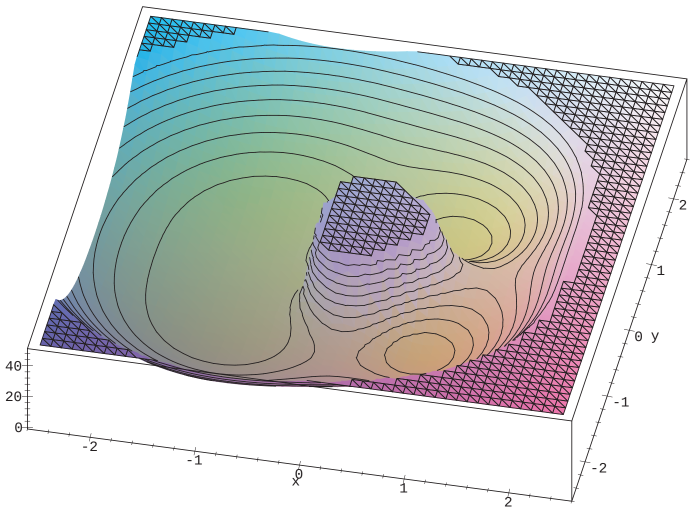
Analysis of Algorithms
July 3-7, 2000
Krynica Morska (near Gdansk)

The Kernel Method, Asymptotics Implications

Cyril Banderier
Algorithms Project
INRIA Rocquencourt
http://algo.inria.fr/banderier/



History

Mathematical folklore since the 1970's
stacks and lattice paths problems (Knuth, 68)
graph enumeration (Cori & Richard, 1972)
($\approx$ quadratic method, Tutte & Brown, 1960)
queuing system (Fayolle & Iasnogorodski, 79)
More recent works
pebbling a chessboard (Odlyzko & al., 1995)
solving bivariate recurrences (Petkovšek & Bousquet-Mélou, 2000)

Our contribution:

Walks over $\mathbb{N}$
Paths under a line of a rational slope
2-dimensional walks (knight moves)
Generalized Dyck Languages
Number of Dyck factors
Uniform random generation

DEK & Dyck Paths

TAOCP 2.2.1 exercice 4 Ballot problem:

$f_{n,k} :=$ # different ballots with n electors, where votes for candidate A always exceed votes for candidate B, and A wins by k votes.

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

1 equation, 2 unknowns!

Solution Cancel the kernel

$K(u,z) = 1 - z\left(u + \frac{1}{u}\right) = 0 \Longrightarrow 2$ roots

$u_1 = \frac{1+\sqrt{1-4z^2}}{2z}$ and $u_2 = \frac{1-\sqrt{1-4z^2}}{2z}$,

substituting u_2 gives $0 = 1 - \frac{z}{u_2} F(z,0)$

$$\text{Thus } F(z,u) = \frac{\frac{1}{u}F(z,0) - \frac{1}{z}}{K(u,z)}$$

Walks over $\mathbb{N}$

Aim: Find the generating function of a "walk on the integers" :

Dyck: jumps $+1, -1 \leftrightarrow P(u) = \frac{1}{u} + u$
Motzkin: jumps $-1, 0, +1 \leftrightarrow P(u) = \frac{1}{u} + 1 + u$

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

$n \leftrightarrow$ length $\qquad k \leftrightarrow$ height

•$f_{n,k}$: # walks with length n which start from the initial point and end at level k
$=$ # nodes with label k at depth n in the tree.
•$F(z,1)$ counts the walks.
•$F(z,0)$ counts the excursions (final height$= 0$).

Algebraicity of Walks GF

Theorem 1. *GF for excursions/paths of walks with [all backward steps and other] steps in a finite set is algebraic and one has*

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

$$F(z,1) = \frac{\prod_{u_i}(1-u_i)}{1-zP(1)}$$

(product over the convergent roots u_i of the kernel K).

Proof:

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

$$F(z,u)K(z,u) = \text{ LC of } \partial_u^j F(z,0)$$

$K(z,u) = 0 \Longrightarrow c$ convergent roots u_i

LC= polynomial (!) in u of degree c

factorization $\Longrightarrow F(z,u) = \prod (u-u_i)/K(z,u)$

$\square$

Generalized Dyck languages

alphabet: $(a_1, \ldots, a_m)$

each letter has a weight $(w_1, \ldots, w_m) \in \mathbb{Z}$

weight of every prefix ≥ 0

total weight of each word $= 0$

Theorem 2. *Generalized Dyck Languages are algebraic.*

Proof: Labelle and Yeh (90) & Duchon (98) make explicit the (cumbersome!) grammar $\square$

Our proof: easy! by the kernel method

Generalized Dyck words=excursions for the walk with the finite set of steps $(w_1, \ldots, w_m)$ whose GF is algebraic $\square$

Remark the two letters case $\Leftrightarrow$ walks under a rational slope line $(y = \frac{a}{b}x)$.

