Combinatoire analytique : application aux marches aléatoires

Cyril Banderier
Cyril.Banderier@inria.fr

Mémoire du DEA algorithmique
Juillet 1998

Travail dirigé et encadré par Philippe Flajolet
Philippe.Flajolet@inria.fr
Projet Algorithmes INRIA

Cette version est celle disponible, le 23/10/98, à l'adresse
http://algo.inria.fr/banderier/index.html
et est susceptible de modifications.

Résumé.

Ce mémoire a pour sujet le dénombrement de marches aléatoires. Nous mettons en place un formalisme, une **méthode de calcul** permettant l'obtention aisée de séries génératrices de marches aléatoires.

La méthode repose sur l'obtention d'un **opérateur linéaire**, qui donne une relation de récurrence entre les marches et qui permet l'obtention d'une équation fonctionnelle, que l'on résout avec la **méthode du noyau**. De telles équations fonctionnelles laissent également la perspective d'un traitement automatique, grâce au choix des opérateurs, par la théorie de l'**holonomie**.

Ce mémoire permet aussi une **unification** de différents résultats disséminés dans la littérature, obtenus jusqu'à présent par des méthodes combinatoires, de théorie des langages (méthodologie de Schützenberger) ou encore par des astuces ponctuelles. Nous appliquerons notre méthode pour trouver des "généralisations" des **théorèmes**

- de Labelle et Yeh [**27**] sur les **marches de cavaliers** (III section 9),
- de Duchon [**10**] sur les **langages de Dyck généralisés** (III théorème 10.1),
- de Denise [**8**] sur les **grammaires infinies** rationnelles (II section 8).

On donnera par ailleurs la série génératrice des

- **marches** et **excursions** à pas bornés (III 3.1 et 3.2),
- marches et excursions **positives** à pas arrière **non bornés** (III 4.1 et 4.2).

Elles sont **algébriques** comme une conjecture de Pinzani sur des grammaires infinies le supposait.

Nous donnons également un résultat **asymptotique** lié à l'équation $y^c = z\phi(y)$ qui généralise le cas déjà connu $y = z\phi(y)$. La démonstration passe par quelques points techniques dont un lemme qui donne une inégalité entre une série à coefficients positifs, sa dérivée et sa dérivée seconde (3.2). On aboutit finalement à une asymptotique en

$$\frac{A}{\rho^n\sqrt{\pi n^3}}.$$

La méthode est également utilisée pour étudier différents paramètres comme

- le nombres d'effleurements (on regardera le comportement **en moyenne**),
- l'aire (on aboutit alors à des q-**analogues**).

On jouera également avec les **conditions de bord** et on traitera en particulier le cas d'une limitation par une droite de **pente rationnelle.**

En conclusion, la méthode dégagée ici, par sa souplesse, permet de traiter les marches aléatoires dans une **grande généralité**. Il ne fait nul doute que son utilisation va permettre de retrouver de façon directe de nombreux résultats de combinatoire. Elle ouvre également un vaste champ de possibilités quant à l'étude de marches jusqu'à présent peu regardées. Il serait intéressant de dégager des comportements de classes de marches en **loi limite**. On peut également imaginer une **théorie des équations fonctionnelles** obtenues par le jeu d'opérateurs mentionné dans ce mémoire, elle devrait englober la théorie des q-analogues comme celles des équations aux différences finies.

Plan de lecture.

Les lecteurs déjà familiarisés avec la combinatoire analytique pourront ignorer le premier chapitre qui contient un chouia de perspective historique et des rappels sur des notions d'analyse et de développements en série.

On trouvera dans le chapitre deux la plupart des définitions clefs de ce mémoire. Il est sans doute éclairant de jeter un coup d'œil sur le tableau de la page 18 qui donne les correspondances entre une marche et son codage sous forme de série génératrice.

Le chapitre trois est constitué de nombreux exemples. On illustrera ainsi les moultes variations que la combinatoire analytique permet sur les marches aléatoires : marche homogène (ou non) en temps, marche homogène (ou non) en espace, marche absorbante, marche réfléchissante, excursion, pont, marche en dimension 2, limitation aux bords par des contraintes plus ou moins exotiques, étude de l'aire. Pour tous ces cas, nous montrerons comment aboutir très simplement à des équations fonctionnelles dont nous exposons pour certaines une méthode résolution.

Le quatrième chapitre a pour but d'établir l'asymptotique des diverses familles de marches étudiées précédemment.

On trouvera en annexe quelques lignes de programmation en Maple.

Le chapitre six permet de replacer quelques uns des intervenants de ce mémoire dans leurs époques.

La bibliographie contient essentiellement quatre classes de références :

- des liens vers des ouvrages et des articles d'horizons différents que j'ai cités dans ce mémoire pour établir quelques perspectives historiques ou pour souligner d'éventuels ponts “inter-disiplinaires” ;
- des livres offrant un panorama fort complet sur la combinatoire analytique ou sur l'analyse (complexe ou asympotique) ;
- différents articles dont les résultats sont retrouvés ou étendus par nos méthodes ;
- les mentions des communications (orales ou électroniques !) avec diverses personnes qui m'ont toutes gentiment aidé.

Table des matières

CHAPITRE 1

La combinatoire analytique

1. Analyse combinatoire versus combinatoire analytique

Quelle différence faire entre ces deux domaines, au moins sémantiquement proches ? On trouve dans [**19**] la définition suivante : "L'analyse combinatoire est l'ensemble des techniques qui servent en mathématique, à *compter* (ou *dénombrer*) certaines structures *finies*, ou à les *énumérer* (établir des listes exhaustives de structures considérées), enfin à dénombrer leur *existence* pour certaines valeurs des paramètres dont elles dépendent."

C'est en effet ce que l'on peut ressentir à la lecture du livre de Comtet (voir [**6**]), qui a fait le point des connaissances en analyse combinatoire à la fin des années 60. Depuis, ce domaine a connu des progrès immenses, grâce en particulier au lien fait avec l'analyse complexe. Un mariage fructueux entre le discret et le continu avait déjà eu lieu entre la théorie des nombres et l'analyse complexe et avait donné naissance à la théorie analytique des nombres. Cette fois-ci, le mariage donne naissance à la combinatoire analytique.

La combinatoire analytique peut traiter ce dont traitait l'analyse combinatoire, mais elle apporte surtout le grand avantage de permettre une étude asymptotique puissante, aisée et précise, c'est-à-dire de répondre à des questions du type "Combien y a-t-il environ d'objets de taille n ayant telle structure, quand n est grand ?". Le parodoxal "environ" obtient une précision satisfaisante avec l'utilisation d'échelles asymptotiques. Par exemple combien y a-t-il d'arbres binaires à n nœuds ? La combinatoire analytique permet aussi des études probabilistes, c'est-à-dire de répondre à des questions du type "Parmi les objets de taille n ayant telle structure, quelle est la proportion des objets vérifiant telle propriété ?". Par exemple, quel est le nombre moyen de cycles dans une permutation de $[1, \ldots, n]$?

Le lien avec la théorie des probabilités n'est pas le moindre, et bien souvent, on s'efforce de trouver des lois limites aux structures combinatoires vérifiant une propriété fixée.

Il reste à savoir ce qu'on appelle une "structure combinatoire". C'est simplement un ensemble (éventuellement infini) d'objets sur lesquels il existe une notion de "taille". Les objets en question doivent être un assemblage de "briques de base" de taille 1 (les atomes). Par exemple, les atomes des arbres sont les noeuds, les atomes des colliers sont les perles, etc. La deuxième question est de savoir quelle règle du jeu on adopte pour jouer avec de telles structures ? Il y en fait toute une kyrielle de constructions "admissibles" (qui sont pour certaines d'entre elles tout à fait comparables aux manipulations usuelles en théorie des langages). Une définition formelle de la notion d'admissibilité est donnée dans [**17**] et dans [**25**].

De façon intuitive, cela permet de jouer avec des structures définies récursivement (comme les arbres, les grammaires), cela devient plus délicat en ce qui concerne des structures n'ayant pas de contraintes structurelles suffisamment fortes (comme les graphes). Toutefois, il n'est pas rare que la combinatoire analytique arrive à dire son mot sur des problèmes de graphes, où elle peut concurrencer la "méthode probabiliste" (confer le modèle d'Erdős-Rényi [**1**]).

L'idée de "marier", de "combiner" la combinatoire et l'analyse n'est pas une idée si récente que les lignes précédentes pourraient le laisser croire. Elle remonte en fait à Euler [**12**], en 1748. D'ailleurs ce dernier est même allé jusqu'à étudier l'asymptotique de certains des problèmes qu'il a considérés. Historiquement, le saut suivant se fait sans doute avec les nombreuses relations de Ramanujan, que l'on classe aujourd'hui dans la théorie des q-analogues (nous donnerons un exemple d'équations fonctionnelles de q-analogues à la section 14 du chapitre 3).

Schématiquement, la combinatoire analytique utilise les domaines suivants :

- analyse
- combinatoire
- analyse complexe
- probabilités
- théorie spectrale
- théorèmes taubériens

et elle a ses applications dans les domaines suivants :

- combinatoire
- théorie des langages
- physique statistique
- théorie des nombres
- algèbre (groupes de permutations, monoïdes et groupes de représentation)
- théorie de la complexité
- analyse d'algorithmes.

En fait les deux derniers points sont bien sûr à regrouper. La combinatoire analytique permet en effet d'étudier la complexité d'un algorithme en moyenne, soit par majorations, soit par une étude asymptotique précise. Que la complexité d'un algorithme puisse être étudiée par la combinatoire analytique est lié au fait suivant : un algorithme a une structure itérative ou récursive, il est ainsi à même d'avoir un paramètre (par exemple le nombre d'itérations, de boucles “pour”, d'affectations effectuées...) tombant sous la coupe de la combinatoire analytique. Hormis les utilisations en sciences via la physique statistique, c'est sans doute la principale application “industrielle” de ce domaine.

2. Les séries génératrices

Expliquons comment on passe si aisément de problèmes issus du discret, du fini, à des mathématiques sur des objets continus, sur lesquels nous allons faire de l'analyse (complexe).

On s'intéresse à dénombrer une suite a_n d'objets de taille n ; l'idée est la suivante : on peut former la série “formelle”

$$\sum_{n\in\mathbb{N}} a_n z^n .$$

“Formelle” signifie qu'on ne se préoccupe pas pour l'instant de savoir si elle est convergente ou non, mais qu'on s'autorise toutes les opérations usuelles sur les séries (addition, multiplication, dérivation, etc). Par exemple les permutations de taille n (*i.e.* de $[1,\ldots,n]$) donnent la série formelle $\sum_{n\geq 0} n! z^n$, qui est divergente[1] pour $z \neq 0$. Ces séries divergentes se révèlent fort utiles dans la résolution de certaines équation différentielles pour lesquelles elles permettent en effet de trouver une solution convergente ! Dans le cas des séries convergentes, on est dans un domaine davantage balisé et on dispose d'une artillerie de théorèmes (voir les ouvrages [**24**], [**9**], [**17**]), par exemple pour exprimer les coefficients d'une série sous forme d'intégrales (comme le théorème de Cauchy sur les fonctions analytiques), en outre le rayon de convergence de la série donne directement une estimation asymptotique des coefficients (théorème d'Hadamard), etc.

Ainsi, on codera souvent des dénombrements sous forme

- de série génératrice ordinaire (ordinary generating function, OGF), c'est-à-dire une série du type
$$\sum_{n\in\mathbb{N}} a_n z^n ;$$
- de série génératrice exponentielle (EGF), c'est-à-dire une série du type
$$\sum_{n\in\mathbb{N}} a_n \frac{z^n}{n!} ;$$

[1] Les séries divergentes ont été manipulées très tôt dans l'histoire de l'analyse, quelquefois pour le grand horreur de mathématiciens célèbres comme Abel qui a écrit en 1828 que les séries divergentes étaient une invention du diable et qu'il était honteux de baser sur elles une démonstration ou quoi que ce soit d'autre. Les bases théoriques en ont été données par Borel [**3**] et Hardy [**23**].

on recourt à cette classe de fonctions génératrices quand les objets (de taille n) ont leurs atomes numérotés de 1 à n, c'est ce qu'on appelle "l'univers étiqueté" ;
- de série génératrice de Dirichlet (DGF), c'est-à-dire une série du type
$$\sum_{n\in\mathbb{N}} \frac{a_n}{n^z};$$
on recourt à cette classe de fonctions génératrices quand on a une suite (a_n) qui a des propriétés multiplicatives, ce qui arrive souvent dans des problèmes de théorie des nombres (voir par exemple [**33**])
- de série génératrice multivariée quand on veut suivre plusieurs paramètres. Par exemple, voici une série génératrice ordinaire bivariée (BGF)
$$\sum_{i,j\in\mathbb{N}} a_{ij}\, z^i u^j\,;$$
- de série génératrice de probabilités (PGF) si l'on souhaite non plus seulement compter mais aussi tenir compte de la probabilité d'arriver à telle configuration.

Les coefficients sont des nombres positifs, entiers dans les premiers cas et réels dans le dernier. Rappelons qu'un point où une série entière cesse d'être dérivable est appelé singularité. Les séries entières vérifient le fait suivant :

Théorème 2.1 (Pringsheim). *Une série entière à coefficients positifs et de rayon de convergence fini ρ a une de ses singularités dominantes (*i.e.* celles de plus petit module) en $z = \rho$.*

Preuve. (due à Pringsheim)
Premièrement, la série a une singularité sur le bord de son disque de convergence car sinon on pourrait la prolonger sur un disque de rayon plus grand et elle aurait ainsi un rayon de convergence $> \rho$, d'où une contradiction. D'autre part, soit z_1 l'une de ces singularités du bord du disque de $f(z) = \sum_n a_n z^n$. Soit $0 < r < 1$; puisque les coefficients sont positifs, on a $|f^{(k)}(rz_1)| \leq f^{(k)}(r\rho)$, donc la série de Taylor de f en $z = r\rho$, à savoir $f(z) = \sum f^{(k)}(r\rho)/k!(z-r\rho)^k$ a un rayon de convergence inférieur à celui de la série de Taylor de f en $z = rz_1$, à savoir $f(z) = \sum f^{(k)}(rz_1)/k!(z-rz_1)^k$. Or le rayon de convergence de cette dernière série est $\rho - r\rho$. Comme ceci est vrai pour tout $0 < r < 1$, si f était régulière en $z = \rho$, son développement en série de Taylor, a un rayon de convergence nul et donc ρ est une singularité. Par conséquent f n'est pas régulière en $z = \rho$ qui est donc une singularité. Ainsi la série a toujours une singularité dominante réelle positive. $\square$

3. Extraction de coefficients dans une série

On notera $[z^n]F(z)$ le coefficient de z^n dans le développement en série de F en zéro. Cette notation laconique couvre un abus de langage ; par exemple (didactique !), si on considère

$$F(x) = 1 + x + x^2 + \cdots = \frac{1}{1-x} = \frac{x^{-1}}{\frac{1}{x}-1} = -\frac{1}{x}\frac{1}{1-\frac{1}{x}} = -\frac{1}{x}\,(1 + \frac{1}{x} + \frac{1}{x^2} + \cdots),$$

on arrive donc à un deuxième développement

$$F(x) = -\frac{1}{x} - \frac{1}{x^2} - \cdots .$$

Si l'on extrait maintenant le coefficient de degré 1, on obtient $[x^1]F = 1$ pour la première série et $[x^1]F = 0$ pour la deuxième. Le paradoxe est levé et l'unicité est rétablie si l'on rappelle que la notation $[z^n]$ est reliée à un contexte (*i.e.* un type de développement choisi en un point fixé).

Dans ce mémoire, nous aurons affaire à des séries génératrices ordinaires et les développements que nous considérerons, sauf mention explicite du contraire, seront des développements de Puiseux (qui se réduisent bien sûr en des développements de Taylor dans divers cas).

Nous utiliserons souvent la propriété suivante :

$$[z^0]F(z) = F(0)$$

et, plus généralement,

$$[z^n]F(z) = \frac{1}{n!}\frac{\partial^n F}{\partial z^n}(0).$$

4. Fonctions algébriques

Pour un polynôme bivarié

$$P(z,u) = p_0(z)u^n + p_1(z)u^{n-1} + \cdots + p_n(z)$$

que nous supposons irréductible dans $\mathbb{C}[z,u]$, nous allons nous intéresser à l'équation $P(z,u)=0$ par rapport à u.

Pour z fixé cette équation a n racines distinctes sauf si z est tel que

- $p_0(z) = 0$ (car alors on a un polynôme de degré $< n$),
- P a des racines multiples.

Dans ces deux cas, on dit que z est une singularité.

En faisant varier z sur tout $\mathbb{C}$ sauf autour de petits disques entourants les singularités, on parvient (voir [**24**]) à définir n fonctions $u_1, \ldots, u_n$ racines de $P(z,u) = 0$. Ces n racines sont telles qu'il existe pour chacunes d'elles des entiers positifs r_i et q_i tels que $t^{\frac{r_i}{q_i}} u_i(t^{q_i})$ soit holomorphe au voisinage de $t = 0$; on a donc

$$u_i(z) = \sum_{n \geq n_i} (c_i)_n z^{\frac{n}{q_i}}.$$

Ainsi, ces racines correspondent aux n branches d'une fonction analytique $u(z)$ dont les singularités z_i sont soit telles que $p_0(z_i) = 0$ (et alors au moins l'une des branches a un pôle en z_i), soit telles que $P(z_i,u) = 0$ a une racine multiple u (et alors au moins l'une des branches a un point de branchement en z_i).

La section suivante va nous permettre d'expliciter le développement en "série" de $u_1, \ldots, u_n$.

5. Diagramme de Newton et développement en série de Puiseux

Une série de Puiseux, ramenée en 0, est une série du type

$$\sum_{n \geq n_0} \alpha_n z^{n/q}$$

pour un entier n_0 et un entier strictement positif q. Afin d'obtenir un développement en série de Puiseux des solutions (que nous appellons aussi "branches") de

$$P(z,u) = 0,$$

on écrit P sous la forme

$$P(z,u) = \sum a_{ij} z^i u^j$$

et on suppose $P(0,0) = 0$ *i.e.* $a_{00} = 0$.

On forme alors le polygone de Newton *i.e.* l'ensemble des points (i,j) tels que $a_{ij} \neq 0$.

Puisque $P(z,u)$ est irréductible on ne peut donc pas le factoriser par u^m ni par z^m pour un entier $m > 0$ et ainsi, le polygone de Newton a toujours des points sur les deux axes. Soit A_0 de coordonnées $(x_0, 0)$ le point le plus proche de l'origine sur l'axe des abscisses et A_k de coordonnées $(0, y_k)$ le point le plus proche de l'origine sur l'axe des ordonnées.

Pour connaître le comportement de différentes branches en $z = 0$, nous allons nous intéresser à la partie $S = [A_0, A_1] \cup \cdots \cup [A_{k-1}, A_k]$ de l'enveloppe convexe du polygone, où les A_i sont des points extrémaux du polygone. Nous noterons (x_i, y_i) les coordonnées du point A_i $(0 \leq i \leq k)$.

Il y a x_0 branches qui tendent vers 0 quand z tends vers 0, que nous appellerons petites "branches". De façon plus précise :

Théorème 5.1. *Notons $px + qy = r$ l'équation de la droite (A_0, A_1) et posons $\mu = \frac{y_1}{q}$. Il y a μp branches équivalentes à $z^{\frac{q}{p}}$ quand $z = 0$.*

PREUVE. On a $p \leq y_0$ et $q \leq x_0$. La droite (A_0, A_1) a $\frac{-p}{q}$ comme coefficient directeur.

En fait $1+\mu$ est le nombre de points (i,j) qui sont (A_0, A_1) et tels que $a_{ij} \neq 0$

En effet si on pose $z = t^p$ et $u = t^q u_2$, on a donc

$$P(z,u) = \sum a_{ij} t^{ip+jq} u_2^j = 0$$

mais pour tous les points (i,j) qui sont sur (A_0, A_1) on a $ip + jq = r$ et pour les autres on a $ip + jq > r$, ainsi on peut diviser la somme par t^r, puis en scindant en deux sommes, l'une sur les points de (A_0, A_1), la deuxième sur les points restants, on obtient :

$$\sum a_{ij} u_2^j + \sum a_{ij} a_{ij} t^{ip+jq-r} u_2^j = 0.$$

On trouvera le restant de la preuve dans [**24**] □

Si $P(0,0) \neq 0$ de deux choses l'une

- soit il existe des a_{0j} non nuls et dans ce cas on retombe dans le cadre étudié ci-dessus en posant $u := u - r$ où r est une solution de $\sum_{j>0} a_{0j} u^j = P(0,0)$;
- soit tous les a_{0j} sont nuls et dans ce cas il est clair que toutes les branches sont divergentes.

Nous aurons besoin de cette remarque dans la section sur les marches positives à pas arrières non bornés.

CHAPITRE 2

Marches aléatoires

1. Méthode probabiliste, analytique ou combinatoire ?

Les marches aléatoires (avec leurs variantes sur les conditions de bords : absorption, réflexion ...) ont été étudiées par les probabilistes. Pour les marches en temps discrets, des outils tels la décomposition de Wiener-Hopf permettent de traiter des paramètres comme le maximum d'une marche.

Les marches aléatoires apparaissent incontournables dans divers domaines allant de problèmes d'automatisme ([**34**]), d'ordonnancement en informatique ([**13**]), de génétique ([**21**]), à des problèmes liés à des réseaux cristallins en physique statistique... Par exemple une fourmi dans un labyrinthe, simulant une marche aléatoire sur un réseau de liens, permet d'étudier la diffusion à deux et trois dimensions dans un milieu poreux ou de déterminer l'exposant de la conductivité ([**32**]).

2. Définition d'une marche en dimension n

Commençons par donner un exemple : celui de l'ivrogne de Pólya. "C'est l'histoire d'un mec qui bu un coup d'trop et qui voudrait bien rentrer chez lui, mais il est tellement envinassé que toutes les dix secondes, il fait au hasard un pas en avant ou un pas arrière..." D'aucuns, à l'âme scientifique, diront que notre ivrogne a ainsi le comportement d'une marche aléatoire unidimensionnelle. On peut aussi regarder la variante dans le plan, où notre ami fait au hasard un pas en avant, un pas en arrière, un pas vers la droite ou un pas vers la gauche. Pólya s'était intéressé à la probabilité pour que l'ivrogne reviennent à son point de départ et on a le résultat surprenant suivant : sur une droite ou dans le plan, cette probabilité est 1, tandis qu'en dimension > 2, cette probabilité est < 1. Nous verrons à la section 12 comment nous pouvons traiter ce genre de problème.

Nous allons maintenant donner une définition un peu formelle des marches aléatoires.

Une marche aléatoire sur $\mathbb{Z}^n$ est une application ψ de $\mathbb{Z}^n$ dans $M(\mathbb{Z}^n)$, la famille des multi-ensembles de $\mathbb{Z}^n$, c'est-à-dire une application qui à tout point de $\mathbb{Z}^n$ associe un multi-ensemble de $\mathbb{Z}^n$ (un multi-ensemble est "un ensemble où les éléments peuvent apparaître plusieurs fois" ainsi $\{0,2,3\} \neq \{0,2,3,2\}$). Donnons une explication plus intuitive : vous êtes à la gare Saint-Lazare, que vous voulez absolument fuir, sur la voie 1 et sur la voie 3, il y un train en direction de Versailles et sur la voie 15 il y a un train pour Évreux, les autres voies sont en grève. On a ainsi ψ(St-Lazare)=$\{$Versailles, Versailles, Évreux$\}$.

Éventuellement, on peut vouloir tenir compte de probabilités, introduire des pondérations, et, dans ce cas, on a une application ψ qui à tout point de $\mathbb{Z}^n$ associe un multi-ensemble constitué de paires (A_i, p_i) où les A_i sont des multi-ensemble de $\mathbb{Z}^n$ et où $\sum_i p_i$ vaut 1.

On appelle "marche de longueur k et partant du point A" une suite de points $A_0, A_1, \ldots, A_k$ où $A_0 = A$ et $A_{i+1} \in \psi(A_i)$. On qualifiera souvent une telle suite de "marche" ou de "chemin" (commençant en A et finissant en A_k). Pour deux points A et B, si $B \in \psi(A)$, on dit qu'il y a un pas entre A et B.

