Differential Equations for Algebraic Functions

Alin Bostan, Frédéric Chyzak, Bruno Salvy, Grégoire Lecerf, Éric Schost

Differential Equations for Algebraic Functions

Alin Bostan, Frédéric Chyzak, Bruno Salvy
Algorithms Project, Inria Rocquencourt 78153 Le
Chesnay (France)

{Alin.Bostan,Frederic.Chyzak,Bruno.Salvy}@inria.fr

Grégoire Lecerf
Laboratoire de Mathématiques (CNRS)
Université de Versailles St-Quentin-en-Yvelines
78035 Versailles (France)
lecerf@math.uvsq.fr

Éric Schost
Department of Computer Science University of
Western Ontario London, Ontario (Canada)
eschost@uwo.ca

ABSTRACT

It is classical that univariate algebraic functions satisfy linear differential equations with polynomial coefficients. Linear recurrences follow for the coefficients of their power series expansions. We show that the linear differential equation of minimal order has coefficients whose degree is cubic in the degree of the function. We also show that there exists a linear differential equation of order linear in the degree whose coefficients are only of quadratic degree. Furthermore, we prove the existence of recurrences of order and degree close to optimal. We study the complexity of computing these differential equations and recurrences. We deduce a fast algorithm for the expansion of algebraic series.

Categories and Subject Descriptors: I.1.2 [**Symbolic and Algebraic Manipulation**]: Algorithms

General Terms: Algorithms, Experimentation, Theory

Keywords: Computer algebra, algebraic series, differential resolvents, creative telescoping, complexity

1. INTRODUCTION

A power series $\alpha(X) \in \mathbb{K}[[X]]$ is called algebraic over the field $\mathbb{K}$ when there is a nonzero polynomial $P \in \mathbb{K}[X, Y]$ such that $P(X, \alpha(X)) = 0$.

The existence of a linear differential equation with polynomial coefficients satisfied by algebraic series seems to have been observed first in 1827 by Abel [1, p. 287] who did not publish it. Cockle gave an algorithm for the computation of such a differential equation, that he called a *differential resolvent* [11]. The method was then presented by Harley [15] and rediscovered by Tannery [21] (see also [14]).

One of the applications of these differential equations is the computation of power series solutions: a linear differ-

ential equation translates into a linear recurrence, with the consequence that the number of operations required to compute the first N coefficients grows only linearly with N. This is useful in combinatorial applications, since the enumeration sequences of context-free languages have algebraic generating series. This method has been described by Comtet [12] and studied from the complexity point of view by Chudnovsky and Chudnovsky [8]. In practice, this procedure will be all the more efficient as its dependency on the size of P is small. This in turn is related to the order and degree of the coefficients of the differential equation.

We say that a linear differential equation or operator vanishing at all the roots $\alpha(X)$ of P is *associated to* P. In this work, we derive precise bounds for the order and degree of coefficients of several such equations. For brevity and clarity, the following hypothesis will be assumed throughout.

HYPOTHESIS (H). *The characteristic of the field $\mathbb{K}$ is 0. The polynomial $P \in \mathbb{K}[X, Y]$ is* separable *(wrt Y), i.e., coprime with its derivative P_Y wrt Y. The total degree of P is denoted D and its degrees in X and Y are written D_X and D_Y. We assume that $D_Y \geq 1$.*

The first work on bounds that we know of in this area is due to Cormier *et alii* [13]. They consider the monic linear differential operator of minimal order associated to P, which we call *the differential resolvent of P*. A bound in $\mathcal{O}(D_X D_Y^4)$ for the degrees of its coefficients can be deduced from their work. Nahay [17, 18] obtained degree bounds in $\mathcal{O}(D^3)$ for the coefficients of an analogous equation satisfied by α^λ with λ transcendent over $\mathbb{K}$. In §2, we prove the following cubic bound (∂_X denotes differentiation wrt X).

THEOREM 1. *Under (H), the differential resolvent of P can be written*

$$\partial_X^r + \frac{M_{r-1}}{M_r}\partial_X^{r-1} + \cdots + \frac{M_1}{M_r}\partial_X + \frac{M_0}{M_r},$$

where $r \leq D_Y$ and $M_0, \ldots, M_r$ in $\mathbb{K}[X]$ have degrees at most

$$\eta = ((2r-1)D_Y + 2r^2 - 4r + 3)D_X - r(r-1)/2 \in \mathcal{O}(rD_X D_Y).$$

Experiments show that a cubic degree is to be expected (see §4.2). This theorem improves and extends the results of [13, 17, 18]. It leads to a linear recurrence of order $\mathcal{O}(D^3)$ with

coefficients of degree $\mathcal{O}(D)$. Using fast multipoint evaluation on an arithmetic progression then leads to a computation of the first N coefficients in $\mathcal{O}(D^2\mathsf{M}(D)N)$ arithmetic operations (see §3.5). As usual, $\mathsf{M}(D)$ denotes a bound on the number of arithmetic operations in the ring of coefficients needed to multiply two polynomials of degree at most D. We make the standard assumption that $\mathsf{M}(d_1) + \mathsf{M}(d_2) \leq \mathsf{M}(d_1 + d_2)$ for all d_1, d_2 and that $D \log D \in \mathcal{O}(\mathsf{M}(D))$.

One reason for this cubic degree growth is the presence of numerous *apparent singularities.* These are points where the differential resolvent is singular (its leading coefficient vanishes) while none of its solutions is. However, the number of apparent singularities decreases dramatically if one does not insist on the order of the equation being minimal. It is actually possible to remove all apparent singularities by *Weyl closure* [22], but we do not know how to get bounds from this algorithm.

In §3.2, we show that a moderate increase of the order leads to the following *quadratic* bound.

THEOREM 2. *Under (H), there exists a nonzero linear differential operator in $\mathbb{K}[X]\langle\partial_X\rangle$ associated to P whose degree in ∂_X is at most $6D_Y$ and with coefficients of degree at most $3D_X D_Y$.*

This operator then translates into a linear recurrence of order at most $3D_Y(D_X + 2)$ with coefficients of degree at most $6D_Y$. Then, a natural question is the order of a minimal linear recurrence, or equivalently the minimal degree in X of linear differential operators in X and $\theta_X = X\partial_X$ associated to P. In §4.2, we explain why quadratic order is optimal. We prove in §3 the following precise bound.

THEOREM 3. *Under (H), there exists a non-zero differential operator in $\mathbb{K}[X]\langle\theta_X\rangle$ associated to P whose degrees in X and in θ_X are bounded above by*

$$2D_X D_Y + D_Y - \Delta^2 - \Delta + 1 \quad with \quad \Delta = D_X + D_Y - D.$$

Experiments showing that this bound is not far from optimal are reported in §4.2. This result implies the existence of a linear recurrence of order $\mathcal{O}(D^2)$ with coefficients of the same degree, whence a computation of the first N coefficients of series solutions in $\mathcal{O}(\mathsf{M}(D^2)N)$, with a small constant in the $\mathcal{O}(\cdot)$ term. In §3.4, we go a bit further: with a degree in X slightly larger, but still quadratic, we obtain a degree in θ_X that is only linear; this results in an algorithm for algebraic series in $\mathcal{O}(D\mathsf{M}(D)N)$ operations only.

2. DIFFERENTIAL RESOLVENT

We write $\alpha_1, \ldots, \alpha_{D_Y}$ for the zeroes of $P(X, \cdot)$ in the algebraic closure $\overline{\mathbb{K}(X)}$ of $\mathbb{K}(X)$. Since P is separable wrt Y, the derivation ∂_X uniquely extends to the splitting field $\mathbb{K}(X)(\alpha_1, \ldots, \alpha_{D_Y})$ of P by $\partial_X \alpha_i = -P_X(X, \alpha_i)/P_Y(X, \alpha_i)$.

