_{-v}(t) = T_{P \setminus B, -v}(t).$$

On écrit le produit du premier membre en sommant sur les empilements $g = e \odot f$ obtenus en faisant tomber un empilement trivial f sur un empilement e dont toutes les pièces minimales sont de type dans B :

$$H_{B,v}(t)T_{-v}(t) = \sum_{g \in \mathcal{H}} \left(\sum_{e \odot f = g} v(e)(-1)^{|f|} v(f) \right).$$

On considère un empilement g et on veut calculer sa contribution à la somme. L'ensemble $\max(g)$ des pièces maximales de g se partitionne en deux :

$$\max_1(g) = \{(p, h) \in \max(g) \mid h > 0 \text{ ou } p \in B\}$$

l'ensemble des pièces maximales de hauteur non nulle ou de type appartenant à B et

$$\max_2(g) = \{(p, h) \in \max(g) \mid h = 0 \text{ et } p \notin B\}$$

l'ensemble des pièces maximales qui sont aussi minimales dans g mais d'un type n'appartenant pas à B . Soit $g = e \odot f$ une factorisation de l'empilement. Les pièces de $\max_2(g)$ appartiennent nécessairement à f sinon e n'est pas un empilement de $\mathcal{H}_{B,v}$. Par contre les pièces de $\max_1(g)$ peuvent indifféremment appartenir à e ou f . Pour obtenir toutes les décompositions $g = e \odot f$ il suffit de choisir un sous-ensemble s de $\max_1(g)$ et de poser $f = s \cup \max_2(g)$. La contribution de l'empilement g est donc

$$\sum_{s \subseteq \max_1(g)} (-1)^{|s|} v(g)(-1)^{|\max_2(g)|} = (1-1)^{|\max_1(g)|} v(g)(-1)^{|\max_2(g)|}.$$

Pour que la contribution de g soit non-nulle il est donc nécessaire que $\max_1(g)$ soit vide ce qui implique que g est un empilement trivial dont aucune pièce n'a un type dans B . Dans ce cas, $f = g$, e est vide et la valuation des pièces de g est $-v$, on retrouve donc la série alternée $T_{P \setminus B, -v}(t)$.

□

Le calcul de la série génératrice alternée des empilements triviaux est souvent plus simple que celui de la série des empilements mais cela reste une difficulté.

1.2.2 Empilements de dominos et animaux dirigés

L'intérêt des empilements pour les animaux dirigés est qu'il existe des bijections entre certains empilements et les animaux dirigés. On présente ici les travaux de Bétréma et Penaud [BP93]. On considère le chemin de longueur n qui est un graphe $P_n = (V, E)$ dont les sommets sont étiquetés par les entiers de 0 à n et les arêtes relient les sommets étiquetés i et $i + 1$ pour $i = 0 \dots n - 1$. L'ensemble des types de pièces est l'ensemble des arêtes de P_n . Dans ce cadre, une pièce est appelée un domino et un domino d'indice i correspond à l'arête $(i, (i + 1))$. Une *demi-pyramide (droite)* est un empilement dont l'éventuelle unique pièce minimale est un domino d'indice 0. Un empilement strict est un empilement tel qu'il n'existe pas deux pièces de la forme (p, h) et $(p, h + 1)$. On note \mathcal{DP}_n l'ensemble des demi-pyramides sur le graphe P_n et \mathcal{DP}_n^s l'ensemble des demi-pyramides strictes.

Proposition 1.2 *Il existe une bijection préservant la taille entre les demi-pyramides sur le graphe P_n et les demi-animaux de largeur au plus n sur réseau triangulaire.*

Il existe une bijection préservant la taille entre les demi-pyramides strictes sur le graphe P_n et les demi-animaux de largeur au plus n sur réseau carré.

Preuve. On fait tourner le réseau d'un huitième de tour dans le sens trigonométrique pour mieux voir la correspondance. On associe au site (i, j) un domino d'indice $i - j$ et de rang $i + j$. On fait tomber ces dominos selon les $i + j$ croissants. La figure 1.11 illustre cette construction sur un exemple.

La bijection sur le réseau carré est une restriction de celle du réseau triangulaire. Elle a une propriété supplémentaire qui est que les niveaux des pièces sont directement liés aux coordonnées des sites du demi-animal. À un demi-animal $a = \{(i, j) \in a\}$ sur le réseau carré, on associe l'empilement $e = \{(i - j, i + j)\}$. Cette construction est illustrée par un exemple sur la figure 1.12. \square

La *largeur d'une demi-pyramide* est le plus grand indice d'un des dominos de la demi-pyramide augmenté de un. La série génératrice usuelle de ces demi-pyramides,

$$DP_n(t) = \sum_{e \in \mathcal{DP}_n} t^{|e|},$$

est liée aux polynômes de Fibonacci $(T_n(t))_{n \geq 0}$ définis dans la partie 1.1.2.

Proposition 1.3 *La série génératrice selon le nombre de dominos des demi-pyramides de largeur au plus n est*

$$DP_n(t) = \frac{T_{n-1}(t)}{T_n(t)}.$$

Preuve. Soit $A_n(t)$ la série génératrice de l'ensemble \mathcal{T}_n des empilements triviaux sur le graphe P_n :

$$A_n(t) = \sum_{e \in \mathcal{T}_n} t^{|e|}.$$

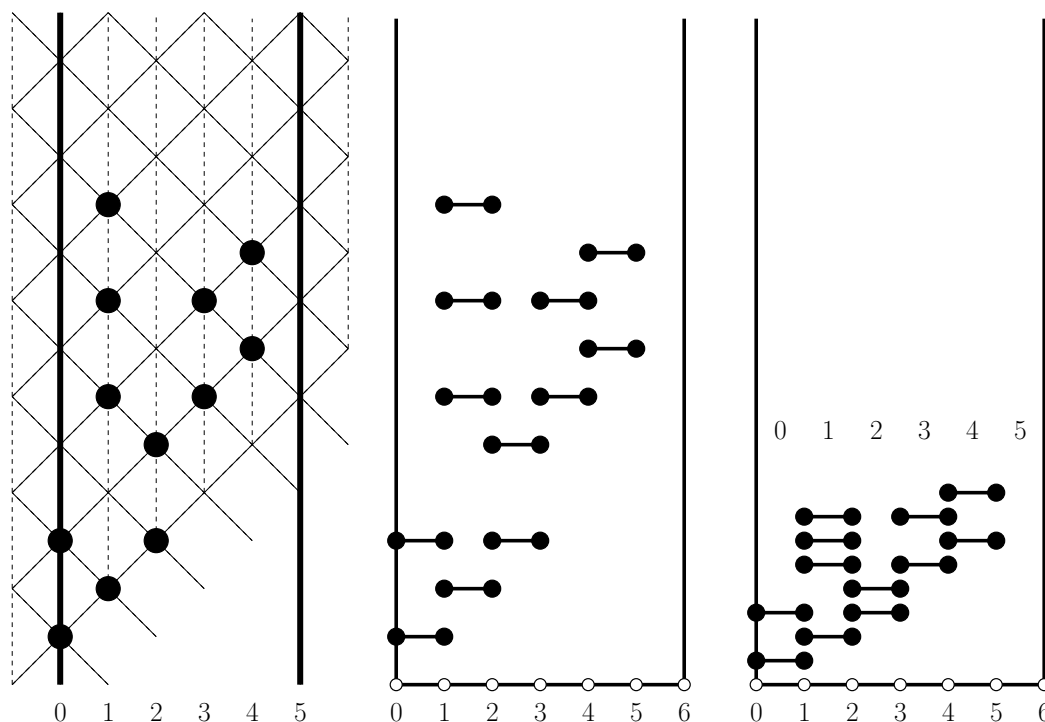


FIG. 1.11 – Demi-animaux sur le réseau triangulaire et demi-pyramides de dominos

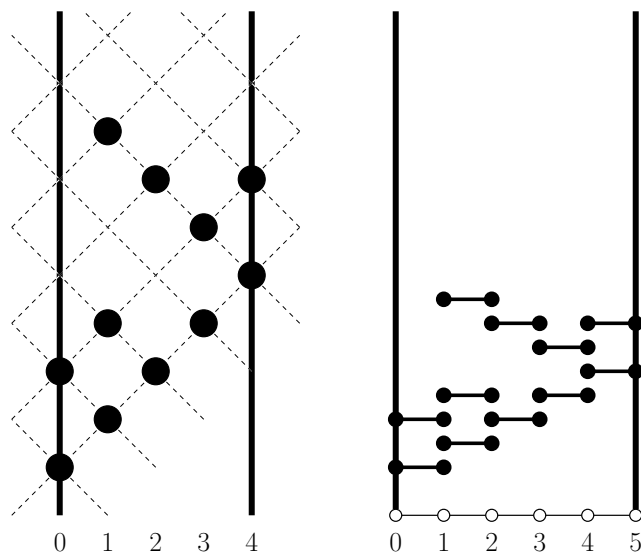
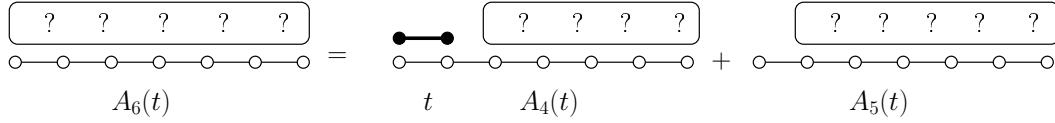


FIG. 1.12 – Demi-animaux sur le réseau carré et demi-pyramides strictes de dominos

Il n'est pas possible de poser des dominos sur les graphes P_0 et $P_{-1} = (\emptyset, \emptyset)$ donc $A_0(t) = 1 = A_{-1}(t)$. Pour obtenir une expression de $A_n(t)$ on partitionne les empilements triviaux du graphe P_n en deux sous-ensembles. La figure 1.13 illustre cette décomposition pour $A_6(t)$. Les empilements qui contiennent un domino d'indice 0 se complètent en un

FIG. 1.13 – Décomposition des empilements triviaux sur P_6

empilement trivial de P_{n-2} dont les indices des dominos sont augmentés de 2. La série génératrice de ce sous-ensemble est donc $tA_{n-2}(t)$. Un empilement sans domino d'indice 0 est aussi un empilement trivial de P_{n-1} dont les indices sont augmentés de 1. La série génératrice de ce sous-ensemble est $A_{n-1}(t)$. Finalement

$$A_n(t) = tA_{n-2}(t) + A_{n-1}(t).$$

Cette récurrence et les premières valeurs de $A_n(t)$ permettent de conclure que $A_n(-t) = T_n(t)$. En appliquant la formule de la proposition 1.1 avec pour ensemble $B = \{(01)\}$, le seul domino d'indice 0, on obtient l'expression de la proposition 1.3. \square

Il existe une relation entre les empilements et les empilements stricts qui permet d'établir un résultat similaire pour l'ensemble \mathcal{DP}_n^s des demi-pyramides strictes de dominos de largeur au plus n .

Proposition 1.4 *La série génératrice selon le nombre de pièces des demi-pyramides strictes de dominos de largeur au plus n est*

$$DP_n^s(t) = \frac{T_{n-1}\left(\frac{t}{1+t}\right)}{T_n\left(\frac{t}{1+t}\right)}.$$

Preuve. On construit un unique empilement strict à partir d'un empilement quelconque en supprimant la pièce $(p, h+1)$ de toute occurrence du doublon $\{(p, h+1), (p, h)\}$ de l'empilement. Cette opération peut s'inverser si sur la pièce (p, h) le nombre de pièces supprimées à été mémorisé. La figure 1.14 illustre cette construction sur un exemple. En termes de séries génératrices chaque domino de l'empilement strict, valué t , peut être remplacé par une colonne de dominos contenant au moins une pièce. La valuation totale de ces divers remplacements est $\sum_{n \geq 1} t^n = \frac{t}{1-t}$. On a donc

$$DP_n^s\left(\frac{t}{1-t}\right) = DP_n(t). \quad (9)$$

En inversant cette relation on obtient

$$DP_n^s(t) = DP_n\left(\frac{t}{1+t}\right) = \frac{T_{n-1}\left(\frac{t}{1+t}\right)}{T_n\left(\frac{t}{1+t}\right)}.$$

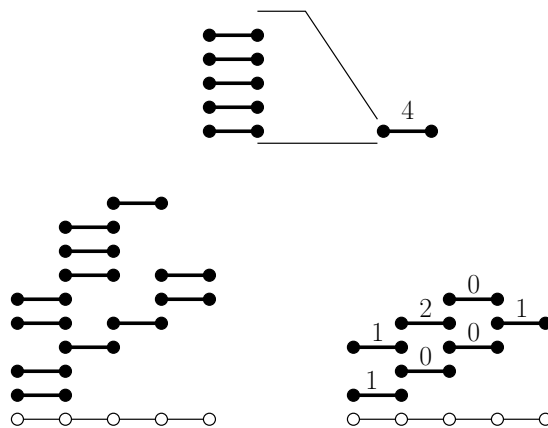


FIG. 1.14 – Lien entre empilements et empilements stricts

□

On a présenté le lien entre les polynômes de Fibonacci et les demi-pyramides sur ruban borné pour éclairer les conjectures à venir. Cependant la notion d'empilement permet d'écrire directement des équations de décomposition algébriques lorsque l'on considère le graphe infini P_∞ dont l'ensemble des sommets est indexé par \mathbb{N} et les arêtes sont de la forme $(i, (i + 1))$. Sur ce graphe la série génératrice des empilements n'est pas définie ; il y a une infinité d'empilements avec un seul domino. Par contre la série génératrice des demi-pyramides est bien définie car il n'y a qu'un nombre fini de demi-pyramides avec n dominos. Pour des raisons techniques on considère $DP(t)$ la série génératrice selon le nombre de dominos des demi-pyramides *non vides*.

Proposition 1.5 *La série génératrice $DP(t)$ des demi-pyramides non vides satisfait l'équation de décomposition*

$$DP(t) = t + 2tDP(t) + tDP(t)^2.$$

Preuve. Pour factoriser une demi-pyramide non-vide on considère le domino minimal d_1 , et on discute de l'existence d'un second domino d_2 d'indice 0 et de niveau minimal et de celle d'un troisième domino d_3 d'indice 1 et de niveau 1. Les différents cas sont illustrés dans la figure 1.15. S'ils existent ces dominos sont la base d'une demi-pyramide, dont les

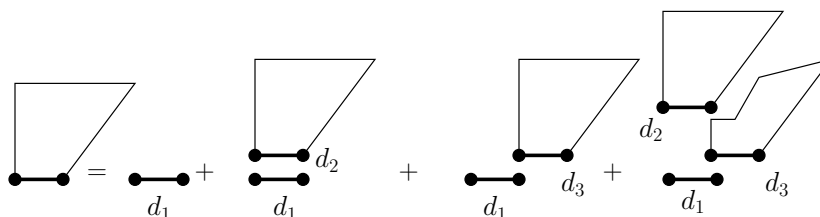


FIG. 1.15 – Décomposition algébrique des demi-pyramides

indices sont décalés de 1 pour une demi-pyramide de pièce minimale d_3 . Si les deux existent simultanément, on extrait d'abord la demi-pyramide de pièce minimale d_2 en soulevant ce domino comme illustré sur la figure 1.16. Cette décomposition des demi-pyramides non-

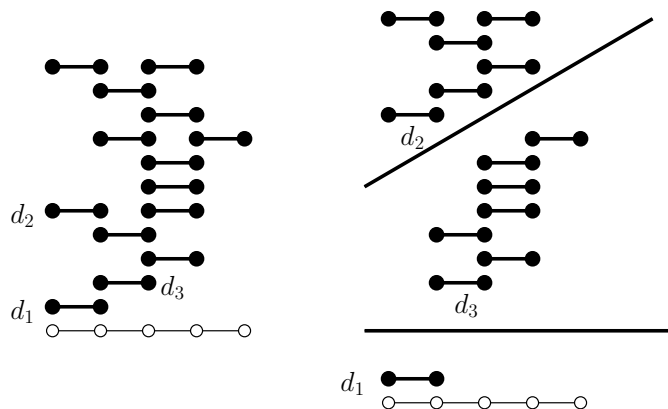


FIG. 1.16 – Factorisation en deux demi-pyramides

vides conduit à l'équation algébrique

$$DP(t) = t + tDP(t) + tDP(t) + tDP(t)^2.$$

□

Corollaire 1.5.1 *La série génératrice $DP^s(t)$ des demi-pyramides strictes satisfait l'équation algébrique*

$$DP^s(t) = t + tDP^s(t) + tDP^s(t)^2.$$

Preuve. Le cas où le domino d_2 existe mais pas le domino d_3 ne se présente pas dans les décompositions des pyramides strictes. De plus si les demi-pyramides de pièces minimale d_2 ou d_3 sont strictes alors les diverses recompositions conduisent à des pyramides strictes. Ces cas conduisent à l'équation de décomposition du corollaire. □

1.3 Conjectures pour le premier moment sur un ruban borné

1.3.1 Une différence entre les réseaux triangulaire et carré ?

La correspondance entre les pyramides et les animaux permet d'énumérer les animaux selon leur aire. Lorsque l'on souhaite ajouter le périmètre de site, une tentative raisonnable

consiste à placer un domino d'une autre nature sur les sites périmétriques. Cette correspondance entre animaux et empilements échoue pour le réseau triangulaire tandis que c'est l'énumération des empilements qui pose problème pour le réseau carré.

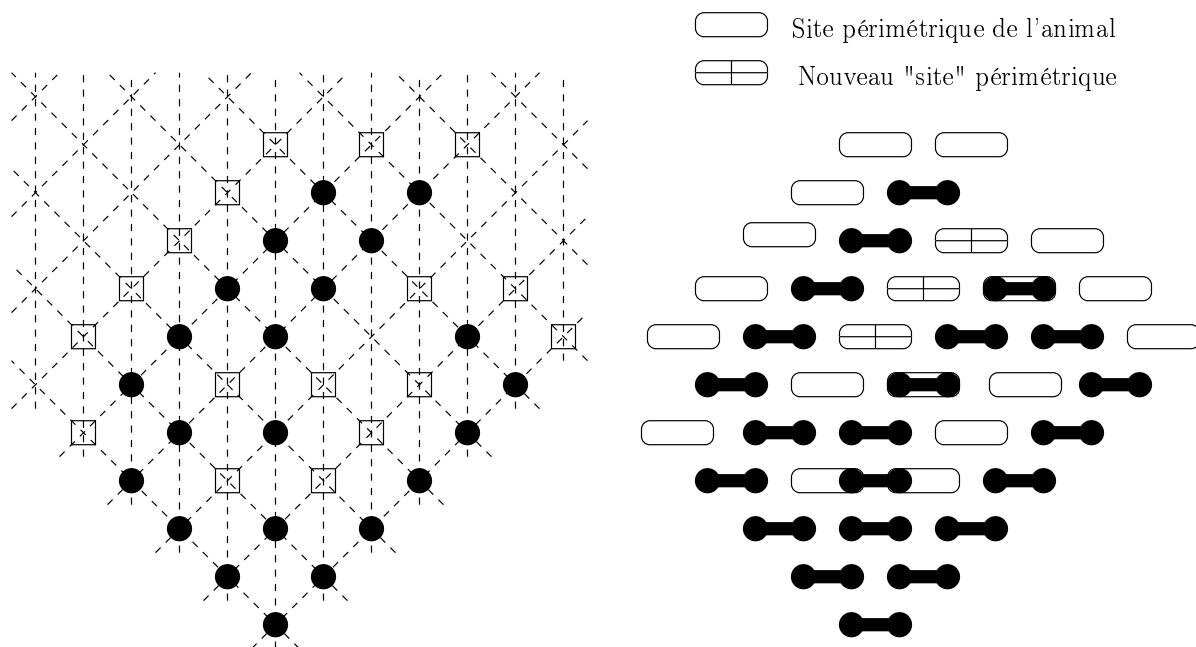


FIG. 1.17 – Perturbation du périmètre de site sur le réseau triangulaire

Sur le réseau triangulaire, il y a des déplacements relatifs des dominos pour passer de l'animal à l'empilement. Comme l'illustre la figure 1.17, les dominos des périmètres de site peuvent apparaître ou disparaître à cause des chutes des pièces.

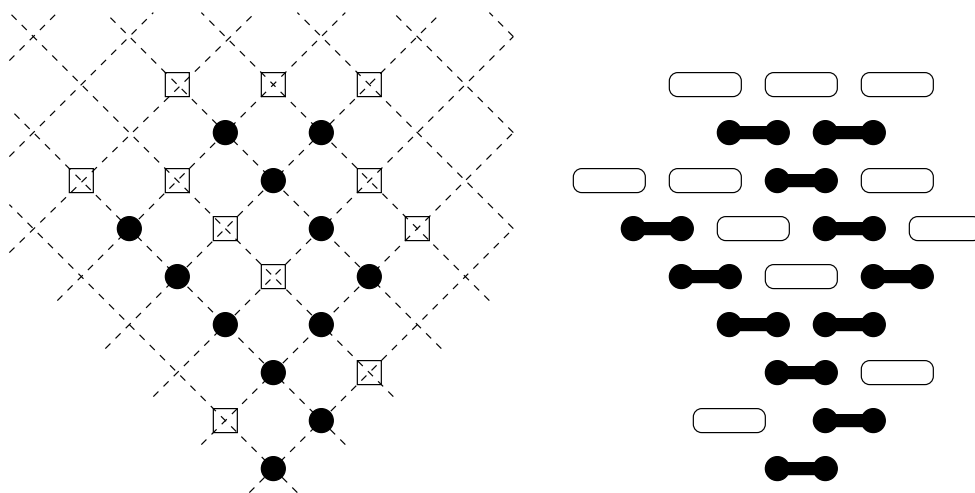


FIG. 1.18 – Préservation du périmètre de site sur le réseau carré

En revanche, la correspondance entre les animaux et pyramides strictes est "rigide",

c'est-à-dire que les sites transformés en dominos ne bougent pas lorsque l'on en déduit l'empilement. La figure 1.18 est un exemple de cette correspondance. Grâce à cette propriété on peut définir le périmètre de site d'une *demi-pyramide stricte* e comme le cardinal de l'ensemble des positions voisines de l'empilement où on peut ajouter un domino sans perturber les niveaux du reste de l'empilement. Plus formellement, les positions périmétriques de l'empilement sont les pièces (p, h) telles que (p, h) n'appartienne pas à e , aucun des types des pièces de niveau h de e ne concurrence le type de pièce p et il existe dans e une pièce de niveau $h - 1$ dont le type concurrence le type p .

Dans le cas des demi-pyramides strictes il peut y avoir des sites périmétriques d'indice -1 comme l'illustre la figure 1.19.

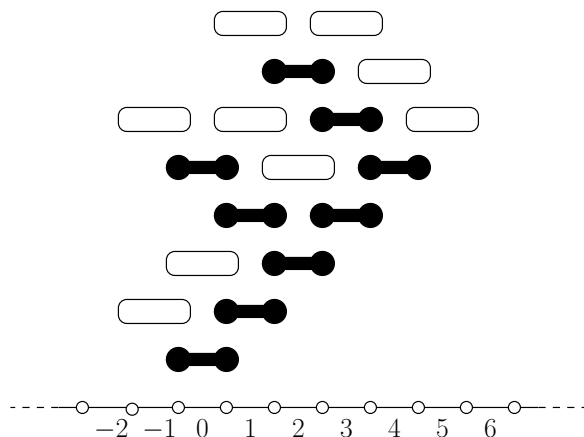


FIG. 1.19 – Périmètre de site des demi-pyramides strictes

Cette comparaison informelle entre les réseaux triangulaire et carré explique peut-être qu'il existe des conjectures comme (7) et (8) sur le premier moment des animaux sur le réseau carré mais pas sur le triangulaire. Pour les animaux dirigés c'est la première fois qu'une différence notable semble exister entre ces deux réseaux.

1.3.2 Matrice de transfert pour le périmètre de site

On va utiliser une approche par matrice de transfert pour énumérer selon l'aire et le périmètre les demi-animaux sur le réseau carré. Les définitions suivantes sont données pour le réseau carré et deviennent naturelles si on tourne le réseau carré d'un huitième de tour dans le sens trigonométrique comme sur la figure 1.20. La *hauteur d'un site* (i, j) est $i + j$. La *hauteur d'un animal* a est le maximum des hauteurs parmi les sites appartenant à a . La *ligne* h est l'ensemble des sites à la hauteur h . Pour un animal a , on code par C_k l'occupation de la ligne de hauteur k en posant

$$C_{2k} = \{(i - j)/2 | (i, j) \in a, i + j = 2k\}$$

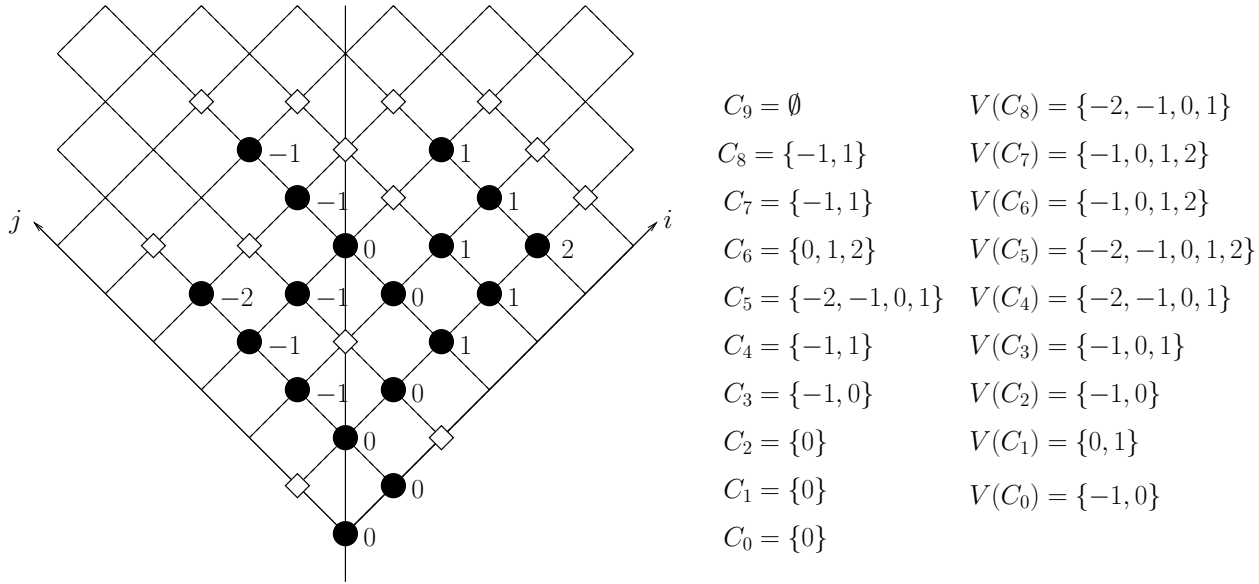


FIG. 1.20 – Codage par ligne d'un animal sur le réseau carré

et

$$C_{2k+1} = \{(i-j-1)/2 \mid (i,j) \in a, i+j = 2k+1\}.$$

Un animal est caractérisé par la suite de ses lignes $C_0, \dots, C_{a(h)+1}$, la dernière ligne $C_{a(h)+1}$ étant vide. On note $V(i,j) = \{(i+1,j), (i,j+1)\}$ le *voisinage* d'un sommet (i,j) . Le voisinage d'un ensemble de sites E est

$$V(E) = \bigcup_{(i,j) \in E} V((i,j)).$$

En traduisant cette notion sur le codage des lignes,

$$V(C_{2k}) = \bigcup_{i \in C_{2k}} \{i-1, i\}$$

et

$$V(C_{2k+1}) = \bigcup_{i \in C_{2k+1}} \{i, i+1\}.$$

Soit $C_0 C_1 \dots C_{a(h)+1}$ la décomposition en lignes de l'animal a . Deux lignes consécutives C_i et C_{i+1} sont *compatibles* si et seulement si $C_{i+1} \subseteq V(C_i)$. Une suite de lignes est compatible si deux lignes consécutives dans la suite sont toujours compatibles. De plus, la ligne de hauteur 0 d'un animal est soit vide soit le seul site $(0,0)$ occupé. Les deux lignes possibles pour C_0 sont \emptyset et $\{0\}$. L'aire de l'animal est

$$\sum_{i=0}^{a(h)} |C_i|$$

et le périmètre de site est

$$\sum_{i=0}^{a(h)} (|V(C_i)| - |C_{i+1}|).$$

On définit par ces formules l'aire et le périmètre de site pour une suite de lignes compatibles. La nature des sites sur la ligne h est déterminée par les lignes C_{h-1} et C_h . On partitionne l'ensemble des suites de lignes compatibles selon la parité de la dernière ligne :

$$H_{pair}(A) = \sum_{C_0 \dots C_{2k-1} A} t^{\sum_{i=0}^{2k-1} |C_i| + |A|} r^{\sum_{i=0}^{2k-2} (|V(C_i)| - |C_{i+1}|) + (|V(C_{2k-1})| - |A|)}.$$

est la série génératrice selon l'aire et le périmètre des suites dont la dernière ligne A est de hauteur paire, et symétriquement

$$H_{impair}(B) = \sum_{C_0 \dots C_{2k} B} t^{\sum_{i=0}^{2k} |C_i| + |B|} r^{\sum_{i=0}^{2k-1} (|V(C_i)| - |C_{i+1}|) + (|V(C_{2k})| - |B|)}$$

est la série génératrice selon l'aire et le périmètre des suites dont la dernière ligne B est de hauteur impaire. En retirant la dernière ligne d'une suite on obtient les équations de décompositions suivantes liant ces séries :

$$\left\{ \begin{array}{l} H_{pair}(\emptyset) = r + \sum_{B \neq \emptyset | \emptyset \subseteq V(B)} r^{|V(B)|} H_{impair}(B), \\ H_{pair}(\{0\}) = t + \sum_{B \neq \emptyset | \{0\} \subseteq V(B)} t r^{|V(B)|-1} H_{impair}(B), \\ H_{pair}(A) = \sum_{B \neq \emptyset | A \subseteq V(B)} t^{|A|} r^{|V(B)|-|A|} H_{impair}(B), \\ H_{impair}(B) = \sum_{A \neq \emptyset | B \subseteq V(A)} t^{|B|} r^{|V(A)|-|B|} H_{pair}(A). \end{array} \right.$$

La série génératrice des animaux se calcule à partir des séries génératrices des suites se terminant par une ligne vide

$$D(t, r) = H_{pair}(\emptyset) + H_{impair}(\emptyset).$$

Proposition 1.6 *La série génératrice $D_{1/2,n}(t, r)$ selon l'aire et le périmètre de site des demi-animaux de largeur au plus n est une fraction rationnelle en t et r . Le premier moment $\frac{\partial}{\partial r} D_{1/2,n}(t, 1)$ selon le périmètre de ces objets est une fraction rationnelle en t .*

Preuve. Sur le réseau carré entier, il y a une infinité d'occupations possibles pour une ligne. Par contre pour les demi-animaux de largeur au plus n , les occupations possibles des lignes paires sont les $2^{\lceil n/2 \rceil}$ sous-ensembles de $\{0, 1, \dots, \lceil n/2 \rceil\}$ et celles des lignes impaires sont les $2^{\lfloor n/2 \rfloor}$ sous-ensembles de $\{0, 1, \dots, \lfloor n/2 \rfloor\}$. Le système linéaire restreint à ces occupations des lignes possède $2^{\lceil n/2 \rceil}$ inconnues $H_{pair}(A)$ ou $H_{impair}(B)$ et autant

d'équations. Ce système possède une solution non-nulle puisque les séries génératrices des suites compatibles sont bien définies et en sont une solution. Les séries solutions de ce système sont des fractions rationnelles en r et t , donc $D_{1/2,n}(r, t)$ qui est la somme de deux de ces séries est une fraction rationnelle en t, r . La dérivée selon r et l'évaluation en $r = 1$ préservant le caractère rationnel $\frac{\partial}{\partial r} D_{1/2,n}(t, 1)$ est une fraction rationnelle en la variable t . \square

Le calcul pratique des fractions rationnelles a été effectué de la façon suivante. On veut calculer H , sachant que $H = MH + B$ où M est la matrice des coefficients des inconnues du système linéaire et B correspond aux entrées constantes. Par itération on a $H = \sum_{n \geq 0}^N M^n B + M^N H$. Chaque produit par M multiplie tous les termes par au moins un coefficient t car chaque ligne, sauf la dernière, contient au moins un site, et les séries sont entières donc en observant les termes de degré au plus $N - 1$ en t dans l'équation précédente on a $H(t) = \sum_{n \geq 0}^N M^n B + \mathcal{O}(t^N)$. Par cette formule, on calcule la série tronquée à l'ordre N de $D_n(t, r)$. Ce calcul est un calcul par matrice de transfert (*cf.* [Sta86]). On en déduit la série tronquée à l'ordre N du premier moment. Une fraction rationnelle $P(t)/Q(t)$ est caractérisée par les $\deg(P) + \deg(Q) + 1$ premiers termes de son développement parmi les fractions rationnelles de degré comparable. En utilisant une approximation de Padé on conjecture la fraction rationnelle en t dont le début du développement coïncide avec le développement tronqué calculé. Pour n allant de 2 à 7 on a obtenu les fractions

$$\begin{aligned} \frac{\partial}{\partial r} D_{1/2,2}(t, 1) &= -\frac{t(-2+t)}{(-1+t)^2}, \\ \frac{\partial}{\partial r} D_{1/2,3}(t, 1) &= -\frac{t(-2-t+t^2+t^3)}{(t+1)(t^2+t-1)^2}, \\ \frac{\partial}{\partial r} D_{1/2,4}(t, 1) &= \frac{t(2-3t-t^2+3t^3)}{(t+1)(2t-1)^2}, \\ \frac{\partial}{\partial r} D_{1/2,5}(t, 1) &= -\frac{t(-2+3t+t^6-8t^4+5t^2-5t^3)}{(t+1)(t^3-t^2-2t+1)^2}, \\ \frac{\partial}{\partial r} D_{1/2,6}(t, 1) &= -\frac{t(-2+7t+2t^5+t^4-3t^2-8t^3)}{(t+1)(-1+t)^2(t^2+2t-1)^2}, \\ \frac{\partial}{\partial r} D_{1/2,7}(t, 1) &= \frac{t(2-7t-15t^5+t^6+4t^7+t^4-t^2+18t^3)}{(t+1)(1-3t+3t^3)^2}, \\ \frac{\partial}{\partial r} D_{1/2,8}(t, 1) &= -\frac{t(-2+11t+5t^5-14t^6-t^7+t^8+22t^4-15t^2-9t^3)}{(t+1)(t^2+t-1)^2(t^2-3t+1)^2}. \end{aligned}$$

Pour prouver ces résultats on pourrait essayer d'inverser la matrice M mais elle est assez volumineuse dès les valeurs $n = 7$ ou 8 .

1.3.3 Conjecture

On étudie les demi-animaux de largeur au plus n sur le réseau carré. En utilisant l'égalité (9) liée à la transformation des demi-pyramides en demi-pyramides strictes puis la correspondance entre animaux et empilements, on constate que

$$D_{1/2,n} \left(\frac{t}{1-t}, 1 \right) = E_{1/2,n}(t) = \frac{T_{n-1}(t)}{T_n(t)}.$$

Cette expression suggère de chercher une expression des fractions rationnelles $\frac{\partial}{\partial r} D_{1/2,n}(\frac{t}{1-t}, 1)$ en fonction de ces polynômes $T_n(t)$. Grâce à Maple, on a conjecturé ces fractions rationnelles pour $n \leq 20$ et l'observation de ces polynômes a conduit à la conjecture suivante.

Conjecture 1.6.1 *Pour $n \geq 2$, la fraction rationnelle*

$$\frac{\partial}{\partial r} D_{1/2,n}(t/(1-t), 1) = \frac{P_n(t)}{Q_n(t)}$$

admet pour dénominateur

$$Q_n(t) = T_n(t)^2 \tag{10}$$

et son numérateur satisfait la récurrence

$$P_{n+2}(t) - t^2 P_n(t) = -t^2 (1-t+t^2) T_{n-1}(t)^2 + t(2-t)(1-t) T_n(t)^2 + t^2 T_{n+1}(t)^2. \tag{11}$$

L'observation expérimentale débute après la substitution $t \rightarrow t/(1-t)$. On remarque assez vite que les dénominateurs $Q_n(t)$ s'expriment simplement en fonction des polynômes de Fibonacci par la formule

$$Q_n(t) = T_n(t)^2.$$

Pour les numérateurs $P_n(t)$, on a dû les observer un peu plus méticuleusement. Les coefficients de ces polynômes sont de signes alternants, on donne donc ici les polynômes évalués en $-t$. Ces premiers polynômes sont

$$\begin{aligned} P_2(-t) &= 2t + 3t^2, \\ P_3(-t) &= 2t + 7t^2 + 8t^3 + 4t^4 + t^5, \\ P_4(-t) &= 2t + 11t^2 + 20t^3 + 12t^4 + t^5, \\ P_5(-t) &= 2t + 15t^2 + 40t^3 + 45t^4 + 23t^5 + 8t^6 + 2t^7, \\ P_6(-t) &= 2t + 19t^2 + 68t^3 + 114t^4 + 93t^5 + 35t^6 + 5t^7, \\ P_7(-t) &= 2t + 23t^2 + 104t^3 + 235t^4 + 281t^5 + 176t^6 + 60t^7 + 16t^8 + 3t^9, \\ P_8(-t) &= 2t + 27t^2 + 148t^3 + 424t^4 + 683t^5 + 624t^6 + 319t^7 + 92t^8 + 14t^9, \\ P_9(-t) &= 2t + 31t^2 + 200t^3 + 697t^4 + 1427t^5 + \dots + 554t^8 + 148t^9 + 30t^{10} + 4t^{11}, \\ P_{10}(-t) &= 2t + 35t^2 + 260t^3 + 1070t^4 + 2673t^5 + 4185t^6 + \dots + 937t^9 + 223t^{10} + 30t^{11}. \end{aligned}$$

Les régularités des coefficients de degrés extrémaux suggèrent des combinaisons linéaires. La forme des dénominateurs incite à observer les polynômes $T_i(t)T_j(t)$. La chance a peut-être aussi sa place dans les tâtonnements qui conduisent à la découverte de la relation de récurrence (11).

Proposition 1.7 *La conjecture 1.6.1 implique que la série génératrice $\frac{\partial}{\partial r}D_{1/2}(t, 1)$ des demi-animaux pointés dans un site du périmètre de site et la série génératrice $D_{1/2}(t, 1)$ des demi-animaux sur le réseau carré sont liées par la relation*

$$\left(\frac{(1+t)^2}{D_{1/2}(t, 1)^2} - t^2 D_{1/2}(t, 1)^2 \right) \frac{\partial}{\partial r} D_{1/2}(t, 1) + \frac{t^2 (1+t+t^2) D_{1/2}(t, 1)^4}{(1+t)^2} - \frac{t(2+t) D_{1/2}(t, 1)^2}{1+t} - t^2 = 0 \quad (12)$$

Preuve. La série génératrice des demi-animaux de largeur non-bornée est la limite de la suite des séries génératrice des demi-animaux de largeur bornée par n . On en déduit que

$$\lim_{n \rightarrow \infty} \frac{T_{n-1}(t)}{T_n(t)} = E_{1/2}(t, 1).$$

Plus généralement

$$\lim_{n \rightarrow \infty} \frac{T_{n-k}(t)}{T_n(t)} = \lim_{n \rightarrow \infty} \prod_{i=0}^{k-1} \frac{T_{n-i-1}(t)}{T_{n-i}(t)} = (E_{1/2}(t, 1))^k.$$

Pour la même raison le premier moment selon le périmètre des demi-pyramides s'obtient par

$$\frac{\partial}{\partial r} D_{1/2}\left(\frac{t}{1-t}, 1\right) = \lim_{n \rightarrow \infty} \frac{P_n(t)}{Q_n(t)}.$$

En divisant l'équation (11) par $Q_n(t)$ on fait apparaître des termes $\frac{\partial}{\partial r} D_{1/2, n}\left(\frac{t}{1-t}, 1\right)$ et $T_n(t)/T_k(t)$ qui, lorsque n tend vers l'infini, donnent une relation entre la série génératrice du premier moment $\frac{\partial}{\partial r} D_{1/2}\left(\frac{t}{1-t}, 1\right)$ et des puissances de $E_{1/2}(t, 1)$:

$$\frac{\frac{\partial}{\partial r} D\left(\frac{t}{1-t}, 1\right)}{E(t, 1)^2} - t^2 \frac{\partial}{\partial r} D\left(\frac{t}{1-t}, 1\right) E(t, 1)^2 + t^2 (1-t+t^2) E(t, 1)^4 - t(2-t)(1-t) E(t, 1)^2 - t^2 = 0. \quad (13)$$

Pour obtenir une équation sur $\frac{\partial}{\partial r} D_{1/2}(t, 1)$ on fait le changement de variable $t \rightarrow t/(1+t)$. Comme ce changement de variable est liée à la transformation d'un empilement général en un empilement strict, on a la relation $E_{1/2}(t/(1+t), 1) = D_{1/2}(t)$ qui fait apparaître les demi-pyramides strictes. On obtient alors la relation de l'énoncé entre le premier moment du périmètre et la série selon l'aire.

□

Remarque 1.7.1 *En se souvenant que $E_{1/2}(t, 1)$ est aussi la série génératrice des demi-pyramides sur le réseau triangulaire on a la relation*

$$E_{1/2}(t, 1) = 1 + t (E_{1/2}(t, 1))^2$$

qui permet de rendre la relation (13) linéaire en $E_{1/2}(t, 1)$. Malheureusement les coefficients grossissent significativement, on a donc préféré laisser cette équation sous sa forme initiale.

Le début de la suite d'énumération du premier moment permet aussi de conjecturer une équation algébrique satisfaite par $\frac{\partial}{\partial r} D(t, 1)$. Cette forme de conjecture correspond à celle qu'a obtenue Conway pour le premier moment des animaux, cf équations (7) et (8).

Conjecture 1.7.1 *Le premier moment des demi-animaux pointés dans le périmètre de site et énumérés selon l'aire satisfait l'équation algébrique*

$$\frac{t(-2 + 3t + 9t^2 + 5t^4 + 7t^3 + 3t^5)}{(t+1)^2(3t-1)} + (t^3 + t^2 + 2t - 1) \frac{\partial}{\partial r} D_{1/2}(t, 1) + t^2(t+1) \left(\frac{\partial}{\partial r} D_{1/2}(t, 1) \right)^2 = 0 \quad (14)$$

et donc

$$\frac{\partial}{\partial r} D_{1/2}(t, 1) = \frac{3t^5 + 5t^4 + 7t^3 - 4t + 1 + (-1 + 3t + t^2 + 2t^3 + 3t^4) \sqrt{(t+1)(1-3t)}}{2t^2(t+1)^2(1-3t)}. \quad (15)$$

Le début du développement du premier moment obtenu par cette formule est

$$\frac{\partial}{\partial r} D_{1/2}(t, 1) = 2t + 3t^2 + 8t^3 + 19t^4 + 51t^5 + 135t^6 + 370t^7 + 1020t^8 + 2849t^9 + 8012t^{10} + 22689t^{11} + 64586t^{12} + 18$$

1.4 Modèles de gaz

Dhar a utilisé deux modèles de gaz différents [Dha82], [Dha83] pour donner les deux premières preuves de la formule d'énumération des animaux dirigés. Le principe de ces preuves est de prouver que les animaux dirigés et la densité d'un modèle de gaz vérifient deux systèmes d'équations équivalents. L'expertise des physiciens sur les modèles de gaz permet de calculer la densité du modèle de gaz et d'en déduire la série génératrice des animaux dirigés. On présente rapidement dans la section 1.4.1 une partie des idées du second travail de Dhar.

Bousquet-Mélou [BM98] a généralisé cette technique ce qui lui a permis de prouver que la série génératrice $D_{droit}(t, u)$ des animaux dirigés sur le réseau carré selon l'aire et le périmètre droit satisfait une équation algébrique de degré 2. Pour l'instant on ne dispose pas de décomposition combinatoire des animaux expliquant cette équation algébrique. Les

animaux dirigés sur le réseau carré sont en bijection avec les préfixes de chemins de Motzkin. On propose dans la section 1.4.2 une interprétation combinatoire sur les chemins de Motzkin de l'équation de décomposition algébrique en introduisant sur les chemins un paramètre qui se comporte comme le périmètre droit des animaux.

Le problème de la percolation dirigée se ramène à calculer

$$D(p, 1-p) = \sum_{a \in \mathcal{A}} p^{\text{aire}(a)} (1-p)^{\text{perimetre}(a)}$$

où \mathcal{A} est l'ensemble des animaux dirigés. On présente dans la section 1.4.3 un sous-ensemble $\mathcal{B} \subseteq \mathcal{A}$ pour lequel on sait calculer la série

$$\sum_{b \in \mathcal{B}} p^{\text{aire}(b)} (1-p)^{\text{perimetre}(b)}.$$

lorsqu'il n'y a pas de percolation donc pour $p < p_c$.

1.4.1 Une énumération de Dhar utilisant un modèle de gaz

Pour énumérer les animaux sur le réseau carré, Dhar étudie un modèle de gaz dont l'espace est de dimension 1. Dans ce modèle, nommé par la suite CGM, des particules identiques peuvent occuper ou non les N sites d'un cercle indexé par $\mathbb{Z}/N\mathbb{Z}$. L'état de ce modèle à l'instant t est donc caractérisé par l'ensemble S_t des sites occupés. L'occupation d'un site i à l'instant $t+1$ dépend de la parité de t , des occupations des sites $i-1$, i , $i+1$ et d'un paramètre aléatoire $p \in [0, 1]$. Si t est pair, respectivement impair, le site i est occupé à l'instant $t+1$ avec une probabilité p si et seulement si les sites i et $i-1$, respectivement i et $i+1$, sont inoccupés à l'instant t . Ces règles de transition sont illustrées par la figure 1.21, les probabilités de transition pour le site i de la première ligne se déduisent de l'occupation à l'instant t des deux sites i et $i-1$ ou $i+1$ de la seconde ligne.

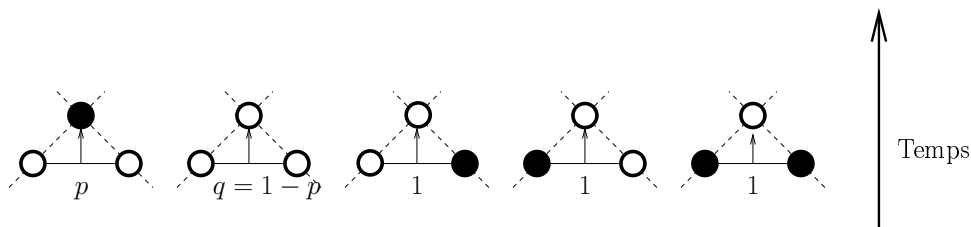


FIG. 1.21 – Règles de transition des modèles de gaz

Ce modèle de gaz a été introduit par des physiciens pour observer l'état d'équilibre d'un modèle de gaz à particules dures sur le réseau carré en utilisant des algorithmes de type Monte-Carlo. Pour faire le lien entre ces deux modèles de gaz, il faut observer le diagramme espace/temps du modèle CGM en décalant les lignes correspondants aux instants impairs

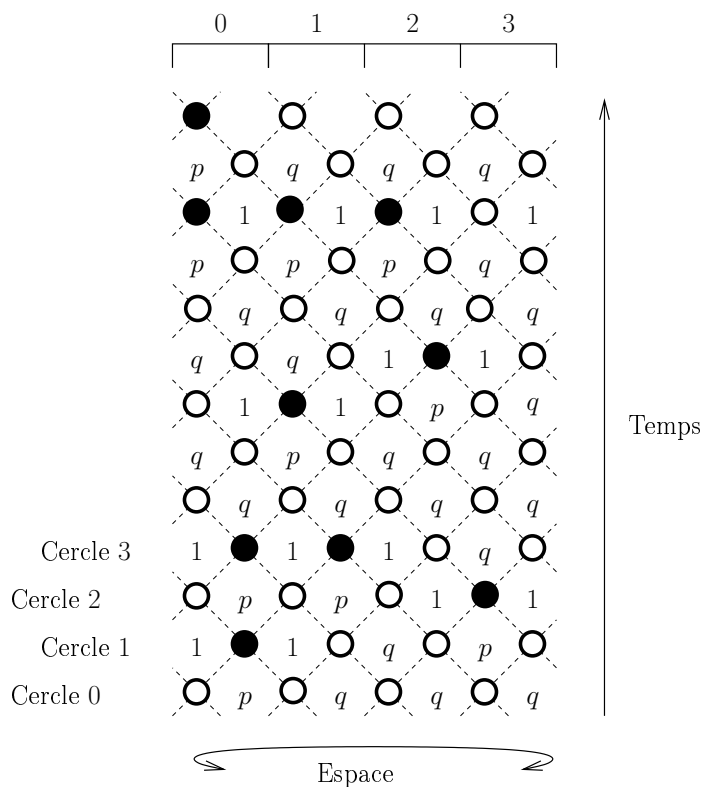


FIG. 1.22 – Diagramme espace-temps du modèle de gaz

d'un demi-inter-site. La figure 1.22 donne une partie d'un diagramme espace/temps de l'évolution d'un gaz sur un réseau carré de périmètre 4.

Les dépendances entre les sites lors des transitions dessinent alors un réseau carré cylindrique de périmètre N . Dans ce réseau, les sites occupés ne sont jamais voisins, ce qui correspond à un état du modèle de gaz à particules dures. Les physiciens programmant des algorithmes de type Monte-Carlo considèrent que lorsque l'on observe les propriétés statistiques de ce diagramme espace/temps pour un temps élevé on retrouve les propriétés à l'équilibre du modèle de gaz sur le réseau carré. La distribution limite du modèle CGM, observée après un temps infini, est très particulière : elle est Boltzmanienne et son Hamiltonien est suffisamment simple pour que Dhar ait pu le conjecturer. Cette propriété permet de calculer explicitement la densité uniforme $\rho_N(p)$ de ce modèle de gaz. Pour un exposé plus détaillé, on peut se référer aux travaux de Bousquet-Mélou [BM98]. Ici, on se contente du fait que l'on sait calculer la densité du modèle de gaz CGM. Dhar a établi un lien entre cette densité et la série génératrice $D_N(t)$ des animaux dirigés sur le réseau cylindrique de périmètre N .

Proposition 1.8 (Dhar 83) *La densité uniforme $\rho_N(p)$ du modèle de gaz CGM sur le*

cercle $\mathbb{Z}/N\mathbb{Z}$ est liée à la série génératrice des animaux dirigés sur le réseau carré cylindrique de périmètre N par

$$D_N(t = -p, 1) = 1 - \rho_N(p).$$

Preuve. On suppose que le gaz est à l'équilibre, ce qui implique que la probabilité d'un état est indépendante du temps. Sans entrer dans les détails, l'existence et l'unicité de cette probabilité sont garanties, en outre, par le fait que l'on travaille sur cercle $\mathbb{Z}/N\mathbb{Z}$ et non la ligne infinie \mathbb{Z} . On note $P(C)$ la probabilité que tous les sites de C soient occupés simultanément. Supposons que tous les sites de C soient occupés à l'instant τ . La notation $V(C)$ est la même que celle employée pour le codage des animaux par une suite de lignes. D'après les probabilités de transition du modèle CGM, la probabilité de cet évènement est aussi la probabilité que premièrement, tous les sites de $V(C)$, le voisinage de C , soient vides à l'instant $\tau - 1$ et que, deuxièmement, tous les $|C|$ tirages aléatoires aient été réussis avec une probabilité p pour chacun. En remarquant que l'instant $\tau - 1$ ne dépend pas de l'instant τ et en utilisant le principe de l'inclusion-exclusion, la probabilité du premier évènement peut s'exprimer en fonction des probabilités $P(D)$ pour $D \in V(C)$; le second évènement est de probabilité $p^{|C|}$. Finalement on obtient l'expression

$$P(C) = p^{|C|} \sum_{D \subseteq V(C)} (-1)^{|D|} P(D).$$

La densité $\rho(p)$ est aussi la probabilité $P(\{0\})$ d'occupation du site 0 car le modèle est invariant par translation.

Pour faire le lien avec les animaux on utilise la décomposition en lignes de la section 1.3. Les lignes sont maintenant des cercles du réseau cylindrique. Soit C l'occupation d'un cercle. Un *animal de source* C sur la ligne de hauteur h est un sous-ensemble fini a des sites du réseau carré tel que l'animal a occupe tous les sites de C du cercle h et que pour tous les autres sites (i, j) de a , $(i - 1, j) \in a$ ou $(i, j - 1) \in a$. L'aire d'un tel animal est le cardinal de l'ensemble a . On note $A(C)$ la série génératrice selon l'aire des animaux de source C . Cette série est indépendante de la hauteur du cercle du réseau carré occupé par les sources C , on ne mentionne donc pas cette hauteur dans la définition de $A(C)$. Les animaux dirigés sont les animaux de source $\{0\}$ (ou \emptyset) sur le cercle de hauteur 0. On décompose un animal de source C en supprimant le cercle C contenant les sources du reste b de l'animal. Le reste b est un animal de source D telle que $D \subseteq V(C)$. On obtient les équations

$$\forall C, A(C) = t^{|C|} \sum_{D \subseteq V(C)} A(D).$$

Les séries $P(C)$ et $A(C)$ satisfont presque les mêmes systèmes d'équations linéaires. En multipliant l'équation de décomposition de $P(C)$ par $(-1)^{|C|}$ on constate que $(-1)^{|C|} P(C)$ et $A(C)$ sont solutions du même système si $t = -p$:

$$(-1)^{|C|} P(C) = (-p)^{|C|} \sum_{D \subseteq V(C)} (-1)^{|D|} P(D).$$

Dans ce cas, l'unicité à une constante multiplicative près des solutions sous forme de série entière permet de prouver que $(-1)^{|C|}P(C) = \lambda A(C)$. Comme d'après leur définition on sait que $P(\emptyset) = 1$ et $A(\emptyset) = 1$, la constante multiplicative λ est 1. La série génératrice des animaux non-vides est donc

$$A(\{0\}) = (-1)^1 P(\{0\}) = -\rho_N(p).$$

Il suffit de rajouter la contribution de l'animal vide pour conclure. \square

Le calcul de la densité du modèle CGM est possible car on peut identifier la distribution invariante par la transition (en pratique on travaille sur deux lignes ou cercles consécutifs). Comme pour beaucoup de modèles de dimension 1 ce calcul s'effectue à l'aide des matrices de transfert. On ne détaille pas cette étape que l'on peut lire dans [BM98]. La suite des séries formelles $(\rho_N(p))_{N \in \mathbb{N}}$ converge formellement vers la série algébrique

$$R(p) = -\frac{1}{2} \left(\sqrt{\frac{1-p}{1+3p}} - 1 \right).$$

Cette série est définie pour $p \in]-1/3, 1[$.

Les animaux de taille inférieure à $N - 1$ sur le réseau cylindrique de périmètre N sont en bijection avec les animaux de taille $N - 1$ sur le réseau carré. (Il n'est pas possible de faire le tour du cylindre avec $N - 1$ sites.) La suite des séries formelles $(D_N(t))_{N \in \mathbb{N}}$ converge formellement vers la série $D(t)$. De ces deux convergences formelles et des égalités de la proposition 1.8 on déduit que

$$D(t) = 1 - R(p).$$

Dans cette preuve on n'évoque jamais la densité d'un modèle de gaz défini sur la ligne infinie \mathbb{Z} . En effet, travailler sur un cercle $\mathbb{Z}/N\mathbb{Z}$ permet de garantir l'existence et l'unicité d'une densité limite. Dans la partie 1.4.3 on travaillera avec le modèle de gaz sur la ligne \mathbb{Z} mais pour une valeur de $p < p_c$ telle que la probabilité de percolation soit nulle. Cette hypothèse permettra de définir une densité de ce modèle. Pour faire le lien entre ces deux hypothèses on peut remarquer que pour tout $p < 1$ la probabilité de percolation est nulle sur le réseau cylindrique.

1.4.2 Interprétation sur les préfixes de chemins de Motzkin de l'équation algébrique pour le périmètre droit

En généralisant la méthode de Dhar [Dha82], Bousquet-Mélou [BM98] a construit un modèle de gaz dont le calcul de la densité lui a permis de déterminer la série génératrice $D_{droit}(t, u)$ des animaux dirigés non-vides sur le réseau carré selon l'aire et le périmètre droit. Elle a montré que cette série satisfait l'équation algébrique

$$(1 - t(2 + u) + t^2(1 - u)) (D_{droit}(t, u) + D_{droit}(t, u)^2) = t. \quad (16)$$

On ne dispose pas de décomposition combinatoire des animaux expliquant cette équation.

Un *chemin quelconque* est un chemin dans le plan, issu de l'origine et formé de pas diagonaux montants $(1, 1)$ et descendant $(1, -1)$ ainsi que de pas horizontaux $(1, 0)$. La longueur d'un chemin est le nombre de pas qui le compose. Un chemin quelconque est un *préfixe de Motzkin* si tous ses pas sont dans le demi-plan supérieur. Un préfixe de Motzkin est un chemin de Motzkin si son dernier sommet est sur l'axe horizontal. Un *pic* dans un chemin quelconque est un sommet précédé d'un pas montant et suivi d'un pas descendant. Le *poids* d'un chemin quelconque est la somme de son nombre de pics et de son nombre de pas horizontaux.

Gouyou-Beauchamps et Viennot [GBV88] ont décrit une bijection entre les préfixes de Motzkin de longueur $n - 1$ et les animaux dirigés d'aire n . Le poids d'un préfixe de Motzkin proposé dans cette thèse est l'analogie du périmètre droit sur les animaux dirigés :

Proposition 1.9 *Soit $P(t, u)$ la série génératrice des préfixes de Motzkin comptés selon la longueur et le poids. Cette série satisfait l'équation algébrique*

$$P(t, u) + tP(t, u)^2 = \frac{1}{1 - (t + ut + (t - t^2 + ut^2))} \quad (17)$$

de sorte que

$$P(t, u) = \frac{1}{t} D_{\text{droit}}(t, u).$$

Preuve. L'équation (17) s'obtiendra en décomposant un chemin quelconque en un ou deux préfixes de Motzkin.

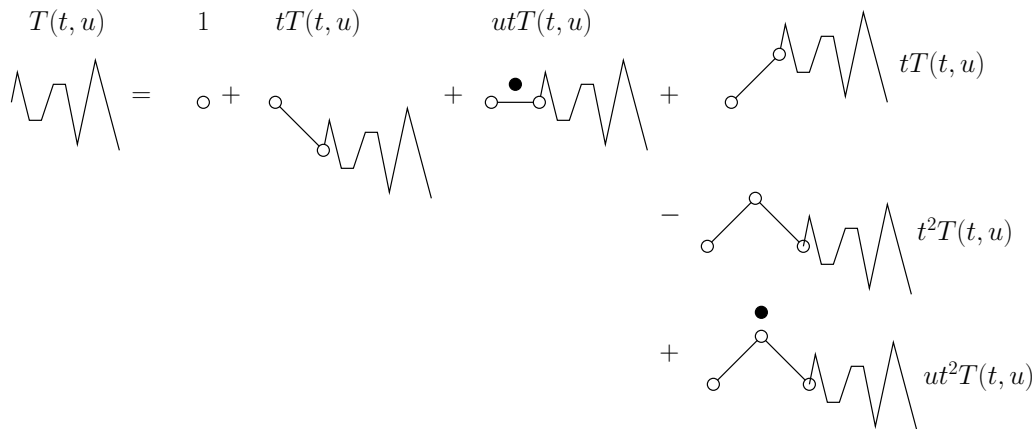


FIG. 1.23 – Décomposition des chemins quelconques

On calcule d'abord la série génératrice $T(t, u)$ des chemins quelconques selon la longueur et le poids. On discute sur les premiers pas du chemin comme l'illustre la figure 1.23 : soit le chemin est vide (1), soit il débute par un pas descendant ($tT(t, u)$), soit par un pas horizontal ($utT(t, u)$) soit par un pas montant et ce cas est un peu plus complexe. La

série des chemins débutant par un pas montant est $tT(t, u)$, sauf que pour les chemins où il existe un deuxième pas qui est descendant ($t^2T(t, u)$), la série de cet ensemble doit prendre en compte le poids du premier pic ce qui donne ($t^2uT(t, u)$). La valuation des chemins débutant par un pas montant est donc $tT(t, u) - t^2T(t, u) + ut^2T(t, u)$. Cette décomposition conduit donc à l'équation

$$T(t, u) = 1 + tT(t, u) + ut^2T(t, u) + (tT(t, u) - t^2T(t, u) + ut^2T(t, u)).$$

On reconnaît que $T(t, u)$ est le second membre de l'équation (17).

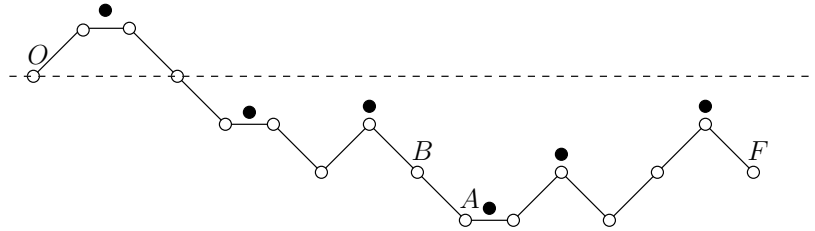


FIG. 1.24 – Factorisation des chemins quelconques en préfixes de Motzkin

On factorise un chemin quelconque en observant la première fois qu'il atteint sa hauteur minimale. Si ce minimum est aussi l'origine du chemin, on a un préfixe de Motzkin et la série de ces chemins est $P(t, u)$. Sinon comme sur l'exemple la figure 1.24, on note A ce premier minimum, B le sommet précédent, O l'origine du chemin et F le dernier sommet. Par choix de A , BA est un pas descendant; entre A et F il y a un préfixe de Motzkin à une translation près; entre O et B il y a un préfixe de Motzkin lu à l'envers. La série génératrice des chemins AF est $P(t, u)$. Le pas descendant BA n'appartient pas à un pic, sinon A ne serait pas le premier minimum, il est donc valué t . D'autre part, le poids d'un chemin est préservé lorsque le chemin est lu à l'envers donc la série des chemins OB est $P(t, u)$. Cette décomposition conduit donc à l'équation

$$T(t, u) = P(t, u) + P(t, u)tP(t, u).$$

□

La factorisation des chemins quelconques en préfixes de Motzkin utilisée dans la preuve précédente fait partie du folklore combinatoire. La difficulté était de trouver la statistique du poids sur les chemins quelconques.

La bijection de Gouyou-Beauchamps et Viennot est plus générale que sa restriction qui associe un préfixe de Motzkin à un animal : elle associe un chemin quelconque à un animal du réseau carré à source compacte. Un *animal à source compacte* $k \in \mathbb{N}$ est un animal dont la source est l'ensemble des sites $\{(-i, i), 0 \leq i \leq k - 1\}$. Dans la partie gauche de la figure 1.25, il y a un exemple d'animal à source compacte. La factorisation des chemins quelconques en préfixes de Motzkin se traduit sur les animaux à source compacte non-vide. Dans la partie droite de la figure 1.25 on illustre cette factorisation. Il suffit de voir les animaux comme des empilements stricts de dominos et de pousser la pièce à l'origine du

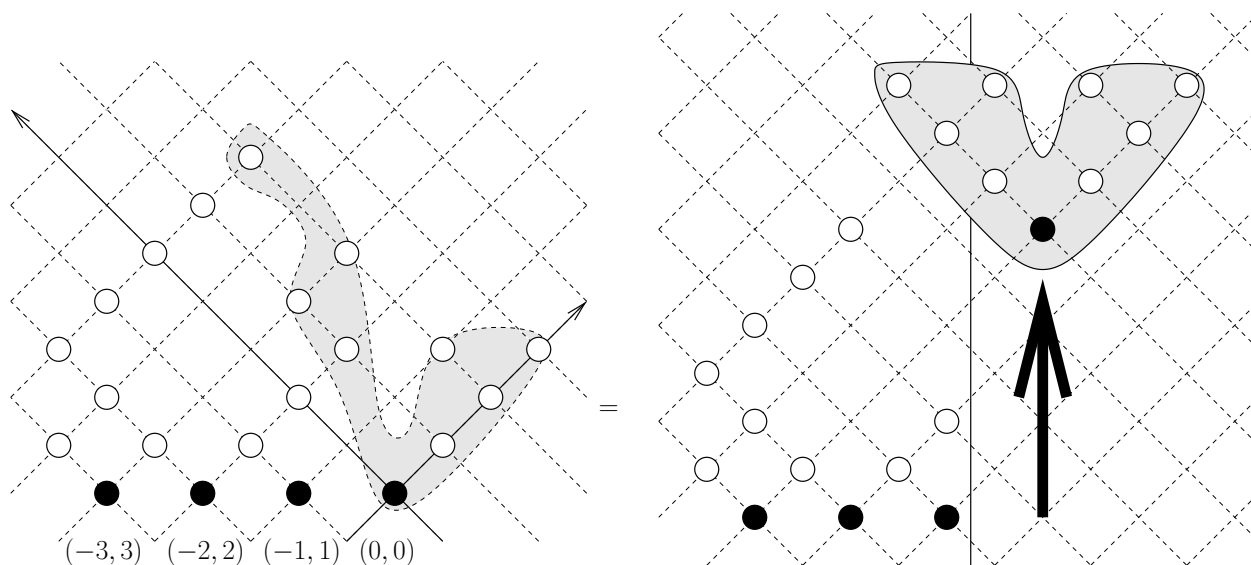


FIG. 1.25 – Un animal à source compacte 4 et sa factorisation

réseau pour faire apparaître un animal non-vide dont elle est la source; le reste est un demi-animal gauche de source compacte éventuellement vide. Un *demi-animal gauche de source compacte* k est un animal de source $\{(i, -i), 1 \leq k - 1\}$ tel que tous ses sites (i, j) soit dans le demi-plan défini par $i \leq j + 2$. Comme le montre sur un exemple la figure 1.26, un demi-animal gauche se factorise en une suite de demi-animaux gauche non-vides; pour cela il suffit de remplacer les sites par des dominos et de pousser les dominos de la source en commençant par celui sur le site $(1, -1)$. Les animaux se factorisent aussi en une suite de demi-animaux gauche, en poussant les dominos de hauteur minimale dans chaque colonne à la droite de la source. On obtient ainsi une bijection entre les animaux et des demi-animaux gauche de source compacte.

On peut donc interpréter en termes d'empilements de dominos l'équation algébrique (16) pour $u = 1$ mais on ne dispose pas de statistique analogue au périmètre droit pour les animaux de source compacte et les demi-animaux gauche de source compacte. Un intérêt de cette factorisation en empilement est que la série génératrice des animaux de source compacte est rationnelle en t et u ce qui laisse espérer qu'il soit possible de trouver un analogue du poids des chemins sur ces objets plus "simples" que les animaux.

