

Publications du **Laboratoire de
Combinatoire et d'
Informatique
Mathématique**

10

Édité par J. Labelle et J.-G. Penaud

Atelier de Combinatoire franco-québécois

6-7 mai 1991
Université Bordeaux I, Bordeaux, France

Actes de l'atelier

Département de mathématiques et d'informatique



Université du Québec à Montréal

L' " Atelier de Combinatoire franco-québécois " a eu lieu à Bordeaux les 6 et 7 mai 1991, à la suite du 3e Colloque " Séries Formelles et Combinatoire Algébrique ". Il a été organisé conjointement par le LaBRI (Université Bordeaux I) et le LACIM (Université du Québec à Montréal).

Dépôt légal, premier semestre 1992, Bibliothèque nationale du Québec.

ISBN 2-89276-101-8 LACIM Montréal

© LACIM, Montréal, Mai 1992.

Laboratoire de combinatoire et d'informatique mathématique
Département de mathématiques et d'informatique
Université du Québec à Montréal
C.P. 8888, Succ. A
Montréal, Qc.
Canada H3C 3P8



AVANT-PROPOS

L' « Atelier de Combinatoire franco-québécois » a eu lieu à Bordeaux les 6 et 7 mai 1991, à la suite du 3e Colloque « Séries Formelles et Combinatoire Algébrique ». Il a été organisé conjointement par le LaBRI (Université Bordeaux I) et le LACIM (Université du Québec à Montréal).

Les sujets développés furent les mêmes que ceux du colloque, c'est-à-dire en liaison avec le thème « séries formelles » : d'une part, la combinatoire énumérative (cartes planaires, animaux dirigés, chemins de Dyck, p, q -nombres de Stirling, graphes 2- et 3-connexes), la combinatoire algébrique (groupe de Magnus, congruences plaxiques, super-algèbres de Lie-Jordan), et d'autre part, des sujets connexes (chaînes d'additions, coefficients de linéarisation de polynômes orthogonaux, probabilités discrètes et codes sémaphores, ensembles hybrides, graphes sur les alphabets).

Montréal, avril 1992

Jacques Labelle, Jean-Guy Penaud

TABLE DES MATIÈRES

Arquès D., Jacques I. (LIB, Besançon), , Orioux K. (LIM, Mulhouse) <i>Une identité sur les graphes et son application aux cartes planaires pointées</i>	1
Barcucci E., Pinzani R. (Univ. de Florence), Sprugnoli R. (Univ. de Padoue) <i>Génération aléatoire des animaux dirigés</i>	17
Bergeron A. (LACIM, Montréal) <i>Symbolic Computation and Discrete Probabilities</i>	27
Bond J. (LaBRI, Bordeaux) <i>Graphs on Alphabets</i>	43
Briek S., Malette R. (LACIM, Montréal) <i>Sur le calcul des chaînes d'additions optimales</i>	71
De Médicis A., Leroux P. (LACIM, Montréal) <i>Généralisations des nombres de Stirling et p, q-analogues</i>	87
Duchamp G. (LITP, Paris, LIR, Rouen) <i>Groupe de Magnus: applications</i>	105
Han G.-H. (ULP, Strasbourg) <i>Symétries trivariées sur les nombres de Genocchi</i>	119
Jacob G. (LIFL, Lille) <i>Lie-Jordan superalgebras and complete free finitely generated superalgebras</i>	135
Labelle G., Laforest L. (LACIM, Montréal) <i>Variations combinatoires autour des arborescences hyperquaternaires</i>	153
Labelle J. (LACIM, Montréal) <i>Quelques problèmes concernant les chemins de Dyck généralisés</i>	167
Lavialle O., Vidal Cl. (LaBRI-ENITA, Bordeaux) <i>Algorithmes d'obtention d'ordres à distance minimum d'un ensemble de graphes de surclassement</i>	177
Loeb D. (LaBRI, Bordeaux), Damiani E., D'Antona O. (Univ. de Milan) <i>Getting Results with Negative Thinking</i>	191
Malvenuto C. (LACIM, Montréal) <i>P-partitions and the Plactic Congruence</i>	215
Robinson R. W. (Univ. of Georgia, Athens), Walsh T.R. (LACIM, Montréal) <i>Cycle Index Series Relations for Counting 2- and 3- Connected Graphs</i>	227
Zeng J. (ULP, Strasbourg) <i>Counting a Pair of Permutations and the Linearization Coefficients for Jacobi Polynomials</i>	243

UNE IDENTITE SUR LES GRAPHE ET SON APPLICATION AUX CARTES PLANAIRES POINTEES

Didier ARQUES & Isabelle JACQUES

Laboratoire d'Informatique de Besançon
Université de Franche-Comté
16 Route de Gray
25030 BESANÇON Cedex.

Karine ORIEUX

Laboratoire d'Informatique de Mulhouse
Université de Haute Alsace
4 Rue des Frères Lumière
68093 MULHOUSE Cedex.

ABSTRACT

For any graph with edges weighted by formal variables, we consider the following "linear evolution" problem. For each n (time) and for x and y (vertices of the graph), $H(n,x,y)$ is, at time 0 the unity value on x , and at time $n+1$ the sum of the $H(n,x,z)$ weighted by the variables associated to the edges (z,y) going in y . There are two different forms of the unique solution of this problem, leading to an identity closely related to the underlying graph. The first form is usual and comes from linear algebra. The second form, very different and using continued fractions of Dyck, is obtained by a totally new approach issued from evolution problems. A particular case of this problem in the case of graph \mathbb{Z} , gives the well known three terms recurrence.

An application in the combinatorics of maps is then proposed, using as underlying graph, the infinite tree, associated to the family of rooted planar maps (well labeled trees), weighted by appropriately chosen formal variables. We then establish a new functional relation for rooted planar maps, linking the generating series of rooted planar maps with the generating series of Dyck words.

RESUME

Etant donné un graphe quelconque dont les arcs sont valués par des variables formelles, on considère le problème d' "évolution linéaire" suivant : déterminer pour tout n (temps), x et y (sommets du graphe), $H(n,x,y)$, sachant qu'à l'instant 0, $H(0,x,y)$ est la valeur unité en x , et qu'à l'instant $n+1$, $H(n+1,x,y)$ est la somme des $H(n,x,z)$ pondérés par les variables associées aux arcs (z,y) entrant en y . On peut donner deux formes différentes de la solution unique de ce problème, d'où une identité intimement associée au graphe sous-jacent au problème. La première forme est classique et issue de l'algèbre linéaire. La seconde forme, totalement différente et impliquant l'apparition de fractions continues de Dyck est obtenue par une approche totalement nouvelle issue des problèmes d'évolution. Un cas particulier de ce problème est le cas où le graphe est \mathbb{Z} , celui de la récurrence à trois termes.

Une application dans le cadre de la combinatoire des cartes est alors proposée, utilisant comme graphe, l'arbre infini sous-jacent à la famille des cartes planaires pointées (arbres bien étiquetés), valué par des variables formelles convenablement choisies. On en déduit une nouvelle relation fonctionnelle sur les cartes planaires pointées, liant la série génératrice des cartes planaires pointées à la série génératrice des mots de Dyck.

INTRODUCTION

De nombreux problèmes en mathématiques concernent des phénomènes d'évolution, c'est-à-dire dépendant du temps. D'une certaine façon, le problème résolu dans cet article, peut être vu comme celui d'une équation d'évolution, traitée sous un aspect calcul formel, totalement discrétisé et combinatoire. La technique utilisée dans cet article peut également être rapprochée de la classique méthode des éléments finis utilisée en analyse numérique pour modéliser un tel problème d'évolution. La différence étant que l'analyse numérique réalisera des calculs sur des valeurs concrètes alors que les résultats obtenus ici sont réalisés au niveau du calcul formel.

Une autre façon d'interpréter les objets combinatoires et les résultats présentés dans cet article est issue de l'algèbre linéaire. On verra en particulier, que la série formelle $\mathfrak{H}(t,x,y)$ définie ci-dessous est le terme général de l'inverse d'une matrice.

Cependant, cet aspect n'est pas celui qui a guidé l'obtention des résultats de cet article qui présente les résultats suivants :

Au Paragraphe I., une identité nouvelle, intimement associée à tout graphe quelconque est d'abord établie. Précisément, étant donné un graphe quelconque dont les arcs sont valués par des lettres d'un alphabet \mathcal{U} , on considère le problème d' "évolution linéaire" ainsi défini :

Déterminer pour tout n (entier représentant le temps), x et y (sommets de G), $H(n,x,y)$ dans $\mathbb{N}[[\mathcal{U}]]$, sachant qu'à l'instant 0, $H(0,x,y)$ est la valeur unité en x , c'est-à-dire $H(0,x,y) = \mathbb{1}_{\{x\}}(y)$ et sachant qu'à l'instant $n+1$, $H(n+1,x,y)$ est la somme des $H(n,x,z)$ pondérés par les variables associées aux arcs (z,y) entrant en y .

Il s'avère que l'on peut donner deux formes différentes de la solution (à l'évidence unique) de ce problème d'évolution, d'où une identité intimement associée au graphe sous-jacent au problème :

La première forme, que l'on trouvera par exemple dans [CF69] ou [V185], est très classique et issue de l'algèbre linéaire. $H(n,x,y)$ est dans ce cadre la somme pour tous les chemins de longueur n allant de x à y dans le graphe G , des monômes associés à ces chemins (ce monôme est le produit des valuations des arcs constituant le chemin considéré).

La seconde forme, totalement différente et impliquant l'apparition de fractions continues de Dyck est obtenue par une approche totalement nouvelle issue des problèmes d'évolution. Ainsi l'interprétation des termes qui constituent cette seconde forme sera de nature "markovienne" (c'est-à-dire que la valuation d'un arc le long d'un chemin sera fonction de l'arc qui le précède dans le chemin) par opposition à la situation de l'algèbre linéaire où la valuation d'un chemin est le produit des valuations de ses arcs (situation de nature "indépendante").

Un cas particulier de ce problème est le cas où le graphe sous-jacent est \mathbb{Z} , celui de la bien connue récurrence à trois termes : l'identité, dans ce cas particulier, se déduit essentiellement de l'identité entre la fraction continue et son développement sous forme d'une série formelle.

Les applications de cette équation sont potentiellement nombreuses. On peut songer à des applications à des réseaux de neurones linéaires par exemple.

Au Paragraphe II., est présentée une application dans le cadre de la combinatoire des cartes : on détermine, dans ce paragraphe, l'identité précitée de ce problème d'évolution associé à l'arbre infini (cf [Ar85]) sous-jacent à la famille des cartes planaires pointées (arbres bien étiquetés) (cf [CV81], [Ar86]), valué par des variables formelles convenablement choisies. On en déduit une nouvelle relation fonctionnelle sur les cartes planaires pointées, liant la série génératrice des cartes planaires pointées à la série génératrice des mots de Dyck.

Les méthodes usuelles dans l'étude des cartes planaires pointées concernent la définition de transformations topologiques sur les cartes à partir desquelles on déduit des équations sur les séries génératrices. Cette approche, initiée par W. T. Tutte, a donné la première équation (cf [Tu68]). Une seconde équation totalement différente dans sa forme mais également obtenue par des techniques topologiques a été obtenue par D. Arquès (cf [Ar85]). La troisième équation obtenue ici pour la série génératrice des cartes planaires est totalement nouvelle dans sa technique d'obtention et dans sa forme.

I. UN PROBLEME D'EVOLUTION

I.1. Fonction d'évolution associée à un graphe

DEFINITION : Graphe muni d'une valuation formelle (G, α) .

1. On se donne un multigraphe non orienté $G = (X, U)$ quelconque (éventuellement infini), où l'on rappelle que :

* X représente l'ensemble des *sommets* du graphe G ,

* U représente l'ensemble des *arcs* du graphe G (les arcs multiples et les boucles sont autorisés).

G étant non orienté, tout arc i de U admet un arc opposé, noté $-i$ dans la suite.

2. On associe à chaque arc i de G une variable formelle $\alpha(i)$.

On note \mathcal{A} l'alphabet constitué par les lettres $\alpha(i)$ codant les arcs de G .

REMARQUES

1. Dans le cadre d'applications, les variables $\alpha(i)$ peuvent formellement représenter le coefficient de transmission électrique si le graphe est un réseau de neurones, le flux autorisé par l'arc dans un problème de transport ou une variable d'une série génératrice intervenant dans un problème d'énumération naturellement lié au graphe. C'est dans ce dernier cadre que les résultats seront appliqués au Paragraphe II.

2. Le cas d'un multigraphe orienté est inclus dans cette étude. En effet, si i est un arc du graphe G et si $-i$ n'est pas arc de G , il suffit de prendre $\alpha(-i) = 0$.

DEFINITION : Fonction d'évolution H associée à (G, α) .

On considère les fonctions : $y \in X \rightarrow H(n, x, y) \in \mathbb{N}[\mathcal{A}]$

où : $n \in \mathbb{N}$ est le temps,

x est un sommet (donné au départ) du graphe G ,

y est un sommet générique du graphe G ,

$\mathbb{N}[\mathcal{A}]$ est l'ensemble des polynômes formels en les variables commutatives de \mathcal{A} à coefficients entiers.

H étant définie par (cf Figure 1) :

$$(I) \quad \left\{ \begin{array}{l} H(0, x, y) = \mathbb{1}_{\{x\}}(y) \\ H(n+1, x, y) = \sum_{j \in A(y)} \alpha(-j) H(n, x, T(j)) , n \geq 0 \end{array} \right.$$

où : $A(s)$ est l'ensemble des arcs issus de s dans G ,

$T(j)$ est l'extrémité terminale de l'arc j .

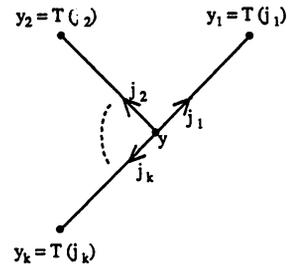


Figure 1

$\mathbb{1}_I(y)$ (notée également $\mathbb{1}_{\{y \in I\}}$) est la valeur prise en y de la fonction indicatrice de l'ensemble I :

$$\mathbb{1}_I(y) = 1 \text{ si } y \in I \quad \text{et} \quad \mathbb{1}_I(y) = 0 \text{ si } y \notin I$$

Intuitivement, $H(n, x, s)$ correspond à la sortie en s à l'instant n , lorsqu'à l'instant 0, on a injecté la valeur 1 en x dans le graphe G .

PROBLEME D'EVOLUTION : Il s'agit de déterminer H.

ASPECT EQUATION FONCTIONNELLE DE CE PROBLEME D'EVOLUTION

Selon une technique classique en combinatoire, on ramène (I) à une équation fonctionnelle en faisant intervenir la variable t dont l'exposant donne le temps n.

On pose :
$$\mathfrak{H}(t, x, y) = \sum_{n \geq 0} H(n, x, y) t^n$$

On en déduit facilement l'équivalent du système (I) sous forme d'une relation fonctionnelle :

(II)
$$\mathfrak{H}(t, x, y) = \mathbb{1}_{(x)}(y) + t \sum_{j \in A(y)} \alpha_{(-j)} \mathfrak{H}(t, x, T(j))$$

THEOREME 1

La relation fonctionnelle (II) admet une solution unique $\mathfrak{H}(t,x,y)$ dans $\mathbb{N}[\mathcal{Q}][[t]]$ (ensemble des séries formelles en la variable commutative t à coefficients dans $\mathbb{N}[\mathcal{Q}]$).

DEMONSTRATION

Evidente : $\mathfrak{H}(t,x,y)$ est le point fixe unique d'une application contractante dans l'espace métrique complet $\mathbb{N}[\mathcal{Q}][[t]]$ muni de la distance ultramétrique classique définie sur les séries.

REMARQUES

1. Le problème a été posé dans le cadre d'un problème d'évolution, cadre qui a guidé sa résolution, en particulier dans la "Situation 2" (cf Paragraphe I.3. ci-dessous).

Dans le cadre de l'algèbre linéaire, une autre définition classique de $\mathfrak{H}(t,x,y)$ aurait pu être donnée : si l'on associe à tout chemin orienté C de x à y, le monôme constitué par la multiplication des valuations formelles $t.\alpha(i)$ de ses arcs i, $\mathfrak{H}(t,x,y)$ n'est autre que la série somme de tous ces monômes (c'est la "Situation 1" décrite au Paragraphe I.3. ci-dessous). Par ailleurs, si \mathcal{A} est la matrice d'adjacence formelle du multigraphe G (le terme $a_{x,y}$ de \mathcal{A} est la somme des valuations des arcs allant de x à y) alors la matrice $(\mathfrak{H}(t,x,y))_{x,y}$ n'est autre que $(Id - t \mathcal{A})^{-1}$. On en déduit l'expression bien classique $\mathfrak{H}(t,x,y) = N_{x,y} / \det (Id - t \mathcal{A})$, où $N_{x,y}$ est le mineur pour (x,y) de la matrice (cf [CF69], [Vi85]).

2. Cette représentation matricielle n'est cependant pas adaptée à la "Situation 2" définie Paragraphe I.3., d'où ce choix de présentation du problème.

I.2. Résolution dans le cas particulier du graphe infini $G = \mathbb{Z}$

Dans ce cas particulier, la relation (I) est une récurrence à trois termes, dont la solution est classiquement donnée par une fraction continue.

DESCRIPTION DU PROBLEME

* On considère le cas où α_+ et α_- étant des variables données,

\forall arc $(y, y+1), y \in \mathbb{Z}$: $\alpha(y,y+1) = \alpha_+$ et $\alpha(y+1,y) = \alpha_-$

sont indépendantes de y.

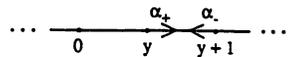


Figure 2

* Il y a alors invariance par translation : $\forall x$ et $y \in \mathbb{Z}$ $\mathfrak{H}(t, x, y) = \mathfrak{H}(t, 0, y - x)$

SOLUTION DE CE PROBLEME SUR \mathbb{Z}

THEOREME 2 : Détermination de \mathfrak{H} .

Si l'on note d la série génératrice des mots de Dyck, solution de l'équation $d(z) = 1 + zd^2(z)$, et :

$$\lambda_- = \lambda(\alpha_-, \alpha_-) = t \cdot \alpha_- \cdot d(t \cdot \alpha_- \cdot t \cdot \alpha_-) = \frac{t \alpha_-}{1 - \frac{t^2 \alpha_- \alpha_-}{1 - \frac{t^2 \alpha_- \alpha_-}{\dots}}}$$

$$\lambda_+ = \lambda(\alpha_+, \alpha_+) = t \cdot \alpha_+ \cdot d(t \cdot \alpha_+ \cdot t \cdot \alpha_+) = \frac{t \alpha_+}{1 - \frac{t^2 \alpha_+ \alpha_+}{1 - \frac{t^2 \alpha_+ \alpha_+}{\dots}}}$$

$\mu = \alpha_- \lambda_+ = \alpha_+ \lambda_-$

On a : 1. $\mathfrak{H}(t, 0, 0) = \frac{1}{1 - 2\mu t}$

2. $\forall i > 0 :$

$$\left. \begin{aligned} \mathfrak{H}(t, 0, i) &= \lambda_+^i \mathfrak{H}(t, 0, 0) \\ \mathfrak{H}(t, 0, -i) &= \lambda_-^i \mathfrak{H}(t, 0, 0) \end{aligned} \right\}$$

DEMONSTRATION

2. Comme rappelé à la remarque précédente, $\mathfrak{H}(t,0,i), i>0$ (démonstration analogue pour $i<0$), est la somme sur tous les chemins allant de 0 à i sur \mathbb{Z} , des monôme constitués pour chacun de ces chemins de la multiplication des valuations formelles $t \cdot \alpha_+$ (resp. $t \cdot \alpha_-$) de ses arcs montants (resp. descendants). Par décomposition en fonction des derniers passages aux niveaux 0, 1, ..., $i-1$, on décompose un tel chemin en un chemin allant de 0 en 0 (terme générique de $\mathfrak{H}(t,0,0)$), puis en une montée (la dernière de 0 en 1 (terme $t \cdot \alpha_+$) suivi d'un chemin allant de 1 en 1 sans jamais passer en dessous du niveau 1 (c'est un mot de Dyck dont les pas montants et descendants sont respectivement codés par $t \cdot \alpha_+$ et $t \cdot \alpha_-$, donc le terme générique de $d(t \cdot \alpha_+ \cdot t \cdot \alpha_-)$) et ainsi de suite jusqu'au niveau i , d'où la formule :

$\mathfrak{H}(t, 0, i) = \mathfrak{H}(t, 0, 0) \cdot (t \alpha_+ \cdot d(t^2 \alpha_- \alpha_+))^i = \mathfrak{H}(t, 0, 0) \lambda_+^i$

1. Pour $i = 0$, le résultat se déduit de 2. et de l'équation :

$\mathfrak{H}(t, 0, 0) = 1 + t \alpha_+ \mathfrak{H}(t, 0, -1) + t \alpha_- \mathfrak{H}(t, 0, 1)$

REMARQUE : L'équation $\mathfrak{H}(t, 0, i) = \lambda_+^i \mathfrak{H}(t, 0, 0)$ est un cas particulier dans \mathbb{Z} de l'identité générale annoncée dans l'introduction (cf § I.3.) : le membre de gauche est la somme classique sur tous les chemins allant de 0 à i , issue de l'algèbre linéaire, alors que le membre de droite peut être interprété autrement. Considérons la valuation de type markovienne suivante : un arc d'un chemin donné est valué λ_+ ou λ_- (resp. 0) s'il est montant ou descendant et précédé,

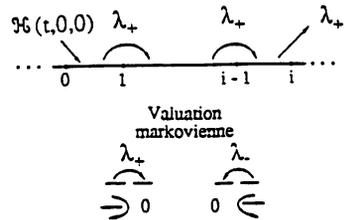


Figure 3

dans ce chemin, par un arc allant dans le même sens (resp. dans le sens contraire). Le premier arc du chemin issu de 0 sera valué par $\mathfrak{H}(t,0,0)$ et le dernier arc du chemin par λ_+ ou λ_- suivant qu'il est montant ou descendant. On constate alors (pour $i > 0$ par exemple) que le seul chemin ayant une contribution non nulle pour cette valuation est le chemin qui monte directement de 0 à i sans jamais faire de retour arrière, et cette contribution est le membre de droite de l'équation interprétée. Il s'avère que ce type de formule peut être généralisée au cas d'un multigraphe valué quelconque :

I.3. Résolution dans le cas général

Généralisant les valuations markoviennes décrites dans le cas de \mathbb{Z} dans la remarque précédente, on a les notations suivantes :

NOTATIONS

Les séries définies ci-dessous, sont formelles en les variables commutatives : $t, \alpha(i), i \in U$.

1. On note $\epsilon_{i,j}$ une série formelle appelée *coefficient de transmission* de l'arc i à l'arc j .
2. On note $v(x)$, une série formelle appelée *ordre de x* , associée à tout sommet x du graphe G .
On note $\phi(i)$, une série formelle associée à tout arc i du graphe G .
3. x et y étant deux sommets du graphe G , on note $\mathfrak{C}(x,y)$, l'ensemble des chemins orientés de x à y dans G , de longueur $n \geq 1$.

On notera : $C = (x_0 = x, x_1, \dots, x_n = y)$ un chemin dans $\mathfrak{C}(x,y)$, constitué des arcs : $i_1 = (x_0, x_1), \dots, i_n = (x_{n-1}, x_n)$ où i_n , dernier arc de C , sera noté $f(C)$.

On associe alors à un tel chemin C , la série formelle de transmission le long de C :

$$\delta(C) = \mathbb{1}_{n=1} + \mathbb{1}_{n \geq 2} \prod_{i=1}^{n-1} \epsilon_{i_i, i_{i+1}}$$

$\delta(C)$ n'est autre que le produit des coefficients de transmission le long des arcs de C joignant x à y dans le graphe G .

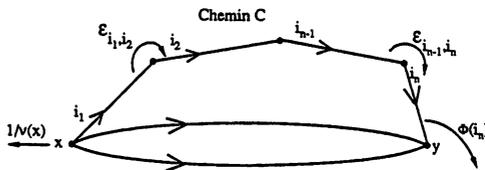


Figure 4

4. La solution sur \mathbb{Z} (cf § I.2.) "pour l'arc i " du graphe G , c'est-à-dire avec :
 $\alpha_+ = \alpha(i)$ et $\alpha_- = \alpha(-i)$ sera notée : $\lambda_+ = \lambda(\alpha(i), \alpha(-i)) = \lambda_i(t)$.
On a alors : $\lambda_- = \lambda_{-i}(t)$ (évident).

THEOREME 3 : La solution unique du problème d'évolution (II) est :

$$\mathfrak{H}(t,x,y) = \frac{1}{v(x)} \left[\mathbb{1}_{\{x\}}(y) + \sum_{C \in \mathfrak{C}(x,y)} \delta(C) \cdot \phi(f(C)) \right]$$

Cette formule étant prise dans l'une ou l'autre des deux Situations 1 et 2 :

<i>SITUATION 1</i> (algèbre linéaire)	$v(x) = 1$	$\phi(i) = t \cdot \alpha(i)$
	$\epsilon_{i,j} = t \cdot \alpha(i)$	si i est prédécesseur de j dans G
	$\epsilon_{i,j} = 0$	sinon

<i>SITUATION 2</i> (markovienne)	$v(x) = 1 - \sum_{j \in A(x)} \alpha(-j) \cdot t \cdot \lambda_j$	$\phi(i) = \alpha(i) (1 - \lambda_i \lambda_{-i}) \frac{t}{v(\Gamma(i))}$
	$\epsilon_{i,j} = \alpha(i) (1 - \lambda_i \lambda_{-i}) \frac{t}{v(\Gamma(i))} = \phi(i)$	si i est prédécesseur de j et $j \neq -i$
	$\epsilon_{i,j} = \alpha(i) (1 - \lambda_i \lambda_{-i}) \frac{t}{v(\Gamma(i))} - \lambda_i = \phi(i) - \lambda_i$	si i est prédécesseur de j et $j = -i$

REMARQUES

1. Dans la Situation 1, on a la situation classique de l'algèbre linéaire, l'expression de $\mathfrak{H}(t,x,y)$ donnée par le second membre du Théorème 3 étant alors la classique série somme des monômes (obtenus par multiplication des variables $t.\alpha(i)$ associées aux arcs i du chemin) associés aux chemins orientés de x à y .

2. Dans la Situation 2, le monôme $\delta(C).\phi(f(C))$ associé au chemin C , n'est pas multiplicatif comme dans la Situation 1, mais de type markovien, la quantité $\epsilon_{i,j}$ associée à l'arc j dépendant de son prédécesseur i dans le chemin C . Ce n'est plus la situation classique de l'algèbre linéaire.

3. S'il existe $i \in G / \alpha(i)$ et $\alpha(-i)$ sont différents de 0 alors λ_i et λ_{-i} sont différents de 0 dans la Situation 2, et donc les deux expressions de $\epsilon_{i,j}$ dans les Situations 1 et 2 sont effectivement profondément différentes. Par suite, ces deux situations conduisent, par le Théorème 3, à deux expressions différentes de la même fonctionnelle \mathfrak{H} .

On en déduit donc une relation fonctionnelle nouvelle associée au graphe G .

4. Dans le cas particulier du Paragraphe I.2. où $G = \mathbb{Z}$, $\mathfrak{H}(t,0,0)$ apparaît comme la fraction continue associée à une récurrence à trois termes particulière. Le Théorème 3 peut donc être vu comme une généralisation à un graphe quelconque de l'égalité entre une fraction continue (expression de $\mathfrak{H}(t,x,y)$ dans la Situation 2) et son développement en série formelle (expression de $\mathfrak{H}(t,x,y)$ dans la Situation 1).

Pour prouver le Théorème 3, le lemme technique ci-dessous sera nécessaire :

LEMME : Décomposition de chemins

Soient x et y deux sommets de G .

Alors l'ensemble des chemins de x à y est : $\mathbb{1}_{\{x\}}(y) + \mathfrak{C}(x,y)$

où $\mathfrak{C}(x,y)$, ensemble des chemins orientés de x à y , de longueur ≥ 1 , se décompose en :

$$\mathfrak{C}(x,y) = \sum_{j \in A(y)} \mathbb{1}_{\{x\}}(y_j) \cdot (-j) + \sum_{j \in A(y)} \mathfrak{C}(x,y_j) \cdot (-j)$$

avec : j l'arc générique ($y, y_j = T(j)$) issu de y ,

$A(y)$ l'ensemble des arcs issus de y dans G ,

" \cdot " l'opération de concaténation des chemins et

\sum l'union ensembliste disjointe.

DEMONSTRATION DU LEMME

Evidente (le premier terme du second membre de cette équation donne les chemins de longueur 1 et le second terme ceux de longueur ≥ 2).

DEMONSTRATION DU THEOREME 3

SITUATION 1. Cette situation est évidente et classique : laissée au lecteur.

SITUATION 2. On garde la notation du lemme précédent : j arc générique ($y, y_j = T(j)$) issu de y .

Montrons que l'équation (II) est vérifiée quand on remplace \mathfrak{H} par sa valeur donnée par le Théorème 3 dans la Situation 2. La quantité :

$$\sigma = \mathbb{1}_{\{x\}}(y) + t \sum_{j \in A(y)} \alpha(-j) \mathfrak{H}(t, x, y_j = T(j)) \quad \text{est égale, par le Théorème 3 à :}$$

$$\sigma = \mathbb{1}_{\{x\}}(y) + t \sum_{j \in A(y)} \alpha(-j) \frac{1}{v(x)} \left[\mathbb{1}_{\{x\}}(y_j) + \sum_{C_j \in \mathfrak{C}(x,y_j)} \delta(C_j) \cdot \phi(f(C_j)) \right]$$

8 Or, par définition de ϕ et du coefficient de transmission ϵ , on a : $\phi(f(C_j)) = \epsilon_{f(C_j), j} + \mathbb{1}_{\{j\}}(f(C_j)) \cdot \lambda_j$

Donc, en notant $C = C_j \cdot (-j)$, où $C_j \in \mathfrak{C}(x, y)$: $\delta(C_j) \cdot \phi(f(C_j)) = \delta(C) + \mathbb{1}_{\{j\}}(f(C_j)) \cdot \delta(C_j) \cdot \lambda_j$

De plus, le dernier terme de l'équation précédente n'est non nul que si : $f(C_j) = j$, c'est-à-dire si $C_j = d \cdot j$, avec d un chemin de x à y .

Donc, si l'on note $C_j = d \cdot j$ dans ce dernier terme, on a, en exprimant $\delta(C_j)$ en fonction de $\delta(d)$, en distinguant les cas où la longueur $|d|$ du chemin d est ≥ 1 (et donc : d appartient à $\mathfrak{C}(x, y)$) ou nulle :

$$\delta(C_j) \cdot \phi(f(C_j)) = \delta(C) + \mathbb{1}_{C_j = d \cdot j, |d| \geq 1} \delta(d) \epsilon_{f(d), j} \cdot \lambda_j + \mathbb{1}_{C_j = j} \lambda_j$$

Donc :
$$\sigma = \mathbb{1}_{\{x\}}(y) + t \sum_{j \in A(y)} \alpha(-j) \frac{1}{v(x)} \dots$$

$$\cdot \left[\mathbb{1}_{\{x\}}(y_j) + \sum_{C_j \in \mathfrak{C}(x, y)} \delta(C) + \sum_{C_j \in \mathfrak{C}(x, y)} \mathbb{1}_{C_j = d \cdot j, |d| \geq 1} \delta(d) \epsilon_{f(d), j} \cdot \lambda_j + \sum_{C_j \in \mathfrak{C}(x, y)} \mathbb{1}_{C_j = j} \lambda_j \right]$$

Or, par le lemme de décomposition de chemins, les deux premiers termes du crochet ci-dessus, se réécrivent, compte tenu de la définition de δ :

$$\frac{1}{v(x)} \sum_{j \in A(y)} \alpha(-j) \left[\mathbb{1}_{\{x\}}(y_j) + \sum_{C_j \in \mathfrak{C}(x, y)} \delta(C) \right] = \frac{1}{v(x)} \sum_{C \in \mathfrak{C}(x, y)} \alpha(f(C)) \delta(C)$$

Par ailleurs, dire que C_j dans $\mathfrak{C}(x, y)$ est de la forme $d \cdot j$ avec $|d| \geq 1$, est équivalent à dire que d appartient à $\mathfrak{C}(x, y)$, ce qui permet de réécrire le troisième terme du crochet :

$$t \sum_{j \in A(y)} \sum_{d \in \mathfrak{C}(x, y)} \alpha(-j) \frac{1}{v(x)} \delta(d) \epsilon_{f(d), j} \lambda_j$$

Enfin, dans le quatrième terme du crochet, C_j dans $\mathfrak{C}(x, y)$ est égal à $j = (y, y)$ si et seulement si $x = y$.

Ce quatrième terme se réécrit donc : $\mathbb{1}_{\{x\}}(y) \cdot t \sum_{j \in A(y)} \alpha(-j) \frac{1}{v(x)} \lambda_j$

On déduit de ces remarques :

$$\begin{aligned} \sigma &= \mathbb{1}_{\{x\}}(y) + t \sum_{C \in \mathfrak{C}(x, y)} \alpha(f(C)) \frac{1}{v(x)} \delta(C) + t \sum_{j \in A(y)} \sum_{d \in \mathfrak{C}(x, y)} \alpha(-j) \frac{1}{v(x)} \delta(d) \epsilon_{f(d), j} \lambda_j \\ &\quad + \mathbb{1}_{\{x\}}(y) t \sum_{j \in A(y)} \alpha(-j) \frac{1}{v(x)} \lambda_j \end{aligned}$$

En réécrivant le dernier terme à l'aide de la définition de $v(x)$ et en remplaçant la notation d par C dans le troisième terme, on a :

$$\sigma = \mathbb{1}_{\{x\}}(y) + \mathbb{1}_{\{x\}}(y) \cdot \frac{1 - v(x)}{v(x)} + \frac{1}{v(x)} \sum_{C \in \mathfrak{C}(x, y)} t \cdot \delta(C) \left[\alpha(f(C)) + \sum_{j \in A(y)} \alpha(-j) \epsilon_{f(C), j} \lambda_j \right]$$

Or : $\phi(f(C)) = \alpha(f(C)) (1 - \lambda_{f(C)} \lambda_{-f(C)}) \frac{1}{v(y)}$, soit : $t = \phi(f(C)) \frac{v(y)}{\alpha(f(C)) (1 - \lambda_{f(C)} \lambda_{-f(C)})}$, d'où :

$$\sigma = \frac{1}{v(x)} \mathbb{1}_{\{x\}}(y) + \frac{1}{v(x)} \sum_{C \in \mathfrak{C}(x, y)} \delta(C) \cdot \phi(f(C)) \cdot \frac{v(y)}{\alpha(f(C)) (1 - \lambda_{f(C)} \lambda_{-f(C)})} \left[\alpha(f(C)) + \sum_{j \in A(y)} \alpha(-j) \epsilon_{f(C), j} \lambda_j \right]$$

qui sera bien égale à l'expression donnant $\mathfrak{H}(t, x, y)$ au Théorème 3, Situation 2 si l'on vérifie l'égalité :

$$(*) \quad \frac{v(y)}{\alpha(f(C)) (1 - \lambda_{f(C)} \lambda_{-f(C)})} \cdot \left[\alpha(f(C)) + \sum_{j \in A(y)} \alpha(-j) \epsilon_{f(C), j} \lambda_j \right] = 1$$

c'est-à-dire :
$$\frac{1 - \lambda_{f(C)} \lambda_{-f(C)}}{v(y)} = 1 + \sum_{j \in A(y)} \frac{\alpha(-j)}{\alpha(f(C))} \varepsilon_{f(C),j} \lambda_j$$

Or $T(f(C)) = y$ (extrémité terminale de $f(C)$), donc :

$$\varepsilon_{f(C),j} = \alpha(f(C)) \left(1 - \lambda_{f(C)} \lambda_{-f(C)}\right) \frac{t}{v(y)} \quad \text{si } j \neq -f(C)$$

$$\varepsilon_{f(C),j} = \alpha(f(C)) \left(1 - \lambda_{f(C)} \lambda_{-f(C)}\right) \frac{t}{v(y)} - \lambda_{f(C)} \quad \text{si } j = -f(C)$$

Donc (*) se réécrit :
$$\frac{1 - \lambda_{f(C)} \lambda_{-f(C)}}{v(y)} = 1 + \left(\sum_{j \in A(y)} \alpha(-j) \left(1 - \lambda_{f(C)} \lambda_{-f(C)}\right) \frac{t}{v(y)} \lambda_j \right) - \lambda_{f(C)} \lambda_{-f(C)}$$

Le dernier terme étant issu du terme en $-\lambda_{f(C)}$ de $\varepsilon_{f(C),j}$ dans le cas où $j = -f(C)$.

Donc (*) se simplifie en :
$$\frac{1 - \lambda_{f(C)} \lambda_{-f(C)}}{v(y)} = \left(1 - \lambda_{f(C)} \lambda_{-f(C)}\right) \left(1 + \sum_{j \in A(y)} \alpha(-j) \frac{t}{v(y)} \lambda_j\right)$$

Soit : $\frac{1}{v(y)} = 1 + \sum_{j \in A(y)} \alpha(-j) \frac{t}{v(y)} \lambda_j$, d'où : $1 = v(y) + \sum_{j \in A(y)} \alpha(-j) \cdot t \cdot \lambda_j$, exact par définition de v .

CQFD.

REMARQUE

La démonstration présentée ici, montre que les deux situations proposées sont bien solutions du problème. Il est bien évident que trouver la forme explicite de ces deux solutions (en particulier, celle de la Situation 2) n'est pas le fruit du hasard, mais a nécessité une autre approche dans laquelle la forme de la solution \mathcal{K} est exprimée à l'aide de séries inconnues intervenant dans des produits de convolution liés aux noeuds du graphe. Cette approche, semble nouvelle dans le cadre du calcul formel (cf [AJ91a]).

II. UNE RELATION FONCTIONNELLE SUR LES CARTES

II.1. Définitions et résultats relatifs aux cartes planaires pointées

On va maintenant étudier le problème d'évolution présenté au Paragraphe I., avec pour graphe G , l'arbre infini I naturellement associé aux cartes planaires pointées (cf Figure 5).

Nous rappelons d'abord quelques définitions et résultats :

Une *carte planaire* est une partition de la sphère de \mathbb{R}^3 en trois ensembles finis de cellules : les sommets qui sont représentés par des points, les arêtes qui sont représentées par des arcs de Jordan ouverts, simples, deux à deux disjoints et dont les extrémités (confondues ou non) sont des sommets, les faces qui sont des domaines simplement connexes dont les frontières sont constituées d'arêtes et de sommets. On choisit sur cette carte un arc (ou brin) appelé *brin distingué*. Deux telles cartes planaires munies d'un brin distingué sont dites *équivalentes* s'il existe un homéomorphisme de la sphère, préservant son orientation qui échange sommets, arêtes, faces et brins distingués des deux cartes. On appellera alors *carte planaire pointée* une telle classe d'équivalence de cartes planaires munies d'un brin distingué.

On notera $R_0(z)$ la série génératrice du nombre de cartes planaires pointées dans laquelle le degré de z décompte le nombre d'arêtes de la carte planaire pointée.

Un *arbre bien étiqueté* est un arbre plongé dans le plan dont les sommets sont étiquetés par des entiers naturels, de racine étiquetée 0 et dont deux sommets adjacents ont des étiquettes qui diffèrent au plus d'une unité.

L'*arbre infini I* associé à la famille des arbres bien étiquetés est l'arborescence infinie plongée dans le plan, de racine r étiquetée 0, dont tout sommet étiqueté 0 a exactement deux fils gauche et droit respectivement étiquetés 0 et 1, et dont tout sommet étiqueté $i \geq 1$ a exactement trois fils étiquetés $i-1, i, i+1$ dans cet ordre de gauche à droite.

On a les résultats suivants (cf [Ar85], [Ar86] et [CV81]) illustrés par la Figure 5 :

- THEOREME 4 :** Les trois ensembles suivants sont en bijection (bijections laissant invariant le nombre d'arcs dans la carte, l'arbre et le circuit associés) :
- l'ensemble des cartes planaires pointées,
 - l'ensemble des arbres bien étiquetés et
 - l'ensemble des circuits allant de la racine à la racine sur l'arbre infini I.

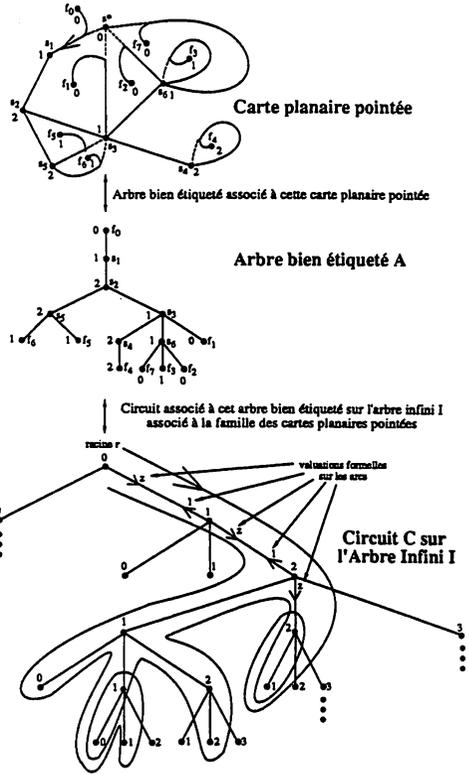


Figure 5

DEMONSTRATION : Se reporter à [Ar85], [Ar86], [CV81].

Nous rappelons seulement ici l'algorithme de passage d'un circuit C à un arbre A :

Lors du parcours du circuit C issu de la racine de I :

- * lorsque l'on descend le long d'un arc a de l'arbre infini I, on crée dans l'arbre bien étiqueté A en construction à partir de la position courante, un nouvel arc a' , dont l'extrémité finale est une feuille d'étiquette celle de l'extrémité finale de a dans I,
- * lorsque l'on remonte le long d'un arc a de l'arbre infini I, on remonte dans A le long de l'arc a' qui lui a été associé lors de la précédente descente de a dans I.

II.2. Fonction \mathfrak{K} associée à l'arbre infini I

Afin d'appliquer les résultats du Paragraphe I. au graphe constitué par l'arbre infini I, on associe (cf Figure 5) à tout arc descendant (resp. montant) i de I, la valuation $\alpha(i) = z$ (resp. 1).

On imposera par ailleurs dans la suite $t = 1$ dans l'expression de la fonction \mathfrak{K} .

La série génératrice $\mathfrak{H}(1,r,r)$ (cf Remarque en fin de Paragraphe I.1. et Situation 1 du Paragraphe I.3.) est la série somme pour tous les circuits C allant de la racine r à la racine r dans I, des monômes multiplicatifs $m(C)$ ($m(C)$ est la multiplication des étiquettes associées à chaque arc de C). De par la valuation choisie sur I, $m(C)$ compte (variable z) le nombre d'arcs descendants de C, c'est-à-dire le nombre d'arêtes de l'arbre bien étiqueté et de la carte associée par les bijections du Théorème 4. Donc :

PROPOSITION 1 : r étant la racine de l'arbre I et $R_0(z)$ la série génératrice des cartes planaires pointées décomptées en fonction du nombre d'arêtes, on a : $\mathfrak{H}(1, r, r) = R_0(z)$

EXEMPLE

Monôme associé au circuit Figure 5 : $m(C) = z.z.z.z.1.z.1.1.z.z.z.1.1.z.z.1.z.1.z.1.1.1.1.1 = z^{13}$
 Le degré de z donne le nombre d'arêtes dans l'arbre bien étiqueté et dans la carte associés.

II.3. Une nouvelle relation fonctionnelle sur les cartes planaires pointées

En utilisant l'expression de \mathfrak{H} dans la Situation 2, on va déduire de la Proposition 1 précédente, une nouvelle expression pour R_0 . Cette nouvelle expression de \mathfrak{H} nécessite de déterminer les expressions de v et des coefficients de transmission ϵ :

II.3.1. Détermination de v et $\epsilon_{i,j}$ pour l'arbre infini I

La simplification des formules donnant v et $\epsilon_{i,j}$ (Lemmes 2 et 3) utilise le Lemme 1 (évident) suivant :

LEMME 1 : Avec $t = 1$, on déduit du Théorème 2 (cf § I.2.) les formules sur $\lambda_+ = \lambda(\alpha_+, \alpha_-)$

1. $\lambda_+ = \frac{\alpha_+}{1 - \alpha_- \lambda_+}$ et $\frac{1}{\alpha_+} \lambda(\alpha_+, \alpha_-) = \frac{1}{\alpha_-} \lambda(\alpha_-, \alpha_+) = \frac{1}{\alpha_+ \alpha_-} \lambda(\alpha_+ \alpha_-, 1) = \lambda(1, \alpha_+ \alpha_-)$
2. $\lambda(1, \alpha_+ \alpha_-) = d(\alpha_+ \alpha_-)$, où d est la série génératrice des mots de Dyck ($d(z) = 1 + z d^2$)

LEMME 2 : Expression de $v(s)$.

On distingue 3 situations pour le sommet s dans l'arbre infini I (cf Figures 5 et 6) donnant des expressions différentes pour $v(s) = 1 - \sum_{j \in A(s)} \alpha(-j) \lambda_j$, (d est la série génératrice des mots de Dyck) :

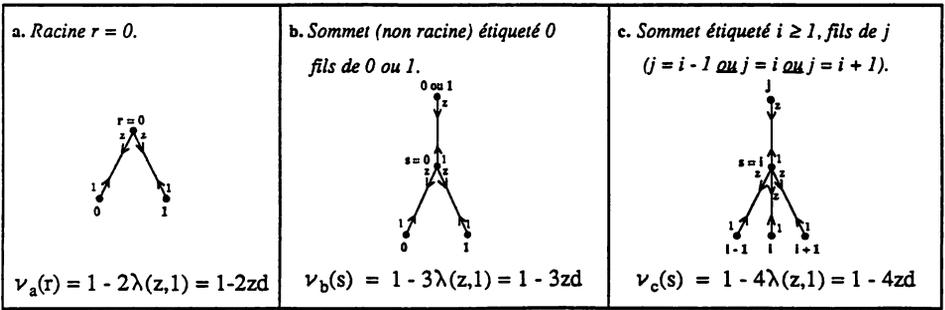


Figure 6

DEMONSTRATION : Evidente à partir du Lemme 1.

LEMME 3 : Soit un circuit C associé à un arbre bien étiqueté A. On donne ci-dessous les différentes situations possibles dans I pour les arcs successifs i et j, donnant des valeurs différentes pour les coefficients $\epsilon_{i,j}$ et créant des sommets de natures différentes dans A par l'algorithme du Théorème 4 :

<p>Cas 1. L'arc i remonte en la racine r depuis k (k = 0 ou k = 1)</p>			
<p>1.1</p>	<p>$\epsilon_{i,-j} = 0 \Rightarrow \delta(C) = 0$</p> <p>Création dans A d'un nouveau fils à la racine d'étiquette égale à celle de son frère gauche. Un arbre bien étiqueté ayant deux fils successifs de la racine de même étiquette aura une contribution 0 dans $\mathcal{N}(1,r)$.</p>	<p>1.2</p>	<p>$\epsilon_{i,j} = \frac{1 - z d^2}{1 - 2zd}$</p> <p>Création dans A d'un nouveau fils à la racine d'étiquette différente de celle de son frère gauche. On note p la variable comptant le nombre de fils à la racine dans A.</p>
<p>Cas 2. L'arc i remonte en un sommet 0 (fils de 0 ou 1, donc non racine) depuis k (k = 0 ou k = 1).</p>			
<p>2.1</p>	<p>$\epsilon_{i,j} = \frac{1 - z d^2}{1 - 3zd}$</p> <p>Création dans A d'un fils d'un sommet non racine étiqueté 0, fils qui est soit fils gauche, soit de frère à gauche d'étiquette différente. On note δ_0 la variable qui décompte le nombre de ces sommets dans A.</p>	<p>2.2</p>	<p>$\epsilon_{i,-j} = \frac{z d^2}{1 - 3zd}$</p> <p>Création dans A d'un fils d'un sommet non racine étiqueté 0, fils dont le frère à gauche a la même étiquette (0 ou 1). Un tel sommet sera appelé <i>sommet dupliqué</i>. On note σ_0 la variable qui décompte le nombre de ces sommets dans A.</p>
<p>Cas 3. L'arc i remonte en un sommet k ≥ 1 depuis l (l = k - 1 ou l = k ou l = k + 1).</p>			
<p>3.1</p>	<p>$\epsilon_{i,j} = \frac{1 - z d^2}{1 - 4zd}$</p> <p>Création dans A d'un fils d'un sommet étiqueté k ≥ 1, fils qui est soit fils gauche, soit de frère à gauche d'étiquette différente. On note δ_1 la variable qui décompte le nombre de ces sommets dans A.</p>	<p>3.2</p>	<p>$\epsilon_{i,-j} = \frac{2z d^2}{1 - 4zd}$</p> <p>Création dans A d'un sommet dupliqué, fils d'un sommet étiqueté k ≥ 1. On note σ_1 la variable qui décompte le nombre de ces sommets dans A.</p>
<p>Cas 4. L'arc i descend en un sommet 0 depuis 0 ou 1.</p>			
<p>4.1</p>	<p>$\epsilon_{i,j} = \frac{z(1 - z d^2)}{1 - 3zd}$</p> <p>Création dans A d'un sommet étiqueté 0 (non racine et non feuille). On note ω_0 la variable qui décompte le nombre de ces sommets dans A.</p>	<p>4.2</p>	<p>$\epsilon_{i,-j} = \frac{z^2 d^2}{1 - 3zd}$</p> <p>Création dans A d'une feuille étiquetée 0. On note φ_0 la variable qui décompte le nombre de ces sommets dans A.</p>
<p>Cas 5. L'arc i descend en un sommet k ≥ 1 depuis l ≥ 0.</p>			
<p>5.1</p>	<p>$\epsilon_{i,j} = \frac{z(1 - z d^2)}{1 - 4zd}$</p> <p>Création dans A d'un sommet étiqueté k ≥ 1 (non feuille). On note ω_1 la variable qui décompte le nombre de ces sommets dans A.</p>	<p>5.2</p>	<p>$\epsilon_{i,-j} = \frac{2z^2 d^2}{1 - 4zd}$</p> <p>Création dans A d'une feuille étiquetée k ≥ 1. On note φ_1 la variable qui décompte le nombre de ces sommets dans A.</p>

Figure 7

DEMONSTRATION

On vérifie facilement que tous les cas possibles ont été envisagés dans le calcul des coefficients de transmission. Il est facile de contrôler les séries formelles donnant ces $\epsilon_{i,j}$. Il reste à justifier (à l'aide de l'algorithme du Théorème 4) l'association à chaque situation d'un paramètre combinatoire dans l'arbre A. Montrons par exemple, que le nombre de couples (i,j) d'arcs successifs de C appartenant au cas 2.2 est bien égal au nombre de sommets dupliqués, fils d'un sommet étiqueté 0 non racine, dans A (c'est-à-dire au degré de la variable σ_0). Les autres cas ont un traitement analogue.

Par la bijection entre circuit et arbre bien étiqueté, permettant de construire A à partir de C, un nouveau sommet s est créé dans l'arbre, lorsque l'on descend le long d'un arc $j = (\alpha, \beta)$ de I avec $\beta = \alpha - 1, \alpha$ ou $\alpha + 1$). Son étiquette sera l'étiquette β du sommet atteint dans I par cette descente. Par ailleurs, ce nouveau sommet créé dans A, sera frère d'un sommet t étiqueté γ de A, si l'arc i précédent j dans C, est un arc remontant dans I (depuis le sommet étiqueté γ) (évident par l'algorithme construisant A à partir de C).

Dès lors, s sera un sommet dupliqué étiqueté 0 (resp. 1) (avec $\beta = \gamma = 0$ (resp. $\beta = \gamma = 1$)) si il est associé aux arcs successifs i et j avec i montant depuis 0 (resp. 1) dans I et j redescendant en 0 (resp. 1), c'est-à-dire $i = -j$ (car i et j sont successeurs dans I). Par ailleurs, s sera fils d'un sommet non racine étiqueté 0 si le sommet final de i (et initial de j) est étiqueté 0 et est différent de la racine r de I.

Le cas 2.2 du Lemme 3 est exactement celui satisfaisant ces conditions.

REMARQUE : Nouveaux paramètres dans l'étude combinatoire des arbres bien étiquetés

Le Lemme 3 a montré que les différentes situations possibles pour deux arcs successifs i, j dans le circuit C sur l'arbre infini I, se regroupent par paquets donnant des expressions identiques pour $\epsilon_{i,j}$, et correspondant à la création d'un type précis de paramètres combinatoires caractérisant l'arbre bien étiqueté (feuilles étiquetées 0, ...) associé à C par l'algorithme du théorème 4. Par ailleurs, le cas 1.1 du Lemme 3 montre que les seuls arbres bien étiquetés (bijectivement associés aux circuits dans I) ayant une contribution non nulle dans $\mathcal{K}(1,r,r)$, sont ceux dont deux fils successifs de la racine sont d'étiquettes différentes. On note F_0 la sous-famille de ces arbres bien étiquetés.

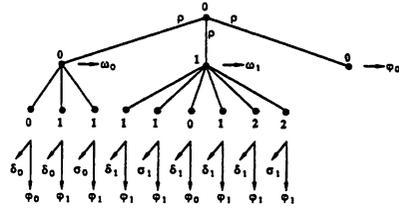
II.3.2. Une nouvelle relation fonctionnelle pour la série génératrice R_0 des cartes planaires pointées

F_0 étant la sous-famille des arbres bien étiquetés dont deux fils successifs de la racine (dans l'ordre où on les rencontre de gauche à droite) n'ont pas la même étiquette, on lui associe sa série génératrice $\Theta(\rho, \sigma_0, \sigma_1, \varphi_0, \varphi_1, \delta_0, \delta_1, \omega_0, \omega_1)$, somme des monômes (en les variables, définies Figure 7 : $\rho, \sigma_0, \sigma_1, \varphi_0, \varphi_1, \delta_0, \delta_1, \omega_0, \omega_1$) associés aux arbres de F_0 .

EXEMPLE

Monôme dans Θ associé à l'arbre de la Figure 8 :

$$m = \rho^3 \sigma_0^1 \sigma_1^2 \varphi_0^3 \varphi_1^7 \delta_0^2 \delta_1^4 \omega_0^1 \omega_1^1$$



REMARQUE

Le monôme $m(A)$ associé à un arbre A de F_0 a pour degré total, la somme des degrés des sommets de A .

Figure 8 : Etiquettes associées aux sommets.

Ceci est évident si l'on remarque que :

$$\begin{cases} d^\circ \delta_0 + d^\circ \omega_0 + d^\circ \sigma_0 = \text{Somme des degrés des sommets étiquetés 0 (non racine et non feuille)} = s_0 \\ d^\circ \delta_1 + d^\circ \omega_1 + d^\circ \sigma_1 = \text{Somme des degrés des sommets étiquetés } \geq 1 \text{ (non feuille)} = s_1 \end{cases}$$

D'où : $d^\circ \rho + d^\circ \sigma_0 + d^\circ \sigma_1 + d^\circ \varphi_0 + d^\circ \varphi_1 + d^\circ \delta_0 + d^\circ \delta_1 + d^\circ \omega_0 + d^\circ \omega_1 = d^\circ \rho + (d^\circ \varphi_0 + d^\circ \varphi_1) + (s_0 + s_1)$
 = Degré racine + Somme degrés des feuilles + Somme degrés des sommets (non racine et non feuille)
 = Somme des degrés des sommets de A = Nombre de brins de A .

THEOREME 5 : Lien entre les séries génératrices des cartes planaires pointées R_0 et de Dyck R_0, Θ et d sont liées par l'équation :

$$R_0(z) = \frac{1}{1-2zd} \Theta \left(\frac{1-zd^2}{1-2zd}, \frac{zd^2}{1-3zd}, \frac{2zd^2}{1-4zd}, \frac{z^2d^2}{1-3zd}, \frac{2z^2d^2}{1-4zd}, \frac{1-zd^2}{1-3zd}, \frac{1-zd^2}{1-4zd}, \frac{z(1-zd^2)}{1-3zd}, \frac{z(1-zd^2)}{1-4zd} \right)$$

DEMONSTRATION

Par la Proposition 1 : $R_0(z) = \mathfrak{H}(1, r, r)$.

Par ailleurs, par la formule du Théorème 3 : $\mathfrak{H}(t, x, y) = \frac{1}{v(x)} \left[\mathbb{1}_{(x)}(y) + \sum_{C \in \mathfrak{C}(x,y)} \delta(C) \cdot \phi(f(C)) \right]$.

On en déduit, en appliquant le Lemme 2 dans la Situation 2 :

$$\mathfrak{H}(1, r, r) = \frac{1}{v(r)} \left[\mathbb{1}_{(r)}(r) + \sum_{C \in \mathfrak{C}(r,r)} \delta(C) \cdot \phi(f(C)) \right] = \frac{1}{v_a} \left[1 + \sum_{C \in \mathfrak{C}(r,r)} \delta(C) \cdot \phi(f(C)) \right]$$

On déduit alors directement du Lemme 3 en exprimant $\delta(C) \cdot \phi(f(C))$ en fonction des coefficients de transmission et des paramètres définis Figure 7 :

$$\begin{aligned} \mathfrak{H}(1, r, r) &= \frac{1}{1-2zd} \left(1 + \sum_{\substack{\text{Arbres de } F_0 \\ \text{non réduits à la racine}}} \left(\frac{1-zd^2}{1-2zd} \right)^{d^\circ \rho} \left(\frac{zd^2}{1-3zd} \right)^{d^\circ \sigma_0} \left(\frac{2zd^2}{1-4zd} \right)^{d^\circ \sigma_1} \right. \\ &\quad \left. \left(\frac{z^2d^2}{1-3zd} \right)^{d^\circ \varphi_0} \left(\frac{2z^2d^2}{1-4zd} \right)^{d^\circ \varphi_1} \left(\frac{1-zd^2}{1-3zd} \right)^{d^\circ \delta_0} \right. \\ &\quad \left. \left(\frac{1-zd^2}{1-4zd} \right)^{d^\circ \delta_1} \left(\frac{z(1-zd^2)}{1-3zd} \right)^{d^\circ \omega_0} \left(\frac{z(1-zd^2)}{1-4zd} \right)^{d^\circ \omega_1} \right) \\ &= \frac{1}{1-2zd} \Theta \left(\frac{1-zd^2}{1-2zd}, \frac{zd^2}{1-3zd}, \frac{2zd^2}{1-4zd}, \frac{z^2d^2}{1-3zd}, \frac{2z^2d^2}{1-4zd}, \frac{1-zd^2}{1-3zd}, \frac{1-zd^2}{1-4zd}, \frac{z(1-zd^2)}{1-3zd}, \frac{z(1-zd^2)}{1-4zd} \right) \end{aligned}$$

On remarquera que le terme $\left(\frac{1-zd^2}{1-2zd}\right)^{d^p}$ est le produit de $\left(\frac{1-zd^2}{1-2zd}\right)^{d^{p-1}}$, terme de $\delta(C)$ issu du cas 1.2 dans le circuit C, par $\frac{1-zd^2}{1-2zd}$, terme égal à $\phi(f(C))$ (évident).

CONCLUSION

L'équation du Théorème 5 est la première du genre faisant apparaître la série génératrice R_0 des cartes planaires pointées comme composée de la série génératrice Θ d'une sous-famille de la famille des arbres bien étiquetés et de fonctions rationnelles de la série génératrice des mots de Dyck.

A cette occasion, sont apparus de nouveaux paramètres sur les arbres bien étiquetés. Il serait intéressant de voir leur interprétation sur les cartes planaires pointées.

Inversement, cette série génératrice $\Theta(\rho, \sigma_0, \sigma_1, \varphi_0, \varphi_1, \delta_0, \delta_1, \omega_0, \omega_1)$ de la famille F_0 peut être exprimée en fonction de R_0 et plus précisément de la série génératrice des arbres bien étiquetés par un système d'équations closes (non détaillé ici, cf [AJ91b]).

Ces résultats, relativement techniques, montrent que tout n'est pas encore totalement compris dans le domaine pourtant bien étudié des cartes planaires pointées.

REFERENCES

- [AJ91a] ARQUES D., JACQUES I. & ORIEUX K., 1991.
Equations fonctionnelles liant la série génératrice des cartes planaires à celle des mots de Dyck.
Article en soumission.
- [AJ91b] ARQUES D. & JACQUES I., 1991.
Une équation de convolution associée aux noeuds d'un graphe pour la résolution de problèmes d'évolution.
Article en préparation.
- [Ar85] ARQUES D., 1985.
Énumération et codage des cartes et hypercartes planaires pointées, Calcul formel sur les fractions multicontinues.
Thèse de Doctorat d'Etat, Université de Haute Alsace.
- [Ar86] ARQUES D., 1986.
Les hypercartes planaires sont des arbres très bien étiquetés.
Discrete Mathematics, N° 58, North-Holland, 11-24.
- [CF69] CARTIER P. & FOATA D., 1969.
Problèmes combinatoires de commutation et réarrangements.
Lecture Notes in Mathematics, N° 85, Springer-Verlag, Berlin, 1-88.
- [CV81] CORI R. & VAUQUELIN B., 1981.
Planar maps are well labeled trees.
Canadian Journal of Mathematical, Volume 33, N° 5, 1023-1042.
- [Tu68] TUTTE W. T., 1968.
On the enumeration of planar maps.
Bull. Amer. Math. Soc., Volume 74, 64-74.
- [Vi85] VIENNOT G. X., 1985.
Heaps of pieces. I : Basic definitions and combinatorial lemmas.
Lecture Notes in Mathematics, N° 1234, Combinatoire énumérative, Springer-Verlag, Berlin, 321-350.

GENERATION ALEATOIRE DES ANIMAUX DIRIGES

E. Barucci¹, R. Pinzani¹, R. Sprugnoli²

¹ Dipartimento di Sistemi e Informatica, Firenze (Italie)

² Dipartimento di Matematica Pura e Applicata, Padova (Italie)

INTRODUCTION

Le concept d' *animal dirigé* a beaucoup d'importance en plusieurs domaines comme les mathématiques, l'informatique et la physique. En mathématiques, les animaux dirigés sont un objet combinatoire ayant des propriétés remarquables; en informatique, on les a proposés comme une structure propre à la mémorisation et à la récupération des informations; enfin, en physique, ils constituent un modèle de divers phénomènes, comme, par exemple, la percolation.

Par conséquent, la génération aléatoire des animaux dirigés et, plus spécifiquement, une méthode pour produire telles structures dans l'ordinateur en sorte que chaque animal ait la même probabilité d'être généré ont une grande importance et nous offrent la capacité de *simuler* le comportement des animaux dirigés et de mesurer expérimentalement les quantités liées à ces structures.

Dans ce papier, on présente une méthode pour la génération aléatoire des animaux dirigés, telle que chaque animal est en moyenne généré dans un temps linéaire lorsque ce temps est mesuré en fonction du nombre de points qui constituent l'animal même. La méthode est très simple mais la linéarité du temps de génération n'est pas du tout évidente. Par conséquent, l'intérêt de cet article est concentré sur les aspects mathématiques qui nous permettent de prouver telle linéarité. Successivement on montre les résultats expérimentaux qui confirment les preuves analytiques précédentes. Enfin, on présente deux exemples de quantité qu'on peut étudier seulement expérimentalement parce que, jusqu'ici, on ne connaît pas des résultats analytiques.

1. LA GENERATION DES FACTEURS GAUCHES DE MOTZKIN

Gouyou-Beauchamps et Viennot [4] ont démontré qu'il y a une correspondance biunivoque entre les animaux dirigés composés par n points et le langage des facteurs gauches de Motzkin de longueur $n-1$. Les facteurs gauches de Motzkin sont les mots sur l'alphabet $\{a, b, c\}$ dont les préfixes ont un nombre de b qui ne dépasse jamais le nombre de a . Par conséquent, le premier pas dans la génération aléatoire des animaux dirigés est de générer aléatoirement un facteur gauche de Motzkin ayant n lettres. L'algorithme est le suivant:

- a) on génère successivement les lettres qui forment le facteur gauche de Motzkin en les tirant uniformément de l'alphabet $\{a, b, c\}$;
- b) avec deux compteurs A et B , on régitre le nombre de lettres a et b tirées, respectivement;
- c) si on génère n lettres avant que B ait dépassé A , la séquence obtenue est le facteur gauche de Motzkin voulu;
- d) si, au contraire, B dépasse A à un certain point, on quitte toutes les lettres déjà tirées et on recommence la génération du mot depuis le début.

A cause de ces recommencements, il n'est pas du tout évident que cet algorithme opère moyennement dans en temps linéaire mais nous ferons une démonstration analytique. Soient $\hat{\mu}_0, \hat{\mu}_1, \hat{\mu}_2, \dots$ les langages des mots de Motzkin de longueur $0, 1, 2, \dots$ de sorte que $|\hat{\mu}_n| = \mu_n$, l' n -ième nombre de Motzkin [3]. Nous définissons alors les langages ν_1, ν_2, \dots des mots négatifs, c'est à dire des mots de longueur k ($k \leq n$ pour ν_n) qui arrêtent la génération d'un facteur gauche de Motzkin. Donc nous avons:

$$\nu_n ::= \hat{\mu}_0 b \mid \hat{\mu}_1 b \mid \dots \mid \hat{\mu}_{n-1} b \quad \forall n \in \mathbb{N}$$

Quand on veut générer un facteur gauche de Motzkin par cet algorithme, en effet on génère un mot de ν_n (eventuellement vide) et puis le facteur gauche voulu. Par conséquent, le langage γ_n des mots qui génèrent un facteur gauche de Motzkin est donné par:

$$\sigma_n ::= \epsilon \mid \nu_n \sigma_n$$

$$\gamma_n ::= \sigma_n \hat{\alpha}_n$$

où $\hat{\alpha}_n$ est le langage des facteurs gauches de Motzkin de longueur n . Par la méthode de Schützenberger [7] on trouve la fonction génératrice:

$$\gamma_n(t) = \sigma_n(t) \hat{\alpha}_n(t) = \alpha_n t^n / (1 - \sum_{k=0}^{n-1} \mu_k t^{k+1})$$

De la fonction $\gamma_n(t)$ on obtient la fonction génératrice de probabilité et en observant que chaque lettre est tirée avec la probabilité $1/3$:

$$P_n(t) = \gamma_n(t/3) = \frac{\alpha_n t^n}{3^n - \sum_{k=0}^{n-1} \mu_k t^{k+1} 3^{n-k-1}} = \frac{\alpha_n t^n}{Q_n(t)}$$

Le nombre moyen d'extractions nécessaires pour générer un facteur gauche de Motzkin est donné par $P'_n(1)$:

$$P'_n(t) = \frac{n \alpha_n t^{n-1} Q_n(t) - \alpha_n t^n Q'_n(t)}{(Q_n(t))^2} = \frac{n \alpha_n t^{n-1}}{Q_n(t)} - \frac{\alpha_n t^n Q'_n(t)}{(Q_n(t))^2}$$

$$avg_n = P'_n(1) = \frac{n\alpha_n^2 - \alpha_n Q'_n(1)}{\alpha_n^2} = n + \frac{1}{\alpha_n} \left(\sum_{k=0}^{n-1} (k+1) \mu_k 3^{n-k-1} \right) = n + \frac{S_n}{\alpha_n}$$

La somme peut être évaluée au moyen des fonctions génératrices: si $S(t)$ indique la $\mathfrak{G}\{S_n\}$, on a:

$$S(t) = \frac{D_t \mu(t)}{1-3t} = \frac{1-t - \sqrt{1-2t-3t^2}}{2t^2(1-3t)\sqrt{1-2t-3t^2}}$$

où $\mu(t) = (1-t - \sqrt{1-2t-3t^2})/2t$ est la fonction génératrice des nombres de Motzkin. En appliquant la méthode de Darboux, on trouve:

$$S_n \sim \frac{2n+3}{2} 3^n \sqrt{\frac{3}{\pi(n+1)}} \left(1 + \frac{3}{16(n+1)} - \frac{107}{512(n+1)^2} \right) - \frac{3^{n+1}}{2}$$

Puisque le nombre des facteurs gauches de Motzkin est

$$\alpha_n \sim 3^n \sqrt{\frac{3}{\pi(n+1)}} \left(1 - \frac{1}{16(n+1)} + \frac{13}{512(n+1)^2} \right)$$

on trouve finalement:

$$avg_n = 2n - \frac{\sqrt{3\pi(n+1)}}{2} + \frac{7}{4} - \frac{\sqrt{3\pi}}{32\sqrt{n+1}} - \frac{3}{32(n+1)} + O(n^{-3/2})$$

et cela prouve la linéarité de la méthode de génération.

En employant la même technique, il est possible trouver aussi la variance, donnée par $var_n = P''_n(1) + P'_n(1) - (P'_n(1))^2$. En effet, on trouve:

$$var_n = \frac{4}{3}(n+1)^2 - \sqrt{3\pi}(n+1)^{3/2} + \frac{4+9\pi}{12}(n+1) - \frac{5\sqrt{3\pi}}{16}(n+1)^{1/2} - \frac{14-3\pi}{32} + O(n^{-1/2})$$

2. DES FACTEURS GAUCHES DE MOTZKIN AUX ANIMAUX DIRIGES

Après avoir obtenu le facteur gauche de Motzkin w , il faut construire l'animal dirigé correspondant. Les correspondances biunivoques trouvées précédemment de Gouyou-Beauchamps et Viennot [4] puis de Penaud [6] sont en effet *constructives* et théoriquement elles nous permettent de passer de w à l'animal dirigé. Malheureusement, leur temps d'exécution n'est pas linéaire avec $n = |w|$ à cause de la construction requise de la lettre c quand la largeur droite de l'animal n'est pas nulle. Dans ce cas, il faut trouver les sommets *dominés* par le sommet le plus en haut à droite de la racine et changer leur position. Tels sommets peuvent être nombreux (et parfois presque tous ceux de l'animal)

et donc il est possible que la construction exige un ou plusieurs pas qui n'ont pas un temps constant d'exécution.

Heureusement, Penaud a introduit le concept d'arbre de guingois, une structure équivalente aux animaux dirigés et dans laquelle l'ensemble des sommets dominés par le fils droit de la racine constitue un sous-arbre bien précis. Donc l'idée est de diviser la construction en deux phases: dans la première, on construit l'arbre de guingois T équivalent au facteur gauche de Motzkin w ; dans la seconde, on passe de T à l'animal dirigé \mathcal{A} équivalent. Puisque les deux constructions peuvent être exécutées dans un temps linéaire, la procédure globale aussi est linéaire.

Si \mathcal{A} est un animal dirigé, chaque sommet ξ de \mathcal{A} est déterminé par le numéro de la colonne $col(\xi)$ dans laquelle il se trouve et par son niveau $lev(\xi)$. La racine ρ est caractérisée par $col(\rho) = lev(\rho) = 0$, en sorte que $col(\xi)$ puisse être positif, négatif ou nul, tandis que $lev(\xi)$ est toujours positif pour $\xi \neq \rho$. La *frontière droite* de \mathcal{A} est l'ensemble de ses sommets ξ pour lesquels $col(\xi) \geq 0$ et ξ est de niveau minimal. La racine de \mathcal{A} appartient à la frontière et elle en est le seul sommet si la demi-largeur droite de \mathcal{A} est nulle.

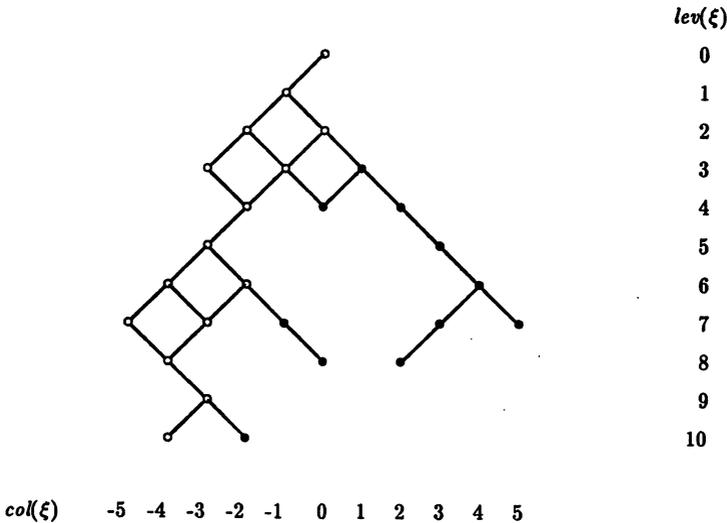


Figure 1 - Un animal dirigé.

Soit T un arbre binaire (orienté et ayant racine) et soient x et z deux de ses sommets, tels que z soit un aïeul de x ; le *déplacement* de x par rapport à z est la différence entre le nombre de branches droites et gauches dans le chemin de z à x . Si dans un arbre T le déplacement de chaque sommet par

rapport à la racine n'est jamais positif, on dit que T est une *équerre*. Si pour chaque sommet double d'un arbre T le sous-arbre gauche est une équerre, on dit T un *arbre de guingois*. La correspondance biunivoque trouvée par Penaud entre animaux dirigés et arbres de guingois nous permet de transférer le concept de frontière droite de l'un à l'autre. La propriété essentielle est que si ξ est un sommet de la frontière droite de \mathcal{A} et x est le sommet associé à ξ dans T (correspondent de \mathcal{A}), l'ensemble $\Delta(\xi)$ des sommets dominés par ξ correspond au sous-arbre de T qui a x comme racine. On définit alors un *arbre de guingois étendu* comme un arbre de guingois dans lequel chaque sommet a un troisième arc (appelé *trame*) qui pointe de chaque sommet x de la frontière droite de T au père du sommet successif y de la même frontière, lorsque y existe. Toutes les autres trames sont vides.

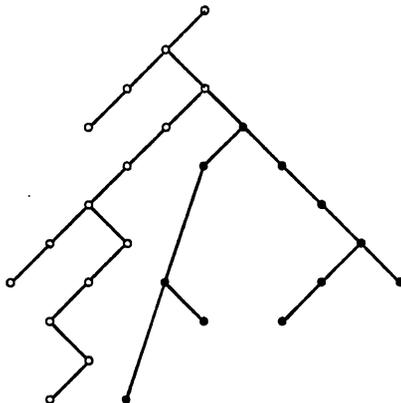


Figure 2 - L'arbre de guingois correspondant à l'animal de la Figure 1.

Maintenant nous pouvons finalement donner la construction de l'arbre de guingois étendu T en partant du facteur gauche de Motzkin w . On scande w de gauche à droite en ajoutant un nouveau sommet à T selon le caractère w_k trouvé:

- 1) $w_k = a$: on ajoute une nouvelle racine x à T de façon que le vieux T devienne le sous-arbre droit de x . Le sous-arbre gauche est vide et on fait pointer la trame de x à x même;
- 2) $w_k = b$ (ce cas peut se produire seulement si la trame de la racine de T n'est pas vide): on ajoute une nouvelle racine x de façon que le vieux T devienne le sous-arbre gauche. Le sous-arbre droit est vide. On fait pointer la trame de x au fils droit du sommet qui était pointé par la trame de la vieille racine de T , qui devient vide.
- 3) $w_k = c$: il y a deux sous-cas:

3 a) si la trame de la racine est vide, on ajoute une nouvelle racine x en sorte que le vieux T en devienne le sous-arbre gauche. Le sous-arbre droit est vide et la trame de x est vide;

3 b) autrement, on ajoute une nouvelle racine x à T , et T est divisé en deux sous-arbres: le sous-arbre dont la racine est le second sommet y dans la frontière droite de T devient le sous-arbre droit de x et le reste de T en devient le sous-arbre gauche. On fait pointer la trame de x à x même; les trames de la vieille racine de T et du père de y deviennent vides.

Chaque pas de cette procédure se compose d'un nombre bien défini d'opérations et donc elle a une complexité linéaire. Le dernier pas est la construction de l'animal dirigé \mathcal{A} à partir de l'arbre de guingois T . En suivant Bétréma et Penaud [2], on définit un vecteur L d'entiers initialisé à zéro et ayant indices de $-n$ à $+n$; par conséquent

1) la source de \mathcal{A} correspond à la racine de T et elle est le sommet défini par $col(\xi) = 0$ et $lev(\xi) = 0$;

2) on détermine les autres sommets de \mathcal{A} en visitant T en préordre et en exécutant les pas suivants:

2 a) le fils gauche y de chaque sommet x de T génère le sommet η de \mathcal{A} défini par $col(\eta) = col(\xi) - 1$ (ξ est le sommet qui correspond à x) et $lev(\eta) = 1 + \max(L(col(\eta) + 1), L(col(\eta) - 1))$; on met aussi $L(col(\eta)) = lev(\eta)$;

2 b) le fils droit de chaque sommet x de T génère le sommet η de \mathcal{A} défini par $col(\eta) = col(\xi) + 1$ et $lev(\eta) = lev(\xi) + 1$; on met aussi $L(col(\eta)) = lev(\eta)$.

L'animal \mathcal{A} peut être représenté par un vecteur de n listes (au plus), dans lequel la k -ième liste contient l'ensemble ordonné des $col(\xi)$, pour chaque sommet ξ de \mathcal{A} ayant $lev(\xi) = k$. De cette représentation, on peut aisément obtenir le dessin de \mathcal{A} . Naturellement, la linéarité de la procédure est évidente.

3. RESULTATS DES SIMULATIONS

Premièrement, nous avons réalisé un programme qui génère les facteurs gauches de Motzkin selon l'algorithme du paragraphe 1. En registrant les appels à la routine *random*, il est possible d'avoir la moyenne et la variance de l'ensemble des générations, qui ont été 100000 pour chaque valeur de $n = 1000, 2000, \dots, 10000$. On a aussi calculé la moyenne de la différence $A - B$; cette quantité est connue analytiquement (v. Barucci, Pinzani, Sprugnoli [1]) et on l'a employée comme indicateur pour la casualité des facteurs gauches de Motzkin générés. Dans la table 1, on a reporté les valeurs attendues et expérimentées et, comme on peut remarquer, leur correspondance est très bonne.

On a fourni une preuve différente de la casualité en employant le test du χ^2 sur 30000 générations de facteurs gauches de Motzkin ayant $n = 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50$. On a reporté les résultats dans la table 2, et en ce cas aussi ils sont très probants.

n	appels à <i>random</i>			deviation standard			difference <i>A-B</i>		
	attendue	expérim.	err. %	attendue	expérim.	err. %	attendue	expérim.	err. %
1000	1953.2	1953.4	0.01	1114.2	1117.9	0.33	31.379	31.266	0.36
2000	3933.1	3934.8	0.04	2251.5	2257.5	0.27	44.777	44.693	0.19
3000	5917.7	5909.3	0.14	3392.8	3411.3	0.55	55.060	54.917	0.26
4000	7904.7	7898.4	0.08	4536.3	4543.1	0.15	63.730	63.685	0.07
5000	9893.2	9871.7	0.22	5681.0	5655.1	0.46	71.368	71.287	0.11
6000	11882.8	11870.9	0.10	6826.8	6822.0	0.07	78.274	78.199	0.10
7000	13873.3	13849.0	0.18	7973.2	7951.2	0.28	84.625	84.525	0.12
8000	15864.4	15829.6	0.22	9120.3	9102.0	0.21	90.536	90.493	0.05
9000	17856.1	17840.1	0.09	10267.7	10286.3	0.18	96.087	95.962	0.13
10000	19848.2	19845.8	0.01	11415.6	11432.2	0.15	101.338	101.358	0.02

Table 1. Résultats des simulations

n	degrés de liberté	Valeurs théoriques de χ^2		Valeurs expérimentales de χ^2	
4	4	0.21	— 14.86	0.79	— 7.24
5	5	0.41	— 16.75	1.20	— 6.64
6	6	0.68	— 18.55	1.25	— 10.90
7	7	0.90	— 20.28	0.83	— 9.31
8	8	1.34	— 21.96	1.75	— 16.01
9	9	1.73	— 23.59	2.64	— 12.80
10	9	1.73	— 23.59	5.53	— 15.95
20	14	4.07	— 31.32	3.30	— 26.37
30	17	5.70	— 35.72	7.85	— 17.16
40	20	7.43	— 40.00	10.50	— 33.76
50	23	9.26	— 44.18	15.67	— 31.83

Table 2. Le test du χ^2

Puis on a réalisé un programme pour dessiner l'animal dirigé correspondant à un facteur gauche de Motzkin généré d'une façon aléatoire.

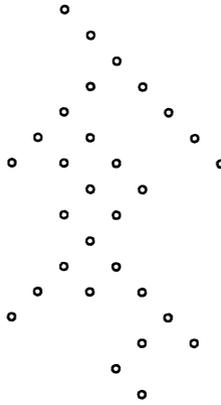


Figure 3 - Un animal aléatoire avec 30 sommets

Enfin, on a cherché à évaluer expérimentalement la longueur moyenne h_n et le chemin intérieur moyen p_n des animaux dirigés ayant n sommets. On a obtenu les meilleurs résultats avec les fonctions:

$$h_n = A_1 n^B + K_1 \qquad p_n = A_2 n^C + K_2 n$$

où $1.21 \leq A_1 \leq 1.27$, $0.816 \leq B \leq 0.823$, $0.61 \leq A_2 \leq 0.65$ et $1.815 \leq C \leq 1.822$. Par conséquent il semble probable que $A_2 = A_1/2$ et $C = 1 + B$. La valeur de B a été précédemment estimé [5] en $9/11 \approx 0.818$ et nos épreuves expérimentales ne contredisent pas cette hypothèse.

REFERENCES

- [1] Barucci E., Pinzani R., Sprugnoli R.: "The Motzkin Family", Report RT 2/89, Dipartimento di Sistemi e Informatica, Università di Firenze, 1989.
- [2] Bétrema J., Penaud J. G.: "Animaux et Arbres Guingois", in Series Formelles et Combinatoire Algébrique (M. Delest, G. Jacob, P. Leroux eds.), Bordeaux, 1991, 85-102.
- [3] Donaghey R., Shapiro L. W.: "Motzkin Numbers", Journal of Combinatorial Theory A, 23 (1977), 291-301.
- [4] Gouyou Beauchamps D., Viennot X. G.: "Equivalence of the Two Dimensional Directed Animal Problem to a One-Dimensional Path Problem", Advances in Applied Mathematics, 9 (1988), 334-357.
- [5] Nadal J. P., Derrida B., Vannimenus J.: "Directed Lattices Animals in 2 Dimension: Numerical and Exact Results", J. Physique, 43 (1982), 1561.

- [6] Penaud J. G.: "Une Nouvelle Bijection pour les Animaux Dirigés", Rapport LaBRI n. 89-45, Université de Bordeaux, 1989.
- [7] Schützenberger M. P.: "Context-free Languages and Pushdown Automata", *Information and Control*, 6 (1963), 246-264.

Symbolic Computation and Discrete Probabilities

Anne Bergeron

Département de mathématiques et d'informatique
UQAM
C.P. 8888 Succursale 'A'
Montréal, Canada, H3C 3P8

Abstract

In this paper, we study the formal and algorithmic aspects of symbolic manipulation of discrete random experiments. Many random experiments can be described with formal series. When these series are rational, we can easily compute generating functions, distributions and probabilities. However, even if they provide a suitable framework for computing, formal representations of random experiments lack closure properties under rational operations. We are investigating classes of random experiments which can be generated from 'elementary' experiments and simple algebraic operations. One such class is the rational semaphore codes on a finite alphabet A , together with Bernoulli probability measure on A^* .

0. Introduction

The problems discussed here arose during the development of a simulation environment in discrete probability theory [Bergeron & Bordier, 91]. In that project, we wanted the user to be able to describe freely 'any' random experiment he might think of. Computer simulations would then be generated and the results analyzed with various numerical tools. A further development was to provide feedback on an experiment before any simulation was attempted: the goal was to use the symbolic computation capabilities of the computer to derive information on an experiment from its description only. An interesting example of feedback would be, for example, the average time needed to complete the simulation. It became increasingly evident that symbolic computations could be applied to solve many other questions relevant to discrete random experiments, giving rise to 'computational probabilities' in the same fashion that other branches of mathematics have their computational variant.

The formalisms to describe a discrete random experiment were found, unsurprisingly, within coding theory. Links between coding theory and probability theory can be traced far back [Shannon, 49] and many results in computing the average length of codes [Berstel & Perrin, 85] provided the first hints towards a more general solution. We will show in this paper that for a large class of experiments the generating functions of many random variables, including the length, are exactly computable.

The theory of formal series provided the representations and algorithms needed to carry out the various computations involved in solving discrete probability problems. All the examples in the text were computed using a prototype developed in Maple [Char & al., 86]. Given a short specification of a discrete experiment, many questions concerning distributions of random variables and probabilities of events can be answered automatically through simple function calls.

1. Rational Random Experiments

In the following, a discrete random experiment will be given by the finite or denumerable set Ω of all its possible outcomes, the *sample space*, and by a function

$$\pi : \Omega \longrightarrow [0, 1]$$

such that

$$\sum_{x \in \Omega} \pi(x) = 1$$

An *event* X is any subset of Ω . The probability $P(X)$ of an event is given by

$$P(X) = \sum_{x \in X} \pi(x)$$

A *integral valued random variable* is a function $V : \Omega \longrightarrow \mathbb{N}$ which associate an integer to each possible outcome of the experiment. The *generating function* $F(s)$ of a random variable V is defined by

$$F(s) = \sum_{k \geq 0} P[V(x) = k] s^k \quad \text{where} \quad P[V(x) = k] = \sum_{\substack{x \in \Omega \\ V(x) = k}} \pi(x)$$

In the development of computer environments allowing specification and manipulation of discrete random experiments, we have to solve two main problems. The first is to construct suitable formal tools for describing random experiments, events and random variables. This construction will involve the identification of 'elementary' experiments and algebraic operations that generate classes of random experiments.