2.1 Cockle's Algorithm

Cockle's algorithm [11] computes the differential resolvent. It reduces the computation to linear algebra in the quotient ring $\mathbb{K}(X)[Y]/(P)$, that is a vector space of dimension D_Y over $\mathbb{K}(X)$. The heart of the idea is to express the successive derivatives of any α_i as a polynomial of degree at most $D_Y - 1$ in itself. In order to prepare the derivation of bounds, we introduce two sequences of polynomials.

First, polynomials $(W_k)_{k\geq 1}$ in $\mathbb{K}[X, Y]$ satisfying

$$\alpha_i^{(k)} = \frac{W_k(X, \alpha_i)}{P_Y(X, \alpha_i)^{2k-1}}, \quad i \in \{1, \ldots, D_Y\}$$

are defined by induction: $W_1 = -P_X$, and for all $k \geq 1$, $W_{k+1} = (P_Y \partial_X W_k - P_X \partial_Y W_k)P_Y - (2k-1)(P_Y \partial_X P_Y - P_X \partial_Y P_Y)W_k$. They satisfy

$$\deg_X(W_k) \leq (2D_X - 1)k - D_X,$$
$$\deg_Y(W_k) \leq 2(D_Y - 1)k - D_Y + 2.$$

Cockle's algorithm computes a second sequence of polynomials, $(V_k)_{k\geq 0}$, such that $V_k(\alpha)$ is the image of W_k/P_Y^{2k-1} in $\mathbb{K}(X)[Y]/(P)$ for $k \geq 1$, while $V_0(\alpha) = \alpha$ is that of Y. The computation is as follows:

1. V_1 is computed by first inverting P_Y modulo P and then reducing the product with $-P_X$ modulo P;

2. for $k = 0, 1, \ldots,$ the polynomial V_{k+1} is computed by taking the remainder of $\partial_X V_k + V_1\partial_Y V_k$ modulo P;

3. the computation stops with the first integer $r \leq D_Y$ such that V_r belongs to the $\mathbb{K}(X)$-vector space spanned by $V_0, \ldots, V_{r-1}$. Then, together with r, Cockle's algorithm computes the relation

$$V_r = A_{r-1}V_{r-1} + \cdots + A_0 V_0, \tag{1}$$

and returns the unique monic differential operator of minimal order $M = \partial_X^r - A_{r-1}\partial_X^{r-1} - \cdots - A_1\partial_X - A_0$ associated to P.

2.2 Degree Bound

We are now going to prove Thm. 1 by establishing a formula for the differential resolvent in terms of the elementary symmetric polynomials of the α_i. This idea originates in Nahay's PhD thesis [16, Thm. 46, p. 96] (published in [18, Thm. 9.1]).

Let $Y_1, \ldots, Y_{D_Y}$ be new indeterminates and $\mathcal{N}$ the matrix:

$$\begin{pmatrix} Y_1 P_Y(X, Y_1)^{2r-1} & \cdots & Y_r P_Y(X, Y_r)^{2r-1} & 1 \\ W_1(X, Y_1)P_Y(X, Y_1)^{2r-2} & \cdots & W_1(X, Y_r)P_Y(X, Y_r)^{2r-2} & \partial_X \\ W_2(X, Y_1)P_Y(X, Y_1)^{2r-4} & \cdots & W_2(X, Y_r)P_Y(X, Y_r)^{2r-4} & \partial_X^2 \\ \vdots & & \vdots & \vdots \\ W_r(X, Y_1) & \cdots & W_r(X, Y_r) & \partial_X^r \end{pmatrix}.$$

By subtracting the jth column of $\mathcal{N}$ to the ith one ($i < j \leq r$) we observe that $Y_i - Y_j$ divides the determinant $\det(\mathcal{N})$ of $\mathcal{N}$, hence $\delta := \prod_{1\leq i<j\leq r}(Y_i - Y_j)$ divides $\det(\mathcal{N})$. The operator $L = \det(\mathcal{N})/\delta$ thus belongs to $\mathbb{K}[X][Y_1, \ldots, Y_r]\langle\partial_X\rangle$ and it is symmetric in $Y_1, \ldots, Y_r$.

By definition of the W_k's, for all σ in the permutation group S_{D_Y} of $\{1, \ldots, D_Y\}$ and all $i \in \{1, \ldots, r\}$, we have

$$L(\alpha_{\sigma(1)}, \ldots, \alpha_{\sigma(r)}, \partial_X)\alpha_i =$$
$$\left(\prod_{j=1}^{r} P_Y(X, \alpha_{\sigma(j)})\right)^{2r-1} \frac{\mathrm{Wr}(\alpha_{\sigma(1)}, \ldots, \alpha_{\sigma(r)}, \alpha_i)}{\delta(\alpha_{\sigma(1)}, \ldots, \alpha_{\sigma(r)})} = 0,$$

where $\mathrm{Wr}(\cdot)$ is the Wronskian determinant. In other words $L(\alpha_{\sigma(1)}, \ldots, \alpha_{\sigma(r)}, \partial_X)$ is associated to P for all $\sigma \in S_{D_Y}$.

For all $k \in \{0, \ldots, r\}$, the entries of the $(k + 1)$th row of $\mathcal{N}$ have degree in X at most $m_k = (2r-1)D_X - k$, except in the last column. Therefore, for all $i \in \{0, \ldots, r\}$, the degree

of the coefficient L_i of ∂_X^i in L is bounded by:

$$\deg_X(L_i) \leq \sum_{k=0,k\neq i}^r m_k \leq r(2r-1)D_X - r(r-1)/2 =: m_X.$$

On the other hand, for all $l \in \{1,\ldots,r\}$, the entries of the lth column of $\mathcal{N}$ have degree in Y_l at most $(2r-1)D_Y - 2r + 2$. Therefore, for all $i \in \{0,\ldots,r\}$, all $j \geq 0$, and all $l \in \{1,\ldots,r\}$, the degree of the coefficient $L_{i,j}$ of X^j in L_i is bounded by:

$$\deg_{Y_l}(L_{i,j}) \leq (2r-1)D_Y - 2r + 2 - (r-1) =: m_Y.$$

By definition of r, there exists a permutation $\sigma \in S_{D_Y}$ such that $L(\alpha_{\sigma(1)},\ldots,\alpha_{\sigma(r)},\partial_X)$ has order r exactly. Therefore, Lemma 1 below provides us with a nonzero operator $\bar{L} \in \mathbb{K}[X][Y_1,\ldots,Y_{D_Y}]\langle\partial_X\rangle$ symmetric in $Y_1,\ldots,Y_{D_Y}$, such that $\bar{L}(\alpha_1,\ldots,\alpha_{D_Y},\partial_X)$ has order r, is associated to P, and $\deg_{Y_i}(\bar{L}) \leq \max(m_Y, D_Y-1) = m_Y$, for all $i \in \{1,\ldots,D_Y\}$.

Let p_i denote the coefficient of Y^i in P seen in $\mathbb{K}[X][Y]$. By Lemma 2 below, there exists $\tilde{L} \in \mathbb{K}[X][Z_1,\ldots,Z_{D_Y}]\langle\partial_X\rangle$ of total degree at most m_Y in $Z_1,\ldots,Z_{D_Y}$, of partial degree in X at most m_X, and such that $\bar{L}(\alpha_1,\ldots,\alpha_{D_Y},\partial_X) = \tilde{L}(p_0/p_{D_Y},\ldots,p_{D_Y-1}/p_{D_Y},\partial_X)$.

Finally, $\tilde{M} = p_{D_Y}^{m_Y}\tilde{L}(p_0/p_{D_Y},\ldots,p_{D_Y-1}/p_{D_Y},\partial_X)$ is in $\mathbb{K}[X]\langle\partial_X\rangle$, has order r and degree at most $m_X + D_X m_Y$ in X, and is associated to P. The conclusion of Thm. 1 thus follows from the uniqueness of the minimal operator.

The rest of the proof consists of elementary properties of symmetric polynomials. If $r \leq n$ then the group S_r of permutations of $\{1,\ldots,r\}$ is naturally embedded into S_n. We write $(S_n/S_r)_l$ for the left quotient of S_n by S_r.

