D-finiteness:
Algorithms and Applications

Alin Bostan, Frédéric Chyzak, Bruno Salvy

Algorithms Project, Inria

March 21, 2007
I Introduction
Theorem (Richardson 68)

*In the class obtained from \( \mathbb{Q}(x), \pi, \log 2 \) by the operations \(+, -, \times\) and composition with \( \exp, \sin \) and \(| \cdot |\), testing for zero-equivalence is undecidable.*

Consequences:

1. “Simplification” is always difficult;
2. Computer algebra isolates classes for which it can provide algorithms.
Example: Algebraic Numbers

An irreducible polynomial is a good data-structure for representing its roots.

\[
x = \frac{\sin \frac{2\pi}{7}}{\sin^2 \frac{3\pi}{7}} - \frac{\sin \frac{\pi}{7}}{\sin^2 \frac{2\pi}{7}} + \frac{\sin \frac{3\pi}{7}}{\sin^2 \frac{\pi}{7}} = 2\sqrt{7}.
\]

\(\mathbb{Q}(\exp(i\pi/7))\) has dim 6 over \(\mathbb{Q}\).

Coordinates of \(x^2\)

**Definition**

A number \(x \in \mathbb{C}\) is *algebraic* when its powers generate a finite-dimensional vector space over \(\mathbb{Q}\).

**Tools:** Euclidean division, (extended) Euclidean algorithm, linear algebra.
Starting point:

\[
\frac{\sin(2a)}{\sin(a)} \frac{\sin(3a)}{\sin(a)} \frac{2 - \sin(a)}{\sin(a)} \frac{2 + \sin(3a)}{\sin(a)} \frac{2}{\sin(2a)^2} - \frac{\sin(a)}{\sin(a)} + \frac{\sin(3a)}{\sin(a)}
\]

Convert into rational function in \( T = \exp \left( \frac{i \cdot \pi}{7} \right) \):

\[
= \frac{21e^{ia}^2}{\left( e^{ia} - 1 \right)^2} - \frac{21}{\left( e^{ia} - 1 \right)^2} + \frac{21}{\left( e^{ia} - 1 \right)^2} + \frac{21}{\left( e^{ia} - 1 \right)^2}
\]

\[
= \frac{21(T^{16} + 5T^{14} + 12T^{12} + T^{11} + 20T^{10} + 3T^9 + 23T^8 + 3T^7 + 20T^6 + T^5 + 12T^4 + 5T^2 + 1)}{T(T^2 - 1)(T^2 + 1)^2(T^4 + T^2 + 1)^2}
\]

Simplify denominator by Bezout

\[
G = \gcdex(denom(newval), polmin, T, 'U', 'V'): U, V, G;
\]

\[
10 \frac{7}{7} T - \frac{5}{7} T^2 - \frac{5}{7} T^3 + \frac{1}{7} T^4 + \frac{10}{7} T^5,
\]

\[
1 + \frac{17}{7} T + \frac{24}{7} T^5 - \frac{25}{7} T^9 + \frac{26}{7} T^3 + \frac{11}{7} T^2 + \frac{25}{7} T^6 + \frac{23}{7} T^4 - \frac{2}{7} T^8 - \frac{26}{7} T^{12} - \frac{24}{7} T^{10} - \frac{11}{7} T^{13} - \frac{23}{7} T^{11} - \frac{5}{7} T^7 - \frac{10}{7} T^{14}, 1
\]

\[
= \text{rem}(U \ast \text{numer}(newval), polmin, T);
\]

Compute square

\[
= \text{rem}(\%^2, polmin, T);
\]
Automatic Univariate Identities

- Cassini
  \[ F_n F_{n+2} - F_{n+1}^2 = (-1)^n \]

- Catalan numbers
  \[ \frac{1 - \sqrt{1 - 4z}}{2z} = \sum_{n=0}^{\infty} \binom{2n}{n} \frac{z^n}{n+1} \]

- Legendre
  \[ {}_2F_1 \left( \begin{array}{c} a, b \\ a + b + 1/2 \end{array} \middle| z \right) = {}_2F_1 \left( \begin{array}{c} 2a, 2b \\ a + b + 1/2 \end{array} \middle| \frac{1 - \sqrt{1 - z}}{2} \right) \]

