A Fast Algorithm for Computing the p-Curvature
Alin Bostan, Xavier Caruso, Éric Schost

To cite this version:
Alin Bostan, Xavier Caruso, Éric Schost. A Fast Algorithm for Computing the p-Curvature. ISSAC 2015, Jul 2015, Bath, United Kingdom. pp.69-76, 10.1145/2755996.2756674 . hal-01164471

HAL Id: hal-01164471
https://hal.archives-ouvertes.fr/hal-01164471
Submitted on 17 Jun 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
A Fast Algorithm for Computing the p-Curvature

Alin Bostan
Inria (France)
alin.bostan@inria.fr

Xavier Caruso
Université Rennes 1
xavier.caruso@normalesup.org

Éric Schost
Western University
eshost@uwo.ca

ABSTRACT

We design an algorithm for computing the p-curvature of a differential system in positive characteristic p. For a system of dimension r with coefficients of degree at most d, its complexity is $O((p^r dr)^3)$ operations in the ground field, whereas the size of the output is about $p^{2r}$. Our algorithm is then quasi-optimal assuming that matrix multiplication is (i.e. $\omega = 2$). The main theoretical input we use is the existence of a well-suited ring of series with divided powers for functions over $k$. It is the so-called p-curvature of a differential system $Y' = AY$. More precisely, at an ordinary point, the system $Y' = AY$ admits a fundamental matrix of power series solutions in $k[[x]]$ if and only if the p-curvature $A_p$ vanishes. In this case, the system $Y' = AY$ even admits a fundamental matrix of solutions which are rational functions in $k(x)$. More generally, the dimension of the kernel of $A_p$ is equal to the dimension of the space of rational function solutions of $Y' = AY$.

The primary importance of the notion of p-curvature relies in its occurrence in one of the versions of the celebrated Grothendieck–Katz conjecture [19, 20, 12, 30]. This conjecture, first formulated by Alexandre Grothendieck in the late 1960s, is a local-global principle for linear differential systems, which states that a linear differential system with rational function coefficients over a function field admits a fundamental matrix of algebraic solutions if and only if its p-curvatures vanish for almost all primes p.

In computer algebra, p-curvature has been introduced by van der Put [22, 23], who popularized it as a tool for factoring differential operators in characteristic p. Cluzeau [13] generalized the approach to the decomposition of differential systems over $k(x)$. The p-curvature has also been used by Cluzeau and van Hoeij [14] as an algorithmic filter for computing exponential solutions of differential operators in characteristic zero.

Computing efficiently the p-curvature is in itself a challenging problem, especially for large values of p. Our initial motivation for studying this question emerged from concrete applications, in lattice path combinatorics [6, 7] and in statistical physics [3]. In this article, we address the question of the computation of $A_p$ in good complexity, with respect to three parameters: the dimension r of the system $Y' = AY$, the maximum degree d of the rational function entries of A, and the characteristic p of the ground field. In terms of these quantities, the arithmetical size of $A_p$ is generically proportional to $pdr^2$ if $r > 1$.

Previous work. Cluzeau [13, Prop. 3.2] observed that the direct algorithm based on recurrence (1) has complexity $O((p^r dr)^3)$, where $\omega$ is the matrix multiplication exponent and the soft-O notation $O(\cdot)$ hides polylogarithmic factors. Compared to the size of the p-curvature, this cost is good with respect to r and d, but not to p. The first subquadratic algorithm in p, of complexity $O(p^{1+\omega/3})$, was designed in [9, 3]. In some special cases, additional partial results were obtained in [9], notably an algorithm of quasi-linear cost $O(p)$ for certain systems of order $r = 2$. However, the question of designing a general algorithm for computing $A_p$ with quasi-linear complexity in p remained open. In a related, but

1. INTRODUCTION

We study in this article algorithmic questions related to linear differential systems in positive characteristic. Let $k$ be an arbitrary field of characteristic $p$, and $A$ be an $r \times r$ matrix with entries in the field $k(x)$ of rational functions over $k$. A simple-to-define, yet very important object attached to the differential system $Y' = AY$ is its so-called p-curvature. It is the p-th iterate $\partial_A^n$ of the map $\partial_A : k(x)^r \rightarrow k(x)^r$ that sends v to $v' - Av$. It turns out that it is $k(x)$-linear.

We design an algorithm for computing the p-curvature of $Y' = AY$. The above recurrence yields an algorithm for computing it, sometimes referred to as Katz’s algorithm.

The p-curvature is related to solutions; it measures to what extent the usual Cauchy–Lipschitz theorem applies in characteristic $p$. More precisely, at an ordinary point, the system $Y' = AY$ admits a fundamental matrix of power series solutions in $k[[x]]$ if and only if the p-curvature $A_p$ vanishes. In this case, the system $Y' = AY$ even admits a fundamental matrix of solutions which are rational functions in $k(x)$. More generally, the dimension of the kernel of $A_p$ is equal to the dimension of the space of rational function solutions of $Y' = AY$.

The primary importance of the notion of p-curvature relies in its occurrence in one of the versions of the celebrated Grothendieck–Katz conjecture [19, 20, 12, 30]. This conjecture, first formulated by Alexandre Grothendieck in the late 1960s, is a local-global principle for linear differential systems, which states that a linear differential system with rational function coefficients over a function field admits a fundamental matrix of algebraic solutions if and only if its p-curvatures vanish for almost all primes p.

In computer algebra, p-curvature has been introduced by van der Put [22, 23], who popularized it as a tool for factoring differential operators in characteristic p. Cluzeau [13] generalized the approach to the decomposition of differential systems over $k(x)$. The p-curvature has also been used by Cluzeau and van Hoeij [14] as an algorithmic filter for computing exponential solutions of differential operators in characteristic zero.

Computing efficiently the p-curvature is in itself a challenging problem, especially for large values of p. Our initial motivation for studying this question emerged from concrete applications, in lattice path combinatorics [6, 7] and in statistical physics [3]. In this article, we address the question of the computation of $A_p$ in good complexity, with respect to three parameters: the dimension r of the system $Y' = AY$, the maximum degree d of the rational function entries of A, and the characteristic p of the ground field. In terms of these quantities, the arithmetical size of $A_p$ is generically proportional to $pdr^2$ if $r > 1$.