LEMMA 1. *Let $\mathbb{E}$ be a field, let $r \leq n$ be nonnegative integers, and let $\alpha_1,\ldots,\alpha_n$ be pairwise distinct elements in $\mathbb{E}$. If $E \in \mathbb{E}[Y_1,\ldots,Y_r]$ is symmetric and does not vanish at $(\alpha_1,\ldots,\alpha_r)$, then there exist nonnegative integers $e_{r+1} \leq r,\ldots,e_n \leq n-1$ such that*

$$\bar{E} = \sum_{\sigma \in (S_n/S_r)_l} E(Y_{\sigma(1)},\ldots,Y_{\sigma(r)})Y_{\sigma(r+1)}^{e_{r+1}}\cdots Y_{\sigma(n)}^{e_n}$$

is a symmetric polynomial in $\mathbb{E}[Y_1,\ldots,Y_n]$ that does not vanish at $(\alpha_1,\ldots,\alpha_n)$.

PROOF. For fixed r, the proof proceeds by induction on n. If $n = r$, then the lemma trivially holds. We now assume that the lemma holds for some $n \geq r$, and prove that it holds for $n+1$. The induction hypothesis provides us with nonnegative integers $e_{r+1} \leq r,\ldots,e_n \leq n-1$ such that $\bar{E}$ (as defined in the lemma) is a symmetric polynomial in $Y_1,\ldots,Y_n$ that does not vanish at $(\alpha_1,\ldots,\alpha_n)$. Since the Vandermonde matrix of $\alpha_1,\ldots,\alpha_{n+1}$ is invertible, there exists $e_{n+1} \leq n$ such that

$$\sum_{\rho \in (S_{n+1}/S_n)_l} \bar{E}(Y_{\rho(1)},\ldots,Y_{\rho(n)})Y_{\rho(n+1)}^{e_{n+1}}$$

is symmetric in $Y_1,\ldots,Y_{n+1}$ and moreover does not vanish at $(\alpha_1,\ldots,\alpha_{n+1})$. Plugging the expression of $\bar{E}$ into the latter sum concludes the proof for $n+1$. $\square$

LEMMA 2. *[24, Thm 6.21] Let $\mathbb{A}$ be a commutative ring, and let $E \in \mathbb{A}[Y_1,\ldots,Y_n]$ be symmetric of partial degree D in each of the Y_i's. Then there exists a unique polynomial $F \in$*

$\mathbb{A}[Z_1,\ldots,Z_n]$ *of total degree D such that $E(Y_1,\ldots,Y_n) = F(\Sigma_1,\ldots,\Sigma_n)$, where Σ_i represents the i-th elementary symmetric polynomial in $Y_1,\ldots,Y_n$.*

2.3 Algorithm and Complexity

With the bounds of the previous section, we now turn to the computation of the differential resolvent. As in [8, 13], the idea is to compute power series expansions of the coefficients and recover them by Padé approximants. For this, we have to choose a point where expansion is easy and determine a bound on the order of the series expansions.

DEFINITION 1. *A point $a \in \mathbb{K}$ that annihilates neither the leading coefficient p_{D_Y} nor the discriminant Δ_Y of P seen in $\mathbb{K}[X][Y]$, and such that $V_0(a),\ldots,V_{r-1}(a)$ are linearly independent in $\mathbb{K}[Y]/(P(a,Y))$ is called a* lucky *point.*

If a is lucky, then the α_i's can be seen as power series in $\overline{\mathbb{K}}[[X-a]]$, and the V_i's as elements of $\overline{\mathbb{K}}[[X-a]][Y]/(P(a,Y))$. Here $\overline{\mathbb{K}}$ represents the algebraic closure of $\mathbb{K}$.

LEMMA 3. *There exists $\Phi \in \mathbb{K}[X] \setminus \{0\}$ of total degree at most η (as defined in Thm. 1) such that a is lucky whenever $p_{D_Y}(a)\Delta_Y(a)\Phi(a) \neq 0$.*

PROOF. For Φ we take the leading coefficient of the operator $\tilde{M}$ introduced at the end of the proof of Thm. 1. If $p_{D_Y}(a)\Delta_Y(a) \neq 0$ then we can define $A(a)$ as the $r \times D_Y$ matrix $(\alpha_j^{(i-1)}(a))_{i,j}$ with entries in $\overline{\mathbb{K}}[[X-a]]$, $B(a)$ as the $D_Y \times r$ matrix in $\mathbb{K}[[X-a]]$ whose columns are the coefficients in Y of $V_0,\ldots,V_{r-1}$, and $C(a)$ as the invertible Vandermonde matrix of the α_i. If $\Phi(a) \neq 0$ then $A(a)$ has rank r by construction, and so has $B(a)$ since $A(a) = C(a)B(a)$. $\square$

Lemma 3 implies that all but a finite set of elements in $\mathbb{K}$ are lucky. A lucky point a can thus be obtained either deterministically by computing the rank of $V_0(a),\ldots,V_{r-1}(a)$ for more than $\deg_X(p_{D_Y}\Delta_Y\Phi) \in \mathcal{O}(D_X D_Y^2)$ points, or probabilistically by means of the Schwartz-Zippel zero test [23, Lemma 6.44]. Once a lucky point is determined (which is very fast by the probabilistic approach) then the computation of the differential resolvent of P can be done efficiently.

PROPOSITION 1. *Given a lucky point $a \in \mathbb{K}$ and the order r of M, the computation of M can be done with $\mathcal{O}((r^\omega + r\mathsf{M}(D_Y) + r\log(D_X))\mathsf{M}(\eta))$ arithmetic operations in $\mathbb{K}$.*

Here $2 < \omega \leq 3$ is the exponent of matrix multiplication, see, e.g., [23].

PROOF. First of all, shifting the variable X by a in P takes $\mathcal{O}(D_Y\mathsf{M}(D_X))$ operations in $\mathbb{K}$ by means of [4, Ch. 1, §2] since the characteristic is zero. In order to reconstruct the coefficients of M from their series expansions via Padé approximants [23, §5.9], a precision $(X-a)^{2\eta+1}$ is sufficient. The idea is to follow the steps of Cockle's algorithm in $\mathbb{K}[[X-a]][Y]/(P)$ instead of $\mathbb{K}(X)[Y]/(P)$. The precision of the series is initially set to $2\eta + 1 + r \in \mathcal{O}(\eta)$ in order to accommodate the precision loss due to differentiation.

In Step 1, the computation of V_1 takes $\mathcal{O}(\mathsf{M}(D_Y)\mathsf{M}(\eta))$ operations by [23, Thms. 15.10 & 15.11]. In Step 2, each V_{k+1} can be obtained from V_k and V_1 by means of $\mathcal{O}(\mathsf{M}(D_Y)\mathsf{M}(\eta))$

operations. In Step 3, we start by computing the approximation of the relation (1) to precision $(X - a)$. This boils down to linear algebra over $\mathbb{K}$ in size $(r + 1) \times D_Y$, hence a cost in $\mathcal{O}(r^{\omega-1}D_Y)$. From this we use Hensel-Newton lifting in order to recover relation (1) to precision $(X - a)^{2\eta+1}$ with $\mathcal{O}(r^\omega \mathsf{M}(\eta))$ operations in $\mathbb{K}$. The rational reconstruction of each A_i and its shift back of X by $-a$ cost $\mathcal{O}(\mathsf{M}(\eta) \log(\eta))$. $\quad\square$

When $r = D_Y$ the cost of the computation of M drops to $\tilde{\mathcal{O}}(D_X D_Y^{\omega+2})$ when fast multiplication is used. (Here the tilde indicates that the factors polynomial in $\log D_X$, $\log D_Y$ and their logarithms have been omitted.) Note that the results presented in this section extend to analyze the bit-complexity via p-adic approximation when $\mathbb{K} = \mathbb{Q}$.

3. CREATIVE TELESCOPING FOR ALGEBRAIC FUNCTIONS

In this section we search for differential operators associated to P in $\mathbb{K}(X)\langle\partial_X\rangle$ with non-minimal orders but with degrees in X much smaller than in the resolvent of P.

