

N° d'ordre : 10/2010 - M/MT

**MINISTÈRE DE L'ENSEIGNEMENT SUPÉRIEUR ET DE LA
RECHERCHE SCIENTIFIQUE
UNIVERSITÉ DES SCIENCES ET DE LA TECHNOLOGIE
HOUARI BOUMEDIENE
FACULTÉ DE MATHÉMATIQUES**



MEMOIRE

Présenté pour l'obtention du diplôme de **MAGISTER**

EN : MATHÉMATIQUES

Spécialité : Recherche Opérationnelle (Génie Mathématiques)

Par : **ABOUD Ammar**

THÈME

**Nombres de Catalan généralisés
et interprétation combinatoire**

Soutenu publiquement, le 21/10/2010, devant le jury composé de :

M.	Benali BENZAGHOU	Professeur	à l'U.S.T.H.B	Président
M.	Hacène BELBACHIR	Maître de Conférences A	à l'U.S.T.H.B	Directeur de mémoire
Mme.	Isma BOUCHEMAKH	Professeur	à l'U.S.T.H.B	Examineur
M.	Sadek BOUROUBI	Professeur	à l'U.S.T.H.B	Examineur
M.	Miloud MIHOUBI	Maître de Conférences A	à l'U.S.T.H.B	Examineur

Table des matières

Introduction	1
1 Nombres de Catalan	3
1.1 Un petit historique des nombres de Catalan	4
1.2 Propriétés et quelques expressions alternatives	6
1.2.1 Les nombres de Catalan et la fonction Gamma	7
1.2.2 Propriétés de divisibilité des nombres de Catalan	10
1.2.3 Le triangle de Pascal et les nombres de Catalan	12
1.2.4 Le triangle de Catalan	15
1.2.5 La matrice de Hankel et les nombres de Catalan	15
2 Interprétations combinatoires	16
2.1 Méthodes d'énumération symbolique	16
2.1.1 Calcul des fonctions génératrices ordinaires avec la méthode symbolique	19
2.2 Triangulations d'un polygone	19
2.2.1 Triangulations d'un polygone et les nombres de Catalan	21
2.3 Parenthésages binaires	24
2.4 Chemins dans la grille du plan	25
2.4.1 Les chemins de Dyck	26
2.4.2 Chemins monotones et chemins monotones de Dyck	27
2.5 Mots de Dyck	33
2.5.1 Mots de Dyck et problème de ballottage	34
2.6 Arbres et nombres de Catalan	36

2.6.1	Définitions	36
2.6.2	Interprétations combinatoires	40
2.6.3	Enumération des arbres par la méthode symbolique	43
3	Nombres de Catalan généralisés	45
3.1	Les nombres k -Catalan	45
3.1.1	Interprétations combinatoires	46
3.2	Les nombres de Catalan multidimensionnels	51
3.2.1	Interprétations combinatoires	52
4	Convolution de Gould-Vandermonde et nombres de Catalan généralisés	60
4.1	Convolution de Gould-Vandermonde	60
4.2	Preuve combinatoire et généralisation de la convolution de Gould-Vandermonde	65
4.2.1	Généralisation multinomiale des nombres de Catalan	66
	Conclusion	72

Remerciements

Ce Mémoire, intitulé "Nombres de Catalan généralisés et interprétation combinatoire", s'est déroulé dans le cadre de l'Ecole Doctorale de Recherche Opérationnelle à l'USTHB durant l'année universitaire 2009/2010 sous la direction de Monsieur Belbachir Hacène Maître de Conférences en Mathématique.

C'est avec un sincère et vif enthousiasme que je voudrai rendre mérite à tous ceux qui m'ont aidé à mener à bien ce travail. Je désire alors exprimer ma profonde gratitude à mon Professeur et Directeur de mémoire Hacène Belbachir, pour m'avoir dirigé patiemment, pour sa disponibilité et son soutien constant. Son ardeur au travail, son ouverture d'esprit et sa bonne humeur ont su me donner une forte motivation et m'ont fait découvrir et aimer passionnément ce beau domaine de la combinatoire.

Je remercie le Professeur Benali Benzaghrou de m'avoir fait l'honneur de présider le jury de mon mémoire.

J'adresse mes remerciements aux Professeurs Isma Bouchemakh, Sadek Bouroubi et Miloud Mihoubi pour avoir accepté d'être examinateurs de mon mémoire.

Je tiens également à remercier tous les professeurs de l'Ecole Doctorale qui n'ont pas lésiné sur les moyens et les efforts pour nous dispenser des cours dans une ambiance intellectuelle formidable, ainsi qu'à tous mes camarades avec qui j'ai travaillé dans la joie et la bonne humeur.

Enfin, c'est avec une grande émotion que j'exprime ma profonde gratitude à ma famille, particulièrement mes chers parents, mon frère et mes sœurs, qui furent toujours présents et ont rempli ma vie, je leur suis infiniment reconnaissant pour leur amour et leurs soutiens.

Introduction

La suite naturelle des nombres de Catalan est l'une des suites les plus curieuses en combinatoire. La diversité des objets combinatoires où elle intervient rend son étude passionnante. Des surveys couvrant les nombres de Catalan ont été écrits, on pourra consulter les deux ouvrages remarquables de Koshy [13] et Stanley [19].

Le thème principal du mémoire concerne les interprétations combinatoires des nombres de Catalan et l'étude de la convolution de Gould-Vandermonde s'y rattachant.

Dans le premier chapitre nous introduisons un petit historique relatif à la suite de Catalan qui a vu sa naissance avec le problème des triangulations du polygone, étudié par Euler. Puis, nous présentons des propriétés remarquables et quelques expressions alternatives aux nombres de Catalan. En particulier, nous étudions les propriétés de divisibilité des nombres de Catalan, les nombres de Catalan et la fonction Gamma, différentes façons de déterminer les nombres de Catalan à partir du triangle de Pascal, le triangle de Catalan et enfin une relation entre la matrice de Hankel et les nombres de Catalan.

Le deuxième chapitre nous l'abordons par la méthode d'énumération symbolique. L'avantage de cette approche nous permet de retrouver la fonction génératrice ordinaire de quelques classes d'objets combinatoires (essentiellement celles définies récursivement) sans passer par la méthode analytique. Cette partie est largement puisée de l'ouvrage remarquable de Flajolet et Sedgewick [8]. Puis, nous entamerons l'étude des interprétations combinatoires par le problème des triangulations du polygone où on retrouve l'expression explicite des nombres de Catalan en calculant leur fonction génératrice ordinaire, d'abord analytiquement puis par la méthode symbolique. Nous définissons les parenthésages binaires en montrant qu'ils sont énumérés par les nombres de Catalan. Les chemins dans la grille du plan est un thème important de la combinatoire et là aussi, on retrouve les nombres de Catalan, liés à

une classe particulière de ces chemins qui est la classe des chemins de Dyck et les chemins monotones de Dyck, auxquels nous donnons une interprétation en termes de mots de Dyck. Enfin, nous définissons quelques classes d'arbres, les arbres binaires en particulier, énumérés par les nombres de Catalan.

Au troisième chapitre, nous passons aux généralisations. La première traite des nombres k -Catalan, avec leurs différentes interprétations combinatoires. Nous introduisons les suites k -Raney, qui sont une extension des mots de Dyck; les parenthésages k -aires et les arbres k -aires. La deuxième généralisation, concerne les nombres de Catalan multidimensionnels dont l'interprétation combinatoire principale se traduit dans les tableaux de Young. Là, nous proposons une bijection combinatoire entre les tableaux de Young à deux colonnes et le triangle de Catalan.

Le quatrième chapitre est consacré à la convolution de Gould-Vandermonde. Nous présentons et détaillons la preuve de Gould [10] puis, après avoir défini les forêts d'arbres k -aires, nous proposerons une preuve combinatoire de la convolution de Gould-Vandermonde. Cette dernière nous permet d'entrevoir une version multinomiale de la convolution de Gould-Vandermonde dont nous tenterons de donner une preuve combinatoire, à partir de la définition des nombres de Catalan de type vectoriel.

Ce modeste panorama sur les nombres de Catalan et les diverses facettes de les présenter et de les étudier nous a permis d'apprécier comment un édifice sur un concept donné est successivement pris en charge par des mathématiciens remarquables à travers les décennies (dont nous avons donné quelques brèves notes biographiques puisées de l'encyclopédie Wikipédia), ce qui nous force à l'humilité et au respect.

Chapitre 1

Nombres de Catalan

Parmi les suites de nombres les plus fascinantes et les plus célèbres en combinatoire on trouve la suite naturelle des *nombres de Catalan*. Ces nombres apparaissent de façon inattendue dans beaucoup de situations combinatoires. En fait, Richard P. Stanley [17, 19] a compilé une liste de quelques 165 différentes interprétations combinatoires. Le $n^{\text{ème}}$ nombre de Catalan est donné explicitement en termes de coefficients binomiaux par

$$C_n = \frac{1}{n+1} \binom{2n}{n} = \frac{(2n)!}{(n+1)!n!}. \quad (1)$$

Ces nombres, dont l'aire d'influence ne cesse de s'élargir, s'appellent comme tels en hommage au mathématicien Franco-belge Eugène Charles Catalan⁽¹⁾, qui a fait une longue

⁽¹⁾Le mathématicien Eugène Charles Catalan (1814-1894) est né et mort en Belgique où il passa une partie importante de sa vie, ayant notamment enseigné à l'Université de Liège. Catalan, obstinément fidèle à ses convictions scientifiques, idéologiques ou politiques, connut une carrière d'enseignant et de chercheur aussi riche que mouvementée. Dans la gamme étendue des problèmes mathématiques abordés par lui, son nom est resté attaché à des conjectures en théorie des nombres et, en géométrie combinatoire, à des entiers dont l'aire d'influence ne cesse de s'élargir. Il a côtoyé les principaux mathématiciens de son temps, avec bon nombre desquels il échangea un abondant courrier, exploré puis exploité par l'auteur, lui-même professeur émérite à l'Université de Liège.

étude sur ces nombres. Toutefois, ils sont apparus bien avant lui d'un siècle dans les travaux d'Euler⁽¹⁾ et Segner⁽²⁾ dans la deuxième moitié du 18^{ème} siècle.

On donnera ci-après une petite genèse de ces nombres.

1.1 Un petit historique des nombres de Catalan

Cette note historique a été tirée du livre de Flajolet et Sedgewick [8].

Le 4 Septembre 1751, Euler a écrit une lettre à son ami Goldbach⁽³⁾ :

"J'ai rencontré récemment une question, qui m'est apparue plutôt remarquable. Elle concerne le nombre de manières de décomposer un polygone [convexe] en triangles par des lignes diagonales."

Puis, Euler décrit le problème (pour un n -gone décomposé en $(n - 2)$ triangles par des diagonales non sécantes) et en conclut :

"Soit maintenant x le nombre de manières [...]. Je suis alors parvenu à conclure qu'en général

$$x = \frac{2.6.10.14\dots(4n - 10)}{2.3.4.5\dots(n - 1)}. \quad (2)$$

[...] Regardant la progression des nombres 1, 2, 5, 14, 42, 132, et ainsi de suite, j'ai aussi observé la propriété suivante :

$$1 + 2a + 5a^2 + 14a^3 + 42a^4 + 132a^5 + etc. = \frac{1 - 2a - \sqrt{1 - 4a}}{2aa}. \quad (3)$$

⁽¹⁾Leonhard Euler (1707-1783), mathématicien suisse, physicien, ingénieur et philosophe. Né à Bâle, Euler y suit à l'université les cours de Jean Bernoulli et obtient sa maîtrise à l'âge de seize ans. Euler a fait d'importantes découvertes dans le domaine du calcul infinitésimal et de la théorie des graphes.

⁽²⁾Johann Andreas Segner (1704-1777), mathématicien allemand, né à Bratislava, la capitale de la Slovaquie. Il a commencé ses études à l'université de Jena, et y a reçu le certificat d'études médicales en 1729. Intéressé par les mathématiques, il est retourné à l'université de Jena en 1732 pour un grade de Master en mathématiques. Un éminent scientifique de son temps, Segner a été membres des académies de Berlin, Londres et Saint Petersburg.

⁽³⁾Christian Goldbach (1690-1764), mathématicien Prusse dont le nom est collé à une fameuse conjecture dite la conjecture de Goldbach. Il a fait ses études à l'université de Königsberg. Il a côtoyé de célèbres mathématiciens, comme Gottfried Leibniz, Leonhard Euler, et Nicholas I Bernoulli.

Ainsi, en 1751 Euler connaissait déjà comment approcher un tel problème par ce qui sera appelé plus tard une *fonction génératrice*. Néanmoins, à partir de son écriture, il n'est pas tout à fait clair s'il a obtenu une preuve complète du problème.

Dans ses cours de l'année 1750, Euler a communiqué le problème à Segner, qui a écrit dans sa publication datant de 1758 ceci: "*Le grand Euler m'a bénévolement communiqué ces nombres; la façon avec laquelle il les a trouvés, et la loi de leur progression demeurent flous pour moi.*" En approchant cette progression par une récurrence, Segner a obtenu la formule remarquable suivante:

$$T_n = \sum_{k=0}^{n-1} T_k T_{n-1-k}, \quad T_0 = T_1 = 1. \quad (4)$$

Durant l'année 1830, Liouville⁽¹⁾ a fait circuler le problème et l'a envoyé à Lamé⁽²⁾, qui lui a répondu le lendemain, avec une preuve basée sur une récurrence similaire à la récurrence précédente, donnant l'expression explicite :

$$T_n = \frac{1}{n+1} \binom{2n}{n}. \quad (5)$$

Les notes de Lamé ont été publiées en 1838 dans le *Journal de mathématiques pures et appliquées* (Journal de Liouville).

Juste après cette publication, une profonde étude a été consacrée à ces nombres par Catalan, qui a observé que T_n intervient dans le nombre de manières de multiplier n nombres. Il a aussi prouvé la validité de la fonction génératrice donnée par Euler:

$$T(z) = \sum_n T_n z^n = \frac{1 - \sqrt{1 - 4z}}{2z}. \quad (6)$$

Et depuis, grâce à cette influente étude de Catalan, ces nombres portent son nom. Ils s'appellent **les nombres de Catalan**, et le $n^{\text{ème}}$ nombre de Catalan est noté par C_n .

On reviendra en détails sur cette interprétation des nombres de Catalan en termes de *triangulations*, ainsi que sur d'autres interprétations combinatoires.

⁽¹⁾Joseph Liouville (1809-1882), mathématicien analyste français qui établit l'existence des nombres transcendants. Né à Saint-Omer, Joseph Liouville fut professeur à l'école Polytechnique puis au Collège de France. Il fut également membre de l'Académie des sciences dans la section d'astronomie. En janvier 1836, il fonda le Journal des mathématiques pures et appliquées.

⁽²⁾Gabriel Lamé (1795-1870), mathématicien français, connu pour sa théorie générale sur les coordonnées curvilignes. Il est aussi connu pour son analyse de l'algorithme d'Euclide.

1.2 Propriétés et quelques expressions alternatives

Cette section est largement puisée des livres de Koshy [13] et de Stanley [19], sauf mention contraire.

Les premiers termes des nombres de Catalan, séquence A000108 dans Sloane [16], sont donnés par:

1, 1, 2, 5, 14, 42, 132, 429, 1430, 4862, 16796, 58786, 208012, 742900, 2674440, 9694845, 35357670, 129644790, 477638700, 1767263190, 6564120420, 24466267020, 91482563640, 343059613650, 1289904147324, 4861946401452,...

Derrière leur formule explicite pourtant toute simple, les nombres de Catalan cachent des propriétés remarquables. Nous verrons certaines d'entre elles tout au long de ce mémoire. D'autres formules explicites existent, et parfois chaque formule est liée à une interprétation combinatoire.

Une expression alternative pour C_n est la suivante :

$$C_n = \frac{1}{2n+1} \binom{2n+1}{n}, \quad (7)$$

dont l'interprétation combinatoire est dans les suites "2-Raney" qu'on verra dans le troisième chapitre.

On peut aussi l'écrire ainsi :

$$C_n = \binom{2n}{n} - \binom{2n}{n-1}. \quad (8)$$

Cette expression montre de manière évidente que C_n est un entier.

Il y a aussi cette expression qui 'pourrait' avoir une interprétation combinatoire intéressante :

$$C_n = \frac{\binom{2n}{2, 2, \dots, 2}}{\binom{2}{2} \binom{3}{2} \cdots \binom{n+1}{2}}. \quad (9)$$

Ils s'écrivent aussi [21]:

$$C_n = \frac{2^n (2n-1)!!}{(n+1)!}, \quad (10)$$

où la double factorielle "!!" désigne le produit de tout les nombres entiers impairs de 1 à $(2n - 1)$.

Les nombres de Catalan satisfont (et sont bien définis par) la *réurrence convolutionnelle* (4) qu'on peut réécrire comme suit [13] :

$$C_0 = 1 \text{ et } C_{n+1} = \sum_{k=0}^n C_k C_{n-k}, \quad n \geq 0. \quad (11)$$

C'est cette convolution qui permet de démontrer, pour beaucoup d'objets combinatoires, qu'ils sont comptés par les nombres de Catalan, comme on le verra dans le deuxième chapitre. On démontrera cette réurrence dans le problème des *triangulations d'un polygone*.

A partir de la convolution (11), on verra comment obtenir la *fonction génératrice ordinaire* des nombres de Catalan [13] :

$$C(z) = \sum_{n \geq 0} C_n z^n = \frac{1 - \sqrt{1 - 4z}}{2z}. \quad (12)$$

Nous avons aussi la relation réursive suivante :

$$C_0 = 1 \text{ et } C_{n+1} = \frac{2(2n + 1)}{n + 2} C_n, \quad (13)$$

Cette relation est plus efficiente pour calculer les nombres de Catalan.

1.2.1 Les nombres de Catalan et la fonction Gamma

La *fonction Gamma* Γ a été introduite par le mathématicien Adrien-Marie Legendre⁽¹⁾. Elle est définie par [13] :

$$\Gamma(z) = \int_0^{\infty} t^{z-1} e^{-t} dt, \quad (14)$$

⁽¹⁾Adrien-Marie Legendre, né le 18 septembre 1752 à Paris et mort le 9 janvier 1833 à Auteuil, est un mathématicien français. Il occupe la chaire de Mathématiques de l'école militaire de Paris de 1775 à 1780. Le 30 mars 1783, il devient membre de l'académie des sciences. Il fit d'importantes contributions à la statistique, à la théorie des nombres, aux algèbres abstraites et à l'analyse. Une grande partie de son travail fut perfectionné par d'autres : son travail sur les racines des polynômes inspira la théorie de Galois ; le travail d'Abel sur les fonctions elliptiques fut construit sur celui de Legendre ; certains travaux de Gauss en statistique et en théorie des nombres complétèrent ceux de Legendre.

où z est une variable complexe avec une partie réelle positive. En utilisant l'intégration par partie on établit

$$\Gamma(z) = (z - 1)\Gamma(z - 1). \quad (15)$$

Quand z est un entier positif, cela correspond à

$$\Gamma(n) = (n - 1)!. \quad (16)$$

Par la formule de réflexion d'Euler, on a

$$\Gamma(1 - z)\Gamma(z) = \frac{\pi}{\sin \pi z}, \quad (17)$$

d'où $\Gamma(1/2) = \sqrt{\pi}$.

De (15) on peut obtenir

$$\Gamma\left(n + \frac{1}{2}\right) = \frac{(2n)!}{4^n n!} \sqrt{\pi}. \quad (18)$$

Ainsi

$$\frac{\Gamma\left(n + \frac{1}{2}\right)}{\Gamma(n + 2)} = \frac{(2n)!}{4^n n!(n + 1)!} \sqrt{\pi} = \frac{C_n}{4^n} \sqrt{\pi}. \quad (19)$$

D'où

$$C_n = \frac{4^n \Gamma\left(n + \frac{1}{2}\right)}{\sqrt{\pi} \Gamma(n + 2)}. \quad (20)$$

Quelques expressions sommatoires relatives aux coefficients binomiaux [13]

$$C_n = \sum_{k=0}^{n-1} C_k 2^{n-2k-1} \binom{n-1}{2k}. \quad (21)$$

$$C_n = \frac{1}{n} \sum_{k=0}^{n-1} C_{n-k+1} \binom{2k+1}{k+1}. \quad (22)$$

$$C_n = \sum_{k=0}^n (-1)^k 2^{n-k} \binom{n}{k} \binom{k}{\lfloor k/2 \rfloor}. \quad (23)$$

$$C_n = \sum_{k=0}^{\lfloor n/2 \rfloor} \left[\frac{n-2k+1}{n-k+1} \binom{n}{n-k} \right]^2. \quad (24)$$

$$C_n = \sum_{p=0}^{\lfloor n/2 \rfloor} \frac{1}{3p+1} \binom{3p+1}{p} \binom{n+p}{3p}. \quad (25)$$

Un produit où intervient C_n

$$C_n = \frac{1}{(n+1)!} \prod_{k=1}^n (4k-2). \quad (26)$$

Une formule asymptotique de C_n peut être trouvée en utilisant l'approximation de Stirling de la factorielle, $n! \approx (n/e)^n \sqrt{2\pi n}$, ainsi

$$C_n \approx \frac{2^{2n}}{(n+1)\sqrt{n\pi}} \approx \frac{2^{2n}}{n\sqrt{n\pi}}, \quad (27)$$

dans le sens où le quotient du $n^{\text{ème}}$ nombre de Catalan avec cette expression tend vers 1 quand $n \rightarrow +\infty$.

