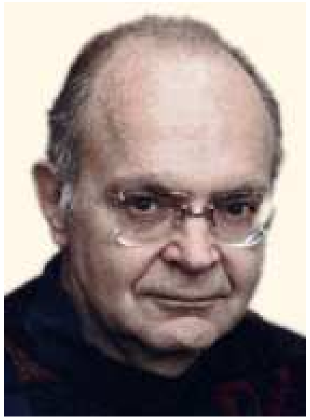
Singularity Analysis --- A Perspective

Philippe Flajolet, INRIA (France)

@ Queen Mary University, London.
FEB 15, 2008



Analysis of Algorithms

↓

Average-Case, Probabilistic

↓

Properties of Random Structures?

- Counting and asymptotics $\boxed{n! \sim n^n e^{-n}\sqrt{2\pi n}}$
- Asymptotic laws $\boxed{\Omega_n \xrightarrow{\mathcal{D}} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2}\, dt.}$ (e.g., Monkey and typewriter!)

— Probabilistic, stochastic

— Analytic Combinatorics: $\boxed{\text{Generating Functions}}$

1. Introduction

“Symbolic” Methods

Rota-Stanley; Foata-Schutzenberger; Joyal and UQAM group; Jackson-Goulden, &c; F.; ca $1980^{\pm}$. F-Salvy-Zimmermann 1991 $\rightsquigarrow$ *Computer Algebra*.

Basic combinatorial constructions admit of direct translations as operators over generating functions (GF’s).

$\mathcal{C}$: class of comb. structures;

C_n : # objects of size n

$$\Downarrow\Downarrow\Downarrow$$

(counting) $\begin{cases} C(z) := \sum C_n z^n \\ \widehat{C}(z) := \sum C_n \dfrac{z^n}{n!} \end{cases}$

(params) $\begin{cases} C(z,u) := \sum C_{n,k} z^n u^k \\ \widehat{C}(z,u) := \sum C_{n,k} u^k \dfrac{z^n}{n!} \end{cases}$

Ordinary GF's for unlabelled structures. Exponential GF's for labelled structures.

"Dictionaries"

= Constructions viewed as Operators over GF's.

Constr.	Operations	
Union	$+$	$+$
Product	$\times$	$\times$
$\mathfrak{S}$equence	$(1-f)^{-1}$	$(1-f)^{-1}$
$\mathfrak{M}$ultiSet	Pólya Exp.	e^f
$\mathfrak{C}$ycle	Pólya Log.	$\log(1-f)^{-1}$
	(unlab.)	(lab.)

$$\mathrm{Exp}(f) := \exp\left(f(z) + \tfrac{1}{2} f(z^2) + \cdots\right)$$

$$\mathrm{Log}(f) := \log \frac{1}{1-f(z)} + \cdots$$

Books: Goulden-Jackson, Bergeron-LL, Stanley, F-Sedgewick

$\Longrightarrow$ How to extract coeff., especially, asymptotically??

$\boxed{\textbf{“Complex–analytic Structures”}}$

Interpret:

♡ *Counting GF* as *analytic transformation of* $\mathbb{C}$;

♡ *Comb. Construction* as *analytic functional.*

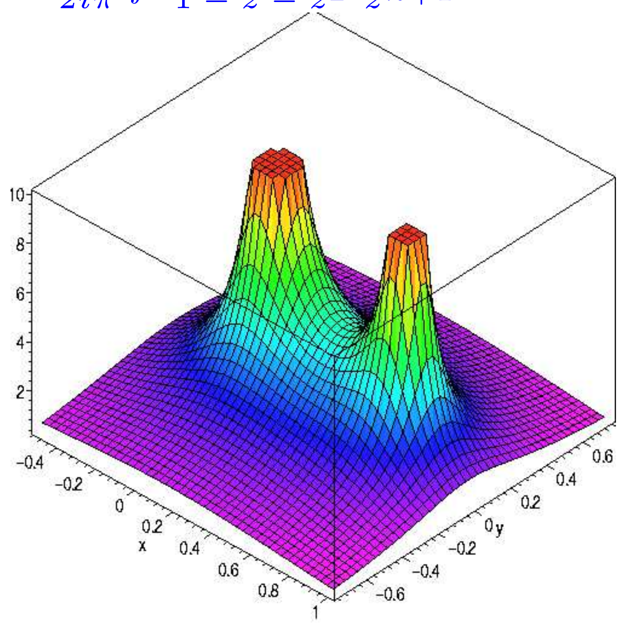
$\boxed{\textbf{Singularities}\text{ are crucial to }\textbf{asymptotic}\text{ prop's!}}$