Souvent, on prend l'origine de $\mathbb{Z}^n$ comme point de départ A de la marche, on peut également prendre un multi-ensemble, les modifications à faire aux définitions précédentes sont triviales.

Une marche qui part de l'origine et qui finit sur $\mathbb{Z}^{n-1} \times \{0\}$ est appelée une "excursion". Ainsi, en dimension 1, une excursion est une marche qui part de 0 et qui y finit, en dimension 2, c'est une marche qui part de (0,0) et qui finit sur l'axe des abscisses.

Définition 2.1. *On code un point $P \in \mathbb{Z}^n$ de coordonnées $(a_1, \ldots, a_n)$ par le monôme*

$$\eta(P) := u_1^{a_1} \cdots u_n^{a_n}.$$

On prolonge ψ à $M(\mathbb{Z}^n)$ par

$$\psi : E \in M(\mathbb{Z}^n) \mapsto \cup_{P \in E} \psi(P)$$

où l'union se fait au sens des multi-ensembles. On prolonge ce codage η à un ensemble de points par

$$\eta : E \in M(\mathbb{Z}^n) \mapsto \sum_{P \in \psi(E)} \eta(P).$$

Définition 2.2. *On code alors une marche aléatoire par la série génératrice*

$$F(z, u_1, \ldots, u_n) := \sum_{k \geq 0} \eta(\psi^k(A)) z^k.$$

Ainsi F est dans l'anneau des séries formelles à coefficients entiers en $n+1$ variables $\mathbb{Z}[[z, u_1, \ldots, u_n]]$. Quand on tient compte des probabilités, les coefficients de la série génératrice F ne sont plus entiers mais réels et on travaillera alors sur $\mathbb{R}[[z, u_1, \ldots, u_n]]$.

Le nombre de chemins de longueur k allant de A au point de coordonnées $(a_1, \ldots, a_n)$ est donné par $f_{k, a_1, \ldots, a_n} := [z^k u_1^{a_1} \cdots u_n^{a_n}] F(z, u_1, \ldots, u_n)$. Le nombre total de chemins de longueur k est $f_k := [z^k] F(z, 1, \ldots, 1)$.

On considère des marches aléatoires ψ telles que les multi-ensembles images des points de $\mathbb{Z}^n$ soient toujours de cardinal fini ; ceci assure l'existence formelle des séries correspondantes.

Nous allons donner dans ce mémoire une méthode pour obtenir explicitement cette fonction génératrice F, pour une classe de marches aléatoires "régulières", c'est-à-dire s'exprimant sous la forme

$$\sum_{i=1}^{\mu} \chi_i \psi_i$$

pour un certain entier μ, avec des fonctions indicatrices χ_i telles que

$$\sum_{A \in \mathbb{Z}^n} \chi_i(A) \in \mathbb{Q}(u_1, \ldots, u_n),$$

c'est-à-dire que ce sont des indicatrices d'ensembles "rationnels", avec les conventions $0 \times E = \emptyset$ et $1 \times E = E$ (pour un multi-ensemble E) et $(\psi + \phi)(A) = \psi(A) \cup \phi(A)$. On demande de plus que les $\eta \circ \psi_i$ appartiennent à l'anneau du jeu d'opérateurs

$$\{\Delta_{0,i}, \Delta_{1,i}, S_i, \partial_i, \sigma_i\}_{i=1,\ldots,n}$$

où

- $\Delta_{0,i}$ est l'opérateur de différence finie en 0 par rapport à la i-ième variable ;
- $\Delta_{1,i}$ est l'opérateur de différence finie en 1 par rapport à la i-ième variable ;
- S_i est l'opérateur de "shift", *i.e.* la multiplication par u_i ;
- ∂_i est l'opérateur qui dérive par rapport à la i-ième variable u_i ;
- σ_i un opérateur de substitution de la i-ième variable ;

L'opérateur σ couvre en fait une infinité d'opérateurs de substitution différents. Nous utiliserons en particulier la substitution de u_i par u_i^α, que nous noterons $\sigma_{i,\alpha}$. Lorsque α est entier, ces opérateurs sont connus sous le nom d'opérateurs de Mahler [**11**].

Nous utiliserons également les opérateurs suivants $\{U_i, Z_i\}_{i=1,\ldots,n}$, où

- U_i est l'opérateur qui remplace la i-ième variable u_i par 1,
- Z_i est l'opérateur qui remplace u_i par 0.

Ainsi, pour $f \in \mathbb{K}[[z, u_1, \ldots, u_n]]$,

$$\begin{cases} U_i f(z, u_1, \ldots, u_n) = f(z, u_1, \ldots, u_{i-1}, 1, u_{i+1}, \ldots, u_n), \\ Z_i f(z, u_1, \ldots, u_n) = f(z, u_1, \ldots, u_{i-1}, 0, u_{i+1}, \ldots, u_n), \\ \Delta_{0,i} f(z, u_1, \ldots, u_n) = f(z, u_1, \ldots, u_n) - Z_i f(z, u_1, \ldots, u_n), \\ \Delta_{1,i} f(z, u_1, \ldots, u_n) = f(z, u_1, \ldots, u_n) - U_i f(z, u_1, \ldots, u_n), \\ \partial_i f(z, u_1, \ldots, u_n) = \frac{\partial f}{\partial u_i}(z, u_1, \ldots, u_n), \\ S_i^n f(z, u_1, \ldots, u_n) = u_i^n f(z, u_1, \ldots, u_n). \end{cases}$$

Quand la variable sur laquelle on veut appliquer l'opérateur sera explicite, nous ne noterons pas l'indice i. On a donc $U = \sigma_0$ et $Z = \sigma_\infty$ (cette dernière égalité a sens non seulement en considérant la limite dans la topologie usuelle de l'anneau des polynômes mais aussi ponctuellement avec la topologie de $\mathbb{R}$ car les séries génératrices -non polynomiales- ont un rayon convergence ≤ 1, en effet une infinité de leurs coefficients sont ≥ 1). Remarquons en outre que $\mathrm{Id} = \sigma_1 = S^0$ et que $\partial_u^{-1} = \int_0^u$. Par ailleurs, le même jeu d'opérateurs appliqué à la variable z permet une attaque des marches non homogènes en temps (*i.e.* les marches pour lesquelles les pas possibles varient suivant le temps, voir l'exemple de la section 6 du chapitre 3).

Nous appelons "opérateur caractéristique de la marche" l'opérateur donné par

$$L := \eta \circ \psi.$$

Nous allons voir dans la suite du mémoire que c'est le fait d'arriver à coder la marche par un opérateur L linéaire qui permet d'obtenir aisément la série génératrice de la marche aléatoire.

3. Les jeux de récriture de Pinzani

Nous allons présenter dans cette section une conjecture qui a été à l'origine de notre intérêt pour les marches aléatoires. Renzo Pinzani & *al.*, dans [**2**], ont considéré des bijections entre des permutations "à motifs exclus" et une classe particulière d'arbres plans. Cela les a amenés à rencontrer des grammaires infinies dont la régularité des membres droits des productions implique l'algébricité (voire parfois la rationalité) de la série génératrice associée.

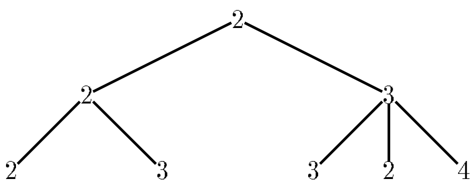
Définition 3.1 (Jeu de récriture de Pinzani)**.** *Soit R une fonction qui, à un entier k dans son domaine de définition, associe un multiplet de $\mathbb{N}^k$. Un jeu de récriture de Pinzani P (aussi appelé "arbre de possibilités") est un couple (i, R) où $i \in \mathbb{N}$ est l'élément initial et où R est l'ensemble des règles de récriture autorisées.*

On lie facilement un arbre à un tel ensemble de règles de récriture : la racine sera étiquetée par i et chaque nœud étiqueté par k aura k fils, étiquetés par les éléments de $R(k)$. Si k n'est pas dans le domaine de définition de R, alors il n'a pas de fils. Ces règles qui définissent un arbre de "possibilités" ou de "génération" sont aussi appelées avec humour[1] par l'équipe florentine des *eco*-systems (enumeration of combinatorial objects).

Par exemple, soit R définie sur $\{n \geq 2, n \in \mathbb{N}\}$ par

$$R(k) := (k, k-1, \ldots, 3, 2, k+1).$$

Prenons $i = 2$, on ainsi un arbre dont le début est :



[1]En italien "Eco" est un tic de langage équivalent à notre "tu vois ?"

Si l'on compte le nombre de nœuds se trouvant à la hauteur n (la racine étant à la hauteur 1), on obtient (voir la section 4.2 du chapitre 3 pour les calculs) les nombres de Catalan $1, 2, 5, 14\dots$ de série génératrice

$$\frac{1-\sqrt{1-4z}}{2z}.$$

Remarque. *Notons que les arbres de possibilités présentent le grand intérêt de procurer, par simple parcours d'arbre, un algorithme de génération aléatoire de basse complexité (au plus cubique à cause des pondérations des branches qui ne sont pas uniformes et qu'il faut donc 'rectifier').*

Peut-on dégager une classe de jeux de récriture menant ainsi à une série génératrice algébrique ? C'est ce que nous allons faire après la formalisation suivante. Sur l'alphabet infini des entiers strictement positifs, où (1) joue le rôle du mot vide, soit la grammaire infinie suivante :

- un entier $i \geq 1$ est choisi comme axiome ;
- on se fixe une suite $(c_n)_{n\in\mathbb{N}^*}$ d'entiers ≥ 1 puis un jeu infini de règles (productions)

  $$(k) \to (c_1)^{f_1(k)} \cdots (c_{m(k)})^{f_{m(k)}(k)} (b_1(k))^{e_1(k)} \cdots \big(b_{n(k)}(k)\big)^{e_{n(k)}(k)}$$

  où les exposants sont des entiers ≥ 0 et où les indices, les c_j et les b_j sont des entiers ≥ 1.

Les productions sont donc du type

$$(k) \to (c_1)^{f_1} \cdots (c_m)^{f_m} (b_1)^{e_1} \cdots (b_n)^{e_n}$$

où les b_i dépendent de k tandis que les c_i en sont indépendants.

Pinzani [**30**] a conjecturé que la série génératrice associée est

- rationnelle quand les $n(k)$ et les e_i sont bornés.

  Alain Denise montre dans [**8**] que la série génératrice est rationnelle si $\sum c_i f_i + \sum b_j e_j$ est linéaire en k.

  Exemple : $(k) \to 2^{k-1}(k+1)$ de série génératrice $\frac{1-z}{z^2-3z+1}$ (on obtient les nombres de Fibonacci de rang pair $1, 2, 5, 13\dots$, voir le détail des calculs à la section 4.3 du chapitre 3).
- algébrique quand les e_i sont bornés.

  Ce sont ces séries et les séries rationnelles du cas précédent que nous serons à même de traiter avec la méthode exposée ci-après.

  Exemple : $(k) \to (3)\dots(k-1)(k)(k+1)^2$ de série génératrice

  $$\frac{1-3z-\sqrt{1-6z+z^2}}{4z^2}$$

  (les premiers coefficients sont $1, 3, 11, 45, 197, \dots$, les nombres de Schroeder, voir le détail des calculs à la section 4.4 du chapitre 3).
- transcendante quand il n'y a pas de telles contraintes, par exemple quand on a des k^k.

  Exemple : $(k) \to (k)^{k-1}(k+1)$ donne les nombres de Bell (1,2,5,15,52...) de série génératrice exponentielle $\exp(\exp(x)-1)$.

  Il est clair que ces dernières règles ne peuvent pas aboutir à une série génératrice algébrique. En effet, l'algébricité implique une croissance géométrique des coefficients or, ici, les règles de type k^k imposent une croissance exponentielle des coefficients. Ainsi, la série génératrice (ordinaire) associée a un rayon de convergence nul. Danièle Gardy et Dominique Gouyou-Beauchamps donnent dans [**8**] un exemple de grammaire de Pinzani qui mène à série génératrice transcendante mais de rayon de convergence non nul

  $$(k) \to (2)^a (\lfloor \frac{k}{2} \rfloor)(\lfloor \frac{k}{2} \rfloor + 1)\dots(k+1)$$

  où a est tel qu'on a toujours k lettres dans le mot de droite.

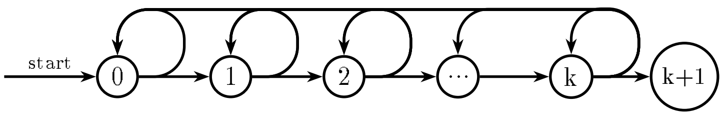
La conjecture de Pinzani sur l'algébricité de la série génératrice d'une grammaire du type $k \to (2)(3)\cdots(k)(k+1)$ (avec des modifications finies au deux extrémités) a été l'un des moteurs de notre intérêt pour les marches aléatoires.

En effet, nous allons maintenant voir comment ces grammaires ne sont qu'un cas particulier de marches aléatoires et nous montrerons qu'on aboutit *effectivement* à des séries génératrices algébriques.

4. Traduction en terme de marches aléatoires

Il se trouve que de telles règles de récriture (i, R) peuvent être englobées dans le problème plus général des marches aléatoires. Cette idée est due à Philippe Flajolet [**16**]. Dans le cas des règles de récriture de Pinzani, il s'agit plus précisément de marches aléatoires en dimension 1 avec conditions de bord. En effet, si l'on considère la "droite" des entiers, un jeu de récriture s'interprête de la façon suivante : de la position k, on peut aller aux positions $R(k)$. On peut même s'affranchir de la contrainte qui consiste à partir de i, en translatant le domaine de définition de R pour qu'il parte de l'origine. Plus précisément, on peut directement, à partir de la grammaire, appliquer la traduction suivante sur les membres droits : dans les exposants, k devient $k + i$ et les mantisses constantes sont décrémentées de i. La condition de bord consiste dans notre cas à nous interdire les entiers négatifs.

Ainsi, la marche aléatoire suivante, translatée de l'exemple donné plus haut, à savoir $(k) \to (2)(3)\cdots(k)(k+1)$ aura la traduction suivante en terme de marche aléatoire



On peut soit faire un pas en avant, soit revenir n'importe où en arrière, on obtient la série génératrice des nombres de Catalan (voir calculs à la section 4.2 du chapitre 3).

Nous voici donc ramenés à dénombrer les marches aléatoires partant de 0 et de longueur n. C'est ce que nous allons faire en utilisant une série génératrice bivariée, ce qui permet par la même occasion d'étudier d'autres propriétés, comme par exemple de dénombrer les excursions (*i.e.* les marches aléatoires qui partent de l'origine et qui y finissent). Nous introduisons la série génératrice

$$F(z, u) := \sum_{n,k} F_{n,k} u^k z^n$$

dénombrant le nombre $F_{n,k}$ de marches de longueur n (la longueur est ainsi marquée par z) partant de 0 et arrivant à k (la position est ainsi marquée par u).

Posons $f_n(u) := [z^n]F(z, u)$. Le nombre total de marches de longueur n est donc $f_n(1)$. Le nombre d'excursions de longueur n est $f_n(0)$.

Nos problèmes sont équivalents à l'étude des séries génératrices de langages donnés par un automate infini dont les états sont numérotés de 0 à $+\infty$ (ou de $-\infty$ à $+\infty$) et pour lesquels les transitions peuvent être "colorées", c'est-à-dire qu'on autorise les multiplicités. Tous les états sont terminaux pour la série génératrice des chemins et seul l'état initial est final pour les excursions. Un autre probème équivalent est l'étude des puissances itérées d'une matrice infinie à coefficients dans $\mathbb{N}$, dont les colonnes ne comportent qu'un nombre fini de termes non nuls.

Retournons aux problèmes des marches. Nous allons maintenant expliquer comment obtenir L, l'opérateur caractéristique de la marche, à partir des membres droits de la k-ième production des grammaires que nous considérons. Les mots de Pinzani sont des produits de mots du type $(ak + b)^{ck+d}$ avec a, b, c, d entiers et tels que les exposants soient ≥ 0 et les mantisses ≥ 1. On a le schéma de traduction suivant :

mot	codé dans la série génératrice par	action associée de L
$(k+b)$	u^{k+b}	S^b
$(k+b)^c$	cu^{k+b}	cS^b
$(ak+b)^c$	cu^{ak+b}	$cS^b\sigma_a$
$(k)^k$	ku^k	$S\partial$
$(k)^{k^n}$	$k^n u^k$	$(S\partial)^n$
$(b)^{ck}$	cku^b	$cS^bU\partial$
$(b)^{ck+d}$	$(ck+d)u^b$	$cS^bU\partial + dS^bU$
$(ak+b)^{ck+d}$	$(ck+d)u^{ak+b}$	$cS^b\sigma_a S\partial + dS^b\sigma_a$

Ainsi, l'ensemble des mots du dernier type pour a fixé donne au total l'opérateur

$$P_a(S)\sigma_a S\partial + Q_a(S)\sigma_a$$

(où P_a et Q_a sont des polynômes de Laurent explicites). On commettra souvent l'abus de notation consistant à écrire u à la place de l'opérateur de shift S et donc on considérera indifféremment les P_a et les Q_a comme des polynômes en S ou en u.

On peut ainsi toujours expliciter un opérateur **linéaire** (sur les monômes u^k)

$$L := \sum_a P_a(S)\sigma_a S\partial + Q_a(S)\sigma_a.$$

On peut alors obtenir la série génératrice par les techniques que nous allons maintenant détailler.

5. Obtention de l'équation fonctionnelle

L'équation fonctionnelle satisfaite par la série génératrice $F(z)$ s'obtient en écrivant, pour chaque n, l'égalité reliant le coefficient de z^n à celui de z^{n+1} et en sommant alors verticalement ces égalités (c'est ici que joue la linéarité de l'opérateur L). Détaillons ce procédé.

Avec l'aide de l'opérateur caractéristique L, on peut définir f_n par récurrence :

$$\begin{cases} f_0(u) = 1 \text{ (on pourrait choisir de partir d'une autre valeur initiale)} \\ f_{n+1}(u) = Lf_n(u) - r_n(u) \end{cases}$$

où $r_n(u)$ est la somme des monômes en u de degré négatif de $Lf_n(u)$, *i.e.* :

$$r_n(u) = \{u^{<0}\}Lf_n(u) := \sum_{j<0} u^j \left([u^j]\, Lf_n(u)\right).$$

comme L est linéaire, on a :

$$z^{n+1}f_{n+1}(u) = zLz^n f_n(u) - zz^n r_n(u)$$

en sommant sur n, on a :

$$\sum_{n\geq 0} z^{n+1}f_{n+1}(u) = z\sum_{n\geq 0} Lz^n f_n(u) - z\sum_{n\geq 0} z^n r_n(u)$$

en ajoutant $z^0 f_0(u) = 1$ aux deux membres, on obtient :

$$F(z,u) = 1 + zLF(z,u) - z\sum_{n\geq 0} z^n r_n(u)$$

On peut écrire L comme la somme d'un opérateur "différentiel" D (où interviennent des dérivées itérées par rapport à u ainsi que l'évaluation en $u = 1$ et en $u = 0$ et autres substitutions) et d'un deuxième opérateur P qui est en fait une multiplication par une fraction rationnelle en u. On a alors

$$F(z,u) = 1 + z((u) + D)F(z,u) - z\sum_{n\geq 0} z^n \sum_{j<0} u^j\,([u^j]\,(P+D)f_n(u))$$

$$(1)\quad F(z,u)\,(1 - zP(u)) = 1 + zDF(z,u) - z\sum_{n\geq 0} z^n \sum_{j<0} u^j\,([u^j]\,(P(u)+D)f_n(u))$$

Le membre de droite s'exprime alors en fonction de $\frac{\partial^i}{\partial u^i}F(z,0)$ et de $\frac{\partial^j}{\partial u^j}F(z,1)$ (ansi que d'autres valeurs du type $F(z,u^c)$ si l'on a utilisé des opérateurs mahleriens, ce qui n'est pas le cas pour les jeux de Pinzani).

6. La méthode du noyau

La méthode que nous utiliserons ici pour résoudre les équations fonctionnelles auxquelles nous venons d'aboutir et qui donnent la série génératrice à déterminer est connue par certains probabilistes sous le nom de "méthode du noyau" ; elle a par exemple été utilisée pour résoudre des équations fonctionnelles apparaissant dans des problèmes de files d'attente [**13**]. Elle fait partie du folklore et on en trouve les premières traces en combinatoire dans [**26**] et [**7**]).

Pour déterminer les fonctions inconnues du membre droit, la méthode du noyau consiste à annuler $1-zP(u)$, ainsi, en résolvant l'équation $1-zP(u) = 0$ par rapport à u, on obtient un certain nombre de solutions $u_1,\ldots,u_r$ que l'ont peut licitement substituer à u dans le membre de droite de l'équation (1), qui est alors nul lui aussi. On a ainsi un système qui, s'il est de rang plein, permet de déterminer les fonctions inconnues $\frac{\partial^i}{\partial u^i}F(z,0)$ et $\frac{\partial^j}{\partial u^j}F(z,1)$. On a alors totalement déterminé $F(z,u)$ qui est algébrique car exprimable en termes de fonctions algébriques.

7. Dégénerescence du noyau

Si on a insuffisament d'équations pour déterminer les inconnues (par exemple quand u n'apparaît pas dans le membre droit), on peut alors combler ce manque en dérivant l'équation fonctionnelle et en la regardant en $u = 1$ pour trouver de nouvelles relations.

Prenons l'exemple de $L := U\partial + (S+S^2+S^9)$, (qui est traité par une tout autre méthode dans [**8**] en tant que dénombrements de nœuds, c'est-à-dire quand $u = 1$, de hauteur n dans un arbre de possibilités). On aboutit à l'équation fonctionnelle

$$F(z,u)(1 - z(u+u^2+u^9)) = 1 + zF'(z,1).$$

Avec $u = 1$, on a la relation

$$F(z,1) = \frac{1+zF'(z,1)}{1-3z}.$$

Maintenant en dérivant l'équation fonctionnelle par rapport à u, on a

$$F'(z,u)(1 - z(u+u^2+u^9)) - F(z,u)z(1+2u+9u^8)) = 0.$$

Avec $u = 1$ on obtient cette fois

$$F'(z,1) = \frac{12zF(z,1)}{1-3z}.$$

Finalement on a

$$F(z,1) = \frac{3z-1}{3z^2+6z-1}$$

et on a du même coup une forme close pour F, à savoir

$$F(z,u) = \frac{1 - \frac{12z^2}{3z^2+6z-1}}{1 - z(u+u^2+u^9)}.$$

8. Un ensemble de marches rationnelles

Nous allons définir le poids d'une marche aléatoire d'opérateur caractéristique

$$L := \sum_{c\neq 0} P_c(S)\sigma_c + R(S)U\partial + Q(S)U$$

comme la fonction $p(k)$ définie par

$$p(k) := \sum_c (ckP_c(1) + P_c'(1)) + kR'(1) + Q'(1).$$

On rappelle que $\sigma_c(u^k) := u^{kc}$, que $U(u^k) = 1$ et que $\partial u^k = ku^{k-1}$.

Pour une grammaire de Pinzani $(k) \to (a_1)^{e_1}\cdots(a_k)^{e_k}$, le poids est donc la somme des $a_i e_i$ et on a alors un joli résultat dû à Alain Denise [**8**], à savoir que deux marches de Pinzani de même poids, si ce dernier est linéaire en k, ont la même série génératrice rationnelle $F(z,1)$.