3.1 Principle

Let $\alpha(X)$ be a solution of $P(X, \alpha(X)) = 0$. If $P \in \mathbb{C}[X, Y]$, then by the residue theorem,

$$\alpha(X) = \frac{1}{2\pi i} \oint F(X,Y)dY, \text{ with } F(X,Y) = \frac{Y P_Y(X,Y)}{P(X,Y)}$$

where the contour is for instance a circle containing $\alpha(X)$ and no other root of P. In this setting, *creative telescoping* can be seen as an algorithm based on differentiating under the integral sign and integrating by parts [2]. It consists in computing a linear differential operator of the form

$$\Lambda = A(X, \partial_X) + \partial_Y B(X, \partial_X, Y, \partial_Y) \qquad (2)$$

that cancels the rational function F. There, ∂_X and ∂_Y denote differentiation with respect to X and Y. If such a Λ exists, integrating $\Lambda(Y P_Y/P) = 0$ shows that A cancels α. We now give an algebraic proof of this property.

PROPOSITION 2. *Under (H), suppose that an element $\Lambda \in \mathbb{K}[X,Y]\langle\partial_X, \partial_Y\rangle$ is of the form (2) and that $\partial_Y^k \Lambda$ cancels the rational function $F = Y P_Y / P$ for some integer $k \geq 0$. Then, the operator A is associated to P.*

PROOF. The element $F = Y P_Y/P$ of $\mathbb{K}(X,Y)$ can be expanded as a Laurent series in $\mathbb{K}(X,\alpha)((Y - \alpha))$:

$$F = \frac{\alpha}{Y - \alpha} + \sum_{i \geq 0} \gamma_i (Y - \alpha)^i$$

with coefficients γ_i in $\mathbb{K}(X, \alpha)$. On the other hand the action of derivations in X and Y on a summand $\gamma(Y - \alpha)^i$, for γ in $\mathbb{K}(X, \alpha)$, is given for any $i \in \mathbb{Z}$ by

$$\partial_Y \left(\gamma(Y - \alpha)^i\right) = i\gamma(Y - \alpha)^{i-1}$$

$$\partial_X \left(\gamma(Y - \alpha)^i\right) = \gamma'(Y - \alpha)^i - i\gamma(Y - \alpha)^{i-1}\alpha'.$$

Thus no term in $(Y - \alpha)^{-1}$ remains after action of ∂_Y. In this way, applying $\partial_Y^k \Lambda$ to F and considering the resulting Laurent series yields

$$0 = \partial_Y^k \Lambda(F) = \partial_Y^k \Lambda \left(\frac{\alpha}{Y - \alpha}\right) \Rightarrow (-1)^k \frac{A(X, \partial_X)(\alpha)}{(Y - \alpha)^{k+1}} = 0,$$

which proves that A annihilates α. $\quad\square$

3.2 Quadratic Bound

We now prove Thm. 2 by considering the degrees of the numerators and denominators of the successive derivatives of the rational function $F = Y P_Y/P$. By Leibniz's rule, for any nonnegative integers N_X and N_∂, the $r_{N_X, N_\partial} := (N_X + 1)\binom{N_\partial + 2}{2}$ rational functions

$$X^i \partial_X^j \partial_Y^k F, \quad 0 \leq i \leq N_X, \ 0 \leq j + k \leq N_\partial$$

are contained in the $\mathbb{K}$-vector space of dimension at most $s_{N_X, N_\partial} := (D_Y(N_\partial + 1) + 1)(N_X + 1 + D_X(N_\partial + 1))$ spanned by the elements

$$\frac{X^i Y^j}{P^{N_\partial + 1}}, \quad 0 \leq i \leq N_X + D_X(N_\partial + 1), \ 0 \leq j \leq D_Y(N_\partial + 1).$$

With the values of N_X and N_∂ given in the theorem, the difference $r_{N_X, N_\partial} - s_{N_X, N_\partial}$ becomes

$$(12D_Y^2 - 7D_Y - 1)D_X + 12D_Y^2 + 8D_Y$$

which is positive for positive D_Y.

This proves the existence of a nonzero differential operator $\Lambda(X, \partial_X, \partial_Y)$ with degree in X at most N_X and total degree in ∂_X and ∂_Y at most N_∂ that annihilates F. Let now $k \leq N_\partial$ be such that Λ rewrites $\partial_Y^k(A(X, \partial_X) + \partial_Y B(X, \partial_X, \partial_Y))$ with $A \neq 0$. The conclusion follows from Prop. 2.

3.3 Refined Bound

We now prove Thm. 3. The bound of the previous section relies on the elimination of Y in the left ideal annihilating F in the Weyl algebra $W_{X,Y}(\mathbb{K})$. In the context of creative telescoping, this is known as Zeilberger's slow algorithm, which relies on holonomic D-modules [25]. A faster algorithm for hyperexponential integrands [2] consists in looking for an operator of the form (2). Indeed, letting B in (2) depend on Y also allows for the existence of operators of smaller order. In this context, the only good bounds [3] we know of do not apply to rational integrands. Instead, we give a construction inspired by one of Takayama [20, Thm. 2.4 and Appendix 6] showing that any differentiably finite function is holonomic.

For each integer $d \geq 0$, we consider the vector space $\mathcal{V}_d$ of rational functions G/P^{d+1}, with polynomials $G \in \mathbb{K}[X,Y]$ satisfying

$$\deg G \leq (d+1)D + d, \quad \deg_X G \leq (d+1)D_X + d,$$
$$\deg_Y G \leq (d+1)D_Y.$$

This is defined in such a way that an induction on d shows that the vector space F_d generated by the rational functions

$$X^i \theta_X^j \cdot F, \quad 0 \leq i, j \leq d \qquad (3)$$

is a subspace of $\mathcal{V}_d$, and that so is $\partial_Y \cdot \mathcal{V}_{d-1}$. We are going to establish that for d large enough,

$$\dim_{\mathbb{K}} F_d + \dim_{\mathbb{K}}(\mathcal{V}_d \cap \partial_Y \cdot \mathcal{V}_{d+1}) > \dim_{\mathbb{K}} \mathcal{V}_d. \qquad (4)$$

From here the result follows: both F_d and $\mathcal{V}_d \cap \partial_Y \cdot \mathcal{V}_{d+1}$ are subspaces of $\mathcal{V}_d$ whose dimensions are such that they have a nonzero intersection. Any element of this intersection is of the form $AF = \partial_Y(G/P^{d+2})$ with $AF \in F_d$. By the same arguments as in the proof of Prop. 2, we deduce that A is associated to P.

We now consider the dimensions in (4) more precisely.

The generators (3) are linearly independent since $P_Y \neq 0$: the reduced form of $\theta_X^i \cdot F$ has denominator P^{i+1}, which cannot be eliminated by multiplication by a polynomial in X only. Therefore $\dim_{\mathbb{K}} F_d = (d+1)^2$.