- Mehler
  \[ \sum_{n=0}^{\infty} H_n(x) H_n(y) \frac{u^n}{n!} = \exp \left( \frac{4u(xy - u(x^2 + y^2))}{1 - 4u^2} \right) \frac{1}{\sqrt{1 - 4u^2}} \]
Automatic Multivariate Identities

\[
\sum_{k=0}^{n} \binom{n}{k}^2 \binom{n+k}{k}^2 = \sum_{k=0}^{n} \binom{n}{k} \binom{n+k}{k} \sum_{j=0}^{k} \binom{k}{j}^3,
\]

\[
\sum_{n=0}^{+\infty} P_n(x) y^n = \frac{1}{\sqrt{1 - 2xy + y^2}}, \quad P_n(x) = \frac{1}{2^n} \sum_{k=0}^{n} \binom{n}{k}^2 x^k
\]

\[
\int_{0}^{+\infty} x J_1(ax) I_1(ax) Y_0(x) K_0(x) \, dx = -\frac{\ln(1 - a^4)}{2\pi a^2},
\]

\[
\int_{0}^{\infty} \frac{(1 + 2xy + 4y^2) \exp \left( \frac{4x^2y^2}{1+4y^2} \right)}{y^{n+1} (1 + 4y^2)^{3/2}} \, dy = \frac{n! H_n(x)}{\lfloor n/2 \rfloor!},
\]

\[
\sum_{k=0}^{n} \frac{q^{k^2}}{(q; q)_k (q; q)_{n-k}} = \sum_{k=-n}^{n} \frac{(-1)^k q^{(5k^2-k)/2}}{(q; q)_{n-k} (q; q)_{n+k}},
\]

\[
\sum_{j=0}^{n} \sum_{i=0}^{n-j} \frac{q^{(i+j)^2+j^2}}{(q; q)_{n-i-j} (q; q)_{i} (q; q)_{j}} = \sum_{k=-n}^{n} \frac{(-1)^k q^{7/2k^2+1/2k}}{(q; q)_{n+k} (q; q)_{n-k}}.
\]
Plan

1. Definitions and closure properties (BS)
2. Fast algorithms for functions (AB)
3. Algorithms for sequences (BS)
4. Multivariate case (FC)
http://algo.inria.fr/esf
II Definitions
## D-finite Series

**Definition**

A series \( f(x) \in \mathbb{Q}[[x]] \) is **differentially finite** (D-finite) when its derivatives generate a finite-dimensional vector space over \( \mathbb{Q}(x) \).

(LDE)

**Equivalent definition** A power series \( y(z) \) is **D-finite** if there exist polynomials \( a_i \in \mathbb{Q}[z] \) such that

\[
a_k(z)y^{(k)}(z) + \cdots + a_0(z)y(z) = 0.
\]

**Examples of functions:** exp, log, sin, cos, sinh, cosh, arccos, arccosh, arcsin, arccsinh, arctan, arctanh, arccot, arccoth, arccsc, arccsch, arcsec, arcsech, \( pF_q \) (includes Bessel \( J, Y, I \) and \( K \), Airy \( Ai \) and \( Bi \) and polylogarithms), Struve, Weber and Anger fcns, the large class of **algebraic functions**, ...  

About 60% of Abramowitz & Stegun.
Coefﬁcients ↔ Series

Theorem

A series is D-ﬁnite if and only if its sequence of coefﬁcients satisﬁes a linear recurrence.

Proof (Dictionary)

\[ y = \sum_n a_n z^n \]
\[ y' = \sum_n n a_n z^{n-1} = \sum_n (n + 1) a_{n+1} z^n \]
\[ z \cdot y = \sum_n a_{n-1} z^n \]
\[ z^i y^{(j)} = z^i D_z^j \cdot y = \sum_n (n - i + 1)(n - i + 2) \cdots (n + j - i) a_{n+j-i} z^n \]
\[ (zD_z)^k z^{-i} \cdot y = \sum_n a_{n+i} n^k z^n . \]

Warning: orders differ
D-finite Sequences

Definition

A sequence $u_n$ is **D-finite** when its shifts $(u_n, u_{n+1}, \ldots)$ generate a finite-dimensional vector space over $\mathbb{Q}(n)$. (LRE)

Examples of sequences: rational sequences, hypergeometric sequences (includes $n!$, multinomials,\ldots), classical orthogonal polynomials.

