Complexity bounds for zero-test algorithms

Joris van der Hoeven

Dépt. de Mathématiques (bât. 425)
Université Paris-Sud
91405 Orsay Cedex
France

John Shackell

Department of Mathematics
University of Kent at Canterbury
Canterbury
England

Abstract

In this paper, we analyze the complexity of a zero test for expressions built from formal power series solutions of first order differential equations with non degenerate initial conditions. We will prove a doubly exponential complexity bound. This bound establishes a power series analogue for "witness conjectures".

1. Introduction

Zero equivalence is a major issue on the analysis side of symbolic computation. Standard mathematical notation provides a way of representing many transcendental functions. However, trivial cases apart, this notation gives rise to the following problems:

- Expressions may not be defined: consider $1/0$, $\log(0)$ or $\log(e^{x+y} - e^x e^y)$.
- Expressions may be ambiguous: what values should we take for $\log(-1)$ or $\sqrt{z^2}$?
- Expressions may be redundant: $\sin^2 x + \cos^2 x$ and 1 are different expressions, but they represent the same function.

⋆ The paper was originally written using GNU TeXmacs (see `www.texmacs.org`). Unfortunately, Elsevier insists on the use of LaTeX with its own style files. Any insufficiencies in the typesetting quality should therefore be imputed to Elsevier.

Email addresses: `vdhoeven@texmacs.org` (Joris van der Hoeven), `jrs@ukc.ac.uk` (John Shackell).

Often, one is interested in expressions which represent functions in a ring. In that case, the third problem reduces to deciding when a given expression represents the zero function.

As to the first two problems, one has to decide where and how we want our functions to be defined. In this paper, we will mainly be concerned with expressions that represent multivariate power series. The expressions will then be formed from the constants and the indeterminates using the ring operations and power series solutions to first-order differential equations. The correctness and non-ambiguity of expressions may then be ensured by structural induction. This may involve zero-testing for the series represented by subexpressions.

In order to evaluate the complexity of algorithms, one needs a reasonable notion for the size of an expression. In this paper, the size of an expression will always be the number of nodes in the corresponding "expression tree". For instance the size of $\sin(\log(x)) + 5x$ is 7. In section 3.1, we will also introduce the alternative notion of the "pseudo-norm" of an expression. Roughly speaking, all expressions in this paper may be represented by polynomials in a tower of field extensions $\mathcal{F}_0 \subseteq \cdots \subseteq \mathcal{F}_h$. Such towers start with a field $\mathcal{F}_0 = \mathcal{C}$ of constants and each $\mathcal{F}_i$ with $i > 0$ is of the form

$$\mathcal{F}_i = \mathcal{F}_{i-1}\left[f_i, \frac{1}{B_{i,1}(f_i)}, \ldots, \frac{1}{B_{i,k_i}(f_i)}\right],$$

where f_i is a solution to an algebraic differential equations over $\mathcal{F}_{i-1}$ and the $B_{i,j}$ are polynomials over $\mathcal{F}_{i-1}$. The pseudo-norm of an element in $\mathcal{F}_i$ is defined recursively in terms of its degree in f_i and the pseudo-norms of its coefficients.

1.1. Zero-tests for constants

As a first step, one would like to be able to solve zero-equivalence for the elementary constants, that is to say the smallest field of constants closed under the application of the exponential, trigonometric and (for non-zero argument) logarithmic functions. Alas no such algorithm is known and it is clear that some formidable problems in transcendental number theory would need to be solved before one was found. In the face of this dilemma implementers have used heuristic methods generally involving floating-point computations.

Theoreticians have often resorted to the use of an oracle; in other words they presupposed a solution to the problem for constants. They have then gone on to develop other algorithms, for example to decide zero equivalence of functions, on this basis. However for elementary constants one can do better than merely invoke an oracle.

> The Schanuel Conjecture may be stated as follows: Let $\alpha_1, \ldots, \alpha_k$ be complex numbers which are linearly independent over the rational numbers $\mathbb{Q}$. Then the transcendence degree of
>
> $$\mathbb{Q}(\alpha_1, \ldots, \alpha_k, \exp(\alpha_1), \ldots, \exp(\alpha_k)) : \mathbb{Q}$$
>
> is at least k. Many special cases of this are well known unsolved conjectures in transcendental number theory. Following work by Lang, [Lang (1971)], algorithms for deciding the signs of elementary constants based on the Schanuel Conjecture have been given by Caviness and Prelle, [Caviness and Prelle (1978)] and Richardson, [Richardson (1997)]. The conjecture has been shown to imply the decidability of the real exponential field, [Macintyre and Wilkie (1995)].

There are definite advantages in using a conjecture from number theory rather than heuristic methods, in that it is clear what is being assumed and any counter example found would be of considerable mathematical interest. However in a practical situation, a zero equivalence method for constants is generally needed very often, and here the algorithms based on the Schanuel Conjecture are really rather slow.

Another limitation of the above approach is that it is very hard to see how to generalize the Schanuel Conjecture to cover constants given by Liouvillian or Pfaffian functions. For the substance of the conjecture is that the relations between exponentials and logarithms of numbers are just the ones we already know about, but it seems impossible to even formulate such a claim when integrals and solutions of differential equations are involved.

In [van der Hoeven (2001b)], the following witness conjectures was made. Let $N \geqslant 3$ and consider the set $\mathcal{E}_N$ of real exp-log expressions such that for each subexpression of the form $\exp f$ or $\log f$, we have $|\hat{f}| \leqslant N$ resp. $N^{-1} \leqslant |\hat{f}| \leqslant N$, where $\hat{f}$ denotes the value of f as a real constant. Then there exists a *witness function* of the form $\varpi(n) = C_N n$ (strong witness conjectures) or $\varpi(n) = \mathrm{e}^{C_n n}$ (weak witness conjecture) such that for any $f \in \mathcal{E}_N$ of size $\sigma(f)$, it suffices to evaluate $\hat{f}$ up to $\varpi(\sigma(f))$ digits in order to determine whether it vanishes.

Earlier versions and variants of witness conjectures appeared in [van der Hoeven (1997, 2001a); Richardson (2001); van der Hoeven (2001b)]. Also, Dan Richardson has accumulated numerical evidence and worked out some number-theoretic consequences. It should be noticed that these conjectures are apparently independent of the Schanuel conjecture. Indeed, there might exist non zero elementary constants, which yet evaluate to extremely small numbers. On the other hand, there might exist counterexamples to the Schanuel conjecture which can be "detected" to be zero by evaluating a reasonable number of digits. The interest of witness conjectures is that they potentially provide us with fast zero tests, if they can be proved to to hold for "reasonably small" witness functions ϖ.