Write P in the form $X^{D_X}\phi + \tilde{P}$ where ϕ depends only on Y. For $d \geq 0$, define the vector space $\mathcal{W}_d$ of rational functions F/P^d, with polynomials $F \in \mathbb{K}[X,Y]$ of the form $F = cX^{(D_X+1)d+1}\phi^d + H$ where c is in $\mathbb{K}$ and H satisfies

$$\deg H \leq (D+1)d+1, \quad \deg_X H \leq (D_X+1)d, \quad (5)$$
$$\deg_Y H \leq D_Y d. \quad (6)$$

The decomposition of F induces $\mathcal{W}_d \subseteq \mathcal{V}_{d+1}$ directly. Then, from

$$P^{d+1}\partial_Y\left(\frac{F}{P^d}\right) = cd\phi^{d-1}X^{(D_X+1)d+1}(\tilde{P}\partial_Y\phi - \phi\partial_Y\tilde{P})$$
$$+P\partial_Y H - dH\partial_Y P,$$

we obtain the inclusion $\partial_Y \cdot \mathcal{W}_d \subseteq \mathcal{V}_d$. It follows that $\partial_Y \cdot \mathcal{W}_d \subseteq \mathcal{V}_d \cap \partial_Y \cdot \mathcal{V}_{d+1}$, and that Eq. (4) holds as soon as

$$\dim_{\mathbb{K}} F_d + \dim_{\mathbb{K}} \partial_Y \cdot \mathcal{W}_d > \dim_{\mathbb{K}} \mathcal{V}_d. \quad (7)$$

In order to compute the dimension of $\partial_Y \cdot \mathcal{W}_d$, we first consider the kernel of $\partial_Y|_{\mathcal{W}_d}$, which is precisely $\mathbb{K}[X] \cap \mathcal{W}_d$. In view of the degree constraints (5–6), this corresponds to numerators $F = P^d Q$ with $Q \in \mathbb{K}[X]$ and

$$\deg_X Q \leq (D_X+1)d+1 - D_X d = d+1,$$

so that $\dim_{\mathbb{K}}(\partial_Y \cdot \mathcal{W}_d) = \dim_{\mathbb{K}} \mathcal{W}_d - (d+2)$.

The dimensions of $\mathcal{V}_d$ and $\mathcal{W}_d$ are given by the possible supports of the numerators of their elements. Lemma 4 below thus provides us with:

$$\dim_{\mathbb{K}}\mathcal{V}_d = (d+1)(D_X+1)((d+1)D_Y+1) - \binom{d\Delta+\Delta+1}{2},$$

$$\dim_{\mathbb{K}}\mathcal{W}_d = ((D_X+1)d+1)(D_Y d+1) - \binom{d\Delta}{2} + 1.$$

Injecting the three dimensions in Eq. (7) finally gives the following condition:

$$d^2 - \left(2D_X D_Y + D_Y - \Delta^2 - \Delta - 1\right)d$$
$$+ \frac{(\Delta-1)\Delta}{2} - (D_X+1)(D_Y+1) > 0.$$

Call ψ the value obtained on substituting the bound of the theorem for d in the left-hand side:

$$\psi = (3Dy-1)Dx + Dy - 3\Delta(\Delta+1)/2 + 2.$$

From $D_Y \geq 1$ and $0 \leq \Delta \leq \min(D_X, D_Y) =: m \geq 1$ it follows that

$$\psi \geq (3m-1)m+m-3m(m+1)/2+2 = 3m(m-1)/2+2 > 0.$$

LEMMA 4. *The number of monomials in X and Y of total degree at most δ, degree in X at most δ_X, and degree in Y at most δ_Y is $(\delta_X+1)(\delta_Y+1) - \binom{\delta_X+\delta_Y-\delta+1}{2}$, whenever $\delta_X+\delta_Y+1 \geq \delta \geq \max(\delta_X,\delta_Y)$.*

3.4 Construction of Differential Operators

Given a polynomial $P \in \mathbb{K}[X,Y]$ satisfying (H), we have proved in the previous sections the existence of several operators in $\mathbb{K}[X]\langle\partial_X\rangle$ associated to P. The aim of this section is to study the efficient construction of such operators. This is done under an additional, but mild, hypothesis ((H') below). As a by-product, we obtain in Thm. 4 an alternative unified proof of slightly weaker versions of Thms. 1, 2 and 3.

Direct Algorithms. The proof of Thm. 2 boils down to a kernel computation of a matrix of size $\sim D_X D_Y^3$. Even using FFT for (univariate and bivariate) polynomial multiplication and a Wiedemann-like approach for the linear algebra step, this computation requires $\sim D_X^2 D_Y^6$ operations in $\mathbb{K}$.

In the sequel we provide an algorithm with better cost, which is based on Padé–Hermite approximation of series expansions of an algebraic solution and its derivatives.

Notation and Assumptions. We assume in the rest of this section that the following hypothesis holds.

HYPOTHESIS (H'). *Besides Hypothesis (H),*

(H_a) $D_Y \geq 2$ and P is absolutely irreducible in $\mathbb{K}[X,Y]$ *(which means that P is irreducible when seen over the algebraic closure $\overline{\mathbb{K}}$ of $\mathbb{K}$), and*

(H_b) $x = 0$ annihilates neither the discriminant $\Delta_Y(X)$ nor the leading coefficient $P_{D_Y}(X)$ of P, seen in $\mathbb{K}[X][Y]$.

To clarify the status of these assumptions, note that (H_a) is a global property of P which is crucial in the subsequent proof that our algorithm works correctly. In contrast, the local property (H_b) is purely of technical nature, being similar with that on *lucky points* in Section 2.3. (H_b) can be ensured deterministically and in good complexity after a change of coordinates $X \leftarrow X + a$ with $a \in \mathbb{K}$ not a zero of $P_{D_Y}\Delta_Y$.

Principle of the Algorithm. Under (H_b), the polynomial $p(Y) := P(0,Y)$ is separable of degree D_Y. By Hensel's lemma, the D_Y roots α_i of $P(X,\alpha_i) = 0$ are all in $\overline{\mathbb{K}}[[X]]$. To avoid the factorization of $p(Y)$, we work in the quotient algebra $\mathbb{A} := \mathbb{K}[Y]/(p(Y))$. Let y denote the residue class of Y in $\mathbb{A}$. By Hensel's lemma again, there exists a unique series $\varphi \in \mathbb{A}[[X]]$ such that $\varphi(0) = y$ and $P(X, \varphi(X)) = 0$. Moreover, the expansion of φ to any prescribed precision can be efficiently computed by Newton iteration.

To give a unified presentation, we denote by ∂ one of the derivation operators ∂_X and θ_X. The principle of our algorithm is to compute enough terms of $\varphi(X)$, so that the coefficients of an associated operator can be recovered by Padé–Hermite (abbrev. PH) approximation of the vector $(\varphi, \partial\varphi, \partial^2\varphi, \ldots)^t$ of the first derivatives of φ wrt ∂. Thus, we obtain first an operator with coefficients in $\mathbb{A}[X]$, from which we extract an operator in $\mathbb{K}[X]\langle\partial\rangle$ associated to P. We give in Figure 1 a detailed presentation of the algorithm. The main difficulty of this approach is to provide suitable bounds on the orders of truncation. This is done in the next theorem, which encapsulates the main results of this section.

THEOREM 4. *Let $P \in \mathbb{K}[X,Y]$ be of bi-degree (D_X, D_Y) and satisfy Hypothesis (H'). The number of ops in $\mathbb{K}$ for*

1. $\mathsf{AlgToDiff}(P, 4D_X D_Y^2, D_Y, \partial)$ *is $\tilde{\mathcal{O}}(D_X D_Y^{\omega+3})$;*

AlgToDiff$(P, B_X, B_\partial, \partial)$

Input: $P \in \mathbb{K}[X, Y]$, $B_X, B_\partial \in \mathbb{N}$, $\partial \in \{\partial_X, \theta_X\}$

Output: an operator $L \in \mathbb{K}[X]\langle\partial\rangle \setminus \{0\}$ associated to P, with $\deg_X(L) \le B_X$ and $\deg_\theta(L) \le B_\partial$.

1. Set $\Sigma := B_X B_\partial + B_X + B_\partial$.
2. Expand $\varphi(X) \in \mathbb{A}[[X]]$ up to precision $\Sigma + B_\partial$.
3. Compute $Z_i = \partial^i \varphi$, for $i = 0, 1, \ldots, B_\partial$ to precision Σ.
4. Compute a PH approximant $(\ell_0(X), \ldots, \ell_{B_\partial}(X)) \ne 0$ in $\mathbb{A}[X]^{B_\partial + 1}$ of $(Z_0, \ldots, Z_{B_\partial})$, of type $(B_X, \ldots, B_X)$.
5. Write $L = \ell_0(X) + \ell_1(X)\partial + \cdots + \ell_{B_\partial}(X)\partial^{B_\partial}$ as $L_0 + yL_1 + \cdots + y^{D_Y - 1}L_{D_Y - 1}$, with $L_i \in \mathbb{K}[X]\langle\partial\rangle$.
6. Return any $L_i \ne 0$.