About 25% of Sloane & Plouffe.

eqn+ini. cond. = data structure
Example: Generalized Hypergeometric Series

\[ y(z) := \left. \begin{pmatrix} a_1, \ldots, a_p \\ b_1, \ldots, b_q \end{pmatrix} \right| _{(a_1)_n \cdots (a_p)_n \cdots (b_1)_n \cdots (b_q)_n n!} z^n, \]

\[ (x)_n := x(x + 1) \cdots (x + n - 1). \]

\[ \iff \text{ first order linear recurrences (hypergeometric sequences)} \]

\[ \frac{u_n}{u_{n-1}} = \frac{(n + a_1 - 1) \cdots (n + a_p - 1)}{n(n + b_1 - 1) \cdots (n + b_q - 1)} \implies ((\theta + a_1 - 1) \cdots (\theta + a_p - 1)z - \theta(\theta + b_1 - 1) \cdots (\theta + b_q - 1)) y(z) = 0, \]

\[ (\theta = zD_z). \]
Special Cases

\[ \exp(z) = 0F_0\left( - \left| \frac{z}{2!} \right. \right) = 1 + z + \frac{z^2}{2!} + \cdots , \]

\[ (1 - z)^{-a} = 1F_0\left( -a \left| \frac{z}{2} \right. \right) = 1 + az + a(a + 1)\frac{z^2}{2!} + \cdots , \]

\[ \log \frac{1}{1 - z} = z \, 2F_1\left( \frac{1, 1}{2} \left| \frac{z}{2} \right. \right) = z + \frac{z^2}{2} + \cdots , \]

\[ \arcsin(z) = z \, 2F_1\left( \frac{1}{2}, \frac{1}{2} \left| \frac{z}{3} \right. \right), \quad \arctan(z) = z \, 2F_1\left( \frac{1}{2}, \frac{1}{3} \left| \frac{z}{2} \right. \right) , \]

\[ J_\nu(z) = \frac{1}{\Gamma(\nu + 1)} \left( \frac{z}{2} \right) ^\nu 0F_1\left( - \left| \frac{z^2}{4} \right. \right) , \]

\[ \text{AGM}(a, b) = \frac{a}{2F_1\left( \frac{1}{2}, \frac{1}{2} \left| 1 - \frac{b^2}{a^2} \right. \right)} , \cdots \]
III Closure Properties
Proof of Identities: $\sin^2 + \cos^2 = 1$

> series(sin(x)^2+cos(x)^2,x,4);

$$1 + O(x^4)$$

Why is this a proof?

1. $\sin$ and $\cos$ satisfy a 2nd order LDE: $y'' + y = 0$;
2. their squares (and their sum) satisfy a 3rd order LDE (basis $\sin^2, \cos^2, \sin \cos$);
3. the constant 1 satisfies a 1st order LDE: $y' = 0$;
4. $\rightarrow \sin^2 + \cos^2 - 1$ satisfies a LDE of order at most 4;
5. it is not singular at 0;
6. Cauchy’s theorem concludes.

Another identity (same idea): $F_n^2 - F_{n+1}F_{n-1} = (-1)^{n+1}$

> for n to 5 do
  fibonacci(n)^2-fibonacci(n+1)*fibonacci(n-1)+(-1)^n od;
Euclidean Division & Finite Dimension

**Theorem (XIXth century)**

*D-finite series and sequences form \( \mathbb{Q} \)-algebras.*

**Proof.** (for product of series)

\[
f^{(n)} = a_0 f + a_1 f' + \cdots + a_{n-1} f^{(n-1)}, \quad g^{(m)} = b_0 g + b_1 g' + \cdots + b_{m-1} g^{(m-1)}.\]

\[
h = fg,\]

\[
h' = f'g + fg'.\]

\[
\vdots\]

\[
h^{(k)} = \sum_{0 \leq i < n} \sum_{0 \leq j < m} c_{ijk} f^{(i)}g^{(j)}\]

\[\Rightarrow h, h', \ldots, h^{(mn)} \text{ linearly dependent.}\]

This proof is an algorithm.

