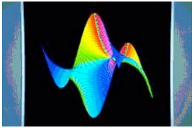
Complexity of the Creative Telescoping for Bivariate Rational Functions

Shaoshi Chen
(joint work with Alin Bostan, Frédéric Chyzak & Ziming Li)

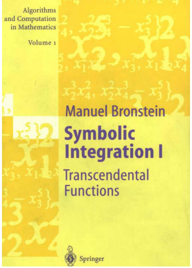
Algorithms Project-Team, INRIA (France)
KLMM, Chinese Academy of Sciences (China)



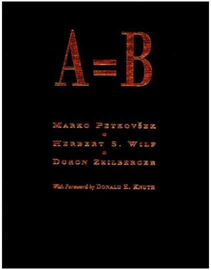
February 24, 2010

Symbolic integration

From differential algebra to creative telescoping



→



Outline

Definite integration for special functions

Definite Integration:

$$F(x) = \int_a^b f(x, y) dy$$

$\xrightarrow{\text{by}}$

Creative telescoping (CT):

$$L(x, D_x)(f) = D_y(g)$$

L: telescoper g: certificate

$$\lim_{y \to a} g(x, y) = \lim_{y \to b} g(x, y) \quad \Longrightarrow \quad L(x, D_x)(F) = 0.$$

Example: An integral of a product of four Bessel functions [GlaMon1994]

$$\int_0^{+\infty} y J_1(xy) I_1(xy) Y_0(y) K_0(y) \, dy = -\frac{\ln(1 - x^4)}{2\pi x^2}$$

$L = x^3(x^4 - 1)D_x^4 + \cdots$ and $g =$ poly. in Bessel functions

Previous works: General special functions

$$\begin{cases} P(x, y, D_x)(f) = 0 \\ Q(x, y, D_y)(f) = 0 \end{cases} \quad + \quad \text{Ini. cond.} = D\text{-finite data structure}$$

Existence: If f is D-finite, then there exists (L, g) s.t. $L(f) = D_y(g)$.

- Holonomic D-modules: Bernstein (1971)
- Closure property of diagonal operation: Lipshitz (1988)

Algorithms and implementations:

- Slow algo. for general holonomic inputs: Zeilberger (1990)
- Fast algo. for hyperexponential inputs: AlmkvistZeilberger (1990)
- Gröbner-bases approach: Takayama (1992)
- Fast algo. for general holonomic inputs: Chyzak (1997)
- Non-holonomic generalization: Chyzak-Kauers-Salvy (2009)
- Mgfun (Chyzak1997, Pech2010), HolonomicFunctions (Koutschan2009)

Motivation for our work

1. No complexity analysis for CT algorithms yet
2. Algorithms for special functions are often slow in practice
3. Interesting applications of rational-function telescoping

   3.1 Differential equations for algebraic functions [BCLSS2007]

$$P(x,\alpha) = 0 \to L\left(\frac{yD_y(P)}{P}\right) = D_y(g) \Rightarrow L(\alpha) = 0$$

   3.2 Differential equations for diagonals [PemantleWilson2008]

$$L\left(\frac{f(y, x/y)}{y}\right) = D_y(g) \Rightarrow L(\mathrm{diag}(f)) = 0$$

Our work: Bivariate rational functions

Problem (CT for bivariate rational functions)

$f \in k(x, y)$, construct $(L, g) \in k(x)\langle D_x \rangle \setminus \{0\} \times k(x, y)$ such that

$$L(x, D_x)(f) = D_y(g) \qquad \textit{(Telescoping equation)}$$

Example: An integral of a bivariate rational function

$$F(x) := \int_0^{+\infty} \frac{dy}{x^2 + y^2 + 1} \rightarrow (L, g) = \left(x + (x^2 + 1)D_x, -\frac{xy}{x^2 + y^2 + 1} \right)$$

$$xF + (x^2 + 1)D_x(F) = 0 \text{ and } F(0) = \frac{\pi}{2} \rightarrow F(x) = \frac{\pi}{2\sqrt{x^2 + 1}}.$$

Focus: Compute a telescoper of minimal order (minimal telescoper).

Main results

Theorem (Complexity for rational-function telescoping)

CT for bivariate rational functions has polynomial complexity.

$$f = \frac{P}{Q} \in k(x, y) \to L(x, D_x)(f) = D_y(g)$$

d : The max. of total degrees of P and Q in x and y.
ω : Any feasible exponent of matrix multiplication ($2 \leq \omega \leq 3$).

