Algorithmic Tools for the Asymptotics of Linear Recurrences

Bruno Salvy

Inria & ENS de Lyon

Computer Algebra in Combinatorics, Schrödinger Institute, Vienna, Nov. 2017

Motivation

$$p_0(n)a_{n+k} + \cdots + p_k(n)a_n = 0, \quad 0 \notin p_0(\mathbb{N}), \quad p_i \in \mathbb{Z}[n]$$

Aim: given $a_0, \ldots, a_{k-1}$, predict the behaviour of a_n as $n \to \infty$.

Simplified version: ``compute'', when they exist,

$$K, \alpha, m, c \neq 0 \qquad \text{such that} \qquad a_n \sim cK^n n^\alpha \log^m n.$$

Message of this talk:

1. there are tools;
2. c can be the hard part (ie, discarding a very small c);
3. a full asymptotic expansion is not more difficult.

Wimp-Zeilberger Approach

$$p_0(n)a_{n+k} + \cdots + p_k(n)a_n = 0, \quad 0 \notin p_0(\mathbb{N}), \quad p_i \in \mathbb{Z}[n]$$

1. Compute a basis of formal asymptotic expansions

$$\phi_1(n), \ldots, \phi_k(n) \quad \text{(generally divergent)}$$

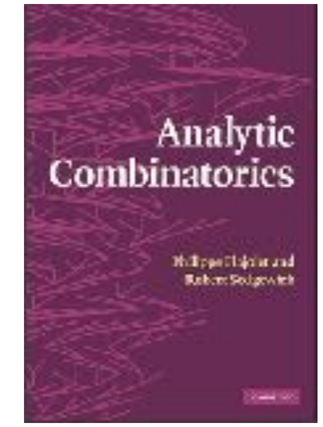
2. Using the initial conditions compute values for large *n* and deduce approximate $c_1, \ldots, c_k$ s.t.

$$a_n \approx c_1\phi_1(n) + \cdots + c_k\phi_k(n)$$

3. In the (many) cases when $\phi_2(n), \ldots, \phi_k(n)$ are $o(\phi_1(n))$ and c_1 is numerically nonzero, conclude

$$a_n \sim c_1\phi_1(n).$$

Singularity Analysis



counts the number of objects of size *n*

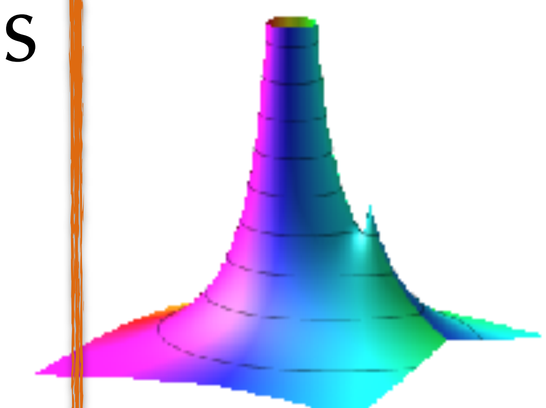
$$(a_n) \mapsto A(z) := \sum_{n \geq 0} a_n z^n$$

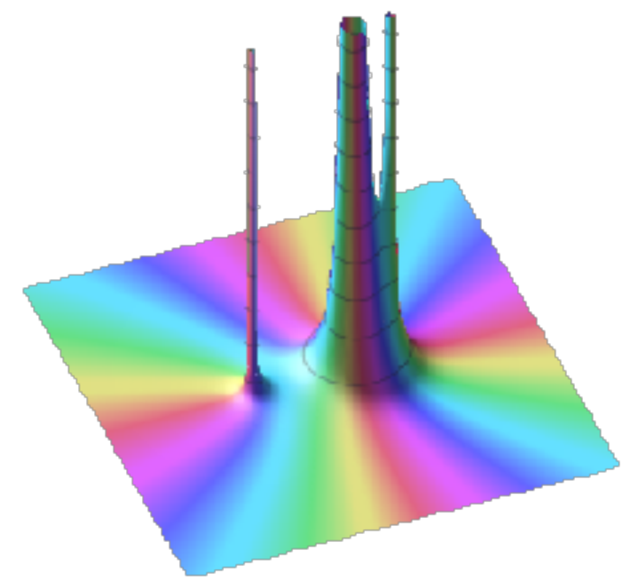
captures some structure

A 3-Step Method:

1. Locate dominant singularities
   a. singularities; b. dominant ones
2. Compute local behaviour
3. Translate into asymptotics

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





$$A(z) \underset{z \to \rho}{\sim} c \left(1 - \frac{z}{\rho}\right)^{\alpha} \log^{m} \frac{1}{1 - \frac{z}{\rho}} \qquad a_n \underset{n \to \infty}{\sim} c\, \rho^{-n} \frac{n^{-\alpha - 1}}{\Gamma(-\alpha)} \log^{m} n$$

$(\alpha \notin \mathbb{N})$

full asymptotic expansion available

Next: computation of ρ, α, m, c

P-recursivity & D-finiteness

$$(a_n) \mapsto A(z) := \sum_{n \geq 0} a_n z^n$$

(a_n) P-recursive $\iff$ **A(z) D-finite**

$$p_0(n)a_{n+k} + \cdots + p_k(n)a_n = 0 \qquad q_0(z)A^{(\ell)}(z) + \cdots + q_\ell(z)A(z) = 0$$

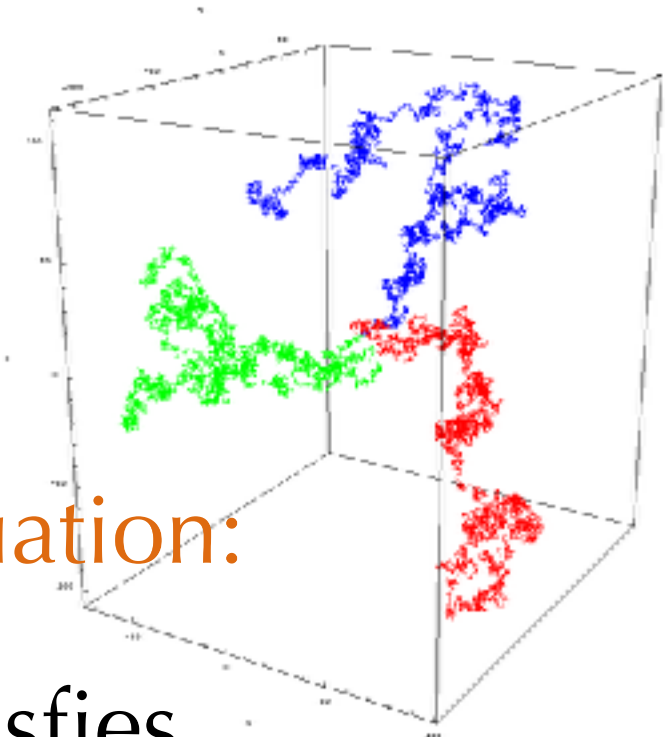
Classical properties of LDEs:

1. singularities satisfy $q_0(\rho) = 0$;
2. one can compute a basis of formal solutions at (regular) singular points, of the form

$$\left(1 - \frac{z}{\rho}\right)^{\alpha} \log^{m}\left(\frac{1}{1-\frac{z}{\rho}}\right)(1 + \cdots), \qquad \alpha \in \overline{\mathbb{Q}}, m \in \mathbb{N}.$$

More recently (M. Mezzarobba's talk on Thursday): *certified analytic continuation* (*→ c numerically*).