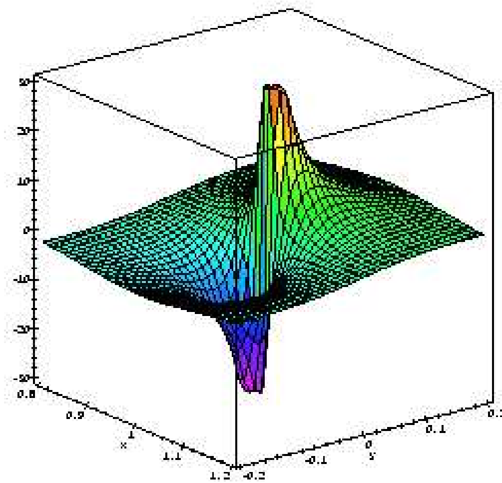
(cf. analytic number theory, complex analysis, etc)

Asymptotic counting via Singularity Analysis (S.A.)

Asymptotic laws via Perturbation + S.A.

$$\frac{1}{2i\pi}\int \frac{1}{1-z-z^2}\,\frac{dz}{z^{n+1}} \qquad\qquad \Im f(z), \quad f(z)=(1-z-z^2)^{-1}.$$





Refs: F–Odlyzko, SIAM A&DM, 1990 $\ll$ FO82 on tree height; Odlyzko's 1995 survey in *Handbook of Combinatorics*

+ Banderier, Fill, J. Gao, Gonnet, Gourdon, Kapur, G. Labelle, Laforest, T. Lafforgue, Noy, Odlyzko, Panario, Poblete, Pouyanne, Prodinger, Puech, Richmond, Robson, Salvy, Schaeffer, Sipala, Soria, Steyaert, Szpankowski, B. Vallée, Viola .

♠ Location of singularity at $z = \rho$: coeff. $[z^n]f(z) = \rho^{-n} \cdot$ coeff. $[z^n]f(\rho z)$

♠ Nature of singularity at $z = 1$:

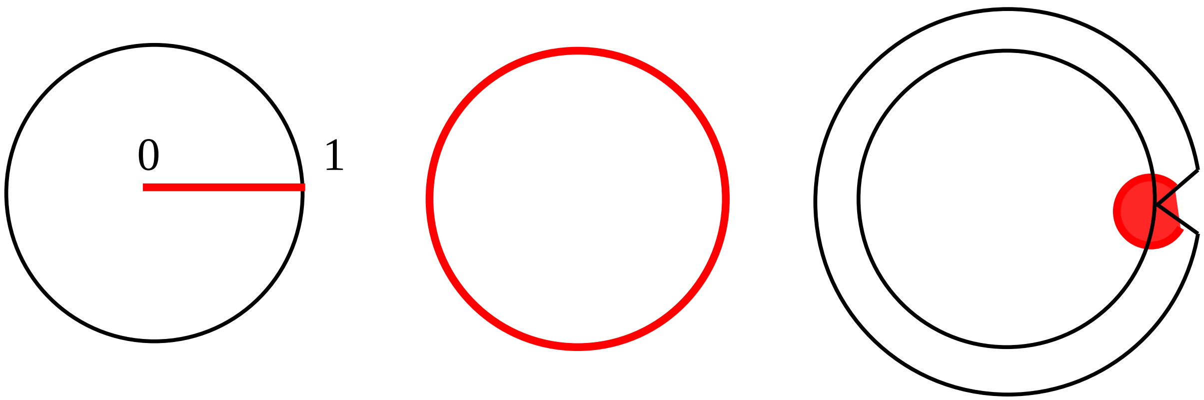
$\dfrac{1}{(1-z)^2}$	$\longrightarrow$	$n+1$	$\sim$	n
$\dfrac{1}{1-z}\log\dfrac{1}{1-z}$	$\longrightarrow$	$H_n \equiv \frac{1}{1} + \ldots + \frac{1}{n}$	$\sim$	$\log n$
$\dfrac{1}{1-z}$	$\longrightarrow$	1	$\sim$	1
$\dfrac{1}{\sqrt{1-z}}$	$\longrightarrow$	$\dfrac{1}{2^{2n}}\dbinom{2n}{n}$	$\sim$	$\dfrac{1}{\sqrt{\pi n}}$

$$\left\{\begin{array}{lll} \text{Location of sing's :} & \text{Exponential factor} & \rho^{-n} \\ \text{Nature of sing's :} & \text{“Polynomial” factor} & \vartheta(n) \end{array}\right.$$

Generating Function $\rightsquigarrow$ Coefficients

Solving a "Tauberian" problem

$\mathbb{R}$eal–Tauberian	Darboux-Pólya	Singularity An.