Remark 1 Recently, Joris van der Hoeven and Dan Richardson found a counterexample to the strong witness conjecture: consider the function

$$f(z) = \log(1 + z) - 2\log(1 + \log(1 + z/2)).$$

The variable z occurs only twice in the function, but f has valuation 3 as a power series. Therefore, the n-th iterate $f^{\circ n}$ of f has size $O(2^n)$, but valuation 3^n. Consequently, the constant $f^{\circ n}(\frac{1}{2})$ yields a counterexample to the strong witness conjecture for sufficiently large n. This counterexample has been generalized in [van der Hoeven (2003)] to all polynomial witness functions. Nevertheless, no counterexamples to the weak witness conjecture are currently known.

1.2. Zero-tests for functions

Although zero-test algorithms for constants are extremely hard to design, more progress has been made on zero-tests for functions [Shackell (1989, 1993); Péladan-Germa (1995)]. Unfortunately, no reasonable complexity bounds (i.e. less that the Ackermann function) for these algorithms were known up till now. In this paper, we both generalize the algorithm from [Shackell (1989, 2004)] to the multivariate setting and provide complexity bounds. A recent survey on the theoretical complexity of calculations involving Pfaffian functions is given in [Gabrielov and Vorobjov (2004)].

Now it is interesting to study the significance of such bounds for the exp-log constant conjecture. Indeed, since number theoretical questions about transcendence or Diophantine approximation are usually very hard, a first step usually consists of formulating analogue questions in the setting of function fields. A deep and well-known theorem of Ax [Ax (1971)] states that the power series version of Schanuel's conjecture does hold.

The exp-log conjecture also admits a natural power series analogue. Given a field $\mathcal{C}$, consider the set $\mathcal{E}_k$ of multivariate power series expressions constructed from $\mathcal{C}$ and $z_1, \ldots, z_k$ using $+, -, \times$ and left composition of infinitesimal series by $1/(1+z)$, $\exp z$ and $\log(1+z)$. Now let $f \in \mathcal{E}_k$ be such an expression of size $\sigma(f)$ and let $\rho(f) \in \mathcal{C}[[z_1, \ldots, z_k]]$ be the power series represented by f. Then we expect that there exists a constant C_k, such that $\rho(f) = 0$ if and only if the coefficient of $z_1^{\alpha_1} \cdots z_k^{\alpha_k}$ in $\rho(f)$ vanishes for all $\alpha_1, \ldots, \alpha_k \in \{0, \ldots, C_k \sigma(f)\}$.

As a side effect of our complexity bounds, we will be able to prove a weaker result: with the above notation, there exists a constant C, such that $\rho(f)$ vanishes if and only if the coefficient of $z_1^{\alpha_1} \cdots z_k^{\alpha_k}$ in $\rho(f)$ vanishes for all $\alpha_1, \ldots, \alpha_k \in \{0, \ldots, k^{C^{\sigma(f)}}\}$. Just as the Ax theorem gives theoretical evidence that for the numerical Schanuel conjecture, our result thereby gives evidence that the numerical witness conjecture might be true.

1.3. Overview

In section 2, we describe our setup of *effective local domains* for doing computations on power series. Such computations may either be effective zero-tests or the extraction of coefficients. We will next consider the extension of effective local domains by solutions of first order partial differential equations. In section 5, we will show that such extensions are again effective local domains.

In section 3, we recall the Bareiss method for Gaussian elimination of matrices with coefficients in an integral domain. This method has the advantage of limiting the expression swell. More precisely, we give bounds in terms of *pseudo-norms* on integral domains. In the sequel, the Bareiss method is applied to the efficient g.c.d. computation of several polynomials. This is an essential improvement with respect to [Shackell (1989)], which allows us to obtain "reasonable" complexity bounds for our zero test.

In section 4, we prove four key lemmas which ensure the correctness of our zero test. We also corrected a small mistake in the original correctness proof in [Shackell (1989)]. In section 5 we present the actual algorithm and complexity bounds. The main idea behind the algorithm is as follows: consider a polynomial $P \in \mathcal{C}[f_1, \ldots, f_n]$, where $f_1, \ldots, f_n$ are solutions to given algebraic differential equations. Then P is zero-equivalent if and only if any differentially algebraic consequence of $P = 0$ and the defining equations of $f_1, \ldots, f_n$ is zero-equivalent. Using g.c.d. computations, we first compute a particularly simple such consequence. Next, it will suffice to check the zero-equivalence of this consequence up to a certain order.

In the case of power series over the real numbers, it is possible to obtain better theoretical bounds using techniques from [Khovanskii (1991)]. Nevertheless, we think that the results of this paper are interesting from several point of views:

- The framework is more general, because we show how to obtain complexity bounds in a relative way for extensions of effective local domains.
- We presented an improved version of an actual zero-test algorithm, which might have a better average complexity than Khovanskii's complexity bounds in non-degenerate cases, although we have not yet proved a better bound for the worst case.

- Our methods are likely to generalize to higher order differential equations, by adapting the algorithms from [Shackell (1993); van der Hoeven (2002a)].

We plan to improve our complexity bounds in a forthcoming paper, with the hope of obtaining bounds closer to those in [Khovanskii(1991), Theorem 1.2], i.e. of the form $2^{\sigma^2/2+o(\sigma^2)}$ instead of $O((4k\sigma)^{7^\sigma})$.

2. The main setup

2.1. *Effective local domains*

Let $\mathcal{C}$ be an effective field of constants, which means that all field operations can be performed algorithmically and that we have an effective zero test. The ring $\mathcal{C}[[z_1,\ldots,z_k]]$ is a differential ring for the partial differentiations $\partial_1,\ldots,\partial_k$ w.r.t. $z_1,\ldots,z_k$ on $\mathcal{C}[[z_1,\ldots,z_k]]$.

We will frequently consider multivariate power series in $\mathcal{C}[[z_1,\ldots,z_k]]$ as recursive power series in $\mathcal{C}[[z_1]]\cdots[[z_k]]$. For this reason, it is convenient to introduce the partial evaluation mappings $\varepsilon_i : \mathcal{C}[[z_1,\ldots,z_j]] \to \mathcal{C}[[z_1,\ldots,z_i]]$ with

$$\varepsilon_i(f(z_1,\ldots,z_j)) = f(z_1,\ldots,z_i,0,\ldots,0)$$

for all $0 \leqslant i \leqslant j \leqslant k$. We re-obtain the total evaluation mappings $\varepsilon = \varepsilon_0$ as special cases. Notice that

$$\partial_i \varepsilon_j(f) = \varepsilon_j(\partial_i f),$$

for every $f \in \mathcal{C}[[z_1,\ldots,z_k]]$ and $1 \leqslant i \leqslant j \leqslant k$.