Nous allons montrer comment obtenir un équivalent de ce théorème pour une classe très générale de marches. Nous allons regarder l'équation fonctionnelle

$$F(z,u) = f_0(u) + z\Big(\sum_c P_c(u)F(z,u^c) + R(u)\partial_u F(z,1)$$

$$+Q(u)F(z,1) - \{u^{<0}\}\sum_{n\geq 0} z^n Lf_n(u)\Big)$$

dans le cas $f_0(u) = 1$, $u = 1$ et quand $\{u^{<0}\}$ n'intervient pas On a donc

$$F(z,1) = 1 + z\Big(\sum_c P_c(1)F(z,1) + R(1)\partial_u F(z,1) + Q(1)F(z,1)\Big)$$

que l'on peut aussi écrire sous la forme

$$F(z,1)(1-za) = 1 + z\big(R(1)F'(z,1)\big)$$

Or en dérivant l'équation fonctionnelle on avait aussi

$$F'(z,u) = z\Big(\sum_c P_c(u)F'(z,u^c)cu^{c-1} + \sum_c P_c'(u)F(z,u^c)$$

$$+R'(u)F'(z,1) + Q'(u)F(z,1)\Big)$$

et donc, en $u=1$:

$$F'(z,1) = z\Big(\sum_c cP_c(1)F'(z,1) + \sum_c P_c'(1)F(z,1) + R'(1)F'(z,1) + Q'(1)F(z,1)\Big)$$

que l'on peut aussi écrire sous la forme

$$F'(z,1)(1-zb) = z\Big(\sum_c P_c'(1)F(z,1) + Q'(1)F(z,1)\Big)$$

On résout le système et finalement on aboutit à

$$F(z,1) = \frac{-1+z\alpha}{-1+z(\alpha+(1-z\alpha)\beta+zR(1)\beta')}$$

où

$$\alpha = \sum_c cP_c(1) + R'(1)$$

$$\beta = \sum_c P_c(1) + Q(1)$$

$$\beta' = \sum_c P_c'(1) + Q'(1)$$

Notons que $\alpha k + \beta'$ n'est rien d'autre le poids le marche. On voit par ailleurs clairement quelles relations imposer aux polynômes P_c, R et Q de deux marches différentes pour que ces dernières aient la même série génératrice, qui sera rationnelle dans le cas que nous avons traité ici. Sans doute les amoureux des preuves bijectives trouveront une autre démonstration de l'égalité de ces séries !

8.1. De la rationalité des nombres premiers... [section ajoutée en novembre 1998]

Nous allons justifier ce titre parodoxal en exhibant deux marches qui jouent avec cet ensemble “transcendant” que sont les nombres premiers mais qui auront néanmoins une série génératrice rationnelle.

La marche

$$\begin{cases} (k) \to (k+1)(k)(k+3)(2)^{k-3} \text{ pour } k \text{ non premier} \\ (k) \to (k+1)(k+1)(k+2)(2)^{k-3} \text{ pour } k \text{ premier} \end{cases}$$

parcourt trivialement l'ensemble des nombres entiers > 1 et a une règle légèrement modifiée pour les nombres premiers mais néanmoins, la série génératrice de cette marche est rationnelle et vaut $F(z,1) = \frac{z-3z^2}{1-5z+2z^2}$ (car le poids de cette marche est $p(k) = 5k - 2$).

Nous allons maintenant donner une règle de récriture qui satisfait aussi les contraintes d'un ecosystème et qui a pour composante connexe les nombres premiers.

Désignons par p_n le n-ième nombre premier, considérons maintenant la marche

$$\begin{cases} (2) \to (3)(5) \\ (3) \to (5)(5)(2) \\ (5) \to (7)^2(2)^3 \\ (7) \to (11)(7)(2)^4 \\ (11) \to (13)^2(2)^9 \\ (13) \to (17)(13)(2)^{11} \\ (17) \to (19)^2(2)^{15} \\ (19) \to (23)(19)(2)^{17} \\ (23) \to (29)(19)(3)(3)(2)^{19} \\ (29) \to (31)^2(2)^{27} \\ (31) \to (37)(29)(2)^{29} \\ (37) \to (41)(37)(2)^{35} \\ \text{ et pour } n > 12 : \\ (p_n) \to (p_{n+1})\ E\ (2)^{p_n - 39}(q_1)\dots(q_m) \end{cases}$$

où E est le mot $(5)^{12}$ (éventuellement légèrement modifié comme expliqué ci-dessous) et où $\{q_1, \dots, q_m\}$ est une partition en nombres premiers de $4p_n - p_{n+1} - 5.12 - 2(p_n - 39) = 2p_n - p_{n+1} + 16$ (c'est toujours un entier positif impair), un théorème de Vaughan permet alors d'affirmer que $m \leq 26$.

Ainsi la somme des membres droits vaut toujours $4p_n$ et quitte à changer quelques (5) en (3)(2) ou (5)(5) en $(2)^5$ quand la partition de Vaughan contient moins de 26 nombres, le nœud étiqueté p_n a bien $1 + 12 + p_n - 39 + 26 = p_n$ fils.

Finalement, cette marche parcourt exactement l'ensemble des nombres premiers (*i.e.* l'ensemble des nombres premiers est la composante connexe de cette marche) mais néanmoins, la série génératrice de cette marche est rationnelle et vaut $F(z,1) = \frac{z-2z^2}{1-4z}$ (car le poids de cette marche est $p(k) = 4k$).

Ceci est surprenant car l'ensemble des nombres premiers est “transcendant” mais la contrainte imposée ici sur le degré des noeuds de l'arbre et le fait que notre règle de récriture soit “hautement irrégulière” contrebalance cette transcendance.

Nous allons maintenant exploiter notre méthode sur diverses marches, en nous affranchissant des contraintes des grammaires de Pinzani, car celles-ci ne font que cacher des résultats un peu plus généraux. Pour les marches que nous allons considérer, la combinatoire analytique apparaît la plus apte à traiter un maximum de paramètres car elle code très bien la structure de ces marches et les séries génératrices obtenues contiennent donc en puissance quantité de renseignements, qui tombent par la simple application de théorèmes classiques d'analyse. Toutefois, certains paramètres comme le maximum s'avèrent revêches (les séries génératrices codant les maxima ne sont pas faciles à obtenir) et il peut s'avérer utile d'utiliser des outils purement probabilistes pour des études générales de tels paramètres ou bien des astuces combinatoires pour des cas très particuliers.

CHAPITRE 3

Applications

Nous allons commencer par étudier deux cas de figure pour lesquels il n'est pas nécessaire de recourir à la méthode du noyau et nous reviendrons ensuite aux marches pour lesquelles cette méthode est la plus rapide.

1. Marches dans un borné de $\mathbb{Z}^n$

Un premier exemple de marche aléatoire aisée à traiter est celui où on n'impose aucune condition de régularité, *i.e.* il peut y avoir autant de règles de déplacements que de points dans l'ensemble borné considéré. Toutefois, puisque nous sommes dans un borné, on a donc un ensemble fini de points et ainsi un ensemble fini de règles et il n'est pas très surprenant que l'on aboutisse alors à une série génératrice rationnelle comme le montre le théorème suivant.

Théorème 1.1. *La série génératrice d'une marche dans un borné de $\mathbb{Z}^n$ est rationnelle.*

Preuve. Un tel ensemble borné est fini, notons m son cardinal. On peut associer à ses m points les m états d'un automate sur un alphabet unaire. On a une transition entre deux états dès lors qu'il y a un pas entre les deux points correspondants, en préservant l'orientation. Si l'on veut compter les marches de longueur k partant d'un sous-ensemble fixé I du borné et finissant dans un sous-ensemble F du borné, il suffit de compter les mots de longueur k dans l'automate où sont pris comme états initiaux les états associés aux points de I et où sont pris comme états finaux les états associés aux points de F (en particulier, si l'on veut compter tous les mots de longueur k, tous les états sont finaux). Dans tous les cas, par le théorème de Chomsky-Schützenberger [**5**], le langage reconnu est rationnel ainsi que sa série génératrice, qui est, par construction de notre automate, la même que celle de la marche. $\square$

Pour de telles séries génératrices rationnelles, une asymptotique exacte se fait à l'aide de la proposition suivante :

Proposition 1.2. *Si F est une fonction rationnelle dont les pôles $\rho_1, \ldots, \rho_m$ sont $\neq 0$, alors il existe m polynômes $P_1, \ldots, P_m$ (où le degré de P_j est l'ordre du pôle ρ_j moins 1) tels que*

$$F_n = [z^n]F(z) = \sum_{j=1}^{m} P_j(n)\rho_j^{-n}.$$

Preuve. On utilise la décomposition en éléments simples de $F(z)$ puis que

$$[z^n](z-\rho)^{-k} = \frac{(-1)^k}{a^{n+k}}\binom{n+k-1}{k-1}$$

ce qui donne un polynôme de degré $k-1$ en n. $\square$

Si l'on ne se limite pas à un terrain borné mais qu'on ne s'autorise que des pas bornés, on tombe également sur une série génératrice des marches rationnelle (la série génératrice des excursions sera quant à elle algébrique si l'on s'autorise des retours en arrière) comme nous allons le voir dans la section suivante.

2. Marches à pas bornés

Pour de telles marches, on a directement $f_n(u) = L^n f_0(u)$, où L est l'opérateur caractéristique. Ainsi, si L est la multiplication par un polynôme (de Laurent, *i.e.* un polynôme pour lequel on s'autorise les degrés négatifs) P en u, c'est-à-dire si l'ensemble des pas autorisés est codé par un polynôme $P(u)$, on a :

$$f_{n+1}(u) = P(u)f_n(u) = P(u)^n f_0(u)$$

et donc, quand $f_0(u) = 1$ (*i.e.* quand on part de l'origine),

$$F(z,u) = \frac{1}{1-P(u)z}.$$

En particulier, la série génératrice des marches est

$$F(z,1) = \frac{1}{1-P(1)z}.$$

Proposition 2.1. *Une marche à pas bornés homogène en espace et sans retour en arrière a une série génératrice des excursions rationnelle.*

Preuve. Il suffit d'évaluer $F(z,u)$ en $u = 0$ et comme $P(0) \neq \infty$ (puisque nous n'avons pas de retour en arrière), la série génératrice des excursions est

$$\frac{1}{1-P(0)z}.$$

Bien sûr, elle se réduit trivialement à $F(z,0) = 1$ si on n'a pas de retour sur place. □

Notons que la singularité dominante est $P(1)^{-1}$ pour les marches et $P(0)^{-1}$ pour les excursions. Ce sont des nombres rationnels. En fait, $P(1)^{-1}$ demeurera une singularité dans les marches des sections suivantes, mais elle ne restera pas forcément dominante.

Dans le cas moins trivial où l'on a non pas que des pas stationnaires ou vers la droite mais également des pas vers la gauche, la série génératrice des excursions devient algébrique :

Proposition 2.2. *Une marche à pas bornés homogène en espace et avec des retours en arrière a une série génératrice des excursions algébrique.*

Preuve. La relation de Cauchy sur les coefficients d'une fonction analytique donne :

$$F(z,0) = [u^0]F(z,u) = [u^0]\frac{1}{1-zP(u)} = \frac{1}{2i\pi}\int_\gamma \frac{F(z,u)}{u}du$$

où γ est une courbe simple contenant 0 et comprise dans un voisinage de l'origine suffisament petit. D'après le théorème des résidus, on a alors

$$F(z,0) = \sum_{u_i \text{ pôle}} \operatorname{Res}\left(\frac{F}{u}, u_i\right) = \sum_{u_i \text{ pôle}} \operatorname{Res}\left(\frac{1}{u-uzP(u)}, u_i\right)$$

Comme les pôles sont d'ordre 1, les résidus sont égaux aux limites suivantes

$$F(z,0) = \sum_{u_i \text{ pôle}} \lim_{u\to u_i} \frac{u-u_i}{u-uzP(u)} = \sum_{u_i \text{ pôle}} \lim_{u\to u_i} \frac{1}{1-z(uP'(u)+P(u))}.$$

La dernière égalité provient de l'application de la règle de l'Hospital. Comme par ailleurs $P(u_i) = \frac{1}{z}$, on a donc

$$F(z,0) = \sum_{u_i \text{ pôle}} \frac{1}{-zu_iP'(u_i)}.$$

Philippe Flajolet m'a indiqué ce résultat en notant sa similitude avec l'assertion de Furstenberg [**20**] : la série obtenue en prenant les coefficients diagonaux d'une série multivariée rationnelle est algébrique (en fait l'article cité a surtout le mérite de montrer que la rationalité est conservée quand le corps de base est fini). □

2.1. Exemple. Pour la marche de "Dyck" où l'on s'autorise un pas en avant et un pas en arrière, on a donc $P(u) = u + \frac{1}{u}$. Pour les marches, on aboutit à

$$F_{\mathbb{Z}}(z,1) = \frac{1}{1-P(1)z} = \frac{1}{1-2z} = 1 + 2z + 4z^2 + 8z^3 + O(z^4)$$

et pour les excursions, on n'a pas la formule rationnelle puisque $P(0) = \infty$. En fait, on peut s'en tirer avec les manipulations suivantes :

$$F_{\mathbb{Z}}(z,0) = \sum_{n\geq 0} [u^0] P(u)^n z^n = \sum_{n \geq 0} [u^0] \sum_k \binom{n}{k} u^{-k} u^{n-k} z^n$$

$$= \sum_{n\geq 0} \binom{2n}{n} z^{2n} = \frac{1}{\sqrt{1-4z^2}} = 1 + 2z^2 + 6z^4 + 20z^6 + O(z^8).$$

Mais il reste bien sûr préfèrable et surtout plus rapide d'appliquer la formule donnée à la proposition précédente, à savoir $F(z,0) = \frac{1}{-zu_iP'(u_i)}$ avec u_i la racine nulle en 0 de $1 - zP(u) = 0$.

Nous allons maintenant utiliser la méthode du noyau pour regarder ce que l'on obtient lorsque l'on s'interdit les entiers négatifs, c'est ce que nous appellerons une "marche positive".

3. Marches positives à pas bornés

Soit c la taille du plus grand pas en arrière admis, lorsque $c = 0$, on a en fait affaire à des marches qui ne reviennent jamais en arrière, la contrainte de positivité ne joue pas, ainsi la section précédente justifie que nous supposions $c > 0$ pour le restant de cette section. Nous notons $\{u^{<0}\}$ pour $\sum_{j<0} u^j [u^j]$. De la relation

$$f_{n+1}(u) = Lf_n(u) - \{u^{<0}\} Lf_n(u),$$

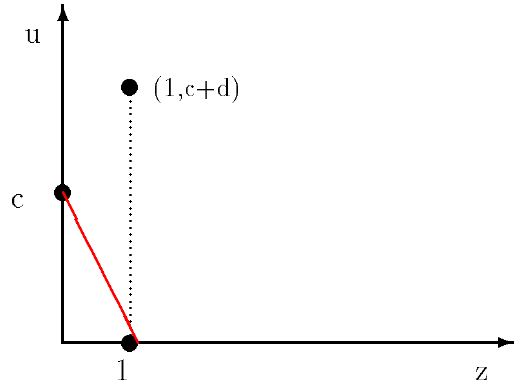
en multipliant par z^{n+1}, puis en sommant pour n allant de 0 à $+\infty$, et enfin en remarquant que l'on peut écrire $\{u^{<0}\}Lf_n(u) = \sum_{j=0}^{c-1} c_j(u) \partial_u^j f_n(0)$ (où $c_j(u)$ est un polynôme de Laurent dont les monômes ont des degrés compris entre $j-c$ et -1), on aboutit à l'équation

$$F(z,u)\big(1 - zP(u)\big) = 1 - z \sum_{j=0}^{c-1} c_j(u) \partial_u^j F(z,0). \tag{2}$$

Nous allons définir le noyau N comme la fonction

$$N(u,z) := u^c - zu^c P(u).$$

Lorsque $N(u,z) = 0$, en utilisant le polygone de Newton pour développer les solutions en série de Puiseux, on obtient $c + d$ solutions $u_1, \ldots, u_{c+d}$.



Notation. *Le diagramme de Newton ci-dessus nous permet d'affirmer l'existence*

- *de c petites branches, équivalentes à $z^{1/c}$ en $z = 0$, que nous noterons $u_1, \ldots, u_c$;*
- *et de d branches infinies en 0, équivalentes à $z^{-1/d}$ en $z = 0$, que nous noterons $v_1, \ldots, v_d = u_{c+1}, \ldots, u_{c+d}$.*

Posons

$$Q(u,z) := u^c\big(1 - z\sum_{j=0}^{c-1} c_j(u)\partial_u^j F(z,0)\big).$$

En général, la stratégie de la méthode du noyau se poursuit en substituant u par les c petites branches, on obtient ainsi un système

$$\begin{cases} u_1^c\big(1 - z\sum_{j=0}^{c-1} c_j(u_1)\partial_u^j F(z,0)\big) = 0 \\ \vdots \\ u_i^c\big(1 - z\sum_{j=0}^{c-1} c_j(u_i)\partial_u^j F(z,0)\big) = 0 \\ \vdots \\ u_c^c\big(1 - z\sum_{j=0}^{c-1} c_j(u_c)\partial_u^j F(z,0)\big) = 0 \end{cases}$$

qui permet de déterminer les c inconnues $\partial_u^j F(z,0)$ pour $j = 0 \ldots c-1$.

Une astuce due à Mireille Bousquet-Mélou [**4**] permet de sauter la fastidieuse résolution d'un tel système. En regardant N et Q comme des polynômes en u et en remplaçant u par u_i (pour $i = 1, \ldots, c$) dans l'équation (2), on voit que les racines $u_1, \ldots, u_c$ de N sont également racines de $Q(u,z)$. Comme le polynôme Q est de degré c en u et puisqu'il est unitaire, on a la factorisation

$$Q(u,z) = \prod_{i=1}^{c}(u - u_i).$$

On a également, avec $s_d := [u^d]P(u)$ (*i.e.* la multiplicité du plus grand pas en avant) :

$$N(u,z) = -zs_d\prod_{i=1}^{c+d}(u - u_i).$$

Ainsi

$$F(z,u) = \frac{Q(u,z)}{N(u,z)} = \frac{\prod_{i=1}^{c}(u-u_i)}{-zs_d\prod_{j=1}^{c+d}(u-u_j)} = \frac{1}{-zs_d\prod_{i=1}^{d}(u-v_i)}$$

et on a prouvé l'algébricité de la série génératrice F. Notons $s_c := [u^{-c}]P(u)$ (s_c est donc la multiplicité du plus grand retour en arrière), on obtient les deux théorèmes suivants :

Théorème 3.1 (Excursions positives à pas bornés). *Une marche aléatoire positive à pas bornés a une série génératrice des excursions algébrique et cette dernière vaut*

$$\frac{(-1)^{c+1}}{s_c z}\prod_{i=1}^{c} u_i$$

Preuve. Pour dénombrer les excursions, c'est-à-dire, rappelons-le, les marches qui finissent en 0, on regarde $[u^0]F(z,u)$, *i.e.* $F(z,0)$. On a donc

$$F(z,0) = \frac{Q(0,z)}{N(0,z)} = \frac{Q(0,z)}{[u^0]N(u,z)} = \frac{Q(0,z)}{[u^0]u^c - z[u^0]u^cP(u)}$$

$$= \frac{Q(0,z)}{-s_c z} = \frac{\prod_{i=1}^{c}(0-u_i)}{-s_c z} = \frac{(-1)^{c+1}}{s_c z}\prod_{i=1}^{c} u_i = \frac{(-1)^{d+1}}{zs_d\prod_{i=1}^{d} v_i}.$$

□

Théorème 3.2 (Marches positives à pas bornés). *Une marche aléatoire positive à pas bornés a une série génératrice des marches algébrique et elle vaut*

$$\frac{1}{1 - zP(1)}\prod_{i=1}^{c}(1 - u_i)$$

Preuve. Pour dénombrer les marches, on regarde $\sum_k [u^k]F(z,u)$, *i.e.* $F(z,1)$. Ainsi, on obtient :

$$F(z,1) = \frac{Q(1,z)}{N(1,z)} = \frac{1 - z\sum_{j=0}^{c-1} c_j(1)\partial_u^j F(z,0)}{1 - zP(1)}$$

$$= \frac{1}{1 - zP(1)} \prod_{i=1}^{c} (1 - u_i) = \frac{1}{-zs_d \prod_{i=1}^{d} (1 - v_i)}.$$

□

Ce résultat a été obtenu en mars 1998, indépendamment de Petkovšek [**29**].

3.1. Exemple : la marche de Dyck. La plus simple de ces marches, qui présente néanmoins son intérêt, est la marche de "Dyck" que nous avons déjà regardée quand nous ne nous limitions pas aux positifs. Rappelons qu'on s'autorise un pas en avant et un pas en arrière, on a donc le polynôme caractéristique $P(u) := u^{-1} + u$. Or $N(u, z) = u - z(u^2 + 1)$ et $N(u, z) = 0$ conduit à

$$\begin{cases} u_1 = \frac{1-\sqrt{1-4z^2}}{2z} = z + z^3 + 2z^5 + 5z^7 + O(z^9) \\ v_1 = \frac{1+\sqrt{1-4z^2}}{2z} = z^{-1} - z - z^3 - 2z^5 - 5z^7 + O(z^9). \end{cases}$$

On a donc $F(z, u) = \frac{-1}{z(u-v_1)}$ et, pour la série génératrice des excursions, on a $F(z, 0) = \frac{u_1}{z}$ ou, avec les branches divergentes, $F(z, 0) = \frac{1}{zv_1}$.

Attardons-nous sur la singularité en $z = \frac{1}{2}$ pour $F(z, 1)$ et $F(z, 0)$. En effet, on a le paradoxe suivant, que l'on considère au choix

- la marche de Dyck usuelle (*i.e.* sans s'interdire les entiers négatifs), dont la série génératrice est
$$F_{\mathbb{Z}}(z, 1) = \frac{1}{1 - P(1)z} = \frac{1}{1 - 2z} = 1 + 2z + 4z^2 + 8z^3 + O(z^4),$$
- les excursions de Dyck, dont la série génératrice est
$$F_{\mathbb{Z}}(z, 0) = \frac{1}{\sqrt{1 - 4z^2}} = 1 + 2z^2 + 6z^4 + 20z^6 + O(z^8),$$
- la marche de Dyck positive, dont la série génératrice est
$$F_{\mathbb{N}}(z, 1) = \frac{2}{1 - 2z + \sqrt{1 - 4z^2}} = 1 + z + 2z^2 + 3z^3 + 6z^4 + 10z^5 + O(z^6),$$
- les excursions de Dyck positives, dont la série génératrice est
$$F_{\mathbb{N}}(z, 0) = \frac{1 - \sqrt{1 - 4z^2}}{2z^2} = 1 + z^2 + 2z^4 + 5z^6 + 14z^8 + O(z^{10}),$$

il est curieux qu'il existe à peu près "autant" (c'est-à dire asymptotiquement 2^n) de marches de chaque type. Par exemple, dans le dernier cas, on a des contraintes très fortes (revenir en 0 et rester > 0) alors que dans le premier on s'autorise toute la droite et la marche peut s'arrêter n'importe où. En fait, on a respectivement, pour les quatre cas envisagés, une asymptotique (avec n pair pour les cas appropriés) en

$$2^n, \sqrt{2}\frac{2^n}{\sqrt{\pi n}}, \frac{1}{\sqrt{2}}\frac{2^n}{\sqrt{\pi n}}, 2\sqrt{2}\frac{2^n}{\sqrt{\pi n^3}},$$

ce qui précise le poids des contraintes imposées.

4. Marches positives à pas avant bornés et arrière non bornés

Un avantage indéniable de la combinatoire analytique pour l'étude des séries génératrices de marches aléatoires est sa souplesse d'adaptation à des variations de règles du jeu. Les méthodes employées demeurent en effet souvent identiques ! Nous allons maintenant illustrer cette souplesse en nous autorisant non plus seulement des pas "bornés" mais également des retours en arrière de longueur arbitraire.

