École Jeunes Chercheurs, LaBRI, mars 1999

Combinatoire analytique

et marches aléatoires

Cyril.Banderier@inria.fr

http://algo.inria.fr/banderier/

Projet Algorithmes
INRIA-Rocquencourt (France)

Du discret au continu... et réciproquement !

discret (structure combinatoire)

$\Downarrow$

continu (série de variable complexe)

$\Downarrow$

analyse (éq. fonctionnelle, forme close)

$\Downarrow$

asymptotique (singularités, cols...)

$\Downarrow$

loi limite

$\Downarrow$

discret

Il existe un dictionnaire qui traduit une structure combinatoire en “série génératrice”.

Arbres binaires et nombres de Catalan

arbres binaires, mots de Dyck (mots bien parenthésés) = marche sur $\mathbb{N}$ à pas ± 1.

b_n nombre d'arbres binaires avec n nœuds (internes) sont comptés par $B(z)$, la série génératrice

$$B(z) = \sum_{n \geq 0} b_n z^n$$

lecture directe de la série génératrice à partir de la relation $\mathcal{B} = \mathcal{F} + \mathcal{B} \times \mathcal{N} \times \mathcal{B}$.

$$B(z) = 1 + zB^2(z)$$

$$B(z) = \frac{1 - \sqrt{1-4z}}{2z} = 1+z+2z^2+5z^3+14z^4+\ldots$$

On obtient les nb de Catalan : $\frac{\binom{2n}{n}}{n+1}$ $\square$

But : généralisation avec plus de pas possibles, en autorisant les multiplicités.

Wanted : \$érie génératrice bivariée

mots engendrés grammaire infinie (“régulière”)

$\approx$ traces automate infini

$\approx$ puissances itérées matrice infinie

$\approx$ arbre de possibilités

$\approx$ marches (aléatoires)

Série génératrice d'une ”marche aléatoire sur les entiers” :

$$F(z,u) = \sum_{n\geq 0, k\geq 0} f_{n,k} u^k z^n$$

n code la longueur $\qquad$ k code la hauteur

•$f_{n,k}$ est le nombre de chemins de longueur n qui vont de l'état initial à l'état k.

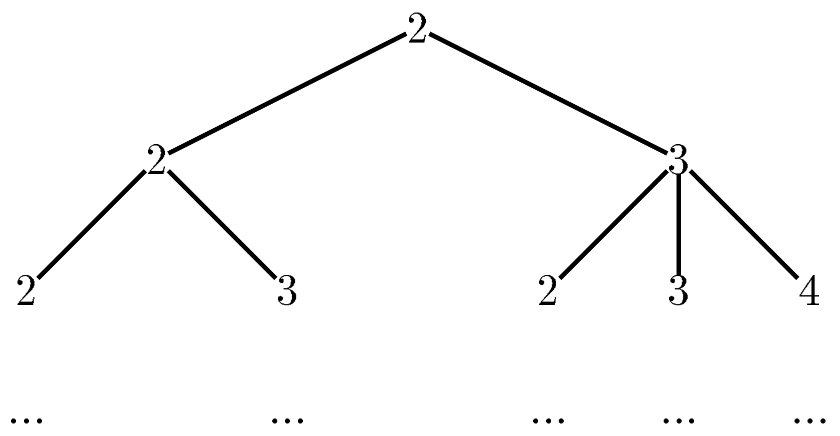
•$F(z,1)$ compte tous les chemins.

•$F(z,0)$ compte les excursions (hauteur finale=0).

Les *ECO*systèmes

confer le papier (FPSAC'99) de C. Banderier, M. Bousquet-Mélou, A. Denise, P. Flajolet, D. Gardy et D. Gouyou-Beauchamps :
"On generating functions on generating trees".

Arbre de possibilités : un nœud étiqueté k a k fils.