Definition 1 *A differential subring $\mathcal{R}$ of $\mathcal{C}[[z_1,\ldots,z_k]]$ is called an **effective power series domain**, if we have algorithms for $+,-,\times,\varepsilon,\partial_1,\ldots,\partial_k$ and an algorithm to test whether $\varepsilon_i(f)=0$ for each $0 \leqslant i \leqslant k$ and $f \in \mathcal{R}$.*

Remark 2 Given an effective power series domain $\mathcal{R} \subseteq \mathcal{C}[[z_1,\ldots,z_k]]$, we observe that $\varepsilon_i(\mathcal{R}) \subseteq \mathcal{C}[[z_1,\ldots,z_i]]$ may be considered as an effective power series domain for each $1 \leqslant i \leqslant k$. Indeed, this follows from the fact that ε_i commutes with $+,-,\times,\varepsilon_0,\ldots,\varepsilon_i,\partial_1,\ldots,\partial_i$.

Let $\mathcal{R}$ be an effective power series domain and let f be a power series in $\mathcal{C}[[z_1,\ldots,z_k]]$, which satisfies partial differential equations

$$\begin{cases} \partial_1 f = \frac{A_1(f)}{B_1(f)} \\ \quad\vdots \\ \partial_k f = \frac{A_k(f)}{B_k(f)} \end{cases} \tag{1}$$

where $A_i, B_i \in \mathcal{R}[F]$ are such that $\varepsilon(B_i(f)) \neq 0$ for each i. Then the ring

$$\mathcal{S} = \mathcal{R}\left[f, \frac{1}{B_1(f)}, \ldots, \frac{1}{B_k(f)}\right]$$

is a differential subring of $\mathcal{C}[[z_1,\ldots,z_k]]$, which is called a *regular D-algebraic extension* of $\mathcal{R}$. The main aim of this paper is to show that $\mathcal{S}$ is also an effective power series domain and to give complexity bounds for the corresponding algorithms.

2.2. Computations in $\mathcal{S}$

Elements in $\mathcal{R}[f]$ are naturally represented by polynomials $\mathcal{R}[F]$ in a formal variable F, via the unique $\mathcal{R}$-algebra morphism $\rho : \mathcal{R}[F] \to \mathcal{R}[f]$ with $\rho(F) = f$. This mapping ρ naturally extends to a mapping $\rho : \check{\mathcal{S}} \to \mathcal{S}$, where

$$\check{\mathcal{S}} = \mathcal{R}\left[F, \frac{1}{B_1(F)}, \ldots, \frac{1}{B_k(F)}\right] \subseteq \mathcal{R}(F)$$

The structure of $\mathcal{S}$ may be transported to $\check{\mathcal{S}}$ in a natural way. The partial differentiations $\partial_1, \ldots, \partial_k$ on $\mathcal{R}$ extend uniquely to $\check{\mathcal{S}}$ by setting $\partial_i F = A_i(F)/B_i(F)$ for each i (so that $\rho \circ \partial_i = \partial_i \circ \rho$). Each partial evaluation mapping $\varepsilon_i : \mathcal{S} \to \mathcal{C}$ induces a natural evaluation mapping $\varepsilon_i \circ \rho$ on $\check{\mathcal{S}}$, which we will also denote by ε_i. Since $\mathcal{R}$ is an effective local domain, the operations $+, -, \times, \varepsilon, \partial_1, \ldots, \partial_k$ can clearly be performed algorithmically on $\mathcal{S}$. Our main problem will therefore be to design a zero-test for $\mathcal{S}$, which amounts to deciding whether $\rho(P/Q) = 0$ for a given rational function $P/Q \in \check{\mathcal{S}}$.

Actually, it is more convenient to work with polynomials in $\mathcal{R}[F]$ instead of rational functions in $\check{\mathcal{S}}$. Our main problem will then be to decide whether $\rho(P) = 0$ for $P \in \mathcal{R}[F]$, since a rational function $P/Q \in \check{\mathcal{S}}$ represents zero if and only if P does. Unfortunately, the ring $\mathcal{R}[F]$ is not necessarily stable under the derivations $\partial_1, \ldots, \partial_k$. For this reason, we introduce the derivations

$$d_i = B_i(F)\partial_i,$$

which do map $\mathcal{R}[F]$ into itself.

In order to determine whether $\rho(P) = 0$ for polynomials $P \in \mathcal{R}[F]$, we will consider the roots of such polynomials in the algebraic closure $\mathcal{R}^{\mathrm{alg}}$ of $\mathcal{R}$. Now it is classical that the algebraic closure of the ring $K[[z]]$ of univariate power series over a field K is the field $K^{\mathrm{alg}}\langle\langle z\rangle\rangle$ of Puiseux series over the algebraic closure K^{alg} of K. Interpreting multivariate power series in $\mathcal{C}[[z_1, \ldots, z_k]]$ as recursive power series in $\mathcal{C}[[z_1]] \cdots [[z_k]]$, we may thus view elements in $\mathcal{R}^{\mathrm{alg}}$ as recursive Puiseux series in $\mathcal{C}^{\mathrm{alg}}\langle\langle z_1\rangle\rangle \cdots \langle\langle z_k\rangle\rangle$.

2.3. Extraction of coefficients

In what follows, it will convenient to use vector notation. We define the anti-lexicographical ordering $\leqslant$ on $\mathbb{Q}^k$ by

$$\begin{aligned}
\boldsymbol{\alpha} \leqslant \boldsymbol{\beta} \Longleftrightarrow & (\alpha_1 = \beta_1 \wedge \cdots \wedge \alpha_k = \beta_k) \vee \\
& (\alpha_1 < \beta_1 \wedge \alpha_2 = \beta_2 \wedge \cdots \wedge \alpha_k = \beta_k) \vee \\
& \vdots \\
& (\alpha_k < \beta_k),
\end{aligned}$$

for $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_k), \boldsymbol{\beta} = (\beta_1, \ldots, \beta_k) \in \mathbb{Q}^k$.