2-dimensional walks

Left to right jumps of knight coded by

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

We are rich $\Longrightarrow$ We have an infinite chessboard

$$F(z,u,v) = 1 + z\big(P(u,v)F(z,u,v)$$

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

"# paths from $(0,0)$ to $(k,0)$":

$$a_k = [u^k v^0]F(1,u,v).$$

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

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

another approach: Labelle and Yeh (89)

Average number of 0's

$$F(z,t) = \sum_{n,j} f_{nk} z^n t^k$$

$z \leftrightarrow$ length, $t \leftrightarrow$ number of 0's.

Excursion = sequence of Arches $\rightarrow E = \frac{1}{1-A}$

$$F(z,t) = \frac{1}{1 - tA(z)} = \frac{1}{1 - t(1 - \frac{1}{E(z)})}$$

Average # of 0's is

$$\mu_n = \frac{[z^n]\partial_t F(z,1)}{[z^n]F(z,1)} = \frac{[z^n]E(z)^2}{[z^n]E(z)} - 1$$

$$= 2b_0 - 1 + O\left(\frac{1}{n}\right)$$

with $E(z) = b_0 - b_1\sqrt{\rho - z} + \cdots$

Example with Dyck words Average # of 0's is

$$\mu_n = 3 - \frac{12}{n} + \frac{48}{n^2} + O(\frac{1}{n^3}).$$

$\square$

Lagrangean Scheme

For $\phi := P(u)u^c$ with nonnegative coefficients

Lagrange :

$$u = z\phi(u)$$

unique singularity $\Rightarrow$ easy!

lot of known results (trees...)

Us :

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

$c > 1 \Rightarrow$ where are the main singulariti(es) ?

Reason: scheme $u = z^{1/c}\phi^{1/c}(u)$ but

$\phi^{1/c}$ no longer has nonnegative coefficients

Solution

Localization of *the* dominant singularity uses

singularity analysis

topology

saddle point method

symmetric properties

A universal property ?

Hypothesis H For any walk, the singularity comes from the real saddle point of $P(u)$.

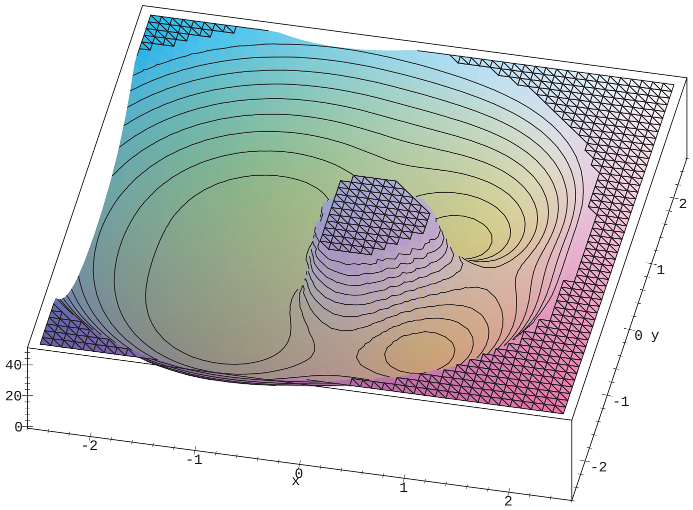
$\tau := \min\{\, u \text{ such that } P'(u) = 0 \text{ and } P(u) \in \mathbb{R}\}$

"no zero or SP between 0 and τ" $\Rightarrow \tau$ real

($\approx$ the highest real saddle point is real)

true for *all* walks with jumps in $-6, \ldots, 6$

and several million examples



graph of $\left|u^{-3} + 2u^{-2} + u + u^2\right|$,
walk with jumps $(-3, -2, -2, +1, +2)$.

Asymptotics of excursions

Theorem 3. *Under the previous hypothesis, the number of excursions of a nonperiodic walk is*

$$[z^n]F(z,0) \sim \frac{A}{\rho^n \sqrt{\pi n^3}}$$

with A and ρ (radius of convergence) algebraic.

Proof (by singularity analysis).