Once we have a formal description of a random experiment, the second problem is to develop algorithms to compute probabilities and generating functions. In this section, we introduce the concept of *rational random experiment*. Given a rational experiment, we will show that the generating functions of many random variables are again rational and readily obtainable from the formal series that represents the experiment. Furthermore, the probability of any rational event can also be computed with simple operations on series.

Section 1.1 presents some basic definitions and results on rational series that can be found in [Berstel & Reutenauer, 84]. We next recall the concept of rational measures [Hansel & Perrin, 89]. Rational random experiments and their properties are discussed in section 1.2.

1.1 Rational Series and Measures

Let $A = \{a_1 \dots a_n\}$ be a finite alphabet and A^* the set of words on A . The concatenation xy of two words $x = x_1 \dots x_k$ and $y = y_1 \dots y_m$ in A^* is defined as $xy = x_1 \dots x_k y_1 \dots y_m$. A^+ represents the set $A^* - \{1\}$ where 1 is the empty word. The length k of a word $x = x_1 \dots x_k$ is denoted by $|x|$, and the number of occurrences of a letter $a \in A$ in x is denoted by $|x|_a$. If X and Y are subsets of A^* , XY denotes the set $\{xy \mid x \in X \text{ and } y \in Y\}$. If $X = \{x\}$, we will simply write x for X .

Rational series and rational sets

Let K be a commutative ring, a *formal series* is a function $F: A^* \rightarrow K$ and we write

$$F = \sum_{x \in A^*} F(x)x$$

where $F(x)$ is the value of F at x . The set of all formal series with values in a ring K is denoted $K\langle\langle A \rangle\rangle$. *Rational operations* on $K\langle\langle A \rangle\rangle$ are the sum and the product:

$$(F + G)(x) = F(x) + G(x) \text{ and } (FG)(x) = \sum_{uv=x} F(u)G(v)$$

and the *star*, defined when $F(1) = 0$:

$$F^* = \sum_{n \geq 0} F^n$$

Each letter $a \in A$ can be seen as the formal series $1a$. A series is rational if it can be obtained through a finite number of rational operations on letters $a \in A$. It is well known that rationality is preserved by many operations, one of which is the Hadamard product of two series F and G defined as:

$$(F \times G)(x) = F(x) G(x)$$

Let x be a word in A^* . The *commutative variant* $c(x)$ of x is:

$$c(x) = a_1^{|x|_{a_1}} \dots a_n^{|x|_{a_n}} \text{ where } a_i \in A$$

For each formal series F , the *commutative variant* $c(F)$ of a formal series F is defined by:

$$c(F) = \sum_{x \in A^*} F(x)c(x) = \sum_{y \in c(A^*)} \left(\sum_{c(x)=y} F(x) \right) y$$

Many computations in this paper are based on the following result:

Proposition 1.1 *If F is rational then $c(F)$ can be written as a quotient of two polynomials in the variables $a_i \in A$:*

$$c(F) = \frac{\mathbb{P}(a_1, \dots, a_n)}{\mathbb{Q}(a_1, \dots, a_n)}$$

where $\mathbb{Q}(0, \dots, 0) = 1$. ■

If X is a subset of A^* , its *characteristic series* is the series:

$$\underline{X} = \sum_{x \in X} x$$

A subset of A^* is *rational* iff its characteristic series is. By a fundamental theorem, the collection of rational sets is the smallest subset of $\mathfrak{P}(A^*)$ containing the singletons and closed under product, union and star, defined by:

$$X^* = \bigcup_{n \geq 0} X^n$$

The complement of a rational set is again rational.

Rational measures

A *probability measure* on A^* is a mapping $\pi: A^* \rightarrow [0, 1]$ such that for each word x of A^* we have:

$$\sum_{a \in A} \pi(xa) = \pi(x) \quad \text{and} \quad \pi(1) = 1$$

A *positive measure* is a measure such that $\pi(x) > 0$ for all x in A^* . Of course, any measure can be extended to subsets S of A^* :

$$\pi(S) = \sum_{x \in S} \pi(x)$$

It follows immediately from the definition of measure that if $n \geq 0$, $\pi(A^n) = 1$.

A *prefix* subset X of A^* is a set that contains no proper prefix of its elements. That is, if $x \in X$ and $|y| \geq 1$, then $xy \notin X$. For prefix subsets of A^* , we have the following:

Proposition 1.2 *If X is a prefix subset of A^* and π is a probability measure on A^* , then $\pi(X) \leq 1$.*

Proof: Suppose first that X is finite, and let $|X|$ denote the sum of all word lengths of words in X . If X contains only words of length 1, we have from the definition of measure that $\pi(X) \leq 1$. If X contains longer words, let X' be constructed by replacing in X all words of the form $\{ma \mid a \in A, ma \in X\}$ by the word m . Then $\pi(X') \geq \pi(X)$, $|X'| < |X|$ and X' is again prefix. If it contains words of length greater than 1, we can reapply the procedure. Since $|X|$ is finite, we eventually obtain a set of words of length 1.

If X is infinite, the sequence $\pi(X_n)$ where $X_n = \{x \in X \mid |x| \leq n\}$ is non-decreasing, bounded by 1, and converges to $\pi(X)$. ■

In the sequel, we will be mainly interested in showing that some prefix subsets have measure 1. The following proposition will prove helpful.

A measure is *multiplicative* on a subset X if for all x, y in X^* we have

$$\pi(xy) = \pi(x)\pi(y)$$

When a measure is multiplicative on a subset X , we can relate the measure of X^* and the measure of X :

Proposition 1.3 *If π is multiplicative on X then*

$$\pi(X^*) \leq \sum_{n \geq 0} (\pi(X))^n$$

in particular, if $\pi(X) < 1$, then $\pi(X^)$ is finite.*

Proof: First, we have that

$$\pi(X^*) \leq \sum_{n \geq 0} \pi(X^n)$$

since X^* is contained in the union of sets of the form X^n . We also have, for $n > 0$

$$\pi(X^n) \leq \sum_{y \in X^{n-1}} \sum_{x \in X} \pi(yx) = \sum_{y \in X^{n-1}} \pi(y) \sum_{x \in X} \pi(x) = \pi(X^{n-1})\pi(X)$$

from which the result follows immediately. ■

A probability measure is *rational* if the formal series

$$\pi = \sum_{x \in A^*} \pi(x) x$$

is rational.

The simplest examples of rational probabilities measures are *Bernoulli measures*. They are obtained by multiplicative extension of measures defined on elements of A such that:

$$\sum_{a \in A} \pi(a) = 1$$

in this case, the formal series representing the measure can be written as:

$$\pi = \left(\sum_{a \in A} \pi(a)a \right)^*$$

1.2 Rational random experiments

Suppose that $A = \{a_1, a_2, \dots, a_n\}$ represents the possible outcomes of an elementary random experiment, or elementary *trial*. We are interested in complex experiments where we repeat an elementary trial until the sequence of results exhibits some characteristic¹. Our main problem is to define properly what is meant by 'until the sequence of results exhibits some characteristic'. First, we want the process to stop with probability 1, eliminating in practice infinite sequences of results. When this occurs, the possible (realistic) outcomes can be viewed as words of A^* . The second problem is the computability or recognizability of the characteristic the sequence should exhibit to complete the experiment. The next definition gives the simplest class of 'computable' experiments:

Definition 1.1 Let $A = \{a_1, a_2, \dots, a_n\}$ be a finite alphabet and π a rational probability measure. A subset Ω of A^* defines a *rational random experiment* with the measure π if Ω is rational and $\pi(\Omega) = 1$.

Example 1.1 For any measure π , the set A^n defines a rational experiment.

Example 1.2 Let $A = \{a, b\}$. Suppose that we repeatedly choose at random a letter from A until we get any letter for the second time. The set of possible outcomes is

$$\Omega = \{aa, bb, aba, abb, bab, baa\}$$

This set is finite, thus rational. Let π be any measure defined on A^* . Applying repeatedly the definition of a probability measure, we get $\pi(\Omega) = 1$.

¹Those experiments are studied in [Feller 57] under the name of *recurrent events*.

Example 1.3 Suppose now that we repeatedly choose at random a letter from the alphabet $A = \{a, b\}$ until we get the letter b . The set $\Omega = a^*b$ is rational. If π is a Bernoulli probability measure such that $\pi(b) > 0$, it is easy to check that $\pi(\Omega) = 1$.

If Ω and π are rational, we can represent them with the following rational series

$$\underline{\Omega} = \sum_{x \in \Omega} x \quad \text{and} \quad \pi = \sum_{x \in A^*} \pi(x)x$$

The series

$$E = \underline{\Omega} \times \pi = \sum_{x \in \Omega} \pi(x)x$$

corresponding to the Hadamard product of $\underline{\Omega}$ and π is rational. We say that the series E represents the rational experiment defined by Ω with the measure π .

Example 1.3 (Continued) The characteristic series of the set $\Omega = a^*b$ is given by:

$$\underline{\Omega} = \sum_{n \geq 0} a^n b$$

and the series that represents the experiment is

$$E = \sum_{n \geq 0} \pi(a)^n \pi(b) a^n b$$

In the next paragraphs we will show that the usual computations of discrete probability theory can be effectively performed using the series E that represents an experiment.

Linear random variables

Definition 1.2 Let Ω be a rational random experiment on A with the measure π . A *linear* random variable on Ω is the restriction to Ω of a mapping

$$V: A^* \longrightarrow \mathbb{N}$$

such that $V(xy) = V(x) + V(y)$.

Simple examples of linear random variables are word length and number of occurrences of a letter in a word. A general linear random variable is a linear combination of the functions $|x|_a$ with positive integer coefficient.

We have the following:

Proposition 1.4 *If Ω defines a rational random experiment with the measure π , and V is a linear random variable defined on Ω , then its generating function $F(s)$ is rational.*

Proof: Let E be the series that represents the random experiment

$$E = \sum_{x \in \Omega} \pi(x)x$$

and let the commutative variant $c(E)$ of E be represented by the following quotient:

$$\frac{\mathbb{P}(a_1, \dots, a_n)}{\mathbb{Q}(a_1, \dots, a_n)}$$

We have

$$F(s) = \sum_{k \geq 0} \left[\sum_{V(x)=k} \pi(x) \right] s^k = \frac{\mathbb{P}(s^{V(a_1)}, \dots, s^{V(a_n)})}{\mathbb{Q}(s^{V(a_1)}, \dots, s^{V(a_n)})} = \frac{P(s)}{Q(s)}$$

which has a development around 0 since

$$\frac{P(0)}{Q(0)} = \sum_{V(x)=0} \pi(x) \leq 1$$

Thus $F(s)$ is rational. ■

Example 1.3 (Continued) Consider the random variable 'length' defined by $V(x) = |x|$ on the experiment that consisted of choosing a letter in $\{a, b\}$ until we get a b . We saw that this experiment was represented by:

$$E = \sum_{n \geq 0} \pi(a)^n \pi(b) a^n b = \frac{1}{(1 - \pi(a)a)} \pi(b)b$$

The generating function $F(s)$ of V is then:

$$F(s) = \frac{\pi(b)s}{(1 - \pi(a)s)} = \sum_{k \geq 0} \pi(a)^k \pi(b) s^{k+1}$$

and the average length is given by

$$F'(1) = 1 + \pi(a)/\pi(b).$$

Probabilities of rational events

The preceding result allows us to compute effectively the distributions of many random variables and the probabilities of events that can be described in terms of those variables. However, many events cannot be described in terms of linear random variables. But if an event is a rational subset of Ω , we have the following:

Proposition 1.5 *Let E be the series that represents a rational random experiment with sample space Ω . Let X be a rational set with characteristic series \underline{X} . The Hadamard product $H = E \times \underline{X}$ of the two series is rational, and if the commutative variant $c(H)$ of H is represented by the quotient:*

$$c(H) = \frac{\mathbb{P}(a_1, \dots, a_n)}{\mathbb{Q}(a_1, \dots, a_n)}$$

then the probability of the event $X \cap \Omega$ is given by

$$\frac{\mathbb{P}(1, \dots, 1)}{\mathbb{Q}(1, \dots, 1)}$$

Proof: We have

$$c(H) = \sum_{x \in X \cap \Omega} \pi(x) c(x)$$

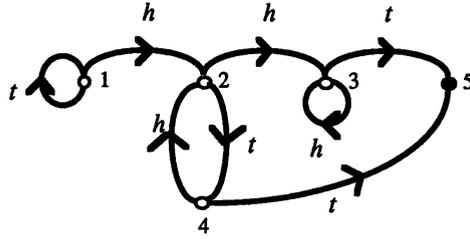
thus

$$\frac{\mathbb{P}(1, \dots, 1)}{\mathbb{Q}(1, \dots, 1)} = \sum_{x \in X \cap \Omega} \pi(x)$$

is the probability of the event $X \cap \Omega$. ■

Example 1.5 Consider the following game [Penney, 74; Graham & al., 89] in which two players repeatedly throw a penny until the sequence hht or htt appears for the first time (h for head, t for tail). If hht appears first, player A wins and if htt appears first, player B wins. Which one, if any, has an advantage in this game?

We first mention (see section 2.2) that the average number of throws in order to obtain hht (or htt) is 8. When we combine the two experiments, we get the surprising result that hht is two times more likely to show up first than htt ! The following diagram is useful in analyzing this situation:



The initial state of the game is at vertex 1. The game moves to another state along the arc labeled h or t according to the result of each successive throw. If the game reaches state 5 through state 3, player A wins, otherwise player B wins. From that finite automaton representation of the game, it is possible to compute (see section 2) the series that represents the experiment:

$$E = \frac{1}{2-t} h \frac{1}{4-th} tt + \frac{1}{2-t} h \frac{1}{4-th} h \frac{2}{2-h} t$$

On the other hand, the characteristic series of the set of sequences that end with the pattern hht is given by:

$$\underline{X} = \frac{1}{1-(t+ht)} hh \frac{1}{1-h} t$$

The Hadamard product of these two series is:

$$H = \frac{1}{2-t} h \frac{1}{4-th} h \frac{2}{2-h} t$$

If we substitute 1 for h and t in this expression, we get $2/3$.

2. Algebras of experiments

The results of the preceding section give us a general framework for symbolic manipulation of discrete random experiments. Unfortunately, rational experiments have very few closure properties with respect to rational operations. In Example 1.5, we stated the stopping condition for the experiment as "... ending with the pattern hht or htt ". There is a natural use of the conjunction 'or' in the formulation of that condition, but the set of possible outcomes of the experiment is not the union of the set of possible outcomes of the experiment where we stop with hht and the one where we stop with htt . Some sequences, like $hhtht$, must be left over.

The next section introduces a class of experiments in which rational operators for expressing characteristics of sequences can be used freely while preserving the rationality of the experiments and the fact that $\pi(\Omega) = 1$.

2.1 Rational semaphore codes and Bernoulli measures

Let s be a word in A^* . Consider the experiment that consists in repeating the elementary trial A up to the first appearance of the pattern s . The set Ω of possible outcomes can be represented by the rational expression

$$\Omega = A^*s - A^*s A^+$$

It corresponds to all words of A^* that end with s , but that have no other occurrences of s in them. If t is a second pattern, we can represent the experiment of repeating the elementary trial up to the first appearance of s or t as

$$\Omega' = A^*(s + t) - A^*(s + t) A^+$$

This representation displays nicely where the conjunction 'or' of the stopping condition is effectively used in the construction of the set of possible outcomes. Sets constructed in this way are studied in code theory under the name of \mathcal{J} codes [Schützenberger, 64] or *semaphore codes* [Berstel & Perrin, 85]. In general, we have the following definition:

Definition 2.1 Let A be an alphabet and S a non-empty subset of A^+ . The set

$$X = A^*S - A^*SA^+$$

is called a *semaphore code*. The set S is the set of *semaphores* of X .

If a semaphore code is constructed using a rational set of semaphores, then the code itself is rational. The main result of this section is that rational semaphore codes are a 'natural' class of rational random experiments and are generated with unrestricted rational sets:

Proposition 2.1 *Any rational semaphore code defines a rational random experiment with any positive Bernoulli measure. Moreover, the class of rational semaphore codes is precisely the class obtained with the operation*

$$X = A^*S - A^*SA^+$$

where S ranges through the rational sets.

Proof: The fact that any rational semaphore code defines a rational experiment with any positive Bernoulli measure is a consequence of more general results from code theory (see, for example, [Hansel & Perrin, 89]). We can also give the following direct proof, using the results of section 1.

First observe that any semaphore code Ω is a prefix set, thus for any Bernoulli measure π , we have $\pi(\Omega) \leq 1$ by Proposition 1.2. Since Bernoulli measures are multiplicative over any subset of A^* , if we can show that $\pi(\Omega^*)$ is infinite, we can use Proposition 1.3 to conclude that $\pi(\Omega) = 1$.

If Ω is a semaphore code, any word of the form us where s is a semaphore of Ω can be written as $w_1w_2\dots w_nd$ where $w_i \in \Omega$ and d is a proper right factor of s . Thus the set A^*s is contained in Ω^*D where D is the finite set of proper right factors of s . But

$$\pi(A^*s) = \sum_{n \geq 0} \pi(A^n s) = \sum_{n \geq 0} \pi(s) \text{ is infinite}$$

and

$$\pi(A^*s) \leq \pi(\Omega^*D) \leq |D|\pi(\Omega^*)$$

thus $\pi(\Omega^*)$ is infinite.

In order to prove the second part of the theorem, we show that if X is a semaphore code, then $X = A^*X - A^*XA^+$. Thus any rational semaphore code has at least one set of semaphores which is rational.

If X is a semaphore code, we have immediately $X \subseteq A^*X - A^*XA^+$ since any word in A^*XA^+ contains as a proper left factor a word that ends with a semaphore of X . On the other hand, since both sides of that inequality are semaphore codes, we have for any positive Bernoulli measure π

$$\pi(X) = \pi(A^*X - A^*XA^+) = 1$$

so we must have $X = A^*X - A^*XA^+$. ■

The most important consequence of Proposition 2.1 is that any formalism for representing rational sets can be used to represent rational random experiments, as long as it is understood that only the semaphores are described. This restriction is not too severe as we will see in the next examples. Let A be a finite alphabet representing the possible outcomes of an elementary trial. In each of the following experiments the set of possible outcomes is a rational semaphore code:

Example 2.1 "Repeat the trial until one of a pattern in a finite set S appears."

In this case the set of semaphores is given by S which, being finite, is rational.

Example 2.2 "Repeat the trial k times."

Here the set of semaphores is simply A^k .

Example 2.3 "Repeat the trial until the k^{th} occurrence of the letter a ."

Let $A_a = A - \{a\}$, the set of semaphores of this experiment is given by $(A_a^*a)^k$.

Example 2.4 "Repeat the trial until we get k different results, $k \leq |A|$."
Let B be an ordered sequence of k distinct elements of A . The set:

$$\prod_{b \in B} (B_b)^* b$$

corresponds to all words where the elements of B appear in order. If we take the finite sum of these sets over all possible ordered sequences of k elements, we obtain a rational set of semaphores for this experiment.

Even if the class of rational semaphore codes can express a wide variety of experiments, it is strictly contained in the class of rational experiments. With any measure, the set $\Omega = \{a, baa, bab, bb\}$ defines a rational experiment. But Ω is not semaphore since the word a appears as a proper factor in the words baa and bab .

2.2 Linear representations and algorithms

In this section, we will sketch the main algorithms that enable us to construct the series representing a rational experiment defined with a Bernoulli measure. We will assume here that Ω is described with a finite deterministic automaton.

We first need the concept of *linear representation* of a rational series. If A is a finite alphabet, then by Kleene's theorem F is a rational series in $K\langle\langle A \rangle\rangle$ iff there is a homomorphism of monoids:

$$\varphi : A^* \longrightarrow K^{n \times n}$$

together with two vectors $\lambda \in K^{1 \times n}$ and $\gamma \in K^{n \times 1}$ such that:

$$F(x) = \lambda \varphi(x) \gamma$$

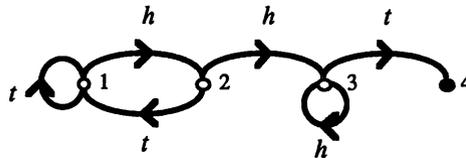
In this case, the triplet $(\lambda, \varphi, \gamma)$ is called a *linear representation* of F of *dimension* n .

If a rational set is represented with a finite deterministic automaton, we can easily derive a linear representation for the characteristic series of the set. Let $N = \{1, \dots, n\}$ be the set of states of the automaton with $I \subseteq N$ the initial states and $F \subseteq N$ the final states. For each $a \in A$, the matrix $\varphi(a)$ is defined by:

$$\begin{aligned} \varphi(a)_{ij} &= 1 \text{ if the automaton goes from state } i \text{ to state } j \text{ with the letter } a, \\ \varphi(a)_{ij} &= 0 \text{ otherwise.} \end{aligned}$$

Vectors λ and γ are defined respectively as characteristic vectors of I and F .

Example 2.5 The automaton



recognizes the (semaphore) set $A^*hht - A^*hhtA^+$ and we have the following linear representation:

$$\varphi(h) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad \varphi(t) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad \lambda = [1 \ 0 \ 0 \ 0], \quad \gamma = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

If $(\lambda, \varphi, \gamma)$ is a linear representation of a set Ω defining a random experiment with a Bernoulli measure π , the series E that represents the experiment admits the linear representation $(\lambda, \varphi, \gamma)$ with:

$$\varphi'(a) = \pi(a)\varphi(a)$$

In fact, if $x = x_1 \dots x_n$ and $x \in \Omega$, we have by direct computation:

$$\begin{aligned} \lambda\varphi'(x)\gamma &= \lambda\varphi'(x_1)\dots\varphi'(x_n)\gamma \\ &= \lambda\pi(x_1)\varphi(x_1)\dots\pi(x_n)\varphi(x_n)\gamma \\ &= \pi(x)\lambda\varphi(x)\gamma \\ &= \pi(x) \end{aligned}$$

Finally, we can obtain an explicit rational expression for a series from a linear representation (see, for example, [Berstel & Reutenauer, 85]. Theorem 1.6.1). The technique is based on the following:

Let M be the formal matrix:

$$M = \sum_{a \in A} \varphi(a)a$$

The successive powers M^n satisfy the identities:

$$M^n = \left[\sum_{a \in A} \varphi(a)a \right]^n = \sum_{x \in A^n} \varphi(x)x$$

Thus we have

$$M^* = \sum_{n \geq 0} M^n = \sum_{n \geq 0} \left[\sum_{x \in A^n} \varphi(x)x \right] = \sum_{x \in A^*} \varphi(x)x$$

And the series E can be expressed as

$$E = \sum_{x \in A^*} (\lambda \varphi(x) \gamma) x = \lambda \left(\sum_{x \in A^*} \varphi(x) x \right) \gamma = \lambda M^* \gamma$$

It can be shown that if I_n is the identity matrix of order n , the matrix M^* satisfies the relation:

$$M^* = I_n + MM^*$$

which gives a linear system of equations for the coefficients of M^* and the commutative variants of the coefficients of M^* can be simply computed by inverting the matrix

$$M^* = (I_n - M)^{-1}$$

Example 2.5 (Continued) The set $A^*hht - A^*hhtA^*$ defines a random experiment for any Bernoulli measure. If we set $\pi(h) = \pi(t) = 1/2$, the matrix M is given by:

$$M = \begin{bmatrix} t/2 & h/2 & 0 & 0 \\ t/2 & 0 & h/2 & 0 \\ 0 & 0 & h/2 & t/2 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

As $\lambda = [1 \ 0 \ 0 \ 0]$ and $\gamma = [0 \ 0 \ 0 \ 1]$, (1,4) is the only coefficient of M^* contributing to the series E that represents the experiment:

$$E = M_{14}^* = \frac{1}{4 - (2t + ht)} hh \frac{1}{2 - h} t$$

The generating function of linear random variable 'length' is obtained by substituting s for h and t :

$$F(s) = \frac{s^3}{(4 - 2s - s^2)(2 - s)}$$

The average length is $F'(1) = 8$.

Notes

An interesting extension of Bernoulli measures are Markov measures where we allow the probability of an elementary trial to be influenced by the preceding result. Since Markov measures are rational [Hansel & Perrin, 89] and Hadamard products are computable from linear representations, we can develop analogs of the algorithms of section 2.2.

Another extension would be to relax the computability conditions on random experiments. 'Algebraic' random experiment could be analyzed similarly but it seems much more difficult to

get simple results in that case. The simplest algebraic experiment can be described as repeatedly throwing a penny until the number of heads equals the number of tails. But it can be shown [Feller, 57; Berstel & Perrin, 85] that the only Bernoulli measure for which this set defines a random experiment is when we put $\pi(p) = \pi(f) = 1/2$.

3. References

Bergeron, A., Bordier, J., An Intelligent Discovery Environment for Probability and Statistics, in *Advanced Research on Computers in Education*, R. Lewis and S. Otsuki, Eds, North Holland, 1991.

Berstel, J., Perrin, D., *Theory of Codes*, Academic Press, 1985.

Berstel, J., Reutenauer, C., *Les séries rationnelles et leurs langages*, Masson, 1984.

Char, B., Geddes, K., Gonnet, G., Monagan, M., Watt, S., *MAPLE Reference Manual*, Symbolic Computation Group, Department of Computer Science, University of Waterloo, 1988.

Graham, R., Knuth, D., Patashnik, O., *Concrete Mathematics*, Addison-Wesley, 1989.

Feller, W., *An Introduction to Probability Theory and its Applications*, John Wiley, 1957.

Hansel, G., Perrin, D., Rational Probability Measures, *Theoretical Computer Science*, 65, 1989, 171-188.

Penney, W., Problem 95: Penney-Ante, *Journal of Recreational Mathematics*, 7, 1974, 321.

Schützenberger, M. P., On the synchronizing properties of certain prefix codes. *Inform. and Control* 7, 1964, 23-36.

Shannon, C., *The Mathematical Theory of Communication*, The University of Illinois Press, Urbana, 1949.

GRAPHS ON ALPHABETS

J. Bond

* LaBRI - Laboratoire Bordelais de Recherche en Informatique
U.R.A. 1304 du C.N.R.S., Université Bordeaux I
F-33405 TALENCE Cedex

Abstract

The problems considered here are motivated by the study of the structure of large interconnection networks. Most of the known families of large graphs of given maximum degree and diameter can be defined as graphs on alphabets. We survey the known families of such graphs and deal with possible generalisations.

Résumé

Les problèmes traités dans cet article sont liés à la structure des grands réseaux d'interconnexion. Les plus grands graphes connus de degré maximum et diamètre donnés appartiennent à des familles de graphes définies comme graphes d'alphabets. Nous donnons une synthèse des familles connues et évoquons certaines propriétés de ces graphes et des généralisations possibles.

* research partially supported by P.R.C. C³

1. Introduction

An important consideration in the design of communication networks as well as distributed computer systems is the interconnection network. This network is usually modeled by a graph or a digraph in which the vertices represent the processors. Communication links are represented by edges if they are bidirectional and by arcs if they are unidirectional. Several factors have to be taken into account in the design of interconnection networks:

- communication delays between processors must be short: the graph must have a small diameter or mean distance.
- the number of processors directly connected to a given processor is limited: the graph has a bounded maximum degree.
- the routing of the messages must be simple and distributed if possible.
- the network must be fault-tolerant (in a large system the probability of a faulty processor or link can not be too small)

Let us call (Δ, D) -graph, a graph with given maximal degree Δ and diameter D .

Constructions giving *good* (Δ, D) -graphs (having many vertices) have been extensively studied in the literature (see the recent article of Bermond, Delorme and Quisquater [2] on these topics).

Here we study some families of graphs constructed by using alphabets and propose some equivalent definitions which give raise to certain useful generalizations. These families are the largest known for given maximum degree and diameter.

2. Preliminaries

We use standard terminology (see for example Berge's book [1]). Let us recall some definitions and notations which might be different from the usual terms :

- $G = (V, E)$ will denote either an undirected graph or a directed graph with vertex set V and edge set E .
- In a directed graph (resp. undirected) the *distance* $d(x, y)$ from x to y (resp. between x and y) is the length of a shortest path from x to y (resp. between x and y).
- The *diameter* D is the maximum of the distance over all the pairs of vertices.
- In a directed graph $\Gamma^+(x)$ (resp. $\Gamma^-(x)$) will denote the set of all vertices y such that there is an arc from x to y (resp. from y to x).
- The cardinality of $\Gamma^+(x)$ (resp. $\Gamma^-(x)$) will be denoted by $d^+(x)$ (resp. $d^-(x)$).
- In an undirected graph, the degree $d(x)$ of a vertex x is the number of edges

incidents with x .

- We will denote by Δ the maximum degree of the graph.
- When G is a digraph, we will denote UG the underlying undirected graph associated to G , which is obtained from G by deleting the loops, the orientation and replacing the double edges, if any, by single edges. In some cases we define first an undirected graph G , and later give a digraph, denoted \vec{G} having G as underlying undirected graph.
- A graph is said to be *connected* if between any two vertices there exists a path. A digraph is *connected* if the underlying undirected graph is connected.
- A vertex r in a digraph is a *root* if there is a dipath from r to any other vertex. Conversely, an *anti-root* is a vertex r such that from any vertex there is a dipath to r .

The line digraph of a digraph G is the digraph whose vertices e_1, \dots, e_m represent the arcs of G and which has an arc from e_i to e_j if and only if the terminal endpoint of the arc of G corresponding to e_i is the initial endpoint of the arc of G corresponding to e_j . It will be denoted by LG and the k -th iterated line digraph will be denoted L^kG .

Line digraphs have been characterized by Heuchenne [14] (see [1]) by the following property:

Theorem : A digraph G is the line digraph if and only if it is simple and for any two vertices $x, y \in V$, $\Gamma^+(x) \cap \Gamma^+(y) = \emptyset$ or $\Gamma^+(x) = \Gamma^+(y)$ (the same condition holds with Γ^- instead of Γ^+).

If G is regular with $d^+(x) = d^-(x) = d$ for every vertex x , has diameter D and n vertices, then LG is regular with $d^+(x) = d^-(x) = d$, has dn vertices and, if G is not a cycle, the diameter of LG is $D+1$.

The Kronecker product $G_1 \otimes G_2$ (also called cartesian product [1] or conjunction [13]) of the digraphs $G_1(X_1, E_1)$ and $G_2(X_2, E_2)$ is defined as follows :

- The vertices are $[x_1, x_2]$ with $x_1 \in X_1$ and $x_2 \in X_2$.
- There is an arc from $[x_1, x_2]$ to $[y_1, y_2]$ if and only if $(x_1, y_1) \in E_1$ and $(x_2, y_2) \in E_2$.

The name Kronecker product is used because the adjacency matrix of the product digraph is the Kronecker product of the adjacency matrices. Let us recall the definition of the Kronecker product : the matrix $A \otimes B$ is obtained by replacing the entry $a_{i,j}$ of A by the block $a_{i,j} \times B$.

3. Definitions of some families of graphs on alphabets

We define here some of the best known graphs as underlying graphs of digraphs. All these graphs have as vertex set some words on a given alphabet and the arcs are obtained by a "shift" introducing any possible new letter.

3.1. De Bruijn graphs

The *de Bruijn digraph* $B(d,D)$ (defined in [5]) with in- and outdegree $d \geq 2$ and diameter D is defined in the following way:

- The vertices are the words (x_1x_2, \dots, x_D) where the x_i 's are letters from an alphabet of size d .
- There is an arc from the vertex (x_1x_2, \dots, x_D) to the vertices $(x_2x_3, \dots, x_D, x_{D+1})$, where x_{D+1} can be any letter of the alphabet.
- This digraph has d^D vertices.
- The underlying undirected graph, denoted $UB(d,D)$, is of maximum degree $2d$ (if $D \geq 3$) and diameter D .

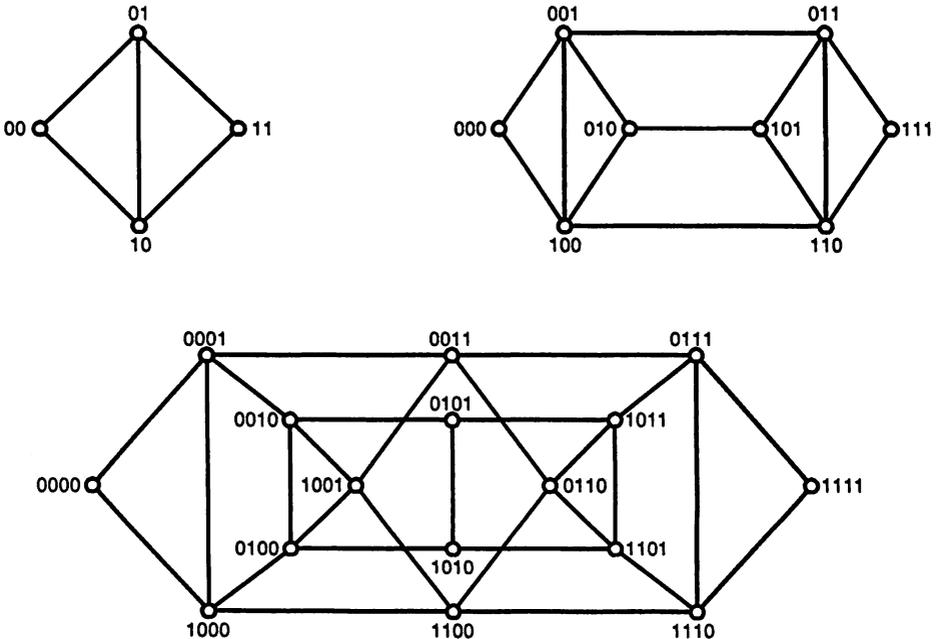


Figure 1 : de Bruijn graphs $UB(2,2)$, $UB(2,3)$ and $UB(2,4)$.

3.2. Kautz graphs

The *Kautz digraph* $K(d,D)$ (defined in [18]) with in- and outdegree $d \geq 2$ and diameter D is the subgraph of the de Bruijn digraph $B(d+1,D)$ induced by the set of vertices without two consecutive identical letters. In other words

- The vertices are labeled with words (x_1, \dots, x_D) , where x_i belongs to an alphabet of $d+1$ letters, and $x_i \neq x_{i+1}$, for $1 \leq i \leq D-1$.
- The vertex (x_1, \dots, x_D) is joined to the d vertices $(x_2, \dots, x_D, x_{D+1})$, where x_{D+1} can be any letter different from x_D .
- This digraph has $d^D + d^{D-1}$ vertices.
- The underlying undirected graph $UK(d,D)$ is of maximum degree $2d$ (if $D \geq 3$) and diameter D .

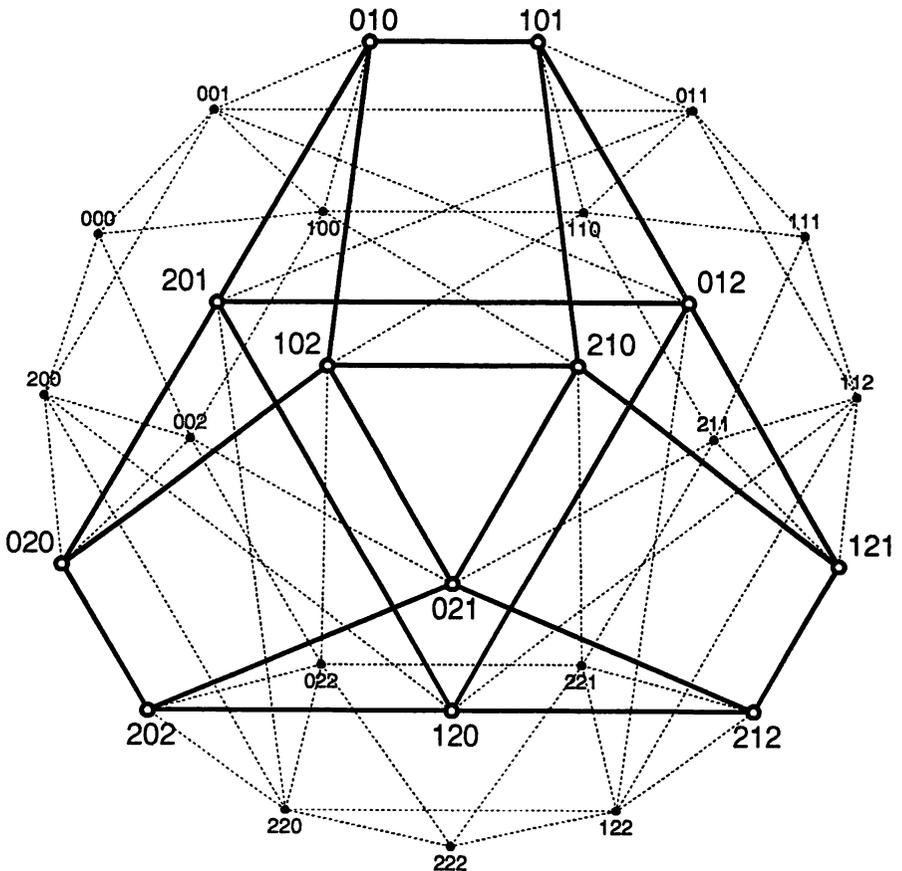


Figure 2 : Kautz graph $UK(2,3)$ as subgraph of $UB(3,3)$

3.3. C_s graphs

The C_s graphs are defined by Delorme and Farhi in [7]. Here we define first a directed version, denoted \vec{C}_s . The digraph $\vec{C}_s(m,d,k)$ with in- and outdegree $d \geq 2$ and diameter $D = k+m-1$ is the digraph such that

- The vertices are the couples $(a ; x_1, x_2, \dots, x_k)$ where a is an integer modulo m and the x_i are letters from an alphabet on d letters.
- There is an arc from a vertex $(a ; x_1, x_2, \dots, x_k)$ to the vertices $(a+1 ; x_2, x_3, \dots, x_k, x_{k+1})$ where x_{k+1} can be any letter of the alphabet.
- This digraph has md^k vertices.
- The underlying undirected graph, denoted $C_s(m,d,k)$, is of maximum degree $2d$ (if $m \geq 3$ or $k \geq 3$). Farhi proved in [8] that the diameter of the C_s graphs is

$$D = k + \sup_{b \in \mathbb{Z}_m} \inf \{ |b-k|, |b+k| \}.$$

De Bruijn graphs can be seen as a subfamily of C_s graphs, namely those with $m=1$.

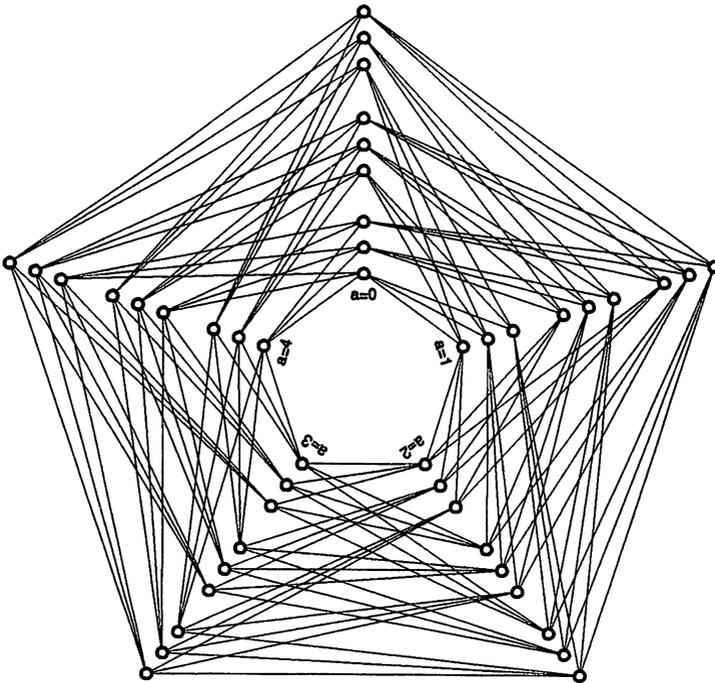


Figure 3 : The graph $C_s(5,3,2)$

The vertices are in the order 00, 01, 02, 10, 11, 12, 20, 21, 22 from inside to outside

3.5. Improved sequence graphs

In [3] we defined the *improved sequence graph* $IS(d_1, d_2, k)$ (an induced subgraph of the sequence graph $S(d_1+1, d_2, k)$, $k \geq 2$) in the following way :

- The vertices are $(x_1, y_1, x_2, y_2, \dots, x_k, y_k)$ where the x_i belongs to an alphabet on d_1+1 letters and $x_i \neq x_{i+1}$, and the y_i are letters of an alphabet on d_2 letters.
- The vertex $(x_1, y_1, x_2, y_2, \dots, x_k, y_k)$ is adjacent to the vertices $(x_{k+1}, y_k, x_k, \dots, y_2, x_2, y_1)$ and $(x_k, y_{k-1}, x_{k-1}, \dots, y_1, x_1, y_0)$ with $x_{k+1} \neq x_k$ and y_0 any letter of the corresponding alphabets.
- The graph $IS(d_1, d_2, k)$ has maximum degree d_1+d_2 , diameter $D = 2k$ and $(d_1+1)d_2 \binom{D-1}{2}^{d_1}$ vertices.

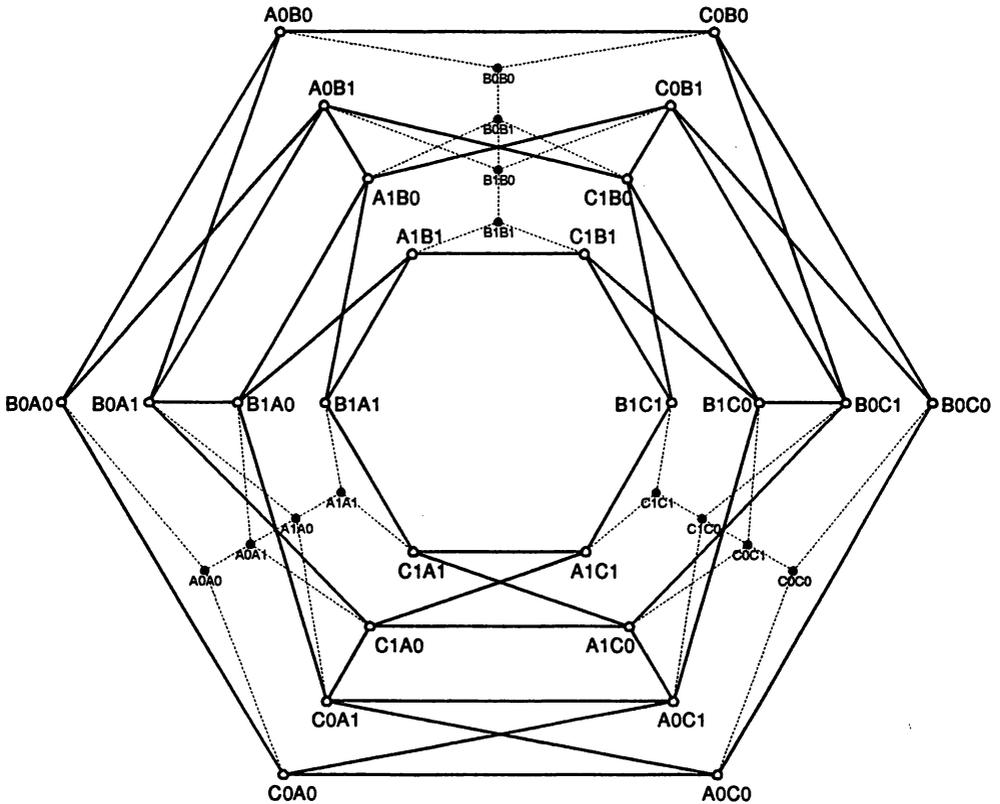


Figure 5 : Improved sequence graph $IS(2,3,2)$ as subgraph of the sequence graph $S(3,3,2)$

3.6. Delorme graphs

The *Delorme graphs* $C'_s(d,k)$ are defined in [6] in the following way :

- The vertices are the words $(\alpha ; x_1, x_2, \dots, x_k)$ with α boolean and the x_i letters of an alphabet on d letters.
- The vertex $(\alpha ; x_1, x_2, \dots, x_k)$ is adjacent to the vertices $(\alpha ; x_{k+1}, x_k, \dots, x_3, x_2)$ and $(\bar{\alpha} ; x_{k-1}, x_{k-2}, \dots, x_1, x_0)$ with x_{k+1} and x_0 any letter of the alphabet.
- The graph $C'_s(d,k)$ has $2d^D$ vertices and maximum degree $2d$ (if $k \geq 3$). Its diameter is k if k is odd and $k+1$ otherwise.

The notation is due to the fact that these graphs are obtained from C_s graphs as follows. The graph $C_s(4,d,k)$ is bipartite by the partition of odd and even prefixes. The diameter is $k+1$ if k is odd and $k+2$ if k is even. A polarity is given by

$$\rho \left[(i ; x_1, x_2, \dots, x_k) \right] = (3-i ; x_k, x_{k-1}, \dots, x_1)$$

By identifying the pairs of vertices of the form $x, \rho(x)$ the diameter is decreased by one, and we obtain the graph $C'_s(d,k)$.

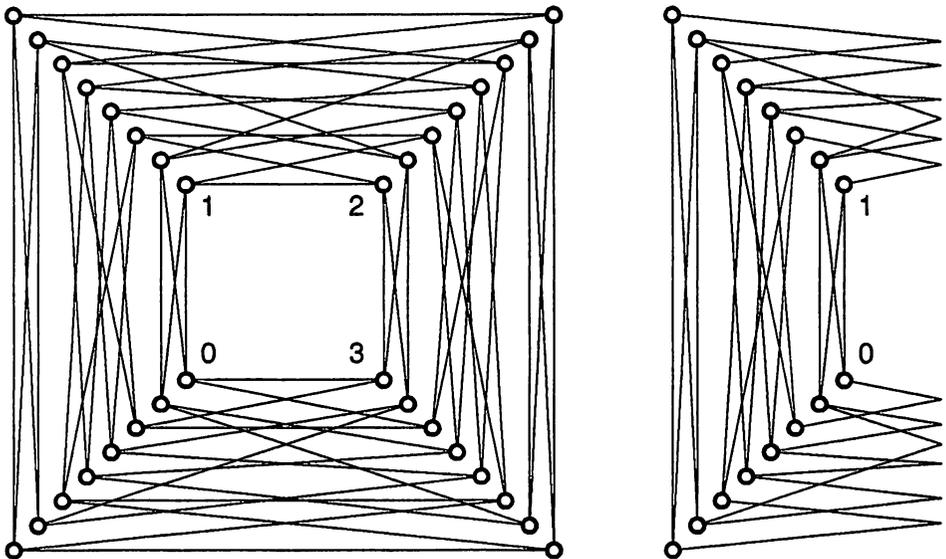


Figure 6 : Delorme graph $C'_s(2,3)$ as quotient by polarity of $C_s(4,2,3)$

The order of the vertices from inside to outside is:

000, 100, 001, 010, 101, 110, 011, 111 for the indices 0 and 2

000, 001, 100, 010, 101, 011, 110, 111 for the indices 1 and 3

3.8. Graph sizes for large degree and diameter

Our main interest is to construct graphs as having as much as possible vertices for given maximum degree Δ and diameter D . As we mentioned in the introduction, the families presented in this section are the largest known. Let us summarize now the number of vertices of the described families.

- If both Δ and D are even, then both the Kautz graph $UK(\frac{\Delta}{2}, D)$ and the improved sequence graph $IS(\frac{\Delta}{2}, \frac{\Delta}{2}, \frac{D}{2})$ have

$$\left(\frac{\Delta}{2}+1\right)\left(\frac{\Delta}{2}\right)^{D-1} \text{ vertices.}$$

- If Δ is even and D is odd, then both the Delorme graph $C'_s(\frac{\Delta}{2}, D)$ and the bisequence graph $T(\frac{\Delta}{2}, \frac{\Delta}{2}, D)$ have

$$2\left(\frac{\Delta}{2}\right)^D \text{ vertices}$$

(and are isomorphic).

- If Δ is odd and D is even, then the improved sequence graph $IS(\frac{\Delta-1}{2}, \frac{\Delta+1}{2}, \frac{D}{2})$ has

$$\left(\frac{\Delta+1}{2}\right)^{\frac{D}{2}+1}\left(\frac{\Delta-1}{2}\right)^{\frac{D}{2}-1} \text{ vertices.}$$

- If both Δ and D are odd, then the bisequence graph $T(\frac{\Delta+1}{2}, \frac{\Delta+1}{2}, D)$ has

$$\Delta\left(\frac{\Delta+1}{2}\right)^{\frac{D-1}{2}}\left(\frac{\Delta-1}{2}\right)^{\frac{D-1}{2}} \text{ vertices.}$$

In this paper we restricted ourselves to these families of graphs. In [11] Gómez, Fiol and Yebra describe some additional graphs obtained by modifications of these graphs on alphabets.

4. Line digraphs

4.1. Some remarks on line digraphs

In what follows we will use the immediate result :

Proposition: Let G be a digraph, and let LG denote its line digraph. Then H is an induced subgraph of LG if and only if there exists H' , subgraph of G , such that $H = LH'$.

Proof: By the theorem of Heuchenne [14] mentioned in section 2, H is a line digraph if and only if for any two vertices x and y of H , $\Gamma^+(x)$ and $\Gamma^+(y)$ are either identical

or disjoint (and as well for Γ^-). By using this result we obtain that any induced subgraph of a line digraph is itself a line digraph. \square

Our aim is to show that each of the graph families described in section 3 can be viewed as the underlying graphs of an iterated line digraph family (or eventually obtained as quotient of such a graph). In order to prove this, we show how each of the digraphs we defined can be obtained as iterated line digraphs. We begin with the C_s graphs, as they are the most general.

4.2. \vec{C}_s digraphs

The digraph $\vec{C}_s(m,d,k)$ defined in section 3.3 can be defined as an iterated line digraph. Indeed let $\vec{C}_{i,j}$ denote the multidigraph obtained by replacing each arc of a circuit of length i by j arcs. Then the following result can be stated :

Proposition : $\vec{C}_s(m,d,k)$ is isomorphic to $L^k \vec{C}_{m,d}$.

Proof : We prove it by induction, using the following two statements:

* $\vec{C}_s(m,d,1)$ is isomorphic to $L \vec{C}_{m,d}$

* $\vec{C}_s(m,d,k+1)$ is isomorphic to $L \vec{C}_s(m,d,k)$

The statements are immediate by denoting the arc of $\vec{C}_s(m,d,k)$ $(i ; x_1 x_2, \dots, x_k) \rightarrow (i+1 ; x_2 x_3, \dots, x_k x_{k+1})$ by $[i ; x_1 x_2, \dots, x_k x_{k+1}]$, which enables us to associate it to the vertex $(i ; x_1 x_2, \dots, x_k x_{k+1})$ of the line digraph. \square

4.3. De Bruijn digraphs

In order to prove that De Bruijn digraphs can be defined as line digraphs we remark that De Bruijn digraphs form a sub-family of C_s digraphs, namely those with $m=1$. In order to obtain the definition of $B(d,D)$ as a line digraph we observe that the circuit on one vertex is a selfloop. Therefore the definition of $B(d,D)$ as an iterated line digraph gives :

Corollary : Let P_i be the multidigraph on one vertex with i self-loops. Then $B(d,D)$ is isomorphic to $L^D P_i$.

We can also define the digraph $B(d,D)$ as $L^{D-1} B(d,1)$, where $B(d,1)$ is in fact the complete symmetric digraph on d vertices, with selfloops in every vertex.

4.4. Kautz digraphs

The observation that Kautz digraphs can be defined as line digraphs is due to Fiol, Yebra and Alegre [9]. As $K(d,D)$ is defined for $D \geq 2$, let us first check for $K(d,2)$. It is exactly the subgraph of $B(d+1,2)$ induced by the vertices (x_1, x_2) with $x_1 \neq x_2$. One can easily check that $K(d,2)$ is a line digraph. We must have a G such that $K(d,2) = LG$. In this way we obtain the following definition of $K(d,D)$ as an iterated line digraph :

Proposition : Let K_i^* be the complete symmetric digraph on i vertices without selfloops. Then $K(d,D)$ is isomorphic to $L^{D-1}K_d^*$.

4.5. Sequence digraphs

Since the sequence graphs are defined as undirected graphs, let us first define the following directed graphs :

- $\vec{S}(d_1, d_2, k)$ has as vertex set the words $(x_1, y_1, x_2, y_2, \dots, x_k, y_k)$ and the words $(y_1, x_1, y_2, x_2, \dots, y_k, x_k)$, where the x_i 's are letters from an alphabet of size d_1 and the y_i 's are letters from an alphabet of size d_2 .
- There is an arc from the vertex $(x_1, y_1, x_2, y_2, \dots, x_k, y_k)$ to the vertices $(y_1, x_2, y_2, \dots, x_k, y_k, x_{k+1})$, where x_{k+1} can be any letter from the first alphabet.
- In the same manner, there is an arc from the vertex $(y_1, x_1, y_2, x_2, \dots, y_k, x_k)$ to the vertices $(x_1, y_2, x_2, \dots, y_k, x_k, y_{k+1})$, where y_{k+1} can be any letter from the second alphabet.

Let now G denote the quotient digraph defined by the polarity identifying a vertex and its opposite (i.e. $(x_1, y_1, x_2, y_2, \dots, x_k, y_k)$ and $(y_k, x_k, y_{k-1}, x_{k-1}, \dots, y_1, x_1)$). In other words, as a vertex is identified to its opposite, now every vertex has two labels, and one of them, which will be used, begins with an x . Then $S(d_1, d_2, k)$ is the underlying undirected graph UG .

$\vec{S}(d_1, d_2, k)$ can be defined as an iterated line digraph, in the following way :

Proposition : Let P_{d_1, d_2} be the multidigraph on two vertices, with d_1 arcs in one direction and d_2 arcs in the other direction. Then $\vec{S}(d_1, d_2, k)$ is isomorphic to $L^{2k}P_{d_1, d_2}$.

As in the case of \vec{C}_s digraphs, this proposition can be proved by induction, namely by proving:

- $\vec{S}(d_1, d_2, 1)$ is isomorphic to $L^2P_{d_1, d_2}$

- $\vec{S}(d_1, d_2, k+1)$ is isomorphic to $L^2\vec{S}(d_1, d_2, k)$

4.6. Improved sequence digraphs

As in the case of sequence graphs we should first define the directed improved sequence graphs. Let $\vec{I}\vec{S}(d_1, d_2, k)$ be the subgraph of $\vec{S}(d_1, d_2, k)$ induced by the vertices associated to words without two equal consecutive x_i 's. The undirected graph is obtained in the same manner as the sequence graph.

One can easily check that as the \vec{C}_s digraphs, the $\vec{I}\vec{S}(d_1, d_2, k)$ are line digraphs for $k \geq 4$. Let $\vec{S}_{i,j}$ denote the digraph obtained from $\vec{S}(i+1, j, 2)$ after removal of all the arcs $(x, y) \rightarrow (y, x)$, where x is a letter of the first alphabet and y a letter of the second. $\vec{I}\vec{S}(d_1, d_2, k)$ can be defined as an iterated line digraph. in the following way :

Proposition : $\vec{I}\vec{S}(d_1, d_2, k)$ is isomorphic to $L^{2k-2}\vec{S}_{d_1, d_2}$.

As in the case of \vec{C}_s digraphs, this equality can be proved by induction, namely by proving:

- $\vec{I}\vec{S}(d_1, d_2, 2)$ is isomorphic to $L^2\vec{S}_{d_1, d_2}$
- $\vec{I}\vec{S}(d_1, d_2, k+1)$ is isomorphic to $L^2\vec{I}\vec{S}(d_1, d_2, k)$

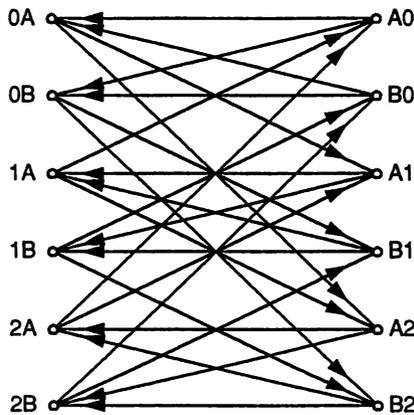


Figure 8 : the digraph $\vec{S}_{2,2}$

4.7. Bisequence digraphs

As in the case of sequence graphs we should first define the directed bisequence graphs :

- $\vec{T}(d_1, d_2, k)$ ($k \equiv 1 \pmod{4}$, $k \geq 5$) has as vertex set the words

$$\begin{aligned} & (x_1, x_2, y_3, y_4, \dots, y_{k-2}, y_{k-1}, x_k), \\ & (x_1, y_2, y_3, x_4, \dots, y_{k-2}, x_{k-1}, x_k), \\ & (y_1, y_2, x_3, x_4, \dots, x_{k-2}, x_{k-1}, y_k) \end{aligned}$$

and

$$(y_1, x_2, x_3, y_4, \dots, x_{k-2}, y_{k-1}, y_k),$$

where the x_i 's are from an alphabet on d_1 letters and the y_i 's from an alphabet on d_2 letters.

- The arcs are obtained by doing a shift to the right, so for example there is an arc from the vertex $(x_1, x_2, y_3, y_4, \dots, y_{k-2}, y_{k-1}, x_k)$ to the vertices $(x_2, y_3, y_4, \dots, y_{k-2}, y_{k-1}, x_{k+1})$, where x_{k+1} can be any letter of the first alphabet.

Let now G denote the quotient digraph defined by the polarity identifying a vertex and its opposite (i.e. $(x_1, x_2, y_3, y_4, \dots, y_{k-2}, y_{k-1}, x_k) \longleftrightarrow (x_k, y_{k-1}, y_{k-2}, \dots, y_4, y_3, x_2, x_1)$). Then $T(d_1, d_2, k)$ is the underlying undirected graph UG .

$\vec{T}(d_1, d_2, k)$ can be defined as an iterated line digraph, in the following way :

Proposition : Let $C_4^{d_1, d_2}$ be the multidigraph on the four vertices 0, 1, 2 and 3 with d_1 arcs from 0 to 1 and from 1 to 2, and d_2 arcs from 2 to 3 and from 3 to 0. Then $\vec{T}(d_1, d_2, k)$ is isomorphic to $L^k C_4^{d_1, d_2}$.

As in the case of \vec{C}_s digraphs, this equality can be proved by induction, namely by proving:

- $\vec{T}(d_1, d_2, 5)$ is somorphic to $L^5 C_4^{d_1, d_2}$
- $\vec{T}(d_1, d_2, k+4)$ is isomorphic to $L^4 \vec{T}(d_1, d_2, k)$

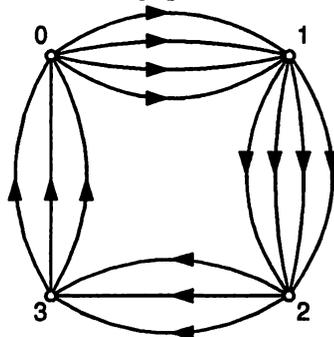


Figure 9 : the digraph $C_4^{4,3}$

5. Algebraic definitions

We give here some algebraic definitions, enabling us to define the de Bruijn and Kautz digraphs algebraically, and also to make some generalizations. Reddy, Pradhan and Kuhl [21] and Imase and Itoh [16] gave a definition of the de Bruijn digraphs in arithmetic terms. They define the digraph $B(d,D)$ in the following way :

The digraph $B(d,D)$ has as its vertex set the integers $\{0, 1, \dots, d^D-1\}$. There is an arc from the vertex x to the vertices $dx+a \pmod{d^D}$, where $0 \leq a \leq d-1$.

This definition gave rise to a generalization of the de Bruijn digraphs denoted $RPK_{d,n}$. The $RPK_{d,n}$ digraph has as its vertex set the integers modulo n and the same adjacency rules as de Bruijn digraphs. The special interest of $RPK_{d,n}$ digraphs is that they exist for any n and not only for powers of d . Reddy, Pradhan and Kuhl prove in [21] the following result about the diameter of the graphs $RPK_{d,n}$:

Theorem : The diameter of $RPK_{d,n}$ is $\lceil \log_d n \rceil$.

A generalisation of Kautz digraph is due to Imase and Itoh [17]. They define the digraph $K(d,D)$ in the following way :

The digraph $K(d,D)$ has as its vertex set the integers $\{0, 1, \dots, d^D+d^{D-1}-1\}$.

There is an arc from the vertex x to the vertices $-dx-a \pmod{d^D+d^{D-1}}$, where $1 \leq a \leq d$.

This definition gave rise to a generalization of Kautz digraphs denoted $II_{d,n}$. A digraph in this class has as its vertex set the integers modulo n and the same adjacency rules as Kautz digraphs. We are interested in these digraphs, because like the generalized de Bruijn digraphs, they exist for any n . Imase and Itoh do not give an exact formula for the diameter of digraphs $II_{d,n}$, but prove the following result :

Theorem : The diameter D of the digraph $II_{d,n}$ verifies $\lceil \log_d n \rceil \geq D \geq \lfloor \log_d n \rfloor$.

The equivalence of these two definitions with those given in section 3 is proved in [9] and [15].

Let us now recall some facts about the adjacency matrix of a digraphs. Let A_G denote the adjacency matrix of the digraph G . Then, it is well-known that the ij entry of A_G^k designates the number of paths of length k from the vertex i to the vertex j . In these terms the diameter of G is

$$\min\{k \mid \sum_{i=0}^k A_G^i \text{ has no null entries} \}.$$

Since the de Bruijn digraph has a unique directed path of length exactly D between each pair of vertices, then for its adjacency matrix A we have $A^D = J$, where J is the matrix where all the entries are 1. In case of Kautz graphs, we obtain the matrix equation $A^D + A^{D-1} = J$.

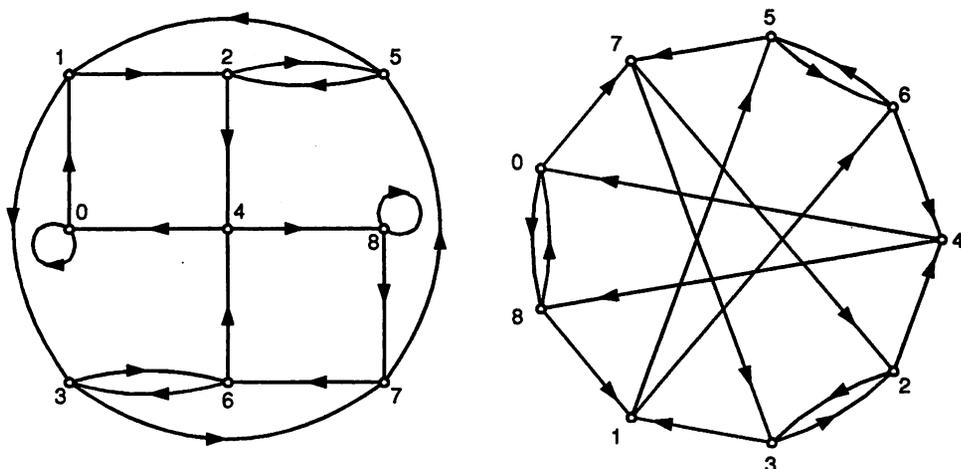


Figure 10 : the digraphs $RPK_{2,9}$ and $H_{2,9}$

Recently we found an earlier paper of Bosák [4] where he rediscovered de Bruijn and Kautz digraphs that he defined in an arithmetic way. The original problem studied by Bosák was not the (Δ, D) problem. He called a digraph a W_a^b -digraph, if for any pair of vertices (x, y) there is a unique path of length at least a and at most b from x to y (x and y are not necessarily distinct). Consequently $\sum_{i=a}^b A_G^i = J$. By using this definition, a W_0^D -digraph is a Moore digraph. As his main result Bosák proved that W_a^b -digraphs exist only for the cases $b=a$ and $b=a+1$, except the trivial case when the digraph is a circuit. The two families of digraphs he proposed are exactly the de Bruijn and Kautz digraphs, defined in the following way.

The digraph $B(d, D)$ has as vertex set the integers $\{0, 1, \dots, d^D - 1\}$. If $D = 0$ then there is only one vertex and d selfloops. In the other cases there is an arc from the vertex x to the vertices $hd^{D-1} + \left\lfloor \frac{x}{d} \right\rfloor$, where $0 \leq h \leq d-1$.

The digraph $K(d, D)$ has as its vertex set the integers $\{0, 1, \dots, d^D + d^{D-1} - 1\}$. If $D = 1$ then the digraph is the complete symmetric digraph on $d+1$ vertices. In the other cases there is an arc from the vertex x to the vertices

$$h(d^{D-1} + d^{D-2}) - \left\lfloor \frac{x}{d} \right\rfloor - 1, \text{ where } 1 \leq h \leq d.$$

We remark that in the digraphs defined by Bosák, the arcs are directed exactly in the opposite direction, as in the definition of Reddy, Pradhan and Kuhl and Imase and Itoh respectively. However, in case of de Bruijn and Kautz digraph, there exists anti-automorphisms (for example $(x_1, x_2, \dots, x_D) \rightarrow (x_D, x_{D-1}, \dots, x_1)$) so we obtain again the same digraphs.

A related problem was also studied by Lam and van Lint [19]. The problem studied by Lam and van Lint is the design of digraphs with a unique path (not necessarily elementary) of given length from any vertex to any other. In terms of matrices, that gives the equation $A^k = J - I$. Digraphs satisfying this equation exist only for odd k . The adjacency matrices of the digraphs for odd k 's are obtained in a special manner, namely by using blocks which are permutation matrices. The digraphs obtained are different from de Bruijn and Kautz digraphs.

However the ideas used by Lam and van Lint can also give rise to a new definition of de Bruijn digraphs.

We first give a method in the case of diameter 2. Let K_d^* the complete symmetric digraph on d vertices. Let us make a partition of the arcs into d 1-regular digraphs G_1, G_2, \dots, G_d . We denote the adjacency matrix of G_i by M_i . Let M be the $d^2 \times d$ matrix having $d \times d$ blocks of size $d \times d$, where the blocks in the i -th column are M_i . Let G_M the digraph having M as adjacency matrix. Then G_M is isomorphic to the de Bruijn digraph $B(d, 2)$.

The matrix M can be defined formally in the following way : let denote by C_i the matrix with null entries except the i -th column, where all the entries are 1. Then

$$M = \sum_{i=1}^d M_i \otimes C_i,$$

where $A \otimes B$ designs the usual Kronecker product of the matrices A and B .

Before giving a general formula, let us present the adjacency matrices of de Bruijn and Kautz digraphs. By using the definitions given by Reddy, Pradhan and Kuhl, there are arcs from the vertex 0 to the vertices $0, 1, \dots, d-1$, then arcs from the vertex 1 to the vertices $d, d+1, \dots, 2d-1$ and so on (see figure 11 for an example).

In the general case we can give the following construction (issued from discussions with J-J. Quisquater) :

- Let I_i denote the the $i \times i$ identity matrix.

- Let P_d^i denote the $d \times d$ matrix with all the entries zero except the i,i entry which is 1.
- Let C_d^i the $d \times d$ matrix with all the entries zero except the i -th column where all the entries are 1.
- Let $M_{d,1}$ be the $d \times d$ matrix with all the entries 1.
- Let $M_{d,k}$ be defined for $k \geq 2$ in the following way :

$$M_{d,k+1} = (M_{d,k} \otimes I_d) (I_{d^{k-1}} \otimes (\sum_{i=1}^d P_d^i \otimes C_d^i))$$

- Let denote $G_{d,k}$ the digraph having $M_{d,k}$ as adjacency matrix.

Then $G_{d,k}$ is isomorphic to $B(d,k)$.

$$\begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix}$$

Figure 11 : the adjacency matrix of $B(2,3)$

It is well known that the de Bruijn graph $UB(d,D)$ is isomorphic to an induced subgraph of the de Bruijn graph $UB(d+1,D)$. There is no such relation between $UB(d,D)$ and $UB(d,D+1)$. Indeed one can show that $UB(d,D)$ cannot be an induced subgraph of $UB(d,D+1)$ as the distance between the vertices of degree $2d-2$ is not the same. The preceding recursive definition of the adjacency matrices links directly the structure of $B(d,D+1)$ to that of $B(d,D)$. In the last section we continue this approach.

By using the Imase and Itoh definition of Kautz digraphs we can also easily describe the adjacency matrix of the Kautz digraph. Once again there are d consecutive ones and then in the next ligne we continue, but in this case from left to right (see figure 12 for an example of adjacency matrix of Kautz digraph).

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Figure 12 : the adjacency matrix of $K(2,2)$

6. Decompositions into partial graphs

Let us first remark that the adjacency matrix of the de Bruijn digraph, $B(d,D)$ as given above can be decomposed into d identical blocks, each one composed of d^{D-1} successive rows of the matrix. One can partition the arc set of $B(d,D)$ into d sets corresponding to the blocks. Let G_i the partial digraph having as arc set the arcs corresponding to the i -th block. In G_i in every vertex there is a unique outgoing arc. Moreover, one can check that G_i is connected.

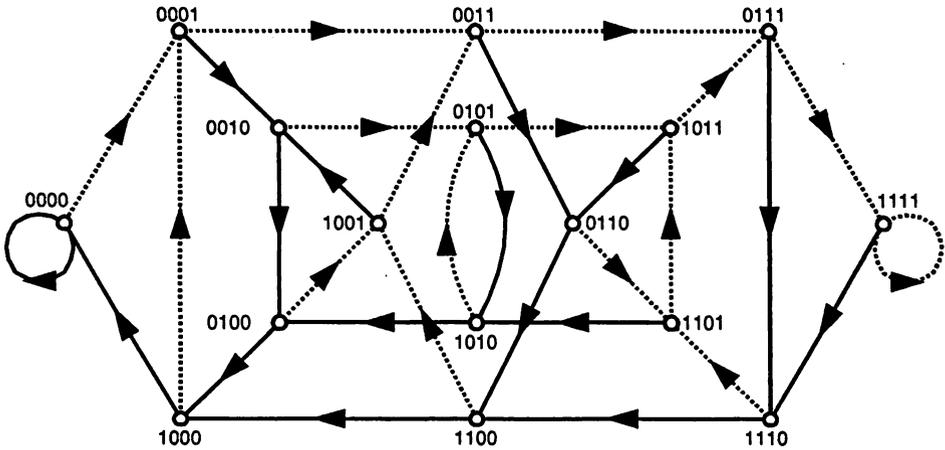


Figure 13 : the decomposition of $B(2,4)$ into two partial digraphs
the two types of arcs correspond to the two partial digraphs

By checking these two properties we obtain the following remark :

Remark : Let G be a connected finite digraph such the out-degree of all the vertices is one. Then there is at least one anti-root in G .

The remark holds also in the case when each vertex has in-degree one and gives the existence of a root.

By this result one can already prove the existence of an "anti-tree" plus a selfloop on the root in the decomposition of $B(d,D)$. In fact this property is a consequence of the line digraph structure. The decomposition property is hereditary on iterated line digraphs.

Theorem : Let G be a regular finite digraph, which can be decomposed into the sum of *connected* partial digraphs in which all the vertices have out-degree one. Then $L^k G$ can also be decomposed in the same way.

Proof : It is enough to prove the statement for LG (for other iterations one can conclude by induction). Let G be a d -regular digraph on n vertices. The partition of the arcs of G can be seen as a coloration with d colors of the arcs in such a way that the arcs leaving a given vertex have different colors. This way the vertices of LG are colored in d colors. Let us now color the arcs of LG by the color of their terminal endpoint. We claim that this coloring has the properties we need.

Let xy be a vertex of LG (denoted by the endpoints of the corresponding arc of G). Let suppose that there are two arcs of the same color issued from xy , (xy, yv) and (xy, yw) for example. That means that the vertices yv and yw are of the same color, so the arcs (y, v) and (y, w) in G are of the same color, a contradiction.

Now let us show that the partial digraph given by each color is connected. In fact we will prove that there exists an anti-root. Let us choose a color γ , and let r be an anti-root of the partial digraph of G formed by the arcs of the color γ . Let (r, t) be the arc of color γ leaving r in G . Then rt is an anti-root in the partial digraph of LG of color γ . Indeed, let xy be a vertex of LG . Let P be dipath from y to r in G , $P = y = y_0, y_1, \dots, y_k = r$. In LG the arcs $(xy, y_1), (y_1, y_1 y_2), \dots, (y_{k-2} y_{k-1}, y_{k-1} r), (y_{k-1} r, rt)$ are of the color γ .

□

The partial digraphs of the line digraph can in fact be obtained in the following way :

- take a partial digraph G_γ of the digraph G
- add $(d-1)n$ new vertices
- for every new vertex add an arc to one of the vertices of G_γ , in such a way that every vertex of G_γ has exactly d predecessors.

The lemma can also be stated in case of in-degree one, and then it gives the existence of a root.

We can remark that the decomposition of the de Bruijn digraph we found by the matrix blocks in the beginning of this section is exactly one of those given by the lemma, obtained from the coloring of the arcs of the digraph on one vertex and d loops by d different colors. In fact, in the case of de Bruijn digraph this decomposition is unique (up to anti isomorphism), as the d loops have to be of different colors. In term of alphabets, the partial digraphs we obtain correspond to arcs which introduce always the same letter.

If we apply this method to the Kautz digraphs there are many possible decompositions. An example of decomposition giving isomorphic partial digraphs is obtained if we first decompose the complete symmetric digraph, and then use the construction of the lemma.

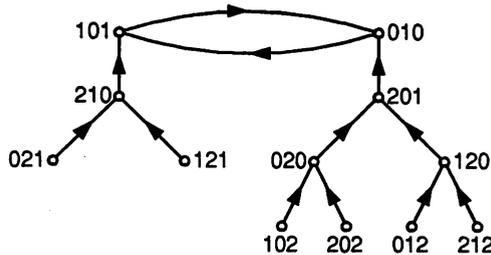


Figure 14 : an example of partial digraph of the decomposition of $K(2,3)$.

The case of $\vec{C}_s(m,d,k)$ digraphs is also simple. As we have shown, $\vec{C}_s(m,d,k) = L^k \vec{C}_{m,d}$. The only way to decompose $\vec{C}_{m,d}$ is to obtain d circuits of length d (and again we obtain isomorphic partial digraphs).

The sequence digraphs can be decomposed this way only in the case they are regular, which means that the two alphabets are of same size. In this case as we have seen the sequence digraph $\vec{S}(d,d,k) = L^{2k} P_{d,d}$. The complete symmetric digraph $P_{d,d}$ can be decomposed example into two trees (or anti-trees) with a double arc between the roots.

In the case of bisequence digraphs the decomposition can only be made if the two alphabets are of the same size, and the result is the same as in case of the corresponding \vec{C}_s digraph.

As in case of sequence digraphs, the decomposition of the improved sequence graphs can only be made if the digraph is regular, in the case $\vec{S}(d+1,d,k)$. However a decomposition into isomorphic partial digraphs seems to be more difficult to obtain, as already the decomposition of $\vec{S}_{d,d}$ is not immediate.

Let us now mention an other kind of decomposition. Let us decompose the arcs of the complete symmetric digraph on $d+1$ vertices into $d+1$ stars. One can check that using methods as in the lemma, one can decompose $K(d,D)$ into $d+1$ partial digraphs, each one consisting of d^{D-1} stars. In other words, the Kautz digraph can be obtained by a special assembling of these stars. It would be interesting to find out if by assembling not trees (as in the given decompositions) but other digraphs one could obtain better results.

7. Projections

Let us first define the notion of *projection*. Let $G_1(V_1, E_1)$ and $G_2(V_2, E_2)$ be two digraphs. Let $f_p : V_1 \rightarrow V_2$ be a surjective mapping, such that $(x_2, y_2) \in E_2$ if and only if there are $x_1, y_1 \in V_1$ such that $f_p(x_1) = x_2$ and $f_p(y_1) = y_2$ with $(x_1, y_1) \in E_1$. Then we will say that G_2 is the projection of G_1 .

In other words G_2 is the projection of G_1 if G_2 can be obtained from G_1 by the identification of some vertices followed by the removal of the multiple arcs and loops.

First let us show that the line digraph structure gives a simple projection. An immediate result can be stated as follows :

Proposition : Let G be a digraph and LG its line digraph. The *standard projection* s_p mapping a vertex xy of LG (obtained from the arc (x, y) of G) into the vertex x of G is a projection if and only if G does not have any vertex of out-degree zero.

Proof : Let (x, y) be an arc of G . As $d^+(y) > 0$, there exist an arc (y, z) leaving y in G . So the arc (xy, yz) of LG gives the arc (x, y) of G by the standard projection. Conversely if (x, y) is an arc of the projection, then there exist some z such that (xy, yz) is an arc of LG , so xy is a vertex of LG obtained from the arc (x, y) of G .

If there exist a vertex x in G with $d^+(x) = 0$, then the vertex x is not obtained by the standard projection.

□

As a corollary we obtain that all the families we described as line digraphs have projections.

In what follows we will propose some projections different from this standard projection. First we remark that the existence of a projection of G_1 onto G_2 does not necessarily imply the existence of a projection of LG_1 onto LG_2 . For example, let us take a large regular digraph G_1 with small degree, which can be projected onto a smaller regular digraph G_2 of larger degree. The number of vertices of $L^k G_2$ grows faster than the number of vertices of $L^k G_1$.

Let us now define a *uniform projection*, that is a projection f_p such that if $(f_p(x), f_p(y))$ is an arc then there exist vertices v and w (not necessarily distinct from x and y) satisfying :

$$* \ x, v \text{ is an arc and } f_p(v) = f_p(y)$$

$$* \ w, y \text{ is an arc and } f_p(w) = f_p(x)$$

The uniform projection property is conserved for line digraphs :

Proposition : Let G_1 and G_2 two digraphs, such that there exist a uniform projection f_p of G_1 onto G_2 . Then there exist a uniform projection of LG_1 onto LG_2 .

Proof : We will prove that the mapping $g : LG_1 \rightarrow LG_2$ defined by $xy \rightarrow f_p(x)f_p(y)$ is a uniform projection. First let us prove that g is a projection.

Let (xy,yz) be an arc of $L(G_1)$. Let a, b, c be the projection of x, y, z . Then clearly (ab,bc) is an arc of $L(G_2)$. Conversely, Let (ab,bc) be an arc of LG_2 . Let x be a vertex of G_1 such that $f_p(x) = a$. As f_p is a uniform projection, there exists a vertex y in G_1 such that (x,y) is an arc of G_1 and $f_p(y) = b$. In the same way, there exists a vertex z of G_1 such that (y,z) is an arc of G_1 and $f_p(z) = c$. Now we are done, as (xy,yz) is an arc of LG_1 mapped into (ab,bc) by g . We leave to the reader the proof that g is also uniform. □

As a first corollary let us propose a uniform projection of the de Bruijn digraph $B(d,D+1)$ onto the de Bruijn digraph $B(d,D)$. We define the following uniform projection of $B(d,2)$ onto $B(d,1)$.

- * the vertices of $B(d,2)$ are labeled by couples of numbers on $0, 1, \dots, d-1$.
- * the vertices of $B(d,1)$ (the complete symmetric digraphs with selfloops) are labeled by the numbers $0, 1, \dots, d-1$.
- * the vertex ij of $B(d,2)$ is mapped to the vertex $j-i \pmod{d}$ of $B(d,1)$.

It is easy to verify that we obtain a uniform projection. By the proposition above, this projection induces a uniform projection of $B(d,D+1)$ onto $B(d,D)$. The projection associates to the vertex (a_1, \dots, a_{D+1}) of $B(d,D+1)$ the vertex $(a_2-a_1, a_3-a_2, \dots, a_{D+1}-a_D)$.

This projection gives rise to a new definition of de Bruijn digraphs :

- The vertices are the arrays $(x ; z_1, z_2, \dots, z_{D-1})$, where x and the z_i are in \mathbb{Z}_d .
- There is an edge between the vertices $(x ; z_1, z_2, \dots, z_{D-1})$ and $(x+z_1 ; z_2, \dots, z_{D-1}, z_D)$, where z_D can take any value in \mathbb{Z}_d .

As an other corollary let us propose a uniform projection of Kautz digraphs onto de Bruijn digraphs.

Let us define the following uniform projection of $K(d,2)$ onto $B(d,1)$.

- * the vertices of $K(d,2)$ are labeled by couples of numbers on $0, 1, \dots, d$.
- * the vertices of $B(d,1)$ are labeled by the numbers $1, 2, \dots, d$.
- * the vertex ij of $K(d,2)$ is mapped to the vertex $j-i \pmod{d+1}$ of $B(d,1)$.

It is easy to verify that we obtain a uniform projection. This projection allows to give a new definition of Kautz digraphs :

- The vertices are the arrays $(x ; z_1, z_2, \dots, z_{D-1})$, where x and the z_i are in \mathbb{Z}_{d+1} , $z_i \neq 0$.

- There is an edge between the vertices $(x ; z_1, z_2, \dots, z_{D-1})$ and $(x+z_1 ; z_2, \dots, z_{D-1}, z_D)$, where z_D can take any value in \mathbb{Z}_{d+1} , except 0.

8. Definitions as graph products

Graph products are often used for construction of new large graphs. We remark for example that the \vec{C}_s digraph can be considered as the Kronecker product of the de Bruijn digraph and a circuit. Here we define a new product, giving as result the sequence graphs and the improved sequence graphs.

Definition : Let $G_1(X_1, E_1)$ and $G_2(X_2, E_2)$ be two digraphs. Let $x_1 \rightarrow x_1$ (resp. $x_2 \rightarrow x_2$) be an involutive antiautomorphism in G_1 (resp. G_2). We recall, that an antiautomorphisme is a mapping of the vertices reversing the arcs. We define the *alternating product* graph $G = G_1 \boxtimes G_2$ as follows :

- The vertices are $[x_1, x_2]$ with $x_1 \in X_1$ and $x_2 \in X_2$
- There is an edge between the vertices $[x_1, x_2]$ and $[y_1, y_2]$ in one the two cases
 - i) $(x_1, y_1) \in E_1$ and $y_2 = x_2$
 - ii) $(y_2, x_2) \in E_2$ and $y_1 = x_1$

Let us check the properties of the graph G . We remark that as the mapping is an antiautomorphism and is involutive the following equalities hold : $\overline{\Gamma^-(x)} = \Gamma^+(x)$ and $\overline{\Gamma^+(y)} = \Gamma^-(y)$ both in G_1 and G_2 . That implies a symmetrie in the definition, so that an edge is "of the same type" when checked from its two endpoints.

Therefore if G_1 is d_1 regular and G_2 is d_2 regular, then the maximum degree will be at most $d_1 + d_2$.

The resulting graph will not necessarily be connected. However, the following result about the connectivity, is immediate (the condition is the same as in the case of cartesian product, in the result due to McAndrews [20]) :

Proposition : If G_1 and G_2 are two strongly connected digraphs with the circuit sets $M(G_1)$ and $M(G_2)$ respectively, then the alternating product $G_1 \boxtimes G_2$ is connected if, and only if, the length of the circuits in $M(G_1) \cup M(G_2)$ are relatively prime.

However as the example shows, the diameter can be very large. Obviously, the diameter will be at least the sum of the diameters of the two digraphs. Let us check the diameter in two special cases :

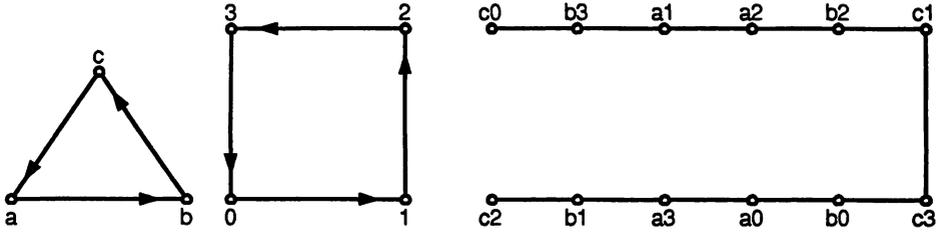


Figure 15 : $\vec{C}_3 \boxtimes \vec{C}_4$, with the involution

mappings $\bar{a} = a$; $\bar{b} = c$; $\bar{0} = 0$; $\bar{1} = 3$; $\bar{2} = 2$;

• In the first case let us suppose, that there exists a dipath of length exactly k from any vertex to any vertex in G_1 and in G_2 , for some k . In this case, there exist a path of length at most $2k$ between any two vertices in $G_1 \boxtimes G_2$. Indeed, let $[x_1, x_2]$ be a vertex of $G_1 \boxtimes G_2$. We obtain

* at distance 1 from $[x_1, x_2]$: vertices $[y_1, y_2]$ such that $y_1 \in \Gamma^+(x_1)$ and $y_2 = x_2$

* at distance 2 from $[x_1, x_2]$: vertices $[y_1, y_2]$ such that $y_1 \in \overline{\Gamma^+(x_1)} = \Gamma^-(x_1)$ and $y_2 \in \Gamma^-(x_2) = \Gamma^-(x_2)$

...