Previous work. Cluzeau [13, Prop. 3.2] observed that the direct algorithm based on recurrence (1) has complexity $O((p^r dr)^3)$, where $\omega$ is the matrix multiplication exponent and the soft-O notation $O(\cdot)$ hides polylogarithmic factors. Compared to the size of the p-curvature, this cost is good with respect to r and d, but not to p. The first subquadratic algorithm in p, of complexity $O(p^{1+\omega/3})$, was designed in [9, 3]. In some special cases, additional partial results were obtained in [9], notably an algorithm of quasi-linear cost $O(p)$ for certain systems of order $r = 2$. However, the question of designing a general algorithm for computing $A_p$ with quasi-linear complexity in p remained open. In a related, but
different direction, the article [4] proposed an algorithm for computing the characteristic polynomial of the $p$-curvature in time essentially linear in $\sqrt{p}$, without computing $A_p$ itself.

**Contribution.** We prove that the $p$-curvature $A_p$ can be computed in quasi-linear time with respect to $p$. More precisely, our main result (Theorem 4.2) states that $O(pdra)$ operations in $k$ are sufficient for this task. This complexity result is quasi-optimal not only with respect to the main parameter $p$, but also to $d$; with respect to the dimension $r$, it is as optimal as matrix multiplication. Moreover the algorithm we obtain is highly parallelizable by design. The key tools underlying the proof of Theorem 4.2 are the notion of divided power rings in characteristic $p$, and a new formula for the $p$-curvature (Propositions 4.3 and 4.4) in terms of divided power series. Crucial ingredients are the fact that a Cauchy–Lipschitz theorem for differential systems holds over divided power rings (Proposition 3.4) and the fact that Newton iteration can be used to efficiently compute (truncations of) fundamental matrices of divided power solutions.

**Structure of the paper.** In Section 2, we recall the main theoretical properties of the basic objects used in this article. Section 3 is devoted to the existence and the computation of solutions of differential systems in divided power rings. In Section 4, we move to the main objective of the article, the computation of the $p$-curvature: after relating $A_p$ to the framework of divided powers, we describe our main algorithm for $A_p$, of complexity $O(pdra)$. We conclude in Section 5 by describing the implementation of our algorithm and some benchmarks.

**Complexity measures.** Throughout this article, we estimate the cost of our algorithms by counting arithmetic operations in the base ring or field at unit cost.

We use standard complexity notations. The letter $\omega$ refers to a feasible exponent for matrix multiplication (i.e. there exists an algorithm for multiplying $n \times n$ matrices over a ring $\mathcal{A}$ with at most $O(n^\omega)$ operations in $\mathcal{A}$; the best known bound is $\omega < 2.3729$ from [15]). The soft-$O$ notation $O(\cdot)$ indicates that polylogarithmic factors are omitted; in particular, we will use the fact that many arithmetic operations on univariate polynomials of degree $d$ can be done in $O(d)$ operations: addition, multiplication, Chinese remaindering, etc. the key to these results being fast polynomial multiplication [27, 26, 11, 18]. A general reference for these questions is [17].

2. **THEORETICAL SETTING**

We introduce and briefly recall the main properties of the theoretical objects we are going to use in this article. All the material presented in this section is classical; a general reference is [24].

**Definitions and notations.** Let $\mathcal{A}$ be a commutative ring with unit. We recall that a **derivation** on $\mathcal{A}$ is an additive map $\partial : \mathcal{A} \rightarrow \mathcal{A}$ satisfying the Leibniz rule ($fg)' = f'g + fg'$ for all $f, g \in \mathcal{A}$. The image $f'$ of $f$ under the derivation is called the **derivative** of $f$. From now on, we assume that $\mathcal{A}$ is equipped with a derivation. A **differential system** with coefficients in $\mathcal{A}$ is an equation of the form $Y' = AY$ where $A$ is a given $r \times r$ matrix with coefficients in $\mathcal{A}$ (for a certain positive integer $r$), the unknown $Y$ is a column vector of length $r$ and $Y'$ denotes the vector obtained from $Y$ by taking the derivative component-wise. The integer $r$ is called the dimension of the system. We recall briefly that a linear differential equation:

$$a_0 y^{(r)} + \cdots + a_1 y' + a_0 y = 0 \quad (\text{with } a_i \in \mathcal{A})$$

(2)

can be viewed as a particular case of a differential system. Indeed, defining the companion matrix

$$C = \begin{pmatrix} 1 & -a_1 & \cdots & -a_{r-1} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & 1 \end{pmatrix}$$

(3)

and $A = C$, the solutions of the system $Y' = AY$ are exactly the vectors of the form $1(y, y', \ldots, y^{(r-1)})$ where $y$ is a solution of (2). In this correspondence, the order of the differential equation agrees with the dimension of the associated differential system.

**Differential modules.** A differential module over $\mathcal{A}$ is a pair $(M, \partial)$ where $M$ is an $\mathcal{A}$-module and $\partial : M \rightarrow M$ is an additive map satisfying a Leibniz-like rule, which is:

$$\forall f \in \mathcal{A}, \forall x \in M, \quad \partial(fx) = f' \cdot x + f \cdot \partial(x).$$

(4)

There exists a canonical one-to-one correspondence between differential systems and differential modules $(M, \partial)$ for which $M = \mathcal{A}^n$ for some $n$: to a differential system $Y' = AY$ of dimension $n$, we attach the differential module $(\mathcal{A}^n, \partial_A)$ where $\partial_A : \mathcal{A}^n \rightarrow \mathcal{A}^n$ is the function mapping $X$ to $X' - AX$. Under this correspondence, the solutions of $Y' = AY$ are exactly vectors in the kernel of $\partial_A$.

To a differential equation as (2), one can associate the differential operator $L = a_0 \partial^r + a_{r-1} \partial^{r-1} + \cdots + a_1 \partial + a_0$; it lies in the non-commutative ring $\mathcal{A}(\partial)$, endowed with the usual addition of polynomials and a multiplication ruled by the relation $\partial f = f \cdot \partial + f'$ for all $f \in \mathcal{A}$ (note that, as often in the literature, we are using $\partial$ to denote either the structure map of a differential module, and a non-commutative indeterminate).