	Method	$\deg_x(L)$	$\deg_{D_x}(L)$	$\deg_x(g)$	$\deg_y(g)$	Cost	Expon.
Minimal	Hermite	$\mathcal{O}(d^3)$	$\leq d$	$\mathcal{O}(d^3)$	$\mathcal{O}(d^2)$	$\tilde{\mathcal{O}}(d^{\omega+4})$	6.80
Telescoper	RatAZ	$\mathcal{O}(d^3)$	$\leq d$	$\mathcal{O}(d^3)$	$\mathcal{O}(d^2)$	$\tilde{\mathcal{O}}(d^{2\omega+3})$	8.61
Non-mini.	Lipshitz	$\mathcal{O}(d^2)$	$\mathcal{O}(d^2)$	$\mathcal{O}(d^3)$	$\mathcal{O}(d^3)$	$\mathcal{O}(d^{6\omega})$	16.8
Telescoper	Cubic	$\mathcal{O}(d^2)$	$\mathcal{O}(d)$	$\mathcal{O}(d^2)$	$\mathcal{O}(d^2)$	$\mathcal{O}(d^{4\omega})$	11.2

(Complexity is in terms of arithmetic operations in k)

Linear systems in different methods

Non-linear problem: $L(f) = D_y(g) \longrightarrow$ Linear problem: $\mathcal{M} \cdot x = 0$

	Method	System size	Coeff. deg.	Cost
Minimal	Hermite	$i \times d$	$\mathcal{O}(id^2)$	$\tilde{\mathcal{O}}(d^{\omega+4})$
Telescoper	RatAZ	$id \times id$	$\mathcal{O}(id)$	$\tilde{\mathcal{O}}(d^{2\omega+3})$
Non-mini.	Lipshitz	$\mathcal{O}(d^6) \times \mathcal{O}(d^6)$	0	$\mathcal{O}(d^{6\omega})$
Telescoper	Cubic	$\mathcal{O}(d^4) \times \mathcal{O}(d^4)$	0	$\mathcal{O}(d^{4\omega})$

(For mini. telescoper, costs take account of a loop over $i = 1, \ldots, d$)

Theorem (StorjohannVillard2005)

Given $M \in k[x]_{\leq d}^{m \times n}$, its rank and a basis of its null space can be computed using $\tilde{\mathcal{O}}(nmr^{\omega-2}d)$ ops. $\tilde{\mathcal{O}}(n^{\omega}d)$

Two approaches for constructing minimal telescopers

Aim: Given $f = P/Q \in k(x,y)$, find $L := \sum_{i=0}^{\rho} \eta_i(x) D_x^i \in k[x]\langle D_x \rangle \setminus \{0\}$ and $g \in k(x,y)$, s.t.

$$L(x, D_x)(f) = D_y(g) \quad \text{and } \deg_{D_x}(L) \text{ is minimal.}$$

Almkvist and Zeilberger's approach: $g = Rf$ for $R \in k(x,y)$

$$L - D_y(R) \equiv 0 \bmod \mathrm{Ann}(f) \rightsquigarrow \text{ ODE in } R \text{ parametrized by } \eta_i$$

Hermite reduction approach:

$$D_x^i(f) \equiv r_i \bmod D_y(k(x,y)) \rightsquigarrow \text{ Linear system in } \eta_i$$

Hermite reduction for indefinite integration

Additive decomposition: For $f \in K(y)$, decompose f into

$$f = D_y(g) + \frac{a}{b}, \quad \deg_y(a) < \deg_y(b) \text{ and } b \text{ square-free.}$$

1. Hermite (1872): Algorithm for computing $(g, a/b)$ by GCD only!

   Key idea: Reduction of pole order for A/Q^m with Q square-free

$$\boxed{\frac{A}{Q^m} = \frac{sQ + tD_y(Q)}{Q^m} = \frac{(1-m)s - D_y(t)}{(1-m)Q^{m-1}} + D_y\left(\frac{t}{(1-m)Q^{m-1}}\right)}$$

2. Ostrogradsky (1845) and Horowitz (1971): Algorithm by linear solver.
3. Yun (1977): Complexity analysis of Hermite reduction (quasi-linear).