(large $\Longrightarrow$ large)	(smooth $\Longrightarrow$ small)	(Full mappings)

Combinatorial constructions $\rightsquigarrow$ Analytic Functionals

$\Longrightarrow$ Analytic continuation prevails for comb. GF's

2. Basic Singularity Analysis

Theorem 1. *Basic scale translates:*

$$\sigma_{\alpha,\beta}(z) := (1-z)^{-\alpha}\left(\tfrac{1}{z}\log\tfrac{1}{1-z}\right)^{\beta}$$

$$\Longrightarrow \qquad [z^n]\sigma_{\alpha,\beta} \underset{n\to\infty}{\sim} \frac{n^{\alpha-1}}{\Gamma(\alpha)}(\log n)^{\beta}.$$

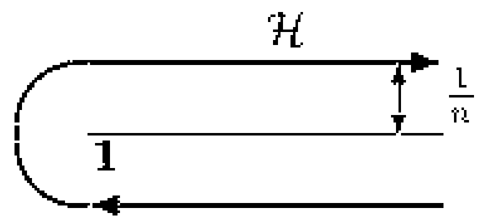
PROOF. Cauchy's coefficient integral, $f(z) = (1-z)^{-\alpha}$

$$[z^n]f(z) = \frac{1}{2i\pi}\int_{\gamma} f(z)\,\frac{dz}{z^{n+1}}$$

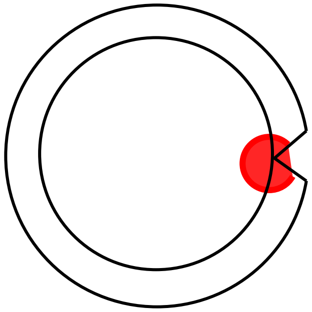
$$\Downarrow \qquad \left(z = 1 + \frac{t}{n}\right) \qquad \Downarrow$$

$$\frac{1}{2i\pi}\int_{\mathcal{H}}\left(-\frac{t}{n}\right)^{-\alpha} e^{-t}\,\frac{dt}{n}$$

$$n^{\alpha-1}\times\tfrac{1}{\Gamma(\alpha)}.$$



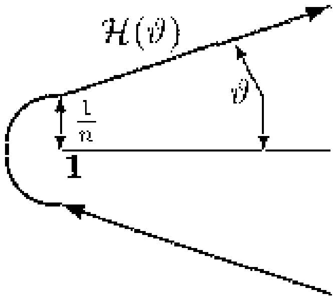
“Camembert”



Theorem 2. *$\mathcal{O}$–transfers: Under continuation in a Δ-domain,*

$$f(z) = O(\sigma_{\alpha,\beta}(z)) \quad \Longrightarrow \quad [z^n]f(z) = O\left([z^n]\sigma_{\alpha,\beta}(z)\right).$$

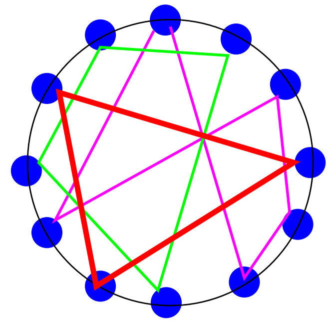
PROOF:



Usage: $\left\{ \begin{array}{c} f(z) = \lambda\sigma(z) + \mu\tau(z) + \ldots + O(\omega(z)) \\ \Longrightarrow \\ f_n = \lambda\sigma_n + \mu\tau_n + \ldots + O(\omega_n). \end{array} \right.$

Similarly: o-transfer.

- Dominant singularity at ρ gives factor ρ^{-n}.
- Finitely many singularities work fine



EXAMPLE **1**. *2-regular graphs* [Comtet] (Originally by Darboux-Pólya.)

$$\mathcal{G} = \mathfrak{M}\left(\frac{1}{2}\mathfrak{C}_{\geq 3}(\mathcal{Z})\right)$$

$$\widehat{G}(z) = \exp\left(\frac{1}{2}\log\frac{1}{1-z} - \frac{z}{2} - \frac{z^2}{4}\right)$$

$$\widehat{G}(z) \underset{z\to 1}{\sim} \frac{e^{-3/4}}{\sqrt{1-z}}$$

$$\frac{G_n}{n!} \underset{n\to\infty}{\sim} \frac{e^{-3/4}}{\sqrt{\pi n}}.$$

$\square$

```
> equivalent(exp(-z/2-z^2/4)/sqrt(1-z),z,n,4); # By SALVY
                  1/2                            3/2                               5/2
exp(-3/4) (1/n)          exp(-3/4) (1/n)                exp(-3/4) (1/n)
------------------ - 5/8 ------------------ + 1/128 ------------------
         1/2                      1/2                          1/2
       Pi                       Pi                           Pi
```