La méthode du noyau va ici marcher car nous pouvons coder dans notre jeu d'opérateurs (en fait en passant à l'espace de Banach de l'espace des shifts) tous ces retours en arrière. En effet, si de la position k, on peut aller aux positions $0, 1, \ldots, k-1$, on doit avoir un opérateur L, qui à u^k, doit associer $u^0 + u^1 + \cdots + u^{k-1}$, *i.e.* $\frac{1-u^k}{1-u} = \frac{u^0}{1-u} - \frac{u^k}{1-u}$; ainsi

$$L := \frac{1}{1-u} U - \frac{1}{1-u} S^0.$$

On peut maintenant facilement donner l'opérateur caractéristique d'une marche dans laquelle on s'on s'autorise tous les retours en arrière (sauf éventuellement quelques-uns des premiers pas en arrière, que nous coderons par un polynôme $A(u)$), plus quelques pas bornés vers l'avant (que nous coderons par un polynôme $B(u)$), en effet en notant $P(u) := B(u) - A(u)$, on a

$$L := \frac{1}{1-u}U - \frac{1}{1-u} + P(u).$$

De la relation

$$f_{n+1}(u) = Lf_n(u) - \{u^{<0}\}Lf_n(u),$$

en multipliant par z^{n+1}, puis en sommant pour n allant de 0 à $+\infty$, on arrive à

$$F(z,u) = 1 + z\big(\frac{F(z,1)}{1-u} - \frac{F(z,u)}{1-u} + P(u)F(z,u) - \{u^{<0}\}\sum_{n\geq 0} z^n Lf_n(u)\big)$$

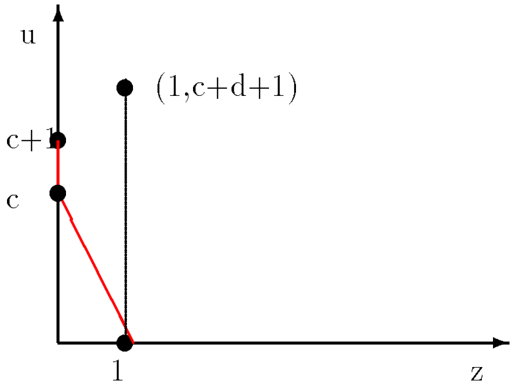
et enfin en remarquant que l'on peut écrire $\{u^{<0}\}Lf_n(u) = \sum_{j=0}^{c-1} c_j(u)\partial_u^j f_n(0)$ (où $c_j(u)$ est un polynôme de Laurent dont les monômes ont des degrés compris entre $j-c$ et -1), on aboutit à l'équation

$$(3) \qquad F(z,u)\big(1 + \frac{z}{1-u} - zP(u)\big) = 1 + \frac{z}{1-u}F(z,1) - z\sum_{j=0}^{c-1} c_j(u)\partial_u^j F(z,0).$$

Définissons le noyau N comme la fonction

$$N(u,z) := -u^c(1-u) - u^c z + zu^c(1-u)P(u).$$

Lorsque $N(u,z) = 0$, en utilisant le polygone de Newton pour développer les solutions en série de Puiseux, on obtient $c+d+1$ solutions $u_0, \ldots, u_{c+d}$.



Notation. *Le diagramme de Newton ci-dessus nous permet d'affirmer l'existence*

- *d'une branche équivalente à 1 en $z = 0$, que nous noterons u_0 ;*
- *de c petites branches, équivalentes à $z^{1/c}$ en $z = 0$, que nous noterons $u_1, \ldots, u_c$;*
- *et de d branches infinies en 0, équivalentes à $z^{-1/d}$ en $z = 0$, que nous noterons $v_1, \ldots, v_d = u_{c+1}, \ldots, u_{c+d}$.*

Posons

$$Q(u,z) := -u^c(1-u)\big(1 + \frac{z}{1-u}F(z,1) - z\sum_{j=0}^{c-1} c_j(u)\partial_u^j F(z,0)\big).$$

Comme pour les marches à pas avant et arrière bornés, on applique l'astuce qui consiste à factoriser N et Q :

$$Q(u,z) = \prod_{i=0}^{c}(u-u_i) \quad \text{et} \quad N(u,z) = -zs_d\prod_{i=0}^{c+d}(u-u_i),$$

ce qui donne

$$F(z,u) = \frac{Q(u,z)}{N(u,z)} = \frac{\prod_{i=0}^{c}(u-u_i)}{-zs_d\prod_{j=0}^{c+d}(u-u_j)} = \frac{1}{-zs_d\prod_{i=1}^{d}(u-v_i)}.$$

On a donc, dans ce cas aussi, une série génératrice algébrique.

On conserve la notation $s_d := [u^d]P(u)$ (s_d est donc la multiplicité du plus grand pas en avant) et on obtient les deux théorèmes suivants :

Théorème 4.1 (Excursions positives à pas arrière non bornés). *Une marche aléatoire positive à pas arrière non bornés et à pas avant bornés a une série génératrice des excursions algébrique et elle vaut*

$$\frac{(-1)^{c+1}}{s_c z}\prod_{i=0}^{c} u_i.$$

Preuve. Pour dénombrer les excursions, on regarde $F(z,0)$ et on obtient :

$$F(z,0) = \frac{Q(0,z)}{N(0,z)} = \frac{\prod_{i=0}^{c}(0-u_i)}{[u^0]N(u,z)}$$
$$= \frac{(-1)^{c+1}\prod_{i=0}^{c} u_i}{[u^0]\big(-u^c(1-u) - u^c z + zu^c(1-u)P(u)\big)} = \frac{(-1)^{c+1}\prod_{i=0}^{c} u_i}{[u^0]zu^cP(u)}$$
$$= \frac{(-1)^{c+1}}{s_c z}\prod_{i=0}^{c} u_i = \frac{(-1)^{d+1}}{zs_d\prod_{i=1}^{d} v_i}.$$

□

Théorème 4.2 (Marches positives à pas arrière non bornés). *Une marche aléatoire positive à pas arrière non bornés et à pas avant bornés a une série génératrice des marches algébrique et elle vaut*

$$-\frac{1}{z}\prod_{i=0}^{c}(1-u_i).$$

Preuve. Pour dénombrer les marches, on regarde $F(z,1)$. Ainsi, on obtient, car $N(1,z) = -z$:

$$F(z,1) = \frac{Q(1,z)}{N(1,z)} = -\frac{1}{z}\prod_{i=0}^{c}(1-u_i) = \frac{1}{-zs_d\prod_{i=1}^{d}(1-v_i)}.$$

□

4.1. Exemple : la marche de Motzkin. La plus simple de ces marches serait la marche de “Motzkin” où l'on s'autorise un pas de longueur 1 en avant et tous les pas en arrière (rester sur place n'est pas considéré comme un pas en arrière), elle a pour opérateur caractéristique

$$L = \frac{-1}{1-u} + u + \frac{1}{1-u}U,$$

où U est l'opérateur de substitution $u = 1$. Ainsi

$$N(u,z) = -(1-u) - z + z(1-u)u,$$

ce qui conduit à

$$u_0 = \frac{1+z-\sqrt{1-2z-3z^2}}{2z} \quad \text{et} \quad v_1 = \frac{1+z+\sqrt{1-2z-3z^2}}{2z}.$$

On a donc $F(z,u) = \frac{1}{u-v_1}$ et pour la série génératrice des excursions, on a au choix $F(z,0) = \frac{u_0}{z}$ ou, avec les branches divergentes, $F(z,0) = \frac{-1}{v_1}$.

Attardons-nous sur la singularité en $z = 1/3$ pour $F(z,1)$ et $F(z,0)$. Ici, il n'y a pas lieu de distinguer quatre cas puisque la condition de bord ne joue pas. Que l'on considère

- les marches de Motzkin de série génératrice

$$F(z,1) = \frac{1-z-\sqrt{1-2z-3z^2}}{2z^2} = 1 + z + 2z^2 + 4z^3 + 9z^4 + 21z^5 + O(z^6),$$

- les excursions de Motzkin de série génératrice

$$F(z,0) = \frac{1+z-\sqrt{1-2z-3z^2}}{2z(1+z)} = 1 + z^2 + z^3 + 3z^4 + 6z^5 + O(z^6),$$

on aboutit à une asymptotique en $\frac{3^n}{n^{3/2}}$, “paradoxe” que nous avons déjà expliqué dans la section sur les marches positives à pas bornés.

4.2. La marche de Catalan. Avec la grammaire $(k) \to (2)(3)\cdots(k)(k+1)$, on a l'opérateur

$$L = \frac{-1}{1-u} + \frac{1}{1-u}U + 1 + u,$$

ainsi $N(u,z) = -(1-u) - z + z(1-u)(1+u)$ d'où $v_1(z) = \frac{1+\sqrt{1-4z}}{2z}$ et donc

$$F(1,z) = \frac{1}{-z(1-v_1)} = 1 + 2z + 5z^2 + 14z^3 + 42z^4 + 132z^5 + O(z^6).$$

4.3. Marche de Fibonacci. Si l'on regarde maintenant la marche liée à la grammaire $(k) \to 2^{k-1}(k+1)$, c'est-à-dire à l'opérateur $L := S + U\partial + U$, on aboutit à l'équation fonctionnelle

$$F(z,u)(1-zu) = 1 + z\big(F(z,1) + \partial_u F(z,1)\big)$$

avec $u = 1$ on obtient

$$\partial_u F(z,1) = \frac{F(z,1)(1-2z) - 1}{z}$$

par ailleurs, en dérivant l'équation fonctionnelle par rapport à u, on a

$$(1-zu)\partial_u F(z,u) - zF(z,u) = 0$$

d'où

$$\partial_u F(z,u) = \frac{zF(z,u)}{1-zu}$$

avec $u = 1$ on obtient de plus que

$$\partial_u F(z,1) = \frac{zF(z,1)}{1-z} = \frac{F(z,1)(1-2z) - 1}{z}$$

et donc

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

On a ainsi, pour les marches, la série génératrice des nombres de Fibonacci de rang pair $(1,2,5,13\ldots)$, résultat déjà obtenu par Dominique Gouyou-Beauchamps [**22**]. De plus, on a déterminé F

$$F(z,u) = \frac{z^2 - z^3u - zu - 2z + 3z^2u + 1}{(z^2 - 3z + 1)(1-zu)}$$

ce qui avec, $u = 0$ mène à la série génératrice des excursions

$$F(z,0) = \frac{z^2 - 2z + 1}{z^2 - 3z + 1}$$

qui génère les nombres de Fibonacci de rang impair $(1,1,3,8,21,\ldots)$.

4.4. Marche de Schroeder. Si l'on regarde cette fois la marche liée à la grammaire

$$(k) \to (3)\ldots(k-1)(k)(k+1)^2$$

c'est-à-dire à l'opérateur

$$L := 2S + S^0 + U\partial + U$$

On a donc

$$F(z,u)(1 - z(2u+1)) = 1 + z(F'(z,1) + F(z,1))$$

avec $u = 1$, on obtient

$$F(z,1)(1-3z) = 1 + z(F'(z,1) + F(z,1))$$

en dérivant par rapport à u et en résolvant le système, on trouve

$$F(z,1) = \frac{1 - 3z - \sqrt{1 - 6z + z^2}}{4z^2}$$

les premiers coefficients sont $1, 3, 11, 45, 197, \ldots$ *i.e.* les nombres de Schroeder.

5. Marches régulières définies modulo un pavé

Cette fois, au lieu de considérer une marche pour laquelle les possibilités de sauts sont toujours les mêmes, quelque soit le point où l'on est, nous allons nous autoriser de varier de règles (*i.e.* d'opérateur caractéristique) suivant les points, toutefois, nous n'allons nous autoriser qu'un nombre fini de variations, réparties de façon uniforme sur notre réseau. Ceci constitue une généralisation de l'exemple suivant en une dimension : “si je suis dans un état pair, je peux faire cela, si je suis dans un état impair, je peux faire ceci.” On a donc affaire à des marches inhomogènes en espace mais néanmoins périodiquement homogènes.

5.1. Méthode combinatoire par agrégation d'états. Notons que ce type de marche peut faire l'objet du traitement “combinatoire” suivant : on regroupe les états par paquets, *i.e.* à chaque pavé on associe un état et on autorise les sauts correspondants.

Cela permet d'obtenir une série génératrice univariée donnant le nombre de chemins de longueur n mais malheureusement, on perd ainsi la possibilité d'obtenir une série bivariée par rapport à la longueur et la position finale (exacte et non pas modulo le pavé) des chemins.

5.2. Méthode analytique par récurrences croisées. Il est parfois possible d'obtenir une série bivariée exacte, on n'utilise bien sûr plus l'agrégation d'états précédente mais des récurrences croisées.

Fixons n nombres $k_1, \dots, k_n$ et considérons une marche dans $\mathbb{Z}^n$ où les sauts sont différents suivant les classes d'équivalences modulo $(k_1, \dots, k_n)$.

Une méthode de résolution consiste à considérer les $(k_1 k_2 \cdot k_n)^{k_1 k_2 \cdots k_n}$ opérateurs qui, aux points d'une classe, associent les points de telle classe obtenus après les sauts.

Par exemple, pour $n = 1$ et $k_1 = 2$, on construit les 4 opérateurs

$$L_{p \to p}, L_{p \to i}, L_{i \to p}, L_{i \to i}$$

qui décrivent quels sont les sauts autorisés allant d'un point de coordonnée paire à un point de coordonnée paire, d'un point de coordonnée paire à un point de coordonnée impaire, etc.

En considérant alors les deux suites $f_n^{[p]}(u)$ (resp. $f_n^{[i]}(u)$) des points à coordonnée paire (resp. impaire) que l'on atteint au bout de n étapes, on a la récurrence :

$$\begin{cases} f_{n+1}^{[p]} = L_{p \to p} f_n^{[p]} + L_{i \to p} f_n^{[i]} \\ f_{n+1}^{[i]} = L_{i \to i} f_n^{[i]} + L_{p \to i} f_n^{[p]} \end{cases}$$

Quand par exemple L est un polynôme, l'obtention de la série génératrice ne pose pas de problèmes particuliers.

6. Marches non homogènes en temps

Jusqu'à présent, nous nous sommes autorisé des pas qui dépendent éventuellement du lieu où l'on est (c'est-à-dire des cas inhomogènes en espace). On peut également prendre en compte la variable temps, *i.e.* z.

Nous allons regarder la marche positive où, de k au temps n, on peut aller en $k-2$ avec la probabilité $\frac{n}{n+1}$ ou en $k+1$ avec la probabilité $\frac{1}{n+1}$.

Écrivons

$$F(z,u) = \sum_{n \geq 0} F_n(z,u).$$

On a donc $F_{n+1}(z,u) = zLF_n(z,u) - z\{u^{<0}\}LF_n(z,u)$ avec

$$L = \frac{1}{u^2} U_z \partial_z^{-1} z \partial_z + u U_z \partial_z^{-1}.$$

i.e

$$F_{n+1}(z,u) = zLF_n(z,u) - z\left(\frac{1}{u}[u^{-1}] + \frac{1}{u^2}[u^{-2}]\right) \frac{1}{u^2} U_z \partial_z^{-1} z \partial_z F_n(z,u)$$

$$F_{n+1}(z,u) = zLF_n(z,u) - z\left(\frac{1}{u}[u^{1}] + \frac{1}{u^2}[u^{0}]\right) U_z \partial_z^{-1} z \partial_z F_n(z,u)$$

$$F_{n+1}(z,u) = zLF_n(z,u) - z\left(\frac{1}{u}[u^1] + \frac{1}{u^2}[u^0]\right)\int_0^1 z\partial_z F_n(z,u)dz$$

$$F(z,u) = 1 + \frac{z}{u^2}\int_0^1 z\partial_z F(z,u)dz + uz\int_0^1 F(z,u)dz - \frac{z}{u}\int_0^1 z\partial_z\partial_u F(z,0)dz$$
$$-\frac{z}{u^2}\int_0^1 z\partial_z F(z,0)dz$$
$$F(z,u) = 1 + \frac{z}{u^2}\left(F(1,u) - \int_0^1 F(z,u)dz\right) + uz\int_0^1 F(z,u)dz$$
$$-\frac{z}{u}\left(\partial_u F(1,0) - \int_0^1 \partial_u F(z,0)\right)dz - \frac{z}{u^2}\left(F(1,0) - \int_0^1 F(z,0)dz\right)$$

En tenant compte des instants initiaux, on a donc, puisque $F(1,0)=1$,

$$F(z,u) = 1 + \frac{z}{u^2}\left(F(1,u) - \int_0^1 F(z,u)dz\right) + uz\int_0^1 F(z,u)dz$$
$$-\frac{z}{u}\left(\partial_u F(1,0) - \int_0^1 \partial_u F(z,0)\right)dz - \frac{z}{u^2}\left(1 - \int_0^1 F(z,0)dz\right)$$

7. La partie de bonneteau progressive

Nous allons donner un exemple de marche non homogène en espace. Prenons par exemple la marche unidimensionnelle qui, lorsque l'on est en k, va en $k+1$ avec la probabilité $\frac{1}{k+1}$ et en 0 avec la probabilité $\frac{k}{k+1}$.

Ce schèma correspond à une "partie de bonneteau" : un magicien introduit sous vos yeux une pièce dans un des k gobelets opaques qu'il va manipuler. Il fait passer la pièce très rapidement d'un gobelet à l'autre. Il vous demande à la fin où se trouve la pièce. Il est pervers et il a très bien pu ne pas conserver la pièce dans l'un des gobelets. En général, le manipulateur est si habile, que vous n'avez pas de meilleure stratégie que de faire une proposition au hasard ! Si vous gagnez, on continue avec un gobelet supplémentaire, sinon on reprend au début avec un seul gobelet. Voyons maintenant à quelle série génératrice ce petit jeu nous fait aboutir.

On a donc
$$L = U\partial^{-1}S\partial + \partial^{-1}$$
ce qui mène à l'équation fonctionnelle
$$F(z,u) = 1 + z\left(\int_0^1 tF'(z,t)dt + \int_0^u F(z,t)dt\right)$$
$$= 1 + z\left(F(z,1) - \int_0^1 F(z,t)dt + \int_0^u F(z,t)dt\right)$$
$$F(z,u) = 1 + z\left(F(z,1) + \int_1^u F(z,t)dt\right) \tag{4}$$
En dérivant par rapport à u, on a donc
$$F'_u(z,u) = zF(z,u)$$
d'où $\ln F(z,u) = uz + C$ et donc $F(z,u) = K\exp(uz)$ (C et K sont des constantes par rapport à u). Nous allons déterminer K en réinjectant dans l'équation (4) la valeur obtenue pour $F(z,u)$.

$$K\exp(uz) = 1 + z\left(K\exp(z) + \left[\frac{1}{z}K\exp(tz)\right]_1^u\right)$$
$$K\exp(uz) = 1 + Kz\left(\exp(z) + \frac{\exp(uz)}{z} - \frac{\exp(z)}{z}\right)$$
$$K\exp(uz) = 1 + K\Big((z-1)\exp(z) + \exp(uz)\Big)$$
$$0 = 1 + K(z-1)\exp(z) \text{ d'où } K = \frac{e^{-z}}{1-z}$$

Ainsi
$$F(z,u) = \frac{\exp((u-1)z)}{1-z}.$$

8. Marches et fractions continuées

[section ajoutée le 26/11/1998]

La correspondance entre les marches et ces objets familiers de la combinatoire que sont les fractions continuées a été introduite par Flajolet (voir [**14**]).

Elle apparaît naturellement quand on songe au fait qu'une marche peut souvent être décrite en terme de suite de suite de suite (...) d'une combinaison de pas d'un type donné ; ceci se transcrit naturellement par la suite itérée de constructeurs **Seq(Seq(...))**, éventuellement avec quelques variations **Seq(A+Seq(B+Seq(...))** et donc à une série génératrice du type

$$\cfrac{1}{1-A(z)-\cfrac{1}{1-B(z)-\ddots}}.$$

8.1. Cas général.

L'équation fonctionnelle de l'EGF

$$\tilde{F}(z,u)=\sum f_n(u)\frac{z^n}{n!}=\sum \tilde{f}_n(u)z^n$$

n'est guère plus difficile à obtenir que l'équation fonctionnelle de l'OGF, en effet, la récurrence est

$$f_{n+1}(u)\frac{z^{n+1}}{(n+1)!}=Lf_n(u)\delta_z^{-1}\frac{z^n}{n!}=\int_z Lf_n(u)\frac{z^n}{n!}$$

ce qui mène à l'équation différentielle pour l'EGF

$$\tilde{F}(z,u)=f_0(u)+\int L\tilde{F}(z,u)dz$$

que l'on récrit

$$\partial_z\tilde{F}(z,u)=L\tilde{F}(z,u).$$

Quand $L=P(u)+Q(u)\partial_u$ (pour des polynôme de Laurent P et Q fixés) les équations

$$\partial_z F(z,u)=P(u)F(z,u)+Q(u)\partial_u F(z,u),$$

font partie d'une classe d'EDP que l'on sait résoudre.

Ainsi, on peut fournir une forme close pour les marches pour lesquelles les multiplicités ne sont plus seulement en $O(1)$, à condition toutefois que la contrainte de positivité ne joue pas.

Rappelons que l'on a $F(1/z,u)=z\mathcal{L}\tilde{F}(z,u)$ où $\mathcal{L}$ désigne la transformée de Laplace en z.

8.2. Les polynômes des arrangements.

La grammaire

$$[2,(k)\to(k+1)^{k-1}]$$

i.e. la marche $[0,(k)\to(k+1)^{k+1}]$ *i.e.*

$$\begin{cases}(0)\to(1)(2)\\(1)\to(1)(2)(2)\\(2)\to(2)(3)(3)(3)\\\dots\end{cases}$$

aboutit à l'EGF

$$\tilde{F}(z,u)=\frac{\exp(z)}{1-uz}.$$

Les coefficients vérifient $f_n(u)=\sum_{k=0}^{n}\frac{n!}{(n-k)!}u^k$.

En particulier on a $F(z,1)=1+2z+5z^2+16z^3+65z^4+\dots$.

8.3. Les polynômes d'Hermite. La grammaire

$$[2, (k) \rightarrow (k-1)^{k-1}(k+1)]$$

i.e. $(1) \rightarrow (2), (2) \rightarrow (1)(3), (3) \rightarrow (2)(2)(4), \ldots$ se renormalise (mène à la même série génératrice) que la marche

$$[1, (k) \rightarrow (k-1)^{k}(k+1)]$$

i.e.

$$\begin{cases} (0) \rightarrow (1) \\ (1) \rightarrow (0)(2) \\ (2) \rightarrow (1)(1)(3) \\ \ldots \end{cases}$$

L'opérateur caractéristique de cette dernière est $L = S + \partial_u$. On a la relation de récurrence

$$f_{n+1}(u) = uf_n(u) + \partial_u f_n(u) \text{ avec } f_0(u) = u,$$

ce qui mène à l'équation fonctionnelle $F(z,u)(1-zu) = u + z\partial_u F(z,u)$ pour l'OGF et à l'équation fonctionnelle suivante pour l'EGF

$$\partial_z \tilde{F}(z,u) = u\tilde{F}(z,u) + \partial_u \tilde{F}(z,u).$$

La solution de cette équation est

$$\tilde{F}(z,u) = \exp(zu + \frac{z^2}{2}).$$

On reconnaît ici la série génératrice des involutions sans point fixe ; en effet le graphe fonctionnel d'une permutation montre clairement qu'une involution est une suite de cycles de longueur 1 ou 2, marquons les cycles de longueur 1 par u, on a donc la structure suivante **Seq(u Cycle(longueur 1) + Cycle(longueur 2))**, qui se traduit classiquement par la série génératrice (EGF) $\tilde{F}(z,u) = \exp(uz + \frac{z^2}{2})$. En particulier $F(z,0)$ compte les involutions sans point fixe. Les coefficients de la série génératrice bivariée sont en fait très intimements liés aux polynômes d'Hermite [1] par la relation suivante

$$f_n(u) = n!\tilde{f}_n(u) = \frac{(-i)^n H_n(\frac{iu}{\sqrt{2}})}{\sqrt{2^n}}.$$

Rappelons que les polynômes d'Hermite sont définis par la récurrence

$$H_n(x) = 2xH_{n-1}(x) - 2(n-1)H_{n-2}(x) \text{ avec } H_1(x) = 2x \text{ et } H_0(x) = 1.$$

Ainsi, on a $f_0(u) = u$, $f_1(u) = 1 + u^2$, $f_2(u) = 3u + u^3$, $f_3(u) = 3 + 6u^2 + u^4$, $f_4(u) = 15u + 10u^3 + u^5 \ldots$

Donnons pour finir, les premiers termes des deux séries qui nous intéressent :

$$F(z,1) = 1 + z + 2z^2 + 4z^3 + 10z^4 + 26z^5 + \ldots$$

$$F(z,0) = \cfrac{1}{1 - \cfrac{z^2}{1 - \cfrac{2z^2}{1 - \cfrac{3z^2}{1 - \cfrac{4z^2}{\ddots}}}}} = 1 + z^2 + 3z^4 + 15z^6 + 105z^8 + \cdots$$

[1]Le h du mot hermite ($\epsilon\rho\eta\mu\iota\alpha$=désert en grec) est une anomalie et à ce titre n'est pas aspiré.