Conjecture de R. Pinzani :
règle de récriture $(k) \to e_1(k) \dots e_k(k)$
"Poids linéaire" $\Rightarrow$ s.g. rationnelle ?
Ex. : $(k) \to (2)^{k-1}(k+1)$
"Pas avant bornés" $\Rightarrow$ s.g. algébrique ?
Ex. : $(k) \to (1)(2)\dots(k-1)(k+1)$

Primalité et rationalité

Définition. Le poids de $[(k) \to e_1(k) \ldots e_k(k)]$ est $p(k) := e_1(k) + e_2(k) + \ldots e_k(k)$.

Théorème 1 (Denise98). *Une règle de poids linéaire a une série génératrice rationnelle.*

Preuve/exemple. La marche

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

a une s.g. rationnelle : $F(z,1) = \frac{z-3z^2}{1-5z+2z^2}$.
$p(k) = \alpha k + \beta = 5k - 2$.
$l_1, l_2, \ldots l_{f_n}$: étiquettes des f_n nœuds de profondeur n, on a

$$f_{n+2} = (\alpha l_1 + \beta) + (\alpha l_2 + \beta) + \ldots + (\alpha l_{f_n} + \beta)$$

$$= \alpha f_{n+1} + \beta f_n$$

d'où

$$F(z) = \frac{z + (f_0 - \alpha)z^2}{1 - \alpha z - \beta z^2}.$$

$\square$

Marches dans un borné

Théorème 2. *La série génératrice d'une marche dans un borné de $\mathbb{Z}^n$ est rationnelle.*

Preuve. Transformation en automate puis théorème de Chomsky-Schützenberger □

On a une asymptotique exacte :

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

Marches à pas bornés

Théorème 3. *La série génératrice d'une marche (dans $\mathbb{Z}$) à pas bornés est rationnelle.*

Preuve. $f_n(u) = P(u)^n f_0(u)$, donc :

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

n.b. : $u = 0 \Rightarrow$ résidus $\Rightarrow$ s.g. algébrique. □

La méthode du noyau

fait partie du folklore des 70's (proba, combinatoire [*e.g.* Knuth I.2.2.1]).

Problème. *On veut déterminer $F(z,u)$, on a $F.N = Q$ où $Q(z,u)$ dépend de n inconnues.*

Solution : on résout $N(z,u) = 0$ en u, ce qui donne n racines, qu'on réinjecte pour obtenir un système $n \times n$. Et alors $F = Q/N$. $\square$

Application. À partir du système infini

$$\begin{cases} f_0(u) = 1 \\ \dots \\ f_{n+1}(u) = M f_n(u) \end{cases}$$

on multiplie par z^{n+1}, or $M = P(u) - \{u^{<0}\}P(u)$ est linéaire, en sommant on a donc :

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

$$F(z,u) = 1 + zP(u)F(z,u) - z\{u^{<0}\}P(u)F$$

Et on applique la méthode du noyau à :

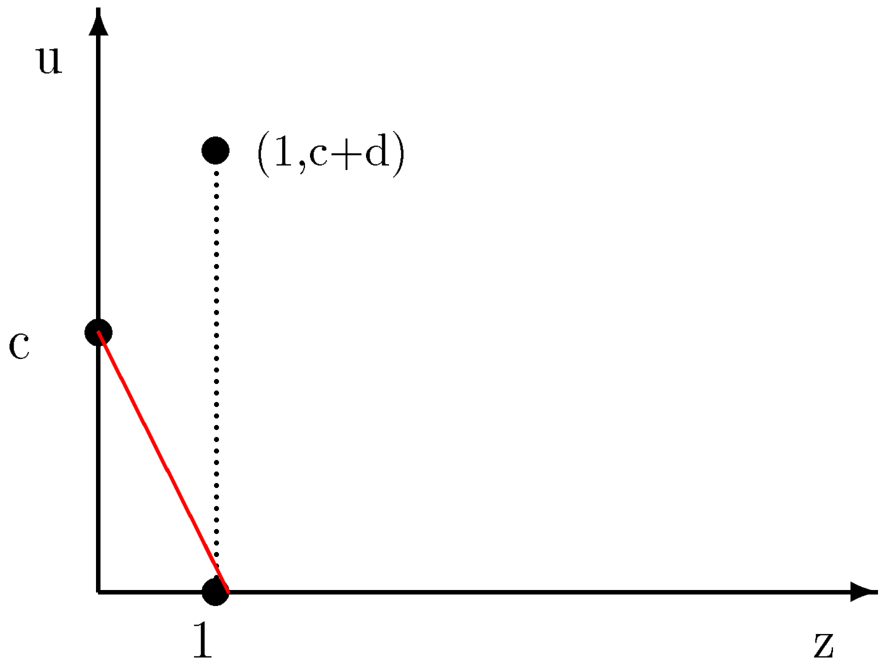
$$F(z,u)(1 - P(u)) = 1 - z \sum_{i=0}^{c-1} p_i(u) \partial_i F(z,0)$$

