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

Creative Telescoping for Bivariate Rational Functions
Outline

Introduction

Minimal telescopes
  Hermite reduction approach
  Almkvist and Zeilberger’s approach

Non-minimal telescopes

Implementation and Application

Conclusion
Introduction

Minimal telescopers
   Hermite reduction approach
   Almkvist and Zeilberger’s approach

Non-minimal telescopers

Implementation and Application

Conclusion
Definite integration for special functions

\[ F(x) = \int_{a}^{b} f(x, y) \, dy \]

Creative telescoping (CT):

\[ L(x, D_x)(f) = D_y(g) \]

\( L \): telescoper \hspace{1cm} g: \text{certificate} \]

\[ \lim_{y \to a} g(x, y) = \lim_{y \to b} g(x, y) \implies L(x, D_x)(F) = 0. \]

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

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

\[ L = x^3(x^4 - 1) D_x^4 + \cdots \quad \text{and} \quad g = \text{poly. in Bessel functions} \]
Previous works: General special functions

\[
\begin{aligned}
P(x, y, D_x)(f) &= 0 \\
Q(x, y, D_y)(f) &= 0 \\
\end{aligned}
\]

+ Ini. cond. = \textit{D}-finite data structure

**Existence:** If \( f \) is \( \textit{D} \)-finite, then there exists \( (L, g) \) s.t. \( L(f) = D_y(g) \).

- Holonomic \( \textit{D} \)-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]

   \[ P(x, \alpha) = 0 \rightarrow L \left( \frac{yD_y(P)}{P} \right) = D_y(g) \Rightarrow L(\alpha) = 0 \]

   3.2 Differential equations for diagonals [PemantleWilson2008]

   \[ L \left( \frac{f(y, x/y)}{y} \right) = D_y(g) \Rightarrow L(\text{diag}(f)) = 0 \]
Our work: Bivariate rational functions

Problem (CT for bivariate rational functions)

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

\[ L(x, D_x)(f) = D_y(g) \quad \text{(Telescoping equation)} \]

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.

\[ f = \frac{P}{Q} \in k(x, y) \rightarrow L(x, D_x)(f) = D_y(g) \]

\( d \): The max. of total degrees of \( P \) and \( Q \) in \( x \) and \( y \).

\( \omega \): Any feasible exponent of matrix multiplication \((2 \leq \omega \leq 3)\).

<table>
<thead>
<tr>
<th>Method</th>
<th>deg(_x)(L)</th>
<th>deg(_{Dx})(L)</th>
<th>deg(_x)(g)</th>
<th>deg(_y)(g)</th>
<th>Cost</th>
<th>Expon.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimal</td>
<td>Hermite</td>
<td>( \mathcal{O}(d^3) )</td>
<td>( \leq d )</td>
<td>( \mathcal{O}(d^3) )</td>
<td>( \mathcal{O}(d^2) )</td>
<td>( \tilde{O}(d^{\omega+4}) )</td>
</tr>
<tr>
<td>Telescoper</td>
<td>RatAZ</td>
<td>( \mathcal{O}(d^3) )</td>
<td>( \leq d )</td>
<td>( \mathcal{O}(d^3) )</td>
<td>( \mathcal{O}(d^2) )</td>
<td>( \tilde{O}(d^{2\omega+3}) )</td>
</tr>
<tr>
<td>Non-mini.</td>
<td>Lipshitz</td>
<td>( \mathcal{O}(d^2) )</td>
<td>( \mathcal{O}(d^2) )</td>
<td>( \mathcal{O}(d^3) )</td>
<td>( \mathcal{O}(d^3) )</td>
<td>( \mathcal{O}(d^{6\omega}) )</td>
</tr>
<tr>
<td>Telescoper</td>
<td>Cubic</td>
<td>( \mathcal{O}(d^2) )</td>
<td>( \mathcal{O}(d) )</td>
<td>( \mathcal{O}(d^2) )</td>
<td>( \mathcal{O}(d^2) )</td>
<td>( \mathcal{O}(d^{4\omega}) )</td>
</tr>
</tbody>
</table>

(Complexity is in terms of arithmetic operations in \( k \))
Linear systems in different methods