- $1 - zP(u) = 0$ has c small branches u_i
- $F = \prod_{i=1}^{c} u_i$
- u_i's are conjugated: $U_i(t) = U_1(w^{j-1}t)$
- graph of $P(u)$ is convex
- $\exists' \tau > 0$ such that $P'(\tau) = 0, \quad \rho := P(\tau)$
- u_1 (the real branch) such that $u_1(\rho) = \tau$
- $\phi(x)\,(x\phi''(x) + \phi'(x)) > x{\phi'}^2(x)$
- u_1 has $\sqrt{}$ singularity in ρ

saddle point method: u_1 ultimately ≥ 0 coefficients
no smaller singularities
no other singularities (level curves' topology)

But singulariti(es) of U_i dictate the growing rate of coeff of F. $\square$

Lemma The radius of convergence of u is

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

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

Proof: singularity if cancelation of the leading term in $K(z,u) := u^c - z\phi$ or if

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

For $c \in [1, \deg(\phi)[$, $c = \frac{u\phi'(u)}{\phi(u)}$ has an unique solution τ. Indeed set $h := \phi(u)/u^c$,

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

$h'(0+) < 0$ and $h(R) > 0$, Rolle theorem $\implies$ $h' = 0$ has an unique solution on $]0, R[$. $\square$

Square-root behavior

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

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

$$\text{i.e.} \quad u - \tau = d_1\sqrt{\rho - z} + d_2(\rho - z) + \ldots$$

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

Technical lemma

ϕ nonnegative coefficients (nonconstant, $\phi(0) \neq 0$)
$\implies \phi(x)(x\phi''(x) + \phi'(x)) > x{\phi'}^2(x)$

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

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

$$\Big(\sum_{n\geq 0} a_n x^n\Big)\Big(x^2 \sum_{n\geq 2} a_n n(n-1)x^{n-2} + x \sum_{n\geq 1} a_n n x^{n-1}\Big)$$

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

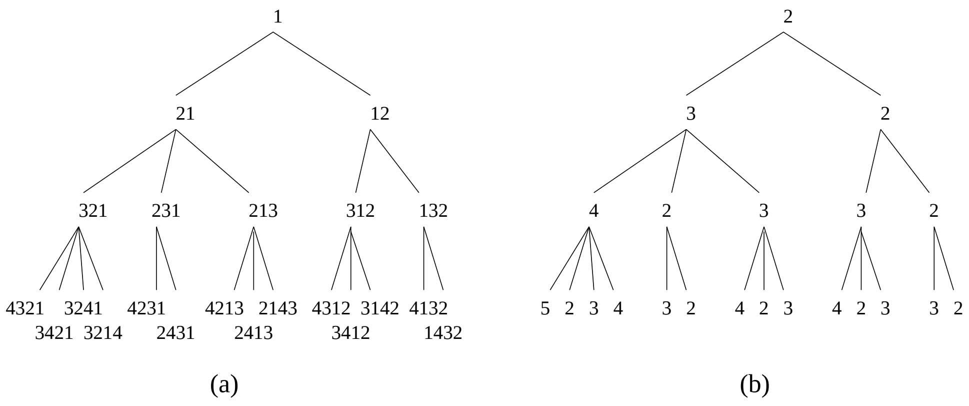
$$\Big(\sum_{n\geq 2} a_n x^n\Big)\Big(\sum_{n\geq 2} a_n n^2 x^n\Big) \geq \Big(\sum_{n\geq 2} a_n n x^n\Big)^2$$

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

boustrophedon:
$((n-k)^2 + k^2)a_{n-k}a_k \geq 2k(n-k)a_k a_{n-k}$,
ok because $(n-2k)^2 a_k a_{n-k} \geq 0$. $\square$

Enumerating Combinatorial Objects

J. West, S. Dulucq, O. Guibert: generating tree of 123-avoiding permutations



(a) nodes labelled by the permutations

(b) nodes labelled by the numbers of children

Basic fact: ECO-method of R. Pinzani & al. $\Longleftrightarrow$ rewriting rules (ecosystem) axiom $= (2)$ and $(k) \to (2)(3)\dots(k+1)$.

Usefulness: Random generation

! non trivial weights for the branches :

e.g. $\Pr(213) = \frac{1}{6}$ but $\Pr(132) = \frac{1}{4}$

$\Longrightarrow$ need enumeration

generating trees

$\approx$ random walks on the integer

Generating Trees

														# of nodes	Σ labels
					2									1 nodes	2
		2						3						2 nodes	5
	2		3			2		3			4			5 nodes	14
2	3	2	3	4	2	3	2	3	4	2	3	4	5	14 nodes	42
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	f_n nodes	s_n

Generating trees: node with label $k \Rightarrow k$ children

$\exists$ associated rewriting rules e.g. (above): axiom = (2) and $(k) \rightarrow (2)(3)\ldots(k+1)$.