Then, if $a_r$ is a unit in $\mathcal{A}$, one can further associate to $L$ the quotient $\mathcal{A}(\partial)/\mathcal{A}(\partial)[L] \simeq \mathcal{A}^n$. The differential structure inherited from $\mathcal{A}(\partial)$ makes it a differential module with structure map $X \mapsto X' - CX$, where $C$ is the companion matrix defined above; in other words, this is the module $(\mathcal{A}(\partial)/\mathcal{A}(\partial)[L], \partial_{C-})$, with the previous notation.

**Scalar extension.** Let $\mathcal{A}$ and $\mathcal{B}$ be two rings equipped with derivations and let $\varphi : \mathcal{A} \rightarrow \mathcal{B}$ be a ring homomorphism commuting with derivation. From a given differential system $Y' = AY$ with coefficients in $\mathcal{A}$, one can build a differential system over $\mathcal{B}$ by applying $\varphi$: it is $Y' = \varphi(A)Y$, where $\varphi(A)$ is the matrix obtained from $A$ by applying $\varphi$ entry-wise.

This operation admits an analogue at the level of differential modules: to a differential module $(M, \partial)$ over $\mathcal{A}$, we attach the differential module $(M_\mathcal{B}, \partial_\mathcal{B})$ over $\mathcal{B}$ where $M_\mathcal{B} = \mathcal{B} \otimes_{\mathcal{A}} M$ and $\partial_\mathcal{B} : M_\mathcal{B} \rightarrow M_\mathcal{B}$ is defined by:

$$\forall f \in \mathcal{B}, \forall x \in M, \quad \partial_\mathcal{B}(f \otimes x) = f' \otimes x + f \otimes \partial(x).$$

It is easily seen that if $(M, \partial)$ is a differential module associated to the system $Y' = AY$ then $(M_\mathcal{B}, \partial_\mathcal{B})$ is that associated to the system $Y' = \varphi(A)Y$.

**The $p$-curvature.** Let $k$ be any field of characteristic $p$. We assume here that $\mathcal{A}$ is the field $k(x)$ — consisting of rational functions over $k$ — equipped with the standard derivation.
3. SERIES WITH DIVIDED POWERS

In this section, we let \( \ell \) be a ring in which \( p \) vanishes. We recall the definition of the divided power ring over \( \ell \) and its main properties — mainly, a Cauchy–Lipschitz theorem that will allow us to compute solutions of differential systems. We show how some approaches that are well-known for power series solutions carry over without significant changes in this context. Most results in this section are not new; those from §3.1 and §3.2 are implicitly contained in [1, 2], while the theoretical basis of §3.3 is similar to [21].

3.1 The ring \( \ell[[t]]^{dp} \)

Let \( \ell[[t]]^{dp} \) be the ring of formal series of the form:

\[
    f = a_0 + a_1 \gamma_1(t) + a_2 \gamma_2(t) + \cdots + a_i \gamma_i(t) + \cdots
\]

(5)

where the \( a_i \)'s are elements of \( \ell \) and each \( \gamma_i(t) \) is a symbol which should be thought of as \( t^i/p \). The multiplication on \( \ell[[t]]^{dp} \) is defined by the rule \( \gamma_i(t) \cdot \gamma_j(t) = (i+j) \cdot \gamma_{i+j}(t) \).

Remark 3.1. The ring \( \ell[[t]]^{dp} \) is not the PD-envelope in the sense of [1, 2] of \( \ell[[t]] \) with respect to the ideal \( (t) \) but its completion for the topology defined by the divided powers makes taking the completion is essential to have an analogue of the Cauchy–Lipschitz Theorem (cf Proposition 3.4).

Invertible elements of \( \ell[[t]]^{dp} \) are easily described: they are exactly those for which the “constant” coefficient \( a_0 \) is invertible in \( \ell \). The ring \( \ell[[t]]^{dp} \) is moreover endowed with a derivation defined by \( f' = \sum_{i=0}^{\infty} a_{i+1} \gamma_i(t) \) for \( f = \sum_{i=0}^{\infty} a_i \gamma_i(t) \). It then makes sense to consider differential systems over \( \ell[[t]]^{dp} \). A significant difference with power series is the existence of an integral operator: it maps \( f \) as above to \( \int f = \sum_{i=0}^{\infty} a_i \gamma_{i+1}(t) \) and satisfies \((\int f)' = f \) for all \( f \).

Divided power ideals. For all positive integers \( N \), we denote by \( \ell[[t]]^{dp}_N \) the ideal of \( \ell[[t]]^{dp} \) consisting of the series of the form \( \sum_{i \geq N} a_i \gamma_i(t) \). The quotient \( \ell[[t]]^{dp}/\ell[[t]]^{dp}_N \) is a free \( \ell \)-module of rank \( N \) and a basis of it is \( \{1, \gamma_1(t), \ldots, \gamma_{N-1}(t)\} \). In particular, for \( N = 1 \), the quotient \( \ell[[t]]^{dp}/\ell[[t]]^{dp}_1 \) is isomorphic to \( \ell \) in the sequel, we shall denote by \( f(0) \in \ell \) the reduction of an element \( f \in \ell[[t]]^{dp} \) modulo \( \ell[[t]]^{dp}_1 \). On the writing (5), it is nothing but the constant coefficient \( a_0 \) in the expansion of \( f \).

We draw the reader’s attention to the fact that \( \ell[[t]]^{dp}_N \) is not stable under derivation, so the quotients \( \ell[[t]]^{dp}/\ell[[t]]^{dp}_N \) do not inherit a derivation.

Relationship with \( \ell[[t]] \). There exists a natural map \( \varepsilon : \ell[[t]]^{dp} \to \ell[[t]]^{dp} \) taking a polynomial \( \sum \gamma_i(t) a_i \) to \( \sum_{i=0}^{\infty} a_i \gamma_i(t) \). The latter sum stops at \( i = p-1 \) because \( ! \) becomes divisible by \( p \) after that. Clearly, the kernel of \( \varepsilon \) is the principal ideal generated by \( t^p \). Hence \( \varepsilon \) factors through \( \ell[[t]]^{dp}/t^p \) as follows:

\[
    \ell[[t]]^{dp} \xrightarrow{\varepsilon} \ell[[t]]^{dp}/t^p \xrightarrow{\ell[[t]]^{dp}/t^p \to \ell[[t]]^{dp}} \ell[[t]]^{dp}/t^p
\]

(6)

where \( pr \) is the canonical projection taking a polynomial to its reduction modulo \( t^p \). We observe moreover that the ideal \( t^p \) \( \ell[[t]] \) is stable under derivation and, consequently, that the quotient ring \( \ell[[t]]^{dp}/t^p \) inherits a derivation. Furthermore, the two mappings in (6) commute with the derivation.