Bivariate Hermite reduction (BHR)

Horowitz-Ostrogradsky's method: Given $f = P/Q \in k(x, y)$,

$$\frac{P}{Q} = D_y\left(\frac{A}{Q^-}\right) + \frac{a}{Q^*}, \quad Q^* := Q_1 \cdots Q_m \text{ and } Q^- := \frac{Q}{Q^*}.$$

$$\text{H.O. System:} \qquad \mathbf{P} = \mathcal{H}\begin{pmatrix}\mathbf{A}\\ \mathbf{a}\end{pmatrix}, \quad \mathcal{H} \in k[x]_{\leq d_x}^{d_y \times d_y}$$

Output size: $d_x := \max\{\deg_x P, \deg_x Q\}$, $d_y := \max\{\deg_y P, \deg_y Q\}$

$$\text{Cramer's rule} \rightarrow \deg_x(A), \deg_x(a) \in \mathcal{O}(d_x d_y)$$

Eval-Interp algorithm: BHR with quasi-optimal complexity $\tilde{\mathcal{O}}(d_x d_y^2)$

$$\text{BHR} = (\text{Eval. } f(x_0, y) + \text{UHR on } f(x_0, y)) \times \mathcal{O}(d_x d_y) + \text{Rat.interp.}$$

Hermite reduction for creative telescoping

Key idea: For $f = P/Q \in k(x, y)$, $d_y^* := \deg_y(Q^*)$,

$$D_x^i(f) \xrightarrow{HR_y} D_y(g_i) + \frac{a_i}{Q^*}, \quad a_i \in k(x)[y] \text{ and } \deg_y(a_i) < d_y^*.$$

Lemma: $a_0, a_1, \ldots, a_{d_y^*}$ are linearly dependent over $k(x)$. Furthermore,

$$\sum_{i=0}^{\rho} \eta_i(x) a_i = 0 \Longleftrightarrow \sum_{i=0}^{\rho} \eta_i(x) D_x^i(f) = D_y\left(\sum_{i=0}^{\rho} \eta_i g_i\right).$$

Theorem

1. *There exists a telescoper for f of order at most d_y^*.*
2. *If $\sum_{i=0}^{\rho} \eta_i a_i = 0$ for smallest $\rho \in \mathbb{N}$, then $\sum_{i=0}^{\rho} \eta_i D_x^i$ is a minimal telescoper for f with certificate $\sum_{i=0}^{\rho} \eta_i g_i$.*

Algorithm and complexity I

Algorithm (HermiteTelescoping)

1. *BHR*: $f = D_y(g_0) + a_0/Q^*$;
2. *For* i *from* 1 *to* d_y^* *do*
   2.1 *BHR*: $D_x^i(f) = D_y(g_i) + a_i/Q^*$;
   2.2 *If* $\sum_{j=0}^{i} \eta_j a_j = 0$ *for* $\eta_j \in k(x)$, *not all zero,*
   then return $(\sum_{j=0}^{i} \eta_j D_x^j, \sum_{j=0}^{i} \eta_j g_j)$.

Incremental strategy: $(g_i, a_i) \rightarrow (g_{i+1}, a_{i+1})$

$$D_x^i(f) = D_y(g_i) + \frac{a_i}{Q^*} \Rightarrow D_x^{i+1}(f) = D_y(D_x(g_i)) + \frac{D_x(a_i)}{Q^*} - \frac{a_i D_x(Q^*)}{{Q^*}^2}$$

$$-\frac{a_i D_x(Q^*)}{{Q^*}^2} = D_y(\tilde{g}_{i+1}) + \frac{\tilde{a}_{i+1}}{Q^*} \Rightarrow D_x^{i+1}(f) = D_y(D_x(g_i) + \tilde{g}_{i+1}) + \frac{D_x(a_i) + \tilde{a}_{i+1}}{Q^*}$$

Algorithm and complexity II

Theorem (Complexity for HermiteTelescoping)

For $f = P/Q \in k(x,y)$ of bidegree (d_x, d_y), HermiteTelescoping computes (L, g) in $\tilde{\mathcal{O}}(d_x d_y^{\omega+3})$ ops.