Everything implemented in gfun [SaZi94]
Euclidean Division & Finite Dimension

**Theorem (XIXth century)**

*D*-finite series and sequences form \( \mathbb{Q} \)-algebras.

**Corollary**

*D*-finite series are closed under Hadamard (termwise) product, Laplace transform, Borel transform (ogf↔egf).

Everything implemented in *gfun* [SaZi94]


**Mehler’s Identity for Hermite Polynomials**

\[ \sum_{n=0}^{\infty} H_n(x) H_n(y) \frac{u^n}{n!} = \exp \left( \frac{4u(xy - u(x^2 + y^2))}{1 - 4u^2} \right) \sqrt{1 - 4u^2} \]

1. Definition of Hermite polynomials (D-finite over \( \mathbb{Q}(x) \)): recurrence of order 2

2. Product by linear algebra: \( H_{n+k}(x)H_{n+k}(y)/(n+k)! \), \( k \in \mathbb{N} \) generated over \( \mathbb{Q}(x, n) \) by

\[ \frac{H_n(x)H_n(y)}{n!} \cdot \frac{H_{n+1}(x)H_n(y)}{n!} \cdot \frac{H_n(x)H_{n+1}(y)}{n!} \cdot \frac{H_{n+1}(x)H_{n+1}(y)}{n!} \]

→ recurrence of order at most 4;

3. Translation into a differential equation
I. Definition

\[ R_1 := \{ H(n + 2) = -2n - 2H(n) + 2H(n + 1)x, H(0) = 1, H(1) = 2x \} : \]

\[ R_2 := \text{subs}(H = H_2, x = y, R_1); \]

\[ R_2 := \{ H_2(0) = 1, H_2(n + 2) = -2n - 2H_2(n) + 2H_2(n + 1)y, H_2(1) = 2y \} \]

II. Product

\[ R_3 := \text{gfun:-poltorec}(H(n) \cdot H_2(n) \cdot v(n), [R_1, R_2, \{v(n + 1) \cdot (n + 1) = v(n), v(1) = 1\}], [H(n), H_2(n), v(n)], c(n)); \]

\[ R_3 := \left\{ c(0) = 1, c(1) = 4xy, c(2) = 8x^2y^2 + 2 - 4y^2 - 4x^2, c(3) = \frac{32}{3}x^3y^3 + 24xy - 16xy^3 - 16x^3y, (16n + 16)c(n) - 16xy\cdot c(n + 1) + (-8n - 20 + 8y^2 + 8x^2)c(n + 2) - 4xc(n + 3)y + (n + 4)c(n + 4) \right\} \]

III. Differential Equation

\[ \text{gfun:-rectodiffeq}(R_3, c(n), f(u)); \]

\[ \left\{ (16u^3 - 16u^2yx - 4u + 8uy^2 + 8ux^2 - 4xy)f(u) + (16u^4 - 8u^2 + 1)\left(\frac{d}{du}f(u)\right), f(0) = 1 \right\} \]

\[ dsolve(%, f(u)); \]

\[ f(u) = \frac{1e^\left(\frac{-4xyu + y^2 + x^2}{(2u - 1)(2u + 1)}\right)}{e^\left((-y^2 - x^2)^2\right)\sqrt{2u + 1}\sqrt{2u - 1}} \]
Theorem (Cockle1860)

*D-finite series composed with algebraic power series are D-finite.*

Proof.

\[ P(x, y) = 0 \text{ and } AP + BP_y = 1 \Rightarrow y' = -\frac{P_x}{P_y} = -BP_x \mod P \]

\[ \Rightarrow y^{(k)} \in \bigoplus_{i < \deg_y P} \mathbb{Q}(x)y^i. \]

\((f \circ y)^{(p)}\) linear combination of \((f^j \circ y)y^k\).