3.2 Computations with divided powers

It turns out that the \( \gamma_n(t) \)'s can all be expressed in terms of only few of them, resulting in a more flexible description of the ring \( \ell[[t]]^{dp} \). To make this precise, we set \( t_i = \gamma_{pi}(t) \) and first observe that \( t_i^n = n! \gamma_{pi}(t) \) for all \( i \) and \( n \); this is proved by induction on \( n \), using the equalities

\[
    t_i^{n+1} = n! \gamma_{pi}(t) \gamma_{pi}(t) = n! \cdot \gamma_{pi}(t) \gamma_{pi+1}(t)
\]

since Lucas’ Theorem shows that \( \gamma_{pi+1}(t) \equiv 1 \mod (p) \). The second equality then follows from the relations \( t_i^n = n! \gamma_{pi+1}(t) \).

Proposition 3.2. Let \( n \) be a positive integer and \( n = \sum_{i=0}^{s} n_i p^i \) its writing in basis \( p \). Then:

\[
    \gamma_n(t) = \gamma_n(t_0) \gamma_{n_1}(t_1) \cdots \gamma_{n_s}(t_s) = \frac{t_0^{n_0}}{n_0!} \cdot \frac{t_1^{n_1}}{n_1!} \cdots \frac{t_s^{n_s}}{n_s!}
\]

Proof. The first equality is proved by induction on \( n \) using the fact that if \( n = a + bp \) with \( 0 \leq a < p \), then Lucas’ Theorem shows that \( \gamma_{n+bp} = \gamma_n \), since \( \gamma_{n+bp} \equiv 1 \mod (p) \). The second equality then follows from the additional rules \( t_i^0 = 0 \) for all \( i \).

A corollary of the above proposition is that elements of \( \ell[[t]]^{dp} \) can be alternatively described as infinite sums of monomials \( a_{n_0} \cdots a_{n_s} \cdot t_0^{n_0} \cdots t_s^{n_s} \) where the \( n_i \)'s are integers in the range \( [0, 1) \) and the coefficient \( a_{n_0} \cdots a_{n_s} \) lies in \( \ell \). The product in \( \ell[[t]]^{dp} \) is then the usual product of series subject to the additional rules \( t_i^0 = 0 \) for all \( i \).

More precisely, restricting ourselves to some given precision of the form \( N = np^s \), we deduce from the above discussion the following corollary.

Corollary 3.3. For \( N = np^s \), with \( s \in \mathbb{N} \) and \( n \in \{1, \ldots, p^s\} \), there is a canonical isomorphism of \( \ell \)-algebras:

\[
    \ell[[t]]^{dp}/\ell[[t]]^{dp}_{\leq N} \simeq \ell[[t_0, \ldots, t_s]/(t_0^{p^0}, \ldots, t_s^{p^s})]
\]

For instance, if we take \( s = 0 \) and \( N = n \in \{1, \ldots, p\} \), we obtain the isomorphism \( \ell[[t]]^{dp}/\ell[[t]]^{dp}_{\leq N} \simeq \ell[[t]]^{dp}/t^N \).

In terms of complexity, the change of bases between left- and right-hand sides can both be done in \( O(N) \) operations in \( \ell \); all the factorials we need can be computed once and for all for \( O(\min(N, p)) \) operations; then each monomial conversion takes \( O(\gamma(N)) \) operations, for a total of \( O(N \log(N)) = O(N) \).

The previous corollary is useful in order to devise a multiplication algorithm for divided powers, since it reduces this question to multivariate power series multiplication (addition takes linear time in both bases). To multiply in \( \ell[[t_0, \ldots, t_s]/(t_0^{p^0}, \ldots, t_s^{p^s})] \), one can use a direct algorithm: multiply and discard unwanted terms. Using for instance
A nice feature of the ring $\ell[[t]]^{dp}$ — which does not hold for $\ell[[t]]$ — notably — is the existence of an analogue of the classical Cauchy–Lipschitz theorem. This property will have a fundamental importance for the purpose of our paper; see for instance [21, Proposition 4.2] for similar considerations.

**Proposition 3.4.** Let $Y' = AY$ be a differential system of dimension $r$ with coefficients in $\ell[[t]]^{dp}$. For all initial data $V \in \ell^r$ (considered as a column vector) the following Cauchy problem has a unique solution in $\ell[[t]]^{dp}$:

$$
\begin{align*}
Y' &= A \cdot Y \\
Y(0) &= V.
\end{align*}
$$

**Proof.** Let us write the expansions of $A$ and $Y$:

$$
A = \sum_{i=0}^{\infty} A_i \gamma_i(t) \quad \text{and} \quad Y = \sum_{i=0}^{\infty} Y_i \gamma_i(t)
$$

where the $A_i$'s and $Y_i$'s have coefficients in $\ell$. The Cauchy problem translates to $Y_0 = V$ and $Y_{i+1} = \sum_{i=0}^{\infty} \left( t^i \cdot A_i \right) \cdot Y_i$. It is now clear that it has a unique solution.

Of course, Proposition 3.4 extends readily to the case where the initial data $V$ is any matrix having $r$ rows. In particular, taking $V = I_r$ (the identity matrix of size $r$), we find that there exists a unique $r \times r$ matrix $Y$ with coefficients in $\ell[[t]]^{dp}$ such that $Y(0) = I_r$ and $Y' = A \cdot Y$. This matrix is often called a fundamental system of solutions.

**Finding solutions using Newton iteration.** In characteristic zero, it is possible to compute power series solutions of a differential system such as $Y' = A \cdot Y$ using Newton iteration; an algorithm for this is presented on [5, Fig. 1].

One can use this algorithm to compute a fundamental system of solutions in our context. For this, we first need to introduce two notations. Given an element $f \in \ell[[t]]^{dp}$ written as $f = \sum_{m=0}^{\infty} a_m \gamma_m(t)$ together with an integer $m$, we set $[f]^m = \sum_{i=0}^{m} a_i \gamma_i(t)$. Similarly, if $M$ is a matrix with coefficients in $\ell[[t]]^{dp}$, we define $[M]^m$ and $\int M$ by applying the corresponding operations entry-wise.