Consider a Puiseux series $\varphi \in \mathcal{C}\langle\langle z_1\rangle\rangle \cdots \langle\langle z_k\rangle\rangle$. We will write

$$\varphi = \sum_{\alpha} \varphi_{\alpha} z_k^{\alpha}$$

for the power series expansion of φ w.r.t. z_k. Each coefficient may recursively be expanded in a similar way w.r.t. $z_{k-1}, \ldots, z_1$. Alternatively, we may expand φ at once w.r.t. all variables using the anti-lexicographical ordering:

$$\varphi = \sum_{\boldsymbol{\alpha}} \varphi_{\boldsymbol{\alpha}} z^{\boldsymbol{\alpha}},$$

where $z^{\boldsymbol{\alpha}} = z_1^{\alpha_1} \cdots z_k^{\alpha_k}$. If $f \neq 0$, then the minimal $\boldsymbol{\alpha}$ with $\varphi_{\boldsymbol{\alpha}} \neq 0$ is called the valuation of φ and we denote it by $\boldsymbol{v}(\varphi)$. If $\boldsymbol{v}(\varphi) \geqslant \mathbf{0}$, then may we define $\varepsilon_i(\varphi)$ to be the coefficient of $z_{i+1}^0 \cdots z_k^0$ in φ, for each $i \in \{0, \ldots, k\}$. As usual, we denote $\varepsilon(\varphi) = \varepsilon_0(\varphi)$.

If $P \in \mathcal{R}[F]$ is a non zero polynomial, then we define the valuation $\boldsymbol{v}(P)$ of P to be the minimum of the valuations of its non zero coefficients. Suppose that $P(\lambda) = P_d \lambda^d + \cdots + P_0$. Then we recall that $P_d, \ldots, P_0$ are power series and define

$$P_{\boldsymbol{\alpha}}(\lambda) = P_{d,\boldsymbol{\alpha}} \lambda^d + \cdots + P_{0,\boldsymbol{\alpha}} \in \mathcal{C}[\lambda],$$

for each $\boldsymbol{\alpha} \in \mathbb{N}^k$.

It should be noticed that the recursive extraction of coefficients can be done effectively in an effective power series domain $\mathcal{R}$, because

$$\varphi_{\alpha_k, \ldots, \alpha_{i+1}} = \frac{1}{\alpha_k! \cdots \alpha_{i+1}!} \varepsilon_i(\partial_k^{\alpha_k} \cdots \partial_{i+1}^{\alpha_{i+1}} \varphi)$$

for all $\varphi \in \mathcal{R}$ and $\alpha_k, \ldots, \alpha_{i+1} \in \mathbb{N}$. More generally, if $P \in \check{\mathcal{S}}$, then we may formally represent the coefficient $\varphi_{\alpha_k, \ldots, \alpha_{i+1}}$ of $\varphi = \rho(P)$ by polynomials in $\varepsilon_i(\check{\mathcal{S}})$. Such representations are best derived through relaxed evaluation of formal power series [van der Hoeven (2002b)], by using the partial differential equations satisfied by f.

3. The Bareiss method and g.c.d. computations

3.1. Pseudo-norms

Let $\mathcal{R}$ be an effective integral domain. In what follows, we will describe algorithms to triangulate matrices with entries in $\mathcal{R}$ and compute g.c.d.s of polynomials with coefficients in $\mathcal{R}$. In order to state complexity bounds, it is convenient to measure the "sizes" of coefficients in $\mathcal{R}$ in terms of a *pseudo-norm*, which is a function $\nu : \mathcal{R} \to \mathbb{N}$ with the following properties:

N1 $\nu(\varphi + \psi) \leqslant \max\{\nu(\varphi), \nu(\psi)\}$.

N2 $\nu(\varphi\psi) \leqslant \nu(\varphi) + \nu(\psi)$.

If $\mathcal{R}$ is actually a differential ring with derivations $\partial_1, \ldots, \partial_k$, then we also assume the existence of a constant $K_{\mathcal{R}} \in \mathbb{N}$ with

N3 $\nu(\partial_i \varphi) \leqslant \nu(\varphi) + K_{\mathcal{R}}$.

As remarked in the introduction, the pseudo-norm of an expression should not be confused with its "natural size", i.e. the number of nodes of the corresponding expression tree.

Example 1 If $\mathcal{R} = \mathcal{C}[z_1, \ldots, z_k]$, then we may take $\nu(P)$ to be the maximum of the degrees of P in $z_1, \ldots, z_k$ and $K_{\mathcal{R}} = 0$.

Example 2 Assume that $\mathcal{R}$ and $\mathcal{S}$ are as in section 2 and assume that we have a pseudo-norm ν on $\mathcal{R}$. Then we define a pseudo-norm on $\check{\mathcal{S}}$ by

$$\nu(P) = \max\left\{\deg_F P, \deg_{B_1(F)^{-1}} P, \ldots, \deg_{B_k(F)^{-1}} P, \max_{P_* \text{ coefficient of } P} \nu(P_*)\right\}.$$

This pseudo-norm induces a pseudo-norm on $\mathcal{S}$ by

$$\nu(\varphi) = \min\left\{\nu(P) | P \in \check{\mathcal{S}}, \varphi = \rho(P)\right\}.$$

Notice that we may take

$$K_{\mathcal{S}} = K_{\check{\mathcal{S}}} = \max\left\{K_{\mathcal{R}}, 2, \max\{\nu(A_1), \ldots, \nu(A_k)\} + \max\left\{\nu\left(\frac{\partial B_1}{\partial F}\right), \ldots, \nu\left(\frac{\partial B_k}{\partial F}\right)\right\}\right\}.$$

3.2. The Bareiss method

Let $\mathcal{R}$ still be an effective integral domain with a pseudo-norm ν and quotient field $\mathcal{F}$. Consider an $m \times n$ matrix M with entries in $\mathcal{F}$ (i.e. a matrix with m rows and n columns). For all indices $1 \leqslant i_1 < \cdots < i_k \leqslant m$ and $1 \leqslant j_1 < \cdots < j_l \leqslant n$, we will also write $M_{[i_1,\ldots,i_k],[j_1,\ldots,j_l]}$ for the $k \times l$ minor of M when we only keep the rows $i_1, \ldots, i_k$ and columns $j_1, \ldots, j_l$.

It is classical that we may upper triangulate M using Gaussian elimination. This leads to a formula

$$T = UM,$$

where U is a matrix with determinant one and T an upper triangular matrix. Unfortunately, this process usually leads to a fast coefficient growth for the numerators of the entries of the successive matrices. In order to remove this drawback, we will rather do all computations over $\mathcal{R}$ instead of $\mathcal{F}$. In this section, we will briefly recall this approach, which is due to Bareiss [Bareiss (1968); Loos (1983)].

