Combinatoire analytique : application aux marches aléatoires

Cyril Banderier

Travail dirigé et encadré par Philippe Flajolet
Projet Algorithmes INRIA

Mémoire du DEA algorithmique
Juillet 1998

La combinatoire analytique

combinatoire analytique

=

(analyse) combinatoire

+

analyse (complexe)

- prêtre : Leonhard Euler (1707-1783) †
- mariage discret (fini) / continu (infini)
- autre exemple de mariage heureux :

analyse

+

théorie des nombres

=

théorie analytiques des nombres

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
- théorie de la complexité
- analyse d'algorithmes

Les séries génératrices

Le lien entre le discret et l'analyse se fait via les séries génératrices.

Pour dénombrer une suite d'objets par rapport à leur taille, on forme la série "formelle"

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

où a_n est le nombre d'objets de taille n.

Suivant les propriétés de ces objets, on a recourt à une

- série génératrice ordinaire (OGF)

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

- série génératrice exponentielle (EGF) "l'univers étiqueté",

$$\sum_{n\in\mathbb{N}} a_n \frac{z^n}{n!};$$

- série génératrice de Dirichlet (DGF) (propriétés multiplicatives, théorie des nombres)

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

- série génératrice multivariée quand on veux suivre plusieurs paramètres. En bivariée (BGF) :

$$\sum_{i,j \in \mathbb{N}} a_{ij} z^i u^j;$$

- (PGF) quand on veut tenir compte de la probabilité d'arriver à telle configuration.

Attirail d'analyse classique

On a affaire à des séries à coefficients **positifs**, elles vérifient le
Théorème de Pringsheim :
Une série à coefficients positifs et de rayon de convergence fini ρ a une de ses singularités dominantes (*i.e.* celles de plus petit module) en $z = \rho$.

Rappelons le "théorème fondamental de l'asymptotique" (Hadamard) :

$$a_n \sim \rho^{-n}$$

On utilisera les notations suivantes :

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

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

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 on a alors un polynôme de degré $< n$),
- P a des racines multiples.

On a n fonctions $u_1, \ldots, u_n$ racines de $P(z,u) = 0$ développables en série de Puiseux avec l'aide du polygone du Newton, *i.e.* $\{(i,j)$ tels que $a_{ij} \neq 0\}$.

marche + aléatoire = marches aléatoires

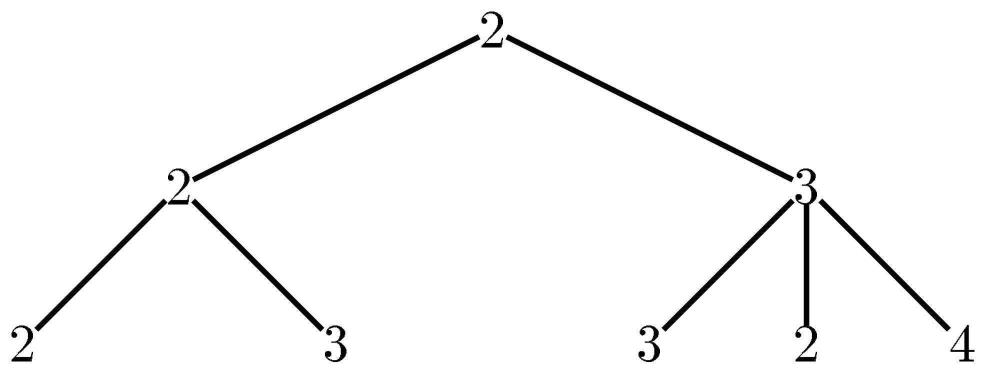
Omniprésentes :

- automatisme (Wiener)
- ordonnancement (files d'attente)
- génétique (Galton-Watson)
- physique statistique (réseaux cristallins)

Exemple 'amusant' de l'ivrogne de Pólya.
Probabilité de repasser par le point de départ ?
$= 1$ si dim $= 1$ ou 2
< 1 si dim ≥ 3.

Les jeux de récriture de Pinzani

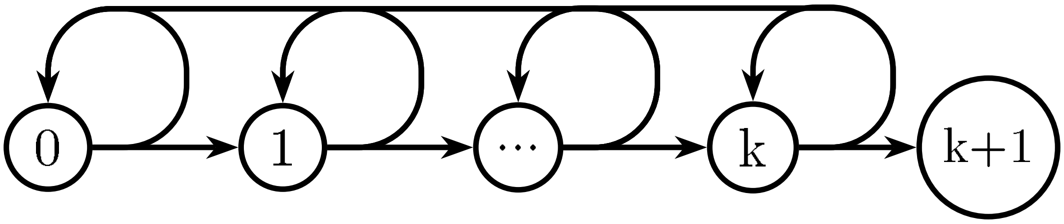
$(k) \rightarrow (2)(3) \cdots (k)(k+1)$



Nombre de nœuds se trouvant à la hauteur n ?
ici : Catalan
Conjecture pour une grammaire quelconque
série génératrice rationnelle, algébrique ?