EXAMPLE **2**. *Richness index of trees* [F-Sipala-Steyaert,90]
= Number of different terminal subtrees. Catalan case:

$$K(z) = \frac{1}{2z} \sum_{k \geq 0} \frac{1}{k+1} \binom{2k}{k} \left(\sqrt{1 - 4z - 4z^{k+1}} - \sqrt{1 - 4z} \right)$$

$$K(z) \underset{z \to 1/4}{\approx} \frac{1}{\sqrt{Z \log Z}}, \qquad Z := 1 - 4z$$

$$\text{Mean index} \underset{n \to \infty}{\sim} C \frac{n}{\sqrt{\log n}}, \qquad C \equiv \sqrt{\frac{8 \log 2}{\pi}}.$$

= Compact tree representations as DAGs = Common Subexpression Pb. $\square$

Extensions

$\heartsuit$ Slowly varying $\Longrightarrow$ slowly varying: Log-log $\Longrightarrow$ Log-Log, ...

$\heartsuit$ Full asymptotic expansions

$\heartsuit$ Uniformity of coefficient extraction $[z^n]\{F_u(z)\}_{u\in\Omega}$ = $\leadsto$ later!.

$\heartsuit$ Some cases with natural boundary [Fl-Gourdon-Panario-Pouyanne]

EXAMPLE **3**. *Distinct Degree Factorization* [DDF] in Polynomial Fact $\leadsto$ Greene–Knuth:

$$[z^n]\prod_{k=1}^{\infty}\left(1+\frac{z^k}{k}\right).$$

Hybrid w/ Darboux: $e^{-\gamma}+\frac{e^{-\gamma}}{n}+\cdots+\star\frac{(-1)^n}{n^3}+\star\frac{\omega^n}{n^3}+\cdots$ $\square$

Cf. Hardy-Ramanujan's partition analysis "without contrast".

3. Closure Properties

Function of S.A.–type = amenable to singularity analysis

- is continuable in a Δ-domain,
- admits singular expansion in scale $\{\sigma_{\alpha,\beta}\}$.

Theorem 3. *Generalized polylogarithms*

$$\mathrm{Li}_{\alpha,k} := \sum (\log n)^k n^{-\alpha} z^n$$

are of S.A.-type.

PROOF. Cauchy-Lindelöf representations

$$\sum \varphi(n)(-z)^n = -\frac{1}{2i\pi}\int_{1/2-i\infty}^{1/2+i\infty} \varphi(s) z^s \frac{\pi}{\sin \pi s}\, ds.$$

+ Mellin transform techniques (Ford, Wong, F.).

EXAMPLE 4. *Entropy of Bernoulli distribution*

$$H_n := -\sum_k \pi_{n,k} \log \pi_{n,k}, \qquad \pi_{n,k} \equiv \binom{n}{k} p^k (1-p)^{n-k}$$

$$\text{involves } \sum \log(k!) z^k = (1-z)^{-1} \operatorname{Li}_{0,1}(z)$$

$$\frac{1}{2} \log n + \frac{1}{2} + \log \sqrt{2\pi p(1-p)} + \cdots .$$

Redundancy, coding, information th.; Jacquet-Szpankowski via Analytic dePoissonization. □

- Elements like $\log n$, $\sqrt{n}$ in combinatorial sums

Theorem 4. *Functions of S.A.-type are closed under integration and differentiation.*

PROOF. Adapt from Olver, Henrici, etc.

Theorem 5. *Functions of S.A.-type are closed under Hadamard product*

$$f(z) \odot g(z) := \sum_n (f_n g_n) z^n.$$

PROOF. Start from Hadamard's formula

$$f(z) \odot g(z) = \frac{1}{2i\pi} \int_\gamma f(t) g\left(\frac{w}{t}\right) \frac{dt}{t}.$$

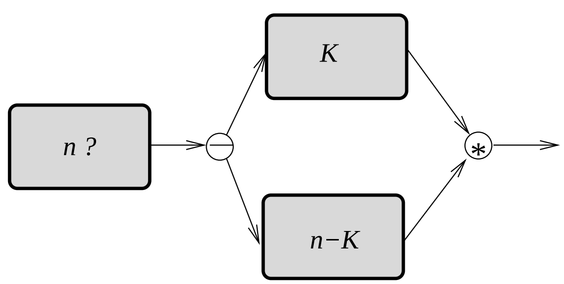
+ adapt Hankel contours [H., Jungen, R. Wilson $\rightsquigarrow$ Fill-F-Kapur]

EXAMPLE 5. *Divide-and -conquer recurrences*

$$f_n = t_n + \sum \pi_{n,k}(f_k + f_{n-k})$$
$$\mathrm{Sing}(f(z)) = \Phi(\mathrm{Sing}(t(z)))$$
$$\mathrm{Asympt}[f_n] = \Psi(\mathrm{Sing}(t)).$$

E.g., Catalan statistics: need $\sum \binom{2n}{n} \log n \cdot z^n$.