So let us now be given an $m \times n$ matrix M with entries in $\mathcal{R}$. For simplicity, we will first assume that the usual triangulation of M as a matrix with entries in $\mathcal{F}$ does not involve row permutations. This usual triangulation of M gives rise to a sequence of identities

$$\bar{T}_k = \bar{U}_k M,$$

with $k \in \{0, \ldots, m\}$, where $\bar{T}_k$ is the matrix obtained from $M = \bar{T}_0$ after k steps. More precisely, $\bar{T}_k$ is obtained from $\bar{T}_{k-1}$ by leaving the first k rows invariant and by adding multiples of the k-th row to the others (in particular, the $\bar{U}_k$ will be lower triangular throughout the process). If there exists a q with $(\bar{T}_k)_{k,q} \neq 0$, then let p_k be the minimal such q, so that $(\bar{T}_k)_{l,r} = 0$ for all $l > k$ and $r < p_k$. Each $\bar{T}_k$ may be rewritten as a product

$$\bar{T}_k = D_k^{-1} T_k,$$

of an invertible diagonal matrix D_k and another matrix T_k with entries in $\mathcal{R}$. Our aim is to show that we may choose the D_k and T_k of small pseudo-norms. In fact, we claim that we may take $D_k = \mathrm{Diag}(1, \delta_1, \ldots, \delta_{k-2}, \delta_{k-1}, \delta_{k-1}, \ldots, \delta_{k-1})$ for each k, where $\delta_k = (T_k)_{k,p_k}$.

In order to see this, let $k \geqslant 1$, $i > k-1$ and $j > p_{k-1}$. Then

$$(\bar{T}_k)_{[1,\ldots,k-1,i],[p_1,\ldots,p_{k-1},j]} = \bar{U}_{k,[1,\ldots,k-1,i],[1,\ldots,k-1,i]} M_{[1,\ldots,k-1,i],[p_1,\ldots,p_{k-1},j]},$$

since $\bar{U}_{k,[1,\ldots,k-1,i],[1,\ldots,k-1,i]}$ is a lower triangular matrix. Moreover, this matrix has only ones on its diagonal, whence

$$\det(\bar{T}_k)_{[1,\ldots,k-1,i],[p_1,\ldots,p_{k-1},j]} = \det M_{[1,\ldots,k-1,i],[p_1,\ldots,p_{k-1},j]}$$

Since $(\bar{T}_k)_{[1,\ldots,k-1,i],[p_1,\ldots,p_{k-1},j]}$ is upper triangular, we also have

$$\det(\bar{T}_k)_{[1,\ldots,k-1,i],[p_1,\ldots,p_{k-1},j]} = (\bar{T}_k)_{1,p_1} \cdots (\bar{T}_k)_{k-1,p_{k-1}} (\bar{T}_k)_{i,j}.$$

By our choice of D_k, we finally have

$$(\bar{T}_k)_{1,p_1} \cdots (\bar{T}_k)_{k-1,p_{k-1}} (\bar{T}_k)_{i,j} = \frac{(T_k)_{1,p_1} \cdots (T_k)_{k-1,p_{k-1}} (T_k)_{i,j}}{1\delta_1 \cdots \delta_{k-1}} = (T_k)_{i,j}.$$

Putting this together, we see that each coefficient of T_k (whence in particular δ_k) may be written as the determinant of a minor of M:

$$(T_k)_{i,j} = \det M_{[1,\ldots,k-1,i],[p_1,\ldots,p_{k-1},j]}. \tag{2}$$

Hence, we have not only shown that the D_k and T_k have coefficients in $\mathcal{R}$, but even that they may be written explicitly as determinants of minors of M. This result remains so (up to a factor ± 1) if row permutations were needed in the triangulation process, since we may always permute the rows of M *a priori*, so that no further row permutations are necessary during the triangulations. This proves the following theorem:

Theorem 1 *There exists an algorithm, which takes an $m \times n$ matrix M with entries in $\mathcal{R}$ on input, and which computes an invertible $m \times m$ diagonal matrix D and an $m \times n$ upper triangular matrix T with entries in $\mathcal{R}$, such that there exists a matrix $\bar{U}$ with entries in $\mathcal{F}$, of determinant one, and so that $D^{-1}T = \bar{U}M$. Moreover, $\nu(D) \leqslant \min(m,n)\nu(M)$ and $\nu(T) \leqslant \min(m,n)\nu(M)$. Here $\nu(M) = \max_{i,j} \nu(M_{i,j})$.*

By way of comment, we note that the actual computation of T involves $O(mn\min(m,n))$ elementary operations. If we do have an algorithm for exact division in $\mathcal{R}$, then this is also the time complexity of the algorithm in terms of operations in $\mathcal{R}$. Otherwise, it may be necessary to compute the entries of the intermediate matrices T_k by formula (2), which yields an overall complexity of $O(mn(\min(m,n))^3)$.

3.3. Computing greatest common divisors of several polynomials

Let $\mathcal{R}$ still be an effective integral domain with a pseudo-norm ν and quotient field $\mathcal{F}$. Consider a finite number $P_1,\ldots,P_k$ of polynomials in $\mathcal{R}[F]$. In this section, we address the question of computing a g.c.d. $G \in \mathcal{R}[F]$ of $P_1,\ldots,P_k$ and a corresponding Bezout relation. Since we are only computing over an integral domain, we call G *a* g.c.d., if G is a scalar multiple of *the* g.c.d. of $P_1,\ldots,P_k$, when considered as polynomials over the quotient field $\mathcal{F}$. Accordingly, a Bezout relation for $P_1,\ldots,P_k$ has the form

$$Q_1P_1 + \cdots + Q_kP_k = cG, \tag{3}$$

where $Q_1,\ldots,Q_k \in \mathcal{R}[F]$ and $c \in \mathcal{R}^*$. From the computational point of view, we are interested in minimizing the pseudo-norms of $Q_1,\ldots,Q_k,c$ and G. As to the degrees, such "small Bezout relations" always exist. In fact quite a lot of research has been done in this area, see [Kollar (1999)] for example. Here we give a relatively simple result which is sufficient for our purposes.

Proposition 1 *Let $P_1,\ldots,P_k \in \mathcal{R}[F]$ be more than one non-zero polynomial. Then there exists a Bezout relation* (3)*, such that* $\max\{\deg Q_iP_i\} < \max\{\deg P_i + \deg P_j | i \neq j\}$.