8.4. Les polynômes de Laguerre. La grammaire

$$[2, (k) \rightarrow (k+1)^{k-1}(k+2)],$$

ou de façon équivalente la marche $[0, (k) \rightarrow (k+1)^{k+1}(k+2)]$ *i.e.*

$$\begin{cases} (0) \rightarrow (1)(2) \\ (1) \rightarrow (2)(2)(3) \\ (2) \rightarrow (3)(3)(3)(5) \\ \dots \end{cases}$$

aboutit à une OGF dont les coefficients vérifient (car $L = u\partial u + u^2 = u + u^2\partial + u^2$),

$$f_{n+1}(u) = (u + u^2)f_n(u) + u^2\partial f_n(u) \text{ avec } f_0(u) = 1.$$

On a l'égalité

$$f_n(u) = n!u^n L_n(-u)$$

où L_n est le n-ième polynôme de Laguerre, ces derniers étant définis par la récurrence

$$L_n(u) = \frac{2n-1-u}{n} L_{n-1}(u) - \frac{n-1}{n} L_{n-2}(u) \text{ avec } L_0(u) = 1 \text{ et } L_1(u) = 1 - u.$$

Ainsi $f_0(u) = 1$, $f_1(u) = u+u^2$, $f_2(u) = 2u^2+4u^3+u^4$, $f_3(u) = 6u^3+18u^4+9u^5+u^6$, $f_4(u) = 24u^4 + 96u^5 + 72u^6 + 16u^7 + u^8 \dots$

La série génératrice exponentielle est

$$\tilde{F}(z, u) = \frac{\exp(\frac{u^2 z}{1-uz})}{1 - uz}.$$

La série génératrice des excursions est bien sûr sans intérêt ici. Donnons les premiers termes de la série génératrice des chemins :

$$F(z, 1) = 1 + 2z + 7z^2 + 34z^3 + 209z^4 + 1546z^5 + \dots$$

8.5. Les polynômes de Stirling. La grammaire

$$[2, (k) \rightarrow (k)^{k-1}(k+1)],$$

ou sous notre point de vue habituel, la marche

$$[0, (k) \rightarrow (k)^{k+1}(k+1)]$$

i.e.

$$\begin{cases} (0) \rightarrow (0)(1) \\ (1) \rightarrow (1)(1)(2) \\ (2) \rightarrow (2)(2)(2)(3) \\ \dots \end{cases}$$

donne l'EGF

$$\tilde{F}(z, u) = \sum_{n \geq 0} S_n(u) \frac{z^n}{n!}.$$

Rappelons que le n-ième polynôme de Stirling est défini comme

$$S_n(u) := \sum_{k=1}^{n} \left\{ {n \atop k} \right\} u^k$$

où $\left\{ {n \atop k} \right\}$ est le nombre de Stirling de deuxième espèce (*i.e.* le nombre de façon de partager les entiers de 1 à n en k ensembles non vides).

La preuve se fait en utilisant le fait que ces nombres de Stirling vérifient

$$\left\{ {n+1 \atop k} \right\} = \left\{ {n \atop k-1} \right\} + k \left\{ {n \atop k} \right\}.$$

Ainsi $f_0(u) = 1$, $f_1(u) = 1 + u$, $f_2(u) = 1 + 3u + u^2$, $f_3(u) = 1 + 7u + 6u^2 + u^3$, $f_4(u) = 1 + 15u + 25u^2 + 10u^3 + u^4 \dots$. En $u = 1$, l'EGF est

$$\tilde{F}(z, 1) = \exp(\exp(z) - 1)) = 1 + 2z + 5z^2 + 15z^3 + 52z^4 + 203z^5 + \dots$$

la série génératrice des nombres de Bell.

8.6. Les Bessel: un cas non holonome. Les systèmes de Pinzani les plus simples sont ceux pour lesquels on a des pas de hauteur +1,-1 ou 0. La plupart mènent à des fonctions holomes (du type Hermite, Laguerre ou arrangements comme vu plus hauts). En dépit de sa simplicité, le système suivant mène à une série génératrice non holonome.

On prend l'axiome (2) et les productions

$$(2) \to (2)(3),\ (3) \to (2)(3)(4),\ \ldots\ ,(k) \to (k-1)(k)^{k-2}(k+1),$$

la première règle est modifiée pour les besoins des systèmes eco. La règle modifiée

$$[(0),(k) \to (k-1)(k)^{k}(k+1),$$

donne une série génératrice bivariée $F(z,u)$ qui satisfait

$$F(z,u)\left(1-z-z(u+u^{-1})\right) - zu\frac{\partial}{\partial u}F(z,u) = 1 + z(1-u^{-1})F(z,0),$$

Pour les excursions, il est plus clair de regarder sa repésentation en fraction continuée

$$F(z,0) = \cfrac{1}{1-z-\cfrac{z^2}{1-z-\cfrac{z^2}{1-2z-\cfrac{z^2}{1-3z-\cfrac{z^2}{\ddots}}}}} = 1 + z + 2z^2 + 4z^3 + 9z^4 + \cdots,$$

pour laquelle seul le premier niveau est irrégulier. Une comparaison avec [**15**] montre que

$$F(z,0) = \frac{1}{1-z-z^2B(z)} \text{ où } B(z) = 1 + z + 2z^2 + 5z^3 + 14z^4 + 43z^5 + 143z^6 + \cdots$$

est la série génératrice des nombres de Bessel, c'est-à-dire, **M1462** dans l'EIS. De [**15**], on sait que

$$z - \frac{1}{z}B(-z) = \frac{J_{z-1}(2)}{J_z(2)},$$

où J_ν est la J-function de Bessel d'ordre ν.

Il reste à vérifier que $F(z,u)$ est non holonome. La croisance rapide de $[z^n]B(z)$ donne

$$[z^n]F(z,0) \sim [z^{n-2}]B(z),$$

et l'asymptotique connue de $[z^n]B(z)$ qui n'est de la forme "holonome" (voir [**35**] pour les types admissibles) implique $F(z,0)$ (et donc $F(z,u)$ n'est pas holonome).

9. Marches en dimension 2 : des chemins de Dyck bien cavaliers

Dans [**27**] et [**28**], Jacques Labelle et Yeong-Nan Yeh ont donné la série génératrice d'une marche de Dyck "modifiée" pour laquelle on ne considère plus des pas "nord-est" et "sud-est" comme c'est le cas traditionnellement mais des pas "de cavalier d'échecs". Leur méthode consiste à trouver une décomposition astucieuse de telles marches, ils arrivent ainsi à six langages dont ils trouvent les séries génératrices, puis ils en déduisent la série génératrice des marches de cavaliers. Ils appliquent ensuite formellement la même idée à un cavalier qui n'aurait plus des pas $(1,2)$ mais des pas (r,s).

Nous allons présenter ici une alternative à la solution de Labelle et Yeh. En effet, notre façon de traiter les marches aléatoires s'applique pleinement ici et de plus, elle nous permet de coller plus précisément à la "structure" de ce problème. Nous pouvons ainsi étudier à loisir différents paramètres.

9.1. Marches d'un coin à l'autre. *On considère un échiquier $n\times\infty$ (*i.e.* de hauteur non bornée et de largeur n). De combien de manières peut-on rejoindre le coin sud-ouest de coordonnées $(0,0)$ au coin sud-est de coordonnées $(n,0)$ par des mouvements de cavalier "gauche-vers-droite", c'est-à-dire de type $(1,2)$, $(1,-2)$, $(2,1)$, $(2,-1)$? La règle du jeu est donc qu'on ne peut pas déborder vers le bas.*

Nous allons en fait tout de suite regarder la série génératrice de la marche. Le polynôme caractéristique de cette marche est $P(u,v) := uv^2+uv^{-2}+u^2v+u^2v^{-1}$. Notons

$$\{u^{<0}v^{<0}\} := \left(\sum_{j<0}\sum_{i\geq 0}+\sum_{i<0}\sum_{j\geq 0}+\sum_{i<0}\sum_{j<0}\right) u^iv^j[u^iv^j].$$

On a

$$f_{n+1}(u,v) := P(u,v)f_n(u,v) - \{u^{<0}v^{<0}\}P(u,v)f_n(u,v)$$

comme on ne fait qu'avancer en u, la contrainte $u^{\geq 0}$ ne joue pas et on a :

$$f_{n+1}(u,v) = P(u,v)f_n(u,v) - \{v^{<0}\}uv^{-2}+u^2v^{-1}f_n(u,v)$$

or, comme les seules puissances négatives de v qui interviennent sont v^{-1} et v^{-2}, on a

$$f_{n+1}(u,v) = P(u,v)f_n(u,v) - (v^{-1}[v^{-1}]+v^{-2}[v^{-2}])uv^{-2}+u^2v^{-1}f_n(u,v)$$

en distribuant l'opérateur d'extraction des coefficients :

$$f_{n+1} = Pf_n - v^{-1}[v^{-1}]uv^{-2}f_n - v^{-1}[v^{-1}]u^2v^{-1}f_n - v^{-2}[v^0]uf_n$$

d'où

$$f_{n+1}(u,v) = P(u,v)f_n(u,v) - v^{-1}[v^1]uf_n(u,v) - v^{-1}[v^0]u^2f_n(u,v) - v^{-2}[v^0]uf_n(u,v)$$

c'est-à-dire

$$f_{n+1}(u,v) = P(u,v)f_n(u,v) - uv^{-1}[v^1]f_n(u,v) - (u^2v^{-1}+uv^{-2})[v^0]f_n(u,v)$$

d'où, en multipliant par z^{n+1},

$$z^{n+1}f_{n+1} = z\big(P(u,v)z^nf_n - uv^{-1}[v^1]z^nf_n - (u^2v^{-1}+uv^{-2})[v^0]z^nf_n\big)$$

et en sommant pour n allant de 0 à $+\infty$, puis en ajoutant $f_0(u,v)$

$$F(z,u,v) = 1+z\big(P(u,v)F(z,u,v) - uv^{-1}\partial_vF(z,u,0) - (u^2v^{-1}+uv^{-2})F(z,u,0)\big)$$

Ainsi, en regroupant les $F(z,u,v)$ à gauche puis en multipliant par v^2, on obtient :

$$F(z,u,v)\big(v^2 - z(uv^4+u+u^2v^3+u^2v)\big) = v^2 - uv\partial_vF(z,u,0) - (u^2v+u)F(z,u,0)\big)$$

En notant v_1 et v_2 les deux solutions de

$$v^2 = zv^2P(u,v)$$

qui tendent vers 0 quand z tend vers 0, à savoir

$$\begin{cases} v_1(u,z) := -\dfrac{z\,u^2-\Delta}{4zu} - \dfrac{1}{4}\sqrt{2}\sqrt{\dfrac{u^3z-u\Delta+2-4zu}{zu}} = +\sqrt{z}\sqrt{u}+\cdots \\ v_2(u,z) := -\dfrac{z\,u^2+\Delta}{4zu} + \dfrac{1}{4}\sqrt{2}\sqrt{\dfrac{u^3z+u\Delta+2-4zu}{zu}} = -\sqrt{z}\sqrt{u}+\cdots \end{cases}$$

où $\Delta := \sqrt{z^2u^4+4zu+8z^2u^2}$. On peut alors résoudre le système

$$\begin{cases} v_1^2 - uv_1\partial_vF(z,u,0) - (u^2v_1+u)F(z,u,0)\big) = 0 \\ v_2^2 - uv_2\partial_vF(z,u,0) - (u^2v_2+u)F(z,u,0)\big) = 0 \end{cases}$$

Cela donne

$$\begin{cases} \partial_vF(z,u,0) = \frac{v_1+v_2+uv_1v_2}{zu} \\ F(z,u,0) = -\frac{v_1v_2}{zu} \end{cases}$$

d'où, au final :

$$F(z,u,v) = \frac{v^2+v_1v_2-vv_1-vv_2}{v^2-z(u^2v^3+uv^4+u+u^2v)}$$

Nous ne donnons pas la forme close obtenue par substitution des formes closes de v_1 et v_2 car celle-ci n'est pas très "compacte".

On peut maintenant retourner à la question initiale, "déterminer le nombre de chemins allant de $(0,0)$ à $(n,0)$" ; la suite recherchée est donnée par

$$a_n = [u^n v^0]F(1,u,v).$$

Ainsi la série génératrice est

$$E(u) := F(1,u,0) = -\frac{v_1(1,u)v_2(1,u)}{u} = 1 + u^2 + 3u^4 + 2u^5 + 12u^6 + O(u^8).$$

Par un calcul de résultants, avec les lignes suivantes en Maple :
>**resultant(**$uE + v1 * v2, v1^2 - (u^2 * v1^3 + u^2 * v1 + u + u * v1^4), v1)$
>**resultant(%,**$v2^2 - (u^2 * v2^3 + u^2 * v2 + u + u * v2^4), v2)$**);**
>**factor(%);**
on obtient que la fonction E vérifie l'équation

$$u^8(uE+1)^4 AB$$

avec $A = u^4E^4 + 2u^2E^3 + u^5E^3 + E^2 + 2u^2E^2 - 2E^2u^4 + 2E + u^3E + 1$
et $B = 1 - E - 2uE + 2uE^2 + 2u^2E^2 + E^2u^4 - u^2E^3 - 2u^3E^3 + u^4E^4$. On trouve aisément le "bon" facteur (puisqu'on a $E(0) = 1$) et ainsi E vérifie

$$1 - E - 2uE + 2uE^2 + 2u^2E^2 + E^2u^4 - u^2E^3 - 2u^3E^3 + u^4E^4 = 0$$

que l'on peut résoudre explicitement et alors déterminer laquelle des 4 racines est la bonne. En posant

$$\Delta := \sqrt{1 - 4u + 4u^2 - 4u^4},$$

on obtient

$$E(u) := \frac{1+2u+\Delta}{4u^2} - \frac{\sqrt{2 + 2\Delta - 4u - (8+8\Delta)u^2 + 16u^3 - (8+4\Delta)u^4 - 16u^5}}{4u^2\sqrt{\Delta}}.$$

Notons au pasage, que la forme close que l'on obtient en faisant $z = 1$ et $v = 0$ dans la forme close donnant $F(z,u,v)$ est bien plus complexe que celle que nous obtenons par ce calcul de résultants. Ainsi, faute d'un algorithme puissant de simplification de radicaux (désemboîtage de racines imbriquées ou "denesting"), le passage par les résultants reste fort appréciable.

La singularité réelle de plus petit module de $E(u)$ est $\frac{-1+\sqrt{3}}{2} = 0.36\ldots$. Ainsi, le théorème de Pringsheim nous donne

$$a_n \sim (1+\sqrt{3})^n \approx 2.73^n.$$

Une analyse de singularité donnerait

$$a_n = \frac{A}{\sqrt{\pi}n^{3/2}}(1+\sqrt{3})^n + O\left(\frac{(1+\sqrt{3})^n}{n^{5/2}}\right)$$

où

$$A = \frac{237 - 137\sqrt{3} - 15\sqrt{-13+8\sqrt{3}} + 9\sqrt{-13+8\sqrt{3}}\sqrt{3}}{(-1+\sqrt{3})^2\sqrt{18-10\sqrt{3}}(-13+8\sqrt{3})} = 0,6703\ldots$$

Remarquons que l'on obtient ici une série génératrice algébrique, ce qui ne contredit nullement le théorème 1.1 qui impliquait la rationalité de la série génératrice quand le terrain de jeu était borné : ici en effet, quoique l'on soit pour chaque n sur un domaine borné, ce n'est pas le cas pour l'ensemble de la marche car "le borné" croît avec n.

Notons de plus que l'on peut ici appliquer la formule d'inversion de Lagrange-Bürmann (à savoir $[z^n]y(z) = \frac{1}{n}[u^{n-1}]\phi(u)^n$ pour toute fonction $y(z)$ qui vérifie $y(z) = z\phi(y(z))$ où ϕ est une série non nulle en zéro) en choisissant

$$\phi(y) := \frac{y^4 - 2y^3 + 3y^2 - 2y + 1}{y^2 - 2y + 1}.$$

Remarquons aussi que les coefficients de ϕ sont facilement explicitables car

$$\phi = 1 + y^2 + y^2\left(\frac{1}{1-y}\right)' = 1 + 2y^2 + 2y^3 + 3y^4 + \cdots.$$

9.2. Marches qui finissent sur le bord. Voici une question parallèle à celle du paragraphe précédent : *on considère un échiquier $\infty \times \infty$. Combien de séquences de n mouvements de cavalier sont-elles possibles, qui partent de $(0,0)$ et qui reviennent sur la ligne de base $y = 0$, en étant toujours positifs ou nuls ?*

La suite recherchée est donnée par $b_n = [z^n v^0] = F(z,1,v)$.
Ainsi la série génératrice est

$$E(z) := F(z,1,0) = -\frac{v_1(z,1)v_2(z,1)}{z} = 1 + 2z^2 + 2z^3 + 11z^4 + 24z^5 + 93z^6 + O(z^8)$$

Par un calcul de résultants, avec les lignes suivantes en Maple :
$>$**resultant**$(zE + v1 * v2, v1^2 - z * (v1^3 + v1 + 1 + v1^4), v1)$
$>$**resultant**$(\%, v2^2 - z * (v2^3 + v2 + 1 + v2^4), v2)$;
$>$**factor(%);**
on obtient que la fonction E vérifie l'équation

$$z^8(1 + zE)^4 AB^2 = 0$$

avec $A = 1 + 2E + zE + E^2 + 2z^2E^3 + E^3z^3 + E^4z^4$
et $B = 1 - E - 2zE + 2zE^2 + 3z^2E^2 - z^2E^3 - 2E^3z^3 + E^4z^4$.

On trouve aisément le "bon" facteur (car $E(0) = 1$) et ainsi E vérifie

$$1 - E - 2zE + 2zE^2 + 3z^2E^2 - z^2E^3 - 2E^3z^3 + E^4z^4 = 0$$

que l'on peut résoudre explicitement et alors déterminer laquelle des 4 racines est la bonne. Nous faisons la même remarque que précédemment, à savoir qu'il vaut mieux appliquer les formules de Ferrari sur un polynôme le plus "simple" possible, afin d'obtenir une forme close "compacte" sans avoir à faire du denesting, c'est pour cette raison que nous sommes encore passés par un calcul de résultants.

On obtient, en notant $\Delta := \sqrt{1-4z}$,

$$E(z) := \frac{1 + 2z + \Delta}{4z^2} - \frac{\sqrt{2 + 2\Delta - 4z - 16z^2 - 12\Delta z^2}}{4z^2\sqrt{\Delta}}.$$

Les singularités réelles sont $\frac{1}{4}$ et $-\frac{4}{9}$, ainsi le théorème de Pringsheim nous donne

$$b_n \sim 4^n.$$

Une analyse de singularité donnerait

$$b_n = \frac{A}{\sqrt{\pi}n^{3/2}}4^n + O\left(\frac{4^n}{n^{5/2}}\right)$$

avec $A = \frac{6}{\sqrt{5}} - 2 = 0,6832\ldots$

10. Limitation par une droite de pente rationnelle

Au lieu de prendre le premier quadrant comme terrain de jeu, on pourrait se limiter à des marches sur les points (à coordonnées entières et positives) compris entre les droites $x = 0$ et $y = ax$. On sait que, pour une droite de pente irrationnelle, la suite des points à coordonnées entières (x_i, y_i) les plus proches de la droite correspondent aux réduites successives $\frac{y_i}{x_i}$ du développement en fraction continuée de a, c'est-à-dire la suite des fractions de plus petit numérateur qui approche a. Si la droite a un coefficient rationnel, il y a des points à coordonnées entières dessus, et l'on peut s'amuser à dénombrer les marches partant de l'origine et finissant sur cette droite.

Comme pour les marches avec conditions de bord usuelles (*i.e.* limitées au premier quadrant), on construit le polynôme caractéristique $P(u,v)$ comme dans la section 9. On se limite à gauche par $x = 0$ et à droite par $y = \frac{p}{q}x$, posons

$$P(u) := P(u^{-p}, u^q).$$

Ainsi, un pas dans notre marche 2-dimensionelle d'origine qui dépasserait la droite frontière correspondra à un pas qui atterrira dans les négatifs pour la marche unidimensionnelle que nous venons de lui associer.

Nous n'avons plus qu'à appliquer la méthode utilisée pour les marches unidimensionnelles pour obtenir $F(z,u)$. Le nombre de marches de longueur n est alors donné par $[z^n]F(z,1)$ où $F(z,u)$ est

$$F(z,u)\big(1-zP(u)\big) = 1 - z\{u^{<0}\}\sum P(u)z^n f_n(u).$$

Le nombre de marches de longueur n finissant sur la droite $y=\frac{p}{q}x$ est quant à lui donné par $[z^n]F(z,0)$.

Remarque. *Soulignons que nous n'obtenons pas la série génératrice de la marche bivariée (qui peut ne pas être algébrique) mais simplement sa "projection" sur l'axe des abscisses et sa "trace" sur la droite "frontière" de pente rationnelle, ce qui nous permet néanmoins de répondre aux questions que nous nous posions.*

10.1. Exemple : Philippe Duchon [**10**] a étudié les "langages de Dyck généralisés" : on prend un alphabet fini, on donne à chaque lettre une valeur entière positive ou négative, et on forme un langage avec des conditions de positivités similaires à celles des mots de Dyck (valeur totale nulle, valeur positive pour tous les facteurs gauches).

Théorème 10.1. *Les langages de Dyck généralisés sont algébriques, ou encore : la série génératrice des chemins de Dyck généralisés est algébrique.*

Preuve. Labelle et Yeh [**28**] ainsi que Duchon [**10**] l'ont démontré en déterminant une grammaire génératrice. Avec 2 lettres, cela revient à considérer des marches limitées par une droite de pente rationnelle et cela tombe donc sous la coupe de la méthode exposée ci-dessus. Pour un alphabet de taille quelconque, on peut sinon utiliser le théorème 3.1 qui nous dit qu'une excursion positive à pas bornés a une série génératrice algébrique. □

Par exemple, si l'on s'autorise un pas vers le haut et un pas vers la droite et pour lequel on se limite à droite par $y=x$, on aboutit à

$$F(z,u)\big(1-z(u^{-1}+u)\big) = 1 - z\{u^{<0}\}\sum_{n\geq 0} P(u)z^n f_n(u)$$

donne

$$F(z,u)\big(u-z(1+u^2)\big) = u - zF(z,0).$$

Ce cas est traité en annexe sous la forme d'une session Maple.

11. Marches sur un réseau hexagonal

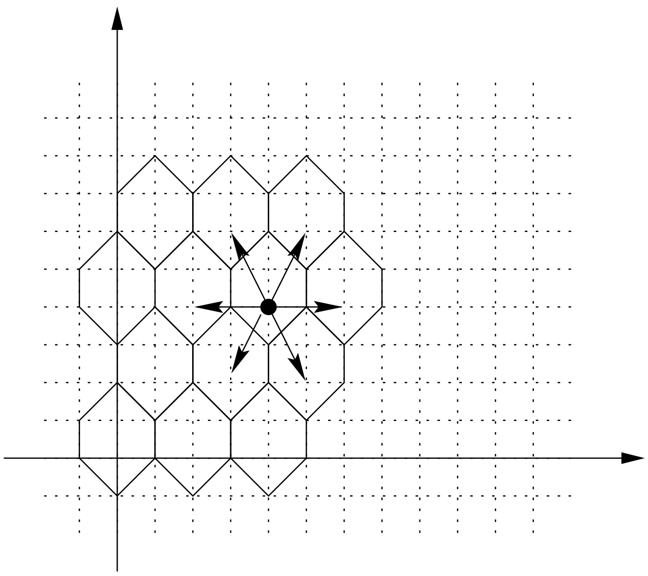


Figure 1. Correspondance entre le réseau hexagonal et réseau carré.

Avec le schéma 1, on voit qu'aller dans l'alvéole supérieure[2] droite correspond à faire un pas $(+1,+2)$ sur un réseau carré.
Aller dans l'alvéole supérieure gauche correspond à faire un pas $(-1,+2)$.
Aller dans l'alvéole de droite correspond à faire un pas $(+2,+0)$.
Aller dans l'alvéole de gauche correspond à faire un pas $(-2,+0)$.
Aller dans l'alvéole inférieure droite correspond à faire un pas $(+1,-2)$.
Aller dans l'alvéole inférieure gauche correspond à faire un pas $(-1,-2)$.