Traduction en terme de marches aléatoires

$(k) \to (2)(3)\cdots(k)(k+1)$ correspond à :



Nous sommes ramenés à dénombrer les marches aléatoires partant de 0 et de longueur n.

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

Le jeu d'opérateurs suivants permet de décrire récursivement les marches

- $\Delta_{0,i}$ 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 ;
- 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.

schéma de traduction suivant :

mot	codé dans s.g.	action de L
$(ak+b)^c$	cu^{ak+b}	$cS^a\sigma_c$
b^{ck+d}	$(ck+d)u^b$	$cS^bU\partial + dS^bU$
$(ak+b)^{ck+d}$	$(ck+d)u^{ak+b}$	non linéaire

un opérateur **linéaire** (sur les monômes u^k)

$$L := \sum_c P_c(S)\sigma_c + R(S)U\partial + Q(S)U$$

La méthode du noyau

$$\begin{cases} f_0(u) = 1 \\ f_{n+1}(u) = Lf_n(u) - r_n(u) \end{cases}$$

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

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 :

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

$$F(z,u)(1 - zS(u)) = 1 - z \sum_i p_i(u) \partial_i F(z,0)$$

Pépin dans le noyau

$L := U\partial + (S + S^2 + S^9)$

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

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 $F(z,u)$.

Un ensemble de marches rationnelles

$$L := \sum_c P_c(S)\sigma_c + R(S)U\partial + Q(S)U$$

poids:= $\sum_c (ckP_c(1) + P'_c(1)) + kR'(1) + Q'(1)$

Pour une grammaire de Pinzani

$(k) \to (a_1)^{e_1} \cdots (a_k)^{e_k}$, poids=somme des $a_i e_i$,

Poids linéaires égaux $\Longrightarrow$ s.g. égales [Denise98]

Nous obtenons avec notre formalisme :

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

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

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

$$\beta' = \sum_c P'_c(1) + Q'(1)$$

Marches dans un borné

Théorème 0.1. *La série génératrice d'une marche dans un borné de $\mathbb{Z}^n$ est rationnelle.*

Preuve. transformation en automate puis théorème Chomsky-Schützenberger $\square$

On a une asymptotique exacte :

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

Marches à pas bornés

$f_n(u) = L^n f_0(u)$ L polynôme donne :

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

La série génératrice des marches est

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

La série génératrice des excursions est

$$F(z,0) = \frac{1}{1 - S(0)z}.$$

Nota bene : la singularité dominante est
$S(1)^{-1}$ pour les marches
et $S(0)^{-1}$ pour les excursions.