Proof Assume the contrary and choose a Bezout relation (3) of minimal degree $d = \max\{\deg Q_i P_i\}$ and such that the number l of indices $i_1 < \cdots < i_l$ with $\deg Q_{i_k} P_{i_k} = d$ is minimal. Since $d > \deg G$, we must have $l > 1$, and modulo a permutation of indices, we may assume that $i_k = k$ for each k. Let λ be the leading coefficient of Q_1 and μ the leading coefficient of P_2. Then for $\delta = d - \deg P_1 - \deg P_2 \geqslant 0$, we have

$$(\mu Q_1 - \lambda x^\delta P_2) P_1 + (\mu Q_2 + \lambda x^\delta P_1) P_2 + \mu Q_3 P_3 + \cdots + \mu Q_k P_k = \mu c G,$$

is again a Bezout relation, which contradicts our minimality hypothesis. □

Let $d = \max\{\deg P_i + \deg P_j | i \neq j\}$. In order to actually find a Bezout relation of degree $< d$, we now consider the matrix M with $m = kd - \deg P_1 - \cdots - \deg P_k$ rows and d columns, which is the vertical superposition of all matrices of the form

$$\begin{pmatrix} P_{i,r_i} & \cdots & P_{i,0} & & & & 0 \\ & P_{i,r_i} & \cdots & P_{i,0} & & & \\ & & \ddots & & \ddots & & \\ & & & P_{i,r_i} & \cdots & P_{i,0} & \\ 0 & & & & P_{i,r_i} & \cdots & P_{i,0} \end{pmatrix},$$

where $r_i = \deg P_i$. Now triangulate M as in the section above

$$D^{-1} T = \bar{U} M \tag{4}$$

and let l be the number of non-zero rows in T. The l-th row of T corresponds to a polynomial linear combination of $P_1, \ldots, P_k$ of minimal degree. In other words, it contains the coefficients of a g.c.d. of $P_1, \ldots, P_k$.

Moreover, we may obtain a Bezout relation when considering the matrix M with an $m \times m$ identity matrix glued at its right hand side. When triangulating this matrix, we obtain a relation of the form

$$D^{-1} (\ T \ | \ T' \ ) = \bar{U} (\ M \ | \ \mathrm{Id}\ ) = (\ \bar{U} M \ | \ \bar{U}\ ),$$

such that the additional part T' of the triangulated matrix gives us the transformation matrix $\bar{U}$ in (4) up to the diagonal matrix D:

$$T' = D \bar{U}.$$

The finite sum which leads to the l-th row in the product $T' M$ now yields the desired Bezout relation. Notice that $c = 1$ in this Bezout relation.

From the complexity point of view, we also notice that at most d rows in M may actually have contributed to the first $l \leqslant d$ rows of T. When replacing M with its restriction to these rows, the above triangulations will therefore yield the same g.c.d. and the same Bezout relation. We have proved

Theorem 2 *Let $P_1, \ldots, P_k \in \mathcal{R}[F]$ be more than one non-zero polynomials. Then there exists an algorithm to compute a g.c.d. of $P_1, \ldots, P_k$ as well as a Bezout relation* (3) *with $c = 1$, such that*

$$\max\{\nu(G), \nu(Q_1), \ldots, \nu(Q_k)\} \leqslant 2(\max\{\nu(P_1), \ldots, \nu(P_k)\})^2.$$

3.4. Making polynomials square-free using pseudo-division

Let $\mathcal{R}$ be an effective integral domain and let $U, V \in \mathcal{R}[F]$ be polynomials over $\mathcal{R}$ with $\deg V \leqslant \deg U$. If $\mathcal{R}$ were actually an effective field, then we might have used the Euclidean division algorithm to obtain the unique expression for U of the form

$$U = QV + R,$$

with $Q, R \in \mathcal{R}[F]$ and $\deg R < \deg V$. However, this algorithm involves divisions and can no longer be used if $\mathcal{R}$ is an integral domain but not a field. Nevertheless, pseudo-division may always be used to obtain the unique expression for U of the form

$$I_V^{\deg U - \deg V + 1} U = QV + R,$$

where I_V is the *leading coefficient* or *initial* of V and $Q, R \in \mathcal{R}[F]$ are such that $\deg R < \deg V$. We also call Q the *pseudo-quotient* and R the *pseudo-remainder* of the division of U by V.

Algorithm `pdiv`
Input: $U, V \in \mathcal{R}[F]$ with $\deg U \geqslant \deg V$.
Output: the pseudo-quotient resp. pseudo-remainder of the division of U by V.

Set $Q := 0$ and $R := U$
For $i := \deg R, \ldots, \deg V$ do
  $Q := I_V Q + R_i F^{\,i - \deg V}$
  $R := I_V R - R_i F^{\,i - \deg V} V$
Return Q

In particular, we may use pseudo-division to make a polynomial P square-free. Namely, if $\deg \texttt{gcd}(P, P') > 0$, then we take $\texttt{sqfree}(P) = \texttt{pdiv}(P, \texttt{gcd}(P, P'))$ to be the square-free part of P.

Proposition 2 *Let $S = \texttt{sqfree}(P)$ of P as above. Then $\nu(S) \leqslant 3\nu(P)^3$.*

Proof Let $G = \texttt{gcd}(P, P')$ and $Q = \texttt{pdiv}(P, G)$. Then $\nu(G) \leqslant 2\nu(P)^2$, by Theorem 2. Also, $I_G^k P = QG$ for $k = \deg P - \deg G + 1 \leqslant \nu(P)$. Hence, $\nu(Q) \leqslant \nu(I_G^k P) \leqslant k\nu(G) + \nu(P) \leqslant 2\nu(P)^3 + \nu(P) \leqslant 3\nu(P)^3$. □

4. Four key lemmas

We recall that derivations on integral domains extend uniquely to their algebraic closures.

Lemma 1 *Let $P \in \mathcal{R}[F]$ be a square-free polynomial and*

$$G = \texttt{gcd}(P, d_1 P, \ldots, d_k P).$$

Consider the factorization

$$G = g(F - h_1) \cdots (F - h_q),$$

with $g \in \mathcal{R}$ and $h_1, \ldots, h_q \in \mathcal{R}^{\mathrm{alg}}$. Then each of the h_p satisfies the partial differential equations (1)*.*

Proof Consider one of the factors $F - h_p$ of G and write

$$P = (F - h_p)Q.$$

For each $i \in \{1, \ldots, k\}$, we have

$$d_i P = (A_i(F) - B_i(F)\partial_i h_p)Q + (F - h_p)d_i Q$$

Since $F - h_p$ both divides $d_i P$ and $(F - h_p)d_i Q$, it also divides $(A_i(F) - B_i(F)\partial_i h_p)Q$. Now P is square-free, so that $F - h_p$ does not divide Q. Therefore $F - h_p$ divides $A_i(F) - B_i(F)\partial_i h_p$. Consequently, $A_i(h_p) - B_i(h_p)\partial_i h_p = 0$ for each i, i.e. h_p satisfies (1). $\square$