Ex: Pólya's 3D Random Walk



Start from the origin in $\mathbf{Z}^d$;
move one step along one of the axes; repeat.
What is the probability p_d of returning to 0?

Numerical approximation by analytic continuation:

1. $u_n :=$ **P**(3D-walk returns to 0 in 2n steps) satisfies

$$(2n+3)(2n+1)(n+1)u_n - 2(2n+3)(10n^2+30n+23)u_{n+1} + 36(n+2)^3 u_{n+2} = 0$$

2. $a_n := \sum_{k=0}^{n} u_k \to c := \frac{1}{1-p_3}$ converges slowly (1 is a singularity)

3. Given a_0, a_1, a_2, `NumGfun` produces 100 digits of c, c_2, c_3 s.t.

$$A(z) \approx c\left(\frac{1}{1-z} + \cdots\right) + c_2\left(\frac{1}{\sqrt{1-z}} + \cdots\right) + c_3\,(1 + \cdots) \text{ in 3 sec.}$$

$$c = \frac{\sqrt{6}}{32\pi^3}\Gamma\left(\frac{1}{24}\right)\Gamma\left(\frac{5}{24}\right)\Gamma\left(\frac{7}{24}\right)\Gamma\left(\frac{11}{24}\right)$$ not accessible to the algorithms presented here.

[Koutschan *et alii* 13,16;Glasser-Zucker77]

Asymptotics of D-Finite Combinatorial Sequences

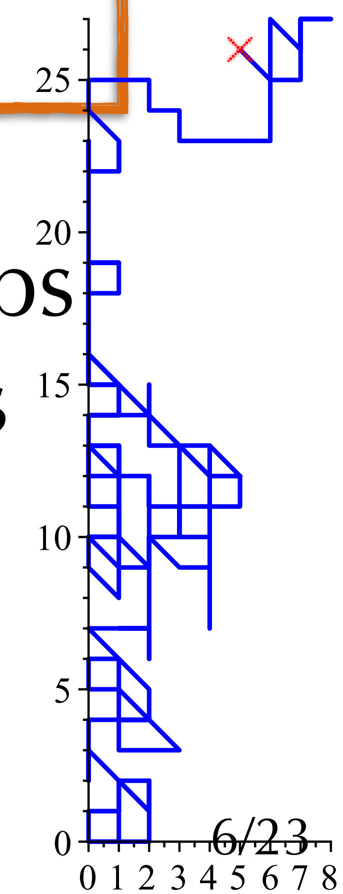
Thm. [Katz70,Chudnovsky85,André00]

$a_0+a_1z+\dots$ D-finite, a_i integers, radius in $(0,\infty)$, then its singular points are regular with **rational** exponents

$$a_n \sim \sum_{\substack{(\lambda,\alpha,k)\in \text{ finite set}\\ \text{in } \overline{\mathbb{Q}}\times\mathbb{Q}\times\mathbb{N}}} \lambda^{-n} n^{\alpha} \log^k(n)\, f_{\lambda,\alpha,k}\left(\frac{1}{n}\right).$$

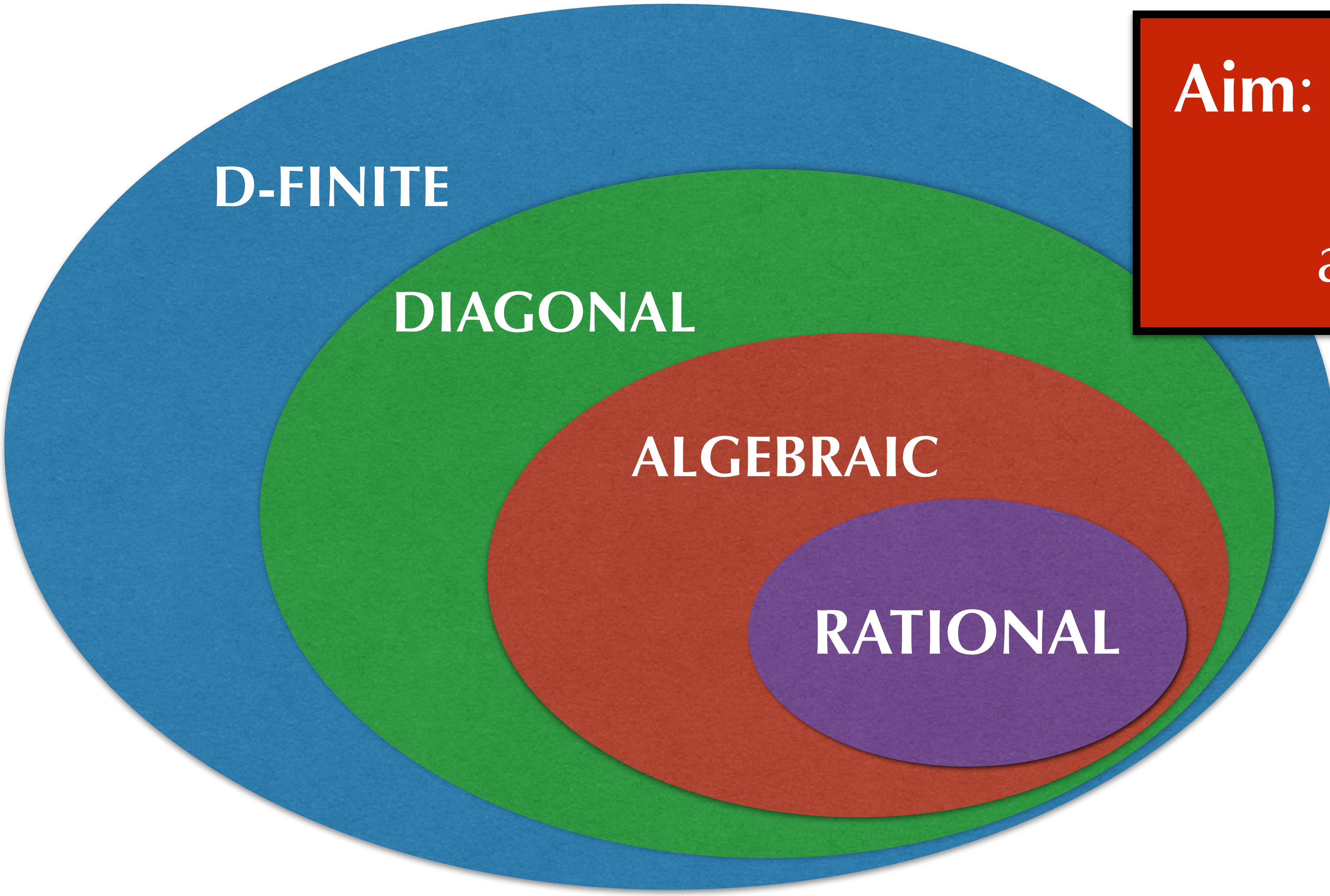
Ex. The number a_n of walks from the origin taking n steps *{N,S,E,W,NW}* and staying in the first quadrant behaves like $C\lambda^{-n}n^{\alpha}$ with $\alpha \notin \mathbb{Q}$ → not D-finite.