Marches sur $\mathbb{N}$ à pas bornés

see also M. Bousquet-Mélou & M. Petkovsek

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

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



polygone de Newton pour développer les solutions de $N = 0$ en série de Puiseux, on obtient $c + d$ solutions :

- c petites branches, $u_1, \ldots, u_c$;
- d branches infinies en 0

Théorème 4. *La s.g. d'une marche sur $\mathbb{N}$ à pas bornés est algébrique.*

Preuve.

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

En $u=0$ (excursions) et en $u=1$ (marches), on obtient ainsi une série génératrice algébrique (prouve une généralisation de la conjecture de Pinzani). $\square$

Marches sur $\mathbb{N}$ à pas avant bornés et arrière "continus"

$[(k) \to (1)(2)(3)\dots(k-4)(k-1)(k+5)]$

$L : u^k \mapsto u^0 + u^1 + \dots + u^{k-1}$ *i.e.* $L = \frac{U}{1-u} - \frac{1}{1-u}$

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

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

Théorème 5. *Une marche sur $\mathbb{N}$ à pas arrière "continus" et à pas avant bornés a une série génératrice bivariée algébrique*

$$\textit{excursions} : F(z,0) = \frac{(-1)^{c+1}}{s_d z} \prod_{i=0}^{c} u_i.$$

$$\textit{chemins} : F(z,1) = \frac{-1}{z} \prod_{i=0}^{c} (1 - u_i).$$

Exemples :

Catalan $[(2),(k) \to (2)(3)\ldots(k)(k+1)]$

arbres t-aires $[(t),(k) \to (t)\ldots(k)\ldots(k+t-1)]$

Motzkin $[(1),(k) \to (1)(2)\ldots(k-1)(k+1)]$

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

Schröder $[(3),(k) \to (3)\ldots(k)(k+1)^2]$

$$F(z,1) = \frac{1 - 3z - \sqrt{1 - 6z + z^2}}{4z^2} = \{1,3,11,45,197\}$$

Marches en dimension 2

pas du cavalier codés par

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

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

"nombre de chemins allant de $(0,0)$ à $(n,0)$" :

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

$$F(1,u,0) = \frac{1+2u+\Delta}{4u} - \frac{\sqrt{2+2\Delta-8u^2+4u\Delta}}{4u}$$

$$\Delta := \sqrt{1-4u+4u^2-4u^4} \qquad a_n \sim (1+\sqrt{3})^n \approx 2.73^n \quad \square$$

"séquences de n mouvements de cavalier partant de $(0,0)$ et qui reviennent sur $y=0$"

$$b_n = [z^n v^0]F(z,1,v)$$

$$F(z,1,0) = \frac{1+2z+\Delta}{4z} - \frac{\sqrt{2+2\Delta-12z^2+4z\Delta}}{4z}$$

$$\Delta := \sqrt{1-4z} \qquad b_n \sim 4^n \quad \square$$

autre approche : Labelle et Yeh

Limitation par une droite de pente rationnelle

Définition (langages de Dyck généralisés) :
chaque lettre de l'alphabet a une valeur $\in \mathbb{Z}$
valeur totale du mot$= 0$
valeur positive pour tous les facteurs gauches.