* at distance $2k$ from $[x_1, x_2]$: vertices $[y_1, y_2]$ such that $y_1 \in \Gamma^{-k}(x_1)$ and $y_2 \in \Gamma^{-k}(x_2)$

• In the second case let us suppose, that there exist a dipath of length k or $k-1$ from any vertex to any vertex in G_1 and a dipath of length exactly k from any vertex to any vertex in G_2 , for some k . In this case also, there exist a path of length at most $2k$ between any two vertices in $G_1 \boxtimes G_2$. We only need to check the case where the distance in G_1 is $k-1$ (the other case is as before). Indeed, let $[x_1, x_2]$ be a vertex of $G_1 \boxtimes G_2$.

We obtain

* at distance 1 from $[x_1, x_2]$: vertices $[y_1, y_2]$ such that $y_1 = x_1$ and $y_2 \in \Gamma^-(x_2)$

* at distance 2 from $[x_1, x_2]$: vertices $[y_1, y_2]$ such that $y_1 = \bar{x}_1 = x_1$ and $y_2 \in \Gamma^-\left[\overline{\Gamma^-(x_2)}\right] = \Gamma^-(\Gamma^+(x_2))$

...

* at distance $2k$ from $[x_1, x_2]$: vertices $[y_1, y_2]$ such that $y_1 \in \Gamma^{-(k-1)}(x_1)$ and $y_2 \in \Gamma^{-k}\left[\overline{\Gamma^+(x_2)}\right]$

We remark that the two cases correspond to the distance in de Bruijn and respectively in Kautz digraphs. Now it is immediate to show that :

Proposition :

i) $S(d_1, d_2, k)$ is isomorphic to $B(d_1, k) \boxtimes B(d_2, k)$

ii) $IS(d_1, d_2, k)$ is isomorphic to $K(d_1, k) \boxtimes B(d_2, k)$

Proof : In order to obtain the isomorphism we associate the couple $[(x_1, \dots, x_k), (y_1, \dots, y_k)]$ to the vertex $(x_1, y_1, \dots, x_k, y_k)$ of the sequence graph. This way the adjacency conditions are verified.

□

References

1. Berge, C., *Graphs and hypergraphs*, North-Holland Publishing Co., New York, 1973.
2. Bermond, J.-C., Delorme, C., and Quisquater, J.-J., "Strategies for interconnection networks," *J. of Parallel and Distributed Computing*, vol. 3, no. 4, pp. 433-449, 1986.
3. Bond, J., "Construction de grands réseaux d'interconnexions," *Thesis*, LRI, Université Paris-Sud, Orsay, 1984.
4. Bosák, J., "Directed graphs and matrix equations," *Math. Slovaca*, no. 2, pp. 189-202, 1978.
5. Bruijn, N.G. de, "A combinatorial problem," *Koninklijke Nederlandse Academie van Wetenschappen Proc.*, Ser A49, pp. 758-764, 1946.
6. Delorme, C., "Two notes on large graphs," *Rapport de Recherche n° 172*, LRI, Orsay, 1984.
7. Delorme, C. and Farhi, G., "Large graphs with given degree and diameter. I," *IEEE Trans. on Computers*, vol. C-33, pp. 857-860, 1984.
8. Farhi, G., "Diamètres dans les graphes et numérotations gracieuses," *Thesis*, LRI, Université Paris-Sud, Orsay, 1981.
9. Fiol, M.A., Yebra, J.L.A., and Alegre, I., "Line digraph iterations and the (d, k) digraph problem," *IEEE Trans. on Computers*, vol. C-33, pp. 400-403, 1984.
10. Fiol, M.A., Yebra, J.L.A., and Fabrega, J., "Sequence graphs and interconnection networks," *Ars Combinatoria*, vol. 16-A, pp. 7-14, 1983.
11. Gómez, J., Fiol, M.A., and Yebra, J.L.A., "Graphs on alphabets as models for large computer networks," Submitted to *J. of Parallel and Distributed Computing*.

12. Gomez, J. and Fiol, M.A., "Una nueva familia de grafos con elevado numero de vertices," *Research report, Polytechnic University of Catalonia*, 1984.
13. Harary, F., *Graph Theory*, Addison-Wesley, 1969.
14. Heuchenne, C., "Sur une certaine correspondance entre graphes," *Bull. Soc. Roy. Sc. Liège*, no. 33, pp. 743-753, 1964.
15. Homobono, N., "Résistance aux pannes de grands réseaux d'interconnexion," *Thesis, Université de Paris Sud*, 1987.
16. Imase, M. and Itoh, M., "Design to minimize diameter on building block network," *IEEE Trans. on Computers*, vol. C-30, pp. 439-442, 1981.
17. Imase, M. and Itoh, M., "A design for directed graphs with minimum diameter," *IEEE Trans. on Computers*, vol. C-32, pp. 782-784, 1983.
18. Kautz, W.H., "Bounds on directed (d,k) graphs," in *Theory of cellular logic networks and machines*, pp. 20-28, AFCRL-68-0668, SRI Project 7258, Final report, 1968.
19. Lam, C.W. and Lint, J.H. van, "Directed graphs with unique paths of fixed length," *J. of Comb. Theory, Ser. B*, no. 24, pp. 331-337, 1978.
20. McAndrew, M.H., "On the product of Directed Graphs," *Proc. Amer. Math. Soc.*, vol. 14, pp. 600-606, 1963.
21. Reddy, S.M., Pradhan, D.K., and Kuhl, J., "Directed graphs with minimal diameter and maximal connectivity," *School of Engineering Oakland Univ., Tech. Rep.* 1980.

Sur le calcul des chaînes d'additions optimales

S. Brlek¹, R. Mallette²

Résumé

Le problème de l'évaluation efficace des puissances d'un élément dans un semi-groupe est lié au problème du calcul de chaînes d'additions optimales. Bien qu'il existe des algorithmes asymptotiquement optimaux, le problème subsiste pour les exposants de taille raisonnable ($\log_2(n) < 1000$). Récemment, des algorithmes sous-optimaux ont été proposés pour les petites valeurs des exposants et pour les comparer avec les chaînes optimales, nous avons produit une table des chaînes optimales pour $n \leq 1024$. Le problème du calcul des chaînes optimales étant vraisemblablement NP-complet, le calcul utilise la programmation dynamique incluant une méthode de "backtracking" limitée par "branch and bound". Nous présenterons l'algorithme de calcul des chaînes optimales, quelques statistiques ainsi qu'un résultat partiel sur leur énumération.

Abstract

The problem of computing x^n using a minimal number of semigroup operations is equivalent to the problem of computing an optimal addition chain for n . Although there exist many asymptotically optimal algorithms in the literature, the problem remains open for exponents of small size ($\log_2(n) < 1000$). Recently, several sub-optimal algorithms have been proposed for small exponent values, and, in order to compare them with optimal chains, we produced a table of all optimal chains for $n \leq 1024$. The problem of computing optimal addition chains seems to be NP-complete. Their calculation uses dynamic programming including a backtracking method with branch and bound techniques. We will present this algorithm, some statistics and a partial result on their enumeration as well.

1. INTRODUCTION

Nous suivons la terminologie de A. Scholz [11] et définissons les chaînes d'additions pour un nombre entier positif n en adoptant la notation de Knuth [10].

Définition 1. Une chaîne d'additions pour un nombre entier positif n est une suite $C = (n_0, n_1, \dots, n_s)$ d'entiers positifs tels que

- I) $n_0 = 1$ et $n_s = n$,
- II) $\forall i, 1 \leq i \leq s, \exists k, j < s$ tels que $n_i = n_j + n_k$.

Il est commode d'écrire la chaîne C de manière croissante, de sorte que la condition II) s'écrive

- II') $\forall i, 1 \leq i \leq s, \exists k, j < i$ tels que $n_i = n_j + n_k$.

Définition 2. Soit C une chaîne d'addition telle qu'il existe un $i > 0, n_i \in C$ tel que $n_i = 2n_{i-1}$ (i.e. lorsque $j = k = i - 1$). On dira alors que le pas i est un pas double.

¹ LACIM, Dépt. de Math. et Info., Univ. du Québec à Montréal, C.P. 8888, Succ. A, Montréal H3C 3P8. Recherche subventionnée par le CRSNG-Canada et FCAR-Québec, e-mail: brlek@lacim.uqam.ca

² Dépt. de Math. et Info., Univ. du Québec à Montréal, C.P. 8888, Succ. A, Montréal H3C 3P8. e-mail: richard@lacim.uqam.ca

Exemple 1. La chaîne $C_9 = (1, 2, 4, 8, 9)$ possède trois pas doubles: $\{1, 2, 3\}$. En effet, le passage de $n_0 = 1$ à $n_1 = 2$ est toujours un pas double. D'autre part, les passages de 2 à 4 et de 4 à 8 sont aussi des pas doubles.

La longueur de la chaîne C est l'entier s , et est dénotée par $|C|$. Soit N les entiers naturels positifs, la fonction $\ell : N \rightarrow N$ détermine la longueur minimale de toutes les chaînes possibles. Pour l'entier n , cette longueur de chaîne C est dénoté par $\ell(n)$. Pour un n donné, les chaînes d'additions génèrent des schémas de multiplication pour x^n . Par exemple, la chaîne $(1, 2, 4, 8, 9, 18, 26, 27, 53)$ nous donne, pour x^{53} , le schéma de multiplication suivant:

$$xx = x^2, x^2x^2 = x^4, x^4x^4 = x^8, xx^8 = x^9, x^9x^9 = x^{18}, x^8x^{18} = x^{26}, xx^{26} = x^{27}, x^{26}x^{27} = x^{53}.$$

Cette correspondance entre les chaînes d'additions et les schémas de multiplication est naturelle. De plus, le nombre de multiplications nécessaires dans un schéma est égal à la longueur de la chaîne d'addition correspondante à ce schéma. Par suite, la plus petite longueur d'un tel schéma multiplicatif pour un entier n donné sera aussi dénotée par $\ell(n)$.

La notion de chaîne d'additions a été souvent étudiée par de nombreux auteurs, et nous incitons le lecteur à consulter Knuth [10], qui donne un compte rendu historique des chaînes d'additions dans son second volume. Plus récemment, Bos et Coster [6] ont étudié les chaînes d'additions dans un contexte un peu plus restreint. En effet, ils sont concernés par les applications en cryptographie, où le calcul d'une puissance intervient dans le célèbre algorithme de Rivest, Shamir et Adleman (RSA): les puissances qu'on y considère sont des nombres tels que $\lambda(n) \leq 512$.

Il existe plusieurs algorithmes pour calculer des chaînes pour un entier n , qui ne sont pas nécessairement optimales. Une chaîne $C = (n_0, n_1, \dots, n_s)$ est clairement optimale lorsque $s = |C| = \ell(n)$. Tout algorithme pouvant générer des chaînes d'additions fixe une borne supérieure pour $\ell(n)$. Ainsi, la méthode binaire décrite dans le volume de Knuth [10] donne la borne supérieure $\ell(n) \leq \lambda(n) + \nu(n) - 1$, où $\lambda(n) = \lfloor \log_2(n) \rfloor$ et $\nu(n)$ est le nombre de 1 dans le développement binaire de n . Dans [2], les auteurs définissent la classe des *cf-chaînes*, qui sont très souvent optimales. Pour suivre la terminologie introduite dans [7], nous désignons ces chaînes sous le vocable de chaînes *euclidiennes*. Leur intérêt repose sur les critères suivants:

- elles sont faciles à calculer
- elles sont souvent optimales
- l'algorithme s'exprime de manière fonctionnelle
- l'algorithme se généralise à des ensembles d'entiers [5].

Cependant, pour un ensemble d'entiers, le problème du calcul des chaînes optimales est NP-complet [9]. Par contre, le même problème pour un entier fixé est encore ouvert. Ainsi, si on doit concevoir un algorithme ne produisant que des chaînes optimales, il faut faire un compromis en temps et utiliser des techniques qui permettent de couper le plus possible l'exploration systématique.

2. PARCOURS DES CHAÎNES

Nous pouvons représenter simultanément toutes les chaînes d'additions par une arborescence A dont la racine $r(A)$ est l'élément générateur 1. La définition de chaîne d'additions se traduit naturellement en une définition d'arborescence.

Définition 3. L'arborescence des chaînes d'additions est définie récursivement par

- i) $r(A) = 1$,
- ii) $\forall i \neq 1, \exists k, j \in Anc(i)$ tels que $v(i) = v(j) + v(k)$,
ou
 $\forall i \neq 1, \exists j \in Anc(i)$ tel que $v(i) = 2v(j)$.

Exemple 2. Arborescence générant toutes les chaînes pour les entiers de 1 à 8. Pour l'entier 8, il existe 19 chaînes possibles dont les longueurs varient de 3 à 7. Nous avons $\ell(8) = 3$ et il n'y a qu'une seule chaîne ayant cette longueur, de plus, elle est optimale. Le lecteur pourra se convaincre que tout nombre de la forme

2^j possède une seule chaîne optimale, à savoir $(1, 2, 4, 8, \dots, 2^j)$. En ce qui concerne l'entier 7, il y a 10 chaînes possibles dont les longueurs varient de 4 à 6. Nous avons $\ell(7) = 4$ et il y a 5 chaînes réalisant cette longueur, qui sont, en outre, optimales.

Un algorithme parcourant une telle arborescence pour un intervalle d'entiers, requiert un temps énorme pour en extraire les chaînes optimales. Ce problème impose donc de trouver des bornes pour réduire le parcours. En examinant l'arbre, on peut établir une correspondance entre les niveaux de l'arbre et la position des éléments dans la chaîne. Si un entier n est rencontré pour la première fois au niveau s , alors $\ell(n) = s$ et la chaîne $C = (n_0, n_1, \dots, n_s)$ est une chaîne optimale. Prenons un tableau de cases représentant le schéma dont la longueur est s :

1	2	n_2	...	n_s
---	---	-------	-----	-------

Nous voulons que toutes les chaînes de longueur s soient optimales. Il n'est évidemment pas nécessaire de commencer par l'entier 1 dans la case représentant n_s , ni par 2.

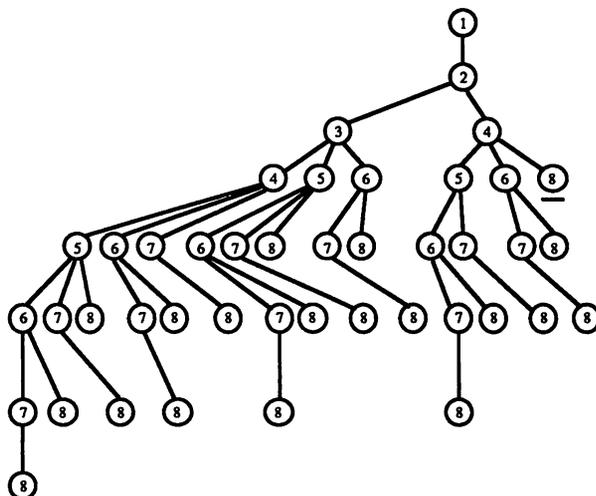


Fig. 1. Arborecence de toutes les chaînes de 1 à 8.

Bornes inférieure et supérieure

Le problème que nous considérons d'abord, est de trouver une borne inférieure pour la longueur des chaînes qui permettra de couper la recherche sans perdre de chaîne optimale. Soit $\ell^{-1}(s) = \{n \in N | \ell(n) = s\}$, alors le plus petit entier non encore marqué par l'algorithme à un niveau inférieur à s est

$$c(s) = \min(\ell^{-1}(s)).$$

Prenons comme exemple $s = 4$, les nombres 1,2,3,4,5,6 sont rencontrés à des niveaux inférieurs alors 7 est le plus petit entier i pouvant avoir $\ell(i) = 4$. Et, par le fait même, pour les chaînes optimales de longueur 4 l'algorithme commencera à vérifier à partir de l'entier 7.

- *Calcul de $c(s)$.* Pendant le parcours, l'algorithme note dans un tableau l'entier rencontré; par la suite, en changeant de niveau l'algorithme lit le tableau de façon séquentielle afin d'y trouver le premier entier qui n'est pas marqué, ce sera le plus petit entier qui peut-être rencontré précisément à ce niveau.

En utilisant les pas doubles uniquement, on obtient la chaîne $(1, 2, 4, 8, 16, \dots, 2^s)$ qui est maximale.

Nous savons donc que les chaînes optimales de longueur s se situent, pour des entiers n entre ces bornes:

$$c(s) \leq n \leq 2^s.$$

En observant la figure 1 on remarque que $3 = \ell(8) \leq \ell(7) = 4$. Ceci indique que certains entiers entre les deux bornes n'ont pas de chaînes optimales de longueur s . Cependant ces bornes permettent la réduction du parcours.

Avec ces bornes, l'algorithme part du niveau s et remonte dans l'arbre vers la racine pour retrouver les éléments de la chaîne. Pour ce faire, nous utilisons un parcours en largeur de l'arbre, en éliminant tout entier à l'extérieur des bornes. On élimine ensuite les entiers marqués dans la table (ceux qui ont une chaîne minimale de longueur inférieure à s .)

À ce stade ci, il ne reste plus qu'à explorer les entiers entre les deux bornes ayant une chaîne optimale de longueur égale ou supérieure à s .

Condition de chaîne

Chaque branche de l'arborescence est une suite croissante d'entiers positifs vérifiant la condition

$$n_i \leq n_j \text{ pour } i \leq j.$$

L'émondage s'effectue ensuite en utilisant la condition II) $n_i = n_j + n_k$ caractérisant les chaînes d'additions. En effet nous savons que n_s doit satisfaire cette condition. Puisque les cases du tableau correspondent aux niveaux de l'arbre, un élément dans la case i représente le i -ème élément de la chaîne au niveau i de l'arbre. Il suffit alors de trouver deux éléments d'indices $j < s$ et $k < s$ tels que $n_s = n_j + n_k$ dans ce tableau.

3. OPTIMISATION DU PARCOURS

Réduisons le "backtracking" pour les chaînes n'ayant que des éléments de la forme $n_i = n_j + n_k$ dans la table. Prenons un tableau de départ, où les d_i indiquent les entiers de départ pour un parcours quelconque de niveau s

1	2	d_2	...	d_s
---	---	-------	-----	-------

Lorsque nous trouvons un n entre les bornes au niveau s , il n'est pas nécessaire de vérifier pour tous les entiers les cases correspondantes aux niveaux inférieurs. Puisque $n_{s-1} < n_s \leq 2(n_{s-1})$, il suffit de vérifier pour le niveau $s-1$ que l'entier n_{s-1} satisfait cette inégalité. En divisant l'entier n_s par 2, nous aurons une borne inférieure pour l'entier n_{s-1} . Prenons $d_{s-1} = \lfloor n_s/2 \rfloor$ comme borne inférieure. De façon récursive nous trouvons pour $i < s$ que $d_{i-1} = \lfloor d_s/2 \rfloor$. Il ne faut pas oublier que le plus petit entier possible dans une case est l'indice de la case lui même, car la chaîne d'addition la plus longue est celle où on augmente de 1 l'élément suivant (i.e. $n_i = n_{i-1} + 1$). Prenons $d_{i-1} = \max(i-1, \lfloor d_i/2 \rfloor)$. Ceci nous donne une branche de départ dont les éléments sont les bornes inférieures pour les éléments des chaînes d'additions correspondantes aux entiers du niveau s .

Maintenant que nous avons un intervalle pour les entiers candidats au niveau s et que nous avons un tableau de départ donnant des bornes inférieures pour chaque élément de niveau inférieur, il suffit, pour exécuter le "backtracking", de simuler un compteur. Ce compteur commence par l'élément d'indice le plus élevé et l'augmente de 1. Par la suite il recherche la condition II) avec les éléments d'indice inférieur. Aussitôt que l'algorithme rencontre une somme de deux éléments qui est inférieure à la valeur de l'élément vérifié ($n_i > n_j + n_k$), il cesse le "backtracking" dans cette branche et augmente de 1 la valeur de l'élément du niveau vérifié. Cette action se fait de façon récursive dans la branche que l'algorithme remonte.

4. COMPLEXITÉ DE L'ALGORITHME

Si nous considérons la fouille comme un compteur, nous avons pour le niveau s , un compteur à $s+1$ cases. Puisque chaque case contient un entier satisfaisant la condition II) des chaînes d'additions, un compteur de i à $2i-1$ pour la case i rencontrera sûrement tous les entiers candidats pour les chaînes d'additions. Le compteur augmente l'élément i de une unité et, lorsqu'il arrive à $2i-1$, remet la case i à la valeur i et

augmente la valeur de la case précédente $i - 1$. De façon récursive nous passons de la case s vers la case 1. Le compteur atteint son maximum lorsque toutes les cases atteignent leur valeurs maximales. Ce compteur passe beaucoup plus de suites d'entiers qu'il n'y a de chaînes optimales ce qui en fait une borne supérieure. Pour trouver la borne supérieure il suffit de compter les façons de placer les entiers dans ces cases. Pour une longueur optimale s une borne possible est :

$$O(\text{chaînes optimales}) \leq \prod_{i=0}^s (2^i - i)$$

Si nous utilisons les bornes inférieures et supérieures calculées pendant le parcours de l'algorithme nous obtenons la borne suivante:

$$O(\text{chaînes optimales}) \leq \prod_{i=0}^s (2^i - \max(i, \frac{c(s)}{2^{s-i}}))$$

Cependant cette dernière borne n'est pas calculable avant le parcours de l'algorithme, car il n'existe pas de formule close pour $c(s)$. On sait que $c(s)$ est asymptotiquement exponentielle ($\log_2(c(s)) \sim s$), mais le fait que la fonction $c(r)$ soit monotone croissante reste une conjecture.

5. STATISTIQUES

Le tableau suivant indique pour chaque longueur minimale r possible de 1 à 13, quel est l'entier minimal $c(r)$ qui réalise cette longueur. La fonction $d(r)$ donne le nombre d'entiers entre 1 et 1024 qui réalisent un chaîne optimale de longueur r :

$$d(r) = \text{Card} \{n \in [1, 1024] : \ell(n) = r\}.$$

Total donne le nombre total de chaînes optimales de longueur r .

r	$c(r)$	$d(r)$	Total
1	2	1	1
2	3	2	2
3	5	3	5
4	7	5	16
5	11	9	63
6	19	15	216
7	29	26	814
8	47	44	2,939
9	71	78	11,824
10	127	136	43,065
11	191	235	154,512
12	379	332	440,779
13	607	137	831,747

Table 1. Nombre total de chaînes optimales : 1,485,983.

On trouvera en appendice un tableau pour les entiers de 241 à 540. Dans ce tableau sont compilées les informations suivantes:

1. l'entier
2. la longueur optimale pour cet entier
3. le nombre de chaînes optimales rencontrées pour cet entier
4. une chaîne optimale pour cet entier (c'est la première chaîne optimale rencontrée pour l'entier en question dans l'ordre lexicographique).

6. ÉNUMÉRATION DES CHAÎNES OPTIMALES

Un examen attentif des données présentées en appendice montre certaines régularités dans l'énumération des chaînes optimales. Nous établissons dans ce qui suit un pas dans cette direction pour les valeurs de n telles que $\nu(n) = 2$, c'est à dire lorsque $n = 2^A + 2^B$ où $A > B \geq 0$.

Rappelons que $\nu(n)$ est le nombre de 1 dans la représentation binaire de n . Nous désignons par F_k , le k -ème nombre de Fibonacci. Nous rappelons d'abord deux résultats classiques qu'on trouve dans Knuth[10].

Lemme 1. (Knuth[10, Th. A, p.448]) *Si la chaîne $C = (c_0, c_1, \dots, c_s)$ comprend d pas doubles et $f = s - d$ pas non doubles, alors*

$$n \leq 2^{d-1} F_{f+3}.$$

Lemme 2. (Knuth[10, p.448]) $n = 2^A + 2^B \Rightarrow \ell(n) = A + 1$.

Nous avons besoin du lemme technique suivant.

Lemme 3. *Soit $n = 2^A + 2^B$ avec $A > B > 0$. Une chaîne optimale pour n possède $f \leq 2$ pas non doubles. De plus, si $A = B + 1$, alors $f = 1$.*

Preuve. Nous avons, par le lemme 1 que $n \leq 2^{d-1} F_{f+3}$, et, par le lemme 2, que $s = \ell(n) = \ell(2^A + 2^B) = A + 1$. Cela entraîne que $2^A + 2^B = n \leq 2^{A-f} F_{f+3}$, puis

$$1 < \frac{2^A + 2^B}{2^A} \leq \frac{F_{f+3}}{2^f},$$

d'où l'on conclue que $F_{f+3} \geq 2^f$. On vérifie aisément que $f \leq 2$.

Soit $d = A - 1$ et $f = 2$ (i.e. $A - 1$ pas doubles et 2 pas non doubles). Alors nous avons l'inégalité suivante:

$$n \leq 2^{A-2} F_5 = 2^{A-2} 5 = 2^{A-2} (2^2 + 1) = 2^A + 2^{A-2}$$

d'où on déduit que $B \leq A - 2$.

Pour $B = A - 1$ on vérifie l'inégalité

$$n \leq 2^{A-1} F_4 = 2^{A-1} 3 = 2^{A-1} (2 + 1) = 2^A + 2^{A-1} = 2^A + 2^B$$

qui montre qu'il n'y a qu'un pas non double. ■

Proposition 1. *Soit $n = 2^A + 2^B$ où $A = B + 1$ et $B > 0$, alors il existe A chaînes d'additions distinctes de longueur optimale pour n .*

Preuve. Puisque $A = B + 1$, nous avons $n = 2^{B+1} + 2^B = 2^B (2 + 1) = 2^B \cdot 3$ et, nous avons aussi $\ell(2^A + 2^B) = A + 1$. Ainsi il y a exactement $A + 1$ étapes dans la chaîne $(c_0 = 1, c_1 = 2, \dots, c_{A+1} = 2^B \cdot 3)$.

Par le lemme 3, il y a A étapes à pas double et une seule étape à pas non double. Puisque l'étape de $c_0 = 1$ à $c_1 = 2$ est toujours un pas double, il reste A étapes où l'on peut placer le pas non double. De plus, la position du terme non double détermine pour $n = 2^B \cdot 3$ une chaîne unique. Aussi, $2^B \cdot 3$ s'écrit sous la forme $(2^{j+1} + 2^j) 2^i$ où $j + i = B$, indiquant que le pas non double peut se placer à toutes les étapes entre c_1 et c_{A+1} . Finalement il est clair que deux positions distinctes donnent deux chaînes distinctes. Une chaîne d'addition optimale pour $n = 2^B \cdot 3$ a donc la forme:

$$(1, 2, \dots, 2^j, 2^{j+1}, 2^j + 2^{j+1}, 2(2^j + 2^{j+1}), \dots, 2^i(2^j + 2^{j+1}) = c_{A+1}).$$

Puisqu'il y a A choix possibles ceci nous donne bien A chaînes distinctes. ■

Proposition 2. Soit $n = 2^A + 2^B$ où $A \neq B + 1$, $A \neq B + 3$ et $A > B \geq 0$ alors
 i) si $B = 0$ il y a exactement 2 chaînes d'additions distinctes de longueur optimale.
 ii) si $B \neq 0$ il y a $2(B + 1)$ chaînes d'additions distinctes de longueur optimale.

Preuve. Cas i). Puisque $B = 0$ nous avons $n = 2^A + 1$ avec $\ell(2^A + 1) = A + 1$. Mais par hypothèse $A \neq B + 1$ donc $A > 1$. Regardons d'abord les chaînes n'ayant qu'un pas non double. Puisque $n = 2^A + 1$ est impair, le pas non double apparaît nécessairement à la fin et la seule chaîne possible est $(1, 2, \dots, 2^A, 2^A + 1)$.

Maintenant, voyons pour les chaînes ayant deux pas non double. Puisque n est impair la dernière étape est un pas non double. Donc $c_{A+1} = c_j + c_k$ où $j \neq k$ et $j, k < A + 1$. et il reste à déterminer j et k . Or, la seule combinaison qui nous amène à ce résultat, avec seulement deux pas non doubles, est $2^A + 1 = 2^{A-1} + (2^{A-1} + 1)$ alors $c_j = 2^{A-1} = c_{A-1}$, et $c_k = 2^{A-1} + 1 = c_A$. Voyons pourquoi: supposons que le pas non double se situe de c_{A-2} à c_{A-1} alors nous aurions une chaîne de la forme :

$$(1, 2, \dots, 2^{A-2}, 2^{A-2} + 1, 2 \cdot (2^{A-2} + 1), (2^{A-2} + 1) + (2 \cdot (2^{A-2} + 1))).$$

Il s'en suivrait que $c_{A+1} = (2^{A-2} + 1) + (2 \cdot (2^{A-2} + 1)) = 2^{A-2} + 2^{A-1} + 2 + 1 = 2^A + 1$, entraînant $A + 3 = B + 3$, ce qui est exclu par hypothèse.

Enfin, si le pas non double se situe avant l'étape c_{A-2} à c_{A-1} . On montre que l'on ne peut atteindre $2^A + 1$. Soit $B < A'$, on considère une chaîne de la forme

$$C(n) = \begin{cases} (1, 2, \dots, 2^{A'}, 2^{A'} + 2^B, 2 \cdot (2^{A'} + 2^B), \dots, 2^C \cdot (2^{A'} + 2^B), 2^C \cdot (2^{A'} + 2^B) + 2^{A'} + 2^B), & \text{si } B \neq 0; \\ (1, 2, \dots, 2^{A'}, 2^{A'} + 1, 2 \cdot (2^{A'} + 1), \dots, 2^C \cdot (2^{A'} + 1), 2^C \cdot (2^{A'} + 1) + 1), & \text{si } B = 0. \end{cases}$$

Dans les deux cas on doit satisfaire l'équation sur les longueurs $A' + C + 1 = A$. Si $B \neq 0$, on doit satisfaire l'équation $2^C \cdot (2^{A'} + 2^B) + 2^{A'} + 2^B = 2^A + 1$. En substituant $A' = A - C - 1$ dans cette dernière, on obtient $2^{B-A'} = 1$ contredisant ainsi l'hypothèse $B < A'$. Le cas $B = 0$ est similaire.

Il n'y a donc que deux chaînes au total.

Cas ii). Puisque $A > B$, posons $j = A - B$. Alors $n = 2^A + 2^B = 2^{B+j} + 2^B = 2^B(2^j + 2^0)$. Ainsi, les pas non doubles ne pouvant être dans le calcul de 2^B apparaissent dans le calcul de $(2^j + 2^0)$. Mais le cas I nous indique qu'il y a deux chaînes possibles pour $(2^j + 2^0)$. D'autre part, comme on peut écrire n sous la forme

$$n = 2^B(2^j + 2^0) = 2^s(2^j + 2^0)2^t \quad \text{avec } s + t = B,$$

nous obtenons des chaînes de la forme:

$$(1, 2, \dots, 2^s, \dots, 2^s(2^{j-1}), 2^s(2^j), 2^s(2^j + 2^0), 2^s(2^j + 2^0)2, \dots, 2^s(2^j + 2^0)2^t = c_{A+1}),$$

ou bien

$$(1, 2, \dots, 2^s, \dots, 2^s(2^{j-1}), 2^s(2^{j-1} + 1), 2^s(2^j + 2^0), 2^s(2^j + 2^0)2, \dots, 2^s(2^j + 2^0)2^t = c_{A+1}).$$

Il suffit alors de compter le nombre de positions possibles pour le calcul de $(2^j + 2^0)$. Puisque $s + t = B$, alors $0 \leq s \leq B$ donnant ainsi $B + 1$ possibilités de positions différentes pour le calcul de $(2^j + 2^0)$. Le compte total nous donne $2(B + 1)$. ■

Proposition 3. Soit $n = 2^A + 2^B$ où $A = B + 3$ et $A > B \geq 0$ alors il y a $\binom{A}{2} + B$ chaînes de longueur optimale.

Preuve. Il suffit de compter les positions où peuvent se placer les pas non doubles. Puisque $A = B + 3$ nous avons $n = 2^{B+3} + 2^B = 2^B(2^3 + 2^0)$.

Regardons d'abord le cas des chaînes optimales n'ayant qu'un seul pas non double. Elles ont alors la forme suivante:

$$(1, 2, \dots, 2^i, 2^i + 2^j, 2(2^i + 2^j), \dots, 2^k(2^i + 2^j)).$$

Il suffit de regarder les solutions pour $2^k(2^i + 2^j) = 2^B(2^3 + 2^0)$ pour compter le nombre de chaînes optimales. Puisque $j < i$ nous avons que $2^k(2^i + 2^j) = 2^{k+j}(2^{i-j} + 2^0) = 2^B(2^3 + 2^0)$. Ainsi, les solutions vérifient les

équations $k + j = B$ et $i - j = 3$. Ceci entraîne que $i \geq 3$. Puisqu'il y a $B + 3$ positions où l'on peut insérer un pas non double et que, parmi celles-ci, deux ($i = 1$ ou 2) ne sont pas acceptables, alors il reste donc $B + 1$ positions donnant $B + 1$ chaînes optimales.

Considérons maintenant celles qui ont deux pas non doubles. La forme que prennent ces chaînes est:

$$(1, 2, \dots, 2^i, 2^i + 2^j, \dots, 2^k(2^i + 2^j), 2^k(2^i + 2^j) + 2^l(2^i + 2^j), \dots, 2^s(2^k(2^i + 2^j) + 2^l(2^i + 2^j))).$$

Il suffit de regarder les solutions pour $2^s(2^k(2^i + 2^j) + 2^l(2^i + 2^j)) = 2^B(2^3 + 2^0)$ pour compter le nombre de chaînes optimales. Puisque $j < i$ nous avons que

$$2^s(2^k(2^i + 2^j) + 2^l(2^i + 2^j)) = 2^{s+j}((2^k + 2^l)(2^{i-j} + 2^0)) = 2^B(2^3 + 2^0).$$

Ainsi les solutions vérifient les équations $2^{s+j}(2^k + 2^l) = 2^B$ et $i - j = 3$. Il s'en suit que $i \geq 3$. Il suffit maintenant de compter le nombre de façons de placer les 2 pas non doubles. Puisqu'il y a $B + 3$ positions possibles, ceci équivaut à choisir 2 positions parmi $B + 3$. Enlevons celles qui ne donnent pas de solution aux équations. Il n'y en a qu'une car pour $i \leq 2$ on ne peut avoir que $i = 2$ et $j = 1$, ce qui donne $\binom{B+3}{2} - 1$ chaînes optimales distinctes.

On a donc un total de $\binom{B+3}{2} - 1 + B + 1 = \binom{B+3}{2} + B$ chaînes optimales distinctes. ■

Théorème 4. Soit $n = 2^A + 2^B$ avec $A > B \geq 0$. Alors le nombre de chaînes d'additions optimales est:

- i) A si $A = B + 1$ et $B > 0$;
- ii) 2 si $B = 0$ et $A \neq B + 3$;
- iii) $2(B + 1)$ si $B \neq 0$ et $A \neq B + 3$;
- iv) $\binom{A}{2} + B$ si $A = B + 3$.

Références

- [1] A. Brauer, On Addition Chains, *Bull. Amer. Math. Soc.* 45, 1939, 736–739.
- [2] F. Bergeron, J. Berstel, S. Brlek, A Unifying Approach to the Generation of Addition Chains, *Proc. XV Latin American Conf. on Informatics, Santiago, Chile, 10-14 July 1989*, 29-38.
- [3] F. Bergeron, J. Berstel, S. Brlek, Efficient Computation of Addition Chains, *Submitted*.
- [4] F. Bergeron, J. Berstel, S. Brlek, C. Duboc, Addition Chains Using Continued Fractions, *J. Algorithms* 10, 1989, 4 03–412.
- [5] F. Bergeron, J. Olivos, Vectorial Addition Chains Using Euclid's Algorithm, *Submitted*.
- [6] J. Bos, M. Coster, Addition Chain Heuristics, *Proceedings of CRYPTO 89*.
- [7] S. Brlek, P. Castéran, R. Strandh, Chaînes d'additions et structures de contrôle, *Journées JFLA 91, Gresse-en Vercors, France, January 28-29, 1991*.
- [8] S. Brlek, P. Castéran, R. Strandh, On Addition Schemes, *Proceedings of TAPSOFT91, April 8–12 1991, Brighton, England, Lecture Notes in Computer Science*, 494, 1991, 379–393.
- [9] P. Downey, B. Leong, R. Sethi, Computing Sequences with Addition Chains, *SIAM J. Computing* 3, 1974, 1–10.
- [10] D.E. Knuth, *The Art of Computer Programming*, vol. 2, Addison-Wesley, 1981.
- [11] A. Scholz, Jahresbericht, *Deutsche Math.-Verein.* 47, 1937, 41.
- [12] A. Schönhage, A Lower Bound for the Length of Addition Chains), *Theoretical Comp. Sci.* 1, 1975, 1–12.
- [13] E.G. Thurber, On Addition Chains $\ell(mn) \leq \ell(n) - b$ and Lower Bounds for $c(r)$, *Duke Math. J.* 40, 1973, 907–013.
- [14] A.C.-C. Yao, On the Evaluation of Powers, *SIAM J. Comp.* 9, 1976, 100–103.

APPENDICES

A. Programme qui génère les chaînes en C.

```

#include <stdio.h>
#define n=14          /* longueur maximale des chaines*/
#define limite=1024  /* borne supérieure des entiers*/
int indice[n],depart[n],quant[n],minim[n],tableau[n],
    precedant[n], chaines[limite],trouve[limite],maxi,
    longueur,boucle,ancien;
int Augmente (longs)
int longs;
{int boucle1,boucle2,long1,test,temp,ok;
long1=longs-1;
ok=1;
  if (longs>2)
    { while (ok == 1)
      { tableau[longs] += 1;
        if ((tableau[longs]>2*tableau[long1])
            || (tableau[longs] > indice[longs]))
          {Augmente (long1);
            if (depart[longs]>tableau[long1]+1)
              || tableau[longs]=depart[longs]
              else tableau[longs]=tableau[long1]+1;}
          if (longueur==longs)
            {if ((chaines[tableau[longs]]==0)
                || (chaines[tableau[longs]]==long1))
              boucle1=long1
            else boucle1=0;}
          else boucle1=long1;
          while (boucle1 > 0 )
            { boucle2=boucle1;
              while (boucle2>0)
                { temp=tableau[boucle2]+tableau[boucle1];
                  if (tableau[longs]==temp)
                    { ok=0;
                      boucle1=1; boucle2=1;
                      if ((longueur==longs) &&
                          (chaines[tableau[longs]]==0))
                        chaines[tableau[longs]]=long1;}
                  test=boucle2;
                  if (tableau[longs]>temp) boucle2=0
                  else boucle2 - = 1;}
                  if (boucle1<=test) boucle1=0
                  else boucle1 - = 1;}}
            else {longueur +=1;
              for (boucle1=2;boucle1<=maxi;boucle1++)
                { if (trouve[boucle1]==0)
                  { minim[longueur-1]=boucle1;
                    boucle1=maxi+1;}}
              tableau[longueur]=minim[longueur-1];
              depart[longueur]=tableau[longueur];
              for (boucle1=longueur-1;boucle1>=3;boucle1--)
                { depart[boucle1]=(depart[boucle1+1]+3)/2;

```

```

        if (depart[boucle1]<boucle1)
            depart[boucle1]=boucle1;}
    tableau[longueur] - =1;}
return ok;});

int Imprime(s)
int s;
{int boucle2,difference;
  difference=0;
  for (boucle2=1; boucle2<=longueur; boucle2++)
    { if (tableau[boucle2] != precedant[boucle2])
      { precedant[boucle2]=tableau[boucle2];
        if (difference == 0) difference = boucle2;}}
  difference=longueur-difference+1;
  printf ("%u",difference);
  for (boucle2=1; boucle2<difference;boucle2++)
    printf ( " %u",tableau[longueur-difference+boucle2]);
  printf (" %u\ n",tableau[longueur]);});

main ()
{for (boucle=1;boucle<=n;boucle++)
  {tableau[boucle]=boucle;
  precedant[boucle]=0;
  minim[boucle]=0;
  quant[boucle]=0;}
maxi=1024;
minim[1]=2;
indice[1]=1;
indice[2]=2;
for (boucle=3;boucle<=n;boucle++)
  indice[boucle]=indice[boucle-1] *2;
for (boucle=1;boucle<=limite;boucle++) trouve[boucle]=0;
for (boucle=1;boucle<=limite;boucle++) chaines[boucle]=0;
trouve[1]=1; trouve[2]=1;
ancien=2;
longueur=2;
Imprime (2);
for (boucle=12;boucle<=n;boucle++)
  indice[boucle]=1024+boucle-11;
while (longueur <= n)
  { Augmente (longueur);
  if (chaines[tableau[longueur]]==longueur-1)
    {if (longueur<=n) { trouve[tableau[longueur]] + =1;
      if (tableau[longueur]<=maxi)
        Imprime (boucle);}}}
return 0;});

```

n	$\ell(n)$	Qte	exemplaire
241	10	165	(1,2,3,5,10,15,30,60,120,121,241)
242	10	329	(1,2,3,5,10,15,30,60,61,121,242)
243	10	93	(1,2,3,5,10,15,30,60,120,123,243)
244	10	457	(1,2,3,5,7,14,28,33,61,122,244)
245	10	48	(1,2,3,5,10,15,30,60,120,125,245)
246	10	194	(1,2,3,5,10,15,30,60,63,123,246)
247	11	2726	(1,2,3,4,7,8,15,30,60,120,127,247)
248	10	609	(1,2,3,4,7,14,17,31,62,124,248)
249	10	8	(1,2,3,5,10,20,40,43,83,166,249)
250	10	168	(1,2,3,5,10,15,25,50,75,125,250)
251	11	1608	(1,2,3,4,7,11,15,30,60,120,131,251)
252	10	700	(1,2,3,4,7,14,21,42,63,126,252)
253	11	3506	(1,2,3,4,7,11,22,33,55,110,143,253)
254	11	9006	(1,2,3,4,7,8,15,30,60,67,127,254)
255	10	24	(1,2,3,5,10,15,30,60,120,135,255)
256	8	1	(1,2,4,8,16,32,64,128,256)
257	9	2	(1,2,4,8,16,32,64,128,129,257)
258	9	4	(1,2,4,8,16,32,64,65,129,258)
259	10	23	(1,2,3,4,8,16,32,64,128,131,259)
260	9	6	(1,2,4,8,16,32,33,65,130,260)
261	10	21	(1,2,3,5,8,16,32,64,128,133,261)
262	10	58	(1,2,3,4,8,16,32,64,67,131,262)
263	11	1059	(1,2,3,4,7,8,16,32,64,128,135,263)
264	9	8	(1,2,4,8,16,17,33,66,132,264)
265	10	21	(1,2,4,8,9,16,32,64,128,137,265)
266	10	50	(1,2,3,5,8,16,32,64,69,133,266)
267	11	1772	(1,2,3,4,7,11,22,33,66,132,135,267)
268	10	101	(1,2,3,4,8,16,32,35,67,134,268)
269	11	982	(1,2,3,4,8,16,32,35,67,134,135,269)
270	10	136	(1,2,3,5,10,15,30,45,90,135,270)
271	11	747	(1,2,3,4,7,11,22,33,66,132,139,271)
272	9	10	(1,2,4,8,9,17,34,68,136,272)
273	10	25	(1,2,4,8,9,17,34,68,136,137,273)
274	10	50	(1,2,4,8,9,16,32,64,73,137,274)
275	11	1451	(1,2,3,4,7,10,17,34,68,136,139,275)
276	10	135	(1,2,3,5,8,16,32,37,69,138,276)
277	11	741	(1,2,3,5,6,11,17,34,68,136,141,277)
278	11	2682	(1,2,3,4,7,10,17,34,68,71,139,278)
279	11	987	(1,2,3,4,7,10,17,34,68,136,143,279)
280	10	448	(1,2,3,4,7,14,21,35,70,140,280)
281	10	4	(1,2,4,8,9,17,34,68,136,145,281)
282	11	5600	(1,2,3,4,7,10,17,34,68,136,146,282)
283	11	493	(1,2,3,4,7,14,21,35,70,140,143,283)
284	11	7548	(1,2,3,4,7,8,16,32,39,71,142,284)
285	11	1200	(1,2,3,4,7,11,22,44,51,95,190,285)
286	11	2781	(1,2,3,4,7,10,17,34,68,75,143,286)
287	11	270	(1,2,3,4,7,14,21,35,70,140,147,287)
288	9	33	(1,2,3,6,9,18,36,72,144,288)
289	10	87	(1,2,3,6,9,18,36,72,144,145,289)
290	10	168	(1,2,3,6,9,18,36,72,73,145,290)
291	10	28	(1,2,3,6,9,18,36,72,144,147,291)
292	10	217	(1,2,3,6,9,18,36,37,73,146,292)
293	10	4	(1,2,4,5,9,18,36,72,144,149,293)
294	10	80	(1,2,3,6,9,18,36,72,75,147,294)
295	11	699	(1,2,3,4,7,9,18,36,72,144,151,295)
296	10	258	(1,2,3,5,8,16,21,37,74,148,296)
297	10	14	(1,2,3,6,9,18,36,72,144,153,297)
298	10	8	(1,2,4,5,9,18,36,72,77,149,298)
299	11	347	(1,2,3,4,7,11,18,36,72,144,155,299)
300	10	264	(1,2,3,5,10,15,25,50,75,150,300)

n	$\ell(n)$	Qte	exemplaire
301	11	1096	(1,2,3,4,7,14,21,35,70,140,161,301)
302	11	2416	(1,2,3,4,7,9,18,36,72,79,151,302)
303	11	748	(1,2,3,4,8,16,32,35,67,134,169,303)
304	10	307	(1,2,3,4,8,11,19,38,76,152,304)
305	11	1339	(1,2,3,4,8,11,19,38,76,152,153,305)
306	10	48	(1,2,3,6,9,18,36,72,81,153,306)
307	11	348	(1,2,3,4,8,11,19,38,76,152,155,307)
308	10	12	(1,2,4,5,9,18,36,41,77,154,308)
309	11	347	(1,2,3,5,7,12,19,38,76,152,157,309)
310	11	2156	(1,2,3,4,7,11,18,36,72,83,155,310)
311	11	61	(1,2,3,5,7,12,19,38,76,152,159,311)
312	10	248	(1,2,3,5,8,13,26,39,78,156,312)
313	11	911	(1,2,3,5,8,13,26,39,78,156,157,313)
314	11	2077	(1,2,3,5,7,12,19,38,76,81,157,314)
315	11	1167	(1,2,3,4,7,14,21,35,70,105,210,315)
316	11	3532	(1,2,3,4,7,9,18,36,43,79,158,316)
317	11	142	(1,2,3,5,8,13,26,39,78,156,161,317)
318	11	1976	(1,2,3,4,8,16,19,35,70,89,159,318)
319	11	52	(1,2,3,6,7,13,26,39,78,156,163,319)
320	9	14	(1,2,3,5,10,20,40,80,160,320)
321	10	33	(1,2,3,5,10,20,40,80,160,161,321)
322	10	66	(1,2,3,5,10,20,40,80,81,161,322)
323	10	4	(1,2,3,5,10,20,40,80,160,163,323)
324	10	194	(1,2,3,5,10,20,40,41,81,162,324)
325	10	8	(1,2,3,5,10,20,40,80,160,165,325)
326	10	8	(1,2,3,5,10,20,40,80,83,163,326)
327	11	381	(1,2,3,4,7,10,20,40,80,160,167,327)
328	10	124	(1,2,3,5,10,20,21,41,82,164,328)
329	11	561	(1,2,3,4,8,16,32,64,67,131,198,329)
330	10	24	(1,2,3,5,10,20,40,80,85,165,330)
331	11	131	(1,2,3,5,10,11,20,40,80,160,171,331)
332	10	12	(1,2,3,5,10,20,40,43,83,166,332)
333	11	418	(1,2,3,5,8,16,21,37,74,111,222,333)
334	11	503	(1,2,3,4,7,10,20,40,80,87,167,334)
335	11	106	(1,2,3,4,8,16,32,35,67,134,201,335)
336	10	369	(1,2,3,4,7,14,21,42,84,168,336)
337	11	1082	(1,2,3,4,7,14,21,42,84,168,169,337)
338	11	2919	(1,2,3,4,7,14,21,42,84,85,169,338)
339	11	430	(1,2,3,4,7,14,21,42,84,168,171,339)
340	10	48	(1,2,3,5,10,20,40,45,85,170,340)
341	11	304	(1,2,3,5,7,14,21,42,84,168,173,341)
342	11	2457	(1,2,3,4,7,14,21,42,84,87,171,342)
343	11	149	(1,2,3,4,7,14,21,42,84,168,175,343)
344	10	16	(1,2,3,5,10,20,23,43,86,172,344)
345	11	262	(1,2,3,4,7,14,28,56,59,115,230,345)
346	11	548	(1,2,3,5,7,14,21,42,84,89,173,346)
347	11	31	(1,2,3,5,10,11,21,42,84,168,179,347)
348	11	3539	(1,2,3,4,7,10,20,40,47,87,174,348)
349	11	53	(1,2,3,5,8,13,21,42,84,168,181,349)
350	11	1060	(1,2,3,4,7,14,21,35,70,105,175,350)
351	11	196	(1,2,3,5,7,14,28,56,61,117,234,351)
352	10	156	(1,2,3,4,7,11,22,44,88,176,352)
353	11	440	(1,2,3,4,7,11,22,44,88,176,177,353)
354	11	1198	(1,2,3,4,7,11,22,44,88,89,177,354)
355	11	62	(1,2,3,4,7,11,22,44,88,176,179,355)
356	11	2108	(1,2,3,4,7,11,22,44,45,89,178,356)
357	11	207	(1,2,3,4,7,14,21,42,84,168,189,357)
358	11	210	(1,2,3,4,7,11,22,44,88,91,179,358)
359	11	13	(1,2,3,4,7,11,22,44,88,176,183,359)
360	10	200	(1,2,3,5,10,15,30,45,90,180,360)

n	$\ell(n)$	Qte	exemplaire
361	11	654	(1,2,3,5,10,15,30,45,90,180,181,361)
362	11	1363	(1,2,3,5,6,11,22,44,88,93,181,362)
363	11	385	(1,2,3,4,7,11,22,44,88,176,187,363)
364	11	2714	(1,2,3,4,7,11,22,44,47,91,182,364)
365	11	241	(1,2,3,5,10,15,30,45,90,180,185,365)
366	11	872	(1,2,3,4,7,11,22,44,88,95,183,366)
367	11	4	(1,2,3,5,10,20,23,43,86,172,195,367)
368	10	20	(1,2,3,5,10,13,23,46,92,184,368)
369	11	247	(1,2,3,5,10,13,23,46,92,184,185,369)
370	11	830	(1,2,3,5,8,16,21,37,74,111,185,370)
371	11	10	(1,2,3,5,10,13,23,46,92,184,187,371)
372	11	2122	(1,2,3,4,7,14,17,31,62,93,186,372)
373	11	18	(1,2,3,5,10,13,23,46,92,184,189,373)
374	11	244	(1,2,3,4,7,11,22,44,88,99,187,374)
375	11	156	(1,2,3,5,10,15,25,50,75,125,250,375)
376	11	1683	(1,2,3,4,7,10,20,27,47,94,188,376)
377	11	8	(1,2,4,5,9,14,23,46,92,184,193,377)
378	11	1072	(1,2,3,4,7,14,21,42,63,126,189,378)
379	12	9076	(1,2,3,4,7,10,17,27,44,88,176,203,379)
380	11	1316	(1,2,3,4,7,11,22,44,51,95,190,380)
381	11	4	(1,2,3,5,10,13,23,46,92,184,197,381)
382	11	4	(1,2,4,5,9,14,23,46,92,184,198,382)
383	12	5169	(1,2,3,4,7,8,15,23,46,92,184,199,383)
384	9	8	(1,2,3,6,12,24,48,96,192,384)
385	10	18	(1,2,3,6,12,24,48,96,192,193,385)
386	10	36	(1,2,3,6,12,24,48,96,97,193,386)
387	10	4	(1,2,3,6,12,24,48,96,192,195,387)
388	10	50	(1,2,3,6,12,24,48,49,97,194,388)
389	11	226	(1,2,3,5,6,12,24,48,96,192,197,389)
390	10	12	(1,2,3,6,12,24,48,96,99,195,390)
391	11	86	(1,2,3,5,7,12,24,48,96,192,199,391)
392	10	60	(1,2,3,6,12,24,25,49,98,196,392)
393	11	213	(1,2,3,4,8,16,32,64,67,131,262,393)
394	11	537	(1,2,3,5,6,12,24,48,96,101,197,394)
395	11	20	(1,2,3,6,12,24,25,49,98,196,199,395)
396	10	24	(1,2,3,6,12,24,48,51,99,198,396)
397	11	75	(1,2,3,6,12,13,24,48,96,192,205,397)
398	11	229	(1,2,3,5,7,12,24,48,96,103,199,398)
399	11	42	(1,2,3,5,8,16,32,64,69,133,266,399)
400	10	126	(1,2,3,5,10,15,25,50,100,200,400)
401	11	344	(1,2,3,5,10,15,25,50,100,200,201,401)
402	11	847	(1,2,3,4,8,16,32,35,67,134,201,402)
403	11	65	(1,2,3,5,10,15,25,50,100,200,203,403)
404	11	1321	(1,2,3,5,6,12,24,48,53,101,202,404)
405	11	182	(1,2,3,5,10,15,25,50,100,200,205,405)
406	11	150	(1,2,3,5,10,15,25,50,100,103,203,406)
407	12	4671	(1,2,3,4,7,9,16,25,50,100,200,207,407)
408	10	40	(1,2,3,6,12,24,27,51,102,204,408)
409	11	126	(1,2,3,5,10,20,40,80,83,163,246,409)
410	11	531	(1,2,3,5,10,15,25,50,100,105,205,410)
411	11	42	(1,2,3,6,12,24,27,48,96,192,219,411)
412	11	575	(1,2,3,5,7,12,24,48,55,103,206,412)
413	11	14	(1,2,3,6,12,13,25,50,100,200,213,413)
414	11	238	(1,2,3,5,8,16,32,37,69,138,207,414)
415	11	16	(1,2,3,5,10,15,25,50,100,200,215,415)
416	10	80	(1,2,3,5,8,13,26,52,104,208,416)
417	11	228	(1,2,3,5,8,13,26,52,104,208,209,417)
418	11	425	(1,2,3,5,8,13,26,52,104,105,209,418)
419	11	30	(1,2,3,5,8,13,26,52,104,208,211,419)
420	11	2336	(1,2,3,4,7,14,21,35,70,105,210,420)

n	$\ell(n)$	Qte	exemplaire
421	11	16	(1,2,3,5,8,13,26,52,104,208,213,421)
422	11	88	(1,2,3,5,8,13,26,52,104,107,211,422)
423	11	10	(1,2,3,6,7,13,26,52,104,208,215,423)
424	11	1391	(1,2,3,5,6,12,24,29,53,106,212,424)
425	11	80	(1,2,3,5,10,15,25,50,100,200,225,425)
426	11	68	(1,2,3,5,8,13,26,52,104,109,213,426)
427	12	2924	(1,2,3,4,7,11,13,26,52,104,208,219,427)
428	11	168	(1,2,3,5,8,13,26,52,55,107,214,428)
429	11	40	(1,2,3,5,8,13,26,52,104,208,221,429)
430	11	72	(1,2,3,5,10,15,25,50,100,115,215,430)
431	12	1461	(1,2,3,4,7,11,15,26,52,104,208,223,431)
432	10	95	(1,2,3,6,9,18,27,54,108,216,432)
433	11	269	(1,2,3,6,9,18,27,54,108,216,217,433)
434	11	547	(1,2,3,6,9,18,27,54,108,109,217,434)
435	11	134	(1,2,3,6,9,18,27,54,108,216,219,435)
436	11	687	(1,2,3,5,8,13,26,52,57,109,218,436)
437	11	22	(1,2,3,6,12,24,48,49,97,194,243,437)
438	11	354	(1,2,3,6,9,18,27,54,108,111,219,438)
439	12	2981	(1,2,3,4,7,9,18,27,54,108,216,223,439)
440	11	1811	(1,2,3,4,7,11,22,33,55,110,220,440)
441	11	84	(1,2,3,6,9,18,27,54,108,216,225,441)
442	11	176	(1,2,3,5,8,13,26,52,104,117,221,442)
443	12	3438	(1,2,3,4,7,11,18,36,54,108,216,227,443)
444	11	682	(1,2,3,5,8,16,21,37,74,111,222,444)
445	12	6297	(1,2,3,4,7,11,22,44,45,89,178,267,445)
446	12	12989	(1,2,3,4,7,9,18,27,54,108,115,223,446)
447	11	8	(1,2,3,6,12,15,27,54,108,216,231,447)
448	10	59	(1,2,3,4,7,14,28,56,112,224,448)
449	11	149	(1,2,3,4,7,14,28,56,112,224,225,449)
450	11	794	(1,2,3,4,7,14,28,56,112,113,225,450)
451	11	25	(1,2,3,4,7,14,28,56,112,224,227,451)
452	11	528	(1,2,3,4,7,14,28,56,57,113,226,452)
453	11	16	(1,2,3,5,7,14,28,56,112,224,229,453)
454	11	68	(1,2,3,4,7,14,28,56,112,115,227,454)
455	11	20	(1,2,3,4,7,14,28,56,112,224,231,455)
456	11	1662	(1,2,3,4,7,14,28,29,57,114,228,456)
457	11	4	(1,2,4,5,9,14,28,56,112,224,233,457)
458	11	36	(1,2,3,5,7,14,28,56,112,117,229,458)
459	11	32	(1,2,3,6,9,18,27,54,108,216,243,459)
460	11	211	(1,2,3,4,7,14,28,56,59,115,230,460)
461	12	2004	(1,2,3,4,7,11,22,44,88,95,183,278,461)
462	11	100	(1,2,3,4,7,14,28,56,112,119,231,462)
463	12	1495	(1,2,3,4,7,14,15,28,56,112,224,239,463)
464	11	1169	(1,2,3,4,7,11,18,29,58,116,232,464)
465	12	6352	(1,2,3,4,7,11,18,29,58,116,232,233,465)
466	11	8	(1,2,4,5,9,14,28,56,112,121,233,466)
467	12	1151	(1,2,3,4,7,11,18,29,58,116,232,235,467)
468	11	692	(1,2,3,5,7,14,28,56,61,117,234,468)
469	12	3393	(1,2,3,4,7,14,21,28,56,112,224,245,469)
470	12	11867	(1,2,3,4,7,10,20,27,47,94,141,235,470)
471	12	2139	(1,2,3,4,7,11,18,29,58,116,232,239,471)
472	11	222	(1,2,3,4,7,14,28,31,59,118,236,472)
473	12	1757	(1,2,3,4,7,11,22,33,55,110,220,253,473)
474	12	10531	(1,2,3,4,7,9,18,36,43,79,158,237,474)
475	12	1757	(1,2,3,4,7,11,18,29,58,116,232,243,475)
476	11	168	(1,2,3,4,7,14,28,56,63,119,238,476)
477	12	2846	(1,2,3,4,7,14,28,29,56,112,224,253,477)
478	12	4192	(1,2,3,4,7,11,18,29,58,116,123,239,478)
479	12	378	(1,2,3,4,7,14,15,29,58,116,232,247,479)
480	10	84	(1,2,3,5,10,15,30,60,120,240,480)

n	$\ell(n)$	Qte	exemplaire
481	11	222	(1,2,3,5,10,15,30,60,120,240,241,481)
482	11	435	(1,2,3,5,10,15,30,60,120,121,241,482)
483	11	84	(1,2,3,5,10,15,30,60,120,240,243,483)
484	11	572	(1,2,3,5,10,15,30,60,61,121,242,484)
485	11	56	(1,2,3,5,10,15,30,60,120,240,245,485)
486	11	308	(1,2,3,5,10,15,30,60,120,123,243,486)
487	12	2394	(1,2,3,4,7,8,15,30,60,120,240,247,487)
488	11	659	(1,2,3,5,7,14,28,33,61,122,244,488)
489	11	8	(1,2,3,5,10,20,40,80,83,163,326,489)
490	11	160	(1,2,3,5,10,15,30,60,120,125,245,490)
491	12	1462	(1,2,3,4,7,11,15,30,60,120,240,251,491)
492	11	394	(1,2,3,5,10,15,30,60,63,123,246,492)
493	12	2674	(1,2,3,4,7,11,18,29,58,116,232,261,493)
494	12	9513	(1,2,3,4,7,8,15,30,60,120,127,247,494)
495	11	24	(1,2,3,5,10,15,30,60,120,240,255,495)
496	11	799	(1,2,3,4,7,14,17,31,62,124,248,496)
497	12	2733	(1,2,3,4,7,14,17,31,62,124,248,249,497)
498	11	24	(1,2,3,5,10,20,40,43,83,166,249,498)
499	12	724	(1,2,3,4,7,14,17,31,62,124,248,251,499)
500	11	368	(1,2,3,5,10,15,25,50,75,125,250,500)
501	12	2126	(1,2,3,4,7,10,20,40,80,87,167,334,501)
502	12	5203	(1,2,3,4,7,11,15,30,60,120,131,251,502)
503	12	699	(1,2,3,4,7,14,17,31,62,124,248,255,503)
504	11	1284	(1,2,3,4,7,14,21,42,63,126,252,504)
505	12	5374	(1,2,3,4,7,14,21,42,63,126,252,253,505)
506	12	11096	(1,2,3,4,7,11,22,33,55,110,143,253,506)
507	12	2903	(1,2,3,4,7,14,21,42,63,126,252,255,507)
508	12	17961	(1,2,3,4,7,8,15,30,60,67,127,254,508)
509	12	957	(1,2,3,4,7,14,28,56,57,113,226,283,509)
510	11	96	(1,2,3,5,10,15,30,60,120,135,255,510)
511	12	944	(1,2,3,4,7,14,21,42,63,126,252,259,511)
512	9	1	(1,2,4,8,16,32,64,128,256,512)
513	10	2	(1,2,4,8,16,32,64,128,256,257,513)
514	10	4	(1,2,4,8,16,32,64,128,129,257,514)
515	11	23	(1,2,3,4,8,16,32,64,128,256,259,515)
516	10	6	(1,2,4,8,16,32,64,65,129,258,516)
517	11	21	(1,2,3,5,8,16,32,64,128,256,261,517)
518	11	58	(1,2,3,4,8,16,32,64,128,131,259,518)
519	12	1009	(1,2,3,4,7,8,16,32,64,128,256,263,519)
520	10	8	(1,2,4,8,16,32,33,65,130,260,520)
521	11	21	(1,2,4,8,9,16,32,64,128,256,265,521)
522	11	50	(1,2,3,5,8,16,32,64,128,133,261,522)
523	12	915	(1,2,3,4,8,11,16,32,64,128,256,267,523)
524	11	101	(1,2,3,4,8,16,32,64,67,131,262,524)
525	12	2712	(1,2,3,4,7,14,21,35,70,105,175,350,525)
526	12	3143	(1,2,3,4,7,8,16,32,64,128,135,263,526)
527	12	545	(1,2,3,4,7,14,17,31,62,124,248,279,527)
528	10	10	(1,2,4,8,16,17,33,66,132,264,528)
529	11	25	(1,2,4,8,16,17,32,64,128,256,273,529)
530	11	50	(1,2,4,8,9,16,32,64,128,137,265,530)
531	12	1754	(1,2,3,4,7,11,22,33,66,132,264,267,531)
532	11	87	(1,2,3,5,8,16,32,64,69,133,266,532)
533	12	1268	(1,2,3,5,6,11,22,33,66,132,264,269,533)
534	12	5736	(1,2,3,4,7,11,22,33,66,132,135,267,534)
535	12	275	(1,2,3,4,7,11,22,33,66,132,264,271,535)
536	11	152	(1,2,3,4,8,16,32,35,67,134,268,536)
537	12	884	(1,2,3,4,7,11,22,44,88,91,179,358,537)
538	12	2739	(1,2,3,4,8,16,32,35,67,134,135,269,538)
539	12	471	(1,2,3,4,7,11,22,33,66,132,264,275,539)
540	11	360	(1,2,3,5,10,15,30,45,90,135,270,540)

GÉNÉRALISATIONS DES NOMBRES DE STIRLING ET p, q -ANALOGUES

par Anne de Médicis
et Pierre Leroux

LACIM

Département de Mathématiques et Informatique
Université du Québec à Montréal

1. INTRODUCTION

DÉFINITION: Un *tableau 0-1* consiste en un couple $\varphi = (\lambda, f)$ tel que $\lambda = (\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_k)$ est un partage d'un entier $m = |\lambda|$ et $f = (f_{ij})_{1 \leq j \leq \lambda_i}$ est un "remplissage" des cellules du diagramme de Ferrers de λ avec des 0 et des 1, de telle sorte qu'il y ait exactement un 1 dans chaque colonne.

Par exemple, la figure suivante représente le tableau 0-1 $\varphi = (\lambda, f)$ où $\lambda = (8, 7, 6, 2)$ et $f_{15} = f_{17} = f_{18} = f_{21} = f_{23} = f_{32} = f_{34} = f_{36} = 1$, $f_{ij} = 0$ ailleurs ($1 \leq j \leq \lambda_i$).

0	0	0	0	1	0	1	1
1	0	1	0	0	0	0	
0	1	0	1	0	1		
0	0						

Nous définissons deux statistiques sur ces objets: d'abord le *nombre d'inversions de* φ , noté $\text{inv}(\varphi)$, qui est égal au nombre de 0 en dessous d'un 1 dans φ ; puis le *nombre de non-inversions de* φ , noté $\text{nin}(\varphi)$, qui est égal au nombre de 0 au-dessus d'un 1 dans φ . Dans l'exemple précédent, $\text{inv}(\varphi) = 7$ et $\text{nin}(\varphi) = 8$.

Définissons ensuite les p, q -*nombres de Stirling de première et de deuxième espèces* comme étant respectivement les polynômes suivants:

$$c_{p,q}[n, k] := \sum_{\varphi \in Td(n-1, n-k)} p^{\text{nin}(\varphi)} q^{\text{inv}(\varphi)} \quad (1.1)$$

$$\text{et } s_{p,q}[n, k] := \sum_{\varphi \in T(k, n-k)} p^{\text{nin}(\varphi)} q^{\text{inv}(\varphi)}, \quad (1.2)$$

où $T(h, r)$ dénote l'ensemble des tableaux 0-1 possédant exactement r colonnes de longueurs plus petites ou égales à h , et $Td(h, r)$ dénote le sous-ensemble de $T(h, r)$ contenant les tableaux 0-1 possédant des colonnes de longueurs deux à deux distinctes.

Ces polynômes vérifient les récurrences suivantes (qui sont des p, q -analogues des récurrences des nombres de Stirling classiques):

si $n \geq 1$,

$$c_{p,q}[n, k] = c_{p,q}[n-1, k-1] + [n-1]_{p,q} c_{p,q}[n-1, k] \quad (1.3)$$

$$\text{et } S_{p,q}[n, k] = S_{p,q}[n-1, k-1] + [k]_{p,q} S_{p,q}[n-1, k], \quad (1.4)$$

avec conditions initiales $c_{p,q}[0, k] = \delta_{0,k} = S_{p,q}[0, k]$ et où $[n]_{p,q} = (p^n - q^n)/(p - q)$.

L'étude combinatoire des tableaux 0-1 a permis des développements intéressants dans la théorie des q -analogues des nombres de Stirling classiques. Ces tableaux ont notamment servi à établir une conjecture de L. Butler sur la q -concavité logarithmique des q -nombres de Stirling [But1, Le2]. De plus, nous avons présenté dans un article récent [dML1] une étude systématique des q et p, q -analogues d'identités faisant intervenir les nombres de Stirling des deux espèces, à l'aide de calculs algébriques et/ou combinatoires. Nous avons ainsi obtenu de manière bijective un certain nombre d'identités remarquables.

Dans le présent article, nous considérons deux généralisations des nombres de Stirling, suggérées par l'approche combinatoire des tableaux 0-1.

D'un côté, la généralisation consiste à restreindre les longueurs de colonnes admises et à remplacer le p, q -comptage du remplissage de colonnes par une pondération sur les colonnes. Plus exactement, si $\mathfrak{A} = (\mathcal{A}, w)$, où $\mathcal{A} = (a_i)_{i \geq 0}$ dénote une suite strictement croissante d'entiers non-négatifs (les longueurs de colonnes admises), et $w : \mathbb{N} \rightarrow K$ dénote une fonction de \mathbb{N} à valeurs dans un anneau K (le poids des colonnes selon leur longueur), on définit un \mathcal{A} -tableau comme une liste φ des colonnes c du diagramme de Ferrers d'un partage d'entier λ (par ordre décroissant de longueur) dont les longueurs $|c|$ apparaissent dans la suite \mathcal{A} , et on pose

$$w_{\mathfrak{A}}(\varphi) = \prod_{c \in \varphi} w(|c|). \quad (1.5)$$

Par convention, φ peut contenir un nombre fini de colonnes de longueur nulle. Pour simplifier l'écriture, nous noterons $w_i := w(a_i)$. On définit alors

$$c^{\mathfrak{A}}(n, k) := \sum_{\varphi \in Td^{\mathcal{A}}(n-1, n-k)} w_{\mathfrak{A}}(\varphi) \quad (1.6)$$

$$\text{et } S^{\mathfrak{A}}(n, k) := \sum_{\varphi \in T^{\mathcal{A}}(k, n-k)} w_{\mathfrak{A}}(\varphi), \quad (1.7)$$

où $T^{\mathcal{A}}(h, r)$ dénote l'ensemble des \mathcal{A} -tableaux possédant exactement r colonnes dont les longueurs sont prises dans l'ensemble $\{a_0, a_1, \dots, a_h\}$, et $Td^{\mathcal{A}}(h, r)$ dénote le sous-ensemble

de $T^{\mathcal{A}}(h, r)$ contenant les \mathcal{A} -tableaux possédant des colonnes de longueurs deux à deux distinctes.

Tout comme pour les p, q -nombres de Stirling, en considérant la longueur maximum possible des colonnes des \mathcal{A} -tableaux ainsi que leur format, on obtient les récurrences suivantes:

$$c^{\mathcal{A}}(n, k) = c^{\mathcal{A}}(n-1, k-1) + w_{n-1}c^{\mathcal{A}}(n-1, k) \quad (1.8)$$

$$\text{et } S^{\mathcal{A}}(n, k) = S^{\mathcal{A}}(n-1, k-1) + w_k S^{\mathcal{A}}(n-1, k), \quad (1.9)$$

avec conditions initiales $c^{\mathcal{A}}(0, k) = \delta_{0,k}$, $S^{\mathcal{A}}(0, k) = \delta_{0,k}$, $c^{\mathcal{A}}(n, 0) = w_0 w_1 \dots w_{n-1}$ et $S^{\mathcal{A}}(n, 0) = w_0^n$.

Plusieurs familles de nombres ou de polynômes classiques apparaissent comme cas particulier. Le tableau suivant en donne un aperçu lorsque $a_i = i$.

$w(i)$	$c^{\mathcal{A}}(n, k)$	$S^{\mathcal{A}}(n, k)$
q^i	$q^{\binom{k}{2}} \begin{bmatrix} n \\ k \end{bmatrix}_q$	$\begin{bmatrix} n \\ k \end{bmatrix}_q$
$q^i - 1$	$(q-1)^{n-k} c_q[n, k]$	$(q-1)^{n-k} S_q[n, k]$
$[i]_{p,q}$	$c_{p,q}[n, k]$	$S_{p,q}[n, k]$
$1 + i\alpha$ où α dénote la cardinalité d'un groupe fini	nombres de Whitney des treillis de Dowling [Le1, Dow]	nombres de Whitney des treillis de Dowling [Le1, Dow]
w_i quelconque	$e_{n-k}(w_0, w_1, \dots, w_{n-1})$	$h_{n-k}(w_0, w_1, \dots, w_k)$

où $\begin{bmatrix} n \\ k \end{bmatrix}_q$ dénote les coefficients q -binomiaux et $e_n(x_1, \dots, x_k)$ et $h_n(x_1, \dots, x_k)$ dénotent respectivement les fonctions symétriques élémentaires et homogènes de degré n en k variables.

Ainsi les familles $(c^{\mathcal{A}}(n, k))_{n,k}$ et $(S^{\mathcal{A}}(n, k))_{n,k}$ sont aussi générales que les fonctions symétriques élémentaires et homogènes. Elles possèdent donc toutes sortes de propriétés combinatoires héritées des fonctions symétriques, telles par exemple certaines propriétés de

concavité logarithmique (cf [Hab, Sa]). Dans ce qui suit, nous allons exhiber les relèvements des identités de l'article [dML1] à ce niveau de généralisation. Nous prolongeons ainsi les travaux de L. Comtet, B. Voigt, S. Toader et G.H. Toader (cf [Co, V1, V2, TT]).

D'un autre côté, nous introduisons des *tableaux 0-1 faibles*, remplissages de diagramme de Ferrers avec des 0 et des 1, en exigeant qu'il y ait au plus un 1 dans chaque colonne. Nous définissons combinatoirement des familles de polynômes, les p, q -*nombre de Stirling partiels*, notés $(\mathcal{P}_{p,q}[n, m, k])$ et $(\mathcal{P}S_{p,q}[n, m, k])$, à l'aide de ces tableaux ayant un nombre fixé $n - m$ de colonnes remplies de 0. Lorsque le nombre de colonnes remplies de 0 n'est pas fixé, on obtient un cas particulier de la première généralisation. Les tableaux 0-1 faibles sont d'ailleurs apparus comme modèle combinatoire lorsqu'on considérait les a_i en progression arithmétique et $w_i = [i]_{p,q}$. Nous faisons une étude exhaustive de ce cas dans [dML2].

Les p, q -nombre de Stirling partiels apparaissent naturellement dans le p, q -analogue de certaines identités sur les nombres de Stirling. Ils correspondent, en terme des modèles combinatoires classiques des nombres de Stirling, à un p, q -comptage de permutations à cycles marqués pour $\mathcal{P}_{p,q}[n, m, k]$ et de partitions partielles pour $\mathcal{P}S_{p,q}[n, m, k]$.

Nous présenterons une étude systématique des identités faisant intervenir ces familles de polynômes et nous terminerons avec une discussion sur les propriétés de concavité logarithmique des nouvelles suites de polynômes introduites.

2. RELÈVEMENT D'IDENTITÉS AU NIVEAU DES \mathcal{A} -TABLEAUX

PROPOSITION: Les familles $(c^{\mathcal{A}}(n, k))_{n,k}$ et $(S^{\mathcal{A}}(n, k))_{n,k}$ vérifient les identités suivantes:

a) fonctions symétriques

$$\begin{aligned} c^{\mathcal{A}}(n, k) &= e_{n-k}(w_0, w_1, w_2, \dots, w_{n-1}) \\ &= \sum_{0 \leq i_0 < i_1 < \dots < i_{n-k} \leq n-1} w_{i_0} w_{i_1} \dots w_{i_{n-k}}, \end{aligned} \quad (2.1)$$

$$\begin{aligned} S^{\mathcal{A}}(n, k) &= h_{n-k}(w_0, w_1, w_2, \dots, w_k) \\ &= \sum_{0 \leq i_0 \leq i_1 \leq \dots \leq i_{n-k} \leq k} w_{i_0} w_{i_1} \dots w_{i_{n-k}}; \end{aligned} \quad (2.2)$$

b) fonctions génératrices: pour $n \geq 0$,

$$\sum_{r \geq 0} c^{\mathcal{A}}(n, r) y^{n-r} x^r = (x + w_0 y) (x + w_1 y) \dots (x + w_{n-1} y), \quad (2.3)$$

pour $k \geq 0$,

$$\sum_{r \geq 0} S^{\mathcal{A}}(k+r, k)x^r = \frac{1}{(1-w_0x)} \frac{1}{(1-w_1x)} \cdots \frac{1}{(1-w_kx)}; \quad (2.4)$$

c) récurrences

$$c^{\mathcal{A}}(n+1, k+1) = \sum_{j=k}^n w_n w_{n-1} \cdots w_{j+1} c^{\mathcal{A}}(j, k), \quad (2.5)$$

$$S^{\mathcal{A}}(n+1, k+1) = \sum_{j=k}^n w_{k+1}^{n-j} S^{\mathcal{A}}(j, k). \quad (2.6)$$

DÉMONSTRATION:

- (a) Construire tous les \mathcal{A} -tableaux dans $Td^{\mathcal{A}}(n-1, n-k)$ pour (2.1), ou dans $T^{\mathcal{A}}(k, n-k)$ pour (2.2), en choisissant leurs colonnes (les w_{i_j} représentent les pondérations des colonnes selon leurs longueurs a_{i_j}).
- (b) découle immédiatement de (c).
- (c) démonstration combinatoire analogue à celle de (2.9) et (2.10) dans [dML]: trouver l'escalier ou le rectangle maximum. \square

PROPOSITION: Pour k fixé, les $S^{\mathcal{A}}(n, k)$ vérifient la récurrence:

$$(n-k)S^{\mathcal{A}}(n, k) = \sum_{j=0}^{n-k-1} S^{\mathcal{A}}(n-1-j, k) \left(w_0^{j+1} + w_1^{j+1} + \cdots + w_k^{j+1} \right). \quad (2.7)$$

DÉMONSTRATION: Nous comptons avec pondération les \mathcal{A} -tableaux dans $T^{\mathcal{A}}(k, n-k)$ dont une colonne est pointée. Pour obtenir le membre de droite de (2.7), il suffit d'isoler la colonne pointée ainsi que toutes les colonnes de même longueur se trouvant à gauche d'elle; il y en a exactement j , $0 \leq j \leq n-k-1$, et le \mathcal{A} -tableau restant possède $n-k-(j+1)$ colonnes dont les longueurs sont dans l'ensemble $\{a_0, a_1, \dots, a_k\}$. \square

PROPOSITION: (Orthogonalité)

Si $s^{\mathcal{A}}(n, k) := (-1)^{n-k} c^{\mathcal{A}}(n, k)$, alors les matrices $\left(S^{\mathcal{A}}(n, k) \right)_{n,k}$ et $\left(s^{\mathcal{A}}(n, k) \right)_{n,k}$ sont l'inverse l'une de l'autre. En particulier,

$$\sum_{k=m}^n (-1)^{n-k} c^{\mathcal{A}}(n, k) S^{\mathcal{A}}(k, m) = \delta_{n,m}, \quad (2.8)$$

$$\sum_{k=m}^n S^{\mathcal{A}}(n, k) (-1)^{k-m} c^{\mathcal{A}}(k, m) = \delta_{n,m}. \quad (2.9)$$

DÉMONSTRATION: C'est algébriquement trivial si l'on considère les fonctions génératrices (comparer l'identité (2.3) avec la substitution $x := 1$ et $y := -x$ et l'identité (2.4)).

Il existe une jolie preuve combinatoire involutive des identités (2.8) et (2.9) consistant à déplacer une colonne de longueur minimum (resp. maximum pour (2.9)) à l'intérieur de paires de \mathcal{A} -tableaux dans les ensembles $Td^{\mathcal{A}}(n-1, n-k) \times T^{\mathcal{A}}(m, k-m)$ (resp. $T^{\mathcal{A}}(k, n-k) \times Td^{\mathcal{A}}(k-1, k-m)$), $m \leq k \leq n$. Les détails sont laissés au lecteur. \square

PROPOSITION: On a:

$$c^{\mathfrak{A}}(n, k) = \sum_{j=k}^n (-1)^{j-k} w_n^{j-k} c^{\mathfrak{A}}(n+1, j+1), \quad (2.10)$$

$$S^{\mathfrak{A}}(n, k) = \sum_{j=0}^{n-k} (-1)^j w_{k+1} w_{k+2} \dots w_{k+j} S^{\mathfrak{A}}(n+1, k+j+1), \quad (2.11)$$

$$\sum_{k=2}^n (1-w_1)(1-w_2) \dots (1-w_{k-1}) S^{\mathfrak{A}}(n, k) = 1 + w_0 + \dots + w_0^{n-2}. \quad (2.12)$$

DÉMONSTRATION: Même genre de preuves involutives que la proposition précédente. \square

Le relèvement de certaines identités dépend parfois de la pondération w choisie. C'est le cas des identités (2.11), (2.12), et (4.5) à (4.10) de l'article [dML1]. Dans ce cas, la pondération a souvent la forme d'une somme canonique finie d'éléments de K . Pour faciliter les démonstrations combinatoires, on peut distribuer les différents termes de la pondération sur les \mathcal{A} -tableaux φ en attribuant à chaque colonne un seul terme de la somme correspondant à sa longueur plutôt que la somme entière. On retrouve d'ailleurs ainsi l'interprétation combinatoire en termes de tableaux 0-1 des p, q -nombres de Stirling des deux espèces obtenus avec $a_i = i$ et $w(i) = [i]_{p,q}$. Plutôt que d'attribuer à une colonne de longueur i un poids unique de la forme $[i]_{p,q} = p^{i-1} + p^{i-2}q + \dots + q^{i-1}$, on lui attribue un seul monôme $p^j q^{i-j-1}$ dans $[i]_{p,q}$. En général, on dira que φ est un \mathcal{A} -tableau w -distribué.

PROPOSITION: (relèvement de (2.11) et (2.12))

Si la pondération de $\mathfrak{A} = (\mathcal{A}, w)$ est de la forme

$$w(a_i) = w_0 + w_0 z + w_0 z^2 + \dots + w_0 z^i \in \mathbb{N}[z],$$

où z est une variable ou une expression formelle et w_0 est une constante, alors

$$c^{\mathfrak{A}}(n+1, k+1) = \sum_{j=k}^n \binom{j}{k} w_0^{j-k} z^{n-j} c^{\mathfrak{A}}(n, j), \quad (2.13)$$

$$S^{\mathfrak{A}}(n+1, k+1) = \sum_{m=k}^n \binom{n}{m} w_0^{n-m} z^{m-k} S^{\mathfrak{A}}(m, k). \quad (2.14)$$

DÉMONSTRATION: On peut supposer sans perte de généralité que $a_i = i$.

Alors pour (2.13), soit $n \geq k$ et φ un \mathcal{A} -tableau w -distribué dans $Td^{\mathcal{A}}(n, n-k)$. Considérons toutes les colonnes de φ qui ont poids w_0 . Il y en a exactement $j-k$, pour un certain j dépendant de φ ($k \leq j \leq n$). On peut enlever ces colonnes, en prenant note de leurs longueurs, sans perte de généralité, obtenant un facteur w_0^{j-k} . Si φ' désigne le \mathcal{A} -tableau obtenu après cette opération, alors φ' est dans $Td^{\mathcal{A}}(n, n-j)$, c'est-à-dire φ' contient $n-j$ colonnes (de longueurs distinctes plus petites ou égales à $n-1$) et les longueurs des $j-k$ colonnes extraites constituent un sous-ensemble des j longueurs de colonnes n'apparaissant pas dans φ' . De plus, chaque colonne de φ' de longueur i possède une pondération $w_0 z^l$, $1 \leq l \leq i-1$. Si l'on remplace chaque colonne de longueur i par une colonne de longueur $i-1$ et qu'on enlève un facteur z à sa pondération, on obtient un \mathcal{A} -tableau w -pondéré général dans $Td^{\mathcal{A}}(n-1, n-j)$ et un facteur z^{n-j} .

Il est facile de voir que cette transformation est réversible; on obtient ainsi l'identité (2.13). L'identité (2.14) s'obtient d'une manière similaire. \square

À l'origine, les q -nombres de Stirling de deuxième espèce ont été introduits par L. Carlitz [Ca] comme étant la famille de polynômes $S_q[n, k]$ vérifiant l'identité suivante:

$$\begin{bmatrix} n \\ k \end{bmatrix}_q = \sum_{j=k}^n \binom{n}{j} (q-1)^{j-k} S_q[j, k]. \quad (2.15)$$

THÉORÈME: (relèvement de la formule de Carlitz)

Soit \mathcal{A} une suite strictement croissante d'entiers non-négatifs et $w : \mathbb{N} \rightarrow K$ et $w^* : \mathbb{N} \rightarrow K$ deux fonctions dans un anneau K telles que $w(i) = w_0 + w^*(i)$. Si $\mathfrak{A} = (\mathcal{A}, w)$, $\mathfrak{A}^* = (\mathcal{A}, w^*)$ et $\mathfrak{A}_0 = (\mathcal{A}, w_0)$ où w_0 dénote la fonction constante égale à w_0 , alors

$$S^{\mathfrak{A}}(n, k) = \sum_{j=k}^n \binom{n}{j} w_0^{n-j} S^{\mathfrak{A}^*}(j, k). \quad (2.16)$$

De plus, les identités suivantes, obtenues en appliquant diverses inversions de Möbius, sont équivalentes à (2.16).

$$c^{\mathfrak{A}}(n, s) = \sum_{j=s}^n \binom{j}{s} w_0^{j-s} c^{\mathfrak{A}^*}(n, j); \quad (2.17)$$

$$S^{\mathfrak{A}^*}(n, s) = \sum_{j=s}^n (-w_0)^{n-j} \binom{n}{j} S^{\mathfrak{A}}(j, s); \quad (2.18)$$

$$c^{\mathcal{A}^*}(n, s) = \sum_{j=s}^n (-w_0)^{j-s} \binom{j}{s} c^{\mathcal{A}}(n, j); \quad (2.19)$$

$$S^{\mathcal{A}_0}(n, s) = \sum_{j=s}^n (-1)^{j-s} c^{\mathcal{A}^*}(j, s) S^{\mathcal{A}}(n, j); \quad (2.20)$$

$$c^{\mathcal{A}_0}(n, s) = \sum_{j=s}^n (-1)^{n-j} c^{\mathcal{A}}(j, s) S^{\mathcal{A}^*}(n, j). \quad (2.21)$$

DÉMONSTRATION: Pour (2.16) et (2.17), il suffit essentiellement de choisir les colonnes des \mathcal{A} -tableaux w -distribués dont la pondération est différente de w_0 .