Also, \(\exp \int y\).
Motzkin Numbers (Unary-Binary Trees)

\[ M(z) - (1 + zM(z) + zM(z)^2) =: P(M) = 0. \]

Bézout: \( AP + BP_M = 1 \Rightarrow M' = -BP_z \mod P =: cM + d1. \)

Vector space of dimension 2

\[
\text{gfun[algeqtodiffeq]}(M = 1 + zM + zM^2, M(z)) ; \\
-1 - z + (-3z + 1) M(z) + (z - 6z^2 + z^3) M'(z)
\]

\[
\text{gfun[diffeqtorec]}(\%, M(z), u(n)) ; \\
\{ nu(n) + (-9 - 6n)u(1 + n) + (3 + n)u(n + 2), u(0) = 1, u(1) = 2 \}
\]

→ fast computation.

Works for arbitrary degree.
Forests of Catalan Trees

\[ Y(z) = \exp \left( \frac{1 - \sqrt{1 - 4z}}{2} \right). \]

Same computation →

\[ \{(n + 2)(n + 1)u(n + 2) = u(n) + 2(n + 1)(2n + 1)u(n + 1), \]
\[ u(0) = 1, u(1) = 1\} \]
Example: Airy Ai at Infinity

\[ \text{Ai}(z) = \frac{\sqrt{z} e^{-\xi}}{2\pi} \int_{-\infty}^{\infty} e^{-\xi[(u-1)(4u^2+4u+1)]} \, dv, \quad \xi = \frac{2}{3}z^{3/2}, \quad u = \sqrt{1 + \frac{v^2}{3}} \]

\[ \sim \frac{1}{2} \pi^{-1/2} z^{-1/4} e^{-\xi} \sum_{n=0}^{\infty} (-1)^n \xi^{-n} \frac{\Gamma(3n + \frac{1}{2})}{54^n n! \Gamma(n + \frac{1}{2})}. \]

Computation:

1. **algebraic** change of variables \( t^2 = (u - 1)(4u^2 + 4u + 1); \)

\[ \rightarrow \int_{-\infty}^{\infty} e^{-\xi t^2} f(t) \, dt, \quad f(t) = \frac{dv}{dt}, \]

2. recurrence satisfied by the coefficients of \( f \) (generating series);

3. termwise integration (Hadamard product).
I. Algebraic change of variables

\[ eq_u := u^2 - \left( 1 + \frac{v^2}{3} \right) ; \]
\[ eq_t := t^2 - (u - 1) \cdot \left( 4 \cdot u^2 + 4 \cdot u + 1 \right) ; \]
\[ res := \text{resultant}(eq_u, eq_t, u) ; \]
\[ t^4 + 2 \cdot t^2 - 3 \cdot v^2 - \frac{8}{3} \cdot v^4 - \frac{16}{27} \cdot v^6 \]
\[ \text{gfun} := \text{algeqtodiffeq} \left( \text{res}, v(t), \left\{ v(0) = 0, D(v)(0) = \sqrt{\frac{2}{3}} \right\} \right) ; \]
\[ \left\{ -4 \cdot v(t) + 9 \cdot t \left( \frac{d}{dt} v(t) \right) + \left( 9 \cdot t^2 + 18 \right) \left( \frac{d^2}{dt^2} v(t) \right), v(0) = 0, (D(v))(0) = \frac{1}{3} \cdot \sqrt{6} \right\} \]

II. Recurrence satisfied by the coefficients of f

\[ \text{gfun} := \text{poltodiffeq}(\text{diff}(v(t), t), [% ], [v(t)], f(t)) ; \]
\[ \left\{ 5 \cdot f(t) + 27 \cdot t \left( \frac{d}{dt} f(t) \right) + \left( 9 \cdot t^2 + 18 \right) \left( \frac{d^2}{dt^2} f(t) \right), f(0) = \frac{1}{3} \cdot \sqrt{6}, (D(f))(0) = 0 \right\} \]
\[ R_f := \text{gfun} := \text{diffeqtorec}(%, f(t), c(n)) ; \]
\[ \left\{ \left( 5 + 18 \cdot n + 9 \cdot n^2 \right) c(n) + \left( 18 \cdot n^2 + 54 \cdot n + 36 \right) c(n + 2), c(0) = \frac{1}{3} \cdot \sqrt{6}, c(1) = 0 \right\} \]