Useful in random tree applications [Fill-F-Kapur, 2004^+, Fill-Kapur] // Neininger-Hwang *et al.* $\ll$ Knuth-Pittel. Moments $\leftrightarrow$ contraction method [Rösler-Rüschendorf-Neininger] $\square$



4. Functional Equations

• Rational functions. Linear system $\mathbb{Q}_{\geq 0}[z]$ implies polar singularities:

$$[z^n]f(z) \approx \sum \omega^n n^k, \qquad \omega \in \overline{\mathbb{Q}}, \quad k \in \mathbb{Z}_{\geq 0}.$$

+ irreducibility: Perron-Frobenius $\Longrightarrow$ *simple dom. pole*.

• ***Word problems*** from regular language models;

• *Transfer matrices* [Bender-Richmond]: dimer in strip, knights, etc.

$\rightsquigarrow$ Vallée's generalization to *dynamical sources* via transfer operators.

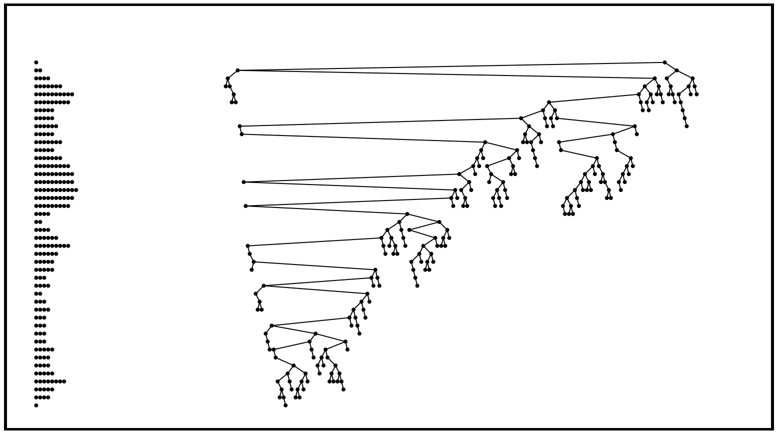
• Algebraic functions, by Puiseux expansions $(Z^{p/q})$ $\ll$ S.A. or Darboux!

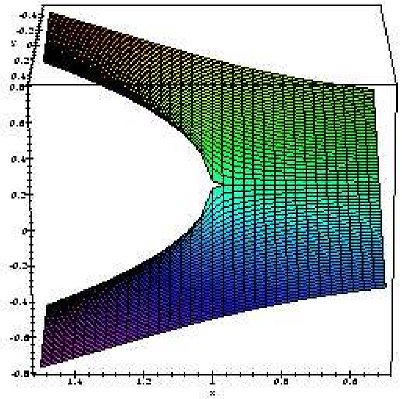
$$[z^n]f(z) \approx \sum \sum \omega^n n^{p/q}, \qquad \omega \in \overline{\mathbb{Q}}, \quad p/q \in \mathbb{Q},$$

Asymptotics of coeff. is decidable [Chabaud-F-Salvy].

• ***Word problems from context-free models***;

• *Trees*; *Geom. configurations* (non-crossing graphs, polygonal triangs.);
Planar Maps [Tutte...]; *Walks* [Banderier Bousquet-M., Schaeffer], ...





Square-root singularity is "**universal**" for many recursive classes = controlled "failure" of Implicit Function Theorem $Z \propto Y^2$

Entails coeff. asymptotic $\approx \omega^n n^{-3/2}$ with critical exponent $-3/2$.

E.g., unbalanced 2–3 trees (Meir-Moon): $f = z\phi(f),\ \phi(u) = 1 + u^2 + u^3.$

Pólya's combinatorial chemistry programme:

$$\boxed{f(z) = z\,\mathrm{Exp}(f(z)) \equiv z e^{f(z) + \frac{1}{2} f(z^2) + \frac{1}{3} f(z^3) + \cdots}}$$

Starting with Pólya 1937; Otter 1949; Harary-Robinson et al. 1970's; Meir-Moon 1978; Bender/Meir-Moon; Drmota-Lalley-Woods thm. 1990^{+}

• "Holonomic" functions. Defined as solutions of linear ODE's with coeffs in $\mathbb{C}(z)$ [Zeilberger] $\equiv$ $\mathcal{D}$-finite.

$$\mathcal{L}[f(z)] = 0, \qquad \mathcal{L} \in \mathbb{C}(z)[\partial_z].$$

• Stanley, Zeilberger, Gessel: Young tableaux and permutation statistics; regular graphs, constrained matrices, etc.