Question: How many nodes at depth n ?

Algebraic Ecosystems

Getting the functional equation.

ex.: $[(0), (k) \to (0)(1) \dots (k+1)]$

$$u^k \to u^0 + \cdots + u^{k-1} + u^k + u^{k+1} = \frac{1-u^k}{1-u} + u^k + u^{k+1}$$

The basic recurrence is

$$f_{n+1}(u) = \frac{f_n(1) - f_n(u)}{1-u} + (1+u) f_n(u)$$

Multiplying by z^{n+1} and summing (for $n \geq 0$) yields

$$F(z,u) = f_0(u) + z\frac{F(z,1) - F(z,u)}{1-u} + z(1+u)F(z,u)$$

The functional equation is

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

The Kernel Method

Theorem 4. *Factorial ecosystem $\Rightarrow$ algebraic GF. More precisely*

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

Solving the functional equation: the kernel method

$(k) \to (0)(1)\cdots(k-3)(k-1)(k+4)(k+15)$ thus $P(u) := -u^{-2}+u^4+u^{15}$

$$F(z,u) = 1 + z\left(\frac{F(z,1)-F(z,u)}{1-u} + P(u)F(z,u) - \{u^{<0}\}\sum_{n\geq 0} z^n L[f_n](u)\right).$$

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

$$F(z,u)K(z,u) = \text{lin. comb. of } (c+1) \text{ unknown functions.}$$

Cancelling the kernel gives $c+1$ convergent branches.

linear system $\Rightarrow$ the $\partial_u^j F(z,0)$ are known $\Rightarrow$ F is algebraic. $\square$

A catalog of ecosystems of combinatorial interest

Ecosystem	Name	Id.	Generating Function
	Rational GF's		(ogf)
$(1),\ (k) \to (k)^{k-1}((k \bmod 2)+1)$	Fibonacci	M0692	$\frac{1}{1-z-z^2}$
$(2),\ (k) \to (2)^{k-1}(k+1)$	odd Fibonacci	M1439	$\frac{1-z}{1-3z+z^2}$
$(3),\ (k) \to (2)^{k-1}(k+1)$	even Fibonacci	M2741	$\frac{1}{1-3z+z^2}$
	Algebraic GF's		
$(1),\ (k) \to (1)\cdots(k-1)(k+1)$	Motzkin numbers	M1184	$\frac{1-z-\sqrt{1-2z-3z^2}}{2z}$
$(2),\ (k) \to (2)\cdots(k)(k+1)$	Catalan numbers	M1459	$\frac{1-2z-\sqrt{1-4z}}{2z}$
$(t{+}1),\ (k) \to (t{+}1)\cdots(k)(k+1)^t$	t=2: Schröder numbers	M2898	$\frac{1-2tz-z+z^t-\sqrt{(1+z-tz)^2-4z}}{2tz}$
$(3),\ (k) \to (3)\cdots(k+2)$	Ternary trees	M2926	equation: $F=1+zF^3$
$(t),\ (k) \to (t)\cdots(k+t-1)$	t-ary trees		equation: $F=1+zF^t$
	Transcendental GF's		(egf)
$(2),\ (k) \to (k-1)^{k-1}(k+1)$	Involutions	M1221	$e^{z+\frac{1}{2}z^2}$
$(2),\ (k) \to (k-1)^{k-2}(k)(k+1)$	Switchboard problem	M1461	$e^{2z+\frac{1}{2}z^2}$
$(2),\ (k) \to (k)(k+1)^{k-1}$	Arrangements	M1497	$e^z/(1-z)$
$(2),\ (k) \to (k-1)^{k-2}(k+1)^2$	Bicolored involutions	M1648	e^{2z+z^2}
$(2),\ (k) \to (k+1)^k$	Factorial numbers	M1675	$1/(1-z)$
$(2),\ (k) \to (k+1)^{k-1}(k+2)$	Increasing subsequences	M1795	$e^{z/(1-z)}/(1-z)$
	Nonholonomic GF's		
$(1),\ (k) \to (k)^{k-1}(k+1)$	Bell numbers	M1484	e^{e^z-1}
$(t),\ (k) \to (k)^{k-t}(k+1)^t$	labelled balls in t-col. urns	M1662	$e^{t(e^z-1)}$
$(2),\ (k) \to (k-1)(k)^{k-2}(k+1)$	Bessel numbers	M1462	—