Théorème 6. *La série génératrice des chemins de Dyck généralisés est algébrique.*

Preuve.
Pour deux lettres $\Leftrightarrow$ marches comprises entre $y = 0$ et $y = \frac{p}{q}x$. Équivalent à une marche en dimension 1 de polynôme caractéristique

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

Cas général : polynôme caractéristique du poids des lettres et théorème "pas bornés". $\square$
see also : P. Duchon

Fractions continuées

naturel (pour sauts -1,0,1) de par la traduction structure combinatoire / analyse :
F=Seq(A+Seq(B+Seq(...))

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

Nos s.g. étaient des s.g. ordinaires, mais pas de problème si on recourt aux s.g. exponentielles :

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

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

Exemples non transcendants avec jolies formes closes :

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

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

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

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

Hermite $[2,(k)\to(k-1)^{k-1}(k+1)]$
(involutions en marquant les points fixes)

$$\widetilde{F}(z,u)=\exp(zu+\frac{z^2}{2})\quad F(z,0)=\cfrac{1}{1-\cfrac{kz^2}{\ddots}}$$

Premiers termes :$\{1,0,3,0,15,0,105,0,\frac{(2n)!}{n!2^n}\}$
Exemples non holonomes :
Stirling $[2,(k)\to(k)^{k-1}(k+1)]$ (nb de partages de $\{1,\dots n\}$ en k ss-ens. non vides)

$$\widetilde{F}(z,u)=e^{ue^z-1}\quad F(z,1)=\cfrac{1}{1-kz-\cfrac{kz^2}{\ddots}}$$

Premiers termes :$\{1,2,5,15,52,203\}$
Bessel $[0,(k)\to(k-1)(k)^k(k+1)]$
(permutations sans croisements)

$$F(z,0)=\frac{1}{1-z-z^2B(z)}=\cfrac{1}{1-z-\cfrac{z^2}{1-kz-\cfrac{z^2}{\ddots}}}$$

Premiers termes :$\{1,1,2,4,9\}$
see also S. Corteel

Nombre moyen d'effleurements

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

$$-[u^0]P(u)f_n + t[u^0]P(u)f_n$$

Exemple avec les mots de Dyck $(P(u) = u + \frac{1}{u})$:

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

Le nombre moyen d'effleurements est

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

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

$$= 3\frac{k}{k+2} = 3\sum_{p\geq 0}\frac{(-1)^p 2^p}{k^p} = 3 - \frac{6}{k} + \frac{12}{k^2} + O(\frac{1}{k^3}).$$

$\square$

Asymptotique des excursions

Combien y a-t-il, pour n “grand”, d’excursions de longueur n ?

Théorème 7. *Asymptotique en* $\frac{A}{\rho^n \sqrt{\pi n^3}}$. *$A$ et ρ (le rayon de convergence) sont algébriques.*

Preuve (par analyse de singularité).
$F = \prod U_i$. Or singularité(s) de U_i dicte(nt) la croissance des coeff de F (th. d’Hadamard).
$c = \frac{u\phi'(u)}{\phi(u)}$ une unique racine positive τ.
De plus $u(\rho) = \tau$. Par ailleurs, lemme :

$$\forall x > 0 \ : \phi(x)(x\phi''(x) + \phi'(x)) > x(\phi'(x))^2$$

(ϕ série à coeff positifs)

$$\text{d'où} \quad u - \tau = d_1\sqrt{\rho - z} + d_2(\rho - z) + \dots$$

Or $[z^n]\sqrt{1 - z} = \frac{-1}{2\sqrt{\pi n^3}}(1 + O(\frac{1}{n}))$ $\square$

Coefficient = intégrale

$$A(z) = a_{-2}z^{-2} + a_{-1}z^{-1} + a_0 z^0 + a_1 z + \cdots$$

$$\int_\gamma A(z) = \int_\gamma a_{-2}z^{-2} + \int_\gamma a_{-1}z^{-1} + \int_\gamma a_0 + \int_\gamma a_1 z + \cdots$$

(γ courbe simple qui entoure zéro)

$$\int_\gamma A(z) = 0 + [a_{-1}\ln(z)]_\gamma + 0 = a_{-1}2i\pi$$

a_{-1} est appelé résidu de A.