Lemma 2 *Let $\varphi \in \mathcal{C}^{\mathrm{alg}}\langle\langle z_1\rangle\rangle \cdots \langle\langle z_k\rangle\rangle$ be a Puiseux series with $\boldsymbol{v}(\varphi) \geqslant \mathbf{0}$. Assume that φ satisfies the same equations* (1) *as f and the same initial condition $\varepsilon(\varphi) = \varepsilon(f)$. Then $\varphi = f$.*

Proof Let us prove by induction over i that $\varepsilon_i(\varphi) = \varepsilon_i(f)$ for all $i \in \{0, \ldots, k\}$. We have $\varepsilon_0(\varphi) = \varepsilon_0(f)$ by assumption. Assume therefore that $i > 0$ and $\varepsilon_{i-1}(\varphi) = \varepsilon_{i-1}(f)$. Then $\psi = \varepsilon_i(\varphi)$ is a Puiseux series in $\mathcal{C}\langle\langle z_1\rangle\rangle \cdots \langle\langle z_i\rangle\rangle$ of the form

$$\psi = \varepsilon_{i-1}(f) + \sum_{\alpha > 0} \psi_\alpha z_i^\alpha. \tag{5}$$

Since ∂_i and ε_i commute, ψ satisfies the partial differential equation

$$\partial_i \psi = \frac{\varepsilon_i(A_i)(\psi)}{\varepsilon_i(B_i)(\psi)}. \tag{6}$$

In particular, extraction of the coefficient in $z_i^{\alpha-1}$ yields

$$\alpha \psi_\alpha = \left(\frac{\varepsilon_i(A_i)(\psi)}{\varepsilon_i(B_i)(\psi)}\right)_{\alpha-1} \tag{7}$$

for every $\alpha > 0$. Now

$$(\varepsilon_i(B_i)(\psi))_0 = \varepsilon_{i-1}(B_i)(\psi_0) = \varepsilon_{i-1}(B_i(f)) \neq 0,$$

since $\varepsilon(\varepsilon_{i-1}(B_i(f))) = \varepsilon(B_i(f)) \neq 0$. Consequently, we may see (7) as a recurrence relation which uniquely determines ψ_α as a function of other ψ_β with $\beta < \alpha$. Hence $\psi = \varepsilon_i(f)$ is the unique solution to (6) of the form (5). The lemma now follows by induction. $\square$

Lemma 3 *Let P be a polynomial in $\mathcal{R}[F]$. Given $\lambda \in \mathcal{C}^{\mathrm{alg}}$ with*

$$P_{\boldsymbol{v}(P)}(\lambda) = 0, \tag{8}$$

there exists a root $\varphi \in \mathcal{C}^{\mathrm{alg}}\langle\langle z_1\rangle\rangle \cdots \langle\langle z_k\rangle\rangle$ of P with $\boldsymbol{v}(\varphi) \geqslant \mathbf{0}$ and $\varepsilon(\varphi) = \lambda$.

Proof Intuitively speaking, this lemma follows from the Newton polygon method: the existence of a solution $\lambda \neq 0$ to (8) implies that the Newton polygon associated to the equation $P(\varphi) = 0$ admits a horizontal slope and that λ is a solution to the associated Newton polynomial. Therefore, λ is the first term of a solution to $P(\varphi) = 0$, the full solution being obtained using the Newton polygon method. If $\lambda = 0$, then the Newton polygon admits a "strictly positive slope" and a similar argument applies.

More precisely, we may apply the results from chapter 3 in [van der Hoeven (2004)] (see also [van der Hoeven (1997)]). We first note that

$$\mathcal{C}^{\text{alg}}\langle\langle z_1\rangle\rangle \cdots \langle\langle z_k\rangle\rangle = \mathcal{C}^{\text{alg}}[[z_1^{\mathbb{Q}}; \ldots; z_k^{\mathbb{Q}}]]$$

is a field of grid-based power series. Now setting $\varphi = \lambda + \psi$, the Newton degree d of

$$P_{+\lambda}(\psi) = P(\lambda + \psi) = 0 \qquad (\boldsymbol{v}(\psi) > \mathbf{0}) \tag{9}$$

is strictly positive. Indeed, this is clear if $\lambda = 0$, and this follows from Lemma 3.6 in [van der Hoeven (2004)] if $\lambda \neq 0$. Now our lemma follows from the fact that the algorithm `polynomial_solve` returns d solutions to (9). □

Lemma 4 *With the notation from Lemma 1, we have*

$$\rho(P) = 0 \iff G_{\boldsymbol{v}(G)}(\varepsilon(f)) = 0.$$

Proof Assuming that $\rho(P) = 0$, we have $\rho(d_i P) = d_i \rho(P) = 0$ for $1 \leqslant i \leqslant k$. It follows that $F - f$ divides both P and $d_1 P, \ldots, d_k P$, whence $F - f | G$ and $\rho(G) = 0$. Since $G(f) = G_{\boldsymbol{v}(G)}(\varepsilon(f)) z^{\boldsymbol{v}(G)} + \cdots$, it follows in particular that $G_{\boldsymbol{v}(G)}(\varepsilon(f)) = 0$.

Assume now that $G_{\boldsymbol{v}(G)}(\varepsilon(f)) = 0$. Lemma 3 implies that G admits a root $\varphi \in \mathcal{C}^{\text{alg}}\langle\langle z_1\rangle\rangle \cdots \langle\langle z_k\rangle\rangle$ with $\boldsymbol{v}(\varphi) \geqslant \mathbf{0}$ and $\varepsilon(\varphi) = \varepsilon(f)$. This root φ satisfies the equations (1), by Lemma 1. Hence $\varphi = f$, by Lemma 2. We conclude that $G(f) = 0$ and $\rho(P) = P(f) = 0$. □

5. The algorithm

5.1. Statement of the algorithm

The lemmas from the previous section yield the following zero-test algorithm:

Algorithm `zero_test`
Input: $P \in \mathcal{R}[F]$.
Output: result of the test $\rho(P) = 0$.

Step 1. [trivial case]
If $P = 0$ then return `true`.
Step 2. [g.c.d. computations]
Replace $P :=$ `sqfree`(P).
Let $G :=$ `gcd`$(P, d_1 P, \ldots, d_k P)$.
Step 3. [compute the valuation $\boldsymbol{\alpha}$ of G]
Denote $G = G_q F^q + \cdots + G_0$.
For $i = k, \ldots, 1$, compute α_i as a function of $\alpha_j, \ldots, \alpha_{i+1}$ as follows:
Expand each coefficient $G_{j,\alpha_k,\ldots,\alpha_{i+1}}$ $(j = 0, \ldots, q)$ w.r.t. z_i.
Stop at the least α_i, such that there exists a p with $G_{p,\alpha_k,\ldots,\alpha_i} \neq 0$.
Step 4. [evaluate and conclude]
Return $G_{\boldsymbol{\alpha}}(\varepsilon(f)) = 0$.