Pour (2.18) et (2.19), le membre de droite compte des \mathcal{A} -tableaux w -distribués dont certaines colonnes de poids w_0 sont distinguées et contribuent à la pondération par un facteur de (-1) . On peut construire sur ces tableaux une involution conservant les poids mais inversant les signes et ayant comme points fixes les \mathcal{A} -tableaux w -distribués dont aucune colonne n'est de poids w_0 . Il suffit, pour ce faire, de "changer le statut" de la colonne de poids w_0 la plus à gauche du \mathcal{A} -tableau w -distribué. On obtient ainsi le membre de gauche.

Pour (2.20) et (2.21), des involutions conservant les poids mais inversant le signe du même genre que celles pour l'orthogonalité font l'affaire. \square

Le cas particulier où $w^*(i) = q^i - 1$, $w_0 = 1$ et $a_i = i$ redonne l'identité de Carlitz ainsi que ses proches parents (cf [dML1]).

3. IDENTITÉS RELATIVES AUX TABLEAUX 0-1 FAIBLES

DÉFINITION: Un *tableau 0-1 faible* Ψ est la donnée d'un triplet $\Psi = (\lambda, l, f)$ où $\lambda = (\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_k)$ désigne un partage d'un entier m , l est un entier tel que $l \geq \lambda_1$ et $f = (f_{ij})_{1 \leq j \leq \lambda_i}$ est un "remplissage" des cellules du diagramme de Ferrers de λ avec des 0 et des 1, de telle sorte qu'il y ait au plus un 1 par colonne. On considérera par convention que Ψ possède l colonnes dont $l - \lambda_1$ sont de longueur nulle et sont remplies de zéros.

Par exemple, la figure suivante représente le tableau 0-1 faible $\Psi = (\lambda, l, f)$ où $\lambda = (5, 4, 2, 2)$, $l = 6$ et $f_{13} = f_{24} = f_{32} = 1$, $f_{ij} = 0$ ailleurs ($1 \leq j \leq \lambda_i$).

Le *nombre d'inversions* de Ψ , noté $\text{inv}(\Psi)$, est égal au nombre de 0 se trouvant en dessous d'un 1 dans la représentation graphique de Ψ . Le *nombre de non-inversions* de Ψ , noté $\text{nin}(\Psi)$, est égal au nombre de 0 se trouvant au-dessus d'un 1 dans la représentation

0	0	1	0	0	
0	0	0	1		
0	1				
0	0				

graphique de Ψ plus le nombre de 0 dans les colonnes remplies de 0. Dans l'exemple précédent, $\text{inv}(\Psi) = 2$ et $\text{nin}(\Psi) = 8$.

Nous noterons $T_f(h, r)$ (resp. $T_f(h, r, s)$) l'ensemble des tableaux 0-1 faibles possédant r colonnes de longueurs plus petites ou égales à h (et possédant exactement s colonnes remplies de 0 resp.) et $Td_f(h, r)$ (resp. $Td_f(h, r, s)$) le sous-ensemble de $T_f(h, r)$ (resp. $T_f(h, r, s)$) contenant les tableaux 0-1 faibles dont les longueurs des colonnes sont deux à deux distinctes.

DÉFINITION: Les p, q -nombres de Stirling partiels sont donnés par:

$$\mathcal{P}S_{p,q}[n, m, k] := \sum_{\Psi \in T_f(k, n-k, n-m)} p^{\text{nin}(\Psi)} q^{\text{inv}(\Psi)}; \tag{3.1}$$

$$\mathcal{P}C_{p,q}[n, m, k] := \sum_{\Psi \in Td_f(n-1, n-k, n-m)} p^{\text{nin}(\Psi)} q^{\text{inv}(\Psi)}; \tag{3.2}$$

$$\mathcal{P}S_{p,q}[n, k] := \sum_{\Psi \in T_f(k, n-k)} p^{\text{nin}(\Psi)} q^{\text{inv}(\Psi)}; \tag{3.3}$$

$$\mathcal{P}C_{p,q}[n, k] := \sum_{\Psi \in Td_f(n-1, n-k)} p^{\text{nin}(\Psi)} q^{\text{inv}(\Psi)}. \tag{3.4}$$

Remarquons que ces polynômes ne sont pas symétriques en p et q .

PROPOSITION: On a:

a) liens et cas particuliers.

$$\mathcal{P}S_{p,q}[n, k] = \sum_{m=k}^n \mathcal{P}S_{p,q}[n, m, k], \tag{3.5}$$

$$\mathcal{P}C_{p,q}[n, k] = \sum_{m=k}^n \mathcal{P}C_{p,q}[n, m, k], \tag{3.6}$$

$$\mathcal{P}S_{p,q}[n, k] = S^{(\mathcal{A}, w)}(n, k) \text{ et } \mathcal{P}C_{p,q}[n, k] = c^{(\mathcal{A}, w)}(n, k) \tag{3.7}$$

où $\mathcal{A} = (i)_{i \geq 0}$ et $w(i) = p^i + [i]_{p,q}$;

b) récurrences.

$$\mathcal{P}S_{p,q}[n + 1, k] = \mathcal{P}S_{p,q}[n, k - 1] + (p^k + [k]_{p,q})\mathcal{P}S_{p,q}[n, k], \tag{3.8}$$

$$\mathcal{P}c_{p,q}[n + 1, k] = \mathcal{P}c_{p,q}[n, k - 1] + (p^n + [n]_{p,q})\mathcal{P}c_{p,q}[n, k], \quad (3.9)$$

avec conditions initiales $\mathcal{P}S_{p,q}[n, k] = 0 = \mathcal{P}c_{p,q}[n, k]$ si $k > n$, $\mathcal{P}S_{p,q}[n, 0] = 1$
 $\mathcal{P}c_{p,q}[n, 0] = (p + [1]_{p,q})(p^2 + [2]_{p,q}) \dots (p^{n-1} + [n - 1]_{p,q})$ et $\mathcal{P}S_{p,q}[0, k] = \delta_{0,k} =$
 $\mathcal{P}c_{p,q}[0, k]$.

DÉMONSTRATION:

a) (3.5) et (3.6) découlent de la définition des $\mathcal{P}S_{p,q}[n, k]$ et $\mathcal{P}c_{p,q}[n, k]$ par rapport à celles de $\mathcal{P}S_{p,q}[n, m, k]$ et $\mathcal{P}c_{p,q}[n, m, k]$.

Pour (3.7), comparer par exemple les récurrences (3.8) et (3.9) données en (b) avec les formules (1.8) et (1.9).

b) Pour démontrer les récurrences, il suffit de considérer la longueur maximum possible des colonnes des tableaux 0-1 faibles dans les ensembles $T_f(k, n - k + 1)$ et $Td_f(n, n - k + 1)$ respectivement (i.e. si $\Psi \in T_f(k, n - k + 1)$ par exemple, soit sa colonne de longueur maximum est de longueur k , soit elle est de longueur strictement plus petite). \square

Les suites $(\mathcal{P}S_{p,q}[n, k])$ et $(\mathcal{P}c_{p,q}[n, k])$ étant des cas particuliers des suites $(S^{\mathcal{A}}(n, k))$ et $(c^{\mathcal{A}}(n, k))$, toutes les identités de la section précédente s'y appliquent. Ce n'est en revanche pas le cas des suites $(\mathcal{P}S_{p,q}[n, m, k])$ et $(\mathcal{P}c_{p,q}[n, m, k])$ pour lesquelles on peut trouver de nouvelles identités.

PROPOSITION: On a:

a) liens et cas particuliers.

$$\mathcal{P}S_{p,q}[n, n, k] = S_{p,q}[n, k] \text{ et } \mathcal{P}S_{p,q}[n, k, k] = \begin{bmatrix} n \\ k \end{bmatrix}_p, \quad (3.10)$$

$$\mathcal{P}c_{p,q}[n, n, k] = c_{p,q}[n, k] \text{ et } \mathcal{P}c_{p,q}[n, k, k] = p^{\binom{n-k}{2}} \begin{bmatrix} n \\ k \end{bmatrix}_p, \quad (3.11)$$

$$\mathcal{P}S_{1,q}[n, m, k] = \binom{n}{m} S_q[m, k], \quad (3.12)$$

$$\mathcal{P}c_{1,q}[n, m, k] = \binom{n - m + k}{k} c_q[n, n - m + k]; \quad (3.13)$$

b) récurrences.

$$\begin{aligned} \mathcal{P}S_{p,q}[n + 1, m + 1, k] &= \mathcal{P}S_{p,q}[n, m, k - 1] + p^k \mathcal{P}S_{p,q}[n, m + 1, k] \\ &\quad + [k]_{p,q} \mathcal{P}S_{p,q}[n, m, k], \end{aligned} \quad (3.14)$$

$$\begin{aligned} \mathcal{P}c_{p,q}[n + 1, m + 1, k] &= \mathcal{P}c_{p,q}[n, m, k - 1] + p^n \mathcal{P}c_{p,q}[n, m + 1, k] + \\ &\quad [n]_{p,q} \mathcal{P}c_{p,q}[n, m, k], \end{aligned} \quad (3.15)$$

avec conditions initiales $\mathcal{P}S_{p,q}[n, m, k] = 0 = \mathcal{P}C_{p,q}[n, m, k]$ si $m > n$ ou si $k > m$, $\mathcal{P}S_{p,q}[0, m, k] = \delta_{0,k}\delta_{0,m} = \mathcal{P}C_{p,q}[0, m, k]$, $\mathcal{P}S_{p,q}[n, 0, k] = \delta_{0,k}$, $\mathcal{P}C_{p,q}[n, 0, k] = p^{\binom{n}{2}}\delta_{k,0}$ et $\mathcal{P}S_{p,q}[n, m, 0] = \delta_{m,0}$, $\mathcal{P}C_{p,q}[n, m, 0] = p^{\binom{n}{2}-m}c_{1,q/p}[n, n - m]$;

c) séries génératrices. Pour $n > 0$,

$$\sum_{r \geq 0} \sum_{m \geq 0} \mathcal{P}C_{p,q}[n, m, r] x^r y^{m-r} z^{n-m} = (x + z)(x + pz + [1]_{p,q}y) \dots (x + p^{n-1}z + [n - 1]_{p,q}y), \quad (3.16)$$

pour $k \geq 0$,

$$\sum_{r \geq 0} \sum_{l \geq 0} \mathcal{P}S_{p,q}[k + r, k + l, k] x^l y^{r-l} = \frac{1}{(1 - y)} \frac{1}{(1 - py - [1]_{p,q}x)} \frac{1}{(1 - p^2y - [2]_{p,q}x)} \dots \frac{1}{(1 - p^k y - [k]_{p,q}x)}. \quad (3.17)$$

DÉMONSTRATION:

a) Lorsque $n = m$, toutes les colonnes contiennent un 1 et $\mathcal{P}S_{p,q}[n, n, k]$ et $\mathcal{P}C_{p,q}[n, n, k]$ comptent des tableaux 0-1 normaux.

Lorsque $m = k$, les tableaux 0-1 faibles considérés ne contiennent que des cases remplies avec des 0 et comptant chacune pour p , qui correspondent à l'interprétation combinatoire des coefficients q -binomiaux en terme de diagrammes de Ferrers.

Pour les deux égalités suivantes, il suffit de montrer que $\binom{n}{m}S_q[m, k]$ et $\binom{n-m+k}{k} \times c_q[n, n - m + k]$ vérifient les récurrences (3.14) et (3.15) de (b) pour $p = 1$.

b) Même raisonnement que la proposition précédente. Remarquons que $\mathcal{P}C_{p,q}[n, m, 0]$ compte avec pondération les tableaux 0-1 faibles $\Psi \in Td_f(n - 1, n, n - m)$, donc les "escaliers" de $n - 1$ à 0 possédant m colonnes avec des 1. Le nombre d'inversions de tels tableaux est égal au nombre d'inversions du tableau 0-1 $\varphi \in Td(n - 1, m)$ obtenu en considérant seulement les colonnes contenant des 1 tandis que le nombre de non-inversions du tableau Ψ est égal à

$$\# \text{ cases dans l'escalier} - \# \text{ de 1 dans le tableau} - \text{inv}(\varphi) = \binom{n}{2} - m - \text{inv}(\varphi).$$

D'où $\mathcal{P}C_{p,q}[n, m, 0] = p^{\binom{n}{2}-m}c_{1,q/p}[n, n - m]$.

c) découle de la définition combinatoire. □

Dans le cas $p = 1$, les nombres de Stirling partiels possèdent des expressions en termes des q -nombres de Stirling ordinaires.

PROPOSITION: On a:

$$\mathcal{P}c_{1,q}[n+1, k+1] = \sum_{j=k}^n \left[\binom{j+1}{k+1} + \binom{j}{k} \right] 2^{j-k-1} q^{n-j} c_q[n, j], \quad (3.18)$$

$$\mathcal{P}S_{1,q}[n+1, k+1] = \sum_{j=0}^n \left[\sum_{i=0}^{n-j} \binom{n-i}{j} 2^{n-j-i} \right] q^{j-k} S_q[j, k], \quad (3.19)$$

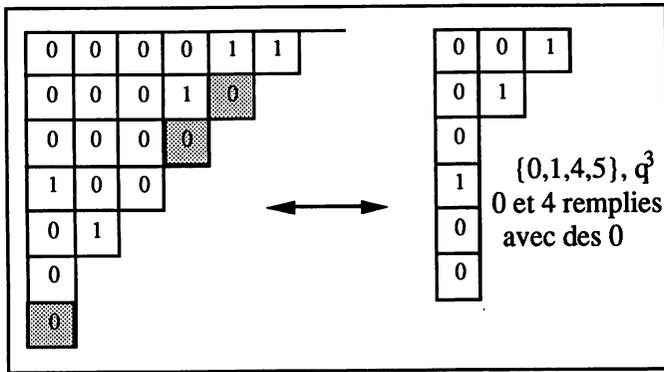
$$\begin{aligned} & \mathcal{P}c_{1,q}[n+1, m+1, k+1] \\ &= \sum_{j=k}^n \left[\binom{j}{k} \binom{j-k}{n-m} + \binom{j}{k+1} \binom{j-k-1}{n-m-1} \right] q^{n-j} c_q[n, j], \end{aligned} \quad (3.20)$$

$$\begin{aligned} & \mathcal{P}S_{1,q}[n+1, m+1, k+1] \\ &= \sum_{j=0}^n \left[\sum_{i=0}^{n-j} \binom{n-i}{j} \binom{n-j-i}{n-m-i} \right] q^{j-k} S_q[j, k]. \end{aligned} \quad (3.21)$$

DÉMONSTRATION: Le principe de démonstration de ces formules est le même que celui de (2.13) et (2.14). Dans ce qui suit, nous démontrons (3.22) transformée à l'aide de l'identité

$$2^{j-k-1} \left[\binom{j+1}{k+1} + \binom{j}{k} \right] = 2^{j-k-1} \binom{j}{j-k-1} + 2^{j-k} \binom{j}{j-k}.$$

Si $\Psi \in Td_f(n, k)$, ôtons-lui ses colonnes qui ne causent aucune inversion, il y en a exactement $j-k$ pour un certain j , $k \leq j \leq n$. D'un côté, il reste un tableau 0-1 normal dont chaque colonne compte au moins pour une inversion. En lui enlevant une case à l'extrémité inférieure de chaque colonne, on obtient un tableau 0-1 général φ dans l'ensemble $Td(n-1, n-j)$, ce qui explique dans la somme le terme $q^{n-j} c_q[n, j]$ rajusté pour tenir compte des inversions perdues. De l'autre côté, les $j-k$ colonnes extraites de Ψ prenaient des longueurs dans les $n - (n-j) = j$ éléments de $[n]$ n'apparaissant pas dans φ (puisque les colonnes de Ψ sont de longueurs deux à deux disjointes) plus éventuellement la longueur zéro. De plus, les colonnes qui ne contribuent pas au nombre d'inversions d'un tableau 0-1 faible sont de deux types, soit elles ne contiennent que des zéros, soit elles contiennent un 1 positionné dans la case inférieure de la colonne. Cependant lorsqu'une colonne est de longueur nulle, elle ne peut être que du premier type, ce qui nous donne donc deux sous-cas: si les colonnes extraites de Ψ sont de longueurs non nulles, elles forment un sous-ensemble à $j-k$ éléments d'un ensemble de j longueurs possibles et chacune d'entre



elles a le choix de deux configurations, ce qui explique le terme $2^{j-k} \binom{j}{j-k}$. De même, si l'une des colonnes extraites de Ψ est de longueur nulle, les autres forment un sous-ensemble à $j - k - 1$ éléments des mêmes j longueurs possibles avec toujours deux possibilités de configurations, ce qui explique le terme $2^{j-k-1} \binom{j}{j-k-1}$.

Raisonnement similaire pour (3.19) à (3.21). □

Les suites $(\mathcal{P}S_{p,q}[n, m, k])$ et $(\mathcal{P}c_{p,q}[n, m, k])$ trouvent la justification de leur existence dans la recherche d'un p, q -analogue d'une identité due à Chen [Ch]:

$$S(m + n, k) = \sum_{i+j \geq k} \binom{m}{j} i^{m-j} S(n, i) S(j, k - i). \tag{3.22}$$

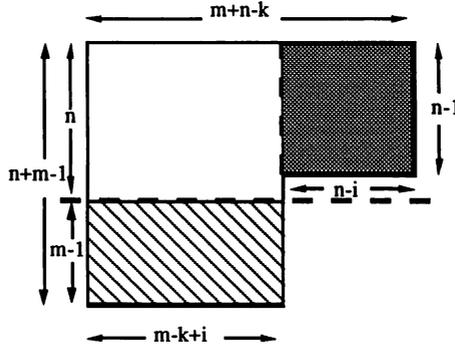
PROPOSITION: On a:

$$c_{p,q}[m + n, k] = \sum_{i+j \geq k} p^{n(i+j-k)} [n]_{p,q}^{m-j} c_{p,q}[n, i] \mathcal{P}c_{q,p}[m, j, k - i], \tag{3.23}$$

$$S_{p,q}[m + n, k] = \sum_{i+j \geq k} p^{i(i+j-k)} [i]_{p,q}^{m-j} S_{p,q}[n, i] \mathcal{P}S_{q,p}[m, j, k - i]. \tag{3.24}$$

DÉMONSTRATION: Soit $\varphi \in Td(m + n - 1, m + n - k)$. φ peut être décomposé uniquement de la manière suivante:

- a) un tableau 0-1 φ' dans l'ensemble $Td(n - 1, n - i)$ en considérant les $n - i$ colonnes de longueurs strictement plus petites que n de φ , correspondant à la partie ombragée de la figure suivante,
- b) un tableau 0-1 faible Ψ_1 dans l'ensemble $Td_f(m - 1, m - k + i, m - j)$ en considérant la partie inférieure des $m - k + i$ colonnes les plus longues qui dépasse la longueur n , correspondant à la partie hachurée de la figure suivante, et finalement



c) un tableau 0-1 faible Ψ_2 rectangulaire de format $(m + i - k) \times n$, contenant $m - j$ colonnes avec un 1 dont les positions correspondent aux positions des $m - j$ colonnes remplies de 0 de Ψ_1 .

Posons Ψ'_1 le tableau 0-1 faible obtenu en appliquant à chaque colonne de Ψ_1 l'involution qui consiste à déplacer le 1 (s'il y en a un) de la position i à la position $l - i + 1$, où l désigne la longueur de la colonne. Il est clair que

$$\begin{aligned} \text{inv}(\varphi) &= \text{inv}(\varphi') + \text{nin}(\Psi'_1) + \text{inv}(\Psi_2) \\ \text{et } \text{nin}(\varphi) &= \text{nin}(\varphi') + \text{inv}(\Psi'_1) + \text{nin}(\Psi_2). \end{aligned}$$

De même, il est clair que

$$\sum_{\Psi_2} p^{\text{inv}(\Psi_2)} q^{\text{nin}(\Psi_2)} = p^{n(i+j-k)} [n]_{p,q}^{m-j},$$

où la somme parcourt l'ensemble des tableaux 0-1 faibles Ψ_2 rectangulaires de format $(m + i - k) \times n$, contenant exactement $i + j - k$ colonnes sans 1 dont les positions sont fixées d'avance, quelques soient ces positions. Par conséquent,

$$\begin{aligned} c_{p,q}[m + n, k] &= \sum_{\varphi \in Td(m+n-1, m+n-k)} p^{\text{inv}(\varphi)} q^{\text{nin}(\varphi)} \\ &= \sum_{(\varphi', \Psi'_1, \Psi_2)} p^{\text{inv}(\varphi') + \text{nin}(\Psi'_1) + \text{inv}(\Psi_2)} q^{\text{nin}(\varphi') + \text{inv}(\Psi'_1) + \text{nin}(\Psi_2)} \\ &= \sum_{i+j \geq k} c_{p,q}[n, i] \mathcal{P}c_{q,p}[m, j, k - i] p^{n(i+j-k)} [n]_{p,q}^{m-j}. \end{aligned}$$

L'identité (3.24) se démontre de manière similaire. □

4. CONCAVITÉ LOGARITHMIQUE ET PROBLÈMES OUVERTS

Soient $f(p, q)$ et $g(p, q) \in \mathbb{N}[p, q]$ des polynômes en p et q . Définissons la relation d'ordre $\leq_{p,q}$ par:

$$f(p, q) \leq_{p,q} g(p, q) \iff g(p, q) - f(p, q) \in \mathbb{N}[p, q].$$

DÉFINITION: Une suite $(a_k(p, q))_{k \in \mathbb{N}} \subseteq \mathbb{N}[p, q]$ est dite p, q -log concave (resp. *fortement* p, q -log concave) si et seulement si

$$a_{k-1}(p, q)a_{k+1}(p, q) \leq_{p,q} a_k^2(p, q) \quad \forall k \geq 1. \quad (4.1)$$

$$\text{(resp. } a_{k-1}(p, q)a_{l+1}(p, q) \leq_{p,q} a_k(p, q)a_l(p, q) \quad \forall l \geq k \geq 1.) \quad (4.2)$$

Plusieurs mathématiciens se sont intéressés récemment aux propriétés d'unimodalité et de (q - et p, q -)concavité logarithmique des coefficients binomiaux et des nombres de Stirling ainsi que leurs q et p, q -analogues (voir [Br, But1, But2, dM, Hab, Le2, OH, Sa1, Sa2, Sta, Ze], etc.). Ces propriétés sont souvent faciles à constater mais difficiles à démontrer.

Nous donnons ici des résultats sur la concavité logarithmique des suites que nous avons introduites dans la section précédente.

THÉORÈME: Les suites suivantes sont fortement p, q -log concaves: $(\mathcal{PS}_{p,q}[n, k])_{n \in \mathbb{N}}$, $(\mathcal{PC}_{p,q}[n, k])_{k \in \mathbb{Z}}$, $(\mathcal{PC}_{p,q}[n, m, k])_{m \in \mathbb{N}}$, $(\mathcal{PS}_{p,q}[n, m, k])_{n \in \mathbb{N}}$ et $(\mathcal{PS}_{p,q}[n, m, k])_{m \in \mathbb{N}}$. Les suites $(\mathcal{PS}_{p,q}[n, n - k])_{n \in \mathbb{N}}$, $(\mathcal{PC}_{p,q}[n, n - k])_{n \in \mathbb{N}}$ et $(\mathcal{PC}_{p,q}[n, k])_{n \in \mathbb{N}}$ ne sont pas p, q -log concaves. De plus, si $n_2 \geq n_1$,

$$\mathcal{PC}_{p,q}[n_1 - 1, k]\mathcal{PC}_{p,q}[n_2 + 1, k] \leq_{p,q} (p + q)^{n_2 - n_1 + 1} \mathcal{PC}_{p,q}[n_1, k]\mathcal{PC}_{p,q}[n_2, k]. \quad (4.3)$$

DÉMONSTRATION: Nous donnons un aperçu du type de preuves que nous avons utilisées pour démontrer chaque résultat.

Puisque

$$\begin{aligned} \mathcal{PC}_{p,q}[n, k] &= e_{n-k}(1, p + [1]_{p,q}, p^2 + [2]_{p,q}, \dots, p^{n-1} + [n-1]_{p,q}) \\ \text{et } \mathcal{PS}_{p,q}[n, k] &= h_{n-k}(p + [1]_{p,q}, p^2 + [2]_{p,q}, \dots, p^k + [k]_{p,q}), \end{aligned}$$

la p, q -log concavité forte de $\mathcal{PS}_{p,q}[n, k]$ en n et de $\mathcal{PC}_{p,q}[n, k]$ en k découle de propriétés semblables sur les fonctions symétriques (cf [dM]). De même, l'identité (4.3) découle d'un résultat de L. Habsieger [Hab] sur les fonctions symétriques élémentaires.

Pour la p, q -log concavité forte des autres suites, nous utilisons des preuves injectives. Ainsi, pour démontrer l'inégalité

$$\mathcal{P}_{\mathcal{C}_{p,q}[n, m_1 - 1, k]} \mathcal{P}_{\mathcal{C}_{p,q}[n, m_2 + 1, k]} \leq_{p,q} \mathcal{P}_{\mathcal{C}_{p,q}[n, m_1, k]} \mathcal{P}_{\mathcal{C}_{p,q}[n, m_2, k]},$$

avec $n \geq m_2 \geq m_1 \geq k$ par exemple, nous avons construit une injection de l'ensemble $Td_f(n-1, n-k, n-m_1+1) \times Td_f(n-1, n-k, n-m_2-1)$ vers l'ensemble $Td_f(n-1, n-k, n-m_1) \times Td_f(n-1, n-k, n-m_2)$, conservant le p, q -comptage.

Finalement,

$$\begin{aligned} \mathcal{P}_{\mathcal{C}_{p,q}[3, 2]} \mathcal{P}_{\mathcal{C}_{p,q}[3, 2]} - \mathcal{P}_{\mathcal{C}_{p,q}[4, 3]} \mathcal{P}_{\mathcal{C}_{p,q}[2, 1]} &= 2p + 2p^2 + 2q + pq + p^2q - q^2 - pq^2 \\ &= \mathcal{P}_{\mathcal{S}_{p,q}[3, 2]} \mathcal{P}_{\mathcal{S}_{p,q}[3, 2]} - \mathcal{P}_{\mathcal{S}_{p,q}[4, 3]} \mathcal{P}_{\mathcal{S}_{p,q}[2, 1]} \end{aligned}$$

$$\text{et } \mathcal{P}_{\mathcal{C}_{p,q}[2, 1]} \mathcal{P}_{\mathcal{C}_{p,q}[2, 1]} - \mathcal{P}_{\mathcal{C}_{p,q}[3, 1]} \mathcal{P}_{\mathcal{C}_{p,q}[3, 1]} = 4 + p - 2p^2 - p^3 - pq - 2q$$

montrent que les suites $(\mathcal{P}_{\mathcal{S}_{p,q}[n, n-k]})_{n \in \mathbb{N}}$, $(\mathcal{P}_{\mathcal{C}_{p,q}[n, n-k]})_{n \in \mathbb{N}}$ et $(\mathcal{P}_{\mathcal{C}_{p,q}[n, k]})_{n \in \mathbb{N}}$ ne sont pas p, q -log concaves. \square

CONJECTURES: les suites $(\mathcal{P}_{\mathcal{C}_{p,q}[n, m, k]})_{k \in \mathbb{N}}$, $(\mathcal{P}_{\mathcal{S}_{p,q}[n, m, k]})_{k \in \mathbb{N}}$ et $(\mathcal{P}_{\mathcal{S}_{p,q}[n, k]})_{k \in \mathbb{N}}$ sont fortement p, q -log concaves. De plus, si $n_2 \geq n_1$, on a

$$\mathcal{P}_{\mathcal{C}_{p,q}[n_1-1, m, k]} \mathcal{P}_{\mathcal{C}_{p,q}[n_2+1, m, k]} \leq_{p,q} (p+q)^{n_2-n_1+1} \mathcal{P}_{\mathcal{C}_{p,q}[n_1, m, k]} \mathcal{P}_{\mathcal{C}_{p,q}[n_2, m, k]}. \quad (4.4)$$

RÉFÉRENCES

- [Br] BRENTI, Francesco, "Unimodal, Log-Concave and Pólya Frequency Sequences in Combinatorics", version révisée de sa thèse de Ph.D., MIT, 1988, 158 p.
- [But1] BUTLER, Lynne M., "A Unimodality Result in the enumeration of Subgroups of a finite Abelian Group", Proceedings of the American Mathematical Society, Vol. 101, No 4, décembre 1987, pp. 771-775.
- [But2] BUTLER, Lynne M., "The q -log concavity of q -binomial coefficients", J. Comb. Theory, Série A 54, 1990, pp. 53-62.
- [Ca] CARLITZ, L., "On abelian fields", Trans. Amer. Math. Soc. 35, 1933, pp. 122-136.
- [Ch] CHEN, W. Y., "Context-free grammars, differential operators and formal power series", Actes du Colloque sur les séries formelles et la combinatoire algébrique, 1991, pp.145-160.
- [Co] COMTET, Louis, "Nombres de Stirling généraux et fonctions symétriques", C. R. Acad. Sc. Paris, t. 275, 1972, pp. 747-750.

- [dM] DE MEDICIS, Anne, "Unimodalité et concavité logarithmique pour des suites de nombres classiques et leurs q -analogue", mémoire de maîtrise de l'UQAM, septembre 1989.
- [dML1] DE MEDICIS, A. et LEROUX, P., "A unified combinatorial approach for q -(and p, q -)Stirling numbers", *J. of Stat. Planning and Inference*, à paraître.
- [dML2] DE MEDICIS, A. et LEROUX, P., "Sur les nombres de Stirling généralisés", en préparation.
- [Hab] HABSIEGER, L., "Inégalités entre fonctions symétriques élémentaires: applications à des problèmes de log concavité", *Discrete Math.*, à paraître.
- [Le1] LEROUX, Pierre, "Catégories triangulaires: exemples, applications et problèmes", Rapport de recherche, Université du Québec à Montréal, juillet 1980, 72 p.
- [Le2] LEROUX, Pierre, "Reduced Matrices and q -log Concavity Properties of q -Stirling Numbers", *J. Comb. Theory*, Série A 54, 1990, pp. 64-84.
- [OH] O'HARA, Kathleen M., "Unimodality of Gaussian coefficients: a constructive proof", *J. Comb. Theory*, Série A 53, 1990, pp. 29-52.
- [ReWa] REMMEL, J.B. et M. WACHS, " p, q -analogues of the Stirling numbers", private communication, 1986.
- [Sa1] SAGAN, Bruce E., "Inductive and Injective Proofs of Log Concavity Results", *Discrete Mathematics* 68 (1988), pp. 281-292.
- [Sa2] SAGAN, Bruce, "Inductive Proofs of q -Log Concavity", en préparation, 1988, 19 p.
- [Sta] STANLEY, Richard P., "Log-Concave and Unimodal Sequences in Algebra, Combinatorics, and Geometry", *Annals of the New York Acad. of Sc.*, à paraître, 79 p.
- [TT] TOADER, Silvia et TOADER, G.H., "Generalized Stirling Numbers", *Studia Univ. Babeş-Bolyai Mathematica*, XXXIII, 1, 1988, pp. 50-53.
- [V1] VOIGT, Bernd, "A local limit theorem for generalized Stirling numbers", *Rev. Roumaine de Math. Pures*, App. 35, no 2, 1990, pp. 161-172.
- [V2] VOIGT, Bernd, "A common generalization of binomial coefficients, Stirling numbers and Gaussian coefficients", *Proc. 11th Winter School on Abstract Analysis (Železná Ruda, 1983)*. *Rend. Circ. Mat. Palermo*, Suppl. no 3, 1984, pp. 339-359.
- [Ze] ZEILBERGER, Doron, "A One-line High School Algebra Proof of the Unimodality of the Gaussian Polynomials $\begin{bmatrix} n \\ k \end{bmatrix}$ for $k < 20$ ", en préparation, février 1988, 6 p.

GROUPES DE MAGNUS: APPLICATIONS

GÉRARD DUCHAMP

*LITP/LIR †, Institut Blaise Pascal
Université Paris 7
2, place Jussieu, 75251 Paris Cedex 05*

Résumé: La transformation de Magnus classique $a \rightarrow 1+a$ a son image dans un groupe de séries formelles non commutatives remarquable: le groupe de Magnus. On étudie ici les propriétés des groupes associés aux monoïdes localement finis et particulièrement celles qui se transmettent à leurs sous groupes. Enfin, on détermine toutes les congruences qui sont compatibles avec la transformation de Magnus.

Abstract: Magnus' classical transformation, that is to say the substitution $a \rightarrow 1+a$ takes its values in a remarkable group: the Magnus group. We study here the properties of the groups associated with locally finite monoids, particularly those who hold for subgroups. We also characterize all the congruences being compatible with Magnus transformations.

1. Introduction

Magnus avait découvert [MKS][Bo1][CM] que la transformation $a \rightarrow 1 + a$ (qui est aussi la partie régulière à l'ordre 1 de l'exponentielle) permet d'identifier le groupe libre à un sous groupe de l'ensemble des séries formelles (non commutatives) de terme constant 1 (le groupe de Magnus). Si la notion de transformation de Magnus ne peut s'étendre, en caractéristique zéro, qu'au cas des commutations partielles (1), en revanche celle de groupe de Magnus se généralise très bien à tous les monoïdes localement finis.

On étudie ici les propriétés héréditaires de ces groupes (plus petite racine, suites centrales, ordonnabilité) en vue de les appliquer à des plongements du type de celui de Magnus. Enfin on montre que, en caractéristique zéro la transformation ne peut s'adapter telle quelle qu'aux commutations partielles.

† Ce travail est aussi supporté par le PRC de Math-Info.

(1) Le monoïde partiellement commutatif a été introduit par P.Cartier et D.Foata [CaFo] pour des raisons combinatoires et X.Viennot [Vi] en a donné une interprétation très suggestive en termes d'empilements.

2. Préliminaires

A) MONOÏDES ET GROUPES PRÉSENTÉS

Soit M un monoïde, une *congruence* sur M est une relation d'équivalence \equiv compatible avec sa loi [cf Eil]. Si A est un alphabet et $\mathbf{R} = (u_i, v_i)_{i \in I} \in (A^* \times A^*)^I$ est une famille de couples de mots de A^* on notera $\equiv_{\mathbf{R}}$ la congruence engendrée par \mathbf{R} , c'est à dire la congruence la plus fine telle que $(\forall i \in I)(u_i \equiv v_i)$. On définit alors $\langle A; \mathbf{R} \rangle$ (ou $\langle A; \mathbf{R} \rangle_{\text{Mon}}$ lorsqu'une confusion est à craindre), le monoïde *présenté par le relateur* \mathbf{R} , comme $A^*/\equiv_{\mathbf{R}}$. Il est immédiat que ce monoïde possède la propriété universelle suivante:

MP) Soit $\phi : A^* \rightarrow M$ (M monoïde quelconque) un morphisme tel que $(\forall i \in I)(\phi(u_i) = \phi(v_i))$ alors il existe une factorisation unique de ϕ à travers $\langle A; \mathbf{R} \rangle$ soit:

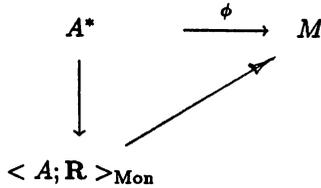


fig.1

Remarque 2.1: La congruence \equiv est dite *homogène* ssi $u \equiv v \implies |u| = |v|$. Il est immédiat que $\equiv_{\mathbf{R}}$ est homogène ssi $(\forall i \in I)(|u_i| = |v_i|)$. Que \equiv soit homogène entraîne aussitôt l'existence d'une fonction longueur à valeurs dans \mathbb{N} (c'est à dire d'un morphisme $\lambda : M \rightarrow \mathbb{N}$ tel que $\lambda^{-1}\{0\} = \{1\}$).

De même soit $F(A)$ le groupe libre engendré par un alphabet A et $\mathbf{R} = (u_i, v_i)_{i \in I} \in (F(A) \times F(A))^I$, une famille de couples dans $F(A)$, on notera $\langle A; \mathbf{R} \rangle$ (ou $\langle A; \mathbf{R} \rangle_{\text{Gr}}$ lorsqu'une confusion est à craindre), le groupe *présenté par le relateur* \mathbf{R} , comme $\langle A; \mathbf{R} \rangle = F(A)/H_{\mathbf{R}}$ où $H_{\mathbf{R}}$ désigne le sous groupe distingué engendré par la famille $(u_i v_i^{-1})_{i \in I}$. On a immédiatement une propriété similaire à (MP) qui est:

GP) Soit $\phi : F(A) \rightarrow G$ (G groupe quelconque) un morphisme tel que $(\forall i \in I)(\phi(u_i) = \phi(v_i))$ alors il existe une factorisation unique de ϕ à travers $\langle A; \mathbf{R} \rangle$ soit:

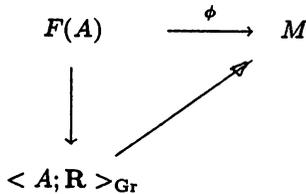


fig.2

Pour ces notions de structures présentées on pourra consulter [Bo2].

B) MONOÏDES LOCALEMENT FINIS

Si k est un anneau et M un monoïde. On peut définir la convolution de deux éléments de k^M par la formule:

$$f * g(w) = \sum_{uv=w} f(u)g(v)$$

cette opération n'est que partiellement définie et n'a de sens, en dehors d'une structure supplémentaire, que si la somme du deuxième membre est finie. Dans la suite on notera $k[[M]]$ au lieu de k^M , fg au lieu de $f * g$ et, pour un élément $f \in k[[M]]$, $(f|m)$ au lieu de $f(m)$, enfin $f = \sum_{m \in M} (f|m)m$ [DK1]. Les monoïdes localement finis sont la classe de monoïdes où l'opération de convolution est partout définie dans k^M et où tout élément de terme constant 1 est inversible (ci-dessous lemme 3.1).

Définition 2.1: [Eil] On dit qu'un monoïde M est *localement fini* ssi:

$$\bigcap_{n \geq 1} (M^+)^n = \emptyset$$

Remarques 2.2: 1) Un monoïde à longueur (à valeurs \mathbb{R}_+ cf rem.2.1.1) est localement fini si sa fonction longueur est bornée inférieurement sur une de ses parties génératrices. Cette condition est nécessaire si M est finiment engendré.

2) Un monoïde localement fini n'admet pas nécessairement de longueur comme le montre l'exemple du monoïde:

$$\langle a, b; a^2 = aba, b^2 = bab \rangle_{Mon}$$

3) En fait, M est localement fini ssi il admet une "fonction d'ordre" ω , à valeurs entières, *suradditive* (LF1) et *séparante* (LF2).

$$\text{LF1) } \omega(fg) \geq \omega(f) + \omega(g) \qquad \text{LF2) } \omega(f) = 0 \iff f = 1$$

4) Si $M = \langle A; \mathbf{R} \rangle$ avec A fini alors:

$$M \text{ est localement fini} \iff \text{les classes de } \equiv_{\mathbf{R}} \text{ sont finies}$$

3. Groupes de Magnus d'un monoïde localement fini

A) DÉFINITIONS

Lorsque M est localement fini, les unités de $k[[M]]$ sont d'une caractérisation comode.

Lemme 3.1. S est inversible dans $k[[M]]$ (pour la convolution) ssi $(S|1)$ est inversible dans k .

Preuve: Immédiat en utilisant $(1 - H)^{-1} = \sum_{n \geq 0} H^n$ \square

D'où la définition suivante:

Définition 3.1. Soit M un monoïde localement fini et k un anneau. On appelle *groupe de Magnus* (à coefficients dans k) associé à M le groupe:

$$1 + \mathfrak{M}_k[[M]] = \{S \in k[[M]] \mid (S|1) = 1\}$$

Remarques 3.1: 1) Avec les notations précédentes on a donc:

$$\mathfrak{M}_k[[M]] = \{S \in k[[M]] \mid (S|1) = 0\}$$

c'est un idéal bilatère de $k[[M]]$ (l'unique idéal maximal lorsque k est un corps).

2) Lorsque $k = \mathbb{Q}$ (ou, plus généralement, que k est une \mathbb{Q} -algèbre) les séries

$$\log(1 + H) = \sum_{n \geq 1} (-1)^{n+1} \frac{H^n}{n} \quad \text{et} \quad \exp(H) = \sum_{n \geq 0} \frac{H^n}{n!}$$

sont définies pour $H \in \mathfrak{M}_k[[M]]$. On a donc deux applications réciproques l'une de l'autre

$$\mathfrak{M}[[M]] \xrightleftharpoons[\log]{\exp} 1 + \mathfrak{M}_k[[M]]$$

qui vérifient, si $HK = KH$, les identités:

$$\log(1 + H)(1 + K) = \log(1 + H) + \log(1 + K) \quad \exp(HK) = \exp(H)\exp(K)$$

ceci permet de démontrer, dans le groupe de Magnus, des propriétés concernant les racines d'un élément.

B) PLUS PETITE RACINE

On définit, sur un monoïde M , les propriétés suivantes:

Définition 3.2. (RU) "racine unique": Pour tout $n \geq 1$, l'application $x \rightarrow x^n$ est injective.

(EB) "exposants bornés": Pour tout $g \in M^+ (= M - \{1\})$, l'ensemble

$$E(g) = \{n \in \mathbb{N} \mid (\exists r \in M) g = r^n\}$$

est borné dans \mathbb{N} .

Remarque 3.2:1) Si M vérifie (RU) ou (EB), il en est de même de chacun de ses sous-monoïdes .

2) (EB) et (RU) sont indépendantes dans la catégorie des groupes (et donc des monoïdes). Par exemple \mathbb{Q} vérifie (RU) mais pas (EB) et $\langle a, b; a^2 = b^2 \rangle_{Gr}$ vérifie (EB) (cf preuve de Prop.4.3, la forme normale vaut aussi dans G , avec $n \in \mathbb{Z}$) mais pas (RU), en effet on a $(aba^{-1})^2 = b^2$ bien que $aba^{-1} \neq b$.

Proposition 3.1. Soit M un monoïde localement fini alors:

- 1) Si k est une \mathbb{Q} -algèbre $1 + \mathfrak{M}_k[[M]]$ vérifie (RU)
- 2) $1 + \mathfrak{M}_{\mathbb{Z}}[[M]]$ vérifie (EB)

Preuve: 1) Immédiat par application du logarithme.

2) Soit ω , la fonction d'ordre sur M (cf rem.2.2.3). Pour tout $g \neq 1$ dans $1 + \mathfrak{M}_{\mathbb{Z}}[[M]]$ notons $n_g = \inf\{\omega(m) \mid m \in M^+ \cap \text{supp}(g)\}$ on a:

$$g = 1 + \sum_{\omega(m)=n_g} (g|m)m + \sum_{\omega(m)>n_g} (g|m)m = 1 + g' + g''$$

et, de même, si $r^n = g$ on a $r \neq 1$ et on écrit $r = 1 + r' + r''$. Alors $g' = nr'$ soit n divise les coefficients $(g|m)$ pour tout $m \in \text{supp}(g')$, ce qui prouve 2). \square

Remarques 3.3: 1) Le point 1) de la proposition précédente ne vaut plus en caractéristique quelconque même pour un monoïde à longueur comme le montre le cas de:

$$k = \mathbb{Z}/p\mathbb{Z} \quad \text{et} \quad M = \langle a, b; a^p b = b a^p \rangle_{Mon}$$

pour p premier.

2) On peut montrer que, dans le cas d'un groupe nilpotent G :

$$G \text{ vérifie (RU)} \iff G \text{ est sans torsion}$$

Ce n'est évidemment pas le cas pour un groupe quelconque puisque le groupe $1 + \mathfrak{M}_k[[M]]$ de l'exemple précédent est sans torsion.

3) $1 + \mathfrak{M}_{\mathbb{Z}}[[M]]$ est un sous-groupe de $1 + \mathfrak{M}_{\mathbb{Q}}[[M]]$ et vérifie donc aussi (RU).

(RU) impose une certaine "rigidité" aux sous groupes cycliques dont fait partie un élément $g \neq 1$ donné. Plus précisément:

Proposition 3.2. Soit G un groupe et, pour tout $g \in G - \{1\}$, soit $\mathcal{H}(g)$, la réunion des sous-groupes cycliques qui contiennent g . Alors:

1) (RU) \iff pour tout $g \neq 1$, $\mathcal{H}(g)$ est un sous-groupe de G isomorphe à un sous-groupe de \mathbb{Q}

2) (RU)+(EB) équivaut à:

$$\text{pour tout } g \neq 1, \mathcal{H}(g) \text{ est isomorphe à } \mathbb{Z}.$$

Preuve: 1) \Rightarrow) Nous nous plaçons sous l'hypothèse (RU).

$E(g) = \{n \in \mathbb{N} | (\exists r \in M) / g = r^n\}$ est alors stable par PGCD. Soit $g \neq 1$ un élément fixé et Γ le plus petit sous-groupe de $\mathbb{Q} \times G$ stable par extraction de racine ($x^n \in \Gamma \Rightarrow x \in \Gamma$) et qui contienne $(1, g)$. On vérifie facilement que $\Gamma = \{(p/q, r) | g^p = r^q\}$ et que Γ est le graphe d'un isomorphisme entre $[\mathbb{Z} : E(g)] = \{p/q\}_{p \in \mathbb{Z}, q \in E(g)}$, sous-groupe de \mathbb{Q} , et $\mathcal{H}(g)$.

\Leftarrow) La réciproque est pratiquement immédiate: si $g = u^n = v^n$ alors u et v commutent. Si on avait $u \neq v$, $\mathcal{H}(uv^{-1})$ serait non trivial, fini et isomorphe à un sous-groupe de \mathbb{Q} ce qui est évidemment impossible.

2) \Rightarrow) (EB) entraîne que $E(g)$ admet un plus grand élément q_0 que divisent tous les autres alors $\mathcal{H}(g) \cong [\mathbb{Z} : E(g)] = \{p/q_0\}_{p \in \mathbb{Z}} \cong \mathbb{Z}$.

\Leftarrow) Si $\mathcal{H}(g) \cong \mathbb{Z}$ on a évidemment (RU). D'autre part, soit r un des deux générateurs de $\mathcal{H}(g)$, on peut écrire $g = r^n$ ($n \in \mathbb{Z} - \{0\}$), une vérification immédiate montre que $E(g)$ est l'ensemble $\{d \in \mathbb{N} | d|n\}$ \square

Définition 3.3. Dans le cas 2) ci-dessus l'unique générateur r de $\mathcal{H}(g)$ tel que $g = r^n$ avec $n \geq 0$, s'appelle la *plus petite racine* de g et est notée \sqrt{g} .

Remarques 3.4: 1) Tout sous-groupe G d'un groupe de Magnus $1 + \mathfrak{M}_{\mathbb{Z}}[[M]]$ vérifie (RU)+(EB) il y a donc existence, sur G , d'une fonction $\sqrt{}$.

2) Plus généralement on dira qu'un élément g d'un semigroupe S admet r pour plus petite racine avec l'exposant e ssi

$$g = r^e \quad (e \in \mathbb{N}^*) \quad \text{et} \quad (\forall x \in S)(x^n = g \Rightarrow n|e \quad \text{et} \quad x = r^{e/n})$$

on voit alors que r et e sont uniques. On pose alors $r = \sqrt{g}$ et $e = e(g)$ (cf [Lot] pour $S = A^*$ et [DK3] pour le monoïde et le groupe partiellement commutatifs libres).

c) SUITES CENTRALES

Sur tout groupe G , on peut définir une deuxième loi, le *commutateur* de deux éléments par:

$$(u, v) = uvu^{-1}v^{-1}$$

On étend cette loi aux sous-groupes de G en posant, pour deux sous-groupes H et K que (H, K) est le sous-groupe engendré par les commutateurs $(u, v)_{(u,v) \in H \times K}$. On dira qu'une suite décroissante $(G_n)_{n \geq 1}$ de sous-groupes de G est *centrale* (N -suite dans [Laz]) si:

$$G_1 = G \quad \text{et} \quad (G_n, G_m) \subseteq G_{n+m}$$

(on pourra également consulter [Bo1] pour un concept un peu plus général dont nous n'aurons pas besoin ici).

L'intérêt des suites centrales est qu'elles offrent un véritable "calcul des commutateurs" permettant de contrôler le "défaut de commutation" des éléments. Ceci permet, en particulier, de linéariser la situation en associant une algèbre de Lie à $(G_n)_{n \geq 1}$. Il est à noter que ce calcul est d'autant plus aisé (et d'autant plus fin) que les éléments du groupe se prêtent à toutes les facilités des "développements limités" en étant justement des séries, c'est bien souvent le cas lorsque l'on plonge un groupe dans un groupe de Magnus approprié, les suites centrales étant alors des suites de sous-groupes de "dimension", c'est à dire induites sur un sous-groupe par $(1 + \mathfrak{M}_k^n[[M]])_{n \geq 1}$ dont nous allons parler maintenant.

Proposition 3.3. Soit k un anneau et M un monoïde localement fini. Alors:

- 1) $\mathfrak{M}_k^n[[M]] = \{S \in k[[M]] \mid \omega(m) < n \Rightarrow (S|m) = 0\}$
- 2) La suite $(1 + \mathfrak{M}_k^n[[M]])_{n \geq 1}$ est centrale
- 3) $1 + \mathfrak{M}_k^n[[M]] / 1 + \mathfrak{M}_k^{n+1}[[M]] \cong \text{lin}_k(M^{+n} - M^{+(n+1)})$ où $\text{lin}_k(X)$ désigne le module librement engendré par X .

Preuve: 1) Récurrence sur n .

2) Soient $H \in \mathfrak{M}_k^p[[M]]$, $K \in \mathfrak{M}_k^q[[M]]$ on a:

$$(1 + H)(1 + K)(1 + H)^{-1} = 1 + K + [H, K] + (1 + H) \sum_{n \geq 2} (-H)^n = 1 + K + [H, K] + R$$

avec $R \in \mathfrak{M}_k^{p+q+1}[[M]]$ d'où

$$(1 + H, 1 + K) = (1 + K + [H, K] + R(1 + K))^{-1} = 1 + [H, K](1 + K)^{-1} + R(1 + K)^{-1}$$

ce qui démontre le résultat.

3) Vient de l'application $\phi : 1 + \mathfrak{M}_k^n[[M]] \rightarrow \text{lin}_k(M^{+n} - M^{+(n+1)})$ définie par $\phi(1 + H) = \sum_{\omega(m)=n} (H|m)m$ dont on vérifie immédiatement qu'elle est surjective et de noyau $1 + \mathfrak{M}_k^{n+1}[[M]]$.

Remarques 3.5: 1) La filtration centrale $(1 + \mathfrak{M}_k^n[[M]])_{n \geq 1}$ n'est pas en général la suite centrale descendante de $1 + \mathfrak{M}_k[[M]]$ même quand $M = A^*$ [Bo1].

2) Certains résultats peuvent se généraliser (sous certaines conditions) à un idéal \mathfrak{M} d'une algèbre \mathcal{A} tel que $\bigcap_{n \geq 1} \mathfrak{M}^n = \{0\}$, mais on a plus nécessairement les propriétés de plus petite racine et seulement $1 + \mathfrak{M}^n / 1 + \mathfrak{M}^{n+1} \cong \mathfrak{M}^n / \mathfrak{M}^{n+1}$.

3) Soit p un nombre premier. Sur un groupe G on définit la topologie p -adique [Hall] par ses voisinages de 1 qui sont les sous-groupes distingués H tels que G/H soit un p -groupe fini. Si G' est un sous-groupe de G et que la topologie p -adique est séparée sur G , il en est de même de la topologie p -adique sur G' . La proposition 3.3.3) montre que lorsque $k = \mathbb{Z}$ ou $\mathbb{Z}/p^n\mathbb{Z}$ la topologie p -adique est séparée dans $1 + \mathfrak{M}_k[[M]]$.

D) ORDONNABILITÉ

Une autre propriété que possèdent certains groupes de Magnus et transmissible aux sous-groupes est l'ordonnabilité.

Définition 3.4. On dira qu'un monoïde M est *ordonné* par une relation d'ordre $<$ si cette relation est compatible avec la loi de M c'est à dire vérifie identiquement:

$$x < y \text{ et } z < t \Rightarrow xz < yt$$

Supposons que M soit muni d'une relation de bon ordre $<$ compatible avec sa structure. Pour $H \in k[[M]] - \{0\}$ posons $\nu(H) = \min(\text{supp}(H))$. On peut alors ordonner le groupe de Magnus à coefficients dans \mathbb{Q} .

Proposition 3.4. Avec les notations ci-dessus la relation d'ordre strict sur $1 + \mathfrak{M}_{\mathbb{Q}}[[M]]$ définie par:

$$1 + H < 1 + K \stackrel{\text{déf}}{\iff} \nu(H) > \nu(K) \text{ ou } (\nu(H) = \nu(K) \text{ et } (H|\nu(H)) < (K|\nu(K)))$$

est totale et compatible avec la structure de groupe de $1 + \mathfrak{M}_{\mathbb{Q}}[[M]]$ [Bo3][Mu].

Preuve: On observe facilement qu'un ordre sur un groupe G est entièrement caractérisé par le *cône positif* c'est à dire l'ensemble

$$P = \{g \in G | g > 1\}$$

(cf [Bo3][Mu]). Il suffit alors de vérifier les propriétés suivantes de P :

$$(OTG) \quad PP \subseteq P; \quad (\forall g \in G)(gPg^{-1} \subseteq P) \quad ; \quad P \cap P^{-1} = \emptyset; \quad P \cup P^{-1} = G - \{1\}$$

On définit alors la partie:

$$P = \{1 + H \in 1 + \mathfrak{M}_{\mathbb{Q}}[[M]] | (H|\nu(H)) > 0\}$$

et on vérifie ensuite facilement les propriétés (OTG) ci-dessus et que $(1 + H)^{-1}(1 + K) \in P \Leftrightarrow 1 + H < 1 + K$ ce qui achève la preuve. \square

Remarques 3.6: 1) On peut évidemment remplacer \mathbb{Q} par un anneau totalement ordonné.

2) Un groupe (totalement) ordonnable vérifie RU) car $x < y \Rightarrow x^n < y^n$ pour tout $n \geq 1$.

3) Dans la construction précédente on peut remplacer le fait que M soit bien ordonné par la condition plus faible que M soit totalement ordonné mais il faut alors se restreindre aux séries à support bien ordonné $1 + \mathfrak{M}_{\mathbb{Q}}^{WO}[[M]]$ ("séries de Malcev") dont on montre aussitôt qu'il forme un groupe grâce aux lemmes de Malcev (cf [Neu]).

Le fait qu'un groupe G soit totalement ordonnable présente un intérêt majeur: $\mathbb{Q}[G]$ (ou $k[G]$ si k est un corps ordonné) peut alors être plongé dans un corps gauche, ce qui permet d'appliquer les techniques de l'algèbre linéaire. Dans la même année 1990, cette technique a été employée deux fois à des questions de décidabilité [HK][Va].

4. Applications

A) PLONGEMENT DU GROUPE PARTIELLEMENT COMMUTATIF

Soit A un alphabet sur lequel on donne une relation ϑ irréflexive et symétrique ($\vartheta \subseteq A^2 - \Delta_A, \vartheta = \vartheta^{-1}$). Le couple (A, ϑ) sera appelé *alphabet à commutations* (cf [DK2]). On considère le monoïde et le groupe partiellement commutatifs libres:

$$M(A, \vartheta) := \langle A; (ab = ba)_{(a,b) \in \vartheta} \rangle_{Mon}$$

$$F(A, \vartheta) := \langle A; (ab = ba)_{(a,b) \in \vartheta} \rangle_{Gr}$$

Soit k un anneau (non trivial $1 \neq 0$). L'application $\mu_k^{alph} : A \rightarrow 1 + \mathfrak{M}_k[[M(A, \vartheta)]]$ définie par $\mu_k^{alph}(a) = 1 + a$ se prolonge d'une seule façon en un morphisme de groupe à $F(A, \vartheta)$ (par compatibilité avec les relateurs de $F(A, \vartheta)$).

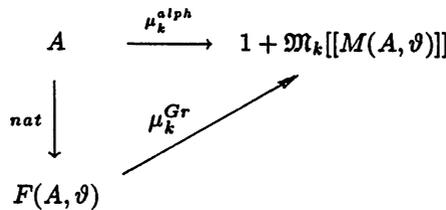


fig.3

Le théorème suivant établit que μ_k^{Gr} est un plongement:

Théorème 4.1: μ_k^{Gr} est injective.

Description de la preuve: Les détails se trouvent dans [DK3] elle repose sur le fait que tout élément de $F(A, \vartheta)$ admet une représentation sous *forme exponentielle réduite* (F.E.R.) c'est à dire $g = a_1^{\alpha_1} a_2^{\alpha_2} \dots a_n^{\alpha_n}$ ($a_i \in A$ et $\alpha_i \in \mathbb{Z}$) unique aux commutations autorisées près quand n est minimal c'est à dire que si on a aussi $g = b_1^{\beta_1} b_2^{\beta_2} \dots b_n^{\beta_n}$ avec n minimal, les deux formes exponentielles réduites s'échangent par une chaîne d'opérations élémentaires du type:

$$a_1^{\alpha_1} \dots a_i^{\alpha_i} a_{i+1}^{\alpha_{i+1}} \dots a_n^{\alpha_n} \longrightarrow a_1^{\alpha_1} \dots a_{i+1}^{\alpha_{i+1}} a_i^{\alpha_i} \dots a_n^{\alpha_n} \quad (Op)$$

avec $(a_i, a_{i+1}) \in \vartheta$

Commentaire: 1) On aura besoin plus bas de la notion d'*alphabet initial (et terminal) sans exposant* qui sont respectivement les ensembles

$$AISE(g) = \{a \in A \mid g \text{ admet une F.E.R. de la forme } a^\alpha a_2^{\alpha_2} \dots a_n^{\alpha_n}\}$$

et

$$ATSE(g) = \{a \in A \mid g \text{ admet une F.E.R. de la forme } a_1^{\alpha_1} a_2^{\alpha_2} \dots a_n^{\alpha_n}\}.$$

2) $F(A, \vartheta)$ peut donc se concevoir comme un sous-groupe des groupes $1 + \mathfrak{M}_k[[M(A, \vartheta)]]$ et en hérite donc certaines propriétés que nous allons maintenant préciser dans chaque cas.

B) PLUS PETITE RACINE

Le plongement $F(A, \vartheta) \hookrightarrow 1 + \mathfrak{M}_k[[M(A, \vartheta)]]$ assure l'existence d'une fonction *plus petite racine* (cf [DK3]) à l'intérieur de $F(A, \vartheta)$. La *forme exponentielle réduite* se prête bien au calcul effectif de cette fonction (mais mal à la démonstration de son existence).

Soit $g \in F(A, \vartheta)$, supposons g donné sous forme d'un produit $g = a_1^{\varepsilon_1} a_2^{\varepsilon_2} \dots a_n^{\varepsilon_n}$ avec $a_i \in A$ et $\varepsilon_i \in \{-1, 1\}$. La forme exponentielle réduite $g = a_1^{\alpha_1} a_2^{\alpha_2} \dots a_n^{\alpha_n}$ se calcule en temps linéaire avec un algorithme évident de regroupement des facteurs. Soit $r = \sqrt{g}$, r a un conjugué h dont la forme exponentielle réduite peut se mettre (avec les opérations élémentaires "Op" comme dans la preuve de Th.4.1) sous la forme $b_1^{\beta_1} b_2^{\beta_2} \dots b_s^{\beta_s} c_1^{\gamma_1} c_2^{\gamma_2} \dots c_t^{\gamma_t}$ tel que $AISE(b_1^{\beta_1} b_2^{\beta_2} \dots b_s^{\beta_s}) \cap ATSE(b_1^{\beta_1} b_2^{\beta_2} \dots b_s^{\beta_s}) = \emptyset$ et que les lettres c_i commutent avec toutes les autres lettres de h . Il est immédiat qu'alors la F.E.R. de h^k est $(b_1^{\beta_1} b_2^{\beta_2} \dots b_s^{\beta_s})^k c_1^{k\gamma_1} c_2^{k\gamma_2} \dots c_t^{k\gamma_t}$. Ceci fournit un algorithme commode de calcul de la plus petite racine de g . On met d'abord g sous la forme:

$$g = W a_1^{\alpha_1} a_2^{\alpha_2} \dots a_s^{\alpha_s} c_1^{\gamma_1} c_2^{\gamma_2} \dots c_t^{\gamma_t} W^{-1} = WUCW^{-1}$$

Ce qui vient d'être dit montre qu'il existe un représentant (équivalent par application des opérations élémentaires "Op" ci-dessus) de $a_1^{\alpha_1} a_2^{\alpha_2} \dots a_s^{\alpha_s}$ qui est périodique et \sqrt{U} est alors sa période. Ensuite $c_1^{\gamma_1} c_2^{\gamma_2} \dots c_t^{\gamma_t}$ étant complètement commutatif, sa racine est $c_1^{\gamma_1/d} c_2^{\gamma_2/d} \dots c_t^{\gamma_t/d}$ où $d = \gcd(\gamma_1, \gamma_2, \dots, \gamma_t)$. Comme \sqrt{U} et \sqrt{C} commutent on a $\sqrt{g} = W \sqrt{U}^{e(U)/k} \sqrt{C}^{e(C)/k} W^{-1}$ avec $k = \gcd(e(U), e(C))$.

B) CONSTRUCTION D'UN ORDRE DÉCIDABLE SUR $F(A, \vartheta)$

On sait que les monoïdes libres sont totalement ordonnables (par un ordre lexicographique par exemple). R.Cori et D.Perrin [CP] ont donné une manière très simple de plonger un monoïde partiellement commutatif dans un produit de monoïdes libres. Cette construction a été reprise par C.Duboc [Dub] sous une forme quelque peu différente que nous reproduisons ici.

Soit (A, ϑ) , un alphabet à commutation et $\vartheta^c = A^2 - \vartheta$, son graphe de non-commutation. Une famille $(C_i)_{i \in I}$ est dite *recouvrir* le graphe ϑ^c si elle en recouvre toute arête. On considère alors les morphismes de projection π_{C_i} définis par les formules:

$$\pi_{C_i}(x) = x \text{ si } x \in C_i \text{ et } \pi_{C_i}(x) = 1 \text{ sinon}$$

L'application $\Phi : M(A, \vartheta) \longrightarrow \prod_{i \in I} C_i^*$ définie par $\phi(m) = (\pi_{C_i}(m))_{i \in I}$ est un morphisme injectif [Dub].

On construit alors un bon ordre sur M par le morphisme $\psi : M(A, \vartheta) \hookrightarrow \mathbb{N} \times \prod_{i \in I} C_i^*$ que définit la formule $\psi(m) = (|m|, \phi(m))$ en prenant l'image réciproque du produit lexicographique des monoïdes ordonnés avec priorité sur le premier facteur. Il est facile de voir que si $a_1^{\alpha_1} a_2^{\alpha_2} \dots a_n^{\alpha_n}$ est la forme exponentielle réduite de $g \in F(A, \vartheta)$ on peut décider si g est dans P en faisant le développement limité de $\mu_{\mathbb{Z}}(g)$ à l'ordre n seulement. Les détails de cette construction peuvent être trouvés dans [DT].

C) SUITES CENTRALES DE $F(A, \vartheta)$

Les sous-groupes de la suite centrale descendante de $F(A, \vartheta)$ sont des sous-groupes de "dimension". Plus précisément

Proposition 4.2: Soit (A, ϑ) , un alphabet à commutations. En identifiant $F(A, \vartheta)$ à son image (par $\mu_{\mathbb{Z}}$) dans $1 + \mathfrak{M}_{\mathbb{Z}}[[M(A, \vartheta)]]$ on a:

- 1) $F(A, \vartheta)_n = F(A, \vartheta) \cap 1 + \mathfrak{M}_{\mathbb{Z}}^n[[M(A, \vartheta)]]$.
- 2) Les topologies p -adiques et profinie sont séparées dans $F(A, \vartheta)$

Preuve:1) Les détails sont dans [DK5]. Elle repose sur ce que l'algèbre de Lie associée à la suite $(F(A, \vartheta)_n)_{n \geq 1}$ est $L_{\mathbb{Z}}(A, \vartheta) = \langle A; ([a, b] = 0)_{(a,b) \in \vartheta} \rangle$.

2) Il résulte immédiatement du plongement $(F(A, \vartheta) \hookrightarrow 1 + \mathfrak{M}_{\mathbb{Z}/p\mathbb{Z}}[[M(A, \vartheta)]]$. Pour la topologie profinie on remarque qu'elle est plus fine que n'importe quelle topologie p -adique \square

D) UN AUTRE EXEMPLE DE PLONGEMENT

Nous verrons tout à l'heure que, en caractéristique zéro, les seules congruences compatibles avec la transformation de Magnus sont, à une section de l'alphabet près, les commutations partielles. Dans les autres caractéristiques, par contre, on peut concevoir d'autres relateurs. En voici un exemple.

Soit $G = \langle a, b; a^2 = b^2 \rangle_{Gr}$ et $M = \langle a, b; a^2 = b^2 \rangle_{Mon}$, il est immédiat que M est un monoïde à longueur pour le quotient de la longueur canonique sur $\{a, b\}^*$. M est donc localement fini. Il résulte de la compatibilité de la transformation de Magnus avec la présentation de G et qu'il existe un unique morphisme de groupes $\mu_G : G \rightarrow 1 + \mathfrak{M}_{\mathbb{Z}/2\mathbb{Z}}[[M]]$ tel que $\mu_G(x) = 1 + x$ pour $x \in \{a, b\}$.

Proposition 4.3: μ_G est injective

Preuve: Il est immédiat que tout élément de G peut s'écrire $g = a^n (ba)^k b^\epsilon$ avec $n \in \mathbb{Z}, k \in \mathbb{N}, \epsilon \in \{0, 1\}$. La représentation étant unique dans M avec $n \in \mathbb{N}$ car c'est la forme normale lexicographique (i.e le plus petit élément de la classe de congruence pour l'ordre lexicographique standard). Ecrivons $n = 2^{\nu_2(n)} n_1$ avec $n_1 \not\equiv 0[2]$ et de même $k = 2^{\nu_2(k)} k_1$ avec $k_1 \not\equiv 0[2]$. Soit $g = a^n (ba)^k b^\epsilon \neq 1$ on a alors $\mu_G(a^n (ba)^k b^\epsilon) = (1 + a^{2^{\nu_2(n)}})^{n_1} (1 + b^{2^{\nu_2(k)}} + a^{2^{\nu_2(k)}} + (ba)^{2^{\nu_2(k)}})^{k_1} (1 + b)^\epsilon$ dans le développement duquel on a immédiatement que le monôme $a^{2^{\nu_2(n)}} (ba)^{2^{\nu_2(k)}} b^\epsilon$ est obtenu une seule fois avec le coefficient 1 et, par suite, $\mu_G(g) \neq 1$. \square

Corollaire 4.4: Le groupe G est résiduellement nilpotent. Les topologies 2-adique et profinie sont séparées sur G .

Preuve: Elle est immédiate grâce au plongement précédent et à ce qui a été dit en remarque 3.5.3. \square

5. Congruences compatibles avec la transformation de Magnus

On l'a vu, la transformation de Magnus est un outil extrêmement commode pour calculer dans certains groupes présentés. L'exemple précédent montre que certains relateurs (ici $a^2 = b^2$) sont compatibles avec cette transformation pour des caractéristiques convenables. Cependant, nous allons le voir, seules les commutations partielles sont compatibles en caractéristique zéro, ceci nous permettra d'en déduire le théorème de limitation [DK4]. Mais formalisons d'abord le problème.

Soit \equiv une congruence sur A^* et k un anneau, le morphisme $\pi_{\equiv} : A^* \rightarrow A^*/\equiv$ s'étend par linéarité aux algèbres en $\pi_{\equiv}^k : k \langle A \rangle \rightarrow k[A^*/\equiv]$. La relation d'équivalence définie par l'application π_{\equiv}^k , prolongeant \equiv on notera $P \equiv Q$ pour $\pi_{\equiv}^k(P) = \pi_{\equiv}^k(Q)$.

Définition 5.1: On dira qu'une congruence est compatible avec la transformation de Magnus à coefficients dans k ssi:

$$u \equiv v \implies \mu_k(u) \equiv \mu_k(v)$$

Où $\mu_k : A^* \rightarrow k \langle A \rangle$ est le morphisme défini par $\mu_k(a) = 1 + a$ pour $a \in A$.

Ceci équivaut à l'existence d'une flèche π_{\equiv} rendant commutatif le diagramme suivant:

$$\begin{array}{ccc} A^* & \xrightarrow{\mu_k} & k \langle A \rangle \\ \pi_{\equiv} \downarrow & & \downarrow \pi_{\equiv}^k \\ A^*/\equiv & \xrightarrow{\mu_{\equiv}^k} & k[A^*/\equiv] \end{array}$$

fig.4

Nous pouvons alors énoncer:

Proposition 5.1: Si k est de caractéristique zéro, une congruence (sur A^*) est compatible avec la transformation de Magnus ssi elle est engendrée par des relateurs du type $(a_i = b_i)$ et $(a_j b_j = b_j a_j)$ ($a_i, b_i, a_j, b_j \in A$). En particulier le monoïde A^*/\equiv est un monoïde partiellement commutatif.

Preuve: Remarquons d'abord qu'on se ramène aussitôt à $k = \mathbb{Z}$ puisque les images de la transformation de Magnus sont des polynômes à coefficients dans $\mathbb{Z} \cdot 1_k$. Montrons alors le sens: compatible \rightarrow relateurs du type $(a_i = b_i)$ et $(a_j b_j = b_j a_j)$; (la réciproque étant une simple vérification) et plaçons nous sous l'hypothèse où \equiv est compatible avec la transformation de Magnus.

La démonstration comprend 6 étapes.

1): La congruence \equiv est homogène.

En effet, si $u \equiv v$ alors $\mu_k(u) \equiv \mu_k(v)$ et donc $\pi_{\equiv}^k(\mu_k(u)) = \pi_{\equiv}^k(\mu_k(v))$ et, en spécialisant tous les mots de A^*/\equiv en $1_{\mathbb{Z}}$, $2^{|u|} = 2^{|v|}$ c'est à dire $|u| = |v|$.

La congruence étant homogène on notera \equiv^n la relation d'équivalence induite sur A^n et $\equiv^{\leq n}$ la congruence engendrée par les relations $u \equiv v$ pour $|u|, |v| \leq n$.

2): Réduction du problème

Considérons un sous-alphabet $S \subseteq A$ qui soit une section de la relation d'équivalence \equiv_1 on a alors un isomorphisme naturel $S^*/\equiv_S \rightarrow A^*/\equiv$ (\equiv_S désigne la congruence induite sur S^*) et on vérifie facilement que \equiv_S est aussi compatible avec la transformation de Magnus. Le résultat sera donc prouvé si on démontre que \equiv_S est une congruence partiellement commutative.

3): \equiv_S est multihomogène.

En effet $\mu_Z(u) = 1 + \sum_{a \in S} |u|_a a + \sum \text{termes de longueur} \geq 2$ où $|u|_a$ désigne le degré partiel en la lettre a . La comparaison avec $\mu_Z(v)$ montre que $\sum_{a \in S} |u|_a a = \sum_{a \in S} |v|_a a$ ce qui, en raison de l'indépendance linéaire des lettres entraîne, la propriété:

$$u \equiv_S v \implies (\forall a \in A)(|u|_a = |v|_a)$$

une telle congruence est dite *multihomogène*.

Dans la suite on notera \equiv au lieu de \equiv_S .

4): \equiv est simplifiable.

En fait, montrons la propriété plus forte que S^*/\equiv se plonge dans son groupe de Magnus. De la commutativité du diagramme (fig.4) on déduit un morphisme $\mu : S^*/\equiv \rightarrow 1 + \mathcal{M}_Z[[S^*/\equiv]]$ dont a une section facilement en considérant l'unique terme de degré dominant dans les transformés des éléments de S^*/\equiv .

5): Contradiction

Il suffit maintenant de montrer que $\equiv \equiv \leq^2$. Supposons que ce ne soit pas le cas, on aurait alors un contreexemple de longueur minimale $u \equiv v$ tel que $u \not\equiv \leq^2 v$ et donc $|u| \geq 3$. On peut écrire $u = u_1 a_n, a_n \in A$ et $v = v_1 a_n w$ avec $a_n \notin \text{Alph}(w)$. On ne peut avoir $w = 1$ sans quoi, comme \equiv est simplifiable, ceci entraînerait $u_1 \equiv v_1$ puis $u_1 \equiv \leq^2 v_1$ en raison de la minimalité du contreexemple et enfin $u \equiv \leq^2 v$, contredisant l'hypothèse. En (6) on va montrer qu'on peut toujours se ramener à ce cas, ce qui entraînera l'impossibilité d'un contreexemple minimal, démontrant ainsi la proposition. 6) a_n commute avec w .

Posons $d = |u|_{a_n} = |v|_{a_n}$ et $w = w_1 w_2 \dots w_k$. Pour toute lettre w_i , le monôme $a_n^d w_i$ qui figure dans $\mu(v)$ figure aussi dans $\mu(u)$. On a donc, en raison de la multihomogénéité, $a_n^d w_i \equiv a_n^{d-j} w_i a_n^j$ et $j \geq 1$ à cause de la forme de w . Ceci entraîne $w_i a_n^j \equiv a_n^j w_i$. En prenant les logarithmes des transformés dans $1 + \mathcal{M}_Q[[S^*/\equiv]]$, on en déduit que $w_i a_n \equiv a_n w_i$ \square

La proposition précédente a un corollaire important, le *théorème de limitation* [DK4] qui montre que les seuls relateurs qui soient à la fois des relateurs monoïdaux (du type $u = v$ avec $u, v \in A^*$) et de type de Lie (du type $P = 0$ avec $P \in L(A)$) sont des commutations partielles. Pour préciser le problème donnons deux définitions.

Définition 5.2: [DK4] Soit k un anneau, A un alphabet, et \mathcal{J} un idéal de $k \langle A \rangle$. On dira que \mathcal{J} est *monoïdal* s'il est engendré par une famille $(u_i - v_i)$ avec $u_i, v_i \in A^*$. On dira que \mathcal{J} est *de type de Lie* s'il est engendré par une famille P_j avec $P_j \in L_k(A)$.

Théorème 5.2:[DK4] On suppose k de caractéristique 0. Alors, si \mathcal{J} est à la fois *monoïdal* et *de type de Lie*, \mathcal{J} est engendré par une famille de différences du type $a_i - b_i$ et $a_j b_j - b_j a_j$ ($a_i, a_j, b_i, b_j \in A$). En particulier l'algèbre $k \langle A \rangle / \mathcal{J}$ est une algèbre de polynômes partiellement commutatifs.

Preuve: Comme \mathcal{J} est de type de Lie le coproduit classique $c : k \langle A \rangle \longrightarrow k \langle A \rangle \otimes k \langle A \rangle$ défini sur les lettres par $c(x) = x \otimes 1 + 1 \otimes x$ passe aux quotients.

Or \mathcal{J} est monoidal et donc engendré par une famille $R = (u_i - v_i)$ de différences de mots. Soit \equiv la congruence engendrée par R . On sait [Thi] que le morphisme naturel $k[A^*/\equiv] \longrightarrow k \langle A \rangle / \mathcal{J}$ est un isomorphisme. Il y a donc un coproduit de bigèbre $c_{\equiv} : k[A^*/\equiv] \longrightarrow k[A^*/\equiv] \otimes k[A^*/\equiv]$ tel que $c_{\equiv}(x) = x \otimes 1 + 1 \otimes x$ pour $x \in A$. On obtient alors la transformation de Magnus de $\langle A; R \rangle_{\text{Mon}} = A^*/\equiv$ à coefficients dans k , en considérant la composition $\alpha \circ c_{\equiv}$ où α est l'application $\alpha : k[A^*/\equiv] \otimes k[A^*/\equiv] \longrightarrow k[A^*/\equiv]$ définie par $\alpha(u \otimes v) = v$ pour $u, v \in A^*/\equiv$. La congruence \equiv est donc compatible avec la transformation de Magnus, ce qui, en caractéristique zéro, entraîne le résultat. \square

Conclusion:

On peut construire d'autres plongements à partir de la transformation de Magnus. Par exemple, en prenant $k = \mathbb{Q}$, on peut diviser indéfiniment les générateurs (équivalent partiellement commutatif du groupe des monômes de Puiseux) ce groupe peut se réaliser à partir des formes exponentielles réduites en considérant des puissances fractionnaires.

D'autre part, en prenant $k = \mathbb{Z}/p\mathbb{Z}$ on peut plonger le groupe de présentation $\langle A; (a^p = 1)_{a \in T}, (ab = ba)_{(a,b) \in \vartheta} \rangle$ (où T est un sous-alphabet de A) dans les unités de $k \langle\langle A, \vartheta \rangle\rangle / J$ où J est un idéal bilatère qui se décrit aisément (voir un premier travail dans [An]). Enfin, on peut remarquer que, dans le cas du groupe partiellement commutatif libre, les images de la transformation de Magnus sont des séries toutes reconnaissables, il serait donc à faire une étude précieuse du rang de $\mu(g)$ en fonction de g .

BIBLIOGRAPHIE

- [An] ANDREAKAKIS S. *On an embedding of a free product of cyclic p -groups*, Bull. Soc. Math. Grèce 10, 19-34 (1969).
- [Bo1] BOURBAKI N., *Groupes et algèbres de Lie* ch.2 et 3, CCLS 1972.
- [Bo2] BOURBAKI N., *Algèbre* ch 1 à 3 HERMANN 1970.
- [Bo3] BOURBAKI N., *Algèbre* ch 6.1 Ex.1 MASSON 1981.
- [CaFo] CARTIER P., FOATA D., *Problèmes combinatoires de permutations et réarrangements*, L.N. in Math., vol 85, Springer Verlag, Berlin 1969.
- [CM] CHANDLER B., MAGNUS W., *The history of combinatorial group theory: a case study in the history of ideas*. Studies in History of Math. and Phys. SPRINGER 1982.
- [CP] CORI R., PERRIN D., *Automates et commutations partielles*, RAIRO Inform.Th., Vol 19,21-32, 1985.
- [Dub] DUBOC C., *Commutations dans les monoïdes libres*. Univ.de Rouen (Thèse 1986)
- [DK1] DUCHAMP G., KROB D., *Partially commutative formal power series* L.N.in Computer Science 469pp 257-276 (1990), rapport LITP 90.64 et LIR.
- [DK2] DUCHAMP G., KROB D., *Free partially commutative structures*, Rapport LITP 90.65 et à paraître au Journal of Algebra.
- [DK3] DUCHAMP G., KROB D., *Partially commutative Magnus transformations* (manuscrit soumis).
- [DK4] DUCHAMP G. et KROB D., *Factorisations dans le monoïde partiellement commutatif libre*, C.R.Acad.Sci. Paris, t. 312, série I (1991), 189-192
- [DK5] DUCHAMP G., KROB D., *The lower central series of the Free Partially Commutative Group*, Rapport LITP 90.32, et soumis à Semigroup Forum.
- [DT] DUCHAMP G., THIBON, J.-Y., *Decidable orderings for free groups*, manuscrit, 1991.
- [Eil] EILENBERG S. *Automata, Languages and Machines*, vol A, Academic Press 1974.
- [Hall] HALL M., *A topology for free groups and related groups*, Ann. of Math, 52 (1950) 127-129.
- [HK] HARJU T., KARHUMAKI J., *Decidability of the multiplicity equivalence of multitape finite automata*, T.C.S 1990.
- [Laz] LAZARD M., *Groupes, anneaux de Lie et problème de Burnside*, Inst. Math. dell. Univ. Roma 1960.
- [Lot] LOTHAIRE M. *Combinatorics on words* ADDISON-WESLEY 1983
- [MKS] MAGNUS W., KARRASS A., SOLITAR D., *Combinatorial group theory*. DOVER 1976.
- [Neu] NEUMANN B.H., *On ordered division rings*, Trans.Amer.Math.Soc., 66, 202-252, 1949.
- [Mu] MURA R.B., RHEMTULLA A., *Orderable groups*, MARCEL DEKKER 1977
- [Thi] THIBON J.-Y., *Intégrité des algèbres de séries formelles sur un alphabet partiellement commutatif*, T.C.S. 1985.
- [Va] VARRICCHIO S., *On the decidability of the equivalence problem for partially commutative rational power series*, rapport LITP 90-97
- [Vi] VIENNOT G.X., *Heaps of pieces, I: basic definitions and combinatorial lemmas*, in Combinatoire énumérative. Proc., Montréal, Québec 1985, LN in Math 1234, SPRINGER.

Interprétations combinatoires des nombres de Genocchi et propriétés de symétrie

Guo-Niu HAN

Université Louis-Pasteur
7, rue René Descartes
67084 Strasbourg

RÉSUMÉ. — Dumont [Dum] a donné une interprétation des nombres de Genocchi en termes de comptages d'applications excédantes et surjectives de l'intervalle $[2n]$ sur le sous-ensemble des entiers pairs $2[n]$. Dans cet article, on propose une méthode nouvelle pour bâtir une famille plus étendue d'interprétations combinatoires de ces mêmes nombres contenant en particulier la précédente. Les nouvelles interprétations ont toutes la propriété de symétrie qu'avaient établie pour la première fois Dumont et Foata [DumFoa]. Enfin, une propriété de symétrie pour la statistique trivariée nombres de points "maximaux", "fixes" et "surfixes" est établie pour la première fois.

ABSTRACT. — Dumont [Dum] has given an interpretation of the Genocchi numbers by showing that they counted the excedent and surjective mappings of the interval $[2n]$ onto the subset $2[n]$ of the even integers. In the present paper a new method is proposed that yields a larger family of combinatorial interpretations of those numbers. The symmetry property of Dumont's interpretation also holds for that new family. Finally another symmetry property is derived for the trivariate statistic number of "maxima", "fixed" and so-called "surfixed" points.

1. Introduction et Notations

Les nombres de Genocchi $(G_{2n})_{n \geq 1}$ peuvent être définis par la fonction génératrice exponentielle :

$$G(t) = \frac{2t}{e^t + 1} = t + \sum_{n \geq 1} (-1)^n \frac{t^{2n}}{(2n)!} G_{2n},$$

les premières valeurs étant reproduites dans le tableau suivant :

n	1	2	3	4	5	6	7
G_{2n}	1	1	3	17	155	2073	38227

Depuis Dumont[Dum], on connaît plusieurs interprétations combinatoires de ces nombres, reposant sur le théorème suivant :

THÉOREME 1.1 (DUMONT). — Notons A_n l'ensemble de toutes les applications excédantes (i.e. pour tout i , $f(i) \geq i$) définies sur $[2n]$ et surjectives sur $2[n]$. Alors, $|A_n| = G_{2n+2}$. \square

Dumont et Foata [DumFoa] ont donné un raffinement de ce théorème. Soient f une application appartenant à A_n et i un élément de $[2n-2]$, on dit que i est un *point saillant* de f , si

$$1 \leq k < i \text{ implique } f(k) < f(i) < 2n;$$

un *point fixe* si $f(i) = i$; et un *point maximal* si $f(i) = 2n$. On désigne par $\text{sai}(f)$ (resp. $\text{fix}(f)$, $\text{max}(f)$) le nombre de points saillants (resp. fixes et maximaux) de l'application f .

THÉOREME 1.2 (DUMONT-FOATA). — Pour tout entier $n \geq 1$, soit $F_n(x, y, z)$ la fonction génératrice du vecteur $(\text{max}, \text{fix}, \text{sai})$ sur A_n , c'est-à-dire,

$$\sum_{f \in A_n} x^{\text{max}(f)} y^{\text{fix}(f)} z^{\text{sai}(f)} = F_n(x, y, z).$$

Alors la suite des polynômes $(F_n(x, y, z))_{n \geq 0}$ est caractérisée par la relation de récurrence suivante :

$$\begin{cases} F_1(x, y, z) = 1; \\ F_n(x, y, z) = (x+y)(x+z)F_{n-1}(x+1, y, z) - x^2 F_{n-1}(x, y, z) \quad (n \geq 2). \end{cases}$$

D'après le théorème 1.1, on a, en particulier,

$$(1.1) \quad F_n(1, 1, 1) = |A_n| = G_{2n+2}.$$

Les premières valeurs des polynômes $F_n(x, y, z)$ sont données par :

$$\begin{aligned} F_1(x, y, z) &= 1; \\ F_2(x, y, z) &= xy + xz + yz; \\ F_3(x, y, z) &= x^2 y^2 + x^2 z^2 + y^2 z^2 + 2x^2 yz + 2xy^2 z + 2xyz^2 \\ &\quad + x^2 y + x^2 z + xy^2 + y^2 z + xz^2 + yz^2 + 2xyz. \end{aligned}$$

On peut vérifier que pour tout $n \geq 1$, le polynôme $F_n(x, y, z)$ est *symétrique* en les trois variables x, y, z ; ou encore, que la distribution des *statistiques* "max", "fix", et "sai" est symétrique. Pour expliquer cette symétrie, Dumont et Foata [DumFoa] avaient construit des involutions

sur A_n qui mettaient en évidence cette propriété. Leur construction était cependant indirecte. Le problème qui reste ouvert est de trouver un modèle dans lequel cette symétrie apparaît de façon immédiate.