1.4.3 Résolution de la percolation pour une sous-famille d'animaux

Pour définir un sous-ensemble des animaux dirigés sur le réseau carré on définit un coloriage en noir et blanc des sites d'un animal a et de ses sites périmétriques. Initialement, les sites périmétriques sont blancs et les sites de l'animal sont non coloriés. Tant qu'il existe des sites non coloriés il existe un site (i, j) non colorié pour lequel les sites de $V((i, j))$ sont

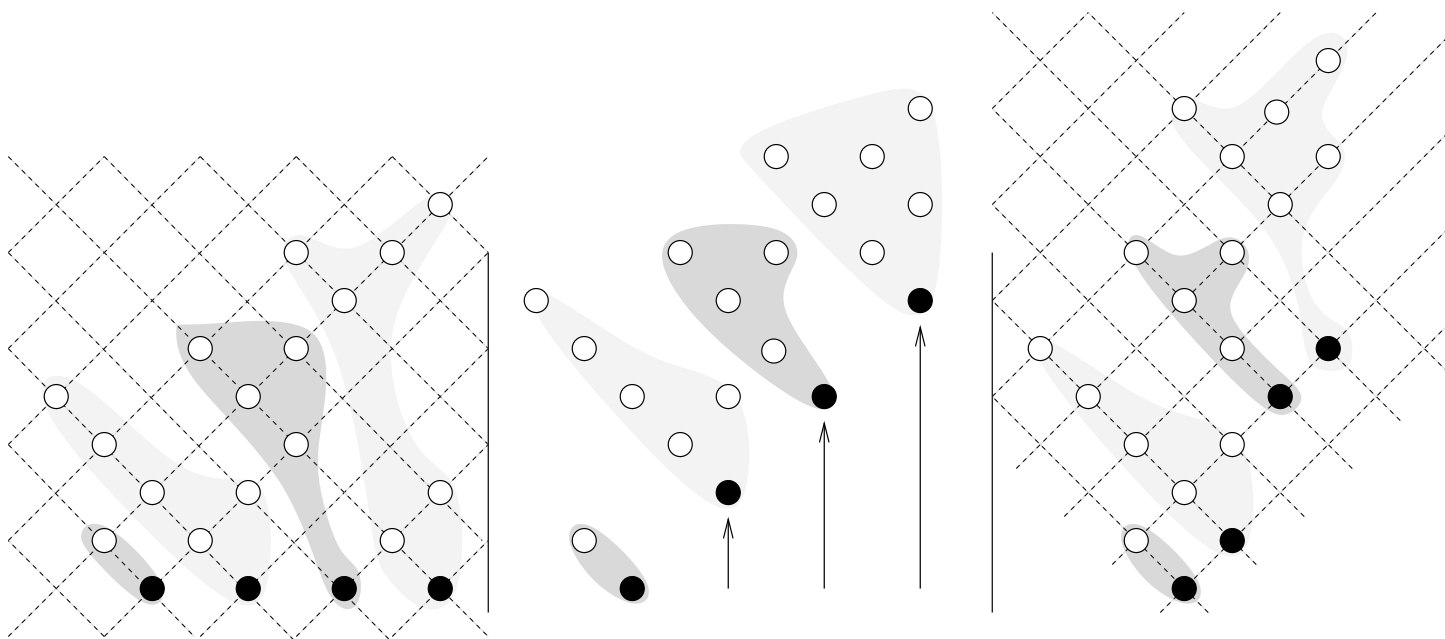


FIG. 1.26 – Demi-animal gauche de source compacte et animal

coloriés. On colorie alors le site (i, j) en noir si les deux sites de $V(i, j)$ sont blancs et sinon en blanc. De proche en proche tout l'animal est colorié comme l'illustre la figure 1.27. Un

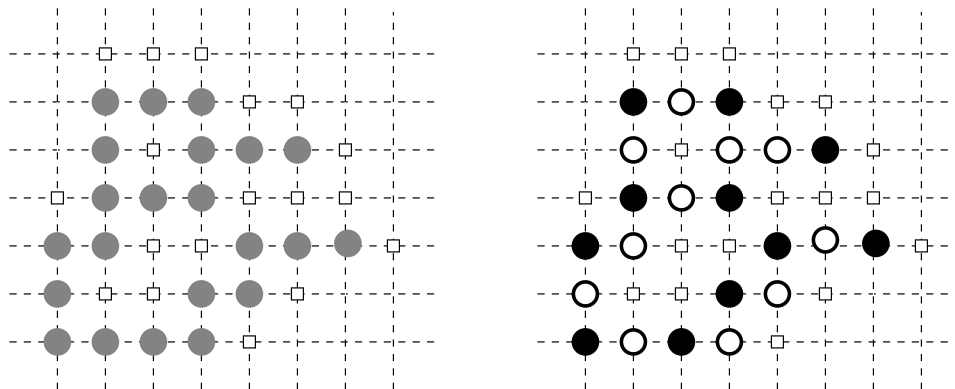


FIG. 1.27 – Coloration et occupation d'un animal

animal occupé est un animal dans lequel le sommet $(0, 0)$ est colorié en noir. Cette notion un peu alambiquée apparaît naturellement dans la preuve de la proposition 1.10.

Proposition 1.10 Soit $B(t, r)$ la série génératrice selon l'aire et le périmètre de site des animaux dirigés occupés sur le réseau carré. On considère le modèle de gaz CGM, sur la ligne infinie indexée par \mathbb{Z} , paramétré par un réel positif p strictement inférieur à la valeur p_c au dessus de laquelle il y a percolation sur le réseau carré.

La densité à l'équilibre $\rho(p)$ du modèle de gaz CGM paramétré par p existe et est unique. De plus on a

$$\rho(p) = B(p, 1 - p). \quad (18)$$

Preuve. Le modèle CGM évolue selon des critères de natures différentes. Le premier, déterministe, autorise l'occupation d'un site i à la date t si à la date $t - 1$ ce site et un de ses voisins $i \pm 1$ sont libres. Le second, probabiliste, autorise l'occupation du site i à la date t si un tirage aléatoire a été réussi avec probabilité p . Sur le diagramme espace/temps de la figure 1.22, on a représenté l'occupation des sites à chaque instant ainsi que la probabilité de la transition aléatoirement sélectionnée. On suppose que pour tous les sites i et à chaque instant t on commence par tirer aléatoirement 1 avec une probabilité p et 0 avec une probabilité $q = 1 - p$ et que, ensuite seulement, on regarde l'occupation des sites à l'instant $t - 1$. Le *diagramme aléatoire* d'un diagramme espace/temps du modèle CGM est alors une fonction qui à chaque site i et à chaque instant t renvoie le résultat du tirage aléatoire, 0 ou 1. Puisque les tirages sont indépendants, la probabilité d'observer une partie finie du diagramme aléatoire avec n valeurs 1 et m valeurs 0 est $p^n q^m$. Sur les schémas, on remplace les entiers 0 et 1 par leur probabilité d'occurrence q et p respectivement. Le premier diagramme de la figure 1.28 est un diagramme aléatoire, il y a 9 valeurs 0 et 23 valeurs 1, la probabilité d'observer ce diagramme est donc $q^9 p^{23}$. Une propriété des transitions du

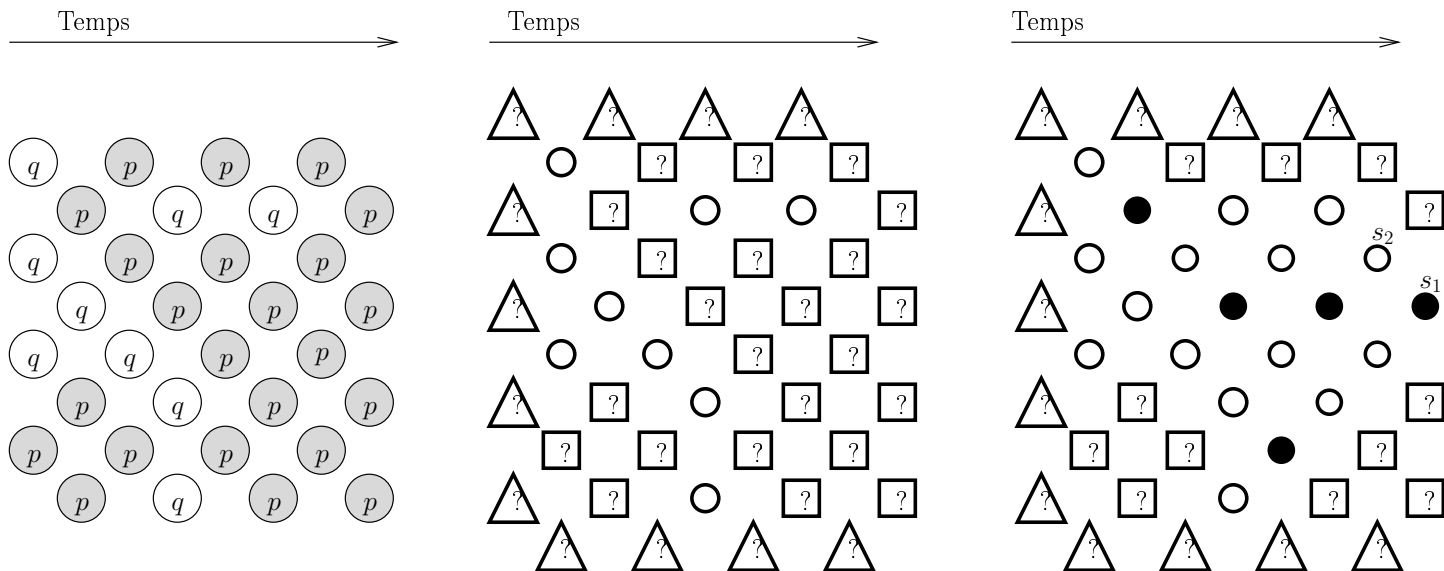


FIG. 1.28 – Dédution à partir du diagramme aléatoire

modèle CGM est que si on a tiré la valeur 0 pour le site i à l'instant t , ce site est libre quelles que soient les autres occupations des sites à l'instant $t - 1$. Le second diagramme de la figure 1.28 représente l'information sur les diagrammes espace/temps cohérents avec le diagramme aléatoire obtenue par cette propriété. Les triangles représentent les sites voisins du rectangle qui influence son contenu sans qu'on ait d'information sur eux. A priori

l'occupation de chaque site est inconnue sauf ceux pour lesquels on a tiré la valeur 0 avec la probabilité q qui sont nécessairement libres. Les règles de transition du modèle CGM permettent alors de déduire certaines occupations d'autres sites. Le troisième diagramme de la figure 1.28 représente toutes les informations sur les diagrammes espace/temps déduite du diagramme aléatoire par les considérations précédentes. On remarque par exemple que le diagramme aléatoire permet de déterminer que le site s_1 est occupé et que le site s_2 est libre. Les sites du diagramme aléatoire où on a tiré 1 avec probabilité p et qui permettent de déduire que s_1 est occupé forment un animal dont la source est s_1 . Plus généralement on connaît l'occupation d'un site s à partir du diagramme aléatoire en observant le plus grand animal de source s dont tous les sites sont valués 1. Comme par hypothèse $p < p_c$, la probabilité d'observer un "animal" infini est nulle donc le plus grand animal de source s existe. Ce plus grand animal ne dépend ni de l'instant ni du site observé, c'est ce qui permet de définir la densité $\rho(p)$ du modèle de gaz CGM.

La figure 1.29 représente le passé du diagramme aléatoire qui détermine l'occupation du site s et donc le plus grand animal. La probabilité d'observer un animal b est $p^{\text{aire}(b)} q^{\text{perimetre}(b)}$. La probabilité que le site s soit occupé est donc la somme de cette quantité sur les animaux dont on déduit par les règles du modèle CGM que la source est occupée. Or ces animaux sont les animaux occupés donc

$$\rho(p) = B(p, 1 - p).$$

□

On conclut cette partie par quatre remarques.

- Il semble possible de généraliser cette construction au modèle de gaz utilisé par Bousquet-Mélou pour déterminer le périmètre droit des animaux dirigés sur le réseau carré.

- Le coloriage déterminant si un animal a est occupé a un sens dans un autre contexte : celui des jeux de type Nim sur les graphes orientés. Ce jeu peut se jouer sur un animal dirigé a non-vide. Un jeton est initialement placé sur le site $(0, 0)$ de l'animal a . Deux joueurs, Alice et Bob, déplacent alternativement ce jeton qui doit rester tout au long de la partie sur un site de a . S'il est sur le site $(i, j) \in a$, le jeton ne peut être déplacé que sur les sites $(i + 1, j)$ ou $(i, j + 1)$. Lorsque qu'un joueur ne peut plus déplacer le jeton sans le faire sortir de l'animal il a perdu. Par convention Alice joue la première. Un site s est *gagnant* si le joueur devant déplacer le jeton situé en s est sûr de gagner en jouant au mieux, sinon le site est *perdant*. Le coloriage en noir ou blanc des sites d'un animal est lié à cette notion de site gagnant ou perdant. Dans un animal a , un site colorié en blanc est gagnant et un site colorié en noir est perdant. La figure 1.30 illustre la preuve de cette propriété. Les sites (i, j) dont aucun successeur $(i + 1, j)$ ou $(i, j + 1)$ n'appartient à l'animal sont perdants et sont noirs d'après la règle de coloriage. D'autre part, les autres sites de l'animal ne sont gagnants (blanc) que s'il est possible de déplacer le jeton sur un site perdant (noir) pour le tour de l'adversaire. Sinon le déplacement conduit à un site gagnant (blanc) pour l'adversaire et le site (i, j) est donc perdant (noir). Comme Alice

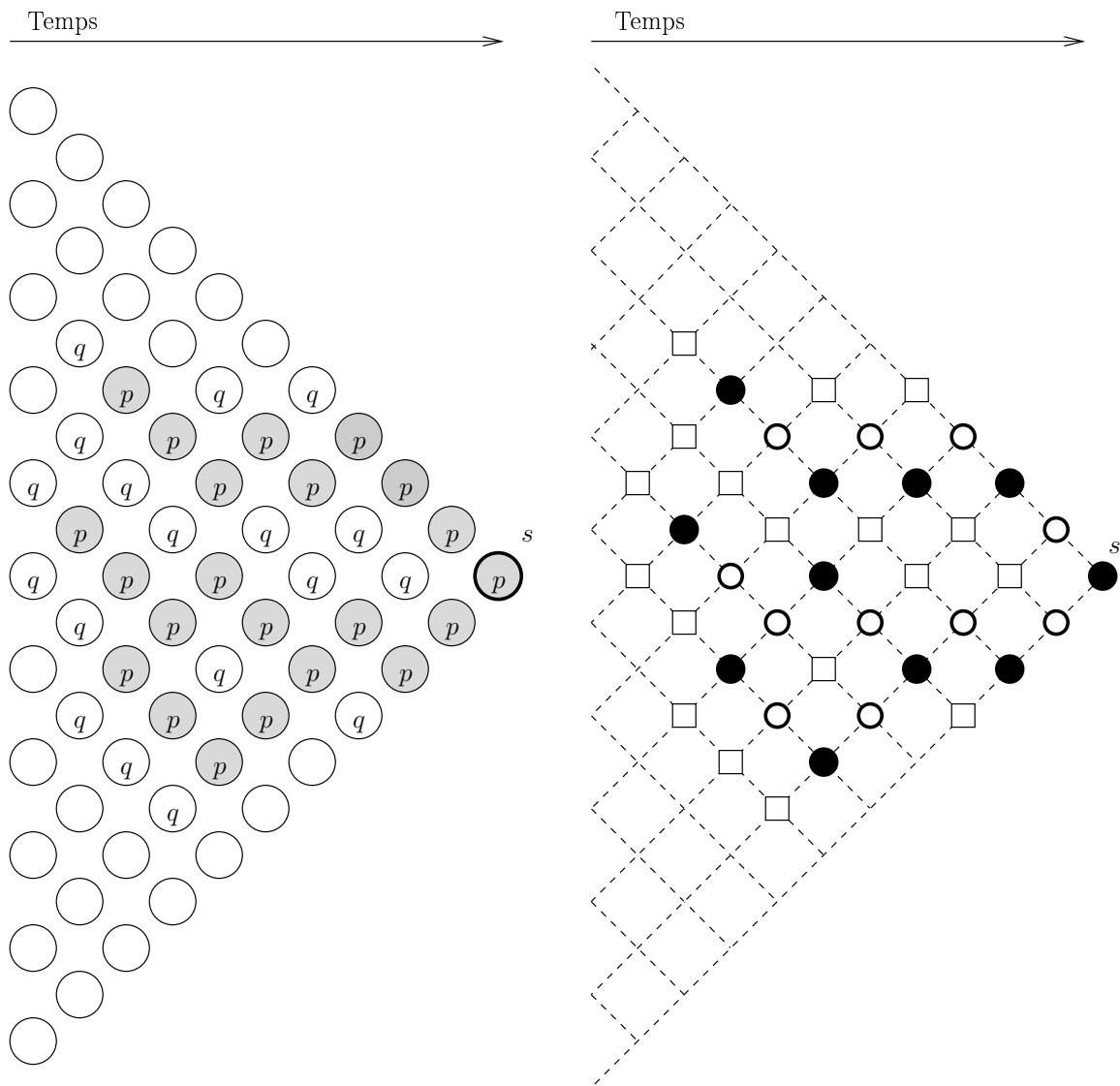


FIG. 1.29 – Triangle de calcul et densité du modèle de gaz

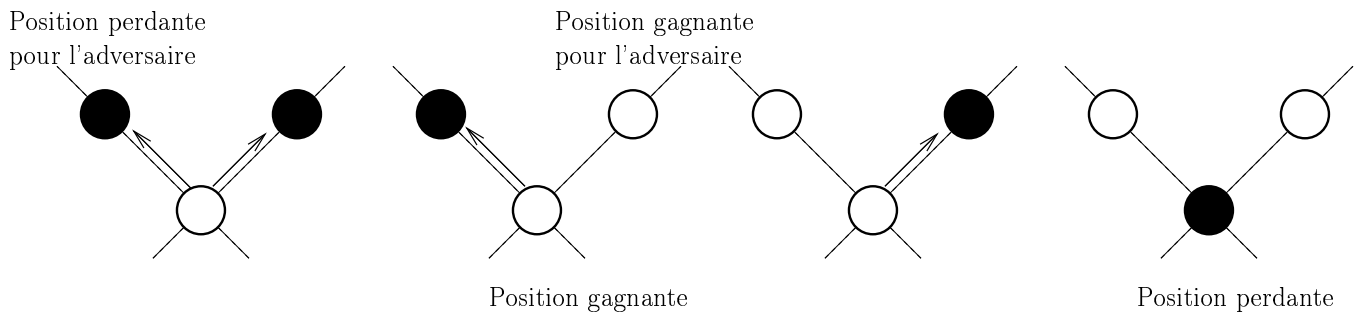


FIG. 1.30 – Stratégie pour le jeu de Nim

début la partie, les animaux occupés sont donc exactement ceux sur lesquels Bob a une stratégie gagnante. Pour s'en convaincre le lecteur peut prendre le rôle d'Alice et jouer contre Bob sur l'animal de la figure 1.27 : Bob déplace toujours le jeton vers un site noir et gagne. Le jeu de Nim sur les graphes a été étudié récemment par Banderier, Le Bars et Ravelomanana dans [BLBR].

- Je conjecture que la taille moyenne $t_A(p)$ des animaux est liée à la taille moyenne $t_B(p)$ des animaux occupés par une relation de la forme

$$K_1 t_B(p) \leq t_A(p) \leq K_2 t_B(p)$$

où K_1 et K_2 sont des constantes.

- L'égalité de la proposition 1.10 implique une égalité entre les séries génératrices de deux familles d'animaux dirigés

$$B(p, 1-p) = \sum_{a \text{ occupe}} p^{\text{aire}(a)} (1-p)^{\text{perimetre}(a)} = - \sum_{a \text{ quelconque}} (-p)^{\text{aire}(a)} = -D(-p, 1). \quad (19)$$

Une explication combinatoire de cette égalité éclairerait peut-être un peu mieux le lien entre les animaux et le modèle de gaz CGM.

Chapitre 2

Des bijections pour les chemins de Dyck

La suite $(c_n)_{n \in \mathbb{N}}$ de terme générique

$$c_n = \frac{1}{2n+1} \binom{2n+1}{n} = \frac{1}{n+1} \binom{2n}{n} \quad (20)$$

est la suite des nombres de Catalan. Les premiers termes de cette suite sont

$$1, 1, 2, 5, 14, 42, \dots \quad (21)$$

Elle est très fréquente en combinatoire, Stanley [Sta86] recense 66 classes combinatoires pour lesquelles il y a c_n objets de taille n . Les chemins de Dyck, les suites quasi-décroissantes et les demi-pyramides de dominos sont trois exemples. La multitude d'objets et les symétries naturelles de ces objets sont la source d'un grand nombre de bijections préservant la taille. On se propose dans ce chapitre d'étudier certaines de ces bijections. Une première approche serait de s'intéresser par exemple au $1!1!2!5!14!42!$ bijections préservant la taille entre les suites quasi-décroissantes de longueur au plus 5 et les chemins de Dyck de taille au plus 5. Ce n'est pas celle que l'on adopte : on cherche des bijections "simples", celles que l'on peut trouver dans la littérature pour expliquer une égalité combinatoire. La notion de simplicité pour une bijection ne veut pas dire grand chose. Pour s'amuser conceptuellement, on pourrait qualifier de simple toute bijection pour laquelle il existe un programme pour une machine de Turing qui, prenant en entrée un codage d'un objet combinatoire, retournerait un codage de son image par la bijection. On pourrait alors s'interroger rapidement sur les applications concrètes pour la combinatoire avant de passer à autre chose...

On ambitionne plus modestement dans ce chapitre de construire un peu artificiellement des parentes d'une bijection déjà connue entre les suites quasi-décroissantes et les chemins de Dyck. Cette bijection est codable par un algorithme proche de la dérivation d'un mot à partir d'une grammaire algébrique. L'algorithme est paramétrable par différentes règles de réécriture. On restreint notre étude à certaines règles de réécriture qui garantissent que l'algorithme paramétré par l'une d'elle codera une application entre les suites quasi-décroissantes et les chemins de Dyck. Dans la première partie de ce chapitre, on prouve

que certaines de ces règles de réécriture conduisent aux mêmes applications. D'autre part il existe une symétrie sur les mots de Dyck et une sur les suites quasi-décroissantes qui se traduisent syntaxiquement sur les règles de réécriture. Ces observations permettent de simplifier la preuve que seules 12 applications bijectives distinctes sont induites par les 210 règles de réécriture que l'on peut utiliser comme paramètre de l'algorithme. De plus chacune d'elles donne un paramètre sur les chemins de Dyck qui a la même distribution que la longueur de la dernière descente. Certains de ces paramètres étaient déjà connus, d'autres peut-être un peu moins.

Dans la seconde partie de ce chapitre, on donne des exemples, souvent issus de la littérature, où apparaissent des bijections appartenant à des généralisations de l'algorithme paramétré par des règles de réécriture. Tout d'abord la symétrie sur les suites quasi-décroissantes est liée à une bijection entre les demi-pyramides strictes et les chemins de Motzkin [Sha]. Ensuite, pour l'énumération des chemins de Dyck, l'équation fonctionnelle écrite dans le cadre de la méthode du noyau [BMP00] peut aussi se résoudre en utilisant le développement de Taylor selon la variable catalytique. Cette résolution passe par un système linéaire liant les dérivées partielles. On propose une interprétation combinatoire de ce système à l'aide de la hauteur de rayon, un (nouveau ?) paramètre sur les chemins de Dyck. On éclaire sa résolution par une bijection entre les chemins de Dyck de hauteur de rayon au plus n et les chemins de Dyck de hauteur (usuelle) $2n + 1$. Cette bijection se code en paramétrant l'algorithme par deux règles de réécriture et en utilisant l'une ou l'autre en fonction de la parité de l'entier de la suite quasi-décroissante. D'autre part, pour interpréter une formule, Haiman et Haglund [Hag03] ont défini des paramètres similaires à l'aire d'un chemin de Dyck. On montre des bijections préservant ces paramètres en utilisant d'autres variantes de l'algorithme paramétré. Ensuite, on donne une relation entre une bijection de Delest et Viennot [DV84] entre les chemins de Dyck et les polyominos parallélogrammes, la réflexion par rapport à la première diagonale des polyominos parallélogrammes et une involution de Kreweras [Lal93] sur les chemins de Dyck. Enfin, on présente une bijection qui n'est pas exprimable par les différentes bijections présentées.

2.1 Une première famille d'applications définies par insertion

2.1.1 Les bijections "classiques"

Dans ce chapitre un *alphabet* A est un ensemble de lettres qui pourra éventuellement être infini. Un *mot* sur l'alphabet A est une suite finie de lettres de A . On note A^* l'ensemble des mots sur l'alphabet A . La longueur du mot u est le nombre de lettres du mot u ; on la note $|u|$. Pour un mot $u \in A^*$ et une lettre $a \in A^*$, on note $|u|_a$ le nombre d'occurrences de la lettre a dans le mot u . Un mot v est un préfixe du mot u si les $|v|$ premières lettres de u forment le mot v . C'est un *suffixe* si les $|v|$ dernières lettres de u forment le mot v .

Un mot u sur l'alphabet $\{x, \bar{x}\}$ est un *mot de Dyck* si $|u|_x = |u|_{\bar{x}}$ et que pour tout préfixe v de u , $|v|_x \leq |v|_{\bar{x}}$. La figure 2.1 contient un exemple de mot de Dyck et illustre la représentation graphique qui suit. Les mots sur l'alphabet $\{x, \bar{x}\}$ s'interprètent graphi-

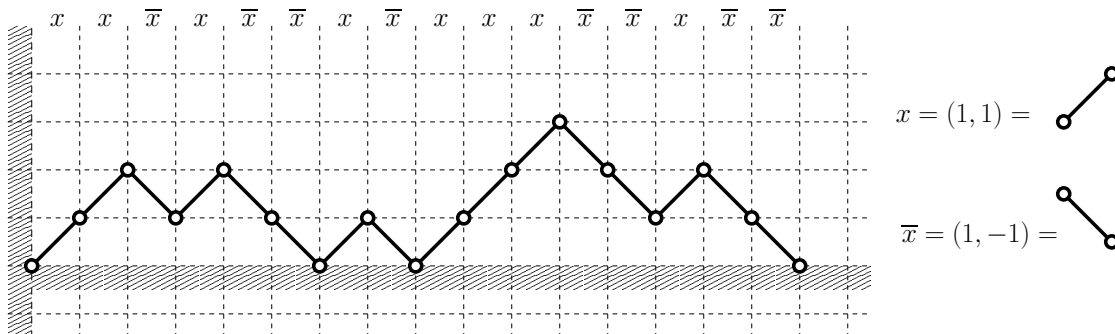


FIG. 2.1 – Un mot de Dyck et son chemin de Dyck associé

quement par un chemin dans le plan formé de pas élémentaires dans $\{(1, 1), (1, -1)\}$. Le chemin associé au mot u est le chemin issu de l'origine du plan dont la suite de pas élémentaires s'obtient à partir des lettres de u en associant $(1, 1)$ à la lettre x et $(1, -1)$ à la lettre \bar{x} . Un *chemin de Dyck* est un chemin issu de l'origine, formé de pas élémentaires $(1, 1)$ et $(1, -1)$, qui se termine sur l'axe des abscisses et ne passe jamais en dessous. En remplaçant les pas élémentaires par les lettres comme ci-dessus on obtient une bijection avec les mots de Dyck. Un *pas montant*, respectivement *descendant*, est un pas $(1, 1)$, respectivement $(1, -1)$.

Les notions de chemins ou de mot de Dyck seront utilisées indifféremment dans ce chapitre. Par exemple, la taille d'un chemin de Dyck u est sa *demi-longueur* $|u|/2 = |u|_x$. (on a utilisé la notion de mot pour définir la taille du chemin)

Une *suite quasi-décroissante* $s = s_1, \dots, s_n$, de taille n , est une suite d'entiers positifs tels que $s_1 = 0$ et pour tout $1 < i \leq n$, $s_i \in [0, s_{i-1} + 1]$.

Un exemple de suite quasi-décroissante est

$$0, 1, 2, 3, 2, 3, 2, 2, 0, 1, 2.$$

Pour les connaisseurs, cette notion de suite est proche de la notion de mots de Lukasiewicz lus à l'envers.

Les *empilements de dominos* ont été présentés dans le chapitre précédent. Ce sont les classes de commutation des suites finies d'entiers selon les relations de commutation $ij = ji$ pour tout i, j tels que $|i - j| \geq 2$. Un entier est *minimal* dans un empilement s'il est le premier élément d'une des suites de la classe de commutation. Les *demi-pyramides de dominos* sont les empilements de dominos dont tous les éléments sont des entiers positifs ou nuls et l'unique éventuel entier minimal est 0. La taille d'une demi-pyramide est son nombre de dominos.

Ces trois classes combinatoires sont énumérées selon leur taille par la suite de Catalan. De plus il existe déjà dans la littérature des bijections entre ces objets. La figure 2.2 illustre les bijections que l'on décrit dans le paragraphe suivant.

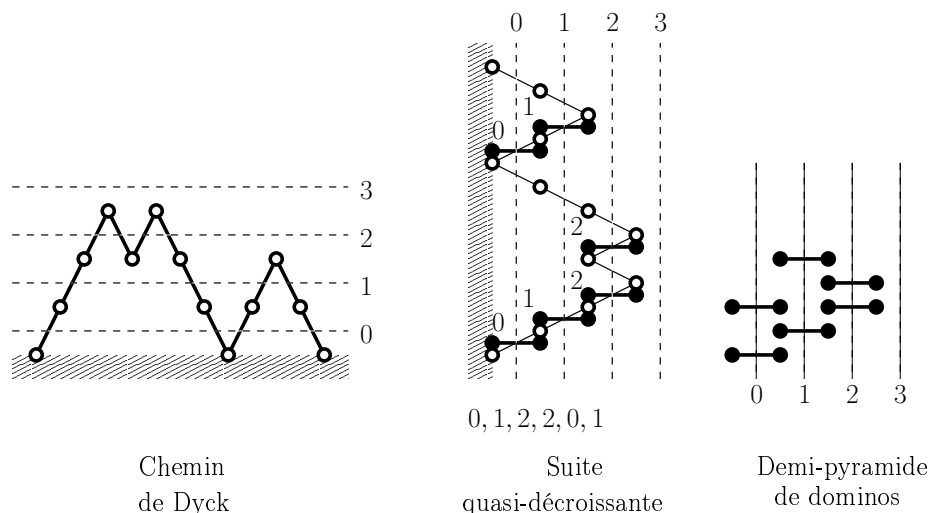


FIG. 2.2 – Bijections entre chemins de Dyck, suites quasi-décroissantes et demi-pyramides de dominos

Pour tout chemin de Dyck, la suite d'entiers formée des hauteurs des sommets origines des pas montants est une suite quasi-décroissante. L'empilement représentant la classe de commutation d'une suite quasi-décroissante s est une demi-pyramide. Cette demi-pyramide s'obtient aussi en faisant tomber des pièces une par une : le i^{eme} domino ajouté est dans la colonne s_i . Étant donné une demi-pyramide on retrouve l'ordre dans lequel sont tombés les dominos dans la construction précédente en retirant récursivement de l'empilement la pièce minimale dans la colonne de plus grand indice. La *dernière descente* d'un chemin de Dyck est le nombre de pas descendants qui ne sont pas suivis d'un pas montant. Pour un chemin non vide u , cela correspond au k du suffixe $x\bar{x}^k$ de u . Partant du chemin vide, on construit le chemin de Dyck associé à la suite quasi-décroissante s en insérant successivement des pics, des facteurs $x\bar{x}$, dans la dernière descente. À la i^{eme} étape, on insère le pic au niveau du sommet à la hauteur s_i de la dernière descente : le suffixe $x\bar{x}^{s_{i-1}+1}$ devient donc $x\bar{x}^{s_{i-1}+1-s_i}x\bar{x}^{s_i+1}$.

La dernière bijection consiste à construire un objet de taille $n + 1$ à partir d'un objet de taille n et d'un atome (de taille 1). Cette approche a été formalisée par Barucci, Del Lungo, Pergola et Pinzani sous le nom de la méthode ECO [BDLPP99]. Dans ce contexte, il y a un arbre sous-jacent dont les sommets de niveau n sont étiquetés par les objets de taille n . En général, la règle d'insertion de l'atome permettant de déterminer un fils dans cet arbre est donnée. Dans ce chapitre on s'intéresse à la façon de découvrir et de coder des règles d'insertions des atomes (les pics) dans les objets (les chemins de Dyck).

2.1.2 Un algorithme pour des applications parentes d'une de ces bijections

La bijection associant un chemin de Dyck à une suite quasi-décroissante peut se décrire autrement. Pour tout entier positif, on nomme *marques* les lettres i et \bar{i} ; i est la *marque montante* d'indice i et \bar{i} est la *marque descendante*. On note M l'alphabet des marques $\bigcup_{i \geq 0} \{i, \bar{i}\}$. Un *mot de Dyck marqué* est un mot sur l'alphabet $\{x, \bar{x}\} \cup M$ tel que le mot obtenu en supprimant les lettres de M soit un mot de Dyck.

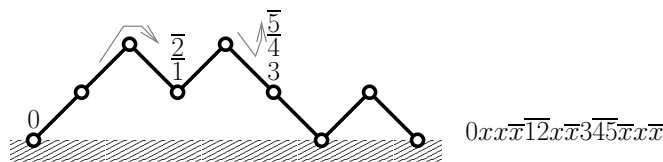


FIG. 2.3 – Représentation graphique d'un chemin marqué

On va utiliser ces chemins marqués comme des objets intermédiaires dans la construction de la bijection entre les suites quasi-décroissantes et les chemins de Dyck. Cette bijection se construit en insérant un pas montant et un pas descendant dans un chemin pour chaque entier de la suite quasi-décroissante. La position où ces pas sont insérés dépend de la valeur lue dans la suite, c'est là qu'interviennent les marques : la marque i , respectivement \bar{i} , indique la position où sera inséré le pas montant, respectivement descendant, si dans la suite quasi-décroissante on lit l'entier i . On ne manipulera que des chemins qui contiennent au plus une marque i et une marque \bar{i} . Après avoir inséré les deux pas, il faut réactualiser les marques pour l'insertion suivante. L'algorithme suivant donne une stratégie possible pour cette réactualisation.

Entrée : une suite quasi-décroissante $s = s_1, \dots, s_n$;

CheminMarque := $0\bar{0}$;

pour i allant de 1 à n **faire**

Supprimer dans *CheminMarque* les marques d'indice strictement supérieur à s_i ;

Substituer dans *CheminMarque* $x.s_i + 1.\overline{s_i + 1}$ à la marque s_i et $\bar{x}.s_i.\bar{s}_i$ à la marque \bar{s}_i ;

fin

Supprimer dans *CheminMarque* toutes les marques.;

Sortie : *CheminMarque*;

Algorithme 1: Codage de la bijection de Viennot par des réécritures de chemins marqués

Pour illustrer le fonctionnement de cet algorithme, on propose d'observer l'exécution de celui-ci avec la suite quasi-décroissante $0, 1, 0, 1, 2, 1$. Sur la figure 2.4, on a représenté les valeurs successives de la variable *CheminMarque* après la lecture de chacun des entiers de la suite.

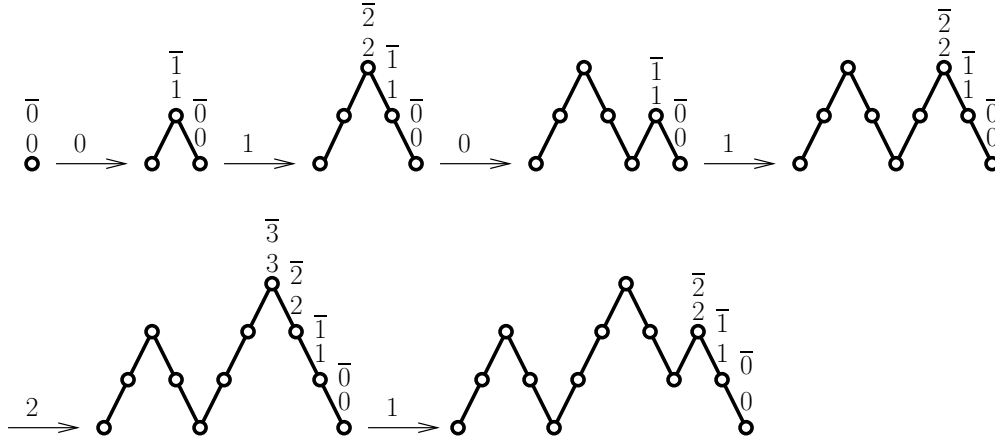


FIG. 2.4 – Constructions de l’algorithme pour la suite 0, 1, 0, 1, 2, 1

Ce codage de la bijection peut sembler compliquer sa compréhension, mais l’intérêt est que l’on peut généraliser la bijection sous cette forme. Un mot w est un *mélange* de deux mots u et v s’il existe une partition (U, V) de l’ensemble des positions de ses lettres telle que u , respectivement v , est le mot formé par les lettres de w dont la position est dans U , respectivement dans V . On note $mel(u, v)$ l’ensemble des mots qui s’obtiennent par mélange de u et v . Un couple de mots (w_1, w_2) est un *mélange de couple* du couple de mots (u_1, u_2) et du mot v si w_1, w_2 est un mélange de mots de u_1, u_2 et v où pour le mélange, la virgule est considérée comme une lettre. Par exemple pour le couple (a, bc) et le mot DE on a

$$\begin{aligned}
 mel((a, bc), DE) = & \{(a, bcDE), (a, bDcE), (a, DbcE), (aD, bcE), (Da, bcE), \\
 & (a, bDEc), (a, DbEc), (aD, bEc), (Da, bEc), (a, DEbc), \\
 & (aD, Ebc), (Da, Ebc), (aDE, bc), (DaE, bc), (DEa, bc)\}.
 \end{aligned}$$

L’ensemble $\{A, \bar{A}, B, \bar{B}\}$ est l’alphabet des *marques génériques*. Un *mode d’insertion* m est un couple de mots qui est un mélange de (x, \bar{x}) , $A\bar{A}$ et $B\bar{B}$. Les couples de mots $(x\bar{B}\bar{B}, \bar{x}A\bar{A})$ et $(Ax, B\bar{x}\bar{B}A)$ sont deux exemples de modes d’insertion.

Un mode d’insertion $m = (m_1, m_2)$ peut être utilisé comme un paramètre d’une généralisation de l’algorithme 1. À la lecture de la valeur s_i dans la suite décroissante on insère m_1 à la place de la marque s_i et m_2 à la place de la marque \bar{s}_i . Ensuite les marques génériques A, \bar{A}, B, \bar{B} deviennent respectivement les marques $s_i, s_i + 1, \bar{s}_i$ et $\bar{s}_i + 1$. Avant de donner l’algorithme avec comme paramètre un mode d’insertion quelconque, on observe d’abord que l’algorithme 2 utilisant le mode $(x\bar{B}\bar{B}, \bar{x}A\bar{A})$ correspond à l’algorithme 1.

Cette version de l’algorithme illustre le rôle des marques d’indices génériques A ou B . Une fois ce mécanisme assimilé on peut donner l’algorithme général 3 utilisant un mode d’insertion quelconque. Cet algorithme est qualifié "d’insertion uniforme" car toutes les réécritures utilisent le même mode d’insertion. Toutes les opérations de l’algorithme portent sur la variable *CheminMarque*.

Entrée : Une suite quasi-décroissante $s = s_1, \dots, s_n$;
 Paramètre : Le mode d'insertion $m = (xB\bar{B}, \bar{x}A\bar{A})$;
CheminMarque := $0\bar{0}$;
pour i allant de 1 à n **faire**
 Supprimer dans *CheminMarque* les marques d'indices supérieur à s_i ;
 Substituer dans *CheminMarque* $xB\bar{B}$ à la marque s_i et $\bar{x}A\bar{A}$ à la marque \bar{s}_i ;
 Substituer dans *CheminMarque* les marques s_i, \bar{s}_i, s_{i+1} et \bar{s}_{i+1} aux marques
 génériques A, \bar{A}, B, \bar{B} ;
fin
 Supprimer les marques de *CheminMarque*;
 Sortie : Le chemin *CheminMarque*;

Algorithme 2: Un premier exemple de paramétrage par un mode d'insertion

Entrée : Une suite quasi-décroissante $s = s_1, s_2, \dots, s_n$;
 Paramètre : Le mode d'insertion $m = (m_1, m_2)$;
CheminMarque := $0\bar{0}$;
pour i allant de 1 à n **faire**
 Supprimer les marques d'indice supérieur à s_i ;
 Substituer m_1 , respectivement m_2 , à la marque s_i , respectivement \bar{s}_i ;
 Substituer les marques d'indices s_i , respectivement $s_i + 1$, aux marques d'indice
 générique A , respectivement B ;
fin
 Supprimer les marques de *CheminMarque*;
 Sortie : Le chemin *CheminMarque* ;

Algorithme 3: Algorithme général avec insertion uniforme

Il est temps de justifier le choix de cet algorithme et celui des modes d'insertions.

Proposition 2.1 *L'algorithme d'insertion uniforme 3 paramétré par un mode d'insertion m donne une application entre les suites quasi-décroissantes de longueur n et les chemins de Dyck de demi-longueur n .*

La preuve est basée sur une récurrence utilisant un sous-ensemble des chemins de Dyck marqués. Une *construction de Dyck de rang k* est un chemin de Dyck marqué qui est un mélange d'un mot de Dyck et des pics $i\bar{i}$ pour tout indice i inférieur ou égal à k .

Preuve. On prouve par récurrence sur la longueur n de la suite quasi-décroissante $s = s_1, \dots, s_n$ passée en paramètre que avant la suppression de toutes les marques, la variable *CheminMarque* est une construction de Dyck de rang $s_n + 1$, ou bien de rang 0 si s est vide. De plus le chemin de Dyck de cette construction est de demi-longueur n .

Pour $n = 0$, la suite s est vide et la boucle **Pour** ne contient pas d'instruction donc *CheminMarque* = $0\bar{0}$ est bien une construction de rang 0.

Pour $n = 1$, le seul élément de la suite s est 0. *CheminMarque* = $0\bar{0}$ donc il contient le

pic $0\bar{0}$. Comme le mode d'insertion $m = (m_1, m_2)$ est tel que $m_1 m_2$ soit une construction de rang 1, l'hypothèse de récurrence est satisfaite.

Pour $n > 1$, soit $s = s_1, \dots, s_{n-1}, s_n$ et on suppose que la propriété est vraie au rang $n - 1$. Par hypothèse, après les $n - 1$ premiers passages dans la boucle **Pour**, la variable *CheminMarque* est une construction de rang $s_{n-1} + 1$ dont le chemin de Dyck est de longueur $n - 1$. Le rang de cette construction correspond exactement à la plus grande valeur possible pour s_n donc quelle que soit sa valeur il existe une seule marque s_n et une seule marque \bar{s}_n . Cette remarque garantit que la suppression des marques d'indice strictement supérieur à s_n conduit à une construction de rang s_n . D'après l'hypothèse de récurrence, la marque s_n apparaît avant la marque \bar{s}_n dans cette construction donc la réécriture de ces marques par m_1 et m_2 conduit bien à un mélange de $m_1 m_2$ et d'une construction de rang $s_n - 1$. C'est aussi un mélange du mot de Dyck de demi-longueur $n - 1$, du pic $x\bar{x}$ apparaissant dans $m_1 m_2$, des pics $i\bar{i}$ et des pics génériques $A\bar{A}$ et $B\bar{B}$. Le mélange des mots sur $\{x, \bar{x}\}$ conduit à un chemin de Dyck de demi-longueur n et l'évaluation des indices générique A , respectivement B , en s_n , respectivement $s_n + 1$, donne une construction de rang $s_n + 1$. \square

2.1.3 Étude des équivalences et symétries de ces applications

On note Φ_m l'application définie par l'algorithme paramétré par le mode d'insertion m . Il existe $\frac{7!}{3!2!2!} = 210$ modes d'insertions distincts. Cependant, certains de ces modes induisent la même application, on parle de *modes équivalents* et on note \equiv_{mode} la relation d'équivalence. On peut observer une partie de ces équivalences directement sur les modes d'insertion.

Une *marque montante isolée* dans le mode $m = (m_1, m_2)$ est une marque générique $Y \in \{A, B\}$ telle que $m_1 = Yx$. Dans ce cas, les modes (Yx, m_2) et (xY, m_2) sont dits *montée-équivalents* et on note cette relation \equiv^{montee} . Symétriquement, une *marque descendante isolée* est une marque générique $\bar{Y} \in \{\bar{A}, \bar{B}\}$ telle que $m_2 = \bar{x}\bar{Y}$. Dans ce cas, les modes $(m_1, \bar{x}\bar{Y})$ et $(m_1, \bar{Y}\bar{x})$ sont *descente-équivalents* et on note cette relation $\equiv^{descente}$. Un *mode de pic* est un mode $m = (m_1, m_2)$ tel que dans la concaténation $m_1 m_2$ apparaissent les facteurs $A\bar{A}$ et $B\bar{B}$. Deux modes de pics (m_1, m_2) et (m'_1, m'_2) tels que $m_1 m_2 = m'_1 m'_2$ sont *pic-équivalents* et on note cette relation \equiv^{pic} .

Proposition 2.2 *Deux modes qui sont montée-équivalents, descente-équivalents ou pic-équivalents induisent la même application lorsqu'ils sont paramètres de l'algorithme de réécriture uniforme 3.*

Preuve.

- $m \equiv_{montee} m' \implies m \equiv_{mode} m'$: Puisque les insertions se font toujours selon le même mode, la marque Y apparaissant lors d'une réécriture selon le mode (Yx, m_2) sera supprimée ou bien conduira à l'insertion d'un certain nombre de pas montants *avant* le

pas x et d'aucun pas descendant. De même, le mode (xY, m_2) conduira à l'insertion de pas montants *après* le pas x . Comme en termes de chemins de Dyck il est équivalent d'insérer un pas montant juste avant ou juste après un pas montant, cela donne toujours le facteur xx . Ainsi les modes (Yx, m_2) et (xY, m_2) donneront les mêmes chemins de Dyck.

• $m \equiv_{descente} m' \implies m \equiv_{mode} m'$: Cette démonstration est symétrique de la précédente.

• $m \equiv_{pic} m' \implies m \equiv_{mode} m'$: On démontre par récurrence sur la longueur des suites quasi-décroissantes que les marques i et \bar{i} sont consécutives dans les constructions de Dyck obtenues par un mode $m = (m_1, m_2)$ tel que $A\bar{A}$ et $B\bar{B}$ sont des facteurs de $m_1 m_2$. Dans ce cas la positionnement de la virgule dans $m_1 m_2$ n'as pas d'influence sur le résultat puisque les deux marques réécrites sont toujours voisines et dans le même ordre.

□

Un exemple d'application de cette proposition est la preuve que les applications $\Phi_{(xB\bar{B}, \bar{x}A\bar{A})}$, $\Phi_{(xB, \bar{B}\bar{x}A\bar{A})}$, $\Phi_{(x, B\bar{B}\bar{x}A\bar{A})}$ et $\Phi_{(Bx, \bar{B}\bar{x}A\bar{A})}$ sont égales : les modes des trois premières applications sont pic-équivalents et le mode de la quatrième est montée-équivalent au mode de la troisième.

Il existe d'autre part des symétries sur les suites quasi-décroissantes et les chemins de Dyck, qui donnent des relations entre différentes modes d'insertion.

Pour tout mot u sur l'alphabet $\{x, \bar{x}\} \cup \bigcup_{i \geq 0} \{i, \bar{i}\}$, son *mot miroir* $mir(u)$ est défini récursivement par

$$\begin{cases} mir(\epsilon) & = \epsilon \\ mir(au) & = mir(u)\bar{a} \quad \forall a \in \{x\} \cup \bigcup_{i \geq 0} i \\ mir(\bar{a}u) & = mir(u)a \quad \forall \bar{a} \in \{\bar{x}\} \cup \bigcup_{i \geq 0} \bar{i}. \end{cases}$$

Pour un mot de Dyck u , le chemin $mir(u)$ est le symétrique de u par rapport à une verticale. Le *mode d'insertion miroir* $mir(m)$ d'un mode d'insertion $m = (m_1, m_2)$ est le mode $(mir(m_2), mir(m_1))$. C'est bien un mode d'insertion.

Pour une suite d'entiers $s = s_1, \dots, s_n$, on note $s \oplus k$ la suite d'entiers $s_1 + k, s_2 + k, \dots, s_n + k$. Toute suite quasi-décroissante se factorise de façon unique sous la forme $0, t_1 \oplus 1, t_2$ où t_1 et t_2 sont des suites quasi-décroissantes. Pour s'en convaincre, on peut remarquer, à l'aide de la figure 2.5, que cette décomposition traduit la décomposition des chemins de Dyck au premier retour sur l'axe : pour t_1 on considère les pas montants du premier facteur et pour t_2 les pas montants du second. On utilise cette décomposition pour définir récursivement la *symétrie d'échange* ech : pour toute suite quasi-décroissante s , $ech(s)$ est la suite quasi-décroissante définie récursivement par

$$\begin{cases} ech(\epsilon) & = \epsilon \\ ech(0, t_1 \oplus 1, t_2) & = 0, ech(t_2) \oplus 1, ech(t_1). \end{cases}$$

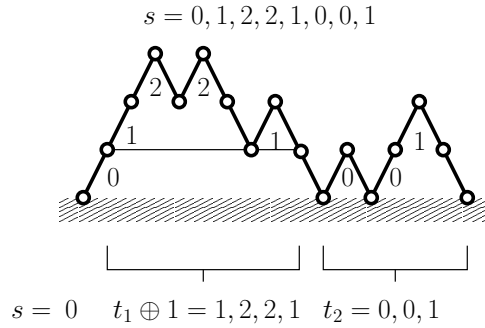


FIG. 2.5 – Factorisation d’une suite quasi-décroissante

Cette application est une bijection car c’est une involution. Ceci se vérifie par récurrence à partir de l’identité :

$$ech(ech(0, t_1 \oplus 1, t_2)) = 0, ech(ech(t_1)) \oplus 1, ech(ech(t_2)).$$

Le *mode d’insertion échangé* $ech(m)$ d’un mode d’insertion m s’obtient en échangeant les indices génériques A et B dans le mode m . Par exemple $ech(Ax, B\bar{B}\bar{x}\bar{A}) = (Bx, A\bar{A}\bar{x}\bar{B})$.

Proposition 2.3 *Pour tout mode d’insertion m on a les égalités*

$$mir \circ \Phi_m = \Phi_{mir(m)}$$

et

$$\Phi_m \circ ech = \Phi_{ech(m)}.$$

La preuve de la seconde relation nécessite une notion supplémentaire. Dans l’algorithme général la variable *CheminMarque* est initialisée à $0\bar{0}$ mais on veut désormais se laisser la possibilité d’initialiser *CheminMarque* à divers chemins marqués. Soit u un chemin marqué, on note Φ_m^u l’application qui à une suite quasi-décroissante s associe un mot sur l’alphabet $\{x, \bar{x}\}$ en initialisant la variable *CheminMarque* au mot marqué u . Comme les marques des constructions de Dyck ne sont pas déplacées, on observe pour tout mode m et toute suite décroissante s , l’égalité $\Phi_m(s) = \Phi_m^0(s)\Phi_m^{\bar{0}}(s)$. D’autre part pour tous mots u, v, w , et toutes lettres a et b on note $u[a := v]$ le mot obtenu en substituant le mot v à toutes les occurrences de la lettre a dans le mot u , $u[a := v, b := w]$ les substitutions *en parallèle* des lettres a et b et enfin $u[a := v][b := w]$ la substitution de la lettre a dans u puis de la lettre b dans $u[a := v]$.

Preuve.

- $mir \circ \Phi_m = \Phi_{mir(m)}$: On constate par récurrence sur la longueur de la suite quasi-décroissante que le miroir de la construction de Dyck selon le mode m est aussi la construction de Dyck selon le mode $mir(m)$.

• $\Phi_m \circ ech = \Phi_{ech(m)}$: On montre par récurrence sur la longueur des suites quasi-décroissantes que pour toute suite s et tout mode m on a $\Phi_m^0(s) = \Phi_{ech(m)}^0(ech(s))$ et $\Phi_m^{\bar{0}}(s) = \Phi_{ech(m)}^{\bar{0}}(ech(s))$. Pour conclure on utilisera le fait que ech est une involution.

Pour $n = 0$, dans tous les cas $\Phi_m^0(\epsilon) = \epsilon = \Phi_{ech(m)}^0(ech(\epsilon))$ et la même chose pour $\Phi_m^{\bar{0}}$.

Pour $n = 1$, dans tous les cas $\Phi_m^0(0) = x = \Phi_{ech(m)}^0(ech(0))$ et $\Phi_m^{\bar{0}}(0) = \bar{x} = \Phi_{ech(m)}^{\bar{0}}(ech(0))$.

Pour $n > 1$, soit s une suite quasi-décroissante de longueur n . On calcule $\Phi_m^0(s)$ et $\Phi_{ech(m)}^0(ech(s))$ en observant la factorisation $s = 0, t_1 \oplus 1, t_2$. On note u la construction obtenue après la lecture selon le mode m de 0, le chemin *CheminMarque* étant initialisé à 0. On a

$$\Phi_m^0(s) = u[1 := \Phi_m^1(t_1 \oplus 1), \bar{1} := \Phi_m^{\bar{1}}(t_1 \oplus 1), 0 := \Phi_m^0(t_2), \bar{0} := \Phi_m^{\bar{0}}(t_2)].$$

Par définition de $ech(m)$, la construction obtenue après la lecture selon le mode $ech(m)$ de 0 est

$$u[0 := 1, \bar{0} := \bar{1}, 1 := 0, \bar{1} := \bar{0}].$$

On a donc

$$\Phi_{ech(m)}^0(ech(s)) = u[0 := 1, 1 := 0, \dots][1 := \Phi_{ech(m)}^1(ech(t_2) \oplus 1), 0 := \Phi_{ech(m)}^0(ech(t_1)), \dots].$$

La valeur des indices n'influence pas les insertions; on a donc l'égalité

$$\Phi_{ech(m)}^1(ech(t_2) \oplus 1) = \Phi_{ech(m)}^0(ech(t_2)).$$

En utilisant l'hypothèse de récurrence sur t_2 on obtient

$$\Phi_{ech(m)}^0(ech(t_2)) = \Phi_m^0(t_2).$$

Un raisonnement utilisant les mêmes arguments permet d'obtenir l'égalité

$$\Phi_{ech(m)}^0(ech(t_1)) = \Phi_m^1(t_1 \oplus 1).$$

Le même type de relation existe pour $\Phi_{ech(m)}^{\bar{0}}(ech(t_2) \oplus 1)$ et $\Phi_{ech(m)}^{\bar{1}}(t_1)$. On obtient donc la relation

$$\Phi_{ech(m)}^0(ech(s)) = u[0 := 1, 1 := 0, \dots][1 := \Phi_m^1(t_2), 0 := \Phi_m^0(t_1 \oplus 1), \dots].$$

On observe que les substitutions se compensent bien pour redonner l'expression de $\Phi_m^0(s)$. On a donc

$$\Phi_{ech(m)}^0(ech(s)) = \Phi_m^0(s).$$

Par un raisonnement similaire on obtient l'égalité

$$\Phi_{ech(m)}^{\bar{0}}(ech(s)) = \Phi_m^{\bar{0}}(s)$$

ce qui termine la récurrence.

Pour conclure on remarque que

$$\Phi_{ech(m)}(ech(s)) = \Phi_{ech(m)}^0(ech(s))\Phi_{ech(m)}^{\bar{0}}(ech(s)) = \Phi_m^0(s)\Phi_m^{\bar{0}}(s) = \Phi_m(s),$$

et on applique ce résultat à $ech(s)$ en se souvenant que ech est une involution.

□

L'usage intensif de l'ordinateur permet de trouver des contre-exemples pour prouver que 178 applications induites par des modes d'insertions ne sont pas bijectives. Les équivalences de la proposition 2.2 et les symétries de la proposition 2.3 réduisent la preuve de la "bijectivité" des 32 modes d'insertions restant à celle de trois d'entre eux.

Lemme 2.3.1 *Parmi les 210 applications définies par les 210 modes d'insertions il y en a 178 qui ne sont pas des bijections avec un contre-exemple de taille inférieure ou égale à 3. Parmi les 32 applications restantes, il y a 12 applications distinctes et il n'y a pas de contre-exemple à leur "bijectivité" de taille inférieure à 11. Une condition suffisante de la "bijectivité" de ces 12 applications est la bijectivité des applications définies par les 3 modes d'insertion suivant :*

$$(x\overline{BB}, \overline{x}A\overline{A}),$$

$$(A\overline{B}Bx, \overline{x}A),$$

$$(B\overline{A}Bx, \overline{A}x).$$

Preuve. Un ordinateur a produit les 178 contre-exemples à la bijectivité de 178 modes d'insertions. On ne reproduit pas dans ce document les résultats mais les contre-exemples sont tous de taille inférieure à 3.

Les 32 modes restant définissent 12 applications différentes compte tenu des relations de la proposition 2.2. Dans la figure 2.6, on donne les modes d'insertions de chacune de ces applications et les symétries liant ces modes.

Il y a 3 classes d'applications reliées par les symétries miroir et d'échange.

Or mir est une involution sur les chemins de Dyck donc d'après la proposition 2.3, $\Phi_{mir(m)}$ est une bijection si et seulement si Φ_m est une bijection. De même ech est une involution sur les suites quasi-décroissantes donc d'après la proposition 2.3, $\Phi_{ech(m)}$ est une bijection si et seulement si Φ_m est une bijection. Pour prouver que les 12 applications restantes sont bijectives il suffit donc d'établir la bijectivité d'un représentant de chacune des trois classes.

□

$$\begin{array}{ccc}
\left\{ \begin{array}{l} \equiv_{pic} (xB\bar{B}, \bar{x}A\bar{A}) \\ \equiv_{pic} (x, B\bar{B}\bar{x}A\bar{A}) \\ \equiv_{pic} (xB, \bar{B}, \bar{x}A\bar{A}) \\ \equiv_{montee} (Bx, \bar{B}, \bar{x}A\bar{A}) \end{array} \right. & \iff ech & \left\{ \begin{array}{l} \equiv_{pic} (xA\bar{A}, \bar{x}B\bar{B}) \\ \equiv_{pic} (x, A\bar{A}\bar{x}B\bar{B}) \\ \equiv_{pic} (xA, \bar{A}, \bar{x}B\bar{B}) \\ \equiv_{montee} (Ax, \bar{A}, \bar{x}B\bar{B}) \end{array} \right. \\
\Downarrow_{mir} & & \Downarrow_{mir} \\
\left\{ \begin{array}{l} \equiv_{pic} (A\bar{A}x, B\bar{B}\bar{x}) \\ \equiv_{pic} (A\bar{A}xB\bar{B}, \bar{x}) \\ \equiv_{pic} (A\bar{A}xB, \bar{B}\bar{x}) \\ \equiv_{descente} (A\bar{A}xB, \bar{x}\bar{B}) \end{array} \right. & \iff ech & \left\{ \begin{array}{l} \equiv_{pic} (B\bar{B}x, A\bar{A}\bar{x}) \\ \equiv_{pic} (B\bar{B}xA\bar{A}, \bar{x}) \\ \equiv_{pic} (B\bar{B}xA, \bar{A}\bar{x}) \\ \equiv_{descente} (B\bar{B}xA, \bar{x}\bar{A}) \end{array} \right. \\
\\
\left\{ \begin{array}{l} \equiv_{descente} (AB\bar{B}x, \bar{x}\bar{A}) \\ \equiv_{descente} (AB\bar{B}x, \bar{A}\bar{x}) \end{array} \right. & \iff ech & \left\{ \begin{array}{l} \equiv_{descente} (BA\bar{A}x, \bar{x}\bar{B}) \\ \equiv_{descente} (BA\bar{A}x, \bar{B}\bar{x}) \end{array} \right. \\
\Downarrow_{mir} & & \Downarrow_{mir} \\
\left\{ \begin{array}{l} \equiv_{montee} (Ax, \bar{x}B\bar{B}\bar{A}) \\ \equiv_{montee} (xA, \bar{x}B\bar{B}\bar{A}) \end{array} \right. & \iff ech & \left\{ \begin{array}{l} \equiv_{montee} (Bx, \bar{x}A\bar{A}\bar{B}) \\ \equiv_{montee} (xB, \bar{x}A\bar{A}\bar{B}) \end{array} \right. \\
\\
\left\{ \begin{array}{l} \equiv_{descente} (BA\bar{B}x, \bar{A}\bar{x}) \\ \equiv_{descente} (BA\bar{B}x, \bar{x}\bar{A}) \end{array} \right. & \iff ech & \left\{ \begin{array}{l} \equiv_{descente} (AB\bar{A}x, \bar{B}\bar{x}) \\ \equiv_{descente} (AB\bar{A}x, \bar{x}\bar{B}) \end{array} \right. \\
\Downarrow_{mir} & & \Downarrow_{mir} \\
\left\{ \begin{array}{l} \equiv_{montee} (xA, \bar{x}B\bar{A}\bar{B}) \\ \equiv_{montee} (Ax, \bar{x}B\bar{A}\bar{B}) \end{array} \right. & \iff ech & \left\{ \begin{array}{l} \equiv_{montee} (xB, \bar{x}A\bar{B}\bar{A}) \\ \equiv_{montee} (Bx, \bar{x}A\bar{B}\bar{A}) \end{array} \right.
\end{array}$$

FIG. 2.6 – Les douze applications regroupées en trois classes

2.1.4 Preuve du caractère bijectif de 12 applications

Les structures des preuves de la bijectivité de chacun des trois modes restants sont similaires. On note $\Phi_m^+(s)$ la construction de Dyck obtenue avant l'ultime effacement des marques lors du calcul de $\Phi_m(s)$ selon l'algorithme d'insertion uniforme 3. La clef de toutes les preuves est une *fonction de marquage* qui étant donné un chemin de Dyck est capable de replacer les marques effacées à la dernière étape de l'algorithme 3. La suite de la preuve consiste à vérifier que les réécritures de l'algorithme sont compatibles avec cette fonction de marquage. Un mode d'insertion m est *compatible* avec une fonction de marquage f si pour toute suite quasi-décroissante s on a

$$\Phi_m^+(s) = f(\Phi_m(s)).$$

Une hypothèse supplémentaire sur le mode d'insertion permet de simplifier cette vérification : un mode d'insertion $m = (m_1, m_2)$ est *localement réversible* si chacun des mots m_1 et m_2 contient au moins une marque d'indice générique. Le mode $m_{rev} = (AxB\bar{A}, \bar{B}\bar{x})$ est un exemple de mode localement réversible et $m_{irrev} = (AxB\bar{A}\bar{B}, \bar{x})$ un contre-exemple. L'intérêt des modes d'insertion localement réversibles est que si on connaît la fonction de

marquage f_m et le mode m d'un chemin de Dyck u on sait retrouver les deux derniers pas insérés dans le chemin : il suffit de repérer dans $f_m(u)$ les marquages des deux indices les plus élevés, une marque au moins est voisine des pas insérés et respecte le motif m . Supposons par exemple qu'une insertion selon le mode m_{rev} donne la construction

$$0x\bar{0}x1x2\bar{1}2\bar{x}\bar{x}\bar{x} = 0x\bar{0}x(1x2\bar{1})(\bar{2}\bar{x})\bar{x}\bar{x}.$$

Les marques d'indices 1 et 2 permettent d'identifier les derniers pas insérés. Par contre, l'insertion selon le mode m_{irrev} ne permet pas de retrouver la position du pas descendant qui peut être un des trois derniers pas descendants de cette construction.

La première fonction de marquage considérée dans ce chapitre est celle qui place les marques $i\bar{i}$ sur le sommet à hauteur i dans la dernière descente du chemin de Dyck. On la note $f_{descente}$. On sait déjà que l'application associée au mode $(xB\bar{B}, \bar{x}A\bar{A})$ est compatible

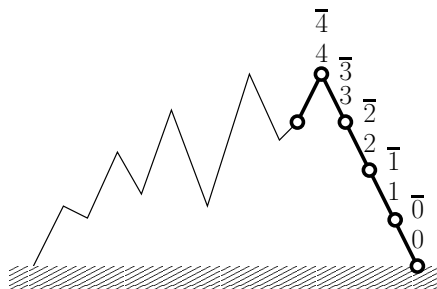


FIG. 2.7 – Fonction de marquage sur la dernière descente

avec ce marquage puisque c'est la bijection que l'on cherche à généraliser. On oublie cette remarque dans la preuve du lemme suivant pour se familiariser avec la technique de preuve.

Lemme 2.3.2 *Le mode d'insertion $(xB\bar{B}, \bar{x}A\bar{A})$ est compatible avec la fonction de marquage $f_{descente}$ et $\Phi_{(xB\bar{B}, \bar{x}A\bar{A})}$ est une bijection.*

Preuve. On prouve cette proposition par récurrence sur la longueur n des suites quasi-décroissantes. Dans cette preuve on note m le mode d'insertion $(xB\bar{B}, \bar{x}A\bar{A})$ et f la fonction de marquage $f_{descente}$.

Hypothèse de récurrence : On suppose que pour toute suite quasi-décroissante de longueur inférieure ou égale à n on a $\Phi_m^+(s) = f(\Phi_m(s))$ et que Φ_m est une bijection lorsqu'elle est restreinte aux suites et chemins de longueur inférieure ou égale à n .

Cas de base : Pour $n = 0$, $\Phi_m^+(s) = 0\bar{0} = f(\epsilon)$ et Φ_m est une bijection qui associe la suite vide au chemin vide.

Hérédité : Soit u un chemin de Dyck de longueur $n + 1$. On cherche par un raisonnement de type analyse/synthèse une suite quasi-décroissante $s = s_1, \dots, s_n, s_{n+1}$ telle que $\Phi_m(s) = u$ et $\Phi_m^+(s) = f(u)$. La figure 2.1.4 illustre le propos.

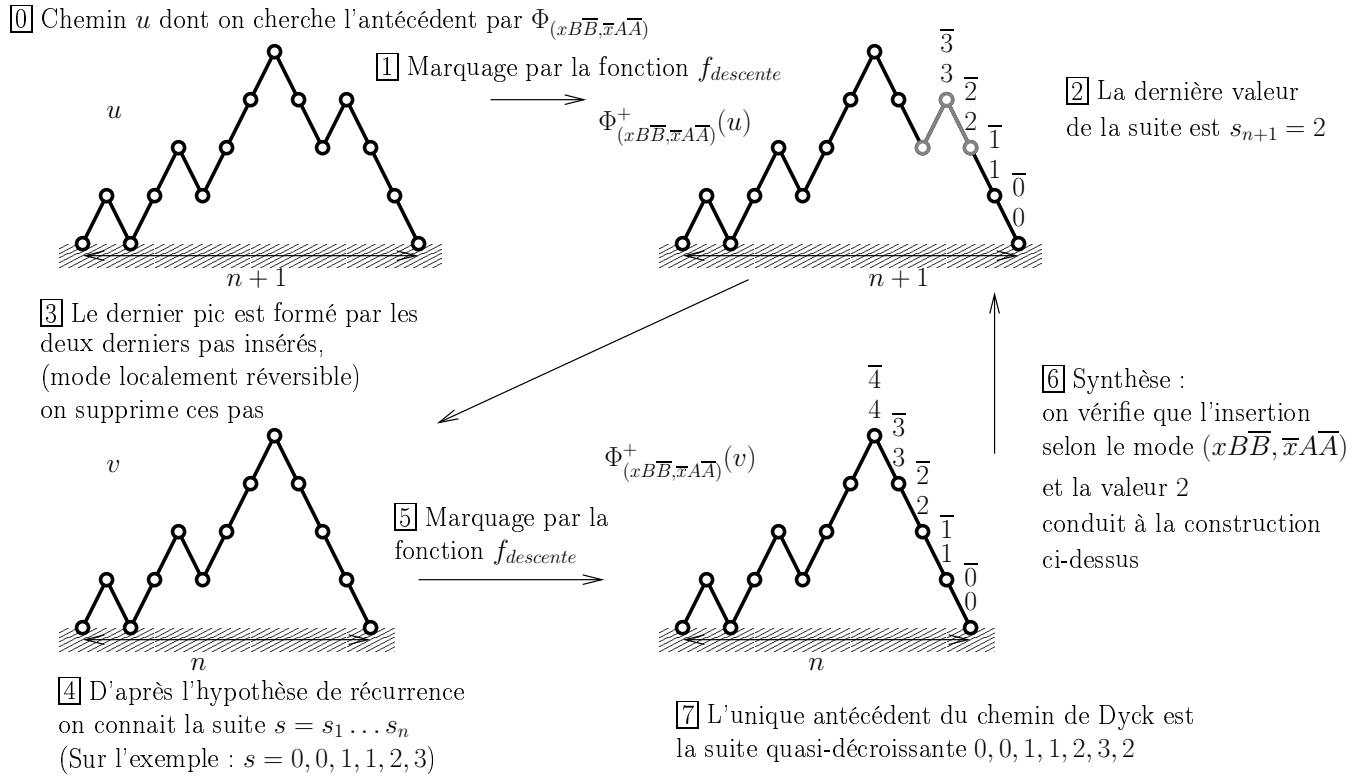


FIG. 2.8 – Un exemple de recherche de l'inverse de $\Phi_{(xB\bar{B}, xA\bar{A})}$

Analyse : Soit s une suite décroissante telle que $\Phi_m(s) = u$ et $\Phi_m^+(s) = f(u)$. On note $l(u)$ la longueur de la dernière descente de u . La construction $f(u)$ est de rang $l(u)$. Puisque $f(u) = \Phi_m^+(s)$, on en déduit que $s_{n+1} = l(u) - 1$. Comme m est localement réversible, on constate que les deux derniers pas insérés selon ce mode pour donner u sont ceux qui forment le dernier pic dont le sommet est à la hauteur $l(u)$. On note v le chemin de Dyck de demi-longueur n obtenu par la suppression du dernier pic de u . D'après l'hypothèse de récurrence, il existe une unique suite décroissante s_1, \dots, s_n telle que $\Phi_m(s_1, \dots, s_n) = v$. D'après cette analyse il existe une unique suite s_1, \dots, s_{n+1} pouvant être associée à u .