La fonction génératrice exponentielle de C_n est donnée par

$$\sum_{n=0}^{+\infty} C_n \frac{z^n}{n!} = e^{2z} [I_0(2z) - I_1(2z)] = 1 + z + z^2 + \frac{5}{6}z^3 + \frac{7}{12}z^4 + \frac{7}{20}z^5 + \dots \quad (28)$$

où $I_n(z)$ est la *fonction de Bessel de première espèce modifiée* :

$$I_n(x) = i^{-n} J_n(iz), \text{ avec } J_n(z) = (z/2)^n \sum_{p=0}^{\infty} \frac{(-1)^p}{2^{2p} p! (n+p)!} z^{2p}. \quad (29)$$

Autres formules remarquables où intervient C_n [17] :

$$\sum_{n=1}^{+\infty} \frac{C_n}{4^n} = 1. \quad (30)$$

$$\sum_{k=0}^n C_{2k} C_{2(n-k)} = 4^n C_n. \quad (31)$$

Cette formule n'a toujours pas de preuve combinatoire comme le signale Stanley [17, Catalan Addendum].

La fonction génératrice ordinaire des nombres de Catalan vérifie l'équation suivante (voir équation (52)) :

$$C(x) = \frac{1}{1 - xC(x)}, \quad (32)$$

d'où l'expression de $C(x)$ avec une élégante *fraction continue* (voir P. Barry [3]) :

$$C(x) = \frac{1 - \sqrt{1 - 4x}}{2x} = \frac{1}{1 - \frac{x}{1 - \frac{x}{1 - \frac{x}{1 - \frac{x}{1 - \dots}}}}}. \quad (33)$$

1.2.2 Propriétés de divisibilité des nombres de Catalan

On présente quelques propriétés de divisibilité des nombres de Catalan. D'abord on identifie les nombres de Catalan qui sont impairs puis ceux qui sont premiers.

La parité des nombres de Catalan

Le tableau suivant présente les 18 premiers nombres de Catalan. Ceux qui sont impairs sont marqués par un astérisque.

n	0	1	2	3	4	5	6	7	8	9	10
C_n	1*	1*	2	5*	14	42	132	429*	1430	4862	16796

n	11	12	13	14	15	16	17
C_n	58786	208012	742900	2674440	9694845*	35357670	129644790

La remarque qu'on peut faire c'est que pour $n \leq 17$, C_n est impair pour $n = 0, 1, 3, 7$ et 15, et tous sont de la forme $2^m - 1$. Quand $m > 0$, les nombres $2^m - 1$ sont appelés *les nombres de Mersenne*.

Le théorème suivant identifie les nombres de Catalan impairs.

Proposition 1.2.1 (*Koshy et Salmasi, 2004 [13]*) *Pour $n > 0$, C_n est impair si et seulement si n est un nombre de Mersenne.*

Preuve. De la formule convolutionnelle (11) on a

$$C_n = 2(C_0C_{n-1} + C_1C_{n-2} + \dots + C_{\frac{n}{2}-1}C_{\frac{n}{2}}) \text{ si } n \text{ est pair;} \quad (34)$$

$$C_n = 2(C_0C_{n-1} + C_1C_{n-2} + \dots + C_{\frac{n-3}{2}}C_{\frac{n+1}{2}}) + C_{\frac{n-1}{2}}^2 \text{ si } n \text{ est impair.} \quad (35)$$

Ainsi, pour $n > 0$, C_n est impair si et seulement si n et $C_{\frac{n-1}{2}}$ sont impairs. Le même argument implique que C_n est impair si et seulement si $\frac{n-1}{2}$ et $C_{\frac{n-3}{4}}$ sont impairs, ou $\frac{n-1}{2} = 0$. On poursuivant cette descente finie il s'ensuit que C_n est impair si et seulement si $C_{\frac{n-(2^m-1)}{2^m}}$ est impair, $m \geq 1$. Mais comme la valeur minimale de k pour laquelle C_k est impair c'est $k = 0$, alors cette descente s'arrête quand $\frac{n-(2^m-1)}{2^m} = 0$, c'est à dire quand $n = 2^m - 1$, le nombre de Mersenne. ■

La primalité des nombres de Catalan

Revenons au tableau précédent. On peut faire une autre observation: exactement deux parmi ces nombres de Catalan sont premiers, C_2 et C_3 .

Le théorème suivant confirme cette observation.

Proposition 1.2.2 (*Koshy and Salmasi, 2004 [13]*) *Les seuls nombres de Catalan premiers sont C_2 et C_3 .*

Preuve. De la formule (13) on a $(n+2)C_{n+1} = (4n+2)C_n$. On suppose que C_n est premier pour un certain n . Alors si $n > 3$, $\frac{n+2}{C_n} < 1$; d'où $C_n > n+2$. Conséquence: $C_n \mid C_{n+1}$, donc $C_{n+1} = kC_n$ pour un certain entier positif k . Alors $4n+2 = k(n+2)$, d'où $1 \leq k \leq 3$ et $n \leq 4$. Alors les seuls nombres de Catalan premiers sont C_2 et C_3 . ■

Autres propriétés de divisibilité des nombres de Catalan

Proposition 1.2.3 (*Koshy and Salmasi, 2004 [13]*) *$C_n \mid (n+2)$ si et seulement si $n \leq 3$.*

Preuve. Si $n \leq 3$, alors on a clairement $C_n \mid (n+2)$.

D'autre part, soit $n > 3$. Alors, de la formule explicite on a

$$C_n = \frac{(n+2)(n+3)\dots(2n)}{n!}, \quad (36)$$

ainsi

$$\frac{n+2}{C_n} = \frac{3 \cdot 4 \cdot 5 \dots (n-1)}{(n+3)(n+4)\dots[n+(n-1)]}. \quad (37)$$

$\frac{n+2}{C_n}$ n'est donc pas un entier, d'où $C_n \mid (n+2)$ si et seulement si $n \leq 3$. ■

1	2	6	20	70	252	...
-		1	6	28	120	...
1	2	5	14	42	132	...

(1.2.4)

Ceci est justifié par le fait que

$$C_{n+1} = \binom{2n}{n} - \binom{2n}{n+2}, n \geq 1. \quad (42)$$

et

				1		↗			C_2										
				1		-1		↗		C_3									
			1		+2		-1		↗	C_4									
		1		3		-3		-1		↗	C_5								
	1		4		+6		-4		-1		↗	C_6							
		1	5		10		-10		-5		-1		↗						
			1	6		15		+20		-15		-6		1		↗			
				1	7		21		35		-35		-21		7		1		
1					8		28		56		+70		-56		28		8		1

(1.2.5)

$1=1$	
$1=2-1$	$1=1$
$2=6-3-1$	$2=3-1$
$5=20-10-4-1$	$5=10-4-1$
$14=70-35-15-5-1$	$14=35-15-5-1$
$42=252-126-56-21-6-1$	$42=126-56-21-6-1$

(1.2.6)

Justification :

En exploitant l'identité suivante (voir Knuth [11]) :

$$\sum_{k \leq m} \binom{r+k}{k} = \binom{r+m+1}{m} \text{ pour } r, m \in \mathbb{N}, \quad (43)$$

on obtient

$$\binom{2n}{n} - \sum_{i=1}^n \binom{2n-i}{n} = \binom{2n}{n} - \sum_{i=1}^n \binom{n+n-i}{n-i} = \binom{2n}{n} - \binom{2n}{n-1} = C_n. \quad (45)$$

1.2.4 Le triangle de Catalan

Le triangle de Catalan est défini comme suit (voir J-C. Aval [2]) :

$n \backslash k$	0	1	2	3	4	5
0	1					
1	1	1				
2	1	2	2			
3	1	3	5	5		
4	1	4	9	14	14	
5	1	5	14	28	42	42

(Sloane A0009766 [16]) dont les éléments sont donnés par :

$$c_{n,k} = \frac{(n+k)! (n-k+1)}{k!(n+1)!} \text{ pour } 0 \leq k \leq n. \quad (46)$$

Chaque élément est égal à la somme de l'élément qui est directement au dessus et son voisin à gauche. La somme sur chaque ligne est égale au dernier élément de la ligne suivante et qui est égal à C_n . En somme : $c_{n,n} = C_n$.

1.2.5 La matrice de Hankel et les nombres de Catalan

Soit H_n la *matrice de Hankel* construite à partir de la suite de Catalan (voir C. Radoux [14]) :

$$H_n = \begin{bmatrix} 1 & 1 & 2 & 5 & 14 & \dots & C_n \\ 1 & 2 & 5 & 14 & 42 & \dots & C_{n+1} \\ 2 & 5 & 14 & 42 & 132 & \dots & C_{n+2} \\ 5 & 14 & 42 & 132 & 429 & \dots & C_{n+3} \\ 14 & 42 & 132 & 429 & 1430 & \dots & C_{n+4} \\ \dots & \dots & \dots & \dots & \dots & \ddots & \dots \\ C_n & C_{n+1} & C_{n+2} & C_{n+3} & C_{n+4} & \dots & C_{2n} \end{bmatrix}, \quad (47)$$

alors quel que soit n , $\det(H_n) = 1$.

Chapitre 2

Interprétations combinatoires

L'essence des nombres de Catalan est combinatoire. Ces nombres, dont la littérature est très vaste, apparaissent dans des objets combinatoires qui n'ont, *à priori*, rien de commun. Les différentes formules et propriétés (formules récurrentes, fonction génératrice,...) citées dans le chapitre premier découlent parfois de l'interprétation combinatoire.

Dans ce chapitre, on va aborder les principales interprétations combinatoires des nombres de Catalan connues dans la littérature.

Mais d'abord, comme ce chapitre est principalement consacré à l'*énumération*; dans le sens où le problème est de déterminer le nombre de *configurations combinatoires* décrites par un nombre fini de règles, il est nécessaire d'introduire une *approche symbolique à l'énumération combinatoire*.

2.1 Méthodes d'énumération symbolique

La combinatoire énumérative est l'art de compter des objets. Des objets qui peuvent être décrits par un nombre fini de règles de constructions. Exemples : les mots, les arbres, les graphes, les permutations, les fonctions d'un ensemble fini vers lui même, les configurations topologiques,...etc. La question est d'*énumérer* ces objets en tenant compte de certains paramètres caractéristiques.

Les deux premiers paragraphes sont largement puisés du livre de P. Flajolet et R. Sedgewick [8].

Définition 2.1.1 Une classe combinatoire \mathcal{A} est un ensemble fini ou dénombrable d'objets sur lequel une fonction est définie, satisfaisant les conditions suivantes :

- (i) l'image de chaque objet est un entier positif;
- (ii) l'image réciproque de chaque entier positif est finie.

On appelle une telle fonction une fonction de taille.

Si \mathcal{A} est une classe, la taille d'un élément $\alpha \in \mathcal{A}$ est noté par $|\alpha|$, ou $|\alpha|_{\mathcal{A}}$. Pour une classe donnée \mathcal{A} , on note par \mathcal{A}_n l'ensemble des objets dans \mathcal{A} qui ont la taille n . On note par A_n le cardinal de \mathcal{A}_n ; i.e. $A_n = \text{card}(\mathcal{A}_n)$ (ou $a_n = \text{card}(\mathcal{A}_n)$). Une présentation axiomatique est la suivante : une classe combinatoire est une paire $(\mathcal{A}, |\cdot|)$ où \mathcal{A} est au plus dénombrable et l'application $|\cdot| \in (\mathcal{A} \mapsto \mathbb{N})$ est telle que l'image inverse de chaque entier est finie.

On note par ϵ un objet de taille nulle.

Définition 2.1.2 La suite d'énumération (ou de comptage) d'une classe combinatoire est la suite d'entiers $(A_n)_{n \geq 0}$ où $A_n = \text{card}(\mathcal{A}_n)$ est le nombre d'objets dans la classe \mathcal{A} qui ont la taille n .

Exemple 2.1.1 Considérons l'ensemble \mathcal{W} des mots binaires, qui sont les suites d'éléments pris à partir d'un alphabet binaire $\mathcal{A} = \{0, 1\}$.

$$\mathcal{W} = \{\epsilon, 0, 1, 00, 01, 10, 11, 000, 001, 010, \dots, 1001101, \dots\},$$

tel que ϵ est le mot vide. On définit la taille comme le nombre de lettres que contient un mot. Il y a deux possibilités pour chaque lettre et donc, au total, la séquence d'énumération (W_n) satisfait

$$W_n = 2^n.$$

Exemple 2.1.2 Une permutation de taille n est par définition une application bijective de l'intervalle d'entiers $\mathcal{I}_n = \{1, \dots, n\}$ sur lui même. Elle peut être représentée par un tableau ou à une séquence $\sigma_1 \sigma_2 \dots \sigma_n$ de ses éléments distincts:

$$\begin{pmatrix} 1 & 2 & \dots & n \\ \sigma_1 & \sigma_2 & \dots & \sigma_n \end{pmatrix},$$

La classe \mathcal{P} des permutations est donnée par

$$\mathcal{P} = \{1, 12, 21, 123, 132, 213, 231, 312, 321, 1234, \dots, 532614, \dots\}.$$

Pour une taille donnée n , il y a $P_n = n!$ permutations possibles.

Définition 2.1.3 Deux classes combinatoires \mathcal{A} et \mathcal{B} sont dites (combinatoirement) isomorphes, qu'on note $\mathcal{A} \cong \mathcal{B}$, si leurs suites d'énumération sont identiques. Cette condition est équivalente à l'existence d'une bijection de \mathcal{A} vers \mathcal{B} qui préserve la taille.

Définition 2.1.4 La fonction génératrice ordinaire (FGO) d'une suite (A_n) est la série formelle

$$A(z) = \sum_{n=0}^{\infty} A_n z^n. \quad (48)$$

La fonction génératrice ordinaire (FGO) d'une classe combinatoire \mathcal{A} est la fonction génératrice de la suite $A_n = \text{card}(\mathcal{A}_n)$.

De façon équivalente, la FGO d'une classe \mathcal{A} admet l'écriture combinatoire

$$A(z) = \sum_{\alpha \in \mathcal{A}} z^{|\alpha|}. \quad (49)$$

Cette écriture combinatoire d'une FGO résulte du fait que le terme z^n apparaît autant de fois que le nombre d'objets dans \mathcal{A} qui sont de taille n .

En fait, toutes les manipulations algébriques basiques restent vraies pour les fonctions génératrices. La chose la plus importante à retenir est que, pour deux séries formelles $A(z)$ et $B(z)$ données, si $A(z) = B(z)$, alors $A_i = B_i$ pour tout entier positif i . Les énoncés suivants sont aussi vrais :

- 1) Si $A(z) + B(z) = C(z)$, alors $C_i = A_i + B_i$.
- 2) Si $A(z) \cdot B(z) = C(z)$, alors $C_i = \sum_{n=0}^i A_n B_{i-n}$.
- 3) Si $A(z) = B(z)$, alors leurs dérivées et leurs intégrales sont aussi égales, ainsi que toute autre opération algébrique est conservée.

4) Une fonction génératrice ordinaire a un inverse multiplicatif si son terme constant n'est pas nul.

- 5) Si $A(z) \cdot B(z) = 0$, alors soit $A(z) = 0$, soit $B(z) = 0$.

2.1.1 Calcul des fonctions génératrices ordinaires avec la méthode symbolique

Par la méthode symbolique, il est possible de trouver directement une expression de la FGO d'une classe combinatoire.

Définition 2.1.5 Soient \mathcal{A} et \mathcal{B} deux classes combinatoires.

-Si \mathcal{A} et \mathcal{B} sont disjointes, alors on définit l'opération $\mathcal{A} + \mathcal{B}$ comme l'union disjointe des objets de \mathcal{A} et \mathcal{B} .

-On définit l'opération $\mathcal{A} \times \mathcal{B}$ comme le produit cartésien des deux ensembles.

Théorème 2.1.1 Soient \mathcal{A} et \mathcal{B} deux classes combinatoires. Si $A(z)$ est la FGO de \mathcal{A} et $B(z)$ la FGO de \mathcal{B} , alors:

- $A(z) + B(z)$ est la FGO de $\mathcal{A} + \mathcal{B}$.

- $A(z)B(z)$ est la FGO de $\mathcal{A} \times \mathcal{B}$.

Preuve. - Si A_n est le nombre d'objets de taille n dans \mathcal{A} , et B_n le nombre d'objets de taille n dans \mathcal{B} , alors $A_n + B_n$ est le nombre d'objets de taille n dans $\mathcal{A} + \mathcal{B}$.

- Pour tout k de 0 à n , on peut coupler chaque objet de \mathcal{A} de taille k avec un objet de \mathcal{B} de taille $n - k$ pour obtenir un objet de taille n de $\mathcal{A} \times \mathcal{B}$. Ainsi, le nombre d'objets de taille n dans $\mathcal{A} \times \mathcal{B}$ est égal à

$$\sum_{0 \leq k \leq n} A_k B_{n-k}. \quad (50)$$

■

Maintenant, en tenant compte de cette symbolique, on peut aborder les différentes interprétations combinatoires des nombres de Catalan.

2.2 Triangulations d'un polygone

Comme cité dans l'historique, la première manifestation des nombres de Catalan était dans le problème des triangulations du polygone abordé par *Euler*.

Définition 2.2.1 On appelle un $(n+2)$ -gone un polygone convexe à $n+2$ sommets ordonnés et arrangés dans le sens contraire des aiguilles d'une montre, conventionnellement numérotés de 0 à $n+1$.

Définition 2.2.2 Une triangulation d'un $(n + 2)$ -gone est une décomposition maximale du convexe $(n + 2)$ -gone en n triangles par des diagonales non sécantes (nécessairement $n - 1$ diagonales).

La taille d'une triangulation d'un $(n + 2)$ -gone est définie par le nombre de triangles qui est n . Une taille n définit bien une triangulation d'un $(n + 2)$ -gone.

On note par \mathcal{T} la classe des triangulations et par \mathcal{T}_n l'ensemble des triangulations dans \mathcal{T} qui sont de taille n .

Exemple 2.2.1 Voici un exemple d'une triangulation d'un 5-gone (un pentagone), donc de taille 3, qui se triangule en T_3 manières, $T_3 = 5$:

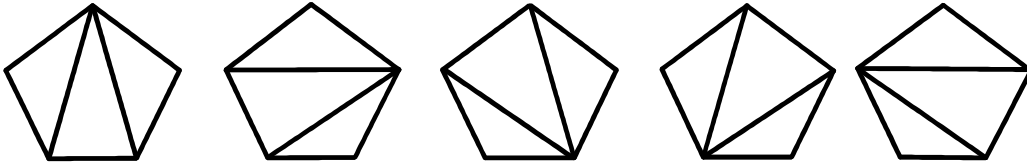


Fig.1: Triangulation d'un 5 - gone.

Remarque 2.2.1 Les triangulations sont considérées dans ce contexte comme des configurations topologiques invariantes par une déformation continue du plan.

Proposition 2.2.1 [8] Le nombre T_n de triangulations de taille n vérifie la récurrence convolutionnelle suivante :

$$T_0 = 1 \quad \text{et} \quad T_n = \sum_{k=0}^{n-1} T_k T_{n-k-1} \quad \text{pour } n \geq 0.$$

Preuve. On appelle la base du $(n + 2)$ -gone le segment dont les extrémités sont les sommets 0 et 1.

Pour chaque triangulation, la base est nécessairement un coté d'un triangle qu'on appelle le triangle de base.