Figure 1: Construction of differential operators by Padé–Hermite approximation.

2. AlgToDiff$(P, 5D_X D_Y, 5D_Y, \partial)$ is $\tilde{\mathcal{O}}(D_X D_Y^{\omega+2})$;

3. AlgToDiff(P, B, B, ∂) with $B = 4D_X D_Y + D_Y - 2D_X - 2$ is $\tilde{\mathcal{O}}(D_X^{\omega+1} D_Y^{\omega+2})$.

All three output a non-zero differential operator in $\mathbb{K}[X]\langle\partial\rangle$ associated to P.

The first algorithm returns an operator of order at most D_Y, but with coefficients of cubic degree. The second one returns an operator whose order is still linear in D_Y, but whose coefficients have only quadratic degree. Finally, the last one returns an operator of quadratic order and degree, whose degree is smaller than the previous one.

First we give a bound on the number of terms needed to reconstruct an operator.

LEMMA 5. *Under (H'), let $L \in \overline{\mathbb{K}}[X]\langle\partial\rangle \setminus \{0\}$ be of order at most B_∂ in ∂, with coefficients of degrees at most B_X.*

(a) *If $\alpha_0 \in \overline{\mathbb{K}}[[X]]$ is such that $P(X, \alpha_0) = 0$ and $L\alpha_0 = 0$, then L is associated to P.*

(b) *Let $\sigma = 4D_X D_Y B_\partial + B_X D_Y - 2D_X B_\partial$. If $\alpha_0 \in \overline{\mathbb{K}}[[X]]$ is such that $P(X, \alpha_0)$ and $L\alpha_0$ are zero modulo X^σ, then L is associated to P.*

PROOF. There exists $Q(X, Y) \in \overline{\mathbb{K}}[X, Y]$ with

$$\deg_X(Q) \le 2B_\partial D_X + B_X - D_X$$
$$\deg_Y(Q) \le 2(D_Y - 1)B_\partial - D_Y + 2,$$

such that $L\alpha = \frac{Q(X, \alpha)}{P_Y(X, \alpha)^{2B_\partial - 1}}$ for any root $\alpha \in \overline{\mathbb{K}(X)}$ of $P(X, \alpha) = 0$. Indeed, if $L = \ell_0(X) + \cdots + \ell_{B_\partial}(X)\partial^{B_\partial}$, we take $Q = \ell_0 Y P_Y^{2B_\partial - 1} + \ell_1 W_1 P_Y^{2B_\partial - 2} + \cdots + \ell_{B_\partial} W_{B_\partial}$, where the $W_i(X, Y)$ are the polynomials introduced in Section 2.

Let $R(X) \in \overline{\mathbb{K}}[X]$ be the resultant of $P(X, Y)$ and $Q(X, Y)$ with respect to Y. To show that L is associated to P, it is enough to prove that R is identically zero. Indeed, since P is irreducible in $\overline{\mathbb{K}}[X, Y]$, the nullity of R implies that P divides Q [23, Cor. 6.20], therefore $Q(X, \alpha) = 0$, and thus $L\alpha = 0$, for any $\alpha \in \overline{\mathbb{K}(X)}$ such that $P(X, \alpha) = 0$.

(a) By definition, there exist polynomials $U, V \in \overline{\mathbb{K}}[X, Y]$ such that $U(X, Y)P(X, Y) + V(X, Y)Q(X, Y) = R(X)$. Evaluating this equality at $Y = \alpha_0$ yields $R(X) = 0$, as wanted.

The same reasoning shows that under the hypothesis of (b) we have $R(X) = 0 \mod X^\sigma$. On the other hand, by Bézout's theorem, one can bound the degree of the resultant R by $D_X \deg_Y Q + D_Y \deg_X Q$. Therefore $\deg(R)$ is at most

$$D_X(2D_Y B_\partial - 2B_\partial - D_Y + 2) + D_Y(2B_\partial D_X + B_X - D_X)$$
$$< 4D_X D_Y B_\partial + B_X D_Y - 2D_X B_\partial = \sigma,$$

since $D_Y \ge 2$. Finally $\deg(R) \le \sigma$ and $R(X) = 0 \mod X^\sigma$ imply that R is the zero polynomial. $\square$

PROOF OF THM. 4. In Step 4, the Padé–Hermite equality $\sum_i \ell_i Z_i = 0 \mod X^\Sigma$ can be viewed as an under-determined homogeneous linear system with Σ equations in the $\Sigma + 1$ unknown coefficients of the polynomials ℓ_i. It thus admits at least one non-trivial solution in $\mathbb{A}^{\Sigma+1}$.

Next let $p = \prod_i p_i$ represent the irreducible factorization of p in $\mathbb{K}[Y]$, and let $\mathbb{A}_i = \mathbb{K}[Y]/(p_i(Y))$, so that $\mathbb{A}$ rewrites into the product of fields $\prod_i \mathbb{A}_i$. Let $\varphi^{(i)}$ be the image of φ in $\mathbb{A}_i[[X]] \subseteq \overline{\mathbb{K}}[[X]]$ and let $L^{(i)}$ be the image of L in $\mathbb{A}_i[X]\langle\partial\rangle \subseteq \overline{\mathbb{K}}[X]\langle\partial\rangle$. By construction $P(X, \varphi^{(i)}) = L^{(i)}\varphi^{(i)} = 0 \mod X^\Sigma$ holds in $\overline{\mathbb{K}}[[X]]$. Since all choices of (B_X, B_∂) as specified in Thm 4 ensure $\Sigma \ge \sigma$, it follows from Lemma 5(b) that $L^{(i)}$ is associated to P for all i. In other words, for any root $a \in \overline{\mathbb{K}}$ of $p(a) = 0$, and any root $\alpha \in \overline{\mathbb{K}}[[X]]$ of $P(X, \alpha) = 0$ we have that $L_0\alpha + aL_1\alpha + \cdots + a^{D_Y - 1}L_{D_Y - 1}\alpha = 0$. We thus deduce that L_i is associated to P for all i, which concludes the proof of correctness.

We are left with the detailed description and complexity study of each step of the algorithm. Operations $(+, \times)$ in $\mathbb{A}$ have cost $\mathcal{O}(\mathsf{M}(D_Y))$ operations in $\mathbb{K}$, while inversion (when possible) takes $\mathcal{O}(\mathsf{M}(D_Y) \log D_Y)$. For simplicity, in what follows we suppose that FFT is used for polynomial multiplication, thus $\mathsf{M}(d) \in \tilde{\mathcal{O}}(d)$ for all d.

For the lifting step 2, we use the Newton iteration. By [7, Algo. 2, Prop. 3], under the hypothesis (H'), the first Σ terms of $\varphi \in \mathbb{A}[[X]]$ cost $\mathcal{O}(\Sigma D_Y^{(\omega+1)/2} + \sqrt{D_Y}\mathsf{M}(\Sigma)\mathsf{M}(D_Y)) \in \tilde{\mathcal{O}}(B_X B_\partial D_Y^{(\omega+1)/2})$ ops in $\mathbb{K}$.

Since differentiating has linear complexity, Step 3 requires $\mathcal{O}(B_\partial^2 B_X)$ ops in $\mathbb{A}$, hence at most $\tilde{\mathcal{O}}(B_\partial^2 B_X D_Y)$ ops in $\mathbb{K}$.

For Step 5, we use the approximation algorithm in [19], which computes a nonzero Padé–Hermite approximant of type $(B_X, \ldots, B_X)$ for $(Z_0, \ldots, Z_{B_\partial})$ within $\tilde{\mathcal{O}}(B_\partial^\omega B_X)$ ops in $\mathbb{A}$, provided that $\mathbb{A}$ is a field, or more generally, if no division by a non-invertible element in $\mathbb{A}$ occurs.