Notre but ici n'est pas de fournir ce modèle symétrique, mais de montrer qu'en réalité le modèle des applications excédantes surjectives recèle plusieurs autres symétries, dont l'une d'entre elles est fournie par le résultat de Dumont-Foata. Il est commode de visualiser les applications excédantes de $[2n]$ dans $2[n]$ comme étant des parties de l'ensemble $\{(i, 2j) \mid 1 \leq i \leq 2j \leq 2n\}$ qu'on appellera *escalier (ordinaire) pair* d'ordre $2n$. Dumont et Foata considéraient les seules applications *surjectives* contenues dans cet escalier. On peut, en fait, introduire d'autres familles de configurations qui ont les mêmes propriétés. On aboutit ainsi au théorème 2.1, qui constitue le résultat principal de cet article. Chaque famille est indiquée par un vecteur $U = (U_1, U_2, \dots, U_{n-1})$ dont les éléments sont pris dans un alphabet à trois lettres $\{X, Y, Z\}$. Lorsqu'on spécialise la suite U à une $(n-1)$ -suite dont tous les termes sont égaux à X , la famille correspondante est simplement A_n . On est amené à introduire une nouvelle statistique sur les éléments f de A_n , qui compte les points *surfices*, c'est-à-dire les points i tels que $1 \leq i \leq 2n-2$ et $f(i) = i+1$. On établit alors (cf. section 3).

THÉORÈME 1.3. — *La statistique (max, fix, sur) a pour fonction génératrice sur A_n le polynôme $F_n(x, y, z)$.*

Lorsque $U = (Y, Y, \dots, Y)$, la famille de configurations est toujours A_n , mais le triplet de statistique devient (sur, max, fix) (dans cet ordre). Comme la symétrie entre "sur" et "fix" est simple (voir section 4), on obtient la propriété de symétrie du triplet (max, fix, sur) sur A_n .

En comparant les énoncés des théorèmes 1.2 et 1.3, on voit qu'il doit exister une bijection de A_n sur A_n envoyant le triplet (max, fix, sai) sur (max, fix, sur). En fait, on construit ici (cf. théorème 4.2) une *involution* $f \mapsto \tilde{f}$ sur A_n ayant ces mêmes propriétés.

2. Les applications excédantes U -surjectives

Une application V définie sur l'escalier pair d'ordre $2n$, dans l'ensemble $\{0, 1, x, y, z\}$ est appelée *évaluation*. Les lettres x, y, z sont des variables commutatives et il sera commode de dire par la suite que y (resp. z, x) est le *successeur* de x (resp. y, z). On écrit, par exemple, $y = \text{succ } x$ et aussi $z = \text{succ} * \text{succ } x$. Soit alors $U = (U_1, U_2, \dots, U_{n-1})$ une suite dont les termes sont pris dans l'ensemble $\{X, Y, Z\}$. On lui associe une évaluation, notée V^U , de l'escalier pair défini par les relations suivantes :

- 1) $V^U(2n - 1, 2n) = V^U(2n, 2n) = 1$;
- 2) $V^U(2i - 1, 2n) = V^U(2i, 2n) = U_i$ pour $1 \leq i \leq n - 1$;
- 3) $V^U(2i, 2i) = \text{succ}(U_i)$, $V^U(2i - 1, 2i) = \text{succ} * \text{succ}(U_i)$
pour $1 \leq i \leq n - 1$;
- 4) pour $1 \leq i \leq n - 1$ et $i + 1 \leq j \leq n - 1$,

$$V^U(i, j) = \begin{cases} 1, & \text{si l'ensemble formé par la ligne et} \\ & \text{la colonne passant par le point } (i, j) \\ & \text{contient les trois lettres } x, y, z; \\ 0, & \text{sinon.} \end{cases}$$

Par exemple, pour $n = 7$ et $U = (X, Y, X, X, Z, Y)$, l'évaluation V^U peut être représentée par le diagramme suivant (les lettres représentant l'évaluation s'écrivent en minuscules) :

$U =$	X	Y	X	X	Z	Y								
14	x	x	y	y	x	x	x	x	z	z	y	y	1	1
12	0	1	1	1	0	1	0	1	1	0	x	z		
10	1	0	0	1	1	0	1	0	y	x				
8	1	1	1	0	1	1	z	y						
6	1	1	1	0	z	y								
4	0	1	x	z										
2	z	y												
1 2 3 4 5 6 7 8 9 10 11 12 13 14														

Escalier pair d'ordre 14 U-évalué

Les conditions 3) et 4) ci-dessus entraînent que les double-cases $((2i - 1, 2j), (2i, 2j))$ ($2 \leq i + 1 \leq j \leq n - 1$) ne sont jamais d'évaluation $(0, 0)$. Elles sont forcément d'évaluation $(0, 1)$, $(1, 0)$ ou $(1, 1)$.

Pour chaque $i = 1, 2, \dots, n$, il est commode de désigner par B_{2i}^U le sous-ensemble de l'escalier pair d'ordre $2n$ constitué par les points de la ligne $2i$ d'évaluation *non nulle*, ainsi que par les points $(2i - 1, 2j)$ (resp. $(2i, 2j)$) de la colonne $(2i - 1)$ (resp. $2i$), dont la case adjacente $(2i, 2j)$ (resp. $(2i - 1, 2j)$) est nulle ($2 \leq i + 1 \leq j \leq n - 1$).

Par exemple, en reprenant la suite U précédente, les ensembles B_{2i}^U sont les cases de l'escalier ci-dessous portant le numéro $2i$. (Noter que certaines cases ont deux numéros, les bandes B_{2i}^U n'étant pas nécessairement disjointes deux à deux.)

$U = \quad X \quad Y \quad X \quad X \quad Z \quad Y$

14	14 14	14 14	14 14	14 14	14 14	14 14	14 14	14 14
12	2,12	12 12	6,12	8,12	10,12	12 12		
10	2,10	4,10	6,10	8,10	10 10			
8	8 8	4,8	8 8	8 8				
6	6 6	4,6	6 6					
4	2,4	4 4						
2	2 2							

1 2 3 4 5 6 7 8 9 10 11 12 13 14

Escalier pair d'ordre 14 U-numéroté

Pour ne pas confondre avec la notion d'escalier pair évalué, on dit qu'un tel escalier pair est *U-numéroté*.

Soit \bar{A}_n l'ensemble des applications f excédantes de $[2n]$ sur $2[n]$ (non nécessairement surjectives), ou encore l'ensemble des configurations de $2n$ points contenues dans l'escalier pair d'ordre $2n$ ayant exactement un point dans chaque colonne.

L'escalier pair *U-évalué* sert à donner une *évaluation* $V^U(f)$ à chaque f . On pose en effet

$$V^U(f) = \prod_{i \in [2n]} V^U(i, f(i)).$$

L'escalier pair *U-numéroté* sert à définir la notion de *U-surjectivité*. On dit, en effet, qu'une application f appartenant à \bar{A}_n est *U-surjective*, si pour tout i ($1 \leq i \leq n$), il existe au moins un point $k \in [2n]$ tel que $(k, f(k)) \in B_{2i}^U$. Autrement dit, f est *U-surjective*, si, lorsqu'on superpose son graphe sur l'escalier pair *U-numéroté*, on trouve au moins une fois les numéros $2, 4, \dots, 2n$ sur les cases occupées par les points du graphe de f . On note A_n^U l'ensemble de toutes les applications f *U-surjectives* d'évaluation $V^U(f)$ non nulle. Soit $i \in [2n]$; on dit que i est un point *U-maximal*, si $V^U(i, f(i)) = x$, que c'est un point *U-fixe*, si $V^U(i, f(i)) = y$ et

qu'enfin c'est un point *U-surfixe*, si $V^U(i, f(i)) = z$. On note $\max^U f, \text{fix}^U f$ et $\text{sur}^U f$, respectivement, le nombre de points *U*-maximaux, *U*-fixes et *U*-surfixes de f . Si $V^U(f) \neq 0$, on a naturellement

$$V^U(f) = x^{\max^U(f)} y^{\text{fix}^U(f)} z^{\text{sur}^U(f)}.$$

On peut alors énoncer notre théorème principal :

THÉOREME 2.1. — *Pour toute suite $U = (U_1, U_2, \dots, U_{n-1})$, la fonction génératrice du vecteur $(\max^U, \text{fix}^U, \text{sur}^U)$ sur A_n^U est indépendante de U et égale à*

$$\sum_{f \in A_n^U} V^U(f) = F_n(x, y, z).$$

En particulier,

$$(2.1) \quad |A_n^U| = F_n(1, 1, 1) = G_{2n+2}.$$

La figure suivante représente, par exemple, deux escaliers pairs d'ordre 6, *U*-évalués et *U*-numérotés comme indiqué.

$U =$	X	Y						
6	x	x	y	y	1	1		
4	0	1	x	z				
2	z	y						
	1	2	3	4	5	6		

$U =$	X	Y						
6	6	6	6	6	6	6		
4	2,4		4	4				
2	2	2						
	1	2	3	4	5	6		

D'après (2.1) il y a donc $G_8 = 17$ applications appartenant à A_3^U . Leurs *U*-évaluations sont représentées dans ce qui suit. On vérifie bien que leur fonction génératrice est $F_3(x, y, z)$.

- Application* : 224466 224666 226466 244466 244666 246466
- Evaluation* : $zyxz11$ $zyxy11$ $zyyz11$ $z1xz11$ $z1xy11$ $z1yz11$
- Application* : 246666 264466 264666 266466 624466 624666
- Evaluation* : $z1yy11$ $zxxz11$ $zxxxy11$ $zxyz11$ $xyxz11$ $xyxy11$
- Application* : 626466 644466 644666 646466 646666
- Evaluation* : $xyyz11$ $x1xz11$ $x1xy11$ $x1yz11$ $x1yy11$

Noter que les applications $f = 644666$ et $g = 246666$ sont bien *U*-surjectives, puisque le point $(2, 4) \in B_2 \cap B_4$.

3. Démonstration du théorème 2.1

Un *escalier gauche* E d'ordre $2n$ est défini comme une suite d'entiers positifs $(E_1, E_2, \dots, E_{2n})$ telle que

- i) $E_{2i-1} = E_{2i}$ pour $1 \leq i \leq n$;
- ii) $E_i - E_{i+1} = 0$ ou 1 pour $1 \leq i \leq 2n - 1$;
- iii) $E_{2n} = 1$.

On représente un tel escalier gauche comme un tableau de Young (à l'anglo-saxonne) ayant une profondeur égale à E_i à l'abscisse i . (Voir diagramme ci-dessous.) Par convention, les cases du tableau ainsi constitué sont repérées par les couples $(i, 2j)$ où $1 \leq i \leq 2n$ et $n - E_i + 1 \leq j \leq n$. Avec cette convention, un escalier gauche d'ordre $2n$ n'est donc rien d'autre qu'un *sous-ensemble* d'un escalier pair $\{(i, 2j) \mid 1 \leq i \leq 2j \leq 2n\}$ d'ordre $2n$, où les cases $(i, 2j)$ satisfont les inégalités $1 \leq i \leq 2n$ et $n - E_i + 1 \leq j \leq n$. En particulier, un escalier gauche d'ordre $2n$ contient toujours la *première ligne* $\{(i, 2n) \mid 1 \leq i \leq 2n\}$ de l'escalier pair.

On associe une *évaluation* V à l'escalier gauche E en donnant à chaque case $(i, 2j)$ une valeur prise dans l'alphabet $\{1, x, y, z\}$ selon les règles suivantes :

- i) $(x, x, \dots, x, x, 1, 1)$ pour la première ligne $\{(i, 2n) \mid 1 \leq i \leq 2n\}$;
- ii) $(1, 1, \dots, 1, 1, z, y)$ pour les autres lignes.

Notons que l'escalier gauche $(n, n, n - 1, n - 1, \dots, 1, 1)$ n'est rien d'autre que l'escalier pair d'ordre $2n$ et que son évaluation V coïncide alors avec l'évaluation V^U de l'escalier pair U -évalué à l'aide de la suite $U = (X, X, \dots, X)$.

On définit enfin la *signature* de E comme étant :

$$\text{sgn}(E) = (-1)^{n-E_1}.$$

Par exemple, $E = (3, 3, 2, 2, 2, 2, 1, 1, 1, 1)$ est un escalier gauche d'ordre 10; sa signature est $(-1)^2 = 1$. L'évaluation associée V peut être représentée par

$$V = \begin{array}{r} 10 \\ 8 \\ 6 \end{array} \begin{array}{|cccccc|} \hline x & x & x & x & x & x & 1 & 1 \\ \hline 1 & 1 & 1 & 1 & z & y & & \\ \hline z & y & & & & & & \\ \hline \end{array} \begin{array}{cccccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \end{array}$$

Notons \mathcal{F}_E l'ensemble de toutes les applications f de \bar{A}_n telles que

$f(i) \geq 2n - 2E_i + 2$, (i.e., telles que tous les points $(i, f(i))$ soient dans le diagramme de E). L'évaluation V de E induit alors une évaluation de f définie par

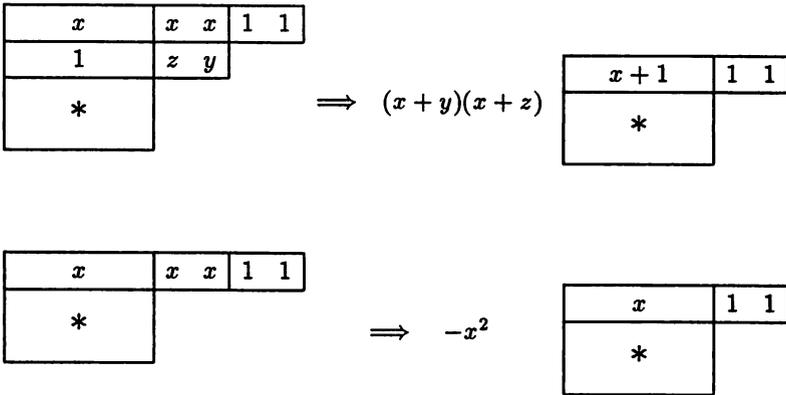
$$V(E; f) = \text{sgn}(E) \prod_{i \in [2n]} V(i, f(i)).$$

LEMME 3.1. — On a

$$F_n(x, y, z) = \sum_E \sum_{f \in \mathcal{F}_E} V(E; f),$$

où la première somme est sur l'ensemble de tous les escaliers gauches d'ordre $2n$.

DÉMONSTRATION. — Le seul escalier gauche possible d'ordre 2 est $(1, 1)$; le lemme est donc vrai pour $n = 1$. Pour $n \geq 2$, on décompose l'ensemble \mathcal{E}_{2n} des escaliers gauches en deux sous-classes \mathcal{E}_{2n}^1 et \mathcal{E}_{2n}^2 distinctes : on pose $E \in \mathcal{E}_{2n}^1$, si $E_{2n-2} = 2$ et $E \in \mathcal{E}_{2n}^2$, si $E_{2n-2} = 1$. Les diagrammes suivants donnent la clé de la démonstration du lemme.



Le premier diagramme, par exemple, montre que tout escalier gauche E d'ordre $2n$ tel que $E_{2n-2} = 2$ est envoyé sur un escalier gauche E' d'ordre $(2n - 2)$ donné par $E' = (E_1 - 1, E_2 - 1, \dots, E_{2n-5} - 1, E_{2n-4} - 1, E_{2n-1}, E_{2n})$.

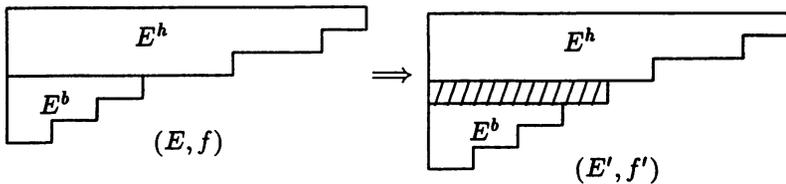
En imposant la valeur $(x + 1, x + 1, \dots, x + 1, x + 1, 1, 1)$ à la première ligne de E' et $(1, 1, \dots, 1, 1, z, y)$ aux autres lignes, on obtient une nouvelle évaluation V' pour E' , d'où une évaluation $V'(E'; g)$ pour toute application g de \bar{A}_{n-1} appartenant à $\mathcal{F}_{E'}$. Comme $\text{sgn } E' = (-1)^{n-1-(E_1-1)} = (-1)^{n-E_1} = \text{sgn } E$, on a bien

$$\sum_{f \in \mathcal{F}_E} V(E; f) = (x + y)(x + z) \sum_{g \in \mathcal{F}_{E'}} V'(E'; g). \quad \square$$

Un escalier gauche E d'ordre $2n$ peut se décomposer en deux parties E^b et E^h : on prend pour E^b le plus grand escalier ordinaire pair dans le coin inférieur gauche de E . Autrement dit, ou bien $E_2 = E_3$ et $E^b = \emptyset$, ou bien, dans le cas contraire, $E^b = (E_1, E_2, \dots, E_{2m})$, où m est le plus petit entier satisfaisant $1 \leq m \leq n - 1$ et $E_{2m+2} = E_{2m+3}$. Naturellement E^h est la partie haute restante : $E^h = E \setminus E^b$.

On décompose alors l'ensemble des couples (E, f) , où E est un escalier gauche d'ordre $2n$ et f un élément de \mathcal{F}_E en trois classes $\mathcal{F}_{2n}^1, \mathcal{F}_{2n}^2$ et \mathcal{F}_{2n}^0 . Dans la première, nous rangeons tous les couples (E, f) tels que $E^h \neq \emptyset$ et f est surjective sur E^b ; dans la seconde, ceux des couples où f n'est pas surjective sur E^b ; et dans la troisième, E est l'escalier ordinaire pair et f une application appartenant à A_n .

On définit une bijection de \mathcal{F}_{2n}^1 sur \mathcal{F}_{2n}^2 indiquée par le diagramme suivant :



L'escalier gauche E' est obtenu en incluant une ligne vide dont la longueur est égale à l'ordre de E^b plus 2 entre E^b et E^h . On obtient bien encore un escalier gauche d'ordre $2n$, puisque la longueur de la dernière ligne de E^h surpasse au moins de quatre unités la longueur de la première ligne de E^b . De la même façon, on obtient le graphe de l'application f' en incluant une ligne vide dans le graphe de f . De plus $(E')^b$ contient strictement E^b et donc (au moins) une ligne vide. L'application f' n'est

donc pas surjective sur $(E')^b$, de sorte que le nouveau couple (E', f') appartient à \mathcal{F}_{2n}^2 .

On passe, réciproquement, de (E', f') à (E, f) par une définition évidente. On a enfin

$$V(E'; f') = -V(E; f),$$

puisque E' a une ligne de plus que E .

Dans la somme (3.1) il ne reste donc plus que les couples (E, f) où E est l'escalier pair et f une application surjective sur E . On a donc démontré le lemme suivant :

LEMME 3.2. — On a

$$F_n(x, y, z) = \sum_{f \in A_n} V^U(f),$$

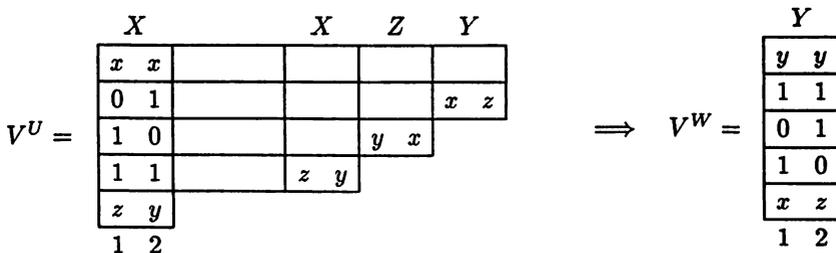
où $U = (X, X, \dots, X)$. \square

LEMME 3.3. — Le changement de la première lettre dans une suite ne change pas la fonction génératrice associée à cette suite.

Nous donnons ci-après la démonstration seulement dans le cas où $U = (X, U_2, U_3, \dots, U_{n-1})$ et $W = (Y, U_2, U_3, \dots, U_{n-1})$. Les autres cas comme $U = (Z, U_2, U_3, \dots, U_{n-1})$ et $W = (X, U_2, U_3, \dots, U_{n-1})$ se démontrent de la même façon.

DÉMONSTRATION. — Pour démontrer le lemme, on construit une bijection $\Phi : f \mapsto g$ de l'ensemble A_n^U sur A_n^W telle que $V^U(f) = V^W(g)$.

Pour les deux premières colonnes, le dessin suivant montre le changement d'évaluation :



On constate donc que pour $4 \leq j \leq 2n - 2$, on a dans tous les cas $V^U(2, j) = V^W(1, j)$.

L'application g est ainsi définie :

- (i) $g(i) = f(i)$ pour $3 \leq i \leq 2n$;
 (ii) si $(1, f(1)) \in B_2^U$, on pose $g(2) = f(1)$ et

$$g(1) = \begin{cases} 2n, & \text{si } f(2) = 2; \\ 2, & \text{si } f(2) = 2n; \\ f(2), & \text{dans les autres cas;} \end{cases}$$

(iii) Si $(1, f(1)) \notin B_2^U$, on a $(2, f(2)) \in B_2^U$ par U -surjectivité, on pose alors

$$g(1) = \begin{cases} 2, & \text{si } f(1) = 2n; \\ f(1), & \text{sinon;} \end{cases}$$

et

$$g(2) = \begin{cases} 2n, & \text{si } f(2) = 2; \\ f(2), & \text{sinon.} \end{cases}$$

L'inverse de cette construction est assurée par les faits que dans le cas (ii) on a toujours $(2, g(2)) \in B_2^U$, et dans le cas (iii) on a $(2, g(2)) \notin B_2^U$.

On vérifie, en distinguant plusieurs cas simples, que $f \mapsto g$ est une bijection ayant bien la propriété $V^U(f) = V^W(g)$ désirée. \square

LEMME 3.4. — *Le changement de deux lettres consécutives dans une suite ne change pas la fonction génératrice associée à cette suite.*

Nous donnons ci-après la démonstration seulement dans le cas où $U = (U_1, U_2, \dots, U_r = X, U_{r+1} = Y, \dots, U_{n-1})$ et $W = (U_1, U_2, \dots, U_r = Y, U_{r+1} = X, \dots, U_{n-1})$. Les autres cas comme $U = (U_1, U_2, \dots, U_r = Z, U_{r+1} = Y, \dots, U_{n-1})$ et $W = (U_1, U_2, \dots, U_r = Y, U_{r+1} = Z, \dots, U_{n-1})$ se démontrent de la même façon.

DÉMONSTRATION. — Pour démontrer le lemme, on construit une bijection $\Psi : f \mapsto g$ de l'ensemble A_n^U sur A_n^W telle que $V^U(f) = V^W(g)$.

Pour $i \geq 2r + 3$, on pose toujours $g(i) = f(i)$. Puisque pour $i \leq 2r - 2$, on a $V^U(i, 2r + 2) = V^W(i, 2r)$ et $V^U(i, 2r) = V^W(i, 2r + 2)$, on peut poser

$$g(i) = \begin{cases} 2r, & \text{si } f(i) = 2r + 2; \\ 2r + 2, & \text{si } f(i) = 2r; \\ f(i), & \text{dans les autres cas.} \end{cases}$$

En résumé, les valeurs de g dans la ligne $2r$ et $2r + 2$ sont déterminées de façon naturelle à partir de ceux de f dans la ligne $2r + 2$ et $2r$

respectivement. Et pour les quatres colonnes restantes, le changement d'évaluation est représenté par le dessin suivant :

$$V^U = \begin{array}{c} \begin{array}{cc} X & Y \end{array} \\ \begin{array}{|cc|cc|} \hline x & x & y & y \\ \hline 0 & 1 & 1 & 1 \\ \hline 1 & 0 & 0 & 1 \\ \hline 1 & 1 & 1 & 0 \\ \hline 0 & 1 & x & z \\ \hline z & y & & \\ \hline \end{array} \\ \begin{array}{ccc} 2r-1 & 2r & 2r+1 \end{array} \end{array} \quad \Rightarrow \quad \begin{array}{c} \begin{array}{cc} Y & X \end{array} \\ \begin{array}{|cc|cc|} \hline y & y & x & x \\ \hline 1 & 1 & 0 & 1 \\ \hline 0 & 1 & 1 & 0 \\ \hline 1 & 0 & 1 & 1 \\ \hline 1 & 0 & z & y \\ \hline x & z & & \\ \hline \end{array} \\ \begin{array}{ccc} 2r-1 & 2r & 2r+1 \end{array} \end{array}$$

(i) Si $f(2r) \neq 2r + 2$, alors les valeurs de g dans les colonnes $2r - 1, 2r, 2r + 1$ et $2r + 2$ sont déterminées de façon naturelle par celles de f dans les colonnes $2r + 1, 2r + 2, 2r - 1$ et $2r$, respectivement.

(ii) Si $f(2r) = 2r + 2$, alors les valeurs de g dans les colonnes $2r - 1, 2r, 2r + 1$ et $2r + 2$ sont déterminées de façon naturelle par celles de f dans les colonnes $2r, 2r + 2, 2r - 1$ et $2r + 1$, respectivement. Remarquons que dans ce cas, les valeurs de f pour les lignes $2r - 1, 2r + 1$ et $2r + 2$ sont assez arbitraires puisque $(2r, f(2r)) \in B_{2r} \cap B_{2r+2}$.

L'inverse de cette construction est assurée par les faits que dans le cas (i) on a toujours $g(2r-1) \neq 2r+2$, et dans le cas (ii) on a $g(2r-1) = 2r+2$.

On vérifie que $f \mapsto g$ est une bijection conservant l'évaluation. \square

Le lemme 3.2 dit que le théorème 2.1 est vrai pour la suite $U = (X, X, \dots, X)$. Par les lemmes 3.3 et 3.4, le théorème 2.1 est alors vrai pour toute suite U arbitraire.

4. La symétrie et le codage de Dumont-Foata

Le fait que les polynômes $F_n(x, y, z)$ sont symétriques en les trois variables x, y, z peut être établi de la façon suivante :

En prenant la suite $U = (X, X, \dots, X)$, on obtient $A_n^U = A_n$ et $\max^U = \max$, $\text{fix}^U = \text{fix}$ et $\text{sur}^U = \text{sur}$. Si l'on prend la suite $U = (Y, Y, \dots, Y)$, on a, en revanche, $A_n^U = A_n$ et $\max^U = \text{sur}$, $\text{fix}^U = \max$ et $\text{sur}^U = \text{fix}$. D'où $F_n(x, y, z) = F_n(y, z, x)$.

D'autre part, soit $f \in A_n$, on définit une involution $f \mapsto f'$ telle que $f'(2i) = f(2i - 1)$ et $f'(2i - 1) = f(2i)$ pour $1 \leq i \leq n$. Cette involution montre que $F_n(x, y, z) = F_n(x, z, y)$. Par conséquent, le polynôme $F_n(x, y, z)$ est symétrique en les trois variables x, y, z .

En comparant le théorème 1.2 et le théorème 2.1 dans le cas particulier où $U = (X, X, \dots, X)$, on voit qu'il doit exister une bijection de A_n sur A_n qui envoie les statistiques "max", "fix" et "sur" sur les statistiques "max", "fix" et "sai". Notre propos final est de construire une telle bijection, qui est, en fait, une involution. Auparavant, nous donnons un *codage* pour les applications excédantes surjectives de $[2n]$ sur $[2n]$.

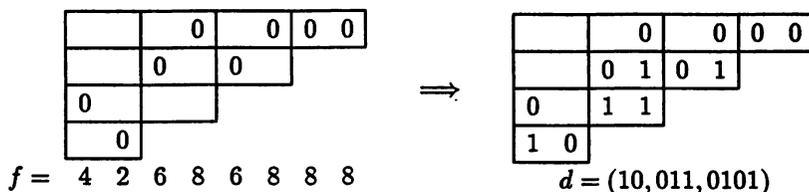
Notons \mathcal{D} l'ensemble de tous les mots ayant les lettres 0 et 1. Soit $m \in \mathcal{D}$, la longueur de m et le nombre de lettres α dans m sont notés respectivement $|m|$ et $|m|_\alpha$.

Soit n un entier positif. Notons \mathcal{D}_n l'ensemble de toutes les suites de mots $(d_1, d_2, \dots, d_{n-1})$ satisfaisant les conditions

- C1) $|d_1| = 2$;
- C2) $|d_i| = |d_{i-1}|_1 + 2$ pour $2 \leq i \leq n-1$;
- C3) $|d_i|_0 \geq 1$ pour $1 \leq i \leq n-1$.

Par exemple, la suite $(10, 011, 0101)$ est un élément de \mathcal{D}_4 . Dans [DumFoa], Dumont et Foata ont construit une bijection $f \mapsto d$ entre A_n et \mathcal{D}_n ; la suite d est appelée *codage* de Dumont-Foata de l'application f . Par exemple, le codage de l'application $f = 4268688$ est la suite $d = (10, 011, 0101)$. Le codage est construit de la façon suivante :

- 1) Dessiner le graphe de l'application f dans un escalier pair en mettant des lettres 0 sur les points $(i, f(i))$;
- 2) Ajouter des lettres 1 de sorte que chaque colonne lue de haut en bas contienne une suite maximale de la forme $011\dots 1$;
- 3) Lire les nombres écrits, ligne par ligne, de bas en haut et de gauche à droite. Pour $1 \leq i \leq n-1$, d_i correspond à la ligne de hauteur i .



Nous remarquons que ce codage peut être étendu à l'ensemble de toutes les applications définies sur l'escalier pair en supprimant simplement la condition C3) dans la définition de l'ensemble \mathcal{D}_n .

D'après cette définition, on a les propriétés suivantes :

LEMME 4.1. — Soit d le codage de l'application f . Alors on a

- 1) $\max(f) = |d_{n-1}|_1$;
 - 2) $\text{fix}(f) = \#\{i \mid d_i = m0\}$;
 - 3) $\text{sur}(f) = \#\{i \mid d_i = m0\alpha\}$;
 - 4) $\text{sai}(f) = \#\{i \mid d_i = 0m\}$;
- où $|\alpha| = 1$ et $m \in \mathcal{D}$. \square

Nous définissons maintenant une involution $m \mapsto \hat{m}$ sur \mathcal{D} . Si $|m| = 2$, on pose $\hat{m} = m$; sinon, m peut s'écrire sous la forme $m = \alpha m' \beta \gamma$ où $|\alpha| = |\beta| = |\gamma| = 1, m' \in \mathcal{D}$, et on pose alors $\hat{m} = \beta m' \alpha \gamma$. Remarquons que cette involution ne change pas la longueur et le nombre de 1. Cette définition peut être prolongée à tout l'ensemble \mathcal{D}_n d'une façon naturelle :

$$\begin{array}{ccc} \mathcal{D}_n & \longrightarrow & \mathcal{D}_n \\ d = (d_1, d_2, \dots, d_{n-1}) & \longmapsto & \hat{d} = (\hat{d}_1, \hat{d}_2, \dots, \hat{d}_{n-1}) \end{array}$$

THÉORÈME 4.2. — Il existe une involution $f \mapsto \hat{f}$ sur A_n telle que :

- 1) $\max(f) = \max(\hat{f})$;
- 2) $\text{fix}(f) = \text{fix}(\hat{f})$;
- 3) $\text{sur}(f) = \text{sai}(\hat{f})$;
- 4) $\text{sai}(f) = \text{sur}(\hat{f})$.

DÉMONSTRATION. — Soient $f \in A_n$ et $d \in \mathcal{D}_n$ le codage de f . Par l'involution précédente, on obtient une autre suite \hat{d} ; l'application \hat{f} vient alors du codage \hat{d} . D'après le lemme précédent, on peut vérifier que les quatre relations pour les statistiques sont satisfaites. \square

Par exemple, si on prend l'application $f = 42686888$, alors, son codage de Dumont-Foata est la suite $d = (10, 011, 0101)$; d'où, $\hat{d} = (10, 101, 0101)$ et $\hat{f} = 62486888$. On peut vérifier que $\max(f) = \max(\hat{f}) = 2$, $\text{fix}(f) = \text{fix}(\hat{f}) = 1$, $\text{sur}(f) = \text{sai}(\hat{f}) = 1$ et $\text{sai}(f) = \text{sur}(\hat{f}) = 2$.

BIBLIOGRAPHIE

- [And] ANDRÉ (D.). — Sur les permutations alternées, *J. Math. Pures Appl.*, t. 7, 1881, p. 167-187.
- [Car1] CARLITZ (L.). — A conjecture concerning Genocchi numbers, *K. Norske Vidensk. Selsk. Sk.*, t. 9, 1972, p. 1-4.
- [Car2] CARLITZ (L.). — Explicit formulas for the Dumont-Foata polynomial, *Discrete Math.*, t. 30, 1980, p. 211-225.
- [Dum] DUMONT (D.). — Interprétations combinatoires des nombres de Genocchi, *Duke math. J.*, t. 41(2), 1974, p. 305-318.
- [DumFoa] DUMONT (D.) et FOATA (D.). — Une propriété de symétrie des nombres de Genocchi, *Bull. Soc. math. France*, t. 104, 1976, p. 433-451.
- [DumVie] DUMONT (D.) et VIENNOT (G.). — A combinatorial interpretation of the Seidel generation of Genocchi numbers, *Annals of Discrete Mathematics*, t. 6, 1980, p. 77-87.
- [FoaSch] FOATA (D.) et SCHÜTZENBERGER (M. -P.). — Nombres d'Euler et permutations alternantes, *A survey of Combinatorial Theory*, (J.N. Srivastava, Ed.), North-Holland, Amsterdam, 1973.
- [Gan] GANDHI (J. M.). — A conjectured representation of Genocchi numbers, *Am. Math. Monthly*, t. 77(1), 1970, p. 505-506.
- [Rio] RIORDAN (J.) et STEIN (P.). — Proof of a conjecture on Genocchi numbers, *Discrete Math.*, Amsterdam, t. 5, 1973, p. 381-388.

Lie-Jordan superalgebras, and complete free finitely generated superalgebras.

G.JACOB

L.I.F.L.(U.A.369), Bât. M3, Université Lille I,
59655 Villeneuve d'Ascq Cedex, FRANCE.

15 Octobre 1990

Abstract.

This paper concerns the Lie and Jordan brackets in superalgebras. It contains two parts. The first part explores the four "Jacobi-like relations", that are the universal "linear" relations of order 3 satisfied by the Lie and Jordan brackets in any associative graded algebra. Using these relations to define abstract Lie-Jordan algebras, we prove that the enveloping algebra of any Lie algebra \mathcal{G} , is nothing else as the free Lie-Jordan algebra over this Lie algebra \mathcal{G} .

In the second part, we study the graded complete free Lie algebras. We transcribe, in a noncommutative powers series setting, the method of R. Ree to characterize the Lie elements as primitive elements (in characteristic 0), with its classical consequences. We define also the operator s of "left Jordan bracketing", and show that, in any characteristic, it acts like the identity on the primitive elements of degree odd, and annihilates the primitive elements of degree even.

Résumé.

Ce texte présente deux études concernant les crochets de Lie et de Jordan dans une superalgèbre graduée. La première étude explore les quatre "relations à la Jacobi", qui sont les seules relations linéaires universelles d'ordre 3 satisfaites par les crochets de Lie et de Jordan dans les algèbres associatives. Inversement, si nous appelons "algèbre de Lie-Jordan" toute algèbre munie de deux opérations binaires, respectivement symétrique et antisymétrique, vérifiant les quatre "relations à la Jacobi", alors l'algèbre enveloppante d'une algèbre de Lie \mathcal{G} n'est autre que l'algèbre de Lie-Jordan libre au dessus de \mathcal{G} .

La seconde étude concerne les algèbre de Lie graduées complètes. Nous reprenons, dans une version moderne des séries formelles non commutatives, la preuve donnée par R. Ree de la caractérisation des éléments de Lie comme éléments primitifs (en caractéristique 0), , avec ses conséquences classiques. Nous définissons aussi l'opérateur s de "crochetage de Jordan à gauche". Nous montrons qu'il agit, quelque soit la caractéristique, comme l'identité sur les éléments primitifs de degré impair, et qu'il annule les éléments primitifs de degré pair.

Contents

1 Lie-Jordan Super-algebras.	2
1.1 Jordan-brackets and Lie-brackets in associative algebras.	3
1.2 Abstract Lie-Jordan algebras.	6
2 Complete free finitely generated superalgebras.	7
2.1 Power Series on an alphabet \mathcal{Z}	8
2.2 Two-fold power series on $\mathcal{Z} \otimes \mathcal{Y}$	9
2.3 Partition's Coalgebra.	10
2.4 Primitive elements.	12
2.5 Analysis of Relations.	13
2.6 Classical consequences.	16

1 Lie-Jordan Super-algebras.

Graded Superalgebras and graded Lie Superalgebras was presented and extensively studied in 1979 by M. Scheunert [8]. (In its book can be found many examples). They have been used also as an algebraic tool for developing Supergravities theories (see Kaç [4], and Choquet-Bruhat [1], and recently Sanchez-Valenzuela and Sternberg [7]).

We present here new results concerning the free Lie and Jordan Super-algebras. But our actual presentation, following R. Ree (1960) [6], will be given in a *more general setting*, involving *graduations in an arbitrary finitely generated abelian semigroup* (and not only in \mathbb{Z}_2). We hope this setting will also someday incite new developments in mathematical physic.

Let K be a field of characteristic 0. We recall that a (not necessary associative) K -algebra \mathcal{A} is defined by

- a K -vector space $(\mathcal{A}, +, 0)$
- a binary operation “.” on \mathcal{A} , left and right K -linear, with neutral element noted “1”,
- satisfying the “scalar’s associativity”: $\forall a, b \in K, \forall u \in \mathcal{A}, a.(b.u) = (ab).u$.

Definition 1.1 Let M be an abelian semigroup, with operation noted additively by “+”. A K -algebra \mathcal{A} is said to be M -graded (or equivalently, (\mathcal{A}, M) is a graded algebra) if \mathcal{A} it is a direct sum of vector subspaces $\mathcal{A} = \coprod_{m \in M} \mathcal{A}_m$, with the additional property:

$$\forall f \in \mathcal{A}_m, \forall g \in \mathcal{A}_n, \quad f.g \in \mathcal{A}_{m+n}.$$

Any element f of \mathcal{A}_m is called homogeneous of degree m . Any element of \mathcal{A} can be written (not uniquely) as a finite sum of homogeneous elements.

In all the sequel, we shall note the degree of an homogenous element by the double bar:

$$f, g, h, \dots \quad \text{have respective degree : } |f|, |g|, |h|.$$

Definition 1.2 (R. Ree) We call bicharacter over the graded algebra (A, M) any application $\chi : M \times M \rightarrow K$ (noted by: $(\alpha, \beta) \mapsto \chi_{\alpha, \beta}$) that satisfies:

$$\begin{aligned} \text{(i)} \quad & \forall \alpha, \beta, \gamma \in M, \quad \begin{cases} \chi(\alpha, \beta + \gamma) = \chi(\alpha, \beta) \chi(\alpha, \gamma) \\ \chi(\alpha + \beta, \gamma) = \chi(\alpha, \gamma) \chi(\beta, \gamma) \end{cases} \\ \text{(ii)} \quad & \forall \alpha, \beta \in M, \quad \chi(\alpha, \beta) \chi(\beta, \alpha) = 1. \end{aligned}$$

Notation 1.1 In all the sequel, we shall adopt the following concise notation:

$$\forall f, g \in \mathcal{A} - \{0\} \quad \text{homogeneous,} \quad \chi_{f, g} = \chi(|f|, |g|).$$

and then we have:

$$\begin{aligned} \chi_{a, b, c} &= \chi_{a, b} \chi_{a, c}, & \chi_{a, b, c} &= \chi_{a, c} \chi_{b, c}, \\ \chi_{a, b} \chi_{b, a} &= 1, & \chi_{a, a} &= \pm 1. \end{aligned}$$

1.1 Jordan-brackets and Lie-brackets in associative algebras.

Let (A, M) be an associative graded algebra, and let χ a bicharacter on (A, M) .

Definition 1.3 The Lie bracketing on A is the binary operation on A defined by:

$$\forall f, g \text{ homogeneous } \in \mathcal{A}, \quad [f, g] = f \cdot g - \chi_{f, g} g \cdot f$$

The Jordan bracketing on A is the binary operation on A defined by:

$$\forall f, g \text{ homogeneous } \in \mathcal{A}, \quad \langle f, g \rangle = f \cdot g + \chi_{f, g} g \cdot f$$

Remark 1.1 We recover the "Lie superalgebras" of Scheunert as a particular case by putting:

$$M = \mathbb{Z}/2 = \{0, +1\}, \quad \text{and} \quad \chi_{f, g} = (-1)^{|f||g|}$$

Lemma 1.1 The following relation of antisymmetry (and symmetry) are satisfied:

$$\begin{aligned} [f, g] + \chi_{f, g} [g, f] &= 0 \\ \langle f, g \rangle - \chi_{f, g} \langle g, f \rangle &= 0 \end{aligned} \tag{1}$$

These are the only universal linear relation of order two between Lie brackets (resp. Jordan brackets).

Now we present the only universal linear relations of order three involving Lie brackets (resp. Jordan brackets):

Lemma 1.2 (“Jacobi-like relations”):

If χ is any bicharacter on any graded algebra A, M , the following relations are satisfied:

$$\begin{aligned} \chi_{c,a} [[a, b], c] + \chi_{a,b} [[b, c], a] + \chi_{b,c} [[c, a], b] &= 0, \\ \chi_{c,a} [\langle a, b \rangle, c] + \chi_{a,b} [\langle b, c \rangle, a] + \chi_{b,c} [\langle c, a \rangle, b] &= 0, \\ \chi_{c,a} [\langle a, b \rangle, c] + \chi_{a,b} [\langle b, c \rangle, a] + \chi_{b,c} \langle [c, a], b \rangle &= 0, \\ \chi_{c,a} \langle \langle a, b \rangle, c \rangle + \chi_{a,b} [[b, c], a] + \chi_{b,c} \langle \langle c, a \rangle, b \rangle &= 0. \end{aligned} \tag{2}$$

Proof

Let $i = \pm 1$. It will be convenient to use the following notation:

Notation 1.2

$$\{ \{ i, f, g \} \} = f.g - i \chi_{f,g} g.f.$$

In order to study extensively the universal linear relations of order three, we inquire for solving the following equation:

$$\alpha \chi_{a,c} \{ \{ i, \{ a, b \}, c \} \} + \beta \chi_{c,b} \{ \{ k, \{ c, a \}, b \} \} + \gamma \chi_{b,a} \{ \{ m, \{ n, b, c \}, a \} \} = 0$$

with the “indeterminated constants” $\alpha, \beta, \gamma \in K$, and $i, j, k, l, m, n, \in \{+1, -1\}$. In fact, each of these brackets can be developed as a linear combination of the six permutations of the word “a.b.c”. The involved coefficients can be presented along two boards as follows:

board of bicharacter’s:

	I	II	III
	$\{ \{ i, \{ a, b \}, c \} \}$	$\{ \{ k, \{ b, c \}, a \} \}$	$\{ \{ m, \{ n, c, a \}, b \} \}$
a.b.c	1	$\chi_{bc,a}$	0
b.c.a	0	1	$\chi_{ca,b}$
c.a.b	$\chi_{ab,c}$	0	1
a.c.b	0	$\chi_{bc,a} \chi_{b,c}$	$\chi_{c,a}$
b.a.c	$\chi_{a,b}$	0	$\chi_{ca,b} \chi_{c,a}$
c.b.a	$\chi_{ab,c} \chi_{a,b}$	$\chi_{b,c}$	0

board of symmetry signs:

	IV	V	VI
	$\{ \{ i, \{ a, b \}, c \} \}$	$\{ \{ k, \{ b, c \}, a \} \}$	$\{ \{ m, \{ n, c, a \}, b \} \}$
a.b.c	+1	-k	0
b.c.a	0	+1	-m
c.a.b	-i	0	+1
a.c.b	0	kl	-n
b.a.c	-j	0	mn
c.b.a	ij	-l	0

We can equalize the three columns of the board of bicharacter's, by multiplying the first column by $\chi_{a,c}$, the second by $\chi_{b,a}$, and the third by $\chi_{c,b}$, as follows:

	$\chi_{c,a}.I$	$\chi_{a,b}.II$	$\chi_{b,c}.III$
$a.b.c$	$\chi_{c,a}$	$\chi_{c,a}$	0
$b.c.a$	0	$\chi_{a,b}$	$\chi_{a,b}$
$c.a.b$	$\chi_{b,c}$	0	$\chi_{b,c}$
$a.c.b$	0	$\chi_{c,a}\chi_{b,c}$	$\chi_{c,a}\chi_{b,c}$
$b.a.c$	$\chi_{a,b}\chi_{c,a}$	0	$\chi_{a,b}\chi_{c,a}$
$c.b.a$	$\chi_{b,c}\chi_{a,b}$	$\chi_{b,c}\chi_{a,b}$	0

Now we want to equalize to 0 the linear combination:

$$\alpha\chi_{a,c} \left\{ \left\{ a, b \right\}, c \right\} + \beta\chi_{c,b} \left\{ \left\{ c, a \right\}, b \right\} + \gamma\chi_{b,a} \left\{ \left\{ b, c \right\}, a \right\}.$$

Let us simplify the common bicharacter factors. Then it suffices to identify the coefficients α, β and γ in order to satisfy the relation:

$$\alpha IV + \beta V + \gamma VI = 0.$$

It is clear that we have (up to a common multiple factor):

$$\alpha, \beta, \gamma \in \{-1, +1\}.$$

We deduce the only two possibilities:

$$\begin{aligned} \alpha = \beta = \gamma = 1 \\ \alpha = \beta = -\gamma = 1, \end{aligned}$$

(This last possibility giving two similary others, by circular permutation of the terms a, b and c). The first possibility gives: $k = m = i = 1$, and $n = l = j$.

1. if $n = j = l = 1$: $\chi_{c,a} [[a, b], c] + \chi_{a,b} [[b, c], a] + \chi_{b,c} [[c, a], b] = 0$,
2. if $n = j = l = -1$: $\chi_{c,a} [\langle a, b \rangle, c] + \chi_{a,b} [\langle b, c \rangle, a] + \chi_{b,c} [\langle c, a \rangle, b] = 0$,

and we have just recovered the two first Jacobi-like relations.

The second possibility gives: $k = -m = -i = 1$, and $n = -l = j$.

1. if $n = -l = j = 1$: $\chi_{c,a} \langle [a, b], c \rangle + \chi_{a,b} \langle [b, c], a \rangle - \chi_{b,c} \langle [c, a], b \rangle = 0$,
2. if $n = -l = j = -1$: $\chi_{c,a} \langle \langle a, b \rangle, c \rangle + \chi_{a,b} [[b, c], a] - \chi_{b,c} \langle \langle c, a \rangle, b \rangle = 0$,

and we have just recovered the two last Jacobi-like relations.



1.2 Abstract Lie-Jordan algebras.

First we define the notion of *abstract Lie-Jordan algebra*, and then we verify that it can always be recovered, by the construction of the subsection 1.1, as an algebra of Jordan-brackets and Lie-brackets in some associative algebra.

Definition 1.4 abstract Lie-Jordan algebras.

An abstract Lie-Jordan algebra is a M -graded K -vector space W , with two binary operations " \langle, \rangle " and " $[\cdot, \cdot]$ " of signature $W \times W \mapsto W$, that satisfy the following properties:

- *symmetry (resp. antisymmetry):*

$$\begin{aligned} \langle f, g \rangle - \chi_{f,g} \langle g, f \rangle &= 0 \\ [f, g] + \chi_{f,g} [g, f] &= 0 \end{aligned} \quad (3)$$

- *Jacobi-like relations:*

$$\chi_{c,a} [[a, b], c] + \chi_{a,b} [[b, c], a] + \chi_{b,c} [[c, a], b] = 0, \quad (4)$$

$$\chi_{c,a} [\langle a, b \rangle, c] + \chi_{a,b} [\langle b, c \rangle, a] + \chi_{b,c} [\langle c, a \rangle, b] = 0, \quad (5)$$

$$\chi_{c,a} \langle [a, b], c \rangle + \chi_{a,b} \langle [b, c], a \rangle - \chi_{b,c} \langle [c, a], b \rangle = 0, \quad (6)$$

$$\chi_{c,a} \langle \langle a, b \rangle, c \rangle + \chi_{a,b} [[b, c], a] - \chi_{b,c} \langle \langle c, a \rangle, b \rangle = 0. \quad (7)$$

Theorem 1 any Lie-Jordan abstract algebra W can be considered as an associative algebra with the associative product:

$$f \dagger g = \frac{1}{2}(\langle f, g \rangle + [f, g]). \quad (8)$$

Moreover, the abstract products " \langle, \rangle " and " $[\cdot, \cdot]$ " are identical to the Jordan and Lie products defined as in the preceding subsection in the associative algebra $(W, \text{"o"})$.

Proof:

- We prove first that " \dagger " is associative. Indeed we have, (by the property 8):

$$\begin{aligned} 4(f \dagger g) \dagger h &= \langle \langle f, g \rangle, h \rangle + \langle [f, g], h \rangle \\ &\quad + [\langle f, g \rangle, h] + [[f, g], h]. \end{aligned}$$

$$\begin{aligned} 4f \dagger (g \dagger h) &= \langle f, \langle g, h \rangle \rangle + \langle f, [g, h] \rangle \\ &\quad + [f, \langle g, h \rangle] + [f, [g, h]] \\ &= \chi_{f,g+h} \langle \langle g, h \rangle, f \rangle + \chi_{f,g+h} \langle [g, h], f \rangle \\ &\quad - \chi_{f,g+h} [\langle g, h \rangle, f] - \chi_{f,g+h} [[g, h], f]. \end{aligned}$$

by symmetry and antisymmetry (equation 3)

by the equalities $\chi_{h,f}\chi_{f,g+h} = \chi_{h,f}\chi_{f,g}\chi_{f,h} = \chi_{f,g}$, (definition 1.2), we deduce:

$$\begin{aligned} 4 \chi_{h,f}[(f \dagger g) \dagger h - f \dagger (g \dagger h)] &= \chi_{h,f} \langle \langle f, g \rangle, h \rangle - \chi_{f,g} \langle \langle g, h \rangle, f \rangle \\ &\quad + \chi_{h,f} \langle [f, g], h \rangle - \chi_{f,g} \langle [g, h], f \rangle \\ &\quad - \chi_{h,f} \langle \langle f, g \rangle, h \rangle + \chi_{f,g} \langle \langle g, h \rangle, f \rangle \\ &\quad - \chi_{h,f} \langle [f, g], h \rangle + \chi_{f,g} \langle [g, h], f \rangle. \end{aligned}$$

Using Jacobi-like relations (2), (for $h = b$, $f = c$ and $f = a$), we obtain:

$$\begin{aligned} 4 [\chi_{h,f}(f \dagger g) \dagger h - \chi_{h,f}f \dagger (g \dagger h)] &= +\chi_{g,h} \langle [h, f], g \rangle \\ &\quad + \chi_{g,h} \langle \langle h, f \rangle, g \rangle \\ &\quad - \chi_{g,h} \langle \langle h, f \rangle, g \rangle \\ &\quad - \chi_{g,h} \langle [h, f], g \rangle \\ &= 0. \end{aligned}$$

thus " \dagger " is associative.

- Now, by equation 8, and elementary verification, we have:

$$\begin{aligned} 2(a \dagger b - \chi_{a,b}b \dagger a) &= \langle a, b \rangle + [a, b] - \chi_{a,b} \langle b, a \rangle - \chi_{a,b} [b, a] \\ &= \langle a, b \rangle - \chi_{a,b} \langle b, a \rangle + [a, b] - \chi_{a,b} [b, a] \\ &= \langle a, b \rangle - \langle a, b \rangle + [a, b] + [a, b] \\ &= 2 [a, b]. \end{aligned}$$

In the same way, we can also verify:

$$2(a \dagger b + \chi_{a,b}b \dagger a) = 2 \langle a, b \rangle.$$

hence the abstract operations " \langle, \rangle " and " $[,]$ " are exactly the same as the Jordan and Lie brackets, defined in the last subsection, in the associative algebra (W, \dagger) . That achieves the proof of this theorem.



This result can be also be restated as follows:

Theorem 2 *Let \mathcal{G} be some Lie Algebra. Then the enveloping algebra $\mathcal{U}(\mathcal{G})$ is also the free Lie-Jordan extension of the Lie Algebra \mathcal{G} .*

2 Complete free finitely generated superalgebras.

This part of the paper is devoted to the complete free algebras generated by a finite set \mathcal{Z} of "letters".

We note \mathcal{Z}^* the set of words on \mathcal{Z} , i.e. the set of finite sequences of letters. The length of a word w is its length as sequence of letters. The empty word, of length 0, is noted indifferently ϵ or 1. \mathcal{Z}^\oplus is the free monoid generated by \mathcal{Z} .

We note $\mathcal{Z}^\&$ is the free abelian monoid generated by \mathcal{Z} .

2.1 Power Series on an alphabet Z .

Definition 2.1 We call power series on Z any formal sum of the form:

$$S = \sum_{w \in Z^*} \langle S|w \rangle w.$$

We note $K \ll Z \gg$ the set of all the power series. Indeed, a power series can be viewed as an application from the monoid Z^* in K . A power series P is called polynomial if its support $supp(P) = \{w \in Z^* | \langle P|w \rangle \neq 0\}$ is finite. The set of all polynomials is noted $K < Z \rangle$.

Definition 2.2 The Cauchy product of two power series S and T is defined as follows, by its coefficient on any $w \in Z^*$:

$$\langle S.T|w \rangle = \sum_{u,v \in Z^*, uv=w} \langle S|u \rangle \langle T|v \rangle.$$

(Remark that for any $w \in Z^*$, the index set $\{u, v \in Z^*, uv = w\}$ is finite.)

Observe that the canonical surjective map of Z^* onto Z^\oplus , called “multidegree”, gives a non trivial example of graded algebra structure on $K \ll Z \gg$ (The length application of Z^* onto \mathbb{N} gives on the same set of power series the usual graded algebra by the “degree”).

Notation 2.1 Duality.

The vector space $K \ll Z \gg$ is the dual vector space of $K < Z \rangle$, for the duality defined by the following bilinear form:

$$\langle S|P \rangle = \sum_{w \in Z^*} \langle S|w \rangle \langle P|w \rangle,$$

for all $S \in K \ll Z \gg$, and for all $P \in K < Z \rangle$.

Definition 2.3 The left internal derivation by some letter $x \in Z$, if the dual linear map of the right product by x . In other words, for any power series S and any polynomial P , we define:

$$\langle \partial_x S|P \rangle = \langle S|P.x \rangle$$

It clearly satisfies :

$$\begin{cases} \partial_x \epsilon = 0, \\ \partial_x(Sy) = \delta_x^y S = \begin{cases} S & \text{if } y = x, \\ 0 & \text{if } y \neq x, \end{cases} \\ \partial_x S = \sum_{w \in Z^*} \langle S|w \rangle \partial_x w. \end{cases}$$

Lemma 2.1 Reconstruction lemma.

Any power series S can be recovered from its left derivatives as follows:

$$S = \langle S|\epsilon \rangle + \sum_{x \in Z} (\partial_x S)x \tag{9}$$

$$eval(S) + \sum_{x \in Z} (\partial_x S)x, \tag{10}$$

where we have adopted the notation: $eval(S) = \langle S|\epsilon \rangle$.

2.2 Two-fold power series on $\mathcal{Z} \otimes \mathcal{Y}$.

Definition 2.4 We call “two-fold” power series on $\mathcal{Z} \otimes \mathcal{Y}$ any formal sum of the form:

$$\mathcal{H} = \sum_{u \in \mathcal{Z}^*, f \in \mathcal{Y}^*} u \otimes f.$$

We note $K \ll \mathcal{Z} \otimes \mathcal{Y} \gg$ the set of the “two-fold” power series. Indeed, a two-fold power series can be viewed as an application from the product monoid $\mathcal{Z}^* \otimes \mathcal{Y}^*$ in K . We note $K < \mathcal{Z} \otimes \mathcal{Y} >$ the set of the “two-fold” polynomials, i.e. “two-fold” power series of finite support.

Definition 2.5 The Cauchy product $\mathcal{H}\mathcal{K}$ of the two-fold power series \mathcal{H} and \mathcal{K} is defined as follows by its coefficients for each $w \otimes h \in \mathcal{Z}^* \otimes \mathcal{K}^*$:

$$\langle \mathcal{H}\mathcal{K} | w \otimes h \rangle = \sum_{\substack{u, v \in \mathcal{Z}^*, uv=w \\ f, g \in \mathcal{Y}^*, fg=h}} \langle \mathcal{H} | u \otimes f \rangle \langle \mathcal{K} | v \otimes g \rangle.$$

(Remark that for any fixed $w \otimes h \in \mathcal{Z}^* \otimes \mathcal{Y}^*$, the index set of the sum is finite).

Definition 2.6 The χ -product $\mathcal{H} \diamond \mathcal{K}$ of two power series \mathcal{H} and \mathcal{K} is defined as follows by its coefficients for each $w \otimes h \in \mathcal{Z}^* \otimes \mathcal{K}^*$:

$$\langle \mathcal{H} \diamond \mathcal{K} | w \otimes h \rangle = \sum_{\substack{u, v \in \mathcal{Z}^*, uv=w \\ f, g \in \mathcal{Y}^*, fg=h}} \chi_{f,v} \langle \mathcal{H} | u \otimes f \rangle \langle \mathcal{K} | v \otimes g \rangle.$$

(Remark that the index set of the sum is finite).

Hence the Cauchy product (resp. the χ -product) satisfies:

$$\begin{aligned} [(u \otimes f) \cdot (v \otimes g)] &= (uv \otimes fg), \\ [(u \otimes f) \diamond (v \otimes g)] &= \chi_{f,v} uv \otimes fg. \end{aligned}$$

We can also define a “ χ -product of order 3” as follows:

$$[(u \otimes f \otimes r) \diamond_3 (v \otimes g \otimes s)] = \chi_{f,v} \chi_{r,v} \chi_{r,g} (uv \otimes fg \otimes rs).$$

(In this formula, the bicharacter’s coefficients compute the “cost” of intercrossing v with rs , and g with s).

Notation 2.2 Duality.

The vector space $K \ll \mathcal{Z} \otimes \mathcal{Y} \gg$ is the dual vector space of $K < \mathcal{Z} \otimes \mathcal{Y} >$, for the duality defined by the following bilinear form:

$$\langle \mathcal{H} | \mathcal{P} \rangle = \sum_{w \otimes h \in \mathcal{Z}^* \otimes \mathcal{P}^*} \langle S | w \otimes h \rangle \langle P | w \otimes h \rangle.$$

for a two-fold power series \mathcal{H} and a polynomial \mathcal{P} .

Definition 2.7 The left internal derivation in $K \ll \mathcal{Z} \otimes \mathcal{Y} \gg$ by some letter $x \in \mathcal{Z}$, noted D_x , is the dual linear map of the right product by $x \otimes \epsilon$. In other words:

$$\langle D_x \mathcal{H} | \mathcal{P} \rangle = \langle \mathcal{H} | \mathcal{P} \cdot (x \otimes 1) \rangle .$$

It clearly satisfies :

$$\begin{cases} D_x(u \otimes f) = (\partial_x u) \otimes f, \\ D_x(\mathcal{H}) = \sum_{u \in \mathcal{Z}^*, f \in \mathcal{Y}^*} \langle \mathcal{H} | u \otimes f \rangle (\partial_x u) \otimes f. \end{cases}$$

Lemma 2.2 Reconstruction lemma for the two-fold power series.

Any two-fold power series \mathcal{H} can be recovered from its left derivatives as follows:

$$\mathcal{H} = \mathcal{H}|_\epsilon + \sum_{x \in \mathcal{Z}} (D_x \mathcal{H})(x \otimes 1)$$

where we set :

$$\mathcal{H}|_\epsilon = eval(\mathcal{H}) = \sum_{f \in \mathcal{Y}^*} \langle \mathcal{H} | \epsilon \otimes f \rangle \epsilon \otimes f$$

2.3 Partition's Coalgebra.

Notation 2.3 We adopt the following notation, for any $x \in \mathcal{Z}$, $w \in \mathcal{Z}^*$ and $S \in K \ll \mathcal{Z} \gg$:

$$\phi_x(w) = \chi_{x,w} w, \tag{11}$$

$$\phi_x(S) = \sum_{w \in \mathcal{Z}^*} \langle S | w \rangle \phi_x(w). \tag{12}$$

Notation 2.4 Similarly we set, for any $x \in \mathcal{Z}$, $u \in \mathcal{Z}^*$, $f \in \mathcal{Y}^*$ and $\mathcal{H} \in K \ll \mathcal{Z} \otimes \mathcal{Y} \gg$:

$$\phi_x(u \otimes f) = \chi_{x,u} u \otimes f, \tag{13}$$

$$\phi_x(\mathcal{H}) = \sum_{(u \otimes f) \in \mathcal{Z}^* \otimes \mathcal{Y}^*} \langle \mathcal{H} | u \otimes f \rangle \phi_x(u \otimes f). \tag{14}$$

Definition 2.8 (Partition's comultiplication.) The partition's comultiplication is the algebra morphism Γ from \mathcal{A} into $\mathcal{A} \otimes \mathcal{A}$ with χ -product generated by the following monoid morphisms:

$$\begin{aligned} \Gamma(1) &= 1 \otimes 1, \\ \text{for } x \in \mathcal{Z} : \quad \Gamma(x) &= 1 \otimes x + x \otimes 1, \\ \text{for } w \in \mathcal{Z}^*, x \in \mathcal{Z} : \quad \Gamma(wx) &= \Gamma(w) \diamond \Gamma(x). \end{aligned}$$

Lemma 2.3 The partition's comultiplication provides on $K \ll \mathcal{Z} \gg$ a structure of associative coalgebra, with as counity the linear map $eval : K \ll \mathcal{Z} \gg \rightarrow K$ defined as follows:

$$\text{for } S \in K \ll \mathcal{Z} \gg \quad eval(S) = \langle S | \epsilon \rangle (= \langle S | 1 \rangle).$$

Proof:

- Let us define the following a 3-coproduct as follows:

$$\begin{aligned} \text{if } x \in \mathcal{Z}, \quad \Gamma_3(x) &= x \otimes 1 \otimes 1 + 1 \otimes x \otimes 1 + 1 \otimes 1 \otimes x. \\ \text{if } v \in \mathcal{Z}^*, \quad \Gamma_3(vx) &= \Gamma_3(vx) \circ_3 \Gamma_3(x). \end{aligned}$$

It suffices to prove the following:

$$(\Gamma \otimes Id) \circ \Gamma = \Gamma_3 = (Id \otimes \Gamma) \circ \Gamma.$$

One verifies easily, for any $x \in \mathcal{Z}$ the equalities :

$$\begin{aligned} \Gamma_3(x) &= \Gamma(x) \otimes 1 + \Gamma(1) \otimes x = [(\Gamma \otimes Id) \circ \Gamma](x) \\ &= x \otimes \Gamma(1) + 1 \otimes \Gamma(x) = [(Id \otimes \Gamma) \circ \Gamma](x). \end{aligned}$$

One verifies also easily that $\Gamma \otimes Id$, Γ_3 and $Id \otimes \Gamma$ are monoid morphisms for the χ -product of order 3.



Definition 2.9 Shuffle product.

The shuffle product is the dual multiplication of the partition's comultiplication Γ . In other words, for any $u, v, w \in \mathcal{Z}^*$:

$$\langle u \sqcup v | w \rangle = \langle \Gamma(w) | u \otimes v \rangle.$$

Lemma 2.4 The shuffle product of two power series $S, T \in K \ll \mathcal{Z} \gg$ can be directly defined as follows:

$$\begin{aligned} \epsilon \sqcup u &= u = u \sqcup \epsilon, \\ \partial_x(S \sqcup T) &= S \sqcup (\partial_x T) + (\partial_x S) \sqcup \phi_x(T). \end{aligned}$$

and satisfies : $v \sqcup u = (-1)^{|u||v|} \chi_{u,v} u \sqcup v.$

Proposition 2.1 (Derivation property).

The left derivation ∂_x by a letter x is a derivation for the shuffle product \sqcup in the following sense:

$$\partial_x(S \sqcup T) = S \sqcup (\partial_x T) + (\partial_x S) \sqcup \phi_x(T)$$

Definition 2.10 (Bialgebras of two-fold power series.)

We define on $K \ll \mathcal{Z} \otimes \mathcal{Y} \gg$ the following comultiplication Π an multiplication \sqcup :

$$\begin{aligned} \Pi(u \otimes f) &= (\Gamma u) \otimes f, \quad (\text{for } u \in \mathcal{Z}^*, f \in \mathcal{Y}^*) \\ (u \otimes f) \sqcup (v \otimes g) &= (u \sqcup v) \otimes f.g, \end{aligned}$$

Lemma 2.5 (Derivation properties of two-fold power series).

The following properties are satisfied:

$$D_x(\mathcal{H} \sqcup \mathcal{K}) = \mathcal{H} \sqcup (D_x(\mathcal{K})) + (D_x \mathcal{H}) \sqcup (\phi_x(\mathcal{K})),$$

2.4 Primitive elements.

Definition 2.11 A power series $S \in K \ll \mathcal{Z} \gg$ is called primitive if it satisfies:

$$\Gamma(S) = S \otimes 1 + 1 \otimes S.$$

Proposition 2.2 Any Lie power series is primitive.

Proof:

- Indeed, the letters are primitive. So it suffices to prove that any Lie bracket of two homogeneous primitive elements is primitive.
- Suppose that f, g are homogeneous primitive elements. Then we have:

$$\begin{aligned} \Gamma(fg) &= \Gamma(f) \diamond \Gamma(g) \\ &= (1 \otimes f + f \otimes 1) \diamond (1 \otimes g + g \otimes 1) \\ &= 1 \otimes (f.g) + \chi_{f,g} g \otimes f + f \otimes g + (f.g)1 \otimes f \\ \Gamma(gf) &= 1 \otimes (g.f) + \chi_{g,f} f \otimes g + g \otimes f + (g.f)1 \otimes f \end{aligned}$$

and we obtain finally:

$$\begin{aligned} \Gamma([f, g]) &= \Gamma(fg) - \chi_{f,g} \Gamma(gf) \\ &= 1 \otimes (f.g - \chi_{f,g} g.f) + (f.g - \chi_{f,g} g.f) \otimes 1 \\ &= 1 \otimes \Gamma([f, g]) + \Gamma([f, g]) \otimes 1. \end{aligned}$$



Proposition 2.3 A two-fold power series \mathcal{H} is primitive if and only if it is orthogonal to any shuffle product $u \sqcup v$, where u and v are non empty words.

Proof:

- By definition of Γ , we have for any power series S :

$$\Gamma(S) = 1 \otimes S + S \otimes 1 + R,$$

with R is 'left and right proper' i.e. satisfies:

$$u \otimes v \in \text{support}(S) \Rightarrow u \neq \epsilon, \text{ and } v \neq \epsilon.$$

From the equality $\langle u \sqcup v | \mathcal{Q} \rangle = \langle \mathcal{Q} | u \otimes v \rangle$, we deduce that S is primitive if and only if $R = 0$, i.e. if and only if $\Gamma(S) = 1 \otimes S + S \otimes 1$.



2.5 Analysis of Relations.

Any function τ from Z^* in $K \ll Z \gg$ can also be described by the two-fold power series:

$$\sum_{w \in Z^*} w \otimes \tau(w).$$

More precisely, the power series description of τ can be viewed as a combinatorial tool allowing some algebraic computations on the involved function, or relation. For example, we set:

Definition 2.12 *The Lie bracketing l (resp. Jordan bracketing s) is inductively defined as follows, for any $x \in Z$ and any $v \in Z^*$:*

$$\begin{aligned} l(x) &= x, & l(vx) &= [l(v), x], \\ s(x) &= x, & s(vx) &= \langle l(v), x \rangle. \end{aligned}$$

Notation 2.5 *We define the following two-fold power series:*

$$\begin{aligned} \mathcal{N} &= \sum_{u \in Z^*} u \otimes u, & \mathcal{E}lm &= \sum_{x \in Z} x \otimes x, \\ \mathcal{L} &= \sum_{v \in Z^*, v \neq 1} v \otimes l(v), & \overline{\mathcal{N}} &= \sum_{u \in Z^*} (-1)^{|u|} u \otimes u, \\ & & \mathcal{S} &= \sum_{v \in Z^*, v \neq 1} v \otimes s(v). \end{aligned}$$

Lemma 2.6 *For any letter $x \in Z$, the following relations are satisfied:*

$$\begin{aligned} \partial_x \mathcal{N} &= \mathcal{N} \cdot (1 \otimes x), & \partial_x \mathcal{L} &= 1 \otimes x + \mathcal{L} \cdot (1 \otimes x) - (1 \otimes x) \cdot \phi_x(\mathcal{L}), \\ \partial_x \overline{\mathcal{N}} &= -\overline{\mathcal{N}} \cdot (1 \otimes x), & \partial_x \mathcal{S} &= 1 \otimes x + \mathcal{S} \cdot (1 \otimes x) + (1 \otimes x) \cdot \phi_x(\mathcal{S}). \end{aligned}$$

Proof:

- This equalities derived directly from the identities:

$$\begin{aligned} \mathcal{N} &= 1 \otimes 1 + \mathcal{N} \cdot \mathcal{E}lm, \\ \overline{\mathcal{N}} &= 1 \otimes 1 - \overline{\mathcal{N}} \cdot \mathcal{E}lm, \\ \mathcal{L} &= \mathcal{E}lm + \mathcal{L} \cdot \mathcal{E}lm - \sum_{x \in Z} (1 \otimes x) \cdot \phi_x(\mathcal{L}) \cdot (x \otimes 1), \\ \mathcal{S} &= \mathcal{E}lm + \mathcal{S} \cdot \mathcal{E}lm + \sum_{x \in Z} (1 \otimes x) \cdot \phi_x(\mathcal{S}) \cdot (x \otimes 1). \end{aligned}$$



Lemma 2.7 *The following identities hold:*

$$\begin{aligned} \mathcal{N} \sqcup \mathcal{L} = \mathcal{N} \wedge \mathcal{S} &= \sum_{w \in Z^*} |w| w \otimes w, \\ \overline{\mathcal{N}} \sqcup \mathcal{L} = \overline{\mathcal{N}} \wedge \mathcal{S} &= \sum_{w \in Z^*} j_w w \otimes w, \end{aligned}$$

where $j_w = \begin{cases} 1 & \text{if } |w| \text{ is odd,} \\ 0 & \text{if } |w| \text{ is even.} \end{cases}$

Proof:

- These identities are proved by simple use of derivation properties of the internal derivations D_x for $x \in \mathcal{Z}$, and by reconstruction lemma:

$$\begin{aligned}
 D_x(\mathcal{N} \sqcup \mathcal{L}) &= \mathcal{N} \sqcup (1 \otimes x) + \mathcal{N} \sqcup [\mathcal{L}.(1 \otimes x)] - \mathcal{N} \sqcup [(1 \otimes x).\phi_x(\mathcal{L})] \\
 &\quad + [\mathcal{N}.(1 \otimes x)] \sqcup \phi_x(\mathcal{L}), \\
 &= \mathcal{N}.(1 \otimes x) + (\mathcal{N} \sqcup \mathcal{L}).(1 \otimes x); \\
 D_x(\overline{\mathcal{N}} \sqcup \mathcal{S}) &= \overline{\mathcal{N}} \sqcup (1 \otimes x) + \overline{\mathcal{N}} \sqcup [\mathcal{S}.(1 \otimes x)] + \overline{\mathcal{N}} \sqcup [(1 \otimes x).\phi_x(\mathcal{S})] \\
 &\quad + [(-\overline{\mathcal{N}}).(1 \otimes x)] \sqcup \phi_x(\mathcal{S}), \\
 &= \overline{\mathcal{N}}.(1 \otimes x) + (\overline{\mathcal{N}} \sqcup \mathcal{S}).(1 \otimes x);
 \end{aligned}$$

After simplification, and by the reconstruction lemma, we deduce that the series $\mathcal{N} \sqcup \mathcal{L}$ satisfy the following recursive equation:

$$\begin{aligned}
 \mathcal{X} &= \mathcal{N}.Elm + \mathcal{X}.Elm \\
 &= (1 - Elm)^{-1}.Elm.(1 - Elm)^{-1} \\
 &= \sum_{k \in \mathbb{N}} k (Elm)^k.
 \end{aligned}$$

Similarly, the series $\overline{\mathcal{N}} \sqcup \mathcal{S}$ satisfy the following recursive equation:

$$\begin{aligned}
 \mathcal{X} &= \overline{\mathcal{N}}.Elm + \mathcal{X}.Elm \\
 &= (1 + Elm)^{-1}.Elm.(1 + Elm)^{-1} \\
 &= Elm.[1 - (Elm)^2]^{-1} \\
 &= \sum_{k \text{ odd} \in \mathbb{N}} n (Elm)^n.
 \end{aligned}$$

hence follows the lemma.



Theorem 3 We have for any primitive homogeneous polynomial Q of degree q :

$$l(Q) = q.Q, \quad s(Q) = j_q Q = \begin{cases} Q & \text{if } |Q| = \text{degree}(Q) \text{ is odd,} \\ 0 & \text{if } |Q| = \text{degree}(Q) \text{ is even.} \end{cases}$$

Proof:

- The preceding lemma can also be write:

$$\begin{aligned}
 \sum_{u,v \in \mathcal{Z}^*} (u \sqcup v) \otimes ul(v) &= \sum_{w \in \mathcal{Z}^*} |w| w \otimes w, \\
 \sum_{u,v \in \mathcal{Z}^*} (-1)^{|u|} (u \sqcup v) \otimes us(v) &= \sum_{w \in \mathcal{Z}^*} j_w w \otimes w.
 \end{aligned}$$

By duality, we obtain, for any polynomial S :

$$\begin{aligned}
 \sum_{u,v \in \mathcal{Z}^*} \langle S|u \sqcup v \rangle ul(v) &= \sum_{w \in \mathcal{Z}^*} |w| \langle S|w \rangle w, \\
 \sum_{u,v \in \mathcal{Z}^*} (-1)^{|u|} \langle S|u \sqcup v \rangle us(v) &= \sum_{w \in \mathcal{Z}^*} j_w \langle S|w \rangle w.
 \end{aligned}$$

Now, if $S = Q$ is primitive, $\langle Q|u\omega v \rangle = 0$ if $u = \epsilon$, or if $v = \epsilon$. Moreover $l(\epsilon) = s(\epsilon) = 0$. So we obtain:

$$\begin{aligned} \sum_{v \in \mathcal{Z}^*} \langle Q|v \rangle l(v) &= \sum_{w \in \mathcal{Z}^*} |w| \langle Q|w \rangle w, \\ \sum_{u, v \in \mathcal{Z}^*} \langle Q|v \rangle s(v) &= \sum_{w \in \mathcal{Z}^*} j_w \langle Q|w \rangle w, \end{aligned}$$

hence follows the result.



From the theorem 3, since any Lie element is primitive, it follows:

Corollary 2.1 For any homogeneous element P , we have:

$$l(l(P)) = |P| P, \quad s(l(P)) = \begin{cases} l(P) & \text{if } |P| \text{ is odd,} \\ 0 & \text{if } |P| \text{ is even.} \end{cases}$$

This last corollary is equivalent to an identity obtain by P. Crouch and F. Lamnabhi-Lagarrigue [2], using the following definition:

Definition 2.13 Let $t : K \langle \mathcal{Z} \rangle \rightarrow K \langle \mathcal{Z} \rangle$ the graded map inductively defined as follows (for any $z \in \mathcal{Z}$ and $u \in \mathcal{Z}^*$):

$$\begin{aligned} t(z) &= z, \\ \partial_x(t(u)) &= \langle t|\partial_x u \rangle z - \langle t|\eta_x u \rangle z, \end{aligned}$$

where the right internal derivation η_x is defined by:

$$\forall z \in \mathcal{Z}, \forall u, v \in \mathcal{Z}^*, \quad \langle \eta_x u|v \rangle = \langle u|zv \rangle.$$

In other words, we have:

$$\begin{aligned} \forall z \in \mathcal{Z}, \quad t(z) &= z, \\ \forall x, y \in \mathcal{Z}, \forall u \in \mathcal{Z}^* \quad \langle t|xyu \rangle &= t(xu)y - \chi_{x,uy} t(uy)x \end{aligned}$$

Lemma 2.8 The graded map “ t ” is the inverse relation of the graded map “ q ”. In other words, we have:

$$\sum_{u \in \mathcal{Z}^*} t(u) \otimes u = \sum_{v \in \mathcal{Z}^*} v \otimes l(v).$$

Proof:

- Let $\mathcal{T} = \sum_{u \in \mathcal{Z}^*} t(u) \otimes u$. It suffices to prove the equality of \mathcal{T} and \mathcal{L} . Now we have clearly:

$$\mathcal{T}|_{\epsilon} = 0 = \mathcal{L}|_{\epsilon}.$$

On the other hand we have, for any $x \in \mathcal{Z}$:

$$\begin{aligned} D_x \mathcal{T} &= 1 \otimes x + \sum_{u \in \mathcal{Z}^*} (\partial_x \langle t|u \rangle) \otimes u \\ &= 1 \otimes x + \sum_{u \in \mathcal{Z}^*} \langle t|\partial_x u \rangle \otimes u - \sum_{u \in \mathcal{Z}^*} \langle t|\eta_x u \rangle \otimes u \\ &= 1 \otimes x + \sum_{v \in \mathcal{Z}^*} \langle t|v \rangle \otimes vx - \sum_{v \in \mathcal{Z}^*} \langle t|v \rangle \otimes xv \\ &= 1 \otimes x + \mathcal{T} \cdot (1 \otimes x) - \chi_{x,v} (1 \otimes x) \mathcal{T}. \end{aligned}$$

Then the two-fold series \mathcal{T} and \mathcal{L} satisfy the same constructive equation, and consequently are equal.

2.6 Classical consequences.

We recall briefly the classical results that outcome quasi-immediately from the results of the preceding subsection.

Theorem 4 *The following properties are equivalent for a power series $H \in K \ll \mathcal{Z} \gg$ (in any characteristic):*

1. H is the exponential of a primitive element,
2. $\Gamma(H) = H \otimes H$,
3. H satisfies the "Friedrich criterium":

$$\forall u, v \in \mathcal{Z}^*, \quad \langle H|u \cup v \rangle = \langle H|u \rangle \langle H|v \rangle .$$

In characteristic 0, these are equivalent to:

4. H is the exponential of a Lie element.

Proof:

– The implication (1) \Leftrightarrow (2) results by equivalence of the following equalities:

$$\begin{aligned} \Gamma(S) &= 1 \otimes S + S \otimes 1 \\ \exp(\Gamma(S)) &= \exp(1 \otimes S + S \otimes 1) = \exp(1 \otimes S) \cdot \exp(S \otimes 1) \\ &\quad \text{(because } 1 \otimes S \text{ and } S \otimes 1 \text{ commute),} \\ \Gamma(\exp(S)) &= [1 \otimes \exp(S)] \cdot [\exp(S) \otimes 1], \\ \Gamma(\exp(S)) &= \exp(S) \otimes \exp(S). \end{aligned}$$

– The implication (1) \Leftrightarrow (2) results by equivalence of the following equalities (for $H \in K \ll \mathcal{Z} \gg$):

$$\begin{aligned} \Gamma(H) &= H \otimes H \\ \forall u, v \in \mathcal{Z}^*, \quad \langle \Gamma(H)|u \otimes v \rangle &= \langle H \otimes H|u \otimes v \rangle \\ \forall u, v \in \mathcal{Z}^*, \quad \langle H|u \cup v \rangle &= \langle H|u \rangle \langle H|v \rangle . \end{aligned}$$



The result on Lie (resp. Jordan) bracketting s is in [5,6] (resp. [2]).

References

- [1] Choquet-Bruhat Y., "Mathematics for classical supergravities", Advances in Mathematics, 26, (1977).
- [2] Crouch P. and Lamnabhi-Lagarrigue F., "Algebraic and multiple integral identities", Acta Mathematica Applicandae vol.25 (1989) pp.235-274.
- [3] Jacobson N., "Structure and representations of Jordan Algebras", Amer. Math. Soc., 1968.
- [4] Kaç, "Lie Superalgebras, Advances in Mathematics", 26, (1977), 8-96.

- [5] Ree. R., "*Lie elements and an algebra associated with shuffle*", Ann. of Math., 68 (1958), p.210-220.
- [6] Ree R., "*Generalized Lie algebras*", Canadian J. Math., 12 (1960), pp.493-502., (1965)
- [7] Sanchez-Valenzuela O.A. and Sternberg S., The automorphism group of a hermitian superalgebra, in "*Differential Geometric Methods in Mathematical Physics*", Proceedings of the 14th International Conference in Salamanca, Spain, June 24-29, 1985.
- [8] Scheunert M., "*The theory of Lie superalgebras*", Lect. Notes in Math. n 716.

VARIATIONS COMBINATOIRES AUTOUR DES ARBORESCENCES HYPERQUATERNAIRES

Gilbert Labelle et Louise Laforest, LACIM-UQAM

1. Introduction

L'étude des arborescences hyperquaternaires* de points [5, 12] (une généralisation des arborescences binaires de fouille [15]) nous amène d'abord naturellement à l'étude de deux autres modèles plus simples, moins contraignants du point de vue combinatoire, soit les arborescences hyperquaternaires de type Catalan et les arborescences hyperquaternaires croissantes. Pour fins de comparaison et de simplification de l'exposé, nous avons choisi de désigner par les mêmes lettres les espèces correspondantes dans les trois modèles étudiés.

Le premier modèle, celui des arborescences hyperquaternaires de type Catalan, est l'espèce T caractérisée par les équations combinatoires

$$T = 1 + A, \quad A = XR(A) \tag{1.1}$$

où l'espèce « enrichissante » R est donnée par $R = (1 + X)^{2^d}$. Il s'agit donc essentiellement d'espèces d'arborescences R -enrichies au sens de Labelle [10].

La deuxième variante est la L -espèce T caractérisée par les équations combinatoires

$$T = 1 + A, \quad A' = R(A) \tag{1.2}$$

où l'espèce enrichissante est encore une fois donnée par $R = (1 + X)^{2^d}$. Dans ce cas, les arborescences considérées sont « croissantes » (i.e., l'ensemble sous-jacent est préalablement muni d'un ordre total et les sommets vont en ordre croissant de la racine aux feuilles). Il s'agit donc essentiellement d'espèces d'arborescences croissantes R -enrichies au sens de Leroux-Viennot [13] (voir aussi [1]).

Nous présentons, pour ces deux variantes, diverses équations combinatoires donnant automatiquement lieu, par passage aux séries génératrices, à des formules explicites ou asymptotiques pour la probabilité d'avoir k noeuds dans un « hyperquadrant » donné, l'espérance et la variance du nombre de feuilles.

* « quadrees » multidimensionnels.

La troisième variante de la notion d'arborescence hyperquaternaire fait appel aux d -tuplets de permutations en considérant l'espèce produit

$$T = L \times L \times \dots \times L \quad (d \text{ facteurs}) \tag{1.3}$$

où L est l'espèce des ordres linéaires. Les concepts de feuille, noeud ayant k enfants, poids d'une sous-arborescence donnée sont codés de façon différente des deux premiers modèles et donnent lieu à d'autres probabilités, espérances et variances. Par exemple, nous donnons plusieurs expressions (explicites et récursives) pour la probabilité $\pi_{n,k}$ qu'une sous-arborescence donnée d'une arborescence hyperquaternaire de n noeuds, à d dimensions ait k noeuds. Nous montrons, entre autres que

$$\pi_{n,k} = \binom{n-1}{k} \sum_{i=0}^{n-1-k} \binom{n-1-k}{i} \frac{(-1)^i}{(k+i+1)^d} \tag{1.4}$$

et que ces nombres sont liés aux nombres harmoniques

$$H_n^{(k)} = \sum_{i=1}^n \frac{1}{i^k} \tag{1.5}$$

Nous donnons aussi des formules explicites et asymptotiques pour l'espérance du nombre de noeuds ayant k enfants dans le cas bidimensionnel, $0 \leq k \leq 4$.

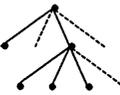
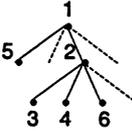
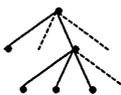
	Catalan	Croissantes	de points
$d = 1$	 132	 6! = 720	 6! = 720 (10)
$d = 2$	 7084	 58 240	 6!^2 = 518 400 (4)
	$T = 1 + A$ $A = XR(A)$ $R = (1 + X)^\delta$ $T = 1 + XT^\delta$	$T = 1 + A$ $A' = R(A)$ $R = (1 + X)^\delta$ $T' = T^\delta$	$T = L \times L \times \dots \times L$ $(d \text{ facteurs})$

Tableau 1

Le tableau 1 nous donne des exemples d'arborescences hyperquaternaires sur six points pour les trois modèles, à une et deux dimensions, ainsi que les équations combinatoires associées à chacun de ceux-ci en dimension d quelconque. Nous avons mis, pour chaque dimension et chaque modèle, le nombre de telles arborescences sur six points. Le nombre entre parenthèses dans le cas des arborescences hyperquaternaires de points correspond au nombre de permutations, dans le cas à une dimension et au nombre de couples de permutations dans le cas à deux dimensions, engendrant l'arborescence donnée en exemple. Nous avons aussi posé $\delta = 2^d$. Cette convention sera utilisée tout au long de cet article.