Marches positives à pas bornés

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

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

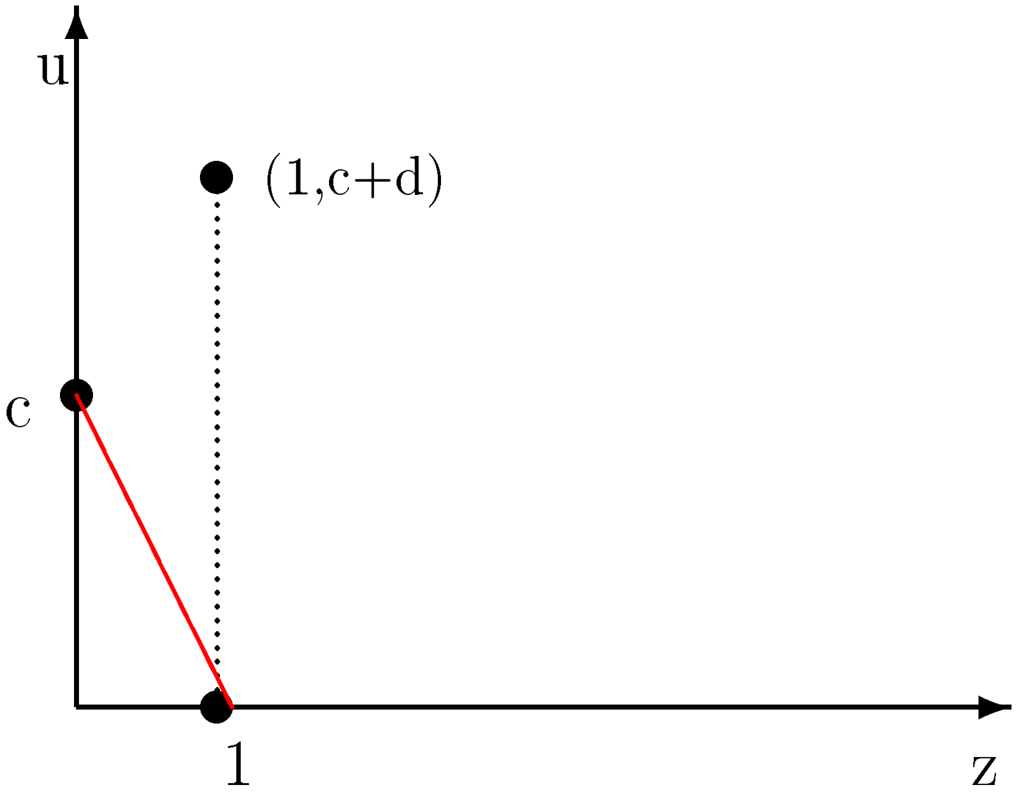
Posons

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

et

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

polygone de Newton pour développer les solutions en série de Puiseux, on obtient $c+d$ solutions $u_0, \ldots, u_{c+d-1}$.



- c petites branches, que nous noterons $u_0, \ldots, u_{c-1}$;
- d branches infinies en 0, que nous noterons $v_0, \ldots, v_{d-1}$.

Raccourci Mebeme

En regardant P et Q comme des polynômes en u, et en remplaçant u par u_i, on voit que les u_i sont aussi racines de $Q(u, z)$.

$$F(z, u) = \frac{Q(u, z)}{P(u, z)}$$

$$= \frac{\prod_{i=0}^{c-1}(u - u_i)}{-zs_d \prod_{j=0}^{c+d-1}(u - u_j)}$$

$$= \frac{1}{zs_d \prod_{i=0}^{d-1}(u - v_i)}$$

En $u = 0$ (excursions) et en $u = 1$ (marches) on obtient toujours une série génératrice algébrique : youpi pour Pinzani !

Marches positives à pas avant bornés et arrière non bornés

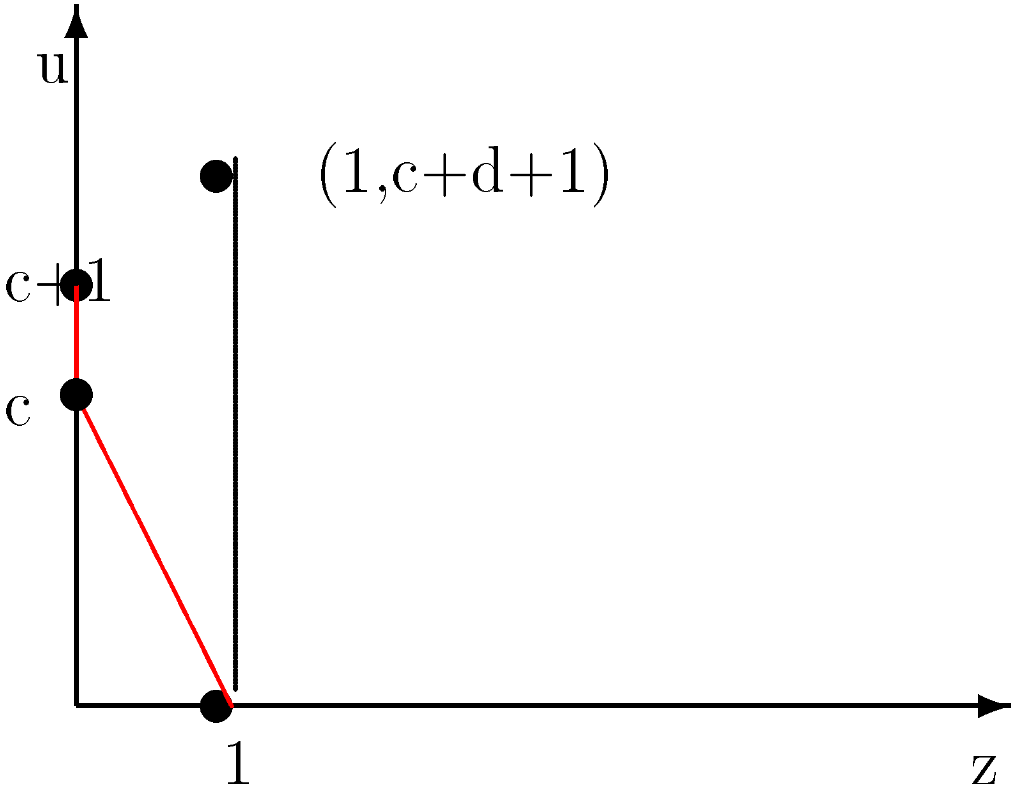
$$L := \frac{1}{1-u}U + \frac{1}{1-u}S^0.$$

$$L : u^k \mapsto u^0 + u^1 + \cdots + u^{k-1}$$

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

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

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



- une branche équivalente à 1 en $z = 0$, que nous noterons 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, que nous noterons $v_0, \ldots, v_{d-1}$.