In the general case, since $\mathbb{A}$ is not a field, we use the following strategy. We begin by applying the approximation algorithm over $\mathbb{A}$. If no forbidden division occurs, this algorithm outputs the desired operator L with coefficients in $\mathbb{A}[X]$. If we encounter a nonzero $q \in \mathbb{A}$ that is not invertible, then by a gcd computation we obtain a proper factor $\tilde{p}$ of p of degree at most $D_Y/2$ such that q is either invertible or zero modulo $\tilde{p}$. This takes $\tilde{\mathcal{O}}(D_Y)$ operations in $\mathbb{K}$ [23, Th. 11.5]. We replace $\mathbb{A}$ by $\mathbb{K}[Y]/(\tilde{p}(Y))$ and restart the approximation algorithm. This procedure is repeated at most $\log D_Y$ times, until all the divisions are doable. The total cost of this computation is $\tilde{\mathcal{O}}(B_\partial^\omega B_X D_Y)$ operations in $\mathbb{K}$.

Finally, Steps 1, 5 and 6 do not use any arithmetic operation. Adding together these costs, and under the assumption $D_Y \le B_\partial^2$ (satisfied by the choices in Thm. 4), we get that AlgToDiff$(P, B_X, B_\partial, \partial)$ uses $\tilde{\mathcal{O}}(B_\partial^\omega B_X D_Y)$ ops in $\mathbb{K}$. $\square$

Probabilistic and Heuristic Versions. We now give a probabilistic version of Algorithm AlgToDiff, called AlgToDiffP. The idea is to compute a Padé–Hermite approximant over $\mathbb{K}$ rather than over $\mathbb{A}$. The probabilistic algorithm AlgToDiffP is obtained by replacing Steps 3–6 in AlgToDiff by:

3′. Let $\varphi(X) = \varphi_0(X) + y\varphi_1(X) + \cdots + y^{D_Y-1}\varphi_{D_Y-1}(X)$. Choose random $a_0, a_1, \ldots, a_{D_Y-1} \in \mathbb{K} \setminus \{0\}$ and compute $Z_0 = \sum_i a_i\varphi_i(X) \bmod X^\Sigma$, $Z_i = \partial^i Z_0, i = 1, \ldots, B_\partial$.

4′. Compute a PH approximant $(\ell_0(X), \ldots, \ell_{B_\partial}(X))$ in $\mathbb{K}[X]^{B_\partial+1}$ of $(Z_0, Z_1, \ldots, Z_{B_\partial})$ of type $(B_X, \ldots, B_X)$.

5′. Return $L = \ell_0(X) + \ell_1(X)\partial + \cdots + \ell_{B_\partial}(X)\partial^{B_\partial}$.

AlgToDiffP$(P, 5D_X D_Y, 5D_Y, \partial)$ uses $\tilde{\mathcal{O}}(D_X D_Y^{(\omega+5)/2})$ ops in $\mathbb{K}$; with B as in Thm. 4, the cost of AlgToDiffP(P, B, B, ∂) is $\tilde{\mathcal{O}}(D_X^2 D_Y^{(\omega+5)/2} + (D_X D_Y)^{\omega+1})$. The exponent of D_Y is better than for the deterministic algorithms.

Another improvement comes from the remark that the algorithms derived from AlgToDiff and AlgToDiffP do not return operators of order and degree as small as predicted by Thms. 2 and 3. The reason for this is technical: we need the values B_X, B_∂ it uses in order to prove that the operator constructed by Padé–Hermite approximation cancels φ and not merely its truncation. Replacing arguments B_X, B_∂ by the better bounds from Thms. 2 and 3 we obtain heuristic modifications AlgToDiffP$(P, 3D_X D_Y + 6D_Y, 6D_Y, \partial)$ and AlgToDiffP(P, B, B, ∂) with B as in Thm. 3. So far, we do not have a proof that these heuristics always return a correct output, but they behave very well in practice (see §4).

3.5 Unrolling the Recurrence

The operators from Thm. 4 let compute series expansions of algebraic functions asymptotically faster than Newton iteration. The following quantifies the improvement.

THEOREM 5. *Let $\psi \in \mathbb{L}[[X]]$ be such that $P(X, \psi(X)) = 0$, where $\mathbb{L} = \mathbb{K}(\alpha)$ is a finite extension of $\mathbb{K} \subset \mathbb{C}$. Then, under (H'), the first N terms of ψ can be computed in $\mathcal{O}(ND_X D_Y([\mathbb{L} : \mathbb{K}] + \mathsf{M}(D_Y)/D_Y))$ ops in $\mathbb{K}$, as $N \to \infty$.*

In the case when $\mathbb{L} = \mathbb{K}$, the complexity we get is therefore linear in N and *quadratic* in the degree if fast multiplication is used. A better complexity $O(ND)$ was announced by Chudnovsky & Chudnovsky in [8, 9]. While the dependency in N is correct, they seem to have overlooked the degree growth of the coefficients of the differential resolvent, that gives a cubic dependency in D for their algorithm.

Our algorithm is presented in the proof below.

PROOF. The first part of the computation is independent of N. We begin by calling AlgToDiff$(P, B_X, B_\partial, \partial)$ with $B_X = 5D_X D_Y$, $B_\partial = 5D_Y$ and $\partial = \theta_X$. We then convert the operator into a recurrence of the form

$$r_0(n)u_n + r_1(n)u_{n+1} + \cdots + r_{B_X}(n)u_{n+B_X} = 0, \quad \text{for } n \geq 0$$

with coefficients $r_i(X) \in \mathbb{K}[X]$ of degree at most B_∂. Let ρ be the largest nonnegative integer root of the polynomial $r_{B_X}(X)$, or -1 if no such root exists. Thus for any $n > \rho$, u_{n+B_X} can be computed from the previous ones using the recurrence. Note that we also have at our disposal the initial segment $\sum_{n=0}^{B_X-1} u_n X^n$ of $\varphi(X)$. For practical purposes ρ can be replaced by an upper bound like a Cauchy bound (see, e.g. [24, Lemma 6.7]). If $\rho \neq -1$, we use Newton iteration to extend this initial segment up to the index $\rho + B_X$ if necessary. Using the minimal polynomial of α, we then deduce the corresponding segment of ψ.

Only the rest of the computation depends on N. The recurrence is unrolled for n from $\rho+1$ to $N - B_X - 1$. Since the coefficients r_i belong to $\mathbb{K}$, so do the values $r_i(n)$ and thus each new coefficient u_n of ψ can be obtained using $\mathcal{O}(B_X[\mathbb{L} : \mathbb{K}]) \in \mathcal{O}(D_X D_Y[\mathbb{L} : \mathbb{K}])$ operations in $\mathbb{K}$, provided that the values taken by the coefficients $r_i(X)$ at the points $\rho + 1, \ldots, N - B_X - 1$ have been computed and stored.

This computation can be done separately for each coefficient $r_i(X)$ using fast multipoint evaluation on $\rho + 1, \rho + 2, \ldots, \rho + B_\partial$ and then fast extrapolation on arithmetic sequences. The cost of all the fast multipoint evaluation steps is $\mathcal{O}(B_X \mathsf{M}(B_\partial) \log B_\partial)$ and that of all the fast extrapolation steps is $\mathcal{O}(B_X N \mathsf{M}(B_\partial)/B_\partial)$, see e.g. [23, 6]. $\square$

We note that Hypothesis (H') is not strictly necessary in practice: the same method handles singular cases as well, provided power series solutions exist at the point of interest.

4. EXPERIMENTS

4.1 Timings

We have implemented the algorithm of §2.3 in Magma V2.11-14 [5]. In Table 1 we report on our experiments with random dense bivariate polynomials over the prime field with 9973 elements on a Pentium (M) 1.8GHz processor. Column "ser." indicates the time (in sec.) of our algorithm, while column "rat." is for Magma's DifferentialOperator, where we removed the characteristic 0 requirement. The latter routine implements Cockle's algorithm in $\mathbb{K}(X)$ and has been contributed by A. van der Waall. Column "η" indicates the value of η and column "$\deg_X(M)$" gives the actual degree in X of the minimal operator. We observe that Cockle's algorithm is always faster when run over the series and that there is still room for improvement on our degree bound η.