$$\alpha = -1 + \frac{\pi}{\arccos(u)}, \quad 8u^3-8u^2+6u-1=0, \quad u>0.$$



[Bostan-Raschel-S.14]

Univariate Generating Functions



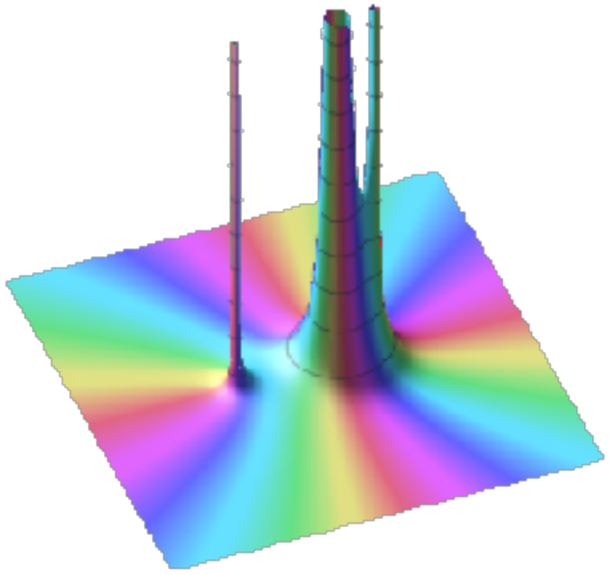
Aim: asymptotics of the coefficients, automatically.

More structure
→
more complete algorithm

Def diagonal: R. Pemantle's talk yesterday.

Christol's conjecture: All differentially finite power series with integer coefficients and radius of convergence in $(0,\infty)$ are diagonals.

I. Rational Generating Functions (Linear Recurrences with Constant Coefficients)



Conway's sequence

1,11,21,1211,111221,…

Generating function for lengths:

$$f(z)=P(z)/Q(z)$$

with deg Q=72.

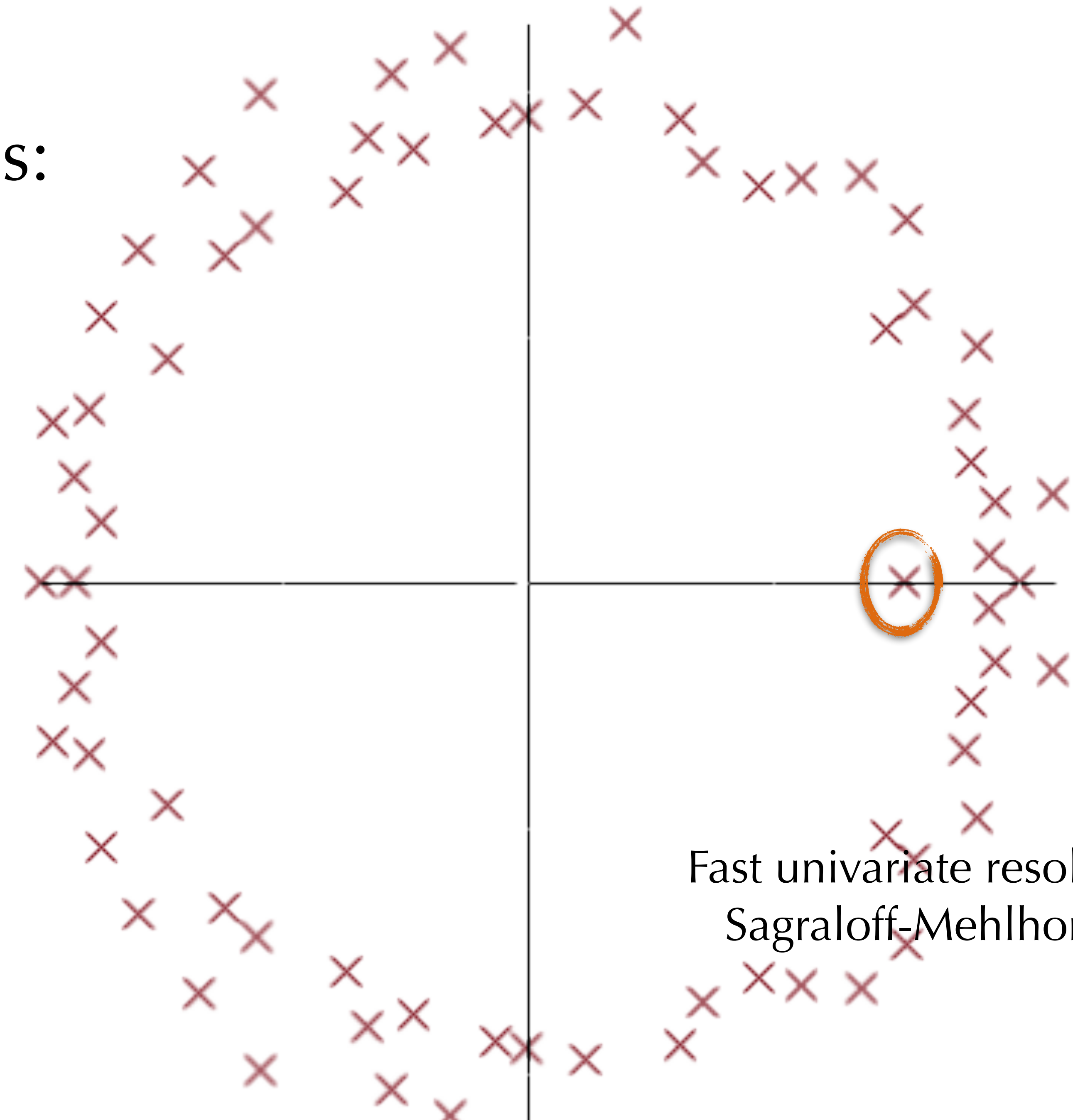
Smallest singularity:

$\rho \simeq 0.7671198507$

$$\ell_n \simeq 2.04216\, \rho^{-n}$$

$c=\rho^{-1}\,\mathrm{Res}(f,\rho)$

algebraic



Fast univariate resolution:
Sagraloff-Mehlhorn16

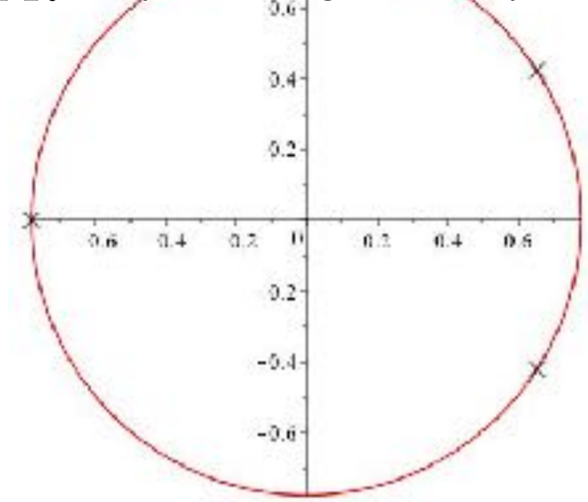
remainder exponentially small

Singularity Analysis for Rational Functions

A 3-Step Method:

1. Locate dominant singularities
   a. singularities; b. dominant ones
2. Compute local behaviour
3. Translate into asymptotics

Ex: $17z^3 - 9z^2 - 7z + 8$



dist 10^{-5}

1. Numerical resolution with sufficient precision + algebraic manipulations
2. Local expansion (easy).
3. Easy.

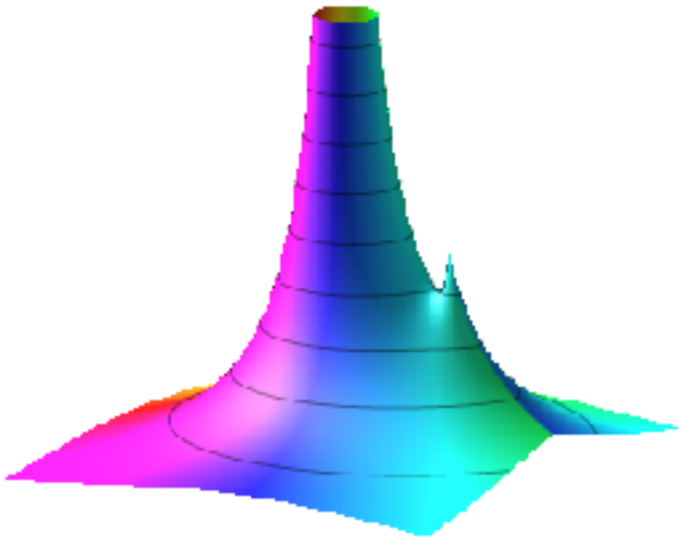
Useful property [Pringsheim Borel]

$a_n \geq 0$ for all $n \Longrightarrow$ real positive dominant singularity.

II. Algebraic Generating Functions

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

with $P(z, y) \in \mathbb{Z}[z, y] \setminus \{0\}$



Algebraic Generating Functions

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

1a. Location of possible singularities
Implicit Function Theorem:

$$P(z, y(z)) = \frac{\partial P}{\partial y}(z, y(z)) = 0 \quad \text{(discriminant)}$$

Numerical resolution
with sufficient precision
+ algebraic manipulations

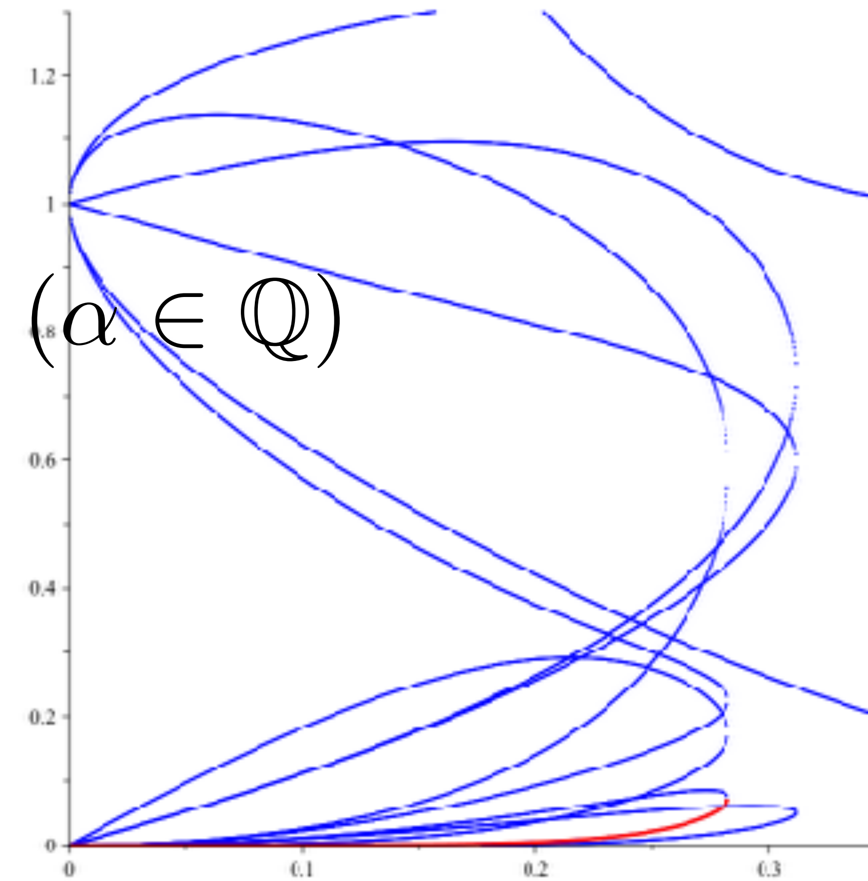
1b. Analytic continuation
finds the dominant ones