Théorème 0.2. *Une marche aléatoire positive à pas arrière non bornés et à pas avant bornés a une série génératrice des excursions algébrique et elle vaut*

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

Théorème 0.3. *Une marche aléatoire positive à pas arrière non bornés et à pas avant bornés a une série génératrice des marches algébrique et elle vaut*

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

Marche régulière définie modulo un pavé

- Méthode combinatoire par passage au quotient
- Méthode analytique par suites imbriquées

Par exemple en dimension 1 suivant la parité :

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

Marches en dimension 2 : des chemins de Dyck bien cavaliers

Les pas du cavalier sont codés par le polynôme caractéristique

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

On a la récurrence

$$f_{n+1} := Sf_n - \{u^{<0}v^{<0}\}Sf_n$$

d'où l'équation fonctionnelle

$$F(z,u,v) = 1 + z\Big(S(u,v)F(z,u,v)$$

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

En annulant le noyau, on obtient

$$\begin{cases} v_0(u,z) = +\sqrt{z}\sqrt{u} + \cdots \\ v_1(u,z) := -\sqrt{z}\sqrt{u} + \cdots \end{cases}$$

et finalement

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

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

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

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

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

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

“séquences de n mouvements de cavalier
partant de $(0,0)$ et qui reviennent sur $y=0$”
$b_n = [z^n v^0] = F(z,1,v)$

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

$\Delta := \sqrt{1-4z}$

$$b_n \sim 4^n$$

Limitation par une droite de pente rationnelle

$y = \frac{p}{q}x$, on pose

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

on est rameé à une marche en dim 2

"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 0.4. *La série génératrice des chemins de Dyck généralisés est algébrique.*

Aire et q-analogues

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

opérateur de substitution $\sigma : v \mapsto vq^a$

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

Avec les chemins de Dyck, on aboutit à l'équation

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

et finalement

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

Pour toutes les marches, on aboutit ici à des q-analogues, la résolution de ces équations dans un cadre général reste à faire.

Asymptotique

marches simples (avance avec pas bornés) mènent au noyau $y - z\phi(y) = 0$ asymptotique connue : $A\frac{\rho^n}{\sqrt{\pi n^3}}$
Cas général $y^c - z\phi(y) = 0$

$$P(y,z) := y^c - z\phi(y) = 0.$$

singularité si $P(y,\tau) = 0$ et $\frac{\partial P}{\partial y}(y,\tau) = 0$,
i.e. $cy^{c-1} - z\phi'(y) = 0$ et $y^c - z\phi(y) = 0$.

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

Primordial de savoir si et quand cette relation est vérifiée

Théorème 0.5. *$d :=$degré de ϕ Pour tout $c < d$ fixé (avec $c > 0$), l'équation*

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

admet une unique racine positive τ.

Lemme technique

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

De plus

$$y - \tau = d_1(\rho - z)^{1/2} + d_2(\rho - z) +$$

$$d_3(\rho - z)^{3/2} + d_4(\rho - z)^2 + \cdots$$

Ainsi

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

Ce mémoire apporte :

- la mise en place d'un **formalisme**, d'une **méthode de calcul** permettant l'obtention aisée de séries génératrices de marches aléatoires.

- une **unification** de différents résultats, des "généralisations" de **théorèmes**
  - J. Labelle et Y.N. Yeh sur les **marches de cavaliers**,
  - P. Duchon sur les **langages de Dyck généralisés**,
  - A. Denise sur les **grammaires infinies** rationnelles,
  - s.g. excursions et marches à pas bornés,
  - s.g. excursions et marches > 0 à pas arrière non bornés, elles sont **algébriques** (conjecture de Pinzani)

- un résultat **asymptotique** sur $y^c = z\phi(y)$ avec obtention au passage de quelques lemmes techniques, et le résultat final

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

- une stratégie d'étude de paramètres (nombres d'effleurements **en moyenne**, aire q**-analogues**).
- montre comment jouer avec les **conditions de bord** (limitation par une droite de **pente rationnelle**).

Perspectives : étude de marches jusqu'à présent peu regardées, dégager des **lois limites**, élaborer une **théorie de résolution des équations fonctionnelles**.