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

<table>
<thead>
<tr>
<th>Method</th>
<th>System size</th>
<th>Coeff. deg.</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimal</td>
<td>Hermite</td>
<td>$i \times d$</td>
<td>$\mathcal{O}(id^2)$</td>
</tr>
<tr>
<td>Telescoper</td>
<td>RatAZ</td>
<td>$id \times id$</td>
<td>$\mathcal{O}(id)$</td>
</tr>
<tr>
<td>Non-mini.</td>
<td>Lipshitz</td>
<td>$\mathcal{O}(d^6) \times \mathcal{O}(d^6)$</td>
<td>$0$</td>
</tr>
<tr>
<td>Telescoper</td>
<td>Cubic</td>
<td>$\mathcal{O}(d^4) \times \mathcal{O}(d^4)$</td>
<td>$0$</td>
</tr>
</tbody>
</table>

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

**Theorem (Storjohann-Villard2005)**

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

Minimal telescopers

   Hermite reduction approach
   Almkvist and Zeilberger’s approach

Non-minimal telescopers

Implementation and Application

Conclusion
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.

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

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

\[
L - D_y(R) \equiv 0 \mod \text{Ann}(f) \leadsto \text{ODE in } R \text{ parametrized by } \eta_i
\]

Hermite reduction approach:

\[
D_x^i(f) \equiv r_i \mod D_y(k(x, y)) \leadsto \text{Linear system in } \eta_i
\]
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

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

2. Ostrogradsky (1845) and Horowitz (1971): Algorithm by linear solver.
Bivariate Hermite reduction (BHR)

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

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

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

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

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 $\tilde{O}(d_x d_y^2)$

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

Key idea: For $f = P/Q \in k(x, y)$, $d^*_y := \text{deg}_y(Q^*)$,

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

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

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

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**: \( D^i_x(f) = D_y(g_i) + a_i/Q^*;\)
   2.2 If \( \sum_{j=0}^{i} \eta_j a_j = 0 \) for \( \eta_j \in k(x) \), not all zero,
      then return \( (\sum_{j=0}^{i} \eta_j D^j_x, \sum_{j=0}^{i} \eta_j g_j) \).

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

\[
D^i_x(f) = D_y(g_i) + \frac{a_i}{Q^*} \Rightarrow D^{i+1}_x(f) = D_y(D_x(g_i)) + \frac{D_x(a_i)}{Q^*} - \frac{a_i D_x(Q^*)}{Q^*^2} - \frac{a_i D_x(Q^*)}{Q^*^2} = D_y(\tilde{g}_{i+1}) + \frac{\tilde{a}_{i+1}}{Q^*} \Rightarrow D^{i+1}_x(f) = D_y(D_x(g_i) + \tilde{g}_{i+1}) + \frac{D_x(a_i) + \tilde{a}_{i+1}}{Q^*}
\]
Algorithm and complexity II

Theorem (Complexity for Hermite Telescoping)

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

Degree bounds on \( g_i \) and \( a_i \):

\[
\deg_x(g_i), \quad \deg_x(a_i) \in \mathcal{O}(id_x d_y), \quad \deg_y(g_i) \leq id_y, \quad \text{and} \quad \deg_y(a_i) \leq d_y^*.
\]

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

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

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

\[
\sum_{j=0}^{i} \eta_j a_j = 0 \leadsto \begin{pmatrix} \deg_x \in \mathcal{O}(id_x d_y) \end{pmatrix} \xleftarrow{SV} \tilde{O}(i^\omega d_x d_y^2).
\]
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: $\deg_y(Q^{-}) + i \deg_y(Q^*) \in O(id_y)$. 

Shaoshi Chen

Creative Telescoping for Bivariate Rational Functions
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 \( \tilde{O}(d_x d_y^{2\omega+2}) \) ops.