**Algorithm fundamental_solutions**

**Input:** a differential system $Y' = AY$, an integer $N$

**Output:** the fund. system of solutions modulo $\ell[[t]]^{dp}$

1. $Y = I_r + t \cdot A(0); \ Z = I_r; \ m = 2$
2. while $m \leq N/2$
3. $Z = Z + \left( Z(I_r - Y Z) \right)^m$
4. $Y = Y - \left( Y \left( \int Z \cdot (Y' - [A]^{2m-1} Y) \right) \right)^{2m}$
5. $m = 2m$
6. return $Y$

Correction is proved as in the classical case [5, Lemma 1].

Let us take $n \in \{2, \ldots, p\}$ and $s \in \mathbb{N}$ such that $n - 1$ is the last digit of $N$ written in basis $p$, and $s$ the corresponding exponent; then, we have $(n-1)p^s \leq N < np^s$. Since we are only interested in costs up to logarithmic factors, we may assume that we do all operations at precision $np^s$ (a better analysis would take into account the fact that the precision grows quadratically).

By Corollary 3.3 and the discussion that follows, arithmetic operations in $\ell[[t]]^{dp}/\ell[[t]]^{dp}$ take time $O(2^{dpN}N^r)$, the case for differentiation and integration, in view of the formulas given in the previous subsection; truncation is free. The total complexity of Algorithm fundamental_solutions is therefore $O(2^{dpN}N^{N^r})$ operations in $\ell$, where $r$ is the dimension of the differential system. If $N = p^{O(1)}$, which is what we need later on, this is $O(N^r)$.

**The case of differential operators.** We now consider the case of the differential system associated to a differential operator $L = a_0 t^{r} + \cdots + a_r t^{0}$, where $\ell$ is the differential system. The Cauchy problem translates to $V_{n+r}$, together with an integer $r$, we get $\xi_{n+r} = \alpha_{r,0} \sum_{i=0}^{r} \left( \sum_{j=0}^{\infty} \alpha_{r,j} \xi_{n+i-j} \right)$.

Letting $m = i-j$, we find $\xi_{n+i} = \sum_{m=0}^{\infty} A_m(n) \xi_{n+m}$ with

$$
A_m(n) = -\frac{1}{\alpha_{r,0}} \sum_{i=0}^{r-m} \left( \sum_{j=0}^{\infty} \alpha_{r,j} \xi_{n+i-j} \right) \frac{\alpha_{i,m}}{\alpha_{r,0}(i-m)!}.
$$

and $\xi_{n+m} = n(n-1) \cdots (n-(i-m+1))$ a falling factorial. The expression above for $A_m$ is well-defined, since we assumed that $d < p$, and shows that $A_m$ is a polynomial of degree at most $d$.

From this, writing the algorithm is easy. We need two subroutines: from_falling_factorial(F), which computes the expansion on the monomial basis of a polynomial of the form $F = \sum_{d \leq n} f_d n^d$, and eval(F, N), which computes the values of a polynomial $F$ at the points $\{0, \ldots, N-1\}$. The former can be done using the divide-and-conquer algorithm of [8, Section 3] in time $O(n)$; the latter by the algorithm of [17, Chapter 10], in time $O\left(\deg(F) + N\right)$. The previous discussion leads to the algorithm solutions_operator below.

**Algorithm solutions_operator**

**Input:** a differential operator $L \in \ell[[t]]^{dp}/\ell[[t]]^{dp}$ of bidegree $(d, r)$, with $d < p$; an integer $N$

**Output:** the solutions $f_0, \ldots, f_{r-1}$ at precision $N$
1. for $m = -d, \ldots, r - 1$:
   \[
   \hat{A}_m = \sum_{0 \leq i \leq r - 1, 0 \leq j - m \leq d} a_{i,j}(t^{i-m})
   \]
   Cost: $O(d(r + d))$

2. $A_m = \text{from_falling_factorial}(\hat{A}_m)$
   Cost: $O'(d(r + d))$

3. Store $\text{eval}(A_m, N - r)$
   Cost: $O'(d(N^2)(r + d))$

4. for $i = 0, \ldots, r - 1$:
   \[
   f_i = [0, \ldots, 0, 1, \ldots, 0] \text{ (ith unit vector of length } r)
   \]
   Cost: $O(r^2)$

5. for $n = 0, \ldots, N - r - 1$:
   \[
   f_{i,n+r} = \sum_{m=-d}^{r-1} A_m(n) f_{i,n+m}
   \]
   Cost: $O((rN(r + d))$ operations in $\ell$.

6. return $f_0, \ldots, f_{r-1}$

Altogether, we obtain the following result, where we use the assumption $N > d$ to simplify slightly the cost estimate.

**Lemma 3.5.** Suppose that $p < d$. Given a positive $N > d$, the classes of $f_0, \ldots, f_{r-1}$ modulo $\ell(t)/\ell^r$ can be computed with at most $O((rN(r + d))$ operations in $\ell$.

In particular, Algorithm **solutions_operator** has a better cost than **fundamental_solutions** when $d = O(r^{p-1})$.

### 4. Computing the P-Curvature

In all this section, we work over a field $k$ of characteristic $p > 0$. We consider a differential system $Y' = AY$ of dimension $r$ and denote by $A_p$ the matrix of its $p$-curvature. We write $A = \frac{1}{p} A$, where $f_A$ is in $k[x]$ and $A$ is a matrix with polynomial entries. Let $d = \max(\deg f_A, \deg \hat{A})$, where $\deg A$ is the maximal degree of the entries of $A$. We recall (13, Prop. 3.2), [9, Lemma 1]) a bound on the size of $A_p$. The bound follows from the recurrence (1), and it is tight.

**Lemma 4.1.** The entries of the matrix $f_A^t A_p$ are all polynomials of degree at most $dp$.

The goal of this section is to prove the following theorem.

**Theorem 4.2.** There exists an algorithm (presented below) which computes the matrix of the $p$-curvature of the differential system $Y' = AY$ in $O'(pdr^2)$ operations in $k$.