Le triangle de base divise le $(n + 2)$ -gone en deux sous-polygones. Le premier est celui qui se trouve dans le sens positif. On suppose qu'il contient $k + 2$ cotés (un $(k + 2)$ -gone).

Donc il a k triangles dans sa triangulation. De l'autre coté il y a le deuxième sous-polygone, qui a nécessairement $n - k + 1$ cotés (un $(n - k + 1)$ -gone), qui est triangulé en $n - k + 1$ triangles. Ainsi, chaque $(n + 2)$ -gone est scindé de cette manière en un $(k + 2)$ -gone et un $(n - k + 1)$ -gone. Et comme k peut varier de 0 jusqu'à $n - 1$; et en admettant que $T_0 = 1$ (le 2-gone est un segment qui donne une triangulation vide de taille 0), on obtient la récurrence convolutionnelle voulue:

$$T_0 = 1 \quad \text{et} \quad T_n = \sum_{k=0}^{n-1} T_k T_{n-k-1} \quad \text{pour } n \geq 0.$$

■

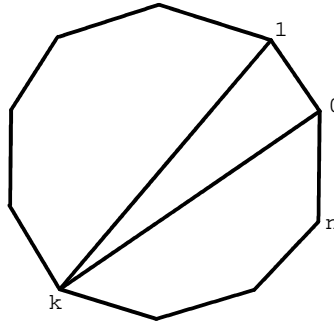


Fig.2 : Décomposition
d'un polygone.

2.2.1 Triangulations d'un polygone et les nombres de Catalan

Maintenant, de la récurrence convolutionnelle (2.1), il est facile de dégager la *fonction génératrice ordinaire* de la suite $(T_n)_{n \geq 0}$ ainsi que sa *formule explicite* qui correspond bien à la formule des nombres de Catalan.

Proposition 2.2.2 *Le nombre $T_n = \text{card}(\mathcal{T}_n)$ de triangulations de taille n est égal au $n^{\text{ème}}$ nombre de Catalan; i.e. $T_n = C_n, n \geq 0$.*

Preuve. Soit $T(z) = \sum_{n=0}^{\infty} T_n z^n$ la fonction génératrice ordinaire de $(T_n)_{n \geq 0}$.

Nous avons (à partir des propriétés des séries formelles) :

$$T(z)^2 = T_0 T_0 z^0 + (T_0 T_1 + T_1 T_0) z^1 + \dots + \left(\sum_{i=0}^n T_i T_{n-i} \right) z^n + \dots \quad (51)$$

Comme on sait que $T_0 = 1$ et $T_n = \sum_{k=0}^{n-1} T_k T_{n-k-1}$, pour $n \geq 0$,

il s'en suit que :

$$\begin{aligned} T(z)^2 &= T_1 z^0 + T_2 z^1 + \dots + T_{n+1} z^n + \dots \\ 1 + zT(z)^2 &= T_0 + T_1 z^1 + T_2 z^2 + \dots + T_n z^n + \dots \end{aligned}$$

On obtient alors l'équation quadratique suivante:

$$1 + zT(z)^2 = T(z). \quad (52)$$

En résolvant l'équation on obtient:

$$zT(z) = \frac{1}{2}(1 \pm \sqrt{1 - 4z}). \quad (53)$$

Mais pour obtenir l'égalité pour $z = 0$, on choisit le signe négatif, d'où

$$T(z) = \frac{1 - \sqrt{1 - 4z}}{2z}. \quad (54)$$

Pour développer $T(z)$, on a recours aux coefficients binomiaux étendus $\binom{m}{k}$, où $m \in \mathbb{R}$ et $k \in \mathbb{N}$:

$$\binom{m}{k} = \frac{m(m-1)\dots(m-k+1)}{k!}, \text{ si } k > 0, \text{ et } \binom{m}{0} = 1. \quad (55)$$

On utilise cette formule pour développer $\sqrt{1 - 4z}$ en une série de puissances de z . Ainsi:

$$\begin{aligned} \binom{\frac{1}{2}}{n} &= \frac{\frac{1}{2}(\frac{-1}{2})(\frac{-3}{2})\dots(\frac{3-2n}{2})}{n!} \\ &= \frac{(1)(-1)(-3)\dots(3-2n)}{2^n \cdot n!} \\ &= \frac{(-1)^{n-1}(2n-3)(2n-5)\dots(3)(1)}{2^n \cdot n!} \\ &= \frac{(-1)^{n-1}(2n-3)!!}{2^n \cdot n!}. \end{aligned}$$

Et donc

$$\begin{aligned}
 \sqrt{1-4z} &= (1-4z)^{1/2} \\
 &= 1 + \sum_{n \geq 1} \binom{\frac{1}{2}}{n} (-4z)^n \\
 &= 1 + \sum_{n \geq 1} \frac{(-1)^{n-1} (2n-3)!!}{2^n \cdot n!} (-4z)^n \\
 &= 1 - \sum_{n \geq 1} \frac{(2n-3)!! \cdot n! \cdot 2^n}{n! \cdot n!} z^n \\
 &= 1 - \sum_{n \geq 1} \frac{(2n-3)!! \cdot n \cdot 2 \cdot (2 \cdot 4 \cdot 6 \dots (2n-2))}{n! \cdot n!} z^n \\
 &= 1 - 2 \sum_{n \geq 1} \frac{(2n-2)!}{n! \cdot (n-1)!} z^n \\
 &= 1 - 2 \sum_{n \geq 1} \frac{\binom{2n-2}{n-1}}{n} z^n.
 \end{aligned}$$

Ainsi donc

$$\begin{aligned}
 T(z) &= \frac{1 - \sqrt{1-4z}}{2z} \\
 &= \frac{1 - (1 - 2 \sum_{n \geq 1} \frac{\binom{2n-2}{n-1}}{n} z^n)}{2z} \\
 &= \frac{2 \sum_{n \geq 1} \frac{\binom{2n-2}{n-1}}{n} z^n}{2z} \\
 &= \sum_{n \geq 1} \frac{\binom{2n-2}{n-1}}{n} z^{n-1} \\
 &= \sum_{n \geq 0} \frac{\binom{2n}{n}}{n+1} z^n.
 \end{aligned}$$

D'où $T_n = C_n$, $n \geq 0$. ■

Calcul de la fonction génératrice des triangulations par la méthode symbolique

Une triangulation est soit vide (pour le cas du 2-gone), soit composée du triangle de base et de sous-triangulations. Donc on peut représenter la classe \mathcal{T} des triangulations avec le

diagramme suivant [8] :

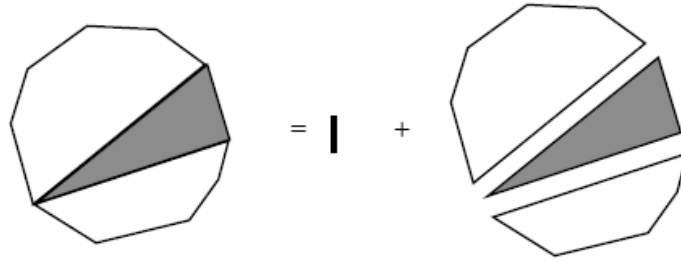


Fig.3: Diagramme spécifiant la classe \mathcal{T}

On peut alors spécifier la classe \mathcal{T} récursivement comme suit

$$\mathcal{T} = \{\epsilon\} + (\mathcal{T} \times \nabla \times \mathcal{T}),$$

où $\{\epsilon\}$ est la classe qui contient l'élément vide, dont la FGO est 1, et ∇ la classe qui contient un seul élément de taille 1 (le triangle), et dont la FGO est égale à z (ceci découle de l'écriture combinatoire de la fonction génératrice).

D'où, la FGO de \mathcal{T} satisfait l'équation

$$T(z) = 1 + zT(z)^2.$$

2.3 Parenthésages binaires

Une autre situation combinatoire où interviennent les nombres de Catalan concerne les *parenthésages binaires*. On se référera ici à l'ouvrage de Comtet [6].

Définition 2.3.1 *Un parenthésage binaire est une manière de parenthéser n lettres d'un ensemble muni d'une loi multiplicative non commutative et non associative, de telle sorte que, à chaque étape, on n'ait droit qu'à la multiplication de deux facteurs adjacents.*

Exemple 2.3.1 *pour le mot $\mu = abcd$, à 4 lettres, on a les 5 parenthésages binaires possibles suivants:*

$$(ab)(cd) \quad , \quad ((ab)c)d \quad , \quad (a(bc))d \quad , \quad a((bc)d) \quad , \quad a(b(cd)).$$

On note par \mathcal{P} la classe des parenthésages binaires, et par \mathcal{P}_n l'ensemble des parenthésages binaires dans \mathcal{P} à $n + 1$ lettres.

Proposition 2.3.1 *Le nombre de parenthésages binaires de $n + 1$ lettres est égal au $n^{\text{ème}}$ nombre de Catalan; i.e. $P_n = |\mathcal{P}_n| = C_n, n \geq 0$.*

Preuve. Pour chaque parenthésage de $n + 1$ lettres $a_1, a_2, \dots, a_n, a_{n+1}$, on considère la dernière multiplication, celle qui achève le calcul du produit. Elle unit $k + 1$ premières lettres aux $n - k$ dernières. Les $k + 1$ premières sont parenthésées de P_k manières et les $n - k$ dernières le sont de P_{n-k-1} manières. On prenant $P_0 = 1$, on retombe sur la formule de récurrence donnée par (4) qui caractérise les nombres de Catalan. ■

2.4 Chemins dans la grille du plan

L'énumération des chemins sur une grille de l'espace Euclidien est un thème important et classique en combinatoire. Beaucoup d'objets combinatoires (les arbres, les mots, les permutations, les cartes, les tableaux de Young...) peuvent être *codés* par des chemins *grillés*. D'où leur utilité dans des domaines qui sont parfois exogènes aux Mathématiques, comme par exemple en physique statistique.

Le problème de l'énumération des chemins se définit ainsi :

Définition 2.4.1 *Etant donné une grille hypercubique \mathbb{Z}^d , et un ensemble fini de pas $S \in \mathbb{Z}^d$, le problème est de déterminer le nombre de chemins à m pas pris dans S , commençant à partir de l'origine et confiné à une certaine région \mathcal{A} de l'espace.*

Plusieurs problèmes relatifs à cette classe restent encore ouverts :

Exemple 2.4.1 *Soit \mathcal{V} la classe des chemins dans la grille \mathbb{Z}^2 , prenant leurs pas dans l'ensemble $S = \{(0, 1), (1, 0), (-1, 0), (0, -1)\}$; qui partent de l'origine et reviennent à leur point de départ sans qu'ils se touchent eux même. La longueur d'un chemin de \mathcal{V} est son nombre de pas. Ces chemins sont appelés polyominoes simplement connexes.*

Combien y a-t-il de chemins dans la classe \mathcal{V}_n (chemins de longueur $2n$) ? Ce problème est encore ouvert.

Dans la suite, on va étudier deux classes importantes de chemins dans la grille \mathbb{Z}^2 , qui sont en rapport étroit avec les nombres de Catalan : *les chemins de Dyck* et *les chemins monotones de Dyck*.

2.4.1 Les chemins de Dyck

Définition 2.4.2 *Un chemin de Dyck de taille n est un chemin de longueur $2n$ (ou à $2n$ pas), sur la grille \mathbb{Z}^2 du plan euclidien, qui a pour point de départ $(0,0)$ et pour point d'arrivée $(2n,0)$, dont les pas sont pris d'un ensemble $S = \{(1,1), (1,-1)\}$, et reste confiné dans le premier quadrant $x \geq 0, y \geq 0$ du plan.*

Exemple 2.4.2 *Voici les chemins de Dyck de taille 0, 1, 2 et 3:*

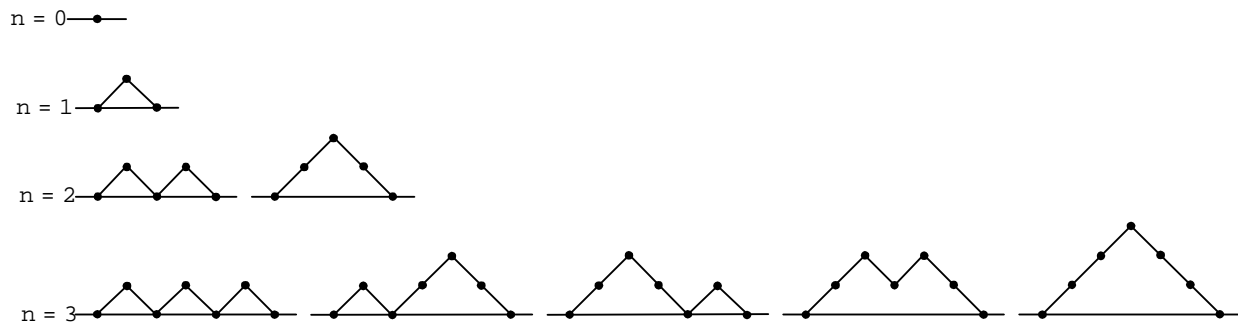


Fig.4: Chemins de Dyck $n = 1, 2, 3$.

Un chemin de Dyck est dit irréductible s'il ne touche l'axe des x qu'en $(0,0)$ et $(2n,0)$.

On note par \mathcal{D} la classe des chemins de Dyck, et par \mathcal{D}_n l'ensemble des chemins dans \mathcal{D} qui sont de taille n ; i.e $D_n := \text{card}(\mathcal{D}_n)$.

Lemme 2.4.1 *Le nombre de chemins de Dyck irréductible de taille k est égal à D_{k-1} .*

Preuve. Puisque un chemin de Dyck irréductible ne touche l'axe des abscisses qu'à ses points extrêmes, alors il suffit de lui supprimer le premier pas et le dernier pas, et de le translater selon le vecteur $\vec{\mu}(0, -1)$ pour obtenir un chemin de Dyck (ordinaire). ■

Proposition 2.4.1 *Le nombre de chemins de Dyck de taille n est égal au $n^{\text{ème}}$ nombre de Catalan; i.e. $D_n = C_n$ pour $n \geq 0$.*

Preuve. Chaque chemin de Dyck de taille n retourne, pour une première fois, sur l'axe des abscisses à un certain point $(2k, 0)$, $1 \leq k \leq n$. Donc on peut décomposer ce chemin de façon unique en deux chemins de Dyck : le premier chemin de $(0, 0)$ à $(2k, 0)$ qui est un chemin de Dyck irréductible de taille k ; et le deuxième chemin de $(2k, 0)$ à $(2n, 0)$ qui est un chemin de Dyck (ordinaire) de taille $n - k$. Le nombre de chemins de Dyck irréductibles de taille k est égal à D_{k-1} et le nombre de chemins de Dyck de taille $n - k$ est égal à D_{n-k} ; et comme k peut varier de 1 à n , il s'ensuit directement que:

$$D_n = \sum_{k=1}^n D_{k-1} D_{n-k},$$

et en prenant $D_0 = 1$, on obtient finalement la formule de récurrence convolutionnelle des nombres de Catalan. ■

La figure suivante représente la décomposition d'un chemin de Dyck de taille n en un chemin de Dyck irréductible de taille k et un chemin de Dyck de taille $n - k$.

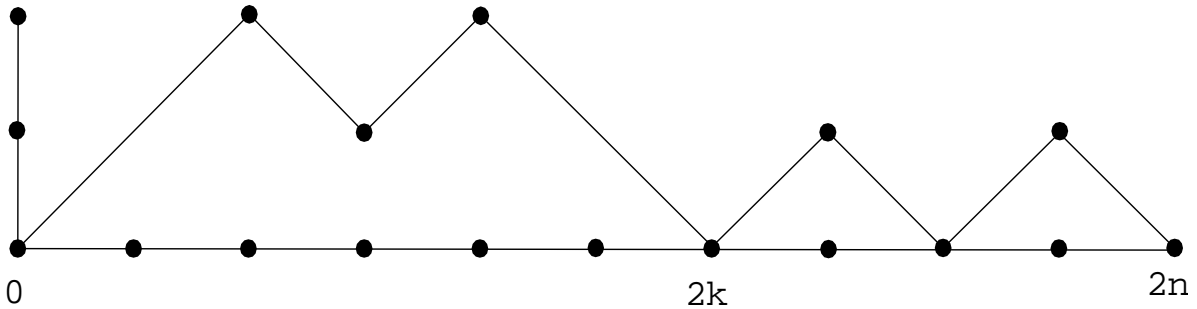


Fig.5: Décomposition d'un chemin de Dyck.

2.4.2 Chemins monotones et chemins monotones de Dyck

Définition 2.4.3 *Etant donné deux points dans le plan discret \mathbb{Z}^2 , $\mathbf{m} = (m_1, m_2)$ et $\mathbf{n} = (n_1, n_2)$, tels que $m_1 \leq n_1$ et $m_2 \leq n_2$; on appelle un chemin monotone de \mathbf{m} à \mathbf{n} un chemin qui prend ses pas dans l'ensemble $S = \{(1, 0), (0, 1)\}$.*

On note par $\mathcal{M}(\mathbf{m}, \mathbf{n})$ l'ensemble des chemins monotones de \mathbf{m} à \mathbf{n} .

Un chemin monotone de \mathbf{m} à \mathbf{n} qui ne passe jamais en dessous de la diagonale $y = x$ (tout en pouvant la toucher) est appelé *un bon chemin*, et l'ensemble des bons chemins de \mathbf{m} à \mathbf{n} est noté par $\mathcal{H}_b(\mathbf{m}, \mathbf{n})$. Un chemin monotone qui ne vérifie pas cette propriété est appelé *un mauvais chemin*. L'ensemble des mauvais chemins est noté par $\mathcal{H}_m(\mathbf{m}, \mathbf{n})$.

Définition 2.4.4 On appelle un chemin monotone de Dyck un chemin qui appartient à l'ensemble $\mathcal{H}_b((0, 0), (n, n))$; tel que $n \in \mathbb{N}$ est sa taille.

On note simplement par \mathcal{H} la classe des chemins monotones de Dyck et par \mathcal{H}_n l'ensemble des chemins dans \mathcal{H} qui sont de taille n , c'est à dire $\mathcal{H}_n = \mathcal{H}_b((0, 0), (n, n))$, $H_n = \text{card}(\mathcal{H}_n)$.

Exemple 2.4.3 Voici tous les chemins de \mathcal{H}_3 .

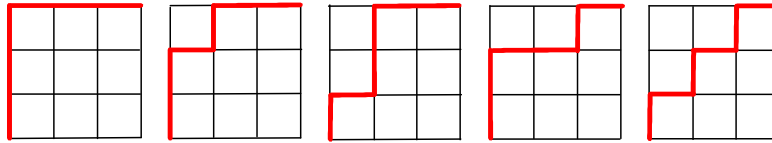


Fig.6: Les chemins monotones de Dyck de taille 3.

Proposition 2.4.2 L'ensemble des chemins monotones de Dyck de taille n est égal au $n^{\text{ème}}$ nombre de Catalan; i.e. $H_n = C_n$, pour $n \geq 0$.

Preuve. (bijective) On pivote un chemin de Dyck de taille n d'un angle $+\frac{\pi}{4}$ on obtient un chemin monotone de Dyck de même taille, et vice versa. ■

Un poset des chemins monotones de Dyck

On définit une relation d'ordre partiel sur l'ensemble \mathcal{H}_n des chemin de Dyck monotones de taille n . On se référera ici à l'article de J. Woodcock [23].

Définition 2.4.5 Un ensemble ordonné est un ensemble d'objets, \mathcal{P} , muni d'une relation d'ordre \leq qui satisfait les trois conditions suivantes :

- \leq est réflexive (on a $x \leq x$ pour tout $x \in \mathcal{P}$)
- \leq est transitive (si pour tous $x, y \in \mathcal{P}$, $x \leq y$ et $y \leq z$ alors $x \leq z$)

Exemple 2.4.5 Ici, d_1 et d_3 sont incomparables.

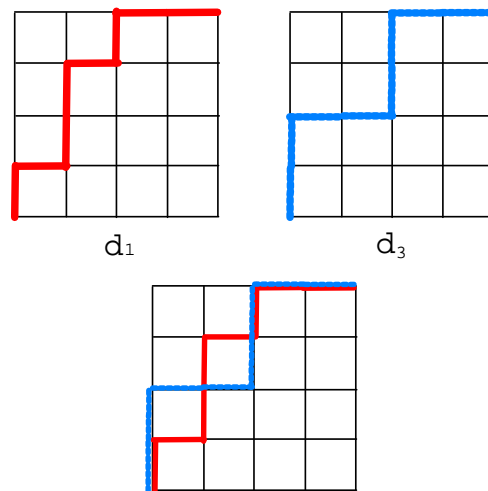


Fig.8: d_1 et d_3 sont incomparables.