2. Local behaviour (Puiseux): $(1 - z/\rho)^{\alpha}$, $(\alpha \in \mathbb{Q})$

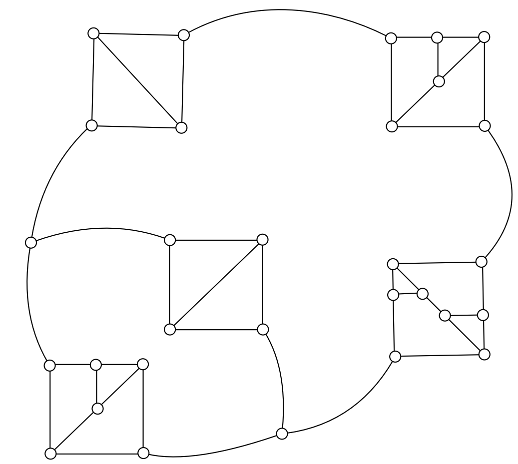
3. Translation: easy:

$$a_n \underset{n\to\infty}{\sim} c\,\rho^{-n} \frac{n^{-\alpha-1}}{\Gamma(-\alpha)}$$

with c, ρ algebraic, α rational.



3-regular 2-connected Planar Graphs



$$U = 2G_3 + T + 2U^2 = \frac{T}{(1-U)^3}, T = z(1+B)^3, B = \frac{G_3 + B^2}{1+B} + z\left(B + \frac{1}{2}B^2\right)$$

define power series $U(z), G_3(z), T(z), B(z)$.

The aim is to compute the asymptotic behaviour of $[z^n]B(z)$.

1. Eliminating U,T,G_3 gives $P = 16B^6z^2 + \cdots + z^2(z^2 + 11z - 1)$.
2. The discriminant has degree 20, but *only one root in (0,1]*:

   $\rho \approx .102$ root of $54z^3 + 324z^2 - 4265z + 432$.
3. At $z = \rho$, *P* has *only 1 (double) real positive* root: $B(\rho)$
4. Computing more terms gives

   $$B(z) = B(\rho) + c_1\left(1 - \frac{z}{\rho}\right) \pm c\left(1 - \frac{z}{\rho}\right)^{3/2} + \cdots$$ with an explicit c
5. Conclusion:

   $$[z^n]B(z) \sim \frac{3c}{4\sqrt{\pi}} n^{-5/2} \rho^{-n}.$$

Analytic continuation exploiting the combinatorial origin.

Singularity Analysis of Algebraic Series

Prop. [Abel1827;Cockle1861;Harley1862;Tannery1875]
Algebraic series are D-finite.

Exact analytic continuation for singularity analysis via LDE:

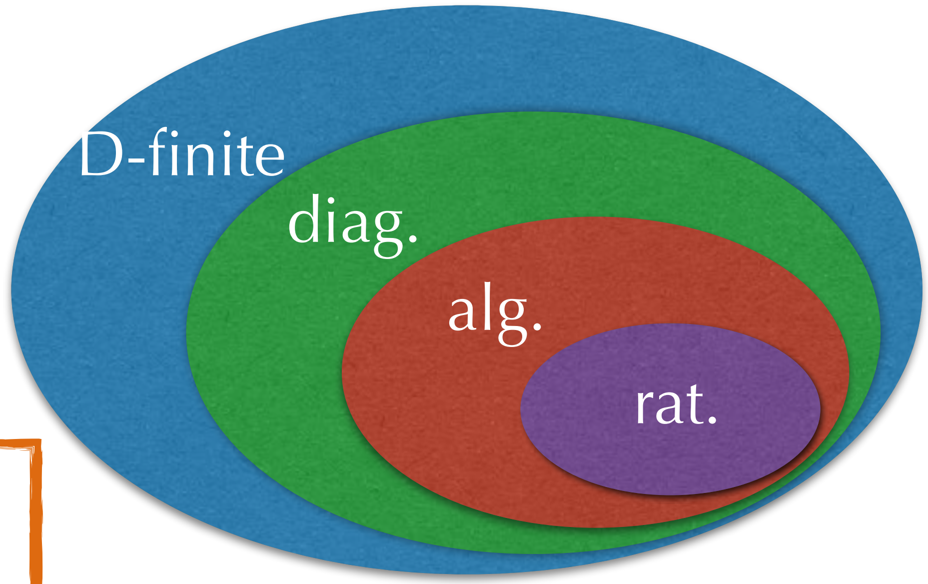
A. Compute a LDE starting from P;

B. For all roots of disc(P), sorted by increasing modulus,

1. compute exactly the local branches;
2. match with numerical continuation (MM's code);
3. if a singular behaviour is encountered, return it.

II. Diagonals

Main Properties



Prop. Algebraic series are the diagonals of bivariate rational functions.
Diagonals are D-finite; they are closed under sum, product, Hadamard product; their coefficients are multiple binomial sums (and conversely).

Christol's conjecture: All D- finite power series with integer coefficients and radius of convergence in $(0,\infty)$ are diagonals.

All these properties are effective, with good bounds and complexity.

→ asymptotics from the LDE

[Pólya21,Furstenberg67,Christol84,BostanLairezS.13,Lairez16,BostanDumontS.17]

LDE for Integrals: Griffiths-Dwork Method

$$\mathrm{I}(t) = \oint \frac{\mathrm{P}(t,\underline{x})}{\mathrm{Q}^m(t,\underline{x})} d\underline{x}$$

Q square-free
Int. over a cycle
where Q≠0.

Basic idea:

1. While m>1, reduce modulo $\mathrm{J} := \langle \partial_1 \mathrm{Q}, \dots, \partial_n \mathrm{Q} \rangle$
   and integrate by parts

$$\frac{P}{Q^m} = \frac{r + v_1 \partial_1 Q + \dots + v_n \partial_n Q}{Q^m} = \frac{r}{Q^m} + \frac{\tilde{P}}{Q^{m-1}} + \text{derivatives}$$

2. Apply to I,I′,I″,… until a linear dependency is found.

Thm. If P/Q has degree *d* in *n* variables, I(t) satisfies a LDE with order ≈d^n, coeffs of degree $d^{O(n)}$.

+Algo in $\tilde{O}(d^{8n})$

Diagonals: $F(\underline{z}) = \frac{G(\underline{z})}{H(\underline{z})} \Rightarrow \Delta F = \left(\frac{1}{2\pi i}\right)^{n-1} \oint F\left(z_1, \dots, z_{n-1}, \frac{t}{z_1 \cdots z_{n-1}}\right) \frac{dz_1 \cdots dz_{n-1}}{z_1 \cdots z_{n-1}}.$

J becomes $\langle z_1 \partial_1 H - z_n \partial_n H, \dots, z_{n-1} \partial_{n-1} H - z_n \partial_n H \rangle$.