It is instructive to compare this cost with the size of the output. By Lemma 4.1, the latter is an $r \times r$ matrix whose entries are rational functions whose numerator and denominator have degree $\geq pd$, so its size is roughly $pdr^2$ elements of $k$. Our result $O'(pdr^2)$ is quasi-optimal if we assume that matrix multiplication can be performed in quasi-optimal time.

### 4.1 A formula for the p-curvature

Let $A_p$ denote the matrix of the $p$-curvature of the differential system $Y' = AY$ (in the usual monomial basis). The expression of $A_p$ given at the very end of §2 is unfortunately not well-suited for fast computation. The aim of this subsection is to give an alternative formula for $A_p$ using the framework of divided powers.

In order to relate $k[x]$ and a ring $\ell(t)/\ell^p$, we pick a separable polynomial $S \in k[x]$ which is coprime with $f_A$ and set $\ell = k[x]/S$ (which is thus not necessarily a field). Let $c \in \ell$ be the class of $x$. We consider the ring homomorphism:

\[
\varphi_S : \ell[t]/\ell^p \to \ell[x]/\ell^p.
\]

Regarding the differential structure, we observe that $\varphi_S$ commutes with the derivation when $\ell[t]/\ell^p$ is endowed with the standard derivation $\frac{\partial}{\partial t}$. We furthermore deduce from the fact that $S$ and $f_A$ are coprime that $\varphi_S$ extends to a homomorphism of differential rings $k[x]/A \to \ell[t]/\ell^p$ that we continue to denote by $\varphi_S$. We set $\varphi_S = r \circ \varphi_S$ where $r$ is the canonical inclusion $\ell[t]/\ell^p \hookrightarrow \ell[t]/\ell^p$ (cf. §3). As before, $\varphi_S$ commutes with the derivation. Finally, because $S$ is separable, we can check that $\varphi_S$ is surjective and its kernel is the ideal generated by $S^p$. Hence $\varphi_S$ induces an isomorphism:

\[
k[x]/S^p = k[x]/k^p \cong \ell[t]/\ell^p.
\]

Let $Y_S$ be a fundamental system of solutions of the differential system $Y' = \varphi_S(A) Y$, i.e. $Y_S$ is an $r \times r$ matrix with coefficients in $\ell[t]/\ell^p$ such that $Y_S(0) = I_r$ and $Y_S^2 = \varphi_S(A) \cdot Y_S$. The existence of $Y_S$ is guaranteed by Proposition 3.4. Moreover, the matrix $Y_S$ is invertible because $Y_S(0) = I_r$ is.

**Proposition 4.3.** Keeping the above notations, we have:

\[
\varphi_S(A_p) = -Y_S^{(p)} \cdot Y_S^{-1}
\]

where $Y_S^{(p)}$ is the matrix obtained from $Y_S$ by taking the $p$-th derivative entry-wise.

**Proof.** We set $Z_S = Y_S^{-1}$ and let $(M, \partial)$ denote the differential module over $\ell[t]/\ell^p$ associated to the differential system $Y' = \varphi_S(A) Y$. Let $y_1, \ldots, y_r$ denote the column vectors of $Y_S$. They are all solutions of the system $Y' = \varphi_S(A) Y$, meaning that $\partial(y_i) = 0$ for all $i$. Furthermore, if $(e_1, \ldots, e_r)$ is the canonical basis of $\ell[t]/\ell^p$, we have the matrix relations: $Y_S \cdot e_i = y_i$ and $e_i = Z_S \cdot y_i$ where $\partial$ is the vector column whose coordinates are the vectors $y_i$'s (resp. $e_i$'s). Applying $\partial$ to the above relation, we find $\partial e_i = -Z_S \cdot e_i = z_S \cdot y_i = Z_S \cdot y_i$ and iterating this $p$ times, we deduce $\partial^p e_i = Z_S^{(p)} \cdot y_i = Z_S^{(p)} \cdot y_i$. On the other hand, the matrix $\varphi_S(A_p)$ of the $p$-curvature is defined by the relation $\partial^p e_i = \varphi_S(A_p) \cdot e_i$. Therefore we get $\varphi_S(A_p) = Y_S \cdot Z_S^{(p)}$. Now differentiating $p$ times the relation $Y_S Z_S = I_r$, we find $Y_S^{(p)} Z_S + Y_S \cdot Z_S^{(p)} = 0$. Combining this with the above formula for $\varphi_S(A_p)$ concludes the proof.

In our setting, the matrix $A_p$ has coefficients in $k[x]/\frac{1}{x^{\frac{1}{p}}}$ (cf. Lemma 4.1), from which we deduce that $\psi_S(A_p)$ has actually coefficients in the subring $\ell[t]/\ell^{2p}$. Therefore, using Eq. (9), one can compute $\psi_S(A_p)$ knowing only $Y_S$ modulo the ideal $\ell[t]/\ell^{2p}$.

One can actually go further in this direction and establish a variant of Eq. (9) giving an expression of $\psi_S(A_p)$ which involves only the reduction of $Y_S$ modulo $\ell[t]/\ell^2$. To make this precise, we need an extra notation. Given an integer $i \in [0, p]$ and a polynomial $f \in \ell[t]/\ell^p$ (resp. a matrix $M$ with coefficients in $\ell[t]/\ell^p$), we write $\text{Coeff}(f, i)$ (resp. $\text{Coeff}(M, i)$) for the coefficient in $t^i$ (resp. in $M$).
**Proposition 4.4.** Keeping the above notations, we have:

\[
\psi_S(A_p) = -\tilde{Y}_S \cdot Y_{S}^{(p)}(0) \cdot \tilde{Y}_S^{-1}
= \tilde{Y}_S \cdot \text{Coeff}(A \cdot \tilde{Y}_S, p-1) \cdot \tilde{Y}_S^{-1}
\]

where we have set \(\tilde{Y}_S = Y_S \mod \ell([t])_{\geq p}^0\).