Morale : coeff=intégrale, c'est ça la formule intégrale de Cauchy :

$$a_{-1} = \frac{1}{2i\pi}\int_\gamma A(z)dz$$

et plus généralement

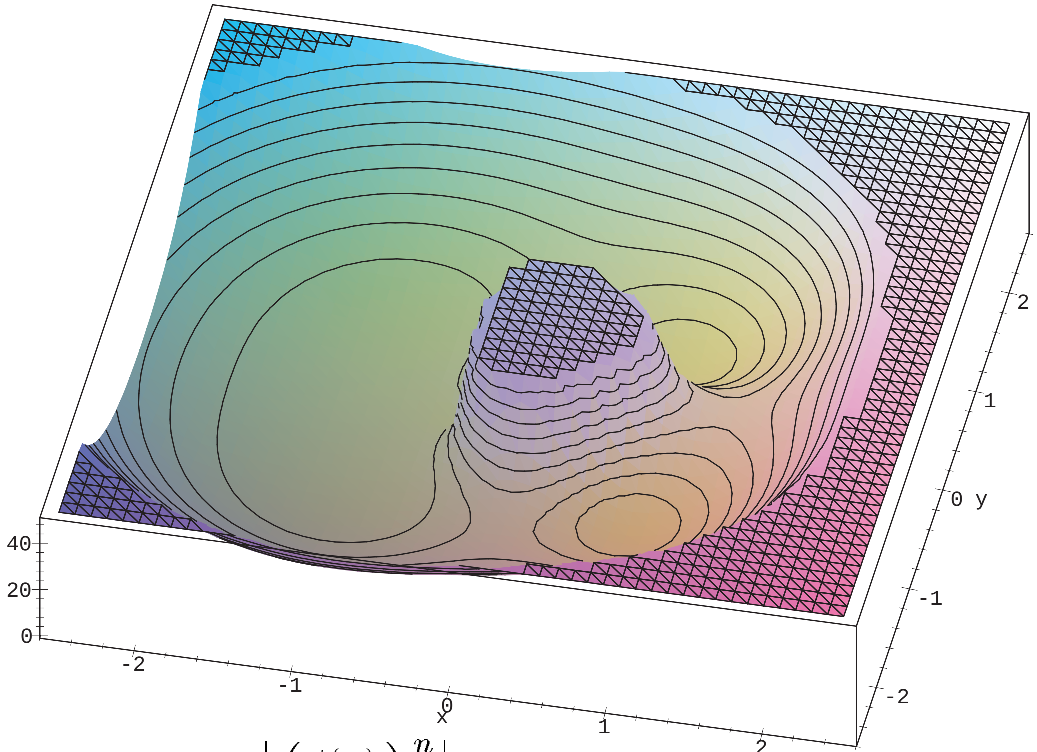
$$a_n = \frac{1}{2i\pi}\int_\gamma A(z)\frac{dz}{z^{n+1}}$$

$\Longrightarrow$ asymptotique de a_n = comportement de l'intégrale pour n grand.

Autre preuve. Méthode du col.

Cauchy nous dit $f_n = \frac{1}{2i\pi} \int_\gamma \frac{F(z)}{z^{n+1}} dz$

Idée : jouer avec le contour γ et obtenir l'asymptotique quand n est grand. $\square$



graphe de $\left| \left(\frac{\phi(u)}{u^3} \right)^n \right|$ avec $\phi = 1 + 2u + u^4 + u^5$,

i.e. $(k) \to (k-3)(k-2)^2(k+1)(k+2)$.

Quand n est grand, tout est concentré autour du point col.

See : D. Gardy, B. Salvy, G. Schaeffer, M. Soria

Perspectives

- résoudre toujours plus d'équations fonctionnelles
- obtention de formes closes (EGF)
- interprétations combinatoires (bijectives)
- génération aléatoire
- étudier l'aire sous une marche (q-analogues)
- étudier le maximum d'une marche
- asymptotique de ces paramètres
- pousser le lien fractions continuées / marches aléatoires
- méthode du col sur classe générale
- obtention de lois limites

Au boulot ! :-)