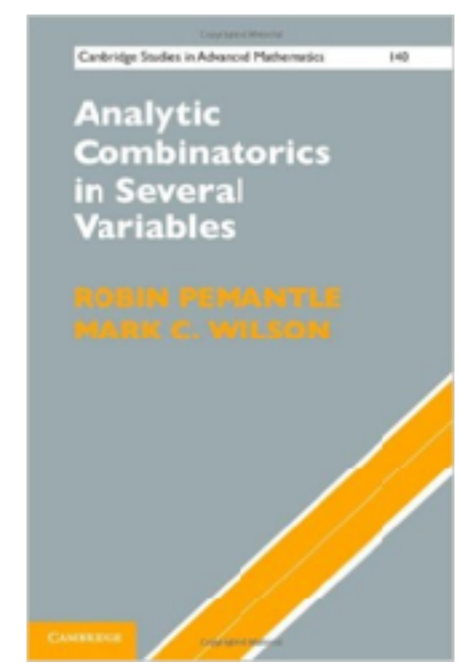
[Griffiths70;Christol84;Bostan-Lairez-S.13;Lairez16]

III. Analytic Combinatorics in Several Variables, with Computer Algebra

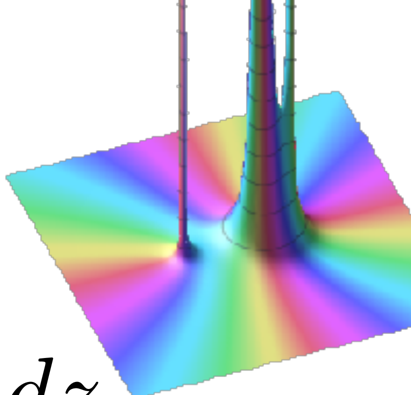
Wanted: complete algorithms, good complexity, more cases with \`explicit' c.

Solution:

1. restrict to simplest class;
2. avoid amoebas and
   deal only with polynomial systems;
3. control all degrees & sizes.



Coefficients of Diagonals



$$F(\underline{z}) = \frac{G(\underline{z})}{H(\underline{z})} \qquad c_{k,\dots,k} = \left(\frac{1}{2\pi i}\right)^n \int_T \frac{G(\underline{z})}{H(\underline{z})} \frac{dz_1 \cdots dz_n}{(z_1 \cdots z_n)^{k+1}}$$

Critical points: minimize $z_1 \cdots z_n$ on $\mathcal{V} = \{\underline{z} \mid H(\underline{z}) = 0\}$

$$\operatorname{rank} \begin{pmatrix} \frac{\partial H}{\partial z_1} & \cdots & \frac{\partial H}{\partial z_n} \\ \frac{\partial (z_1 \cdots z_n)}{\partial z_1} & \cdots & \frac{\partial (z_1 \cdots z_n)}{\partial z_n} \end{pmatrix} \leq 1 \qquad \text{i.e.} \qquad z_1 \frac{\partial H}{\partial z_1} = \cdots = z_n \frac{\partial H}{\partial z_n}$$

J from G-D method

Minimal ones: on the boundary of the domain of convergence.

A 3-step method

1a. locate the critical points (**algebraic** condition);
1b. find the minimal ones (**semi-algebraic** condition);
2. translate (easy in simple cases).

Def. $F(z_1,\dots,z_n)$ is **combinatorial** if every coefficient is ≥ 0.

Prop. [PemantleWilson] In the combinatorial case, one of the minimal critical points has positive real coordinates.

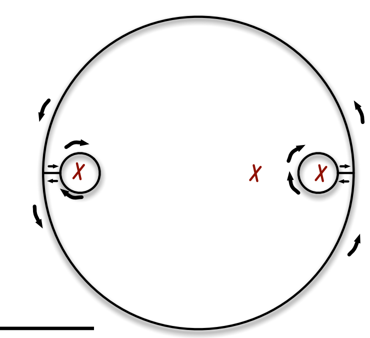
Ex.: Central Binomial Coefficients

$\binom{2k}{k}$: $\quad \frac{1}{1-x-y} = 1 + x + y + 2xy + x^2 + y^2 + \cdots + 6x^2y^2 + \cdots$

(1). Critical points: $1 - x - y = 0, x = y \Longrightarrow x = y = 1/2.$

(2). Minimal ones. Easy. In general, this is the difficult step.

(3). Analysis close to the minimal critical point:



$$a_k = \frac{1}{(2\pi i)^2} \iint \frac{1}{1-x-y} \frac{dx\,dy}{(xy)^{k+1}} \approx \frac{1}{2\pi i} \int \frac{dx}{(x(1-x))^{k+1}}$$

$$\approx \frac{4^{k+1}}{2\pi i} \int \exp(4(k+1)(x-1/2)^2)\,dx \approx \frac{4^k}{\sqrt{k\pi}}.$$

residue

saddle-point approx

Kronecker Representation for the Critical Points

Algebraic part: ``compute'' the solutions of the system

$$H(\underline{z}) = 0 \qquad z_1 \frac{\partial H}{\partial z_1} = \cdots = z_n \frac{\partial H}{\partial z_n}$$

If $\quad \deg(H) = d, \qquad \max \operatorname{coeff}(H) \leq 2^h \qquad D := d^n$

Under genericity assumptions, a probabilistic algorithm running in $\tilde{O}(hD^3)$ bit ops finds:

$$\left.\begin{aligned} P(u) &= 0 \\ P'(u)z_1 - Q_1(u) &= 0 \\ &\vdots \\ P'(u)z_n - Q_n(u) &= 0 \end{aligned}\right\} \quad \begin{aligned} &\text{Degree } \leq D \\ &\text{Height } \leq \tilde{O}(hD) \end{aligned}$$