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

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

\[
\begin{pmatrix}
\deg_x \leq O(id_x) \\
O(id_y) \times O(id_y)
\end{pmatrix} \xleftarrow{SV} \tilde{O}(i^{\omega+1} d_x d_y^\omega)
\]

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

Minimal telescopers
  Hermite reduction approach
  Almkvist and Zeilberger’s approach

Non-minimal telescopers

Implementation and Application

Conclusion
From \(y\)-free annihilators to telescopers

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

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

\[
i + j + \ell \leq N \Rightarrow x^i D_x^j D_y^\ell (f) \in \text{Vect}_k \left\{ \frac{x^m y^n}{Q^{N+1}} \mid m + n \leq N + (N + 1)d \right\}
\]
\[
\phi : \mathcal{W}_N \rightarrow \mathcal{W}_N(f), \quad \text{dim}(\mathcal{W}_N) \in \Theta(N^3), \quad \text{dim}(\mathcal{W}_N(f)) \in O(N^2).
\]

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: \(\text{deg}(A) \leq 3(d + 1)^2\);
- Order splitting filtration: \(\text{deg}_x(A) \leq 4d^2\) and \(\text{deg}_D(A) \leq 4d\).
Bounds by counting dimensions

**Total degree filtration:** Used in [Lipshitz1988] for diagonals

\[
\mathcal{W}_N := \text{span}_k \{ x^i D_x^j D_y^\ell \mid i + j + \ell \leq N \} \quad \left( \binom{N+3}{3} \right)
\]

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

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

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

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

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

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 \( \phi \) is not injective. So there exists a telescoper \( L(x, D_x) \) of cubic size.
Introduction

Minimal telescopers
   Hermite reduction approach
   Almkvist and Zeilberger’s approach

Non-minimal telescopers

Implementation and Application

Conclusion
Implementation and experiments

Random examples:

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

\( P \) and \( Q_i \) are generated by \texttt{randpoly()} in Maple. There are 49 cases.

<table>
<thead>
<tr>
<th>nb</th>
<th>AZ</th>
<th>Abr</th>
<th>RAZ</th>
<th>H1</th>
<th>H2</th>
<th>H0</th>
<th>EI</th>
<th>MG</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>64</td>
<td>80</td>
<td>48</td>
<td>36</td>
<td>60</td>
<td>44</td>
<td>672</td>
<td>624</td>
</tr>
<tr>
<td>2</td>
<td>72</td>
<td>96</td>
<td>40</td>
<td>36</td>
<td>48</td>
<td>56</td>
<td>1116</td>
<td>936</td>
</tr>
<tr>
<td>8</td>
<td>56</td>
<td>76</td>
<td>32</td>
<td>20</td>
<td>24</td>
<td>28</td>
<td>604</td>
<td>496</td>
</tr>
<tr>
<td>9</td>
<td>348</td>
<td>272</td>
<td>140</td>
<td>84</td>
<td>188</td>
<td>128</td>
<td>3876</td>
<td>2976</td>
</tr>
<tr>
<td>29</td>
<td>44</td>
<td>72</td>
<td>32</td>
<td>28</td>
<td>36</td>
<td>20</td>
<td>608</td>
<td>528</td>
</tr>
<tr>
<td>43</td>
<td>52</td>
<td>76</td>
<td>36</td>
<td>20</td>
<td>24</td>
<td>32</td>
<td>652</td>
<td>584</td>
</tr>
<tr>
<td>49</td>
<td>474269</td>
<td>34694</td>
<td>20977</td>
<td>10336</td>
<td>36254</td>
<td>22417</td>
<td>( \infty )</td>
<td>652968</td>
</tr>
</tbody>
</table>

(Timing in ms.)
Application to diagonals

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

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

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

**Example:** [FlaHaSo2004].

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

<table>
<thead>
<tr>
<th>$d$</th>
<th>AZ</th>
<th>Abr</th>
<th>RAZ</th>
<th>H1</th>
<th>H2</th>
<th>HO</th>
<th>RR</th>
<th>GHP</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>176</td>
<td>136</td>
<td>100</td>
<td>116</td>
<td>208</td>
<td>108</td>
<td>220</td>
<td>956</td>
</tr>
<tr>
<td>8</td>
<td>3032</td>
<td>4244</td>
<td>4380</td>
<td>1976</td>
<td>5344</td>
<td>4396</td>
<td>10336</td>
<td>154409</td>
</tr>
<tr>
<td>10</td>
<td>11740</td>
<td>12816</td>
<td>7108</td>
<td>7448</td>
<td>24565</td>
<td>7076</td>
<td>46882</td>
<td>1118313</td>
</tr>
</tbody>
</table>

(Timing in ms.)
Introduction

Minimal telescopers
   Hermite reduction approach
   Almkvist and Zeilberger’s approach

Non-minimal telescopers

Implementation and Application

Conclusion
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;
Thanks!