Définition 2.4.7 Un élément y couvre un élément x dans le poset \mathcal{P} si $x < y$ (i.e. $x \leq y$ et $x \neq y$) et il n'existe pas d'élément $z \in \mathcal{P}$ tel que $x < z < y$.

Cette notion de couverture permet d'associer à tout élément de \mathcal{P} une représentation graphique dans le plan, appelée *diagramme de Hasse*.

Pour le tracer :

- on associe à tout élément de \mathcal{P} un point du plan;
- on place l'élément y couvrant x au dessus de x et on relie les deux points par un segment de droite ascendant.

Exemple 2.4.6 Les diagrammes ci-dessous montrent les posets \mathcal{H}_3 et \mathcal{H}_4 représentés par leurs diagrammes de Hasse respectifs.

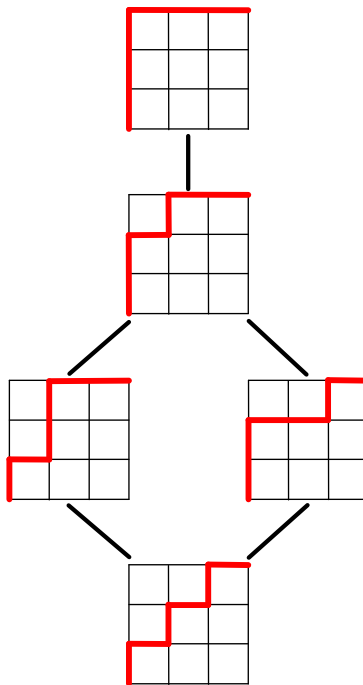


Fig.9: Le poset \mathcal{H}_3

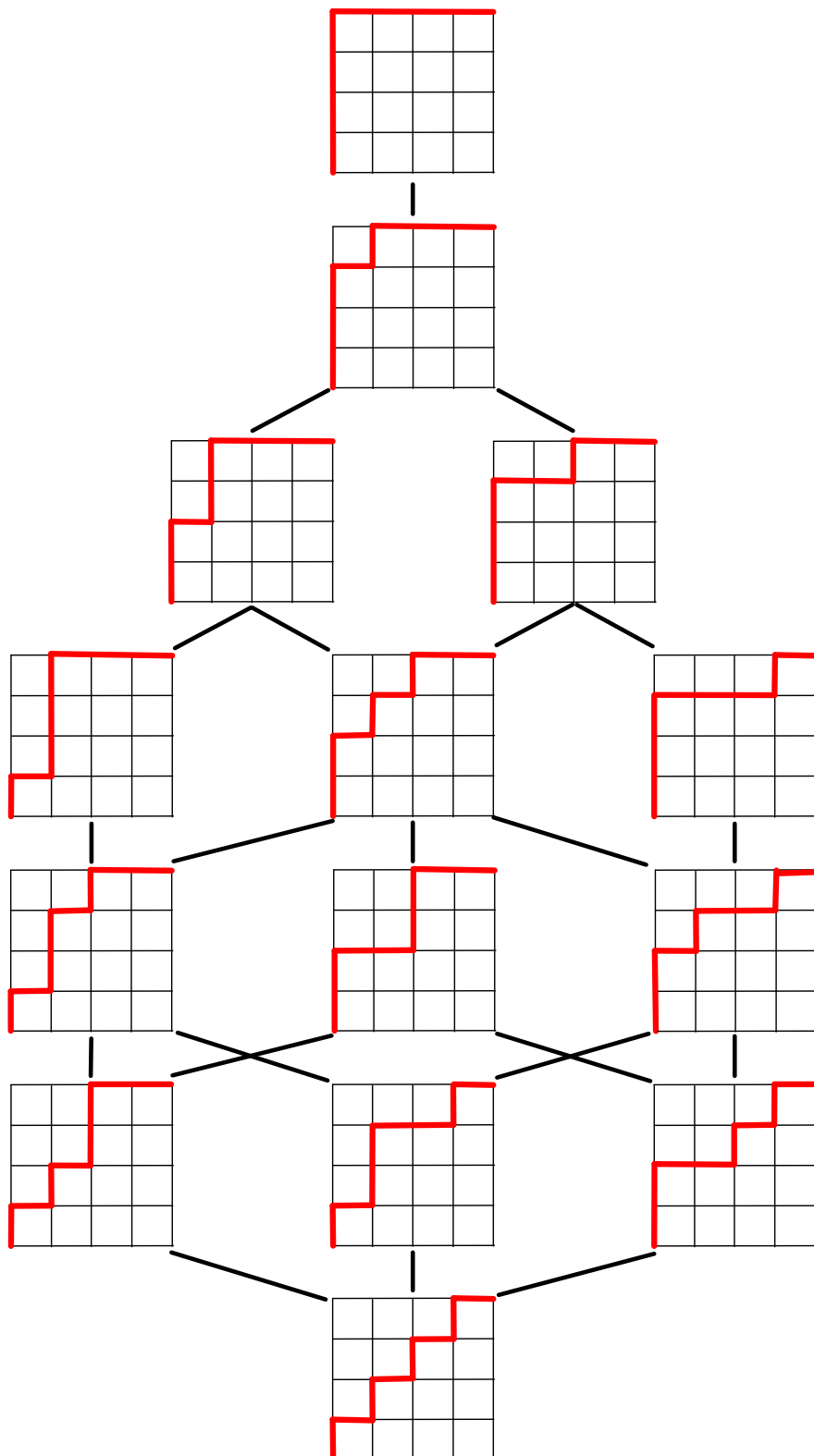


Fig.10: Le poset \mathcal{H}_4 .

2.5 Mots de Dyck

Toujours dans cette "famille des objets de Dyck" énumérés par les nombres de Catalan, on retrouve la classe des mots de Dyck, qui n'est, enfin de compte, qu'une variante, en termes de langages, des chemins de Dyck et des chemins monotones de Dyck.

Définition 2.5.1 *Un mot de Dyck est une séquence de lettres prises à partir d'un alphabet binaire $\mathcal{A} = \{X, Y\}$, vérifiant les deux conditions suivantes :*

- *la séquence contient autant de X que de Y et*
- *aucun préfixe (un mot obtenu en supprimant les dernières lettres à partir d'un rang quelconque) ne contient plus de Y que de X .*

Autrement dit, lorsque nous parcourons un mot de Dyck de gauche à droite, le nombre de X rencontrés est toujours supérieur ou égal au nombre de Y .

On note par \mathcal{M} la classe des mots de Dyck, et par \mathcal{M}_n l'ensemble des mots dans \mathcal{M} qui sont de longueur $2n$; c'est à dire qu'ils contiennent n lettres X et n lettres Y . On dit qu'ils sont de taille n .

Exemple 2.5.1 *Voici tous les mots de Dyck de taille $n = 3$:*

$XXXYYY$, $XXYXY$, $XXYYXY$, $XYXXYY$, $XYXYXY$.

En l'occurrence, $M_3 = 5$.

Proposition 2.5.1 [1] *Le nombre de mots de Dyck de taille n est égal au $n^{\text{ème}}$ nombre de Catalan; i.e. $M_n = \text{card}(\mathcal{M}_n) = C_n$, pour $n \geq 0$.*

Preuve. (bijective). La bijection est immédiate avec les chemins de Dyck de taille n : en assimile chaque lettre X d'un mot de Dyck de taille n avec un pas ascendant $(1, 1)$ d'un chemin de Dyck de même taille; et chaque lettre Y avec un pas descendant $(1, -1)$. La première condition dans la définition d'un mot de Dyck est équivalente au fait qu'un chemin de Dyck commence à partir du point $(0, 0)$ et finit au point $(2n, 0)$; ce qui fait qu'il y a autant de pas ascendants que de pas descendants. La deuxième condition est équivalente au fait qu'un chemin de Dyck ne passe jamais en dessous de l'axe des abscisses. D'où :

$$M_n = D_n = C_n, \text{ pour } n \geq 0. \blacksquare$$

Exemple 2.5.2 Le chemin de Dyck de taille $n = 5$ dans la figure suivante correspond au mot de Dyck : $X X Y X Y Y X X Y Y$.

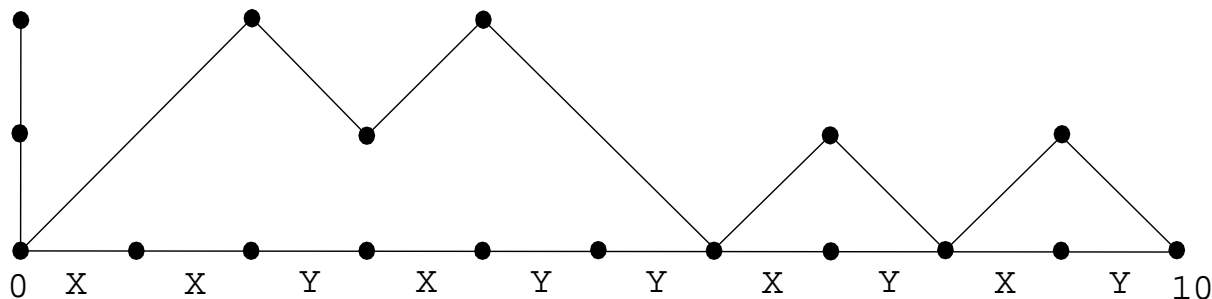


Fig.11 : bijection entre mot de Dyck et chemin de Dyck.

2.5.1 Mots de Dyck et problème de ballottage

Une manière pour retrouver directement la formule explicite de M_n existe, en faisant une correspondance avec le problème de ballottage (voir l'article de D. André [1]).

Problème 2.5.1 On suppose que deux candidats X et Y soient soumis à un scrutin de ballottage. Le nombre des votants est $\alpha + \beta$. X obtient α suffrages et est élu, Y en obtient β . On demande le nombre de manières pour que, pendant le dépouillement du scrutin, le nombre des voix de X ne cesse pas une seule fois de surpasser celles de son concurrent. C'est à dire qu'à chaque étape du dépouillement, le nombre de voix obtenu par X est strictement supérieur à celles de Y .

C'est ce problème là, appelé le problème du scrutin de ballottage, qu'on va adapter au problème des mots de Dyck, de la façon suivante : notons qu'il y a une bijection entre l'ensemble \mathcal{M}_n des mots de Dyck formés de n lettres X et n lettres Y , et l'ensemble \mathcal{M}'_n des mots formés de $n + 1$ lettres X et n lettres Y tels que dans chaque préfixe d'un mot de cet ensemble, le nombre des X est **strictement** supérieur à celui des Y .

Voici tous les mots de \mathcal{M}'_3 :

$XXXXYYY$, $XXXYYXY$, $XXXYYXY$, $XXYXXYY$, $XXYXYXY$.

La bijection est immédiate : chaque mot de \mathcal{M}_n , en lui ajoutant une lettre X à sa gauche, on obtient un mot de \mathcal{M}'_n , et chaque mot de \mathcal{M}'_n , en lui supprimant la première lettre (qui est nécessairement X), on obtient un mot de Dyck.

Maintenant, il suffit de compter les mots de \mathcal{M}'_n .

On peut remarquer qu'il s'agit bien d'un cas particulier du problème de ballottage où $\alpha = n + 1$ et $\beta = n$, en considérant le dépouillement comme une séquence des X et des Y , où chaque lettre X signifie une voix obtenue par le candidat X et chaque lettre Y une voix obtenue par le candidat Y .

Voici la solution de *André* :

Solution 2.5.1 *Le nombre total de mots que l'on peut former avec $n + 1$ lettres X et n lettre Y est égal à $\binom{2n + 1}{n}$. On note le nombre de mots qui ne sont pas dans \mathcal{M}'_n par N_n . Ainsi*

$$|\mathcal{M}'_n| = \binom{2n + 1}{n} - N_n.$$

On cherche N_n (les permutations défavorables). Les permutations qui leurs correspondent sont de deux sortes: celles qui commencent par Y et celles qui commencent par X .

Les permutations défavorables commençant par Y sont en nombre égal au nombre des permutations que l'on peut former avec $n + 1$ lettres X et $n - 1$ lettres Y , car il suffit évidemment d'y supprimer la lettre initiale Y pour obtenir ces dernières.

Les permutations défavorables commençant par X sont aussi nombreuses que les précédentes, car on peut les associer, chacune à chacune, aux permutations formées avec $n + 1$ lettres X et $n - 1$ lettres Y .

Cette règle se compose de deux parties :

1^o Etant donnée une permutation défavorable commençant par X , on y supprime la première lettre Y qui enfreint 'la propriété d'être de \mathcal{M}'_n ', puis on échange les deux groupes séparés par cette lettre entre eux : on obtient ainsi une permutation, parfaitement déterminée, de $n + 1$ lettres X et $n - 1$ lettres Y . Soit, par exemple, la permutation défavorable $XXYYXYX$, de 4 lettres X et 3 lettres Y ; en supprimant le premier Y qui enfreint la 'propriété d'être de \mathcal{M}'_n ' (qui est le Y écrit en caractère foncé), on sépare les deux groupes XYX et XXY ; en échangeant ces groupes entre eux, on obtient la permutation $XYXXXY$, formés de 4 lettres X et 2 lettres Y .

²⁰ Etant donnée une permutation de $n + 1$ lettres X et $n - 1$ lettres Y , on la parcourt de droite à gauche, jusqu'à ce qu'on obtienne un groupe où le nombre des X dépasse d'une unité celui des Y ; on intervertit ces deux groupes, en plaçant entre eux une lettre Y : on forme ainsi une permutation défavorable commençant par X et parfaitement déterminée. Soit, par exemple, la permutation $XYXXXXY$, formés de 4 lettres X et 2 lettres Y ; en opérant comme on vient de le dire, on la partage en deux groupe XXY et XYX ; en intervertissant ces groupes en plaçant entre eux une lettre Y , on forme la permutation défavorable $XXYYXYX$.

Il résulte de tout ce qui précède que le nombre total des événements défavorables est le double du nombre des permutations que l'on peut former avec $n + 1$ lettres X et $n - 1$ lettres Y ; c'est à dire que

$$N_n = 2 \binom{2n}{n+1}, \text{ ainsi } |\mathcal{M}'_n| = \frac{1}{n+1} \binom{2n}{n}, n \geq 0.$$

2.6 Arbres et nombres de Catalan

2.6.1 Définitions

Les arbres sont parmi les objets les plus importants en combinatoire. Là aussi, les nombres de Catalan se manifestent. Mais avant d'en donner les différentes interprétations, il est utile de rappeler quelques définitions sur les arbres. On se référera ici à l'ouvrage de Stanley [18].

On donne quelques définitions basiques sur les graphes :

Définition 2.6.1 [13]

- *Un graphe G est un ensemble non vide de points, appelés sommets, liés par des lignes, appelées arêtes.*
- *Un chemin entre deux sommets v_0 et v_n est une séquence $v_0 - v_1 - \dots - v_n$ de sommets v_i liés par des arêtes.*
- *Un cycle est un chemin $v_0 - v_1 - \dots - v_n$ avec la condition que $v_0 = v_n$.*
- *Un graphe est dit connexe s'il existe un chemin entre chaque deux sommets du graphe.*

Définition 2.6.2 *Un arbre A est un graphe connexe et sans cycle. Un arbre enraciné est un arbre dont un sommet particulier est désigné comme racine.*

Dans toute la suite, on ne va considérer que les arbres enracinés. Alors on omet de dire "enraciné".

Un arbre peut être décrit de façon récursive comme suit :

Définition 2.6.3 [18] *Un arbre A est un ensemble fini d'éléments, qu'on appelle sommets, tel que :*

- *il existe un sommet particulier dans A appelé la racine de A , et*
- *le reste des sommets, en excluant la racine, est partitionné en $k \geq 0$ ensembles disjoints et non vides A_1, A_2, \dots, A_k , tels que chacun de ses ensemble est un arbre. Les arbres A_1, \dots, A_k sont appelés les sous-arbres de la racine. On appelle les racines des sous-arbres A_1, \dots, A_k les successeurs de la racine de A .*

Les sommets qui n'ont pas de successeurs sont appelés *les feuilles* de A . Le reste des sommets sont appelés *sommets internes*.

Dans une représentation graphique d'un arbre dans le plan, il est convenu de mettre la racine en haut, et de la relier à ses successeurs par une arête (implicitement orientée de haut en bas).

Dans la figure ci-dessous, un arbre dont l'ensemble des sommets est $\{6, 2, 7, 3, 5, 1, 8, 4, 9\}$. La racine de cet arbre c'est le sommet 6. Les sous-arbres de la racine 6 sont : $\{2, 7\}$, $\{3, 5, 1, 8, 4, 9\}$.

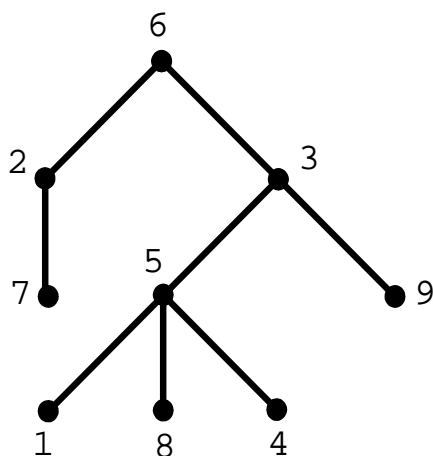


Fig.12 : un arbre à 9 sommets.

Dans un arbre, l'ordre des sous-arbres n'est pas pris en compte. C'est à dire, dans la figure ci-dessus, si on interchange les deux sous-arbres $\{2, 7\}$ et $\{3, 5, 1, 8, 4, 9\}$ entre eux (en mettant $\{2, 7\}$ à droite et l'autre à gauche), l'arbre qu'on obtiendra est le même que l'arbre initial.

Définition 2.6.4 [18] *Un arbre plan (ou arbre ordonné) est défini de façon récursive comme suit :*

- *il existe un sommet particulier appelé la racine de A , et*
- *le reste des sommets, en excluant la racine, est partitionné en une partition **ordonnée** d'ensembles disjoints non vides, chacun d'eux est un arbre plan.*

Dans une présentation graphique d'un arbre plan, on met la racine en haut, et la partition ordonnée en sous arbres (A_1, A_2, \dots, A_k) est ordonnée en la dessinant de gauche à droite dans le plan.

En fait, un arbre plan n'est qu'une représentation planaire d'un arbre (général). Par exemple, dans la figure ci-après, on a 4 arbres plans différents. Mais comme arbres *ordinaires* (non ordonnés), ils n'en font que deux : de gauche à droite, le premier, le deuxième et le quatrième sont un même arbre; le troisième est un autre arbre.

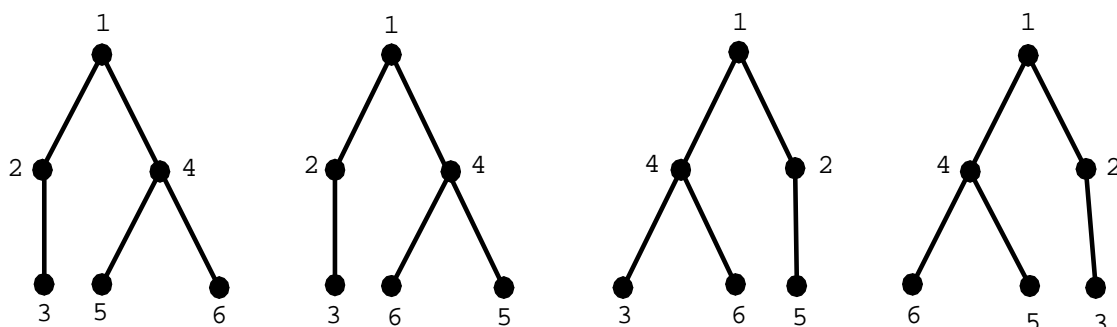


Fig.13: 4 arbres plans différents

Définition 2.6.5 [18] *Un arbre libre est un arbre où on ignore la désignation des sommets (tous les sommets sont indiscernables).*

Ainsi, dans un arbre (respectivement dans un arbre plan), on ne distingue les sous-ensemble A_1, A_2, \dots, A_k que par leur cardinal.

Par exemple, dans la figure 13, les 4 arbres représentent un seul arbre libre. Et comme arbres plans libres, ils n'en représentent que 2 (les deux premiers et les deux derniers).