Proof. Differentiating \(p\) times the relation \(Y_S^p = \psi_S(A)\cdot Y_S\), we observe that \(Y_S^{(p)}\) is solution of the same differential system \(Y' = \psi_S(A)Y\). Hence, thanks to uniqueness in Cauchy-Lipschitz Theorem, we have the relation \(Y_S^{(p)} = \psi_S(A)Y_S^{(p)}(0)\). The first part of the Proposition follows by plugging this in Eq. (9) and reducing the result modulo \(\ell([t])_{\geq p}^0\). To establish the second part, it is now enough to notice that the relation \(Y_S = \psi_S(A)\cdot Y_S\) implies:

\[
Y_S^{(p)}(0) = (A \cdot Y_S)^{(p-1)}(0) = -\text{Coeff}(A \cdot \tilde{Y}_S, p-1)
\]

the minus sign coming from \((p-1)! \equiv -1 \mod p\). \(\square\)

**Remark 4.5.** We can rephrase Proposition 4.4 as follows: letting \(y_1, \ldots, y_n\) denote the column vectors of \(Y_S\) and \(\bar{y}_i \in \ell([t]/\ell^p)\) be the reduction of \(y_i\), the \(p\)-curvature of an \(A\) modulo \(\ell^p\) is the linear endomorphism of \(\ell([t]/\ell^p)\) whose matrix in the basis \((\bar{y}_1, \ldots, \bar{y}_n)\) is \(\text{Coeff}(A \cdot \bar{Y}_S, p-1)\). It is worth remarking that the latter matrix has coefficients in the subring \(\ell\) of \(\ell([t]/\ell^p)\).

**Remark 4.6.** In the case where \(S = x^m - c\) (where \(c \in k\) and \(p\) does not divide \(m\)), there actually exists a quite simple explicit formula for \(\varphi_S^{-1}\): it takes \(t\) to \(t^p\) and \(x\) to \(c^px^m\), where \(n\) and \(q\) are integers satisfying the Bézout’s relation \(pn + qm = 1\). Using this, one can compute \(\varphi_S^{-1}(g)\) in \(O(pm)\) operations in \(k\) for the power \(p\).
To conclude with, it is worth remarking that implementing the algorithm \texttt{local\_p\_curvature} can be done using usual power series arithmetic: indeed, we only need to perform computations in the quotient \( \ell[t]/t^p \) which is isomorphic to \( \ell[t]/t^p \) by Corollary 3.3. Furthermore, we note that if we are using the algorithm \texttt{fundamental\_solutions} at line 2, then \( Y_S^{-1} \) can be computed by performing an extra loop in \texttt{fundamental\_solutions}; indeed the matrix \( Z \) we obtain this way is exactly \( Y_S^{-1} \).

### 4.3 Gluing

We recall that we have started with a differential system \( Y' = AY \) (with \( A = \frac{1}{\lambda A} \hat{A} \)) and that our goal is to compute the matrix \( A_\ell \) of its p-curvature. Lemma 4.1 gives bounds on the size of the entries of \( A_\ell \). We need another lemma, which ensures that we can find enough small “evaluation points” (lying in a finite extension of \( k \)). Let \( F_p \) denote the prime subfield of \( k \).

#### Lemma 4.7

Given a positive integer \( D \) and a nonzero polynomial \( f \in k[x] \), there exist pairwise coprime polynomials \( S_1, \ldots, S_n \in F_p[x] \) with \( n \leq D \) such that:
- \( \sum_{i=1}^n \deg S_i \geq D \)
- for all \( i \), the polynomial \( S_i \) is coprime with \( f \) and has degree at most \( 1 + \log_2(D + \deg f) \).

#### Proof.

Let \( m \) be the smallest positive integer such that \( p^m \geq D + \deg f \). Clearly \( m \leq 1 + \log_2(D + \deg f) \leq 1 + \log_2(D + \deg f) \).

Let \( F_p[m] \) be an extension of \( F_p \) of degree \( m \) and \( K \) be the composition of \( k \) and \( F_p[m] \). Let \( S_1, \ldots, S_n \) be the minimal polynomials over \( F_p \) (without repetition) of all elements in \( F_p[m] \subset K \) which are not a root of \( f \). We then have \( \deg S_i \leq m \) for all \( i \) and \( \sum_{i=1}^n \deg S_i \geq p^m - \deg f \).

It remains to define \( n \) as the smallest integer such that \( \sum_{i=1}^n \deg S_i \geq D \). Minimality implies \( \sum_{i=1}^{n-1} \deg S_i < D \) and thus \( n \leq D \). Therefore \( S_1, \ldots, S_n \) satisfy all the requirements of the lemma. \( \square \)

The above proof yields a concrete algorithm for producing a sequence \( S_1, \ldots, S_n \) satisfying the properties of Lemma 4.7: we run over elements in \( F_p[m] \) and, for each new element, append its minimal polynomial over \( F_p \) to the sequence \( S_i \) unless it is not coprime with \( f \). We continue this process until the condition \( \sum_{i=1}^n \deg S_i \geq D \) holds. Keeping in mind the logarithmic bound on \( m \), we find that the complexity of this algorithm is at most \( O(D + \deg f) \) operations in \( k \). Let us call \texttt{generate\_points} the resulting routine: it takes as inputs the parameters \( f \) and \( D \) and return an admissible sequence \( S_1, \ldots, S_n \).

We are now ready to present our algorithm for computing the p-curvature:

#### Algorithm \texttt{p\_curvature}

**Input:** a matrix \( \hat{A} \) written as \( A = \frac{1}{\lambda A} \hat{A} \)

**Output:** the p-curvature of the differential system \( Y' = AY \)

1. \( S_1, \ldots, S_n = \texttt{generate\_points}(f_A, d + 1) \)
2. for \( i = 1, \ldots, n \) do
   - \( A_i,p = \texttt{local\_p\_curvature}(S_i, A \mod S_i^p) \)
3. compute \( B \in M_n(k[x]) \) with entries of degree \( \leq pD \) such that \( B \equiv f_A \cdot B_i \pmod{S_i^p} \) for all \( i \)

**Cost:** \( O(pd^2) \) operations in \( k \)

4. return \( \frac{1}{\lambda A} \cdot B \)

In view of the previous discussion and Lemma 4.1, the correctness and the cost analysis of the algorithm \texttt{p\_curvature} are both straightforward. Hence, Theorem 4.2 is proved.