In Table 2, we compare the practical behaviour of our heuristic algorithm from §3.4 for algebraic series defined by polynomials of bi-degree $(1, D_Y)$ with that of Newton's iteration (column **Newton**). Column **AlgToRec** contains the cumulated timings of the heuristic AlgToDiffP(P, B, B, ∂) (with B as in Th. 3) and of the conversion of a differential operator into a recurrence. Column N gives the number of terms of the expansion, while column **unroll** shows the timings for unrolling the recurrence. One motivation for the computation when $D_X = 1$ is its application in univariate polynomial root finding by homotopy in [10, § 10].

These experiments show that there is an initial cost due to the size of the recurrence being constructed. Once the number of desired coefficients is high enough compared to this size, then the method is very effective.

4.2 On the Optimality of the Bounds

Conjectures. Intensive experiments suggest that the degree of the differential resolvent is bounded by $D(D^2 - 5D/2 + 5/2)$ for polynomials of total degree D and $D(2D^2 - 3D + 3)$ for polynomials of bi-degree (D, D), within a factor 2 of the bound of Thm. 1. Also, the bound from Thm. 3 seems only slightly pessimistic: for a polynomial of total degree D, there often exists a recurrence of order $D^2 - 2$ with coefficients of degree at most $D^2 - 1$. For a polynomial of bi-degree (D_X, D_Y), experiments suggest the bound $2D_X D_Y - 2 - (D_X - D_Y)$ if $D_Y > 1$ and $D_X + 1$ if $D_Y = 1$.

Table 1: Cockle's algorithm

(D_X, D_Y)	ser.	rat.	η	$\deg_X(M)$
(1,1)	0.002	0.002	2	2
(2,2)	0.003	0.004	17	10
(3,3)	0.02	0.03	69	36
(4,4)	0.10	0.16	182	92
(5,5)	0.47	0.98	380	190
(6,6)	1.86	4.56	687	342
(7,7)	5.80	16.5	1127	560
(8,8)	15.5	49.9	1724	856
(9,9)	38.0	138	2502	1242
(10,10)	72.7	340	3485	1730

Table 2: Newton's algorithm vs. AlgToDiffP

(D_X, D_Y)	N	AlgToRec	unroll	Newton
(1,8)	2^{10}	6.66	0.22	1.75
(1,8)	2^{11}	6.66	0.44	3.91
(1,8)	2^{12}	6.66	0.89	8.83
(1,8)	2^{13}	6.66	1.8	18.84
(1,8)	2^{14}	6.66	3.67	42.53
(1,8)	2^{15}	6.66	7.38	94.99
(1,9)	2^{10}	10.54	0.26	2.14
(1,9)	2^{11}	10.54	0.44	4.35
(1,9)	2^{12}	10.54	1.07	9.31
(1,9)	2^{13}	10.54	2.17	20.64
(1,9)	2^{14}	10.54	4.36	43.44
(1,9)	2^{15}	10.54	8.75	97.64

Lower Bound. We note also that $D(D-1)$ is the degree of
the discriminant of a generic polynomial P of total degree D.
Generically, this discriminant is square-free. The differential
equation satisfied by all solutions of P has to be singular
at the roots of the discriminant and therefore its leading
coefficient cannot have a smaller degree. Thus our bound
on the order of the recurrence is asymptotically optimal.

An example illustrating $D(D-1)$ as a lower bound is
provided by $P(X,Y) = Y^D - Y + X^D$. Then $\alpha = X^D + X^{D^2} + \mathcal{O}(X^{D^2})$ is a root of P seen in $\mathbb{Q}[[X]][Y]$. It has
a sequence of $D(D-1) - 1$ zero coefficients before a non-
zero one. Thus, the order of the recurrence satisfied by the
coefficients of α is at least $D(D-1)$.

Acknowledgments

5. REFERENCES

[1] N. H. Abel. *Œuvres complètes. Tome II.* Éd. J. Gabay,
1992. Reprint of the 2nd (1881) ed. Available at
http://gallica.bnf.fr.

[2] G. Almkvist and D. Zeilberger. The method of
differentiating under the integral sign. *J. Symbolic
Comput.*, 10(6):571–591, 1990.

[3] M. Apagodu and D. Zeilberger. Multi-variable
Zeilberger and Almkvist-Zeilberger algorithms and the
sharpening of Wilf-Zeilberger theory. *Adv. in Appl.
Math.*, 37(2):139–152, 2006.

[4] D. Bini and V. Y. Pan. *Polynomial and matrix
computations. Vol. 1.* Birkhäuser, 1994.

[5] W. Bosma, J. Cannon, and C. Playoust. The Magma
algebra system I: The user language. *J. Symbolic
Comput.*, 24(3-4):235–265, 1997.

[6] A. Bostan, P. Gaudry, and É. Schost. Linear
recurrences with polynomial coefficients and
computation of the Cartier-Manin operator on
hyperelliptic curves. *SIAM J. Comput.*,
36(6):1777–1806, 2007.

[7] G. Chèze and G. Lecerf. Lifting and recombination
techniques for absolute factorization. *J. Complexity*,
doi:10.1016/j.jco.2007.01.008, in press, 2007.

[8] D. V. Chudnovsky and G. V. Chudnovsky. On
expansion of algebraic functions in power and Puiseux
series. I. *J. Complexity*, 2(4):271–294, 1986.

[9] D. V. Chudnovsky and G. V. Chudnovsky. On
expansion of algebraic functions in power and Puiseux
series. II. *J. Complexity*, 3(1):1–25, 1987.

[10] D. V. Chudnovsky and G. V. Chudnovsky. Computer
algebra in the service of mathematical physics and
number theory. In *Computers in mathematics
(Stanford, CA, 1986)*, 109–232, 1990. Dekker.

[11] J. Cockle. On transcendental and algebraic solution.
Philosophical Magazine, XXI:379–383, 1861.

[12] L. Comtet. Calcul pratique des coefficients de Taylor
d'une fonction algébrique. *Enseignement Math. (2)*,
10:267–270, 1964.

[13] O. Cormier, M. F. Singer, B. M. Trager, and
F. Ulmer. Linear differential operators for polynomial
equations. *J. Symbolic Comput.*, 34(5):355–398, 2002.

[14] J. Gray. *Linear differential equations and group theory
from Riemann to Poincaré.* Birkhäuser, 1986.

[15] R. Harley. On the theory of the transcendental
solution of algebraic equations. *Quart. J. of Pure and
Applied Math*, 5:337–361, 1862.

[16] J. M. Nahay. *Linear differential resolvents.* PhD
thesis, Rutgers University, 2000.

[17] J. M. Nahay. Linear relations among algebraic
solutions of differential equations. *J. Differential
Equations*, 191(2):323–347, 2003.

[18] J. M. Nahay. Differential resolvents of minimal order
and weight. *Int. J. Math. and Math. Sciences*,
2004(53-56):2867–2893, 2004.

[19] A. Storjohann. Notes on computing minimal
approximant bases. Dagstuhl preprint, 2006.

[20] N. Takayama. An approach to the zero recognition
problem by Buchberger algorithm. *J. Symbolic
Comput.*, 14(2-3):265–282, 1992.

[21] J. Tannery. Propriétés des intégrales des équations
différentielles linéaires à coefficients variables. *Annales
scientifiques de l'ENS Sér. 2*, 4:113–182, 1875.

[22] H. Tsai. Weyl closure of a linear differential operator.
J. Symbolic Comput., 29(4-5):747–775, 2000.

[23] J. von zur Gathen and J. Gerhard. *Modern computer
algebra.* Cambridge University Press, 2nd ed., 2003.

[24] C. K. Yap. *Fundamental Problems in Algorithmic
Algebra.* Oxford University Press, New York, 2000.

[25] D. Zeilberger. The method of creative telescoping. *J.
Symbolic Comput.*, 11(3):195–204, 1991.