Synthèse : On vérifie que la suite dégagée par l'analyse convient. D'après l'hypothèse de récurrence on a $\Phi_m^+(s_1, \dots, s_n) = f(v)$ avec $s_n = l(v) - 1$ où $l(v)$ est la longueur de la dernière descente de v . La suppression du dernier pic de u diminue d'au plus un la longueur de la dernière descente donc $l(v) \geq l(u) - 1$. Cela prouve que s_1, \dots, s_{n+1} est une suite quasi-décroissante puisque s_1, \dots, s_n est une suite quasi-décroissante et que $s_{n+1} = l(u) - 1 \leq l(v) = s_n + 1$. D'après la définition de f , la construction $f(v) = \Phi_m^+(v)$ contient ses deux marques $l(u) - 1$ et $l(u) - 1$

sur le sommet à hauteur $l(u) - 1$ de la dernière descente de v . Pour la construction $\Phi_m^+(v)$, l'insertion selon le mode m à la lecture de la valeur $s_{n+1} = l(u) - 1$ conduit à l'insertion d'un pic dont le sommet est à la hauteur $l(u)$ dans la dernière descente de v . Par définition de v on a donc $\Phi_m(s_1, \dots, s_{n+1}) = u$. De plus les marques effacées de $\Phi_m^+(v)$ lors de l'insertion sont exactement celles qui précèdent le nouveau pic inséré donc $\Phi_m^+(u) = f(u)$.

Pour tout chemin de u de longueur $n + 1$, on a trouvé une unique suite quasi-décroissante s_1, \dots, s_{n+1} telle que $\Phi_m(s_1, \dots, s_{n+1}) = u$ donc Φ_m est bijective lorsqu'elle est restreinte aux chemins de longueur inférieure ou égale à $n + 1$.

□

La seconde fonction de marquage f_{axe} que l'on considère est liée aux retours sur l'axe d'un chemin de Dyck. Un chemin de Dyck se factorise de façon unique en une suite de facteurs $(xu_k\bar{x})_{k=1..K}$ où les u_k sont des chemins de Dyck. Étant donné un chemin de Dyck u dont K sommets sont sur l'axe, la construction $f_{axe}(u)$ est de rang $K - 1$ et s'obtient en insérant les marques montantes dans l'ordre croissant au début du mot, puis la marque descendante $\bar{K} - k$ sur le k^{eme} sommet sur l'axe.

Exemple : Le chemin de Dyck

$$u = x\bar{x}xx\bar{x}\bar{x}\bar{x}xx\bar{x}\bar{x}$$

donne la construction de rang 3

$$f_{axe}(u) = 0123\bar{3}x\bar{x}\bar{2}xx\bar{x}\bar{x}\bar{x}\bar{1}xx\bar{x}\bar{0}.$$

La figure 2.9 illustre aussi cet exemple.

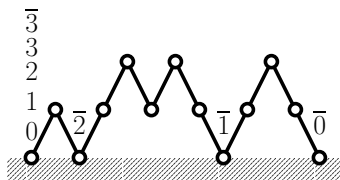


FIG. 2.9 – Fonction de marquage sur l'axe

Lemme 2.3.3 *Le mode d'insertion $(AB\bar{B}x, \bar{x}\bar{A})$ est compatible avec la fonction de marquage f_{axe} et $\Phi_{(AB\bar{B}x, \bar{x}\bar{A})}$ est une bijection.*

Preuve. On prouve cette proposition par récurrence sur la longueur n des suites quasi-décroissantes. Dans cette preuve on note m le mode d'insertion $(AB\bar{B}x, \bar{x}\bar{A})$ et f la fonction de marquage f_{axe} .

Hypothèse de récurrence : On suppose que pour toute suite quasi-décroissante de longueur inférieure ou égale à n on a $\Phi_m^+(s) = f(\Phi_m(s))$ et que Φ_m est une bijection lorsqu'elle est restreinte aux suites et chemins de longueur inférieure ou égale à n . La figure 2.10 illustre le propos.

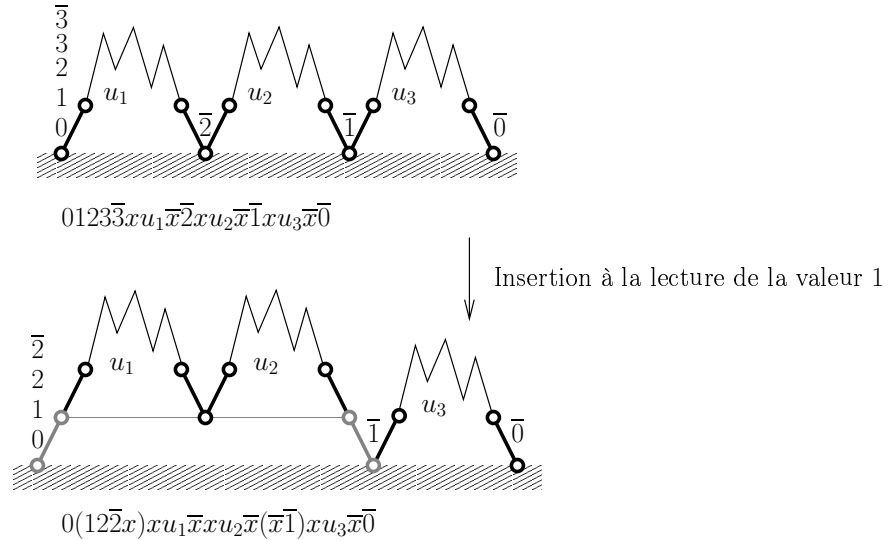


FIG. 2.10 – Fonction de marquage sur l'axe pour le mode $(AB\bar{B}x, \bar{x}\bar{A})$

Cas de base : Pour $n = 0$, $\Phi_m^+(s) = 0\bar{0} = f(\epsilon)$ et Φ_m est une bijection qui associe la suite vide au chemin vide.

Hérédité : Soit u un chemin de Dyck de longueur $n + 1$. On cherche par un raisonnement de type analyse/synthèse une suite quasi-décroissante $s = s_1, \dots, s_n, s_{n+1}$ telle que $\Phi_m(s) = u$ et $\Phi_m^+(s) = f(u)$.

Analyse : Soit s une suite décroissante telle que $\Phi_m(s) = u$ et $\Phi_m^+(s) = f(u)$. On note $c(u)$ le nombre de sommet sur l'axe horizontal du chemin u . On numérote les sommets sur l'axe de u de droite à gauche et de 0 à $c(u) - 1$. La construction $f(u)$ est de rang $c(u) - 1$. Puisque $f(u) = \Phi_m^+(s)$, on en déduit que $s_{n+1} = c(u) - 2$. Comme m est localement réversible, on constate que les deux derniers pas insérés selon ce mode pour donner u sont le premier pas montant et le pas descendant correspondant au retour sur l'axe du chemin sur le sommet numéroté $c(u) - 2$. On note v le chemin de Dyck de demi-longueur n obtenu par la suppression de u de ces deux pas. D'après l'hypothèse de récurrence, il existe une unique suite décroissante s_1, \dots, s_n telle que $\Phi_m(s_1, \dots, s_n) = v$. D'après cette analyse il existe une unique suite s_1, \dots, s_{n+1} pouvant être associée à u .

Synthèse : On vérifie que la suite dégagée par l'analyse convient. D'après l'hypothèse de récurrence on a $\Phi_m^+(s_1, \dots, s_n) = f(v)$ avec $s_n = c(v) - 2$ où $c(v)$ est le nombre de sommet sur l'axe de v . La suppression du premier pas de u et du pas descendant précédant le premier retour sur l'axe diminue d'au

plus un le nombre de sommets sur l'axe de u (dans le cas où u débute par $x\bar{x}$). On a donc la relation $c(v) \geq c(u) - 1$. Cela prouve que s_1, \dots, s_{n+1} est une suite quasi-décroissante puisque s_1, \dots, s_n est une suite quasi-décroissante et que $s_{n+1} = c(u) - 2 \leq c(v) - 1 = s_n + 1$. On numérote les sommets sur l'axe de v de droite à gauche et de 0 à $c(v) - 1$. D'après la définition de f , la construction $f(v) = \Phi_m^+(v)$ contient sa marque s_{n+1} sur le premier sommet du chemin (numéroté $c(v) - 1$) et sa marque $\overline{s_{n+1}}$ sur le sommet numéroté s_{n+1} . Pour la construction $\Phi_m^+(v)$, l'insertion selon le mode m à la lecture de la valeur $s_{n+1} = c(u) - 2$ conduit à l'insertion d'un pas montant au début du chemin et d'un pas descendant aboutissant au sommet numéroté $c(u) - 2$ dans v . Par définition de v on a donc $\Phi_m(s_1, \dots, s_{n+1}) = u$. D'une part, les marques effacées de $\Phi_m^+(v)$ lors de l'insertion sont exactement celles qui correspondent aux sommets sur l'axe de v numéroté de $c(v) - 2$ à $c(u) - 1$ et qui ont été surélevé par l'insertion des deux pas menant à u . D'autre part les marques de pas montants sur le premier sommet du chemin restent ordonnées. Finalement $\Phi_m^+(u) = f(u)$.

Pour tout chemin de u de longueur $n + 1$, on a trouvé une unique suite quasi-décroissante s_1, \dots, s_{n+1} telle que $\Phi_m(s_1, \dots, s_{n+1}) = u$ donc Φ_m est bijective lorsqu'elle est restreinte aux chemins de longueur inférieure ou égale à $n + 1$.

□

La troisième fonction de marquage fait appel à une notion apparemment moins connue que les sommets de la dernière descente ou les sommets sur l'axe. Un *parapluie* p de taille $k \geq 1$ du mot de Dyck d est un facteur de d tel que $d = upv$ avec v ne débutant pas par une descente \bar{x} et $p \in (x(\bar{x}^*))^{k-1}x(\bar{x})^k$. La *pointe* du parapluie est le sommet entre le dernier x de p et les \bar{x}^k . De façon plus géométrique, tout sommet qui est un pic est aussi la pointe d'un parapluie. La taille de ce parapluie est le nombre k de pas descendants qui suivent ce sommet. Le parapluie est formé de ces k pas descendants, des k derniers pas montants qui précèdent la pointe et des éventuels pas descendants qui apparaissent entre ces pas montants et la pointe.

Pour tout chemin de Dyck u il existe un plus long suffixe qui soit la concaténation de parapluies. La *couverture* du chemin de Dyck est le nombre de parapluies qui forment ce suffixe. Les parapluies qui sont concaténés pour former le suffixe sont les *parapluies ouverts* ce qui les distingue des autres parapluies qui pourraient être, par exemple, facteur du suffixe. Le plus long suffixe couvert peut se calculer par un algorithme glouton en ajoutant un parapluie au plus long suffixe trouvé lorsque il est précédé dans le chemin d'un pas descendant. La couverture d'un chemin de Dyck est l'analogue de la longueur de la dernière descente ou du nombre de sommets sur l'axe. Le *sommet critique* d'un chemin de Dyck est le premier sommet du suffixe (cf. figure 2.11).

La fonction de marquage $f_{\text{parapluie}}$ associe à un chemin de Dyck u de couverture k une construction de rang k . Le marquage est illustré sur la figure 2.12 Sur la pointe du dernier parapluie ouvert, le plus à droite, on place la marque $\bar{0}$. Pour $i \geq 2$, sur la pointe du $i^{\text{ème}}$

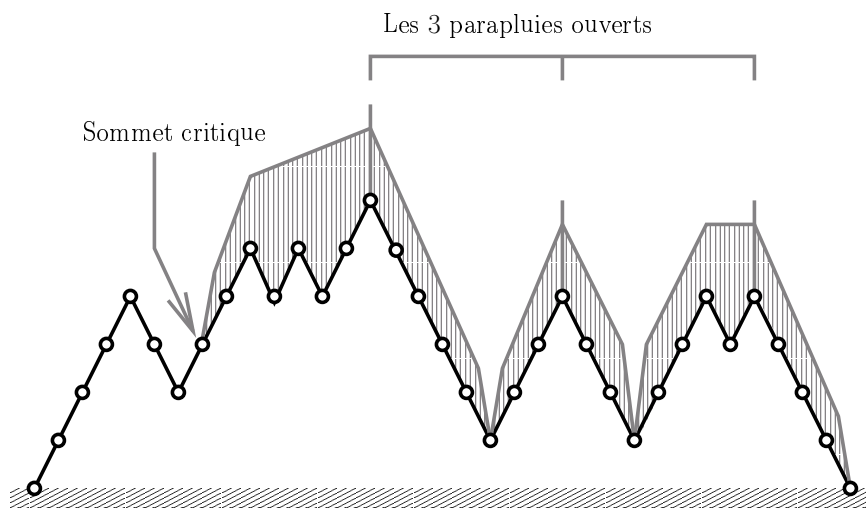


FIG. 2.11 – Un exemple de couverture avec des parapluies

parapluie ouvert en partant de la fin du chemin, on insère le facteur $(i - 2)\overline{i - 1}$. Sur le sommet critique on insère le facteur $k.\overline{k} - \overline{1.k}$.

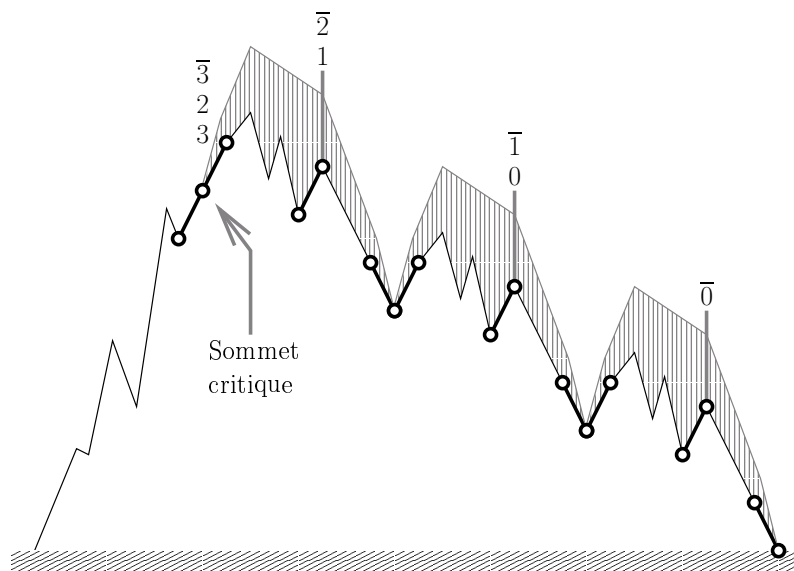


FIG. 2.12 – Fonction de marquage des parapluies pour le mode $(BA\overline{B}x, \overline{A}\overline{x})$

Lemme 2.3.4 *Le mode d'insertion $(BA\overline{B}x, \overline{A}\overline{x})$ est compatible avec la fonction de marquage $f_{\text{parapluie}}$ et $\Phi_{(BA\overline{B}x, \overline{A}\overline{x})}$ est une bijection.*

Preuve. On prouve cette proposition par récurrence sur la longueur n des suites quasi-décroissantes. Dans cette preuve on note m le mode d'insertion $(BA\overline{B}x, \overline{A}\overline{x})$ et f la fonction de marquage $f_{\text{parapluie}}$.

Hypothèse de récurrence : On suppose que pour toute suite quasi-décroissante de longueur inférieure ou égale à n on a $\Phi_m^+(s) = f(\Phi_m(s))$ et que Φ_m est une bijection lorsqu'elle est restreinte aux suites et chemins de longueur inférieure ou égale à n .

Cas de base : Pour $n = 0$, $\Phi_m^+(s) = 0\bar{0} = f(\epsilon)$ et Φ_m est une bijection qui associe la suite vide au chemin vide.

Hérédité : Soit u un chemin de Dyck de longueur $n + 1$. On cherche par un raisonnement de type analyse/synthèse une suite quasi-décroissante $s = s_1, \dots, s_n, s_{n+1}$ telle que $\Phi_m(s) = u$ et $\Phi_m^+(s) = f(u)$.

Analyse : Soit s une suite décroissante telle que $\Phi_m(s) = u$ et $\Phi_m^+(s) = f(u)$. On note $c(u)$ la couverture du chemin u . On numérote les pointes des parapluies ouverts de droite à gauche et de 0 à $c(u) - 1$. La construction $f(u)$ est de rang $c(u)$. Puisque $f(u) = \Phi_m^+(s)$, on en déduit que $s_{n+1} = c(u) - 1$. Comme m est localement réversible, on constate que les deux derniers pas insérés selon ce mode pour donner u appartiennent au parapluie dont la pointe est numérotée $c(u) - 1$: il y a le pas montant le plus à gauche et le pas descendant suivant la pointe. Dans le cas particulier où le parapluie ouvert numéroté $c(u) - 1$ est un simple pic $x\bar{x}$ alors on supprime le pic qui suit le sommet critique et le parapluie ouvert disparaît. On note v le chemin de Dyck de demi-longueur n obtenu par la suppression de u de ces deux pas. D'après l'hypothèse de récurrence, il existe une unique suite décroissante s_1, \dots, s_n telle que $\Phi_m(s_1, \dots, s_n) = v$. D'après cette analyse il existe une unique suite s_1, \dots, s_{n+1} pouvant être associée à u .

Synthèse : On vérifie que la suite dégagée par l'analyse convient. D'après l'hypothèse de récurrence on a $\Phi_m^+(s_1, \dots, s_n) = f(v)$ avec $s_n = c(v) - 1$ où $c(v)$ est la couverture de v . Comme les parapluies dont la pointe est numérotée de 0 à $c(u) - 2$ ne sont pas modifiés lors de la suppression des pas donnant v , on en déduit que la couverture de v est au moins $c(u) - 1$ (l'égalité a lieu lorsque le parapluie numéroté $c(u) - 1$ dans u est un simple pic $x\bar{x}$). On a donc la relation $c(v) \geq c(u) - 1$. Cela prouve que s_1, \dots, s_{n+1} est une suite quasi-décroissante puisque s_1, \dots, s_n est une suite quasi-décroissante et que $s_{n+1} = c(u) - 1 \leq c(v) = s_n + 1$. On numérote les pointes des parapluies ouverts de v de droite à gauche et de 0 à $c(v) - 1$. On discute alors selon les valeurs relatives de s_n et s_{n+1} . La figure 2.13 illustre le propos.

Si $s_{n+1} = s_n + 1$: Dans ce cas il y a $c(u) - 1$ parapluies ouverts dans v et donc pas de parapluie ouvert numéroté $c(u) - 1$. D'après la définition de f , la construction $f(v) = \Phi_m^+(v)$ contient ses marques s_{n+1} et $\overline{s_{n+1}}$ sur le sommet critique de v . Pour la construction $\Phi_m^+(v)$, l'insertion selon le mode m à la lecture de la valeur $s_{n+1} = c(u) - 2$ conduit à l'insertion d'un pic sur le sommet critique aboutissant à la création d'un parapluie ouvert que l'on numérote $c(u) - 1$. Par définition de v dans le cas particulier où les pas supprimés forment un pic dans u , on a donc $\Phi_m(s_1, \dots, s_{n+1}) = u$. D'une part, le parapluie numéroté $c(u) - 1$ précède le sommet critique de

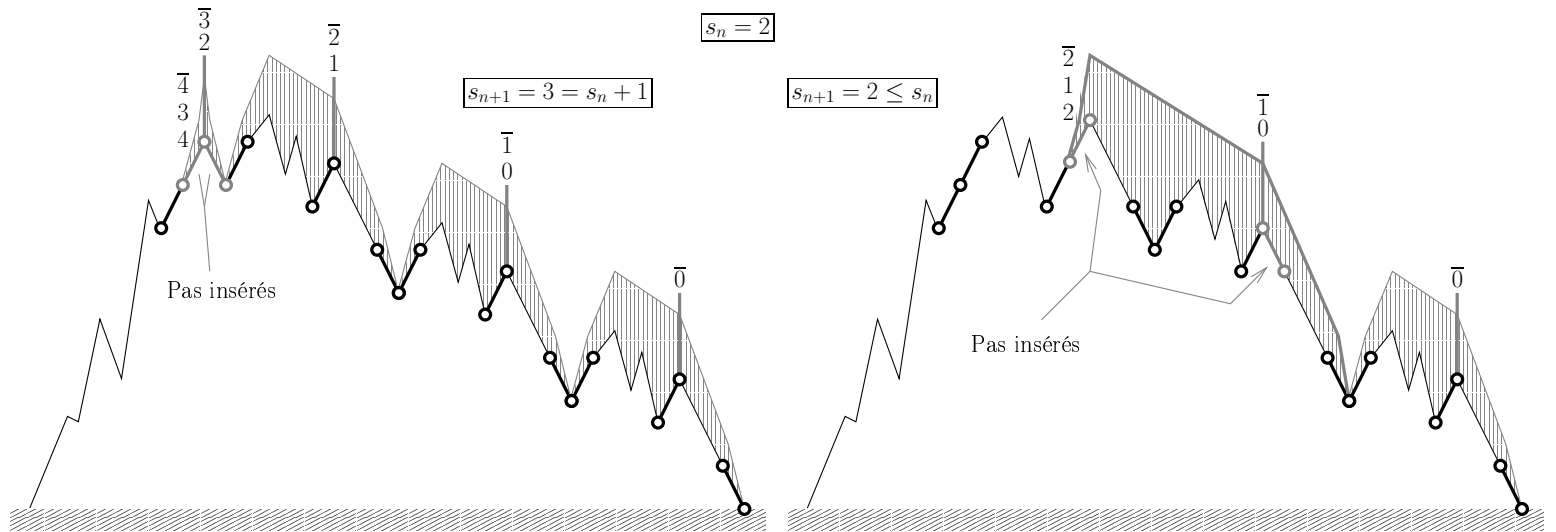


FIG. 2.13 – Les deux types d'insertion dans les factorisations "avec parapluies"

u où se placent les marques $s_{n+1}s_n\overline{s_{n+1}}$. D'autre part, la marque d'indice $s_{n+1} - 1$ initialement placée sur le sommet critique se retrouve sur la pointe du parapluie numéroté $c(u) - 1$. Finalement $\Phi_m^+(u) = f(u)$.

Sinon $s_{n+1} \leq s_n$: Dans ce cas il existe dans v un parapluie numéroté $c(u) - 1$. D'après la définition de f , la construction $f(v) = \Phi_m^+(v)$ contient sa marque s_{n+1} sur le sommet critique si $s_{n+1} = s_n$ et sur la pointe du parapluie numéroté $c(u)$ sinon, la marque $\overline{s_{n+1}}$ se trouve sur le parapluie numéroté $c(u) - 1$. Pour la construction $\Phi_m^+(v)$, l'insertion selon le mode m à la lecture de la valeur $s_{n+1} = c(u) - 2$ conduit à l'insertion d'un pas descendant juste après la pointe du parapluie numéroté $c(u) - 1$. Cela augmente d'une unité la taille du parapluie et lui permet de contenir un pas de plus qui se trouve être, dans les deux cas, le pas montant inséré à la place de la marque s_{n+1} . Les pas ajoutés sont donc le pas montant le plus à gauche du parapluie numéroté $c(u) - 1$ et le pas descendant qui suit sa pointe. Par définition de v on a donc $\Phi_m(s_1, \dots, s_{n+1}) = u$. Dans $\Phi_m^+(v)$, la marque s_{n+1} n'est jamais précédée d'une descente : il y a un pas montant menant à la pointe d'un parapluie ou du sommet critique ou bien rien du tout. Le pas montant inséré est donc le premier pas du plus long suffixe de parapluies ouverts ce qui permet de positionner le sommet critique de u . Les marques effacées de $\Phi_m^+(v)$ lors de l'insertion (d'indice $c(v)$ à $s_{n+1} + 1$) sont donc exactement celles précédant le sommet critique de u . Finalement $\Phi_m^+(u) = f(u)$.

Pour tout chemin de u de longueur $n + 1$, on a trouvé une unique suite quasi-décroissante s_1, \dots, s_{n+1} telle que $\Phi_m(s_1, \dots, s_{n+1}) = u$ donc Φ_m est bijective lorsqu'elle est restreinte aux chemins de longueur inférieure ou égale à $n + 1$.

□

Les 12 applications pour lesquelles l'ordinateur (cf. lemme 2.3.1) ne donne pas de contre-exemple à la bijectivité de taille inférieure à 11 sont donc bien des bijections. Les deux symétries, miroir et d'échange, permettent de regrouper ces applications en 3 classes de 4 applications.

Proposition 2.4 *Les 210 modes d'insertions n'induisent que 12 applications bijectives différentes. Les 6 modes d'insertion suivants conduisent à 6 de ces bijections, les 6 autres en sont les symétries miroir :*

$$\begin{aligned} & (xB\bar{B}, \bar{x}A\bar{A}), \\ & (AB\bar{B}x, \bar{x}\bar{A}), \\ & (BA\bar{B}x, \bar{A}\bar{x}), \\ & (xA\bar{A}, \bar{x}B\bar{B}), \\ & (BA\bar{A}x, \bar{x}\bar{B}), \\ & (AB\bar{A}x, \bar{B}\bar{x}). \end{aligned}$$

Les trois preuves de bijectivité ont donné trois statistiques qui ont la même distribution sur les chemins de Dyck : le nombre de pas descendants après le dernier pas montant, le nombre de retours sur l'axe et la couverture. Les modes qui sont miroirs des trois modes étudiés ci-dessus ne donnent pas de statistique sensiblement différente car ils s'obtiennent par la symétrie miroir sur les mots de Dyck. Par contre, les modes qui sont obtenus par échange des marques génériques sont plus liés à une symétrie sur les suites quasi-décroissantes que sur les chemins de Dyck ; ils donnent donc trois nouvelles statistiques sensiblement différentes. On a réussi à identifier deux de ces statistiques sans parvenir à décrire la troisième. On ne rédige pas les preuves de la compatibilité de ces fonctions de marquage avec leur mode car elles sont similaires aux trois précédentes.

Un sommet dans une *double descente* est un sommet qui est précédé d'une descente et qui n'est pas suivi d'une montée. La fonction de marquage $f_{double-descente}$ compatible avec le mode d'insertion $(xA\bar{A}, \bar{x}B\bar{B})$ se définit en considérant le nombre de pics précédant la première double descente. La figure 2.14 illustre ce marquage. Le rang k de la construction $f_{double-descente}$ est ce nombre de pics. Pour obtenir cette construction, on insère le facteur $(i-1)\bar{i}-1$ dans le i^{eme} pic et le facteur $k\bar{k}$ dans le sommet dans la première double descente.

Un *préfixe croissant* d'un mot de Dyck est un préfixe tel que tout sommet est à une hauteur supérieure ou égale à celle du dernier creux qui le précède. Pour tout mot de Dyck il existe un *plus long préfixe croissant*. La fonction de marquage $f_{préfixe-croissant}$ compatible avec le mode d'insertion $(xB, \bar{x}A\bar{A}\bar{B})$ se définit en considérant les doubles descentes dans le plus long préfixe croissant. Elle est illustrée sur un exemple par la figure 2.15. Pour un chemin de Dyck d , le rang k de la construction est le nombre de doubles descentes dans le plus long préfixe croissant. Pour obtenir cette construction, on insère le facteur $k-1\bar{k}-1\bar{k}$

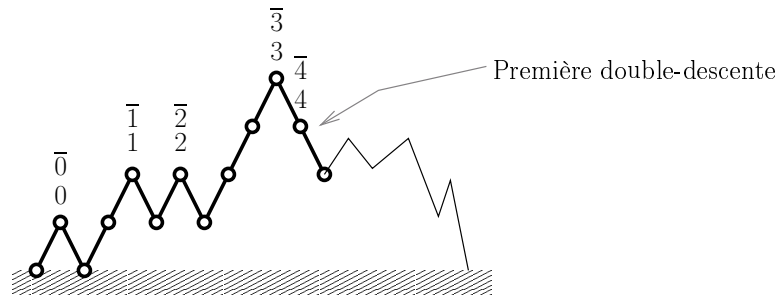


FIG. 2.14 – Fonction de marquage dans les pics avant la première double-descente pour le mode $(xA\bar{A}, \bar{x}B\bar{B})$

dans le dernier sommet du plus long préfixe croissant. Le dernier pic avant ce sommet contient la marque k . Pour $i = 0 \dots k - 2$, on insère le facteur $i\bar{i}$ dans la $i + 1^{eme}$ double descente.

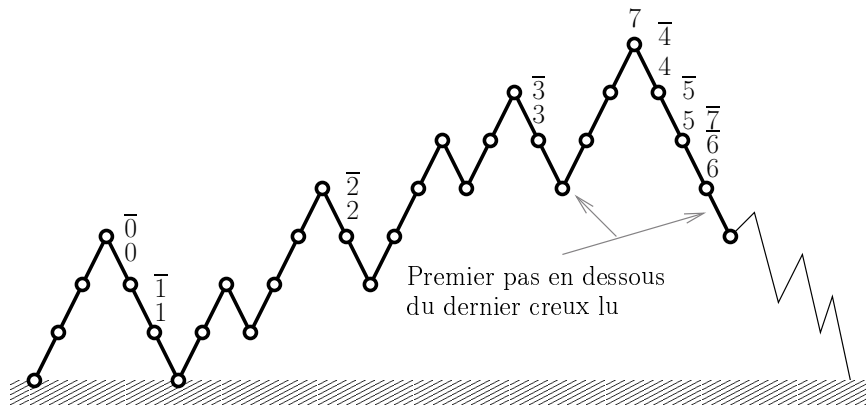


FIG. 2.15 – Fonction de marquage des préfixes croissants pour le mode $(BA\bar{A}x, \bar{x}B\bar{B})$

L'étude des modes d'insertion a permis de prouver que 5 statistiques sur les chemins de Dyck ont la même distribution. Certaines égalités étaient déjà connues, comme le lien entre les sommets sur l'axe et les sommets après le dernier pas montant, d'autres sont nouvelles. Il manque dans cette étude la statistique associée au mode d'insertion $(AB\bar{A}x, \bar{B}\bar{x})$ qui est relié par la symétrie d'échange au mode s'expliquant par les parapluies.

Proposition 2.5 *Les 5 statistiques suivantes ont la même distribution sur les chemins de Dyck :*

- le nombre de pas descendants après le dernier pas montant,
- le nombre de pas descendants dont un sommet est sur l'axe,
- le nombre de parapluies dans le plus long suffixe qui est une concaténation de parapluies,
- le nombre de pics avant la première double descente,

– le nombre de double descentes dans le plus long préfixe croissant.

Preuve. La distribution de ces statistiques est aussi celle de la valeur du dernier élément sur les suites quasi-décroissantes. La bijection dont la fonction de marquage est liée à une des statistiques prouve que cette statistique a la même distribution que la valeur du dernier élément des suites quasi-décroissantes. \square

2.2 Exemples de bijections définies par insertion

La partie précédente était dédiée à l'étude d'un ensemble de variations sur un algorithme codant une bijection classique entre les suites quasi-décroissantes et les chemins de Dyck. L'étude systématique de ces variations n'était pas spécialement motivée par un problème combinatoire, l'intérêt était de comprendre cette méthode. On s'attache maintenant à présenter des questions combinatoires où apparaissent des bijections de ce cadre ou bien de généralisations naturelles.

2.2.1 Demi-pyramides strictes et chemins de Motzkin

On a vu dans la section 2.1.1 que les suites quasi-décroissantes sont en bijection avec les demi-pyramides de dominos. Les demi-pyramides strictes, déjà présentée le chapitre 1, forment un sous-ensemble des demi-pyramides. Une *suite quasi-décroissante sans palier* $s = s_1, \dots, s_n$ est une suite quasi-décroissante telle que pour $i = 2 \dots n$, $s_{i-1} \neq s_i$. On peut caractériser les suites quasi-décroissantes qui sont associées aux demi-pyramides strictes par la suppression récursive de la pièce minimale dans la colonne de plus grand indice : ce sont les suites quasi-décroissantes sans palier. Sur la partie gauche de la figure 2.16, il y a un exemple de demi-pyramide stricte ; les nombres \boxed{i} indiquent l'ordre dans lequel les pièces sont récursivement enlevées pour obtenir la suite quasi-décroissante sans palier.

Un *mot de Motzkin* est un mot u sur l'alphabet $\{x, y, \bar{x}\}$ tel que $|u|_x = |u|_{\bar{x}}$ et pour tout préfixe v de u , $|v|_x \geq |v|_{\bar{x}}$. La longueur d'un mot de Motzkin est le nombre de lettres qui le compose. Comme pour les mots de Dyck, il existe une représentation des mots de Motzkin sous la forme d'un chemin en rajoutant un pas élémentaire $(1, 0)$ associé à chaque occurrence de la lettre y . Sur la droite de la figure 2.16, il y a un exemple de chemin de Motzkin qui a été tourné. Un tel chemin de longueur n peut aussi se coder par la suite des hauteurs de ses $n + 1$ sommets. Une telle suite des hauteurs $s = s_1, \dots, s_{n+1}$ est une suite quasi-décroissante telle que $s_1 = 0 = s_{n+1}$ et pour tout $i = 2, \dots, n + 1$, $|s_{i-1} - s_i| \leq 1$. Réciproquement, toute suite quasi-décroissante satisfaisant ces contraintes est la suite des hauteurs des sommets d'un chemin de Motzkin.

Proposition 2.6 *Pour tout $n \in \mathbb{N}$, il y a autant de chemins de Motzkin de longueur n que de demi-pyramides strictes de taille $n + 1$.*

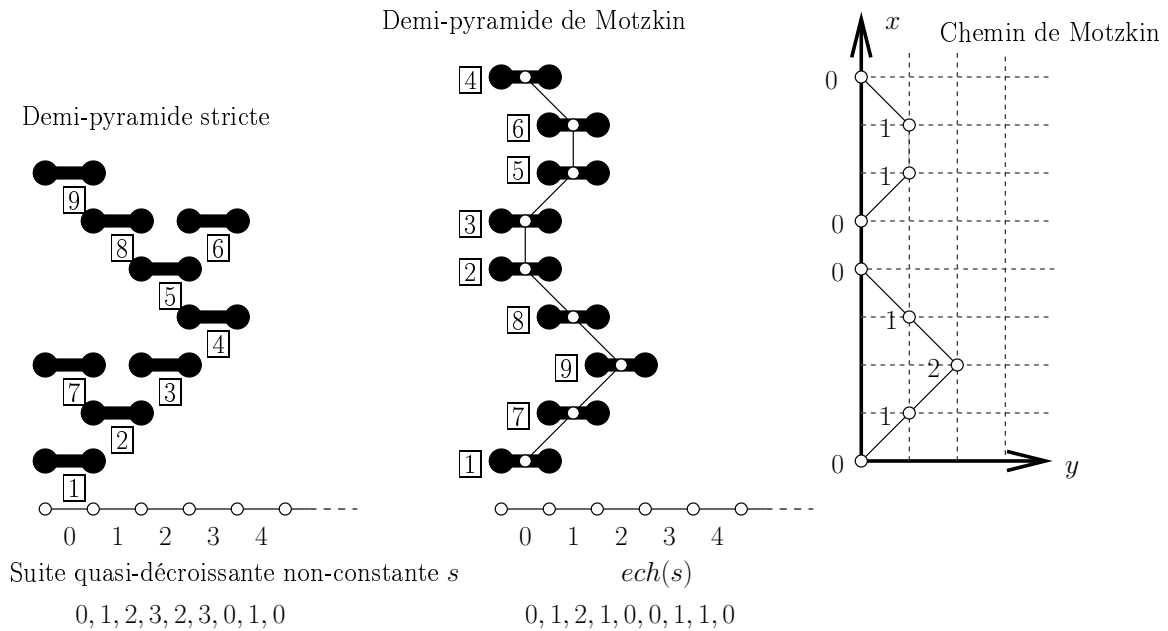


FIG. 2.16 – Demi-pyramides strictes et Chemins de Motzkin : un exemple

Cette proposition est déjà connue [Sha], l'intérêt de la preuve bijective que l'on présente est qu'elle est basée sur l'involution ech définie sur les suites quasi-décroissantes.

Preuve. On associe bijectivement une demi-pyramide stricte d de taille $n + 1$ à un chemin de Motzkin m de longueur n .

La demi-pyramide d est associée à la suite quasi-décroissante sans palier s_d . La clef de cette bijection est que $ech(s_d)$ est une suite de hauteurs de $n + 1$ sommets qui code un chemin de Motzkin de longueur n .

Pour s'en convaincre on considère la décomposition des suites quasi-décroissantes non-vides s :

$$s = 0 + 0.(s_1 \oplus 1) + 0.s_2 + 0.(s_1 \oplus 1).s_2 \tag{22}$$

où s_1 et s_2 sont des suites quasi-décroissantes non vides. Pour obtenir la décomposition des suites quasi-décroissantes sans paliers il suffit de supprimer le cas $0.s_2$ et on obtient la décomposition

$$s = 0 + 0.(s_1 \oplus 1) + 0.(s_1 \oplus 1).s_2 \tag{23}$$

où s_1 et s_2 sont des suites quasi-décroissantes sans palier. Cette décomposition étant compatible avec celle utilisée pour définir récursivement $ech(s)$ on obtient une décomposition des images de ces suites par l'involution ech :

$$ech(s) = 0 + 0.ech(s_1) + 0.(ech(s_2) \oplus 1).ech(s_1). \tag{24}$$

On montre alors grâce à une induction sur la longueur des suites quasi-décroissantes sans palier s que $ech(s)$ est la suite des hauteurs d'un chemin de Motzkin. Toute suite $ech(s)$

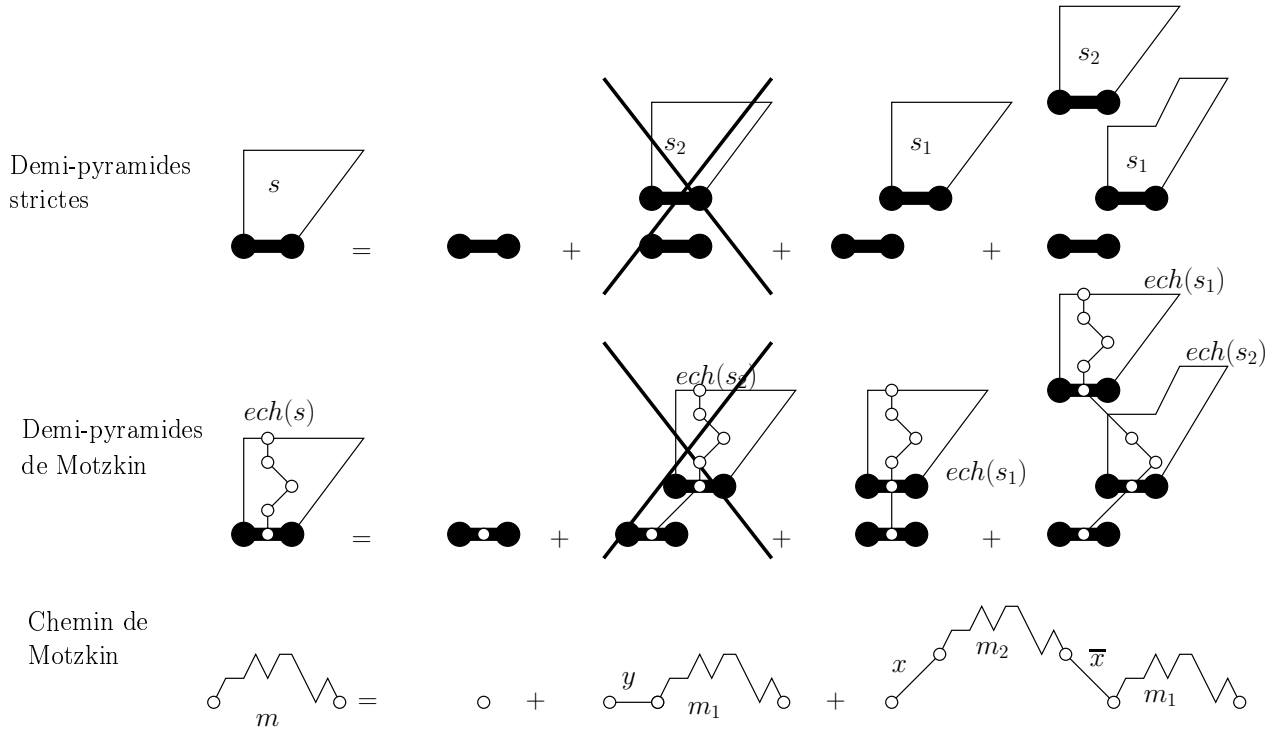


FIG. 2.17 – Décomposition des demi-pyramides strictes et des chemins de Motzkin

est donc associée à un chemin de Motzkin et la réciproque est vrai car

$$m = \epsilon + ym_1 + xm_2\bar{x}m_1 \quad (25)$$

est l'équation de décomposition des chemins de Motzkin au premier retour sur l'axe.

□

2.2.2 Méthode du noyau et développement de Taylor

On commence par présenter la méthode du noyau, une technique de résolution de certaines équations fonctionnelles, en l'illustrant par l'énumération des chemins de Dyck selon leur demi-longueur. Ensuite on montre que, sur cet exemple, une alternative à la méthode du noyau est de travailler avec un développement de Taylor d'une des fonctions inconnues. On obtient ainsi un système infini d'équations linéaires dont les inconnues sont les dérivées partielles d'une même fonction. Les solutions des systèmes linéaires tronqués suggèrent l'existence d'une interprétation combinatoire de ces systèmes. On donne cette interprétation en définissant un nouveau paramètre sur les chemins de Dyck : la hauteur de rayon. On résout le système à l'aide d'une bijection qui s'exprime simplement dans le cadre d'une généralisation de l'algorithme 3 d'insertion uniforme.

Présentation sur un exemple de la méthode du noyau

Pour écrire une équation de décomposition pour les chemins de Dyck on introduit un paramètre supplémentaire : la longueur de la dernière descente $l(u)$ d'un chemin de Dyck u , *i.e.* le nombre de pas descendants après le dernier pas montant. On introduit la série génératrice $D(t; s)$ des chemins de Dyck où la variable t compte la demi-longueur et s la longueur de la dernière descente :

$$D(t; s) = \sum_{u \in Dyck} t^{|u|_x} s^{l(u)}$$

où la variable u parcourt l'ensemble des chemins de Dyck.

Pour écrire une équation sur $D(t, s)$ on partitionne l'ensemble des chemins de Dyck selon que le chemin contiennent au moins deux pics ou non. Cette partition et la suite du raisonnement sont illustrées par la figure 2.18. La série génératrice des chemins de Dyck

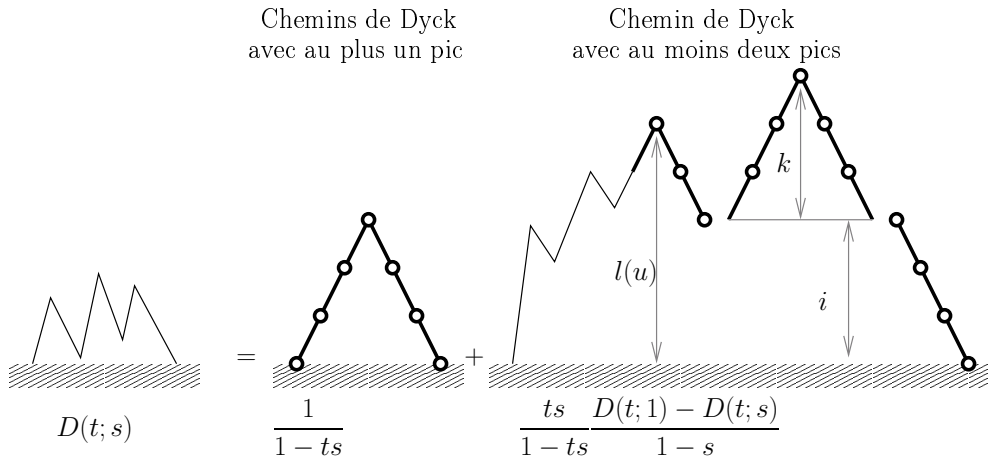


FIG. 2.18 – Décomposition des chemins de Dyck "catalysée" par la longueur de la dernière descente

ne contenant pas plus d'un pic est $\frac{1}{1-ts}$. Tout chemin de Dyck u contenant au moins deux pics se factorise en

$$u = v\bar{x}x^k\bar{x}^k\bar{x}^{l(u)-k}, \text{ avec } k \geq 1$$

qui distingue les pas après et au dessus du dernier creux. Réciproquement, on peut obtenir tous les chemins de Dyck contenant au moins deux pics en ajoutant un facteur $x^k\bar{x}^k$ après un pas descendant de la dernière descente d'un chemin de Dyck non-vide. Soit u un chemin de Dyck de demi-longueur $n > 0$. La série génératrice des chemins avec au moins deux pics obtenus par ajout à ce chemin d'un facteur $x^k\bar{x}^k$ dans la dernière descente est

$$t^{|u|_x} \sum_{i=0}^{l(u)-1} s^i \left(\sum_{k \geq 1} (ts)^k \right) = t^{|u|_x} \frac{ts}{1-ts} \frac{1-s^{l(u)}}{1-s}$$

où i code la hauteur où est inséré le facteur et k la demi-longueur de ce facteur. Dans la sommation sur les chemins de Dyck non-vides on reconnaît des expressions de $D(t; s)$ et $D(t; 1)$:

$$\sum_{u \in \text{Dyck}, u \neq \epsilon} (t^{|u|} - t^{|u|} s^{l(u)}) \frac{ts}{(1-ts)(1-s)} = \frac{ts}{(1-ts)(1-s)} ((D(t; 1) - 1) - (D(t; s) - 1)).$$

Finalement cette décomposition des chemins de Dyck conduit à l'équation fonctionnelle

$$D(t; s) = \frac{1}{1-ts} + \frac{ts}{1-ts} \frac{D(t; 1) - D(t; s)}{1-s}. \quad (26)$$

Le rôle de la variable supplémentaire s est de permettre, de "catalyser", l'écriture d'une équation fonctionnelle. La fonction inconnue recherchée est la série $D(t; 1)$ donc s ne fait pas partie de la solution, des "produits de la réaction". Pour cette raison, Zeilberger [Zei00] qualifie de *catalytique* la variable s comptant la longueur de la dernière descente.

La première tentative pour résoudre l'équation fonctionnelle (26) est de poser $s = 1$; le terme $1/(1-s)$ impose de multiplier tout d'abord l'équation par $1-s$, et on obtient

$$0 = 0 + \frac{t}{1-t} (D(t; 1) - D(t; 1)) \quad (27)$$

ce qui n'est pas faux mais n'apporte pas d'information sur la solution $D(t; 1)$.

La méthode du noyau [BMP00] consiste à rassembler les termes $D(t; s)$ dans l'équation (26) :

$$\left(1 + \frac{ts}{(1-ts)(1-s)}\right) D(t; s) = \frac{1}{1-ts} \left(1 + \frac{ts}{1-s} D(t; 1)\right). \quad (28)$$

Le noyau de cette équation est $K(t; s) = 1 + \frac{ts}{(1-ts)(1-s)}$, le coefficient de $D(t; s)$. On cherche alors $\sigma(t)$ une série de Laurent en t telle que $K(t; \sigma(t)) = 0$. Le numérateur de $K(t; s)$ est un polynôme de degré deux en s qui admet deux racines

$$\sigma_1(t) = \frac{1 - \sqrt{1-4t}}{2t}, \quad \sigma_2(t) = \frac{1 + \sqrt{1-4t}}{2t}.$$

Puisque $\sigma_1(t)$ est finie en $t = 0$, on peut remplacer s par $\sigma_1(t)$, dans (28). On en déduit que

$$D(t; 1) = \frac{1 - \sqrt{1-4t}}{2t}. \quad (29)$$

Utilisation du développement de Taylor

La tentative de poser $s = 1$ dans l'équation (26) peut se raffiner en remarquant que le terme

$$\frac{D(t; 1) - D(t; s)}{1-s}$$

tend vers la dérivée partielle $\frac{\partial}{\partial s}D(t; 1)$ lorsque s tend vers 1. Cette dérivée partielle est une série formelle en t car cela correspond à pointer un pas descendant dans la dernière descente. À la limite, on obtient la relation

$$D(t; 1) = \frac{1}{1-t} + \frac{t}{1-t} \frac{\partial}{\partial s} D(t; 1). \quad (30)$$

Cette relation ne permet pas de conclure mais l'équation fonctionnelle (26) contient une information en $s = 1$.

Le développement de Taylor en $s = 1$ de la série $D(t; s)$ est

$$D(t; s) = \sum_{k \geq 0} \frac{(s-1)^k}{k!} \frac{\partial^k}{\partial s^k} D(t; 1) \quad (31)$$

où chaque $\frac{\partial^k}{\partial s^k} D(t; 1)$ est une série formelle en t car c'est la série génératrice des chemins de Dyck dont k pas descendants distincts, après le dernier pas montant, sont pointés. En injectant le développement de Taylor dans l'équation fonctionnelle (26) on obtient la relation

$$\sum_{k \geq 0} \frac{(s-1)^k}{k!} \frac{\partial^k}{\partial s^k} D(t; 1) = \frac{1}{1-ts} + \frac{ts}{1-ts} \left(\frac{1}{1-s} D(t; 1) - \sum_{k \geq 0} \frac{(s-1)^{k-1}}{k!} \frac{\partial^k}{\partial s^k} D(t; 1) \right). \quad (32)$$

Dans le membre droit de cette équation les termes $D(t; 1)$ se simplifient. Le but est d'observer le comportement de cette équation au voisinage de $s = 1$. On pose donc $u = s - 1$:

$$\sum_{k \geq 0} \frac{u^k}{k!} \frac{\partial^k}{\partial s^k} D(t; 1) = \frac{1}{1-t-ut} - \frac{t(1+u)}{1-t-ut} \left(\sum_{k \geq 1} \frac{u^{k-1}}{k!} \frac{\partial^k}{\partial s^k} D(t; 1) \right). \quad (33)$$

L'extraction du coefficient de degré u^n pour $n \in \mathbb{N}$ donne la relation

$$\frac{1}{n!} \frac{\partial^n}{\partial s^n} D(t; 1) = \frac{t^n}{(1-t)^{n+1}} + \sum_{k=1}^n \frac{t^{n-k+1}}{(1-t)^{n-k+2}} \frac{1}{k!} \frac{\partial^k}{\partial s^k} D(t; 1) + \frac{t}{1-t} \frac{1}{(n+1)!} \frac{\partial^{n+1}}{\partial s^{n+1}} D(t; 1). \quad (34)$$

On note (S_∞) le système linéaire infini d'inconnues $(X_i)_{i \geq 0}$ défini par les équations

$$X_n = \frac{t^n}{(1-t)^{n+1}} + \sum_{k=1}^n \frac{t^{n-k+1}}{(1-t)^{n-k+2}} X_k + \frac{t}{1-t} X_{n+1} \quad (35)$$

valables pour tout $n \in \mathbb{N}$. Ce système a été défini de sorte que $X_i = \frac{1}{i!} \frac{\partial^i}{\partial s^i} D(t; 1)$ en soit une solution. Si on considère le système d'équations formé par les N équations obtenues

pour $n = 0, \dots, N - 1$ on obtient un système linéaire liant les $N + 1$ inconnues $(X_n)_{n=0\dots N}$. Il manque donc une équation à ce système pour pouvoir le résoudre. On remarque que

$$\frac{\partial^N}{\partial s^N} D(t; 1) = O(t^N) \quad (36)$$

car tout chemin contenant N pas descendants distincts pointés dans la dernière descente est de demi-longueur au moins N . La suite des dérivées partielles converge donc vers 0 au sens des séries formelles en t . Cette remarque motive la décision d'altérer le système précédent en ajoutant l'équation

$$X_N = 0. \quad (37)$$

On note (S_N) ce système qui est une "altération" du système (S_∞) . Par exemple on a le système (S_2) qui est

$$\begin{cases} X_0 &= \frac{1}{1-t} + \frac{t}{1-t} X_1, \\ X_1 &= \frac{t}{(1-t)^2} + \frac{t}{(1-t)^2} X_1 + \frac{t}{1-t} X_2, \\ X_2 &= 0. \end{cases}$$

L'ajout d'une équation supplémentaire a permis d'obtenir un système linéaire fini de $N + 1$ équations avec $N + 1$ inconnues. Pour les premières valeurs de N , $N \leq 6$, on peut calculer numériquement la solution, qui s'avère unique, de ce système de $N + 1$ équations et $N + 1$ inconnues. On obtient en particulier

$$X_0 = \frac{T_{2N-2}(t)}{T_{2N-1}(t)} \quad (38)$$

où $T_k(t)$ est le k^{eme} polynôme de Fibonacci. Cette série est connue pour être la série génératrice des chemins de Dyck de hauteur au plus $2N - 1$.

Cela mène à la conjecture suivante.

Proposition 2.7 *Pour tout $N \geq 1$, le système linéaire (S_N) en les $N + 1$ fonctions inconnues $(X_n)_{n=0\dots N}$ admet une unique solution où tous les X_i sont des séries formelles. De plus*

$$X_0 = \frac{T_{2N-2}(t)}{T_{2N-1}(t)}. \quad (39)$$

Pour prouver que ce résultat est vrai pour tout N on va utiliser une interprétation combinatoire des équations du système linéaire (S_N) et en particulier de l'équation supplémentaire (37). Dans cette interprétation X_0 sera la série génératrice d'une certaine classe de chemins de Dyck formée par les chemins où un certain paramètre est strictement inférieur à N . Pour conclure il suffira de construire une bijection entre les chemins de cette classe et les chemins de Dyck de hauteur inférieure ou égale à $2N - 1$ (dont la série génératrice selon la demi-longueur est $T_{2N-2}(t)/T_{2N-1}(t)$).

On commence par proposer une interprétation combinatoire du système (S_∞) . Un *chemin à descente marquée* (d'ordre k) est un chemin de Dyck dont k pas distincts de la dernière descente sont marqués. On note par le symbole \bar{X} un pas descendant marqué. Des exemples de chemins à descente marquée sont $x\bar{x}$ et $x\bar{x}xxxx\bar{x}\bar{X}\bar{X}$. Pour un entier $k \in \mathbb{N}$, l'ensemble \mathcal{D}_k des chemins à descente marquée d'ordre k admet une série génératrice $D_k(t)$ selon la demi-longueur qui vaut $\frac{1}{k!} \frac{\partial^k}{\partial s^k} D(t; 1)$. En effet la série $\frac{\partial^k}{\partial s^k} D(t; 1)$ correspond au pointage de k pas distincts étiquetés par des entiers de 1 à k ; le facteur $1/k!$ traduit le fait que le marquage des pas descendants ne tient pas compte de ces entiers.

Lemme 2.7.1 *La suite des séries génératrices $(D_k(t))$ des chemins à descente marquée d'ordre k satisfait le système d'équations (S_∞) . Plus précisément, la décomposition des chemins à descentes marquées par la suppression du dernier pic conduit au système d'équations (S_∞) .*

Preuve. On peut écrire une équation de décomposition pour les chemins d'ordre k .

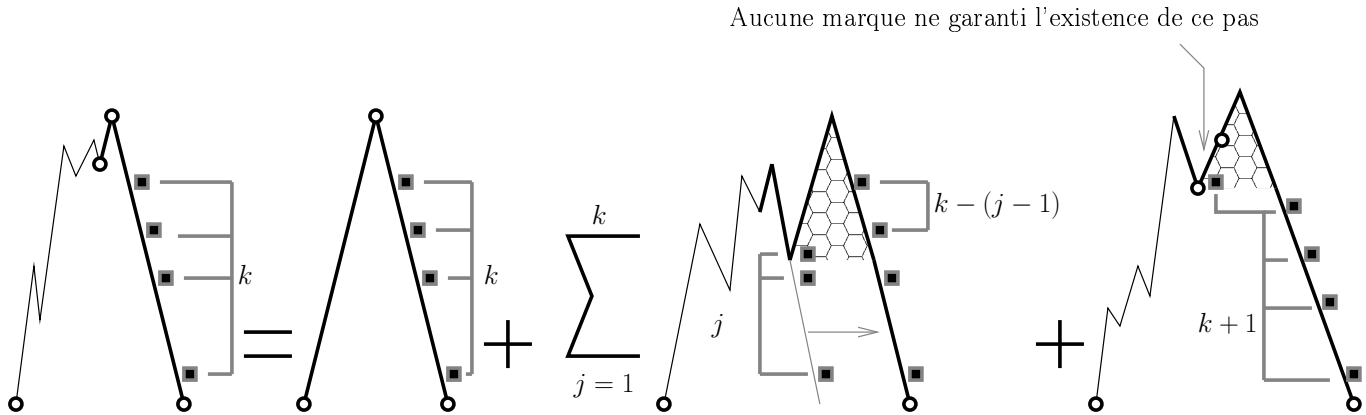


FIG. 2.19 – Décomposition des chemins à descente marquée d'ordre k

On traite d'abord les chemins qui ne contiennent pas de creux. La série génératrice de ces chemins est $\frac{t^k}{(1-t)^{k+1}}$. Pour les autres chemins, il est possible de les factoriser en un sous-chemin d'ordre j , $1 \leq j \leq k + 1$ et un sous-chemin sans creux *non-vide* d'ordre $k - (j - 1)$. Pour cela il suffit de marquer le pas descendant précédent le dernier creux du chemin puis d'isoler le facteur $x^i \bar{x}^i$ qui suit ce creux tout en préservant le marquage. Par exemple le chemin d'ordre 3 $x\bar{x}xxxxx\bar{x}\bar{x}\bar{x}\bar{X}\bar{x}\bar{X}\bar{x}\bar{x}\bar{X}$ se décompose en le sous-chemin d'ordre 2 $x\bar{x}xxxxx\bar{x}\bar{x}\bar{x}\bar{X}$ et le sous-chemin sans creux d'ordre 2 $xxx\bar{x}\bar{X}$. Cette décomposition est réversible car pour reconstruire le chemin initial il suffit d'insérer le sous-chemin sans creux juste après le premier pas descendant marqué du sous-chemin d'ordre j . Un chemin sans creux d'ordre k , $k \geq 1$, est nécessairement non-vide au contraire d'un chemin sans creux d'ordre 0 ce qui explique que dans l'équation de décomposition suivante, le cas d'une factorisation en un sous-chemin d'ordre $k + 1$ et d'un sous-chemin sans creux non-vide

d'ordre 0 contient un facteur t supplémentaire :

$$D_k(t) = \frac{t^k}{(1-t)^{k+1}} + \sum_{j=1}^k D_j(t) \frac{t^{k-(j-1)}}{(1-t)^{k-(j-1)+1}} + D_{k+1}(t) \frac{t}{1-t}. \quad (40)$$

On retrouve les équations (35) qui définissent le système linéaire infini, ce qui est cohérent avec le fait que \mathcal{D}_0 est la classe des chemins de Dyck. \square

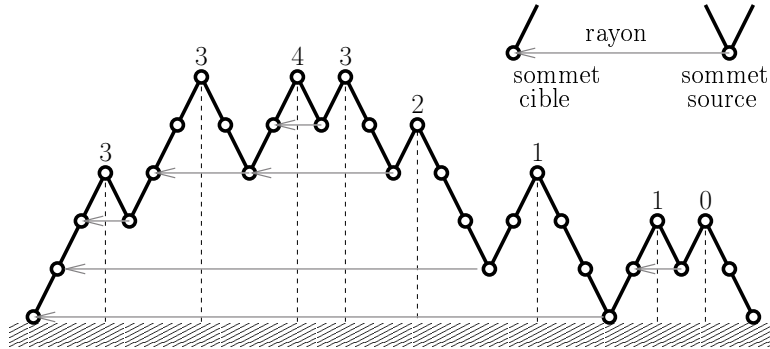


FIG. 2.20 – Définitions des rayons et de la hauteur de rayon

On s'intéresse maintenant à l'interprétation du système (S_N) et en particulier à l'équation

$$X_N = 0.$$

Pour cela on introduit un paramètre sur les chemins de Dyck, la *hauteur de rayon* dont la définition nécessite quelques étapes. On illustre ces définitions sur la figure 2.20. Un *rayon* dans un chemin de Dyck est un segment horizontal orienté issu d'un *sommets source* et s'achevant sur un *sommets cible*. La *source* d'un rayon est un sommet dans un creux, la *cible* de ce rayon est le dernier sommet du chemin qui précède le sommet source et qui est à la même hauteur. La *hauteur de rayon d'un pic* est le nombre maximal de rayons intersectant la verticale issue du sommet au centre du pic. La *hauteur de rayon d'un chemin* est le maximum des hauteurs de rayon de ses pics. On note \mathcal{E}_N l'ensemble des chemins de Dyck dont la hauteur de rayon est au plus N .

La hauteur de rayon est définie pour les chemins de Dyck (d'ordre 0). On l'étend comme suit aux chemins marqués : la source d'un *rayon marqué* est soit un creux soit le sommet ou se termine un pas descendant marqué. La figure 2.21 illustre cette définition. La hauteur de rayon marqué d'un pic est le nombre maximal de rayons marqué intersectant la verticale issue du sommet au centre du pic. On note $\mathcal{E}_{N,k}$ l'ensemble des chemins de hauteur de rayon marqué au plus N et d'ordre k . Pour $k = 0$ on retrouve \mathcal{E}_N puisque sans marque, tous les rayons ont un creux pour source et on retrouve la notion de rayon. Les équations de décomposition des chemins vont donner une interprétation combinatoire du système (S_{N+1}) .

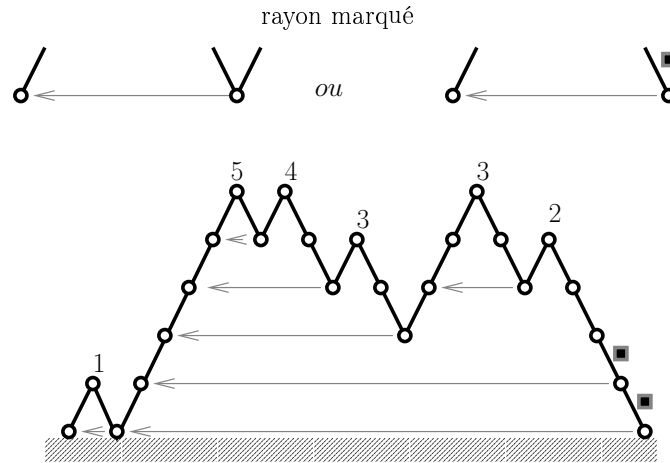


FIG. 2.21 – Rayons marqués et hauteur de rayon marqué

Lemme 2.7.2 Pour $N \in \mathbb{N}$, les séries génératrices $(E_{N,k}(t))_{0 \leq k \leq N}$ des chemins de hauteur de rayon marqué au plus N et de descente marquée d'ordre k satisfont le système d'équations linéaires (S_{N+1}) . En particulier la série formelle X_0 de la solution de (S_{N+1}) est la série génératrice de l'ensemble \mathcal{E}_N des chemins de hauteur de rayon au plus N .

Preuve. Soit u un chemin de Dyck (marqué d'ordre 0) dont les k pics sont étiquetés par

- Pas descendant marqué dans un sous-chemin
- ∕ Base d'une marque

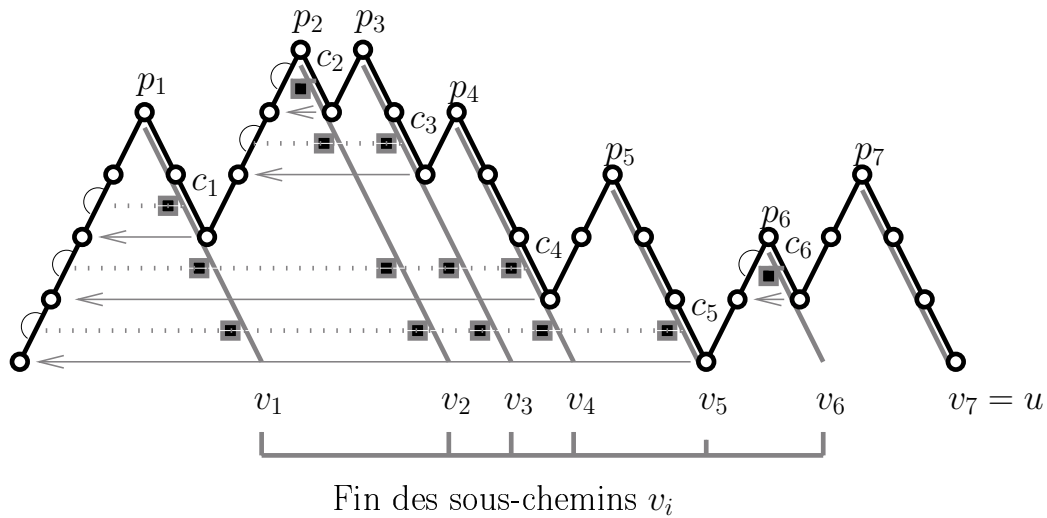


FIG. 2.22 – L'histoire d'un chemin de Dyck

p_1, \dots, p_k de gauche à droite et les creux par c_1, \dots, c_{k-1} comme sur la figure 2.22. On peut voir la décomposition de la figure 2.22, utilisée pour interpréter le système (S_∞) , comme une façon de construire un chemin de Dyck en insérant successivement un nouveau pic à

la fin du chemin. Dans le processus de construction de u il apparaît une suite v_1, \dots, v_k de sous-chemins marqués de u , l'*histoire* de u , qui se termine avec $v_k = u$: le chemin v_i est le chemin qui coïncide avec u jusqu'au pic p_i puis ne contient plus que des pas descendants. Les marques apparaissent dans les chemins v_i lors de l'insertion d'un facteur $x^n(\bar{x} + \overline{X})^n$; un pas descendant marqué est donc créé en même temps qu'un pas montant à la même hauteur que l'on nomme la *base* de la marque. On s'intéresse au devenir d'une marque à la hauteur h dans l'histoire de u après sa création. Lors des insertions de facteurs $x^n(\bar{x} + \overline{X})^n$ au dessus de la marque, le pas descendant marqué est translaté vers la droite mais la base de la marque ne change pas et entre la base et le pas marqué le chemin reste strictement au dessus de la hauteur h . Sur la figure 2.22 on a représenté par des lignes pointillées le lien entre la base d'une marque et les différentes positions du pas descendant marqué. Une marque disparaît lorsqu'un facteur $x^n(\bar{x} + \overline{X})^n$, $n \geq 1$, est inséré dans le chemin juste après le pas descendant marqué formant un creux. Le chemin u est d'ordre 0 donc toutes les marques de l'histoire de u ont disparu pour former un creux. Réciproquement, dans ce processus de construction la seule façon de créer un creux est de faire disparaître une marque. Il y a donc une bijection entre les creux c_i et les marques (ou leurs bases) de l'histoire de u . Le creux où disparaît une marque est la source d'un rayon dont la cible est le sommet précédant la base de cette marque. Les marques de l'histoire de u sont donc en bijection avec les rayons de u . Une marque est présente dans les chemins v_j, \dots, v_l où p_j est le premier pic après la base de la marque et p_l le dernier pic avant le creux c_l associé à la marque. En d'autres termes, la marque est présente dans les chemins v_i dont le dernier pic p_i est compris entre les deux extrémités du rayon associé à la marque. Le nombre de marques dans le chemin v_i est donc le nombre de rayons sous le pic p_i ce qui est la définition de la hauteur de rayon de p_i . La hauteur de rayon d'un chemin u est donc l'ordre maximal des sous-chemins v_1, \dots, v_k .

Cette preuve s'étend à la hauteur de rayon marquée en autorisant les pas descendants marqués à être les sources de rayons. Dans (S_{N+1}) , l'équation $X_{N+1} = 0$ (et implicitement $X_{N+k} = 0$, pour $k > 1$), équivaut à interdire les chemins de hauteur de rayon marquée supérieure ou égale à $N + 1$ dans l'histoire des chemins de Dyck. Les autres équations correspondent à la décomposition des chemins marqués de hauteur de rayon marquée au plus N . Les séries formelles $(E_{N,k}(t))_{k=0..N}$ sont donc solution du système (S_{N+1}) . En particulier, $X_0 = E_{N,0}(t)$ est la série génératrice des chemins de Dyck de hauteur de rayon au plus N .