We conclude this subsection with three remarks. First, when applying Chinese Remainder Theorem (CRT) on line 3 of Algorithm \texttt{p\_curvature}, we notice that all moduli \( S_i^p \) are polynomials in \( x^p \). This allows the following optimization. Writing \( f_A \cdot B \equiv \sum_{j=0}^{p-1} B_{ij} x^j \pmod{S_i^p(x)} \) and denoting by \( C_i \) the unique solution of degree at most \( d \) to the congruence system:

\[ B_j(x) \equiv B_{ij}(x) \pmod{T_i(x)} \]

we have \( B = \sum_{i=0}^{p-1} B_i(x) x^i \). This basically allows us to replace one CRT with polynomials of degree \( pd \) by \( p \) CRT with polynomials of degree \( d \). We save this way the polynomial factors in \( \log(p) \) in the complexity.

Second, instead of working with \( n \) polynomials \( S_i \), one may alternatively choose a unique polynomial \( S \) of the form \( S = X^m - a \) where \( m \geq d \) is an integer not divisible by \( p \) and \( a \in k \) are such that \( S \) and \( f_A \) are coprime. This avoids the use of Chinese Remainder Theorem and the resulting complexity stays in \( O(pd^2) \) provided that we use Remark 4.6 in order to compute the inverse of \( \varphi_S \).

Third, we observe that the algorithm \texttt{p\_curvature} is very easily parallelizable. Indeed, each iteration of the main loop (on line 2) is completely independent from the others. Thus, they all can be performed in parallel. Moreover, according to the first remark (just above), the application of the Chinese Remainder Theorem (on line 3) splits into \( pr^2 \) smaller independent problems and can therefore be efficiently parallelized as well.

### 4.4 The case of differential operators

To conclude with, we would like to discuss the case of a differential operator \( L = a_i \partial^i + a_{i-1} \partial^{i-1} + \cdots + a_0 \partial + a_0 \) with \( a_i \in k[x] \) for all \( i \), of maximal degree \( d \).

Recall that the p-curvature of \( L \) is that of the differential module \( (\partial/\partial x)L, \partial, c) \), where \( c \) is the companion matrix associated to \( L \) as in (3). Applying directly the formulas in Proposition 4.4 requires the knowledge of the solutions of the system \( Y' = -CY \). It is in fact easier to compute solutions for the system \( Y' = CX \), since we saw that these solutions are the vectors of the form \( (y, y', \ldots, y^{(r-1)}) \), where \( y \) is a solution of \( L \). This is however harmless: the p-curvatures \( A_p \) and \( B_p \) of the respective systems \( Y' = -CY \) and \( Y' = CX \) (which are so-called adjoint) satisfy \( A_p = -A_{p+1} \). Thus, we can use the formulas given above to compute \( \varphi_S(B_p) \), and deduce \( \varphi_S(A_p) \) for a negligible cost. Equivalently, one may notice that the fundamental matrices of solutions of our two systems are transpose of one another, up to sign.

Moreover, instead of using the second formula of Proposition 4.4 to compute the local p-curvatures, we recommend using the first one, which is \( \varphi_S(B_p) = -X_S \cdot X_S^{(p)}(0) \cdot X_S^{-1} \) where \( X_S \) is a fundamental system of solutions of \( Y' = CX \) and \( X_S \) denotes its reduction in \( \left[ \frac{1}{\lambda A} \right]/(t^p) \). If \( f_0, \ldots, f_{r-1} \) are solutions of the system (7), the \( (i,j) \)-th entry of \( X_S \) is just \( f_j^{(i)} \). Hence the matrices \( X_S \) and \( X_S^{(p)}(0) \) can be obtained from the knowledge of the image of \( f_i \)’s modulo
\[ \ell(\ell(t)^d)_{d+p,r} \]
just by reorganizing coefficients (and possibly multiplying by some factorials depending on the representation of elements of \( \ell(\ell(t)^d) \), we are using).

As for the \( f_i \)'s, they can be computed by the algorithm solutions_operator (provided its assumptions are satisfied).

We need finally to compute \( X_{-1} \): since \( X_p(0) \) is the identity matrix, this can be done either using Newton iterator, a divide-and-conquer approach or a combination of both, which computes the inverse of \( X_S \) at a small precision, and uses divide-and-conquer techniques for higher ones (the latter being the most efficient in practice). All these remarks do speed up the execution of our algorithms when \( d \) is not too large compared to \( r \).

Last but not least, we notice that, in the case of differential operators, the matrix \( A_p \) is easily deduced from its first column. Indeed, writing \( A_p = (a_{i,j})_{0 \leq i,j \leq r} \) and letting
\[ c_j = a_{r-1,j} \partial^{r-1} + \cdots + a_{1,j} \partial + a_{0,j} \in k[x](\partial) \]
be the differential operator obtained from the \( j \)-column of \( A_p \), it is easily checked that \( c_{j+1} \) is the remainder in the Euclidean division of \( \partial c_j \) by \( L \). Comparing orders, we further find
\[ c_{j+1} = c_j L - l(c_j) \]
where \( l(c_j) \) is the leading coefficient of \( c_j \). This remark is interesting because it permits to save memory: indeed, instead of storing all local \( p \)-curvatures \( A_p, \ell \), we can just store their first column. Doing this, we can reconstruct the first column of \( A_p \) using the Chinese Remainder Theorem (cf \S 4.3) and then compute the whole matrix \( A_p \) using the recurrence.

5. IMPLEMENTATION AND TIMINGS

We implemented our algorithms in Magma in the case of differential operators; the source code is available at https://github.com/schost. Figure 1 gives running times for random operators of degrees \( (d,r) \) in \( k[x](\partial) \) and compares them with running times of a (fraction free version of) Katz’s algorithm which consists in computing the recursive sequence \( (A_i) \) until \( i = p \). In each cell, the first line (resp. the second line) corresponds to the running time obtained with our algorithm (resp. Katz’s algorithm): a dash indicates that the corresponding running time exceeded one hour. Our benchmarks rather well reflect the predicted dependence with respect to \( p \): quasi-linear for our algorithm and quadratic for Katz’s algorithm.

Larger examples (than those presented in Fig. 1) are also reachable: for instance, we computed the first column of the \( p \)-curvature of a “small” multiple of the operator \( \phi^5(\partial) \) considered in [10, Appendix B.3] modulo the prime 27449. This operator has bidegree \( (d,r) = (108,28) \). The computation took about 19 hours and the size of the output in human-readable format is about 1GB (after bzip2 compression, it decreases to about 300MB).

6. REFERENCES