2. Arborescences hyperquaternaires de type Catalan

Ces arborescences sont dénombrées par la fonction génératrice exponentielle

$$T(x) = 1 + A(x) \quad \text{où} \quad A(x) = \sum_{n \geq 1} a_n \frac{x^n}{n!} = xR(A(x)). \quad (2.1)$$

En utilisant la formule d'inversion de Lagrange, avec l'enrichissement $R = (1+X)^\delta$, $\delta = 2^d$, on obtient l'expression suivante pour le nombre de telles arborescences (non-étiquetées) sur n sommets

$$\frac{a_n}{n!} = \frac{1}{(\delta-1)n+1} \binom{\delta n}{n}, \quad n > 0. \quad (2.2)$$

Pour pouvoir calculer les quantités qui nous intéressent, à savoir la probabilité $\pi_{n,k}$ que la première sous-arborescence contienne k des n points, l'espérance et la variance du nombre de feuilles, nous introduisons les nouvelles espèces suivantes:

$B = A^\clubsuit$ = l'espèce des arborescences hyperquaternaires de type Catalan non-vides à une feuille pointée,

$C = A^{\clubsuit\clubsuit}$ = l'espèce des arborescences hyperquaternaires de type Catalan non-vides à deux feuilles pointées (les deux feuilles étant ordonnées et possiblement égales),

T_k = l'espèce des arborescences hyperquaternaires de type Catalan sur k points, ($k \geq 0$),

$T_{(k)}$ = l'espèce des arborescences hyperquaternaires de type Catalan ayant k points dans la première sous-arborescence ($k \geq 0$).

Les relations suivantes se déduisent facilement

$$T_{(0)} = 1 + XT_0 T^{\delta-1} = 1 + A_{(0)} \quad \text{et} \quad T_{(k)} = XT_k T^{\delta-1} = A_{(k)} \quad \text{pour } k > 0,$$

$$A_{(0)} = X(1+A)^{\delta-1}, \quad T_0 = 1 \quad \text{et} \quad A_0 = 0,$$

$$A_{(k)} = XA_k(1+A)^{\delta-1}, \quad T_k = A_k \quad \text{pour } k > 0.$$

Convenons de désigner les séries génératrices des espèces B et C par

$$B(x) = \sum_{n \geq 0} \frac{a_n^\spadesuit}{n!} x^n \quad \text{et} \quad C(x) = \sum_{n \geq 0} \frac{a_n^{\spadesuit\spadesuit}}{n!} x^n.$$

Théorème 1 Pour un enrichissement $R = R(X)$ quelconque, les trois espèces $A = A(X)$, $B = A^\spadesuit(X)$ et $C = A^{\spadesuit\spadesuit}(X)$, sont caractérisées par les équations combinatoires

$$A = XR(A), \quad (2.3)$$

$$B = r_0X + XR'(A)B + \frac{r_0X}{1 - XR'(A)}, \quad (2.4)$$

$$C = B + \frac{XR''(A)}{1 - XR'(A)} B^2, \quad (2.5)$$

où r_0 désigne le nombre de R -structures sur l'ensemble vide.

Preuve : Il suffit d'appliquer les techniques habituelles de la théorie des espèces [10]. ■

En passant aux séries génératrices, on pose, pour $k > 0$,

$$T_{(0)}(x) = 1 + A_{(0)}(x) = 1 + \sum_{n \geq 1} \frac{a_{n,0}}{n!} x^n \quad \text{et} \quad T_{(k)}(x) = A_{(k)}(x) = \sum_{n \geq 1} \frac{a_{n,k}}{n!} x^n,$$

$$A_{(0)}(x) = x(1+A(x))^{\delta-1} \quad \text{et} \quad A_{(k)}(x) = x \frac{a_k}{k!} x^k (1+A(x))^{\delta-1}.$$

Pour calculer $(1+A(x))^{\delta-1}$, on utilise la version composée [10] de l'inversion de Lagrange: $A(x) = xR(A(x))$ entraîne que

$$F(A(x)) = F(0) + \sum_{n \geq 1} \left(\frac{d^{n-1}}{dt^{n-1}} F'(t)(R(t))^n \right)_{t=0} \frac{x^n}{n!},$$

avec $R(x) = (1+x)^\delta$ et $F(x) = (1+x)^{\delta-1}$. On trouve

$$\pi_{n,k} = \begin{cases} 1 & \text{si } n = k = 0, \\ a_{n,k}/n! & \text{si } n > k \geq 0, \\ a_n/n! & \end{cases} \quad (2.6)$$

$$\frac{a_k}{k!} = \begin{cases} 0 & \text{si } k=0, \\ \frac{1}{(\delta-1)k+1} \binom{\delta k}{k} & \text{si } k>0, \end{cases} \quad \frac{a_{n,k}}{n!} = \begin{cases} \frac{(\delta-1)(n\delta-2)^{(n-2)}}{(n-1)!} & \text{si } k=0, \\ \frac{a_k (\delta-1)((n-k)\delta-2)^{(n-k-2)}}{k! (n-k-1)!} & \text{si } k>0, \end{cases} \quad (2.7)$$

où $x^{(n)} = x(x-1)(x-2)\dots(x-n+1)$ désigne la $n^{\text{ème}}$ puissance factorielle descendante de x . En appliquant le théorème 1 et les méthodes de Gilbert Labelle [10], on obtient l'espérance mathématique du nombre de feuilles

$$\varepsilon_{n,d} = \frac{a_n^*/n!}{a_n/n!} = r_0 n \frac{r_{n-1}(n-1)}{r_{n-1}(n)} \quad (2.8)$$

où

$$(R(x))^\lambda = \sum_{n \geq 0} r_n(\lambda) \frac{x^n}{n!}, \quad R(x) = (1+x)^\delta, \quad r_n(\lambda) = (\lambda\delta)^{(n)}.$$

Après calculs, on obtient la formule explicite

$$\varepsilon_{n,d} = n \left(1 - \frac{\delta}{n\delta}\right) \left(1 - \frac{\delta}{n\delta-1}\right) \dots \left(1 - \frac{\delta}{n\delta-n+2}\right).$$

De même, la variance d'une variable aléatoire étant donnée en fonction des deux premiers moments factoriels $\mu^{(1)}$ et $\mu^{(2)}$ par la formule

$$\sigma^2 = \mu^{(2)} + \mu^{(1)} - (\mu^{(1)})^2,$$

on obtient la variance du nombre de feuilles comme suit: on calcule tout d'abord

$$\mu_{n,d}^{(2)} = \frac{a_n^{**}/n! - a_n^*/n!}{a_n/n!} = r_0^2 n(n-1) \frac{r_{n-1}(n-2)}{r_{n-1}(n)} = n(n-1) \left(1 - \frac{2\delta}{n\delta}\right) \dots \left(1 - \frac{2\delta}{n\delta-n+2}\right)$$

et la variance cherchée peut s'écrire sous la forme

$$\sigma_{n,d}^2 = n(n-1) \left(1 - \frac{2\delta}{n\delta}\right) \dots \left(1 - \frac{2\delta}{n\delta-n+2}\right) + n \left(1 - \frac{\delta}{n\delta}\right) \left(1 - \frac{\delta}{n\delta-1}\right) \dots \left(1 - \frac{\delta}{n\delta-n+2}\right) - n^2 \left(1 - \frac{\delta}{n\delta}\right)^2 \left(1 - \frac{\delta}{n\delta-1}\right)^2 \dots \left(1 - \frac{\delta}{n\delta-n+2}\right)^2. \quad (2.9)$$

En utilisant les formules de Stirling et sommatoire d'Euler-Maclaurin [8] on en déduit les résultats asymptotiques

$$\frac{a_n}{n!} \sim \left(\frac{\delta}{2\pi(\delta-1)^3}\right)^{1/2} \left(\frac{\delta^\delta}{(\delta-1)^{\delta-1}}\right)^n n^{-3/2}, \quad (2.10)$$

$$\epsilon_{n,d} \sim \left(1 - \frac{1}{\delta}\right)^\delta n, \quad \sigma_{n,d}^2 \sim \left[1 - \left(1 - \frac{1}{\delta}\right)^\delta \left(\frac{2\delta - 1}{\delta - 1}\right)\right] \left(1 - \frac{1}{\delta}\right)^\delta n, \quad n \rightarrow \infty. \quad (2.11)$$

3. Arborescences hyperquaternaires croissantes

Les arborescences croissantes pourraient être vue comme des arborescences de type Catalan dont on numérote les n sommets, de 1 à n , de façon telle que les étiquettes rencontrées de la racine à une feuille forment une suite croissante, et ce, pour toutes les feuilles, d'où le nom d'arborescences croissantes. Ces arborescences sont dénombrées de la façon suivante : $T(x) = 1 + A(x)$ où

$$A(x) = \sum_{n \geq 1} a_n \frac{x^n}{n!} \text{ satisfait } A'(x) = R(A(x)) = (1 + A(x))^\delta.$$

En résolvant cette équation différentielle on trouve

$$T(x) = (1 - (\delta - 1)x)^{\frac{-1}{\delta - 1}}, \quad (3.1)$$

d'où

$$a_n = \delta(2\delta - 1)(3\delta - 2) \dots ((n - 1)\delta - (n - 2)). \quad (3.2)$$

En utilisant des L -espèces auxiliaires analogues à celles introduites pour les arborescences de type Catalan, on vérifie que

$$T_{(0)} = 1 + \int_0^X T_0 T^{\delta - 1} dX, \quad T_{(k)} = \int_0^X T_k T^{\delta - 1} dX, \quad \text{pour } k > 0.$$

Après calculs on trouve

$$\pi_{n,k} = \begin{cases} 1 & n = k = 0, \\ (\delta - 1)^{n - k - 1} \frac{1}{k!n} \frac{1}{(1 + k(\delta - 1))} \dots \frac{1}{(1 + (n - 1)(\delta - 1))} & n > k \geq 0. \end{cases} \quad (3.3)$$

Théorème 2 Pour un enrichissement $R = R(X)$ quelconque, les trois L -espèces $A = A(X)$, $B = A^*(X)$ et $C = A^{**}(X)$, sont caractérisées par les équations combinatoires

$$A' = R(A), \quad (3.4)$$

$$B' = r_0 + R'(A)B = r_0 + \frac{A''}{A'} B, \quad (3.5)$$

$$C' = r_0 + R''(A)B^2 + R'(A)C = r_0 + \frac{A'A''' - A''^2}{A'^3} B^2 + \frac{A''}{A'} C, \quad (3.6)$$

où r_0 désigne le nombre de R -structures sur l'ensemble vide.

Preuve : Il suffit d'appliquer les techniques habituelles de la théorie des L -espèces [13]. ■

En résolvant à l'aide de Maple les équations différentielles du théorème 2, on trouve les séries

$$B(x) = \sum_{n \geq 1} \frac{a_n^\star}{n!} x^n = \frac{1}{2\delta-1} \left((1-(\delta-1)x)^{\frac{-\delta}{\delta-1}} - (1-(\delta-1)x) \right),$$

$$C(x) = \sum_{n \geq 1} \frac{a_n^{\star\star}}{n!} x^n = \frac{-(\delta-1)}{(3\delta-1)(2\delta-1)} Z^{\frac{-\delta}{\delta-1}} - \frac{1}{2\delta-1} Z + \frac{\delta}{(2\delta-1)^2} Z^{\frac{-2\delta+1}{\delta-1}}$$

$$- \frac{\delta(\delta-1)}{(3\delta-1)(2\delta-1)^2} Z^{\frac{2\delta-1}{\delta-1}} + \frac{2(\delta-1)}{(2\delta-1)^2},$$

où $Z = 1 - (\delta-1)x$. On obtient, après calculs les espérances et variances suivantes

$$\epsilon_{n,d} = \frac{a_n^\star}{a_n} = \begin{cases} 1 & n=1, \\ \frac{(\delta-1)n+1}{2\delta-1} & n > 1, \end{cases} \quad (3.7)$$

$$\sigma_{n,d}^2 = \frac{a_n^{\star\star}}{a_n} - \left(\frac{a_n^\star}{a_n} \right)^2 = \begin{cases} 0 & n \leq 2, \\ \frac{\delta(\delta-1)}{(3\delta-1)(2\delta-1)^2} \left[(\delta-1)n+1 + \prod_{i=1}^{n-3} \frac{i(\delta-1)-1}{i(\delta-1)+(2\delta-1)} \right] & n > 2, \end{cases} \quad (3.8)$$

En utilisant encore une fois la formule sommatoire d'Euler-Maclaurin et la méthode des singularités dominantes on en déduit les résultats asymptotiques

$$a_n \sim \frac{\sqrt{2\pi}}{\Gamma\left(\frac{1}{\delta-1}\right)} \left(\frac{\delta-1}{e} \right)^n n^{n + \frac{3-\delta}{2(\delta-1)}}, \quad (3.9)$$

$$\epsilon_{n,d} \sim \frac{\delta-1}{2\delta-1} n, \quad \sigma_{n,d}^2 \sim \frac{\delta(\delta-1)^2}{(3\delta-1)(2\delta-1)^2} n, \quad n \rightarrow \infty. \quad (3.10)$$

4. Arborences hyperquaternaires de points

Les équations combinatoires (1.1) et (1.2) permettant de calculer récursivement les restrictions T_n aux diverses cardinalités pour les deux premiers modèles, correspondent, dans le cas à deux dimensions à

$$T_{n+1} = X \sum_{n_1+n_2+n_3+n_4=n} T_{n_1} T_{n_2} T_{n_3} T_{n_4}, \quad T_{n+1} = \int \sum_{n_1+n_2+n_3+n_4=n} T_{n_1} T_{n_2} T_{n_3} T_{n_4} dX. \quad (4.1)$$

La première de ces deux équations revient à une équation fonctionnelle tandis que la deuxième équivaut à une équation différentielle. Cependant, le troisième modèle étant défini par

$$T = L \times L \times \dots \times L = T_0 + T_1 + T_2 + \dots, \tag{4.2}$$

on a plutôt le schéma récursif suivant

$$T_n = \int \sum_{n_1+n_2+n_3+n_4=n-1} \binom{n_1+n_2}{n_1} \binom{n_1+n_3}{n_3} \binom{n_2+n_4}{n_2} \binom{n_3+n_4}{n_4} T_{n_1} T_{n_2} T_{n_3} T_{n_4} dX \tag{4.3}$$

qui se démontre par des arguments semblables à ceux donnés dans Laforest [12]. Ce schéma ne correspond ni à une équation fonctionnelle, ni à une équation différentielle et son analogue en d dimensions est donné par

Théorème 3 La L -espèce $T = L \times L \times \dots \times L$ (d facteurs) satisfait le schéma récursif $T_0 = 1$ et, pour $n > 0$,

$$T_n = \int \sum_{\varepsilon} \prod_{\substack{\alpha=0,1 \\ 1 \leq i \leq d}} \mu_i^\alpha(\nu) \prod_{\varepsilon} T_{\nu_\varepsilon} dX, \quad (\varepsilon : [d] \rightarrow 2) \tag{4.4}$$

$$\text{où } \mu_i^\alpha(\nu) = \kappa_i^\alpha / \prod_{\varepsilon_i=\alpha} \nu_\varepsilon! \quad \text{et} \quad \kappa_i^\alpha = \sum_{\varepsilon_i=\alpha} \nu_\varepsilon, \quad \nu : 2^{[d]} \rightarrow \mathbb{N}. \tag{4.5}$$

Preuve : Dans le cas où $d = 2$, l'expression (4.4) est une réécriture de (4.3) où on a simplement remplacé n_1, n_2, n_3, n_4 par $\nu_{00}, \nu_{01}, \nu_{10}, \nu_{11}$ respectivement. Les quatre valeurs de $\mu_1^0(\nu), \mu_1^1(\nu), \mu_2^0(\nu), \mu_2^1(\nu)$ correspondent alors aux quatre coefficients binomiaux rencontrés dans (4.3). Le cas général (en dimension d quelconque) découle d'une analyse semblable basée sur le fait que la racine de l'arborescence divise l'espace en 2^d hyperquadrants et que les $n - 1$ points restants doivent être placés selon une distribution $\nu : 2^{[d]} \rightarrow \mathbb{N}$ satisfaisant la contrainte $\sum \nu_\varepsilon = n - 1$. ■

Théorème 4 En dimension d , les probabilités $\pi_{n,k,d}$ qu'une arborescence aléatoire hyperquaternaire de n points possède exactement k noeuds dans un hyperquadrant donné sont fournies par les formules explicites et récursives suivantes, pour $0 \leq k < n, d \geq 0$,

$$\pi_{n,k,0} = \delta_k^{n-1}, \quad (\text{où } \delta_j^i \text{ est le delta de Kronecker}), \tag{4.6}$$

$$\pi_{n,k,d+1} = \sum_{k < i \leq n} \frac{\pi_{n,i-1,d}}{i}, \tag{4.7}$$

$$\pi_{n,k,d+1} = \frac{1}{n} \sum_{k < i_1 \leq i_2 \leq \dots \leq i_d \leq n} \frac{1}{i_1 i_2 \dots i_d}, \tag{4.8}$$

$$\pi_{n,k,d} = [z^{d-1}] \frac{1}{n} \prod_{i=k+1}^n \frac{1}{1-z/i}, \quad d > 0, \quad (4.9)$$

$$\pi_{n,k,d+1} = \frac{1}{n} \sum_{n_1+2n_2+\dots+dn_d=d} \frac{\Delta_{k,1}^{n_1} \Delta_{k,2}^{n_2} \dots \Delta_{k,d}^{n_d}}{1^{n_1} n_1! 2^{n_2} n_2! \dots d^{n_d} n_d!}, \quad (4.10)$$

$$\text{où} \quad \Delta_{i,j} = H_n^{(j)} - H_i^{(j)} = \frac{1}{n^j} + \frac{1}{(n-1)^j} + \dots + \frac{1}{(i+1)^j}, \quad (4.11)$$

$$\pi_{n,k,d} = \binom{n-1}{k} \int_0^1 \dots \int_0^1 (u_1 \dots u_d)^k (1-u_1 \dots u_d)^{n-1-k} du_1 \dots du_d, \quad d > 0, \quad (4.12)$$

$$\pi_{n,k,d} = \binom{n-1}{k} \int_0^1 t^k (1-t)^{n-1-k} \frac{(-\ln t)^{d-1}}{(d-1)!} dt, \quad d > 0, \quad (4.13)$$

$$\pi_{n,k,d} = \binom{n-1}{k} \sum_{i=0}^{n-1-k} \binom{n-1-k}{i} \frac{(-1)^i}{(k+i+1)^d}, \quad (4.14)$$

$$\left\{ \begin{array}{l} \pi_{n,n-1,d} = \frac{1}{n^d}, \\ \pi_{n,k-1,d} = \frac{(n-1)\pi_{n-1,k-1,d} - k\pi_{n,k,d}}{n-k} \quad \text{pour } k = n-1, \dots, 1. \end{array} \right. \quad (4.15)$$

Preuve : Les formules (4.6) et (4.7) sont démontrées dans Laforest [12]. La formule (4.8) découle de (4.7) par itération. L'expression (4.9) s'obtient en développant le produit et en comparant avec (4.8). La formule (4.10) découle du produit (4.9) en remarquant que chaque facteur peut s'écrire $(1-z/i)^{-1} = \exp(z/i + \frac{1}{2}(z/i)^2 + \frac{1}{3}(z/i)^3 + \dots)$. Les formules (4.12) à (4.14) sont démontrées dans [12] en utilisant des méthodes probabilistes et analytiques. Le schéma récursif (4.15) découle de (4.13) par une manipulation des exposants apparaissant dans l'expression de $\pi_{n,k,d} / \binom{n-1}{k}$ sous forme d'intégrale. ■

Il est intéressant de remarquer que (4.14) peut se réécrire sous la forme

$$\pi_{n,k,d} = \frac{1}{n!^d} \binom{n-1}{k} \sum_{i=0}^{n-1-k} (-1)^i \binom{n-1-k}{i} \frac{n!^d}{(k+i+1)^d} \quad (4.16)$$

et que cette dernière formule peut aussi se démontrer par le principe combinatoire d'inclusion-exclusion en procédant comme suit : Considérons une liste de d permutations aléatoires des entiers $1, 2, \dots, n$ (i.e. une T -structure aléatoire sur l'ensemble ordonné $\{1 < 2 < \dots < n\}$). Il s'agit de calculer la probabilité qu'il y ait exactement k entiers qui soient simultanément à gauche de la « racine » représentée par les « 1 » de la figure 1.

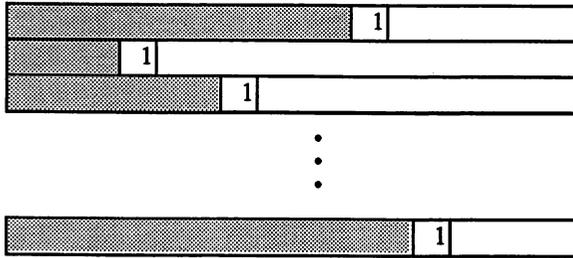


Figure 1

Il y a $\binom{n-1}{k}$ possibilités pour choisir ces k entiers parmi les nombres $2, 3, \dots, n$. Il y a $n!^d / (k+i+1)^d$ possibilités de d -tuplets de permutations tels que $k+i$ éléments fixés soient à gauche des « 1 » (et possiblement d'autres). La somme alternée dans (4.16) représente (part le principe d'inclusion-exclusion) le nombre de façons que ces k éléments soient $2, 3, \dots, k+1$.

Le schéma récursif (4.15) est particulièrement efficace pour dresser des tables des $\pi_{n,k,d}$ pour une dimension d fixée. C'est en utilisant ces probabilités que nous allons calculer l'espérance $\epsilon_{n,k,d}$ du nombre de noeuds ayant k enfants dans une arborescence hyperquaternaire aléatoire de n points en dimension d .

Pour les besoins de la présentation, introduisons quelques notations : posons, pour une fonction quelconque à δ variables $f(z_{00\dots 0}, z_{00\dots 1}, \dots, z_{11\dots 1})$ et pour un sous-ensemble T de l'ensemble des indices $\{ 00\dots 0, 00\dots 1, \dots, \epsilon_1 \epsilon_2 \dots \epsilon_d, \dots, 11\dots 1 \}$,

$$f[T] = f(\chi_T(00\dots 0), \chi_T(00\dots 1), \dots, \chi_T(11\dots 1))$$

où χ_T désigne la fonction caractéristique de l'ensemble T . Pour une variable t , nous écrirons $t^{<0>} = t$ et $t^{<1>} = 1-t$.

Théorème 5 Pour k et d fixés, les espérances $\epsilon_{n,k,d}$ satisfont le schéma de récurrence suivant:

$$\epsilon_{n,k,d} = p_{n,k,d} + 2^d \sum_{i=0}^{n-1} \pi_{n,i,d} \epsilon_{i,k,d} \tag{4.17}$$

où $p_{n,k,d}$ = probabilité que la racine d'une arborescence hyperquaternaire de n points ait k enfants ($0 \leq k \leq 2^d = \delta$)

$$= \sum_{v=0}^k (-1)^{k-v} \binom{\delta-v}{k-v} \sum_{|T|=v} f[T] \tag{4.18}$$

où

$$f(z_{00\dots 0}, \dots, z_{11\dots 1}) = \int_0^1 \dots \int_0^1 \left(\sum_{\epsilon: \{d\} \rightarrow 2} t_1^{<\epsilon_1>} \dots t_d^{<\epsilon_d>} z_{\epsilon_1 \dots \epsilon_d} \right)^{n-1} dt_1 \dots dt_d. \tag{4.19}$$

Preuve : La formule (4.17) est bien connue [7,9,12]. Il suffit de distinguer le cas où la racine a k enfants et d'utiliser le fait que les 2^d sous-arborescences de la racine sont indépendantes et de même nature. Le schéma récursif en découle immédiatement en faisant appel aux probabilités $\pi_{n,i,d}$. Le coefficient de $z_{00\dots 0}^{m_{00\dots 0}} z_{00\dots 1}^{m_{00\dots 1}} \dots z_{\varepsilon_1 \varepsilon_2 \dots \varepsilon_d}^{m_{\varepsilon_1 \varepsilon_2 \dots \varepsilon_d}} \dots$ dans f est égal à la probabilité qu'il y ait $m_{\varepsilon_1 \varepsilon_2 \dots \varepsilon_d}$ points dans l'hyperquadrant indexé par $\varepsilon_1 \varepsilon_2 \dots \varepsilon_d$ pour chaque $\varepsilon : [d] \rightarrow 2$. La probabilité cherchée est donc égale à la somme des coefficients des monômes dans f où exactement k variables, parmi les $z_{\varepsilon_1 \varepsilon_2 \dots \varepsilon_d}$, apparaissent. La formule finale découle du principe d'inclusion-exclusion. ■

La formule (4.17) s'exprime dans le langage des séries génératrices comme suit [7] : pour k, d fixés, posons

$$\varepsilon(z) = \sum_{n \geq 0} \varepsilon_{n,k,d} z^n, \quad p(z) = \sum_{n \geq 0} p_{n,k,d} z^n. \quad (4.20)$$

Alors (4.17) équivaut à l'équation intégrale

$$\varepsilon(z) = p(z) + 2^d J^{d-1} I \varepsilon(z) \quad (4.21)$$

$$\text{où} \quad I f(z) = \int_0^z f(t) \frac{dt}{1-t}, \quad J f(z) = \int_0^z f(t) \frac{dt}{t(1-t)}. \quad (4.22)$$

À son tour, cette équation intégrale se traduit par l'équation différentielle linéaire, non-homogène, d'ordre d

$$\{((1-z)D)(z(1-z)D)^{d-1} - 2^d\} w(z) = 2^d p(z), \quad w(z) = \varepsilon(z) - p(z). \quad (4.23)$$

Corollaire 1 En dimension $d = 2$, on a les formules explicites et asymptotiques suivantes, pour les $p_{n,k,2}$, $0 \leq k \leq 4$:

$$\text{Pour } n \geq 1, \quad p_{n,0,2} = \delta_1^n, \quad (\text{delta de Kronecker}), \quad (4.24)$$

$$\text{pour } n \geq 2, \quad p_{n,1,2} = \frac{4}{n^2}, \quad (4.25)$$

$$\text{pour } n \geq 3, \quad p_{n,2,2} = \frac{4}{n} - \frac{12}{n^2} + \frac{2}{n} \sum_{i=0}^{n-1} \frac{1}{2^i(n-i)} \sim \frac{4}{n} - \frac{8}{n^2}, \quad (4.26)$$

$$\text{pour } n \geq 4, \quad p_{n,3,2} = \frac{4}{n} H_n - \frac{8}{n} + \frac{12}{n^2} - \frac{4}{n} \sum_{i=0}^{n-1} \frac{1}{2^i(n-i)} \sim \frac{4}{n} H_n - \frac{8}{n} + \frac{4}{n^2}, \quad (4.27)$$

$$\text{pour } n \geq 5, \quad p_{n,4,2} = 1 - \frac{4}{n} H_n + \frac{4}{n} - \frac{4}{n^2} + \frac{2}{n} \sum_{i=0}^{n-1} \frac{1}{2^i(n-i)} \sim 1 - \frac{4}{n} H_n + \frac{4}{n}, \quad (4.28)$$

où $H_n = 1 + \frac{1}{2} + \dots + \frac{1}{n}$ désigne le $n^{\text{ème}}$ nombre harmonique.

Preuve : Il suffit d'évaluer les intégrales du théorème 5 et d'utiliser le fait que

$$\sum_{i=0}^{n-1} \frac{1}{2^i(n-i)} = \frac{1}{n} \sum_{i=0}^{n-1} \frac{1}{\binom{n-1}{i}} \sim \frac{2}{n} \quad \text{si } n \rightarrow \infty. \quad \blacksquare$$

Corollaire 2 En dimension $d = 2$, on a la récurrence suivante, pour les $\varepsilon_{n,k,2}$, $0 \leq k \leq 4$, $n \geq 2$:

$$\begin{aligned} \varepsilon_{n,k,2} = & p_{n,k,2} + \frac{4}{n^2} p_{n-1,k,2} + \left(2 - \frac{3}{n} + \frac{5}{n^2}\right) (\varepsilon_{n-1,k,2} - p_{n-1,k,2}) \\ & - \left(1 - \frac{3}{n} + \frac{2}{n^2}\right) (\varepsilon_{n-2,k,2} - p_{n-2,k,2}). \end{aligned} \quad (4.29)$$

En particulier, le nombre moyen de feuilles $\varepsilon_{n,0,2}$ et de noeuds ayant un enfant $\varepsilon_{n,1,2}$ sont donnés par les formules exactes et asymptotiques

$$\varepsilon_{n,0,2} = 8(3n+1)H_n^{(2)} - 39n + 11 - \frac{4}{n} \sim (4\pi^2 - 39) \left(n + \frac{1}{3}\right) - \frac{4}{3n^3} + \dots, \quad (4.30)$$

$$\begin{aligned} \varepsilon_{n,1,2} = & 8(3n+1) \left[\sum_{i=1}^n \frac{H_i}{i^2} - H_n^{(3)} - \frac{13}{2} H_n^{(2)} \right] + 228n + 24H_n - 56 - \frac{4}{n} H_n \\ & + \frac{16}{n} + \frac{4}{n^2} \sim 12(2\zeta(3) - 13\zeta(2) + 19) \left(n + \frac{1}{3}\right) - \frac{4}{3n^2} + \dots \end{aligned} \quad (4.31)$$

où $\zeta(s) = \sum_{k=1}^{\infty} \frac{1}{k^s}$ désigne la fonction zêta de Riemann.

Preuve : Il suffit, pour la récurrence (4.29), d'identifier les coefficients dans les deux membres de l'équation différentielle (4.23) avec $d = 2$. Les formules (4.30) et (4.31) découlent naturellement de la théorie générale des équations linéaires aux différences finies d'ordre 2 [17] et de la formule sommatoire d'Euler-Maclaurin [8]. \blacksquare

Il est à noter que nous avons obtenu, dans Laforest [12] et Labelle, Laforest [11] une forme moins explicite de la constante multiplicative de n dans la formule asymptotique (4.31), à savoir

$$12(2\alpha_1 + 19 - 2\zeta(3) - 13\zeta(2)) \quad \text{où} \quad \alpha_1 = \sum_{i=1}^{\infty} H_i / i^2. \quad (4.32)$$

Flajolet nous a fait remarquer [2,6,14] que la constante α_1 est en fait égale à $2\zeta(3)$ (identité remontant à Euler). Ceci explique les deux écritures différentes pour le même résultat asymptotique.

Le corollaire 2 peut être généralisé comme suit :

Corollaire 3 Soit P une propriété applicable aux noeuds d'une arborescence hyperquaternaire aléatoire. Supposons que la probabilité p_ν qu'un noeud quelconque possède la propriété P ne dépende que de la cardinalité ν de la sous-arborescence déterminée par le noeud, $p_\nu \rightarrow 0$ lorsque $\nu \rightarrow \infty$. En dimension $d = 2$, l'espérance mathématique $\varepsilon_n = \varepsilon_n(P)$ du nombre de noeuds d'une arborescence quaternaire aléatoire de n points possédant la propriété P est donnée par :

$$\varepsilon_n = p_n + 12 \left(n + \frac{1}{3} \right) \left(\sum_{\nu=0}^{n-1} \left(\frac{3\nu+1}{4} - \frac{\nu(3\nu+1)}{2} H_\nu^{(2)} - \frac{9\nu}{4} \right) p_\nu - \left(\frac{1}{4n} - \frac{1}{2} H_n^{(2)} - \frac{9}{4(3n+1)} \right) \sum_{\nu=1}^{n-1} \nu(3\nu+1) p_\nu \right) \quad (4.33)$$

où p_n est la probabilité que la racine d'une arborescence de n points ait la propriété P . On a aussi la formule asymptotique

$$\varepsilon_n(P) \sim \gamma(P) \left(n + \frac{1}{3} \right) \quad (4.34)$$

$$\text{où } \gamma(P) = \sum_{\nu=0}^{\infty} \left(3 - 18\nu + \nu(3\nu+1)(\pi^2 - 6H_\nu^{(2)}) \right) p_\nu. \quad (4.35)$$

Preuve : Pour (4.33), il suffit d'adapter les récurrences (4.17) et (4.29) au présent contexte et d'utiliser, encore une fois, les techniques du calcul des différences finies. La formule (4.34) et l'expression (4.35) pour la constante $\gamma(P)$ s'obtiennent de (4.33) en remarquant que

$$-\left(\frac{1}{4n} - \frac{1}{2} H_n^{(2)} - \frac{9}{4(3n+1)} \right) = \frac{\pi^2}{12} + O\left(\frac{1}{n^4} \right) \quad (4.36)$$

et que $\sum_{\nu=0}^{n-1} \nu(3\nu+1) p_\nu \leq \sum_{\nu=0}^{n-1} \nu(3\nu+1) = O(n^3)$. ■

En $d = 3$ dimensions, le schéma récursif pour les $\varepsilon_{n,k,d}$ prend la forme plus complexe

$$\begin{aligned} \varepsilon_{n,k,3} = & p_{n,k,3} + \frac{8}{n^3} p_{n-1,k,3} + \left(3 - \frac{6}{n} + \frac{4}{n^2} + \frac{7}{n^3} \right) (\varepsilon_{n-1,k,3} - p_{n-1,k,3}) \\ & - 3 \left(1 - \frac{4}{n} + \frac{5}{n^2} - \frac{2}{n^3} \right) (\varepsilon_{n-2,k,3} - p_{n-2,k,3}) \\ & + \left(1 - \frac{6}{n} + \frac{11}{n^2} - \frac{6}{n^3} \right) (\varepsilon_{n-3,k,3} - p_{n-3,k,3}). \end{aligned} \quad (4.37)$$

Le comportement asymptotique de $\epsilon_{n,k,3}$ est décrit $\gamma_{k,3} \cdot \left(n + \frac{1}{7}\right)$. Dans le cas général pour une dimension d quelconque, le comportement asymptotique de $\epsilon_{n,k,d}$ devrait être de la forme

$$\gamma_{k,d} \cdot \left(n + \frac{1}{2^d - 1}\right), \quad k < 2^d \quad (4.38)$$

pour certaines constantes $\gamma_{k,d}$ dont la nature précise est présentement à l'étude.

Références

- [1] F.Bergeron, P.Flajolet, B.Salvy, "Varieties of Increasing Trees," 1991, (à paraître)
- [2] B.C.Berndt, *Ramanujan's notebook, Part I*, Springer-Verlag, 1985.
- [3] B.W. Char, K.O.Geddes, G.H. Gonnet, M.B. Monagan, S.M. Watt, *MAPLE Reference Manual*, 5th ed., Watcom Publ. Ltd., Waterloo, Ont., Canada, 1988.
- [4] L.Devroye, L. Laforest, "An Analysis of Random d -dimensional quadrees," *SIAM J. Comput.*, 19 (5) 1990, pp. 821-832.
- [5] R.A. Finkel, J.L. Bentley, "Quad trees: a data structure for retrieval on composite keys," *Acta Informatica*, vol. 4, pp. 1-9, 1974.
- [6] P. Flajolet, communication par courrier-électronique, avril 1991.
- [7] P.Flajolet, G.Gonnet, C.Puech, J.M.Robson, "Analytic Variations on Quad-Trees," 1991 (à paraître)
- [8] R. L. Graham, D. E. Knuth, O. Patashnik, *Concrete Mathematics, A Foundation for Computer Science*, Addison-Wesley, mai 1989.
- [9] M.Hoshi, P. Flajolet, "Page Usage in Quadtree Indexes," 1991, (à paraître)
- [10] G. Labelle, "Une nouvelle démonstration des formules d'inversion de Lagrange," *Adv. in Math.*, 42(3), 1981, pp. 217-247.
- [11] G.Labelle, L.Laforest, "Étude asymptotique du nombre moyen de noeuds à un enfant dans un arbre quaternaire," Rapport de recherche no. 134, octobre 1990, LACIM/UQAM, Montréal (Québec).
- [12] L. Laforest, *Étude des arbres hyperquaternaires*, Thèse de doctorat, Université McGill, Juin 1990, Publications du LACIM/UQAM, Montréal, 1990.
- [13] P. Leroux, X.G.Viennot, Combinatorial Resolution of Systems of Differential Equations I : Ordinary Differential Equations, dans *combinatoire énumérative*, Montréal (Québec), 1985, Proceedings, G.Labelle et P.Leroux éditeurs, Lectures Notes in Math., no. 1234, Springer-Verlag, Heidelberg (1986), pp 210-245.
- [14] L.Lewin, *Polylogarithms and Associated Functions*, North-Holland, New York, 1981.
- [15] H.M. Mahmoud, "The expected distribution of degrees in random binary trees," *The Computer Journal*, vol. 29, pp. 36-37, 1986.
- [16] H. Samet, "The quadtree and related hierarchical data structures," *Computing Surveys*, vol. 16, pp. 187-260, 1984.
- [17] J. Wimp, *Computation with recurrence relations*, Pitman Advanced Publishing Program, Pitman Publ. Ltd., 1984.

Quelques problèmes concernant les chemins de Dyck généralisés

par Jacques LABELLE*

LACIM / Math.-Info., Université du Québec à Montréal,

C.P. 8888, Succ.A (Québec), Canada H3C 3P8.

Résumé La notion de chemin de Dyck généralisé [13] s'obtient lorsqu'on remplace dans les chemins de Dyck classiques l'unique pas $(1, 1)$ et son symétrique $(1, -1)$ par des pas (et leurs symétriques) provenant d'un sous ensemble fini \mathbf{S} avec répétitions de $\mathbb{N}^+ \times \mathbb{N}$. Ces chemins ainsi que d'autres (\mathbf{S} -grand Dyck, facteurs gauches de \mathbf{S} -Dyck, etc.) correspondent à des langages algébriques intéressants. La méthodologie, dite de Schützenberger, est employée pour trouver des systèmes d'équations et des "formes closes" pour les séries génératrices à $|\mathbf{S}|$ variables non-commutatives (resp. commutatives) de ces langages (resp. ces familles de chemins). Les cas $\mathbf{S} = \{\mathbf{h}, \mathbf{u}\}$ et $\mathbf{S} = \{\mathbf{h}, \mathbf{u}, \mathbf{v}\}$, où $\mathbf{h} = (r_0, 0)$, $\mathbf{u} = (r_1, 1)$ et $\mathbf{v} = (r_2, 2)$, sont entièrement résolus; dans le cas général les équations deviennent vite de degré très grand. L'étude des paires de \mathbf{S} -chemins sans croisement (resp. sans intersection) se ramène à l'étude des $\mathbf{T}(\mathbf{S})$ -chemins de Dyck (resp. $\mathbf{T}(\mathbf{S})$ -chemins de Dyck connexes) où $\mathbf{T}(\mathbf{S})$ est décrit [11] à partir de \mathbf{S} . De plus le lien avec les fractions continuées ainsi que la relation entre le cas "avec paliers" et "sans palier" sont étudiés.

Abstract Generalized Dyck paths [13] are obtained from classical Dyck paths when the unique step $(1, 1)$ is replaced by steps from a finite subset \mathbf{S} (with repetitions) of $\mathbb{N}^+ \times \mathbb{N}$. These families of paths and others (" \mathbf{S} -grand Dyck", left factors of \mathbf{S} -Dyck paths, etc.) correspond to interesting algebraic languages. The so-called "Schützenberger methodology" is used to find systems of algebraic equations and closed forms for the $|\mathbf{S}|$ non-commutative (resp. commutative) variables generating functions of these languages (resp. families of paths). The cases $\mathbf{S} = \{\mathbf{h}, \mathbf{u}\}$ and $\mathbf{S} = \{\mathbf{h}, \mathbf{u}, \mathbf{v}\}$, where $\mathbf{h} = (r_0, 0)$, $\mathbf{u} = (r_1, 1)$ and $\mathbf{v} = (r_2, 2)$, are completely solved; in the general case the degree of the equations becomes rapidly very high. Pairs of non-crossing (resp. non-intersecting) \mathbf{S} -paths lead naturally to $\mathbf{T}(\mathbf{S})$ -Dyck paths (resp. connected $\mathbf{T}(\mathbf{S})$ -Dyck paths) where $\mathbf{T}(\mathbf{S})$ is easily described [11] from \mathbf{S} . The relation with continued fractions and a formula relating the presence or not of horizontal steps are considered.

* Avec le support financier du FCAR (Québec) et du CRSNG (Canada).

1. Chemins de Dyck généralisés (voir [13]).

Soit \mathbf{S} un sous-ensemble fini (avec répétitions possibles) de $\mathbb{N}^+ \times \mathbb{N}$ où $\mathbb{N}^+ = \{1,2,\dots\}$ et $\mathbb{N} = \{0,1,2,\dots\}$. Pour $\mathbf{u} = (r,s) \in \mathbf{S}$, posons $\bar{\mathbf{u}} = (r,-s)$, $\pi_1(\mathbf{u}) = r$ et $\pi_2(\mathbf{u}) = s$.
 Soit $\bar{\mathbf{S}} = \{ \bar{\mathbf{u}} \mid \mathbf{u} \in \mathbf{S} \}$.

Définition 1:

- Un \mathbf{S} -chemin est un chemin dans $\mathbb{Z} \times \mathbb{Z}$ dont tous les pas sont pris dans $\mathbf{S} \cup \bar{\mathbf{S}}$.
- Un \mathbf{S} -chemin de Dyck (nous dirons aussi un D-chemin) est un \mathbf{S} -chemin qui:
 - i) part de (0,0) et finit sur l'axe des x;
 - ii) reste dans $\mathbb{N} \times \mathbb{N}$, i.e. au-dessus de l'axe des x.
- Un \mathbf{S} -chemin de Dyck connexe (nous dirons aussi un D_c -chemin) est un \mathbf{S} -chemin de Dyck qui ne visite l'axe des x qu'à ses extrémités.
- Pour $i, j \geq 0$, un $D_{i,j}$ -chemin est un \mathbf{S} -chemin dans $\mathbb{N} \times \mathbb{N}$ qui va de (0,i) à (n, j).
- Pour $i > 0$, un D_i -chemin est un $D_{0,i}$ -chemin. On pose $D_0 = D$.
- Pour $i > 0$, un $(D_i)_c$ -chemin est un D_i -chemin ne visitant l'axe des x qu'à l'origine. On pose $(D_0)_c = 1$.

En considérant $\mathbf{S} \cup \bar{\mathbf{S}}$ comme un alphabet où chaque lettre a une hauteur définie par $h(\mathbf{u}) = \pi_2(\mathbf{u})$, $h(\bar{\mathbf{u}}) = -h(\mathbf{u})$, et en remplaçant les chemins par des mots, on obtient les langages D , D_c , D_i , $(D_i)_c$ et $D_{i,j}$. Suivant la méthodologie de monsieur Schützenberger (voir [2, 3, 7, 16, 18, 19]), ces langages correspondent à des séries en les variables non-commutatives $\{u, \bar{u} \mid \mathbf{u} \in \mathbf{S}\}$. En rendant ces variables commutatives et en posant $u = \bar{u}$, on obtient la série à $|\mathbf{S}|$ variables de chacun de ces langages. En posant ensuite $u = t$, pour tout $\mathbf{u} \in \mathbf{S}$, on obtient les séries génératrices de ces familles de chemins suivant le nombre de pas.

En particulier, pour $\mathbf{S} = \{\mathbf{u}_0, \mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_m\}$ où $\mathbf{u}_j = (r_j, j)$, $r_j \in \mathbb{N}^+$, on considère donc la série génératrice à $m+1$ variables (commutatives) $u_0, u_1, u_2, \dots, u_m$ une pour chaque vecteur de \mathbf{S} :

$$D = D(u_0, u_1, u_2, \dots, u_m) = \sum_{i_j \geq 0} d_{i_0 i_1 i_2 \dots i_m} u_0^{i_0} u_1^{i_1} u_2^{i_2} \dots u_m^{i_m}$$

où $d_{i_0 i_1 i_2 \dots i_m}$ est le nombre de D-chemins tels que pour $0 \leq j \leq m$, \mathbf{u}_j et $\bar{\mathbf{u}}_j$ apparaissent un total de i_j fois. De même on a les séries génératrices à $m+1$ variables (commutatives) $D_c, D_i, i > 0, (D_i)_c, i > 0$, et $D_{i,j}, i, j \geq 0$.

Le calcul des séries dans ce cas est suffisamment général car, à partir de $D(u_0, \dots, u_m)$, on retrouve le cas où \mathbf{S} est arbitraire en posant $u_i = \sum u$, ($u \in \mathbf{S}$, $\pi_2(u) = i$).

Proposition 1:

Les séries D , D_c , D_i , $(D_i)_c$, $1 \leq i$, et, $D_{i,j}$, $1 \leq i, j$ satisfont aux équations suivantes:

$$D = 1 + D_c \cdot D \quad , \quad D_c = u_0 + \sum_{1 \leq i, j \leq m} u_i u_j D_{i-1, j-1} \quad , \quad D_i = D \cdot (D_i)_c$$

$$(D_i)_c = \sum_{1 \leq j \leq m} u_j D_{j-1, i-1} \quad , \quad D_{i,j} = \sum_{k=0}^{\min(i,j)} D_{i-k} \cdot (D_{j-k})_c$$

Preuve: Au niveau des langages ou chemins, on a: $D = 1 + D_c \cdot D$. De même les égalités:

$$D_c = u_0 + \sum_{1 \leq i, j \leq m} u_i D_{i-1, j-1} \bar{u}_j \quad , \quad D_i = D \cdot (D_i)_c \quad \text{et} \quad (D_i)_c = \sum_{1 \leq j \leq m} u_j D_{j-1, i-1}$$

Pour obtenir la dernière équation, il suffit de considérer la droite horizontale la plus basse, $y = k$, visitée par le $D_{i,j}$ -chemin, et la dernière fois où elle est visitée.

Corollaire: (voir [13])

Les m séries D et D_i , $1 \leq i < m$, satisfont aux m équations quadratiques suivantes:

(où on pose $D_0 = D$ et $D_{-n} = 0$ pour $n > 0$)

$$(*) \quad D = 1 + u_0 D + \sum_{1 \leq i, j \leq m} u_i u_j \left(\sum_{k \geq 1} D_{i-k} D_{j-k} \right)$$

$$(**) \quad D_i = \sum_{j=1}^m u_j \left(\sum_{k \geq 1} D_{i-k} D_{j-k} \right) \quad , \quad 1 \leq i < m$$

Preuve: Pour obtenir (*), il suffit de remplacer D_c et $D_{i,j}$ par leurs expressions et d'utiliser $D_i = D \cdot (D_i)_c$. Pour obtenir (**), on part de $D_i = D \cdot (D_i)_c$ puis on remplace $(D_i)_c$ et $D_{j,i}$ par leurs expressions.

Corollaire:

La série $D(u_0, u_1, u_2, \dots, u_m)$ satisfait à une équation polynomiale de degré 2^m à coefficients dans $\mathbb{Z}[u_0, u_1, \dots, u_m]$ de coefficient dominant u_m à la puissance 2^m .

Exemple 1: Chemins de Dyck et de Motzkin classiques, (voir [8, 15, 17, 18]).

Si $S = \{u_0, u_1\}$, $u_i = (r_i, i)$, $0 \leq i \leq 1$, on a (en posant $u_0 = h$ et $u_1 = u$) l'équation quadratique suivante satisfaite par $D = D(h, u)$: $u^2 D^2 - (1-h)D + 1 = 0$.

D'où l'on tire: $D(h, u) = \frac{1}{2u^2} \left\{ 1-h - \sqrt{(1-h)^2 - 4u^2} \right\}$; $D_c = 1-D^{-1}$; $G = (1-G_c)^{-1}$; $G_c = 2D_c$.
Puisque $D_i = DuD_{i-1}$, on trouve également:

$$D_i = u^i D^{i+1}, \quad D_{i,j} = \sum_{k=0}^{\min(i,j)} D_{i-k} (D_{j-k})_c \quad \text{et} \quad (D_i)_c = u^i D^i.$$

Exemple 2: (voir [11,12,13])

Si $S = \{u_0, u_1, u_2\}$, $u_i = (r_i, i)$, $0 \leq i \leq 2$, on a (en posant $u_0 = h$, $u_1 = u$, $u_2 = v$) l'équation polynomiale suivante de degré quatre satisfaite par $D = D(h, u, v)$:

$$v^4 D^4 - v^2(2v+1-h)D^3 + (u^2+2v^2+2v-2hv)D^2 - (2v+1-h)D + 1 = 0.$$

En effet pour $m = 2$, les équations (*) et (**) deviennent:

$$D = 1 + hD + u^2 D^2 + 2uvD \cdot D_1 + v^2((D_1)^2 + D^2) \quad \text{et}$$

$$D_1 = uD^2 + vD \cdot D_1 \quad \text{d'où} \quad D_1 = uD^2(1-vD)^{-1}.$$

On en déduit l'équation polynomiale après calcul.

En utilisant un logiciel de calcul formel tel MAPLE [1], MACSYMA [14] ou MATHEMATICA [21], on tire:

$$D(h, u, v) = \frac{1}{4u^2} \left\{ 1-h+2u+\Delta - \sqrt{2((1-h)^2 - 4u^2 - 2v^2 + (1-h+2u)\Delta)} \right\}$$

$$\text{où} \quad \Delta = \sqrt{(1-h)^2 - 4u^2 - 4v + 4hv + 4v^2}$$

$$\text{et} \quad D(t) = \frac{1}{4t^2} \left\{ 1+t+\Delta - \sqrt{2(1-2t-5t^2 + (1+t)\Delta)} \right\} \quad \text{où} \quad \Delta = \sqrt{(1-t)(1-5t)}.$$

On a également la récurrence suivante:

Proposition 2:

On a: $D_0 = D$, $D_1 = uD^2(1-vD)^{-1}$, $D_{k+2} = uD(1-vD)^{-1} \cdot D_{k+1} + vD \cdot D_k$

Preuve: Ici (**) s'écrit $D_1 = uD^2 + vD \cdot D_1$; d'où l'on tire $D_1 = uD^2(1-vD)^{-1} = D \cdot (D_1)_c$

Au niveau des chemins on a: $D_{k+2} = DuD_{k+1} + DvD_{1,k+1}$, de plus:

$D_{1,k+1} = D_k$ (si le chemin ne visite pas l'axe des x) + $(D_1)_c \cdot D_{k+1}$ (si le chemin visite l'axe des x).

On en tire $D_{k+2} = D \cdot (u + v(D_1)_c) \cdot D_{k+1} + vD \cdot D_k$, d'où l'expression pour D_{k+2} .

Corollaire:

On a:
$$\mathbf{D}_n = \sum_{k=0}^{\lfloor n/2 \rfloor} \binom{n-k}{k} v^k \mathbf{D}^{k+1} (\mathbf{D}_1)_c^{2n-2k} \quad \text{où} \quad (\mathbf{D}_1)_c = u\mathbf{D}(1-v\mathbf{D})^{-1} .$$

Pour terminer cette section, notons que la *symétrie totale* de l'ensemble des pas n'est pas indispensable à l'étude des chemins de Dyck généralisés.

Exemple 3:

Soit $\{u, v, \bar{u}\}$ l'ensemble des pas, où $u = (r_1, 1)$ et $v = (r_2, 2)$.

Au niveau des langages on a: $\mathbf{D} = 1 + (\mathbf{D}_c \cdot \mathbf{D})$; $\mathbf{D}_c = u \mathbf{D} \bar{u} + v \mathbf{D} \bar{u} \mathbf{D} \bar{u}$.

D'où: $\mathbf{D}_c = u^2 \mathbf{D}(1 + v\mathbf{D})$; $\mathbf{D} = (1 - \mathbf{D}_c)^{-1} = (1 - u^2 \mathbf{D} - u^2 v \mathbf{D}^2)^{-1}$.

Ce qui conduit à l'équation de degré trois: $u^2 v \mathbf{D}^3 + u^2 \mathbf{D}^2 - \mathbf{D} + 1 = 0$.

Plus généralement, si l'ensemble des pas est $\{u_1, u_2, \dots, u_n, \bar{u}_1\}$.

Au niveau des langages on trouve:

$$\mathbf{D} = 1 + \mathbf{D}u_1\mathbf{D}\bar{u}_1 + \mathbf{D}u_2(\mathbf{D}\bar{u}_1)^2 + \dots + \mathbf{D}u_n(\mathbf{D}\bar{u}_1)^n$$

Ceci conduit en variables commutatives à l'équation polynomiale suivante de degré $n + 1$ satisfaite par $\mathbf{D} = \mathbf{D}(u_1, u_2, \dots, u_n)$:

$$u_n u_1^n \mathbf{D}^{n+1} + u_{n-1} u_1^{n-1} \mathbf{D}^n + \dots + u_2 u_1^2 \mathbf{D}^3 + u_1^2 \mathbf{D}^2 - \mathbf{D} + 1 = 0$$

et a

$$t^{n+1} \mathbf{D}^{n+1} + t^n \mathbf{D}^n + \dots + t^2 \mathbf{D}^2 - \mathbf{D} + 1 = 0$$

pour la série génératrice $\mathbf{D} = \mathbf{D}(t)$ selon le nombre de pas.

2. Chemins de grand Dyck.

Définition 2:

- Un **S**-grand Dyck (nous dirons aussi un **G**-chemin) est un **S**-chemin qui commence en (0,0) et finit sur l'axe des x.
- Un **S**-grand Dyck connexe (nous dirons aussi un **G_c**-chemin) est un **S**-grand Dyck qui ne visite l'axe des x qu'à ses extrémités. Notez qu'un tel chemin peut cependant traverser l'axe des x avec des pas **u** et \bar{u} tels que $\pi_2(\mathbf{u}) > 1$.

Considérons les séries génératrices à m+1 variables (commutatives) **G** et **G_c**.

Proposition 3:

Si $\mathbf{S} = \{\mathbf{u}_0, \mathbf{u}_1, \mathbf{u}_2\}$, $\mathbf{u}_i = (r_i, i)$, $0 \leq i \leq 2$, on a (en posant $\mathbf{u}_0 = \mathbf{h}$, $\mathbf{u}_1 = \mathbf{u}$, $\mathbf{u}_2 = \mathbf{v}$) les équations suivantes: $\mathbf{G} = 1 + \mathbf{G}_c \cdot \mathbf{G}$; $\mathbf{G}_c = 2(1 - \mathbf{D}^{-1} + \mathbf{u}^2 \mathbf{v} \mathbf{D}^2 (1 - \mathbf{v} \mathbf{D})^{-3})$.

Preuve: La première équation est évidente. Pour la seconde, considérons un **G_c**-chemin: S'il ne traverse jamais l'axe des x avec un pas **v** ou \bar{v} , c'est un **D_c**-chemin ou bien au-dessus ou bien au-dessous de l'axe des x; d'où le terme $2\mathbf{D}_c = 2(1 - \mathbf{D}^{-1})$. S'il traverse l'axe des x k+1 fois et commence positivement, c'est un chemin dans $(\mathbf{D}_1)_c \bar{\mathbf{v}} \mathbf{D} \mathbf{v} \mathbf{D} \bar{\mathbf{v}} \dots \mathbf{D} \bar{\mathbf{v}} (\mathbf{D}_1)_c$, si k est pair, ou dans $(\mathbf{D}_1)_c \bar{\mathbf{v}} \mathbf{D} \mathbf{v} \dots \mathbf{D} \mathbf{v} \mathbf{D} \bar{\mathbf{v}} (\mathbf{D}_1)_c$, si k est impair; d'où le terme $2 \sum_{k=0} (\mathbf{D}_1)_c \mathbf{v} (\mathbf{v} \mathbf{D})^k (\mathbf{D}_1)_c$ qui, en utilisant $(\mathbf{D}_1)_c = \mathbf{u} \mathbf{D} (1 - \mathbf{v} \mathbf{D})^{-1}$, devient $2\mathbf{u}^2 \mathbf{v} \mathbf{D}^2 (1 - \mathbf{v} \mathbf{D})^{-3}$.

Corollaire:

On a: $\mathbf{G} = \mathbf{D} (1 - \mathbf{v} \mathbf{D})^3 / (2 - (1 + 6\mathbf{v}) \mathbf{D} + 3\mathbf{v} (1 + 2\mathbf{v}) \mathbf{D}^2 - \mathbf{v} (2\mathbf{u}^2 + 3\mathbf{v} + 2\mathbf{v}^2) \mathbf{D}^3 + \mathbf{v}^3 \mathbf{D}^4)$
 où $\mathbf{D} = \mathbf{D}(\mathbf{h}, \mathbf{u}, \mathbf{v})$ est donné à l'exemple 2.

Preuve: Découle des deux équations de la proposition précédente.

3. Paires de chemins sans intersection ou sans croisement (voir [6, 9, 11]).

Soit \mathbf{S} un multi-ensemble fini de vecteurs dans $\{1\} \times \mathbb{N}$.

Définition 3:

Soient c et c' deux \mathbf{S} -chemins partant respectivement de $(0,0)$ et $(0,k)$, $0 \leq k$, et se terminant respectivement en (n,j) et $(n,j+\ell)$. On dit qu'ils sont sans croisement (resp. sans intersection) si c est au-dessous (resp. strictement au-dessous, sauf aux extrémités si $k = 0$ ou $\ell = 0$) de c' .

Proposition 4: (voir [11])

Il y a une bijection entre l'ensemble des paires de \mathbf{S} -chemins partant respectivement de $(0,0)$ et $(0,k)$ et se terminant respectivement en (n,j) et $(n,j+\ell)$, et l'ensemble des $\mathbf{T}(\mathbf{S})$ -chemins de $(0,k)$ à (n,ℓ) , où pour tout i , $0 \leq i$, le nombre de copies de $\mathbf{u}_i = (1, i)$ dans $\mathbf{T} = \mathbf{T}(\mathbf{S})$ est :

$$|\{(v,w) \mid v, w \in \mathbf{S} \cup \overline{\mathbf{S}}, \pi_2(v) - \pi_2(w) = i\}|.$$

Preuve: Il suffit de considérer l'écart entre les deux chemins; on obtient ainsi un $\mathbf{T}(\mathbf{S})$ -chemin. Par cette bijection, une paire sans croisement correspond à un $\mathbf{T}(\mathbf{S})$ - $D_{k,\ell}$ -chemin.

Corollaire:

Il y a bijection entre l'ensemble des paires de \mathbf{S} -chemins de $(0,0)$ jusqu'à un même point et l'ensemble des $\mathbf{T}(\mathbf{S})$ -grand Dyck; si les deux chemins de la paire sont sans croisement (resp. sans intersection) alors le chemin correspondant est un $\mathbf{T}(\mathbf{S})$ -chemin de Dyck (resp. $\mathbf{T}(\mathbf{S})$ -chemin de Dyck connexe).

Exemple 4:

Pour $\mathbf{S} = \{\mathbf{u}_1\}$, chemins de Dyck classiques, on a $\mathbf{T}(\mathbf{S}) = \{2\mathbf{u}_0, \mathbf{u}_2\}$, chemins de Motzkin à paliers bicolores. De plus on a $\mathbf{T}(\mathbf{T}(\mathbf{S})) = \{6\mathbf{u}_0, 4\mathbf{u}_2, \mathbf{u}_4\}$ avec comme série génératrice des $\mathbf{T}(\mathbf{T}(\mathbf{S}))$ -chemins de Dyck:

$$\frac{1}{4t^2} \{1 - 4t + \Delta - \sqrt{2(1 - 12t + \Delta)}\} \quad \text{où } \Delta = \sqrt{1 - 16t} \quad .$$

4. Lien avec les fractions continuées.

Rappelons que dans le cas $m = 1$, $\mathbf{S} = \{\mathbf{h}, \mathbf{u}\}$, $\mathbf{h} = (r_0, 0)$, $\mathbf{u} = (r_1, 1)$, on a:

$$\mathbf{D} = (1 - \mathbf{D}_c)^{-1} = (1 - h - u^2 \mathbf{D})^{-1} .$$

D'où le développement en fractions continuées:

$$\mathbf{D}(h, u) = \frac{1}{1-h} - \frac{u^2}{1-h} - \frac{u^2}{1-h} - \dots$$

Ce qui permet de *plafonner* les \mathbf{S} -chemins de Dyck par la droite $y = k$ en *tronquant* la fraction continuée au k^e niveau (i.e. en prenant son k^e convergent).

Pour $m \geq 2$, l'approche fractions continuées ne s'applique pas (du moins pas trivialement).

Par exemple, dans le cas $m = 2$, $\mathbf{S} = \{\mathbf{h}, \mathbf{u}, \mathbf{v}\}$, $\mathbf{h} = (r_0, 0)$, $\mathbf{u} = (r_1, 1)$, $\mathbf{v} = (r_2, 2)$, on a:

$$\mathbf{D} = (1 - \mathbf{D}_c)^{-1} \quad \text{où} \quad \mathbf{D}_c = h + (u^2 + v^2)\mathbf{D} + \frac{2u^2 v \mathbf{D}^2}{1 - v\mathbf{D}} + \frac{u^2 v^2 \mathbf{D}^3}{(1 - v\mathbf{D})^2} = \frac{P(\mathbf{D})}{Q(\mathbf{D})} = R(\mathbf{D}).$$

$$\text{On en tire: } \mathbf{D} = \frac{1}{1 - R(\mathbf{D})} = \frac{1}{1 - R\left(\frac{1}{1 - R(\mathbf{D})}\right)} = \dots$$

$$\text{où} \quad R(x) = \frac{(1 - vx)^2}{h + (u^2 + v^2 - 2hv)x + v^2(h - 2v)x^2 + v^4 x^3} .$$

Conclusion.

Dans l'étude des chemins de Dyck généralisés, et autres familles de chemins semblables, les difficultés proviennent de l'altitude, $\pi_2(\mathbf{u})$, des pas \mathbf{u} autorisés. S'il s'agit de \mathbf{S} -chemins où pour tout $\mathbf{u} \in \mathbf{S}$, $\pi_2(\mathbf{u}) \leq 1$, les choses sont relativement simples; cependant l'ajout d'autres contraintes comme une altitude maximum pour les chemins peuvent compliquer les choses (voir [3, 10]). Dans ce cas (i.e. pour tout $\mathbf{u} \in \mathbf{S}$, $\pi_2(\mathbf{u}) \leq 1$) l'approche fractions continuées permet également de pondérer les pas suivant l'altitude où ils ont lieu (voir [4, 8, 20]). Comme on l'a vu jusqu'ici la présence de pas d'altitude deux ou plus complique rapidement la situation.

Dans le même ordre d'idée notons que la présence ou non de paliers ne présente pas de problèmes. On a la proposition suivante:

Proposition 5:

Soit $\mathbf{S} = \{h, u_1, u_2, \dots, u_m\}$ où $\pi_2(h) = 0$ et $\pi_2(u_i) = i$, $1 \leq i \leq m$, et $\mathbf{S}' = \mathbf{S} - \{h\}$. Les séries génératrices $D_{\mathbf{S}}(h, u_1, u_2, \dots, u_m)$ et $D_{\mathbf{S}'}(u_1, u_2, \dots, u_m)$ sont liées par les identités suivantes:

$$D_{\mathbf{S}'}(u_1, u_2, \dots, u_m) = D_{\mathbf{S}}(0, u_1, u_2, \dots, u_m),$$

$$D_{\mathbf{S}}(h, u_1, u_2, \dots, u_m) = \frac{1}{1-h} D_{\mathbf{S}'}\left(\frac{u_1}{1-h}, \frac{u_2}{1-h}, \dots, \frac{u_m}{1-h}\right).$$

Preuve: Bien sûr en posant $h = 0$ on interdit la présence de paliers. Par ailleurs un $D_{\mathbf{S}}$ -chemin est une suite de paliers (de série $1/(1-h)$) puis une suite de "vrais pas" chacun suivi d'une suite de paliers.

Références.

- [1] Char B.W., K.O. Geddes, G.H. Gonnet, M.B. Monagan et S.M. Watt, (1988). *MAPLE Reference Manuel*, version 5, Watcom Publ. Ltd., Waterloo, Ont., Canada.
- [2] Delest M.P. et G.X. Viennot (1984). Algebraic languages and polyominoes enumeration, *Theor. Computer Science*, 34, 169-206.
- [3] Fedou J.M. (1989). Grammaires et q-énumérations de polyominos, thèse de Ph.D., Université de Bordeaux I.
- [4] Flajolet P. (1980). Combinatorial aspects of continued fractions, *Discrete Math.* 32, 125-161.
- [5] Gessel I. (1980). A noncommutative generalization and q-analog of the Lagrange inversion formula, *Trans. Amer. Math. Soc.* 257, 455-482.
- [6] Gessel I. and G.X. Viennot (1985). Binomial determinants, paths, and hook length formulae, *Advances in Math.* 58, 300-321.
- [7] Goldman J. (1978). Formal languages and enumeration, *J. of Comb. Th.; A*, 24, 318-338.
- [8] Goulden I.P. et D.M. Jackson (1983). *Combinatorial Enumeration*, John Wiley, New York.
- [9] Gouyou-Beauchamps D. (1986). Chemins sous-diagonaux et tableaux de Young. Dans: G. Labelle and P. Leroux, Eds., *Combinatoire Énumérative, Montréal, Québec 1985, Proceedings*, Lecture Notes in Mathematics, no. 1234, Springer-Verlag, 112-125.

- [10] Kreweras G. (1970). Sur les éventails de segments, *Cahiers du Bur. Univ. de Rech. Opér.*, 15, 1-41.
- [11] Labelle J. (1991). On pairs of nonintersecting generalized Dyck paths, *J. of Statistical Planning and Inference Theory*, (à paraître).
- [12] Labelle J. and Y.-N. Yeh. (1989), Dyck paths of knight moves, *Discrete Applied Mathematics*, 24, 213-221.
- [13] Labelle J. and Y.-N. Yeh. (1990), Generalized Dyck Paths, *Discrete Math.*, 82, 1-6.
- [14] Mathlab Group. (1983). *MACSYMA Reference Manuel*, version 10, (2 vols.), Lab. for Comp. Sc., MIT, Cambridge, Mass., USA.
- [15] Penaud J.-G. (1990). Une nouvelle bijection pour les animaux dirigés, Dans: R. König, éd., *Séminaire Lotharingien de combinatoire, 22^e session, Hesselberg, Gerolfinden, Actes du séminaire*, Publications Univ. Strasbourg., 93-130.
- [16] Schützenberger, M.P. (1962). Certain elementary families of automata, *Proc. Symp. on Mathematical Theory of Automata*, Polytechnic Institute of Brooklyn, pp 139-153.
- [17] Stanley R. (1986). *Enumerative Combinatorics, Vol I*, Wadsworth & Brooks/Cole, Advanced Books & Software, Monterey, California.
- [18] Viennot G.X. (1981-82). Théorie combinatoire des nombres d'Euler et de Genocchi, *Séminaire de théorie des nombres*, Publications Univ. Bordeaux I.
- [19] Viennot G.X. (1985). Enumerative combinatorics and algebraic languages, Dans: L. Budach éd. *Proceedings FCT'85*, Lecture Notes in Computer Science no199, Springer-Verlag, New York/Berlin, 450-464.
- [20] Viennot G.X. (1983). Théorie combinatoire des polynômes orthogonaux, Notes de séminaires de l'Université du Québec à Montréal.
- [21] Wolfram S. (1988). *MATHEMATICA, A System for Doing Math. by Computer*, Addison-Wesley, New York.
-

ALGORITHMES D'OBTENTION D'ORDRES A DISTANCE MINIMUM D'UN ENSEMBLE DE GRAPHS DE SURCLASSEMENT.

O. LAVIALLE¹ - Cl. VIDAL²

1 : ENITA de BORDEAUX - LaBRI Unité de Recherche CNRS N°1306

2 : Ministère de l'Agriculture et de la Forêt

RESUME

Nous présentons un nouvel algorithme permettant d'obtenir un ordre total sur plusieurs objets évalués par plusieurs juges.

INTRODUCTION

Soient n objets (ou actions) évalués par plusieurs juges selon plusieurs critères. Nous présentons ici plusieurs méthodes permettant de classer ces objets en construisant un ordre minimisant une fonction objectif donnée.

Dans un premier temps, l'utilisation d'une méthode de surclassement due à Roy [9] nous permet de construire une relation entre chaque couple d'objets, ramenant ainsi la structure des données à une structure proche de celle de comparaisons par paires.

Pour chacun des juges, on construit à l'aide de cette relation un graphe orienté appelé graphe de surclassement dont les sommets sont les objets à comparer. De la même façon, un ordre peut être représenté par un graphe transitif complet. En considérant l'ensemble des ordres possibles sur les objets, nous chercherons celui ou ceux qui sont les plus proches (en un sens à définir) de l'ensemble des graphes de surclassement.

Pour évaluer la proximité d'un ordre avec l'ensemble des jugements, nous définirons une fonction objectif cherchant à minimiser un critère donné. Nous présentons ici la minimisation selon quatre critères différents. Chacun d'eux conduit à un problème d'optimisation différent.

[Programme de recherche financé par le ministère de l'Agriculture et de la Forêt].

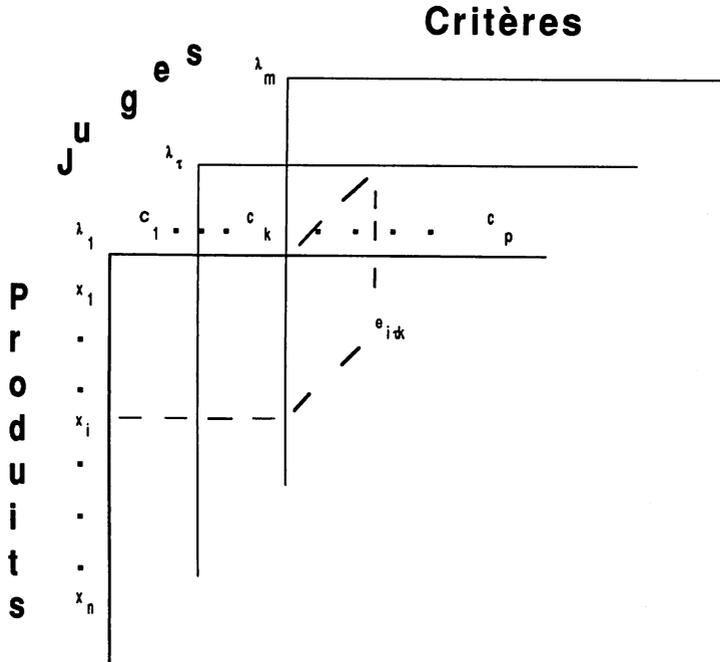
1. CONSTRUCTION D'UNE RELATION BINAIRE

1-1. Méthode de surclassement.

Soit $X = \{ x_1, x_2, \dots, x_n \}$ un ensemble de n objets évalués par m juges $\lambda_1, \lambda_2, \dots, \lambda_m$ selon p critères c_1, c_2, \dots, c_p . Chacun de ces critères est affecté d'une pondération, fonction de l'importance que le décideur lui accorde, soit respectivement $\beta_1, \beta_2, \dots, \beta_p$ (avec $\sum_{i=1}^p \beta_i = 1$).

Pour un objet x_i , chaque juge λ_τ effectue une évaluation selon chacun des critères c_k ($1 \leq k \leq p$). Cette évaluation se traduit par une note $e_{i\tau k}$.

On obtient ainsi m tableaux de données (un par juge) :



A l'aide de chacun des m tableaux, on construit un graphe de surclassement au sens de Roy [8] en reliant les objets par une relation prenant en compte l'ensemble des critères. Il existe plusieurs relations de ce type [10] que l'on peut présenter de manière intuitive par : *un objet x_i surclasse un objet x_j , si x_i est au moins aussi bon que x_j pour une majorité de critères sans être nettement plus mauvais sur les critères restants.*

Nous présentons ici une de ces relations connue sous le nom d'**ELECTRE I** :

Soit λ_τ un juge. Pour chaque paire (x_i, x_j) d'objets, partitionnons, pour chaque juge, l'ensemble des critères en trois sous-ensembles :

- $K_\tau^+(x_i, x_j) = \{ k/e_{i\tau k} > e_{j\tau k} \}$, ensemble des critères favorables à x_i ,
- $K_\tau^-(x_i, x_j) = \{ k/e_{i\tau k} < e_{j\tau k} \}$, ensemble des critères favorables à x_j ,
- $K_\tau^=(x_i, x_j) = \{ k/e_{i\tau k} = e_{j\tau k} \}$, ensemble des critères pour lesquels x_i et x_j sont indiscernables.

En faisant intervenir les coefficients de pondération, on introduit des poids de préférence :

$$P_\tau^+(x_i, x_j) = \sum_{k \in K_\tau^+(x_i, x_j)} \beta_k \quad P_\tau^-(x_i, x_j) = \sum_{k \in K_\tau^-(x_i, x_j)} \beta_k \quad P_\tau^=(x_i, x_j) = \sum_{k \in K_\tau^=(x_i, x_j)} \beta_k$$

Ces coefficients permettent de construire un indice de concordance :

$$c_\tau(x_i, x_j) = P_\tau^+(x_i, x_j) + P_\tau^=(x_i, x_j).$$

On définit enfin un indice de discordance par :

$$d_\tau(x_i, x_j) = \begin{cases} 0 & \text{si } K_\tau^-(x_i, x_j) = \emptyset \\ \text{Max}_{k \in K_\tau^-(x_i, x_j)} \frac{|e_{i\tau k} - e_{j\tau k}|}{L_k} & \text{Sinon} \end{cases}$$

$$\text{où } L_k = \text{Max}_{1 \leq i, j \leq n} |e_{i\tau k} - e_{j\tau k}|$$

L'indice de concordance représente le poids de la somme des critères pour lequel x_i est au moins aussi bon que x_j . L'indice de discordance est calculé sur le critère pour lequel il y a le plus "d'effet contraire", c'est-à-dire pour lequel x_j est meilleur que x_i avec un écart relatif maximum. La part relative de cet écart est mesurée par L_k , que l'on peut interpréter comme la façon de noter de chaque juge. L'indice de discordance permet ainsi de pénaliser un objet trop mal noté pour un critère, même mineur, cette faiblesse pouvant être considérée comme rédhibitoire.

A partir de ces deux indices, on construit une relation de surclassement S_τ , par :

$$x_i S_\tau x_j \Leftrightarrow \begin{cases} c_\tau(x_i, x_j) \geq q_1 \\ d_\tau(x_i, x_j) \leq q_2 \end{cases}$$

où q_1 et q_2 sont des valeurs fixées, indépendantes du juge λ_τ et comprises entre 0 et 1, appelées respectivement seuil de concordance et seuil de discordance.

On obtient ainsi un graphe orienté $G_\tau = (X, U_\tau)$, appelé graphe de surclassement du juge λ_τ , où X est l'ensemble des objets et U_τ l'ensemble des arcs défini par :

$$(x_i, x_j) \in U_\tau \Leftrightarrow x_i S_\tau x_j.$$

1-2. Matrice de préférences

Pour chaque juge λ_τ ($1 \leq \tau \leq m$), on élabore, avec les mêmes seuils q_1 et q_2 fixés, un graphe de surclassement G_τ , dont la matrice des valuations est notée M_τ . Soit :

$$M = \sum_{1 \leq \tau \leq m} M_\tau$$

Le graphe valué G associé à $M = [m_{ij}]$ (pour $1 \leq i, j \leq n$) est appelé graphe de surclassement généralisé, et peut être interprété comme une superposition de l'ensemble des graphes G_τ . La valeur m_{ij} représente le nombre de juges pour lesquels x_i surclasse x_j .

Les éléments diagonaux de cette matrice sont constants et égaux à m (nombre de juges). Ils correspondent à des boucles de poids m sur chacun des sommets du graphes.

En supprimant ces arcs, on obtient un nouveau graphe orienté appelé graphe de préférences dont la matrice des valuations $A = [a_{ij}]$ ($1 \leq i, j \leq n$) est définie par :

$$a_{ij} = \begin{cases} 0 & \text{si } i = j \\ m_{ij} & \text{sinon} \end{cases}$$

A est appelée matrice de préférences.

2. Recherche d'un classement des objets :

Le problème de choix multicritères par plusieurs juges a donc été ramené à un problème faisant appel à des notions analogues à des comparaisons par paires en considérant la matrice $A=[a_{ij}]$. L'objectif que nous nous sommes fixés est de trouver un ordre total (ou un ensemble d'ordres totaux) sur l'ensemble X , compatible (en un sens à définir) avec la matrice A .

2-1. Problème de Slater

Le problème de Slater [11] consiste à ajuster un ordre total à un tournoi donné ou plus généralement à agréger un profil (i.e. un ensemble) de tournois en un ordre total.

La première phase décrite plus haut permet d'aboutir, pour chacun des juges, à une matrice de préférences qui n'est pas, en général, celle d'un tournoi. Toutefois et dans la mesure où chaque objet est confronté un même nombre de fois à un autre, l'assimilation de notre problème à un problème de Slater est envisageable.

Soit un ordre total ρ et le profil des m jugements sur les objets. Nous dirons que l'ordre le plus proche du profil au sens de Slater est l'ordre tel que le nombre des classements incompatibles entre le profil et l'ordre soit minimum. On cherche donc à minimiser le poids des arcs à inverser dans G afin d'obtenir un graphe transitif.

La fonction objectif à minimiser s'écrit :

$$f_1(\rho) = \sum_{\rho(x_i) > \rho(x_j)} a_{ij}$$

Il est bien connu que la recherche des ordres médians est un problème NP-Complet. Il existe un certain nombre d'algorithmes pour résoudre ce problème. La plupart sont des heuristiques qui, partant d'une solution construite à partir d'un critère donné (méthodes des scores par exemple), permettent d'obtenir des solutions approchées. Toutefois il existe des méthodes arborescentes dont une développée récemment par Barthélémy, Guenoche et Hudry [1], permettant d'énumérer toutes les solutions exactes. La complexité est, dans ce cas, exponentielle.

2-2. Pondération par les rangs

Le problème de Slater ne tient aucun compte de la différence entre les rangs, dans l'ordre recherché, des deux sommets d'un arc à inverser. Il peut donc donner, par exemple, comme ordre total le plus proche, celui qui inverse le premier objet et le dernier.

Pour pallier cet inconvénient, nous avons "pénalisé" chaque arcs à inverser par la différence des rangs entre les deux objets concernés, Vidal-Yehia [12].

Cette méthode consiste à trouver une bijection ρ de X dans $I=\{1,2,\dots, n\}$ (i.e. un ordre) minimisant la fonction objectif suivante :

$$f(\rho) = \sum_{\rho(x_j) < \rho(x_i)} a_{ij} (\rho(x_i) - \rho(x_j))$$

Soit $\sigma \in S_n$, la permutation associée à ρ par $x_{\sigma(i)} = \rho(x_i)$, la fonction f définie précédemment s'écrit :

$$f(\sigma) = \sum_{i,j} a_{ij} d_{\sigma(i)\sigma(j)}, \quad \text{où } d_{st} = \begin{cases} t-t \text{ si } t > s \\ 0 \text{ sinon} \end{cases}$$

Ce problème est un problème d'affectation quadratique décrit comme étant NP-Complet par Beghin-Picavet et Hansen [2]. Ce problème étudié dans [5], est approché par un algorithme du type programmation dynamique dont les formules de récurrence sont les suivantes : soit $S \subset X$

$$\begin{cases} f_s = \min_{k \in S} \{ f_{s-(k)} \} + \sum_{\substack{j \in S \\ i \in X-S}} a_{ij} & \text{si } |s| \neq 0 \\ f_s = 0 & \text{sinon} \end{cases}$$

Soit k_0 l'élément de S minimisant f_s , alors le classement partiel de S , noté classe(S) est classe($S - \{k_0\}$), k_0 .

Le classement obtenu par cette méthode est équivalent à celui donné par la méthode des scores dans le cas de tournois ou de tournois généralisés (c'est à dire pour $a_{ij} + a_{ji} = \text{cte}$), voir [3]. Nous verrons que cette propriété revêt une importance toute particulière pour notre problématique.

2-3. Recherche de l'ordre médian à distance minimum d'un profil de relations binaires

Les matrices de préférences que nous avons construites ne sont pas des matrices de tournois. Le problème posé est donc légèrement différent du problème de Slater. Pour tenir compte de cette particularité dans la recherche d'un ordre, nous considérons comme "incohérences" de jugements à la fois les arcs allant d'un objet à un autre mieux classé mais aussi les arcs manquants qui devraient apparaître entre un objet et ceux qui, dans l'ordre, sont moins bien classés que lui. Le critère est celui consistant à rechercher l'ordre à distance minimum du profil des relations associés aux graphes de surclassement au sens de la distance de la différence symétrique.

2-3-1. Codage d'une Relation Binaire

Soit le codage caractéristique α sur un préordre ρ . La coordonnée de α indexée par le couple (x_i, x_j) est notée $\alpha(x_i, x_j)$. (Monjardet-Giamoukis [7]).

$$\left| \begin{array}{l} \alpha(x_i, x_j) = 1 \text{ si } x_i \rho x_j \\ \alpha(x_i, x_j) = 0 \text{ sinon} \end{array} \right.$$

ρ étant un préordre total, on en déduit que $\alpha(x_i, x_j) = 1$ si x_i est au moins aussi bien placé que x_j dans ρ .

A toute relation binaire τ , on peut associer une matrice $A^\tau = [a^{\tau_{ij}}]$. On peut alors généraliser la définition du codage α à une telle matrice :

$$\left| \begin{array}{l} \alpha^\tau(x_i, x_j) = 1 \iff a^{\tau_{ij}} = 1 \\ \alpha^\tau(x_i, x_j) = 0 \iff a^{\tau_{ij}} = 0 \\ \alpha^\tau(x_i, x_i) = 0 \end{array} \right.$$

Remarque : un préordre total ρ peut être représenté par la matrice A^ρ où $a^{\rho_{ij}} = 1$ si $\rho(x_i) < \rho(x_j)$ et 0 sinon. A^ρ est un cas particulier de matrice de préférences.

2-3-2. Distance entre deux relations binaires

A partir de la notion de codage d'une matrice de préférence, on définit un écart entre deux relations en construisant une distance entre les deux matrices associées à ces relations. On utilise généralement la distance euclidienne.

En ce qui concerne des données binaires, on peut utiliser également la distance de Hamming. Elle est définie comme étant le nombre de coordonnées différentes entre les deux vecteurs codant les relations binaires.

Dans le cadre d'un codage α , les deux distances sont équivalentes. Elles sont aussi égales à une distance connue sous le nom de distance de la différence symétrique (Wakabayashi [13]).

Nous utiliserons donc cette distance, en envisageant le cas de deux préordres totaux et en généralisant ensuite à des relations quelconques.

Soit ρ , un préordre total codé par α . Il induit une partition de X^2 en deux sous-ensembles :

$$\begin{aligned}\pi &= \{(x_i, x_j) \in X^2 / \rho(x_i) < \rho(x_j)\} \\ \bar{\pi} &= \{(x_i, x_j) \in X^2 / \rho(x_i) \geq \rho(x_j)\}\end{aligned}$$

Deux préordres totaux ρ et ρ' induisent donc une partition de X^2 en quatre sous-ensembles suivant le schéma ci-dessous :

	$\rho(x_i) < \rho(x_j)$ $\alpha(x_i, x_j) = 1$	$\rho(x_i) \geq \rho(x_j)$ $\alpha(x_i, x_j) = 0$
$\rho'(x_i) < \rho'(x_j)$ $\alpha'(x_i, x_j) = 1$	a	b
$\rho'(x_i) \geq \rho'(x_j)$ $\alpha'(x_i, x_j) = 0$	c	d

a, b, c et **d** sont les cardinaux des quatre sous-ensembles ainsi définis :

a est le nombre de couples (x_i, x_j) pour lesquels x_i est mieux classé que x_j dans ρ et ρ' .

b est le nombre de couples (x_i, x_j) pour lesquels x_i est mieux classé que x_j dans ρ' mais pas dans ρ .

c est le nombre de couples (x_i, x_j) pour lesquels x_i est mieux classé que x_j dans ρ mais pas dans ρ' .

d est le nombre de couples (x_i, x_j) pour lesquels x_i est moins bien classé que x_j dans ρ et dans ρ' .

Soit δ la distance de la différence symétrique. On a :

$$\delta(\rho, \rho') = |\pi \cap \bar{\pi}'| + |\pi' \cap \bar{\pi}|$$

$$\text{d'où } \delta(\rho, \rho') = \mathbf{b} + \mathbf{c}$$

Remarque : dans le cas où ρ et ρ' sont deux ordres totaux, si $x_i = x_j : \rho(x_i) \geq \rho(x_j)$ implique $\rho(x_i) > \rho(x_j)$. On en déduit facilement qu'alors, $b=c$ et $a=d$.

Cette approche peut être aisément généralisée à deux matrices de préférences A et A' . Ces matrices permettent d'effectuer une partition de X^2 en quatre sous-ensembles, comme ci-dessus.

La distance de la différence symétrique est égale au nombre de couples (x_i, x_j) tels que :

$$\text{soit } \begin{cases} \alpha(x_i, x_j) = 1 \\ \alpha'(x_i, x_j) = 0 \end{cases} \quad \text{soit } \begin{cases} \alpha(x_i, x_j) = 0 \\ \alpha'(x_i, x_j) = 1. \end{cases}$$

2-3-3. Recherche d'un ordre médian

Nous venons de définir la notion de distance entre deux relations binaires. Cette définition nous permet de poser le problème de la recherche du classement optimum des objets de la manière suivante :

à partir des m matrices de préférences individuelles A^r , on recherche un consensus sous la forme d'un ordre tel que la distance entre l'ensemble des relations binaires et cet ordre soit minimum.

On a vu que la matrice associée à un ordre pouvait être considérée comme un cas particulier de matrice de préférences. On peut donc définir comme ci-dessus la distance de la différence symétrique entre une matrice A^p codée par α et associée à un ordre total p et la matrice de préférences du juge λ_τ , soit A^τ , codée par α^τ .