□

D'après ce lemme, la série génératrice $E_N(t)$ des chemins de Dyck de hauteur de rayon au plus N est la valeur de $D(t; 1)$ dans la solution de (S_{N+1}) . La conjecture menant à la proposition 2.7 équivaut donc à l'égalité

$$E_N(t) = \frac{T_{2N}(t)}{T_{2N+1}(t)}.$$

On a vu que ce quotient de polynômes de Fibonacci est aussi la série génératrice des chemins

de Dyck de hauteur (usuelle) au plus $2N + 1$. Pour prouver la formule de la proposition 2.7 il suffit donc d'établir une bijection entre les chemins de hauteur au plus $2N + 1$ et les chemins de hauteur de rayon au plus N .

Proposition 2.8 *Il existe une bijection préservant la demi-longueur entre les chemins de Dyck de hauteur de rayon N et les chemins de Dyck de hauteur $2N + 1$.*

Pour définir simplement cette bijection, on va utiliser une généralisation de l'algorithme 3. Jusqu'à présent un unique mode d'insertion paramétrait toutes les réécritures dans l'algorithme, on pouvait parler de réécriture "uniforme". La généralisation consiste à utiliser un mode lorsque le terme s_i de la suite décroissante est pair et un autre mode lorsqu'il est impair. Cela conduit à l'algorithme 4.

```

Entrée : Une suite quasi-décroissante  $s = s_1, s_2, \dots, s_n$  ;
Paramètre : Les modes d'insertions  $m_1 = (m_a, m_b)$  et  $m_2 = (m_c, m_d)$  ;
CheminMarque :=  $0\bar{0}$  ;
pour  $i$  allant de 1 à  $n$  faire
    Supprimer les marques d'indice supérieur à  $s_i$  ;
    si  $s_i$  est pair alors
        Substituer  $m_a$ , resp.  $m_b$ , à la marque d'indice  $s_i$ , resp.  $\bar{s}_i$ ;
    sinon
        Substituer  $m_c$ , resp.  $m_d$ , à la marque d'indice  $s_i$ , resp.  $\bar{s}_i$ ;
    fin
    Substituer les marques d'indice  $s_i$ , respectivement  $s_i + 1$ , aux marques d'indice
    générique  $A$ , respectivement  $B$ ;
fin
Supprimer les marques de CheminMarque;
Sortie : Le chemin CheminMarque ;
    
```

Algorithme 4: Algorithme général de période 2

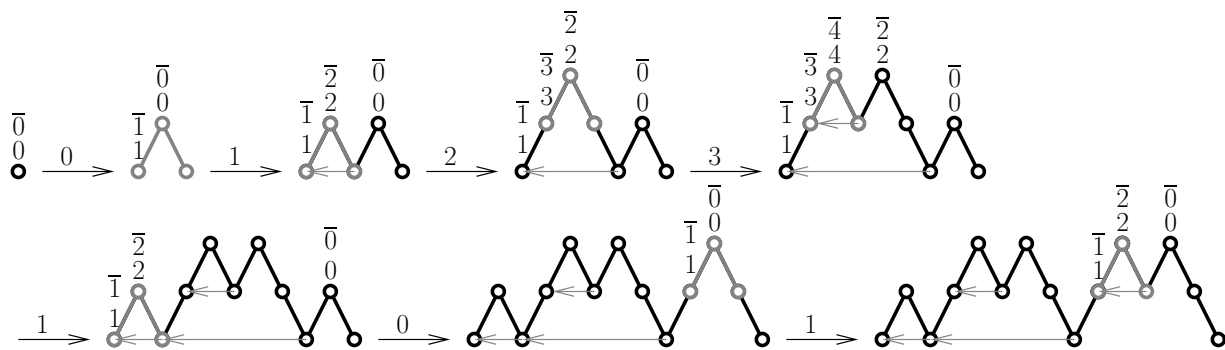


FIG. 2.23 – Un exemple de calcul de l'application $\Phi_{\langle (B\bar{B}x, A\bar{A}\bar{x}); (A\bar{A}x, B\bar{B}\bar{x}) \rangle}$

Le paramétrage de l'algorithme par les modes m_1 et m_2 est noté $\langle m_1; m_2 \rangle$. L'application obtenue par ce paramétrage est noté $\Phi_{\langle m_1; m_2 \rangle}$. Cette application associe un chemin

de Dyck à toute suite quasi-décroissante : en effet, la proposition 2.1 se généralise à ce type d'algorithme car les pas et marques sont toujours insérés selon un mode d'insertion. La figure 2.23 donne un exemple d'exécution de cet algorithme lorsqu'il est paramétré par $\langle (\overline{BBx}, \overline{AAx}); (\overline{AAx}, \overline{BBx}) \rangle$.

Une bijection satisfaisant la proposition 2.8 est l'application de paramétrage

$$\langle (\overline{BBx}, \overline{AAx}); (\overline{AAx}, \overline{BBx}) \rangle .$$

En effet le pic inséré à la lecture de la valeur $2p - 1$ ou $2p$ est de hauteur de rayon p . En avant-goût, on peut observer cette propriété sur l'exemple de la figure 2.23. La preuve du caractère bijectif de cette application et de son lien avec la hauteur de rayon est une récurrence sur la longueur des suites quasi-décroissantes. Elle est similaire aux preuves pour les applications utilisant des réécritures uniformes mais la définition de la fonction de marquage est un peu plus complexe.

Preuve.(Proposition 2.8)

Les deux modes d'insertion admettent \overline{AA} et \overline{BB} comme facteur donc les marques d'indice égaux seront toujours consécutives et on n'insérera que des pics dans les constructions de Dyck. Pour cette application, la fonction de marquage f_{rayon} se décrit en explicitant l'existence et la position des marques d'indice i en fonction de la position des marques d'indice $i - 1$. Ces règles de placement sont illustrées par la figure 2.24.

Ce placement vérifie deux propriétés qui garantissent la cohérence de la définition : les marques d'indices $2p$, $p \geq 0$ sont placées sur un pic et s'il existe des marques d'indice $2p + 2$ alors les marques d'indice $2p + 1$, $p \geq 0$ sont placées sur le sommet cible d'un rayon. Comme cette définition du marquage est récursive on vérifiera et utilisera ces deux propriétés au cours de la définition.

- **Marques d'indice 0** : Les marques $0\overline{0}$ se placent sur le dernier pic du chemin. Si le chemin est vide, la construction est de rang 0. Sinon, il existe une marque d'indice 1. Ces marques sont bien placées sur un pic (sauf dans le cas où le chemin est vide).
- **Marques d'indice $2p + 1$** : On suppose que les marques d'indice $2p$ sont placées sur un pic de hauteur h . On nomme A le sommet précédant celui portant les marques d'indice $2p$. Dans le chemin de Dyck, soit il existe un dernier pas montant y , parmi ceux qui précèdent ce sommet, qui soit de hauteur inférieure au sommet A , dans ce cas on nomme B le sommet après y , sinon B est le premier sommet du chemin. C'est sur le sommet B que se placent les marques d'indice $2p + 1$. Si ce sommet est le sommet précédant celui portant les marques d'indice $2p$ ($A = B$) alors la construction de Dyck est de rang $2p + 1$. Sinon d'une part il existe des marques d'indice $2p + 2$ et d'autre part le sommet B contenant ces marques est la cible d'un rayon (il y a une suite de rayons débutant par celui issu de A).
- **Marques d'indice $2p$, $p > 0$** : On suppose que les marques d'indice $2p - 1$ sont placées sur un sommet qui est la cible d'un rayon. On considère le rayon qui arrive sur le sommet portant les marques d'indice $2p - 1$ et on considère le dernier pic avant le

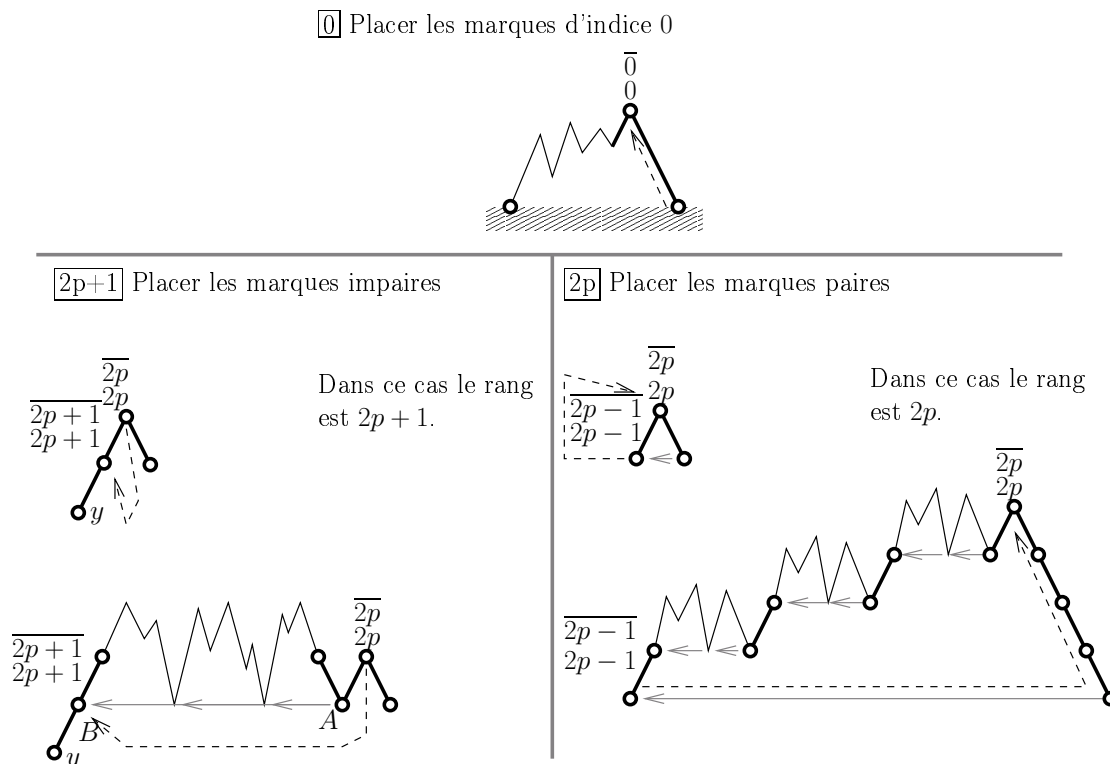


FIG. 2.24 – Règles pour le marquage associé à la hauteur de rayon

sommet source de ce rayon. C'est sur ce sommet que se place le facteur $2p\overline{2p}$. Si le sous-chemin entre les deux extrémités du rayon est un simple pic $x\bar{x}$, alors la construction de Dyck est de rang $2p$, sinon il existe des marques d'indice $2p + 1$. Par définition, les marques d'indice $2p$ sont sur un pic.

La figure 2.25 est un exemple de cette fonction de marquage.

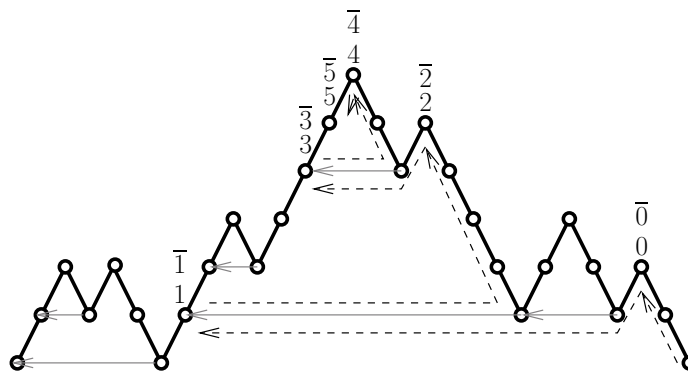


FIG. 2.25 – Fonction de marquage pour la hauteur de rayon

On montre maintenant par récurrence sur la longueur de la suite quasi-décroissante que

le mode d'insertion $\langle (A\bar{A}x, B\bar{B}\bar{x}); (B\bar{B}x, A\bar{A}\bar{x}) \rangle$ est compatible avec cette fonction de marquage.

Soit u un chemin de Dyck de demi-longueur n . On note m le mode d'insertion

$$\langle (B\bar{B}x, A\bar{A}\bar{x}); (A\bar{A}x, B\bar{B}\bar{x}) \rangle .$$

Cas de base $n \leq 1$: Pour le chemin vide on a

$$\Phi_m^+(\epsilon) = 0\bar{0} = f_m(\epsilon).$$

Pour le pic on a

$$\Phi_m^+(x\bar{x}) = 1\bar{1}x0\bar{0}\bar{x} = f_m(x\bar{x}).$$

Dans les autres cas on pourra donc supposer que les marques d'indices 0 sont sur un pic.

Hérédité pour $n + 1 \geq 2$ et $f_{\text{rayon}}(u)$ de rang impair : On suppose l'hypothèse de récurrence satisfaite au rang n . Soit u un chemin de Dyck de demi-longueur $n + 1$ dont la construction $f_m(u)$ est de rang $2p + 1$.

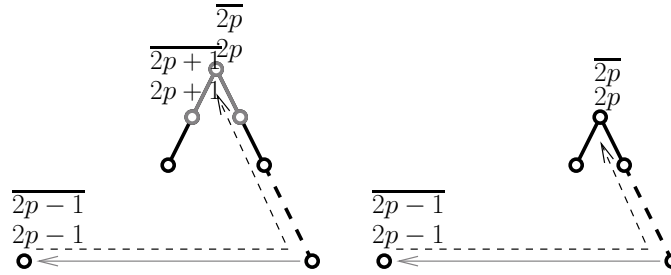


FIG. 2.26 – Insertion pour les rayons (Cas pair)

Analyse : Soit $s = s_1, \dots, s_{n+1}$ une suite quasi-décroissante telle que $\Phi_m^+(s) = f_m(u)$. Nécessairement $s_{n+1} = 2p$ et le mode d'insertion utilisé pour s_{n+1} est donc $(B\bar{B}x, A\bar{A}\bar{x})$. Or les marques d'indice $2p + 1$ sont celles du rang de la construction donc d'après le critère du marquage, cf. figure 2.24, il y a dans $f_m(u)$ un facteur $(2p + 1)\overline{(2p + 1)}x2p\bar{2p}\bar{x}$. Ce facteur correspond à l'insertion selon le mode localement réversible $(B\bar{B}x, A\bar{A}\bar{x})$ donc les deux derniers pas insérés sont ceux qui forment le pic sur lequel se trouvent les marques d'indice $2p$. Le chemin v obtenu par suppression de ce pic est de longueur n donc d'après l'hypothèse de récurrence il existe une unique suite quasi-décroissante s_1, \dots, s_n associée à v . Finalement il n'y a que la suite $s_1, \dots, s_n, 2p$ qui puisse donner u par l'application Φ_m .

Synthèse : Le marquage de u et v est identique jusqu'aux marques d'indice $2p - 1$ car la suppression du pic de u n'a pas d'influence jusque là. (Si $p = 0$, le marquage est identique car il débute à la fin du chemin aussi bien pour u que pour v .)

Puisque $f_m(u)$ est une construction de rang $2p + 1$, les marques $2p$ sont sur un pic double (un facteur $xx\bar{x}$) dans $f_m(u)$. La figure 2.26 montre que dans ce cas là après la suppression du pic il reste un pic sur lesquelles se trouvent les marques d'indice $2p$ de $f_m(v)$. Donc $f_m(v)$ est de rang au moins $2p$. Donc $s_1, \dots, s_n, 2p$ est une suite quasi-décroissante. D'autre part l'insertion selon le mode $(B\bar{B}x, A\bar{A}\bar{x})$ conduit bien à $f_m(u)$. Donc $f_m(u) = \Phi_m^+(s_1, \dots, s_n, 2p)$.

Hérédité pour $n + 1 \geq 2$ et $f_{\text{rayon}}(u)$ de rang pair : On suppose l'hypothèse de récurrence satisfaite au rang n . Soit u un chemin de Dyck de demi-longueur $n + 1$ dont la construction $f_m(u)$ est de rang $2p + 2$.

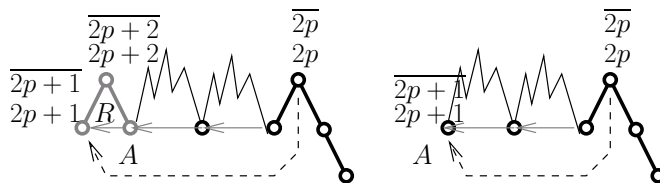


FIG. 2.27 – Insertion pour les rayons (Cas impair)

Analyse : Soit $s = s_1, \dots, s_{n+1}$ une suite quasi-décroissante telle que $\Phi_m^+(s) = f_m(u)$. Nécessairement $s_{n+1} = 2p + 1$ et le mode d'insertion utilisé pour s_{n+1} est donc $(A\bar{A}x, B\bar{B}\bar{x})$. Or les marques d'indice $2p + 2$ sont celles du rang de la construction donc d'après le critère du marquage, cf. figure 2.24, il y a dans $f_m(u)$ un facteur $(2p + 1)(2p + 1)x(2p + 2)(2p + 2)\bar{x}$. Ce facteur correspond bien à l'insertion selon le mode $(A\bar{A}x, B\bar{B}\bar{x})$ donc les deux derniers pas insérés forme le pic qui apparaît juste après les marque d'indice $2p + 1$. Le chemin v obtenu par suppression de ce pic est de longueur n donc d'après l'hypothèse de récurrence il existe une unique suite quasi-décroissante s_1, \dots, s_n associée à v . Finalement il n'y a que la suite $s_1, \dots, s_n, (2p + 1)$ qui puisse donner u par l'application Φ_m .

Synthèse : Le marquage de u et v est identique jusqu'aux marque d'indice $2p$ car la suppression du pic de u n'a pas d'influence jusque là. Puisque $f_m(u)$ est une construction de rang $2p + 2$, le rayon R qui arrive sur le sommet contenant les marques d'indice $2p + 1$ passe sous un simple pic $x\bar{x}$. On note A la source de ce rayon qui est creux de u . Après la suppression de du pic pour obtenir v , le rayon R a disparu, comme le montre la figure 2.27 donc c'est le sommet A qui porte les marques d'indice $2p + 1$. (Ces marques existent car on a pas supprimé le pic portant les marques d'indice $2p$.) Donc $f_m(v)$ est de rang au moins $2p + 1$. Donc $s_1, \dots, s_n, 2p + 1$ est une suite quasi-décroissante. D'autre part l'insertion selon le mode $(A\bar{A}x, B\bar{B}\bar{x})$ conduit bien à $f_m(u)$. Donc $f_m(u) = \Phi_m^+(s_1, \dots, s_n, 2p + 1)$.

De plus on montre aussi que les marques d'indices $2p - 1$ et $2p$ créent des pics de hauteur de rayon p en cas d'insertion. Pour cela il suffit de suivre la définition de la fonction de marquage, cf. figure 2.24. Cette preuve a des aspects récurrents.

Marques d'indice 0 : Le pic créé par des marques d'indice 0 forme le dernier pic qui est bien de hauteur de rayon 0. De plus cette insertion ne modifie pas les autres rayons donc elle préserve la hauteur de rayon des autres pics.

Marques d'indice $2p + 1$, $p \geq 0$: Comme les marques d'indice $2p$ sont sur un pic de hauteur de rayon p , l'insertion d'un pic dans les marques d'indice $2p + 1$ conduit à la création d'un pic sous lequel passe un rayon (car les marques d'indice $2p + 1$ sont toujours suivies d'un pas montant). Ce rayon donne une hauteur de rayon de $p + 1$ pour le pic créé et ne modifie pas les autres hauteurs de rayon.

Marques d'indice $2p$, $p \geq 1$: Comme les marques d'indice $2p - 1$ sont de hauteur de rayon p , l'insertion d'un pic dans les marques d'indice $2p$ ne crée pas de nouveau rayon donc le nouveau pic est de hauteur de rayon p . Cette insertion ne perturbe pas les autres rayons.

□

La combinaison du lemme 2.7.2 et la proposition 2.8 conduit à la proposition 2.7. La hauteur de rayon est donc un paramètre qui se comporte comme le double de la hauteur usuelle.

2.2.3 Les bistatistiques de Haiman et Haglund

Haiman et Garcia [GH96] ont conjecturé une formule, définie par sommation sur les partitions d'entiers, qui donnerait une série de Hilbert d'un certain espace bigradué (alternantes diagonales harmoniques). Dans le but d'interpréter combinatoirement cette formule par une sommation sur les chemins de Dyck, Haiman et Haglund [Hag03] ont donné différentes statistiques sur les chemins de Dyck.

L'ambition de ce paragraphe est de montrer dans quelle mesure ces statistiques sont liées à la démarche utilisée dans ce chapitre pour coder des bijections.

L'*aire* $aire(d)$ d'un chemin de Dyck d est le nombre maximal de carrés de côté $\sqrt{2}$ qui peuvent être placés entre le chemin de Dyck et l'axe horizontal. Pour le chemin de la

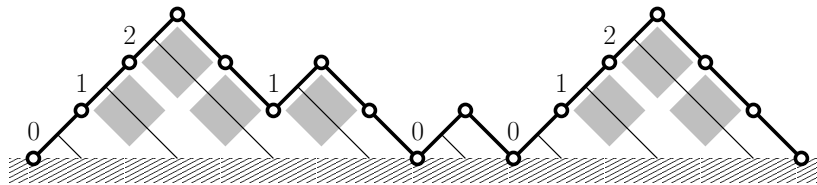


FIG. 2.28 – Aire d'un chemin de Dyck

figure 2.28, on a placé les carrés qui prouvent que l'aire de ce chemin est 7. Si on regroupe les carrés selon les diagonales parallèles aux pas descendants on constate que l'aire d'un chemin de Dyck est aussi la somme des hauteurs des premiers sommets des pas montants.

Le *poids* d'une suite quasi-décroissante $s = s_1, \dots, s_n$ est la somme $\sum_{i=1}^n s_i$. Par la bijection de Viennot on en déduit que l'aire d'un chemin de Dyck est aussi le poids de la suite quasi-décroissante associée. Sur l'exemple, la suite quasi-décroissante est $0, 1, 2, 1, 0, 0, 1, 2$ et l'aire du chemin est bien $0 + 1 + 2 + 1 + 0 + 0 + 1 + 2 = 7$.

La statistique de Haiman

Haiman, cité par Haglund et Loehr dans [HL02], a proposé une seconde statistique sur le chemin de Dyck d qui se définit aussi en utilisant la suite $s = s_1, s_2, \dots, s_n$ des hauteurs des premiers sommets des pas montants. La *statistique de Haiman* d'un chemin de Dyck d , notée $div(d)$, est le nombre d'indices i, j tels que $i < j$ et $s_i \in \{s_j, s_j + 1\}$. Sur la figure 2.29 on représente les paires de pas montants comptabilisées par la statistique div .

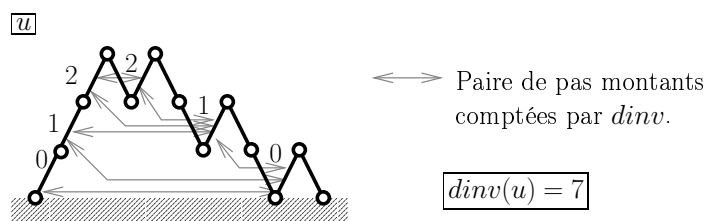


FIG. 2.29 – Couplage des pas montants définissant la statistique de Haiman

L'aire et la statistique de Haiman s'ajoutent à la demi-longueur pour donner la série génératrice

$$C(t; q, p) \equiv \sum_{d \in Dyck} t^{|d|_x} q^{aire(d)} p^{div(d)}. \quad (41)$$

Une des définitions de la formule interprétée par $C(t; q, p)$ est symétrique en q et p , cette série est donc symétrique en q et p :

$$C(t; q, p) = C(t; p, q). \quad (42)$$

Une question ouverte dans ce domaine est de donner une preuve bijective de la symétrie de cette série en utilisant l'interprétation combinatoire sur les chemins de Dyck dont la définition semble *a priori* asymétrique.

Ces deux statistiques ont la même distribution sur les chemins de Dyck. Cette distribution est aussi celle du poids des suites quasi-décroissantes. La bijection de Viennot, codée par l'insertion uniforme selon le mode $(x\overline{B\overline{B}}, \overline{x}A\overline{A})$, prouve bijectivement que l'aire des chemins de Dyck et le poids des suites quasi-décroissantes ont la même distribution. On propose une correspondance entre les suites quasi-décroissantes et les chemins de Dyck qui transforme le poids en la statistique de Haiman. Dans un premier temps on présente la fonction de marquage utilisée, on montre son lien avec la statistique de Haiman puis on présente un mode d'insertion compatible avec la fonction de marquage.

Dans un chemin de Dyck de hauteur h , un pas montant est *de hauteur maximale à droite* si c'est un pas se finissant à la hauteur h ou si c'est un pas montant de hauteur $h - 1$ à la droite duquel il n'y a pas de pas à la hauteur h . Sur la figure 2.30 on a marqué par un triangle les pas de hauteur maximale à droite. On va voir que le nombre de pas de hauteur maximale à droite dans un chemin est un analogue de la dernière descente.

On va ordonner les sommets situés après un pas de hauteur maximale à droite et le sommet situé après le pas descendant qui suit le pas de hauteur maximale (usuelle) le plus à droite. Tout d'abord il y a les sommets de hauteur maximale (usuelle), ordonnés de droite à gauche, puis les sommets de hauteur inférieure également ordonnés de droite à gauche, le dernier de ces sommets étant celui qui succède au pas descendant ci-dessus. La fonction de marquage f_{haiman} est la suivante : sur le i^{eme} sommet on place le facteur $i - \overline{1i - 1}$. Ce marquage est illustré par l'exemple de la figure 2.30.

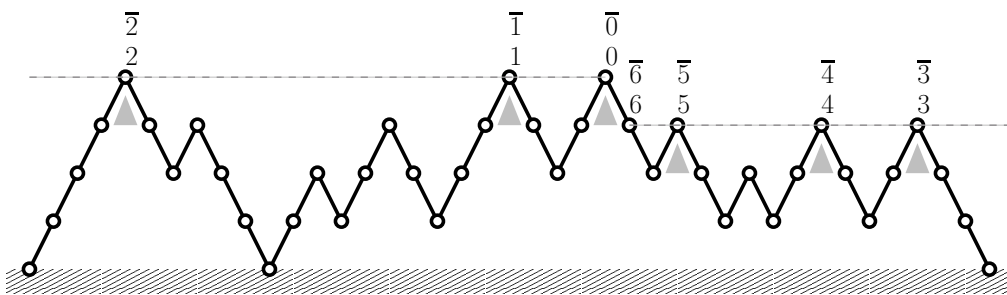


FIG. 2.30 – Fonction de marquage la statistique de Haiman dinv

Les sommets marqués par la fonction de marquage f_{haiman} apparaissent dans la littérature et en particulier dans un article basée sur la méthode ECO [BDLPP98]. Il suffit d'y remarquer que les arbres planaires présentés par Barcucci, Del Lungo, Pergola et Pinzani peuvent être vus comme des chemins de Dyck.

On pointe ici une propriété supplémentaire de cette fonction de marquage ; elle se comporte bien par rapport à la statistique dinv .

Lemme 2.8.1 *Soit u un mot de Dyck et k le rang de la construction $f_{\text{haiman}}(u)$. Pour tout $i \leq k$, l'insertion des pas x et \bar{x} dans les marques i et \bar{i} conduit à un chemin de Dyck v_i tel que*

$$\text{dinv}(v_i) = \text{dinv}(u) + i.$$

Preuve. L'insertion d'un pic dans le chemin u ne modifie pas les hauteurs et les positions relatives des pas montants de u ; les paires de pas montants initialement dans u contribuant à la statistique dinv contribuent aussi à la statistique dinv dans v_i . Le pas montant ajouté dans la marque i forme une paire avec exactement tous les pas montants menant à un pic contenant les marques d'indice j inférieur à i . La statistique dinv est donc augmentée de i dans le chemin v_i par rapport au chemin u . \square

Pour trouver un mode d'insertion compatible avec la fonction de marquage f_{haiman} il faut généraliser la notion de mode d'insertion. La généralisation tient en une opération de *permutation circulaire des indices* des marques : pour une construction de rang $s_i + 1$, les marques d'indices k deviennent des marques d'indice $k + 1 \pmod{s_i + 1}$. Par exemple, la construction de Dyck $1x0\bar{1}\bar{x}23x\bar{3}\bar{x}\bar{2}0$ devient la construction $2x1\bar{2}\bar{x}30x\bar{0}\bar{x}\bar{3}\bar{1}$. La variation du mode d'insertion uniforme de l'algorithme 3 consiste à permuter les marques après chaque insertion, on obtient l'algorithme 5.

Entrée : Une suite quasi-décroissante $s = s_1, s_2, \dots, s_n$;

Paramètre : Le mode d'insertion $m = (m_1, m_2)$;

CheminMarque := $0\bar{0}$;

pour i allant de 1 à n **faire**

Supprimer les marques d'indice supérieur à s_i ;

Substituer m_1 , respectivement m_2 , à la marque d'indice s_i , respectivement \bar{s}_i ;

Substituer les marques d'indices s_i , respectivement $s_i + 1$, aux marques d'indice générique A , respectivement B ;

Permuter les indices des marques ($k := k + 1 \pmod{s_i + 1}$) ;

fin

Supprimer les marques de *CheminMarque* ;

Sortie : Le chemin *CheminMarque* ;

Algorithme 5: Algorithme d'insertion uniforme avec permutation des indices

Tant tout ce chapitre σ désigne les permutations cycliques définies sur tout ensemble $\{0, \dots, n\}$ par $\sigma(i) = i + 1 \pmod n$. Pour des raisons qui apparaîtront dans la partie 2.2.4, on note $\sigma \circ \bar{\sigma} \circ m_1$, l'insertion selon le mode m_1 , suivie d'une permutation cyclique $\bar{\sigma}$ des marques de pas descendants et d'une permutation cyclique σ des marques de pas montants (l'insertion est une application entre constructions).

Proposition 2.9 *Le mode d'insertion $\sigma \circ \bar{\sigma} \circ (xB\bar{B}, \bar{x}A\bar{A})$ est compatible avec la fonction de marquage f_{haiman} et $\Phi_{\sigma \circ \bar{\sigma} \circ (xB\bar{B}, \bar{x}A\bar{A})}$ est une bijection qui transforme le poids d'une suite décroissante en la statistique de Haiman din_v sur le chemin de Dyck.*

Preuve. On prouve cette proposition par récurrence sur la longueur n des suites quasi-décroissantes. Dans cette preuve on note m le mode d'insertion $\sigma \circ \bar{\sigma} \circ (xB\bar{B}, \bar{x}A\bar{A})$ et f la fonction de marquage f_{haiman} .

Hypothèse de récurrence : On suppose que pour toute suite quasi-décroissante de longueur inférieure ou égale à n on a $\Phi_m^+(s) = f(\Phi_m(s))$ et que Φ_m est une bijection lorsqu'elle est restreinte aux suites et chemins de longueur inférieure ou égale à n .

Cas de base : Pour $n = 0$, $\Phi_m^+(s) = 0\bar{0} = f(\epsilon)$ et Φ_m est une bijection qui associe la suite vide au chemin vide. Pour $n = 1$ on a

$$\Phi_m^+(0) = x0\bar{0}\bar{x}1\bar{1} = f(x\bar{x}).$$

Hérédité : Soit u un chemin de Dyck de longueur $n + 1$ et de hauteur h_u . On cherche par un raisonnement de type analyse/synthèse une suite quasi-décroissante $s = s_1, \dots, s_n, s_{n+1}$ telle que $\Phi_m(s) = u$ et $\Phi_m^+(s) = f(u)$. La figure 2.31 illustre le propos.

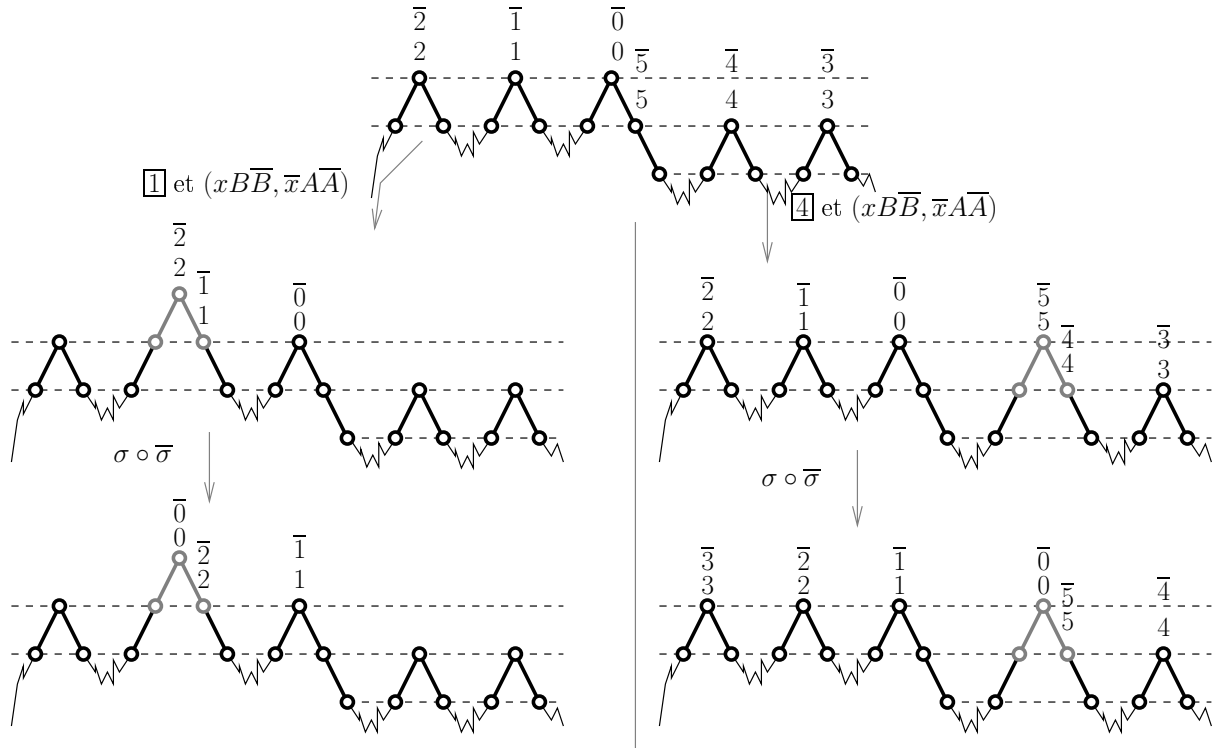


FIG. 2.31 – Un exemple de recherche de l'inverse de $\Phi_{\sigma \circ \bar{\sigma}(xB\bar{B}, \bar{x}A\bar{A})}$

Analyse : Soit s une suite décroissante telle que $\Phi_m(s) = u$ et $\Phi_m^+(s) = f(u)$. On note $maxd(u)$ le nombre de pas de hauteur maximale à droite. La construction $f(u)$ est de rang $maxd(u)$. Puisque $f(u) = \Phi_m^+(s)$, on en déduit que $s_{n+1} = maxd(u) - 1$. Après l'insertion à la lecture de s_{n+1} on a permuté les marques $(\sigma \circ \bar{\sigma})$ donc pour trouver les deux derniers pas insérés il faut observer les marques d'indice 0 et $maxd(u)$. Comme $(xB\bar{B}, \bar{x}A\bar{A})$ est localement réversible, on constate que les deux derniers pas insérés selon le mode m pour donner u sont ceux qui forment le dernier pic où le chemin de Dyck atteint sa hauteur maximale. On note v le chemin de Dyck de demi-longueur n et de hauteur h_v obtenu par la suppression du dernier pic de hauteur maximale de u .

D'après l'hypothèse de récurrence, il existe une unique suite décroissante s_1, \dots, s_n telle que $\Phi_m(s_1, \dots, s_n) = v$. D'après cette analyse il existe une unique suite s_1, \dots, s_{n+1} pouvant être associée à u .

Synthèse : On discute suivant le nombre de pics k de hauteur maximale h dans le chemin u .

$k = 1$: Ce cas est illustré par l'exemple de la figure 2.31 pour le choix de $s_{n+1} = 1$. Le pic, noté p , que l'on supprime de u pour donner v est le seul à la hauteur h , il y a donc $\max d(u) - 1$ pics à la hauteur $h - 1$ à la droite de p . La suppression du pic p donne un chemin v qui est de hauteur $h - 1$ et possède donc au moins $\max d(u) - 1$ pics de hauteur maximale. À ceux-ci s'ajoute au moins le nouveau pic p' à la hauteur $h - 1$ créé par la suppression de p (Puisque p est seul à la hauteur h dans u , il ne peut y être précédé d'un pas descendant). Dans v la fonction de marquage marque le dernier maximum, maintenant à la hauteur $h - 1$, et que p' possède $\max d(u) - 1$ pics de hauteur $h - 1$ à sa droite, p' est marqué par les marques d'indice $\max d(u) - 1$. Donc $\Phi_m(s_1, \dots, s_{n+1}) = u$. De plus, la construction $\Phi_m^+(s_1, \dots, s_{n+1})$ est bien $f(u)$ car la permutation $\sigma \circ \bar{\sigma}$ permet de décaler les indices pour, par exemple, placer les marques d'indice 0 sur le nouveau maximum.

$k > 1$: Ce cas est illustré par l'exemple de la figure 2.31 pour le choix de $s_{n+1} = 4$. On nomme p le pic correspondant au dernier maximum de u . Il y a $\max d(u) - k$ pics à la hauteur $h - 1$. Dans $f(v)$, le pic ou la double descente où a été supprimé p porte les marques d'indice $\max d(u) - 1$ car les pics à la hauteur h portent les indices 0 à $k - 2$ et les $\max d(u) - k$ pics à la hauteur $h - 1$, déjà marqués dans u , les indices de $k - 1$ à $\max d(u) - k + (k - 1)$. Donc $\Phi_m(s_1, \dots, s_{n+1}) = f_m(u)$. De plus $\Phi_m^+(s_1, \dots, s_{n+1}) = f_m(u)$ car si l'insertion selon le mode $(xB\bar{B}, \bar{x}A\bar{A})$ conduit au marquage de p par $x(\max d(u))(\overline{\max d(u)})\bar{x}(\max d(u) - 1)(\overline{\max d(u) - 1})$, la permutation $\sigma \circ \bar{\sigma}$ conduit à $x00\bar{x}(\max d(u))(\overline{\max d(u)})$ et les autres marques sont décalées pour prendre en compte le pic p .

Pour tout chemin de u de longueur $n + 1$, on a trouvé une unique suite quasi-décroissante s_1, \dots, s_{n+1} telle que $\Phi_m(s_1, \dots, s_{n+1}) = u$ donc Φ_m est bijective lorsqu'elle est restreinte aux chemins de longueur inférieure ou égale à $n + 1$.

□

Dans la preuve précédente, la bijectivité de l'application semble pouvoir être garantie grâce à un lemme plus général car la permutation circulaire des marques est liée à une bijection sur les suites quasi-décroissantes. N'ayant pu formaliser l'argument, ce lemme reste une conjecture.

Conjecture 2.9.1 *Soit m un mode d'insertion. L'application définie par le paramètre $\sigma \circ \bar{\sigma} \circ m$ est bijective si et seulement si l'application définie par le paramètre m est bijective.*

Grâce à cette conjecture et à la proposition 2.4 on établirait la bijectivité de douze nouvelles applications et il resterait à comprendre les paramètres analogues de la dernière descente qui en découlerait.

La statistique de Haglund

Haglund a proposé un autre paramètre sur les chemins de Dyck dont la distribution jointe à celle de l'aire est la même que la bivariate de l'aire et de la statistique de Haiman. On s'inspire du nom proposé par Haglund pour nommer cette statistique le *moment des ricochets* d'un chemin de Dyck. Pour définir cette statistique sur un chemin de Dyck, on imagine que l'on lance une balle à partir du dernier sommet du chemin de Dyck et parallèlement à la dernière descente; on suppose que cette balle rebondit, "ricoché", sur un plan horizontal à chaque fois qu'elle rencontre un sommet du chemin précédé d'un pas montant ou bien l'axe horizontal. La trajectoire de la balle est un chemin de Dyck de la même longueur que le chemin de Dyck d , on le nomme la *trajectoire* du chemin. On considère maintenant les distances entre un ricochet sur l'axe horizontal et le point d'arrivée de la balle; le premier et le dernier sommet de la trajectoire ne sont pas des ricochets. Le moment d'un contact, ricochet, sur l'axe est la moitié de sa distance à la fin du chemin. La somme des moments des ricochets est le moment des ricochets du chemin de Dyck.

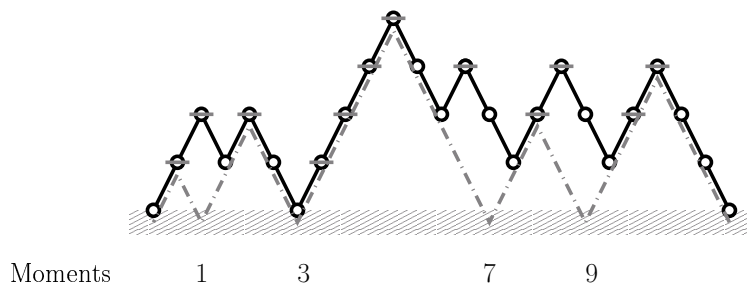


FIG. 2.32 – Moments des ricochets d'un chemin de Dyck

Sur l'exemple de la figure 2.32, on a un chemin de Dyck dont le moment des ricochets est $1 + 3 + 7 + 9 = 20$.

On définit pour le moment des ricochets un paramètre qui sera l'analogue de la longueur de la dernière descente pour l'aire des chemins de Dyck. La figure 2.33 illustre la définition de ce paramètre. Pour définir les ricochets on a autorisé la balle à rebondir sur tous les sommets après un pas montant. On va définir une autre trajectoire, la balle ne rebondissant maintenant plus que sur les sommets qui sont des pics du chemin. Cette balle peut donc éventuellement croiser le chemin sur un sommet qui est une double montée. Le *nombre de ricochets sur les pics* d'un chemin de Dyck est le nombre de ricochets de cette trajectoire sur un pic du chemin.

Proposition 2.10 *Il existe une bijection qui transforme une suite quasi-décroissante de poids p et de dernier élément k en un chemin de Dyck de moment de ricochet p et de nombre de ricochets sur les pics $k + 1$.*

Pour définir cette bijection on va utiliser uniquement une fonction de marquage. Une fonction de marquage f est *complète* si pour tout chemin de Dyck u dont la construction

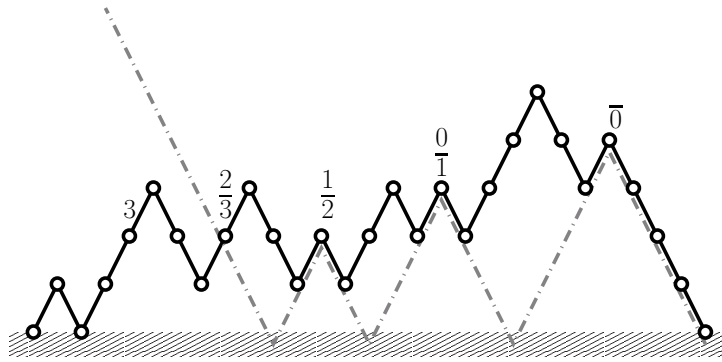


FIG. 2.33 – Ricochets sur les pics d'un chemin de Dyck

$f(u)$ est de rang k , pour tout $i \leq k$, l'insertion des pas x et \bar{x} dans les marques i et \bar{i} conduit à un chemin de Dyck v_i tel que le rang de la construction $f(v_i)$ est $i + 1$. À l'aide d'une fonction de marquage complète f on peut définir une application I_f entre les suites quasi-décroissantes et les chemins de Dyck qui préserve la taille des objets. L'algorithme 6 est celui qui accepte en entrée la suite s et retourne $I_f(s)$.

Entrée : Une suite quasi-décroissante $s = s_1, s_2, \dots, s_n$;

Paramètre : Une fonction de marquage f complète.

Chemin := ϵ ;

pour i allant de 1 à n **faire**

Insérer dans *Chemin* les pas x et \bar{x} indiqués par les marques d'indice s_i dans la construction $f(\text{Chemin})$;

fin

Sortie : Le chemin *Chemin* ;

Algorithme 6: Algorithme paramétré par une fonction de marquage complète

La caractère complet de la fonction de marquage suffit à garantir que cet algorithme associe un chemin de Dyck de demi-longueur n à toute suite quasi-décroissante de longueur n . La différence principale avec les algorithmes précédents est que ici on ne se soucie pas de réactualiser localement les marques. Cette algorithme est une généralisation des algorithmes précédents lorsque ceux-ci conduisent à des bijections. En effet, pour un mode d'insertion m compatible avec la fonction de marquage f tel que Φ_m soit bijectif on a $\Phi_m = I_f$ puisque pour toute suite quasi-décroissante s , $\Phi_m^+(s) = f(s)$.

On définit une fonction de marquage f_{haglund} pour décrire les insertions de la bijection. À la différence des bijections précédentes, cette fonction de marquage n'est pas codée par l'insertion de marques au voisinage des deux pas insérés. Pour marquer un chemin de Dyck on considère les k ricochets sur les pics de droite à gauche. Sur le sommet du premier ricochet, on place la marque $\bar{0}$. Pour les pics suivants, on place le facteur $(i - 2)(i - 1)$ sur le i^{eme} sommet. Sur la dernière intersection de la trajectoire et du chemin de Dyck on place le facteur $\bar{k}k - 1$. Pour placer la dernière marque k , on suppose que la balle ricoche encore

une fois sur le chemin au lieu de sortir, la prochaine intersection avec le chemin de Dyck est le sommet où on insère le facteur k . La figure 2.33 est un exemple de la construction de ce marquage.

Lemme 2.10.1 *Soit u un chemin de Dyck dont la construction $f_{haglund}(u)$ est de rang k . L'insertion de pas dans les marques d'indice i de $f_{haglund}$ conduit à un chemin v_i dont le moment des ricochets est augmenté de i par rapport à u . De plus $f_{haglund}$ est une fonction de marquage complète.*

Preuve. (Lemme 2.10.1) On observe d'abord sur un exemple que l'insertion des deux pas à la lecture de la valeur i dans la suite quasi-décroissante augmente la valeur du moment des ricochets de i . L'exemple de la figure 2.34 traite de l'insertion de pas correspondant à la lecture de la valeur 2 dans la suite quasi-décroissante. Plus formellement, l'insertion

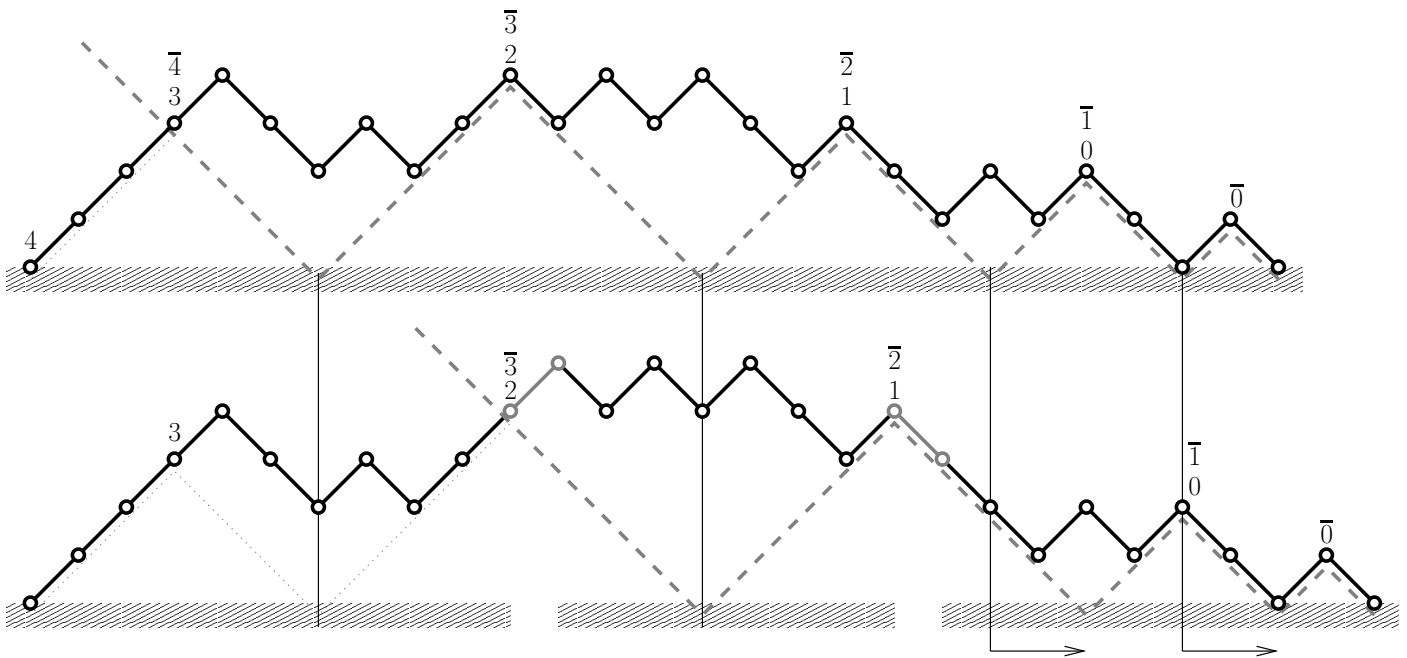


FIG. 2.34 – Une insertion de pas pour le moment des ricochets

conduit à un décalage des i premiers ricochets de la balle sur l'axe de 2 unités vers la droite et laisse la fin de la trajectoire invariante. Donc le moment des ricochets augmente de i . D'autre part, la trajectoire des ricochets sur les pics est modifiée puisque le sommet où se trouvait la marque i est maintenant une double-descente; le marquage du nouveau chemin conduit à une construction de Dyck de rang $i + 1$. \square

À ce stade, une simple récurrence sur la taille des objets, permet de constater que l'application $I_{f_{haglund}}$ transforme le poids d'une suite quasi-décroissante s en le moment des ricochets du chemin de Dyck $I_{f_{haglund}}(s)$. Pour conclure il reste à prouver que $I_{f_{haglund}}$ est bijective.

Lemme 2.10.2 *L'application $I_{f_{haglund}}$ est bijective.*

Preuve. On le montre par récurrence sur la longueur n des suites quasi-décroissantes. Dans cette preuve f est la fonction de marquage $f_{haglund}$.

Hypothèse de récurrence : L'application I_f est bijective pour l'ensemble des suites quasi-décroissantes de longueur inférieure à n . Pour toute suite s_1, \dots, s_n , $n \geq 1$, $I_f(s_1, \dots, s_n) = v$ est un mot de Dyck tel que $f(v)$ est une construction de rang $s_n + 1$.

Cas de base : Pour $n = 0$ on a $I_f(\emptyset) = \epsilon$ et $f(\epsilon) = 0\bar{0}$. Pour $n = 1$ on a $I_f(0) = x\bar{x}$ et $f(x\bar{x}) = 1\bar{1}0x\bar{0}\bar{x}$. L'hypothèse de récurrence est donc satisfaite pour $n \leq 1$.

Hérédité : On suppose que l'hypothèse de récurrence est satisfaite au rang $n \geq 1$. Soit u un chemin de Dyck de demi-longueur $n + 1$. On note k le rang de la construction $f(u)$ qui est au moins 1 puisque $n \geq 1$.

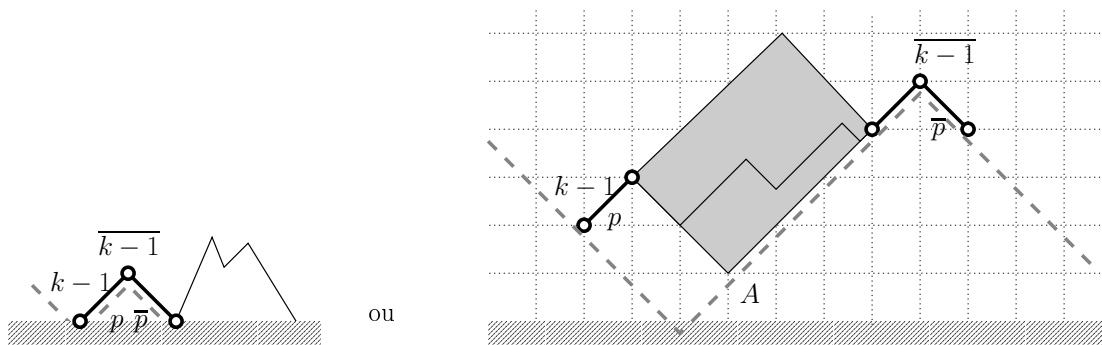


FIG. 2.35 – Les derniers pas insérés dans un chemin par $f_{haglund}$

La figure 2.35 tente d'illustrer la discussion sur le choix des deux pas à supprimer de u . Dans $f(u)$, la marque $\overline{k-1}$ est sur le k^{eme} rebond sur un pic de la balle donc il y a un pas descendant, noté \bar{p} , après cette marque. D'autre part, la marque $k-1$ est sur le sommet de u où pour la première fois la balle ne rebondit pas sur un pic : soit c'est le premier sommet du chemin u , soit c'est une double montée ; dans les deux cas il y a un pas montant p après la marque $k-1$. On considère le chemin v obtenu par la suppression des pas p et \bar{p} de u . Le chemin v est bien un chemin de Dyck car soit $p\bar{p}$ est un pic de u (et alors ce pic est au début du chemin) soit la trajectoire de la balle entre p et \bar{p} impose que la hauteur minimale soit 1 (le sommet A sur la figure 2.35 représente le pire cas). Comme v est de longueur n il existe, d'après l'hypothèse de récurrence, une suite quasi-décroissante $s = s_1, \dots, s_n$ telle que $I_f(s) = v$ et $s_n + 1$ est le rang de la construction $f(v)$. Pour finir la preuve on doit discuter selon que les deux pas p et \bar{p} sont consécutifs dans u ou non.

p et \bar{p} sont consécutifs dans u : La trajectoire de la balle impose que le pic $p\bar{p}$ soit un préfixe de u . Dans ce cas, v est une construction de rang $k-1$ car la disparition de $p\bar{p}$ a supprimé un rebond par rapport à u . Donc d'après l'hypothèse de

réurrence $s_n = k - 2$. Donc $s_1, \dots, s_n, (k - 1)$ est une suite quasi-décroissante. De plus dans ce cas, la trajectoire des rebonds sur pics de v se termine sur le premier sommet de v et les marques d'indices $k - 1$ sont sur ce sommet. Donc $I_f(s_1, \dots, s_n, (k - 1)) = u$ donc la surjectivité de I_f est établie pour u . Comme le rang de u est k la seconde partie de l'hypothèse de récurrence est vérifiée.

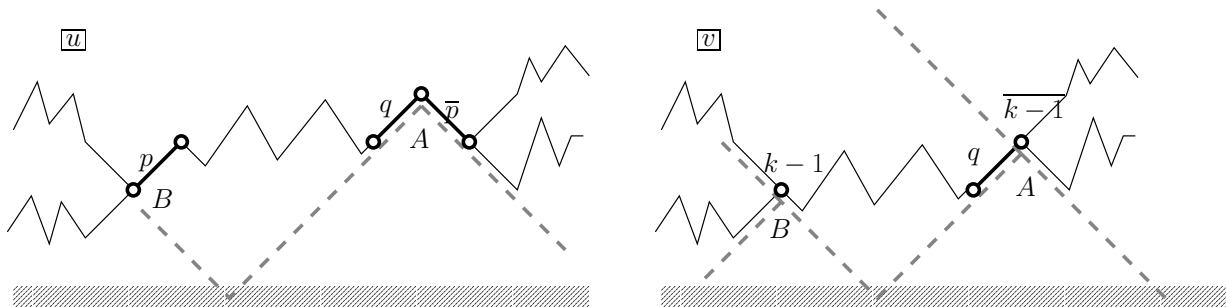


FIG. 2.36 – p et \bar{p} non consécutifs

p et \bar{p} ne sont pas consécutifs dans u : La figure 2.36 tente de représenter le raisonnement. La trajectoire des rebonds sur les pics n'est pas modifiée jusqu'au pic que le pas \bar{p} formait avec le pas montant q qui, dans ce cas, est différent de p . On nomme B le sommet avant le pas p dans u et A le sommet avant le pas \bar{p} . Ce pic $q\bar{p}$ étant le $k^{\text{ème}}$ ricochet sur un pic dans u , il y a au moins $k - 1$ ricochet sur un pic dans v donc $f(v)$ est une construction de rang au moins $k - 1$. Donc $s_n \geq k - 2$ d'après l'hypothèse de récurrence et la suite $s_1, \dots, s_n, (k - 1)$ est une suite quasi-décroissante. Ensuite dans $f(v)$ les marques d'indice $k - 1$ se placent sur les sommets B et A donc on a $I_f(s_1, \dots, s_n, k - 1) = u$ et on rappelle que u est une construction de rang k . Donc l'hypothèse de récurrence est vérifiée hormis l'injectivité de I_f .

On a montré que I_f est surjective pour les objets de taille $n + 1$, or il y a autant d'objets de taille $n + 1$, I_f est bijective et l'hypothèse de récurrence est maintenant complètement vérifiée. □

La réactualisation des marques des constructions après la lecture de la valeur i ne correspond pas à un mode d'insertion. Certes la marque i est remplacée par le facteur $\overline{i+1}x$ et la marque \bar{i} par le facteur $\overline{i}x$ ce qui permet de dire que cette insertion est localement réversible. Mais la marque $i + 1$ n'est pas placée au voisinage d'un pas inséré, il faut utiliser un ricochet supplémentaire sur l'axe et l'intersection avec le chemin pour obtenir sa position. Pour coder cette insertion par un algorithme paramétré par des réécritures locales, on peut proposer d'ajouter des marques supplémentaires \bar{i} marquant les ricochets oubliés mais on préfère d'abord étudier les algorithmes déjà proposés.

L'algorithme paramétré par les règles de réécriture permet d'obtenir un grand nombre

de statistiques distribuées comme l'aire sous les chemins de Dyck. Une réponse à la symétrie de la bistatistique proposée par Haiman et Haglund pourrait passer par la découverte d'une bistatistique naturellement symétrique. L'étude systématique des bistatistiques est envisagée mais n'a pas été réalisée.

2.2.4 Codage de bijections entre classes combinatoires comptées par la suite de Catalan

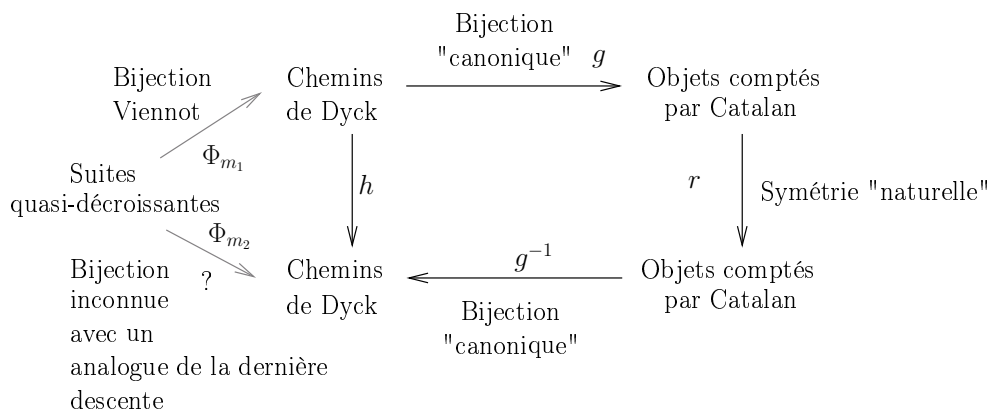


FIG. 2.37 – Une symétrie naturelle peut induire un paramètre analogue à la dernière descente

Pour des objets comptés par la suite de Catalan il existe souvent une bijection "canonique" g avec les chemins de Dyck. Ces objets possèdent parfois des symétries naturelles qui sont autant de bijections "simples", notées r . Grâce aux bijections canoniques on en déduit la bijection $h = g^{-1} \circ r \circ g$ entre les chemins de Dyck. Pour coder la bijection h sur les mots de Dyck on cherche deux bijections Φ_{m_1} et Φ_{m_2} entre les suites décroissantes et ces chemins telles que $h = \Phi_{m_2} \circ \Phi_{m_1}^{-1}$.

Pour illustrer ce principe on s'intéresse d'abord au codage de la symétrie par rapport à la diagonale des polyominos parallélogrammes puis à une involution de Kreweras sur les chemins de Dyck. Dans un premier temps on cherche la fonction de marquage complète des chemins de Dyck utilisée pour coder ces deux bijections. Ensuite on propose une nouvelle variante de l'algorithme 3 permettant de trouver des modes d'insertions compatibles avec ces fonctions de marquage.

Codage d'une symétrie sur les polyominos parallélogrammes

Un *polyomino parallélogramme* est une surface bornée sur le réseau carré dont la frontière est définie par deux chemins issus de l'origine, formés de pas élémentaire Nord et Est, tels que les extrémités des deux chemins soient les seules intersections entre les deux chemins. Sur la droite de la figure 2.38 les deux chemins entre A et B définissent un

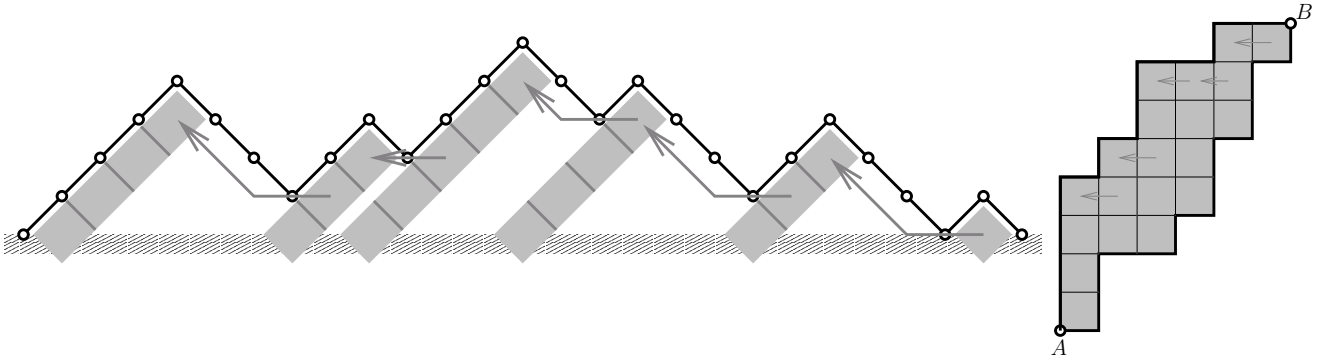


FIG. 2.38 – Chemins de Dyck et polyominos parallélogrammes

polyomino parallélogramme. Le *périmètre* d'un polyomino parallélogramme est le nombre de pas formant les deux chemins. Sur la figure, le polyomino est de périmètre 28. Delest et Viennot [DV84] ont prouvé que les polyominos de périmètre $2p + 2$ sont en bijection avec les chemins de Dyck de demi-longueur p . La figure 2.38 illustre leur bijection g : partant d'un chemin de Dyck de demi-longueur p , les colonnes du futur polyomino se placent sous chaque pic et leur longueur est la hauteur de ces pics. On forme un polyomino en collant ces colonnes de gauche à droite, de telle sorte que la hauteur du creux entre deux pics code la position relative des deux colonnes : la case la plus haute de la colonne de droite est à la même hauteur que la case de la colonne de gauche sous le pas montant formant le creux. Autrement dit, si le creux est de hauteur h , les deux colonnes ont $h + 1$ cases adjacentes. On remarque que la longueur du chemin supérieur entre A et B est $p + 1$ donc le périmètre du polyomino parallélogramme est $2p + 2$.

L'ensemble des polyominos parallélogrammes est stable par la réflexion par rapport à la première diagonale. Dans le coin inférieur gauche de la figure 2.39 on donne un exemple de polyomino, en dessous de la diagonale en pointillés, et de son image, au dessus de la diagonale, par la réflexion. Cette réflexion r est donc une involution sur l'ensemble des polyominos parallélogrammes. On en déduit que $g^{-1} \circ r \circ g$ est une involution sur les chemins de Dyck. On cherche maintenant à coder cette bijection sous la forme $\Phi_{m_2} \circ \Phi_{(xB\bar{B}, \bar{x}A\bar{A})}^{-1}$ avec deux bijections entre les chemins de Dyck et les suites décroissantes. Redisons cela : on veut trouver une bijection Φ_{m_2} telle que

$$g^{-1} \circ r \circ g = \Phi_{m_2} \circ \Phi_{(xB\bar{B}, \bar{x}A\bar{A})}^{-1} \quad (43)$$

ce qui implique par composition à droite par $\Phi_{(xB\bar{B}, \bar{x}A\bar{A})}$ que Φ_{m_2} est défini par

$$\Phi_{m_2} = g^{-1} \circ r \circ g \circ \Phi_{(xB\bar{B}, \bar{x}A\bar{A})}. \quad (44)$$

La bijection Φ_{m_2} est donc définie, il reste à décider si cette bijection est engendrée par insertion et si oui à trouver des modes d'insertion compatibles. Pour conjecturer l'éventuelle fonction de marquage associée à Φ_{m_2} on fait le calcul, illustré par la figure 2.39, pour la

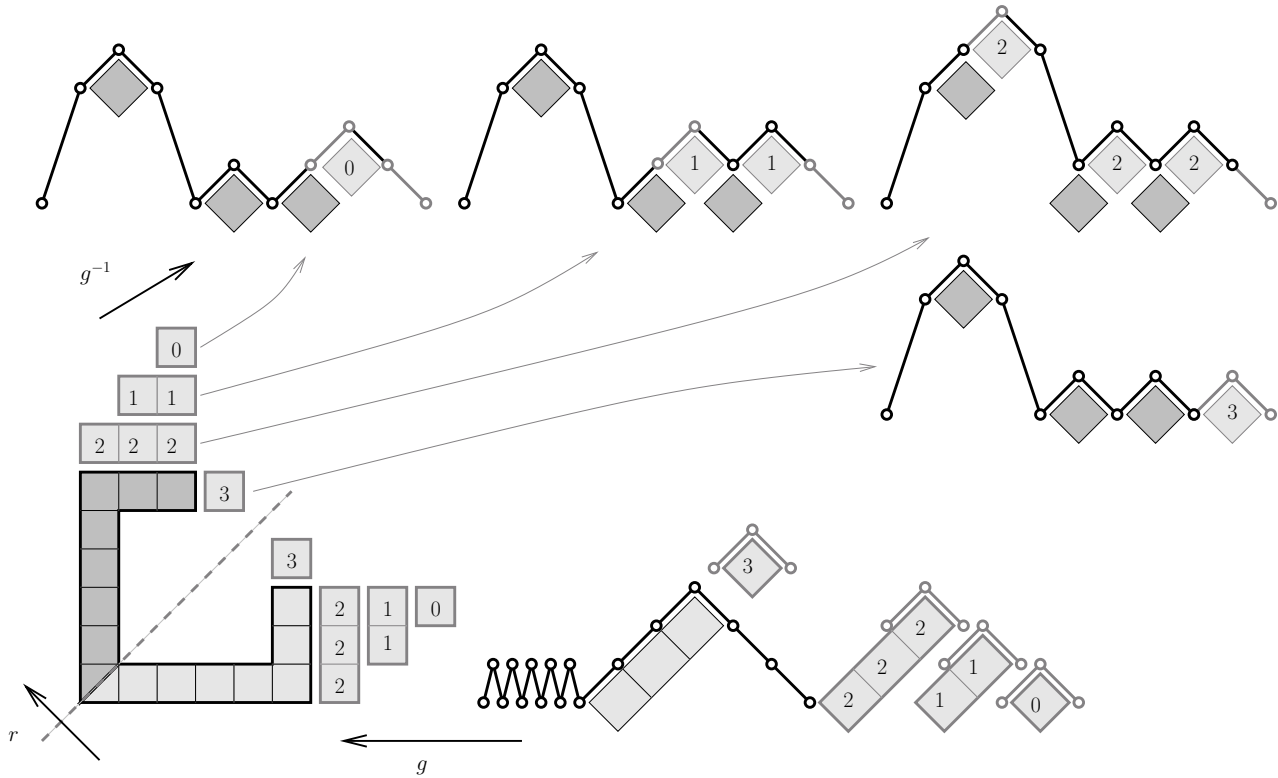


FIG. 2.39 – Symétrie selon la diagonale principale des polyominoes parallélogrammes

suite quasi-décroissante 0, 0, 0, 0, 0, 0, 1, 2 et ses prolongements par les valeurs 0, 1, 2 et 3. Sur la figure, on étiquete par i les cases des colonnes ajoutées lors de la lecture de la valeur i dans la suite quasi-décroissante.