Degree bounds on g_i and a_i:

$$\deg_x(g_i), \deg_x(a_i) \in \mathcal{O}(i d_x d_y), \quad \deg_y(g_i) \le i d_y, \text{ and } \deg_y(a_i) \le d_y^*.$$

Cost estimate for Step $i \geq 1$: Hermite reduction $+$ linear system solving

2.1 Hermite reduction on $D_x^i(f)$: $\tilde{\mathcal{O}}(i^2 d_x d_y^2)$;

2.2 Finding linear dependence of a_i's: $\tilde{\mathcal{O}}(i^\omega d_x d_y^2)$.

$$\sum_{j=0}^{i} \eta_j a_j = 0 \rightsquigarrow \begin{pmatrix} \quad \deg_x \in \mathcal{O}(i d_x d_y) \quad \end{pmatrix}_{i \times d_y^*} \xleftarrow{SV} \tilde{\mathcal{O}}(i^\omega d_x d_y^2).$$

Differential Gosper algorithm

Problem: Given H with $D_y(H)/H \in K(y)$, determine T with

$$\frac{D_y(T)}{T} \in K(y), \quad \text{s.t.} \quad H = D_y(T).$$

Differential Gosper form: A triple $(p, q, r) \in K[y]^3$ for $f \in K(y)$, s.t.

$$f = \frac{D_y(p)}{p} + \frac{q}{r}, \quad \gcd(r, q - \tau D_y(r)) = 1, \quad \text{for all } \tau \in \mathbb{N}.$$

Algorithm (DiffGosper)

1. *Compute a differential Gosper form (p, q, r) of $D_y(H)/H$;*
2. *Determine whether $p = rD_y(z) + (q + D_y(r))z$ has a sol. in $K[y]$;*
3. *If there exists a poly. sol. $s \in K[y]$, then return $T := srH/p$.*

Almkvist and Zeilberger's approach for CT

Algorithm (AZ for hyperexponential functions)

For $i = 0, 1, \ldots$ *do*

1. *Solve* $\sum_{j=0}^{i} \eta_j D_x^j(f) = D_y(T)$ *by a variant of DiffGosper.*
2. *If there exist* $\eta_j \in k(x)$*, not all zero, and* $D_y(T)/T \in k(x, y)$*, then return* $(\sum_{j=0}^{i} \eta_j D_x^j, T)$*.*

AZ for rational functions (RatAZ): $f = P/Q \in k(x, y)$

1. Prediction of diff. Gosper form: $H := -D_y(Q)/Q^{-} - iD_y(Q^{*})$

$$F := \sum_{j=0}^{i} \eta_j D_x^j(f) = \frac{N}{QQ^{*i}} \Rightarrow \frac{D_y(F)}{F} = \frac{D_y(N)}{N} + \frac{H}{Q^{*}}.$$

2. Degree bound on poly. sols: $\deg_y(Q^{-}) + i \deg_y(Q^{*}) \in \mathcal{O}(id_y)$.

Complexity analysis of RatAZ

Theorem (Complexity for RatAZ)

For $f = P/Q \in k(x, y)$ of bidegree (d_x, d_y), RatAZ computes (L, g) in $\tilde{\mathcal{O}}(d_x d_y^{2\omega+2})$ ops.

Cost estimate for Step $i \geq 0$: $\tilde{\mathcal{O}}(i^{\omega+1} d_x d_y^{\omega})$

$$\sum_{j=0}^{i} \eta_j D_x^j(f) = D_y(g) \xrightarrow{diff.G-form} N = Q^* D_y(z) + (D_y(Q^*) + H)z$$

$$\left(\quad \deg_x \leq \mathcal{O}(i d_x) \quad \right)_{\mathcal{O}(i d_y) \times \mathcal{O}(i d_y)} \xleftarrow{SV} \tilde{\mathcal{O}}(i^{\omega+1} d_x d_y^{\omega})$$

Total cost: $\sum_{i=0}^{d_y} \tilde{\mathcal{O}}(i^{\omega+1} d_x d_y^{\omega}) = \tilde{\mathcal{O}}(d_x d_y^{2\omega+2})$.

From y-free annihilators to telescopers

$A(x, D_x, D_y)(f) = 0$ $A := D_y^m(L(x, D_x) - D_y R)$	$\rightsquigarrow$	$L(x, D_x)(f) = D_y(g)$ $g \in R(f) + k(x)[y]_{\leq m}$

Key idea: Truncating $\mathcal{W} := k[x]\langle D_x, D_y\rangle$ + Counting dimension

$$i + j + \ell \leq N \Rightarrow x^i D_x^j D_y^\ell(f) \in \mathrm{Vect}_k \left\{ \frac{x^m y^n}{Q^{N+1}} \mid m + n \leq N + (N+1)d \right\}$$

$$\phi : \mathcal{W}_N \longrightarrow \mathcal{W}_N(f), \quad \dim(\mathcal{W}_N) \in \Theta(N^3)\ , \quad \dim(\mathcal{W}_N(f)) \in \mathcal{O}(N^2).$$

Theorem (Existence and degree bounds of A)

Given $f = P/Q \in k(x, y)$, there exists $A(x, D_x, D_y) \neq 0$ s.t. $A(f) = 0$.