	$\alpha^\tau(x_i, x_j) = 1$	$\alpha^\tau(x_i, x_j) = 0$
$\rho(x_i) < \rho(x_j)$ $\alpha(x_i, x_j) = 1$	a	b
$\rho(x_i) > \rho(x_j)$ $\alpha(x_i, x_j) = 0$	c	d

$$\delta(\alpha^\tau, \alpha) = \mathbf{b} + \mathbf{c}$$

Interprétation en terme d'arcs

* **c** représente le nombre de couples (x_i, x_j) tels que $\alpha(x_i, x_j) = 0$ et $\alpha^\tau(x_i, x_j) = 1$.

c est donc le nombre d'arcs du graphe G^τ allant d'un objet à un autre mieux classé dans p (i.e. le nombre d'arcs à inverser au sens de Slater).

$$\mathbf{c} = \sum_{\rho(x_j) < \rho(x_i)} \alpha^\tau(x_i, x_j) = \sum_{\rho(x_j) < \rho(x_i)} a^{\tau_{ij}}$$

* **b** est le nombre de couples (x_i, x_j) pour lesquels $\alpha(x_i, x_j) = 1$ et $\alpha^\tau(x_i, x_j) = 0$.

b est donc le nombre d'arcs manquant du graphe G^τ allant d'un objet à un autre, moins bien classé dans p .

$$\mathbf{b} = \sum_{\rho(x_j) > \rho(x_i)} (1 - \alpha^\tau(x_i, x_j)) = \sum_{\rho(x_j) > \rho(x_i)} (1 - a^{\tau_{ij}})$$

ou encore (Lavialle, Qannari, Vidal-[5]) :

$$\mathbf{b} = \sum_{\rho(x_j) < \rho(x_i)} (1 - \alpha^\tau(x_j, x_i)) = \sum_{\rho(x_j) < \rho(x_i)} (1 - a^{\tau_{ji}})$$

Soit A la matrice de préférences globales. On définit la distance de A à la matrice A^ρ associée à un ordre ρ , comme la somme des distances entre les matrices élémentaires A^τ et A^ρ .

$$d(A, A^\rho) = \sum_{1 \leq \tau \leq m} \delta(A^\tau, A^\rho) = \sum_{1 \leq \tau \leq m} \delta(\alpha^\tau, \alpha)$$

Chercher l'ordre ρ qui minimise $d(A, A^\rho)$ revient donc à chercher l'ordre qui minimise :

$$\begin{aligned} \sum_{1 \leq \tau \leq m} b_\tau + c_\tau &= \sum_{1 \leq \tau \leq m} \left(\sum_{\rho(x_j) < \rho(x_i)} (1 - a^{\tau_{ji}}) + \sum_{\rho(x_j) < \rho(x_i)} a^{\tau_{ij}} \right) \\ &= \sum_{1 \leq \tau \leq m} \left(\sum_{\rho(x_j) < \rho(x_i)} (a^{\tau_{ij}} + 1 - a^{\tau_{ji}}) \right) \\ &= \sum_{\rho(x_j) < \rho(x_i)} (a_{ij} + m - a_{ji}) \end{aligned}$$

Soit la matrice $A' = [a'_{ij}]$ telle que $a'_{ij} = a_{ij} + m - a_{ji}$. A' est la matrice d'un tournoi généralisé.

Minimiser la distance $d(A, A^\rho)$ revient donc à chercher le minimum de la fonction $f(\rho)$ définie par :

$$f(\rho) = \sum_{\rho(x_j) < \rho(x_i)} a'_{ij}$$

La recherche d'un ordre médian minimisant la somme des distances de la différence symétrique avec un ensemble de matrices de préférences consiste donc à minimiser le critère de Slater sur la matrice A' , (Lavialle, Vidal [6]). Le nombre d'arcs à inverser dans A' représente, pour la matrice A le minimum de la somme des arcs à inverser et de ceux à rajouter pour que le graphe associé à A soit transitif et complet.

2-4. Distance pondérée

Soit la partition induite par un ordre ρ et une matrice A^τ , (voir 2-3-3)

Les couples (x_i, x_j) appartenant aux ensembles de cardinal \mathbf{b} et \mathbf{c} sont les couples pour lesquels ρ et A^τ sont en désaccord.

La distance de Hamming est augmentée de 1 à chaque fois qu'un désaccord intervient.

En suivant une démarche parallèle à celle de Frey-Yehia [4], nous proposons de tenir compte de l'importance du désaccord en pénalisant un ordre d'autant plus fortement que les désaccords avec A^τ portent sur des objets dont la différence des rangs est importante.

Nous pondérons donc chaque désaccord par la différence, dans ρ , entre les rangs des deux objets concernés.

Ainsi notons :

$$\begin{aligned} C &= \sum_{\rho(x_i) > \rho(x_j)} a^{\tau_{ij}} (\rho(x_i) - \rho(x_j)) \\ &= \sum_{x_i, x_j} a^{\tau_{ij}} (1 - \alpha_{ij}) (\rho(x_i) - \rho(x_j)) \end{aligned}$$

C est la somme des "pénalités" infligées aux arcs à inverser

$$\begin{aligned} b &= \sum_{\rho(x_i) < \rho(x_j)} (1 - a^{\tau_{ij}}) (\rho(x_j) - \rho(x_i)) \\ &= \sum_{x_i, x_j} (1 - a^{\tau_{ij}}) \alpha_{ij} (\rho(x_j) - \rho(x_i)) \end{aligned}$$

b est la somme des "pénalités" infligées aux arcs à ajouter

Soit δ_r , définie par :

$$\delta_r(A^\tau, \rho) = C + b$$

δ est une combinaison linéaire de distances. C'est donc, par définition, une distance.

On cherche l'ordre médian au sens de cette distance : c'est l'ordre qui minimise :

$$\sum_{\tau=1}^m \delta_r(A^\tau, \rho) = d_r(A, \rho)$$

$$\begin{aligned}
 d_r(A, \rho) &= \sum_{x_i, x_j} a_{ij} (1 - \alpha_{ij}) (\rho(x_i) - \rho(x_j)) + \sum_{x_i, x_j} (1 - a_{ij}) \alpha_{ij} (\rho(x_j) - \rho(x_i)) \\
 &= \sum_{x_i, x_j} (a_{ij} - \alpha_{ij}) (\rho(x_i) - \rho(x_j)) \\
 &= \sum_{x_i, x_j} a_{ij} (\rho(x_i) - \rho(x_j)) - \sum_{x_i, x_j} \alpha_{ij} (\rho(x_i) - \rho(x_j))
 \end{aligned}$$

où, pour un ordre ρ donné : $\sum_{x_i, x_j} \alpha_{ij} (\rho(x_i) - \rho(x_j))$ est une constante.

Minimiser $d_r(A, \rho)$ revient à minimiser $\sum_{x_i, x_j} a_{ij} (\rho(x_i) - \rho(x_j))$.

$$\begin{aligned}
 \min_{x_i, x_j} \sum_{x_i, x_j} a_{ij} (\rho(x_i) - \rho(x_j)) &= \min \left(\sum_{x_i, x_j} a_{ij} \rho(x_i) - \sum_{x_i, x_j} a_{ij} \rho(x_j) \right) \\
 &= \min \left(\sum_{x_i, x_j} a_{ij} \rho(x_i) - \sum_{x_i, x_j} a_{ji} \rho(x_i) \right) \\
 &= \min \left(\sum_{x_i, x_j} (a_{ij} - a_{ji}) \rho(x_i) \right) \quad (1)
 \end{aligned}$$

minimiser (1) revient à minimiser $\sum s(i) \rho(x_i)$ où $s(i)$ est le score de l'objet x_i dans la matrice de terme général $[a_{ij} - a_{ji}]$. C'est donc le nombre d'arcs allant de l'objet x_i vers les $(n-1)$ autres objets.

L'ordre qui minimise la distance tenant compte des rangs est donc l'ordre des scores calculé sur la matrice A' définie au point 2-3-3 (A' est de terme général $a_{ij} - a_{ji} + m$ où m est le nombre de juges).

Cette dernière fonction objectif permet d'obtenir un problème qui n'est plus un problème NP-complet. De plus, on reproche souvent aux méthodes de recherche d'ordre médian de donner des ordres solutions qui peuvent être très éloignés les uns des autres; la pondération par les rangs permet également de lever cet inconvénient.

Bibliographie

- [1]-**BARTHELEMY JP., GUENOCHE A., HUDRY O.** (1989)-Median linear order : heuristics and a branch and bound algorithm, *European Journal of Operational Research* 42, 1, pp 313-325 , North-Holland.
- [2]-**BEGHIN-PICAVET M., HANSEN P.** (1982)-Deux problèmes d'affectation non linéaires. *R.A.I.R.O. Recherche opérationnelle* n°3, pp. 235-267.
- [3]-**DAVID H.** (1987)-Ranking from unbalanced paired comparison data. *Biometrika* vol 74, n°2 pp. 432-436.
- [4]-**FREY J.J., YEHIA ALCOUTLABI A.** (1986) - Comparaisons par paires : une interprétation et une généralisation de la méthode des scores, *RAIRO - Recherche opérationnelle* , Vol 20 n°3 pp 213-227.
- [5]-**LAVIALLE O., QANNARI E.M., VIDAL C.** (1990) - Agrégation d'ordres sous contrainte : classement de produits à partir de données sensorielles, *Revue de Statistique Appliquée XXXVIII(3)*, pp 83-91.
- [6]-**LAVIALLE O., VIDAL C.** (1991) - Median order of a set of binary relations : application to multicriteria psycho-sensorial judgments. *Rapport interne Laboratoire Bordelais de Recherche en Informatique* n°90-96.
- [7]-**MONJARDET B., GIAMOUKIS V.** (1987) - Coefficients d'accord entre deux préordres totaux, *Statistiques et Analyse des données* Vol 12 n°1 et 2 - pp 46-99.
- [8]-**ROY B.** (1985) - Méthodologie multicritère d'aide à la décision, *Economica-collection Gestion, Paris*.
- [9]-**ROY B.** (1968) - Classement et choix en présence de points de vue multiples. (la méthode ELECTRE). *RIRO* n°8, pp 57-75.
- [10]-**SCHÄRLIG A.** (1985) - Décider sur plusieurs critères. *Presses polytechniques romandes*, Lausanne.
- [11]-**SLATER P.** (1961) - Inconsistencies in schedule of paired comparisons, *Biometrika* Vol 48, pp 303-312.
- [12]-**VIDAL C., YEHIA ALCOUTLABI A.** (1990) - Méthode d'aide à la décision sur des évaluations multicritères, *MISH* n°112, pp 27-36.
- [13]-**WAKAYASHI Y.** (1986) - Aggregation of binary relations : algorithmic and polyhedral investigations, *Dissertation für erlangung des Doktorgrades der Naturwissenschaftlichen Fakultät der Universität Augsburg*.

Getting Results with Negative Thinking

D. Loeb
LaBRI, URA CNRS 1304
Université de Bordeaux I
33405 Talence, France

E. Damiani and O. D'Antona
Dipartimento di Scienze dell'Informazione
Università di Milano
Milano, Italy

February 14, 1992

Abstract

Given a universe of discourse U , a *multiset* can be thought of as a function M from U to the natural numbers N . In this paper, we define a *hybrid set* to be any function from the universe U to the integers Z . These sets are called hybrid since they contain elements with either a positive or negative multiplicity.

Our goal is to use these hybrid sets *as if* they were multisets in order to adequately generalize certain combinatorial facts which are true classically only for nonnegative integers. However, the definition above does not tell us much about these hybrid sets; we must define operations on them which provide us with the needed combinatorial structure.

In section 1, we will define an analog of a set which can contain either a positive or negative number of elements. We will allow sums to be calculated over an arbitrary hybrid set and in particular over "improper" intervals. This will lead us in section 2 to the calculation of the *complementary symmetric functions* $\text{comp}_n(H)$ whose argument is a hybrid set of variables which is at the same time a generalization of elementary and complete symmetric functions.

In particular, as we will see in section 2.2, the complementary symmetric function generalizes the following two classic results concerning sequences of polynomials with *persistent roots*, that is to say, having the form $p_n(x) = (x - a_1)(x - a_2) \cdots (x - a_n)$.

$$p_n(x) = \sum_{k=0}^{\infty} e_k(-a_1, \dots, -a_n) x^{n-k} \quad (1)$$

$$x^n = \sum_{k=0}^{\infty} h_k(a_1, \dots, a_{n-k+1}) p_{n-k}(x). \quad (2)$$

The complementary symmetric function not only allows us to express the connection constants between two arbitrary sequences of polynomials with persistent roots, but to also expand arbitrary rational functions in terms of "extended" sequences with persistent roots and persistent poles.

Many sequences of combinatorial interest c_{nk} can be defined in terms of connection constants. The complementary symmetric function therefore allows us to simultaneously study them: extending them to the case where n and k might be negative, demonstrating their recursions, characterizing them with simple formulas, and providing general inversion formulas. Along these lines, in section 3, we calculate generalized binomial coefficients and give them a combinatorial interpretation in terms of a partial order on the hybrid sets. Similarly, we generalize the Gaussian coefficients, and explain their combinatorial significance. Finally, in section 5, we study a generalization of the p, q -Stirling numbers of the first and second kind.

1 Hybrid Sets

1.1 Notation

One usually thinks of a multiset M as a function from some universe U to the natural numbers \mathbf{N} . $M(x)$ is called the *multiplicity* of x in the multiset M . Sets then are a special case of multisets in which all x have multiplicity either 1 or 0.

In this paper, we generalize these concepts to new "sets" and new "multisets" which may have negative integers as multiplicities as well as nonnegative integers.

Definition 1 (Hybrid Sets) Given a universe U , any function $f : U \rightarrow \mathbf{Z}$ is called a hybrid set.

The value of $f(x)$ is said to be the multiplicity of the element x . If $f(x) \neq 0$ we say x is a member of f and write $x \in f$; otherwise, we write $x \notin f$. If the support of f is finite, define the number of elements $\#f$ to be the sum $\sum_{x \in U} f(x)$. f is said to be an $\#f$ -element hybrid set.

Given two hybrid sets f and g , their sum or union is written as the hybrid set $f + g$ for which the multiplicity of x is the sum of the multiplicities in f and g : $(f + g)(x) = f(x) + g(x)$. Similarly, the hybrid set difference of f and g is denoted $f - g$ and is given by $(f - g)(x) = f(x) - g(x)$.

We will denote hybrid sets by using the usual set braces $\{\}$ and inserting a bar in the middle $\{\bar{\}\}$. Elements occurring with positive multiplicity are written on the left of the bar, and elements occurring with negative multiplicity are written on the right. Order is completely irrelevant.

For example, if $f = \{a, b, c, b|d, e, e\}$ then $f(a) = 1$, $f(b) = 2$, $f(c) = 1$, $f(d) = -1$, $f(e) = -2$, and $f(x) = 0$ for $x \neq a, b, c, d, e$.

So far, we haven't accomplished much. These hybrid sets are equivalent to elements of the free group over the universe in question, so one might wonder if there is anything new here. However, we are planning on treating these hybrid sets as if they were multisets. That is to say, we will consider the subsets of hybrid sets; we will perform sums over hybrid sets; and we will enumerate certain classes of combinatorial objects with them. We begin by considering which hybrid sets correspond to the classical notion of a set.

Definition 2 (New Sets) A positive or classical set is a hybrid set taking only values 0 and 1. A negative set is a hybrid set taking only values 0 and -1 . A new set is either a positive or a negative set.

For example, the set $S = \{a, b, c\}$ can be written as the classical or positive set $S = \{b, a, c\}$ while $-S = \{\bar{a}, b, c\}$ is a negative set. Both S and $-S$ are new sets. However, $\{a|b\}$ is not a new set.

The empty set $\emptyset = \{\bar{\}\}$ is the unique hybrid set for which all elements have multiplicity zero. It is thus simultaneously a positive set and a negative set.

1.2 Sums and Products over a Hybrid Set

One often wishes to sum an expression over all the elements of a *set*. For example, we usually write

$$\phi(n) = \sum_{i \in \text{fact}(n)} i$$

for the Euler phi function where $\text{fact}(n)$ is the set of positive factors of an integer n .

When dealing with *multisets*, it is often convenient to include multiplicities in the sum or product. One defines $\sum_{x \in M} F(x)$ and $\prod_{x \in M} F(x)$ to be $\sum_{x \in U} M(x)F(x)$ and $\prod_{x \in U} F(x)^{M(x)}$ respectively. For example, we can expand a monic polynomial $p(x)$ by

$$p(x) = \prod_{\lambda \in \text{Roots}(p)} (x - \lambda)$$

where $\text{Roots}(p)$ is the multiset of roots of the polynomial p .

This notation can be immediately generalized in the obvious way to *hybrid sets*.

Definition 3 (Sums and Products over Hybrid Sets) For any function F and hybrid set h we write $\sum_{x \in h} F(x)$ for $\sum_{x \in U} h(x)F(x)$, and we write $\prod_{x \in h} F(x)$ for $\prod_{x \in U} F(x)^{h(x)}$.

For example, consider the following extension of the definition of $\text{Roots}(p)$.

Definition 4 (Hybrid Set of Roots) Let $f(x)$ be a rational function. Then $f(x)$ can be written as

$$f(x) = c \frac{(x - a_1)(x - a_2) \cdots (x - a_i)}{(x - b_1)(x - b_2) \cdots (x - b_j)}$$

We then call

$$\text{Roots}(f) = \{a_1, a_2, \dots, a_i | b_1, b_2, \dots, b_j\}$$

the Hybrid Set of Roots of $f(x)$

For $f(x)$ a monic rational function, we thus have

$$f(x) = \prod_{\lambda \in \text{Roots}(f)} (x - \lambda).$$

1.3 Sums and Products with Limits

Often, sets are written using ellipsis for missing elements. For example, $\{a_1, \dots, a_5\}$ stands for the set $\{a_1, a_2, a_3, a_4, a_5\}$. In the case of hybrid sets, this notation turns out to be rather powerful.

Definition 5 (Ellipsis) Let $(a_n)_{n \in \mathbb{Z}}$ be a sequence of constants indexed by an integer n . Their ellipsis is defined by the recursion:

$$\begin{aligned} \{a_1..a_0\} &= \emptyset = \{\} \\ \{a_1..a_n\} &= \{a_1..a_{n-1}\} + \{a_n\}. \end{aligned}$$

It therefore follows that $\{a_i..a_j\}$ is explicitly given by:

$$\{a_i..a_j\} = \begin{cases} \{a_i, a_{i+1}, \dots, a_{j-1}, a_j\} & \text{if } i \leq j, \\ \emptyset & \text{if } i = j + 1, \text{ and} \\ \{\} & \text{if } i \geq j + 2; \end{cases}$$

In the first case, we are reduced to the standard usage of ellipsis. However, in the other cases, the interval being considered is "improper" classically, but the meaning is clear here. Reference is being made to a certain $j - i + 1$ -element negative set. Our ellipsis are written with only two dots to distinguish them from normal ellipsis.

One often uses the notation $\sum_{n=1}^i A(n)$ where n is a nonnegative integer to denote the sum of $A(n)$ over the set $\{1, 2, \dots, n\}$. Using the ellipsis notation, we can define the sum or product of a quantity from an arbitrary integer to another:

Definition 6 (Sums and Products with Limits) Let $(A(n))_{n \in \mathbb{Z}}$ be a sequence of constants indexed by an integer n , and let i and j be arbitrary integers. The sum and product of $A(n)$ from i to j is given by

$$\begin{aligned} \sum_{n=i}^j A(n) &= \sum_{n \in \{i..j\}} A(n) \\ \prod_{n=i}^j A(n) &= \prod_{n \in \{i..j\}} A(n). \end{aligned}$$

Just as with integrals, the sum is shift invariant, the sum over an empty interval is zero, and the sum over a positive expression over an "improper" interval $j < i$ is negative.

For example, $\sum_{i=1}^4 i = 1 + 2 + 3 + 4 = 10$ since $\{1..4\} = \{1, 2, 3, 4\}$, and $\sum_{i=1}^{-4} = -(-3) - (-2) - (-1) - 0 = 6$ since $\{1..-4\} = \{0, -1, -2, -3\}$. In general, $\sum_{i=1}^n i = n(n+1)/2$ for all n regardless of sign. In fact,

Proposition 7 Let $p(x)$ be any polynomial, then there is a polynomial $q(x)$ such that $q(n) = \sum_{i=1}^n p(i)$ for all n regardless of sign. \square

Proof: Classical proof by induction can be applied here *mutatis mutandis*. \square

Finally, note the following reciprocity formula useful in calculations with improper intervals:

Proposition 8 Let $(a_i)_{i \in \mathbb{Z}}$ be a sequence of constants indexed by an integer i , then for all integers n

$$\{a_1..a_n\} = -\{b_1..b_{-n}\}$$

where $b_i = a_{n+i}$. \square

2 The Complementary Symmetric Function

2.1 Introduction

The complementary symmetric function is so called because it *complements* or extends the *complete* and *elementary* symmetric functions. Recall that $e_n(V)$ and $h_n(V)$, the elementary and complete symmetric functions of degree n over the set of variables V , are given by the following generating functions:

$$\sum_{n \geq 0} h_n(V)t^n = \prod_{x \in V} (1 - xt)^{-1}$$

$$\sum_{n \geq 0} e_n(V)t^n = \prod_{x \in V} (1 + xt).$$

The *complementary symmetric function* combines these two definitions by using our notion of the product over a hybrid set.

Definition 9 Let V be a hybrid set of variables. Then the generating function for the complete symmetric function of degree n over these variables is given by the product:

$$\sum_{n \geq 0} \text{comp}_n(V)t^n = \prod_{x \in V} (1 - xt).$$

In other words, we have

$$\sum_{n \geq 0} \text{comp}_n(a_1, a_2, \dots | b_1, b_2, \dots)t^n = \frac{\prod_{i=1}^{\infty} (1 - a_i t)}{\prod_{i=1}^{\infty} (1 - b_i t)} = \frac{\sum_{n=0}^{\infty} e_n(-a_1, -a_2, \dots)t^n}{\sum_{n=0}^{\infty} h_n(a_1, a_2, \dots)t^n}.$$

In the case where V is a negative or positive set, it follows that

$$e_n(a, b, c, \dots) = \text{comp}_n(-a, -b, -c, \dots) \quad (3)$$

$$h_n(a, b, c, \dots) = \text{comp}_n(|a, b, c, \dots). \quad (4)$$

It is now immediate that the transformation

$$* : e_n \mapsto h_n$$

used in the theory of symmetric functions is an involution. $*$ acts on the set of symmetric variables, inverting the sign and the multiplicity of each variable.

The next few results are based on the following convolution identity.

Lemma 10 Suppose V and W are two (hybrid) sets of variables. Then

$$\text{comp}_n(V + W) = \sum_{k=0}^n \text{comp}_k(V) \text{comp}_{n-k}(W).$$

Proof: Using the definition of $V + W$ (Definition 1), we can proceed as follows

$$\begin{aligned} \sum_{n=0}^{\infty} \text{comp}_n(V + W)t^n &= \prod_{x \in V+W} (1 - xt) \\ &= \left(\prod_{x \in V} (1 - xt) \right) \left(\prod_{x \in W} (1 - xt) \right) \\ &= \left(\sum_{n=0}^{\infty} \text{comp}_n(V)t^n \right) \left(\sum_{n=0}^{\infty} \text{comp}_n(W)t^n \right) \\ &= \sum_{n=0}^{\infty} \left(\sum_{k=0}^n \text{comp}_k(V)\text{comp}_{n-k}(W) \right) t^n, \end{aligned}$$

and identify coefficients of t^n . \square

We have the following explicit characterization of the complementary symmetric function.

Theorem 11 *The complementary symmetric function $\text{comp}_n(a_1, a_2, \dots | b_1, b_2, \dots)$ is the sum of all monomials in the a 's and the b 's in which the a 's may appear with no exponent two or greater, and are associated with a minus sign.*

For example,

$$\begin{aligned} \text{comp}_2(a, b, c | x, y, z) &= ab + ac + bc \\ &\quad -ax - ay - az - bx - by - bz - cx - cy - cz \\ &\quad +x^2 + xy + xz + y^2 + yz + z^2. \end{aligned}$$

Proof of Theorem 11: First, by Lemma 10,

$$\text{comp}_n(a_1, a_2, \dots | b_1, b_2, \dots) = \sum_{k=0}^n \text{comp}_k(a_1, a_2, \dots) \text{comp}_{n-k}(b_1, b_2, \dots).$$

Then by equations 3 and 4, we have in turn

$$\text{comp}_n(a_1, a_2, \dots | b_1, b_2, \dots) = \sum_{k=0}^n e_k(-a_1, -a_2, \dots) h_{n-k}(b_1, b_2, \dots)$$

where as we recall the elementary symmetric function is the sum of monomials with no repeated variables, and the complete symmetric function is the sum of monomials possibly with repeated variables. \square

The complementary symmetric function satisfies an interesting recursion:

Theorem 12 *For all n, m, k positive integers, we have*

$$\begin{aligned} \text{comp}_n(a_1, \dots, a_m | b_1, \dots, b_k) &= \text{comp}_n(a_1, \dots, a_{m-1} | b_1, \dots, b_{k-1}) \\ &\quad + (b_k - a_m) \text{comp}_{n-1}(a_1, \dots, a_{m-1} | b_1, \dots, b_k). \end{aligned} \square$$

Proof: Either apply Lemma 10 first with $W = \{a_m|b_k\}$ and then again with $W = \{a_m\}$, or reason with Equations 3 and 4. \square

Note that Theorem 12 holds equally well for $m, k \leq 0$ provided that $\text{comp}_n(a_1, \dots, a_m|b_1, \dots, b_k)$ is read as the complementary symmetric function $\text{comp}_n(\{a_1..a_m\} - \{b_1..b_k\})$ whose argument is the difference of two hybrid sets each given by ellipsis.

The following expansion can be found in [3].

Theorem 13 *The complementary symmetric function is given by the following sum*

$$\text{comp}_n(a_1, a_2, \dots | b_1, b_2, \dots) = \sum_{0 < \alpha_1 < \alpha_2 < \dots < \alpha_n} \prod_{i=1}^n (b_{\alpha_i+1-i} - a_{\alpha_i})$$

over strictly increasing sequences of positive integers.

Proof: Apply Theorem 12 until all complementary symmetric function disappear either by reducing their degree to zero, or by eliminating all their variables. \square

2.2 Connection Constants

In order to state the main result of this section, we must first introduce the concept of persistent roots.

Definition 14 *Let $(b_n)_{n \in \mathbb{Z}}$ be a sequence of constants indexed by an integer, and let*

$$q_n(x) = \prod_{i=1}^n (x - b_i)$$

for all n positive or negative. In other words,

$$\begin{aligned} & \vdots & \vdots \\ q_2(x) &= (x - b_1)(x - b_2) \\ q_1(x) &= (x - b_1) \\ q_0(x) &= 1 \\ q_{-1}(x) &= 1/(x - b_0) \\ q_{-2}(x) &= 1/(x - b_0)(x - b_{-1}) \\ & \vdots & \vdots \end{aligned}$$

Then $(q_n(x))_{n \in \mathbb{Z}}$ is called the sequence of rational functions with persistent roots $(b_n)_{n \in \mathbb{Z}}$.

We can now state the main result concerning the complementary symmetric function.

Theorem 15 Let $f(x)$ be a monic rational function of degree n , and $(b_k)_{k \leq n}$ a sequence of constants indexed by an integer. Then $f(x)$ can be expanded in terms of the sequence of rational functions $(q_k(x))_{k \leq n}$ with persistent roots $(b_k)_{k \leq n}$ as follows:

$$f(x) = \sum_{k=0}^{\infty} \text{comp}_k(\text{Roots}(f) - \{b_1..b_{n-k+1}\})q_{n-k}(x). \tag{5}$$

Notes to Theorem 15:

1. The degree of a rational function $f(x) = p(x)/q(x)$ is defined to be the degree of the numerator $p(x)$ minus the degree of the denominator $q(x)$.
2. For $0 \leq k \leq n$, the coefficient of $q_k(x)$ does not depend on $b_{k+2}, b_{k+3}, \dots, b_n$ even though initial consideration of the problem might seem to indicate that it should.
3. To interpret this theorem when $k > n$, recall the definition of improper intervals from section 3.
4. The coefficient of $q_n(x)$ does not depend on the non-existent " b_{n+1} ," since $\text{comp}_0(V)$ is identically one.
5. The case in which $f(x)$ is a polynomial (see Corollary 17) is much simpler and is a direct consequence of equations 1, 2, 3, 4, and Lemma 10.

Proof of Theorem 15: We proceed by induction on the number of roots and poles which $f(x)$ possesses. The base case is thus $f(x) = 1$. In this case, the right hand side of equation 5 is

$$\sum_{k=0}^{\infty} \text{comp}_k(-\{b_1..b_{n-k}\})q_{n-k}(x).$$

For $k = 0$, the summand is one, since the complementary symmetric function of degree zero is identically one and $q_0(x) = 1$. The remaining terms are all of the form

$$\frac{e_k(-b_0, -b_{-1}, \dots, -b_{2-k})}{(x - b_0)(x - b_{-1}) \dots (x - b_{1-k})}$$

which is zero for lack of sufficiently many arguments to the elementary symmetric function. Thus, the sum is one as required.

Next, towards induction assume Eq. 5 holds for $g(x)$. Now, consider $f(x) = (x - a)g(x)$.

$$\begin{aligned} f(x) &= (x - a) \sum_{k=0}^{\infty} \text{comp}_k(\text{Roots}(g) - \{b_1..b_{n-k}\})q_{n-k}(x) \\ &= \sum_{k=0}^{\infty} \text{comp}_k(\text{Roots}(g) - \{b_1..b_{n-k}\})[(x - b_{n-k}) + (b_{n-k} - a)]q_{n-k-1}(x) \\ &= q_n(x) + \sum_{k=0}^{\infty} [\text{comp}_{k+1}(\text{Roots}(g) - \{b_1..b_{n-k-1}\}) \\ &\quad + (b_{n-k} - a)\text{comp}_k(\text{Roots}(g) - \{b_1..b_{n-k}\})]q_{n-k-1}(x) \\ &= q_n(x) + \sum_{k=0}^{\infty} \text{comp}_k(\text{Roots}(f) - \{b_1..b_{n-k}\})q_{n-k-1}(x) \\ &= \sum_{j=0}^{\infty} \text{comp}_j(\text{Roots}(f) - \{b_1..b_{n-k-1}\})q_{n-k}(x) \end{aligned}$$

where the fourth equality results from the recursion theorem (12).

The proof for $f(x) = g(x)/(x - a)$ can be found by reading the above sequence of equalities in reverse order. □

Theorem 15 gives rise to the following inversion formula.

Corollary 16 Let $(a_n)_{n \in \mathbb{Z}}$, $(b_n)_{n \in \mathbb{Z}}$, $(c_n)_{n \in \mathbb{Z}}$, and $(d_n)_{n \in \mathbb{Z}}$ be sequences of real numbers. Then

$$c_n = \sum_{k=0}^{\infty} \text{comp}_k(\{a_1..a_n\} - \{b_1..b_{n-k+1}\})d_{n-k}$$

for all n , if and only if

$$d_n = \sum_{k=0}^{\infty} \text{comp}_k(\{b_1..b_n\} - \{1..a_{n-k+1}\})c_{n-k}$$

for all n . \square

Theorem 15 is particularly easy to state in the case of polynomials.

Corollary 17 Given constants $(a_i)_{i=1}^{\infty}$ and $(b_i)_{i=1}^{\infty}$, we have

$$(x - a_1) \cdots (x - a_n) = \sum_{k=0}^n \text{comp}_k(a_1, \dots, a_n | b_1, \dots, b_{k+1})(x - b_1) \cdots (x - b_k). \square$$

As special cases, we have Eqs. 1 and 2. Other special cases of particularly interesting connection constants along with their combinatorial interpretation in the section below and throughout the rest of this paper.

2.3 Example of Connection Constants

We begin with an example of the expansion of an inverse formal power series since this lies outside of the usual theory. Let us expand $1/x$ in terms of the sequence of rational functions $q_n(x)$ with persistent poles $b_0, b_{-1}, b_{-2}, \dots$; that is, expand $1/x$ with respect to

$$\begin{aligned} q_{-1}(x) &= 1/(x - b_0), \\ q_{-2}(x) &= 1/(x - b_0)(x - b_{-1}), \\ q_{-3}(x) &= 1/(x - b_0)(x - b_{-1})(x - b_{-2}), \text{ etc.} \end{aligned}$$

By theorem 15, we can write

$$\begin{aligned} \frac{1}{x} &= \sum_{k=0}^{\infty} \text{comp}_k(\{0, b_0, b_{-1}, \dots, b_{1-k}\})q_{-1-k}(x) \\ &= \sum_{k=0}^{\infty} e_k(-b_0, -b_{-1}, \dots, b_{1-k})q_{-1-k}(x) \\ &= \sum_{k=0}^{\infty} (-1)^k b_0 b_{-1} \cdots b_{1-k} q_{-1-k}(x). \end{aligned}$$

So in particular,

$$\frac{1}{x} = \sum_{n=1}^{\infty} \frac{1}{(x+1)^n}$$

for $x > 0$ as the convergent sum of a geometric series. Similarly,

$$\frac{1}{x} = \sum_{n=1}^{\infty} \frac{(k-1)!}{(x+1)(x+2) \cdots (x+k)}. \quad (6)$$

Using the classical method of telescoping series, equation 6 can be validated for x a positive integer. However, to demonstrate this identity in general for $x > 0$, the preceding theory of connection constants is necessary.

Other examples of connection constants can be found in the following sections.

3 Binomial Coefficients

3.1 Connection Constants

Another classical example of connection constants is that of the binomial coefficient $\binom{n}{k}$ which is given by the coefficients of the polynomial $(x + 1)^k$, or more generally the Gaussian coefficients $\binom{n}{k}_q$ which is given by the coefficients of the polynomial $(x - 1)(x - q)(x - q^2) \dots (x - q^{n-1})$ (see Figure 2 and 3). These constants then have well known combinatorial interpretations. $\binom{n}{k}$ is the number of k -element subsets of a given n -element set whereas the coefficient of q^t in $\binom{n}{k}_q$ counts the number of partitions $\lambda \vdash t$ of length k with all parts distinct and less than or equal to n where a partition of t is defined as a nonincreasing sequence of nonnegative integers summing to t .¹

Now, that we are capable of calculating these connection constants via Theorem 15, we can equally well interpret them for n negative or positive. We are then calculating the coefficient of x^k in an inverse Laurent series.

By Theorem 12, we have the recurrences

$$\binom{n}{k}_q = \binom{n-1}{k-1}_q + q^{n-1} \binom{n-1}{k}_q \tag{7}$$

$$\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}. \tag{8}$$

Moreover, by theorem 15, we have

$$\begin{aligned} \binom{n}{k}_q &= \text{comp}_{n-k}(-1, -q, \dots, -q^{n-1}) \\ &= \begin{cases} e_{n-k}(1, q, \dots, q_{n-1}) & \text{for } 0 \leq k \leq n, \text{ and} \\ (-1)^{n-k} h_{n-k}(q^{-1}, q^{-2}, \dots, q^n) & \text{for } k \leq n < 0. \end{cases} \tag{9} \\ &= \frac{\prod_{i=1}^k q^{n+1-i} - 1}{(q^k - 1)(q^{k-1} - 1) \dots (q - 1)}. \end{aligned}$$

And in particular for $q = 1$,

$$\begin{aligned} \binom{n}{k} &= \text{comp}_{n-k}(\underbrace{-1, -1, \dots, -1}_{n \text{ terms}}) \\ &= \begin{cases} e_{n-k}(\underbrace{1, 1, \dots, 1}_{n \text{ terms}}) & \text{for } 0 \leq k \leq n, \text{ and} \\ (-1)^{n-k} h_{n-k}(\underbrace{1, 1, \dots, 1}_{-n \text{ terms}}) & \text{for } k \leq n < 0 \end{cases} \\ &= \frac{n(n-1)(n-2) \dots (n-k+1)}{k!}. \end{aligned}$$

¹Our definition of the Gaussian coefficient differs from with certain other definitions (for example, see [10]) by a factor of $q^{(n-k)(n-k-1)/2}$. Under this alternate definition, the coefficients of $\binom{n}{k}_q$ then count the number of partitions $\lambda \vdash t$ of length k with all parts less than or equal to $n - k$. Moreover for q prime, $\binom{n}{k}_q$ counts the number of k -dimensional subvector spaces in an n -dimension vector spaces over a q -element field.

Setting all the variables of the elementary or complete symmetric function to one gives the number of terms. In the case of the elementary symmetric function, this is the number of ways of choosing k variables among n without replacement. Thus, we have rediscovered the classical combinatorial interpretation of $\binom{n}{k}$ for n positive. For the complete symmetric function, the number of terms is the number of ways of choosing k variables among n with replacement. In other words, up to sign, $\binom{-n}{k}$ is the number of k -element multisets based on a given n -element set. Hence, we have combinatorial interpretations of $\binom{n}{k}$ for n positive or negative.

3.2 Roman Coefficients

What is the connection between these two interpretations? And can we permit k to be negative as well?

We can for example, adopt the following definition of $\binom{n}{k}$ (consistent with the definition above) for all values of n and k , positive or negative (see Figure 1).

$$\binom{n}{k} = \lim_{\epsilon \rightarrow 0} \frac{\Gamma(n+1+\epsilon)}{\Gamma(k+1+\epsilon)\Gamma(n-k+1+\epsilon)} \tag{10}$$

Table 1: Binomial Coefficients, $\binom{n}{k}$

$n \backslash k$	-4	-3	-2	-1	0	1	2	3	4	5	6
6	0	0	0	0	1	6	15	20	15	6	1
5	0	0	0	0	1	5	10	10	5	1	0
4	0	0	0	0	1	4	6	4	1	0	0
3	0	0	0	0	1	3	3	1	0	0	0
2	0	0	0	0	1	2	1	0	0	0	0
1	0	0	0	0	1	1	0	0	0	0	0
0	0	0	0	0	1	0	0	0	0	0	0
-1	-1	1	-1	1	1	-1	1	-1	1	-1	1
-2	3	-2	1	0	1	-2	3	-4	5	-6	7
-3	-3	1	0	0	1	-3	6	-10	15	-21	28
-4	1	0	0	0	1	-4	10	-20	35	-56	84
-5	0	0	0	0	1	-5	15	-35	70	-126	210

Proposition 18 (The Six Regions) Let n and k be integers. Depending on what region of the Cartesian plane (n, k) is in, the following formulas apply:

1. If $n \geq k \geq 0$, then $\binom{n}{k} = \frac{n!}{k!(n-k)!}$.
2. If $k \geq 0 > n$, then $\binom{n}{k} = (-1)^k \binom{-n+k-1}{k}$.
3. If $0 > n \geq k$, then $\binom{n}{k} = (-1)^{n+k} \binom{-k-1}{n-k}$.
4. If $k > n \geq 0$, $\binom{n}{k} = 0$.
5. If $0 > k > n$, $\binom{n}{k} = 0$.
6. If $k \geq 0 > n$, $\binom{n}{k} = 0$. □

Note that in regions 4–6 there is an extra factor of ϵ in the numerator of the limit, so we are left with zero. Next, region 1 is the classical case, so we have the usual binomial coefficients. Finally, regions 2 and 3 are simply related to region 1.

Most of the usual properties of binomial coefficients hold true in all six regions.

Proposition 19 (Complementation) For all integers n and m , $\binom{n}{m} = \binom{n}{n-m}$. \square

Proposition 20 (Iteration) For all integers i, j , and k ,

$$\binom{i}{j} \binom{j}{k} = \binom{i}{k} \binom{i-j}{j-k}. \square$$

By equation 8,

Proposition 21 (Pascal) Let n and k be integers not both zero, then

$$\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1}. \square$$

Nevertheless, $\binom{0}{0} = 1$ while $\binom{-1}{-1} + \binom{-1}{0} = 1 + 1 = 2$.

The coefficients in Regions 1 and 3 have already been seen to be connection constants; that is to say, they are the coefficients of x^{n-k} in the inverse power series $(1+x)^n$. Those in regions 4–6 are zero. This leaves only region 3, which as we see below, can be interpreted as a set of connection constants in terms of formal power series.

Proposition 22 For all integers n , $(x+1)^n$ is given by

$$\sum_{k=-\infty}^n \binom{n}{k} x^{n-k}$$

as a formal power series.

Proof: The $(n-k)$ th derivative of $(1+x)^n$ evaluated at zero is $(n-k)! \binom{n}{k}$. \square

3.3 Inclusion of Hybrid Sets

These new binomial coefficients $\binom{n}{k}$ are always integers, but what do they count? In region 1, they count the number of k -element subsets of a given n -element set. We claim that given a suitable generalization of the notion of a subset, this is true in general (up to sign). There is no need for a separate interpretation of $\binom{n}{k}$ in regions 2 and 3.

We define a partial order on these sets which is a generalization of the usual ordering of classical sets and multisets by inclusion.

Definition 23 (Subsets) *Informally, we say f is a subset of g (and write $f \subseteq g$) if one can remove elements one at a time from g (never removing an element that is not a member of g) and thus either achieve f or have removed f .*

For example, we might start with the hybrid set $f = \{a, b, c|d, e\}$. We will remove a few of its elements one at a time. Suppose we remove b ; this leaves $\{a, c|d, e\}$. Now, b is no longer an element, so we can not remove it again. Instead, we might remove d leaving $\{a, c|d, e\}$. Obviously, we can remove d as many times as we choose. Finally, suppose we remove c leaving $\{a, c|d, e\}$. Hence, we have proven two things. Since we were able to remove $\{b, d, c\}$, we know that $\{b, d, c\} \subseteq f$. Also, since we were left with $\{a, c|d, e\}$, we know that $\{a, c|d, e\} \subseteq f$.

Proposition 24 [14] *The ordering " \subseteq " defined above is a well defined partial order.* \square

Proof: See [14].

Note however that the ordering \subseteq does not form a lattice. For example, hybrid sets $f = \{a, b\}$ and $g = \{a|b\}$ have lower bounds $\{a\}$, $\{a, b\}$, and $\{a, b, b\}$, but no greatest lower bound. Conversely, f and g possess no upper bounds.

Theorem 25 *The subsets of a classical set f correspond to the classical subsets of a classical set. The subsets of a multiset f correspond to the classical submultisets of a multiset.*

Proof: To construct a conventional subset of a set S , we merely remove some elements subject to the conditions that we only remove elements of S and we don't remove an element twice. The order of removal is not relevant.

To construct a conventional submultiset of a multiset M , we merely remove some elements subject to the conditions that we do not remove any element more times than its multiplicity in M . \square

Now, we have the desired interpretation of the binomial coefficients.

Theorem 26 *Let n and k be arbitrary integers. Let f be an n -element new set. Then $|\binom{n}{k}|$ counts the number of k -element hybrid sets which are subsets of f .* \square

Let us consider this result in each of the six regions mentioned in Proposition 18.

1. This is the only classical case. In this region, one might count the number of 2-element subsets of the set $\{a, b, c, d\}$. By Theorem 25, we enumerate the usual subsets: $\{a, b\}$, $\{a, c\}$, $\{a, d\}$, $\{b, c\}$, $\{b, d\}$, and $\{c, d\}$ but no others.
2. In this region, one might count the number of 2-element subsets of $f = \{[a, b, c]$. These subsets correspond to what we can remove from f , since what we would have left over after a removal would necessarily contain a negative number of elements. We can remove any of the three elements any number of times, so we have: $\{a, a\}$, $\{b, b\}$, $\{c, c\}$, $\{a, b\}$, $\{b, c\}$, and $\{a, c\}$.
3. Here, we are interested in -5 -element subsets of f . Since f contains -3 elements, we must start with f and remove 2 elements. Thus, there is one subset here for each subset in the corresponding position in region 2. In this case they are: $\{[a, a, a, b, c]$, $\{[a, b, b, b, c]$, $\{[a, b, c, c, c]$, $\{[a, a, b, b, c]$, $\{[a, b, b, c, c]$, and $\{[a, a, b, c, c]$.
4. By Theorem 25, there are no 6-element subsets of the set $\{a, b, c, d\}$. Once you remove 4 elements, you can not remove any more.
5. There are no -2 -element subset of a -3 -element set f . If we remove elements from f , we are left with less than -3 elements, and have removed a positive number of elements. In neither case have we qualified a -2 -element hybrid set to be a subset of f .
6. Again by Theorem 25, there are no -2 -element subsets of the set $\{a, b, c, d\}$, since you are not allowed to introduce elements with negative multiplicities.

4 Linear Partitions

Table 2: Gaussian Coefficients, $\binom{n}{k}_q$ —Region 1

$n \setminus k$	0	1	2	3	4
5	q^{10}	$q^6 + q^7 + q^8$	$q^3 + q^4 + 2q^5 + 2q^6$	$q + q^2 + 2q^3 + 2q^4$	$1 + q + q^2$
4	q^6	$q^3 + q^4 + q^5 + q^6$	$q + q^2 + 2q^3 + q^4 + q^5$	$1 + q + q^2 + q^3$	1
3	q^3	$q + q^2 + q^3$	$1 + q + q^2$	1	0
2	q	$1 + q$	1	0	0
1	1	1	0	0	0
0	1	0	0	0	0

Table 3: Gaussian Coefficients, $\binom{n}{k}_q$ —Region 2

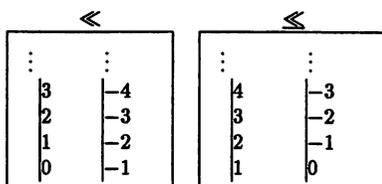
$n \setminus k$	-5	-4	-3	-2	-1
-1	q^{-4}	$-q^{-3}$	q^{-2}	$-q^{-1}$	1
-2	$-q^{-6} - q^{-5} - q^{-4} - q^{-3}$	$q^{-4} + q^{-3} + q^{-2}$	$-q^{-2} - q^{-1}$	1	0
-3	$q^{-6} + q^{-5} + 2q^{-4} + q^{-3} + q^{-2}$	$-q^{-3} - q^{-2} - q^{-1}$	1	0	0
-4	$-q^{-4} - q^{-3} - q^{-2} - q^{-1}$	1	0	0	0
-5	1	0	0	0	0

Just as we did with the binomial coefficients, we must now interpret our generalized Gaussian coefficients according to their combinatorial properties. We will do so in terms of *linear partitions*. One usually defines a (linear) partition or a (linear) partition with distinct parts to be a finite sequence λ of nonnegative integers with $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_k$ or $\lambda_1 > \lambda_2 > \dots > \lambda_k$ respectively. Since we are now giving negative integers a status equal to that of nonnegative ones, the classical definition becomes inappropriate. We now procede to generalize the notion of a partition by first generalizing the relations \geq and $>$.

Definition 27 (Hybrid Inequalities) Given integers i and j . We say that $i \ll j$ or equivalently $j \gg i$ if and only if $i \in \{0, 1, \dots, j-1\}$. Similarly, we say that $i \leq j$ or equivalently $j \geq i$ if and only if $j \in \{1, 2, \dots, j\}$.

Note that these relations have their usual sense for positive integers, but behave rather oddly for negative integers. For example, $-2, -1, 0$ are the integers ≤ -3 , and $-3, -2, -1$ are the integers $\ll -3$. These relations are not partial orders since they do not obey reflexivity. However, they do obey antisymmetry and transitivity, so we can still discuss their Hasse diagrams. As you can see in figure 4, the negative integers

Figure 1: Hasse diagrams of \ll and \leq



are now unrelated to the positive integers,² and the usual order of negative integers has been reversed.

Definition 28 (New Partitions) Given a sequence λ of integers, we say that λ is an n -partition of length k (and width n) if $(n \gg) \lambda_1 \gg \lambda_2 \gg \dots \gg \lambda_k$. We say that λ is a d -partition of length k (and width n) if $(n \gg) \lambda_1 \gg \lambda_2 \gg \dots \gg \lambda_k$.

If t is the sum of the k parts of λ then we write $\lambda \vdash t$ or $|\lambda| = t$, and we say λ is an n -partition or d -partition of t . The length of λ is denoted $\ell(\lambda)$.

By considering the case t positive, we see that the n -partition generalizes the idea of a normal partition and the d -partition generalizes the idea of a partition with distinct parts.³

For example, $(5, 3, 2, 0)$ is a d -partition of 10 with length 4 and minimum width 6. $(5, 5, 3, 2)$ is an n -partition of 15 with length 4 and minimum width 5. $(-32, -7, -7)$ is a d -partition of -46 with length 3 and minimum (maximum?) width -32 . $(-12, -7, -6, -3, 0)$ is an n -partition of -22 with length 5 and minimum (maximum?) width -13 . Note also that a sequence can be at the same time a d -partition and an n -partition, for example $(5, 3, 2)$.

It is well known (see for example [9]) that for n and k nonnegative, $\binom{n}{k}_q$ is a monic polynomial. Its coefficient of q^t is the number of partitions λ of t of length $n - k$ (or equivalently k) with all parts less than or equal to n .

In general, we now have:

Theorem 29 Let k be a nonnegative integer, and n and t be arbitrary integers. Then $\binom{n}{k}_q = \sum_t c_t q^t$ where c_t is the number of d -partitions λ of t with length k , and width n .

Proof: By equation 9.□

We can now express the q -binomial coefficients of regions 1 and 3 in terms of each other.

Corollary 30 Let n and k be nonnegative integers, then

$$\binom{n}{k}_q = (-1)^{n+1} q^{\binom{n+1}{2} - \binom{k+1}{2}} \binom{-k-1}{-n-1}_q. \square$$

² \ll and \leq disagree on whether zero is positive or negative.

³ Note however that for t negative, it is the n -partitions which have distinct parts instead of the d -partitions.

5 Stirling Numbers

5.1 Classical Region

Our next example of connection constants is that of Stirling number of the first and second kind, $s(n, k)$

Table 4: Stirling Numbers of the First Kind, $s(n, k)$ —Region 1

$n \setminus k$	0	1	2	3	4	5	6
6	0	-120	274	-225	85	-15	1
5	0	24	-50	35	-10	1	0
4	0	-6	11	-6	1	0	0
3	0	2	-3	1	0	0	0
2	0	-1	1	0	0	0	0
1	0	1	0	0	0	0	0
0	1	0	0	0	0	0	0

and $S(n, k)$ (see Figures 4 and 5), relating the nonnegative powers of x (x^n) $_{n \geq 0}$ and the lower factorial

Table 5: Stirling Numbers of the Second Kind, $S(n, k)$ —Region 1

$n \setminus k$	0	1	2	3	4	5	6
6	0	1	31	90	65	15	1
5	0	1	15	25	10	1	0
4	0	1	7	6	1	0	0
3	0	1	3	1	0	0	0
2	0	1	1	0	0	0	0
1	0	1	0	0	0	0	0
0	1	0	0	0	0	0	0

polynomials $((x)_n)_{n \geq 0} = (x(x - 1) \cdots (x - n + 1))_{n \geq 0}$

$$(x)_n = \sum_{k=0}^n s(n, k)x^k,$$

$$x^n = \sum_{k=0}^n S(n, k)(x)_k.$$

Their best known combinatorial interpretations is that $S(n, k)$ counts the number of partitions of an n -element set into k blocks while (up to sign) $s(n, k)$ counts the number of partitions of an n -element set containing k -cycles.

When generalizing these constants to negative values of n , we define the lower factorial sequence of rational functions $((x)_n)_{n \in \mathbb{Z}}$ for n an arbitrary integer as the sequence of rational functions with persistent roots $(i - 1)_{i \in \mathbb{Z}}$. Next, we must decide as we did in Proposition 22 whether to consider expansion of $(x)_n$ as a formal power series or as an inverse Laurent series. In Proposition 22, the difference was not critical; the binomial coefficients in regions 2 and 3 are related by Proposition 19. Each region is the reflection of the other.

Table 6: Stirling Numbers of the First Kind $s(n, k)$ —Region 2

$k \setminus n$	-5	-4	-3	-2	-1
6	$\frac{228,576,031,859}{5,588,720,000,000}$	$\frac{11,679,855}{71,663,616}$	$\frac{137,345}{279,938}$	$\frac{127}{128}$	1
5	$-\frac{3,673,451,957}{83,312,000,000}$	$-\frac{952,525}{5,871,988}$	$-\frac{22,631}{46,856}$	$-\frac{63}{64}$	-1
4	$\frac{58,087,611}{1,555,200,000}$	$\frac{76,111}{497,664}$	$\frac{3661}{7776}$	$\frac{31}{32}$	1
3	$-\frac{874,853}{25,920,000}$	$-\frac{5845}{41,472}$	$-\frac{575}{1296}$	$-\frac{15}{16}$	-1
2	$\frac{12,019}{432,000}$	$\frac{419}{3456}$	$\frac{85}{216}$	$\frac{7}{8}$	1
1	$-\frac{187}{7200}$	$-\frac{25}{288}$	$-\frac{11}{36}$	$-\frac{3}{4}$	-1
0	$\frac{1}{120}$	$\frac{1}{24}$	$\frac{1}{6}$	$\frac{1}{2}$	1

5.2 Formal Power Series Region

Here, as we will see, there is no such symmetry. First, let us examine regions 1 and 2 defined by formal power series (see table 6). These constants were considered in [15, 13]. However, except for the classical region 1, we no longer have integers, so there is no combinatorial interpretation to be expected. However, we do have several interesting recursions and identities which prove to be useful in [11] such as

Theorem 31 [13, Porism 4.2 and Proposition 4.12] For all integers n (positive or negative) and nonnegative integers k ,

$$s(n, k) = \lim_{\epsilon \rightarrow 0} \Gamma(1 - n + \epsilon)^{-1} \text{comp}_k \left(\frac{1}{\epsilon}, \frac{1}{-1 + \epsilon}, \frac{1}{-2 + \epsilon}, \dots, \frac{1}{-n + \epsilon} \right). \square$$

We also have the recursion

Theorem 32 [13, Theorem 3.2] For all integers n and for all positive integers k ,

$$s(n + 1, k) = s(n, k - 1) - ns(n, k). \square$$

In region 2, we have the following result due to Knuth.

Proposition 33 [13, Proposition 4.3] Let n and k be nonnegative integers (not both zero). Then $s(n, k)$ is given by the following finite sum:

$$s(-n, k) = \frac{(-1)^{k+1}}{n!} \sum_{m=1}^n \binom{n}{m} (-1)^m m^{-k}. \quad (11)$$

Note that Proposition 33 is the analog of the following classical result [2] involving Stirling numbers of the second kind $S(k, n)$.

$$S(k, n) = \frac{(-1)^n}{n!} \sum_{m=1}^k \binom{n}{m} (-1)^m m^k.$$

Thus, in some sense to be made clear later, we can say that

$$S(k, n) = (-1)^{n+k+1} s(-n, -k).$$

Table 7: Stirling Numbers of the First Kind, $s(n, k)$ —Region 3

$n \setminus k$	-6	-5	-4	-3	-2	-1
-1	-1	1	-1	1	-1	1
-2	31	-15	7	-3	1	0
-3	-90	25	-6	1	0	0
-4	65	-10	1	0	0	0
-5	-15	1	0	0	0	0
-6	1	0	0	0	0	0

Table 8: Stirling Numbers of the Second Kind, $S(n, k)$ —Region 3

$n \setminus k$	-6	-5	-4	-3	-2	-1
-1	120	24	6	2	1	1
-2	274	50	11	3	1	0
-3	225	35	6	1	0	0
-4	85	10	1	0	0	0
-5	15	1	0	0	0	0
-6	1	0	0	0	0	0

5.3 Inverse Laurent Series Region

On the other hand, in the case of inverse Laurent series, we can apply the results of section 2 to calculate $s(n, k)$ and $S(n, k)$. See tables 7 and 8.

$$\begin{aligned}
 s(n, k) &= \text{comp}_{n-k}(0..n-1) \\
 &= \begin{cases} (-1)^{n-k} e_{n-k}(1, 2, \dots, n-1) & \text{for } n \geq k \geq 0, \\ h_{n-k}(-1, -2, \dots, n) & \text{for } 0 > n \geq k. \end{cases} \\
 S(n, k) &= \text{comp}_{n-k}(-\{0..k\}) \\
 &= \begin{cases} h_{n-k}(1, 2, \dots, k) & \text{for } n \geq k \geq 0, \\ (-1)^k e_{n-k}(-1, -2, \dots, k+1) & \text{for } 0 > n \geq k. \end{cases}
 \end{aligned}$$

By Theorem 12, the numbers in regions 1 and 3 obey the identities

$$\begin{aligned}
 s(n, k) &= s(n-1, k-1) - ns(n-1, k) \\
 S(n, k) &= S(n-1, k-1) + kS(n-1, k).
 \end{aligned}$$

These Stirling numbers are always integers, so we can hope to give an extended combinatorial interpretation. We will not only do that, but in fact, we will give a combinatorial interpretation of the extended p, q -Stirling numbers. The p, q -Stirling numbers of the first kind (see tables 9 and 10) and second kind $s_{pq}(n, k)$ and $S_{pq}(n, k)$ (see tables 11 and 12) are the coefficients connecting the powers of x with the p, q -lower factorial $(x; p, q)_n$ where the p, q -lower factorials $((x; p, q)_n)_{n \in \mathbb{Z}}$ form the sequence of rational functions with persistent roots given by the p, q -numbers $(\sqrt{pq}(p^i - q^i)/(p - q))_{i \in \mathbb{Z}}$.⁴

⁴The p, q -numbers are sometimes defined as $(p^i - q^i)/(p - q)$ or simply $p^i - q^i$. These alternate definitions lead to quite similar results.

Table 9: p, q -Stirling Numbers of the First Kind, $s_{pq}(n, k)$ —Region 1

$n \setminus k$	0	1	2	3	4
4	0	$-\sqrt{pq^3}[q^3 + 2q^2p + 2qp^2 - p^3]$	$pq[(q + q^2 + q^3) + (1 + q + 2q^2)p + (1 + 2q)p^2 + p^3]$	$-\sqrt{pq}(1 + q + q^2)$	1
3	0	$pq(q + p)$	$-\sqrt{pq}(1 + p + q)$	1	0
2	0	$-\sqrt{pq}$	1	0	0
1	0	1	0	0	0
0	1	0	0	0	0

Table 10: p, q -Stirling Numbers of the First Kind, $s_{pq}(n, k)$ —Region 3

$n \setminus k$	-4	-3	-2	-1
-1	$-1/\sqrt{pq^3}$	$1/pq$	$-1/\sqrt{pq}$	1
-2	$[p^{-2} + (q^{-1} + 2)p^{-1} + (1 + q^{-1} + q^{-2})]/pq$	$-[p^{-1} + (q^{-1} + 1)]/\sqrt{pq}$	1	0
-3	$[-p^{-2} + (q^{-1} + 1)p^{-1} + (q^{-2} + q^{-1} + 1)]/\sqrt{pq}$	1	0	0
-4	1	0	0	0

Table 11: p, q -Stirling Numbers of the Second Kind, $S_{pq}(n, k)$ —Region 1

$n \setminus k$	0	1	2	3	4
4	0	$\sqrt{pq^3}$	$pq[(1 + q + q^2) + (1 + 2q)p + p^2]$	$\sqrt{pq}(1 + q + q^2)$	1
3	0	pq	$\sqrt{pq}(1 + q + p)$	1	0
2	0	\sqrt{pq}	1	0	0
1	0	1	0	0	0
0	1	0	0	0	0

Table 12: p, q -Stirling Numbers of the Second Kind, $S_{pq}(n, k)$ —Region 3

$n \setminus k$	-4	-3	-2	-1
-1	$[p^{-3} + 2q^{-1}p^{-2} + 2q^{-2}p^{-1} + q^{-3}]/\sqrt{pq^3}$	$[p^{-1} + q^{-1}]/pq$	$1/\sqrt{pq}$	1
-2	$[p^{-3} + (2q^{-1} + 1)p^{-2} + (2q^{-2} + q^{-1} + 1)p^{-1} + (q^{-3} + q^{-2} + q^{-1})]/pq$	$[p^{-1} + (q^{-1} + 1)]/\sqrt{pq}$	1	0
-3	$[p^{-2} + (q^{-1} + 1)p^{-1} + (q^{-2} + q^{-1} + 1)]/\sqrt{pq}$	1	0	0
-4	1	0	0	0

The p, q -Stirling number obey the following properties:

$$\begin{aligned}
 s_{pq}(n, k) &= \text{comp}_{n-k}(\{[0]_{pq} \dots [n-1]_{pq}\}) \\
 &= \begin{cases} (-1)^{n-k} e_{n-k}(\{[1]_{pq}, [2]_{pq}, \dots, [n-1]_{pq}\}) & \text{for } n \geq k \geq 0, \\ h_{n-k}(\{[-1]_{pq}, [-2]_{pq}, \dots, [n]_{pq}\}) & \text{for } 0 > n \geq k. \end{cases} \tag{12}
 \end{aligned}$$

$$\begin{aligned}
 S_{pq}(n, k) &= \text{comp}_{n-k}(-\{[0]_{pq} \dots [k]_{pq}\}) \\
 &= \begin{cases} h_{n-k}(\{[1]_{pq}, [2]_{pq}, \dots, [k]_{pq}\}) & \text{for } n \geq k \geq 0, \\ (-1)^k e_{n-k}(\{[-1]_{pq}, [-2]_{pq}, \dots, [k+1]_{pq}\}) & \text{for } 0 > n \geq k. \end{cases} \tag{13}
 \end{aligned}$$

$$\begin{aligned}
 s_{pq}(n, k) &= s_{pq}(n-1, k-1) - [n]_{pq} s_{pq}(n-1, k) \\
 S_{pq}(n, k) &= S_{pq}(n-1, k-1) + [k]_{pq} S_{pq}(n-1, k).
 \end{aligned}$$

By definition, the p, q -Stirling numbers transform into the classical Stirling numbers by setting p and q equal to one. If we merely set p equal to one, we have the usual q -Stirling numbers $s_q(n, k)$ (see tables

Table 13: q -Stirling Numbers of the First Kind, $s_q(n, k)$ —Region 1

$n \backslash k$	0	1	2	3	4
5	0	$q^2 + 3q^3 + 5q^4 + 6q^5$	$-[4q^{3/2} + 9q^{5/2} + 12q^{7/2} + 12q^{9/2} + 8q^{11/2} + 4q^{13/2} + q^{15/2}]$	$6q + 9q^2 + 9q^3 + 7q^4 + 3q^5 + q^6$	$-[4q^{1/2} - 3q^{3/2} + 2q^{5/2} + q^{7/2}]$
4	0	$-[q^{3/2} + 2q^{5/2} + 2q^{7/2} + q^{9/2}]$	$3q + 4q^2 + 3q^3 + q^4$	$-[3q^{1/2} + 2q^{3/2} + q^{5/2}]$	1
3	0	$q + q^2$	$-[2q^{1/2} + q^{3/2}]$	1	0
2	0	$-q^{1/2}$	1	0	0
1	0	1	0	0	0
0	1	0	0	0	0

Table 14: q -Stirling Numbers of the First Kind, $s_q(n, k)$ —Region 3

$n \backslash k$	-5	-4	-3	-2	-1
-1	q^{-2}	$-\sqrt{q}^{-5}$	q^{-1}	$-\sqrt{q}^{-1}$	1
-2	$-[q^{-9/2} + 4q^{-7/2} + 6q^{-5/2} + 4q^{-3/2}]$	$q^{-3} + 3q^{-2} + 3q^{-1}$	$-[q^{-3/2} - 2q^{-1/2}]$	1	0
-3	$q^{-5} + 3q^{-4} + 7q^{-3} + 8q^{-2} + 6q^{-1}$	$-[4q^{-5/2} + 2q^{-3/2} + 3q^{-1/2}]$	1	0	0
-4	$-[q^{-7/2} - 2q^{-5/2} - 3q^{-3/2} - 4q^{-1/2}]$	1	0	0	0
-5	1	0	0	0	0

13 and 14) and $S_q(n, k)$ (see tables 15 and 16) defined as the coefficients connecting the powers of x with the q -lower factorial $(x; q)_n$ where the q -lower factorials $((x; q)_n)_{n \in \mathbb{Z}}$ form the sequence of rational functions with persistent roots given by the q -numbers $\sqrt{q}(q^i - 1)/(q - 1)$.

Note also that the p, q -numbers are invariant under the interchange of p with q . Thus, the p, q -Stirling numbers are symmetric in p and q .

This interpretation will be made in terms of Ferrers diagrams.

Table 15: q -Stirling Numbers of the Second Kind, $S_q(n, k)$ —Region 1

$n \setminus k$	0	1	2	3	4
5	0	q^2	$4q^{3/2} + 6q^{5/2}$	$6q + 8q^2 + 7q^3$	$4q^{1/2} + 3q^{3/2}$
4	0	$q^{3/2}$	$4q^{7/2} + q^{9/2}$	$+3q^4 + q^5$	$+2q^{5/2} + q^{7/2}$
3	0	q	$2q^{1/2} + q^{3/2}$	1	0
2	0	$q^{1/2}$	1	0	0
1	0	1	0	0	0
0	1	0	0	0	0

Table 16: q -Stirling Numbers of the Second Kind, $S_q(n, k)$ —Region 3

$n \setminus k$	-5	-4	-3	-2	-1
-1	$q^{-8} + 3q^{-7} + 5q^{-6} + 6q^{-5}$	$q^{-9/2} + 2q^{-7/2}$	$q^{-2} + q^{-1}$	$q^{-1/2}$	1
-2	$+5q^{-4} + 3q^{-3} + q^{-2}$	$+2q^{-5/2} + q^{-3/2}$	$q^{-3/2}$	1	0
-3	$q^{-15/2} + 4q^{-13/2} + 8q^{-11/2} + 12q^{-9/2}$	$q^{-4} + 3q^{-3}$	$+4q^{-2} + 3q^{-1}$	$+q^{-1/2}$	0
-4	$+12q^{-7/2} + 9q^{-5/2} + 4q^{-3/2}$	$q^{-5/2} + 2q^{-3/2}$	1	0	0
-5	$q^{-6} + 3q^{-5} + 7q^{-4}$	$+3q^{-1/2}$	0	0	0
-6	$+9q^{-3} + 9q^{-2} + 6q^{-1}$	1	0	0	0
-7	$q^{-7/2} + 2q^{-5/2} + 3q^{-3/2} + 4q^{-1/2}$	0	0	0	0
-8	1	0	0	0	0

Definition 34 (Ferrers Diagram) The Ferrers diagram of a n -partition or d -partition λ with length k is defined as the following hybrid set of squares (i, j) :

$$\mathcal{H}(\lambda) = \sum_{i=1}^k \{(1, 1)..(1, \lambda_i)\}.$$

The cardinality of $\mathcal{H}(\lambda)$ is $|\lambda|$.

In the case of classical partitions, the above corresponds to the traditional concept of a Ferrers diagram. For example, the Ferrers diagram of $(3, 2, 0)$ is the positive set $\{(1, 1), (1, 2), (1, 3), (2, 1), (2, 2)\}$. On the other hand, the Ferrers diagram of $(-3, -2)$ is the negative set $\{(1, 0), (1, -1), (1, -2), (2, 0), (2, -1)\}$.

Definition 35 (0-1 Tableaux) A 0-1 p -tableau (resp. a 0-1 n -tableau) is a pair $\alpha = (\lambda, f)$ where λ is a p -partition (resp. n -partition), and f is a filling of that diagram. That is to say it is a collection of constants equal to 0 or 1 and denoted f_{ij} for each $(i, j) \in \mathcal{H}(\lambda)$ subject to the condition that there is exactly one 1 in each row. ie: for all $1 \leq i \leq \ell(\lambda)$, there is a unique j such that $f_{ij} = 1$.

The inversion number $inv(\alpha)$ of a 0-1 tableau $\alpha = (\lambda, f)$ is one half plus the number of zeros to the left of a one in the diagram, ie: the number of triples (i, j_1, j_2) such that $j_1 \ll j_2$, $f_{ij_1} = 0$, and $f_{ij_2} = 1$ while taking into account possible negative multiplicities. Similarly, the non-inversion number $nin(\alpha)$ of a 0-1 tableau is one half plus the number of zeros to the right of a one in the diagram.

We then have the following generalization of [17, eq. (1.4)]:

Theorem 36 Let n and k be integers. Then the Stirling numbers of the second kind of degree n and order k are given by sums

$$s_{pq}(n, k) = \sum_{\alpha} q^{\text{inv}(\alpha)} p^{n \text{in}(\alpha)}$$

$$s_q(n, k) = \sum_{\alpha} q^{\text{inv}(\alpha)}$$

over 0-1 tableaux α of d -partitions of width n and length $n - k$. Thus, $s(n, k)$ is the number of such tableaux.

The Stirling numbers of the first kind of degree n and order k are given by similar sums

$$S_{pq}(n, k) = \sum_{\alpha} q^{\text{inv}(\alpha)} p^{n \text{in}(\alpha)}$$

$$S_q(n, k) = \sum_{\alpha} q^{\text{inv}(\alpha)} p^{n \text{in}(\alpha)}$$

this time over 0-1 tableaux α of n -partitions of width k , and length $n - k$. Thus, $S(n, k)$ is the number of such tableaux.

Proof: Equation 12. \square

Corollary 37 Let n and k be integers. Then we have the following relations

$$S_{pq}(n, k) = (-1)^{n+k} s_{p-1, q-1}(-k, -n)$$

$$S_q(n, k) = (-1)^{n+k} s_{q-1}(-k, -n)$$

$$S(n, k) = (-1)^{n+k} s(-k, -n). \square$$

References

- [1] P. L. BUTZER, M. HAUSS AND M. SCHMIDT, *Factorial Functions and Stirling Numbers of Fractional Orders*, Results in Mathematics, **16**, 16-48 (1989).
- [2] S. CHARALAMBIDES AND J. SINGH, *A Review of Stirling Numbers: Their Generalization and Statistical Application*, Communications in Statistical Theory: Theory and Methods **17(A)** 2533-95 (1988).
- [3] L. COMTET, "Advanced Combinatorics," Reidel, 1974.
- [4] E. DAMIANI, O. D'ANTONA, AND D. LOEB, *The Complementary Symmetric Function: Connection Constants Using Negative Sets*, submitted to European Journal of Combinatorics.
- [5] O. D'ANTONA AND E. DAMIANI, *Connection Constants*, Rapport Interno no. 54/89, Universita di Milano, Dipartimento di Scienze dell'Informazione.
- [6] H. W. GOULD, *Note on Recurrence Relations for Stirling Numbers*, Publications de l'Institut Mathématique, Nouvelle série, tome 6 (20) 1966, 115-119.
- [7] H. W. GOULD, *The q -Stirling Numbers of the First and Second Kinds*, Duke Math. J. , **28** (1961) 281-289.
- [8] J. G. HAGEN, *Synopsis der hoheren Mathematik*, Vol. 1, Berlin, 1891.
- [9] D. KNUTH, *Subspaces, Subsets, and Partitions*, Journal of Combinatorial Theory **10** (1971) 178-180.

- [10] P. LEROUX, *Reduced Matrices and q -Log-Concavity Properties of q -Stirling Numbers*, *Journal of Combinatorial Theory A*, **54** (1991) 64–84.
- [11] D. LOEB, *The Iterated Logarithmic Algebra*, *Advances in Mathematics*, **86** (1991) 155–234.
- [12] D. LOEB, *A Generalization of the Binomial Coefficients*, *SIAM Journal of Discrete Mathematics*, To appear.
- [13] D. LOEB, *A Generalization of the Stirling Numbers*, *SIAM Journal of Discrete Mathematics*, To appear.
- [14] D. LOEB, *Sets with A Negative Number of Elements*, *Advances in Mathematics*, **91** (1992) 64–74.
- [15] D. LOEB AND G.-C. ROTA, *Formal Power Series of Logarithmic Type*, *Advances in Mathematics*, **75** (1989), 1–118.
- [16] I. G. MACDONALD, “Symmetric Functions and Hall Polynomials,” *Oxford Mathematical Monographs*, Clarendon Press, Oxford, 1979.
- [17] A. MÉDICIS AND P. LEROUX, *A Unified Combinatorial Approach for q - (and p, q -)Stirling Numbers*, Rapport de recherche no. 126, Département de mathématiques et d’informatique, Université du Québec à Montréal.
- [18] F. REGONATI, *Sui numeri di Whitney di certi reticoli supersolibi finiti*.
- [19] S. ROMAN, *A Generalization of the Binomial Coefficients*, To Appear.

P -Partitions and the Plactic Congruence

Claudia Malvenuto¹

Département de Mathématiques et Informatique - Université du Québec à Montréal
Case Postale 8888, Succ. A - Montréal - Québec - H3C 3P8

Abstract

A conjecture of Stanley states that if the generating function of a poset P is symmetric, then in fact P belongs to the family of posets induced by some skew shape λ/μ . In this paper we show that if the set $L(P)$ of the linear extensions of a poset P is plactic-closed, then P is a poset induced by a skew shape.

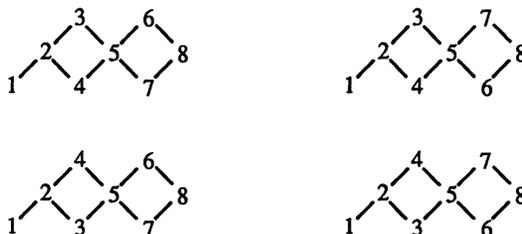
Let P be an alphabet consisting of n letters, with a total order \leq , usually thought of as $[n]$, and let \leq_P be a partial order on P . As usual P^* indicates the free monoid on the alphabet P . Let $L(P)$ be the set of words of P^* which are a linear refinement (or linear extension) of P : that is, $w \in L(P)$ if

- w is a **standard** word of length n (i.e. in w all the letters of P occur without repetition);

- whenever $x \leq_P y$ then x appears on the left of y in $w = \dots x \dots y \dots$

Given a skew shape λ/μ of weight $|\lambda/\mu| = n$ and a totally ordered alphabet P , a family $\mathbf{P}_{\lambda/\mu}$ of posets can be defined by filling the skew shape λ/μ with the letters of P in increasing order (with respect to \leq) on the rows, and in decreasing order on the columns, and then by rotating it 45° counterclockwise.

For example, if $P = [8]$ with the natural total order, and $\lambda/\mu = 443/21$ then $\mathbf{P}_{\lambda/\mu}$ contains the following posets:



For a standard word w on P define its composition $\mathcal{C}(w)$ as the composition

$$(|u_1|, \dots, |u_k|)$$

¹Supported by La.C.I.M.-UQAM and Consiglio Nazionale delle Ricerche.

where $w = u_1 \dots u_k$, each u_i is an increasing word and k is minimal. For a composition $C = (c_1, \dots, c_k)$ of n , define the function of infinitely many totally ordered variables:

$$F_c = \sum x_1 x_2 \dots x_n$$

where the sum is extended to all monomials satisfying the inequalities

$$x_1 \leq \dots \leq x_{c_1} < x_{c_1+1} \leq \dots \leq x_{c_1+c_2} < x_{c_1+c_2+1} \leq \dots \leq x_n,$$

the strict inequality being required between the variables $x_{c_1+\dots+c_i}$ and $x_{c_1+\dots+c_{i+1}}$, $i = 1, \dots, k - 1$. For example, if $w = 312$ then $u_1 = 3$, $u_2 = 12$, $C(w) = (1, 2)$ and $F_{(12)} = \sum_{x < y \leq z} xyz$.

To any subset L of permutations, we can associate the function

$$F_L = \sum_{w \in L} F_{C(w)}$$

and to the poset P , the function

$$\Gamma(P) = F_{L(P)} = \sum_{w \in L(P)} F_{C(w)}$$

Remark 1 This definition of $\Gamma(P)$ is equivalent to that of the generating function for the poset (P, \leq_P) given via P -partitions, and is defined as follows [G]: denote by $A(P)$ the set of all P -partitions, i.e. the functions $f : P \rightarrow \mathbf{N}$ such that, if $i \leq_P j$ then $f(i) \leq f(j)$, and if $i \leq_P j$ but $i > j$ (in the total order of P) then $f(i) < f(j)$. Then

$$\Gamma(P) = \sum_{f \in A(P)} x_{f(1)} x_{f(2)} \dots x_{f(m)}.$$

In [S] Stanley conjectures that if $\Gamma(P)$ is a symmetric function, then there exist λ, μ such that $P \in \mathbf{P}_{\lambda/\mu}$.

Recall [L-S] that the plactic congruence \sim on the set of words in P^* is defined by Knuth's relations [K]:

$$\begin{aligned} \text{if } x < y < z \text{ then } & yzx \sim yxz \\ & \text{and } zxy \sim xzy \\ \text{if } x < y \text{ then } & yxx \sim xyx \\ & \text{and } yyx \sim yxy \end{aligned}$$

That is, we interchange the order of two letters whenever they are preceded or followed by a letter that is in between.

The following facts motivate the consideration of the plactic congruence:

Fact 1 If $P \in \mathbf{P}_{\lambda/\mu}$, then $L(P)$ is plactic-closed.

This may be seen directly from the definitions, and is proved in [G-R].

Fact 2 *If L is a set of permutations which is plactic-closed, then F_L is a symmetric function.*

This result may be obtained by a standard application of the Robinson-Schensted correspondence.

Now, suppose that Stanley's Conjecture is true. If $L(P)$ is plactic closed, then $\Gamma(P)$ is a symmetric function by fact 2, and the Conjecture states that $P \in \mathbf{P}_{\lambda/\mu}$. In this paper we give a direct proof of this result:

Theorem 1 *If $L(P)$ is plactic-closed, then $P \in \mathbf{P}_{\lambda/\mu}$ for some skew diagram λ/μ , i.e. it is one of the orders induced by a skew diagram λ/μ .*

To prove this theorem, we will study the Hasse diagram of the poset locally, giving some restrictions to its subposets.

It is easy to see that every linear refinement of a subset Q of P appears as a subword of some linear refinement of P , but in general not as a factor. Recall that a subposet Q of a poset P is *convex* if, whenever $p \in P$ is such that $q_1 \leq_P p \leq_P q_2$, for some $q_1, q_2 \in Q$, then $p \in Q$.

In this particular case we can state the following:

Lemma 1 *If $Q \subseteq P$ is a convex subposet of a poset P then, for each $w \in L(Q)$, there exists $\bar{w} \in L(P)$ such that w is a factor of \bar{w} . That is, any linear extension of Q appears as a factor of some linear extension of P .*

Proof. By induction on $|P \setminus Q|$. Let A, B, C be the following subsets of P/Q

$$A = \{x \in P \setminus Q \mid \exists q \in Q \text{ such that } x \leq_P q\}$$

$$B = \{x \in P \setminus Q \mid \exists q \in Q \text{ such that } x \geq_P q\}$$

$$C = \{x \in P \setminus Q \mid x \text{ is not comparable to any element of } Q\}$$

Clearly $A \cup B \cup C = P \setminus Q$ and $A \cap C = B \cap C = \emptyset$. Furthermore we have also $A \cap B = \emptyset$: if not, let $y \in A \cap B$, then there exist $q \in A, q' \in B$ with $q' \leq_P y \leq_P q$, and the convexity of Q implies $y \in Q$, a contradiction. So A, B, C form a partition of $P \setminus Q$.

If $|P \setminus Q| = 0$, then $Q = P$ and the statement is trivially true.

Now suppose $|P \setminus Q| > 0$. (i) If $C \neq \emptyset$, let $c \in C$. Now $Q' = Q \cup \{c\}$ is still convex in P : indeed there is no element $p \in P, q \in Q$ such that $c \leq_P p \leq_P q$ or $q \leq_P p \leq_P c$, otherwise c would be comparable with Q . We have $cw \in L(Q')$; since $|P \setminus Q'| < |P \setminus Q|$, by induction, there exist u, v such that $\bar{w} = u(cw)v = (uc)wv \in L(P)$.

(ii) If $C = \emptyset$, and $A \neq \emptyset$, let $a \in A$ be maximal with respect to the partial order \leq_P . In this case $Q' = Q \cup \{a\}$ is convex in P : otherwise there exist $p \in P \setminus Q, q \in Q$ with $a \leq_P p \leq_P q$, so that p would be in A and $p \geq_P a$, contradicting the maximality of a . Furthermore we cannot have $q \leq_P p \leq_P a$, because by definition of A there exists $q' \in Q$ such that $a \leq_P q'$, so $q \leq_P a \leq_P q'$ and the convexity of Q says $a \in Q$. Now observe that

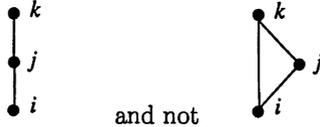
$aw \in L(Q')$: if not, there exists $q \in Q$ with $q \leq_P a$, meaning that $a \in A \cap B$, impossible. Since $|P \setminus Q'| < |P \setminus Q|$, by induction there exist u, v such that $\bar{w} = u(aw)v \in L(P)$.

(iii) Finally suppose $C = \emptyset$, $A = \emptyset$, and $B \neq \emptyset$, let b be a minimal element in B . A similar argument as above shows that $Q' = Q \cup \{b\}$ is convex in P and $wb \in L(Q')$, so that wb is a factor of some word of $L(P)$, and the lemma is proved. \square

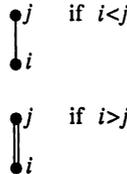
Corollary 1 *If Q is convex in P and $L(P)$ is plactic-closed, then $L(Q)$ is plactic-closed.*

Proof: Denote by $\pi : P^* \rightarrow Q^*$ the canonical projection of the set of words on the alphabet P onto the set of words on Q , which to any $u \in P^*$ associates the word $\pi(u)$ of Q^* obtained by erasing in u the letters of $P \setminus Q$. It is clear, by the definition of a linear refinement, that if $u \in L(P)$ then $\pi(u) \in L(Q)$. Let $w \in L(Q)$ and $w \sim w'$: we want to show that $w' \in L(Q)$. By Lemma 1, we can find $\bar{w} = uvw \in L(P)$. We have $\bar{w} \sim uw'v$ (as \sim is a congruence) and $uw'v \in L(P)$ since $L(P)$ is plactic-closed; in particular its projection $\pi(uw'v) = w'$ is in $L(Q)$. \square

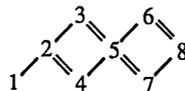
Denote by \prec_P the covering relation of the partial order \leq_P , that is $i \prec_P j$ if and only if $i \leq_P j$ and no element $k \in P$ satisfies $i \leq_P k \leq_P j$. Recall that the **Hasse diagram** of a poset P is the graph whose vertices are the elements of P and whose edges are the covering relations, such that if $i \leq_P j$ then j is drawn above i . For instance, if P is given by $i \leq_P j, j \leq_P k, i \leq_P k$, then its Hasse diagram is



We adopt the following convention: if $i \prec_P j$ we draw



So the Hasse diagram of the first poset in the example will look like

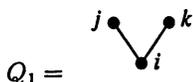


From now on we suppose that $L(P)$ is plactic-closed.

Lemma 2 *None of the following configurations appear in the Hasse diagram of P :*



Proof: First remark that all of the Q_i 's are convex subsets of P , because we are looking at the Hasse diagram. Now suppose that



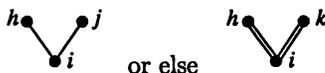
is in the Hasse diagram of P . This means $i \leq_P j$ with $i < j$ and $i \leq_P k$ with $i < k$. If $j < k$ [resp. $k < j$] we have $i < j < k$ [resp. $i < k < j$] so $ikj \sim kij$ [resp. $ijk \sim jik$] with $ikj \in L(Q_1)$, but $kij \notin L(Q_1)$ [resp. $ijk \in L(Q_1)$, but $jik \notin L(Q_1)$], a contradiction, as $L(Q_1)$ is plactic closed by Corollary 1. Similarly, one can prove that Q_2, Q_3 and Q_4 do not appear in P . \square

Lemma 3 For any $i \in P$, there can be at most two elements covering i and at most two elements which are covered by i .

Proof: Suppose that $j, k \succ_P i$; then by Lemma 2 we must have



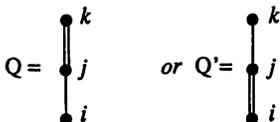
For any other $h \in P$ which covers i , we would have



as subset of P , against Lemma 2.

The same argument shows that i covers at most 2 elements. \square

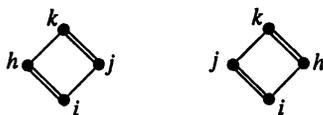
Lemma 4 If the subposets



appear in the Hasse diagram of P , then they are not convex in P .

Proof: Suppose that Q is convex in P , then $L(Q)$ is plactic-closed (Corollary 1). We have $i < j, j > k$. If $i < k$, then $i < k < j$ and $ijk \sim jik \in L(Q)$ [resp. $i > k, k < i < j, ijk \sim ikj \in L(Q)$], a contradiction, because $L(Q) = \{ijk\}$. The case of Q' can be proven in the same way, using Knuth's relations. \square

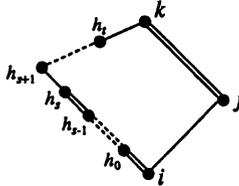
Corollary 2 In the hypothesis of Lemma 4 there exists a unique $h \neq j$ with $i \prec_P h \prec_P k$ and in P we have the following configurations (corresponding respectively to Q and Q')



Proof: Existence of h : if Q appears in P , by Lemma 4, there exists a chain

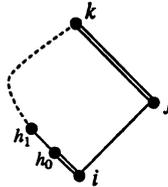
$$\underline{h} = (i \prec_P h = h_0 \prec_P h_1 \prec_P \dots \prec_P h_t \prec_P h_{t+1} = k)$$

from i to k different than $i \prec_P j \prec_P k$. Observe that $h_1 \neq j$, because if $i \prec_P h_0 \prec_P h_1 = j$ then j does not cover i . Note also that by virtue of Lemma 2, $i > h = h_0$ since $i < j$. We want to show that $t = 0$. For any such chain \underline{h} define $l(\underline{h})$ to be the integer $0 \leq s \leq t$ such that $h_0 > h_1 > \dots > h_s$ and $h_s < h_{s+1}$.