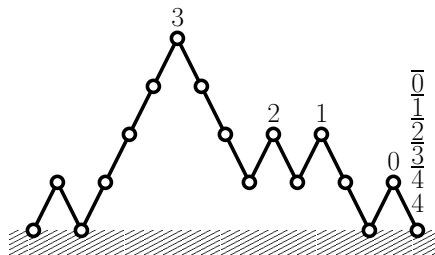


FIG. 2.40 – La fonction de marquage $f_{reflexion}$

Ces calculs, et d'autres de cet acabit dont on dispense le lecteur, permettent de conjecturer la fonction de marquage $f_{reflexion}$ pour le mode Φ_{m_2} . Le rang k de la construction est le nombre de pics après la dernière double-montée du chemin. On numérote ces pics en partant de la fin du chemin. Sur le pic numéroté i on place la marque $i - 1$. Sur le dernier sommet du chemin on place les marques $\overline{k k k - 1} \dots \overline{10}$. La figure 2.40 donne un exemple.

Lemme 2.10.3 *La fonction de marquage $f_{reflexion}$ est complète.*

Preuve.

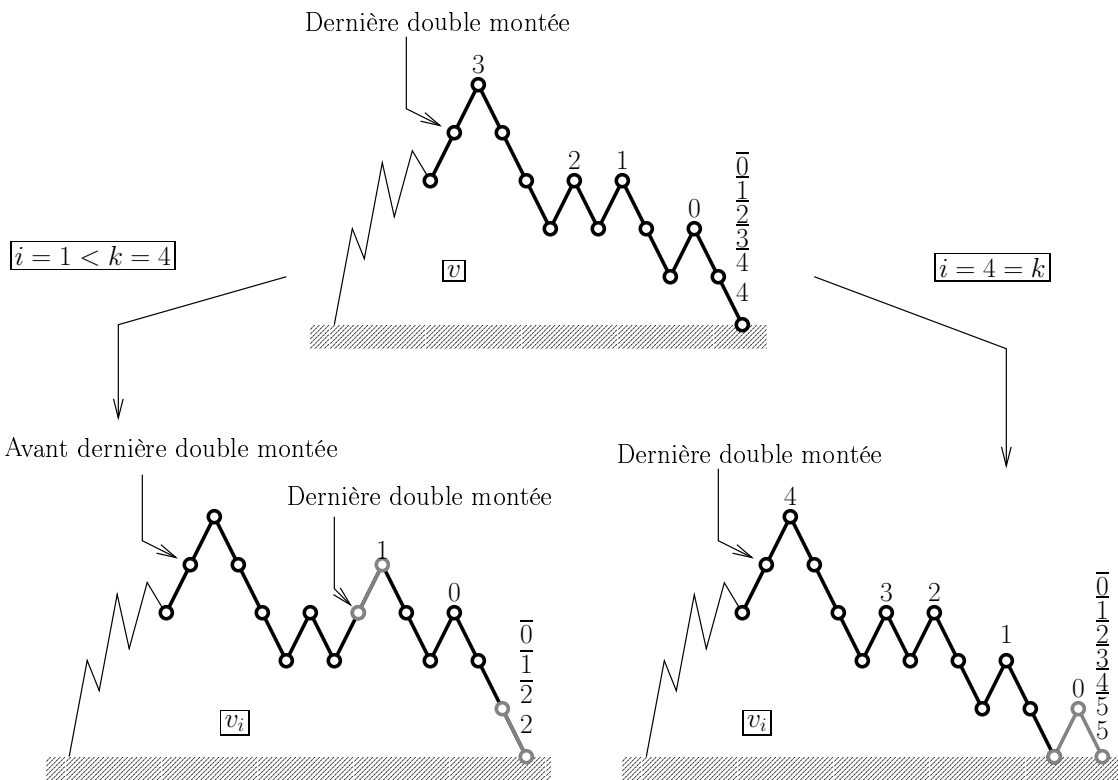


FIG. 2.41 – Une insertion selon la fonction de marquage $f_{reflexion}$.

Soit v un chemin de Dyck dont la construction $f_{reflexion}(v)$ est de rang k . On considère l'insertion de pas sur les marques d'indice $i \leq k$ qui conduit au chemin v_i . La figure 2.41 illustre les deux cas discutés :

- $i = k$: On insère un pic à la fin du chemin, $f(v_i)$ est donc une construction de rang $k + 1$ puisqu'on a ajouté un pic après la dernière double montée
- $i < k$: On insère un pas montant avant le pic portant la marque i ce qui crée une double montée. Les pics marqués par $j > i$ dans v ne sont donc plus marqués dans v_i qui est une construction de rang $i + 1$.

□

La fonction de marque $f_{reflexion}$ est complète donc l'application $I_{f_{reflexion}}$ est bien définie. Le lemme suivant compare cette application à $g^{-1} \circ r \circ g \circ \Phi_{(xB\bar{B}, \bar{x}A\bar{A})}$:

Lemme 2.10.4 *L'application $I_{f_{reflexion}}$ est égale à $g^{-1} \circ r \circ g \circ \Phi_{(xB\bar{B}, \bar{x}A\bar{A})}$.*

Preuve. La preuve est une récurrence sur la longueur n de suites quasi-décroissantes.

Hypothèse de récurrence : Pour toute suite quasi-décroissante $s = s_1, \dots, s_n$ de longueur n on a $g^{-1} \circ r \circ g \circ \Phi_{(xB\bar{B}, \bar{x}A\bar{A})}(s) = I_{f_{reflexion}}(s)$.

Cas de base : Pour $n = 0$ et $n = 1$, un petit calcul permet de vérifier l'hypothèse de récurrence.

Hérédité : On suppose l'hypothèse de récurrence satisfaite au rang n . Soit $s = s_1, \dots, s_n$ une suite quasi-décroissante de longueur n et u le chemin $g^{-1} \circ r \circ g \circ \Phi_{(xB\bar{B}, \bar{x}A\bar{A})}(s) = I_{f_{reflexion}}(s)$. On choisit une valeur s_{n+1} inférieure ou égale à $s_n + 1$ et on discute, selon la valeur relative de s_n et s_{n+1} , des insertions par $g^{-1} \circ r \circ g \circ \Phi_{(xB\bar{B}, \bar{x}A\bar{A})}$ et $I_{f_{reflexion}}$. Une remarque clé est que la bijection g^{-1} transforme la longueur de la ligne supérieure d'un polyomino p en le nombre de pics après la dernière double-montée du chemin $g^{-1}(p)$.

Si $s_{n+1} < s_n + 1$: On crée un creux à la hauteur s_{n+1} dans le chemin de Dyck $\Phi_{(xB\bar{B}, \bar{x}A\bar{A})}(s_1, \dots, s_n)$. Cela a pour conséquence d'ajouter une colonne de hauteur $s_{n+1} + 1$ à la droite du polyomino parallélogramme que l'on obtient par g . Cette colonne devient une ligne supplémentaire de taille $s_{n+1} + 1$ au-dessus du polyomino résultant de la réflexion r . L'ajout de cette ligne correspond à la surélévation d'une hauteur 1 des $s_{n+1} + 1$ derniers pics sur le chemin u . En termes d'insertion, on insère bien un pas montant dans le sommet marqué par s_{n+1} et un pas descendant à la fin du chemin ce qui correspond à la fonction de marquage $f_{reflexion}$.

Si $s_{n+1} = s_n + 1$: On augmente de 1 la longueur de la dernière descente du chemin $\Phi_{(xB\bar{B}, \bar{x}A\bar{A})}(s_1, \dots, s_n)$. Cela agrandit d'une case vers le haut la dernière colonne à droite du polyomino obtenu par g . Sur le polyomino obtenu par la réflexion r , la ligne supérieure augmente d'une case vers la droite. Sur le chemin u , cette transformation conduit bien à l'insertion d'un pic à la fin du chemin et correspond à la fonction de marquage $f_{reflexion}$.

□

La prochaine étape dans la recherche de Φ_{m_2} est de trouver un algorithme de réécriture des marques qui corresponde aux insertions selon la fonction de marquage $f_{reflexion}$. Pour des raisons qui apparaîtront dans la partie suivante on retarde de quelques pages cette étude.

Codage d'une bijection de Kreweras

Dans [Lal93], Lalanne présente une involution sur les mots de Dyck due à Kreweras. On rappelle brièvement cette bijection puis on cherche à la coder elle aussi par deux applications définies par insertion dont une est $\Phi_{(xB\bar{B}, \bar{x}A\bar{A})}$.

La figure 2.42 donne un exemple de cette involution h . On considère un chemin de Dyck u . On définit comme suit le chemin $h(u)$. En chaque sommet d'une double montée

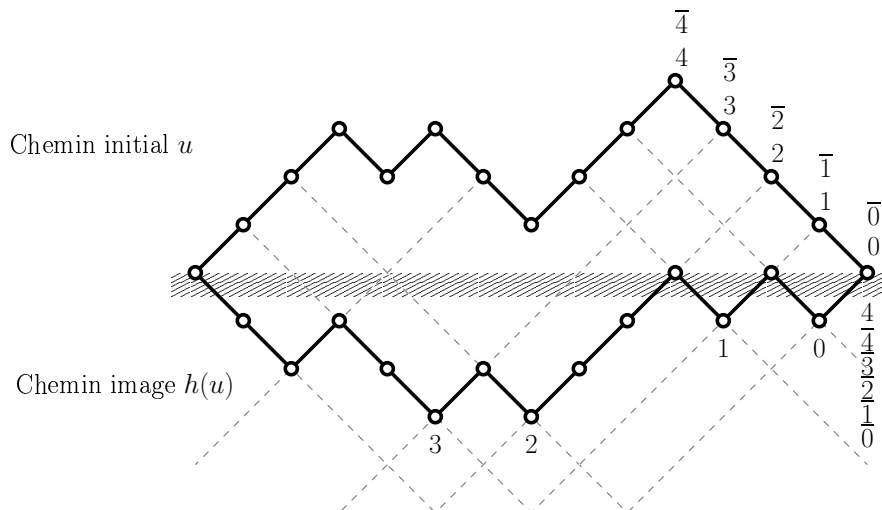


FIG. 2.42 – L’involution de Kreweras sur les chemins de Dyck

ou d’une double descente du chemin u on trace la demi-droite normale aux pas incidents au sommet qui intersecte l’axe horizontal. Ces demi-droites sont en pointillés gris sur la figure, le début et la fin du chemin sont des cas dégénérés de double montée et double descente. Partant du sommet au début du chemin u , on obtient le chemin $h(u)$ en suivant la demi-droite et à chaque intersection entre demi-droites on change de demi-droite : on s’éloigne, respectivement se rapproche, de l’axe horizontal sur une demi-droite issue d’une double-montée, respectivement une double-descente.

Des tâtonnements similaires à ceux de l’étude de la réflexion sur les polyominos parallélogrammes conduisent à conjecturer la fonction marquage correspondant aux insertions dans $h(u)$ à la suite d’une insertion d’un pic dans la dernière descente du chemin u . Cette fonction de marquage est $f_{reflexion}$ celle que l’on a également utilisée pour coder la réflexion ! La proposition suivante confirme cette coïncidence.

Lemme 2.10.5 *L’application $I_{f_{reflexion}}$ définie par insertion selon la fonction de marquage $f_{reflexion}$ est égale à la bijection $h \circ \Phi_{(xB\bar{B}, \bar{x}A\bar{A})}$.*

Preuve.

La preuve est une récurrence sur la longueur des suites quasi-décroissantes.

Hypothèse de récurrence : Pour toute suite quasi-décroissante s_1, \dots, s_n on a

$$I_{f_{reflexion}}(s) = h \circ \Phi_{(xB\bar{B}, \bar{x}A\bar{A})}(s).$$

Cas de base : Pour $n = 0$ et $n = 1$ un petit calcul suffit.

Hérédité : On suppose l’hypothèse de récurrence valable pour n . Soit $s = s_1, \dots, s_n$ une suite quasi-décroissante et $v = I_{f_{reflexion}}(s) = h \circ \Phi_{(xB\bar{B}, \bar{x}A\bar{A})}(s)$. On compare les insertions de pas selon $h \circ \Phi_{(xB\bar{B}, \bar{x}A\bar{A})}(s)$ et $f_{reflexion}$. Soit $s_{n+1} \leq s_n + 1$, on discute des valeurs relatives de s_n et s_{n+1} . La figure 2.43 illustre le propos. Sur ces schémas

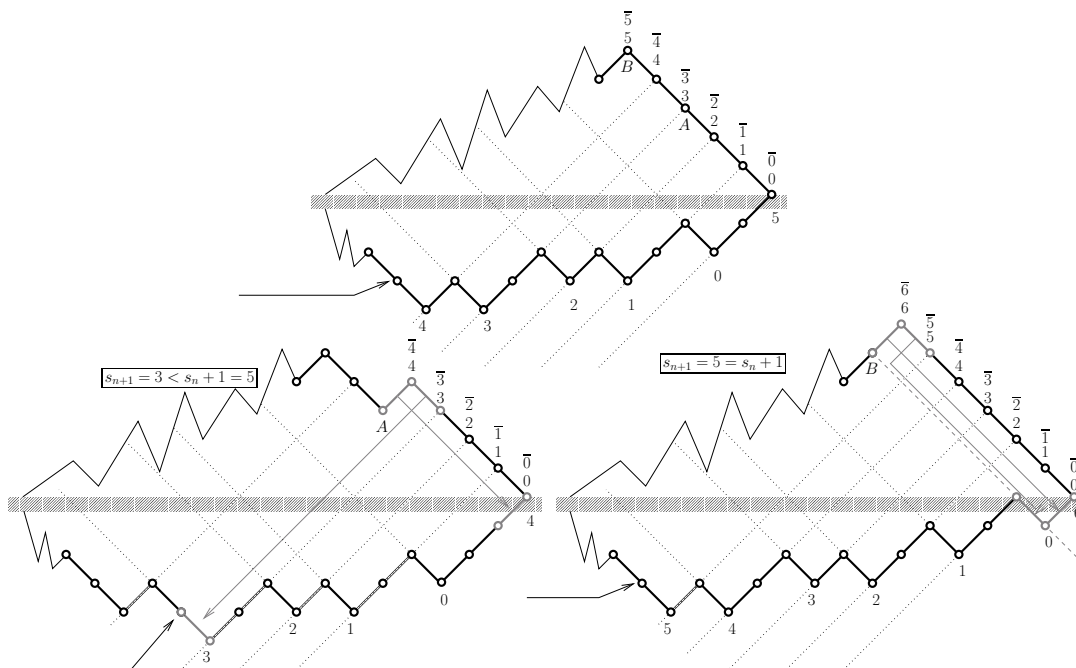


FIG. 2.43 – Insertion de pas pour une involution de Kreweras

le chemin inférieur est v et le chemin supérieur est $h^{-1}(v)$. À la lecture de s_{n+1} selon $h \circ \Phi_{(x\overline{B}\overline{B}, \overline{x}A\overline{A})}$ on insère un pic à la hauteur s_{n+1} dans la dernière descente de $h^{-1}(v)$. On veut montrer que $h \circ \Phi_{(x\overline{B}\overline{B}, \overline{x}A\overline{A})}(s_1, \dots, s_{n+1})$ coïncide avec l'insertion des pas x et \overline{x} dans v aux positions d'indice s_{n+1} .

Si $s_{n+1} < s_n + 1$: L'insertion du pic décale la double descente issue du sommet A portant les marques d'indice s_{n+1} dans $h^{-1}(v)$. Sur la figure on obtient le chemin dans le coin supérieur droit en choisissant $s_{n+1} = 3$. Ce décalage correspond à insérer dans v un pas montant avant le pic numéroté s_{n+1} et un pas descendant à la fin du chemin. Cela correspond à la fonction de marquage $f_{\text{reflexion}}$.

Sinon $s_{n+1} = s_n + 1$: L'insertion du pic crée une double montée B dans le chemin $h^{-1}(v)$. Sur la figure, le résultat de l'insertion s'obtient sur l'exemple en choisissant $s_{n+1} = 4$. La demi-droite issue de cette nouvelle double-montée B implique l'apparition d'un pic à la fin du chemin v . Cela correspond à la fonction de marquage $f_{\text{reflexion}}$.

□

Il y a donc un lien entre l'involution de Kreweras sur les chemins de Dyck, la bijection canonique de Delest et Viennot entre chemins de Dyck et polyominos parallélogrammes et la réflexion par rapport à la première diagonale des polyominos parallélogrammes.

Proposition 2.11 *L'involution de Kreweras h est égale à l'involution $g^{-1} \circ r \circ g$ où g est*

la bijection "canonique" entre les chemins de Dyck et les polyominos parallélogrammes et r est la réflexion par rapport à la première diagonale de ces polyominos.

Preuve. D'après les lemmes 2.10.5 et 2.10.4 on a

$$h \circ \Phi_{(xB\bar{B}, \bar{x}A\bar{A})} = I_{f_{reflexion}} = g^{-1} \circ r \circ g \circ \Phi_{(xB\bar{B}, \bar{x}A\bar{A})}.$$

Comme $\Phi_{(xB\bar{B}, \bar{x}A\bar{A})}$ est une bijection on en déduit par composition à droite par $\Phi_{(xB\bar{B}, \bar{x}A\bar{A})}^{-1}$ que

$$h = g^{-1} \circ r \circ g.$$

□

Variante de l'algorithme pour la fonction de marquage

La bijection $h \circ \Phi_{(xB\bar{B}, \bar{x}A\bar{A})}$ entre les suites quasi-décroissantes et les chemins de Dyck peut se définir grâce à la fonction de marquage $f_{reflexion}$; elle peut aussi se définir à l'aide de réécriture locales. Pour mettre à jour les marques de la construction de Dyck on doit utiliser l'algorithme 7, une nouvelle variante de l'algorithme général 3. Cet algorithme introduit une nouvelle distinction, on utilise deux modes d'insertions différents selon que $s_i \leq s_{i-1}$ ou que $s_i = s_{i-1} + 1$.

Entrée : Une suite quasi-décroissante $s = s_1, s_2, \dots, s_n$;

Paramètre : Deux modes d'insertions m_1 et m_2

CheminMarque := $0\bar{0}$;

pour i allant de 1 à n **faire**

Supprimer les marques d'indices supérieurs à s_i ;

si $i > 1$ et $s_i = s_{i-1} + 1$ **alors**

Insérer les marques selon le mode m_1 ;

sinon

Insérer les marques selon le mode m_2 ;

fin

fin

Supprimer les marques de *CheminMarque* ;

Sortie : Le chemin *CheminMarque* ;

Algorithme 7: Algorithme discutant la position relative de s_i et s_{i-1}

Pour obtenir des constructions qui correspondent à la fonction de marquage $f_{reflexion}$ on paramètre cet algorithme par $m_1 = \sigma \circ (xB, \bar{x}A\bar{B}\bar{A})$ et $m_2 = (xA, \bar{x}B\bar{B}\bar{A})$. Comme pour les algorithmes périodiques on définit une syntaxe pour ce paramétrage que l'on note $[m_1|m_2]$.

Proposition 2.12 *Le mode $[\sigma \circ (xB, \bar{x}A\bar{B}\bar{A})|(xA, \bar{x}B\bar{B}\bar{A})]$ est compatible avec la fonction de marquage $f_{reflexion}$:*

$$I_{f_{reflexion}} = \Phi_{[\sigma \circ (xB, \bar{x}A\bar{B}\bar{A})|(xA, \bar{x}B\bar{B}\bar{A})]}.$$

Preuve. On fait une récurrence sur la longueur n des suites quasi-décroissantes. Dans cette

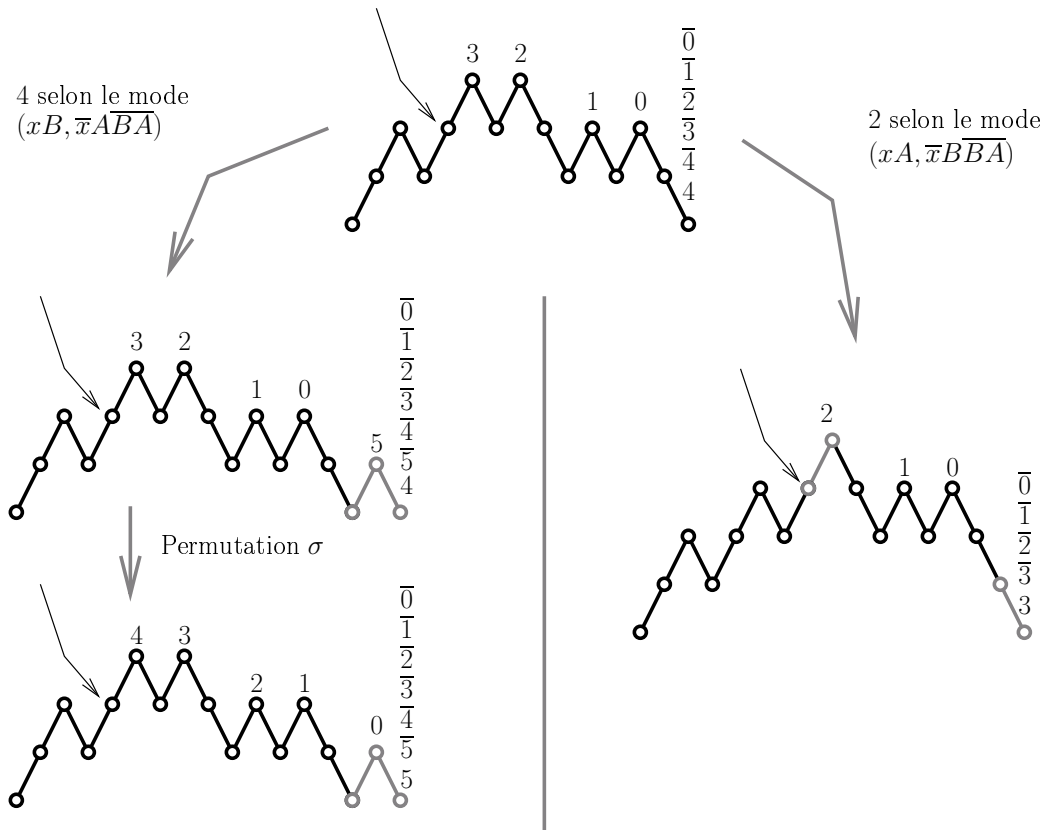


FIG. 2.44 – Compatibilité du $[\sigma \circ (xB, \bar{x}A\bar{B}\bar{A})|(xA, \bar{x}B\bar{B}\bar{A})]$ avec $f_{reflexion}$.

preuve on pose $m = [\sigma \circ (xB, \bar{x}A\bar{B}\bar{A})|(xA, \bar{x}B\bar{B}\bar{A})]$ et $f = f_{reflexion}$.

Hypothèse de récurrence : Pour toute suite $s = s_1, \dots, s_n$ de longueur n , on a $\Phi_m^+(s) = f(\Phi_m^+(s))$.

Cas de Base : Pour $n = 0$ et $n = 1$ un calcul explicite permet de vérifier l'hypothèse de récurrence.

Hérédité : On suppose l'hypothèse de récurrence satisfaite au rang n . Soit $s = s_1, \dots, s_n$ une suite quasi-décroissante. On note v le chemin $\Phi_m(s_1, \dots, s_n)$. D'après l'hypothèse de récurrence $\Phi_m^+(s_1, \dots, s_n) = f(v)$. On choisit s_{n+1} , on note u le chemin $\Phi_m(s_1, \dots, s_{n+1})$ et on compare les constructions $\Phi_m^+(s_1, \dots, s_{n+1})$ et $f(u)$. Il y a deux cas à distinguer suivant le mode d'insertion utilisé. La figure 2.44 illustre le propos.

Si $s_{n+1} = s_n + 1$: Dans ce cas on utilise le mode d'insertion (xB, \overline{xABA}) ce qui crée un pic à la fin du chemin v . La permutation σ permet ensuite de décaler les indices des marques de pas montants sur les pics pour prendre en compte le nouveau pic après la dernière double montée. On retrouve bien la construction $f(u)$.

Sinon $s_{n+1} \leq s_n$: L'insertion selon de mode (xA, \overline{xBBBB}) crée une double montée juste avant le pic portant la marques de pas montant d'indice s_{n+1} . Les marques effacées correspondent à celle placées avant cette nouvelle dernière double montée donc on retrouve bien la construction $f(u)$.

□

2.2.5 Bijections non codables

Dans ce chapitre on a essayé d'accumuler des exemples de bijections qui utilisent des constructions de Dyck et certains algorithmes de réécritures de ces constructions. Cependant il existe des bijections qui ne rentrent pas dans ce cadre. Trouver de telles bijections peut aider à cerner les limites de cadre. On commence par donner une bijection artificielle qui illustre sous la forme d'un contre-exemple une propriété satisfaite par toutes les bijections définies précédemment dans ce chapitre. Ensuite on présente une bijection plus naturelle qui semble ne pas être codable par deux bijections avec les suites quasi-décroissantes.

Un chemin de Dyck u est à *un mélange de pic du chemin* v si u est le mélange de v et d'un pic. Une application g entre les suites quasi-décroissantes et les chemins de Dyck est *une application par mélange de pics* si pour toute suite quasi-décroissante $s = s_1, \dots, s_n, s_{n+1}$, le chemin $g(s)$ est à un mélange de pic du chemin $g(s_1, \dots, s_n)$. Toutes les applications obtenues à partir des algorithmes paramétrés sont des applications par mélange de pics. L'insertion de pas pour le moment des ricochets de Haglund en est un autre exemple (cf. la partie 2.2.3). Une bijection f entre les chemins de Dyck est *définissable par mélange de pic* s'il existe deux bijections g et h qui sont des applications par mélange de pics telles que $f = h \circ g^{-1}$. Par exemple, la bijection f associant un chemin de Dyck de hauteur de rayon n à un chemin de Dyck de hauteur $2n$ est définissable par mélange de pics avec g l'application de l'algorithme paramétré par $\langle (B\overline{B}x, A\overline{A}\overline{x}); (A\overline{A}x, B\overline{B}\overline{x}) \rangle$ et h celle paramétrée par $(x\overline{B}\overline{B}, \overline{x}A\overline{A})$ (cf. la partie 2.2.2).

Proposition 2.13 *Soit f une bijection sur les mots de Dyck, préservant la taille. Si*

$$\begin{cases} f(x\overline{x}x\overline{x}) = x\overline{x}x\overline{x}, \\ f(xx\overline{x}\overline{x}) = xx\overline{x}\overline{x}, \\ f(xx\overline{x}\overline{x}\overline{x}) = x\overline{x}x\overline{x}\overline{x}, \end{cases}$$

alors f n'est pas définissable par mélange de pics.

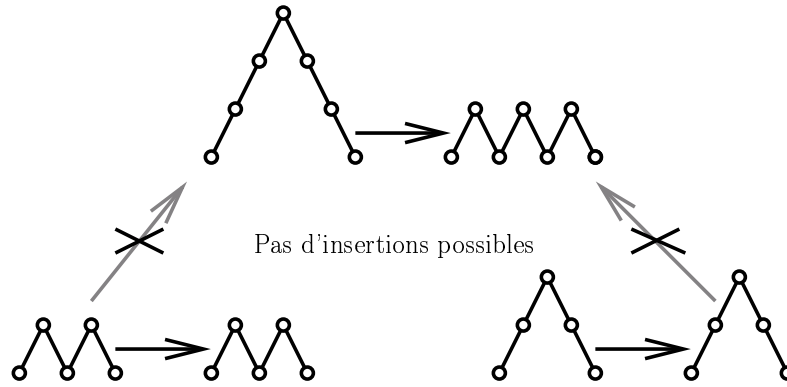


FIG. 2.45 – Un exemple d'application artificielle non codable par mélange de pics

Preuve. Supposons qu'il existe deux bijections par mélange de pics g et h entre les suites quasi-décroissantes et les chemins de Dyck telles que $f = h \circ g^{-1}$. On discute sur la suite quasi-décroissante $s = s_1, s_2, s_3 = g^{-1}(xxx\bar{x}\bar{x})$.

- Si $g(s_1, s_2) = x\bar{x}\bar{x}$, alors g n'est pas définie par mélange de pic car $g(s_1, s_2, s_3) = xxx\bar{x}\bar{x}$ n'est pas à un mélange de pic de $g(s_1, s_2) = x\bar{x}\bar{x}$.
- Si $g(s_1, s_2) = xx\bar{x}$, on utilise la relation $f = h \circ g^{-1}$ pour en déduire que $h = f \circ g$ et donc $h(s_1, s_2) = xx\bar{x}$ et $h(s_1, s_2, s_3) = x\bar{x}x\bar{x}\bar{x}$. h n'est donc pas une application par mélange de pic.

Ces contradictions impliquent que f n'est pas définissable par mélange de pics. \square

Cette proposition ne traite pas d'une bijection "naturelle" sur les mots de Dyck. On propose maintenant une bijection $g_{cyclique}$ basée sur le lemme cyclique découvert entre autres par Raney. On conjecture que cette bijection n'est pas définissable par mélange de pic sans avoir pu le prouver. Comme une variante de cette bijection sera utilisée dans la partie 3.5.1, cela donne une seconde raison de présenter cette bijection.

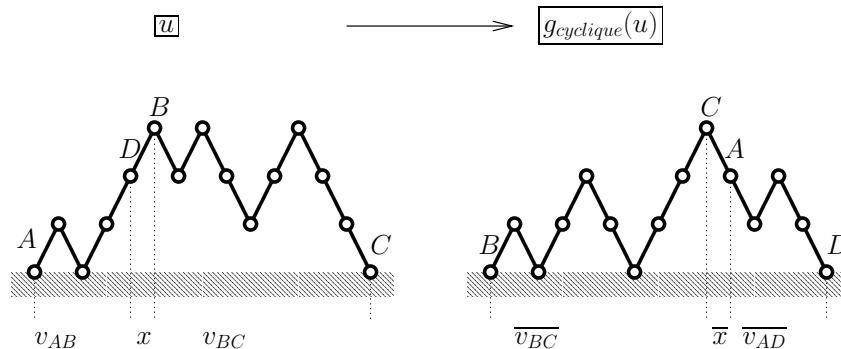


FIG. 2.46 – Définition de $g_{cyclique}$

La figure 2.46 illustre la définition de $g_{cyclique}$. Soit u un chemin de Dyck de hauteur h .

Si u est le chemin vide alors $g_{cyclique}(u)$ est le chemin vide. Sinon on nomme A le premier sommet du chemin, C le dernier sommet, B le sommet où le chemin atteint pour la première fois sa hauteur maximale et D le sommet précédant le sommet B . On remarque que A et D peuvent être égaux mais que cela n'induit pas de cas particulier dans la construction. On note v_{AD} , respectivement v_{BC} , le sous-chemin de u dont les extrémités sont A et D , respectivement B et C . On a

$$u = v_{AD} x v_{BC}.$$

Pour tout mot w sur l'alphabet $\{x, \bar{x}\}$ on note \bar{w} l'image de w par le morphisme défini par $x \rightarrow \bar{x}$ et $\bar{x} \rightarrow x$. En termes de chemin ce morphisme correspond à une symétrie par rapport à un axe horizontal. On pose

$$g_{cyclique}(u) = \overline{v_{BC}} \bar{x} \overline{v_{AD}}.$$

Il reste à vérifier que le chemin $g_{cyclique}$ est bien un mot de Dyck. D'une part v_{AD} est un préfixe de Dyck se terminant à sa hauteur maximale $h - 1$ donc $\overline{v_{AD}}$ est un suffixe de Dyck débutant à sa hauteur maximale $h - 1$. D'autre part v_{BC} est un suffixe débutant à sa hauteur maximale h donc $\overline{v_{BC}}$ est un préfixe se terminant à sa hauteur maximale h . Le pas \bar{x} de la définition de $g_{cyclique}(u)$ relie donc bien les sommets C et D pour former un chemin de Dyck.

Pour se convaincre que $g_{cyclique}$ est une bijection on va construire $g_{cyclique}^{-1}$.

Proposition 2.14 *L'application $g_{cyclique}$ est une bijection sur les mots de Dyck.*

Preuve. Soit u un chemin de Dyck de hauteur h . Si u est vide, la propriété est satisfaite. Sinon on peut retrouver les sommets A, B, C, D de $g_{cyclique}(u)$. En effet B est le premier pas de $g_{cyclique}(u)$, D le dernier pas, C est sur le dernier sommet du chemin atteignant la hauteur maximale et A le sommet suivant C . Seul le positionnement de C et A nécessite une remarque : $\overline{v_{BC}}$ est un préfixe de Dyck se terminant à sa hauteur maximale h , $\overline{v_{AD}}$ est un suffixe de Dyck débutant à sa hauteur maximale h donc C est bien le dernier sommet où le chemin u atteint sa hauteur maximale h . Une fois les sommets A, B, C et D reconnus on peut reconstruire le mot u pour obtenir $g_{cyclique}^{-1}(g_{cyclique}(u)) = u$. \square

La remarque suivante est en rapport avec la partie 3.5.1. On a découvert la bijection $g_{cyclique}$ par la construction géométrique illustrée sur la figure 2.47. On décrit cette construction rapidement. Certaines des affirmations suivantes nécessitent peut-être d'être un peu plus formel. Soit $u = u_0$ un chemin de Dyck non-vide de demi-longueur n . On considère le chemin $u_1 = xu_0xu_0$. On peut coller sous ce chemin une droite de pente $1/(2n + 1)$ qui soit en contact avec 3 sommets du chemin. Cette idée est intimement liée au lemme cyclique pour les chemins de Dyck d'où le nom de $g_{cyclique}$. D'autre part on peut poser sur ce chemin une droite de pente $1/(2n + 1)$ qui est en contact avec exactement deux sommets B_1 et B_2 de u_1 . On définit u_2 comme le chemin entre B_1 et B_2 . On pose $u_3 = \bar{u}_2$, c'est-à-dire le symétrique de u_2 par rapport à un axe horizontal. Ce chemin se termine par un pas descendant que l'on supprime pour obtenir le chemin de Dyck $g_{cyclique}(u)$. Pour le

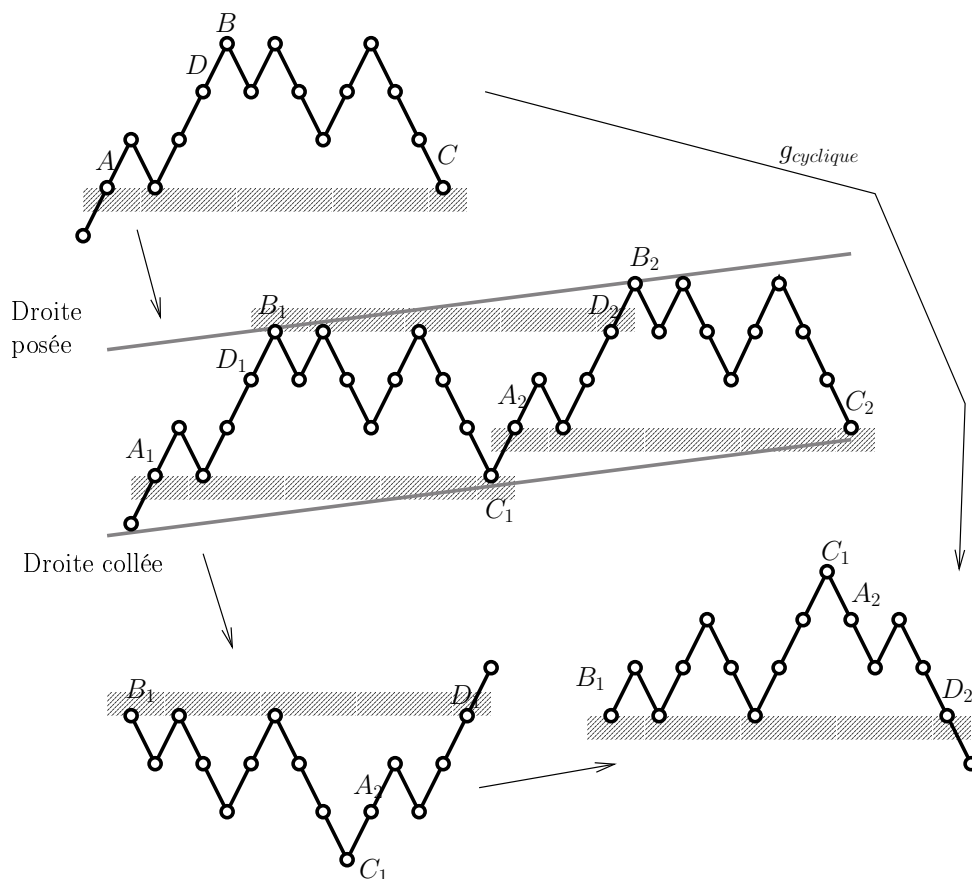


FIG. 2.47 – Le lemme cyclique et $g_{cyclique}$

vérifier le lecteur pourra observer l'adéquation de cette construction avec la précédente en observant sur le chemin u_1 les sommets A_i, B_i, C_i et D_i utilisés pour définir $g_{cyclique}$.

Chapitre 3

Modèle de polymères dirigés, interactions dans les chemins de Dyck

Des physiciens comme Rajesh, Dhar, Giri, Kumar et Singh [RDD⁺02] ont étudié un modèle de polymères dirigés au voisinage d'une surface où l'énergie de chaque monomère ne dépend que de son voisinage le plus proche. Pour le simplifier, sans toutefois le trivialisier, on code la configuration spatiale adoptée par le polymère par un chemin de Dyck. Le calcul de l'énergie fait intervenir deux types d'interactions sur les chemins de Dyck. Dans un chemin de Dyck deux sommets à la même hauteur sont en interaction supérieure si le sous-chemin dont ils sont les extrémités contient exactement un creux et aucun pic. Symétriquement on peut définir les interactions inférieures. Pour les interactions inférieures seules, la série génératrice a été établie par Denise et Simion [DS95] et elle est algébrique. Les interactions supérieures seules ne semblent pas se prêter à des décompositions algébriques car deux sommets en interaction appartiennent à des facteurs différents des décompositions usuelles des chemins de Dyck. Cette difficulté est atténuée par le fait que le chemin entre deux sommets en interaction supérieure est fortement contraint.

On se propose d'exposer trois approches qui semblent liées pour calculer cette série génératrice :

- En s'inspirant de travaux de Bousquet-Mélou et Rechnitzer [BMR03], on introduit la longueur de la dernière descente, un paramètre supplémentaire qualifié de catalytique par Zeilberger [Zei00]. Cela permet d'écrire une équation de décomposition liant trois spécialisations de la série génératrice en des valeurs distinctes de la variable catalytique. L'élimination de deux des spécialisations inconnues nécessite d'une part l'application de la méthode du noyau qui fait apparaître un terme algébrique et d'autre part une itération faisant apparaître des q -factorielles. L'extraction de la série recherchée implique une division. On remarque que le terme algébrique est aussi la série génératrice des chemins dont on compte les interactions inférieures.

- En exploitant la régularité d'un chemin entre deux sommets en interaction, on construit une relation d'équivalence sur les chemins. On travaille alors sur un ensemble

de représentants des classes d'équivalences, les chemins de Dyck à petits creux. L'ajout d'une autre variable catalytique permet d'écrire une équation q -algébrique pour ces représentants correctement valués. La résolution de cette équation nécessite deux étapes : une linéarisation par un changement de fonction inconnue puis une résolution de l'équation q -linéaire obtenue. Pour la linéarisation on s'inspire de travaux de Brak et Prellberg [BP95], et Van Rensburg [JvR00a]. Pour la résolution de l'équation q -linéaire, lorsqu'elle n'est pas directe, on utilise un algorithme proposé par Abramov, Petkovsek et Paule [APP98].

- Les chemins à petits creux sont valués dans chaque creux par la somme de deux termes de nature différente : une constante et une valeur dépendant géométriquement de la hauteur. On développe la somme de ces creux pour obtenir des chemins dont les creux sont valués par l'une ou l'autre. La sous-classe des chemins dont tous les creux sont valués par des constantes donne une interprétation combinatoire du terme algébrique de la formule finale. Une bijection, basée sur le lemme cyclique, explique alors le lien avec l'énumération des interactions inférieures. La sous-classe des chemins dont tous les creux sont valués par le terme "géométrique" correspond aux chemins dont tous les creux sont sur l'axe, modèle résolu par Van Rensburg [JvR00a] à l'aide d'une équation q -algébrique. On retrouve cette formule comme une instance du lemme d'inversion des empilements de segments, étudiés par Viennot et Bousquet-Mélou [BMV92]. Il reste à marier les deux énumérations "combinatoires" de ces deux sous-classes pour trouver la série génératrice initiale. On expliquera pourquoi elles sont déjà fiancées.

3.1 Présentation des polymères et des chemins

3.1.1 Modèles de polymères

Les polymères sont des agrégats de molécules, les monomères. La structure de polymère est maintenue grâce aux interactions entre ces molécules. Pour simplifier on étudie les polymères dont la structure est un fil de monomères. On suppose que ces polymères sont plongés dans un milieu, le solvant, formé de molécules isolées. Pour simplifier encore on travaille en dimension 2. Pour modéliser l'espace on utilise donc le réseau carré. Le *réseau carré* est le graphe infini dont les sommets sont indexés par \mathbb{Z}^2 et où il y a une arête entre les sommets (x_a, y_a) et (x_b, y_b) si et seulement si $|x_a - x_b| + |y_a - y_b| = 1$. Deux sommets formant une arête sont *voisins*.

Chaque sommet de la grille représente un emplacement qui peut être occupé soit par une molécule du solvant, soit par un monomère du polymère. La structure filaire du polymère permet de coder par un chemin la configuration spatiale adoptée par celui-ci. Un *chemin* de longueur n dans le réseau carré est une suite de sommets v_0, v_1, \dots, v_n telle que pour tout $i < n$, v_i et v_{i+1} sont voisins. L'ensemble des chemins que l'on va étudier dans ce chapitre sera défini par des contraintes sur les chemins. Les violations de ces contraintes sont illustrées par le chemin de la figure 3.1.

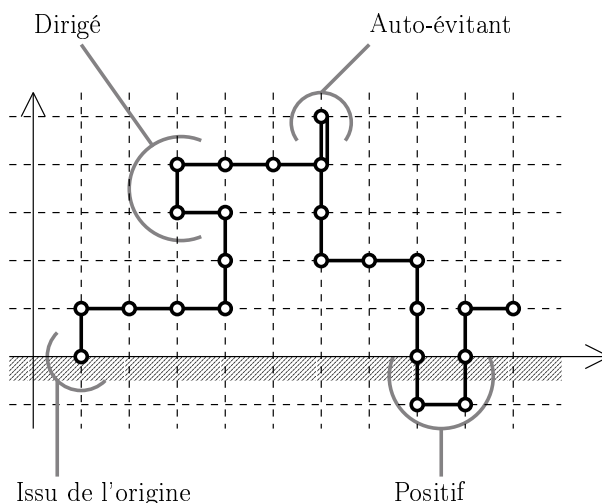


FIG. 3.1 – Un chemin contre-exemple aux polymères dirigés

Deux molécules ne pouvant occuper le même sommet, le chemin qui représente un polymère est naturellement auto-évitant : le chemin est *auto-évitant* si tout sommet apparaît au plus une fois dans la suite. Un couple de sommets voisins définit un vecteur nommé ici un *pas*. Il existe quatre types de pas : le *pas Est* $\mathbf{e} \equiv (1, 0)$, le *pas Nord* $\mathbf{n} \equiv (0, 1)$, le *pas Ouest* $\mathbf{o} = (-1, 0)$ et le *pas Sud* $\mathbf{s} = (0, -1)$. La *suite des pas du chemin* est la suite des pas définis par la suite des couples $(v_0, v_1), (v_1, v_2), \dots, (v_{n-1}, v_n)$. On suppose que le demi-plan inférieur strict est un milieu solide formé d'atomes de sorte que les emplacements de cette partie de l'espace ne peuvent pas être occupés par des molécules. Cette hypothèse se traduit pour les chemins par une contrainte de positivité. Le chemin est *positif* si tous ses sommets appartiennent à $\mathbb{Z} \times \mathbb{N} \subset \mathbb{Z}^2$.

Les deux dernières contraintes ont des motivations plus combinatoires. Le caractère auto-évitant d'un chemin étant difficile à prendre en compte combinatoirement on fait une hypothèse de direction pour simplifier le problème. Le chemin est *dirigé* si la suite de ses pas est formée de pas Nord, Sud ou Est mais pas Ouest. Les symétries du modèle permettent, au prix de quelques calculs, de se contenter d'énumérer les chemins dont une extrémité est à l'origine. Pour simplifier on supposera que les chemins sont issus de l'origine. Le chemin est *issu de l'origine* si son premier sommet v_0 est $(0, 0)$.

Dans ce chapitre, un *polymère dirigé* est un chemin issu de l'origine, dirigé, auto-évitant et positif. Un exemple de polymère dirigé est proposé par la figure 3.2.

Comme les polymères dirigés sont issus de l'origine ils sont caractérisés par leur suite de pas. On peut donc coder naturellement ces polymères par un mot sur l'alphabet $\{\mathbf{n}, \mathbf{s}, \mathbf{e}\}$. Le polymère de la figure 3.2 est codé par le mot $\mathbf{n.e.n.e.s.s.e.n.n.n.e.s.e.n.n.e.e.s.s.e.s.e.e.n.n.e.n.e}$. De la même façon que pour les chemins de Dyck on peut traduire les contraintes sur les chemins par des contraintes sur leurs mots. Un mot code un chemin auto-évitant s'il évite les facteurs $\mathbf{n.s}$ et $\mathbf{s.n}$. On voit à

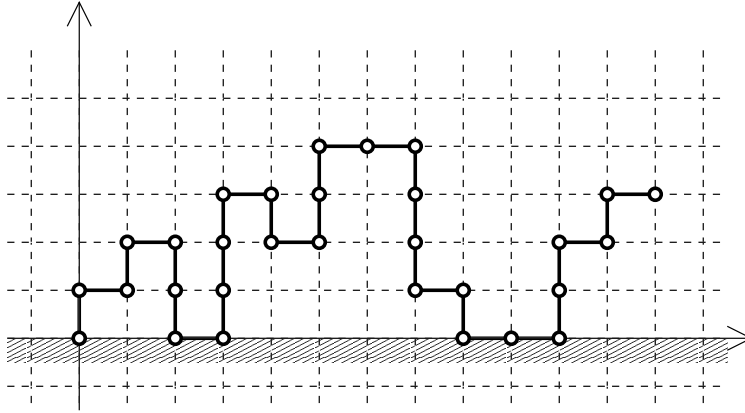


FIG. 3.2 – Exemple de polymère dirigé

cette occasion comment l’hypothèse de direction pour les chemins permet de simplifier la reconnaissance de la nature auto-évitante des chemins. Un mot code un chemin positif si tout préfixe u du mot satisfait $|u|_n \geq |u|_s$.

Pour résumer, les polymères dirigés sont en bijection avec les mots sur l’alphabet $\{n, e, s\}$ qui évitent les facteurs $n.s$ et $s.n$ et dont tout préfixe u satisfait $|u|_n \geq |u|_s$. Comme pour les chemins de Dyck on confond donc ces notions de chemin et de mot.

Les polymères dirigés définissent une classe combinatoire \mathcal{P} dont la taille des éléments, les chemins, est la longueur. On définit donc pour cette classe une série génératrice

$$P(t) = \sum_{c \in \mathcal{P}} t^{l(c)}$$

où $l(c)$ est la longueur du chemin c .

Pour ce modèle issu de la physique, on calcule l’énergie d’un polymère dans une configuration spatiale donnée, en fonction de paramètres supplémentaires. La figure 3.3 illustre la définition de ces paramètres.

L’énergie d’un polymère est la somme des énergies des molécules qui le composent. L’énergie d’une molécule est fonction de ses interactions avec son voisinage. Si on restreint son voisinage aux sommets du réseau carré voisin, son énergie ne dépend que de l’occupation des 4 sommets qui sont à distance 1 de cette molécule. La structure de polymère impose la présence constante dans ce voisinage des molécules précédente et suivante dans le polymère. Les variations d’énergie ne proviennent donc pas de ces deux molécules, une de moins pour les extrémités du polymère. Cependant il peut exister dans le voisinage des molécules supplémentaires du polymère qui sont donc à distance 1 de la molécule dans l’espace, sans être consécutive à la molécule dans le polymère. Dans ce cas on dit que les molécules sont en interaction. Plus formellement, deux sommets v_i et v_j du chemin v_0, \dots, v_n sont en *interaction* s’ils sont voisins sans être consécutifs dans le chemin, c’est-à-dire $j \notin \{i-1, i+1\}$. Pour les polymères dirigés les couples de sommets en interaction sont nécessairement

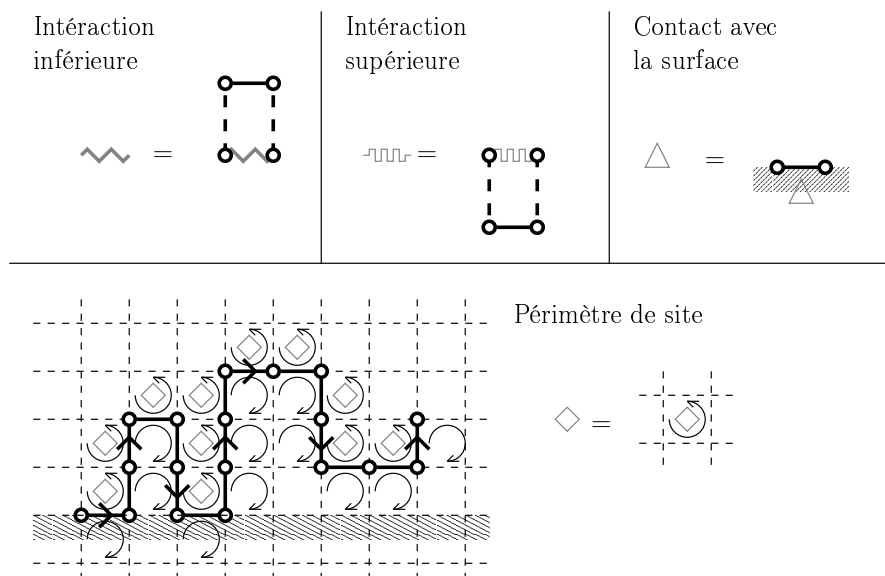


FIG. 3.3 – Définition des paramètres étudiés sur les polymères dirigés

de la forme $((x, y), (x + 1, y))$. De plus dans le sous-chemin débutant en (x, y) et finissant en $(x + 1, y)$ il y a un unique pas Est. Les contraintes de l'énumération vont imposer de distinguer deux types d'interactions : les inférieures et les supérieures. Deux sommets sont *en interaction inférieure*, respectivement *en interaction supérieure*, si leur ordonnée commune est inférieure, respectivement supérieure, à celle des sommets formant le pas Est associé. Dans nos séries génératrices on compte les interactions supérieures par la variable u et les inférieures par la variable v .

Hormis les molécules du solvant, les atomes formant le solide du demi-plan inférieur sont les derniers voisins possibles d'une molécule du polymère. Dans ce cas on parle d'un contact de la molécule avec le sol. Un sommet $v_i = (x_i, y_i)$ d'un chemin est un *contact* si $y_i = 0$. Dans la série génératrice, on compte les contacts d'un chemin par la variable w .

On présente maintenant deux paramètres dont l'origine est plus proche de la combinatoire même s'il doit être possible de leur trouver un intérêt physique. Le réseau carré permet de partitionner le plan \mathbb{R}^2 en *cellules* qui sont des carrés de la forme $[i, i + 1[\times [j, j + 1[$, $i, j \in \mathbb{Z}$. On peut orienter un polymère dirigé de v_0 à v_n . Cette orientation induit une orientation sur les cellules dont au moins deux coins sont des sommets consécutifs du polymère. Un *site périmétrique* du chemin est une cellule dont l'orientation est dans le sens trigonométrique. Le nombre de sites périmétriques définit le *périmètre de site* du polymère dirigé et il est compté par la variable p dans la série génératrice. L'énumération selon ce paramètre sur les polymères dirigés contraints à se terminer par une molécule en contact avec le sol à conduit Bousquet-Mélou et Rechnitzer [BMR03] à utiliser la technique d'énumération "à la Temperley" que nous allons suivre scrupuleusement dans une prochaine partie. La *hauteur finale* d'un polymère est l'ordonnée de son dernier sommet. Cette hauteur est comptée par

la variable h dans la série génératrice.

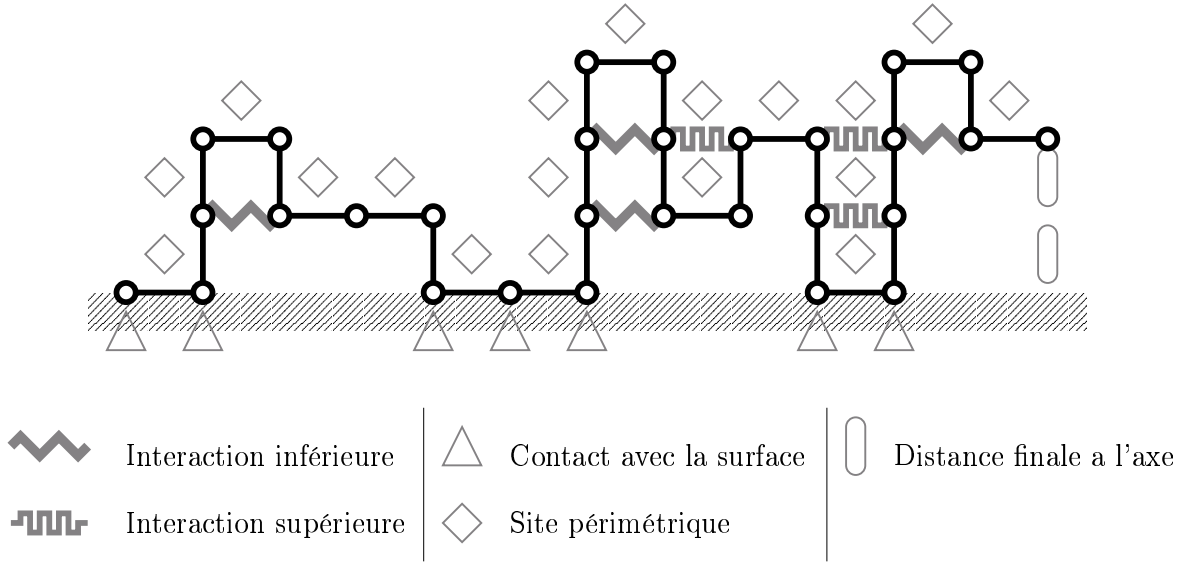


FIG. 3.4 – Récapitulation des paramètres étudiés sur les polymères dirigés

La série génératrice prenant en compte ces paramètres est donc

$$P(t, u, v, w, p, h) = \sum_{c \in \mathcal{P}} t^{\text{longueur}(c)} u^{\text{superieures}(c)} v^{\text{inferieures}(c)} w^{\text{contacts}(c)} p^{\text{perimetre}(c)} h^{\text{hauteur}(c)}. \quad (45)$$

La contribution du chemin de la figure 3.4 à la série génératrice est donc $t^{28}u^3v^4w^7p^{18}h^2$.

À un détail technique mais trivial près, Bousquet-Mélou et Rechnitzer [BMR03] ont énumérés les polymères dirigés contraints à se terminer sur l'axe selon la longueur et le périmètre de site. Ils ont de plus distingué les pas verticaux et les pas horizontaux. La spécialisation de leur formule donne une évaluation partielle de la série génératrice $P(t, u, v, w, p, h)$: la série des polymères dirigés énumérés selon leur longueur et périmètre de site

$$P(t, 1, 1, 1, p, 0) = -\frac{1}{t^2} \cdot \frac{\sum_{n \geq 0} \frac{t^{2n} \xi^n p^n q^{\binom{n+2}{2}} (1-p)^n}{(1+t-tp)^n (1-q)^n (1-pt)^n (q)_n (\xi q)_n}}{\sum_{n \geq 0} \frac{t^{2n} \xi^n p^n q^{\binom{n+2}{2}} (1-p)^n (1+t-\xi q^{n+1}(p+tp-t))}{(1+t-tp)^n (1-q)^n (1+t-\xi q^n)(1+t-\xi q^{n+1})(q)_n (\xi q)_n}}. \quad (46)$$

où

- $q \equiv tp$,

- ξ est la seule série formelle en t qui soit une racine du polynôme en s de degré 2

$$(1-q)(1-s)(1-sqp)+tp\left((1-q)(1-p)(1+s^2pq)+sp\left((1-q)(1+pq-2q)+tpq^2(1-p)\right)\right), \quad (47)$$

- La notation $(a)_n$ désigne comme d'habitude le produit $\prod_{k=0}^{n-1}(1-q^k a)$.

3.1.2 Interactions sur les chemins de Dyck

Les polymères dirigés ont été définis pour être la classe la plus large possible que l'on sache énumérer selon les paramètres que l'on a présenté. On peut définir sur les chemins de Dyck des paramètres analogues aux interactions et aux contacts. Ils forment une classe \mathcal{D} plus simple d'un point de vue combinatoire dont les formules d'énumérations sont similaires et d'écriture plus compacte. Dans ce chapitre, on étudie en détail le modèle des chemins de Dyck.

La figure 3.5 illustre la définition sur les chemins de Dyck des paramètres analogues à ceux des polymères dirigés. Un *contact* du chemin de Dyck est un sommet de ce chemin qui est sur l'axe des abscisses. Deux sommets distincts sont en *interaction inférieure*, respectivement en *interaction supérieure*, s'ils sont à la même hauteur et que le sous-chemin dont ils sont les extrémités contient exactement un pic et aucun creux, respectivement un creux et aucun pic.

Pour énumérer selon ces paramètres on introduit la série génératrice

$$D(t, u, v, w) = \sum_{c \in \mathcal{D}} t^{\text{demilongueur}(c)} u^{\text{superieures}(c)} v^{\text{inferieures}(c)} w^{\text{contacts}(c)}. \quad (48)$$

La contribution à cette série génératrice du chemin de Dyck de de la figure 3.5 est donc $t^{14}u^7v^{10}w^4$.

Le calcul de la série $D(t, u, v, w)$ est *a priori* plus difficile que celui de certaines de ses spécialisations. Par exemple, la série $D(t, 1, 1, w)$ compte les chemins de Dyck selon leur demi-longueur et le nombre de contacts. Pour abrégé les notations on note cette série $D(t, w)$. Plus généralement on supprime dans nos notations les paramètres dont on ne tient pas compte.

3.1.3 Traitement des interactions inférieures

Le fait d'ignorer les interactions supérieures permet d'écrire pour les chemins de Dyck des équations de décomposition algébrique. Denise et Simion [DS95] ont énuméré les chemins de Dyck selon leur longueur et leurs interactions inférieures. On profite des calculs de $D(t, v)$ et $D(t, v, w)$ pour proposer deux décompositions combinatoires. Dans le premier cas, pour $D(t, v)$ on s'intéresse au premier creux du chemin. Ensuite, pour $D(t, v, w)$, on

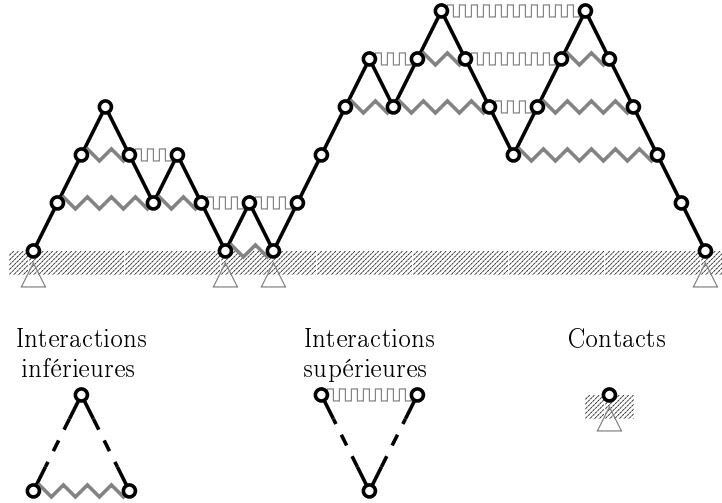


FIG. 3.5 – Paramètres étudiés sur les chemins de Dyck

discute si le premier sommet est en interaction inférieure ou non avec le premier retour sur l'axe.

Proposition 3.1 *La série génératrice $D(t, v)$ des chemins de Dyck non vides dont la demi-longueur est comptée par t et le nombre d'interactions inférieures par v satisfait l'équation de décomposition algébrique*

$$D(t, v) = \frac{tv}{1 - tv} + \frac{tv}{1 - tv} \frac{D(t, v)}{1 - t(1 + D(t, v))}. \quad (49)$$

Cette série est donc

$$D(t, v) = \frac{(1 - t)(1 - 2tv) - \sqrt{(1 - t)(1 - (1 + 4v)t + 4v^2t^2)}}{2t(1 - tv)}. \quad (50)$$

Preuve. On donne une décomposition de ces chemins qui respecte les interactions inférieures : on factorise un chemin en repérant son premier creux. La figure 3.6 illustre cette factorisation.

Plus formellement, on partitionne l'ensemble des chemins de Dyck en :

- Ceux qui ne contiennent pas de creux. Leur série génératrice est $\frac{tv}{1 - tv}$.
- Ceux qui ont au moins un creux. Avant ce creux il y a un chemin de Dyck sans creux. Après le premier creux il existe un chemin de Dyck non vide dont le premier pas forme le creux. Si le premier creux est à la hauteur n , en factorisant sur les premières descentes à chaque hauteur après le premier creux il apparaît n facteurs qui sont soit vides, soit des chemins de Dyck. La série génératrice de ces chemins est donc

$$\frac{tv}{1 - tv} D(t, v) \sum_{n \geq 0} (t(1 + D(t, v)))^n = \frac{tv}{1 - tv} \frac{D(t, v)}{1 - t(1 + D(t, v))}.$$

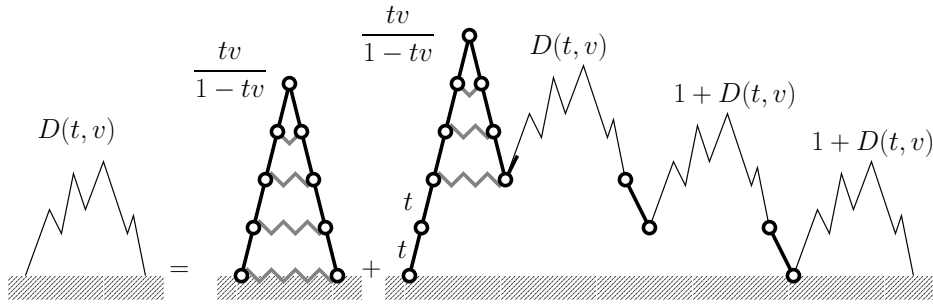


FIG. 3.6 – Factorisation des chemins de Dyck préservant les interactions inférieures

La série génératrice des chemins de Dyck dont on compte les interactions inférieures satisfait l'équation (49) qui conduit à une équation algébrique de degré 2 satisfaite par $D(t, v)$. La résolution de cette équation conduit à deux solutions sous forme de séries de Laurent en t à coefficients polynomiaux en v . Une seule parmi les deux est une série formelle en t à coefficients positifs. C'est donc elle la série génératrice. \square

Proposition 3.2 *La série génératrice $D(t, v, w)$ des chemins de Dyck non vides dont la demi-longueur est comptée par t , les interactions inférieures par v et les contacts par w , satisfait l'équation de décomposition*

$$D(t, v, w) = \frac{wtv}{1-tv} \left(w + D(t, v, w) \right) + wt \left(D(t, v) - \frac{tv}{1-tv} \right) \left(w + D(t, v, w) \right) \quad (51)$$

où $D(t, v)$ est la série des même chemins pour laquelle on ne tient pas compte des contacts. Elle vaut donc

$$D(t, v, w) = \frac{1}{2} \frac{1 + w^2 \left(2wt^2v - 2wtv - t - \sqrt{-(-1+t)(4t^2v^2 - 4tv - t + 1)} \right)}{1 - t^2w^2v - tv + w^2tv + wt - w}. \quad (52)$$

Preuve. Pour factoriser un chemin de Dyck on discute cette fois sur l'existence d'une interaction inférieure entre le premier sommet du chemin et le premier retour sur l'axe. La figure 3.1.3 illustre le propos.

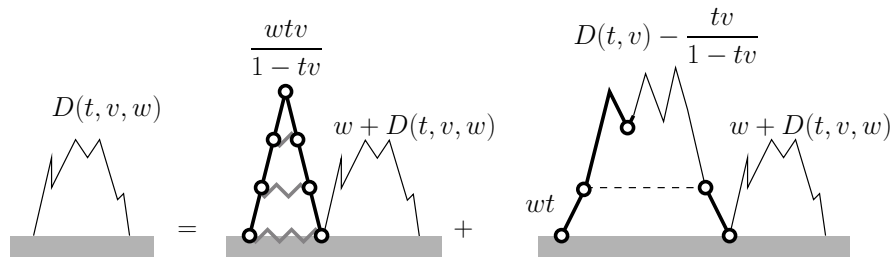


FIG. 3.7 – Factorisation pour les interactions inférieures et les contacts

On partitionne donc les chemins de Dyck en

- Ceux dont le premier sommet est en interaction inférieure avec le premier retour. Le premier facteur ne contient pas de creux, la série génératrice de ces facteurs est $\frac{tv}{1-tv}$. Si le premier retour est aussi le dernier sommet, on compte ce dernier contact, sinon il y a un nouveau chemin de Dyck dont on compte les contacts. La série génératrice de ce second facteur est donc $w + D(t, v, w)$.
- Ceux dont le premier sommet n'est pas en interaction inférieure avec le premier retour. Le premier pas et le premier contact sont valués wt . Ensuite le premier facteur est un chemin de Dyck *qui contient au moins un creux*. On ne compte pas les contacts pour ce facteur puisqu'ils ne correspondent pas à des contacts du chemin factorisé. On obtient la série génératrice de ces chemins par différence : l'ensemble des chemins avec au moins un creux est l'ensemble des chemins privé de l'ensemble des chemins sans creux. La série génératrice du premier facteur est donc $D(t, v) - \frac{tv}{1-tv}$. Le second facteur est le même que dans le premier cas, sa série génératrice est donc $w + D(t, v, w)$.

Cette décomposition conduit à une expression rationnelle de $D(t, v, w)$ en fonction de $D(t, v)$. En utilisant l'expression de $D(t, v)$ de la proposition 3.1 on obtient l'expression de $D(t, v, w)$. \square

La suite de ce chapitre est consacrée aux interactions supérieures. Les titres des sous-chapitres évoquent les méthodes employées pour énumérer les chemins selon ce paramètre.

3.2 Interactions supérieures par équation à la Temperley

Les factorisations usuelles des chemins de Dyck préservent les interactions inférieures. Pour les interactions supérieures c'est plus délicat. Par exemple pour la factorisation usuelle au premier retour sur l'axe, il existe des interactions supérieures où les sommets appartiennent aux deux sous-chemins de Dyck (*cf.* Figure 3.2). Ainsi le nombre d'interactions à cheval entre les deux sous-chemins est le minimum de la longueur de la dernière descente du premier sous-mot et de la longueur de la première montée du second sous-mot.

A priori il semble donc difficile d'écrire des équations de décomposition algébrique pour ce modèle. Les décompositions par tranches, liées à une approche à la Temperley, sont une alternative aux décompositions algébriques. Au lieu de découper récursivement les objets en deux ou plusieurs sous-objets, on ambitionne plus modestement dans cette approche d'effeuiller récursivement un objet : on le décompose en un morceau que l'on sait énumérer et un seul sous-objet de même nature. Cependant pour décrire cette décomposition il est nécessaire de rajouter un paramètre sur les objets que l'on énumère. Comme ce paramètre n'est pas pertinent dans le résultat mais qu'elle permet la mise en équation, Zeilberger le qualifie de *catalytique* [Zei00].