Dans toute la suite, on ne va considérer que les arbres libres (respectivement les arbres plans libres). Donc, les mots *arbre* et *arbre plan* serviront, implicitement, à désigner un arbre libre (respectivement un arbre plan libre).

La figure 14 représente deux arbres plans (libres), mais qui ne font qu'un seul arbre (libre).

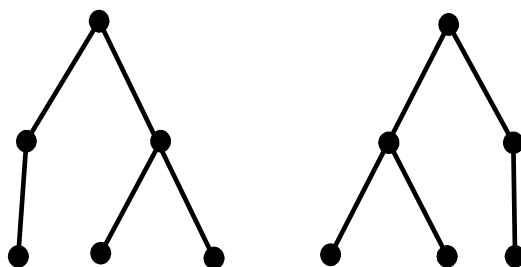


Fig.14 : Deux arbres plans libres.

Définition 2.6.6 [18] *Soit $k \geq 2$. On dit que A est un arbre k -aire si :*

- A est vide, ou un sommet particulier de A est désigné comme racine, et
- le reste des sommets (en excluant la racine) sont mis en une partition ordonnée d'**exactement** k ensembles disjoints (A_1, A_2, \dots, A_k) (**éventuellement vides**), chacun d'eux est un arbre k -aire.

Dans une représentation graphique d'un arbre k -aire dans le plan, on prend le soin de dessiner les arêtes joignant un sommet v aux racines de ses sous-arbres A_1, \dots, A_k avec des angles égaux et symétriques par rapport à l'axe vertical (en prenant en considération même les sous-arbres vides).

Dans la figure 13, on a 5 arbres 3-aires (appelés aussi arbres ternaires), avec 4 sommets.

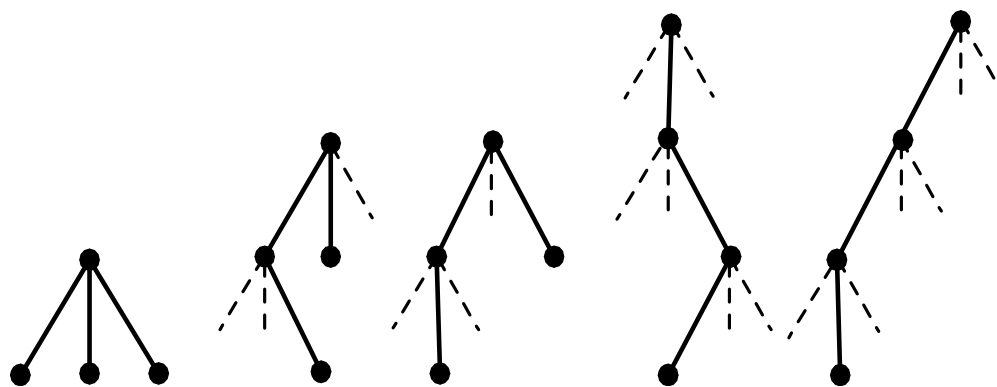


Fig.15 : 5 arbres 3-aires avec 4 sommets.

Un arbre k -aire est dit *complet* si chaque sommet qui n'est pas une feuille a exactement k sous-arbres *non vides* (c'est à dire k successeurs). Dans la figure 15, seul le premier arbre est complet.

2.6.2 Interprétations combinatoires

Les nombres de Catalan énumèrent quelques classes d'arbres plans.

Soient les classes d'arbres suivantes:

- \mathcal{G} la classe des arbres plans et \mathcal{G}_n l'ensemble des arbres dans \mathcal{G} qui ont n sommets,

- \mathcal{B}^2 la classe des arbres 2-aires (appelés binaires) et \mathcal{B}_n^2 l'ensemble des arbres dans \mathcal{B}^2 qui ont n sommets,
- \mathcal{R}^2 la classe des arbres binaires complets et \mathcal{R}_n^2 l'ensemble des arbres dans \mathcal{R}^2 qui ont n sommets internes.

Proposition 2.6.1 [8] $\mathcal{P}_n \cong \mathcal{R}_n^2 \cong \mathcal{B}_n^2 \cong \mathcal{G}_{n+1}$ pour $n \geq 0$.

(Rappelons que \mathcal{P}_n c'est la classe des parenthésages binaires de $n + 1$ lettres).

On va donner quelques applications illustratives entre ces classes :

$\mathcal{P}_n \cong \mathcal{R}_n^2$: soit A un arbre binaire complet de \mathcal{R}_n^2 . Il a donc $n + 1$ feuilles. On affecte à chaque feuille de A une lettre d'un mot de \mathcal{P}_n , puis on affecte à chaque sommet le produit de ses deux successeurs. On obtient, à la racine, un mot parenthésé de $n + 1$ lettres. L'exemple ci-après illustre bien la bijection:

A un arbre de \mathcal{R}_4^2 , et $M = ((ab)c)(de)$ un mot de \mathcal{P}_4^2 :

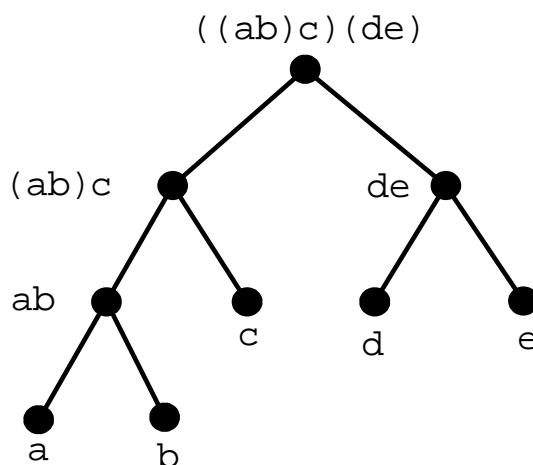


Fig.16 : bijection entre \mathcal{P}_4 et \mathcal{R}_4 .

$\mathcal{R}_n^2 \cong \mathcal{B}_n^2$: Pour obtenir un arbre de \mathcal{B}_n^2 à partir d'un arbre de \mathcal{R}_n^2 , il suffit de supprimer toutes les feuilles ainsi que leurs arêtes correspondantes de l'arbre binaire complet. L'opération est inversible; on colle à chaque feuille d'un arbre de \mathcal{B}_n^2 deux feuilles et à chaque sommet de degré 1 une feuille. La figure 14 illustre la bijection entre tous les arbres de \mathcal{R}_3^2

et \mathcal{B}_3^2 :

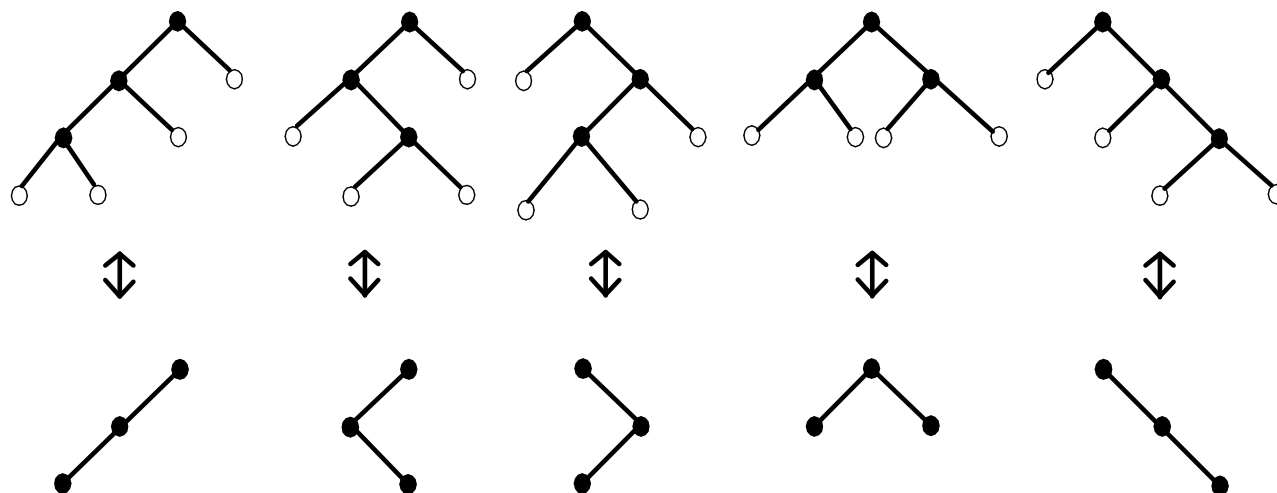


Fig.17 : Bijection entre \mathcal{R}_3^2 et \mathcal{B}_3^2 .

$\mathcal{B}_n^2 \cong \mathcal{G}_{n+1}$: Tout arbre de \mathcal{G}_{n+1} peut être représenté de telle sorte que chaque sommet a deux types de successeurs: le premier pointant vers son successeur naturel le plus à gauche, et l'autre pointant vers son voisin droit dans le même niveau. Avec cette représentation, si on omet la racine de l'arbre, chaque sommet est lié à au plus deux successeurs, ce qui nous donne un arbre de \mathcal{B}_n^2 . Le processus est réversible. La figure suivante illustre la bijection

entre \mathcal{G}_9 et \mathcal{B}_8^2 .

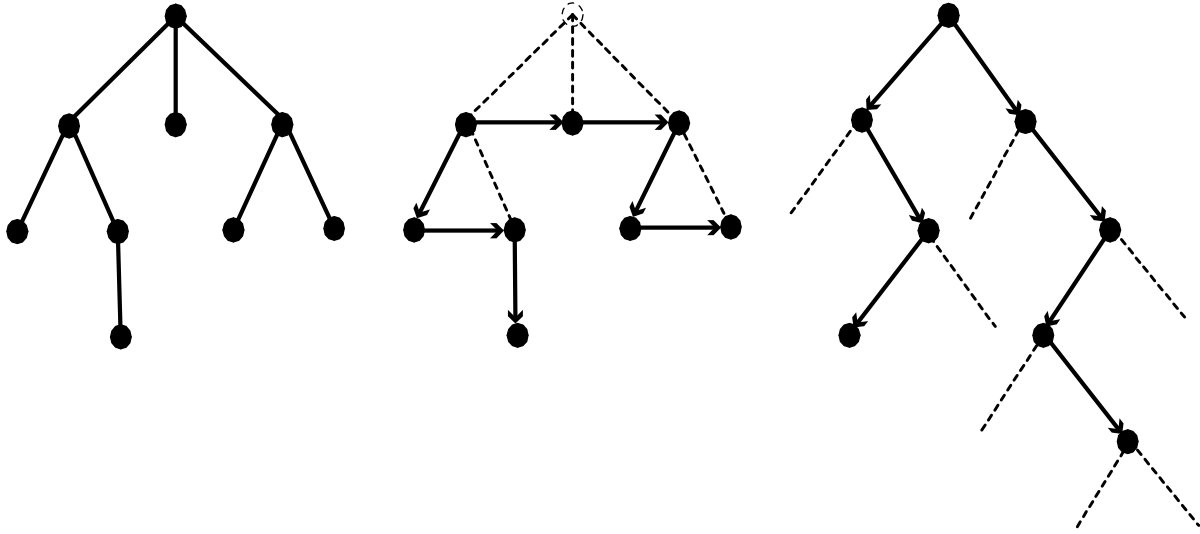


Fig.18: Bijection entre \mathcal{G}_9 et \mathcal{B}_8 .

2.6.3 Enumération des arbres par la méthode symbolique

La définition récursive des arbres plans \mathcal{G} à n sommets, les arbres k -aires $\mathcal{B}^{(k)}$ à n sommets et les arbres k -aires complets $\mathcal{R}^{(k)}$ à n sommets internes permet de retrouver facilement leur FGO et de les énumérer de façon directe avec la méthode symbolique (voir Dromata [7]).

Arbres plans à n sommets

D'après la définition 2.6.4 , un arbre plan à $n \geq 1$ sommets peut être décrit comme suit

$$\mathcal{G} = \{o\} + \{o\} \times \mathcal{G} + \{o\} \times \mathcal{G}^2 + \{o\} \times \mathcal{G}^3 + \dots \quad (56)$$

où $\{o\}$ désigne la classe d'arbres à un seul sommet. La description symbolique (56) découle du fait qu'un arbre plan a soit un seul sommet qui est la racine, soit la racine et un sous-arbre plan, soit la racine et deux sous-arbres plan, et ainsi de suite. Ainsi, la FGO $G(z)$ de \mathcal{G} vérifie

$$G(z) = z + zG(z) + zG(z)^2 + zG(z)^3 + \dots = \frac{z}{1 - G(z)}. \quad (57)$$

D'où,

$$G(z) = \frac{1 - \sqrt{1 - 4z}}{2} = zC(z), \quad (58)$$

tel que $C(z)$ est la FGO des nombres de Catalan. Ce qui nous donne

$$G_n = C_{n-1} = \frac{1}{n} \binom{2n-2}{n-1}. \quad (59)$$

Arbres k -aires à n sommets

Pour $k, n \geq 0$, les arbres k -aires $\mathcal{B}^{(k)}$ à n sommets peuvent être décrits comme suit

$$\mathcal{B}^{(k)} = \{\epsilon\} + \{\circ\} \times (\mathcal{B}^{(k)})^k, \quad (60)$$

où ϵ désigne un arbre vide. Alors la FGO de $\mathcal{B}^{(k)}$ vérifie l'équation

$$B^{(k)}(z) = 1 + zB^{(k)}(z)^k, \quad (61)$$

qui est une généralisation de l'équation quadratique (52) pour $k = 2$. Mais pour $k > 2$ il n'est pas facile de déduire l'expression explicite de $B^{(k)}(z)$ de l'équation (61). Pour cela, on utilise le Théorème d'inversion de Lagrange (voir l'ouvrage de M. Dromata [7]).

Théorème 2.6.1 (Théorème d'inversion de Lagrange). *Soit $\phi(z)$ une série formelle avec $\phi(0) \neq 0$ et $y(z)$ la série formelle solution de l'équation*

$$y(z) = z\phi(y(z)).$$

Alors $y(z)$ est inversible et le $n^{\text{ème}}$ coefficient de $g(y(z))$ (où $g(z)$ est une série formelle arbitraire) est donné par

$$[z^n]g(y(z)) = \frac{1}{n}[u^{n-1}]g'(u)\phi(u)^n \quad (n \geq 1). \quad (62)$$

Maintenant, on posant $\tilde{B}^{(k)}(z) = B^{(k)}(z) - 1$, on obtient

$$\tilde{B}^{(k)}(z) = z(1 + \tilde{B}^{(k)}(z))^k \quad (63)$$

et par le Théorème d'inversion de Lagrange (pour $n \geq 1$)

$$\begin{aligned} B_n^{(k)} &= [z^n]\tilde{B}^{(k)}(z) = \frac{1}{n}[u^{n-1}](1+u)^{kn} \\ &= \frac{1}{n} \binom{kn}{n-1} = \frac{1}{(k-1)n+1} \binom{kn}{n}. \end{aligned} \quad (64)$$

Ce nombre est l'une des généralisations du nombre de Catalan pour $k = 2$, et qui sera l'objet du chapitre suivant.

Chapitre 3

Nombres de Catalan généralisés

Les nombres de Catalan sont un cas particulier d'une classe plus large de suites classiques. Dans ce chapitre, on va explorer deux façons de les généraliser.

3.1 Les nombres k -Catalan

Les nombres k -Catalan, notés $C(n, k)$, sont définis par [13] :

$$C(n, k) = \frac{1}{(k-1)n+1} \binom{kn}{n}, \quad (65)$$

où $n \geq 0$ et $k \geq 1$. Il est clair que $C(n, 2) = C_n$.

Tout comme les nombres de Catalan ordinaires, tout nombre de Catalan généralisé $C(n, k)$ est un entier. Pour le prouver, il suffit de noter que :

$$\binom{kn}{n} - \binom{kn}{n-1} = \frac{(k-2)n+1}{(k-1)n+1} \binom{kn}{n}. \quad (66)$$

Par la relation de Bézout, on montre que $(k-1)n+1$ et $(k-2)n+1$ sont premiers entre eux $(-(k-2)((k-1)n+1) + (k-1)((k-2)n+1))$, alors $((k-1)n+1) \mid \binom{kn}{n}$.

Les nombres k -Catalan s'écrivent aussi sous la forme

$$C(n, k) = \frac{1}{kn+1} \binom{kn+1}{n}, \quad (67)$$

ou sous la forme

$$C(n, k) = \frac{1}{n} \binom{kn}{n-1}. \quad (68)$$

Voici quelques premiers termes de $C(n, k)$:

$n \setminus k$	1	2	3	4	5	6
1	1	1	1	1	1	1
2	1	2	3	4	5	6
3	1	5	12	22	35	51
4	1	14	55	140	570	1012
5	1	42	273	969	2530	5481
6	1	132	1428	7084	23751	62832

(69)

3.1.1 Interprétations combinatoires

Comme les nombres de Catalan ordinaires, les nombres de Catalan généralisés $C(n, k)$ admettent de nombreuses interprétations combinatoires qui sont des généralisations des objets comptés par les nombres de Catalan.

Lemme de Raney et mots de Dyck

Dans la sous-section 2.5.1 on a vu qu'un mot de Dyck de taille n dans un alphabet $\mathcal{A} = \{X, Y\}$ peut être défini de façon équivalente comme une séquence de $n + 1$ lettres X et n lettres Y tel que dans chaque préfixe de ce mot, le nombre des X est strictement supérieur à celui des Y . Si on remplace X par $+1$ et Y par -1 , le mot de Dyck devient une séquence de longueur $n + 1$, constituée de $+1$ et -1 telle que toute somme partielle sur la séquence à partir de la gauche est strictement positive, et la somme totale est égale à $+1$.

Maintenant, en exploitant un fameux lemme, appelé le Lemme de Raney (voir G. N. Raney [15]), on peut énumérer les mots de Dyck de taille n .

Lemme 3.1.1 (Lemme de Raney). *Si (x_1, x_2, \dots, x_m) est une séquence d'entiers dont la somme totale est égale à $+1$, alors il existe une seule séquence parmi toutes les permutations circulaires*

$$(x_1, x_2, \dots, x_m), (x_2, \dots, x_m, x_1), \dots, (x_m, x_1, \dots, x_{m-1})$$

qui a toutes ses sommes partielles strictement positives.

Par exemple, considérons la suite $(1, -3, 2, 4, -5, 0, 2)$. Ses permutations circulaires sont

$$\begin{array}{ll} (1, -3, 2, 4, -5, 0, 2) & (-5, 0, 2, 1, -3, 2, 4) \\ (-3, 2, 4, -5, 0, 2, 1) & (0, 2, 1, -3, 2, 4, -5) \\ (2, 4, -5, 0, 2, 1, -3) & (2, 1, -3, 2, 4, -5, 0) \\ (4, -5, 0, 2, 1, -3, 2) & \end{array}$$

La permutation cyclique qui a ses sommes partielles toutes positives est $(2, 4, -5, 0, 2, 1, -3)$.

Maintenant, pour chaque séquence parmi les $\binom{2n+1}{n}$ séquences constituées de $n + 1$ fois le nombre $+1$ et n fois le nombre -1 , il y a une seule permutation cyclique qui constitue un mot de Dyck. Et comme chaque suite a $2n + 1$ permutations cycliques, alors parmi les $\binom{2n+1}{n}$ séquences il y en a $\frac{1}{2n+1} \binom{2n+1}{n} = C_n$ qui sont des mots de Dyck.

Généralisation On considère maintenant les suites à $kn + 1$ éléments, constituées de n fois le nombre $(1 - k)$ et $(k - 1)n + 1$ fois le nombre $(+1)$, dont les sommes partielles sont toutes positives et la somme totale est égale à 1. On les appellera *séquences k -Raney* de taille n .

Il est clair qu'il y a C_n séquences *2-Raney* de taille n .

Proposition 3.1.1 (Raney [15]) *Le nombre de séquences k -Raney de taille n est égal à*

$$C(n, k) = \frac{1}{kn + 1} \binom{kn + 1}{n}.$$

Preuve. La preuve est immédiate en appliquant le Lemme de Raney. Pour chaque séquence parmi les $\binom{kn+1}{n}$ séquences constituées de n fois le nombre $(1 - k)$ et $(k - 1)n + 1$ fois le nombre $(+1)$, il y a une seule permutation cyclique qui constitue une séquence *k -Raney*. Et comme chaque séquence a $kn + 1$ permutations cycliques, alors parmi les $\binom{kn+1}{n}$ séquences il y en a $\frac{1}{kn+1} \binom{kn+1}{n} = C(n, k)$ qui sont des séquences *k -Raney*. ■

Proposition 3.1.2 (Raney [15]) *Pour $n \geq 1, k \geq 0$, on a*

$$C(n, k) = \sum_{n_1+n_2+\dots+n_k=n-1} C(n_1, k)C(n_2, k)\dots C(n_k, k). \quad (70)$$

Preuve. La séquence triviale $(+1)$, de longueur 1, est bien sur une séquence k -Raney. Si on ajoute le nombre $(1 - k)$ à droite d'une concaténation de k séquences k -Raney, on obtient une nouvelle séquence k -Raney.