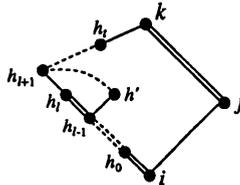


This number is well defined because we have $h_t < h_{t+1} = k$. Indeed, since $j > k = h_{t+1}$, we could not have also $h_t > h_{t+1} = k$, as this would lead to one of the configurations forbidden by Lemma 2.

Let \underline{h} be a chain such that $l(\underline{h}) = l$ is minimal: we claim that $l = 0$. If so, the subposet $i \prec_P h_0 \prec h_1$ satisfies the hypothesis of Lemma 4 and then it is not convex in P :



So there exists $j' \neq h_0$ with $i \prec_P j' \leq_P h_1$ and by Lemma 3, $j' = j$. This implies $j \leq_P h_1 \leq_P k$, but k covers j : we deduce $h_1 = k$ and $t = 0$, as we wanted. Assume $l > 0$. In this case the chain $h_{l-1} \prec_P h_l \prec_P h_{l+1}$, by Lemma 4, is not convex in P , because $h_{l-1} > h < h_{l+1}$: let $h' \in P$ be such that $h_{l-1} \prec_P h' \prec_P \dots h_{l+1}$. Again, since $h_{l-1} > h_l$, we have, by Lemma 2, $h_{l-1} < h'$.

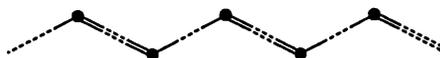


Now define a new chain \underline{h}' by replacing in \underline{h} the subchain $h_{l-1} \prec_P h_l \prec_P h_{l+1}$ by $h_{l-1} \prec_P h' \dots \prec_P h_{l+1}$. We have $l(\underline{h}') = l - 1 < l(\underline{h})$, against minimality of l .

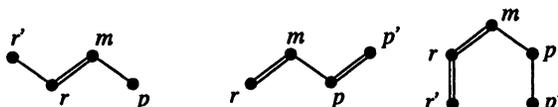
Uniqueness of h : any other element $h' \neq h$ with $i \prec_P h' \prec_P k$, would lead to a contradiction with Lemma 3. \square

Definition 1 A stair-like path is a sequence (p_1, p_2, \dots, p_t) of points of P , possibly empty, connected in the Hasse diagram and such that $p_1 < p_2 < \dots < p_t$.

Note that a stair-like path has the form



Lemma 5 *If one of the three following configurations are subsets of the Hasse diagram of P*



then we cannot find in P a stair-like path joining r' and p , r and p' , r' and p' respectively.

Proof: It is easily seen that $r' > p$ [resp. $r > p'$, $r' > p'$]. If a stair-like path exists, then we would have $r' < p$ [resp. $r < p'$, $r' < p'$], a contradiction. \square

We are now ready to reconstruct the poset P in the discrete plane $\mathbb{N} \times \mathbb{N}$ with its natural partial order

$$(x, y) \leq (x', y') \iff x \leq x' \text{ and } y \leq y'.$$

Then the covering relation \prec in $\mathbb{N} \times \mathbb{N}$ is given by

$$(x, y) \prec (x', y') \iff x' = x, y' = y + 1 \quad \text{or} \quad x' = x + 1, y' = y.$$

Definition 2 *We call a skew tableau of weight n any finite convex subposet T of $\mathbb{N} \times \mathbb{N}$, labelled with the elements of P , such that the labelling, with respect to the total order on P , is row-increasing reading from left to right and column-decreasing reading from the bottom to the top.*

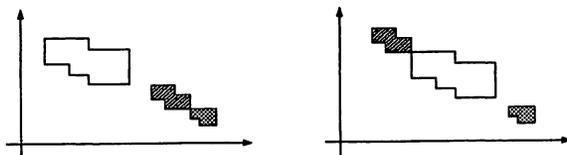
Definition 3 *In $\mathbb{N} \times \mathbb{N}$ a path from (x, y) to (x', y') is a sequence of points*

$$((x, y) = (x_1, y_1); (x_2, y_2); \dots; (x_k, y_k) = (x', y'))$$

such that, for $i = 1, \dots, k - 1$ either $(x_i, y_i) \prec (x_{i+1}, y_{i+1})$ or $(x_{i+1}, y_{i+1}) \prec (x_i, y_i)$.

Definition 4 *Let T be a subposet of $\mathbb{N} \times \mathbb{N}$. Two points (x, y) and (x', y') are connected in T if there exists a path joining them whose elements are all in T .*

Remark 2 In a skew tableau T two comparable points (x, y) and (x', y') are connected in T : indeed any saturated chain between (x, y) and (x', y') is a path whose points are all in T , by convexity of T . **Remark 3** In the discrete plane, "shifting" the connected components of a skew tableau T does not change the partial order given by T .



So the Theorem will take the following form: if $L(P)$ is plactic-closed, then there exists an embedding $e : P \rightarrow \mathbb{N} \times \mathbb{N}$ which respects the order \leq_P and such that $e(P)$, with the labelling ω defined as $\omega(e(p)) = p$, is a skew tableau.

Proof of Theorem 1: By induction on $|P|$. If $|P| \leq 2$, then the statement is true as we can have the following possibilities for P

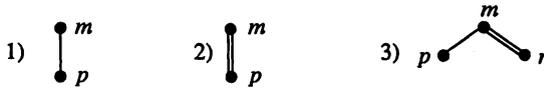


which correspond respectively to the diagrams of shape (1), (2) and (1 1).

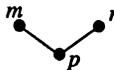
Now suppose $|P| \geq 3$, and let m be a maximal element of (P, \leq_P) . It is easily seen that $Q = P \setminus \{m\}$ is still a convex subposet of P , hence (by Corollary 1) plactic-closed. Then by the induction hypothesis, there is an embedding $e : Q \rightarrow \mathbb{N} \times \mathbb{N}$ and $e(Q)$ is a skew tableau, with the labelling ω defined by: $\omega(e(q)) = q$.

We wish to extend e "correctly" to all of P , i.e. to glue m to the skew shape given by $e(Q)$ in a way such that the result is still a skew shape.

As shown in Lemma 3, m can cover at most two elements, so that there are three possible cases in P :

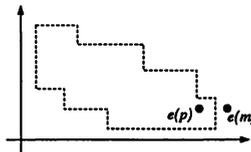


Case 1): Observe that there is nothing at the right of and on the same line as $e(p)$ in $e(Q)$, otherwise there exists $r \neq m$ such that



is a subposet of P which contradicts Lemma 2. Hence we can extend e to P by

$$e(m) = e(p) + (1, 0) = (x_p + 1, y_p)$$



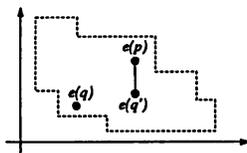
where we write (x_p, y_p) for the coordinates of $e(p)$ in $\mathbb{N} \times \mathbb{N}$. Note that e respects the covering relations in P , hence e respects \leq_P . Thus to show that $e(P)$ is a skew tableau, we need only to verify the convexity.

Claim: for any $q \in Q$ such that $e(q)$ is in the same connected component as $e(p)$, the inequality $y_q \geq y_p$ holds: if not, there exists an element $e(q) = (x_q, y_q)$ with $y_q < y_p$, that is $e(q)$ is below $e(p)$. We can always choose $e(q) = (x_p, y_p - 1)$, in which case the diagram

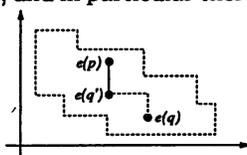


is a subsubset of P , and by Corollary 2 there would be another point covered by m , against assumption. Indeed, if $e(q)$ is to the left of $e(p)$, i.e. $x_q < x_p$, then

$$e(q) \leq (x_p, y_p - 1) \leq e(p)$$



and by convexity of $e(Q)$ there exists $q' \in Q$ with $e(q') = (x_p, y_p - 1)$. If $e(q)$ is to the right of $e(p)$, i.e. $x_q > x_p$, and since $e(p)$ and $e(q)$ are in the same connected component, there exists a path in $e(Q)$ joining them, and in particular there is $q' \in Q$ with $e(q') = (x_p, y_p - 1)$.



In fact, we can assume that the inequality $y_q \geq y_p$ holds as well when $e(q)$ is not connected to $e(p)$: we can define another embedding e' , shifting in the plane the connected component of $e(q)$ above that of $e(p)$ (see Remark 3), and the induction hypothesis on $e'(Q)$ will still apply.

Suppose now that $e(P)$ is not convex: then there exists a point $(x, y) \notin e(P)$ and $q \in P$ with $e(q) \leq (x, y) \leq e(m)$ (the case $e(m) \leq (x, y) \leq e(q)$ has not to be considered, as it contradicts the maximality of m). Remark that (x, y) cannot be comparable with $e(p)$: indeed, if $e(p) < (x, y)$, then

$$e(p) < (x, y) \leq e(m),$$

but in $\mathbb{N} \times \mathbb{N}$, $e(p)$ is covered by $e(m)$; if $(x, y) < e(p)$, then

$$e(q) \leq (x, y) < e(p)$$

against the convexity of $e(Q)$, which was assumed by induction.

Now we have $y_q \leq y \leq y_m = y_p$ and also, by the claim, we have the inequality $y_q \geq y_p$: this implies $y_q = y_p = y$, which means that (x, y) is comparable with (x_p, y_p) , a contradiction.

Case 2) Analogous to case 1.

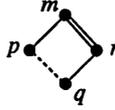
Case 3) Two possibilities have to be considered:

(i) p and r are not connected in Q : by induction we have an embedding $e : Q \rightarrow \mathbb{N} \times \mathbb{N}$ such that $e(Q)$ is a skew tableau. We aim to take another embedding e' of Q , shifting in the plane the components containing $e(p)$ and $e(r)$ such that

$$e'(p) = (x_r - 1, y_r + 1).$$

In order to be able to do so, we have to show that there is nothing to the left and over $e(r)$ in $e(Q)$, i.e. $e(r)$ is in the upper left corner of its connected component in $e(Q)$. Symmetrically, we can show that there is nothing to the right and below $e(p)$ in $e(Q)$,

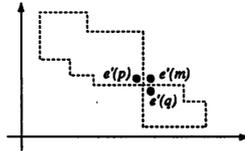
i.e. $e(p)$ is at the bottom right corner of its connected component. Indeed, if there exists $q \in Q$ with $e(q)$ to the left of $e(r)$, then back in P we would have the following



as a subposet of P , so p and r would be connected in Q , by Corollary 2. If there is $q \in Q$ with $e(q)$ over $e(r)$, then in P we would have the subposet

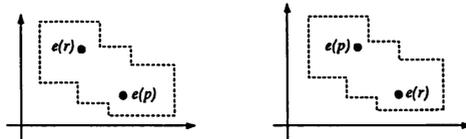


against Lemma 2. So we can extend e' to P by $e'(m) = (x_r, y_p)$



and $e'(P)$ is again a skew tableau.

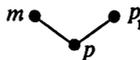
(ii) p and r are connected in Q : as p and r are not comparable, in $\mathbb{N} \times \mathbb{N}$ we can only have one of the following embeddings:



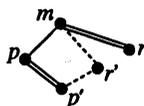
In the discrete plane, as $e(p)$ and $e(r)$ are contained in the same connected component of a skew tableau, we can always choose a stair-like path $(e(r), e(p_1), e(p_2), \dots, e(p_t), e(p))$, $t \geq 1$, joining them.

The first situation is impossible: if the path starts with a horizontal step, i.e. if $e(p_1) = (x_r + 1, y_r)$, then in P we would have the first subposet of Lemma 5 with the stair-like path (p_2, \dots, p_t) (possibly empty) joining p_1 to p , a contradiction. Analogously, if the last step of the path is vertical, we would have the second configuration of Lemma 5. So the only possible path joining $e(r)$ to $e(p)$ has to start with $e(p_1) = (x_r, y_{r-1})$ and end with $e(p_t) = (x_{p-1}, y_p)$ (with possibly $p_1 = p_t$) but this, back in P , leads to the third subposet of Lemma 5, with the stair-like path (p_2, \dots, p_{t-1}) joining p_1 to p_t : again a contradiction.

Let us analyze now the second situation. If $e(p_1) = (x_p + 1, y_p)$, then in P we have



contradicting Lemma 2. So the path starts with the vertical step $e(p_1) = (x_p, y_{p-1})$ and in P we have the subposet



In this case, by Corollary 2, there exists $r' \neq p$ with $p_1 \prec_P r' \prec_P m$, then $r' = r$ (otherwise m covers three elements) and this means in particular that the path chosen has length 1.

Hence we can extend e to all of P by $e(m) = (x_r, y_p)$ and $e(P)$ is still a skew tableau. \square

Acknowledgement: I would like to thank Professor C. Reutenauer, who conjectured Theorem 1, for his useful comments and suggestions.

References

- [G] I. Gessel, *Multipartite P -partitions and inner products of skew Schur functions* - Contemporary Mathematics, Vol. 34 (1984)
- [G-R] A. M. Garsia, J. B. Remmel, *Shuffles of permutations and the Kronecker product* - Graphs and Combinatorics, Vol. 1, n. 3, 217-263 (1985)
- [K] D. E. Knuth, *Permutation matrices and generalized Young tableaux* - Pac. J. Math. N. 34, 709-727 (1970)
- [L-S] A. Lascoux, M.- P. Schützenberger, *Le monoïde plaxique* - "Non commutative structures in algebra and geometric combinatorics", Quaderni de "La ricerca scientifica" C.N.R., n. 109 (1981)
- [M] I. G. Macdonald, *Symmetric Functions and Hall Polynomials* - Clarendon Press, Oxford (1979)
- [S] R. Stanley, *Ordered structures and partitions* - Memoirs of the Am. Math. Soc., n. 119 (1972)

Cycle Index Series Relations for Counting 2- and 3-Connected Graphs

R. W. Robinson, *Computer Science Dept., University of Georgia, Athens, GA 30602, USA.*

T. R. Walsh, *Département de Mathématiques et d'Informatique, Université du Québec à Montréal, Montréal, Québec, H3C3P8, Canada*

Abstract. Cycle index series relations for counting 2-connected graphs were established by the first author, and for 3-connected and homeomorphically irreducible 2-connected graphs by the second author. The 2-connected case was reworked in terms of species of structures by Pierre Leroux; here we make an attempt to do the same for the other two cases. An algebraic method of inverting these relations was established by the first author and applied to 2-connected graphs; and we have jointly applied this method to the other two cases. This method is much more efficient than the iterative method used previously, so that these graphs could be counted with up to 18 vertices instead of only 9. An article deriving these results has been accepted for publication; here we present the 2-connected case for quick accessibility.

1. Introduction

Counting graphs with prescribed properties usually involves decomposing a graph into a core and components. Counting graphs up to isomorphism also involves keeping track of the number of automorphisms of the core with a given cycle decomposition by means of a cycle index [2, ch. 2.2]. In the case of counting all n -vertex graphs, the core is the complete graph on n vertices, in which each edge is replaced by a component with 2 vertices and 0 or 1 edge to obtain the graphs to be counted; in the case of counting connected graphs, the core is a set, in which each element is replaced by a component, which is a connected graph, to obtain an arbitrary graph. In both cases, the cycle structures of the automorphisms of the core are known. Thus one can compute the sum of the cycle indices of the automorphism groups of all possible cores (*cycle index series* for cores) and then use Polya's theorem, which relates this cycle index series and the counting series for components and compositions [2, ch. 4.1 and 4.2]. When counting graphs with up to n nodes on

a computer, one stores only counting series and sums once over a cycle index series; thus the space-complexity is polynomial in n , and the time-complexity is bounded by the number of terms in the cycle index series (which is $p'(n)$, the number of partitions of all the numbers up to n) multiplied by a polynomial in n .

In the case of counting 2-connected graphs, the core is a set of rooted 2-connected graphs joined at the root; every other node has a rooted connected graph attached to it, to obtain another rooted connected graph. The unknown is now the cycle index series for cores, from which the counting series can be obtained. However, this cycle index series cannot be obtained from the counting series for components and compositions; thus it was necessary to generalize Polya's theorem to relate not counting series but cycle index series for components and compositions to those for cores [6 and 2, ch. 8.6]. The enumeration of 3-connected graphs and homeomorphically irreducible 2-connected graphs also involve solving for the cycle index series for cores [11]. A straightforward solution of such a relation involves storing and computing with the coefficients of cycle index series, so that the space-complexity is bounded below by $p'(n)$ and the time-complexity by a power of $p'(n)$.

An algebraic method of extracting counting series from systems of equations involving cycle index series was presented in [1], where it was applied to counting 2-edge-connected graphs. The first author applied these methods to 2-connected graphs in 1978, and we have since applied them to 3-connected graphs and 2-connected graphs without vertices of degree ≤ 2 . Once the equations have been solved by hand, extracting the numbers by computer involves storing only counting series and summing $O(n)$ times over the coefficients of a cycle index series. The time- and space-complexity of such a computation is essentially the same as for Polya's theorem. This makes these methods so efficient that we were able to count these graphs with up to 18 vertices; with the iterative methods used in [6] and [11] we could go only as far as 9 vertices. These results are to appear in [7], and extensive tables of numbers are available in a technical report [8]; therefore, we give the details only for the simplest of these problems: the enumeration of two-connected graphs.

The problem of counting 2-connected graphs has been re-worked by Pierre Leroux [4,5] in terms of species of structures, introduced by André Joyal in [3], and the defining equations for the species of two-connected graphs (but not the method for inverting these equations, which involves virtual species) are reported here. For 3-connected graphs and homeomorphically irreducible 2-connected graphs, the problem of re-working the defining equations in terms of species is still open. Here we present preliminary work done in that direction by the second author, involving a new type of substitution of species in which a component structure replaces a core element which has automorphisms inducing automorphisms in the replacing structure - in this case a 2-pole network replacing an edge of a graph - in the hope that a more experienced taxonomist can generalize the theory of species of structures sufficiently to incorporate the concepts we present, and can rework both the derivation and the inversion of these equations in terms of this generalized theory.

2. Two-Connected Graphs

Since the 2-connected components of a connected graph are attached to each other by a single vertex, it is not necessary to consider the edge-cycles of the graph-automorphisms; it suffices to consider the vertex-cycles and the number of edges. Therefore we weight a cycle index in the variables x_1, x_2, \dots by $y^{\text{number of edges}}$, so that the coefficient of the term $x_1^{n_1} x_2^{n_2} \dots y^j$ is the total number of labelled j -edge graphs left fixed by the permutations of the vertices with n_1 cycles of length 1, n_2 cycles of length 2, ..., divided by $n!$, where n is the number of vertices.

Let \mathbf{G} , \mathbf{C} and \mathbf{B} be the species of graphs, connected graphs and 2-connected graphs, and g , c and b be the corresponding (weighted) cycle index series $Z(\mathbf{G})$, $Z(\mathbf{C})$ and $Z(\mathbf{B})$. Then g is given by

$$g = \sum_{n_1, n_2, \dots} \Lambda(n_1, n_2, \dots) \prod_i x_i^{n_i} / (n_i! i^{n_i}), \quad (1)$$

where the sum runs through all $n_1, n_2, \dots \geq 0$, the product runs through all i from 1 to $n = n_1 + n_2 + \dots$, and

$$\Lambda(n_1, n_2, \dots) = \prod_{i < j} (1 + y^{\text{LCM}(i,j)})^{\text{GCD}(i,j) n_i n_j} \prod_i (1 + y^i)^{(i(n_i(n_i-1)/2) + \text{floor}((i-1)/2)n_i + n_2 i)}.$$

The following equations were derived in [4,5]:

$$G = E \circ C; \quad (2)$$

$$C^* = X E \circ (B' \circ C^*); \quad (3)$$

$$C^* + B \circ C^* = C + C^* \circ B' \circ C^*. \quad (4)$$

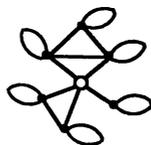
From these three equations, the following equations, originally derived in [6], follow automatically, using the observations that the species operators ' and \circ , which mean distinguishing a vertex and, in the case of ' , removing its label, correspond to the cycle index series operators $x_1(\partial/\partial x_1)$ and $\partial/\partial x_1$, and that the species substitution operator \circ corresponds to the cycle index series operator [] defined as follows: if a and d are cycle index series, then $a[d]$ means the cycle index series obtained by replacing each x_i in a by $d(x_i, x_{2i}, \dots, y^i)$.

$$g = \exp(\sum_i (1/i) x_i [c]); \quad (2z)$$

$$x_1(\partial c / \partial x_1) = x_1 \exp(\sum (1/i) x_i [(\partial b / \partial x_1) [x_1(\partial c / \partial x_1)]]); \quad (3z)$$

$$c = (x_1 + b - x_1(\partial b / \partial x_1)) [x_1(\partial c / \partial x_1)]. \quad (4z)$$

The reasoning behind equations (2) and (2z) is that any graph is a set of its connected components. For equations (3) and (3z), it is that any rooted connected graph can be decomposed into a core, consisting of a set of rooted 2-connected graphs joined by their roots, and components, each one a rooted connected graph attached by its root to one of the non-root vertices of the core (see the diagram below).



The following method was used in [6] for determining b . We know g from (1). We derive c from (2z) and differentiate to obtain $x_1(\partial c / \partial x_1)$, then we

solve (3z) to obtain $(\partial b/\partial x_1)[x_1(\partial c/\partial x_1)]$ and (4z) to obtain $b[x_1(\partial c/\partial x_1)]$. From this last cycle index series we extract $b(x_1, x_2, \dots; y)$ iteratively, and then replace each x_i by x^i to obtain the counting series for 2-connected graphs. This method involves storing and calculating with cycle index series, and since there are as many terms in the cycle index series for graphs with n vertices as there are partitions of n , it is so costly, both in terms of memory and execution time, that the tables in [6] go up only to 9 vertices.

The algebraic method of inverting cycle index series relations, introduced in [1] and used there to count bridgeless graphs, avoids storing and calculating with cycle index series by working directly with power series: the only cycle index series ever referred to is g , and its terms can be calculated from (1) whenever they are needed. We want to determine $b[x] = b(x, x^2, x^3, \dots; y)$, but b occurs in (4z) as $b[x(\partial c/\partial x_1)]$; so we introduce a new undetermined power series $\alpha(x, y)$, and we equate $x_1(\partial c/\partial x_1)[\alpha(x, y)]$ to x so that $b[x_1(\partial c/\partial x_1)[\alpha(x, y)]]$ becomes the desired $b[x]$.

If we replace each x_i in both sides of (3z) by the power series $\alpha(x^i, y^i)$, we obtain

$$x = \alpha(x, y) \exp(\sum (1/i)(\partial b/\partial x_1)[x^i]).$$

Taking logs and using Möbius inversion, we obtain

$$(\partial b/\partial x_1)[x] = -\sum (\mu(k)/k) \ln (\alpha(x^k, y^k)/x^k). \quad (5)$$

Replacing each x_i in (2z) by $\alpha(x_i, y_i)$ and then taking logs and using Möbius inversion, we obtain

$$c[\alpha(x, y)] = \sum (\mu(k)/k) \ln g[\alpha(x^k, y^k)]. \quad (6)$$

From (4z), with each x_i replaced by $\alpha(x^i, y^i)$, and from (5) and (6) we obtain

$$b[x] = -x + \sum (\mu(k)/k) \{ \ln g[\alpha(x^k, y^k)] - x \ln (\alpha(x^k, y^k)/x^k) \}. \quad (7)$$

From (2z) we obtain

$$x_1(\partial g/\partial x_1) = g * x_1(\partial c/\partial x_1).$$

Replacing x_i with $\alpha(x^i, y^i)$, we obtain

$$x_1(\partial g/\partial x_1)[\alpha(x, y)] = xg[\alpha(x, y)]. \quad (8)$$

We solve (8) successively for the coefficients of x, x^2, x^3, \dots in $\alpha(x, y)$ by running through the terms of (1). This yields successive coefficients of $g[\alpha(x, y)]$, the right side of (8), at the same time. Substituting these two power series into (7), we obtain $b[x]$.

We illustrate the method by finding $b[x]$ up to the coefficient of x^3 . Expanding (1) up to the terms of order 3, we find that

$$g = 1 + x_1 + (1/2)(x_1^2(1+y) + x_2(1+y)) + \\ + (1/6)(x_1^3(1+y)^3 + 3x_1x_2(1+y)(1+y^2) + 2x_3(1+y^3)) + \dots$$

Substituting into (8), the right side becomes

$$x + x\alpha(x, y) + (1/2)x\{\alpha^2(x, y)(1+y) + \alpha(x^2, y^2)(1+y)\} + \\ + (1/6)x\{\alpha^3(x, y)(1+3y+3y^2+y^3) + 3\alpha(x, y)\alpha(x^2, y^2)(1+y+y^2+y^3) + \\ + 2\alpha(x^3, y^3)(1+y^3)\} + \dots,$$

and the left side becomes

$$\alpha(x, y) + \alpha^2(x, y)(1+y) + \\ (1/2)\{\alpha^3(x, y)(1+3y+3y^2+y^3) + \alpha(x, y)\alpha(x^2, y^2)(1+y+y^2+y^3)\} + \dots$$

Equating coefficients of x^0 , we find that $\alpha(x, y) = 0 + \dots$

Equating coefficients of x^1 , we find that $\alpha(x, y) = x + \dots$ and that $g[\alpha(x, y)] = 1 + \dots$

Equating coefficients of x^2 , we find that $\alpha(x, y) = x + (-y)x^2 + \dots$ and that $g[\alpha(x, y)] = 1 + x + \dots$

Equating coefficients of x^3 , we find that $\alpha(x,y) = x + (-y)x^2 + (-y^3)x^3 + \dots$ and that $g[\alpha(x,y)] = 1+x+x^2\dots$

Equating coefficients of x^4 , we find that $g[\alpha(x,y)] = 1+x+x^2+x^3 + \dots$

We take formal logarithms using formula (1.2.8) of [2]: if $f(x) = 1+A_1x+A_2x^2+\dots$ and $\log f(x) = a_1x+a_2x^2+\dots$, then

$$a_n = A_n - (1/n)\sum_{m=1,\dots,n-1} m a_m A_{n-m}.$$

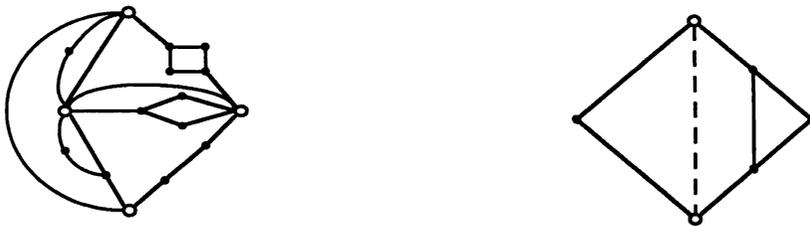
Then (7) becomes $b[x] = x^2y+x^3y^3 + \dots$, which enumerates the two smallest two-connected graphs: the link and the triangle.

To carry these calculations up to n vertices requires $O(n)$ space and $O(n^3)p'(n)$ time, where $p'(n)$ is the number of partitions of all the numbers from 1 to n , if we are counting by number of vertices alone and using fixed precision. If we are counting by number of edges as well as vertices in fixed precision, which we did up to 18 vertices, or by number of vertices alone in multiple precision (assuming a straightforward multiplication algorithm), which we did up to 25 vertices, then the space complexity becomes $O(n^3)$ and the time-complexity $O(n^7)p'(n)$. If we are counting by number of edges and vertices in multiple precision, then the space complexity rises to $O(n^5)$ and the time complexity to $O(n^{11})p'(n)$, but in practice we didn't do this because 16-byte precision was sufficient to go up to 18 vertices and multiple precision became too costly for more vertices than 18.

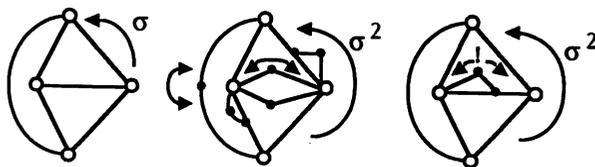
3. Three-Connected and Homeomorphically Irreducible Two-Connected Graphs

To handle these two cases we introduce another kind of graph decomposition: the components have two distinguished vertices rather than one, and each component replaces an edge of the core instead of being attached to a vertex.

To illustrate this decomposition, we restate the folkloric theorem stated and proved in [11] to count 2-connected homeomorphically irreducible graphs. Every 2-connected graph G can be expressed as a 2-connected graph G_r without vertices of degree ≤ 2 with each edge replaced by a series-parallel network. If G_r is not a single edge then the decomposition is unique; otherwise, distinguishing and deleting any edge $\{u,v\}$ of G and calling u and v the poles yields a series-parallel network. See the drawing below for an example.



An edge-cycle of length k of a vertex-permutation σ of the core, applied k times, fixes the poles of the networks that replace the edges in the cycle if σ^k preserves the orientation of the edges, and exchanges the poles if σ^k reverses the orientation of the edges; thus, in the latter case, the networks which replace the edges in the cycle must have automorphisms which reverse their poles or else σ cannot be an automorphism of the whole graph (see the diagram below, where σ^2 , which preserves 2 edges but reverses their orientation, is an automorphism of the middle graph since the networks that replace these edges have automorphisms that exchange their poles, but not of the rightmost graph, since the network which replaces the horizontal edge has no such automorphism).



This sort of decomposition was implicit in the enumeration of all n -vertex graphs given the cycle index of the complete graph, but in that case the edges of the complete graph could be treated as elements since the components that replaced them - an edge or no edge - are invariant under automorphisms of the core which preserve an edge but reverse its orientation. Thus the species

operator \circ , representing the substitution of a component structure for a core element, could be used, with an edge rather than a vertex representing an element. A general series-parallel network, on the other hand, may or may not have automorphisms that exchange its poles; so we have to invent a new species substitution operator to account for them.

In [11] we used and generalized an idea used in [2, p127] for counting orientations of a graph. We keep 2 families of variables for edge-cycles: for an edge-cycle of length k of a vertex-permutation σ , we use y_k if σ^k preserves the orientation of the edges and z_k otherwise. For example, in the leftmost graph in the figure above, the cycle type of σ is $x_4y_4z_1^2$. And we keep two cycle index series for networks: Z^+ for automorphisms which preserve the poles and Z^- for automorphisms which exchange them. For example, for the leftmost network in the figure below, $Z^+ = (x_1^2y_1^4 + x_2y_2^2)/2$ and $Z^- = (x_1^2y_2^2 + x_2y_2^2)/2$, (the poles do not count in the enumeration) and in each of the other two networks, $Z^+ = x_1^2y_1^4$ and $Z^- = 0$. In [11] we derived the defining relations for the cycle index series for 3-connected graphs and for homeomorphically irreducible 2-connected graphs directly; here we make a first attempt to express this derivation in terms of species.

A *two-pole network* is a graph in which two of its vertices have been distinguished as its source and sink. Motivated by the decomposition stated above, where the components are series-parallel networks, we define a new species operator \uparrow . If S is a species of graphs or networks and T is a *pole-symmetric* species of networks (if a network is in T then so is its mate obtained by exchanging the names of the poles), then $S\uparrow T$ is the species of graphs or networks obtained by replacing each edge of a member of S by a member of T . The symbol \uparrow is chosen instead of \circ because it looks like an oriented edge instead of a vertex. By the arguments used in [11], the cycle index series $Z(S\uparrow T)$ is obtained from $Z(S)$ by leaving each x_i alone, replacing each y_i by $Z^+(T)(x_i, x_{2i}, \dots; y_i, y_{2i}, \dots; z_i, z_{2i}, \dots)$ and each z_i by $Z^-(T)(x_i, x_{2i}, \dots; y_i, y_{2i}, \dots; z_i, z_{2i}, \dots)$. Motivated by the above-mentioned conversion of a graph into a series-parallel network, we define another new species operator \wedge . If S is a species of graphs, then $S\wedge$ is the (pole-symmetric) species of networks obtained by distinguishing, orienting and removing an edge and calling its incident vertices the poles, with the deleted edge oriented from source to sink. We choose the symbol \wedge

because the edge is deleted, leaving nothing but its orientation. In the diagram below, S consists of a single graph, and we show this graph and the 3 networks in S^\wedge ; the leftmost network is obtained by deleting the horizontal edge, and each of the other two by deleting one of the other edges after orienting it in one of two distinct ways.



By the arguments used in [11], we have

$$Z^+(S^\wedge) = (2/x_1^2)(\partial Z(S)/\partial y_1) \text{ and } Z^-(S^\wedge) = (2/x_2)(\partial Z(S)/\partial z_1).$$

Equations (1) and (2z)-(4z) can easily be expressed in terms of these three families of variables. Let K be the species of complete graphs. Then

$$Z(K) = \sum_{n_1, n_2, \dots} \Lambda(n_1, n_2, \dots) \prod_i x_i^{n_i} / (n_i! i^{n_i}),$$

where the sum runs through all $n_1, n_2, \dots \geq 0$, the product runs through all i from 1 to $n = n_1 + n_2 + \dots$, and

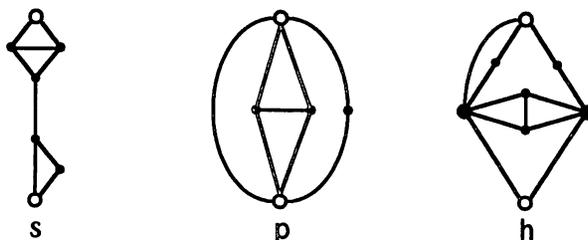
$$\Lambda(n_1, n_2, \dots) = \prod_{i < j} (y_{\text{LCM}(i,j)})^{\text{GCD}(i,j)n_i n_j} * \prod_i (y_i)^{(i(n_i(n_i-1)/2) + \text{floor}((i-1)/2)n_i)} * \prod_i (z_i)^{n_i 2i};$$

The z_i replaces the y_i of the well-known formula in [2] because it represents the diameters of even vertex-cycles. If Y represents the species consisting of a single edge, then $G = K \uparrow (\emptyset + Y)$; so that $Z(G)$ is obtained from $Z(K)$ by replacing each y_i with $1 + y_i$ and each z_i with $1 + z_i$. With this cycle index series replacing g in (1), equations (2z)-(4z) are still valid, except that now $a[d]$ means the cycle index series obtained by replacing each x_i in a by $d(x_i x_{2i} \dots; y_i y_{2i} \dots; z_i z_{2i} \dots)$.

We recall that (3z) was obtained by decomposing a rooted connected graph into a set of rooted connected graphs attached by their roots (the core) with a rooted connected graph (a component) attached by its root to every non-root vertex of the core. The following 2-pole version of that decomposition was

stated and proved in its present form (in Russian) by V. I. Trakhtenbrot [10] - a proof of this theorem in English and a derivation from it of one form of decomposition of a 2-connected graph into 3-connected components [9] was given in an unpublished version of [11] and is available from the second author.

Any two-pole network (possibly with parallel edges) which has ≥ 2 edges and which becomes 2-connected if an edge is added between its poles belongs to exactly one of the 3 following classes: *s*-networks, admitting a series decomposition, with a chain of length ≥ 2 as its core; *p*-networks, admitting a parallel decomposition, with a set of ≥ 2 parallel edges as its core; and *h*-networks, admitting a decomposition in which the core has ≥ 3 edges and is indecomposable in the sense that the components of any decomposition are all single edges. Further, these decompositions are unique under the assumption that none of the components of an *s*-network are themselves *s*-networks and that none of the components of a *p*-network are themselves *p*-networks (see the diagrams below).



We assume now that any network is without parallel edges and becomes 2-connected if an edge is added between its poles. Let \mathbf{B} and \mathbf{F} be the species of 2- and 3-connected graphs, respectively. Then \mathbf{B}^\wedge is the species of networks with non-adjacent poles, whence

$$Z^+(\mathbf{B}^\wedge) = (2/x_1^2)(\partial b/\partial y_1) \quad \text{and} \quad Z^-(\mathbf{B}^\wedge) = (2/x_2)(\partial b/\partial z_1). \quad (9)$$

The species \mathbf{D} of all networks with at least one edge is $(\emptyset + \mathbf{Y})\mathbf{B}^\wedge - \emptyset$, whence

$$d^+ = Z^+(\mathbf{D}) = (1+y_1)Z^+(\mathbf{B}^\wedge) - 1 \quad \text{and} \quad d^- = Z^-(\mathbf{D}) = (1+z_1)Z^-(\mathbf{B}^\wedge) - 1. \quad (10)$$

The species S of s -networks is given by $(\sum_{k \geq 2} \text{species of } k\text{-chain}) \uparrow (\mathbf{D}-S)$.
 For a k -chain we have

$$Z^+ = x_1^{k-1} y_1^k \text{ and } Z^- = \begin{cases} x_2^{(k-1)/2} y_2^{(k-1)/2} z_1 & \text{if } k \text{ is odd,} \\ x_1 x_2^{(k-2)/2} y_2^{k/2} z_1 & \text{if } k \text{ is even.} \end{cases}$$

It follows by the same arguments as in [11] that

$$s^+ = Z^+(S) = x_1 d^+(d^+ - s^+) \text{ and } s^- = Z^-(S) = (x_1 + x_2 d^-)(d^+_{(2)} - s^+_{(2)}), \tag{11}$$

where $d^+_{(2)}$ the cycle index series obtained from d^+ multiplying all the indices by 2 (*inflation* by 2).

Let P be the species of p -networks + Y (the one-edge network). Then $\mathbf{D}-P$ consists of all the non- p -networks with non-adjacent poles, so that \mathbf{B}^\wedge , the species of all networks with non-adjacent poles, is just the species of parallel unions of 0, 1, 2, ... of the networks in $\mathbf{D}-P$. We recast E as the species of 0 or more parallel edges, so that now $\mathbf{B}^\wedge = E \uparrow (\mathbf{D}-P)$. By the arguments used in [11] we have

$$Z^+(E) = \exp \sum (1/k) y_k \text{ and } Z^-(E) = \exp \{ \sum_{k \text{ odd}} (1/k) z_k + \sum_{k \text{ even}} (1/k) y_k \},$$

so that

$$Z^+(\mathbf{B}^\wedge) = \exp \sum (1/k) (d^+_{(k)} - p^+_{(k)}); \tag{12}$$

$$Z^-(\mathbf{B}^\wedge) = \exp \{ \sum_{k \text{ odd}} (1/k) (d^-_{(k)} - p^-_{(k)}) + \sum_{k \text{ even}} (1/k) (d^+_{(k)} - p^+_{(k)}) \}.$$

Since an indecomposable network with ≥ 3 edges is just a 3-connected graph with an edge distinguished, oriented and deleted and its incident vertices called poles, if H is the species of h -networks, then

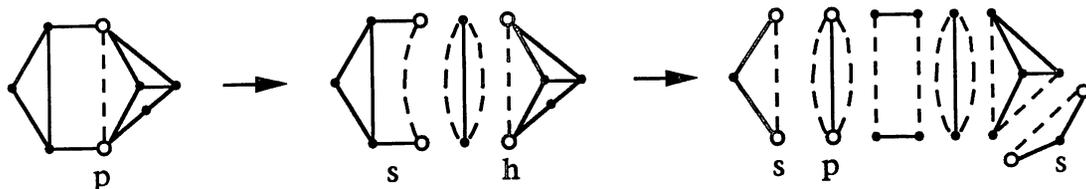
$$H = F^\wedge \uparrow D, \tag{13}$$

and by Trakhtenbrot's theorem we have

$$S+P+H=D.$$

(14)

The decomposition of a 2-connected graph G into 3-connected components is essentially the repeated decomposition of the network obtained by deleting any edge from G . When edges are added between the poles of all the cores, they become the 3-connected components of G : the chains become polygons, the sets of ≥ 2 parallel edges become sets of ≥ 3 parallel edges (bonds) and the indecomposable networks become 3-connected graphs. Two components are adjacent if in some network decomposition, one was a core before its edges were joined, and the other was the core of one of its components. This means, firstly, that in G two adjacent components are joined by identifying an edge in each and deleting the common edge (the edges added to turn networks into graphs), and secondly, by the uniqueness condition for network decomposition, that two bonds can never be adjacent and two polygons can never be adjacent (see the diagram below). This is the uniqueness condition stated in the decomposition used in [9].



Now these 3-connected components together with their adjacency relation form a tree, whose centre could be either a vertex (a component) or an edge (a virtual edge of G - one which was removed when the two central components were joined). Then B is obtained by rooting this tree at the centre.

If we root the tree at an arbitrary vertex, we fix some three-connected component of G . If it is a 3-connected graph, we have obtained the species $F \uparrow D$, from which we can extract F . If it is a polygon, the components cannot be s -networks; so we have (the species of polygons) $\uparrow (D-S)$. If it is a bond, then the components cannot be p -networks, and in addition there must be at least 3 components. The species $P-Y(\emptyset+H+S)$ fails on only two counts: its members have poles, and may have exactly 2 components (neither of which, however, is

a single edge). To eliminate the poles we introduce the operator $\bar{\cdot}$, which removes the special status of the poles of a network, converting it into a graph. For any species N of networks, we have

$$Z(N^{\bar{\cdot}}) = (1/2)(x_1^2 Z^+(N) + x_2 Z^-(N)).$$

The species of graphs which are the parallel unions of 2 components is $L\uparrow(H+S)$, where L consists of a two-vertex two-edge multigraph (a lune). So the species of 2-connected graphs rooted at some 3-connected component is

$$F_{15} = F\uparrow D + (\text{the species of polygons})\uparrow(D-S) + (P-Y(\emptyset+H+S))^{\bar{\cdot}} - L\uparrow(H+S). \quad (15)$$

If we root the tree at an arbitrary edge, we have a graph which is the parallel union of two networks, not both s-networks and not both p-networks. So the species of 2-connected graphs rooted at some virtual edge is

$$F_{16} = L\uparrow D - L\uparrow S - L\uparrow(P-Y). \quad (16)$$

The equation relating a species T of trees to the corresponding species T^{\bullet} of rooted trees is [4,5]

$$T^{\bullet} + E_2^{\circ} T^{\bullet} = T + (T^{\bullet})^2. \quad (17)$$

Therefore, $F_{15} + F_{16} = B + F_{15}^2$. It would be possible to convert this equation into cycle index series form and cancel; however, $(F_{15})^2$ is unnecessarily complicated. It is easy to show that $Z(E_2^{\circ} T^{\bullet}) - Z((T^{\bullet})^2) = -Z(E_2^{\circ} T^{\bullet}) + Z(T^{\bullet})_{(2)}$, and the analogue of $Z(T^{\bullet})_{(2)}$ is the cycle index series of a graph which is the parallel union of two isomorphic h-networks, restricted to those automorphisms which exchange the components. Thus we have

$$Z(F_{15}) - Z(F_{16}) = Z(B) - (1/2)(x_1^2 + x_2)Z^+(H)_{(2)}. \quad (18)$$

From (14) we get the equation for $Z(F\uparrow D)$ derived in [11] and stated (minus the typos) and inverted in [7] and [8].

Let I be the species of homeomorphically irreducible 2-connected graphs and R be the species of graphs which become series-parallel networks when an edge is selected, oriented and deleted and its incident vertices called poles. Then from the theorem quoted at the beginning of this section we have

$$B = I\uparrow(\text{species of series-parallel networks}) + R. \quad (19)$$

Now series-parallel networks are never h -networks, and the graphs in R are distinguished by having no 3-connected graphs among their 3-connected components. If we set F and H to zero in (10)-(16) and (18), then D becomes the species of series-parallel networks, S and P become those that are s -networks and p -networks, and B becomes R . With these modifications, we solve (10), (11) and (12) for D , S and P , and then (15), (16) and (18) for $Z(R)$, and then substitute in (19) to get $Z(I\uparrow D)$, as derived in [11] and inverted in [7] and [8].

REFERENCES

1. P. Hanlon and R. W. Robinson, Counting bridgeless graphs, *J. Combin. Theory Ser. B* **33** (1982), 276-305.
2. F. Harary and E. M. Palmer, "Graphical Enumeration", Academic Press, New York, 1973.
3. A. Joyal, Une théorie combinatoire des séries formelles, *Advances in Mathematics* **42** (1981), 1-82.
4. P. Leroux, Methoden der Anzahlbestimmung für einige Klassen von Graphen, *Beyreuther Mathematische Schriften* **26** (1988), 1-36.
5. P. Leroux and B. Miloudi, Généralisations de la formule d'Otter, research report 161, Université du Québec à Montréal, Sept. 1991, to appear in *Annales des Sciences Mathématiques du Québec*.
6. R. W. Robinson, Enumeration of non-separable graphs, *J. Combin. Theory Ser. B* **9** (1970), 327-356.
7. R. W. Robinson and T. R. Walsh, Inversion of Cycle Index Sum Relations for 2- and 3-Connected Graphs, accepted by *J. Combin. Theory Ser. B*.
8. R. W. Robinson and T. R. Walsh, *Inversion of Cycle Index Sum Relations for 2- and 3-Connected Graphs*, technical report UGA-CS-TR-91-008. Computer Science Dept., Univ. of Georgia, Athens, GA, 1991.
9. R. Tarjan and J. Hopcroft, Finding the triconnected components of a graph, *SIAM J. Comput.* **2** (1973), 136-158.
10. B. A. Trakhtenbrot, Towards a theory of non-repeating contact schemes, *Trudi Mat. Inst. Akad. Nauk SSSR* **51** (1958), 226-259. [Russian]
11. T. R. Walsh, Counting unlabelled three-connected and homeomorphically irreducible two-connected graphs, *J. Combin. Theory Ser. B* **32** (1982), 12-32.

**COUNTING A PAIR OF PERMUTATIONS AND
THE LINEARIZATION COEFFICIENTS
FOR JACOBI POLYNOMIALS**

by

Jiang ZENG

Abstract. — A theorem involving products of pairs of permutations recently published by Andrews, Goulden and Jackson [1] is proved by generalizing Gessel's combinatorial argument. A multiextension of this problem is also studied in connexion with the linearization coefficients of Jacobi polynomials.

1. Introduction. — Let \mathfrak{S}_n be the symmetric group of permutations of $[n] = \{1, 2, \dots, n\}$. For each $\sigma \in \mathfrak{S}_n$, we denote by *cyc* σ (resp. *fix* σ) the number of cycles (resp. fixed points) of σ and weight each pair (σ, τ) of permutations of $[n]$ by $w(\sigma, \tau) = \alpha^{\text{cyc } \sigma} \beta^{\text{cyc } \tau} p^{\text{fix } \sigma} q^{\text{fix } \tau} \gamma^{\text{fix } \sigma \tau^{-1}}$. The corresponding generating polynomial of $\mathfrak{S}_n \times \mathfrak{S}_n$ is then defined by :

$$(1.1) \quad L_n(\alpha, \beta, p, q, \gamma) = \sum_{\sigma, \tau \in \mathfrak{S}_n} w(\sigma, \tau).$$

Recently, by using their generalization of Cauchy's summation theorem for Schur functions and some group representation theories, ANDREWS, GOULDEN and JACKSON [1] found the following explicit formula for the exponential generating function of $L_n(\alpha, \beta, p, q, \gamma)$.

THEOREM 1. — *We have*

$$(1.2) \quad 1 + \sum_{n \geq 1} \frac{t^n}{n!} L_n(\alpha, \beta, p, q, \gamma) = e^{\alpha\beta t(2-p-q-\gamma+pq\gamma)} \\ \sum_{m \geq 0} \frac{t^m}{m!} \frac{(\alpha)_m (\beta)_m}{[1 + \beta(1-q)t]^{m+\alpha} [1 + \alpha(1-p)t]^{m+\beta} [1 + (1-\gamma)t]^{m+\alpha\beta}},$$

where and throughout this paper $(x)_n = x(x+1)\cdots(x+n-1)$.

It is remarkable that Theorem 1 includes several well-known results as its special cases. For instance, the Charlier polynomials $C_n(x; a)$ may be written as

$$a^n C_n(x; a) = \sum_{k=0}^n \binom{n}{k} (-x)_k a^{n-k} = \sum_{\sigma \in \mathfrak{S}_n} (-x)^{\text{cyc } \sigma} \left(1 - \frac{a}{x}\right)^{\text{fix } \sigma}.$$

Hence, setting $\alpha = -x, \beta = -y, \gamma = 1, p = 1 - a/x,$ and $q = 1 - b/y$ in (2), we obtain Meixner's bilinear generating function for Charlier polynomials [13] :

$$(1.3) \quad \sum_{n \geq 0} \frac{a^n b^n}{n!} t^n C_n(x; a) C_n(y; b) = e^{abt}(1 - at)^x(1 - bt)^y \sum_{m \geq 0} \frac{(-x)_m(-y)_m}{m!} \left[\frac{t}{(1 - at)(1 - bt)} \right]^m.$$

Another interesting example is counting three-line Latin rectangle. Recall that a reduced $3 \times n$ Latin rectangle may be identified with the pair (σ, τ) of permutations of $[n]$ subject to the restriction that for each $i \in [n], \sigma(i) \neq i, \tau(i) \neq i$ and $\sigma(i) \neq \tau(i)$, in other words, σ, τ and $\sigma\tau^{-1}$ are *derangements*. Then clearly, $L_n(1, 1, 0, 0, 0)$ gives the number of reduced $3 \times n$ Latin rectangles and its generating function is given by

$$(1.4) \quad e^{2t} \sum_{m \geq 0} \frac{m! t^m}{(1 + t)^{3(m+1)}},$$

which had been proved earlier by GOULDEN and JACKSON [11] and, in an equivalent form, also by BOGART and LONGYEAR [2].

In his combinatorial approach to the problem of counting three-line Latin rectangles, GESSEL [8] refined the above result by counting the number of cycles of σ and τ . More generally, he proved the $p = q = 0$ case of (1.2).

The first goal of this paper is to show (cf. Section 2) that Gessel's method can be generalized to establish Theorem 1. As mentioned by GESSEL [8], this approach, except the part of *umbral calculus*, is also similar to that taken by FOATA and others (cf. [4, 5]) in their combinatorial study of orthogonal polynomials.

In the next place, we will consider a multiextension of the above problem. Let $N = N_1 + N_2 + \dots + N_m$ be an ordered partition of a set N with $|N_i| = n_i$ ($i = 1, 2, \dots, m$). Let \mathfrak{S}_N be the set symmetric group of permutations of N . For each permutation π of N , an element x of N is said to be *c-fixed* by π , if both x and $\pi(x)$ belong to the same component N_i . Denote by $\text{Inc } \pi$ the set of all *c-fixed* elements of N by π . As a natural multiextension of (1.1), we define the generating function of $\mathfrak{S}_N \times \mathfrak{S}_N$ by

$$(1.5) \quad L_N(\alpha, \beta, p, q, \gamma) = \sum_{\sigma, \tau \in \mathfrak{S}_N} \alpha^{cyc \sigma} \beta^{cyc \tau} p^{|\text{Inc } \sigma|} q^{|\text{Inc } \tau|} \gamma^{|\text{Inc } \sigma\tau^{-1}|}.$$

Our consideration of the multiextension of the results of GESSEL and ANDREWS et al. was motivated by a combinatorial interpretation of the linearization coefficients

for Jacobi polynomials due to Foata and Zeilberger [7]. To be more precise, let $\mathbf{n} = (n_1, \dots, n_m)$, $|\mathbf{n}| = n_1 + \dots + n_m$ and $\mathbf{n}! = n_1! \dots n_m!$. By the linearization coefficients of Jacobi polynomials we mean the following integral of a product of Jacobi polynomials :

$$(1.6) \quad I_{\mathbf{n}}(\alpha, \beta) = \frac{\mathbf{n}!(\alpha + \beta + 2)_{|\mathbf{n}|} \Gamma(\alpha + \beta + 2)}{2^{\alpha + \beta + 1} \Gamma(\alpha + 1) \Gamma(\beta + 1)} \\ \times \int_{-1}^{+1} (1-x)^{\alpha} (1+x)^{\beta} \prod_{i=1}^m P_{n_i}^{(\alpha, \beta)}(x) dx,$$

where $P_{n_i}^{(\alpha, \beta)}(x)$ ($1 \leq i \leq m$) are the Jacobi polynomials :

$$P_n^{(\alpha, \beta)}(x) = \sum_{j=0}^n \binom{n+\alpha}{n-j} \binom{n+\beta}{j} \left(\frac{x-1}{2}\right)^j \left(\frac{x+1}{2}\right)^{n-j}.$$

One of the results of FOATA and ZEILBERGER [7] can be stated as follows :

PROPOSITION 1. — *We have*

$$I_{\mathbf{n}}(\alpha, \beta) = \sum (-1)^{|N \setminus K|} (\alpha + 1)^{\text{cyc } \sigma} (\beta + 1)^{\text{cyc } \tau},$$

where the summation is over all triples (σ, τ, K) with σ and τ permutations of N , and K a subset of N with the property that $K = \text{Inc } \sigma$ and $N \setminus K = \text{Inc } \tau$.

The second goal of this paper is to show (cf. Section 3) that the linearization coefficients for Jacobi polynomials is actually a limit case of $L_n(\alpha, \beta, p, q, \gamma)$ and to obtain some partial results of the evaluation of (1.5). We conclude this paper with some remarks and open problems in Section 4.

2. Proof of Theorem 1. — As usual, we identify each permutation σ of $[n]$ with the digraph on $[n]$ having an edge from i to $\sigma(i)$ for each $i \in [n]$. We define a P -configuration π on $[n]$ to be a digraph on $[n]$ with edges in three colors, red, blue, and green, with the following properties :

- i) The red (resp. blue) edges form a permutation σ (resp. τ) on $[n]$.
- ii) If there is a green edge from i to j , then there must be both a red and blue edge from i to j .

The weight of a P -configuration π is defined to be

$$v(\pi) = \alpha^{\text{cyc } \sigma} \beta^{\text{cyc } \tau} p^{\text{fix } \sigma} q^{\text{fix } \tau} \gamma^g,$$

where g is the number of green edges of π . Let $L_n^*(\alpha, \beta, \gamma, p, q)$ be the sum of weights of all P -configurations on $[n]$. It is then readily seen that

LEMMA 1. — We have

$$L_n^*(\alpha, \beta, \gamma, p, q) = L_n(\alpha, \beta, \gamma + 1, p, q).$$

Therefore, it suffices to evaluate the generating function of $L_n^*(\alpha, \beta, \gamma, p, q)$:

$$(2.1) \quad L^*(t) = \sum_{n \geq 0} L_n^*(\alpha, \beta, \gamma, p, q) \frac{t^n}{n!}.$$

First we may split the vertices of a P -configuration into two classes : those in green cycles and all others. A green cycle is coexistentive with a blue and a red cycle, and therefore can have no edges connecting it with any other vertices. The set of green cycles of a P -configuration constitutes itself a permutation on their vertices.

LEMMA 2. — Let $P_n(\alpha, p) = \sum_{\sigma \in \mathfrak{S}_n} \alpha^{cyc} \sigma p^{fix} \sigma$, then

$$(2.2) \quad \sum_{n \geq 0} P_n(\alpha, p) \frac{t^n}{n!} = e^{\alpha p t} \left(\frac{e^{-t}}{1-t} \right)^\alpha.$$

PROOF. — Define the weight of each permutation $\sigma \in \mathfrak{S}_n$ by $w(\sigma) = \alpha^{cyc} \sigma p^{fix} \sigma$. Clearly, this weight is multiplicative with respect to the cycle decomposition (cf. [3]) and each cycle of length 1 is weighted αp , while each cycle of length ≥ 2 is weighted α . Therefore, by the exponential formula (cf. [3]) we have

$$\sum_{n \geq 0} P_n(\alpha, p) \frac{t^n}{n!} = \exp\{\alpha p t + \sum_{n \geq 2} \alpha \frac{t^n}{n}\} = e^{\alpha p t} \left(\frac{e^{-t}}{1-t} \right)^\alpha. \quad \square$$

Therefore, the generating function for green cycles, together with their associated blue and red cycles, is $e^{\alpha \beta p q \gamma t} (e^{-\gamma t} / (1 - \gamma t))^{\alpha \beta}$, for every edge is weighted γ and every cycle is weighted $\alpha \beta$. Thus

$$(2.3) \quad L^*(t) = e^{\alpha \beta p q \gamma t} \left(\frac{e^{-\gamma t}}{1 - \gamma t} \right)^{\alpha \beta} R(t),$$

where $R(t)$ is the generating function for green-acyclic P -configurations. Hence it suffices to count the green-acyclic P -configurations.

Call a skew-green cycle any cycle whose all edges but one are green and a P^* -configuration on $[n]$ any pair (π, Ω) , where π is a P -configuration on $[n]$ and Ω a subset of the skew green cycles of π . Define the weight of a P^* -configuration (π, Ω) by

$$(2.4) \quad w(\pi, \Omega) = v(\pi) p^{|\nu|} q^{|\delta|} u^{|\bar{\nu}|} v^{|\bar{\delta}|},$$

where $|y|$ (resp. $|b|$) is the number of red (resp. blue) cycles in Ω , and $|\bar{y}|$ (resp. $|\bar{b}|$) the number of red (resp. blue) skew-green cycles not in Ω . Let $R_n^*(u, v)$ be the sum of the weights of all green-acyclic P^* -configurations on $[n]$.

LEMMA 3. — *The sum of the weights of all green-acyclic P^* -configurations on $[n]$ is equal to $R_n^*(1-p, 1-q)$.*

PROOF. — According to the definition, we have

$$\begin{aligned} R_n^*(1-p, 1-q) &= \sum_{(\pi, \Omega)} v(\pi) p^{|y|} q^{|b|} (1-p)^{|\bar{y}|} (1-q)^{|\bar{b}|} \\ &= \sum_{\pi} v(\pi) \sum_{\Omega} p^{|y|} (1-p)^{|\bar{y}|} \sum_{\Omega} q^{|b|} (1-q)^{|\bar{b}|}. \end{aligned}$$

By the binomial formula or the principle of inclusion-exclusion, the two interior summations are both equal to 1. Therefore, the lemma follows by the definition. \square

The problem turns out then to count the green-acyclic P^* -configurations with $u = 1-p$ and $v = 1-q$. In other words, we have

$$(2.5) \quad R(t) = \sum_{n \geq 0} R_n^*(1-p, 1-q) \frac{t^n}{n!}.$$

We shall count green-acyclic P^* -configurations by constructing them in two steps, with each step translating into a generating function operation. We first insert the green edges, obtaining a set of *green paths* and a set of *isolated vertices*. We then “mark” certain vertices with indeterminates. Finally, based on the marks, we insert the red and blue edges and form a set of skew-green cycles.

The following two observations are the basis for the derivation of our generating function.

- i) The green edges constitute a set of disjoint paths, each of which has at least two vertices.
- ii) If the green edges are contracted (together with their associated red and blue edges), what remains is a digraph consisting of a red permutation and a blue permutation. Any red or blue loop in the contracted graph must be attached to a vertex that was contracted.

Now we do the construction. We start with a set of green paths, each of at least two vertices, and a set of isolated vertices. We mark each isolated vertex with A and B . We mark the head of each path with either A or A_1 and B or B_1 . Note that the generating function of the configurations we have constructed so far is

$$(2.6) \quad \exp\left(\frac{(A + A_1)(B + B_1)\gamma t^2}{1 - \gamma t} + ABt\right),$$

where each green edge is weighted γ .

Next, we put in the red and blue edges and “mark” skew-green cycles as follows :

1. Construct a red permutation through all the vertices marked A and a blue permutation through all the vertices marked B .
2. For each vertex x which is the head of a green path, do the following : Let y be the tail of the path, if x is marked A , there is a red edge (z, x) , where z may be equal to x . Replace this edge with the red edge (z, y) . If x is marked A_1 , add the red edge (x, y) .
3. Repeat Step 2 with A, A_1 and “red” replaced by B, B_1 and “blue”.
4. Add a red edge and blue edge parrallel to every green edge.
5. Mark the skew-green cycles coming from a green path having a head marked A or B . These skew-green cycles constitute the subset Ω of the resulted P^* -configuration.

It is clear that only green-acyclic P^* -configurations are obtained, and each is obtained exactly once.

We now describe the operation on the generating function which corresponds to the above operations. The green-acyclic P^* -configurations coming from a term $A^i B^j A_1^k B_1^l \gamma^m t^n / n!$ should have total weight

$$P_i(\alpha, p) P_j(\beta, q) (\alpha u)^k (\beta v)^l \gamma^m t^n / n!$$

This is because each vertex marked A_1 yields a non marked red skew-green cycle, each vertex marked B_1 yields a non marked blue skew-green cycle, and the contribution from the vertices marked A and B is $P_i(\alpha, p) P_j(\beta, q)$. It is convenient to adopt the “symbolic” or “umbral” convention :

$$(2.7) \quad \begin{aligned} A^i &= P_i(\alpha, p), & A_1^k &= (\alpha u)^k, \\ B^j &= P_j(\beta, q), & B_1^l &= (\beta v)^l, \end{aligned}$$

by which we mean that henceforth after an expression is expanded, any occurence of A^i is to be replaced by $P_i(\alpha, p)$ and any occurence of B^j is to be replaced by $P_j(\beta, q)$, etc. As shown by ROTA [17], this means formally apply to all our formulas a linear operator

$$\Phi \left(A^i B^j A_1^k B_1^l \gamma^m \frac{t^n}{n!} \right) = P_i(\alpha, p) P_j(\beta, q) (\alpha u)^k (\beta v)^l \gamma^m \frac{t^n}{n!}.$$

Thus we need only evaluate (2.6) with this interpretation. Furthermore, by (2.5) we may suppose hereafter $u = 1 - p$ and $v = 1 - q$. As we will see in what follows, only

in this case the evaluation of (2.6) seems to have a reasonable simple expression. We begin with some elementary properties of the "umbral variables".

By Lemma 2 and (2.7), we have

$$e^{At} = e^{\alpha pt} \left(\frac{e^{-t}}{1-t} \right)^\alpha.$$

Let a and b be the umbral variables defined by $a = A + A_1$ and $b = B + B_1$ with the convention that

$$a^n = (A + A_1)^n \quad \text{and} \quad b^n = (B + B_1)^n$$

for all n . Therefore,

$$e^{A_1 t} = \sum_{n=0}^{\infty} (\alpha(1-p))^n \frac{t^n}{n!} = e^{(1-p)\alpha t}.$$

It follows that

$$\sum_{n=0}^{\infty} a^n \frac{t^n}{n!} = e^{at} = e^{At} e^{A_1 t} = (1-t)^{-\alpha} = \sum_{n=0}^{\infty} (\alpha)_n \frac{t^n}{n!}.$$

Thus $a^n = (\alpha)_n$, and similarly, $b^n = (\beta)_n$.

LEMMA 4. — For all k ,

$$e^{at} a^k = (1-t)^{-\alpha} \left(\frac{a}{1-t} \right)^k.$$

PROOF. — The umbral convention implies

$$\begin{aligned} e^{at} a^k &= \sum_{n \geq 0} a^{n+k} \frac{t^n}{n!} = \sum_{n \geq 0} (\alpha)_{n+k} \frac{t^n}{n!} \\ &= (\alpha)_k \sum_{n \geq 0} (\alpha+k)_n \frac{t^n}{n!} = (1-t)^{-\alpha} \left(\frac{a}{1-t} \right)^k. \quad \square \end{aligned}$$

LEMMA 5. — Let $f(a, b) = \sum_{n \geq 0} f_n(a, b) t^n$ be a formal power series with $f_n(a, b)$ a polynomial of a and b , then

$$(2.8) \quad e^{at+bt} f(a, b) = (1-t)^{-\alpha} (1-t)^{-\beta} f\left(\frac{a}{1-t}, \frac{b}{1-t}\right).$$

PROOF. — By Lemma 4, we have

$$e^{at} f(a, b) = (1 - t)^{-\alpha} f\left(\frac{a}{1 - t}, b\right),$$

and

$$e^{bt} f(a, b) = (1 - t)^{-\beta} f\left(a, \frac{b}{1 - t}\right).$$

Putting these together, we get (2.8). \square

It follows from Lemma 5 that

$$\begin{aligned} (2.9) \quad & e^{-(1-q)\beta at - (1-p)\alpha bt} \exp\left(\frac{abt}{1 - \gamma t}\right) \\ &= (1 + (1 - q)\beta t)^\alpha (1 + (1 - p)\alpha t)^\beta \\ & \quad \exp\left(\frac{abt}{(1 + \beta(1 - q)t)(1 + \alpha(1 - p)t)(1 - \gamma t)}\right) \\ &= \sum_{n \geq 0} \frac{t^n}{n!} \frac{(\alpha)_n (\beta)_n}{(1 + \beta(1 - q)t)^{\alpha+n} (1 + \alpha(1 - p)t)^{\beta+n} (1 - \gamma t)^n}. \end{aligned}$$

On the other hand, (2.6) may be rewritten as

$$(2.10) \quad \exp\left(\frac{abt}{1 - \gamma t} + A_1 B_1 t - a B_1 t - b A_1 t\right) = e^{(1-p)(1-q)\alpha\beta t} e^{-(1-q)\beta at - (1-p)\alpha bt} \exp\left(\frac{abt}{1 - \gamma t}\right).$$

Combining (2.3), (2.9) and (2.10), we obtain the generating function

$$L^*(t) = e^{\alpha\beta t(1-p-q+pq+p\gamma-q\gamma)} \sum_{n \geq 0} \frac{t^n}{n!} \frac{(\alpha)_n (\beta)_n}{(1 + \beta(1 - q)t)^{\alpha+n} (1 + \alpha(1 - p)t)^{\beta+n} (1 - \gamma t)^{n+\alpha\beta}}.$$

Finally, thanks to Lemma 1, formula (1.2) follows by changing γ to $\gamma - 1$ in the above formula.

REMARK. — *The same method can be used to count a pair of involutions with the same weight. More precisely, let*

$$I_n(\alpha, \beta, p, q, \gamma) = \sum_{(\alpha, \tau)} \alpha^{\text{cyc } \sigma} \beta^{\text{cyc } \tau} p^{\text{fix } \sigma} q^{\text{fix } \tau} \gamma^{\text{fix } \sigma \tau^{-1}},$$

where the summation is over all pairs (α, τ) of involutions on $[n]$, then we have the following formula :

$$1 + \sum_{n \geq 0} I_n(\alpha, \beta, p, q, \gamma) \frac{t^n}{n!} = e^{\alpha\beta pq(\gamma-1)t + \frac{1}{2}(\gamma^2-1)\alpha\beta t^2} \times (1 - \alpha\beta t^2)^{-1/2} \exp\left\{ \frac{2pqt + (\alpha q^2 + \beta p^2)t^2}{2(1 - \alpha\beta t^2)} \right\}.$$

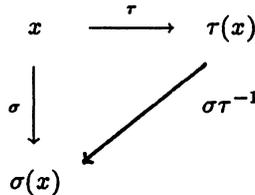
We note that in case $\gamma = 1$, the above formula reduces to Mehler's bilinear generating function for Hermite polynomials (cf. [4]).

3. The linearization coefficients of Jacobi polynomials. — Let us first state some properties of the c -fixed points of a multipermutation, which are crucial to relate $I_n(\alpha, \beta)$ to $L_n(\alpha, \beta, p, q, \gamma)$.

LEMMA 6. — For a pair of permutations (σ, τ) of \mathfrak{S}_N , the two following conditions are equivalent :

- (1) $\text{Inc } \sigma \cap \text{Inc } \tau = \emptyset$;
- (2) $\text{Inc } \sigma\tau^{-1} = \emptyset$.

PROOF. — For every $x \in N$, the following diagramme is commutative :



Hence, $x \in \text{Inc } \sigma \cap \text{Inc } \tau$ if and only if $\tau(x) \in \text{Inc } \sigma\tau^{-1}$. \square

LEMMA 7. — For a pair of permutations (σ, τ) of \mathfrak{S}_N , the two following conditions are equivalent :

- (1) $\text{Inc } \sigma \cup \text{Inc } \tau = N$ and $\text{Inc } \sigma \cap \text{Inc } \tau = \emptyset$,
- (2) $|\text{Inc } \sigma| + |\text{Inc } \tau| = |N|$ and $\text{Inc } \sigma\tau^{-1} = \emptyset$.

Hence, we can replace the condition of summation in Proposition 1 by the second one in Lemma 7.

PROPOSITION 2. — *We have*

$$I_{\mathbf{n}}(\alpha, \beta) = \sum (-1)^{|\text{Inc } \tau|} (\alpha + 1)^{\text{cyc } \sigma} (\beta + 1)^{\text{cyc } \tau},$$

where the summation is over all pairs (σ, τ) of permutations of N with the property that $|\text{Inc } \sigma| + |\text{Inc } \tau| = |N|$ and $\text{Inc } \sigma\tau^{-1} = \emptyset$.

Comparing Proposition 2 with (1.5), we get the following relation :

PROPOSITION 3

$$I_{\mathbf{n}}(\alpha, \beta) = \lim_{p \rightarrow 0} p^{|\mathbf{n}|} L_{\mathbf{n}}(\alpha + 1, \beta + 1, p^{-1}, -p^{-1}, 0).$$

If $\mathbf{n} = (1, \dots, 1)$, we may identify N with $[m]$ and Proposition 3 reduces to

$$I_m := I_{\mathbf{n}}(\alpha, \beta) = \lim_{p \rightarrow 0} p^m L_m(\alpha + 1, \beta + 1, p^{-1}, -p^{-1}, 0).$$

Therefore, it follows from (1.2) that

$$(3.1) \quad \sum_{m \geq 0} I_m \frac{t^m}{m!} = (1 - (\alpha + 1)t)^{-(\beta+1)} (1 + (\beta + 1)t)^{-(\alpha+1)}.$$

It follows that $I_1 = 0$, $I_2 = (\alpha + 1)(\beta + 1)(\alpha + \beta + 2)$, and

$$I_3 = 2(\alpha + 1)(\beta + 1)(\alpha + \beta + 2)(\alpha - \beta).$$

In case $\alpha = \beta$, formula (3.1) reduces to

$$\sum_{m \geq 0} I_m \frac{t^m}{m!} = (1 - (\alpha + 1)^2 t^2)^{-(\alpha+1)}.$$

In other words, we have

$$(3.2) \quad I_m = \begin{cases} (m/2 + 1)_{m/2} (\alpha + 1)_{m/2} (\alpha + 1)^m, & \text{if } m \text{ is even;} \\ 0, & \text{if } m \text{ is odd.} \end{cases}$$

It should note that a direct combinatorial proof of (3.2) is not difficult.

Some other special values of $I_{\mathbf{n}}(\alpha, \beta)$, evaluated combinatorially by FOATA and ZEILBERGER in [7], may be derived more directly from the original meaning of the linearization of Jacobi polynomials.

PROPOSITION 4. — If $n_1 > n_2 + \dots + n_m$, then $I_{\mathbf{n}}(\alpha, \beta) = 0$. If $n_1 = n_2 + \dots + n_m$, then $I_{\mathbf{n}}(\alpha, \beta) = n_1!(\alpha + 1)_{n_1}(\beta + 1)_{n_1} \prod_{i=2}^m (\alpha + \beta + n_i + 1)_{n_i}$.

PROOF. — In view of the orthogonality of Jacobi polynomials, $I_{\mathbf{n}}(\alpha, \beta)$ are the coefficients in the following expansion :

$$(3.3) \quad P_{n_2}^{(\alpha, \beta)}(x) \cdots P_{n_m}^{(\alpha, \beta)}(x) = \sum_{n_1 \geq 0} J_{\mathbf{n}}^{-1}(\alpha, \beta) I_{\mathbf{n}}(\alpha, \beta) P_{n_1}^{(\alpha, \beta)}(x),$$

where $J_{\mathbf{n}}(\alpha, \beta) = n_2! \cdots n_m! (1 + \alpha + \beta + n_1)_{n_1} (\alpha + 1)_{n_1} (\beta + 1)_{n_1}$. Note that $P_n^{(\alpha, \beta)}(x)$ is a polynomial of degree n with the leading coefficient being

$$(3.4) \quad K_n := K_n(\alpha, \beta) = 2^{-n} \sum_{k=0}^n \binom{n+\alpha}{k} \binom{n+\beta}{n-k} = 2^{-n} \binom{2n+\alpha+\beta}{n}.$$

Therefore, $I_{\mathbf{n}}(\alpha, \beta) = 0$ if $n_1 > n_2 + \dots + n_m$, for the left-hand side of (3.3) is a polynomial of degree $n_2 + \dots + n_m$. Equating the leading coefficients of the two sides of (3.3) yields then

$$J_{\mathbf{n}}(\alpha, \beta) K_{n_2} \cdots K_{n_m} = I_{\mathbf{n}}(\alpha, \beta) K_{n_1}.$$

So that $I_{\mathbf{n}}(\alpha, \beta) = n_1!(\alpha + 1)_{n_1}(\beta + 1)_{n_1} \prod_{i=2}^m (\alpha + \beta + n_i + 1)_{n_i}$. \square

Let e_k be the k^{th} elementary symmetric function of x_1, \dots, x_m and $\mathbf{x}^{\mathbf{n}} = x_1^{n_1} \cdots x_m^{n_m}$ for each $\mathbf{n} = (n_1, \dots, n_m)$.

THEOREM 2. — Let

$$(3.5) \quad P(\mathbf{n}; \alpha, p) = \sum_{\pi \in \mathfrak{S}_N} \alpha^{\text{cyc } \pi} p^{|\text{Inc } \pi|},$$

then

$$(3.6) \quad \sum_{\mathbf{n}} P(\mathbf{n}; \alpha, p) \frac{\mathbf{x}^{\mathbf{n}}}{\mathbf{n}!} = \left(1 - \sum_{k=1}^m (1-p)^{k-1} (p+k-1) e_k \right)^{-\alpha},$$

PROOF. — The β -extension of MacMahon's Master Theorem [6] implies that

$$(3.7) \quad \sum_{\mathbf{n}} P(\mathbf{n}; \alpha, p) \frac{\mathbf{x}^{\mathbf{n}}}{\mathbf{n}!} = \{(-1)^m x_1 \cdots x_m D_m(\mathbf{x})\}^{-\alpha},$$

where $D_m(\mathbf{x})$ is the following determinant

$$\begin{vmatrix} p - x_1^{-1} & 1 & \dots & 1 \\ 1 & p - x_2^{-1} & \dots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \dots & p - x_m^{-1} \end{vmatrix}.$$

It is well-known (cf. [14, p. 440]) that

$$(3.8) \quad \begin{vmatrix} c_1 & a & \dots & a \\ b & c_2 & \dots & a \\ \vdots & \vdots & \ddots & \vdots \\ b & b & \dots & c_n \end{vmatrix} = \frac{a\phi(b) - b\phi(a)}{a - b},$$

where $\phi(x) = (c_1 - x)(c_2 - x) \cdots (c_n - x)$. If $a = b$, (3.8) reduces to

$$(3.9) \quad \begin{vmatrix} c_1 & a & \dots & a \\ a & c_2 & \dots & a \\ \vdots & \vdots & \ddots & \vdots \\ a & a & \dots & c_n \end{vmatrix} = \phi(a) - a\phi'(a).$$

Applying (3.9) to $D_m(\mathbf{x})$, we have

$$\begin{aligned} D_m(\mathbf{x}) &= \prod_{i=1}^m (p - 1 - x_i^{-1}) + \sum_{i=1}^m \frac{\prod_{j=1}^m (p - 1 - x_j^{-1})}{(p - 1 - x_i^{-1})} \\ &= \sum_{k=0}^m (-1)^k (p - 1)^{m-k} \bar{e}_k + \sum_{i=1}^m \sum_{k=0}^{m-1} (-1)^k (p - 1)^{m-1-k} \bar{e}_k^{(i)}, \end{aligned}$$

where \bar{e}_k (resp. $\bar{e}_k^{(i)}$) is the k th elementary symmetric function of $\{x_1^{-1}, \dots, x_m^{-1}\}$ (resp. $\{x_1^{-1}, \dots, x_m^{-1}\} \setminus \{x_i^{-1}\}$). Since $\sum_{i=1}^m \bar{e}_k^{(i)} = (m - k)\bar{e}_k$, we can rewrite $D_m(\mathbf{x})$ as

$$(3.10) \quad D_m(\mathbf{x}) = (-1)^m \bar{e}_m + \sum_{k=0}^{m-1} (-1)^k (p - 1)^{m-k-1} (p - 1 + m - k) \bar{e}_k.$$

Note that $x_1 \cdots x_m \bar{e}_k = e_{m-k}$. Therefore, (3.6) follows by substituting (3.10) into (3.7). \square

REMARK 1. — *The above theorem is a multiextension of Lemma 2. The latter can be derived from Theorem 2 by applying an appropriate transformation (cf. [11, p. 233]).*

By Theorem 2 of [5], we can also express $P(\mathbf{n}; \alpha, p)$ in terms of the integral of a product of the Laguerre polynomials :

$$(3.11) \quad P(\mathbf{n}; \alpha + 1, p) = \frac{\mathbf{n}!(p-1)^{|\mathbf{n}|}}{\Gamma(\alpha+1)} \int_0^\infty \prod_{i=1}^m L_{n_i}^{(\alpha)}\left(\frac{x}{1-p}\right) x^\alpha e^{-x} dx,$$

where $L_{n_i}^{(\alpha)}(x)$ ($1 \leq i \leq m$) are Laguerre polynomials defined by

$$n_i! L_{n_i}^{(\alpha)}(x) = \sum_{k=0}^{n_i} (-1)^k \binom{n_i}{k} (\alpha + 1 + k)_{n_i-k} x^k \quad (i = 1, \dots, m).$$

Inversely, the equation (3.11) provides also a combinatorial interpretation for the integral of its right side, which implies that the integral above is actually a polynomial in $\alpha + 1$ and p with *nonnegative integral coefficients*.

PROPOSITION 5. — *For all $\mathbf{n} = (n_1, \dots, n_m)$, we have*

$$(3.12) \quad L_{\mathbf{n}}(1, 1, 1, 1, \gamma) = |\mathbf{n}|! P(\mathbf{n}; 1, \gamma),$$

$$(3.13) \quad L_{\mathbf{n}}(\alpha, \beta, p, q, 1) = P(\mathbf{n}; \alpha, p) P(\mathbf{n}; \beta, q).$$

PROOF. — For each $\pi \in \mathfrak{S}_N$, there are exactly $|\mathbf{n}|!$ pairs of the permutations (σ, τ) of \mathfrak{S}_N satisfying $\sigma\tau^{-1} = \pi$. So that

$$\begin{aligned} L_{\mathbf{n}}(1, 1, 1, 1, \gamma) &= \sum_{\sigma, \tau \in \mathfrak{S}_N} \gamma^{|\text{Inc } \sigma\tau^{-1}|} \\ &= \sum_{\pi \in \mathfrak{S}_N} \gamma^{|\text{Inc } \pi|} \sum_{\sigma\tau^{-1}=\pi} 1 \\ &= |\mathbf{n}|! P(\mathbf{n}; 1, \gamma). \end{aligned}$$

The formula (3.13) is obvious. \square

4. **Concluding remarks.** — The method used in Section 2 seems not to be generalized to count a pair of generalized permutations. The essential difference is that an element $a \in \text{Inc } \sigma\tau^{-1}$ is no longer equivalent to $\sigma\tau^{-1}(a) = a$. Instead of representing the permutations as *digraphs*, we can also represent them as *bipartite graphs*, as GESSEL [9] does for counting general Latin rectangles, and eventually

to get some counting formula. However, as far as the linearization problem is concerned, it is essential to find a *nonnegative* expression. In this regard, even in the case of $m = 3$, the nonnegative expression of $I_n(\alpha, \beta)$ given by RAHMAN [15] is already so complicated to quote here. However, if $\alpha = \beta$, RAHMAN's formula greatly simplifies and reduces to HSU's formula for the ultraspheric polynomials [12] :

$$(4.1) \quad I_{s+j, n-s, n}(\alpha, \beta) = \binom{s+j}{s, j/2, j/2} s! \left(\frac{j}{2}\right)! \binom{n}{s+j/2} (n-s)! \\
(\alpha + 1 + j/2)_{n-j/2} (\alpha + 1 + s + j/2)_{j/2} \\
(\alpha + 1 + n - s - j/2)_{j/2} (\beta + 1)_{n+j/2} \\
(\alpha + \beta + 1 + s + j)_s (\alpha + \beta + 1 + n - s)_{n-s-j} \\
(\alpha + \beta + 1)_j (\alpha + \beta + n + 1)_{j/2},$$

where we assume $n_3 = \max\{n_1, n_2, n_3\}$ and let $n_1 = s + j, n_2 = n - s, n_3 = n$ with $0 \leq s \leq n, j$ even and $0 \leq j \leq n - s$. It is still an open problem to give a nice combinatorial proof of the above formula.

In the special cases where $\alpha = -1/2, \alpha = 1/2$, and $\alpha = 0$, the values of $I_n(\alpha, \alpha)$ correspond respectively to the linearization coefficients of Tchebycheff polynomials of the first, second kind and Legendre polynomials. A combinatorial interpretation of $I_n(-1/2, -1/2)$ in the context of rook polynomials has been given by RIORDAN [15, pp. 202-203]. A combinatorial interpretation of $I_n(1/2, 1/2)$ has been given by DE SAINTE-CATHERINE and VIENNOT [18] in terms of *Dyck words* or *paths*. Both interpretations above imply directly the positivities of the corresponding coefficients. More recently, GILLIS, JEDWAB and ZEILBERGER [10] came up with a combinatorial interpretation of $I_n(0, 0)$ in terms of *colored graph*. However, the derivation of the positivity from the latter interpretation is not obvious, for the *weight is not positive*. Therefore, the problem to find out the positive "invariant" in the linearization coefficients of Jacobi polynomials remains open, even in the case of Legendre polynomials.

Acknowledgement. — Part of this work was done while the author was visiting Institute for Advanced Study as a member in 1989-1990 and supported by NSF Grant DMS-8610730.

REFERENCES

- [1] ANDREWS (G.), GOULDEN (I. P.) and JACKSON (D. M.). — Generalization of Cauchy's summation theorem for Schur functions, *Trans. Amer. Math. Soc.*, t. 310, 1988, p. 805–820.
- [2] BOGARD (K. P.) and LONGYEAR (J. Q.). — Counting 3 by n Latin rectangles, *Proc. Amer. Math. Soc.*, t. 54, 1976, p. 463–467.
- [3] FOATA. — La série génératrice exponentielle dans les problèmes d'énumération, *Presses de l'Université de Montreal*, 1978.
- [4] FOATA. — A combinatorial proof of the Mehler formula, *J. Combinatorial Theory, Ser. A*, t. 24, 1978, p. 367–376.
- [5] FOATA (D.) and STREHL (V.). — Combinatorics of Laguerre polynomials, *Enumeration and Design*, ed. D. M. Jackson and S. A. Vanstone, *Academic Press*, 1984, p. 123–140.
- [6] FOATA (D.) and ZEILBERGER (D.). — Laguerre polynomials, weighted derangements and positivity, *SIAM J. Discrete Math.*, t. 1, 1988, p. 423–433.
- [7] FOATA (D.) and ZEILBERGER (D.). — Linearization coefficients for Jacobi polynomials, *Actes 16^{ème} Séminaire Lotharingien, n° 341/S-16, Publ. I.R.M.A., Strasbourg*, 1987, p. 73–86.
- [8] GESSEL (Ira. M.). — Counting three-line Latin rectangles, *Lectures Notes in Math.*, t. 1234, 1985, p. 106–111.
- [9] GESSEL (Ira. M.). — Counting Lating rectangles, *Bull. Amer. Math. Soc.*, t. 16, 1987, p. 79–82.
- [10] GILLIS (J.), JEDWAB (J.) and ZEILBERGER (D.). — A combinatorial interpretation of the integral of the product of Legendre polynomials, t. 19, 1988, p. 1455–1461.
- [11] GOULDEN (I. P.) and JACKSON (D. M.). — *Combinatorial Enumeration*. — New York, Wiley, 1983.
- [12] HSU (H. Y.). — Certain integrals and infinite series involving ultraspheric polynomials and Bessel functions, *Duke Math. J.*, t. 4, 1938, p. 374–383.
- [13] MEIXNER (J.). — Erzeugende Functionen der Charlierschen Polynome, *Math. Z.*, t. 44, 1938, p. 531–535.
- [14] MUIR (T.). — *A treatise on the theory of determinants*. — New York, Dover, 1960.
- [15] RAHMAN (M.). — A non-negative representation of the linearization coefficients of the product of Jacobi polynomials, *Canad. J. Math.*, t. 33, 1981, p. 915–928.
- [16] RIORDAN (J.). — *An introduction to combinatorial analysis*. — New York, Wiley, 1967.
- [17] ROTA (G. -C). — The number of partitions of a set, *Amer. Math. Monthly*, t. 71, 1964, p. 498–504.
- [18] DE SAINTE-CATHERINE (M.) and VIENNOT (G.). — Combinatorial interpretation of integrals of products of Hermite, Laguerre and Tchebycheff polynomials, *Lecture Notes in Math.*, n° 1171, 1985, p. 120–128.

UFR de Mathématiques et Informatique
 Université Louis Pasteur
 7, rue René Descartes
 67000 Strasbourg Cedex, France
 Email : A18643@FRCCSC21.BITNET