Le polynôme caractéristique est donc

$$P(u,v) := uv^2 + uv^{-2} + u^{-1}v^2 + u^{-1}v^{-2} + u^2 + u^{-2}.$$

La série génératrice des marches sans contrainte de positivité est immédiate c'est $\frac{1}{1-6z}$; la série génératrice des excursions (on part et on finit au point $(0,0)$) sans contrainte de positivité s'obtient comme à la proposition 2.2 du chapitre III. On a donc

$$F(z,0,0) = [u^0][v^0]\frac{1}{1-zP(u,v)} = [u^0]\sum_{v_i}\frac{1}{-zv_i\partial_v P(v_i)}$$

où les v_i sont les zéros de $1-zP$, petits autour de (0,0). Posons $\tilde{P}_i := -zv_i\partial_v P(v_i)$ et soient u_j les zéros de $\tilde{P}_i$ tendant vers 0 quand z tend vers 0 (en fait les $\tilde{P}_i$ ont les mêmes zéros), on a alors

$$F(z,0,0) = \sum_{v_i}\sum_{u_j}\frac{1}{u_i\partial_u\tilde{P}_i(u_i)}.$$

Regardons maintenant la marche pour laquelle on peut aller sur n'importe quelle alvéole adjacente tout en restant dans le premier "quadrant", elle donne la récurrence :

$$f_{n+1}(u,v) := P(u,v)f_n(u,v) - \{u^{<0}v^{<0}\}P(u,v)f_n(u,v).$$

On a donc, avec la notation $\{u^av^b\} = u^av^b[u^av^b]$,

$$\{u^{<0}v^{<0}\} = \sum_{i\geq 0}\left(\{u^iv^{-1}\}+\{u^iv^{-2}\}\right)+\sum_{j\geq 0}\left(\{u^{-1}v^j\}+\{u^{-2}v^j\}\right)+\{u^{-1}v^{-1}\}+\{u^{-1}v^{-2}\}$$

12. L'abeille bourrée

Nous allons ici montrer comment obtenir la série génératrice avec une contrainte de positivité stricte.

Soit $\tilde{F}(z) = \sum_{n\geq 0}\tilde{f}_n z^n$ la série génératrice des marches strictement positives. Notons F la série génératrice des marches positives et E la série génératrice des excursions positives.

Nous définerons la probabilité de "retour à l'origine" comme la limite

$$\lim_{n\to\infty}\frac{\tilde{f}_n}{f_n}.$$

L'ensemble des marches de longueur $n+1$ strictement positives (c'est-à-dire qui n'ont jamais touché la droite $y=0$), c'est l'ensemble de toutes les marches de longueur $n+1$, moins celles qui ont touché la dernière fois l'axe au bout de n pas, moins celles qui ont touché la dernière fois l'axe au bout de $n-1$ pas, ... , moins celles qui ont touché la dernière fois l'axe lors du premier pas.

Ce qui s'écrit aussi : $\tilde{f}_{n+1} = f_{n+1} - e_{n+1} - e_n\tilde{f}_1 - e_{n-1}\tilde{f}_2 - \cdots - e_1\tilde{f}_n$

En multipliant par z^{n+1}, en sommant pour $n>0$, et en ajoutant $\tilde{f}_0 = 0$ et $z\tilde{f}_1 = z(f_1-e_1)$, on obtient :

$$\tilde{F} = z(f_1-e_1) + (F - f_0 - zf_1) - (E - e_0 - ze_1) + \sum_{n\geq 1} z^{n+1}\sum_{k=1}^{n} e_k\tilde{f}_{n+1-k}$$

$$\tilde{F} = F - E + z^2\sum_{n\geq 1} z^{n-1}[z^{n-1}]\frac{E-e_0}{z}\frac{\tilde{F}-\tilde{f}_0}{z} = F - E - (E-e_0)\tilde{F}$$

[2]Les Académiciennes à la lanterne !

$$\tilde{F} = \frac{F}{E} - 1$$

Pour les marches de Dyck (voir la section 3.1), on obtient bien sûr $\tilde{F} = zF$.

13. Le cavalier de Petkovšek

Une suite f_n qui vérifie une récurrence linéaire à coefficients constants a une série génératrice rationnelle. C'est pourquoi on pouvait immédiatement imaginer qu'une marche aléatoire à pas bornés allait donner une série génératrice rationnelle en z tant pour les chemins que pour les excursions, car on avait les deux récurrences linéaires à coefficients constants suivantes : $f_n(1) = P(1)f_n(1)$ et $f_n(0) = P(0)f_n(0)$.

Attention ! Un tel argument n'est plus valable en dimension ≥ 2, en effet, une suite vérifiant une récurrence linéaire à coefficients constants peut très bien mener à une série génératrice transcendante, comme c'est le cas pour l'exemple suivant [**4**] :

$$a_{i,j} := a_{i-2,j+1} + a_{i+1,j-2} + a_{i-1+j-1}$$

avec les conditions de bord suivantes

$$a_{0,i} = a_{1,i} = a_{i,0} = a_{i,1} = 1.$$

C'est d'ailleurs en recherchant la nature (algébrique ou transcendante), d'une suite dénombrant les marches de cavalier de Petkovšek (voir [**29**]) que Mireille Bousquet-Mélou [**4**] a trouvé le contre-exemple précédent.

La récurrence originale de Petkovšek était

$$a_{i,j} := a_{i-2,j+1} + a_{i+1,j-2}$$

avec les conditions de bord suivantes

$$a_{0,i} = a_{1,i} = a_{i,0} = a_{i,1} = 1.$$

On peut la voir comme une généralisation de la définition du triangle de Pascal. En effet, ici chaque terme du triangle est la somme des voisins des deux nombres situés au dessus et en diagonal (ainsi, à la onzième ligne : $a_{4,7} := a_{5,5} + a_{2,8}$ *i.e.* 22=14+8).

$$\begin{array}{c}
1 \\
1 \quad 1 \\
1 \quad 1 \quad \mathbf{1} \\
1 \quad 1 \quad \mathbf{1} \quad 1 \\
1 \quad 1 \quad \mathbf{2} \quad 1 \quad 1 \\
1 \quad 1 \quad \mathbf{2} \quad 2 \quad 1 \quad 1 \\
1 \quad 1 \quad \mathbf{3} \quad 2 \quad 3 \quad 1 \quad 1 \\
1 \quad 1 \quad \mathbf{3} \quad 4 \quad 4 \quad 3 \quad 1 \quad 1 \\
1 \quad 1 \quad \mathbf{5} \quad 5 \quad 6 \quad 5 \quad 5 \quad 1 \quad 1 \\
1 \quad 1 \quad \mathbf{6} \quad 7 \quad 10 \quad 10 \quad 7 \quad 6 \quad 1 \quad 1 \\
1 \quad 1 \quad \mathbf{8} \quad 11 \quad 16 \quad \underline{14} \quad 16 \quad 11 \quad \underline{8} \quad 1 \quad 1 \\
1 \quad 1 \quad \mathbf{12} \quad 17 \quad 22 \quad 27 \quad 27 \quad \underline{22} \quad 17 \quad 12 \quad 1 \quad 1 \\
1 \quad 1 \quad \mathbf{18} \quad 23 \quad 39 \quad 44 \quad 44 \quad 44 \quad 39 \quad 23 \quad 18 \quad 1 \quad 1 \\
1 \quad 1 \quad \mathbf{24} \quad 40 \quad 62 \quad 67 \quad 83 \quad 83 \quad 67 \quad 23 \quad 18 \quad 24 \quad 1 \quad 1 \\
1 \quad 1 \quad \mathbf{a_{12,2}} \; a_{11,3} \; a_{10,4} \qquad \ldots \qquad a_{4,10} \; a_{3,11} \; a_{2,12} \quad 1 \quad 1 \\
\vdots
\end{array}$$

Cette récurrence décompte le nombre de façon qu'a un cavalier situé en (i,j) de rejoindre les bords avec des pas $(-2,+1,)$ et $(+1,-2)$, et où les bords sont absorbants. Indiquons au passage que les marches absorbantes rentrent dans le cadre de notre méthode : en fait on peut distinguer le cas où l'on compte les marches qui atteignent le bord (et alors il faudrait regarder ce que donne l'équation des effleurements en $t = 0$, voir la section 8 du chapitre 4) du cas où on les rejette, il suffit alors dans la définition par récurrence des f_n de remplacer $\{u^{<0}\}$ par $\{u^{\leq 0}\}$.

En fait, il s'avère que l'algébricité de

$$\sum_{i \geq 0, j \geq 0} a_{i,j} u^i v^j$$

est équivalente à l'algébricité de

$$\sum_{j\geq 0} a_{2,j} v^j$$

(série dont les coefficients sont en gras dont le triangle ci-dessus). On sait simplement que cette dernière est irrationnelle, malheureusement, je n'ai pas réussi à traiter complètement ce cas avec la méthode exposée précédemment. La nature de cette série reste donc un problème ouvert.

14. Aire et q-analogues

Dans le plan, on se donne un jeu de pas (a_i, b_i) $(a_i, b_i \in \mathbb{Z})$. Est-il possible d'obtenir une série génératrice qui tiendrait compte de l'aire comprise entre le graphe d'un chemin et l'axe des abscisses ? On comptera négativement l'aire correspondant à un pas vers la gauche.

Pour étudier cette question, remarquons que si la marche se trouve au point (i, j) avec à cet instant une aire k, et que l'on fasse le pas (a, b), on arrive en $(i+a, j+b)$ et nous avons alors une aire $k + (j+b)a - \frac{ab}{2}$. Ainsi, pour obtenir une série génératrice avec 4 variables (u et v marquant les coordonnées, q marquant l'aire et z marquant comme précédemment la longueur de la marche), il nous faut trouver un opérateur L tel que

$$L(u^i v^j q^k) = u^{i+a} v^{j+b} q^{k+aj+\frac{ab}{2}}.$$

Il semble à priori difficile d'obtenir un opérateur **linéaire** (à cause de l'exposant de q), toutefois, avec l'aide de l'opérateur de substitution suivant :

$$\sigma : v \mapsto vq^a,$$

c'est finalement une chose aisée et on obtient

$$L := S_q^{\frac{ab}{2}} S_u^a S_v^b \sigma.$$

14.1. Les chemins de Dyck. Si l'on prend la marche de Dyck où l'on s'autorise les pas $(1,1)$ et $(1,-1)$, on obtient l'opérateur

$$L := S_q^{\frac{1}{2}} S_u S_v \sigma + s_q^{-\frac{1}{2}} S_u S_v^{-1} \sigma$$

et donc l'équation suivante :

$$F(z,u,v,q) = 1 + z\big(uv\sqrt{q}F(z,u,vq,q) + \frac{u}{v\sqrt{q}}F(z,u,vq,q) \\ -\{v^{<0}\} \sum_{n\geq 0} Lz^n f_n(u,v,q)\big)$$

que l'on écrit encore

$$F(z,u,v,q) = 1 + z\big(uv\sqrt{q}F(z,u,vq,q) + \frac{u}{v\sqrt{q}}F(z,u,vq,q) \\ -\frac{1}{v}[v^0] \sum_{n\geq 0} \frac{u}{\sqrt{q}} z^n f_n(u,vq,q)\big)$$

soit

$$F(z,u,v,q) = 1 + z\big(uv\sqrt{q}F(z,u,vq,q) + \frac{u}{v\sqrt{q}}F(z,u,vq,q) - \frac{u}{v\sqrt{q}}F(z,u,0,q)\big).$$

$q = 0$ donne $F(z,u,v,0) = 1$
$q = 1$ donne

$$F(z,u,v,1) = 1 + z\big(uv + \frac{u}{v}F(z,u,v,1) - \frac{u}{v}F(z,u,0,1)\big)$$

$v = 1$ donne

$$F(z,u,1,q) = 1 + z(u\sqrt{q}F(z,u,q,q) + u\sqrt{q}F(z,u,q,q) - F(z,u,0,q))$$

Faire tendre v vers 0 donne

$$F(z,u,0,q) = 1 + zu\sqrt{q}\partial_v F(z,u,0,q)$$

dériver par rapport à v donne

$$\begin{aligned}\partial_v F(z,u,v,q) = z\big(&u\sqrt{q}F(z,u,vq,q) + uvq\sqrt{q}\partial_v F(z,u,vq,q) + \frac{u\sqrt{q}}{v}\partial_v F(z,u,vq,q)\\ &-\frac{u}{v^2\sqrt{q}}F(z,u,vq,q) + \frac{u}{v^2\sqrt{q}}F(z,u,0,q)\big)\end{aligned}$$

et faire alors $v = 1$ donne

$$\begin{aligned}\partial_v F(z,u,1,q) = z\big(&u\sqrt{q}F(z,u,q,q) + uq\sqrt{q}\partial_v F(z,u,q,q) + u\sqrt{q}\partial_v F(z,u,q,q)\\ &-\frac{u}{\sqrt{q}}F(z,u,q,q) + \frac{u}{\sqrt{q}}F(z,u,0,q)\big)\end{aligned}$$

tandis que faire tendre v vers 0 donne

$$\begin{aligned}\partial_v F(z,u,0,q) = z\big(&u\sqrt{q}F(z,u,0,q) + \frac{u\sqrt{q}}{v}\partial_v F(z,u,vq,q)\\ &-\frac{u}{v^2\sqrt{q}}F(z,u,vq,q) + \frac{u}{v^2\sqrt{q}}F(z,u,0,q)\big)\end{aligned}$$

On obtient ainsi une équation de q-analogues. L'étude des équations fonctionnelles obtenues pour les aires devrait faire l'objet de recherches ultérieures.

CHAPITRE 4

Asymptotique

1. Introduction

Le cas des marches où l'on ne fait qu'avancer par des pas bornées et reculer d'un pas de longueur 1 conduit à un noyau $u - z\phi(u) = 0$ (où ϕ est le polynôme caractéristique de la marche multiplié par u). Ce sont typiquement ces équations qui définissent les séries génératrices de diverses familles d'arbres (Cayley, binaire, unaire-binaire ...). La fonction u est ainsi définie de façon implicite. Néanmoins, l'asymptotique de ces coefficients est connue ; on trouvera dans **[17]** le résultat de Meir et Moon en $\frac{A}{\rho^n\sqrt{\pi n^3}}$ où A et ρ sont des nombres algébriques explicitables.

Nous allons montrer que ce résultat se maintient pour une plus large classe d'équations, à savoir $u^c - z\phi(u) = 0$ (où ϕ est une série entière), et il se trouve en particulier que les marches positives à pas bornés mènent à ce type d'équation.

Je tiens à remercier Bruno Salvy et Philippe Flajolet qui ont su m'expliquer les diverses insuffisances de la première rédaction de ce chapitre.

2. Existence d'une singularité positive

Nous allons traiter l'asymptotique des marches pour les fonctions suivantes :

Définition 2.1. *Nous qualifierons d'***admissibles** *les fonctions appartenant à la (large) classe de séries entières ϕ suivantes :*

- *ϕ est à coefficients positifs ;*
- *$\phi(0) \neq 0$;*
- *ϕ n'est pas constante ;*
- *ϕ a un rayon de convergence R soit infini, soit fini sous la réserve que ϕ' ait un pôle en $z = R$.*

Si l'on pose $\phi(u) := u^c P(u)$ où P est le polynôme de Laurent caractéristique de la marche (de degré minimal c et de degré maximal d), on voit que nous avons une fonction admissible au sens précédent. Si l'opérateur caractéristique de la marche est non plus un polynôme de Laurent mais une série de Laurent (c'est-à-dire que le degré minimal c est un entier et que le degré maximal d est infini), il faut alors veiller à ce que la dernière condition soit remplie (les autres sont triviales) pour que les théorèmes que nous allons énoncer s'appliquent.

Ainsi, dans tous les cas, on ramène l'étude du noyau $1 - zP(u) = 0$ à l'étude de

$$N(u,z) := u^c - z\phi(u) = 0.$$

Parmi les $e := c + d$ racines (e est donc infini quand d est infini) de cette équation, il y en a c qui tendent vers 0 en 0 (car pour les branches finies en 0, l'équation devient $u(z)^c = 0$). Nous notons ces "petites branches" $u_1, \ldots, u_c$.

Quand u_1 peut-elle avoir des singularités ? Comme on l'a vu dans le premier chapitre, on a un point critique si $z = 0$ (annulation du coefficient de tête) ou bien si

$$N(u,z) = 0 \;\text{ et }\; \frac{\partial N}{\partial u}(u,z) = 0,$$

c'est-à-dire si $cu^{c-1} - z\phi'(u) = 0$ en même temps que $u^c - z\phi(u) = 0$. Ce dernier système donne, si $\phi(u) \neq 0$ et $\phi'(u) \neq 0$ (ces deux conditions étant équivalentes quand $z \neq 0$), $z = \frac{cu^{c-1}}{\phi'(u)} = \frac{u^c}{\phi(u)}$ et finalement, on trouve

$$c = \frac{u\phi'(u)}{\phi(u)}.$$

Il est donc primordial de savoir si et quand cette relation est vérifiée, c'est ce à quoi répond le théorème suivant.

Lemme 2.1. *Si ϕ est une fonction admissible, alors, pour tout entier c tel que $1 \leq c < e = \deg(\phi)$, l'équation en u*

$$c = \frac{u\phi'(u)}{\phi(u)}$$

admet une unique racine positive τ.

Preuve. On pose $h(u) := \frac{\phi(u)}{u^c}$. Ainsi $h'(u) = \frac{\phi'(u)u^c - \phi(u)cu^{c-1}}{u^{2c}}$. Or, pour $\phi(u) \neq 0$, on a les équivalences suivantes

$$h'(u) = 0 \iff \phi'(u)u^c - \phi(u)cu^{c-1} = 0 \iff \frac{u\phi'(u)}{\phi(u)} = c$$

$$\iff u \text{ est un point col de } h$$

Nous allons pour le reste de cette démonstration prendre u réel positif (et on a alors $\phi(u) > 0$). Pour l'indiquer plus clairement au lecteur, nous notons maintenant la variable par x. Nous nous sommes donc ramenés à démontrer l'existence et l'unicité d'une solution à $h'(x) = 0$.

Rappelons que nous notons R le rayon de convergence de ϕ. Si celui-ci est fini, toutes les évaluations en $x = R$ sont à interpréter comme une limite quand $x \to R^-$.

La fonction h est définie et continue sur $]0, R[$. Elle tend vers $+\infty$ en 0 à droite donc h décroît en 0 à droite) et ainsi h' est négative en 0 à droite (en fait, on a même $h'(0^+) = -\infty$ puisque $\phi(0) > 0$).

Si le rayon de convergence R est infini, h tend vers $+\infty$ en R (parce que $c < e$). Si R est fini, alors, puisque ϕ est une fonction admissible, h' a un pôle en $x = R$ (hérité de ϕ') et dans ce cas h aussi et donc $h'(x) \to +\infty$ quand $z \to R^-$ (l'éventuel pôle de ϕ en R apportant une contribution négative, à gauche de R, de toute façon moindre que la contribution positive de ϕ', et il en est de même quand ϕ a un point de branchement).

Ainsi, dans tous les cas, h' est négative en 0 (à droite) et positive en R.

Donc, par le théorème des valeurs intermédiaires, h' s'annule au moins une fois sur $]0, R[$. Si l'on suppose qu'elle s'annule au moins deux fois, on applique le théorème de Rolle et on a donc l'existence d'un point de $]0, R[$ sur lequel h'' s'annule ce qui impossible car (en posant $\sum_{n=0}^{e} a_n x^n := \phi(x)$), on a

$$h''(x) = \sum_{n=0}^{e} a_n(n-c)(n-c-1)x^{n-c-2}$$

est une série à coefficients positifs (non tous nuls car $e > c \geq 1$). Finalement, la fonction h' ne s'annule qu'une seule fois et fournit une unique solution à $h'(x) = 0$ sur $]0, R[$.

Remarque : Nous allons ci-après donner un lemme (3.2) qui permet d'affirmer la croissance de $\frac{x\phi'(x)}{\phi(x)}$. Ainsi cette fonction croît de 0 en 0 vers e en R, et on a alors une démonstration immédiate de l'existence et de l'unicité de la racine τ pour tout c *réel* et tel que $0 \leq c < e$. □

Notons aussi que la généralisation du théorème à toutes les séries à coefficients positifs (non nulles en zéro) n'a pas lieu comme le montre le contre-exemple suivant :

$$\phi(x) = \frac{3}{2}x + (1-x)^{3/2}$$

en effet $\frac{x\phi'(x)}{\phi(x)}$ croît de 0 en $x = 0$ à 1 en $x = R = 1$ et ainsi l'équation $c = \frac{x\phi'(x)}{\phi(x)}$ n'a pas de solution pour $c > 1$. En revanche, par exemple, le théorème s'applique bien à $\phi(x) = 1 - (1-x)^{1/2}$.

En fait, nous voyons bien que la dernière condition d'admissibilité pour les séries à rayon de convergence R fini (que ϕ' ait un pôle en $z = R$) n'est qu'un cas particulier de la véritable condition qui nous importe : que ϕ croisse plus vite que tout polynôme à gauche de R.

3. Développement d'une branche en série de Puiseux

Maintenant que nous avons obtenu l'existence et l'unicité d'un réel positif τ tel que

$$\begin{cases} N(u,\tau) = 0 \\ \frac{\partial N}{\partial u}(u,\tau) = 0 \end{cases}$$

nous allons caractériser le rayon de convergence de u.

Théorème 3.1. *Soit u une fonction vérifiant $u^c = z\phi(u)$ (avec ϕ admissible). Le rayon de convergence ρ de u est donné par $u(\rho) = \tau$ où τ vérifie $c = \frac{\tau\phi'(\tau)}{\phi(\tau)}$. De plus, on a*

$$\rho = \frac{\tau^c}{\phi(\tau)} = \frac{c\tau^{c-1}}{\phi'(\tau)}.$$

Preuve. Ainsi $u(z)$ est la fonction réciproque de $\psi(u) := \frac{u^c}{\phi(u)}$. Rappelons qu'une fonction analytique ψ est inversible (au sens de la composition de fonction) au point z_0 si et seulement si $\psi'(z_0) \neq 0$. On montre que u est une fonction analytique près de l'origine, que $u(0) = 0$ et que ses coefficients sont positifs. Ainsi $u(z)$ croît de $z = 0$ à $z = \rho$. On a établi dans la section précédente l'existence du réel τ tel que $\psi'(\tau) = 0$. Si on avait $u(\rho) < \tau$, on pourrait prolonger analytiquement u puisque ψ' ne s'annule pas. Si on avait $u(\rho) > \tau$, on aurait un $r < \rho$ tel que $u(r) = \tau$ et donc u serait singulière en un point $z = r$ situé à l'intérieur de son disque de convergence ! Dans les deux cas, on a une contradiction avec le fait que le rayon de convergence est ρ.

Par ailleurs, de

$$z = \frac{u(z)^c}{\phi(u(z))}$$

on tire, quand $z = \rho$, puisque $u(\rho) = \tau$ et car τ vérifie $\partial_u N(u,\tau) = 0$ et donc $\phi(\tau) = \frac{\tau\phi'(\tau)}{c}$, l'expression voulue. □

Nous aurons besoin du lemme suivant dans la prochaine proposition.