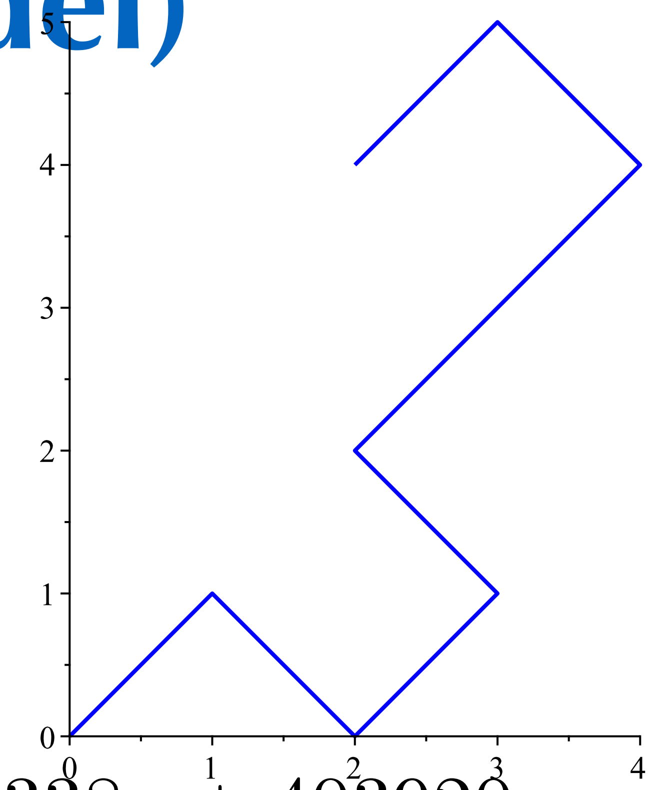
History and Background: see Castro, Pardo, Hägele, and Morais (2001)

System reduced to a univariate polynomial.

[Giusti-Lecerf-S.01;Schost02;SafeySchost16]

Example (Lattice Path Model)

The number of walks from the origin taking steps *{NW,NE,SE,SW}* and staying in the first quadrant is



$$\Delta F, \quad F(x,y,t) = \frac{(1+x)(1+y)}{1-t(1+x^2+y^2+x^2y^2)}$$

Kronecker representation of the critical points:

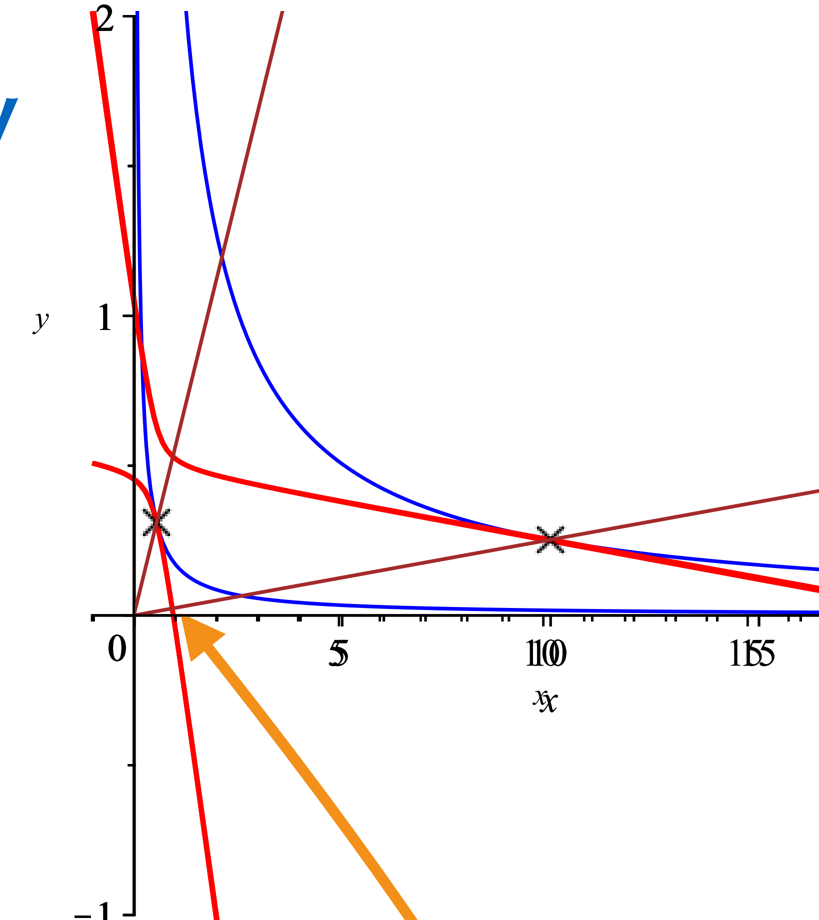
$$P(u) = 4u^4 + 52u^3 - 4339u^2 + 9338u + 403920$$
$$Q_x(u) = 336u^2 + 344u - 105898$$
$$Q_y(u) = -160u^2 + 2824u - 48982$$
$$Q_t(u) = 4u^3 + 39u^2 - 4339u/2 + 4669/2$$

ie, they are given by:

$$P(u) = 0, \quad x = \frac{Q_x(u)}{P'(u)}, \quad y = \frac{Q_y(u)}{P'(u)}, \quad t = \frac{Q_t(u)}{P'(u)}$$

Which one of these 4 is minimal?

Testing Minimality



$$F = \frac{1}{H} = \frac{1}{(1-x-y)(20-x-40y)-1}$$

Critical point equation $x\frac{\partial H}{\partial x} = y\frac{\partial H}{\partial y}$:

$$x(2x+41y-21) = y(41x+80y-60)$$

→ 4 critical points, 2 of which are real:

$$(x_1, y_1) = (0.2528, 9.9971), \quad (x_2, y_2) = (0.30998, 0.54823)$$

Add $H(tx, ty) = 0$ and compute a Kronecker representation:

$$P(u) = 0, \quad x = \frac{Q_x(u)}{P'(u)}, \quad y = \frac{Q_y(u)}{P'(u)}, \quad t = \frac{Q_t(u)}{P'(u)}$$

Solve numerically and keep the real positive sols:

$$(0.31, 0.55, 0.99), (0.31, 0.55, 1.71), (0.25, 9.99, 0.09), (0.25, 0.99, 0.99)$$

(x_1, y_1) is not minimal, (x_2, y_2) is.

Algorithm and Complexity

Thm. If $F(\underline{z})$ is combinatorial, then under regularity conditions, the points contributing to dominant diagonal asymptotics can be determined in $\tilde{O}(hd^5D^4)$ bit operations. Each contribution has the form

$$A_k = \left(T^{-k}k^{(1-n)/2}(2\pi)^{(1-n)/2}\right)(C + O(1/k))$$

T, C can be found to $2^{-\kappa}$ precision in $\tilde{O}(h(dD)^3 + D\kappa)$ bit ops.

explicit algebraic numbers

half-integer

This result covers the easiest cases.
All conditions hold generically and can be checked within the same complexity, except combinatoriality.