Réciproquement, on peut montrer que, pour $n > 0$, toute séquence k -Raney, $(a_0, a_1, \dots, a_{kn})$, peut se décomposer en k séquences k -Raney de cette manière. En effet, le dernier terme a_{kn} est forcément égal à $(1 - k)$. De plus, les sommes partielles $S_j = a_0 + a_1 + \dots + a_{j-1}$ sont strictement positives pour tout $1 \leq j \leq kn$, et $S_{kn} = k$ car $S_{kn} + a_{kn} = 1$. Soit m_1 le plus grand indice $\leq kn$ tel que $S_{m_1} = 1$, m_2 le plus grand indice tel que $S_{m_2} = 2$ et ainsi de suite. Alors $S_{m_j} = j$ et $S_m > j$ pour tous j et m tels que $m_j < m \leq kn$ et $1 \leq j \leq k$. Il s'ensuit que $m_k = kn$.

On peut bien vérifier que chacune des sous-séquences $(a_0, a_1, \dots, a_{m_1-1})$, $(a_{m_1}, \dots, a_{m_2-1})$, \dots , $(a_{m_{k-1}}, \dots, a_{m_k-1})$ est une séquence k -Raney. Il existe donc des entiers positifs ou nuls n_1, n_2, \dots, n_k tels que $m_1 = kn_1 + 1$, $m_2 - m_1 = kn_2 + 1$, \dots , $m_k - m_{k-1} = kn_k + 1$.

D'où la convolution (70). ■

Les chemins monotones de Dyck k -aires (voir S. Heubach, Nelson Y. Li, T. Mansour [12])

Définition 3.1.1 *Un chemin monotone de Dyck k -aire de taille n est un chemin monotone de $(0, 0)$ à $(n, (k - 1)n)$ et qui ne passe jamais en dessous de la ligne $y = (k - 1)x$.*

Notons par $H(n, k)$ le nombre de chemins monotones de Dyck k -aires de taille n .

Proposition 3.1.3 $H(n, k) = C(n, k)$, $n \geq 0$, $k \geq 1$.

Preuve. On peut voir une séquence k -Raney comme une suite de n fois le nombre $(1 - k)$ et $n(k - 1)$ fois le nombre $(+1)$ telle les somme partielles sont positives ou nulles et la somme totale est nulle. D'où l'équivalence avec les chemins monotones de Dyck k -aires en représentant chaque élément $(+1)$ de la séquence k -Raney par un pas vertical $(0, 1)$, et chaque élément $(1 - k)$ par un pas horizontal $(1, 0)$. ■

Les arbres k -aires et les arbres k -aires complets (voir S. Heubach, Nelson Y. Li, T. Mansour [12])

Proposition 3.1.4 $B_n^{(k)} = R_n^{(k)} = C(n, k)$, $n \geq 0$, $k \geq 1$.

Preuve. On a vu à la fin du chapitre précédent que la FGO de $B_n^{(k)}$ vérifie l'équation (61), et que par conséquent

$$B_n^{(k)} = \frac{1}{(k-1)n+1} \binom{kn}{n}.$$

Et d'après la définition récursive de la classe \mathcal{R}^k des arbres k -aires complets, la FGO de $R_n^{(k)}$ vérifie aussi l'équation (61), d'où

$$R_n^{(k)} = \frac{1}{(k-1)n+1} \binom{kn}{n}.$$

Mais une bijection directe entre les classes \mathcal{R}_n^k et \mathcal{B}_n^k existe. Pour obtenir un arbre de \mathcal{B}_n^k à partir d'un arbre de \mathcal{R}_n^k , il suffit de supprimer toutes les feuilles ainsi que leurs arêtes correspondantes de l'arbre binaire complet. L'opération est inversible; on colle à chaque feuille d'un arbre de \mathcal{B}_n^k k feuilles et on colle à chaque sommet interne de degré inférieur à k le nombre nécessaire de feuilles pour qu'il devienne k -aire.

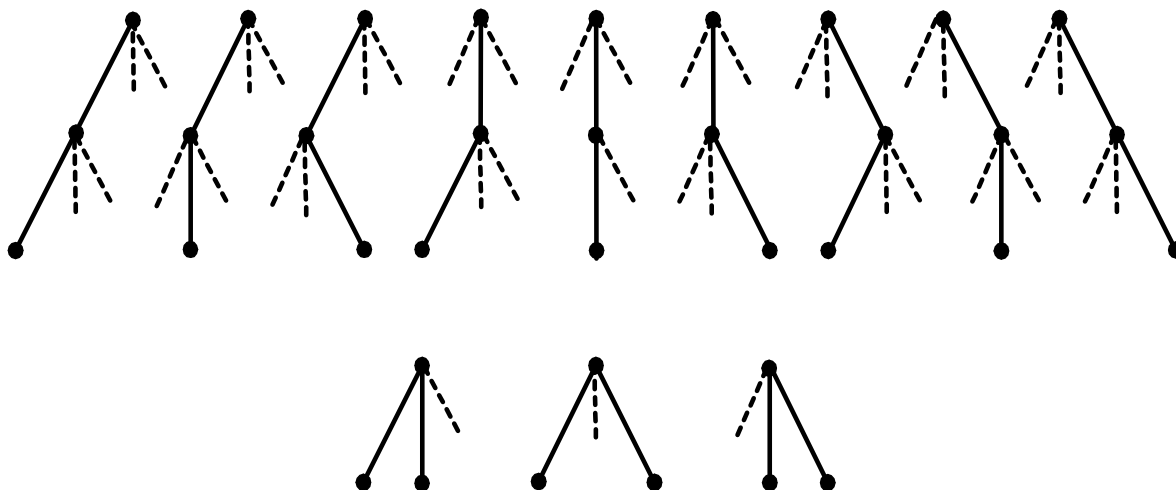
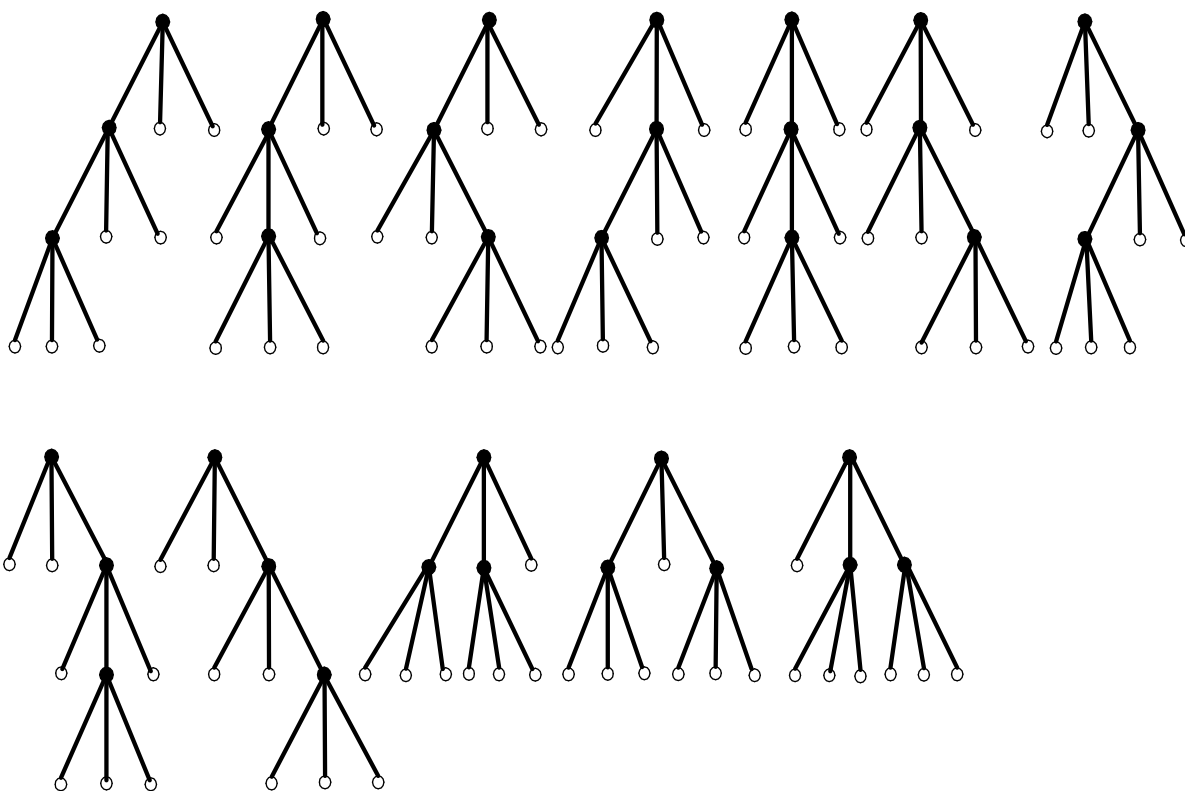


Fig.19: Les arbres de \mathcal{B}_3^3 .

Fig.20: Les arbres de \mathcal{R}_3^3 .

■

Parentésages k -aires (voir S. Heubach, Nelson Y. Li, T. Mansour [12])

Définition 3.1.2 *Un parentésage k -aire est une manière de parenthésier $(k-1)n+1$ lettres d'un ensemble muni d'une loi multiplicative non commutative et non associative, de telle sorte que, à chaque étape, on n'ait droit qu'à la multiplication de k facteurs adjacents.*

On note par $\mathcal{P}^{(k)}$ la classe des parentésages k -aires et par $P_n^{(k)}$ le nombre de parentésages k -aire de $(k-1)n+1$ lettres.

Proposition 3.1.5 $P_n^{(k)} = C(n, k)$, $n \geq 0$, $k \geq 2$.

Preuve. On établit une bijection entre la classe des parentésages k -aires de \mathcal{P}_n^k et les arbres k -aires complets de \mathcal{R}_n^k .

Un arbre de \mathcal{R}_n^k a n sommets internes, et par conséquent $(k-1)n+1$ feuilles (on peut le prouver par récurrence). On affecte à chaque feuille de l'arbre une lettre d'un mot de $\mathcal{P}_n^{(k)}$, puis on affecte à chaque sommet le produit de ses trois successeurs. On obtient, à la racine, un mot parenthésé de $(k-1)n+1$ lettres. ■

L'exemple ci-après illustre la bijection

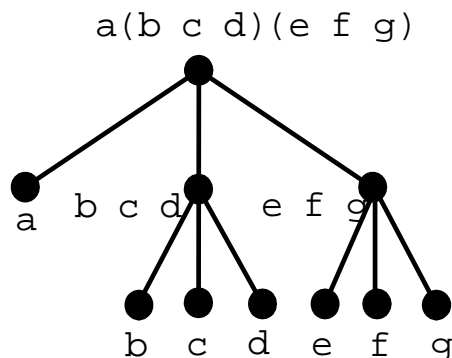


Fig.21: Bijection entre \mathcal{R}_3^3 et \mathcal{P}_3^3 .

Il y a $P_3^{(3)} = C(3,3) = 12$ mots parenthésés possibles : $ab(cd(efg))$, $ab(c(def)g)$, $ab((cde)fg)$, $a(bc(def))g$, $a(b(cde)f)g$, $a(bc(def))g$, $(ab(cde))fg$, $(a(bcd)e)fg$, $((abc)de)fg$, $(abc)(def)g$, $(abc)d(efg)$, $a(bcd)(efg)$.

3.2 Les nombres de Catalan multidimensionnels

Une autre généralisation des nombres de Catalan c'est *les nombres de Catalan multidimensionnels*.

On se référera dans cette partie à Frame, Robinson et Thrall [9], Stanley [20] et Sloane A060854.

Définition 3.2.1 *Les nombres de Catalan multidimensionnels sont donnés par*

$$C_n^m = \frac{0!1!2!\dots(n-1)!(mn)!}{m!(m+1)!\dots(m+n-1)!}, \quad (71)$$

où $n \geq 0$, $m \geq 0$.

Le cas $m = 2$ correspond au nombres de Catalan classiques $C_n^2 = C_n$.

Proposition 3.2.1 *Expression des nombres de Catalan multidimensionnels en termes de coefficients binomiaux et multinomiaux*

$$C_n^m = \frac{\binom{mn}{m, m, \dots, m}}{\binom{m}{m} \binom{m+1}{m} \cdots \binom{m+n-1}{m}}, \quad m \geq 0, \quad n \geq 0. \quad (72)$$

Preuve. Un calcul direct mène au résultat. ■

Ecriture de C_n^m avec les nombres triangulaires

On peut réécrire C_n^m avec les nombres triangulaires

$$T(n) = \frac{(n-1)n}{2} = \binom{n}{2} = 1 + 2 + \dots + (n-1), \quad n \in \mathbb{N}. \quad (73)$$

Proposition 3.2.2 *Pour $m, n \in \mathbb{N}$, on a*

$$C_n^m = \frac{\binom{T(m+n)}{1, 2, \dots, m+n-1}}{\binom{T(m)}{1, 2, \dots, m-1} \binom{T(n)}{1, 2, \dots, n-1} \binom{T(m+n)}{T(m), T(n), mn}}. \quad (74)$$

Corollaire 3.2.1 *Les nombres de Catalan multidimensionnels sont symétriques, i.e. $C_n^m = C_m^n$, $m, n \geq 0$.*

3.2.1 Interprétations combinatoires

Beaucoup d'objets comptés par les nombres de Catalan ordinaires sont généralisables aux nombres de Catalan multidimensionnels.

Tableaux de Young Standards

Les tableaux de Young jouent un rôle prépondérant en combinatoire. Dans cette partie, nous allons considérer les tableaux de Young classiques qui sont en relation étroite avec les nombres de Catalan: *les tableaux de Young standards* (voir Stanley [20]).

Définition 3.2.2 *Une partition d'un entier n est une façon d'écrire l'entier n comme une somme d'entiers positifs en ignorant l'ordre de la somme.*

Exemple 3.2.1 $3+4+2+1+1+4$ est une partition de 15, et $4+4+3+2+1+1$ représente la même partition.

Il est convenu d'écrire les éléments d'une partition dans l'ordre décroissant.

Définition 3.2.3 Soit λ une partition de l'entier n dont les éléments sont $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_l$. Un tableau de Young standard (TYS) de la forme λ est un tableau justifié à gauche d'entiers positifs, avec λ_i entiers dans la $i^{\text{ème}}$ ligne, satisfaisant les conditions suivantes:

- 1) Les entrées sont les entiers $1, 2, \dots, n$, chacun d'eux apparait une et une seule fois, et
- 2) Les entrées dans chaque ligne et chaque colonne sont croissantes.

Exemple 3.2.2 Un exemple d'un tableau de Young de taille $(4, 3, 2)$ est donnée par

1	3	4	6
2	7	8	
5	9		

Voici les 10 tableaux de Young standards à 4 éléments:

1	2	3	4	1	2	3	1	2	4	1	3	4	1	2
				4			3			2			3	4
1	3	1	2	1	3	1	4	1		2			2	
		3		2		2		3		3			3	
		4		4		3		4		4			4	

Les tableaux de Young standards ont, eux même, de nombreuses interprétations qui leur donnent une grande importance dans une variété de problèmes algébriques, combinatoires et probabilistiques. Ici, nous présentons un problème classique appelé *le problème de ballottage*, qui a de nombreuses applications en théorie des probabilités.

Problème de ballottage généralisé Soit une partition donnée $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_l)$ telle que $\lambda_1 + \lambda_2 + \dots + \lambda_l = n$.

On suppose que l candidats A_1, A_2, \dots, A_l sont soumis à une élection (un scrutin de ballottage). A la fin de l'élection, le candidats A_i obtient λ_i voix. Les votants votent

successivement, un par un. On considère les votes des votants comme une suite a_1, a_2, \dots, a_n , telle que $a_j = i$ si le $j^{\text{ème}}$ votant vote pour le candidat A_i . La suite a_1, a_2, \dots, a_n est appelée *la suite de ballotage* (de taille λ) si à aucun moment durant le scrutin un candidat A_j ne dépasse un candidat A_i en nombre de voix avec $i < j$.

Par exemple, la suite 1, 2, 1, 3, 1, 3, 4, 2 n'est pas une suite de ballotage du moment qu'au 6^{ème} vote, A_3 obtient plus de voix que A_2 . Par contre la suite 1, 2, 1, 1, 3, 1, 2, 2, 3 est une suite de ballotage.

Proposition 3.2.3 (Stanley et Bjorner [20]) *Il y a une bijection entre les suites de ballotage de taille λ et les tableaux de Young standards de taille λ .*

Preuve. Soit T un tableau de Young standard.

On pose $a_j = i$ si j apparait dans la $i^{\text{ème}}$ ligne de T. Il est facile de voir que la suite a_1, a_2, \dots, a_n est une suite de ballotage.

A l'inverse, si on prend une suite de ballotage a_1, a_2, \dots, a_n , pour chaque $a_j = i$, on met j dans la $i^{\text{ème}}$ ligne, dans l'ordre de la suite a_1, a_2, \dots, a_n . ■

Exemple 3.2.3 *Le tableau suivant:*

1	3	4	6
2	7	8	
5	9		

correspond à la suite de ballotage 1, 2, 1, 1, 3, 1, 2, 2, 3.

Enumération des tableaux de Young standards On se référera dans cette partie à Stanley [20] et Frame, Robinson et Thrall [9].

Il est naturel de se poser la question suivante: combien y a-t-il de tableaux de Young standards de taille λ ?

Le nombre de tableaux de Young standards de type λ est noté f^λ .

Par exemple, il y a 9 tableaux de Young standards de taille $(4, 2)$, et on écrit $f^{4,2} = 9$. Ces neuf tableaux sont donnés par

1	2	3	4	1	2	3	5	1	2	3	6	1	2	4	5	1	2	4	6
5	6			4	6			4	5			3	6			3	5		

1	2	5	6	1	3	4	5	1	3	4	6	1	3	5	6
3	4			2	6			2	5			2	4		

La formule de l'équerre La formule de f^λ a été donnée par James Sutherland Frame, Robinson and Robert McDowell Thrall en 1954 [9], et est appelée *la formule de la longueur de l'équerre*, ou *la formule de Frame-Robinson-Thrall*.

Pour établir cette formule, on définit le *diagramme de Young* de type λ comme un tableau justifié à gauche, avec λ_i cases dans la $i^{\text{ème}}$ ligne.

Par exemple, voici le diagramme de Young de type $(5, 5, 2)$:

Un tableau de Young de type λ peut être alors regardé comme une insertion des nombres $1, 2, \dots, n$ (chacun apparaît une seule fois) dans les cases du diagramme de Young de type λ , de telle sorte que chaque colonne et chaque ligne est croissante.

Si s est une case du diagramme de Young, alors on définit la longueur-équerre de s comme le nombre de cases qui se trouvent à droite de s sur la même ligne et les cases qui sont en dessous de s sur la même colonne, en comptant s une seule fois.

Dans le diagramme suivant, on a inséré dans chaque case du diagramme de Young de type $(5, 5, 2)$ sa longueur-équerre.

7	6	4	3	2
6	5	3	2	1
2	1			

Le produit de l'équerre H_λ d'une partition λ c'est le produit de toutes les longueurs-équerres du diagramme de Young.

Par exemple, à partir du diagramme précédent on a:

$$H_{5,5,2} = 7.6.4.3.2.6.5.3.2.1.2.1 = 362880.$$

La formule de Frame-Robinson-Thrall stipule ceci (voir Frame, Robinson et Thrall [9]):

Théorème 3.2.1 *Le nombre de tableaux de Young Standard f^λ de type λ est égal à*

$$f^\lambda = \frac{n!}{H_\lambda}, \tag{75}$$

où λ est une partition en entier de n .

Par exemple,

$$f^{5,5,2} = \frac{12!}{362880} = 1320.$$

Il est remarquable qu'une simple formule de f^λ existe mais aucune preuve combinatoire simple n'est encore connue.

Les nombres de Catalan multidimensionnels et les tableaux de Young

Considérons un cas particulier des tableaux de Young standards: les tableaux de Young à m lignes et n colonnes; les rectangles.