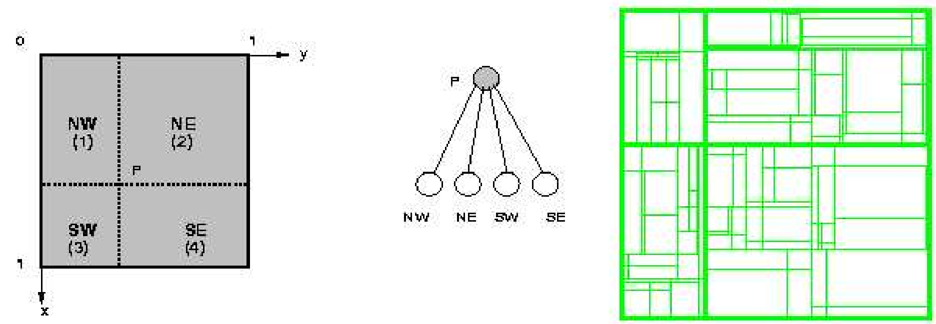
Fuchsian case (or "regular" singularity) ($Z^\beta \log^k Z$):

$$[z^n]f(z) \approx \sum \omega^n n^\beta (\log n)^k, \qquad \omega, \beta \in \overline{\mathbb{Q}}, \quad k \in \mathbb{Z}_{\geq 0}.$$

S.A. applies automatically to classical classification.

Asymptotics of coeff is decidable
— general case: modulo oracle for connection problem;
— strictly positive case: "usually" OKay.

QTrees:



EXAMPLE 6. *Quadtrees—Partial Match* [FGPR'92]

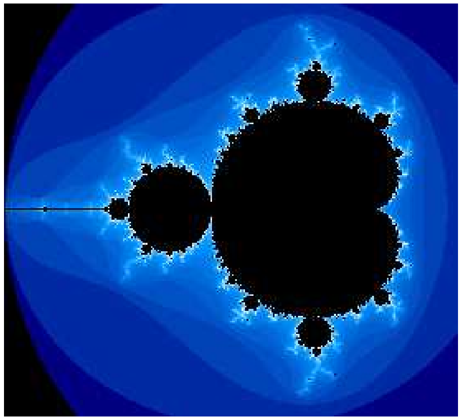
Divide-and-conquer recurrence with coeff. in $\mathbb{Q}(n)$

Fuchsian equation of order d (dimension) for GF

$$Q_n^{(d=2)} \asymp n^{(\sqrt{17}-3)/2}.$$

E.g., $d = 2$: Hypergeom ${}_2F_1$ with algebraic arguments. $\square$

Extended by Hwang et al. Cf also Hwang's *Cauchy ODE* cases.
Panholzer-Prodinger+Martinez, ...



- Functional Equations and Substitution.

- Early example of *balanced 2–3 trees* by Odlyzko, 1979.

$$T(z) = z + T(\tau(z)), \qquad \tau(z) := z^2 + z^3.$$

Infinitely many exponents with common real part implies periodicities: $T_n \sim \dfrac{\phi^n}{n}\Omega(\log n)$.

• Singular iteration for *height of trees* (binary and other simple varieties; F-Gao-Odlyzko-Richmond; cf Renyi-Szekeres):

$$y_h = z + y_{h-1}^2, \qquad y_0 = z.$$

— Moments and convergence in law; Local limit law of ϑ-type.

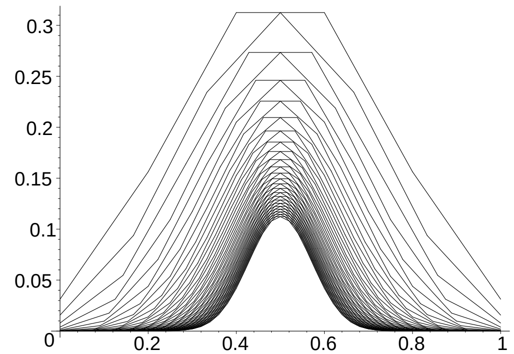
Applies to branching processes conditioned on total progeny.

Cf Chassaing-Marckert for // probabilistic approaches $\rightsquigarrow$ **width**

• *Digital search trees* via q–hypergeometrics: singularities accumulate geometrically $\rightsquigarrow$ periodicities [F-Richmond]:

$$\partial_z^k f(z) = t(z) + 2e^{z/2} f(\frac{z}{2}).$$

• *Order of binary trees* (Horton-Strahler, Register function; F-Prodinger) via Mellin tr. of GF and & singularities.