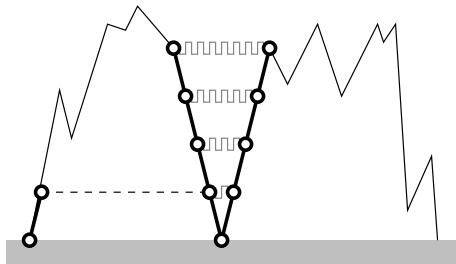


FIG. 3.8 – Difficulté des interactions supérieures

Grâce à une décomposition à la Temperley, Bousquet-Mélou et Rechnitzer [BMR03] ont pu énumérer les polymères dirigés selon le périmètre de site. La difficulté de ce paramètre provient des sites périmétriques qui sont bordés par des interactions supérieures. Sans surprise, leur méthode s'applique donc à l'énumération des interactions supérieures. Dans cette partie on reprend quasi-littéralement la méthode qu'ils ont proposée pour traiter le cas plus simple des chemins de Dyck.

3.2.1 Mise en équation

On veut calculer la série génératrice des chemins de Dyck dont on compte le nombre d'interactions supérieures. Pour la mise en équation on introduit un paramètre catalytique, la longueur de la dernière descente, comptée par la variable s . Pour un chemin de Dyck c on note $descente(c)$ la longueur de sa dernière descente.

On travaille donc avec la série génératrice

$$D(t, u; s) = \sum_{c \in \mathcal{D}} t^{\text{demilongueur}(c)} u^{\text{superieur}(c)} s^{\text{descente}(c)}.$$

Proposition 3.3 *La série génératrice $D(t, u; s)$ des chemins de Dyck dont la demilongueur est comptée par t , le nombre d'interactions supérieures par u et la longueur de la dernière descente par s satisfait l'équation :*

$$\begin{aligned} D(t, u; s) &= \frac{ts}{1-ts} + \frac{ut}{1-ut} \frac{ts}{1-ts} \left(D(t, u; s) - D(t, u; uts) \right) \\ &+ \frac{ut}{1-ut} \left(\frac{s}{1-s} \left(D(t, u; 1) - D(t, u; s) \right) - \frac{uts}{1-uts} \left(D(t, u; 1) - D(t, u; uts) \right) \right). \end{aligned} \quad (53)$$

Preuve. On effeuille les chemins de Dyck qui contiennent au moins un creux en supprimant le facteur de Dyck apparaissant après le dernier creux. On discute ensuite selon la hauteur relative des deux derniers pics (voir figure 3.9). Cette décomposition nécessite donc la partition des chemins en deux classes :

- ceux qui n'ont pas de creux. Leur série génératrice est $\frac{ts}{1-ts}$.

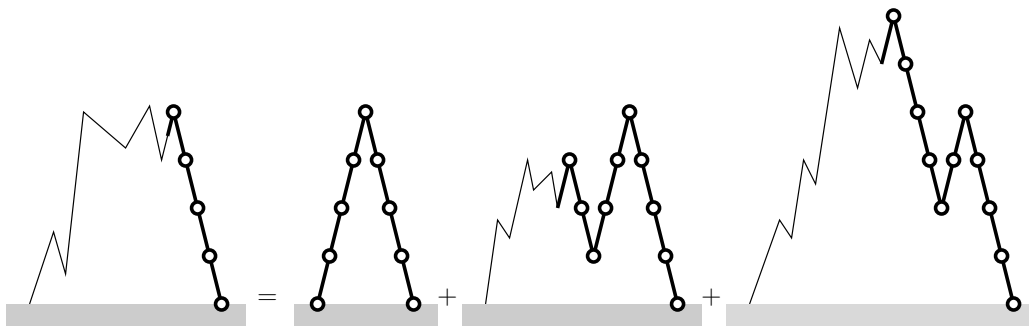


FIG. 3.9 – Décomposition à la Temperley pour les chemins de Dyck

- ceux qui ont au moins un creux. Ces chemins se construisent de façon unique en insérant un facteur de la forme $x^j \bar{x}^j$ à la hauteur i dans un sous-chemin c dont k , la longueur de la dernière descente, vaut au moins $i + 1$. Dans ce cas le nombre d'interactions supérieures créée est $\min(j, k - i)$. On distingue donc deux cas
 - ◇ l'avant dernier pic du chemin est strictement au-dessous du dernier.

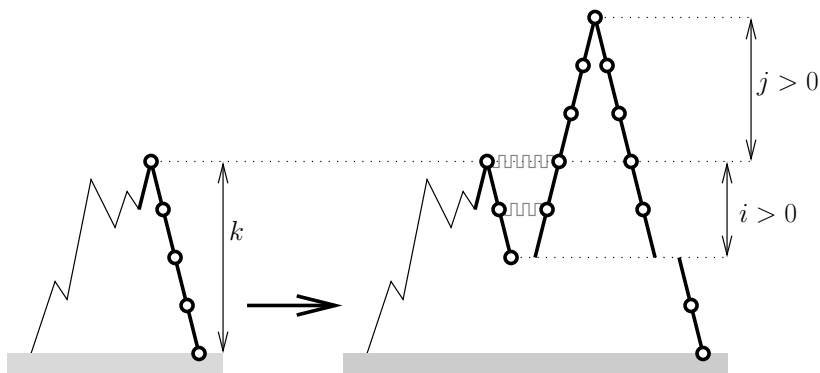


FIG. 3.10 – Insertion strictement croissante de pic

Soit $i < k$ la différence de hauteur entre l'avant-dernier pic et le dernier creux et $j > 0$ la différence de hauteur entre les deux derniers pics. Dans ce cas il y a exactement i interactions supérieures supplémentaires, la demi-longueur augmente de $j + i$ et la longueur de la dernière descente est maintenant $k + j$. On note $f_c(t, u)s^k$ la valuation du sous-chemin c . La série génératrice de cette classe est donc

$$f_c(t, u)s^k \sum_{i=1}^{k_c} (ut)^i \sum_{j=1}^{\infty} (ts)^j = f_c(t, u) \frac{ut}{1-ut} \frac{ts}{1-ts} (s^{k_c} - (uts)^{k_c}).$$

En sommant sur tous les sous-chemins on voit apparaître des évaluations de la

série génératrice principale :

$$\begin{aligned}
 & \sum_{c \in \mathcal{D}} f_c(t, u) \frac{ut}{1-ut} \frac{ts}{1-ts} (s^{k_c} - (uts)^{k_c}) \\
 &= \frac{ut}{1-ut} \cdot \frac{ts}{1-ts} \left(\sum_{c \in \mathcal{D}} f_c(u, t) s^{k_c} - \sum_{c \in \mathcal{D}} f_c(u, t) (uts)^{k_c} \right) \\
 &= \frac{ut}{1-ut} \frac{ts}{1-ts} (D(t, u; s) - D(t, u; uts)).
 \end{aligned}$$

◇ l'avant dernier pic du chemin est au dessous du dernier.

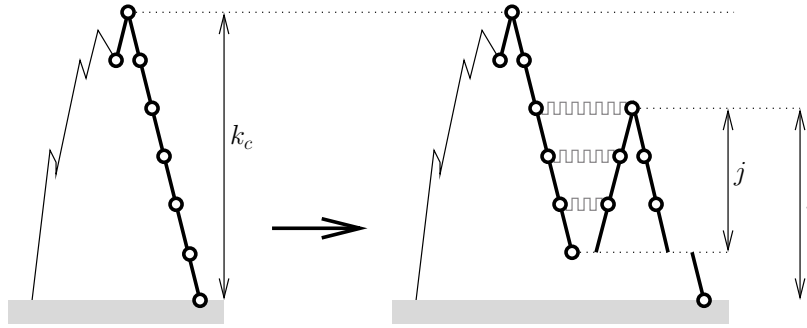


FIG. 3.11 – Insertion décroissante de pic

Dans ce cas, soit $i < k$ la hauteur du dernier pic, et $j \leq i$ la longueur de la dernière montée du chemin formé. Il y a j interactions supérieures créées et la longueur de la dernière descente est i . La série génératrice de cette classe est donc

$$\begin{aligned}
 & f_c(t, u) \sum_{i=1}^k s^i \sum_{j=1}^i (ut)^j \\
 & f_c(t, u) \sum_{i=1}^k s^i \frac{ut}{1-ut} (1 - (ut)^i) \\
 &= f_c(t, u) \frac{ut}{1-ut} \left(\frac{s}{1-s} (1 - s^k) - \frac{uts}{1-uts} (1 - (uts)^k) \right).
 \end{aligned}$$

On reconnaît à nouveau dans la sommation sur tous les sous-chemins de Dyck des évaluations de la série génératrice $D(t, u; s)$:

$$\frac{ut}{1-ut} \left(\frac{s}{1-s} (D(t, u; 1) - D(t, u; s)) - \frac{uts}{1-uts} (D(t, u; 1) - D(t, u; uts)) \right).$$

En regroupant toutes ces partitions on obtient l'équation 53 de la proposition 3.3 □

3.2.2 Résolution de l'équation

La série génératrice $D(t, u)$ des chemins de Dyck dont la demi-longueur est comptée par t et les interactions supérieures par u se calcule à partir de l'équation 53 de la proposition 3.3.

Proposition 3.4 *On pose $q \equiv ut$. Soit $\sigma \equiv \sigma(t, u)$ la série formelle en t satisfaisant*

$$1 + \left(\frac{q-t}{1-q} - 1 \right) \sigma + t\sigma^2 = 0. \quad (54)$$

On a

$$\sigma(t, u) = \frac{1+t-2q - \sqrt{(-1+t)(t-1+4q-4q^2)}}{2t(q-1)} \quad (55)$$

et

$$D(t, u) = - \frac{t \sum_{n \geq 0} \left(\frac{(q-t)\sigma}{1-q} \right)^n \frac{q^{\binom{n+2}{2}-1}}{(q)_n (qt\sigma^2)_n}}{q \sum_{n \geq 0} \left(\left(\frac{(q-t)\sigma}{1-q} \right)^n \frac{q^{\binom{n+2}{2}-1}}{(q)_n (qt\sigma^2)_n} \right) \frac{(1-tq^n\sigma)}{(1-q^n\sigma)(1-q^{n+1}\sigma)}}. \quad (56)$$

L'équation de décomposition à la Temperley (53) est une combinaison linéaire des séries inconnues $D(t, u; s)$, $D(t, u; 1)$ et $D(t, u; uts)$. Elles sont autant d'évaluations de la série $D(t, u; s)$ pour différentes valeurs de la variable catalytique s . On cherche à calculer $D(t, u; 1)$. On élimine la série $D(t, u; uts)$ par une itération de l'équation qui fait apparaître dans l'expression finale des termes en $(a)_n$. La série $D(t, u; s)$ disparaît ensuite par application de la méthode du noyau; elle est responsable, à titre posthume, des occurrences de la série σ . Pour combiner harmonieusement ces deux sources, on utilise les relations entre le produit des racines et un coefficient du polynôme en σ de l'équation (54). Enfin la résolution de l'équation qui ne contient plus que $D(t, u; 1)$ comme série inconnue nécessite une division qui explique que l'expression finale de $D(t, u) \equiv D(t, u; 1)$ soit un quotient. La structure de cette preuve est identique à celle proposée par Bousquet-Mélou et Rechnitzer [BMR03].

Preuve.

Les évaluations de la série génératrice apparaissant dans l'équation de décomposition ne diffèrent que sur la variable s . On utilise donc la notation plus compacte $D(s) \equiv D(t, u; s)$. D'autre part, la présence de $D(tus)$ suggère de poser $q \equiv ut$. Dans un premier temps on cherche une solution sous la forme d'une série formelle en t à coefficients rationnels en u et s . On peut récrire l'équation de décomposition en regroupant les termes correspondant aux mêmes évaluations de la série génératrice. On obtient une équation de la forme

$$D(s) = a(s) + b(s)D(1) + c(s)D(qs) + d(s)D(s),$$

avec

$$a(s) = \frac{ts}{1-ts}, \quad (57)$$

$$b(s) = \frac{q}{1-q} \left(\frac{s}{1-s} - \frac{qs}{1-qs} \right), \quad (58)$$

$$c(s) = \frac{q}{1-q} \left(\frac{qs}{1-qs} - \frac{ts}{1-ts} \right), \quad (59)$$

$$d(s) = \frac{q}{1-q} \left(\frac{ts}{1-ts} - \frac{s}{1-s} \right). \quad (60)$$

$$(61)$$

L'évaluation de cette équation pour $s \rightarrow q^k s$ donne la relation

$$D(q^k s) = a(q^k s) + b(q^k s)D(1) + c(q^k s)D(q^{k+1} s) + d(q^k s)D(q^k s). \quad (62)$$

On constate que $1 - d(q^k s)$ est inversible car son terme de degré 0 en t est 1. On en déduit une expression de $D(q^k s)$ en fonction de $D(q^{k+1} s)$:

$$D(q^k s) = \frac{1}{1-d(q^k s)} (a(q^k s) + b(q^k s)D(1) + c(q^k s)D(q^{k+1} s)). \quad (63)$$

Cette relation permet d'itérer l'équation de décomposition N fois :

$$\begin{aligned} (1-d(s))D(s) &= a(s) + b(s)D(1) + c(s) \frac{1}{1-d(qs)} (a(qs) + b(qs)D(1) + c(qs)D(q^2 s)) \\ &= \sum_{n=0}^N \left(\prod_{k=0}^{n-1} \frac{c(q^k s)}{1-d(q^{k+1} s)} \right) (a(q^n s) + b(q^n s)D(1)) + \left(\prod_{k=0}^N c(q^k s) \right) D(q^{N+1} s). \end{aligned}$$

Comme $\frac{c(s)}{1-d(qs)} = ust + \mathcal{O}(st^2)$, on a

$$\prod_{k=0}^N \frac{c(q^k s)}{1-d(q^{k+1} s)} = \prod_{k=0}^N (q^{k+1} s + \mathcal{O}(t^{k+2})) = q^{\binom{N}{2}} s^N + \mathcal{O}(t^{\binom{N}{2}}).$$

Cette remarque prouve la convergence des itérations, au sens des séries formelles en t , vers la relation

$$(1-d(s))D(s) = \sum_{n=0}^{\infty} \left(\prod_{k=0}^{n-1} \frac{c(q^k s)}{1-d(q^{k+1} s)} \right) (a(q^n s) + b(q^n s)D(1)). \quad (64)$$

La q -itération permet donc d'éliminer $D(qs)$, une des évaluations de la série génératrice.

Pour éliminer $D(s)$ on va utiliser la méthode du noyau [BMP00]. Cela consiste à choisir une valeur pour s de sorte que $1 - d(s) = 0$. Ici on a

$$1 - d(s) = 1 - \frac{q}{1-q} \left(\frac{ts}{1-ts} - \frac{s}{1-s} \right) = \frac{(1-q) + (2q-t-1)s + (1-q)s^2 t}{(1-s)(1-q)(1-ts)}.$$

Cette annulation de $1 - d(s)$ se réduit à l'annulation du polynôme

$$K(s) \equiv K(t, u; s) = (1 - q) + (2q - t - 1)s + (1 - q)ts^2 \quad (65)$$

que l'on nomme le *noyau*.

Ce noyau est un polynôme de degré 2 en s et admet donc deux racines sous la forme d'une série de Laurent en la variable t . Une seule de ces racines est une série formelle en t , on la note $\sigma \equiv \sigma(t, u)$:

$$\sigma := \frac{1 + t - 2ut - \sqrt{(1-t)(1 - (1+4u)t + 4u^2t^2)}}{2t(1-ut)}. \quad (66)$$

Le début du développement de cette série est

$$\sigma = 1 + ut + 2u^2t^2 + (u^2 + 4u^3)t^3 + (u^2 + 5u^3 + 8u^4)t^4 + \mathcal{O}(t^5). \quad (67)$$

On cherche maintenant $D(t, u)$ sous la forme d'une série formelle en t à coefficients rationnels en u . On souhaite substituer σ à s dans l'équation (64) mais en toute rigueur le terme $\frac{s}{1-s}$ qui apparaît dans $1 - d(s)$ et dans $b(s)$ pourrait être source de divergence. Cependant ce n'est pas le cas car d'après (67),

$$1 - \sigma = ut + \mathcal{O}(t^2) \quad (68)$$

donc

$$\frac{q}{1-q} \frac{\sigma}{1-\sigma} = \sigma + \mathcal{O}(t). \quad (69)$$

Le terme σ est donc substituable à s dans (64) et l'équation résultant de la q -itération donne ainsi

$$0 = \sum_{n \geq 0} \left(\prod_{k=0}^{n-1} \frac{c(q^k \sigma)}{1 - d(q^{k+1} \sigma)} \right) (a(q^n \sigma) + b(q^n \sigma) D(1)). \quad (70)$$

Grâce à une division, on en déduit l'expression de $D(1)$:

$$D(1) = - \frac{\sum_{n \geq 0} \left(\prod_{k=0}^{n-1} \frac{c(q^k \sigma)}{1 - d(q^{k+1} \sigma)} \right) a(q^n \sigma)}{\sum_{n \geq 0} \left(\prod_{k=0}^{n-1} \frac{c(q^k \sigma)}{1 - d(q^{k+1} \sigma)} \right) b(q^n \sigma)}. \quad (71)$$

Pour compacter l'écriture du terme issu de n itérations on pose

$$I_n := \prod_{k=0}^{n-1} \frac{c(q^k \sigma)}{1 - d(q^{k+1} \sigma)}. \quad (72)$$

L'expression précédente se récrit

$$D(1) = -\frac{\sum_{n \geq 0} I_n a(q^n \sigma)}{\sum_{n \geq 0} I_n b(q^n \sigma)}. \quad (73)$$

On simplifie alors les expressions de $c(q^k s)$ et $1 - d(q^{k+1} s)$:

$$c(q^k s) = \frac{q}{1-q} \frac{(q-t)q^k s}{(1-q^{k+1} s)(1-tq^k s)} \quad (74)$$

$$1 - d(q^{k+1} s) = \frac{K(q^{k+1} s)}{(1-q)(1-q^{k+1} s)(1-tq^{k+1} s)}. \quad (75)$$

où $K(s)$ est le noyau défini par 65. Si on note σ_2 la seconde racine de $K(s)$ on a la factorisation

$$K(s) = (1-q)t(s-\sigma)(s-\sigma_2). \quad (76)$$

Le coefficient de degré 0 en s donne la relation $(1-q)t\sigma\sigma_2 = (1-q)$. On a donc $\sigma_2 = 1/(t\sigma)$. On rassemble ces simplifications pour obtenir

$$\frac{c(q^k s)}{1 - d(q^{k+1} s)} = \frac{q}{1-q} \frac{(q-t)q^k s}{(1-q^{k+1} s)(1-tq^k s)} \frac{(1-q^{k+1} s)(1-tq^{k+1} s)}{t(q^{k+1} s - \sigma)(q^{k+1} s - 1/(t\sigma))}. \quad (77)$$

En $s = \sigma$ il y a quelques simplifications supplémentaires et finalement on obtient :

$$\frac{c(q^k \sigma)}{1 - d(q^{k+1} \sigma)} = \frac{q-t}{1-q} \frac{q^{k+1} \sigma}{(1-q^{k+1})(1-tq^{k+1} \sigma^2)} \frac{1-tq^{k+1} \sigma}{1-tq^k \sigma}. \quad (78)$$

On utilise la notation $(a)_n$, la simplification des termes $1-tq^k \sigma$ et l'égalité $\sum_{k=0}^{n-1} (k+1) = \binom{n+1}{2}$ pour exprimer le produit

$$I_n \equiv \prod_{k=0}^{n-1} \frac{c(q^k \sigma)}{1 - d(q^{k+1} \sigma)} = \left(\frac{(q-t)\sigma}{1-q}\right)^n \frac{q^{\binom{n+1}{2}}}{(q)_n (qt\sigma^2)_n} \frac{1 - q^n t\sigma}{1 - t\sigma}. \quad (79)$$

Comme on a

$$a(q^n \sigma) = \frac{tq^n \sigma}{1 - tq^n \sigma} \quad (80)$$

et

$$b(q^n \sigma) = \frac{q^{n+1} \sigma}{(1 - q^n \sigma)(1 - q^{n+1} \sigma)} \quad (81)$$

l'expression de $D(1) \equiv D(t, u; 1)$ donnée par (71) devient

$$D(t, u; 1) = -\frac{\sum_{n \geq 0} \frac{\sigma}{1-t\sigma} \left(\frac{(q-t)\sigma}{1-q}\right)^n \frac{q^{\binom{n+2}{2}-1}}{(q)_n (qt\sigma^2)_n} t}{\sum_{n \geq 0} \frac{\sigma}{1-t\sigma} \left(\frac{(q-t)\sigma}{1-q}\right)^n \frac{q^{\binom{n+2}{2}-1}}{(q)_n (qt\sigma^2)_n} \frac{q(1-tq^n \sigma)}{(1-q^n \sigma)(1-q^{n+1} \sigma)}}. \quad (82)$$

Il reste à simplifier par $\frac{\sigma}{1-t\sigma}$ pour obtenir l'expression de la proposition 3.4. \square

On observe que la bonne racine σ de $K(s)$ est une série entière en t dont les coefficients semblent être des polynômes en u à coefficients positifs. Plus précisément pour $u = v$, on retrouve en ajoutant la contribution du chemin vide la série qui énumère les chemins de Dyck selon le nombre d'interactions *inférieures*, cf proposition 3.1. D'autre part, si on considère σ comme une variable supplémentaire, le numérateur et le dénominateur sont deux séries *hypergéométriques basiques*, notion que l'on présentera plus tard. En $u = 1$, le terme $(q - t) = t(u - 1)$ s'annule, la formule se simplifie en

$$D(t, u) = -\frac{t}{\frac{(1-t\sigma(t, 1))}{t(1-\sigma(t, 1))(1-t\sigma(t, 1))}} = \sigma(t, 1) - 1. \quad (83)$$

et on retrouve bien la série génératrice des chemins de Dyck non-vides. Les premiers termes du développement selon t de $D(t, u)$ sont

$$\begin{aligned} & t + (1 + u)t^2 + (u^2 + 3u + 1)t^3 + (7u^2 + u^3 + 5u + 1)t^4 \\ & + (19u^2 + 14u^3 + u^4 + 7u + 1)t^5 + (1 + 9u + 36u^2 + 26u^4 + 59u^3 + u^5)t^6 \\ & + (1 + 11u + 58u^2 + 162u^4 + u^6 + 150u^3 + 46u^5)t^7 \\ & + (1 + 13u + 85u^2 + 543u^4 + 79u^6 + 300u^3 + 408u^5 + u^7)t^8 + O(t^9). \end{aligned}$$

3.3 Interactions supérieures par équation q -algébrique

Les interactions supérieures semblent résister à la mise en équation algébrique de la section 3.1.3 car elles sont formées par des sommets qui apparaissent dans des facteurs différents des décompositions usuelles. Cependant lorsque deux sommets d'un chemin sont en interaction supérieure le sous-chemin qu'ils définissent est fortement contraint : il est de la forme $\bar{x}^k x^k$ avec k entre 1 et la hauteur des deux sommets en interactions. Cette régularité du chemin entre deux sommets en interaction supérieure va permettre de contourner partiellement la difficulté empêchant l'écriture d'une équation algébrique. En effet après un changement de modèle utilisant la régularité, et l'ajout d'une variable catalytique supplémentaire on va pouvoir écrire une équation de décomposition q -algébrique. La résolution de cette équation q -algébrique est encore une aventure. On propose deux approches : la première est l'objet de la section 3.3, la seconde de la section 3.4.

Des équations q -algébriques de cette forme ont été étudiées par Brak et Prellberg [BP95]. Ils suggèrent un changement de fonction inconnue qui permet de linéariser l'équation (en espérant que l'équation q -linéaire obtenue soit "simple" à résoudre). En s'inspirant de travaux de Janse Van Rensburg [JvR00a], on propose un changement de fonction inconnue un peu plus général, mais qui ne conduit pas à une équation aussi "simple" qu'espérée. Cependant, en utilisant un algorithme de Abramov, Petkovšek et Paule [APP98] on peut la résoudre, et retrouver l'expression de la proposition 3.4.

3.3.1 Changement de modèle pour la mise en équation q -algébrique

Voici le modèle que l'on va adopter. Un *chemin de Dyck à petits creux* est un chemin de Dyck évitant le facteur $\bar{x}\bar{x}.x.x$. Le chemin le plus à droite de la figure 3.12 est un exemple de chemin à petits creux, tous les autres sont des contre-exemples. La valuation d'un chemin à petits creux sera le produit de la valuation associée à sa demi-longueur et d'une valuation *ad hoc* de ses creux. Les pas montants sont valués par t . Un creux de hauteur k est valué

$$v(k) \equiv u \frac{1 - (ut)^{k+1}y}{1 - ut} \quad (84)$$

où u est toujours la variable qui compte les interactions supérieures dans le modèle initial et y est une nouvelle variable catalytique.

On note \mathcal{E} l'ensemble des chemins de Dyck à petits creux. La série génératrice $E(t, u; y)$ selon la valuation précédente vaut donc par définition

$$E(t, u; y) \equiv \sum_{d \in \mathcal{E}} \left(t^{\text{demilongueur}(d)} \prod_{c \text{ creux de } d} v(h(c)) \right) \quad (85)$$

où $h(c)$ est la hauteur du creux c .

Ce changement de modèle est pertinent comme le montre la proposition suivante.

Proposition 3.5 *La série $D(t, u)$ des chemins de Dyck dont la demi-longueur est comptée par t et le nombre d'interactions supérieures par u et la série $E(t, u; y)$ des chemins de Dyck à petits creux dont la demi-longueur est valuée par t et les creux de hauteur k par $u(1 - (ut)^{k+1}y)/(1 - ut)$ sont liées par la relation*

$$D(t, u) = E(t, u; 1).$$

La variable y est une variable catalytique puisqu'elle va permettre d'obtenir une équation q -algébrique mais qu'on l'ignore pour le résultat.

Un *creux double* dans un chemin de Dyck est une occurrence du facteur $\bar{x}\bar{x}.x.x$. La *réduction d'un creux double* consiste à récrire le facteur $\bar{x}\bar{x}.x.x$ en $\bar{x}.x$. Cette opération sur les chemins de Dyck préserve la structure de chemin de Dyck tout en diminuant de 1 la demi-longueur du chemin et supprimant une interaction supérieure. La *suite des hauteurs des pics* $shp(d)$ d'un chemin de Dyck d est la suite de la hauteur de ses pics lorsque le chemin est lu de gauche à droite. La suite des hauteurs de pic est un invariant de l'opération de réduction des creux doubles. On utilise cet invariant pour regrouper en classes les chemins de Dyck. Les chemins à petits creux forment un système de représentants de ces classes. La valuation des petits creux dans la classe \mathcal{E} permet de retrouver la valuation de la classe représentée par un chemin à petits creux.

Preuve.(Proposition 3.5) Deux chemins de Dyck sont *pics-équivalents* si leurs suites des hauteurs des pics sont identiques. Cette relation est bien une relation d'équivalence. On

partitionne l'ensemble des chemins de Dyck en classes d'équivalence. La figure 3.12 donne la classe des chemins dont la suite des hauteurs des pics est 3, 3, 2, 1.

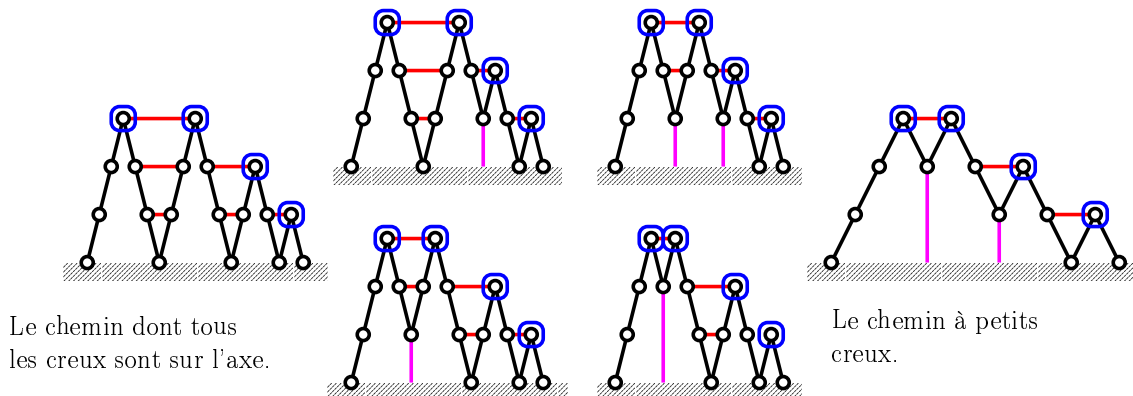


FIG. 3.12 – Une classe de chemins de Dyck pics-équivalents

Dans chaque classe de pic-équivalence il y a un unique chemin à petit creux. Il s'obtient à partir de n'importe quel chemin de la classe en réduisant les creux doubles tant qu'il en existe. L'unicité de ce chemin provient du fait que le seul chemin à petits creux dont la suite des hauteurs des pics est $(h_i)_{1 \leq i \leq n}$ est le chemin dont la suite des hauteurs des creux est $(\min(h_i, h_{i+1}) - 1)_{1 \leq i \leq n-1}$. L'ensemble des chemins à petits creux est donc bien un système de représentants des classes de chemins pics-équivalents. Cela se traduit en termes de série génératrice par l'égalité

$$D(t, u) = \sum_{e \in \mathcal{E}} \left(\sum_{d | \text{shp}(d) = \text{shp}(e)} t^{\text{demilongueur}(d)} u^{\text{superieures}(d)} \right).$$

On considère la classe représentée par le chemin de Dyck à petits creux e . On remarque que les réductions de deux creux différents sont indépendantes. Les chemins de cette classe peuvent donc s'obtenir en indiquant combien de réductions de creux doubles ont été nécessaires pour aboutir à chaque petit creux. Un petit creux à la hauteur k peut s'obtenir par 0 à k réductions de creux doubles. Soit A et B deux sommets en interaction au-dessus d'un petit creux à la hauteur k . La figure 3.3.1 illustre la remarque suivante pour $k = 3$. La somme des valuations possibles pour les chemins initiaux entre les deux sommets A et B est donc

$$\sum_{i=0}^k (ut)^{i+1} = ut(1 - (ut)^{k+1}) / (1 - ut).$$

Toutes les interactions supérieures sont traitées par ce procédé puisque dans un chemin à petits creux toutes les interactions supérieures sont juste au-dessus d'un petit creux. Pour ne pas compter deux fois le pas montant formant le petit creux on divise par t la valuation des creux. La valuation d'un petit creux est la somme de deux termes de nature différente. Le terme $u/(1 - ut)$ est une constante alors que $-u(ut)^{k+1}/(1 - ut)$ dépend

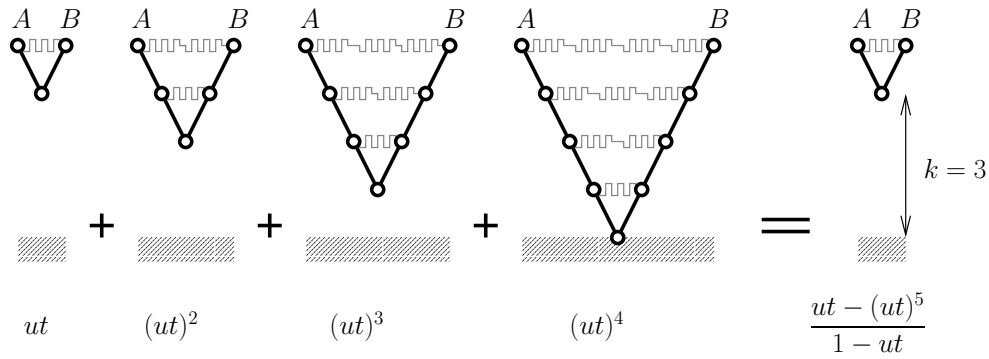


FIG. 3.13 – Valuation d’un petit creux q d’un représentant

géométriquement de la hauteur du creux. Pour pouvoir prendre en compte la dépendance géométrique du second terme dans les équations de décomposition on introduit la variable catalytique y et on aboutit à la valuation $v(k)$ donnée par (84). \square

3.3.2 Mise en équation q -algébrique

Avant de s’intéresser aux chemins à petits creux dont les creux à la hauteur k sont valués par $v(k)$ on va s’intéresser au comptage des interactions supérieures dans les chemins de Dyck à petits creux. Cela permet d’apprécier pourquoi ces chemins sont plus simples que les chemins de Dyck quelconques lorsque l’on s’intéresse aux interactions supérieures.

Il est possible d’écrire une équation de décomposition algébrique pour les interactions supérieures sur les chemins à petits creux *non valués*. En effet, l’absence du facteur $\bar{x}.x.x$ implique une bijection entre les interactions supérieures et les petits creux. Les chemins à petits creux se factorisent selon le premier retour à l’axe.

Proposition 3.6 *La série génératrice $F(t, u)$ des chemins de Dyck à petits creux dont la demi-longueur est comptée par t et le nombre d’interactions supérieures par u satisfait l’équation*

$$F(t, u) = t + tF(t, u) + ut^2F(t, u) + tuF(t, u) + (tu)^2F(t, u)^2. \quad (86)$$

Cette série satisfait donc

$$F(t, u) = \frac{1 - t - t^2u - ut - \sqrt{(t - 1)(u^2t^3 - u^2t^2 + 2t^2u + t + 2ut - 1)}}{2u^2t^2} \quad (87)$$

Preuve. Pour factoriser on considère le premier retour sur l’axe. L’évitement du facteur $\bar{x}x.x$ impose parfois la présence d’un pic $x\bar{x}$ après le premier retour.

La figure 3.14 illustre la décomposition qui donne directement l’équation algébrique satisfaite par $F(t, u)$. La résolution de cette équation et le choix de la seule solution qui

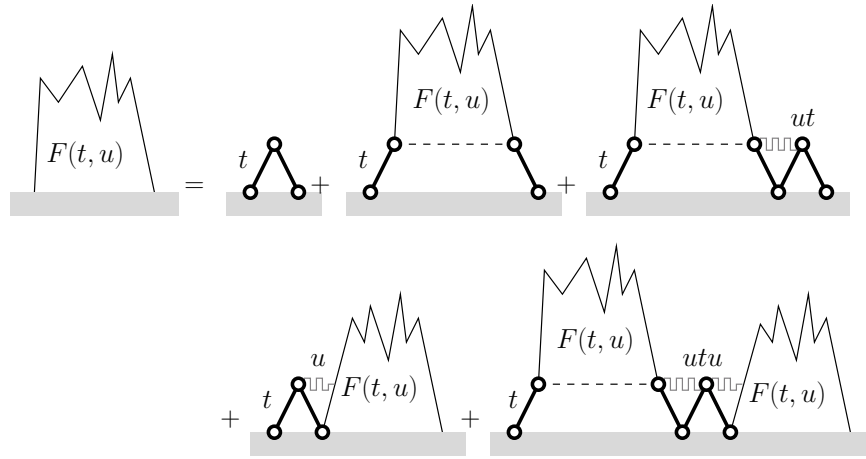


FIG. 3.14 – Décomposition algébrique des chemins à petits creux

soit une série formelle en t mène à la solution. □

La mise en équation des chemins à petits creux valué $v(k)$ correspond à la même décomposition des chemins de Dyck à petits creux. Il faut juste s'assurer que l'on peut suivre la dépendance géométrique du terme $-u(ut)^{k+1}y/(1-ut)$ grâce à la variable catalytique y . À nouveau on utilise la notation $q \equiv ut$.

Proposition 3.7 *La série génératrice $E(t, u; y) \equiv E(y)$ des chemins de Dyck à petits creux dont la demi-longueur est valuée par la variable t et les creux de hauteur k par $q(1 - q^{k+1}y)/t/(1 - q)$ satisfait l'équation q -algébrique :*

$$E(y) = t + t \left(1 + q \frac{1 - qy}{1 - q} \right) E(qy) + q \frac{1 - qy}{1 - q} E(y) + \left(q \frac{1 - qy}{1 - q} \right)^2 E(qy) E(y). \quad (88)$$

Preuve. La factorisation au premier retour sur l'axe des chemins à petits creux, illustrée par la figure 3.14, permet d'écrire une équation q -algébrique. Au premier retour sur l'axe on discute suivant que ce retour est précédé ou non d'une double-descente. Si avant le premier retour il n'y a qu'un pic de hauteur 1, la suite du chemin peut être vide ou un chemin de Dyck à petits creux. Dans ce second cas, il y a un creux à la hauteur 0 entre le pic et le reste du chemin. La série génératrice de ces cas est donc $t + t \frac{q(1-xy)}{t(1-q)} E(y)$. Si le chemin ne débute pas par un pic de hauteur 1, il apparaît un facteur qui est un mot de Dyck à petits creux, translaté de 1 vers le haut. Pour prendre en compte la modification des valuations des creux de ce facteur il suffit de remplacer y par qy . Le retour sur l'axe après une double descente implique que s'il y a un second sous-mot de Dyck il débute nécessairement par un pic. En effet sinon il existerait un motif $\bar{x}.\bar{x}.x.x$ dans le chemin de Dyck. On prend aussi en compte la valuation des creux formés dans ces différents cas pour obtenir la série génératrice $tE(qy) \left(1 + \frac{q(1-xy)}{t(1-q)} t + \left(\frac{q(1-xy)}{t(1-q)} \right)^2 tE(y) \right)$.

Au final on obtient l'équation q -algébrique :

$$E(y) = t + t \frac{q(1-xy)}{t(1-q)} E(y) + tE(xy) + tE(xy) \frac{q(1-xy)}{t(1-q)} t + tE(xy) \left(\frac{q(1-xy)}{t(1-q)} \right)^2 tE(y). \quad (89)$$

Après quelques coquetteries sur l'expression précédente on obtient l'expression de la proposition. \square

3.3.3 Résolution de l'équation q -algébrique par linéarisation

Il n'existe pas de technique générale pour la résolution des équations q -algébriques, même dans le cas particulier d'une combinaison linéaire de 1, $D(y)$, $D(xy)$ et $D(xy)D(y)$. Cependant, dans ce cadre, Brak et Prellberg ont suggéré un changement de fonction inconnue de la forme $D(y) = f(H(y), H(xy))$. La recherche des solutions sous cette forme se réduit à la recherche des solutions d'une équation q -linéaire satisfaite par $H(y)$. Dans l'idéal, on espère que cette équation q -linéaire ait une solution hypergéométrique basique. Soit \mathbb{A} un corps. Une série $B(y) = \sum_{n \geq 0} b_n y^n$ de $\mathbb{A}[[y]]$ est *hypergéométrique basique* (de base $q \in \mathbb{A}$) si il existe une fraction rationnelle $r_B(X) \in \mathbb{A}(X)$, telle que pour tout $n \geq 0$,

$$\frac{b_{n+1}}{b_n} = r_B(q^n).$$

La fraction $r_B(X)$ est la *raison* de cette série hypergéométrique. Une telle série est caractérisée par sa raison $r_B(X)$ et son premier terme $B(0) = b_0$, car la raison permet d'exprimer b_n en fonction de b_0 . Par exemple la seule série de raison $r_B(X) = X/(1-qX)$ telle que $B(0) = 1$ est

$$B(y) = \sum_{n \geq 0} \left(B(0) \prod_{k=0}^{n-1} r_B(q^k) \right) y^n = \sum_{n \geq 0} y^n \prod_{k=0}^{n-1} \frac{q^k}{1-q^{k+1}} = \sum_{n \geq 0} \frac{q^{\binom{n}{2}}}{(q)_n} y^n.$$

Pour mettre en valeur les discussions sur les choix de linéarisation on présente dans l'annexe A des méthodes connues permettant de trouver des solutions hypergéométriques basiques d'équation q -linéaires. Ces trois méthodes utilisent les mêmes idées, elles sont de plus en plus perfectionnées.

Rappelons d'abord la forme proposée par Brak et Prellberg [BP95] :

$$E(y) = \alpha \frac{H(xy)}{H(y)} + \beta(y). \quad (90)$$

Inspiré par un exemple de Van Rensburg [JvR00a], on généralise un peu cette forme pour obtenir une linéarisation qui permette de conclure le calcul. En effet, pour être totalement

satisfaisante, l'équation q -linéaire doit posséder des solutions d'une certaine forme : des séries hypergéométriques basiques. Ce premier changement de fonction inconnue conduit à une équation q -linéaire dont la résolution nécessite la méthode la plus complexe, qui a été proposée par Abramov, Petkovsek et Paule [APP98].

Résolution par une linéarisation inspirée de Brak, Prellberg et Van Rensburg

Le changement de modèle de la partie 3.3.2 a permis d'écrire une équation de décomposition q -algébrique de degré 2. Pour les équations algébriques il existe une méthode de résolution au moins jusqu'au degré 4. Pour les équations q -algébriques, même le cas du degré 1, les équations q -linéaires, reste difficile et on ne sait pas le résoudre en toute généralité, cf. l'annexe A. Cependant, l'équation q -algébrique (88) possède une forme particulière :

$$f_1(y)E(qy)E(y) + f_2(y)E(qy) + f_3(y)E(y) + f_4(y) = 0. \quad (91)$$

Brak et Prellberg [BP95] ont proposé un changement de fonction inconnue qui permet de linéariser cette équation : la nouvelle fonction inconnue satisfait une équation q -linéaire. Une variation de cette technique, en partie inspirée d'un article de Janse Van Rensburg [JvR00a], permet d'obtenir une solution de l'équation q -algébrique.

Proposition 3.8 *La série génératrice $E(t, u; y)$ des chemins de Dyck à petits creux, solution de l'équation q -algébrique (88), est*

$$E(t, u; y) = \frac{(1-q)^2 t\sigma H(qy)}{q^2 (1-qq)^2 H(y)} - \frac{t(1-q)(1-q^2y)}{q^2 (1-qq)^2} \quad (92)$$

avec

$$H(t, u; y) = \sum_{n \geq 0} \left(\frac{(q-t)\sigma}{1-q} \right)^n \frac{q^{\binom{n+1}{2}}}{(q)_n (qt\sigma^2)_n} \frac{q - t\sigma q^n + (t\sigma q^n)^2}{q - t\sigma + t^2\sigma^2} y^n. \quad (93)$$

et

$$\sigma = \frac{1+t-2q - \sqrt{(-1+t)(t-1+4q-4q^2)}}{2t(q-1)}. \quad (94)$$

Le terme constant $q-t\sigma+t^2\sigma^2$ de $H(t, u; y)$ n'est présent que pour garantir que $H(0) = 1$ par convention.

Preuve.

La preuve de cette proposition tient en deux étapes ; le choix d'un changement de fonction inconnue permettant de linéariser le problème puis la résolution de l'équation q -linéaire. On explique d'abord le cheminement qui a conduit à un changement de fonction inconnue donnant une équation q -linéaire dont on peut trouver une solution.

Pour la résolution d'une équation q -algébrique de la forme (91), Brak et Prellberg suggèrent de chercher des solutions sous la forme

$$E(y) = \alpha \frac{H(qy)}{H(y)} - \frac{f_2(y)}{f_1(y)} \quad (95)$$

où α est indépendant de y . La nouvelle fonction inconnue $H(y)$ est une série formelle en y telle que $H(0) = 1$. Elle satisfait l'équation q -linéaire

$$\alpha^2 H(q^2 y) + \alpha \left(\frac{f_3(y)}{f_1(y)} - \frac{f_2(qy)}{f_1(qy)} \right) H(qy) + \left(\frac{f_4(y)}{f_1(y)} - \frac{f_3(y)f_2(y)}{f_1(y)^2} \right) H(y). \quad (96)$$

L'hypothèse sur α permet de déterminer sa valeur par l'évaluation en $y = 0$ de l'équation q -linéaire. Idéalement, ce changement de fonction inconnue mène à une équation q -linéaire de la forme présentée dans la proposition A.1, en particulier de degré 1 en y .

Pour l'équation (88), cette linéarisation conduit à une équation q -linéaire de degré 6 en y . Cette équation q -linéaire n'est donc pas de la forme permettant d'appliquer la proposition A.1. L'évaluation en $y = 0$ de cette équation impose que α soit une racine de

$$P(\alpha) \equiv (q-1)^4 t + (q-1)(t+1-2q)q^2 \alpha + q^4 \alpha^2. \quad (97)$$

On remarque que ce polynôme est lié au polynôme $K(s)$ de l'équation (65) qui est le noyau de la méthode du noyau utilisée lors du traitement de l'équation fonctionnelle à la Temperley. Plus précisément

$$P\left(\frac{ts(q-1)^2}{q^2}\right) = t(q-1)^3 K(s). \quad (98)$$

Par définition de σ , $K(\sigma) = 0 = K(\sigma_2)$ donc les deux racines du polynôme P sont $\frac{t\sigma(q-1)^2}{q^2}$ et $\frac{(q-1)^2}{q^2\sigma}$. On en déduit que α doit être une de ces deux valeurs.

On note $\sum_{n \geq 0} h_i y^i$ le développement selon la variable y de la série formelle $H(y)$. Le fait que α ne soit pas une fraction rationnelle en t et q est une difficulté supplémentaire pour la solution de l'équation q -linéaire. Il est notamment plus difficile de décider expérimentalement si la suite (h_i) de fractions en t , q et σ correspond aux termes d'une suite hypergéométrique basique. Sur cet exemple l'auteur n'a pas réussi. Les deux premières méthodes de résolution de l'équation q -linéaire semblent peu prometteuses, on se tourne alors vers l'algorithme de Abramov, Paule et Petkovšek. Les variables indéfinies qui apparaissent dans les calculs sont celles de l'algorithme A présenté dans l'annexe A. On ne rapporte pas ici les calculs mais toutes les combinaisons de diviseurs de $ContrainteA(X)$ et $ContrainteB(X)$ conduisent à des polynômes $ContrainteDegréC(W)$ qui n'ont pas de racine de la forme q^n avec $n \in \mathbb{N}$. L'équation q -linéaire obtenue par la linéarisation de Brak et Prellberg n'a donc pas de solution hypergéométrique basique.

On a épuisé les techniques pour résoudre l'équation q -linéaire obtenue par le changement de fonction inconnue proposé par Brak et Prellberg. Janse Van Rensburg a linéarisé une q -équation algébrique grâce à un changement de fonction inconnue légèrement différent, de la forme

$$E(y) = \frac{\alpha H(qy)}{y H(y)} - \frac{\beta(y)}{y} \quad (99)$$

où $\beta(y)$ est un polynôme en y [JvR00a]. On retient de cette variation que le coefficient de $H(qy)/H(y)$ peut dépendre de y . Dans ce cas précis, ce coefficient est une fraction dont le numérateur α est indépendant de y et le dénominateur est aussi le dénominateur du terme constant. Dans la partie suivante, on donnera une explication combinatoire à cette égalité des dénominateurs dans le cas traité par Janse Van Rensburg. Pour l'équation (88), cette remarque suggère d'étudier un changement de fonction inconnue de la forme

$$E(y) = \frac{\alpha H(qy)}{q^2 (1 - qy)^2 H(y)} - \frac{t(1 - q)(1 - q^2y)}{q^2 (1 - qy)^2} \quad (100)$$

Le changement de variable (100) conduit à l'équation q -linéaire

$$l_0(y)H(y) + l_1(y)H(qy) + l_2(y)H(q^2y) = 0 \quad (101)$$

où

$$\begin{aligned} l_0(y) &= -t(-1 + q)^4 (-1 + q^2y)^2 \\ l_1(y) &= \alpha(1 - q)(-2q + t + 1 - (-4q^2 + tq^2 + q + 2t)qy \\ &\quad + (-2q^3 + 2tq^2 - q^2 + t)q^2y^2 - q^5(-q + t)y^3) \\ l_2(y) &= -\alpha^2(-1 + qy)^2 \end{aligned}$$

En $y = 0$ on retrouve le polynôme (97) si on remplace $q^2\alpha$ par α . On en déduit les deux valeurs possibles de α en fonction de σ :

$$\alpha \in \left\{ t\sigma(1 - q)^2, \frac{(q - 1)^2}{\sigma} \right\}.$$

L'équation q -linéaire est de degré 3 en y , on ne peut donc pas appliquer la première technique de résolution de la proposition A.1. On ne peut pas non plus conjecturer simplement l'éventuelle raison de la série hypergéométrique basique. On se tourne donc vers l'algorithme de Abramov, Petkovšek et Paule. On cherche à déterminer par une analyse une solution de l'équation q -algébrique. Lorsqu'un nombre fini de choix sont possibles on ne présente ci-dessous que celui qui mène au résultat. Le premier de ces choix est de poser $\alpha = t\sigma(1 - q)^2$. Le changement de fonction inconnue que l'on adopte est donc

$$E(y) = \frac{(1 - q)^2 t\sigma H(qy)}{q^2 (1 - qy)^2 H(y)} - \frac{t(1 - q)(1 - q^2y)}{q^2 (1 - qy)^2} \quad (102)$$

L'équation q -linéaire (101) devient

$$\begin{aligned} l_0(y) &= t(1 - q)^2 (1 - q^2y)^2, \\ l_1(y) &= t(1 - q)\sigma \left((2q - t - 1 + (-4q^2 + tq^2 + q + 2t)qy \right. \\ &\quad \left. - (-2q^3 + 2tq^2 - q^2 + t)q^2y^2 + q^5(-q + t)y^3) \right), \\ l_2(y) &= t^2\sigma^2 (1 - q)^2 (1 - qy)^2 \end{aligned}$$

À partir de cette équation on calcule selon les indications de l'algorithme les polynômes

$$\text{Contrainte}A(X) \equiv t(1-q)\sigma q^5(q-t)X \quad (103)$$

et

$$\text{Contrainte}B(X) \equiv t(q-1)((1-q) + \sigma(1+t-2q)qX + t\sigma^2(1-q)q^2X^2) \quad (104)$$

On remarque que $\text{Contrainte}B(X)$ est lié au noyau $K(s)$ de l'équation (65) par la relation

$$K(X) = \frac{1}{t(q-1)}\text{Contrainte}B(X/q/\sigma). \quad (105)$$

Comme les deux racines de $K(X)$ sont σ et $1/(t\sigma)$, les deux racines de $\text{Contrainte}B(X)$ sont $1/q$ et $1/(qt\sigma^2)$. Donc

$$\text{Contrainte}B(X) = t^2(q-1)^2\sigma^2q^2(X-1/q)(X-1/(qt\sigma^2)) = t(q-1)^2(1-qX)(1-qt\sigma^2X). \quad (106)$$

Les polynômes $a(X)$ et $b(X)$ divisent respectivement $\text{Contrainte}A(X)$ et $\text{Contrainte}B(X)$ ce qui laisse un nombre fini de possibilités. On pose donc

$$a(X) \equiv \text{Contrainte}A(X) \quad (107)$$

et

$$b(X) \equiv \text{Contrainte}B(X). \quad (108)$$

On utilise pour $\text{Contrainte}B(X)$ l'écriture de (104) et non celle de l'équation (106). Pour montrer l'équivalence de ces deux écritures on doit utiliser la relation satisfaite par σ et la suite va montrer que cela peut avoir une influence sur le calcul.

Le polynôme $P_{A,B}(X, Z, C(X))$ qui découle de ce choix de $a(X)$ et $b(X)$ est un peu trop volumineux pour qu'on le reproduise ici. C'est un polynôme de degré 4 en X dont le terme de degré minimal est de degré 0. Avec les notations de l'algorithme de Abramov, Petkovšek et Paule on a donc $K = 4$ et $L = 0$. Le coefficient de degré minimal est

$$q^4ZC(qX) + C(X)$$

et le coefficient de plus haut degré est

$$t^2\sigma^4q^6(q^2ZC(qX) + C(X)).$$

On en déduit donc les polynômes

$$\text{AnnulationDegr}eMin(Z) \equiv t^2(1-q)^4(Zq^4 + 1) \quad (109)$$

et

$$\text{AnnulationDegr}eMax(Z, W) \equiv t^4\sigma^4q^6(1-q)^4(ZWq^2 + 1) \quad (110)$$

L'annulation du polynôme $AnnulationDegreMin(Z)$ permet de déduire que

$$z = -\frac{1}{q^4}. \quad (111)$$

L'usage du résultant est ici inutile puisque z est une fraction rationnelle en q et t et donc il n'y a pas de calcul dans une extension algébrique de $\mathbb{Q}(t, q)$. Pour illustrer la méthode on calcule tout de même

$$ContrainteDegreC(W) = t^6 (1 - q)^8 \sigma^4 q^8 (q^2 - W). \quad (112)$$

Si on souhaite respecter scrupuleusement l'algorithme il faut tester si ce polynôme s'annule pour les q^i avec i allant de 0 à $T = 18$ (lemme A.2.2). Il s'avère que q^2 est la seule racine de $ContrainteDegreC(W)$ parmi ces valeurs. S'il existe le polynôme $C(X)$ est donc de degré 2 en X . On pose donc

$$ContrainteC(X) \equiv c_0 + c_1 X + X^2. \quad (113)$$

La substitution $C(X) = ContrainteC(X)$ dans le polynôme $P_{A,B}(X, Z, C(X))$ conduit à un polynôme dont l'annulation des coefficients donne un système linéaire en les inconnues c_0 et c_1 . Cette linéarité provient de la linéarité de $P_{A,B}(X, Z, C(X))$ en les fonctions $C(q^i X)$. Les coefficients de ce système sont des fractions rationnelles en t, q, Z et σ . On remplace Z par sa valeur $-1/q^4$. Une propriété remarquable de ce système est que l'on peut le résoudre sans utiliser la relation algébrique liant σ à t et q . On trouve

$$c(X) = \frac{q}{t^2 \sigma^2} - \frac{1}{t\sigma} X + X^2. \quad (114)$$

En regroupant les divers résultats intermédiaires on obtient

$$r(X) = \frac{(t - q)q\sigma X}{(q - 1)(1 - qX)(1 - qt\sigma^2 X)} \frac{q - qt\sigma X + q^2 t^2 \sigma^2 X^2}{q - t\sigma X + t^2 \sigma^2 X^2}. \quad (115)$$

On en déduit que

$$H(y) = \sum_{n \geq 0} \left(\frac{(t - q)\sigma}{q - 1} \right)^n \frac{q^{\binom{n+1}{2}}}{(q)_n (qt\sigma^2)_n} \frac{q - t\sigma q^n + (t\sigma q^n)^2}{q - t\sigma + t^2 \sigma^2} y^n. \quad (116)$$

On utilise la relation (54) satisfaite par σ pour récrire le terme $q - t\sigma + t^2 \sigma^2$ en $(q - t)(1 - \frac{t\sigma}{1 - q})$ lorsque $n \geq 1$. Cela permet d'écrire $H(y)$ sous la forme

$$H(y) = 1 + \sum_{n \geq 1} (q - t)^{n-1} \left(\frac{\sigma}{q - 1} \right)^n \frac{q^{\binom{n+1}{2}}}{(q)_n (qt\sigma^2)_n} \frac{q - t\sigma q^n + (t\sigma q^n)^2}{1 - \frac{t\sigma}{1 - q}} y^n. \quad (117)$$

Sous cette forme, il suffit de se rappeler que σ est une série formelle en t de paramètre u pour observer que chacun des termes de degré n en y est une série formelle en t de paramètre u . Le facteur $q^{\binom{n+1}{2}}$ permet de dire que $H(y)$ la somme de ces termes est aussi une série formelle en t dont le terme constant est 1. En poussant un peu plus les calculs on obtient les développements limités à l'ordre 2 en t

$$\begin{aligned} H(y) &= 1 + \mathcal{O}(t^2), \\ H(qy) &= 1 + \mathcal{O}(t^2), \\ \sigma &= 1 + q + \mathcal{O}(t^2). \end{aligned}$$

En utilisant ces développements dans le changement de variable (102) on obtient

$$\begin{aligned} E(y) &= \frac{(1-q)^2 t(1+q + \mathcal{O}(t^2))(1 + \mathcal{O}(t^2))}{q^2(1-qqy)^2(1 + \mathcal{O}(t^2))} - \frac{t(1-q)(1-q^2y)}{q^2(1-qqy)^2} \\ &= \frac{(1-q)t((1-q^2 + \mathcal{O}(t^2))(1 + \mathcal{O}(t^2)) - (1-q^2y))}{q^2(1-qqy)^2(1 + \mathcal{O}(t^2))} \\ &= \mathcal{O}(t) \end{aligned}$$

On en déduit que $E(y)$ obtenue en fonction de $H(y)$ par le changement de fonction inconnue de l'équation (102) est aussi une série formelle en t . Or d'après l'équation q -algébrique (88) il existe une unique solution de $E(y)$ qui soit une série formelle en t . L'expression de $E(y)$ obtenue à partir de $H(y)$ est donc bien la série des chemins de Dyck à petits creux. \square

Le début du développement de $E(y)$ est

$$\begin{aligned} &t + (1+u)t^2 + (2u^2 - yu^2 + 3u + 1)t^3 + (4u^3 - 3yu^3 - 2yu^2 + 9u^2 + 5u + 1)t^4 \\ &+ (8u^4 - 8yu^4 + u^4y^2 + 25u^3 - 11yu^3 + 21u^2 - 2yu^2 + 7u + 1)t^5 + \mathcal{O}(t^6). \end{aligned}$$

3.4 Interactions supérieures en une page

Le changement de modèle pour l'écriture de l'équation q -algébrique n'a que modestement modifié le modèle initial, seule la variable y est apparue. Les perturbations induites par cette variable sur la mise en équation à la Temperley sont mineures ce qui permet de trouver une autre expression développée selon y . La forme de la solution précédente suggère un autre changement de fonction inconnue pour la linéarisation de l'équation q -algébrique. Celui-ci ne semble pas avoir été proposé par Brak et Prellberg bien qu'il convienne pour la linéarisation pour les mêmes raisons. De plus l'équation q -linéaire obtenue est, cette fois-ci, de la forme "simple" espérée par Brak et Prellberg. Sa résolution conduit directement à la solution obtenue par l'équation de décomposition à la Temperley sans recours aux techniques de Abramov, Petkovšek et Paule.

3.4.1 Résolution utilisant une équation fonctionnelle à la Temperley

On a réussi à résoudre l'équation q -algébrique. Mais une autre approche consiste à remarquer que les chemins à petits creux peuvent aussi se traiter à l'aide d'une équation à la Temperley. Par rapport au modèle initial la variable y est la principale différence. En effet la notion de petit creux est compatible avec l'écriture de l'équation de décomposition par la suppression du dernier pic. On ajoute donc à nouveau la variable catalytique s comptant la longueur de la dernière descente.

Proposition 3.9 *La série $E(t, u; y, s) \equiv E(s)$ des chemins de Dyck à petits creux dont la demi-longueur est évaluée par la variable t , les creux à hauteur k par $\frac{q(1 - q^{k+1}y)}{t(1 - q)}$ et la longueur de la dernière descente par s satisfait l'équation :*

$$E(s) = \frac{ts}{1 - ts} + \frac{ts}{1 - ts} \frac{q}{1 - q} \left(E(s) - yE(qs) \right) + \frac{q}{1 - q} \left(\frac{s}{1 - s} \left(E(1) - E(s) \right) - y \frac{qs}{1 - qs} \left(E(1) - E(qs) \right) \right).$$

Preuve. On partitionne l'ensemble des chemins de Dyck à petits creux. On regroupe ceux qui ne contiennent pas de creux, ceux dont le dernier pic est strictement au-dessus de l'avant dernier et ceux dont le dernier pic est au-dessous de l'avant dernier. Cette preuve est similaire à celle de la proposition 3.3, on pourra s'en inspirer si on souhaite refaire les calculs intermédiaires.

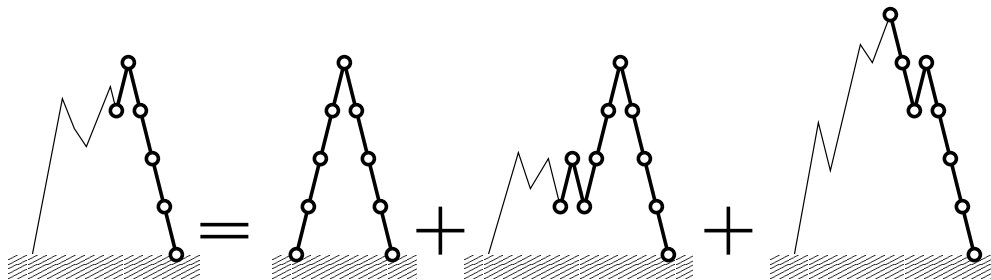


FIG. 3.15 – Décomposition pour les chemins de Dyck à petits creux

De la même façon que pour le modèle initial on peut exprimer la série génératrice de ces sous-ensembles en fonction de séries qui sont des spécialisations de la série $E(t, u, y; s)$ pour différentes valeurs de s . Pour les chemins sans creux on a

$$\frac{ts}{1 - ts}.$$

Pour les chemins dont le dernier pic est strictement au-dessus de l'avant dernier, le dernier creux appartient à un facteur $x\bar{x}xx$ et on a

$$\frac{ts}{1-ts} \frac{q}{1-q} (E(s) - yE(qs)).$$

Pour les chemins dont le dernier pic est au-dessous au sens large de l'avant-dernier, le dernier creux et le dernier pic appartiennent à un facteur $\bar{x}x\bar{x}$ et on a

$$\frac{q}{1-q} \left(\frac{s}{1-s} (E(1) - E(s)) - y \frac{qs}{1-qs} (E(1) - E(qs)) \right).$$

Il n'est pas surprenant que la mise en équation se passe bien car le changement de modèle effectué consiste à anticiper une des sommations géométriques et à la décorer grâce à variable y . Les sommations anticipées portent sur la hauteur du dernier creux du chemin. \square

Proposition 3.10 *Soit $E(t, u; y)$ la série génératrice des chemins de Dyck à petits creux dont la demi-longueur est comptée par t , et les petits creux à la hauteur h sont valués $\frac{q(1 - q^{h+1}y)}{1 - q}$. On a*

$$E(t, u; y) = - \frac{\sum_{n \geq 0} \frac{(q-t)^n \sigma^{n+1} q^{\binom{n+1}{2}} q^n t}{(1-q)^n (q)_n (qt\sigma^2)_n} y^n}{\sum_{n \geq 0} \frac{(q-t)^{n-1} \sigma^n q^{\binom{n+1}{2}} (q^n t\sigma - q - q^{2n} t^2 \sigma^2)}{(1-q)^n (q)_n (qt\sigma^2)_n} y^n}. \quad (118)$$

Preuve. La série inconnue satisfait l'équation de la proposition 3.9. À nouveau on retrouve une équation de la forme

$$E(s) = a(s) + b(s)E(1) + c(s)E(qs) + d(s)E(s)$$

avec

$$\begin{cases} a(s) &= \frac{ts}{1-ts} \\ b(s) &= \frac{q}{1-q} \left(\frac{s}{1-s} - y \frac{qs}{1-qs} \right) \\ c(s) &= \frac{q}{1-q} \left(y \frac{qs}{1-qs} - y \frac{ts}{1-ts} \right) \\ d(s) &= \frac{q}{1-q} \left(\frac{ts}{1-ts} - \frac{s}{1-s} \right). \end{cases}$$

Pour $y = 1$ on retrouve l'équation de décomposition (53) des chemins de Dyck. Les techniques d'élimination que l'on a employées utilisaient principalement les coefficients $c(s)$ pour l'itération et $d(s)$ pour la méthode du noyau. On observe que y se factorise dans

$c(s)$: chaque itération rajoute donc simplement un facteur y . D'autre part y n'apparaît pas dans $d(s)$, donc le noyau $1 - d(s)$ de cette équation est le même que celui de l'équation (64). La série formelle σ telle que $1 - d(\sigma) = 0$ est donc inchangée. La forme

$$E(1) = - \frac{\sum_{n \geq 0} \left(\prod_{k=0}^{n-1} \frac{c(q^k \sigma)}{1 - d(q^{k+1} \sigma)} \right) a(q^n \sigma)}{\sum_{n \geq 0} \left(\prod_{k=0}^{n-1} \frac{c(q^k \sigma)}{1 - d(q^{k+1} \sigma)} \right) b(q^n \sigma)}$$

reste correcte. Par rapport au cas précédent il suffit d'ajouter la variable y qui au vu des remarques précédentes se comporte poliment.

On se contente de cette discussion des variations sur les calculs précédents qui conduisent à la formule

$$E(1) = - \frac{\sum_{n \geq 0} \left(\frac{(q-t)\sigma}{1-q} \right)^n \frac{t\sigma q^{\binom{n+2}{2}-1}}{(q)_n (qt\sigma^2)_n (1-t\sigma)} y^n}{\sum_{n \geq 0} \left(\frac{(q-t)\sigma}{1-q} \right)^n \frac{t\sigma q^{\binom{n+2}{2}-1}}{(q)_n (qt\sigma^2)_n} \frac{1-tq^n\sigma}{1-t\sigma} y^n \frac{q}{1-q} \left(\frac{q^n\sigma}{1-q^n\sigma} - y \frac{q^{n+1}\sigma}{1-q^{n+1}\sigma} \right)}.$$

Le numérateur est déjà développé selon la variable y . Le dénominateur est presque sous cette forme : pour $n \geq 1$ le coefficient de y^n est

$$\frac{(q-t)^{n-1} \sigma^n q^{\binom{n+1}{2}+1} (q^{n-1} t\sigma - 1 - q^{2n-1} t^2 \sigma^2)}{(1-q)^n (q)_n (qt\sigma^2)_n (1-t\sigma)}. \quad (119)$$

Le calcul de ce terme a nécessité de nombreuses simplifications. Un terme constant du dénominateur est

$$\frac{q}{1-q} \frac{\sigma}{1-\sigma} \quad (120)$$

et ne correspond pas à première vue à l'évaluation en $n = 0$ du terme (119) qui donne

$$\frac{t\sigma - q - t^2\sigma^2}{(q-t)(1-t\sigma)}. \quad (121)$$

Toutefois, en utilisant l'équation algébrique (54) satisfaite par σ on montre que ces deux termes sont égaux. Une dernière simplification par $1 - t\sigma$ conduit à la formule de la proposition. \square

On remarque que l'expression de $E(y)$ de l'équation (118) est un quotient de deux séries hypergéométriques basiques de raison

$$r_{Num}(X) = \frac{(q-t)\sigma q^2 X}{(1-q)(1-qX)(1-qt\sigma^2 X)} \quad (122)$$

pour le numérateur et

$$r_{Den}(X) = \frac{(q-t)\sigma qX}{(1-q)(1-qX)(1-qt\sigma^2 X)} \frac{q-qt\sigma X - q^2 t^2 \sigma^2 X^2}{q-t\sigma X - t^2 \sigma^2 X^2} \quad (123)$$

pour le dénominateur. Cela suggère d'introduire la série hypergéométrique

$$J(y) \equiv \sum_{n \geq 0} \left(\frac{(q-t)\sigma}{1-q} \right)^n \frac{q^{\binom{n+1}{2}}}{(q)_n (qt\sigma^2)_n} y^n \quad (124)$$

dont la raison est le plus petit "diviseur" commun des deux séries $r_{Num}(X)$ et $r_{Den}(X)$. La série $J(y)$ est pertinente pour ce problème car d'une part elle satisfait une équation q -linéaire simple et que d'autre part $E(y)$ admet une expression simple en fonction de $J(y)$ et $J(qy)$. Ces remarques forment la proposition suivante.

Proposition 3.11 *Soit $J(t, u; y) \equiv J(y)$ la série formelle en y définie par $J(0) = 1$ et*

$$J(y) + \left(\frac{1-xy}{1-q} (q-t) - 1 \right) \sigma J(qy) + t\sigma^2 J(q^2 y) = 0. \quad (125)$$

La série génératrice $E(t, u; y) \equiv E(y)$ des chemins à petits creux dont la demi-longueur est évaluée par t et les petits creux à la hauteur k par $\frac{q(1-q^{k+1}y)}{t(1-q)}$ vaut

$$E(y) = \frac{t\sigma J(qy)}{J(y) - \frac{1-xy}{1-q} t\sigma J(qy)}. \quad (126)$$

Preuve.