- *Total degree filtration:* $\deg(A) \leq 3(d+1)^2$;
- *Order splitting filtration:* $\deg_x(A) \leq 4d^2$ *and* $\deg_D(A) \leq 4d$.

Bounds by counting dimensions

Total degree filtration: Used in [Lipshitz1988] for diagonals

$$\mathcal{W}_N := \text{span}_k\{x^i D_x^j D_y^\ell \mid i+j+\ell \leq N\} \qquad \binom{N+3}{3}$$

$$\mathcal{W}_N(f) \subset \text{span}_k\left\{\frac{x^m y^n}{Q^{N+1}} \mid m+n \leq N+(N+1)d\right\} \qquad \binom{N+(N+1)d+2}{2}$$

Taking $N = 3(d+1)^2$, $\dim(\mathcal{W}_N) > \dim(\mathcal{W}_N(f))$, then ϕ is not injective.

Order splitting filtration: Used in [BCLSS2007] for algebraic functions

$$\mathcal{W}_{N_x,N_D} := \text{span}_k\{x^i D_x^j D_y^\ell \mid i \leq N_x, j+\ell \leq N_D\} \qquad (N_x+1)\binom{N_D+2}{2}$$

$$\mathcal{W}_{N_x,N_D}(f) \subset \text{span}_k\left\{\frac{x^m y^n}{Q^{N+1}} \mid m+n \leq N_x+(N_D+1)d\right\} \qquad \binom{N_x+(N_D+1)d+2}{2}$$

Taking $N_x = 4d^2$ and $N_D = 4d$, $\dim(\mathcal{W}_{N_x,N_D}) > \dim(\mathcal{W}_{N_x,N_D}(f))$, then ϕ is not injective. So there exists a telescoper $L(x, D_x)$ of cubic size.

Implementation and experiments

Random examples:

$$f = \frac{P}{Q_1 \cdots Q_5^5}, \qquad \sum_{i=1}^{5} i \deg_x(Q_i) = \sum_{i=1}^{5} i \deg_y(Q_i) = 5$$

P and Q_i are generated by randpoly() in Maple. There are 49 cases.

nb	AZ	Abr	RAZ	H1	H2	HO	EI	MG
1	64	80	48	36	60	44	672	624
2	72	96	40	36	48	56	1116	936
8	56	76	32	20	24	28	604	496
9	348	272	140	84	188	128	3876	2976
29	44	72	32	28	36	20	608	528
43	52	76	36	20	24	32	652	584
49	474269	34694	20977	10336	36254	22417	∞	652968

(Timing in ms.)

Application to diagonals

Definition. $\operatorname{diag}(f) := \sum_{i=0}^{\infty} f_{i,i} x^i$ for $f = \sum_{i,j\geq 0} f_{i,j} x^i y^j \in k[[x,y]]$.

Diagonal computation via CT: $F := f(y, x/y)/y$,

$$L(x, D_x)(F) = D_y(G) \Rightarrow L(\operatorname{diag}(f)) = 0.$$

Example: [FlaHaSo2004].

$$f = \frac{1}{1 - x - y - xy(1 - x^d)}, \quad d \in \mathbb{N}.$$

d	AZ	Abr	RAZ	H1	H2	HO	RR	GHP
4	176	136	100	116	208	108	220	956
8	3032	4244	4380	1976	5344	4396	10336	154409
10	11740	12816	7108	7448	24565	7076	46882	1118313

(Timing in ms.)

Conclusion

Summary:

1. First complexity analysis of CT algorithms (Rational case);
2. New and faster algorithm (Hermite reduction approach)
   - separates the computation of L and that of g;
   - good at some applications;
3. Non-minimal = smaller sizes.

Future:

1. Complexity taking into account
   - the multiplicity of denominators;
   - the structure of matrices;
2. Hyperexponential case;
3. Multivariate rational case.

Thanks!