5. Limit Laws

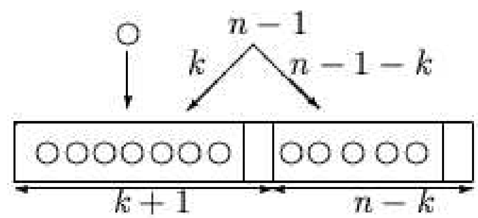
• Moment pumping from bivariate GF

Early theories by Kirschenhofer-Prodinger-Tichy (1987)

Factorial moment of order k: $[z^n]\left(\frac{\partial}{\partial^k}F(z,u)\right)_{u=1}$

EXAMPLE **7**. *Airy distribution of areas* shows up in area below paths, path length in trees, Linear Probing Hashing, inversions in increasing trees, connectivity of graphs.

$$\frac{\partial}{\partial z}F(z,q) = F(z,q)\cdot\frac{F(z,q)-qF(qz,q)}{1-q}$$



Louchard-Takács[Darboux]; Knuth; F-Poblete-Viola // Chassaing-Marckert □

Classical probability theory: sums of Random Variables $\rightsquigarrow$ powers of fixed function (PGF, Fourier tr.) $\rightsquigarrow$ $\mathcal{N}$ormal Law.

For problems expressed by Bivariate GF (BGF): field founded by E. Bender *et al.* + developments by F, Soria, **Hwang**, ...

Idea: BGF $F(z,u) = \sum f_n(u) z^n$, where $f_n(u)$ describes parameter on objects of size n. If (for u near 1)

$$f_n(u) \approx \omega(u)^{\kappa_n}, \qquad \kappa_n \to \infty,$$

then speak of $\boxed{\text{Quasi-Powers}}$ approximation. Recycle continuity theorem, Berry-Esseen, Chernov, etc. $\Longrightarrow$ $\mathcal{N}$ormal law and many goodies...

(speed of convergence, large deviation fn, local limits)

Two important cases:

- Movable singularity:

$$F(z,u) \approx \left(1 - \frac{z}{\rho(u)}\right)^{-\alpha} \Longrightarrow \frac{f_n(u)}{f_n(1)} \approx \left(\frac{\rho(1)}{\rho(u)}\right)^{n}.$$

- Variable exponent:

$$F(z,u) \approx \left(1 - \frac{z}{\rho}\right)^{-\alpha(u)} \Longrightarrow \frac{f_n(u)}{f_n(1)} \approx \begin{cases} n^{\alpha(u)-\alpha(1)} \\ \left(e^{\alpha(u)-\alpha(1)}\right)^{\log n} \end{cases}.$$

Requires *uniformity* afforded by *Singularity Analysis*
($\neq$ Tauber or Darboux).

Singularity Perturbation analysis (smoothness)

$\downarrow$

Uniform Quasi-Powers for coeffs

$\downarrow$

$\mathcal{N}$ormal limit law

EXAMPLE **8**. *Polynomials over finite fields.*

```
> Factor(x^7+x+1) mod 29;
   3    2                 2                  2
 (x  + x  + 3 x + 15) (x  + 25 x + 25) (x  + 3 x + 14)
```

- $\mathcal{P}$olynomial is a $\mathfrak{S}$*equence* of coeffs: $\mathcal{P}$ has Polar singularity.
- By unique factorization, $\mathcal{P}$ is also $\mathfrak{M}$*ultiset of* $\mathcal{I}$*rreducibles*:

  $\mathcal{I}$ has log singulariy.

$\Longrightarrow$ Prime Number Theorem for Polynomials $I_n \sim \dfrac{q^n}{n}$.

- Marking number of $\mathcal{I}$–factors is approx uth power:

$$P(z,u) \approx \left(e^{I(z)}\right)^u.$$

Variable Exponent $\Longrightarrow$ $\mathcal{N}$ormality of # of irred. factors.
(cf Erdős-Kac for integers.) □

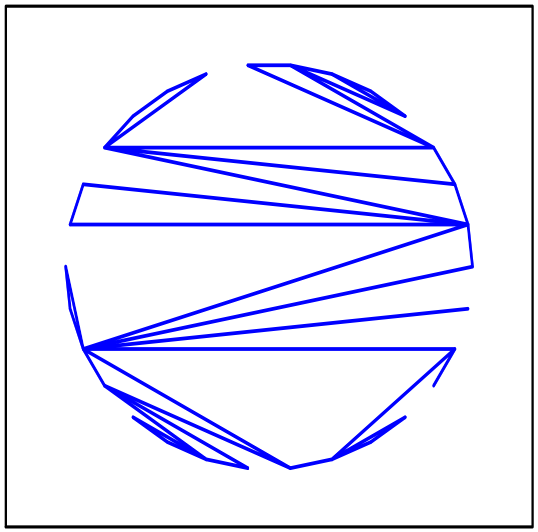
(Analysis of polynomial fact. algorithms, [F-Gourdon-Panario])