Lemme 3.2. *Si ϕ est une série à coefficients positifs (non constante et non nulle en 0) alors pour tout $x > 0$ et à l'intérieur du disque de convergence, on a*

$$\phi(x)(x\phi''(x) + \phi'(x)) > x{\phi'}^2(x)$$

Preuve. Notons

$$\phi(x) = \sum_{n\geq 0} a_n x^n.$$

Nous allons en fait montrer l'inégalité

$$x\phi(x)(x\phi''(x) + \phi'(x)) > x^2{\phi'}^2(x).$$

On a d'une part

$$\begin{aligned}
\phi(x)(x^2\phi''(x) + x\phi'(x)) &= \Big(\sum_{n\geq 0} a_n x^n\Big)\Big(x^2\sum_{n\geq 2} a_n n(n-1)x^{n-2} + x\sum_{n\geq 1} a_n n x^{n-1}\Big) \\
&= \Big(a_0 + a_1 x + \sum_{n\geq 2} a_n x^n\Big)\Big(a_1 x + \sum_{n\geq 2} a_n n^2 x^n\Big) \\
&= a_0\Big(a_1x + \sum_{n\geq 2} a_n n^2 x^n\Big) + (a_1x)^2 + a_1 x\sum_{n\geq 2} a_n n^2 x^n + \Big(\sum_{n\geq 2} a_n x^n\Big)\Big(a_1 x + \sum_{n\geq 2} a_n n^2 x^n\Big) \\
&> (a_1x)^2 + a_1 x\sum_{n\geq 2} a_n n^2 x^n + \Big(\sum_{n\geq 2} a_n x^n\Big)\Big(a_1 x + \sum_{n\geq 2} a_n n^2 x^n\Big)
\end{aligned}$$

(cette dernière minoration a lieu car, conformément aux hypothèses, on a a_0 et au moins un autre coefficient qui sont différents de 0) et d'autre part

$$x^2{\phi'}^2(x) = \Big(a_1 x + \sum_{n\geq 2} a_n n x^n\Big)^2 = (a_1x)^2 + 2a_1 x\sum_{n\geq 2} a_n n x^n + \Big(\sum_{n\geq 2} a_n n x^n\Big)^2$$

or

$$\sum_{n\geq 2} a_n n^2 x^n \geq \sum_{n\geq 1} 2n a_n x^n$$

et sous réserve que l'on montre

$$\Big(\sum_{n\geq 2} a_n x^n\Big)\Big(\sum_{n\geq 2} a_n n^2 x^n\Big) \geq \Big(\sum_{n\geq 2} a_n n x^n\Big)^2$$

on aura bien l'inégalité voulue. Or en développant le produit et en comparant les séries termes à terme, on voit que les coefficients de la série de droite minorent ceux de la série de gauche, *i.e.* :

$$\sum_{k=2}^{n} a_{n-k}(n-k)^2 a_k \geq \sum_{k=2}^{n} a_{n-k}(n-k) a_k k.$$

Pour le prouver, regroupons les termes par deux (en boustrophédon), il suffit de prouver

$$((n-k)^2+k^2)a_{n-k}a_k \geq 2k(n-k)a_k a_{n-k},$$

ce qui est trivial car on reconnaît là le développement de

$$(n-2k)^2 a_k a_{n-k} \geq 0.$$

On a donc montré l'inégalité suivante pour toute série ϕ à coefficients positifs :

$$\phi(x)(x\phi''(x)+\phi'(x)) > x{\phi'}^2(x).$$

Remarque :
En posant $f := \phi$ et $g := x\phi'$, l'inégalité démontrée dans ce lemme est équivalente à $fg' > f'g$ *i.e.* $(g/f)' > 0$ ce qui implique la croissance de $\frac{x\phi'}{\phi}$ et justifie donc notre remarque à la fin de la preuve du lemme 2.1. ☐

Nous allons maintenant regarder le développement en série de Puiseux de u.

Proposition 3.3. *Soit u une fonction définie par $u^c = z\phi(u)$ (avec ϕ admissible) de rayon de convergence ρ et $u(\rho) = \tau$. On a alors $u \in \mathbb{C}[[\sqrt{\rho - z}]]$, c'est-à-dire qu'on a un développement en série de Puiseux du type suivant :*

$$u - \tau = d_1(\rho-z)^{1/2} + d_2(\rho-z) + d_3(\rho-z)^{3/2} + d_4(\rho-z)^2 + \cdots$$

Preuve. Posons

$$H(u) := \frac{\tau^c}{\phi(\tau)} - \frac{u^c}{\phi(u)}.$$

On a donc $H(u) = \rho - z$ d'une part et si l'on développe H en série de Taylor au point $u = \tau$, on obtient :

$$\rho - z = H(u) = H(\tau) + \frac{1}{1!}H'(\tau)(u-\tau) + \frac{1}{2!}H''(\tau)(u-\tau)^2 + \frac{1}{3!}H'''(\tau)(u-\tau)^3 + \cdots$$

Or on a $H(\tau) = 0$ et $H'(\tau) = 0$. On a aussi $H''(\tau) \neq 0$ car

$$H''(u) = \frac{u^c}{\phi(u)^3}\Big(-\frac{c^2\phi(u)^2}{u^2} + \frac{c\phi(u)^2}{u^2} + \frac{2c\phi'(u)\phi(u)}{u} - 2\phi'(u)^2 + \phi''(u)\phi(u)\Big)$$

et donc pour $u > 0$

$$H''(\tau) = 0 \Longrightarrow -c^2\phi(\tau)^2 + c\phi(\tau)^2 + 2c\tau\phi(\tau)\phi'(\tau) - 2\tau^2\phi'(\tau)^2 + \tau^2\phi''(\tau)\phi(\tau) = 0$$

comme de plus $c = \frac{\tau\phi'(\tau)}{\phi(\tau)}$ on aurait alors, après simplifications,

$$-\tau\phi'(\tau)^2 + \tau\phi''(\tau)\phi(\tau) + \phi(\tau)\phi'(\tau) = 0$$

ce qui est impossible d'après le lemme précédent. Ainsi

$$\rho - z = \frac{1}{2!}H''(\tau)(u-\tau)^2 + \frac{1}{3!}H'''(\tau)(u-\tau)^3 + \cdots$$

d'où on tire

$$\pm\sqrt{\rho - z} = \sqrt{\frac{H''(\tau)}{2}}\,(u-\tau)\,\sqrt{1 + \alpha_1(u-\tau) + \alpha_2(u-\tau)^2 + \cdots}$$

En fait il faut prendre $-\sqrt{\rho - z}$ car $u(z)$ croît quand $z \to \rho^-$ (cela se voit bien par étude des deux points d'intersections des courbes u^c et $z\phi(u)$ pour u réel positif et plus petit que τ). Après quelques manipulations triviales (et valides car $H''(\tau) \neq 0$

et car le radical de la somme infinie est non nul puisqu'équivalent à $(1+o(u))^{-1/2}$ pour u au voisinage de τ), on obtient :

$$u-\tau = -\sqrt{\frac{2}{H''(\tau)}}\sqrt{\rho-z}\,\frac{1}{\sqrt{1+\alpha_1(u-\tau)+\alpha_2(u-\tau)^2+\cdots}}.$$

En prenant maintenant z tel que $u(z)$ soit au voisinage de τ, c'est-à-dire pour z au voisinage de ρ, on peut encore écrire

$$u-\tau = -\sqrt{\frac{2}{H''(\tau)}}\sqrt{\rho-z}\,\left(1+\beta_1(\rho-z)+\beta_2(\rho-z)^2+\cdots\right)$$

soit encore

$$u-\tau = d_1(\rho-z)^{1/2}+d_2(\rho-z)+d_3(\rho-z)^{3/2}+d_4(\rho-z)^2+\cdots$$

avec $d_1 := -\sqrt{\frac{2}{H''(\tau)}}$ etc. □

4. Asymptotique en $A\rho^{-n}(\pi n^3)^{-1/2}$

Une série h est dite p-périodique s'il existe un entier m et une série S telle que h vérifie

$$h(z) = z^m S(z^p),$$

p étant le plus grand entier tel que cette propriété soit satisfaite. S'il n'existe pas de $p>1$ tel que h admette une telle décomposition, alors h est dite apériodique. Pour nos marches, puisque $\phi(0)\neq 0$ (resp. $F(0,0)\neq 0$), la périodicité de ϕ (resp. de F), s'écrit donc $\phi(u)=S(u^p)$ (resp.$F(z,0)=S(z^p)$).

Rappelons que notre série génératrice des excursions est donnée par

$$F(z,0) = \frac{(-1)^{c+1}}{s_c z}\prod_{i=1}^{c}(u_i(z))$$

(le produit de l'ensemble des branches finies en 0 de $u^c = z\phi(u)$). Nous noterons ci-dessous F pour $F(z,0)$ (les excursions).

Lemme 4.1. *La série génératrice F des excursions est périodique si et seulement si ϕ est périodique et que c n'est pas un multiple de p.*

Preuve. La preuve est facilitée quand on utilise la vision combinatoire de ϕ et de F.

- ϕ périodique implique F périodique. Considérons les degrés des monômes de ϕ. L'ensemble des différences entre ces degrés sont des valeurs multiples de p (la plus petite différence) puisque ϕ est est p-périodique. C'est clairement équivalent à dire que les différences entre les longueurs des différents sauts sont des multiples de p. Soit l_g la longueur (éventuellement nulle !) du plus petit saut possible vers la gauche. Tous les sauts vers la gauche sont donc de longueur $l_g - kp$ (pour un entier positif k dépendant du saut) et tous les sauts vers la droite sont donc de longueur $-l_g+kp$ (pour un entier positif k dépendant du saut). Qu'un chemin soit une excursion implique que les n_g pas vers la gauche et les n_d pas vers la droite se compensent, *i.e.* : $-n_g l_g - k_1 p - (n_d l_g + k_2 p) = 0$. Ainsi, en passant modulo p, on obtient $l_g(n_g+n_d)\equiv 0[p]$. On a alors deux possibilités.
  Soit n_g+n_d est un multiple de p et alors toutes les excursions ont une longueur multiple de p et ainsi F est p-périodique (car F est la fonction qui, rappelons-le, donne le dénombrement des excursions par rapport à leur longueur).
  Soit l_g est un multiple de p (ce qui est équivalent c multiple de p) et alors $\phi(u^{1/p})$ est un polynôme apériodique, caractéristique d'une marche qui donnera la même série génératrice des excursions (et cette dernière est apériodique comme le montre le point suivant).

- Soit ϕ apériodique et supposons F p-périodique. Puisque ϕ est apériodique, alors il existe un saut vers la gauche de longueur s_g et un saut vers la droite de longueur s_d qui ne sont pas dans la même classe d'équivalence modulo p. Ainsi $s_d = -s_g + kp + q$ avec $0 < q < p$. Or s_g sauts consécutifs de longueur s_d suivi de s_d sauts consécutifs de longueur s_g constitue une excursion. Sa longueur est $s_g + s_d = kp + q$, nombre qui n'est pas congru à 0 modulo p, ce qui contredit le fait que F soit p-périodique. On a donc "ϕ apériodique" implique "F apériodique.".

$\square$

Lemme 4.2. *Si F, définie comme précédemment, est apériodique alors F a une unique singularité dominante.*

Preuve. La fonction F est à coefficients positifs donc pour tout réel $r > 0$ fixé et dans le disque de convergence, on a

$$|F(z)| < F(r) \text{ pour } |z| = r, z \neq r.$$

En effet, il suffit d'appliquer, pour un tel z, l'inégalité triangulaire à $F(z) = \sum_n f_n z^n$ et comme F est apériodique, on a au moins un terme complexe (non réel), d'où une inégalité stricte.

Ainsi $F(z)$ a une singularité en ρ si et seulement si $\tilde{F}(z) := zF(z^c)$ a une singularité en ρ, ce qui est équivalent à ce que le produit des U_i ait une singularité en ρ ce qui implique finalement qu'au moins l'un des U_i a une singularité en ρ. Par ailleurs, "réciproquement", si l'un des U_i a une singularité en un point ρ, alors celle-ci ne peut pas être "tuée" par le produit des U_i restants, car sinon, ce point de branchement devrait également apparaître dans un autre U_i, mais alors de part les invariances par substitution par tw^j, on aurait plusieurs singularités dominantes, ce qui, on l'a vu, est incompatible avec l'apériodicité.

$\square$

Théorème 4.3. *Pour les séries génératrices F de nos marches (avec ϕ admissible et apériodique), on a*

$$[z^n]F(z) \sim \frac{A}{\rho^n \sqrt{\pi n^3}}$$

où ρ est le rayon de convergence de $F(z)$ et où A est un nombre réel explicitable. De plus, quand ϕ est algébrique (e.g. *ϕ polynôme), A et ρ sont des nombres algébriques.*

Preuve. L'asymptotique s'exprime en fonction des singularités dominantes or l'apériodicité de ϕ entraîne par le lemme précédent que $F(z)$ n'a qu'une seule singularité dominante. Or nous avons vu à la proposition 3.3 que ces u_i étaient du type

$$u_i - \tau = d_1(\rho - z)^{1/2} + d_2(\rho - z) + d_3(\rho - z)^{3/2} + d_4(\rho - z)^2 + \cdots .$$

Il suffit alors d'appliquer le résultat usuel

$$[z^n](1-z)^a = \frac{n^{-a-1}}{\Gamma(a)} + O(n^{-a-2})$$

en $a = \frac{1}{2}$ (on sait que $\Gamma(\frac{1}{2}) = \sqrt{\pi}$) pour obtenir que le premier terme du développement de Puiseux de la proposition précédente détermine le comportement asymptotique de $[z^n]F(z)$. Nous rappelons que $\rho = \frac{\tau^c}{\phi(\tau)}$ et que τ est donné par $c\phi(\tau) - \tau\phi'(\tau) = 0$; ainsi, quand ϕ est un polynôme, τ est algébrique et donc ρ et A aussi.

$\square$

5. Asymptotique via le point col

Une autre possibilité pour établir l'asymptotique de nos marches aléatoires consiste, non pas à localiser la singularité dominante et l'utiliser pour y établir le développement de Puiseux, mais à calculer une intégrale qui passe par le point col positif, ce qui donne l'aymptotique en $A\rho^{-n}(\pi n^3)^{-1/2}$ et *a posteriori* montrer une contradiction s'il y avait une ou plusieurs singularités dominantes de module plus petit que ρ. Ce raisonnement est dû à Bruno Salvy.

Lemme 5.1. *Posons $U_1(t) := u_1(t^c)$, où u_1 est la branche réelle sur l'axe réel positif. Les coefficients de F sont donnés par :*

$$[z^n]F(z) = [t^{(n+1)c}]\frac{(-1)^{c+1}}{s_c}U_1(t)U_1(wt)\dots U_1(w^{c-1}t).$$

Preuve. Rappelons l'équation de départ

$$u^c = z\phi(u).$$

Le changement de variable $z := t^c$, aboutit, si l'on pose de plus

$$U_i(t) := u_i(t^c)$$

et en prenant w une racine c-ième primitive de l'unité, à :

$$U_i := w^{i-1}t\phi^{1/c}(U_i).$$

Cette nouvelle équation définit bien une unique fonction analytique U_i, car la fonction $U \mapsto t\phi^{1/c}(U)$ est contractante en tant que série de $\mathbb{C}[[t]][[U]]$. Les U_i sont des fonctions analytiques en 0. L'étude asymptotique directe des coefficients de U_i à partir de l'égalité précédente n'est *a priori* pas aisée car $\phi^{1/c}$ n'est plus forcément à coefficients positifs (par exemple $\phi = 1 + u + u^2 + u^3$ conduit à $\phi^{1/2} = 1 + \frac{1}{2}u + \frac{3}{8}u^2 + \frac{5}{16}u^3 - \frac{29}{128}u^4 + O(u^5)$).

Notons par ailleurs que $U_i(wt) = U_{i+1}(t)$, puisque $V = U_i(wt)$ vérifie $V = w^it\phi^{1/c}(V)$ et on vient de voir que cette équation a une seule solution, à savoir U_{i+1}. Ainsi

$$U_1(t).U_2(t)\cdots U_c(t) = U_1(t).U_1(wt)\dots U_1(w^{j-1}t).$$

Par ailleurs on a

$$[z^n]F(z) = [t^{cn}]F(t^c) = [t^{cn}]\frac{(-1)^{c+1}}{s_ct^c}\prod_{i=1}^{c}u_i(t^c) = [t^{nc}]\frac{(-1)^{c+1}}{s_ct^c}U_1(t).U_2(t)\cdots U_c(t)$$

□

Nous sommes donc ramenés à examiner les coefficients de U_1 un rang sur c. C'est ce que nous allons faire avec une technique similaire à celle employée pour la démonstration de la formule d'inversion de Lagrange.

Lemme 5.2. *On a le résultat asymptotique suivant :*

$$[t^{nc}]U_1(t) = \rho^{-n}\frac{n^{-1/2}}{\sqrt{2\eta\pi}}(1 + o(1)).$$

Preuve. D'après la formule de Cauchy, puisque $U(t) := U_1(t)$ est analytique dans un voisinage de 0 et en particulier dans le cercle (que nous noterons γ) de centre 0 et de rayon τ (τ comme défini au lemme 2.1), nous avons que

$$[t^{nc}]U = \frac{1}{2i\pi}\int_\gamma \frac{U(t)}{t^{nc+1}}dt.$$

Le changement de variable $t = \frac{U}{\phi^{1/c}(U)}$ implique

$$dt = \frac{dU}{\phi^{1/c}(U)} - \frac{U\phi'(U)dU}{c\phi^{1+1/c}(U)}$$

ce qui donne

$$\begin{aligned}[t^{nc}]U &= \frac{1}{2i\pi}\int_\gamma \frac{\phi^n(U)}{U^{nc}}\left(1 - \frac{U\phi'(U)}{c\phi(U)}\right)dU \\ &= \frac{1}{2i\pi}\int_\gamma \left(\frac{\phi(U)}{U^c}\right)^n dU - \frac{1}{c}\frac{1}{2i\pi}\int_\gamma \frac{\phi^{n-1}(U)\phi'(U)}{U^{nc-1}}dU \\ &= [U^{nc-1}]\phi^n(U) - \frac{1}{c}[U^{nc-2}]\frac{1}{n}(\phi^n(U))'\end{aligned}$$

Or on a $[z^n]g'(z) = (n+1)[z^n]g(z)$ donc

$$\begin{aligned}[t^{nc}]U &= [U^{nc-1}]\phi^n(U) - \frac{nc-1}{nc}[U^{nc-1}]\phi^n(U) = [U^{nc-1}]\frac{1}{cn}\phi^n(U) \\ &= \frac{1}{cn}\frac{1}{2i\pi}\int_\gamma \frac{\phi^n(U)}{U^{nc}}dU\end{aligned}$$

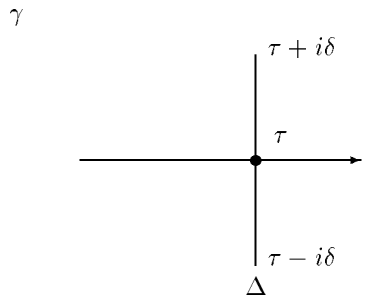


FIGURE 1. Le contour d'intégration peut être ramené à un petit segment.

On utilise alors la méthode du point col avec

$$G_n(U) := \left(\frac{\phi(U)}{U^c}\right)^n = \exp(n(\log\phi(U) - c\log U))$$

Rappelons que τ est l'unique point col réel positif de $\frac{\phi(u)}{u^c}$. Le développement de Taylor de $H := \log\frac{\phi(U)}{U^c}$ en ce point τ donne

$$H(U) = H(\tau) + H'(\tau)(U-\tau) + H''(\tau)\frac{(U-\tau)^2}{2} + \cdots$$

mais comme $H'(U) = \frac{U\phi'(U) - c\phi(U)}{U\phi(U)}$ et $H''(U) = \frac{c\phi(U)^2 - U^2\phi'(U)^2 + U^2\phi''(U)\phi(U)}{U^2\phi^2(U)}$, on a donc $H'(\tau) = 0$ et $\eta := H''(\tau) \neq 0$ (pour établir ce dernier point, il suffit d'appliquer le lemme 3.2 en $u = \tau$ et en se souvenant qu'on a aussi $u\phi'(\tau) = c\phi(\tau)$). Ainsi, pour U au voisinage de τ

$$G_n(U) = \exp\left(n\left(\log\phi(\tau) - c\log(\tau) + \eta\frac{(U-\tau)^2}{2} + \cdots\right)\right).$$

L'asymptotique de $[t^{nc}]U$ découle de celle de $\int_\gamma G_n(U)$ or l'intégrale de $G_n(U)$ sur γ est concentrée autour du point col τ, nous n'avons pas écarté la possibilité qu'il existe d'autres points cols sur ce cercle, mais, en tout état de cause, ceux-ci seraient moins hauts (voir 4.2) que $|\frac{\phi(\tau)}{\tau^c}|$ et donc leur contribution serait bien moindre que celle du point col réel positif, quand n est grand. Ainsi, on ne modifie pas la valeur asymptotique de l'intégrale en se limitant à un petit arc cercle de γ, centré en τ, ni même en prenant un petit segment Δ, centré en τ et perpendiculaire à l'axe réel (voir la figure 5.2), dans lequel le développement de Taylor précédent est valide.

Ainsi,

$$\int_\gamma G_n(U)dU = \left(\frac{\phi(\tau)}{\tau^c}\right)^n \int_\Delta \exp(n\eta\frac{(U-\tau)^2}{2} + \cdots)dU$$

$$= \rho^{-n}\int_\Delta \exp(n\eta\frac{(U-\tau)^2}{2})dU(1+o(1)).$$

On pose alors $U = \tau + it$. Or, pour δ tel que $\delta \to 0$ et $\delta\sqrt{n} \to \infty$, on a :

$$\int_{-\delta}^{\delta} \exp(-n\eta t^2/2)dt = \frac{1}{\sqrt{n}}\int_{-\delta\sqrt{n}}^{\delta\sqrt{n}} \exp(-\eta t^2/2)dt$$

$$= \frac{1}{\sqrt{n}}\int_{-\infty}^{\infty} \exp(-\eta t^2/2)dt(1+o(1)) = \sqrt{\frac{2\pi}{\eta n}}(1+o(1))$$

Ainsi, puisque $dU = idt$, on a

$$[t^{nc}]U(t) = \frac{1}{cn}\frac{1}{2i\pi}\int_\gamma G_n(U)dU = \frac{1}{cn}\frac{1}{2i\pi}\rho^{-n} i\sqrt{\frac{2\pi}{\eta n}}(1+o(1))$$
$$[t^{nc}]U = \rho^{-n}\frac{n^{-1/2}}{\sqrt{2\eta\pi}}(1+o(1))$$

□

Jusqu'ici, avec cette méthode nous ne savions pas pas où se trouvai(en)t la (les) singularité(s) dominante(s), maintenant nous pouvons déduire du résultat précédent que $z = \rho$ est bien une singularité dominante.

Corollaire 5.3. *La fonction U (et donc F) n'a pas de singularité de module inférieur à ρ.*

Preuve. Si l'on suppose que ρ n'est pas une singularité dominante de U, on a alors m singularités dominantes de module $r < \rho$ et l'asymptotique des coefficients est alors

$$[t^n]U(t) = r^{-n}n^{-\alpha}\left(C_1\exp(in\theta_1) + \cdots + C_m\exp(in\theta_m)\right) + o(r^{-n}n^{-\alpha})$$

(avec au moins un C_i non nul)

On devrait alors avoir coïncidence entre les deux asymptotiques trouvées (du moins pour n multiple de c), c'est-à-dire :

$$\begin{aligned}[t^{cn}]U(t) &= r^{-cn}(cn)^{-\alpha}\left(C_1\exp(icn\theta_1) + \cdots + C_m\exp(icn\theta_m)\right) + o(r^{-cn}(cn)^{-\alpha})\\ &= \frac{1}{cn}\left(\frac{\phi(\tau)}{\tau^c}\right)^n\frac{n^{-1/2}}{\sqrt{2\eta\pi}}(1+o(1)) =: B\rho^{-n}n^{-3/2}(1+o(1))\end{aligned}$$

d'où, en multipliant par $r^{cn}(cn)^\alpha$,

$$C_1\exp(icn\theta_1) + \cdots + C_m\exp(icn\theta_m) = o(1).$$

Donc pour tout $\epsilon > 0$, à partir d'un certain n_0, on a, pour $n > n_0$,

$$(C_1\exp(icn\theta_1) + \cdots + C_m\exp(icn\theta_m)) < \epsilon.$$

En posant $K_k := C_k\exp(in_0c\theta_k)$ et $x_k := \exp(ic\theta_k)$, on aboutit au système

$$\begin{cases} K_1x_1 + \cdots + K_kx_m < \epsilon \\ \vdots \\ K_1x_1^2 + \cdots + K_kx_m^2 < \epsilon \\ \vdots \\ K_1x_1^n + \cdots + K_kx_m^n < \epsilon \\ \vdots \end{cases}$$

On a donc en particulier : il existe $\epsilon_1, \ldots, \epsilon_m$ des réels positifs $< \epsilon$ tels que :

$$\begin{pmatrix} x_1 & \ldots & x_m \\ x_1^2 & \ldots & x_m^2 \\ & \vdots & \\ x_1^m & \ldots & x_m^m \end{pmatrix} \cdot \begin{pmatrix} K_1 \\ K_2 \\ \vdots \\ K_m \end{pmatrix} = \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_m \end{pmatrix}$$

On montre par récurrence qu'une telle matrice "à la Vandermonde" a pour déterminant $\Delta = x_1 \ldots x_m \prod_{1\le i<j\le m}(x_j - x_i) \neq 0$ d'après nos hypothèses (puisque les θ_k sont distincts car correspondant à des singularités d'amplitudes distinctes). Notons que ce déterminant Δ est indépendant de ϵ. La méthode de Cramer donne

$$K_k = \frac{\Delta_k}{\Delta}$$

où Δ_k est le déterminant de la matrice de Vandermonde ci-dessus où l'on a remplacé la k-ième colonne par par le vecteur colonne $\begin{pmatrix} \epsilon_1 \\ \vdots \\ \epsilon_m \end{pmatrix}$. Bien qu'il soit également possible d'exhiber une factorisation "esthétique" de Δ_k, nous nous contenterons de la majoration triviale :

$$|\Delta_k| < \epsilon_1 M_1 + \cdots + \epsilon_m M_m < \epsilon m(m-1)!$$

où les M_i sont les valeurs absolues des m sous-déterminants (notons que les M_i sont majorés par $(m-1)!$ puisque les coefficients des matrices associées sont tous de module 1) Ainsi ce déterminant Δ_k tend vers 0 quand ϵ tend vers 0. Finalement, puisque ϵ peut être choisi aussi petit que l'on veut, les K_k peuvent être bornés par des nombres arbitrairement petits, d'où $K_1 = \cdots = K_m = 0$ et on a alors obtenu une contradiction avec l'existence d'au moins un $C_i \neq 0$. On ne peut donc pas avoir une ou plusieurs singularités dominantes de module inférieur à ρ. □

Au final, la méthode du col donne une deuxième démonstration du résultat suivant.