• On cherche d'abord l'équation q -linéaire dont $J(y)$ est solution. Guidé par les termes $(q)_n (qt\sigma^2)_n$ on s'intéresse à la somme

$$S(y) = \left(J(y) - J(qy) \right) - t\sigma^2 \left(J(qy) - J(q^2 y) \right).$$

En utilisant l'expression explicite de $J(y)$ on a

$$S(y) = \sum_{n \geq 0} \frac{(q-t)^n \sigma^n q^{\binom{n+1}{2}}}{(1-q)^n (q)_n (t\sigma^2)_n} \left((1-q^n) - t\sigma^2 (q^n - q^{2n}) \right) y^n.$$

Pour $n = 0$, le terme en y^0 s'annule ; on peut donc restreindre la somme aux $n \geq 1$. De plus la combinaison $S(y)$ a été choisie de sorte que $\left((1-q^n) - t\sigma^2 (q^n - q^{2n}) \right)$ se factorise en $(1-q^n)(1-q^n t\sigma^2)$ ce qui permet une simplification avec les derniers termes de $(q)_n (qt\sigma^2)_n$. On a donc

$$S(y) = \sum_{n \geq 1} \frac{(q-t)^n \sigma^n q^{\binom{n+1}{2}}}{(1-q)^n (q)_{n-1} (t\sigma^2)_{n-1}} y^n$$

où on reconnaît l'expression

$$S(y) = \frac{(q-t)\sigma qy}{1-q} \sum_{n \geq 1} \frac{(q-t)^{n-1} \sigma^{n-1} q^{\binom{n}{2}}}{(1-q)^{n-1} (q)_{n-1} (t\sigma^2)_{n-1}} q^{n-1} y^{n-1}.$$

La réindexation de la somme par le changement de variable $n = n - 1$ nous permet de reconnaître l'expression de $J(qy)$, ce qui mène à la relation

$$S(y) = \frac{(q-t)\sigma qy}{(1-q)} J(qy).$$

Cela donne, après quelques simplifications, l'équation q -linéaire (125) satisfaite par $J(y)$. Comme cette équation q -linéaire est de degré 1 en y , la série $J(y)$ est défini par cette relation et sa valeur en $y = 0$: $J(y = 0) = 1$.¹

• On cherche une expression de $E(y)$ en fonction de $J(y)$ en partant de l'expression (118). Le numérateur s'écrit

$$G(y) = \sum_{n \geq 0} \frac{(q-t)^n \sigma^{n+1} q^{\binom{n+1}{2}} q^n t}{(1-q)^n (q)_n (qt\sigma^2)_n} = t\sigma J(qy).$$

Le dénominateur s'écrit

$$H(y) = \sum_{n \geq 0} \frac{(q-t)^{n-1} \sigma^n q^{\binom{n+1}{2}} (q^n t\sigma - q - q^{2n} t^2 \sigma^2)}{(1-q)^n (q)_n (qt\sigma^2)_n} y^n = \frac{(t\sigma J(qy) - qJ(y) - t^2 \sigma^2 J(q^2 y))}{q-t}.$$

On utilise la relation q -linéaire (125) satisfaite par $J(y)$ pour exprimer

$$t^2 \sigma^2 J(q^2 y) = -tJ(y) + \left(t\sigma - \frac{t(q-t)(1-xy)\sigma}{1-q} \right) J(qy).$$

Après la simplification d'un terme en $t\sigma J(qy)$ et la simplification par $q-t$ le dénominateur se réécrit

$$H(y) = -J(y) + \frac{t\sigma(1-xy)}{1-q} J(qy).$$

On a donc

$$E(y) = \frac{t\sigma J(qy)}{J(y) - \frac{1-xy}{1-q} t\sigma J(qy)}.$$

□

¹De plus on remarque que l'évaluation de l'équation q -linéaire en $y = 0$ correspond exactement au noyau.

Lors de la résolution de l'équation q -algébrique, (proposition 3.8) on a prouvé que l'expression obtenue par l'analyse était la bonne car c'était une série formelle en t . On a utilisé le fait que l'équation q -algébrique définit une unique série formelle. Ce même type d'argument a été utilisé dans les preuves basées sur des équations de décomposition à la Temperley. Ici, l'équation (124) permet de constater que $J(y)$ est une série formelle en t de terme constant 1 et l'expression (126) permet d'en déduire de $E(y)$ est aussi une série formelle en t . L'argument est donc aussi applicable mais on peut faire un peu mieux. Suite à la proposition 3.11 on est en mesure de prouver par un calcul direct que l'expression (126) satisfait l'équation q -algébrique. Cela peut donner l'impression de mieux comprendre le lien entre cette solution et l'équation q -algébrique initiale.

Remarque 3.11.1 *On peut montrer que l'expression (126) satisfait l'équation q -algébrique (88) en utilisant une seule fois une relation de nature algébrique, la relation (125).*

Preuve. On injecte l'expression (126) de $E(y)$ en fonction de $J(y)$ dans le second membre de l'équation q -algébrique (88)

$$t + q \frac{1-xy}{1-q} \frac{t\sigma}{J(y) - \frac{1-xy}{1-q} t\sigma J(qy)} + t \left(1 + \frac{1-xy}{1-q}\right) \frac{t\sigma J(q^2y)}{J(qy) - \frac{1-q^2y}{1-q} t\sigma J(q^2y)} \\ + q^2 \frac{(1-xy)^2}{(1-q)^2} \frac{t\sigma J(qy) t\sigma J(q^2y)}{\left(J(y) - \frac{1-xy}{1-q} t\sigma J(qy)\right) \left(J(qy) - \frac{1-q^2y}{1-q} t\sigma J(q^2y)\right)}.$$

On réduit cette expression au dénominateur

$$\left(J(y) - \frac{1-xy}{1-q} t\sigma J(qy)\right) \left(J(qy) - \frac{1-q^2y}{1-q} t\sigma J(q^2y)\right)$$

et après des simplifications de nature rationnelle on obtient comme numérateur

$$tJ(y)J(qy) + (q-t)t\sigma \frac{1-xy}{1-q} J(qy)^2 - t^2\sigma^2 q \frac{1-xy}{1-q} J(qy)J(q^2y).$$

Il se factorise en

$$J(qy) \left(tJ(y) + (q-t) \frac{1-xy}{1-q} J(qy) - t\sigma q \frac{1-xy}{1-q} J(q^2y) \right) \quad (127)$$

et si le monde est bien fait, $E(y)$ satisfait l'équation q -algébrique, le second facteur doit être égal à

$$t\sigma \left(J(qy) - \frac{1-q^2y}{1-q} t\sigma J(q^2y) \right).$$

Cela suggère d'utiliser l'équation q -linéaire (125) satisfaite par $J(y)$ pour exprimer $J(y)$ en fonction de $J(qy)$ et $J(q^2y)$:

$$tJ(y) = t\sigma \left(1 - \frac{(q-t)(1-xy)}{1-q} \right) J(qy) - t^2\sigma^2 q \frac{1-xy}{1-q} J(q^2y).$$

Après une simplification par $t\sigma(q-t)\frac{1-xy}{1-q}J(xy)$ le second facteur du numérateur (127) devient :

$$t\sigma\left(J(xy) - t\sigma\left(1 + q\frac{1-xy}{1-q}\right)J(x^2y)\right).$$

Une dernière simplification de nature rationnelle permet de conclure. □

3.4.2 Une seconde linéarisation

La proposition 3.11 suggère un changement de fonction inconnue qui mène à une équation q -linéaire simple. Il est donc naturel de se demander si un tel changement de fonction inconnue permet de retrouver cette équation q -linéaire par linéarisation de l'équation q -algébrique. Confronté à une équation q -algébrique de la forme

$$f_1(y)E(xy)E(y) + f_2(y)E(xy) + f_3(y)E(y) + f_4(y) = 0 \quad (128)$$

on observe que le changement de fonction inconnue

$$E(y) = \frac{J(xy)}{\alpha J(y) - \frac{f_2(q^{-1}y)}{f_4(q^{-1}y)}J(xy)} \quad (129)$$

conduit à l'équation q -linéaire

$$f_4(y)\alpha^2 J(y) + \left(f_3(y) - \frac{f_2(q^{-1}y)f_4(y)}{f_4(q^{-1}y)}\right)\alpha J(xy) + \left(f_1(y) - \frac{f_2(y)f_3(y)}{f_4(y)}\right)J(x^2y) = 0. \quad (130)$$

Ce changement de fonction inconnue est fructueux pour les mêmes raisons que celui proposé par Brak et Prellberg (équation (95)). L'expression (128) devient une fraction rationnelle en $J(y)$ et $J(xy)$ et $J(x^2y)$. L'annulation de cette fraction est impliquée par l'annulation de son numérateur Num qui est une combinaison linéaire de $J(y)J(xy)$, $J(xy)^2$, $J(x^2y)J(xy)$ et $J(y)J(x^2y)$. Le choix de la valeur $-f_2(q^{-1}y)/f_4(q^{-1}y)$ dans le changement (129) annule le terme $J(y)J(x^2y)$ dans Num . La factorisation de $J(xy)$ donne l'équation q -linéaire.

Ce changement de variable permet d'obtenir une seconde preuve de la proposition 3.11 en utilisant maintenant l'équation q -algébrique et ce nouveau changement de variable.

Preuve. Dans le cas de l'équation q -algébrique (88), le changement de fonction inconnue (129) devient

$$E(y) = \frac{J(xy)}{\alpha J(y) - \frac{1-xy}{1-q}J(xy)} \quad (131)$$

où α , supposé indépendant de y , est un paramètre encore à déterminer. On obtient après quelques simplifications l'équation q -linéaire

$$t\alpha^2 J(y) - \left(1 + \frac{(t-q)(1-xy)}{1-q}\right) \alpha J(qy) + J(q^2y) = 0. \quad (132)$$

En $y = 0$ cette équation q -linéaire donne la condition sur α

$$t\alpha^2 - \left(1 + \frac{q-t}{1-q}\right) \alpha + 1 = 0 \quad (133)$$

que l'on identifie comme étant le noyau (54). Comme le montre l'étude de ce polynôme déjà présente dans la preuve de la proposition 3.4, les deux valeurs possibles pour α sont σ ou $1/(t\sigma)$. Pour retrouver le cadre de la proposition 3.11 on pose $\alpha = 1/t\sigma$, le changement de fonction inconnue est exactement celui que l'on a obtenu à partir de l'équation fonctionnelle à la Temperley. Comme l'équation q -linéaire est de degré 1 en y et satisfait par construction l'hypothèse de la proposition de Dutour on en déduit la valeur de $J(y)$. \square

3.5 Interprétations combinatoires

La valuation $\frac{q(1-q^{k+1}y)}{t(1-q)}$ des petits creux à la hauteur k est la somme de deux termes de nature différente. Le premier, $\frac{q}{t(1-q)}$ est un terme constant alors que $-\frac{q^{k+2}y}{t(1-q)}$ dépend géométriquement de la hauteur k . On se propose de séparer ces deux termes en codant la valuation des petits creux par la somme des valuations de deux carrés blanc et noir tels que

$$\square = \frac{q}{t(1-q)} \quad (134)$$

et

$$\blacksquare = -\frac{q^{k+2}y}{t(1-q)}. \quad (135)$$

Par construction la série génératrice de ces chemins à petits creux contenant un carré blanc et un carré noir est $E(y)$.

Les *chemins à petits creux bicolores* sont les chemins à petits creux dont chaque creux contient un carré blanc ou un carré noir.

Lemme 3.11.1 *La série génératrice des chemins à petits creux bicolores est aussi la série $E(y)$.*

Preuve. Par définition on a

$$E(y) = \sum_{d \in \mathcal{D}_{pc}} t^{\text{demi-longueur}(d)} \prod_{c \text{ creux de } d} (\square_c + \blacksquare_c)$$

où \mathcal{D}_{pc} est l'ensemble des chemins de Dyck à petits creux. On développe les sommes des creux, en choisissant un carré dans chaque creux, on obtient

$$E(y) = \sum_{d \in \mathcal{D}_{pc}} t^{\text{demi-longueur}(d)} \left(\prod_{c \text{ creux contenant } \square_c} \square_c \right) \left(\prod_{c \text{ creux contenant } \blacksquare_c} \blacksquare_c \right).$$

□

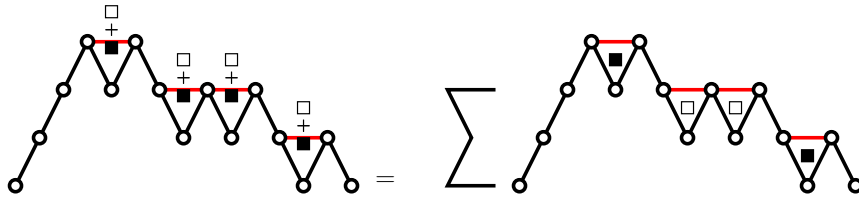


FIG. 3.16 – Développement des sommes $\square + \blacksquare$

L'ensemble des chemins à petits creux bicolores contient deux sous-ensembles remarquables : *les chemins à petits creux blancs*, respectivement *noirs*, dont tous les petits creux contiennent des carrés blancs, respectivement noirs. Avant d'envisager le traitement des chemins bicolores, on considère ces deux sous-classes combinatoires. Il s'avère que les chemins à petits creux blancs permettent de donner une interprétation combinatoire pour la racine du noyau σ . D'autre part, à un changement de variable mineur près, les chemins à petits creux noirs correspondent à un modèle de chemins de Dyck dont tous les creux sont sur l'axe. Ce modèle a été traité par Janse Van Rensburg [JvR00a] en utilisant une équation de décomposition q -algébrique. On présente une nouvelle résolution utilisant un modèle d'empilements. Divers éléments des preuves précédentes possèdent des interprétations combinatoires lorsque l'on restreint le modèle étudié. On tente de combiner ces interprétations dans l'espoir de mieux comprendre l'énumération des chemins à petits creux bicolores.

3.5.1 Interprétation de σ et lien avec les interactions inférieures

On considère les chemins à petits creux blancs.

Proposition 3.12 *La série génératrice $E_{\square}(t, u)$ des chemins non vides à petits creux blancs, valués $u/(1 - ut)$, est*

$$E_{\square}(t, u) = \frac{(1 - ut) \left(1 - t - 2ut - \sqrt{(1 - t)(4u^2t^2 - 4ut - t + 1)} \right)}{2u^2t^2}. \quad (136)$$

Preuve. On note toujours $q = ut$. En utilisant la décomposition des chemins à petits creux de la proposition 3.7 et en valuant les creux par $\frac{q}{t(1-q)}$ on obtient l'équation algébrique

$$E_{\square}(t, u) = t + tE_{\square}(t, u) + \frac{tq}{1-q}E_{\square}(t, u) + \frac{q}{1-q}E_{\square}(t, u) + \frac{q^2}{(1-q)^2}E_{\square}(t, u)^2. \quad (137)$$

La résolution de cette équation et le choix de l'unique série formelle en t solution conduit à l'expression de la proposition. \square

Le même radical $\sqrt{(1-t)(4u^2t^2 - 4ut - t + 1)}$ apparaît dans les expressions (136) de $E_{\square}(t, u)$ et (66) de $\sigma(t, u)$. Pour préciser la relation entre les chemins à petits creux blancs et la racine du noyau on va considérer un sous-ensemble des chemins. Un *pic double* est une occurrence du facteur $xx\bar{x}$.

Proposition 3.13 *La série génératrice des chemins à petits creux blancs débutant par un pic double est $t^2\sigma$ où σ est la racine du noyau.*

Preuve. La figure 3.17 illustre la décomposition des chemins à petits creux débutant par un pic double. La contrainte des petits creux impose qu'un facteur $x\bar{x}$ suive le pic double.

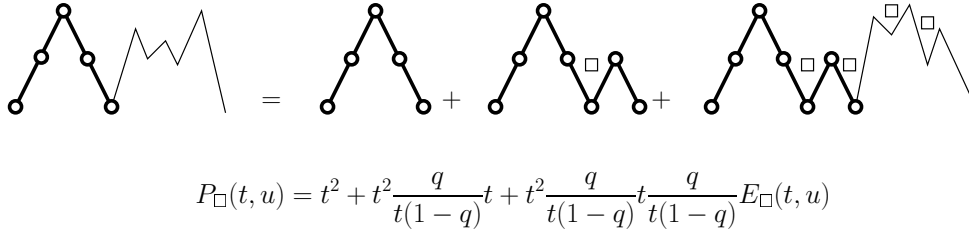


FIG. 3.17 – Chemins à petits creux débutant par un pic double et σ

Après ce facteur il y a un petit creux et un chemin à petits creux quelconque. La série génératrice $P_{\square}(u, t)$ des chemins à petits creux blancs débutant par un pic double vaut donc

$$P_{\square}(t, u) = t^2 + t^2 \frac{q}{t(1-q)} t + t^2 \frac{q}{t(1-q)} t \frac{q}{t(1-q)} E_{\square}(t, u). \quad (138)$$

On en déduit la forme close de $P_{\square}(t, u)$ et on constate qu'elle est égale à $t^2\sigma$:

$$P_{\square}(t, u) = \frac{1 + t - 2ut - \sqrt{(1-t)(4u^2t^2 - 4ut - t + 1)}}{2u^2(1-ut)} = t^2\sigma(t, u). \quad (139)$$

\square

Pour faire apparaître un peu plus naturellement le sous-ensemble des chemins à petits creux débutant par un pic double, on va utiliser une autre interprétation combinatoire de σ : les chemins de Dyck dont on compte les interactions *inférieures*. En effet l'observation

des deux formes closes pour σ , équation (66), et $1 + D(t, v = u)$, équation (50), permet de conclure à l'égalité de ces deux séries. Dans la preuve de la proposition suivante on donne une décomposition récursive des chemins de Dyck dont la traduction en équation donne une des formes du noyau.

Proposition 3.14 *La série génératrice $F(t, u)$ des chemins de Dyck dont on compte la demi-longueur par t et les interactions inférieures par u est, à une constante près, la racine du noyau σ :*

$$\sigma = 1 + F(t, u). \quad (140)$$

Preuve.(Proposition 3.14) Cette preuve consiste à récrire l'équation définissant σ de façon à reconnaître l'équation de décomposition satisfaite par les chemins de Dyck. L'équation du noyau est

$$1 = \frac{q}{1-q} \left(\frac{ts}{1-ts} - \frac{s}{1-s} \right). \quad (141)$$

L'égalité $\sigma = 1 + F(t, v = u)$ se conjecture à la vue des formes closes. Elle suggère le changement de variable $s = 1 + X$ dans l'équation menant au noyau :

$$1 = \frac{q}{1-q} \left(\frac{t(1+X)}{1-t(1+X)} + \frac{1+X}{X} \right) \frac{1}{X}. \quad (142)$$

La réduction du second membre au même dénominateur donne

$$1 = \frac{q}{1-q} \frac{1+X-t-tX}{X(1-t-tX)} = \frac{q}{1-q} \left(\frac{X}{1-t-Xt} + 1 \right). \quad (143)$$

Une multiplication par X permet de conclure que X satisfait

$$X = \frac{q}{1-q} + \frac{q}{1-q} \frac{X}{1-t(1+X)}. \quad (144)$$

Cette équation est exactement celle que satisfait $D(t, v = u)$ comme énoncé dans la proposition 3.1. \square

On est maintenant armé pour donner une seconde preuve plus combinatoire, de la proposition 3.13 :

Preuve.(Proposition 3.13)

D'après la proposition 3.14, la série σ est la série génératrice des chemins de Dyck éventuellement vides énumérés selon leur demi-longueur et le nombre d'interactions inférieures. On propose une suite de bijections et de sommations combinatoires qui prouvent que σ est aussi, à un facteur t^2 près, la série génératrice des chemins à petits creux blancs débutant par un pic double.

• Un *chemin à petits pics étiquetés* est un chemin de Dyck évitant le facteur $xx\bar{x}$ dont tous les pics sont étiquetés par un entier. La valuation d'un tel chemin est $t^i u^j (ut)^k$ où

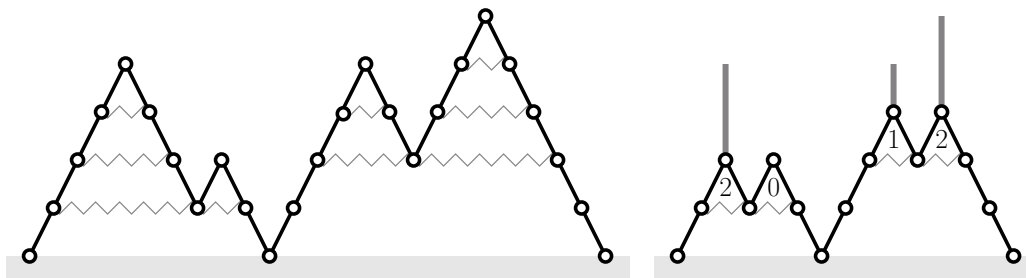


FIG. 3.18 – Un chemin de Dyck et son chemin à petits pics

i est sa demi-longueur, j son nombre de pics et k la somme des étiquettes sur les pics. Ces chemins sont en bijection avec les chemins de Dyck. Un chemin de Dyck w_1 peut se transformer en un chemin à petits pic w_2 en réécrivant récursivement les facteurs $xx\bar{x}$ en $x\bar{x}$. L'étiquette sur un pic de w_2 est le nombre de réécritures de ce pic. La figure 3.18 donne un exemple. Chaque réécriture supprime un pas montant et une interaction inférieure du chemin tout en augmentant de un la somme des étiquettes donc cette bijection préserve la valuation.

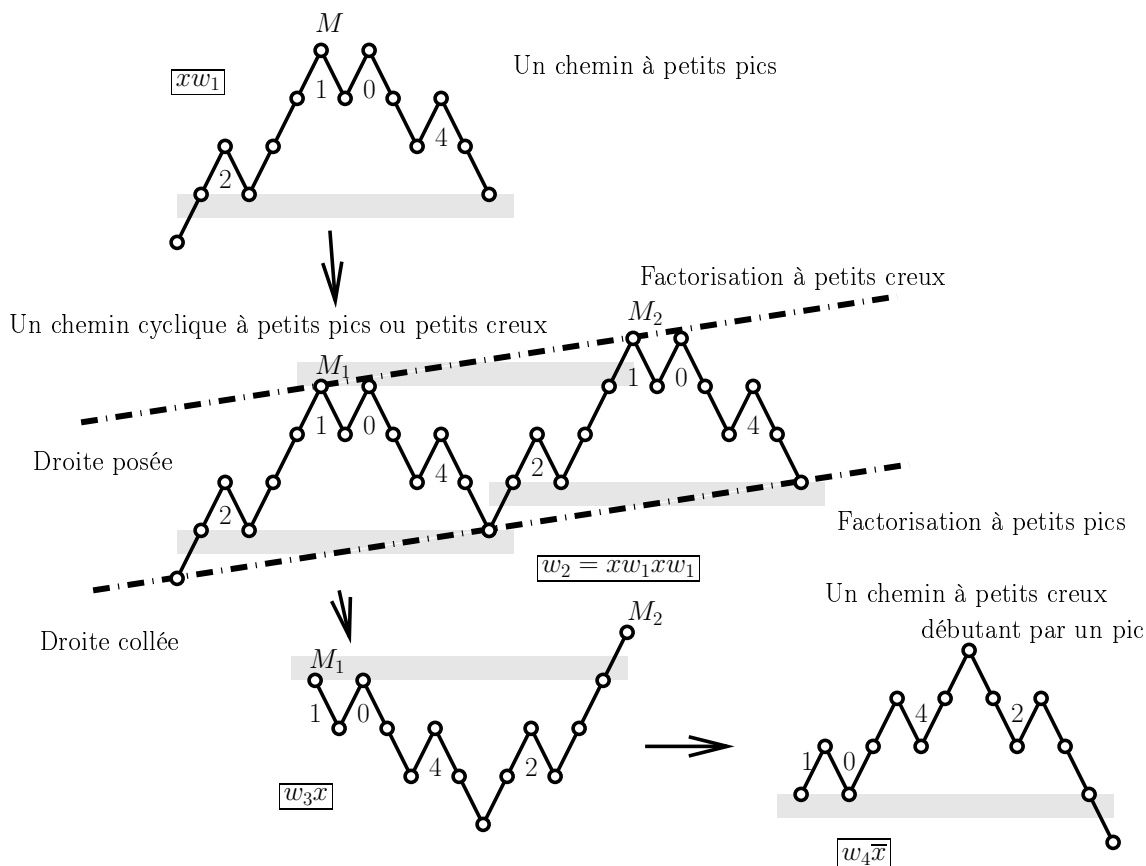


FIG. 3.19 – La bijection liée au lemme cyclique pour les petits creux ou pics

• Un *chemin à petits creux étiquetés débutant par pic* est un chemin de Dyck à petits creux, éventuellement vide, dont les creux et le premier sommet (si le chemin n'est pas vide), sont étiquetés par un entier. La valuation d'un tel chemin est $t^i u^{j+1} (ut)^k$ où i est sa demi-longueur, j son nombre de creux et k la somme des étiquettes sur les creux et le premier sommet. Il existe une bijection entre les chemins à petits pics étiquetés et les chemins à petits creux étiquetés débutant par un pic qui préserve les valuations. Un exemple de cette bijection est proposé sur la figure 3.19. Le cas général de cette bijection traite d'un chemin à petits pics étiquetés w_1 dont la demi-longueur n est au moins 2. Le chemin $w_2 = xw_1xw_1$ est aussi un chemin à petits pics (mais pas de Dyck) car comme w_1 est un chemin de Dyck, l'ajout de x à w_1 ne peut créer un pic double et la concaténation des deux facteurs xw_1 créer un creux. À partir de w_2 on peut retrouver la factorisation xw_1xw_1 en collant une droite de pente $1/(2n)$ sous le chemin w_2 . Entre les deux premiers contacts du chemin avec la droite on retrouve le premier facteur xw_1 . On peut également poser une droite de pente $1/(2n)$ au dessus du chemin w_2 et on observe qu'il y a deux contacts M_1 et M_2 qui correspondent dans chaque facteur xw_1 à la première fois que le chemin w_1 atteint son ordonnée maximale. Le sommet M_2 est donc précédé d'un pas montant. On considère alors le facteur w_3x du chemin entre les sommets M_1 et M_2 , sans prendre en compte l'étiquette sur le sommet M_2 . On pose w_4 l'image de w_3 par le morphisme défini $x \rightarrow \bar{x}$ et $\bar{x} \rightarrow x$, c'est à dire la réflexion par rapport à un axe horizontal. Le chemin w_4 est un chemin à petits creux et premier sommet étiquetés, de demi-longueur n . De plus ce chemin débute par un pic car sinon M_1 est précédé d'un facteur $x\bar{x}$ et cela contredit le fait que M_1 soit le premier contact de la droite posée sur w_2 . Le chemin w_2 contient deux exemplaires de chacun des sommets ou pics valués de w_1 et w_4 contient la moitié de ces exemplaires dont la valuation de w_1 est aussi celle de w_4 . Les deux cas dégénérés cette bijection sont illustrés sur la figure 3.20.

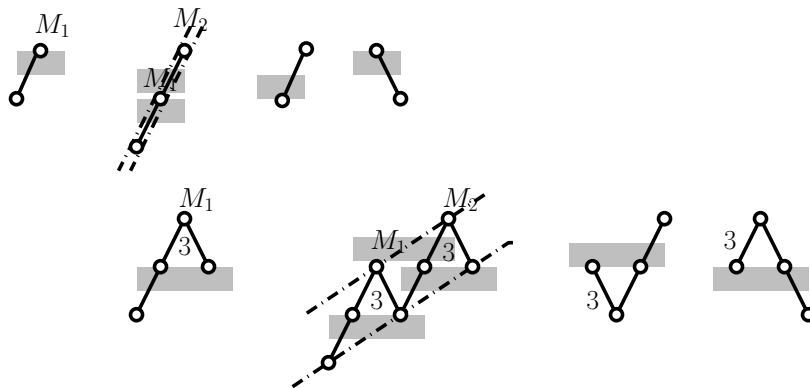


FIG. 3.20 – Les cas dégénérés de la bijection précédente

• Les *chemins à petits creux étiquetés débutant par un pic double* sont en bijection avec les chemins à petits creux étiquetés débutant par un pic. On ajoute au second type de chemin un pic double au début ce qui crée un creux où se place l'étiquette du premier sommet (si le chemin est non vide). La série génératrice des chemins à petits creux étiquetés débutant

par un pic double est donc $t^2\sigma$.

• Soit w un chemin à petits creux non-étiquetés débutant par un pic double. La série génératrice des chemins (w_i) à petits creux étiquetés dont la suppression des étiquettes conduit au chemin w est la valuation du chemin w dont les creux sont blancs. En effet la sommation des valuations des étiquetages possibles d'un creux donne $\sum_{k \geq 0} (ut)^k = 1/(1-ut)$ auquel s'ajoute le poids u de l'interaction supérieure du petit creux ; on retrouve le poids $\square = \frac{q}{t(1-q)}$. La série génératrice des chemins à petits creux blancs débutant par un pic double est donc $t^2\sigma$.

□

Cette interprétation de σ est aussi l'ensemble des facteurs qui apparaissent dans un chemin de Dyck après une double descente $\bar{x}\bar{x}$. De ce fait σ apparaît naturellement dans des factorisations usuelles des chemins de Dyck. Un premier exemple de ces factorisations est donnée par un second calcul de la série génératrice des chemins à petits creux blancs.

Proposition 3.15 *La série génératrice $E_{\square}(t, q)$ des chemins à petits creux blancs est*

$$E_{\square}(t, q) = \frac{t(1-q)(\sigma-1)}{q(1-t\sigma)}. \tag{145}$$

Preuve. Cette preuve consiste à donner une interprétation combinatoire à l'égalité

$$\frac{q}{t(1-q)} E_{\square}(t, q) = \frac{\sigma-1}{1-t\sigma}. \tag{146}$$

Étant donné un chemin w à petit creux blancs non-vide, on place la valuation \square supplémentaire sur le sommet précédent le premier pic du chemin. On factorise le chemin sous la forme $w = x^k \bar{x} \bar{x} u \bar{x} v_1 \bar{x} v_2 \bar{x} \dots \bar{x} v_k$ où u et les $(v_i)_{1 \leq k}$ sont des mots de Dyck à petits creux. La figure 3.21 illustre cette factorisation.

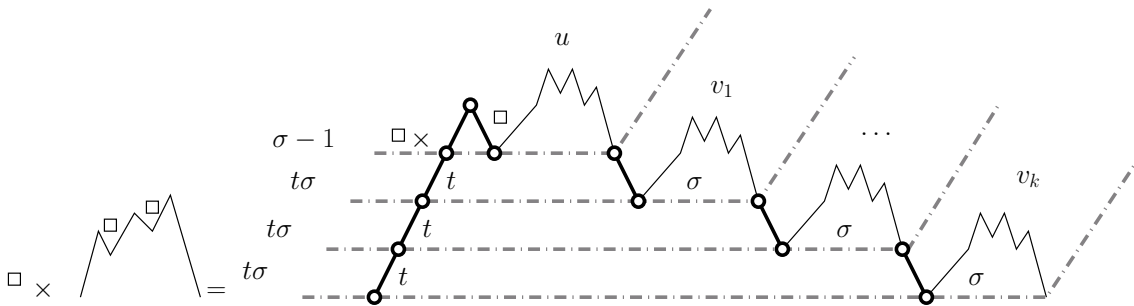


FIG. 3.21 – La factorisation selon σ des chemins à petits creux blancs

Comme chaque v_i est précédé d'une double descente, ce chemin doit débuter par un petit pic s'il n'est pas vide. Les v_i sont donc en bijection avec les chemins à petits creux blancs

débutant par un pic double (en supprimant le pic double). De plus cette bijection préserve les valuations en t et u . En comptant le pas montant x responsable de l'apparition de ce facteur on a une valuation $t\sigma$. Au niveau du premier pic on a un facteur $x\bar{x}u$, c'est donc un chemin à petits creux non vide débutant par un pic. Les chemins w sont en bijection avec les chemins à petits creux blancs débutant par un pic double de demi-longueur au moins 3 (toujours en supprimant le pic double). L'ajout de la valuation \square au premier sommet permet de garantir que cette bijection préserve aussi les valuations. La valuation de ce facteur est donc $\sigma - 1$, car on a soustrait la valuation du chemin vide.

□

3.5.2 Chemins à creux sur l'axe et empilements de segments

Au lieu de traiter directement des chemins à petits creux noirs on va s'intéresser aux chemins à creux sur l'axe. D'une part les séries génératrices de ces deux modèles sont liées par un changement de variable simple (proposition 3.16). D'autre part, Janse Van Rensburg [JvR00a] a étudié ces chemins à creux sur l'axe en donnant une énumération basée sur la résolution d'une équation q -algébrique, la méthode employée dans la partie 3.3.3. On propose ici d'énumérer cette classe combinatoire en construisant une bijection avec des demi-pyramides de segments (proposition 3.17). La difficulté principale est à nouveau de résoudre une équation q -linéaire. Dans ce contexte elle définit la série génératrice alternée des empilements triviaux ; un "miracle calculatoire" permet de la calculer (proposition 3.19). La série génératrice des chemins à petits creux noirs s'obtient alors comme une instance du lemme d'inversion pour les demi-pyramides (corollaire 3.19.1). Un autre intérêt de ce lien avec les empilements de segments est de fournir un exemple d'interprétation combinatoire d'un changement de fonction inconnue permettant de linéariser une équation q -algébrique.

Les *chemins à creux sur l'axe* sont les chemins de Dyck dont tous les creux sont sur l'axe. Comme les chemins à petits creux, ces chemins sont définis par la suite des hauteurs de leurs pics et cette suite donne une bijection naturelle entre les deux types de chemins.

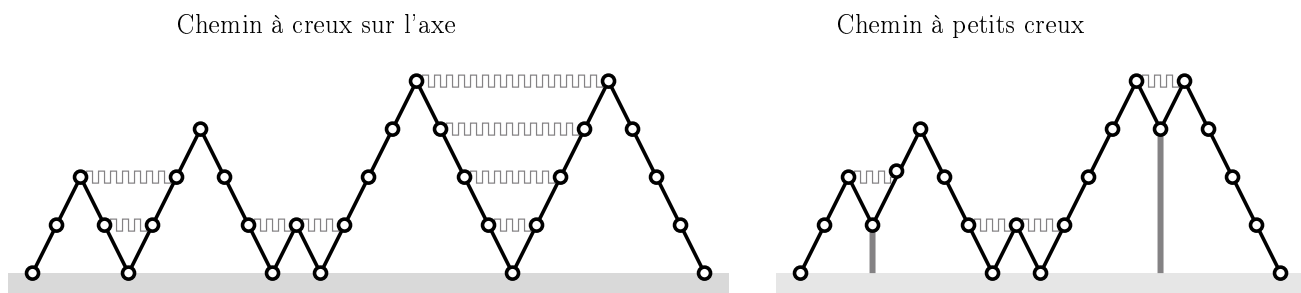


FIG. 3.22 – Chemin à creux sur l'axe et chemin à petits creux

La figure 3.22 donne les chemins des deux types dont la suite des hauteurs des pics est

2, 3, 1, 4, 4. Pour la classe des chemins à creux sur l'axe, le comptage des interactions supérieures est plus simple que pour les chemins de Dyck. En effet entre chaque paire de sommets en interaction supérieure on contraint le chemin à repasser par l'axe. À un changement de variable près ce modèle de chemin est lié aux chemins à petits creux noirs.

Proposition 3.16 *La série génératrice $C(t, u; y)$ des chemins à creux sur l'axe où t compte la demi-longueur, u les interactions supérieures et y les creux, est*

$$C(t, u; y) = E_{\blacksquare}(t, u; -(1-q)y/q) \quad (147)$$

où $E_{\blacksquare}(t, u; y)$ est la série génératrice des chemins à petits creux noirs.

Corollaire 3.16.1 *La série $C(y) \equiv C(t, u; y)$ satisfait l'équation q -algébrique*

$$C(y) = t + (t + tqy)C(qy) + qyC(y) + q^2y^2C(qy)C(y). \quad (148)$$

Preuve. Un creux sur l'axe surmonté de k interactions supérieures est associé à un petit creux à la hauteur $k - 1$ par la bijection préservant la suite des hauteurs de pics. La série génératrice $C(t, u; y)$ est donc aussi celle des chemins à petits creux de hauteur h valué par yq^{h+1}/t . La valuation d'un petit creux noir de hauteur h dans $E_{\blacksquare}(t, u, y)$ est $\blacksquare = -y \frac{q^{h+1}}{t(1-q)}$. Dans cette valuation, le changement de variable $y = -(1-q)y/q$ conduit bien à la valuation précédente.

Pour le corollaire, on peut reprendre la décomposition des chemins à petits creux de la proposition 3.7. \square

Il existe déjà des bijections entre les chemins de Dyck et des empilements de pièces. Viennot a donné une première bijection avec des empilements de dominos. Ensuite Bousquet-Mélou et Viennot [BMV92] ont proposé de "coller" certains de ces dominos pour former des segments et énumérer les chemins de Dyck selon, par exemple, la différence entre la somme des hauteurs des pics et celle des hauteurs des creux. On propose ici une nouvelle variante de ces modèles dans le but d'énumérer grosso-modo la somme des hauteurs des creux des chemins de Dyck à *petits creux*. C'est cette contrainte de petits creux qui est la principale source de nouveauté.

On commence par présenter les segments marqués qui seront les pièces de nos empilements. La figure 3.23 contient quelques exemples de segments marqués. Un *segment marqué* est un couple $([i, j], m)$, où $[i, j]$ est un segment de \mathbb{N}^* et m , le *marquage du segment marqué*, est un sous-ensemble de $]i, j[$. Si $[i, j]$ n'est pas un singleton alors i est la *base* (du segment marqué), j est l'*extrémité* (du segment marqué), les k appartenant à m sont les *éléments lourds* (du segment marqué) et les autres éléments de $]i, j[$ sont les *éléments légers*. Pour les singletons $[i, i]$ on parle de l'*élément* (du segment marqué). Pour ces singletons ainsi que les doublons on remarque que le marquage m est toujours vide. Deux pièces $([i_1, j_1], m_1)$ et

$([i_2, j_2], m_2)$ sont en concurrence si l'intersection $[i_1, j_1] \cap [i_2, j_2]$ est non-vide. La valuation d'un segment marqué est le produit des valuations des éléments de ce segment. On value chaque élément $h \in [i, j]$ selon sa nature et sa valeur en tant qu'entier. La valuation de l'élément h est

- $q^h y$ s'il est l'élément d'un singleton,
- $q^{2h} y^2$ s'il est la base d'un segment,
- t s'il est l'extrémité d'un segment ou un élément léger,
- $tq^h y$ s'il est un élément lourd.

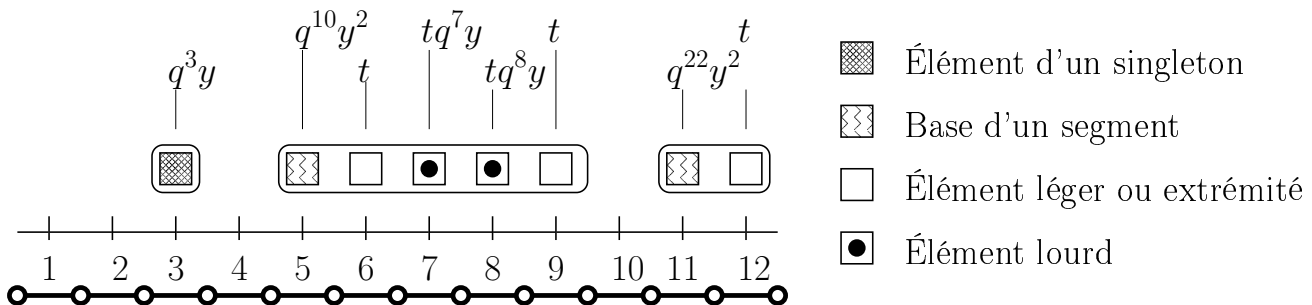


FIG. 3.23 – La valuation des segments marqués

Une demi-pyramide de segments est un empilement dont l'unique pièce minimale est un segment marqué de la forme $([1, j], m)$. Par convention l'empilement vide est aussi une demi-pyramide. On note $P(t, u; y)$ la série génératrice des demi-pyramides de segments marqués valués comme ci-dessus.

Proposition 3.17 *Soit $G(t, u; y)$ la série génératrice des chemins à creux sur l'axe débutant par un pic double et dont le second pic, s'il existe, est à la hauteur 1, où t compte la demi-longueur, u le nombre d'interactions supérieures et y le nombre de creux.*

Il existe une bijection entre ces chemins et les demi-pyramides de segments marqués qui, à un facteur près, préserve les valuations.

Plus précisément, soit $P(t, u; y)$ la série génératrice des demi-pyramides. La valuation des demi-pyramides a été choisie de sorte que

$$G(t, u; y) = t^2 P(t, u; y). \quad (149)$$

Preuve. On décrit la bijection en construisant un empilement à partir d'un chemin. Cette bijection comporte deux étapes, la première consiste à transformer un chemin à creux sur l'axe en chemin à petits creux en récrivant récursivement $\bar{x}x x x$ en $\bar{x}x$. Pour préserver la valuation des chemins, les petits creux à la hauteur h sont valués par yq^{h+1}/t pour le chemin à petits creux. On détaille maintenant la seconde étape de la bijection qui transforme ces

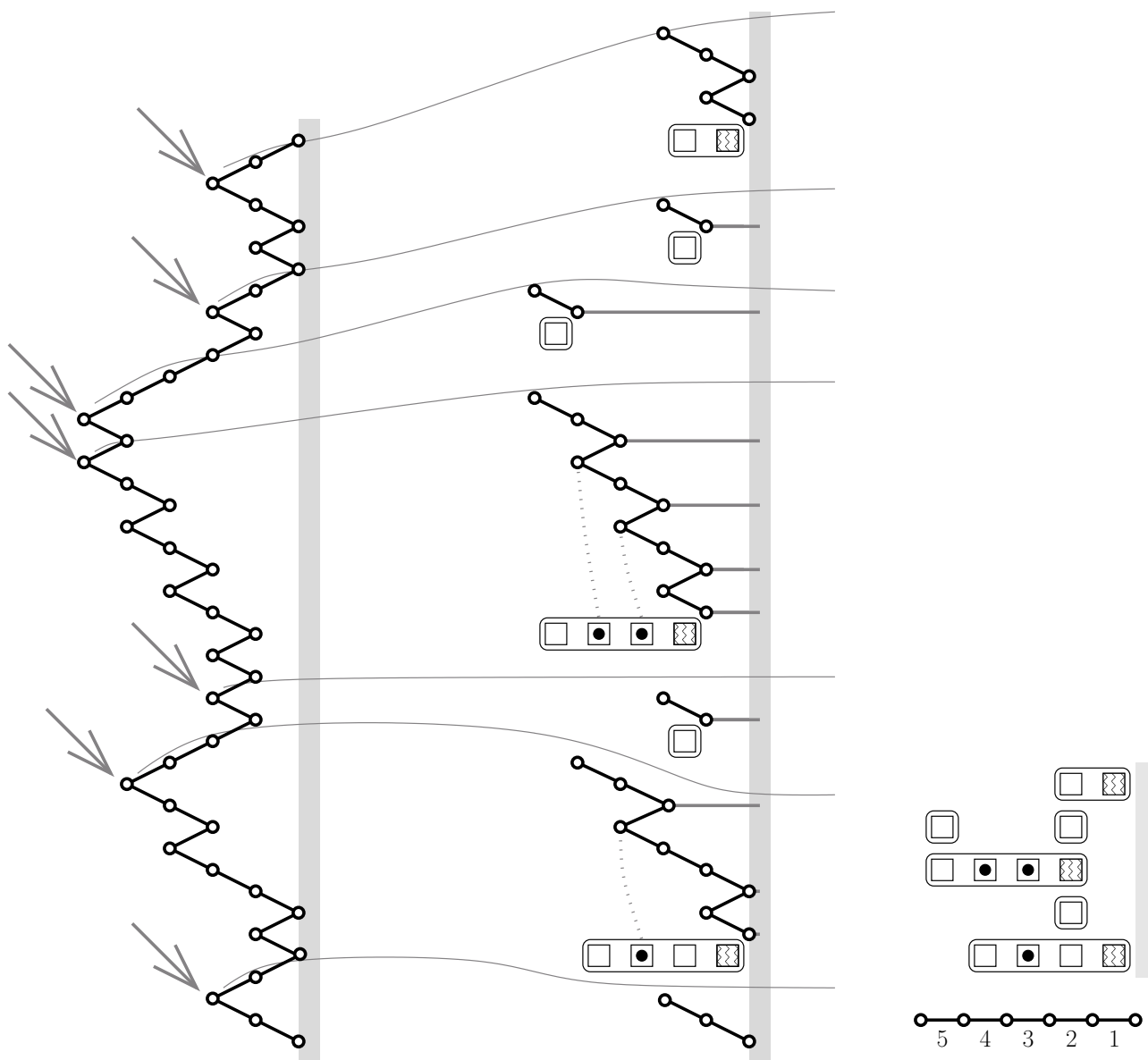


FIG. 3.24 – Bijection entre des chemins à petits creux débutant par un pic double et les empilements de segments marqués

chemins à petits creux en demi-pyramides de segments marqués. La figure 3.24 donne un exemple de cette seconde étape.

Un *pic de factorisation* dans un chemin de Dyck à petit creux est un pic qui n'est pas suivi d'une double montée, ce n'est pas le pic d'un facteur $x\bar{x}xx$. Un *bloc (de pas montants)* dans un chemin de Dyck est un ensemble des pas montants situés entre deux pics de factorisation successifs. On distribue la valuation d'un chemin sur les blocs. La valuation d'un bloc est le produit de celles des pas montants et des creux entre les deux pics de factorisation qui

définissent le bloc. L'hypothèse d'un pic double initial garantit que tout creux est entouré de deux pics de factorisation et que seuls deux pas montants (ceux formant le pic double initial) ne sont pas entre deux pics de factorisation. Le produit des valuations des blocs est donc la valuation du chemin à un facteur t^2 près.

L'ensemble des blocs est en bijection préservant la valuation avec l'ensemble des segments marqués. Dans cette preuve la *hauteur d'un pas montant* est la hauteur du sommet après ce pas. Comme le pic d'un facteur $x\bar{x}$ est un pic de factorisation, l'ensemble des hauteurs des pas montants d'un bloc forme un segment $[i, j]$: c'est le segment à marquer associé au bloc. Pour définir le marquage de ce segment on considère le multi-ensemble des hauteurs des pas montants du bloc en discutant sur la nature du segment.

- Si $i = j$, $[i, j]$ est un singleton. Comme le premier pic d'un facteur $x\bar{x}x\bar{x}$ est un pic de factorisation, le bloc ne contient qu'un pas montant et un creux à la hauteur i pour une valuation totale de $q^i y$ qui correspond bien à celle d'un singleton $([i, i], \emptyset)$.
- Sinon on observe que comme le premier pic d'un facteur $x\bar{x}x\bar{x}$ ou le pic d'un facteur $x\bar{x}$ est un pic de factorisation il y a au plus deux pas de chaque hauteur. Le premier facteur permet aussi de constater qu'il y a un seul pas de hauteur j . C'est en accord avec la valuation t de l'extrémité d'un segment marqué. Il y a deux pas de hauteurs i car après le premier pas d'un bloc est le premier d'un facteur $\bar{x}x\bar{x}xx$. De plus ce facteur forme deux creux à la hauteur $i - 1$ pour une valuation totale de $t^2 q^{2i} y^2$ qui correspond à la valuation de la base d'un segment. Pour les hauteurs de l'intervalle $[i + 1, j - 1]$ il y a soit un pas montant dans un facteur xxx soit deux pas dans un facteur $xx\bar{x}xx$. On définit le marquage du segment $[i, j]$ comme l'ensemble des hauteurs de l'intervalle $]i, j[$ qui apparaissent exactement deux fois dans le multi-ensemble des hauteurs des pas montants d'un bloc. Un élément k marqué est donc associé à un facteur $xx\bar{x}xx$ qui contient deux pas montants à la hauteur k et un creux à la hauteur $k - 1$ pour une valuation $q^k t y$ qui est celle d'un élément marqué. Un élément k non marqué est associé à un facteur xxx pour une valuation t .

La fin de cette correspondance est similaire à celle des travaux de Bousquet-Mélou et Viennot sur les empilements de segments. Pour obtenir l'empilement de segment associé à un chemin on fait tomber les segments marqués dans l'ordre d'apparition des blocs de pas montants. Comme le chemin débute par un pic double le premier pas montant du premier bloc est de hauteur 1, donc cet empilement admet une pièce minimale de segment $[1, j]$. Comme toutes les autres pièces sont des segments dont le premier élément est au plus égal à l'extrémité du segment précédent, on observe que la première pièce est la seule à être minimale. L'empilement obtenu par cette construction est donc bien une demi-pyramide de segments marqués.

Pour montrer que l'on a bien une bijection, on résume les points clefs de l'application inverse. Étant donné une demi-pyramide on ordonne ses segments marqués en retirant récursivement le segment minimal contenant les plus grands entiers. On construit le chemin en partant d'un pic double $x\bar{x}x\bar{x}$ et on ajoute des blocs de pas montants en considérant les segments marqués dans l'ordre. On place le facteur x ou $x\bar{x}$ associé à l'élément e d'un

segment à la hauteur e en lisant le segment dans l'ordre croissant de ces éléments. On complète les trous par des pas descendants. Un élément base d'un segment ou un élément lourd code un facteur $x\bar{x}$. Un singleton, une extrémité de segment ou un élément léger donne un facteur x .

□

Dans un modèle d'empilement, un empilement trivial est un empilement où il n'y a aucune concurrence entre les pièces. L'empilement de la figure 3.23 est un exemple d'empilement trivial. La *série génératrice alternée de l'ensemble $Triv$ des empilements triviaux de segments marqués* est la série

$$T(t, u; y) = \sum_{e \in Triv} (-1)^{|e|} v(e) \quad (150)$$

où $v(e)$ est la valuation de l'empilement e et $|e|$ est le nombre de pièces dans l'empilement e . À l'aide de cette série on peut exprimer simplement la série génératrice des demi-pyramides en appliquant le lemme d'inversion de Viennot [Vie86], que l'on a énoncé dans la proposition 1.1.

Proposition 3.18 *La série génératrice $P(t, u; y)$ des demi-pyramides marquées est*

$$P(t, u; y) = \frac{T(t, u; qy)}{T(t, u; y)} \quad (151)$$

où $T(t, u; y)$ est la série alternée des empilements triviaux de segments marqués.

Preuve. Avant de s'intéresser au modèle des empilements de segments marqués on rappelle le lemme d'inversion valable pour des empilements quelconques.

On veut calculer la série génératrice $Q(t) \in \mathbb{A}[[t]]$ des empilements d'un ensemble de pièces \mathcal{P} dont les pièces minimales sont dans un ensemble de pièces \mathcal{Q} . On introduit la série génératrice alternée $T_{\mathcal{P}}(t)$ des empilements triviaux et la série génératrice alternée $T_{\mathcal{P}-\mathcal{Q}}(t)$ des empilements triviaux formés de pièces de $\mathcal{P} - \mathcal{Q}$. D'après le lemme d'inversion de Viennot, la série génératrice $Q(t)$ s'exprime par

$$Q(t) = \frac{T_{\mathcal{P}-\mathcal{Q}}(t)}{T_{\mathcal{P}}(t)}. \quad (152)$$

Dans le cas des demi-pyramides de segments on peut parfois améliorer cette expression. Pour les demi-pyramides de segments marqués l'ensemble \mathcal{Q} est l'ensemble des segments marqués de la forme $([1, k], m)$. L'ensemble des empilements triviaux impliqués dans le numérateur de (152) est donc celui des empilements dont aucun segment ne contient l'élément 1. Cet ensemble est en bijection avec l'ensemble des empilements triviaux par une simple translation de 1 de tous les segments et marquages. Il reste à étudier le comportement des valuations lors de cette transformation. Or ici on remarque que la valuation du segment

marqué translaté de un, $([i + 1, j + 1], m + 1)$, est liée à celle du segment $([i, j], m)$ par la substitution $y \rightarrow qy$. Autrement dit le décalage de un d'un segment modifie la valuation du segment par la substitution de y par qy . Si on note $T(y)$ la série génératrice alternée des empilements de segments triviaux alors la série génératrice alternée des empilements triviaux de segments ne contenant pas 1 est $T(qy)$. L'application du lemme d'inversion dans ce cas particulier donne donc la formule (151).

□

La principale difficulté est le calcul de la série alternée des empilements triviaux. En général on ne sait pas la calculer. Lorsqu'on travaille avec des empilements de segments, il est en général possible de montrer qu'elle satisfait une équation q -linéaire. Le noeud du problème est, à nouveau dans ce chapitre, de résoudre cette équation q -linéaire. Dans le cas étudié c'est particulièrement simple car après une simplification, l'équation à une forme adaptée à sa résolution.

Proposition 3.19 *La série génératrice alternée $T(t, u; y) \equiv T(y)$ des empilements triviaux de segments marqués satisfait l'équation q -linéaire*

$$T(y) = \left(1 - qy + t/q^2 + ty\right)T(qy) - t/q^2T(q^2y). \quad (153)$$

Elle vaut donc

$$T(y) = \sum_{n \geq 0} \frac{(t - q)^n q^{\binom{n}{2}}}{(q)_n (t/q)_n} y^n. \quad (154)$$

Corollaire 3.19.1 *La série génératrice $P(t, u; y)$ des demi-pyramides de segments marqués est*

$$P(t, u; y) = \frac{\sum_{n \geq 0} \frac{(t - q)^n q^{\binom{n+1}{2}}}{(q)_n (t/q)_n} y^n}{\sum_{n \geq 0} \frac{(t - q)^n q^{\binom{n}{2}}}{(q)_n (t/q)_n} y^n}. \quad (155)$$

Preuve. On commence par donner une équation de décomposition des empilements triviaux en observant le voisinage de 1. La plupart des cas sont classiques et se traitent en reconnaissant dans la factorisation un empilement trivial translaté. Cette décomposition est illustrée par la figure 3.25. Cependant le dernier cas de cette décomposition est un peu plus délicat et son traitement est illustré par la figure 3.26. Ensuite on constate que l'équation q -linéaire a des coefficients de degré au plus 1 en y ce qui permet de trouver directement une solution hypergéométrique de cette équation.

On ajoute un facteur -1 à la valuation de toutes les pièces pour calculer une série génératrice *alternée*. On décompose un empilement trivial en observant la nature des éléments 1 :

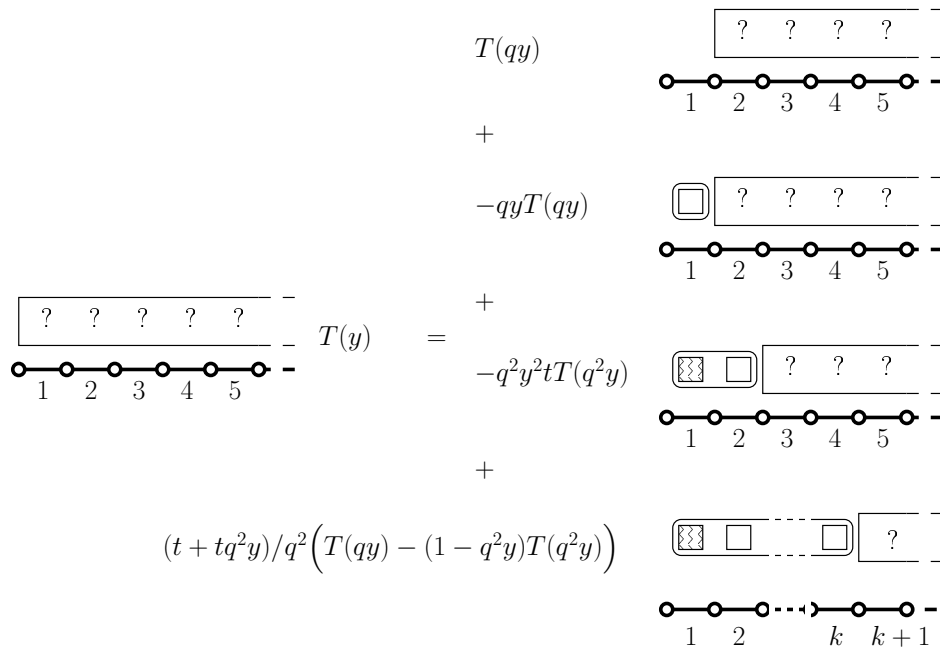


FIG. 3.25 – Décomposition des empilements triviaux de segments marqués

- Si aucun segment ne contient de 1 alors l’empilement s’obtient de façon unique en décalant de 1 les segments d’un empilement quelconque. La série génératrice alternée de ces empilements est donc $T(qy)$.
- Si l’empilement contient le singleton $[1, 1]$ alors il se décompose comme le produit d’un singleton $[1, 1]$ et d’un empilement dont aucun segment ne contient de 1. La série génératrice alternée de ces empilements est donc $-qyT(qy)$.
- Si l’empilement contient le segment $[1, 2]$ alors il se décompose comme le produit d’un segment $[1, 2]$ et d’un empilement dont aucun segment ne contient ni de 1 ni de 2. Ce dernier empilement s’obtient de façon unique en décalant cette fois ci de 2 les segments d’un empilement quelconque. La série génératrice alternée de ces empilements est donc $-q^2y^2tT(q^2y)$.
- Si l’empilement trivial commence par un segment de la forme $[1, k]$ avec $k \geq 3$ le calcul de la série génératrice alternée $T_{\geq 3}(y)$ est plus délicat. La figure 3.26 illustre le propos. On va construire une bijection entre ces empilements et les empilements qui débutent par un segment de la forme $[1, k]$ avec $k \geq 2$. La série génératrice alternée $T_{\geq 2}(y)$ de ces derniers empilements s’obtient par différence en remarquant que ce sont des empilements quelconques privés de ceux dont aucun segment ne contient de 1 et de ceux débutant par le singleton $[1, 1]$. En termes de série génératrice cela donne

$$T_{\geq 2}(y) = T(y) - T(qy) - (-qy)T(qy). \tag{156}$$

On décompose un empilement débutant par un segment de longueur au moins 3 et d’extrémité gauche 1 en un élément 2 qui est lourd ou léger et un empilement

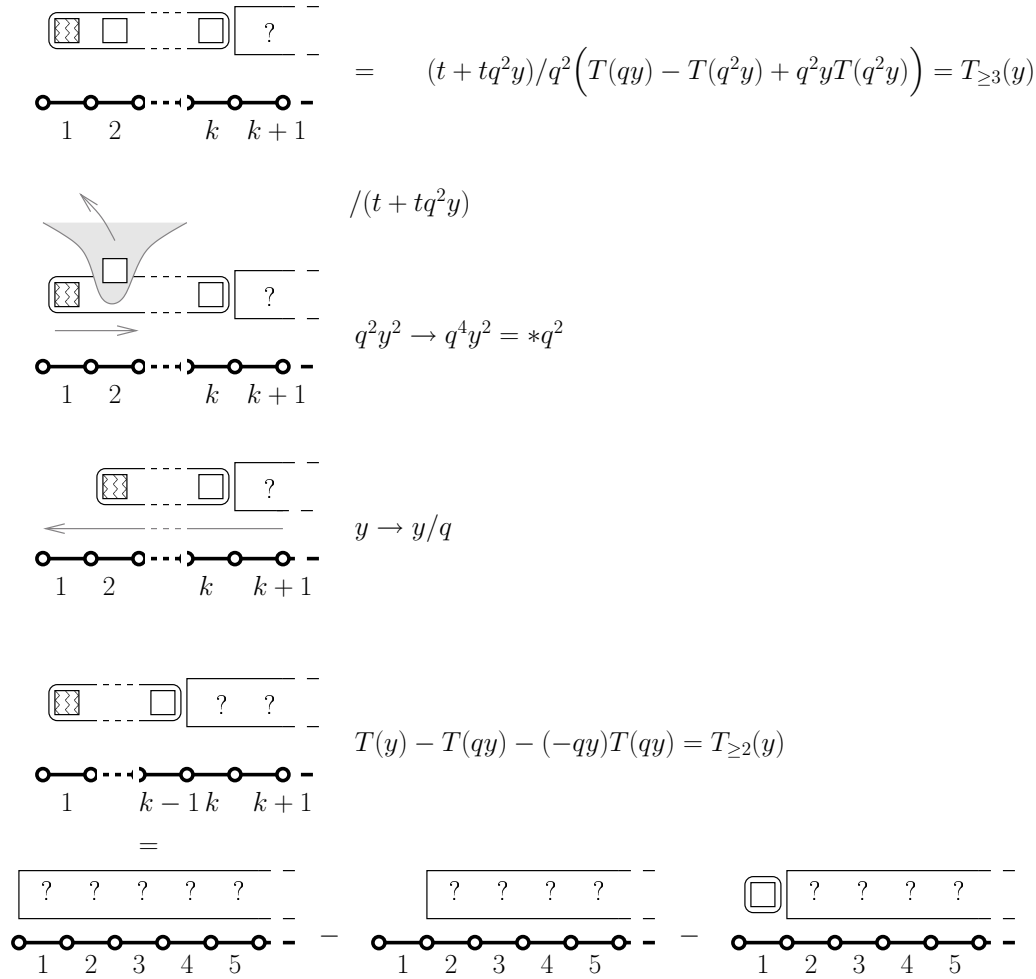


FIG. 3.26 – Décomposition d’un empilement trivial débutant par un segment d’au moins 3 éléments

débutant par un segment de longueur au moins 2 et d’extrémité gauche 1 de la façon suivante : on retire l’élément 2 lourd ou léger du segment initial, on translate de +1 uniquement la base de ce segment pour retrouver un segment de longueur au moins 2 et on translate cet empilement de -1 pour obtenir un empilement dont le segment initial contient 1. Pour calculer la série $T_{\geq 3}(y)$ on considère la réciproque de cette bijection. Partant de l’empilement de $T_{\geq 2}(y)$ on translate cet empilement de +1 ce qui se traduit par la substitution $y \rightarrow qy$. La translation de -1 de l’élément base du premier segment divise par q^2 la valuation de cette base. L’insertion de l’élément 2 rajoute un facteur t ou tq^2y à la valuation selon que l’élément est léger ou lourd. Au final on a la relation

$$T_{\geq 3}(y) = (t + tq^2y)/q^2 T_{\geq 2}(qy). \tag{157}$$

La décomposition précédente conduit à l'équation q -linéaire

$$T(y) = T(qy) - qyT(qy) - tq^2y^2T(q^2y) + (t + tq^2y)/q^2 (T(qy) - T(q^2y) + q^2yT(q^2y)). \quad (158)$$

Les coefficients de cette équation sont *a priori* de degré 2 en y mais une simplification se produit entre ces termes pour donner l'équation (153) dont les coefficients sont de degré 1 en y .

On résout cette équation. En isolant les termes de degré 0 ou 1 en y on a

$$yq^2(q - t)T(qy) = t(T(qy) - T(q^2y)) - q^2(T(y) - T(qy)). \quad (159)$$

Si on développe sous la forme $T(y) = \sum_{n \geq 0} t_n y^n$, que l'on injecte dans l'équation et que l'on extrait les coefficients de y^n . Pour $n = 0$ on obtient la tautologie

$$0 = t(t_0 - t_0) - q^2(t_0 - t_0).$$

Pour $n \geq 1$ on obtient une relation entre t_n et t_{n-1}

$$(q - t)q^{n+1}t_{n-1} = t(q^n t_n - q^{2n} t_n) - q^2(t_n - q^n t_n). \quad (160)$$

Cela donne la relation

$$\frac{t_n}{t_{n-1}} = \frac{(t - q)q^{n-1}}{(1 - q^n)(1 - q^{n-2}t)}. \quad (161)$$

La série $T(y)$ est donc hypergéométrique basique de raison

$$r_T(X) = \frac{(t - q)X}{(1 - qX)(1 - t/qX)}. \quad (162)$$

Comme la valuation de l'empilement vide est 1 on a $t_0 = 1$. Par itération de la raison $r_T(X)$ on trouve le développement de $T(y)$ de la formule (154).

Pour prouver le corollaire il suffit de combiner les résultats des propositions 3.19 et 3.18. \square

À ce stade on sait calculer une expression pour la série $P(t, u; y)$ et donc, d'après la proposition 3.17, on connaît aussi la série génératrice des chemins à creux sur l'axe débutant par un pic double puis éventuellement un pic. Pour conclure, il reste à faire le lien entre ces chemins et les chemins à creux sur l'axe sans contrainte sur leur début. La proposition suivante donne une décomposition combinatoire décrivant ce lien. Elle donne un changement de fonction inconnue permettant de linéariser l'équation q -algébrique satisfaite par $C(t, u; y)$. La forme de ce changement de fonction inconnue est identique à celle suggérée par Brak et Prellberg. Le dénominateur commun aux deux coefficients a, sur cet exemple, une explication combinatoire.

Proposition 3.20 *La série génératrice $C(t, u; y)$ des chemins non vides à creux sur l'axe et la série génératrice $G(t, u; y)$ des chemins à creux sur l'axe débutant par un pic double puis éventuellement un pic sont liées par la relation*

$$G(t, u; y) = t^2 \left(1 + qy + \frac{q^2 y^2}{t} C(t, u; y) \right). \quad (163)$$

Soit $T(t, u; y) \equiv T(y)$ la série génératrice alternée des empilements triviaux de segments marqués donnée par (154). On a

$$C(t, u; y) = \frac{t}{q^2 y^2} \left(\frac{T(qy)}{T(y)} - 1 - qy \right). \quad (164)$$

De façon plus explicite,

$$C(t, u; y) = - \frac{t \sum_{n=0}^{\infty} \frac{(t-q)^{n+1} q^{\binom{n+1}{2}} y^n}{(q)_n (t/q)_{n+2}}}{q \sum_{n=0}^{\infty} \frac{(t-q)^n q^{\binom{n}{2}} y^n}{(q)_n (t/q)_n}}. \quad (165)$$

Corollaire 3.20.1 *La série génératrice $E_{\blacksquare}(t, u; y)$ des chemins à petits creux noirs vaut*

$$E_{\blacksquare}(t, u; y) = - \frac{t \sum_{n=0}^{\infty} \frac{(t-q)^{n+1} q^{\binom{n+2}{2}-1} y^n}{(q-1)^n (q)_n (t/q)_{n+2}}}{q \sum_{n=0}^{\infty} \frac{tq(t-q)^n q^{\binom{n+1}{2}} y^n}{(q-1)^n (q)_n (t/q)_n}}. \quad (166)$$

Preuve. La relation (163) correspond à la décomposition des chemins à creux sur l'axe débutant par un pic double puis éventuellement un pic. Cette décomposition est illustrée sur la figure 3.27.

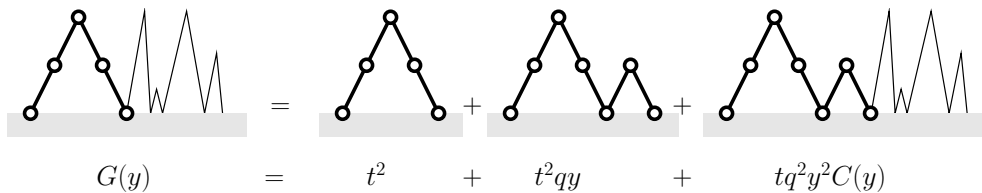


FIG. 3.27 – Décomposition des chemins à creux sur l'axe débutant par un pic double puis éventuellement un pic

En utilisant les résultats des propositions 3.17 et 3.18 et l'équation (163) on peut exprimer $C(y)$ en fonction de $T(y)$ pour obtenir l'expression de l'équation (164) qui est de la forme

$$C(y) = \frac{\delta}{tq^2 y^2} \frac{H(qy)}{H(y)} + \frac{\gamma(y)}{tq^2 y^2}. \quad (167)$$

Sur cet exemple on voit que l'égalité des dénominateurs des coefficients s'explique par le fait qu'ils proviennent tous les deux du coefficient de $C(y)$ dans l'équation (163). Ce phénomène laisse espérer que cette heuristique pour le choix de la linéarisation, notamment dans la proposition 3.8 est peut-être généralisable.