### III. Hadamard product

Assume \( \xi > 0 \); \( s := \text{Int}(\exp(-\xi \cdot t^2) \cdot t^n, t = -\infty .. \infty); s = \text{student}[\text{intpart}](s, \exp(-\xi \cdot t^2))\)

\[
\int_{-\infty}^{\infty} e^{-\xi \cdot t^2} t^n \, dt = -\int_{-\infty}^{\infty} -2 \frac{\xi \cdot t \cdot e^{-\xi \cdot t^2}}{n + 1} \cdot t^{(n + 1)} \, dt
\]

\[
R_i := \begin{cases} 
  c(n) = \frac{2 \cdot \xi}{(n + 1)} \cdot c(n + 2), c(0) = \text{value} \left( \text{eval}(s, n = 0) \right), c(1) = \text{value} \left( \text{eval}(s, n = 1) \right) \\
  \left\{ 
  c(n) = \frac{2 \xi \cdot c(n + 2)}{n + 1}, c(0) = \frac{\sqrt{\pi}}{\sqrt{\xi}}, c(1) = 0 \right\}
\end{cases}
\]

> FinalRec := gfun:-`rec*rec`\((R[i], R[f], c(n))\);

\[
\text{FinalRec} := \begin{cases} 
  \left\{ 5 + 18 n + 9 n^2 \right\} c(n) + (36 \xi \cdot n + 72 \xi \cdot) c(n + 2), c(1) = 0, c(0) = \frac{1}{3} \frac{\sqrt{\pi} \sqrt{6}}{\sqrt{\xi}} \\
  \end{cases}
\]

\[
\text{Sol} := \text{rsolve}(\text{FinalRec}, c(n));
\]

\[
\left\{ 
\begin{array}{l}
  \frac{1}{3} \left( \frac{1}{2} n \right)^2 \frac{1}{2} \left( -1 - \frac{1}{2} n \right) \Gamma \left( \frac{1}{2} n + \frac{5}{6} \right) \Gamma \left( \frac{1}{2} n + \frac{1}{6} \right) \xi \left( -\frac{1}{2} n \right) \sqrt{6} \\
  \sqrt{\pi} \Gamma \left( \frac{1}{2} n + 1 \right) \sqrt{\xi}
\end{array}ight. \\
\begin{array}{l}
  n : \text{even} \\
  0
\end{array}
\]

\[
\begin{array}{l}
  n : \text{odd}
\end{array}
\]
D-finite Series in Arithmetic

**Definition (For \( k \in \mathbb{Z} \), modular form of weight \( k \))**

\( f \) defined on \( \mathbb{H} z > 0 \) such that 
\[ f((az + b)/(cz + d)) = (cz + d)^k f(z), \]
for all matrices \( (a \ b) \\
(\ c \ d) \) in 
\( SL(2, \mathbb{Z}) \) or one of its subgroups of finite index.

**Modular function**: modular form of weight 0.

**Theorem (XIXth century)**

Let \( f(z) \) be a meromorphic modular form of weight \( k > 0 \) and \( t(z) \) a modular function. Then \( F(t) \) defined by \( F(t(z)) = f(z) \) satisfies 
a LDE of order \( k + 1 \) with algebraic coefficients.

**Ex.** [Apéry]
\[ t(z) = \left( \frac{\eta(z)\eta(6z)}{\eta(2z)\eta(3z)} \right)^{12}, \]
\[ f(z) = \frac{(\eta(2z)\eta(3z))^7}{(\eta(z)\eta(6z))^5}, \]
\[ F(t) = 1 + 5t + 73t^2 + \cdots = \sum_{n \geq 0} \sum_{k} \binom{n}{k}^2 \binom{n + k}{k}^2 t^n. \]
LDEs and LREs can be viewed as a data-structure;

algorithms for $+$, $\times$, translation $\text{LDE} \leftrightarrow \text{LRE}$;

several sources of D-finiteness (algebraic series, hypergeometric series, series of arithmetical origin);

applications to proofs of identities, to computations of coefficients of expansions.

Next episodes: efficient algorithms, more applications, generalization to the multivariate case.