Théorème 3.2.2 *Le nombre de tableaux de Young standards à m lignes et n colonnes (donc de type $\lambda = (\underbrace{n, n, \dots, n}_{m \text{ fois}})$) est égal à $C_n^m = \frac{0!1!2!\dots(n-1)!(mn)!}{m!(m+1)!\dots(m+n-1)!}$.*

Preuve. Soit le diagramme de Young à m lignes et n colonnes. Ces longueurs-équerres sont de la forme :

$m+n-1$	$m+n-2$	$m+n-3$...	m
$m+n-2$	$m+n-3$	$m+n-4$...	$m-1$
$m+n-3$	$m+n-4$	$m+n-5$...	$m-2$
...	\ddots	...
n	$n-1$	$n-2$...	1

$$H_{m \times n} = \frac{(m!) (m+1)! (m+2)! \dots (m+n-1)!}{0! 1! 2! \dots (n-1)!} = \frac{m!(m+1)! \dots (m+n-1)!}{0!1!2! \dots (n-1)!}.$$

$$f^{m \times n} = \frac{(mn)!}{H_{m \times n}} = \frac{0!1!2! \dots (n-1)! (mn)!}{m!(m+1)! \dots (m+n-1)!} = C_n^m.$$

■

Une façon combinatoire d'établir la symétrie des nombres de Catalan multidimensionnels, c'est à dire que $C_n^m = C_m^n$ pour $m, n \geq 0$, c'est de remarquer qu'il suffit de considérer le transposé d'un tableau de Young standards à m ligne et n colonnes pour obtenir un tableau de Young standard à n lignes et m colonnes.

Les tableaux de Young standards à deux colonnes et le triangle de Catalan

Dans ce qui suit, nous allons déterminer le triangle de Catalan à partir des tableaux de Young standards à deux colonnes.

Théorème 3.2.3 *Le triangle de Catalan est parfaitement déterminé par les tableaux de Young standards à deux colonnes.*

Preuve. Le problème c'est de chercher combien y a-t-il de tableaux de Young standards à 2 colonnes et n lignes.

Pour cela, il suffit de chercher quelles sont toutes les possibilités de remplir juste la première colonne, car la deuxième est déterminée en ordonnant le reste des nombres de façon croissante.

Dans chaque case de la première colonne, il est facile de déterminer quels sont les nombres qui peuvent y être. On obtient un triangle des possibilités, tel que chaque ligne correspond aux possibilités de la case correspondante de la première colonne du diagramme de Young.

Prenons pour exemple $n = 5$:

1				
2	3			
3	4	5		
4	5	6	7	
5	6	7	8	9

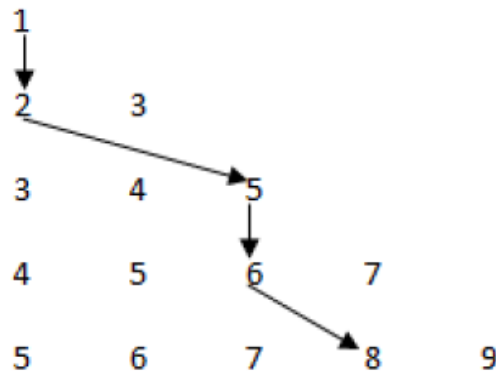
(76)

Ainsi par exemple, dans la 3^{ème} case de la première colonne, les possibilités sont 3, 4 et 5.

Le problème revient donc à choisir de chaque ligne du triangle précédent un seul nombre, de telle sorte à ce que la suite obtenue soit strictement croissante.

Ainsi, le problème revient à choisir un chemin de la première ligne à la dernière ligne du triangle en descendant soit verticalement ou à droite.

Voici un chemin possible:



Ce qui correspond au tableau de Young suivant:

1	3
2	4
5	7
6	9
8	10

On peut, de façon récurrente, déterminer tous les chemins possibles.

Considérons le triangle (76), dans chaque case, on met le nombre de chemins possibles qui aboutissent à cette case, on obtient le triangle de Catalan:

1
 1 1
 1 2 2
 1 3 5 5
 1 4 9 14 14

■

Les mots de Dyck généralisés

Une autre interprétation des nombres de Catalan multidimensionnels en termes de langage c'est la classe des mots de Dyck généralisés.

Définition 3.2.4 *Un mot de Dyck m -aire est une séquence de lettres prises à partir d'une alphabet m -aire $\mathcal{A} = \{X_1, X_2, \dots, X_m\}$, et qui vérifie les deux conditions suivantes:*

- 1) *chaque lettre apparaît le même nombre de fois n , qui est la taille du mot, et*
- 2) *aucun préfixe (un mot obtenu en supprimant les dernières lettres à partir d'un rang quelconque) ne contient plus de X_j que de X_i , avec $j > i$.*

Proposition 3.2.4 *Le nombre de mots de Dyck m -aires de taille n est égal à C_n^m .*

Preuve. Il suffit de faire une correspondance avec un cas particulier du problème de ballottage, où on considère m candidats, chaque candidat obtient n voix à la fin du scrutin.

Il y a donc une bijection entre les mots de Dyck m -aire de taille n et les suites de ballottage de taille $\lambda = \underbrace{(n, n, \dots, n)}_{m \text{ fois}}$. ■

Les chemins monotones de Dyck multidimensionnels

Définition 3.2.5 *On considère le plan discret de dimension m , \mathbb{Z}^m . On appelle un chemin monotone de Dyck de taille n et de dimension m dans \mathbb{Z}^m , un chemin de $(0, \dots, 0)$ à (n, \dots, n) , dont les pas sont pris de $S = \{(1, 0, \dots, 0), (0, 1, 0, \dots, 0), \dots, (0, 0, \dots, 1)\}$ et qui reste confinés à la partie du plan $x_1 \geq x_2 \geq \dots \geq x_m \geq 0$.*

Proposition 3.2.5 *Le nombre de chemins de Dyck de taille n et de dimension m est égal à C_n^m .*

Chapitre 4

Convolution de Gould-Vandermonde et nombres de Catalan généralisés

4.1 Convolution de Gould-Vandermonde

Une des identités combinatoires les plus élégantes est la *convolution de Vandermonde*:

$$\sum_{i=0}^n \binom{\alpha_1}{i} \binom{\alpha_2}{n-i} = \binom{\alpha_1 + \alpha_2}{n}, \quad (77)$$

où α_1 , α_2 et n entiers. Cette identité doit son nom au fait qu'Alexandre Vandermonde⁽¹⁾ a écrit, à la fin du dix-huitième siècle, un important article la concernant. Il s'avère toutefois qu'un chinois, Chu Shih-Chieh⁽²⁾, la connaissait bien avant, en 1303.

⁽¹⁾Alexandre-Théophile Vandermonde (parfois appelé Alexis-Théophile), né à Paris le 28 février 1735 et mort à Paris le 1er janvier 1796, est un mathématicien français. Il fut aussi économiste, musicien et chimiste, travaillant notamment avec Étienne Bézout et Antoine Lavoisier. Son nom est maintenant surtout associé à un déterminant. Il commence à faire des mathématiques en 1770. Son Mémoire sur la résolution des équations (1771), qui préfigure la théorie de Galois, porte sur les fonctions symétriques et la solution des polynômes cyclotomiques. Dans les Remarques sur des problèmes de situation (1771), il étudie le problème du cavalier. Son Mémoire sur des irrationnelles de différents ordres avec une application au cercle (1772) porte sur la combinatoire, et son Mémoire sur l'élimination (1772) sur les fondations de la théorie des déterminants. Ces communications, présentées à l'Académie des sciences, constituent toute son œuvre mathématique.

⁽²⁾Chu Shih-Chieh, 1280-1303, est un mathématicien chinois. Il a contribué à l'étude des séries arithmétiques et géométrique ainsi qu'aux différences finies. Ses deux ouvrages mathématiques, Introduction to Mathematical Studies et Precious Mirror of the Four Elements, n'ont été découverts qu'au 19ème siècle.

On peut démontrer cette convolution avec une jolie interprétation combinatoire. Dans le membre droit de l'identité (77), $\binom{\alpha_1 + \alpha_2}{n}$ représente le nombre de façons de choisir n personnes parmi α_1 hommes et α_2 femmes. Le membre gauche compte le nombre de façons de choisir i hommes parmi α_1 hommes et $n - i$ femmes parmi α_2 femmes, pour tout i , $0 \leq i \leq n$.

Gould, dans [18], a remarqué que l'identité (77) n'est qu'un cas particulier de l'identité générale suivante pour $k = 0$:

Proposition 4.1.1 *Pour $\alpha_1, \alpha_2, k, n \in \mathbb{N}$ on a :*

$$\frac{\alpha_1 + \alpha_2}{kn + \alpha_1 + \alpha_2} \binom{kn + \alpha_1 + \alpha_2}{n} = \sum_{i=0}^n \frac{\alpha_1}{ki + \alpha_1} \binom{ki + \alpha_1}{i} \frac{\alpha_2}{k(n-i) + \alpha_2} \binom{k(n-i) + \alpha_2}{n-i}. \quad (78)$$

Une identité qui s'appelle dès lors *identité de Gould-Vandermonde*. En posant

$$C(n, k, \alpha) = \frac{\alpha}{kn + \alpha} \binom{kn + \alpha}{n}, \quad (79)$$

on peut réécrire l'identité (78) comme suit

$$C(n, k, \alpha_1 + \alpha_2) = \sum_{i=0}^n C(i, k, \alpha_1) C(n-i, k, \alpha_2). \quad (80)$$

Preuve. Pour prouver (78), on démontre d'abord certains résultats préliminaires.

La différence Δf d'une fonction f au point x est

$$\Delta f(x) = f(x+1) - f(x).$$

Si on applique à nouveau l'opérateur Δ on obtient la différence seconde

$$\begin{aligned} \Delta^2 f(x) &= \Delta f(x+1) - \Delta f(x) = (f(x+2) - f(x+1)) - (f(x+1) - f(x)) \\ &= f(x+2) - 2f(x+1) + f(x), \end{aligned}$$

analogue à la dérivée seconde. De manière similaire, on a

$$\begin{aligned} \Delta^3 f(x) &= f(x+3) - 3f(x+2) + 3f(x+1) - f(x); \\ \Delta^4 f(x) &= f(x+4) - 4f(x+3) + 6f(x+2) - 4f(x+1) + f(x); \end{aligned}$$

et ainsi de suite. Ces expressions font apparaître des coefficients binomiaux à signes alternés.

Voici la formule générale pour la nième différence :

$$\Delta^n f(x) = \sum_{i=0}^n (-1)^{n-i} \binom{n}{i} f(x+i), \quad n \geq 0 \text{ entier.}$$

Si $f(x)$ est un polynôme de degré d , la différence $\Delta f(x)$ est un polynôme de degré $d-1$. Par conséquent, $\Delta^d f(x)$ est une constante et $\Delta^n f(x) = 0$ pour tout $n > d$. C'est un résultat extrêmement important qui permet de simplifier beaucoup de formules.

En prenant la $n^{\text{ème}}$ différence du polynôme

$$f(x) = \binom{kx}{n} \frac{\alpha}{kx},$$

de degré $\leq n-1$, et en prenant comme cas particulier $x = \frac{\alpha}{k}$

$$\sum_{i=0}^n (-1)^i \binom{n}{i} \binom{ki+\alpha}{n} \frac{\alpha}{ki+\alpha} \quad (81)$$

$$= (-1)^n C(n, k, \alpha) + \sum_{i=0}^{n-1} (-1)^i \binom{n}{i} \binom{ki+\alpha}{n} \frac{\alpha}{ki+\alpha} = 0, \quad n \geq 0. \quad (82)$$

D'où

$$C(i, k, \alpha) = \sum_{j=0}^{i-1} (-1)^{j+i+1} \binom{i}{j} \binom{kj+\alpha}{n} \frac{\alpha}{kj+\alpha}, \quad i \geq 1. \quad (83)$$

Nous avons donc ceci

$$\begin{aligned} \sum_{i=0}^{\infty} C(i, k, \alpha) z^i &= 1 + \sum_{i=1}^{\infty} C(i, k, \alpha) z^i \\ &= 1 + \sum_{i=1}^{\infty} z^i \sum_{j=0}^{i-1} (-1)^{j+i+1} \binom{i}{j} \binom{kj+\alpha}{n} \frac{\alpha}{kj+\alpha} \\ &= 1 + \sum_{j=0}^{\infty} (-1)^j \frac{\alpha}{kj+\alpha} \sum_{i=j}^{\infty} (-1)^i \binom{i+1}{j} \binom{kj+\alpha}{i+1} z^{i+1} \\ &= 1 + \sum_{j=0}^{\infty} (-1)^j \binom{kj+\alpha}{j} \frac{\alpha}{kj+\alpha} z^j \sum_{i=1}^{\infty} (-1)^i \binom{kj-j+\alpha}{i+1-j} z^{i+1} \\ &= 1 - \sum_{j=0}^{\infty} \binom{kj+\alpha}{j} \frac{\alpha}{kj+\alpha} z^j \sum_{i=1}^{\infty} (-1)^i \binom{kj-j+\alpha}{i} z^i \\ &= 1 - \sum_{j=0}^{\infty} C(j, k, \alpha) z^j \sum_{i=0}^{\infty} (-1)^i \binom{kj-j+\alpha}{i} z^i + \sum_{j=0}^{\infty} C(j, k, \alpha) z^j. \end{aligned}$$

On obtient ainsi

$$\sum_{j=0}^{\infty} C(j, k, \alpha) z^j \sum_{i=0}^{\infty} (-1)^i \binom{kj - j + \alpha}{i} z^i = 1,$$

ce qui donne

$$\sum_{j=0}^{\infty} C(j, k, \alpha) z^j (1 - z)^{kj - j + \alpha} = 1.$$

En prenant

$$z = \frac{x - 1}{x^k},$$

on obtient

$$\sum_{j=0}^{\infty} C(j, k, \alpha) z^j = x^\alpha. \quad (84)$$

Et par un argument de différence finie, on a

$$\sum_{j=0}^n (-1)^j \binom{n}{j} \binom{kj + \alpha}{n} = (-1)^n k^n, \quad n \geq 0. \quad (85)$$

Et à partir de cette identité on trouve

$$(-1)^n \binom{kn + \alpha}{n} = (-1)^n k^n - \sum_{j=0}^{n-1} (-1)^j \binom{n}{j} \binom{kj + \alpha}{n}, \quad n \geq 1.$$

D'où

$$\begin{aligned} \sum_{n=0}^{\infty} (-1)^n \binom{kn + \alpha}{n} z^n &= 1 + \sum_{n=1}^{\infty} (-1)^n \binom{kn + \alpha}{n} z^n \\ &= 1 + \sum_{n=1}^{\infty} (-kz)^n - \sum_{n=1}^{\infty} z^n \sum_{j=0}^{n-1} (-1)^j \binom{n}{j} \binom{kj + \alpha}{n} \\ &= \frac{1}{kz + 1} - \sum_{j=0}^{\infty} (-1)^j \sum_{n=j}^{\infty} \binom{n+1}{k} \binom{kj + \alpha}{n+1} z^{n+1} \\ &= \frac{1}{kz + 1} - \sum_{j=0}^{\infty} (-1)^j \binom{kj + \alpha}{j} \sum_{n=j}^{\infty} \binom{kj - j + \alpha}{n+1-j} z^{n+1} \\ &= \frac{1}{kz + 1} - \sum_{j=0}^{\infty} (-1)^j \binom{kj + \alpha}{j} \left(\sum_{n=0}^{\infty} \binom{kj - j + \alpha}{n} z^{n+j} - z^j \right) \\ &= \frac{1}{kz + 1} - \sum_{j=0}^{\infty} (-1)^j \binom{kj + \alpha}{j} z^j (z + 1)^{kj - j + \alpha} + \sum_{j=0}^{\infty} (-1)^j \binom{kj + \alpha}{j} z^j. \end{aligned}$$

Donc nous avons

$$\sum_{j=0}^{\infty} (-1)^n \binom{kj + \alpha}{j} z^j (1+z)^{kj-j+\alpha} = \frac{1}{kz+1},$$

et que nous pouvons réécrire de la forme

$$\binom{kj + \alpha}{j} z^j = \frac{x^{\alpha+1}}{(1-k)x+k}, \quad (86)$$

où

$$z = \frac{x-1}{x^k}.$$

A partir de (84) on a

$$\begin{aligned} x^{\alpha_1} . x^{\alpha_2} &= x^{\alpha_1+\alpha_2} = \sum_{j=0}^{\infty} C(j, k, \alpha_1 + \alpha_2) z^j \\ &= \left(\sum_{m=0}^{\infty} C(m, k, \alpha_1) \right) \left(\sum_{l=0}^{\infty} C(l, k, \alpha_2) \right) \\ &= \sum_{j=0}^{\infty} z^j \sum_{i=0}^j C(i, k, \alpha_1) C(j-i, k, \alpha_2). \end{aligned}$$

D'où

$$C(n, k, \alpha_1 + \alpha_2) = \sum_{i=0}^n C(i, k, \alpha_1) C(n-i, k, \alpha_2).$$

■

Nous donnons une autre preuve, proposée par M. Mihoubi, basée sur les suites binomiales

:

Preuve. Soit $(f_n(x))_{n \in \mathbb{N}}$ une suite binomiale ($f_n(x)$ un polynôme de degré n , $f_0(x) = 1$, et vérifie $f_n(x+y) = \sum_{k=0}^n \binom{n}{k} f_k(x) f_{n-k}(y)$.)

La suite

$$h_n(x) = \frac{x}{an+x} f_n(an+x)$$

est binomiale (voir M. Mihoubi, thèse).

Il est facile de vérifier que la suite

$$g_n(x) = n! \binom{x}{n}$$

est binomiale (Convolution de Vandermonde).

Alors, la suite

$$h_n(x) = n! \binom{an+x}{n} \frac{x}{an+x}$$

est aussi binomiale. D'où

$$\frac{\alpha_1 + \alpha_2}{kn + \alpha_1 + \alpha_2} \binom{kn + \alpha_1 + \alpha_2}{n} = \sum_{i=0}^n \frac{\alpha_1}{ki + \alpha_1} \binom{ki + \alpha_1}{i} \frac{\alpha_2}{k(n-i) + \alpha_2} \binom{k(n-i) + \alpha_2}{n-i}$$

Pour $\alpha_1, \alpha_2, k, n \in \mathbb{N}$. ■

4.2 Preuve combinatoire et généralisation de la convolution de Gould-Vandermonde

Revenons au nombre

$$C(n, k, \alpha) = \frac{\alpha}{kn + \alpha} \binom{kn + \alpha}{n}.$$

Il est clair que $C(n, k, \alpha)$ est une généralisation du nombre k -Catalan

$$C(n, k) = C(n, k, 1).$$

Tout comme les nombres k -Catalan, les $C(n, k, \alpha)$ ont une interprétation combinatoire en termes d'arbres plans.

Théorème 4.2.1 (Stanley [19]) *Le nombre de forêts composées de α arbres k -aires ordonnés avec n sommets internes est égal à*

$$C(n, k, \alpha) := \frac{\alpha}{kn + \alpha} \binom{kn + \alpha}{n}.$$

On peut démontrer la convolution (80) avec un simple argument combinatoire.

Pour $\alpha_1, \alpha_2, k, n \in \mathbb{N}$, le nombre de forêts ordonnées de $\alpha_1 + \alpha_2$ arbres k -aires avec au total n sommets internes est égal à $C(n, k, \alpha_1 + \alpha_2)$. Mais chaque forêt de ce type est en fait une concaténation de deux forêts, une a α_1 arbres k -aires et n_1 sommets internes, et l'autre a α_2 arbres k -aires et n_2 sommets internes, tel que $n_1 + n_2 = n$. Il en découle directement que

$$C(n, k, \alpha_1 + \alpha_2) = \sum_{n_1+n_2=n} C(n_1, k, \alpha_1) C(n_2, k, \alpha_2). \quad (87)$$

D'où la convolution de Gould-Vandermonde (80).