[Melczer-S.16]

Example: Apéry's sequence

$$\frac{1}{1-t(1+x)(1+y)(1+z)(1+y+z+yz+xyz)} = 1 + \cdots + 5xyzt + \cdots$$

Kronecker representation of the critical points:

$$P(u) = u^2 - 366u - 17711$$

$$x = \frac{2u-1006}{P'(u)}, \quad y = z = -\frac{320}{P'(u)}, \quad t = -\frac{164u+7108}{P'(u)}$$

There are two real critical points, and one is positive. After testing minimality, one has proved asymptotics

```
> A, U := DiagonalAsymptotics(numer(F),denom(F),[t,x,y,z], u, k):
> evala(allvalues(subs(u=U[1],A)));
```

$$\frac{\left(17+12\sqrt{2}\right)^k \sqrt{2}\sqrt{24+17\sqrt{2}}}{8k^{3/2}\pi^{3/2}}$$

Example: Restricted Words in Factors

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

words over {0,1} without 10101101 or 1110101

```
> A,U:=DiagonalAsymptotics(numer(F),denom(F),indets(F),u,k,true,u-T,T):
> A;
```

$$\Bigg(\Bigg(\frac{84u^{20} + 240u^{19} - 285u^{18} - 1548u^{17} - 2125u^{16} - 1408u^{15} + 255u^{14} + 756u^{13} + 2509u^{12} + 2856u^{11} - 605u^{10} + 2020u^{9} + 1233u^{8} - 1760u^{7} + 924u^{6} - 492u^{5} - 675u^{4} + 632u^{3} - 249u^{2} + 24u + 16}{-12u^{20} + 30u^{19} + 258u^{18} + 500u^{17} + 440u^{16} - 102u^{15} - 378u^{14} - 1544u^{13} - 2142u^{12} - 550u^{11} - 2222u^{10} - 1644u^{9} + 2860u^{8} - 1848u^{7} + 1230u^{6} + 2160u^{5} - 2686u^{4} + 1494u^{3} - 228u^{2} - 320u + 84}\Bigg)^k$$

$$\sqrt{2}\sqrt{\frac{84u^{20} + 240u^{19} - 285u^{18} - 1548u^{17} - 2125u^{16} - 1408u^{15} + 255u^{14} + 756u^{13} + 2509u^{12} + 2856u^{11} + 605u^{10} - 2020u^{9} + 1233u^{8} - 1760u^{7} + 924u^{6} - 492u^{5} - 675u^{4} + 632u^{3} - 249u^{2} + 24u + 16}{-162u^{18} - 612u^{17} - 902u^{16} - 616u^{15} + 254u^{14} + 548u^{13} + 2054u^{12} + 2156u^{11} + 898u^{10} - 2268u^{9} + 2462u^{8} - 2088u^{7} + 1312u^{6} - 540u^{5} - 1410u^{4} + 1188u^{3} - 290u^{2} + 32}}$$

$$(12u^{20} + 36u^{19} - 21u^{18} - 170u^{17} - 255u^{16} - 190u^{15} - 19u^{14} + 46u^{13} + 461u^{12} + 628u^{11} + 133u^{10} + 374u^{9} + 161u^{8} - 384u^{7} + 146u^{6} - 138u^{5} - 285u^{4} - 40u^{3} + 91u^{2} - 30u + 32)\Bigg) \Bigg/ \Big(2\sqrt{k}\sqrt{\pi}\,(84u^{20}$$

$$+ 240u^{19} - 285u^{18} - 1548u^{17} - 2125u^{16} - 1408u^{15} + 255u^{14} + 756u^{13} + 2509u^{12} - 2856u^{11} - 605u^{10} + 2020u^{9} - 1233u^{8} - 1760u^{7} + 924u^{6} - 492u^{5} - 675u^{4} + 632u^{3} - 249u^{2} + 24u + 16)\Big)$$

```
> U;
```

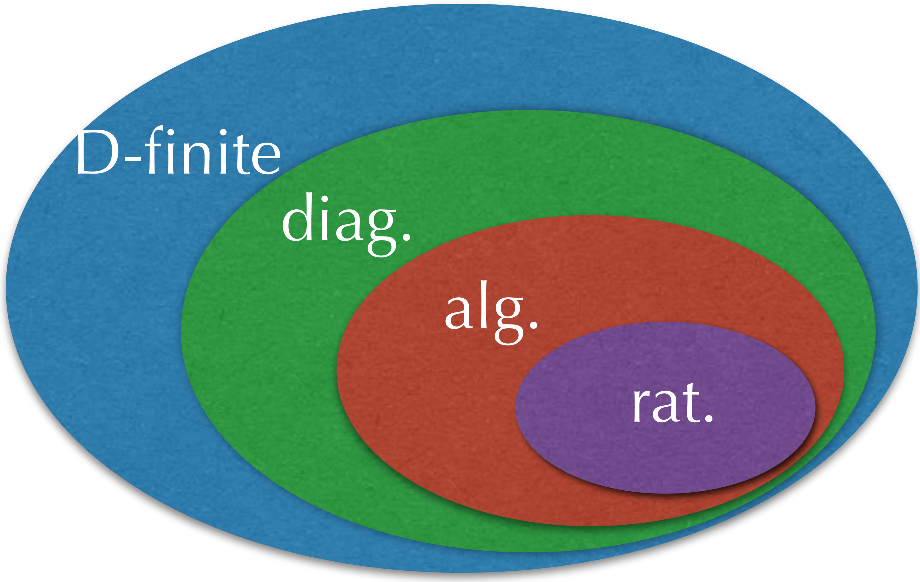
$$[\mathit{RootOf}(4\_Z^{21} + 12\_Z^{20} - 15\_Z^{19} - 86\_Z^{18} - 125\_Z^{17} - 88\_Z^{16} + 17\_Z^{15} + 54\_Z^{14} + 193\_Z^{13} + 238\_Z^{12} + 55\_Z^{11} + 202\_Z^{10} + 137\_Z^{9} - 220\_Z^{8} + 132\_Z^{7} - 82\_Z^{6} - 135\_Z^{5} + 158\_Z^{4} - 83\_Z^{3} + 12\_Z^{2}$$

$$+ 16\_Z - 4, 0.25574184)]$$

```
> evalf(subs(u=U[1],A));
```

$$\frac{0.60294599\ 3.9101932^k}{\sqrt{k}}$$

Summary & Conclusion



- In many cases, LDE + certified analytic continuation works.
- Don't miss Marc's talk (and bring your computer).
- Diagonals are a nice and important class of generating functions for which we now have many good algorithms.
- ACSV can be made effective (at least in simple cases) and recovers explicit constants.
- Complexity issues become clearer.

Work in progress: extend beyond some of the assumptions (see Melczer's talk & thesis).

The End