Théorème 5.4. *L'asymptotique du nombre d'excursions est $A\rho^{-n}n^{-3/2}$.*

PREUVE. En posant $\alpha := \prod_{j=2}^{c} U(w^j\rho)$, on a

$$[z^n]F(z) = [t^{cn}]F(t^c) = [t^{nc}]\frac{(-1)^{c+1}}{s_c t^c}U(t)\alpha = \alpha\frac{1}{cn}\left(\frac{\phi(\tau)}{\tau^c}\right)^n \frac{n^{-1/2}}{\sqrt{2\eta\pi}}(1+o(1))$$

$$[z^n]F(z) = \frac{A}{\rho^n\sqrt{\pi n^3}}\,(1+o(1))$$

avec $A = \frac{\alpha}{c\sqrt{2\eta}}$. □

6. La formule de Soria-Flajolet

Les coefficients de f quand f est définie de manière implicite par l'équation $f(z) = z\phi(f(z))$, où $\phi(y)$ est une fonction analytique ($\phi(0) \neq 0$), sont donnés par la formule d'inversion de Lagrange-Bürmann :

$$[z^n]\,f(z) = \frac{1}{n}[y^{n-1}]\phi^n(y).$$

Michèle Soria et Philippe Flajolet **[18]** ont montré qu'on a la généralisation suivante :

Théorème 6.1. *Si $\phi(z,y)$ est un polynôme bivarié tel que*

$$\phi(0,0) = 0\ ,\ \phi'_y(0,0) = 0\ \textit{et}\ \phi(z,0) \neq 0$$

alors les coefficients de Taylor de f définie par $f(z) = \phi(z, f(z))$ sont

$$[z^n]\,f(z) = \sum_{m\geq 1}\frac{1}{m}[z^n y^{m-1}]\phi^m(z,y).$$

PREUVE. Voir le papier de Flajolet et Soria **[18]**. □

Ainsi toute fonction algébrique a des coefficients qui s'expriment en fait sous la forme d'une somme finie de multinomiaux, multipliés par des monômes. C'est donc le cas pour les fonctions que nous obtenons, toutefois une telle formule permet d'obtenir une forme close qui n'est pas toujours la plus utile pour une étude asymptotique.

7. Accélération de convergence : la méthode de Romberg

La formule précédente ne permet toutefois pas une étude asymptotique simplifiée. Pour nos marches, nous avons montré que l'asymptotique était en $\frac{A}{\rho^n\sqrt{\pi n^3}}$ où A est un nombre algébrique, qui peut éventuellement être difficile à expliciter ou à approcher.

Une autre stratégie pour en obtenir quelques décimales serait, comme bien souvent on a les premiers termes du développement de la série génératrice , d'utiliser la méthode de Romberg d'accélération de convergence.

En effet, lorsque l'on a les premiers termes d'une suite dont on sait que le terme principal vaut asymptotiquement

$$u_n = \lambda + \frac{a}{n} + \frac{b}{n^2} + \cdots$$

sans toutefois connaître ni λ, ni a, ni b, le principe de Romberg repose sur le fait que la suite $u_n^{(1)} = 2u_{2n} - u_n$ a un terme principal qui vaut asymptotiquement

$$u_n = \lambda + \frac{b'}{n^2} + \cdots .$$

Ainsi, la suite $(u_n^{(1)})$ converge plus vite vers λ que la suite (u_n). On peut bien sûr "itérer" le procédé (on a $u_n^{(2)} = \frac{4}{3}u_{2n} - \frac{1}{3}u_n$).

Si l'on pose

$$u_n := \frac{a_n}{\rho^n \sqrt{\pi n^3}},$$

on peut par exemple appliquer cette méthode à la série génératrice donnant les excursions positives de Dyck (on s'autorise un pas en avant et un pas en arrière, on reste toujours positif et on finit sur l'origine, voir la section 3.1).

On a $F(z,0) = \frac{u_1}{z}$ de rayon de convergence $\frac{1}{2}$. Bien sûr, sur cet exemple, on a une forme close de la série génératrice sur laquelle on pourrait faire de l'analyse de singularités, on pourrait également appliquer une variante du théorème 4.3 lorsqu'il y a plusieurs singularités dominantes, mais jouons pour l'instant les numériciens et essayons de déterminer la constante A (que l'on attend algébrique), *i.e.* λ.

Si l'on connaît les 256 premiers coefficents de la série génératrice (en fait il nous faut 512 coefficients car un sur deux est nul), on a $u_{256} = 2,816046212$, avec une première passe de Romberg on obtient $u_{128}^{(1)} = 2,828330470$ et avec une deuxième passe on a $u_{64}^{(2)} = 2,828425637$, tiens, tiens... Mais si ma mémoire est bonne on a $2\sqrt{2} = 2,828427124$, la conjecture $A = 2\sqrt{2}$ s'impose et la méthode de Romberg nous a donc permis de "gratter" quatres décimales.

8. Nombre moyen d'effleurements

Nous appellerons "effleurement" le fait de passer par l'origine (ou par l'axe des abscisses en dimension 2, etc). Nous allons utiliser une nouvelle variable t dans nos séries génératrices pour pourvoir obtenir un décompte des passages par l'origine.

On a donc cette fois une récurrence entre f_{n+1} et f_n qui subit la modification suivante, pour justement tenir compte de ce qui ce passe quand on arrive en u^0 :

$$\begin{gathered} f_{n+1} = P(u)f_n - \{u^{<0}\}P(u)f_n - [u^0]P(u)f_n + t[u^0]P(u)f_n \\ f_{n+1}(u,t) = P(u)f_n(u,t) + (t-1)[u^0]P(u)f_n(u,t) - \{u^{<0}\}P(u)f_n(u,t) \\ F(z,u,t)(1 - zP(u)) = 1 + z(t-1)\sum_1 - z\sum_2 \end{gathered}$$

où $\sum_1$ et $\sum_2$ font intervenir des dérivées de F en $u = 0$, on résout alors cette équation fonctionnelle comme expliqué précédemment.

Le nombre moyen d'effleurements est

$$\mu_n = \frac{[z^n]\partial_t F(z,1,1)}{[z^n]F(z,1,1)}$$

pour les chemins de longueur n et

$$\mu_n = \frac{[z^n]\partial_t F(z,0,1)}{[z^n]F(z,0,1)}$$

pour les excursions de longueur n.

Par exemple avec $P(u) = u + \frac{1}{u}$, on a :

$$F(z,u,t)(u - z(1+u^2)) = u + zu(t-1)F'(z,0,t) - zuu^{-1}[u^{-1}]\sum_n Pz^n f_n$$

$$F(z,u,t)(u - z(1+u^2)) = u + zu(t-1)F'(z,0,t) - z[u^0]\sum_n z^n f_n$$

$$F(z,u,t)(u - z(1+u^2)) = u + zu(t-1))F'(z,0,t) - zF(z,0,t) \tag{5}$$

en annulant le noyau, on obtient, en fonction de la racine $u_1(z)$ du noyau,

$$zF(z,0,t) = u_1 + zu_1(t-1)F'(z,0,t)$$

en dérivant l'équation (5) par rapport à u, puis avec $u = 0$, on a aussi

$$F'(z,0,t) = \frac{F(z,0,t)-1}{zt}$$

d'où

$$F(z,0,t) = \frac{u_1}{(1-t)u_1 + zt}$$

et ainsi on a

$$F(z,u,t) = \frac{u^2(1-t)(u_1-z) + zu_1}{((1-t)u_1+zt)(zu^2-u+z)}$$

Remarquons que l'on retrouve en passant (quand $u = 0$ et $t = 1$) que le nombre d'excursions de longueur n est

$$E_n := [z^n]F(z,0,1) = [z^n]\frac{u_1}{z}.$$

Dans cet exemple, on a $u_1(z) = \frac{1-\sqrt{1-4z^2}}{2z}$ et le nombre moyen d'effleurements pour une excursion de longueur $2k$ est donc :

$$\mu_{2k} = \frac{[z^{2k}]\partial_t F(z,0,1)}{[z^{2k}]F(z,0,1)} = \frac{[z^{2k}]\frac{u_1(z)(u_1(z)-z)}{z^2}}{[z^{2k}]\frac{u_1(z)}{z}}$$

$$= \frac{\frac{\binom{2(k+1)}{k+1}}{k+2} - \frac{\binom{2k}{k}}{k+1}}{\frac{\binom{2k}{k}}{k+1}} = \frac{(k+1)(2k+1)(2k+2)}{(k+2)(k+1)^2} - 1$$

$$= 3\frac{k}{k+2} = 3\sum_{p\geq 0}\frac{(-1)^p 2^p}{k^p} = 3 - \frac{6}{k} + \frac{24}{k^2} + O(\frac{1}{k^3}).$$

Nous finirons donc sur cet apophtegme mémorable : "L'ivrogne frappe toujours trois fois."

CHAPITRE 5

Annexes : sessions Maple

Il sera fait usage du package GFUN de Bruno Salvy et Paul Zimmermann qui se trouve dans la share library de Maple, se référer à l'aide en ligne de Maple où à l'article [**31**] pour des renseignements sur ce package.

Exemple de dénombrements de marches unidimensionnelles avec pas bornés

```
> restart:with(share):with(gfun):

See ?share and ?share,contents for information about the share library
Share Library:  gfun
Authors: Salvy, Bruno and Zimmermann, Paul.
Description:  package for the manipulation
and discovery of generating functions
```

On s'autorise les pas

```
> P:=u^(-1)+u^2;
```

$$P := \frac{1}{u} + u^2$$

et notre série génératrice est donnée par le $F(z,1)$ de la marche $P(u)$ unidimensionnelle. Voici les racines du noyau

```
if ldegree(P)>=0 then c:=0 else
c:=-ldegree(P) fi;
eq:=expand((1-z*P)*u^(c));
preci:=100:L:=convert(algeqtoseries(eq,z,u,preci),polynom):
```

$$c := 1$$

$$eq := u - z - z\,u^3$$

dont on ne retient que les divergentes :

```
ig:=-1:ip:=-1:
Pr:=1:
for k from 1 to nops(L) do if
evalb(abs(limit(L[k],z=0,right))=infinity) then
ig:=ig+1;v[ig]:=L[k]:#print(v[ig]);
P:=indets(v[ig],function);
if P<>{} then
Pr:=Pr*map(`*`,seq(product(subs(P[i]=a,u-v[ig]),
a=P[i]),i=1..nops(P)));
else Pr:=Pr*(u-v[ig]) fi;
# Je ne me sers pas des petites branches
# mais l'obtention de ces branches ne pose pas de pb :
#else ip:=ip+1;u[ip]:=L[k]:
fi od:
```

d'où, avec la formule sous forme de produit des branches divergentes :

```
F1:=1/(coeff(eq,u^degree(eq,u))*subs(u=1,Pr)):
series(F1,z,preci/2);
```

$$\begin{aligned}
&1 + z + 2\,z^2 + 4\,z^3 + 7\,z^4 + 14\,z^5 + 28\,z^6 + 53\,z^7 + 106\,z^8 + 212\,z^9 + 412\,z^{10} + 824\,z^{11} + \\
&\quad 1648\,z^{12} + 3241\,z^{13} + 6482\,z^{14} + 12964\,z^{15} + 25655\,z^{16} + 51310\,z^{17} + 102620\,z^{18} + \\
&\quad 203812\,z^{19} + 407624\,z^{20} + 815248\,z^{21} + 1622744\,z^{22} + 3245488\,z^{23} + 6490976\,z^{24} \\
&\quad + 12938689\,z^{25} + 25877378\,z^{26} + 51754756\,z^{27} + 103262837\,z^{28} + 206525674\,z^{29} \\
&\quad + 413051348\,z^{30} + 824671981\,z^{31} + 1649343962\,z^{32} + 3298687924\,z^{33} + \\
&\quad 6588961208\,z^{34} + 13177922416\,z^{35} + 26355844832\,z^{36} + 52661622556\,z^{37} + \\
&\quad 105323245112\,z^{38} + 210646490224\,z^{39} + 420992149876\,z^{40} + 841984299752\,z^{41} \\
&\quad + 1683968599504\,z^{42} + 3366114432488\,z^{43} + 6732228864976\,z^{44} + \\
&\quad 13464457729952\,z^{45} + 26917790704240\,z^{46} + 53835581408480\,z^{47} + \\
&\quad 107671162816960\,z^{48} + 215273996878961\,z^{49} + \mathrm{O}(z^{50})
\end{aligned}$$

Le dénombrement des effleurements est donné par

```
F0:=1/(coeff(eq,u^degree(eq,u))*subs(u=0,Pr)):
series(F0,z,preci/2);
```

$$1 + z^3 + 3\,z^6 + 12\,z^9 + 55\,z^{12} + 273\,z^{15} + 1428\,z^{18} + 7752\,z^{21} + 43263\,z^{24} + 246675\,z^{27} + 1430715\,z^{30} + 8414640\,z^{33} + 50067108\,z^{36} + 300830572\,z^{39} + 1822766520\,z^{42} + 11124755664\,z^{45} + 68328754959\,z^{48} + \mathrm{O}(z^{51})$$

Exemple de dénombrements de marches en dimension 2 limitées à droite par une droite de pente rationnelle

```
> restart:with(share):with(gfun):
```

```
See ?share and ?share,contents for information about the share library
```

On prend y=(p/q) x comme limite à droite avec

```
> p:=1; q:=1;
```

$$p := 1$$

$$q := 1$$

On s'autorise les pas

```
> P:=v+u;
```

$$P := v + u$$

d'ou

```
> P:=subs({u=u^(-p),v=u^q},P):
```

et notre série génératrice est donnée par le $F(z, 1)$ de la marche $P(u)$ unidimensionnelle. Voici les racines du noyau

```
if ldegree(P)>=0 then c:=0 else
c:=-ldegree(P) fi;
eq:=expand((1-z*P)*u^(c));
preci:=20:
L:=convert(algeqtoseries(eq,z,u,preci),polynom):
```

$$c := 1$$

$$eq := u - z\,u^2 - z$$

dont on ne retient que les divergentes :

```
ig:=-1:ip:=-1:
Pr:=1:
for k from 1 to nops(L) do if
evalb(abs(limit(L[k],z=0,right))=infinity) then
ig:=ig+1;v[ig]:=L[k]:print(v[ig]);
P:=indets(v[ig],function);
if P<>{} then
Pr:=Pr*map(`*`,seq(product(subs(P[i]=a,u-v[ig]),
a=P[i]),i=1..nops(P)));
else Pr:=Pr*(u-v[ig]) fi;
#Je ne me sers pas des petites branches
# mais l'obtention de ces branches ne pose pas de pb :
#else ip:=ip+1;u[ip]:=L[k]:
fi od:
```

d'où, avec la formule sous forme de produit des branches divergentes :

```
F1:=1/(coeff(eq,u^degree(eq,u))*subs(u=1,Pr)):
series(F1,z,preci/2);
```

$$1 + z + 2\,z^2 + 3\,z^3 + 6\,z^4 + 10\,z^5 + 20\,z^6 + 35\,z^7 + 70\,z^8 + 126\,z^9 + \mathrm{O}(z^{10})$$

Le dénombrement des effleurements est donné par

```
F0:=1/(coeff(eq,u^degree(eq,u))*subs(u=0,Pr)):
series(F0,z,preci/2);
```

$$1 + z^2 + 2\,z^4 + 5\,z^6 + 14\,z^8 + \mathrm{O}(z^{10})$$

CHAPITRE 6

Nanobiographies

Avec la participation de :

Niels Henrik Abel (1802–1829)
Alexander Craig Aitken (1895–1967)
Stefan Banach (1892–1945)
Eric Temple Bell (1883–1960)
Joseph Bertrand (1822–1900)
Friedrich Wilhelm Bessel (1784–1846)
Émile Félix Édouard Justin Borel (1871–1956)
Eugène Charles Catalan (1814–1894)
Augustin-Louis Cauchy (1789–1857)
Arthur Cayley (1821–1895)
Noam Chomsky (1928)
Gabriel Cramer (1704–1752)
Peter Gustav Dirichlet (1805–1859)
Paul Erdős (1913–1996)
Leohard Euler (1707–1783)
Lodovico Ferrari (1522–1565)
Leonardo Pisano *alias* Fibonacci (1174–1250)
Jacques Hadamard (1865–1963)
Godfrey Harold Hardy (1877–1947)
Charles Hermite (1822–1901)
Heinz Hopf (1894–1971)
Joseph Louis Lagrange (1736–1813)
Edmond Nicolas Laguerre (1834–1886)
Pierre Alphonse Laurent (1813–1854)
Guillaume de L'Hospital (1661–1704)
Kurt Mahler (1903–1988)
Isaac Newton (1642–1727)
George Pólya (1887–1985)
Alfred Pringsheim (1850–1941)
Victor Puiseux (1820–1883)
Srinivasa Ramanujan (1887–1920)
Alfréd Rényi (1921–1970)
Michel Rolle (1652–1719)
Werner Romberg première moitiée du XX^{e} siècle
Friedrich Wilhelm Karl Ernst Schroeder (1841–1902)
Marcel-Paul Schützenberger (1920–1996)
James Stirling (1692–1770)
Brook Taylor (1685–1731)
Alexandre Vandermonde (1735–1796)
Robert Charles Vaughan (1945)
Karl Weierstraß(1815–1897)
Norbert Wiener (1894–1964)

Bibliographie

[1] Alon (Noga) et Spencer (Joel H.). – *The probabilistic method.* – John Wiley & Sons Inc., New York, 1992, *Wiley-Interscience Series in Discrete Mathematics and Optimization.* With an appendix by Paul Erdős, A Wiley-Interscience Publication.

[2] Barcucci (Elena), Del Lungo (Alberto), Pergola (Elisa), et Pinzani (Renzo). – From Motzkin to Catalan permutations: a “discrete continuity”. *Disc. Math (submitted)*, 1998.

[3] Borel (Émile). – *Leçons sur les séries divergentes.* – Gauthier-Villars, 1928.

[4] Bousquet-Mélou (Mireille). – Communication personnelle, juin 1998.

[5] Chomsky (Noam) et Schützenberger (Marcel-Paul). – The algebraic theory of context-free languages. *In : Computer programming and formal systems*, pp. 118–161. – North-Holland, Amsterdam, 1963.

[6] Comtet (Louis). – *Analyse combinatoire.* – P.U.F, 1970, vol. 1 et 2.

[7] Cori (Robert) et Richard (Jean). – Énumération des graphes planaires à l'aide des séries formelles en variables non commutatives. *Discrete Mathematics*, vol. 2, 1972, pp. 115–162.

[8] Denise (Alain), Gardy (Danièle), et Gouyou-Beauchamps (Dominique). – ECO methodology and generating series. *Preprint*, juin 1998.

[9] Dieudonné (Jean). – *Calcul infinitésimal.* – Hermann, Paris, 1968.

[10] Duchon (Philippe). – On the enumeration and generation of generalized Dyck words. *Actes du colloque SFCA/FPSAC '98*, 1998.

[11] Dumas (Philippe). – *Récurrences Mahlériennes, suites automatiques, études asymptotiques.* – Doctorat de Mathématiques, Université de Bordeaux I, 1993.

[12] Euler (Leonard). – *Introductio in analysin infinitorum.* – 1748.

[13] Fayolle (Guy) et Iasnogorodski (Rudolph). – Solutions of functional equations arising in the analysis of two-server queueing models. *In : Performance of computer systems (Proc. Fourth Internat. Sympos. Modelling Performance Evaluation Comput. Systems, Vienna, 1979)*, pp. 289–303. – North-Holland, Amsterdam, 1979.

[14] Flajolet (P.). – Combinatorial aspects of continued fractions. *Discrete Mathematics*, vol. 32, n° 2, 1980, pp. 125–161.

[15] Flajolet (P.) et Schott (R.). – Non–overlapping partitions, continued fractions, Bessel functions and a divergent series. *European Journal of Combinatorics*, vol. 11, 1990, pp. 421–432.

[16] Flajolet (Philippe). – Excursion generating functions are algebraic, mars 1998.

[17] Flajolet (Philippe) et Sedgewick (Robert). – *The average case of algorithms.* – INRIA, 1993.

[18] Flajolet (Philippe) et Soria (Michèle). – Combinatorial formulæ for coefficients of algebraic generating functions. *Preprint*, 1998.

[19] Foata (Dominique). – *Article "Analyse combinatoire".* – Encyclopædia Universalis, 1974.

[20] Furstenberg (Harry). – Algebraic functions over finite fields. *Journal of Algebra*, vol. 7, 1967, pp. 271–277.

[21] Galton (F.) et Watson (H.W.). – On the probability of the extinction of families. *J. Roy. Anthropol. Inst.*, vol. 4, 1874, pp. 138–144.

[22] Gouyou-Beauchamps (Dominique). – Communication personnelle, avril 1998.

[23] Hardy (G.H.). – *Divergent series.* – Oxford University Press, 1963.

[24] Hille (Einar). – *Analytic function theory. Vol. II.* – Ginn and Co., Boston, Mass.-New York-Toronto, Ont., 1962. Introductions to Higher Mathematics.

[25] Joyal (André). – Une théorie combinatoire des séries formelles. *Advances in Mathematics*, vol. 42, n° 1, 1981, pp. 1–82.

[26] Knuth (Donald E.). – *The art of computer programming. Vol. 1: Fundamental algorithms.* – Addison-Wesley, 1968. exercises 4 and 11, section 2.2.1.

[27] Labelle (Jacques) et Yeh (Yeong-Nan). – Dyck paths of knights moves. *Discrete Applied Mathematics*, vol. 24, 1989, pp. 213–221.

[28] Labelle (Jacques) et Yeh (Yeong-Nan). – Generalized Dyck paths. *Discrete Mathematics*, vol. 82, 1990, pp. 1–6.

[29] Petkovšek (Marko). – The irrational chess knight. *FPSAC '98*, 1998.

[30] Pinzani (Renzo). – Communication personnelle, avril 1998.

[31] Salvy (Bruno) et Zimmermann (Paul). – Gfun: a Maple package for the manipulation of generating and holonomic functions in one variable. *ACM Transactions on Mathematical Software*, vol. 20, n° 2, 1994, pp. 163–177. – ftp://ftp.inria.fr/INRIA/publication/publi-ps-gz/RT/RT-0143.ps.gz.

[32] Stauffer (D.). – *Introduction to Percolation Theory.* – Taylor & Francis, 1992.

[33] Vallée (Brigitte). – The complete analysis of the binary Euclidean algorithm. *In : Proceedings ANTS'98*. – 1998.

[34] Wiener (Norbert). – *The Extrapolation, Interpolation and Smoothing of Stationary Times Series*. – M.I.T. Press, 1949.

[35] Wimp (J.) et Zeilberger (D.). – Resurrecting the asymptotics of linear recurrences. *Journal of Mathematical Analsis and Applications*, vol. 111, 1985, pp. 162–176.