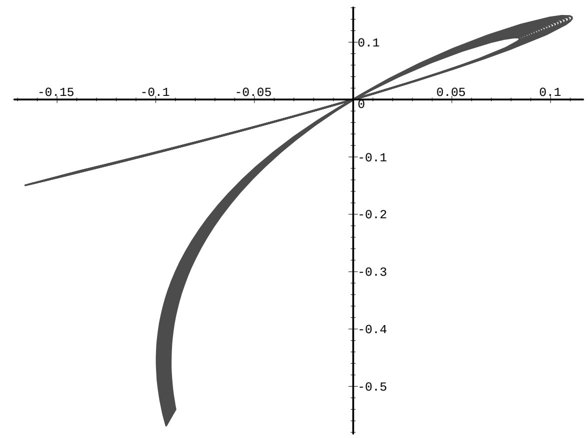
ACCGATCATTAGCAGATTATCATTTACTGAGAGTACTTAACATGCCA

EXAMPLE **9**. *Patterns in Random Strings* = Perturbation of linear system of eqns. (& many problems with finite automata, paths in graphs)

Linear system $X = X_0 + \mathbf{T}X$ w/ Perron-Frobenius. Auxiliary mark u induces smooth singularity displacement. For “natural” problems: $\mathcal{N}$ormal limit law. cf [Régnier & Szpankowski], ... $\square$

Also sets of patterns; similarly for patterns in increasing labelled trees, in permutations, in binary search trees [F-Gourdon-Martinez].
Generalized patterns and/or sources by Szpankowski, Vallée, ...



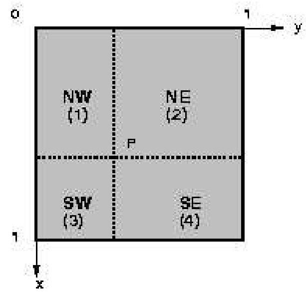


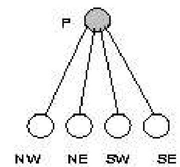
EXAMPLE **10**. *Non crossing graphs.* [F-Noy]
= Perturbation of algebraic equation.

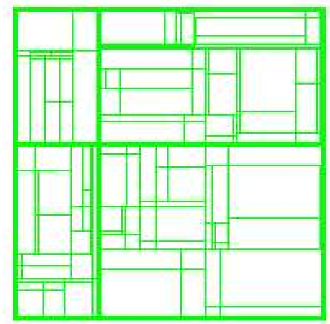
$$G^3 + (2z^2 - 3z - 2)G^2 + (3z + 1)G = 0$$
$$G^3 + (2u^3z^2 - 3u^2z + u - 3)G^2 + (3u^2 - 2u + 3)G + u - 1 = 0$$

Movable singularity scheme applies: $\mathcal{N}$ormality.

+ Patterns in context-free languages, in combinatorial tree models, in functional graphs: Drmota's version of Drmota-Lalley-Woods. □







EXAMPLE **11**. *Profile of Quadtrees.*

$$F(z,u) = 1 + 2^3 u \int_0^z \frac{dx_1}{x_1(1-x_1)} \int_0^{x_1} \frac{dx_2}{1-x_2} \int_0^{x_2} F(x_3,u) \frac{dx_3}{1-x_3}.$$

Solution is of the form $(1-z)^{-\alpha(u)}$ for algebraic branch $\alpha(u)$;
Variable Exponent $\Longrightarrow$ $\mathcal{N}$ormality of search costs. □

Applies to many linear differential models that behave like *cycles-in-perms*.

<u>EXAMPLE **12**</u>. *Urn models.* 2×2–balanced.

$$(u^5 z - u)\frac{\partial G}{\partial z} + (1 - u^6)\frac{\partial G}{\partial u} + u^5 G = 0$$

[FGP'03] $\leftrightarrow$ Janson, Mahmoud, Puyhaubert, Panholzer-Prodinger, . . . $\square$

• ***Coalescence of singularities and/or exponents***: e.g. **Maps** = Airy Law $\equiv$ Stable($\frac{3}{2}$) [BFSS'01]. Cf Pemantle, Wilson, Lladser, ...

Conclusions

For combinatorial counting and limit laws:
Modest technical apparatus & *generic technology*.
High-level for applications, esp., *analysis of algorithms*.

Plug-in on *Symbolic Combinatorics & Symbolic Computation*.

Discussion of *Schemas & "Universality"* in metric aspects of random discrete structures.

E.g. Borges' theorem for words, trees, labelled trees, mappings, permutations, increasing trees, maps, etc.

Thank you!
♡♡♡