Avec le même argument combinatoire, on généralise la convolution (80) comme suit:

Corollaire 4.2.1 Pour $\alpha_1, \alpha_2, \dots, \alpha_r, k, n \in \mathbb{N}$ on a

$$C(n, k, \alpha_1 + \alpha_2 + \dots + \alpha_r) = \sum_{n_1 + n_2 + \dots + n_r = n} C(n_1, k, \alpha_1) C(n_2, k, \alpha_2) \dots C(n_r, k, \alpha_r). \quad (88)$$

Une identité impliquant les arbres k -aires et un certain type de forêts ordonnées d'arbres k -aires

Théorème 4.2.2 Le nombre d'arbres k -aires à $n + 1$ sommets internes est égal au nombre de forêts ordonnées de k arbres k -aires avec un total de n sommets internes. i.e.,

$$C(n + 1, k) = C(n, k, k).$$

Preuve. La preuve est immédiate en utilisant l'identité (70) et en prenant dans l'identité (88) $\alpha_i = 1$ pour $1 \leq i \leq r$. ■

4.2.1 Généralisation multinomiale des nombres de Catalan

Avant de donner la généralisation multinomiale des nombres de Catalan, nous donnons quelques préliminaires (voir Stanley [19]).

Définition 4.2.1 On dit qu'une forêt σ est de type $\mathbf{r} = (r_0, \dots, r_m) \in \mathbb{N}^{m+1}$ si elle a r_i sommets qui ont i successeurs.

Pour énumérer ce type de forêts, il convient de travailler sur les mots dans un *monoïde libre* qu'on va définir ci-après.

monoïde libre

Définition 4.2.2 Soit \mathcal{A} un ensemble fini qu'on appellera *alphabet*. Un mot de cet alphabet est une séquence finie d'éléments $a_1 a_2 \dots a_n$ de \mathcal{A} . On note par ϵ le mot vide. L'ensemble de tout les mots dans l'alphabet \mathcal{A} est noté par \mathcal{A}^* .

On définit le produit de deux mots $u = a_1 \dots a_n$ et $v = b_1 \dots b_m$ comme la concaténation de u et v ,

$$uv = u = a_1 \dots a_n b_1 \dots b_m.$$

En particulier, $\epsilon u = u \epsilon = u$ pour tout u de \mathcal{A}^* .

L'ensemble \mathcal{A}^* muni du produit est appelé monoïde libre sur l'alphabet \mathcal{A} .

Si $u = a_1 \dots a_n \in \mathcal{A}^*$ avec $a_i \in \mathcal{A}$, alors la longueur de u c'est $l(u) = n$. En particulier, $l(\epsilon) = 0$. On note par \mathcal{A}_n^* le sous-monoïde de \mathcal{A}^* dont les mots sont de longueur n .

Définition 4.2.3 Soit \mathcal{B} un sous ensemble de \mathcal{A}^* . Le sous-monoïde \mathcal{B}^* engendré par \mathcal{B} est l'ensemble de tout les mots $u_1 \dots u_n$ où $u_i \in \mathcal{B}$.

On dit que \mathcal{B}^* est librement engendré par \mathcal{B} si chaque mot $u \in \mathcal{B}^*$ peut être écrit de façon unique comme $u = u_1 \dots u_n$ où $u_i \in \mathcal{B}$.

Si \mathcal{B}^* est librement engendré par \mathcal{B} , alors on dit que \mathcal{B}^* est "très pur" si la condition suivante est satisfaite:

Si $u = u_1 \dots u_n \in \mathcal{B}^*$, avec $u_i \in \mathcal{A}$, alors nous avons une unique séquence

$$0 < n_1 < n_2 < \dots < n_\alpha < n,$$

telle que

$$u_1 u_2 \dots u_{n_1} \in \mathcal{B}, u_{n_1+1} u_{n_1+2} \dots u_{n_2} \in \mathcal{B}, \dots, u_{n_\alpha+1} u_{n_\alpha+2} \dots u_n \in \mathcal{B}.$$

Si \mathcal{B}^* est très pur, alors on appelle conjugué de $u = u_1 \dots u_n \in \mathcal{B}_n^*$, avec $u_i \in \mathcal{A}$, tout mot $u_i \dots u_n u_1 \dots u_{i-1} \in \mathcal{A}_n^*$.

Remarque 4.2.1 Il y a autant de conjugués d'un mot $u = u_1 \dots u_n \in \mathcal{B}_n^*$ que de permutations cycliques de u .

monoïde libre associé aux forêts de degré maximum m

Considérons un alphabet $\mathcal{A} = \{x_0, x_1, \dots, x_m\}$. On définit le poids d'une lettre x_i par la fonction suivante

$$\phi(x_i) = i - 1.$$

On étend ϕ à \mathcal{A}^* comme suit

$$\phi(w_1 w_2 \dots w_j) = \phi(w_1) + \phi(w_2) + \dots + \phi(w_j),$$

où $w_i \in \mathcal{A}$ pour tout i .

On définit l'ensemble $\mathcal{B} \in \mathcal{A}^*$ par

$$\mathcal{B} = \{w \in \mathcal{A}^* : \phi(w) = -1; \text{ et si } w = uv \text{ où } v \neq \epsilon, \text{ alors } \phi(u) \geq 0\}.$$

Lemme 4.2.1 (Stanley [19]) *Le monoïde \mathcal{B}^* engendré par \mathcal{B} est très pur.*

Lemme 4.2.2 (Stanley [19]) *Un mot $w \in \mathcal{A}^*$ est un conjugué d'un mot de \mathcal{B}^* si et seulement si $\phi(w) < 0$.*

Maintenant, on associe à une chaque forêt σ de type $\mathbf{r} = (r_0, r_1, \dots, r_m) \in \mathbb{N}^{m+1}$, à α arbres et n sommets, un mot $w(\sigma)$ de \mathcal{A}^* de longueur n et $\phi(w(\sigma)) = -\alpha$. Pour faire ainsi, on définit sur la forêt σ un certain ordre $ord(\sigma)$ de ses sommets comme suit :

(a) Si σ a $\alpha \geq 2$ arbres ordonnés $\tau_1, \tau_2, \dots, \tau_\alpha$, alors

$$ord(\sigma) = ord(\tau_1), ord(\tau_2), \dots, ord(\tau_\alpha) \text{ (concaténation de mots).}$$

(b) Si σ a un seul arbre et $\tau_1, \tau_2, \dots, \tau_j$ les sous arbre de sa racine ν , alors

$$ord(\sigma) = \nu, ord(\tau_1), ord(\tau_2), \dots, ord(\tau_j) \text{ (concaténation de mots).}$$

Si σ est une forêt et $ord(\sigma) = (v_1, \dots, v_n)$, alors on définit le mot $w(\sigma) \in \mathcal{A}^*$ par

$$w(\sigma) = x_{\delta_1} x_{\delta_2} \dots x_{\delta_n},$$

où δ_i est le nombre de successeurs de v_i .

Lemme 4.2.3 *Il y a une bijection entre l'ensemble des forêts d'arbres ordonnés et \mathcal{B}^* .*

On peut maintenant donner le résultat suivant:

Théorème 4.2.3 (Stanley [19]) *Soient $\mathbf{r} = (r_0, \dots, r_m) \in \mathbb{N}^{m+1}$, avec $\sum_{i=0}^m r_i = n$ et $\sum_{i=0}^m (1-i)r_i = \alpha > 0$. Alors le nombre $P(\mathbf{r})$ de forêts à k arbres ordonnés avec n sommets (sommets internes et feuilles) tels qu'il y a r_i sommets qui ont i successeurs est donnée par*

$$P(\mathbf{r}) = \frac{\alpha}{n} \binom{n}{r_0, \dots, r_m}.$$

Preuve. Par le Lemme 4.2.3, $P(\mathbf{r})$ est égal au nombre de mots dans \mathcal{B}^* avec r_i fois x_i pour chaque i . On note l'ensemble de ces mot par $\mathcal{B}_{\mathbf{r}}^*$. On définit $\mathcal{A}_{\mathbf{r}}^*$ comme l'ensemble des mots dans \mathcal{A}^* qui ont r_i fois le terme x_i . On définit l'application $\Psi : \mathcal{B}_{\mathbf{r}}^* \times \{1, \dots, n\} \rightarrow \mathcal{A}_{\mathbf{r}}^* \times \{1, \dots, \alpha\}$ comme suit

soit $w = w_1 w_2 \dots w_n = u_1 u_2 \dots u_\alpha \in \mathcal{B}_r^*$, où $w_i \in \mathcal{A}$ et $u_i \in \mathcal{B}$. Pour $i \in \{1, \dots, n\}$, w_i est une lettre de u_j .

Prenons alors

$$\Psi(w, i) = (w_i w_{i+1} \dots w_{i-1}, j).$$

Par le Lemme 4.2.1 Ψ est injective, et par le Lemme 4.2.2 (et le fait que $\phi(w) = -\alpha$ si $w \in \mathcal{B}^*$) Ψ est surjective. D'où

$$nP(\mathbf{r}) = \alpha \text{card}(\mathcal{A}_r^*).$$

Et comme

$$\text{card}(\mathcal{A}_r^*) = \binom{n}{r_0, \dots, r_m},$$

on a

$$P(\mathbf{r}) = \frac{\alpha}{n} \binom{n}{r_0, \dots, r_m}.$$

■

Voici une autre généralisation des nombres de Catalan:

Définition 4.2.4 Soient $\mathbf{n} = (n_1, \dots, n_t) \in \mathbb{N}^t$; $\beta = (\beta_1, \dots, \beta_t) \geq \mathbf{0}$, $\beta_i \neq \beta_j$ pour $i \neq j$; $\alpha \in \mathbb{N}$. On définit les nombres de Catalan multinomiaux comme suit:

$$Q(\mathbf{n}, \beta, \alpha) = \frac{\alpha}{\langle \beta, \mathbf{n} \rangle + \alpha} \binom{\langle \beta, \mathbf{n} \rangle + \alpha}{\mathbf{n}, \langle \beta - \mathbf{1}, \mathbf{n} \rangle + \alpha}. \quad (89)$$

Théorème 4.2.4 $Q(\mathbf{n}, \beta, \alpha)$ est le nombre de forêts à α arbres ordonnés composée de $\sum_{i=1}^t n_i$ sommets internes repartis en n_i sommets internes β_i -aires, pour $1 \leq i \leq t$.

Preuve. Il suffit de montrer que dans une forêt à α arbres ordonnés composée de $\sum_{i=1}^t n_i$ sommets internes repartis en n_i sommets internes β_i -aires pour $1 \leq i \leq t$, le nombre total des sommets (sommets internes + feuilles) est égal à $n = \langle \beta, \mathbf{n} \rangle + \alpha$. Considérons un arbre j de cette forêts. Supposons qu'il a $\sum_{i=1}^t n_i^j$ sommets internes repartis en n_i^j sommets internes β_i -aires, pour $1 \leq i \leq t$. On peut déconnecter l'arbre j en $\sum_{i=1}^t n_i^j$ sous arbres dont les racines sont les sommets internes de l'arbre j . Le nombre total des sommets dans cet ensemble d'arbres est égal à $\sum_{i=1}^t (\beta_i + 1)n_i^j$. On peut déduire alors que dans l'arbre initial j il y a $(\sum_{i=1}^t (\beta_i + 1)n_i^j) - (\sum_{i=1}^t n_i^j) + 1 = \langle \beta, \mathbf{n}^j \rangle + 1$ sommets où $\mathbf{n}^j = (n_1^j, \dots, n_t^j)$. Et comme il y a α arbres dans la forêts, alors le nombre total des sommets dans la forêt est égal à $\langle \beta, \mathbf{n} \rangle + \alpha$. Par le Théorème 4.2.3 on obtient le résultat voulu. ■

Généralisation multinomiale de la convolution de Gould-Vandermonde

Théorème 4.2.5 Soient $\alpha_1, \alpha_2, \dots, \alpha_r \in \mathbb{C}$, $\alpha = \sum_{i=1}^r \alpha_i$, $\beta_1, \beta_2, \dots, \beta_t \in \mathbb{C}$. Posons $\beta = (\beta_1, \beta_2, \dots, \beta_t)$, $\mathbf{n}_i = (n_{i1}, \dots, n_{it})$, $i = 1, \dots, r$. On se donne $\mathbf{n} = (n_{\cdot 1}, \dots, n_{\cdot t})$. On pose aussi $\tilde{n}_i = \langle \beta - \mathbf{1}, \mathbf{n}_i \rangle + \alpha_i$ et $\tilde{n} = \sum_{i=1}^r \tilde{n}_i$. On a alors

$$\sum_{\substack{\mathbf{n}_{ij}/ \\ \sum_{i=1}^r n_{ij} = n_{\cdot j}}} \prod_{i=1}^r \frac{\alpha_i}{\langle \beta, \mathbf{n}_i \rangle + \alpha_i} \binom{\langle \beta, \mathbf{n}_i \rangle + \alpha_i}{\mathbf{n}_i, \tilde{n}_i} = \frac{\alpha}{\langle \beta, \mathbf{n} \rangle + \alpha} \binom{\langle \beta, \mathbf{n} \rangle + \alpha}{\mathbf{n}, \tilde{n}}. \quad (90)$$

Il est clair que, d'après la Définition 4.2.4, on peut réécrire la formule (90) comme suit :

$$Q(\mathbf{n}, \beta, \alpha) = \sum_{\substack{\mathbf{n}_{ij}/ \\ \sum_{i=1}^r n_{ij} = n_{\cdot j}}} \prod_{i=1}^r Q(\mathbf{n}_i, \beta, \alpha_i). \quad (91)$$

Preuve. Comme les deux membres de l'identité (90) sont des polynômes en $\alpha_1, \alpha_2, \dots, \alpha_r$ et $\beta_1, \beta_2, \dots, \beta_t$, alors il suffit de prouver pour $\alpha_1, \alpha_2, \dots, \alpha_r \in \mathbb{N}$ et $\beta_1, \beta_2, \dots, \beta_t \in \mathbb{N}$. Soit une forêt à α arbres ordonnés composée de n sommets internes repartis en $n_{\cdot 1}$ sommets β_1 -aires, $n_{\cdot 2}$ sommets β_2 -aires, ..., $n_{\cdot t}$ sommets β_t -aires. Le nombre de telles forêts est égal à $Q(\mathbf{n}, \beta, \alpha)$. Le côté droit de l'identité (91) compte le nombre de façons de sectionner cette forêt en r forêts composées de α_1 arbres ordonnés, α_2 arbres ordonnés, ..., α_r arbres ordonnés respectivement, de sorte que l'on somme sur toutes les possibilités d'avoir dans la forêt i , n_{i1} sommets β_1 -aires, n_{i2} sommets β_2 -aires, n_{it} sommets β_t -aires pour tous $i = 1, \dots, r$. Ainsi, $\sum_{i=1}^r n_{i1} = n_{\cdot 1}$, $\sum_{i=1}^r n_{i2} = n_{\cdot 2}$, ..., $\sum_{i=1}^r n_{it} = n_{\cdot t}$. D'où le Théorème. ■

Conclusion

Tout au long de ce mémoire, nous avons fait une étude combinatoire des nombres de Catalan.

Dans le premier chapitre nous avons donné les principales propriétés des nombres de Catalan. Nous avons d'abord énuméré quelques expressions classiques qui les font intervenir. Puis, nous avons présenté quelques propriétés arithmétiques, leur relation avec le triangle de Pascal, ainsi qu'une propriété intéressante relative à la matrice de Hankel.

Dans le deuxième chapitre, nous avons mis en relief des objets combinatoires qui sont en connexions avec les nombres de Catalan. Nous avons vu que le nombre de triangulations d'un $(n + 2)$ -gone, que le nombre de parenthésages binaires de $n + 1$ lettres, que le nombre des chemins de Dyck et le nombre des chemins de Dyck monotones de longueur n est égal au $n^{\text{ème}}$ nombre de Catalan. A la fin du chapitre, nous avons montré qu'il y a un isomorphisme combinatoire entre les arbres plans à $n + 1$ sommets, les arbres binaires à n sommets et les binaires complets à n sommets internes, et que leur nombre est égal à C_n .

Au troisième chapitre, nous avons présenté deux généralisations de ces nombres: Les nombres k-Catalan, notés $C(n, k)$, avec interprétations combinatoires, puis les nombres de Catalan multidimensionnels C_n^k qui se lisent dans les tableaux de Young. Nous avons proposé une bijection combinatoire entre les tableaux de Young à deux colonnes et le triangle de Catalan. Il est à prévoir qu'il est possible de généraliser cette bijection combinatoire aux tableaux de Young à plus de deux colonnes.

Au dernier chapitre, nous avons étudié le nombre $C(n, k, \alpha) = \frac{\alpha}{kn + \alpha} \binom{kn + \alpha}{n}$, où α , k , $n \in \mathbb{N}$, c'est une autre généralisation des nombres k-Catalan. Nous avons présenté la convolution de Gould-Vandermonde qui stipule

$$C(n, k, \alpha_1 + \alpha_2) = \sum_{i=0}^n C(i, k, \alpha_1) C(n - i, k, \alpha_2).$$

4.2. Preuve combinatoire et généralisation de la convolution de Gould-Vandermonde

Nous avons proposé une preuve combinatoire de ce résultat. Enfin, nous avons énoncé et prouvé le résultat suivant

$$Q(\mathbf{n}, \beta, \alpha) = \sum_{\substack{n_{ij}/ \\ \sum_{i=1}^r n_{ij} = n \cdot j}} \prod_{i=1}^r Q(\mathbf{n}_i, \beta, \alpha_i).$$

où

$$Q(\mathbf{n}, \beta, \alpha) = \frac{\alpha}{\langle \beta, \mathbf{n} \rangle + \alpha} \binom{\langle \beta, \mathbf{n} \rangle + \alpha}{\mathbf{n}, \langle \beta - \mathbf{1}, \mathbf{n} \rangle + \alpha}$$

pour $\mathbf{n} = (n_1, \dots, n_t) \in \mathbb{N}^t$; $\beta = (\beta_1, \dots, \beta_t) \geq \mathbf{0}$, $\beta_i \neq \beta_j$ pour $i \neq j$; $\alpha \in \mathbb{N}$.

Bibliographie

- [1] D. André, Solution directe du problème résolu par M. Bertrand, Comptes Rendus de l'Académie des Sciences, Paris 105 (1887) 436-437.
- [2] J-C. Aval, Multivariate Fuss-Catalan Numbers, Discrete Math. 308 (2008) 4660-4669.
- [3] P. Barry, Journal of Integer Sequences, Vol 12 (2009). Article 09.5.4
- [4] M. Bona, A Walk Through Combinatorics, 2nd edition, World Scientific, New Jersey, 2006
- [5] R. Chen, A refinement of the formula for k-ary trees and the Gould-Vandermonde convolution, The electronic Journal of Combinatorics 15 (2008), #52.
- [6] L. Comtet, Advanced Combinatorics, Reidel, Dordrecht, Boston, 1974.
- [7] M. Dromata, Random Trees, Springer Wien New York, 2009.
- [8] P. Flajolet and R. Sedgewick, Analytic Combinatorics, Cambridge University Press, 2009.
- [9] J. S. Frame, G. de B. Robinson, and R. M. Thrall. The hook graphs of the symmetric groups. Canadian J. Math., 6:316–324, 1954.
- [10] H. W. Gould, Some Generalization of Vandermonde's Convolution, Amer. Math. Monthly 63 (1956) 84-91.
- [11] R. L. Graham, D. E. Knuth et O. Patashnik, Mathématiques concrètes, Fondations pour l'informatique, 2^e édition.

- [12] S. Heubach, Nelson Y. Li, T. Mansour, Staircase tilings and k -Catalan structures, *Discrete Mathematics* 308 (2008) 5954–5964.
- [13] T. Koshy, *Catalan Numbers with applications*, Oxford University Press 2009.
- [14] C. Radoux, Nombres de Catalan généralisés, *Bull. Belg. Math. Soc.* 4 (1997), 289-292.
- [15] G. N. Raney, “Functional composition patterns and power series reversion”, *Transactions of the American Mathematical Society* 94 (1960), 441-451.
- [16] N.J.A. Sloane, The On-Line Encyclopedia of Integer Sequences, <http://www.research.att.com/~njas/sequences>
- [17] R. P. Stanley, *Catalan Addendum*, version of 11 August 2009.
- [18] R. P. Stanley, *Enumerative Combinatorics 1*, Cambridge University Press, Cambridge, 1999.
- [19] R. P. Stanley, *Enumerative Combinatorics 2*, Cambridge University Press, Cambridge, 1999.
- [20] R. P. Stanley, A. Björner, *A Combinatorial Miscellany*, September 13, 2004.
- [21] R. P. Stanley, Richard et E. Weisstein, "Catalan Number." From *MathWorld*—A Wolfram Web Resource.
- [22] V. S. Sunder, *Catalan numbers Wonders of Science*, CESCOI, Madurai, August 25 2009.
- [23] J. Woodcock, *Properties of the poset of Dyck paths ordered by inclusion*. August 2008, Math 6002, York University.