La substitution de l'expression de $T(qy)/T(y)$ calculée dans le corollaire 3.19.1 donne comme expression pour $C(t, u; y)$

$$C(t, u; y) = \frac{t}{q^2 y^2} \left(\frac{\sum_{n \geq 0} \frac{(t-q)^n q^{\binom{n+1}{2}}}{(q)_n (t/q)_n} y^n}{\sum_{n \geq 0} \frac{(t-q)^n q^{\binom{n}{2}}}{(q)_n (t/q)_n} y^n} - 1 - qy \right). \quad (168)$$

On remarque que la réduction au même dénominateur de cette expression conduit à l'annulation des coefficients de y^0 et y^1 du numérateur. Une factorisation par y^2 du numérateur, une simplification avec le y^2 du dénominateur et une réindexation de la somme du numérateur permettent d'obtenir l'expression (165).

Pour le corollaire il suffit d'utiliser la proposition 3.16. □

Chapitre 4

Sandpile model and Tutte polynomials

A new explicit bijection between spanning trees and recurrent configurations of the sandpile model is given. This mapping is such that the difference between the number of grains on a configuration and the external activity of the associate tree is the number of edges of the graph. It is a bijective proof of a result of Merino-López that express the generating function of recurrent configurations as an evaluation of the Tutte polynomial.

RÉSUMÉ. On donne une nouvelle bijection entre les arbres couvrants et les configurations récurrentes du modèle du tas de sable de Dhar. Cette application est construite de sorte que la différence entre le nombre de grains de la configuration et l'activité externe de l'arbre associé est le nombre d'arêtes du graphe. Cela donne une preuve bijective d'un résultat de Merino-López identifiant la distribution des configurations récurrentes selon leur nombre de grains à une évaluation partielle du polynôme de Tutte du graphe.

Introduction

The sandpile model was introduced as a model in physics ([BTW88], [DRSV95], [DM92]) and was also considered by combinatorialists as the chip firing game ([BLS91],[Big99],[ML97],[BW97],[CR00]).

One of the interesting results about this model from a combinatorial point of view is that the Tutte polynomial of the underlying graph enumerates the configurations with respect to their level. This result was conjectured by Biggs and proved by Merino-López. His proof is based on the recursive definition of Tutte polynomials.

Our aim is to give an explicit bijection between spanning trees of the graph and recurrent configurations, which maps a tree with p externally active edges to a configuration of level p , giving a bijective proof of the above result.

Our paper is organized as follows. In section 1, we recall definitions of the sandpile model and the Tutte polynomial of a graph and then the result of Merino-López. In section 2, we

define decreasing edge-vertex traversals of a graph which are intermediate objects between spanning trees and recurrent configurations. Indeed in section 3 we give a bijection between them and spanning trees and in section 4 between them and recurrent configurations. Those bijections preserve a parameter defined on a traversal with the external activity of the mapped tree, respectively the mapped configuration, so we recover the result of Merino-López in a bijective way. In section 5 we use the precedent bijection to recover bijectively the fact that the number of spanning subgraphs of a given graph is an evaluation of its Tutte polynomials.

4.1 Preliminaries

We present our notation and some definitions.

Let $G = (V, q, E)$ be a connected rooted graph, where V is the set of vertices, $q \in V$ the root and E the multi-set of the edges. The endpoints of an edge (ab) are a and b . We assume there are no loops, edges (aa) , in E . For any vertex a , $deg(a)$ denotes its degree. $|A|$ denotes the cardinality of the set (or multi-set) A .

The Sandpile model

A *configuration* u on $G = (V, q, E)$ is a mapping $V \rightarrow \mathbb{N}$ (where \mathbb{N} denotes the set of non-negative integers). An interpretation is that the vertex $a \in V$ contains $u(a)$ grains of sand. A vertex a is *unstable* in u if $u(a) \geq deg(a)$; then this vertex can *topple* giving the configuration v defined by :

$$\begin{cases} v(a) = u(a) - |\{(at) | t \neq a, (at) \in E\}| \\ v(b) = u(b) + |\{(ba) | (ba) \in E\}| \text{ for } b \neq a \end{cases}$$

In other words, the vertex u passes one grain of sand along each of its edges to the vertex on the other end of the edge. We will denote by $u \xrightarrow{a} v$ the toppling of vertex a . A configuration is *stable* if all vertices different from q are stable. A stable configuration u is *recurrent* if $u(q) = deg(q)$ and there exists an order $a_1 <_u a_2 <_u \dots <_u a_{|V|}$ on the vertices such that $u \xrightarrow{a_1} \dots \xrightarrow{a_i} \dots \xrightarrow{a_{|V|}} u$. This means that if q is toppled then all vertices of V could be toppled getting back to the initial configuration. $Rec(G)$ denotes the set of recurrent configurations of G . The *level* of a recurrent configuration u is given by

$$level(u) = \left(\sum_{i \in V} u_i \right) - |E|.$$

Let $\mathcal{R}(G, y)$ be the polynomial which enumerates the distribution of recurrent configurations according to their levels,

$$\mathcal{R}(G, y) = \sum_{u \in Rec(G)} y^{level(u)}.$$

The Tutte polynomial

Given a graph $G = (V, E)$ and a order $<_e$ over E , Tutte ([Tut54]) introduced the two-variable polynomial $\mathcal{T}(G, x, y)$ defined by :

$$\mathcal{T}(G, x, y) = \sum_{T \in Sp(G)} x^{i(T)} y^{j(T)}$$

where $Sp(G)$ is the set of spanning trees of G and $i(T)$ (respectively $j(T)$) is the internal (respectively external) activity of T for which we recall the definition : Let T be a spanning tree T of $G = (V, E)$.

An edge $e \in T$ is *internally active* in T if it is the lowest edge, according to $<_e$, in the unique cocycle contained in $(E \setminus T) \cup \{e\}$ (A *cocycle* is a minimal disconnecting subset of E). The *internal activity* of the spanning tree T is the number of edges internally active in T .

An edge $e \in E \setminus T$ is *externally active* in T if it is the lowest edge, according to $<_e$, in the unique cycle of $(V, T \cup \{e\})$. The *external activity* of the spanning tree T is the number of edges externally active in T .

A first non-trivial observation is that this polynomial does not depend on the arbitrary order $<_e$ even if it is the case for the activities of a given spanning tree.

The relation between the two presented notions

Answering to a conjecture of Biggs, Merino-López [ML97] proved that :

Proposition 4.1 *For a graph $G = (V, E)$, a vertex $q \in V$ and an order $<_e$ on E ,*

$$\mathcal{R}(G, y) = \mathcal{T}(G, 1, y).$$

This means that there are as many recurrent configurations of level i as spanning trees of external activity i .

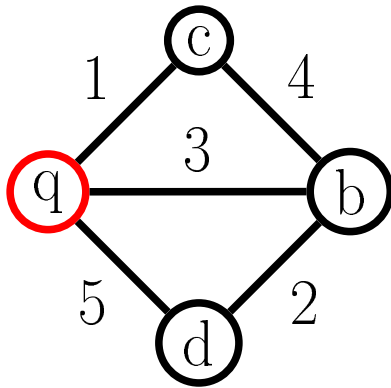
His proof is based on a recursive characterization of the Tutte polynomial, which does not gives a one-to-one correspondence between trees and configurations. Note that Majumdar and Dhar ([DM92]) already gave a bijection between spanning trees and recurrent configurations. The aim of this note is to build a sharper bijection mapping recurrent configurations of level i and spanning trees of external activity i . We introduce a third set of objects, decreasing edge-vertex traversals, in bijection with the two precedents sets and a parameter on these objects which is translated into external activity, respectively level, by the bijections.

4.2 Decreasing traversals

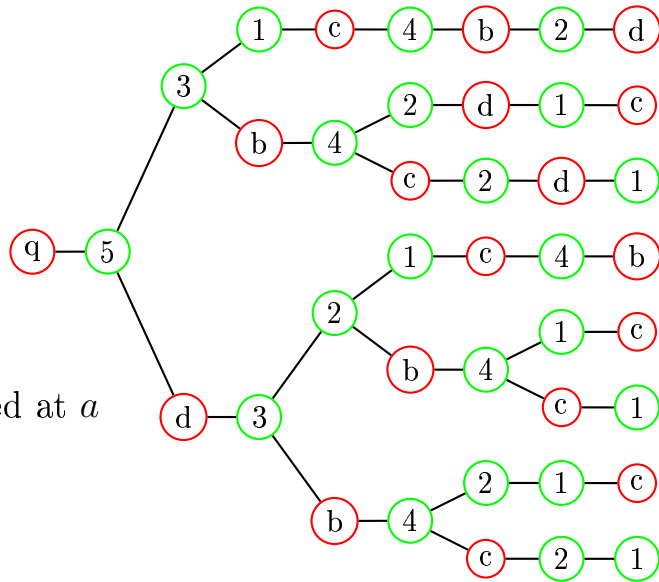
An *edge-vertex traversal* $\sigma = (\sigma_i)_{1 \leq i \leq |V|+|E|}$ of G is a sequence in which each vertex and edge appears exactly only once. We note $\sigma^{<i} = \{\sigma_j | j < i\}$ to formally describe the elements that appears before the element σ_i in σ . A *decreasing traversal* σ is an edge-vertex traversal such that :

- (i) If σ_i is a vertex of $V \setminus \{q\}$ then σ_{i-1} is an edge incident to σ_i .
- (ii) If σ_i is an edge, then it is maximal, according to $<_e$, among all the edges not in $\sigma^{<i}$ and incident to a vertex of $\sigma^{<i}$.

We denote by $\mathcal{D}(G)$ the set of decreasing traversals of G .



A graph G with an order on edges rooted at a



The set of traversals $\mathcal{D}(G)$ represented as a tree

An edge $\sigma_i = (\sigma_j \sigma_k)$ is *strong* in the traversal σ if it appears after its endpoints (ie $j < i$ and $k < i$). The *strength* $\Sigma(\sigma)$ of the traversal is its number of strong edges.

Lemma 1 *Assume σ and τ are two different decreasing traversals. Let k be the minimal index such that $\sigma_k \neq \tau_k$. Then one of σ_k and τ_k is an edge and the other a vertex.*

Proof. By choice of k , we have $\sigma^{<k} = \tau^{<k}$ and $\sigma^{<k-1} = \tau^{<k-1}$.

- Assume σ_k and τ_k are both vertices. Then they are distinct endpoints of the edge $\sigma_{k-1} = \tau_{k-1}$. The existence of a third endpoint in $\sigma^{<k-1} = \tau^{<k-1}$ from condition (ii) gives a contradiction.

- σ_k and τ_k can not be both edges since there is only one maximum among the edges with one endpoint in $\sigma^{<k} = \tau^{<k}$.

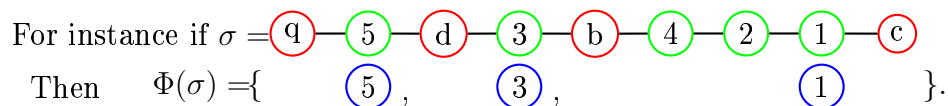
The only possible case is that one of σ_k and τ_k is an edge and the other a vertex.

□

Remark that the first element of a decreasing traversal is q . When the k first elements are fixed, σ_{k+1} is either the maximal edge of $\{e = (vw) \in E \setminus \sigma^{<k} \mid v \in \sigma^{<k} \text{ or } w \in \sigma^{<k}\}$ or the possible vertex incident to σ_k not in $\sigma^{<k+1}$. Hence a decreasing traversal is characterized by the position of the vertices different from q .

4.3 Decreasing traversals and spanning trees

Let Φ be the function which maps the decreasing traversal σ with the set of edges $\Phi(s) = \{\sigma_{k-1} \mid \sigma_k \in V \setminus \{q\}\}$.



Lemma 2 Φ is a bijection between $\mathcal{D}(G)$ and $Sp(G)$.

Proof. From (i), $\Phi(\sigma)$ is a set of $|V| - 1$ edges which cover V hence it is a spanning tree. (from (ii), q is covered by the edge in $\Phi(\sigma)$ of minimal index in σ)

- Φ is an injection :

Let σ and τ be two different decreasing traversals and k the minimal index such that $\sigma_k \neq \tau_k$. From lemma 1 we could assume w.l.o.g that σ_k is a vertex and τ_k an edge. Then $\sigma_{k-1} = \tau_{k-1}$ is an edge of $\Phi(\sigma)$ but not of $\Phi(\tau)$. Hence $\Phi(\sigma) \neq \Phi(\tau)$.

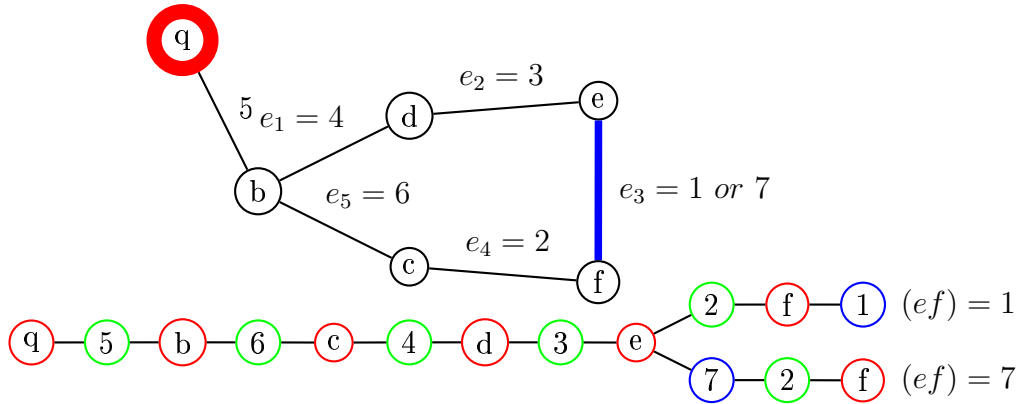
- Φ is a surjection :

Let T be a spanning tree and σ the decreasing traversal defined by : σ_{i+1} is a vertex if and only if σ_i is an edge of T . If σ is well defined then $\Phi(\sigma) = T$. Assume it is not the case, and let k be the last index of the partial traversal. Let $v \notin \sigma^{<k+1}$ a vertex at minimal distance of q in T and $e = (vw)$ the first edge along this path from v to q . By choice of v , $w \in \sigma^{<k+1}$ so from (ii), $e \in \sigma^{<k_0+1}$; $e = \sigma_i$ then $v = \sigma_{i+1} \in \sigma^{<k+1}$. Contradiction. □

Lemma 3 For any decreasing traversal σ , an edge is strong in σ if and only if it is externally active in $\Phi(\sigma)$. So

$$\Sigma(\sigma) = j(\Phi(\sigma)).$$

Proof. This example illustrate the following proof :



- If e is externally active in $\Phi(\sigma)$ then it is strong in σ .

Let $e = \sigma_i$ be an edge externally active in $\Phi(\sigma)$. Assume that a vertex of the cycle created by e and the tree $\Phi(\sigma)$ appears after e in σ . Let σ_j the first vertex of the cycle which appears after e . Then σ_{j-1} is an edge of the cycle and there is no vertex of the cycle between e and this edge, since vertices of the cycle follow an edge of the cycle, then σ_{j-1} was among the edges with at least one endpoint in $\sigma^{<i}$ as e was added to σ . Then $e > \sigma_{j-1}$ and there is a contradiction since we assume e was the minimal edge of the cycle.

- If e is strong in σ then $\Phi(\sigma)$ is externally active in $\Phi(\sigma)$.

Let e be a strong edge of σ . e is external since there is no endpoint after it in σ . Then e create a cycle γ when it is added to the tree $\Phi(\sigma)$. The edges of the cycle in the tree $\Phi(\sigma)$ are by definition immediatly followed by one of their endpoint. Conversely, except for the first vertex of the cycle in σ , all vertices of the cycle are successors of an edge of the cycle. Using those remarks starting from the endpoint of the edge e we state that the edge e is the last element of the cycle in σ . The minimality of e in the cycle comes from the fact we execute the following algorithm when we restrict construction of σ to the edges of the cycle and while the external edge of the cycle has not be added to σ :

- ◊ $i := 1, j := n, k := 1$.
- ◊ While $i \neq j$ do
 - If $e_i > e_j$
 - then $\sigma_k := e_i; i := i + 1;$
 - else $\sigma_k := e_j; j := j - 1;$
 - $k := k + 1$

When this algorithm ends, the edge $e_i = e_j$ is minimal among the others. In our case we show that the edge e is the edge which appears after all the others thus the algorithm ends and e is minimal in the cycle.

□

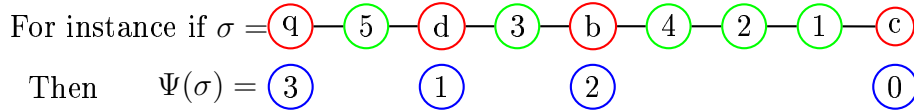
4.4 Decreasing traversals and recurrent configurations

For a vertex v , $I(v)$ denotes the set of edges incident to v .

Let Ψ the function which maps the decreasing traversal σ with the configuration u defined by :

$$u(\sigma_i) = |I(\sigma_i) \setminus \sigma^{<i}|.$$

In other words, there are as many grains on a vertex as incident edges that appear after it in the decreasing traversal.



Lemma 4 Ψ is a bijection between $\mathcal{D}(G)$ and $Rec(G)$.

Proof. • Ψ is a function from $\mathcal{D}(G)$ to $Rec(G)$.

From (i), there is at least one edge ($\sigma_{k-1} \in \sigma^{<k}$) which is not in the set of edges incident to σ_k that bring one grain on σ_k . Thus there are at most $|I(\sigma_k)| - 1$ grains on any vertex $\sigma_k \neq q$ and $\Psi(\sigma)$ is a stable configuration.

$\Psi(\sigma)$ is recurrent because the order induced on the vertices by σ is an order of toppling :

- ◇ Since $q = \sigma_1$, all edge incident to q brings one grain on q . So $\Psi(\sigma)(q) = deg(q)$.
- ◇ Assume all vertices of $\sigma^{<k}$ have been toppled, we count the grains on the vertex σ_k . Initially there were $|I(\sigma_k) \setminus \sigma^k|$ grains on σ_k . All edges of $I(\sigma_k) \cap \sigma^{<k}$ have an endpoint which is a vertex of $\sigma^{<k}$ whose toppling brings on grain per edge to σ_k . Thus there are at least $|I(\sigma_k) \cap \sigma^{<k}|$ added on σ_k during the toppling of vertices of $\sigma^{<k}$. Hence there is at least $|I(\sigma_k) \setminus \sigma^{<k}| + |I(\sigma_k) \cap \sigma^{<k}| = |I(\sigma_k)|$ grains on σ_k which is unstable when it is toppled.

• Ψ is an injection :

Let σ and τ be two different decreasing traversals. Let k be the smallest index such that $\sigma_k \neq \tau_k$, because of lemma 1 we assume w.l.o.g that $\sigma_k = v$ is a vertex and $\tau_k = e$ an edge. Let f be the edge that precede v in τ .

$$\begin{aligned} \sigma &= \sigma^{<k}, v, \dots, f \dots e \dots \\ \tau &= \sigma^{<k}, e, \dots, f, v, \dots \end{aligned}$$

There are at least one more edge that appears before v in τ than in σ since incident edges in $\sigma^{<k}$ are also before v in τ and the edge f is before v in τ not in σ . Thus $\Psi(\sigma)(v) < \Psi(\tau)(v)$. Hence $\Psi(\sigma) \neq \Psi(\tau)$.

• Ψ is a surjection : Given a recurrent configuration u we define the traversal σ as follow :

- ◊ $\sigma_1 = q$
- ◊ If σ_{i-1} is an edge with one endpoint v not in $\sigma^{<i}$ and $|I(v) \setminus \sigma^{<i}| = u(v)$ then $\sigma_i = v$, else σ_i is the maximal edge according to $<_e$ among edges not in $\sigma^{<i}$ but with at least one endpoint in.

Thus we build a complete decreasing traversal such that $\Psi(\sigma) = u$. □

Lemma 5 *For any decreasing traversal σ ,*

$$\Sigma(\sigma) = \text{level}(\Psi(\sigma)).$$

Proof. A strong edge $(\sigma_i \sigma_j)$ (w.l.o.g $i < j$) appears after its both endpoints so it brings two grains, one on each endpoint, whereas a weak edge brings a grain only in one endpoint. Hence the configuration $\Psi(\sigma)$ is made of $|E| + \Sigma(\sigma)$ grains and its level is $\Sigma(\sigma)$. □

4.5 Counting spanning subgraphs

In this section we use decreasing traversals in order to derive a bijective proof of a well known result concerning the enumeration of spanning subgraphs with a given number of edges.

Let $G = (V, E)$ be a connected graph, a spanning subgraph of G is a subset $E' \subset E$ of edges such that (V, E') is connected, such a subgraph satisfies :

$$|V| - 1 \leq |E'| \leq |E|$$

Proposition 1 *The number γ_k of spanning subgraphs of G containing $n - 1 + k$ edges is given by :*

$$\gamma_k = \sum_{i \geq k} \binom{i}{k} \alpha_i$$

where α_i is the number of spanning trees of G with external activity i .

As a consequence the total number of spanning subgraphs of G is

$$\sum_{k \geq 0}^{|E| - |V| + 1} 2^k \alpha_k = \mathcal{T}_G(1, 2)$$

Proof. We define a surjection π from the set of spanning subgraphs of G with $|V| - 1 + k$ edges to the set of decreasing traversals with $i \geq k$ strong edges, then we show that each traversal with i strong edges is the image by π of $\binom{i}{k}$ spanning subgraphs with $|E| - |V| + k$ edges.

Consider a spanning subgraph of $G = (V, E)$ given by the subset E' of E , the traversal $\pi(E')$ is built as follows :

- ◇ $\sigma_1 = q$
- ◇ if $\sigma_i \in E'$ and x is an endpoint of σ_i not visited then $\sigma_{i+1} = x$
- ◇ else (*i. e.* $\sigma_i \in V$, or $\sigma_i \in E \setminus E'$ or $\sigma_i \in E'$ and its two endpoints are visited) σ_{i+1} is the largest reachable and non visited edge.

It is clear from the algorithm that each edge in E' is either in the tree $\Phi(\sigma)$ or is strong. Moreover π is a surjection, since E' of E_s for a traversal σ with set of strong edges is E_s and for each subset we have $\sigma = \pi(E' \cup \Phi(\sigma))$, giving the result.

□

Perspectives

En guise de conclusion on présente les perspectives à court ou moyen terme qui se dégagent de cette thèse.

Dans le chapitre 1, on propose la conjecture 1.6.1 sur la série génératrice des demi-animaux pointés dans leur périmètre de site et inclus dans un ruban borné. L'apparition des polynômes de Fibonacci dans les récurrences laisse espérer une interprétation combinatoire en termes d'empilements.

- D'autre part, la notion d'animal occupé a permis d'exhiber une classe d'animaux \mathcal{B} pour laquelle on sait calculer la série formelle liée à la probabilité de percolation

$$\sum_{a \in \mathcal{B}} p^{\text{aire}(a)} (1-p)^{\text{perimetre}(a)} = \sum_{a \in \mathcal{A}} (-p)^{\text{aire}(a)} = \frac{1}{2} \left(\sqrt{\frac{1+p}{1-3p}} - 1 \right)$$

où \mathcal{A} est la classe des animaux dirigés. Une explication combinatoire de cette égalité pourrait éclairer un peu plus le lien entre modèles de gaz et animaux.

- Enfin, la correspondance entre les diagrammes aléatoires du modèle CGM et les animaux dirigés semble pouvoir se généraliser à des modèles de gaz où certaines transitions permettent d'"oublier" partiellement la configuration précédente. Quel paramètres supplémentaires, comme le périmètre droit déjà étudié par Bousquet-Mélou, peuvent se traiter par cette approche ?

Dans le chapitre 2, on a présenté une famille d'algorithmes permettant de coder un grand nombre d'applications entre suites quasi-décroissantes et chemins de Dyck. D'abord on a montré qu'il est possible de comprendre ces applications, d'en prouver le caractère bijectif ou non, d'en déduire des analogues de la dernière descente et de l'aire pour les chemins de Dyck. Ensuite on a essayé d'accumuler des exemples de contextes combinatoires dans lesquels certains de ces bijections, *a priori* artificielles, apparaissent.

- Pour poursuivre cette démarche on peut s'intéresser à la reconnaissance des bijections f entre chemins de Dyck qui peuvent s'écrire sous la forme $f = \Phi_{m_1} \circ \Phi_{m_2}^{-1}$. En effet si on fixe une des applications, Φ_{m_2} par exemple, il semble possible d'écrire un algorithme déterminant, dans un grand nombre de cas, l'application Φ_{m_1} . L'existence d'un tel algorithme permettrait d'envisager la construction semi-automatique d'une base de données répertoriant une partie des bijections entre les nombreux objets comptés par la suite de Catalan.

- Les différentes familles d'algorithmes correspondent parfois à des combinaisons de différents modes d'insertions. Certaines de ces combinaisons préservent peut-être le caractère bijectif. Par exemple, si Φ_{m_1} et Φ_{m_2} sont bijectifs, peut-on en déduire que $\Phi_{\sigma\bar{\sigma}\circ m_1}$ ou $\Phi_{\langle m_1; m_2 \rangle}$ le sont aussi ? On peut envisager l'étude du groupe de bijections engendré par les bijections de la forme $\Phi_{m_2} \circ \Phi_{m_1}^{-1}$. On a pu constater expérimentalement certaines relations non-triviales entre ces bijections.
- D'autre part les bijections proposées sont toutes liées aux suites quasi-décroissantes. On peut essayer de s'affranchir de cet aspect en définissant des bijections à l'aide d'une procédure de choix d'un pas montant et d'un pas descendant à supprimer dans tout chemin de Dyck. Cette procédure suffit à définir un arbre (ECO ?) dont les sommets de profondeur n sont occupés par les chemins de Dyck de taille n . Apparaît-il dans ces arbres des distributions de degrés sensiblement différentes de celle liée aux suites quasi-décroissantes ?
- Enfin, on a borné l'étude aux chemins de Dyck, mais on peut envisager de généraliser cette démarche à d'autres objets. Les chemins du grand Dyck, de Schröder sont deux exemples. Un troisième serait la classe des permutations dans laquelle les statistiques eulérienne et mahonienne semblent avoir un rôle similaire à la longueur de la dernière descente et l'aire des chemins de Dyck. Comme certaines classes de permutations évitant un motif sont aussi comptées par la suite de Catalan, par exemple pour le motif 132, le lien pourrait être assez significatif. Plus précisément, on pourrait commencer par s'intéresser aux paires de tableaux de Young associées aux permutations évitant le motif 132 par l'algorithme RSK.
- Une perspective annexe de ce chapitre provient de la résolution à l'aide d'un développement de Taylor selon la variable catalytique de l'équation fonctionnelle écrite pour la méthode du noyau. Cette technique peut-elle s'appliquer à des équations fonctionnelles plus complexes et faire apparaître des polynômes analogues à ceux de Fibonacci ?

Dans le chapitre 3, on a proposé des preuves de plus en plus combinatoires d'une formule énumérant les interactions supérieures des chemins de Dyck.

- La dernière de ces preuves est incomplète, la compléter pourrait permettre de voir la formule recherchée comme une instance du lemme d'inversion pour des empilements de segments.
- D'autre part, la comparaison des différentes preuves a permis de mettre à jour un second changement de variable permettant de linéariser des équations q -algébriques de la forme

$$a(q, y) + b(q, y)D(y) + c(q, y)D(qy) + d(q, y)D(y)D(qy) = 0.$$

Le premier changement de variable a déjà permis de résoudre un certain nombre d'équations q -algébriques comme celle de l'aire sous les chemins de Dyck. On peut s'intéresser à la comparaison de ces solutions avec celles obtenues par le nouveau changement de variable.

- Enfin, aucune question autour du comportement asymptotique de la formule n'a été traitée. On peut s'y attaquer avec ou sans ces nouveaux résultats.

Dans le chapitre 4, pour tout graphe G on a fait le lien entre l'activité externe d'un arbre couvrant et le nombre de grains de la configuration récurrente associée.

- Depuis les travaux de Dhar et Majumdar on sait associé un groupe $SP(G)$ aux configurations récurrentes du graphe G . Cori et Rossin ont montré que si G est un graphe planaire le groupe $SP(G^*)$ associé au dual G^* de G est isomorphe à $SP(G)$. La dualité des graphes planaires est une bijection entre les arbres couvrants de G et G^* qui échange les activités externes et internes. Une interprétation de l'activité interne dans le cadre du modèle du tas de sable reste à trouver.
- D'autre part, quelques expériences suggèrent que la distribution de l'activité externe sur les arbres couvrants est log-concave : si a_k est le nombre d'arbres couvrants d'activité externe k alors $a_{k-1}a_{k+1} \leq a_k^2$. Les configurations récurrentes, du fait de la structure de groupe sous-jacente, ou les parcours d'arêtes et sommets, sont peut-être des objets plus adaptés à la construction d'une injection entre les couples de configurations de niveaux $k-1$ et $k+1$ et les couples de configurations de niveau k . L'existence d'une telle injection impliquerait la log-concavité de la distribution.

Annexe A

Résolution d'équations q -linéaires

Dans cette annexe on présente des approches de plus en plus perfectionnées pour trouver des solutions hypergéométriques basiques d'équations q -linéaires. On présente tout particulièrement un article de Abramov, Petkovšek et Paule [APP98] qui contient la méthode générale que l'on utilise dans les calculs de la section 3.3.

Soit \mathbb{B} un corps, et $\mathbb{A} = \mathbb{B}(q)$, le corps des fractions rationnelles en q avec q transcendant sur \mathbb{B} . Une série $B(y) = \sum_{n \geq 0} b_n y^n$ de $\mathbb{A}[[y]]$ est *hypergéométrique basique* (de base $q \in \mathbb{A}$) si il existe une fraction rationnelle $r_B(X) \in \mathbb{A}(X)$, telle que pour tout $n \geq 0$,

$$\frac{b_{n+1}}{b_n} = r_B(q^n).$$

La fraction $r_B(X)$ est la *raison* de cette série hypergéométrique.

Une telle série est caractérisée par sa raison $r_B(X)$ et son premier terme $B(0) = b_0$, car la raison permet d'exprimer b_n en fonction de b_0 . Par exemple la seule série de raison $r_B(X) = X/(1 - qX)$ telle que $B(0) = 1$ est

$$B(y) = \sum_{n \geq 0} \left(B(0) \prod_{k=0}^{n-1} r_B(q^k) \right) y^n = \sum_{n \geq 0} y^n \prod_{k=0}^{n-1} \frac{q^k}{1 - q^{k+1}} = \sum_{n \geq 0} \frac{q^{\binom{n}{2}}}{(q)_n} y^n.$$

Une *équation q -linéaire d'ordre K* d'inconnue $H(y) \in \mathbb{A}[[y]]$ est une équation de la forme

$$\sum_{k=0}^K a_k(y) H(q^k y) = 0$$

où les $a_k(y)$ sont des fractions rationnelles de $\mathbb{A}(y)$.

Trouver les solutions hypergéométriques d'une équation q -linéaire est d'une difficulté variable selon la forme des équations. Le cas le plus favorable est celui envisagé par Brak et Prellberg [BP95], il a été généralisé par Dutour [Dut96] puis par Duchon [Duc98]. On présente ici la généralisation de Dutour.

Proposition A.1 *On suppose que l'équation q -linéaire s'écrit sous la forme*

$$\sum_{k=0}^K a_k H(q^k y) + y \left(\sum_{k=0}^K b_k H(q^k y) \right) = 0 \quad (169)$$

avec les a_k et b_k indépendant de y . De plus, on suppose que $\sum_{k=0}^K a_k = 0$.

Les solutions de cette équation sont les séries hypergéométriques basiques de raison

$$r_H(X) = - \frac{\sum_{k=0}^K b_k X^k}{\sum_{k=0}^K a_k (qX)^k}. \quad (170)$$

Preuve. En évaluant l'équation q -linéaire en $y = 0$, on obtient l'égalité

$$\left(\sum_{k=0}^K a_k \right) H(0) = 0 \quad (171)$$

qui est satisfaite par hypothèse.

En injectant le développement $H(y) = \sum_{n \geq 0} h_n y^n$ dans l'équation et en extrayant le coefficient de y^{n+1} , on obtient pour $n \geq 0$

$$h_{n+1} \sum_{k=0}^K a_k q^{k(n+1)} + h_n \sum_{k=0}^K b_k q^{kn} = 0 \quad (172)$$

qui donne bien la relation $r_H(q^n) = \frac{h_{n+1}}{h_n}$.

□

En plus de la contrainte pour $y = 0$, cette méthode n'est applicable que si les coefficients de l'équation q -linéaire sont de degré au plus 1 en y . En effet, si le degré est plus grand on obtient une relation entre plus de deux h_{n+k} et cela ne permet pas de conclure systématiquement. Cependant dans ces cas on peut tenter de *conjecturer* la raison d'une éventuelle solution hypergéométrique. On illustre cette technique par un exemple tiré de [APP98].

Exemple : La série $A(y)$ de $\mathbb{A}[[y]]$ satisfaisant l'équation q -linéaire

$$q^2 y^2 A(q^3 y) + (1+q)yA(q^2 y) + (1-y)A(qy) - A(y) = 0 \quad (173)$$

telle que $A(0) = 1$ est la série hypergéométrique basique

$$A(y) = \sum_{n \geq 0} \frac{q^{n^2}}{(q)_n} y^n. \quad (174)$$

Preuve. On développe $A(y)$ sous la forme

$$A(y) = 1 + \sum_{n \geq 1} a_n y^n. \quad (175)$$

On injecte cette forme dans l'équation q -linéaire (173). En extrayant de cette expression les termes de degré n en y , on obtient une relation de récurrence sur les $(a_i)_{i \geq 1}$:

$$(1 - q^n)a_n = -q^{n-1}a_{n-1}(1 - q^{n-1}(1 + q)) + q^{3n-4}a_{n-2}. \quad (176)$$

Pour $n = 0$, on obtient

$$1 - 1 = 0 \quad (177)$$

qui est une tautologie. Les coefficients de l'équation q -linéaire étant polynomiaux en y le système linéaire correspondant à la récurrence (176) est triangulé ce qui permet de calculer facilement les premiers a_i :

$$\begin{aligned} a_1 &= \frac{q}{1 - q} \\ a_2 &= \frac{q^4}{(1 - q)^2(1 + q)} \\ a_3 &= \frac{q^9}{(1 - q)^3(1 + q + q^2)(1 + q)} \\ a_4 &= \frac{q^{16}}{(1 - q)^4(1 + q^2)(1 + q + q^2)(1 + q)^2}. \end{aligned}$$

Les facteurs de a_i apparaissant aussi dans a_{i+1} on s'intéresse au quotient de deux termes successifs :

$$\begin{aligned} \frac{a_1}{a_0} &= \frac{q}{1 - q} &= \frac{q}{(1 - q)} \\ \frac{a_2}{a_1} &= \frac{q^3}{(1 - q)(1 + q)} &= \frac{q^3}{(1 - q^2)} \\ \frac{a_3}{a_2} &= \frac{q^5}{(1 - q)(1 + q + q^2)} &= \frac{q^5}{(1 - q^3)} \\ \frac{a_4}{a_3} &= \frac{q^7}{(1 - q)(1 + q)(1 + q^2)} &= \frac{q^7}{(1 - q^4)}. \end{aligned}$$

La régularité de ces quotients incite à conjecturer que la série hypergéométrique de raison

$$r_A(X) = \frac{qX^2}{1 - qX} \quad (178)$$

est solution de l'équation q -linéaire. On pose donc

$$A(y) = \sum_{n \geq 0} \left(\prod_{k=0}^{n-1} r_A(q^k) \right) y^n = \sum_{n \geq 0} \frac{q^{n^2}}{(q)_n} y^n. \quad (179)$$

On vérifie ensuite que cette expression satisfait bien la relation (173) ou bien que la suite $(q^{n^2}/(q)_n)$ satisfait la récurrence (176). \square

Cette technique est donc basée sur une analyse des quotients a_{n+1}/a_n , dont on tire une conjecture sur la raison de la série hypergéométrique basique solution. Cependant cette conjecture n'est pas toujours aussi aisée. On peut remarquer par exemple que toutes les fractions de la forme

$$r(X) = F(X) \prod_{k=0}^4 (q^k - X) + \frac{qX^2}{1 - qX} \quad (180)$$

satisfont les valeurs qui ont permis de conjecturer la valeur de $r_A(X)$. De façon plus générale cette simple analyse ne permet pas de prouver l'absence de solution hypergéométrique basique. Abramov, Paule et Petkovsek [APP98] ont proposé une analyse plus détaillée pour la recherche des éventuelles raisons des solutions hypergéométriques basiques. La première clef de leur algorithme est d'écrire l'hypothétique raison sous une forme normale dont la proposition suivante montre l'existence pour toute fraction rationnelle non nulle. Soit $P(X)$ et $Q(X)$ deux polynômes, leur *plus grand diviseur commun* se note $P(X) \wedge Q(X)$.

Proposition A.2 (*Forme normale pour les fractions rationnelles*) *Toute fraction rationnelle $r(X)$ non nulle à coefficients dans le corps $\mathbb{A} = \mathbb{B}(q)$ s'écrit de façon unique sous la forme :*

$$r(X) = z \frac{a(X)}{b(X)} \frac{c(qX)}{c(X)}$$

où $a(X)$, $b(X)$, $c(X)$ sont des polynômes en X de coefficient dominant 1 tels que :

$$\begin{cases} a(X) \wedge b(q^k X) = 1 \quad \forall k \in \mathbb{N} \\ b(X) \wedge c(qX) = 1 \\ a(X) \wedge c(X) = 1 \\ c(0) \neq 0 \end{cases}$$

et z une constante, c'est-à-dire un élément de $\mathbb{A}(q)$.

Avant de se lancer dans les preuves, on fait quelques rappels d'algèbre linéaire qui ont été nécessaires au rédacteur. Soit $P(X)$ et $Q(X)$ deux polynômes de $\mathbb{A}[X]$ de degré en X $\deg(P)$ et $\deg(Q)$. Pour savoir si ces deux polynômes ont un facteur commun, on va utiliser le résultant de ces deux polynômes. On considère les deux familles de polynômes $(X^k P(X))_{k=0.. \deg(Q)-1}$ et $(X^k Q(X))_{k=0.. \deg(P)-1}$. L'union de ces deux

familles est une famille de $\deg(P) + \deg(Q)$ vecteurs de l'espace vectoriel des polynômes de degré au plus $\deg(P) + \deg(Q) - 1$ en X . Le *résultant* $Res(P(X), Q(X))$ des polynômes $P(X)$ et $Q(X)$ est le déterminant de la famille (ordonnée) des vecteurs $(X^k P(X))_{k=0.. \deg(Q)-1}, (X^j Q(X))_{j=0.. \deg(P)-1}$.

Lemme A.2.1 *Les polynômes $P(X)$ et $Q(X)$ ont un facteur commun si et seulement si leur résultant est nul. Autrement dit*

$$P(X) \wedge Q(X) \neq 1 \iff Res(P(X), Q(X)) = 0.$$

Preuve.

- Si $P(X) \wedge Q(X) = R(x)$ avec $\deg(R) \geq 1$ alors la famille $(X^k(P(X)/R(X)), (X^j(Q(X)/R(X)))$ est formée de $\deg(P) + \deg(Q)$ polynômes de degré au plus $\deg(P) + \deg(Q) - 1 - \deg(R)$. La dimension de l'espace vectoriel des polynômes bornés par un tel degré étant $\deg(P) + \deg(Q) - \deg(R)$, elle est strictement inférieure au cardinal de la famille qui est donc liée. On en déduit l'existence d'une combinaison linéaire représentable par deux polynômes $M_P(X)$ et $M_Q(X)$ de degré inférieur à $\deg(Q) - \deg(R) - 1$, respectivement $\deg(P) - \deg(R) - 1$ tels que

$$M_P(X)(P(X)/R(X)) + M_Q(X)(Q(X)/R(X)) = 0. \quad (181)$$

En remultipliant par $R(X)$ cette relation, on trouve une combinaison linéaire nulle, non triviale, de la famille initiale. Le déterminant de cette famille, c'est-à-dire le résultant, est donc nul.

- Si le résultant est nul, c'est qu'il existe une combinaison linéaire entre les polynômes de la famille définissant le résultant. Celle-ci se représente par deux polynômes $M_P(X)$ et $M_Q(X)$ de degré au plus $\deg(Q) - 1$, respectivement $\deg(P) - 1$ tels que

$$M_P(X)P(X) = M_Q(X)Q(X).$$

Comme le degré de $M_Q(X)$ est strictement inférieur à celui de $P(X)$, $P(X)$ ne divise pas $M_Q(X)$. On en déduit que $P(X)$ et $Q(X)$ ont un facteur commun, ie $P(X) \wedge Q(X) \neq 1$. \square

Lors de l'analyse, on sera parfois amené à chercher pour un polynôme $P(X)$, l'ensemble des entiers $n \in \mathbb{N}$ tels que $P(q^n) = 0$. Le lemme suivant garantit que l'on dispose d'un algorithme permettant effectivement de trouver cet ensemble.

Lemme A.2.2 *On suppose que q est transcendant par rapport à \mathbb{B} . Il existe un algorithme qui pour tout polynôme $P(X) \in \mathbb{A}[X] = \mathbb{B}(q)[X]$ détermine l'ensemble des entiers $n \in \mathbb{N}$ tels que $P(q^n) = 0$.*

Preuve. Le but est de calculer une borne supérieure sur n à partir du polynôme $P(X)$. À une réduction au même dénominateur près, on peut supposer que le polynôme s'écrit sous

la forme

$$P(X) = \sum_{i=0}^{\deg(P)} \sum_{j=0}^{\deg_q} a_{i,j} q^j X^i$$

où $a_{i,j} \in \mathbb{A}$. On note $T = \max\{j \mid \exists i, a_{i,j} \neq 0\}$. Pour tout $n \in \mathbb{N}$, on a donc

$$P(q^n) = \sum_{i=0}^{\deg(P)} \sum_{j=0}^{\deg_q} a_{i,j} q^j q^{ni}.$$

Pour n suffisamment grand ($n > T$), le plus grand degré en q de $P(q^n)$ est *a priori* de la forme $t + ns$ où $s = \max\{i \mid \sum_j a_{i,j} \neq 0\}$ et $t = \max\{j \mid a_{s,j} \neq 0\}$. Si $P(q^n) = 0$, alors les termes de degré $t + ns$ en q s'annulent car q est transcendant sur \mathbb{A} . Il existe donc nécessairement un second coefficient de degré $t + ns$. Par définition de s , celui-ci provient d'un terme $q^{t'} X^{s'}$ avec $s' < s$. On a donc $t + ns = t' + ns'$ et comme $s - s' \geq 1$ et $t' \leq T$ on en déduit que $n \leq T$. Les valeurs de n telles que $P(q^n) = 0$ sont donc bornées par une constante T dépendant des coefficients du polynôme. Pour trouver l'ensemble des valeurs, il suffit donc de tester un nombre fini de valeurs ce qui garantit la terminaison de cet algorithme naïf. \square

Preuve.(Forme normale des fractions rationnelles, proposition A.2)

On veut écrire la fraction $r(X) = f(X)/g(X)$ avec $f(X), g(X) \in \mathbb{A}[X]$ sous la forme normale présentée dans la proposition A.2. Voici un algorithme qui la calcule :

$a(X) := f(X)$; $b(X) := g(X)$; $c(X) := 1$; $z := 1$;

tant que il existe un $n \in \mathbb{N}$ tel que $a(X) \wedge b(q^n X) \neq 1$ **faire**

Soit n le plus petit entier positif pour lequel $e(X) := a(X) \wedge b(q^n X) \neq 1$;

$a(X) := a(X)/e(X)$; $b(X) := b(X)/e(q^{-n} X)$; $c(X) := c(X) \prod_{k=-n}^{-1} e(q^k X)$;

fin

Soit ld_c le plus petit degré de $c(X)$;

$c(X) := c(X)/X^{ld_c}$; $z := q^{ld_c} z$;

Soit l_a le coefficient dominant de $a(X)$ et l_b celui de $b(X)$;

$a(X) := a(X)/l_a$; $b(X) := b(X)/l_b$; $z := z l_a / l_b$;

D'après la propriété du résultant présentée dans le lemme A.2.1, les entiers k tels que $a(X) \wedge b(q^k X) \neq 1$ sont aussi les racines de la forme q^k du polynôme en w , $Res(a(X), b(wX))$. On peut effectivement calculer ces racines en utilisant le lemme A.2.2. Cet algorithme est donc implémentable.

Un invariant de cet algorithme est que

$$r(X) = z \frac{a(X) c(qX)}{b(X) c(X)}.$$

Pour le vérifier, on constate que lors des affectations dans la boucle principale on a bien

$$z \frac{\frac{a(X)}{e(X)} c(qX) \prod_{k=-n}^{-1} e(q^{k+1}X)}{\frac{b(X)}{e(q^{-n}X)} c(X) \prod_{k=-n}^{-1} e(q^k X)} = z \frac{a(X) c(qX)}{b(X) c(X)}$$

après une simplification de $e(X)$ et $e(q^{-n}X)$.

De plus, $a(X) \wedge c(X) = 1$ et $b(X) \wedge c(qX) = 1$ sont deux autres invariants de l'algorithme. Les preuves étant symétriques, on se contente de le prouver pour $a(X)$ et $c(X)$. On suppose que $a_i(X)$ et $c_i(X)$ les valeurs, de $a(X)$ et $c(X)$ lors du i^{eme} passage dans la boucle, satisfont $a_i(X) \wedge c_i(X) = 1$. On montre que les nouvelles valeurs de $a_{i+1}(X)$ et $c_{i+1}(X)$ le satisfont aussi. Par choix de n_{i+1} , on a pour tout $k < n_{i+1}$,

$$a_i(X) \wedge b_i(q^k X) = 1$$

ce qui se récrit

$$(a_{i+1}(X)e_{i+1}(X)) \wedge (b_{i+1}(q^k X)e_{i+1}(q^{k-n_{i+1}}X)) = 1.$$

Cela implique que $a_{i+1}(X) \wedge e_{i+1}(q^j X) = 1$ pour $j (= k - n_{i+1})$ allant de -1 à $-n_{i+1}$. Comme $a_{i+1}(X)$ divise $a_i(X)$, on a aussi $a_{i+1}(X) \wedge c_i(X) = 1$. On utilise ces deux remarques pour conclure que $a_{i+1}(X)$ est premier avec $c_{i+1}(X) = c_i(X) \prod_{j=-n_{i+1}}^{-1} e_{i+1}(q^j X)$.

La terminaison de cet algorithme est garantie par la décroissance stricte du degré des polynômes $a(X)$ et $b(X)$. Puisque l'algorithme a terminé, le test de la boucle "tant que" a échoué donc

$$a(X) \wedge b(q^k X) = 1 \quad \forall k \in \mathbb{N}.$$

Si le polynôme $c(X)$ est de la forme $X^k d(X)$ avec $d(0) \neq 0$, alors grâce à une simplification par X^k on a $c(qX)/c(X) = q^k d(qX)/d(X)$. Cette remarque explique la division de $c(X)$ par X^{ldc} . Après cette opération, on a $c(0) \neq 0$.

Les dernières affectations normalisent les coefficients dominants des polynômes $a(X)$ et $b(X)$.

□

On détaille ici l'analyse permettant de trouver les éventuelles raisons de séries hypergéométriques basiques solutions d'une équation q -linéaire. L'algorithme correspondant à cette analyse sera utilisé lors de la résolution de la première équation q -linéaire liée au calcul des interactions supérieures (Proposition 3.8).

Proposition A.3 *Soit une équation q -linéaire (E)*

$$\sum_{i=0}^I \sum_{j=0}^J e_{i,j} y^j H(q^i y). \quad (182)$$

où $e_{i,j} \in \mathbb{A}(q)$. Sous réserve que l'on puisse trouver tous les diviseurs des polynômes rencontrés, il existe un algorithme déterminant les solutions hypergéométriques basiques de l'équation q -linéaire (182).

Preuve.

On cherche la raison $r(X)$ d'une série hypergéométrique basique $H(y) = \sum_{n \geq 0} h_n y^n$ solution de l'équation q -linéaire (182) avec $h_{n+1}/h_n = r(q^n)$. Le caractère hypergéométrique basique de la série permet de factoriser h_n dans le coefficient de y^{n+J} de l'équation q -linéaire :

$$0 = [y^{n+J}] \left(\sum_{i=0}^I \sum_{j=0}^J e_{i,j} y^j H(q^i y) \right) = \sum_{i=0}^I \sum_{j=0}^J e_{i,j} h_{n+J-j} q^{i(n+J-j)} \quad (183)$$

$$= \sum_{i=0}^I \sum_{j=0}^J e_{i,j} \left(h_n \prod_{k=0}^{J-j-1} r(q^{n+k}) \right) q^{i(n+J-j)} = 0. \quad (184)$$

On introduit la fraction rationnelle $F(X)$ suivante :

$$F(X) := \sum_{i=0}^I \sum_{j=0}^J e_{i,j} \left(\prod_{k=0}^{J-j-1} r(q^k X) \right) (q^{J-j} X)^i. \quad (185)$$

La relation (184) implique que pour tout $n \in \mathbb{N}$ on a

$$h_n F(q^n) = 0. \quad (186)$$

On ne s'intéresse qu'aux solutions $H(y)$ non-polynomiales. Les solutions polynomiales sont traitées par Abramov, Petkovšek et Paule mais ce ne sont pas celles qui sont recherchées dans la section 3.3. D'après cette hypothèse il existe une infinité de h_n non-nuls. Si une telle solution existe $F(X)$ s'annule en q^n pour une infinité de n . Par hypothèse q est transcendant donc tous les q^n sont des racines distinctes de $F(X)$. La fraction rationnelle $F(X)$ possédant une infinité de racines distinctes, on en déduit qu'elle est nulle.

Dans l'expression (185) de $F(X)$, on écrit $r(X)$ sous sa forme normale $z \frac{a(X) c(qX)}{b(X) c(X)}$, on simplifie des termes de la forme $\frac{c(q^k X)}{c(q^k X)}$, on réduit au même dénominateur et on extrait le numérateur $P(X)$. On obtient :

$$P(X) := \sum_{j=0}^J \left(\sum_{i=0}^I e_{i,j} (q^{J-j} X)^i \right) \left(\prod_{k=0}^{J-j-1} z a(q^k X) \right) c(q^{J-j} X) \left(\prod_{k=J-j}^{J-1} b(q^k X) \right) = 0. \quad (187)$$

Grâce aux propriétés de la forme normale on va déduire de l'annulation de $P(X)$ des informations sur $a(X)$ et $b(X)$. On isole dans $P(X)$ les termes qui n'admettent pas $a(X)$

Soient $e_{i,j}$ les coefficients de l'équation $\sum_{i=0}^I \sum_{j=0}^J e_{i,j} y^j H(q^i y) = 0$;

ResultatAnalyse := \emptyset ;

ContrainteA(X) := $\sum_{i=0}^I e_{i,J} X^i$; *ContrainteB*(X) := $\sum_{i=0}^I e_{i,0} (qX)^i$;

/* Grâce à l'écriture sous la forme normale on prouve que $a(X)$ divise *ContrainteA*(X) et que $b(X)$ divise *ContrainteB*(X). */ ;

pour tout $a(X)$ diviseur de *ContrainteA*(X) et $b(X)$ de *ContrainteB*(X) **faire**

$$P_{a,b}(X, Z, C(X)) := \sum_{j=0}^J \left(\sum_{i=0}^I e_{i,j} (q^{J-j} X)^i \right) \left(\prod_{k=0}^{J-j-1} Z a(q^k X) \right) c(q^{J-j} X) \left(\prod_{k=J-j}^{J-1} b(q^k X) \right);$$

$$P_{a,b}(X, Z, C(X)) := P_{A,B}(X, Z, C(X)) / a(X) / b(q^{J-1} X);$$

Soient $f_{k,j}(Z)$ les coefficients de $P_{A,B}(X, Z, C(X))$ sous la forme

$$\sum_{j=0}^J \left(\sum_{k=L}^K f_{k,j}(Z) X^k \right) C(q^j X);$$

$$\textit{AnnulationDegreMin}(Z) := \sum_{j=0}^J f_{L,j}(Z);$$

$$z := \textit{solution}(\textit{AnnulationDegreMin}(Z) = 0);$$

$$\textit{AnnulationDegreMax}(Z, W) := \sum_{j=0}^J f_{K,j}(Z) W^j;$$

$$\textit{ContrainteDegreC}(W) :=$$

$$\textit{Resultant}(\textit{AnnulationDegreMin}(Z), \textit{AnnulationDegreMax}(Z, W), Z);$$

si il n'existe pas de $n \in \mathbb{N}$ tel que $\textit{ContrainteDegreC}(q^n) = 0$ **alors**

Les choix de $a(X)$ et $b(X)$ ne donnent pas de solution.

sinon

pour tout entier n tel que $\textit{ContrainteDegreC}(q^n) = 0$ **faire**

$$\textit{ContrainteC}(X) := \sum_{i=0}^{n-1} C_i X^i + X^n;$$

Soit $(c_i)_{i=0..n-1}$ les éventuelles solutions du système d'équation "quasi-linéaire" d'inconnue $\{Z, (C_i)_{i=0..n-1}\}$ obtenu par l'annulation des coefficients de X^k du polynôme $P_{A,B}(X, Z, \textit{ContrainteC}(X))$. ;

$$c(X) := \sum_{i=0}^{n-1} c_i X^i + X^n ;$$

$$r(X) := z \frac{a(X) c(qX)}{b(X) c(X)};$$

$$\textit{ResultatAnalyse} := \textit{ResultatAnalyse} \cup \{r(X)\};$$

fin

fin

fin

retourner *ResultatAnalyse*

comme facteur (c'est-à-dire le cas $j = J$) :

$$P(X) = \left(\sum_{i=0}^I e_{i,J} X^i \right) c(X) \left(\prod_{k=0}^{J-1} b(q^k X) \right) \\ + a(X) z \sum_{j=0}^{J-1} \left(\sum_{i=0}^I e_{i,j} (q^{J-j} X)^i \right) \left(\prod_{k=1}^{J-j-1} z a(q^k X) \right) c(q^{J-j} X) \left(\prod_{k=J-j}^{J-1} b(q^k X) \right) = 0.$$

D'après la forme normale, $a(X)$ est premier avec $c(X)$ et tous les $b(q^k X)$. On en déduit que $a(X)$ divise $\sum_{i=0}^I e_{i,J} X^i$, un polynôme noté *ContrainteA*(X) dans l'algorithme. Un raisonnement similaire isolant le terme $b(q^{J-1} X)$ prouve que $b(X)$ divise le polynôme *ContrainteB*(X) $\equiv \sum_{i=0}^I e_{i,0} (qX)^i$.

L'algorithme considère ensuite tous les choix possibles pour $a(X)$ et $b(X)$ qui sont bien en nombre fini. Dans le cas général, on ne sait pas comment calculer tous les diviseurs d'un polynôme. Pour la suite de cette preuve on fixe un choix pour $a(X)$ et $b(X)$. Par choix des polynômes, $a(X)$ et $b(q^{J-1} X)$ divise le polynôme $P(X)$. La connaissance de ces deux polynômes permet d'écrire le polynôme $P(X)/a(X)/b(q^{J-1} X)$ sous la forme

$$P_{a,b}(X) = \sum_{j=0}^J \left(\sum_{k=L}^K f_{k,j}(Z) X^k \right) c(q^j X) \quad (188)$$

où les $f_{k,j}(Z)$ sont des polynômes en Z explicites une fois que le choix de $a(X)$ et $b(X)$ est fait. On reconnaît le polynôme $P_{A,B}(X, Z, C(X))$ de l'algorithme. Si on sait que ce polynôme est nul, on ne connaît pas à ce stade le polynôme $c(X)$, ni même son degré.

Le coefficient de plus petit degré en X de $P_{A,B}(X, Z, C(X))$ est $\sum_{j=0}^J f_{L,j}(Z) C(0)$. D'après les propriétés de la forme normal $C(0)$ est non nul on en déduit l'annulation du polynôme

$$\text{AnnulationDegrMin}(Z) \equiv \sum_{j=0}^J f_{L,j}(Z).$$

La constante z est une des racines de ce polynôme en Z .

Supposons que le degré de $C(X)$ soit n , d'après la forme normale le coefficient dominant de ce polynôme est 1. Le coefficient de plus grand degré en X de $P_{A,B}(X, Z, C(X))$ est $\sum_{j=0}^J f_{K,j}(Z) q^{nj}$. En posant $W = q^n$ on en déduit l'annulation du polynôme

$$\text{AnnulationDegrMax}(Z, W) = \sum_{j=0}^J f_{K,j}(Z) W^j.$$

Pour éviter de travailler avec la valeur algébrique de Z on peut utiliser le résultant selon l'inconnue Z de $\text{AnnulationDegrMin}(Z)$ et $\text{AnnulationDegrMax}(Z, W)$ pour obtenir un polynôme en la seule inconnue W . On cherche les racines de ce polynôme de la forme

$W = q^n$ en utilisant le lemme A.2.2. Ainsi on trouve un nombre fini de degrés possibles pour $C(X)$.

Pour chaque degré n on écrit $C(X)$ sous la forme $C(X) = \sum_{i=0}^{n-1} c_i X^i + X^n$. L'annulation de chacun des coefficients de degré non extrémal de $P_{A,B}(X, Z, C(X))$ conduit à un système d'équations linéaires en les variables $(c_i)_{0 \leq i \leq n-1}$. Il y a au moins $n - 1$ équations et on peut résoudre ce système en utilisant l'algorithme de Gauss. La seule difficulté particulière est le calcul de l'inverse du pivot car ce calcul doit se faire en utilisant la valeur algébrique z .

Pour un choix de $a(X)$ et $b(X)$ puis de z , du degré de $c(X)$ et enfin des coefficients de $c(X)$ on peut éventuellement avoir une fraction rationnelle $r(X)$ qui soit la raison d'une série hypergéométrique solution de l'équation q -linéaire considérée. \square

Bibliographie

- [AB03] J.-C. Aval and N. Bergeron. Catalan paths and quasi-symmetric functions. *Proc. Amer. Math. Soc.*, 131(4) :1053–1062 (electronic), 2003.
- [And86] G. E. Andrews. *q-series : their development and application in analysis, number theory, combinatorics, physics, and computer algebra*, volume 66 of *CBMS Regional Conference Series in Mathematics*. Published for the Conference Board of the Mathematical Sciences, Washington, DC, 1986.
- [APP98] S. A. Abramov, P. Paule, and M. Petkovšek. q -hypergeometric solutions of q -difference equations. *Discrete Math.*, 180(1-3) :3–22, 1998.
- [Bax82] R. J. Baxter. *Exactly solved models in statistical mechanics*. Academic Press Inc. [Harcourt Brace Jovanovich Publishers], London, 1982.
- [BDLPP98] E. Barucci, A. Del Lungo, E. Pergola, and R. Pinzani. A methodology for plane tree enumeration. In *Proceedings of the 7th Conference on Formal Power Series and Algebraic Combinatorics (Noisy-le-Grand, 1995)*, volume 180, pages 45–64, 1998.
- [BDLPP99] E. Barucci, A. Del Lungo, E. Pergola, and R. Pinzani. ECO : a methodology for the enumeration of combinatorial objects. *J. Differ. Equations Appl.*, 5(4-5) :435–490, 1999.
- [Big99] N. Biggs. The Tutte polynomial as a growth function. *J. Algebraic Combin.*, 10(2) :115–133, 1999.
- [BLBR] C. Banderier, J.-M. Le Bars, and V. Ravelomanana. Generating functions for kernels of digraphs (enumeration and asymptotics for nim games).
- [BLS91] A. Björner, L. Lovász, and P. W. Shor. Chip-firing games on graphs. *European J. Combin.*, 12(4) :283–291, 1991.
- [BM96] M. Bousquet-Mélou. Percolation models and animals. *European J. Combin.*, 17(4) :343–369, 1996.
- [BM98] M. Bousquet-Mélou. New enumerative results on two-dimensional directed animals. In *Proceedings of the 7th Conference on Formal Power Series and Algebraic Combinatorics (Noisy-le-Grand, 1995)*, volume 180, pages 73–106, 1998.

- [BMP00] M. Bousquet-Mélou and M. Petkovšek. Linear recurrences with constant coefficients : the multivariate case. *Discrete Math.*, 225(1-3) :51–75, 2000. Formal power series and algebraic combinatorics (Toronto, ON, 1998).
- [BMR03] M. Bousquet-Mélou and A. Rechnitzer. The site-perimeter of bargraphs. *Adv. in Appl. Math.*, 31(1) :86–112, 2003.
- [BMV92] M. Bousquet-Mélou and X. G. Viennot. Empilements de segments et q -énumération de polyominos convexes dirigés. *J. Combin. Theory Ser. A*, 60(2) :196–224, 1992.
- [BP93] J. Bétréma and J.-G. Penaud. Animaux et arbres guingois. *Theoret. Comput. Sci.*, 117(1-2) :67–89, 1993. Conference on Formal Power Series and Algebraic Combinatorics (Bordeaux, 1991).
- [BP95] R. Brak and T. Prellberg. Critical exponents from nonlinear functional equations for partially directed cluster models. *J. of Stat. Phys.*, 78(3-4) :701–730, 1995.
- [BTW88] P. Bak, C. Tang, and K. Wiesenfeld. Self-organized criticality. *Phys. Rev. A* (3), 38(1) :364–374, 1988.
- [BW97] N. Biggs and P. Winkler. Chip-firing and the chromatic polynomial. Technical report, CDAM, Lodon School of Economics and political science, 1997.
- [CF69] P. Cartier and D. Foata. *Problèmes combinatoires de commutation et réarrangements*. Lecture Notes in Mathematics, No. 85. Springer-Verlag, Berlin, 1969.
- [CLB03] R. Cori and Y. Le Borgne. The sand-pile model and Tutte polynomials. *Adv. in Appl. Math.*, 30(1-2) :44–52, 2003. Formal power series and algebraic combinatorics (Scottsdale, AZ, 2001).
- [Con96] A. Conway. Some exact results for moments of 2D directed animals. *J. Phys. A*, 29(17) :5273–5283, 1996.
- [CR00] R. Cori and D. Rossin. On the sandpile group of dual graphs. *European J. Combin.*, 21(4) :447–459, 2000.
- [DFR04] E. Duchi, J.-M. Fedou, and S. Rinaldi. From object grammars to ECO systems. *Theoret. Comput. Sci.*, 314(1-2) :57–95, 2004.
- [Dha82] D. Dhar. Equivalence of the two-dimensional directed-site animal problem to Baxter’s hard-square lattice-gas model. *Phys. Rev. Lett.*, 49(14) :959–962, 1982.
- [Dha83] D. Dhar. Exact solution of a directed-site animals-enumeration problem in three dimensions. *Phys. Rev. Lett.*, 51(10) :853–856, 1983.
- [DM92] D. Dhar and S. N. Majumdar. Equivalence the Abelian sandpile model and the $q \rightarrow 0$ limit of the Potts model. *Physica A*, 185 :129–145, 1992.
- [DPB82] D. Dhar, M. K. Phani, and M. Barma. Enumeration of directed site animals on two-dimensional lattices. *J. Phys. A*, 15(6) :L279–L284, 1982.

- [DRSV95] D. Dhar, P. Ruelle, S. Sen, and D.-N. Verma. Algebraic aspects of abelian sandpile models. *J. Phys. A*, 28(4) :805–831, 1995.
- [DS95] A. Denise and R. Simion. Two combinatorial statistics on Dyck paths. *Discrete Math.*, 137(1-3) :155–176, 1995.
- [Duc98] P. Duchon. *Q-grammaires : un outil pour l'énumération*. PhD thesis, Université Bordeaux I, 1998.
- [Dut96] I. Dutour. *Grammaires d'objets : énumération, bijections et génération aléatoire*. PhD thesis, Université Bordeaux I, 1996.
- [DV84] M.-P. Delest and G. Viennot. Algebraic languages and polyominoes enumeration. *Theoret. Comput. Sci.*, 34(1-2) :169–206, 1984.
- [FS95] P. Flajolet and R. Sedgewick. *An Introduction to the Analysis of Algorithms*. Addison-Wesley Publishing Compagny, Inconnu, 1995.
- [FS04] P. Flajolet and R. Sedgewick. *Analytic Combinatorics*. Addison-Wesley Publishing Compagny, Inconnu, A paraitre 2004.
- [GBV88] 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.
- [GH96] A. M. Garsia and M. Haiman. A remarkable q, t -Catalan sequence and q -Lagrange inversion. *J. Algebraic Combin.*, 5(3) :191–244, 1996.
- [GV85] I. Gessel and G. Viennot. Binomial determinants, paths, and hook length formulae. *Adv. in Math.*, 58(3) :300–321, 1985.
- [Hag03] J. Haglund. Conjectured statistics for the q, t -Catalan numbers. *Adv. Math.*, 175(2) :319–334, 2003.
- [HL02] J. Haglund and N. A. Loehr. A conjectured combinatorial formula for the hilbert series of diagonal harmonics. In *Proceedings of the Conference on Formal Power Series and Algebraic Combinatorics (Melbourne, 2002)*, 2002. Preprint available at www.math.upenn.edu/~jhaglund/.
- [Hug96] B. D. Hughes. *Random walks and random environments. Vol. 2*. Oxford Science Publications. The Clarendon Press Oxford University Press, New York, 1996. Random environments.
- [JvR00a] E. J. Janse van Rensburg. Interacting columns : generating functions and scaling exponents. *J. Phys. A*, 33 :7541–7554, 2000.
- [JvR00b] E. J. Janse van Rensburg. *The statistical mechanics of interacting walks, polygons, animals and vesicles*, volume 18 of *Oxford Lecture Series in Mathematics and its Applications*. Oxford University Press, Oxford, 2000.
- [Knu00] D. E. Knuth. *Selected papers on analysis of algorithms*, volume 102 of *CSLI Lecture Notes*. CSLI Publications, Stanford, CA, 2000.
- [Lal93] J.-C. Lalanne. Sur une involution sur les chemins de Dyck. *Theoret. Comput. Sci.*, 117(1-2) :203–215, 1993. Conference on Formal Power Series and Algebraic Combinatorics (Bordeaux, 1991).

- [ML97] C. Merino López. Chip firing and the Tutte polynomial. *Ann. Comb.*, 1(3) :253–259, 1997.
- [MS93] N. Madras and G. Slade. *The self-avoiding walk*. Probability and its Applications. Birkhäuser Boston Inc., Boston, MA, 1993.
- [ONGJ01] W.P. Orrick, B. G. Nickel, A.J. Guttman, and I. Jensen. The susceptibility of the square lattice ising model : new developments. *J. Stat. Phys.*, 102 :795–841, 2001.
- [RDD⁺02] Rajesh R., Dhar D., Giri D., Kumar S., and Singh Y. Adsorption and collapse transitions in a linear polymer chain near an attractive wall. *Physical Review E*, 65 :056134, 2002.
- [SA92] D. Stauffer and A. Aharony. *Introduction to Percolation Theory*. Taylor and Francis, London, 1992.
- [Sha] L. Shapiro. Directed animals and motzkin paths. Communication personnelle (via Bousquet-Mélou).
- [Slo03] N. J. A. Sloane. The on-line encyclopedia of integer sequences. *Notices Amer. Math. Soc.*, 50(8) :912–915, 2003.
- [Sta80] R. P. Stanley. Differentiably finite power series. *European J. Combin.*, 1(2) :175–188, 1980.
- [Sta86] R. P. Stanley. *Enumerative combinatorics. Vol. I*. The Wadsworth & Brooks/Cole Mathematics Series. Wadsworth & Brooks/Cole Advanced Books & Software, Monterey, CA, 1986. With a foreword by Gian-Carlo Rota.
- [Sta99] R. P. Stanley. *Enumerative combinatorics. Vol. 2*, volume 62 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1999. With a foreword by Gian-Carlo Rota and appendix 1 by Sergey Fomin.
- [Tut54] W. T. Tutte. A contribution to the theory of chromatic polynomials. *Canadian J. Math.*, 6 :80–91, 1954.
- [Vie86] G. X. Viennot. Heaps of pieces. I. Basic definitions and combinatorial lemmas. In *Combinatoire énumérative (Montreal, Que., 1985/Quebec, Que., 1985)*, volume 1234 of *Lecture Notes in Math.*, pages 321–350. Springer, Berlin, 1986.
- [Zei00] D. Zeilberger. The umbral transfer-matrix method. I. Foundations. *J. Combin. Theory Ser. A*, 91(1-2) :451–463, 2000. In memory of Gian-Carlo Rota.