Remark 3 The expansion of $G_{j,\alpha_k,\ldots,\alpha_{i+1}}$ in step 3 may be done efficiently using the technique of relaxed evaluation [van der Hoeven (2002b)].

5.2. Complexity bounds

In order to derive complexity bounds, we will have to assume that we have a pseudo-norm ν on $\mathcal{R}$ and that there exists a function $\xi_{\mathcal{R}} : \mathbb{N} \to \mathbb{N}$, which gives a bound $|\boldsymbol{v}(\varphi)| \leqslant \xi_{\mathcal{R}}(\nu(\varphi))$ on the valuation $\boldsymbol{v}(\varphi)$, for each $\varphi \in \mathcal{R}\backslash\{0\}$. Here we understand that

$$|(\alpha_1, \ldots, \alpha_k)| = \max\{\alpha_1, \ldots, \alpha_k\},$$

for each $\boldsymbol{\alpha} \in \mathbb{N}^k$. It is also reasonable to also assume that $\xi_{\mathcal{R}}$ is increasing and that it grows sufficiently fast such that $\xi_{\mathcal{R}}(c) \geqslant c$, $\xi_{\mathcal{R}}(c+d) \geqslant \xi_{\mathcal{R}}(c) + \xi_{\mathcal{R}}(d)$ and $\xi_{\mathcal{R}}(cd) \geqslant \xi_{\mathcal{R}}(c)\xi_{\mathcal{R}}(d)$ for all $c, d \in \mathbb{N}$. Notice that we may take $\xi_{\mathcal{C}}(c) = c$ for all $c \in \mathbb{N}$ when $\mathcal{R} = \mathcal{C}$.

Theorem 3 *Let*

$$C = \max\{K_{\mathcal{R}} + \max\{\nu(B_1), \ldots, \nu(B_k)\}, \max\{\nu(A_1), \ldots, \nu(A_k)\}, 1\}.$$

Then for any $\varphi \in \mathcal{S}\backslash\{0\}$, we have

$$|\boldsymbol{v}(\varphi)| \leqslant \xi_{\mathcal{R}}((2kC\nu(\varphi))^7).$$

Proof Assume first that $\varphi \in \mathcal{S}\backslash\{0\}$ can be represented by a square-free polynomial $P \in \mathcal{R}[F]$. Then P does not change in step 2 of `zero_test` and for $i \in \{1, \ldots, k\}$, we have

$$\nu(d_i P) \leqslant \nu(P) + C,$$

Then Theorem 2 yields

$$\nu(G) \leqslant 2(\nu(P) + C)^2.$$

Since $|\boldsymbol{v}(G_p)| \leqslant \xi_{\mathcal{R}}(\nu(G))$ for each non-zero coefficient G_p of G, it follows that

$$|\boldsymbol{\alpha}| \leqslant \xi_{\mathcal{R}}(2(\nu(P) + C)^2)$$

in step 3 of `zero_test`. Since we assumed $\varphi \neq 0$, we must have $G_{\boldsymbol{\alpha}}(\varepsilon(f)) \neq 0$ in the last step of `zero_test`. Considering the Taylor series expansion

$$\begin{aligned} G(f) &= G(\varepsilon(f)) + G'(\varepsilon(f))\delta + \frac{1}{2}G''(\varepsilon(f))\delta^2 + \cdots \\ &= G_{\boldsymbol{\alpha}}(\varepsilon(f)) z_1^{\alpha_1} \cdots z_k^{\alpha_k} + o(z_1^{\alpha_1} \cdots z_k^{\alpha_k}) \end{aligned}$$

in the infinitesimal power series $\delta = f - \varepsilon(f)$, we observe that $\boldsymbol{v}(\rho(G)) = \boldsymbol{\alpha}$.

By Theorem 2, G also satisfies a Bezout relation of the form

$$G = SP + Q_1 d_1 P + \cdots + Q_k d_k P.$$

Now $|\boldsymbol{v}(\rho(d_i P))| = |\boldsymbol{v}(\rho(B_i)\partial_i \rho(P))| = |\boldsymbol{v}(\partial_i \rho(P))| \geqslant |\boldsymbol{v}(\rho(P))| - 1$ for all i (recall that $\varepsilon(B_i) \neq 0$), so that

$$|\boldsymbol{v}(\rho(SP + Q_1 d_1 P + \cdots + Q_k d_k P))| \geqslant |\boldsymbol{v}(\rho(P))| - 1.$$

We conclude that

$$|\boldsymbol{v}(\rho(P))| \leqslant \xi_{\mathcal{R}}(2(\nu(P) + C)^2 + 1).$$

This gives a bound for $\boldsymbol{v}(\varphi)$ in the case when P is square-free.

Let us now turn to the more general situation in which φ is represented by a polynomial $P \in \mathcal{R}[F]$ which is no longer square-free. Setting $P_* := \mathtt{pdiv}(P, \mathtt{gcd}(P, \partial P/\partial F))$, the above discussion yields the bound

$$|\boldsymbol{v}(\rho(P_*))| \leqslant \xi_{\mathcal{R}}(2(3\nu(P)^3 + C)^2 + 1),$$

since $\nu(P_*) \leqslant 3\nu(P)^3$ by Proposition 2. Now P divides $P_*^{\deg P}$ when we understand these polynomials to have coefficients in the quotient field of $\mathcal{R}$. If c is the leading coefficient of P, we thus have $P | c^{(\deg P)^2} P_*^{\deg P}$ in $\mathcal{R}[F]$, as is seen by pseudo-dividing $P_*^{\deg P}$ by P. It follows that

$$\begin{aligned} |\boldsymbol{v}(\rho(P))| &\leqslant \nu(P)|\boldsymbol{v}(\rho(P_*))| + \nu(P)^2|\boldsymbol{v}(c)| \\ &\leqslant \nu(P)\xi_{\mathcal{R}}(2(3\nu(P)^3 + C)^2 + 1) + \nu(P)^2\xi_{\mathcal{R}}(\nu(P)) \\ &\leqslant \xi_{\mathcal{R}}(2\nu(P)(3\nu(P)^3 + C + 1)^2), \end{aligned}$$

since $|\boldsymbol{v}(c)| \leqslant \xi_{\mathcal{R}}(\nu(P))$.

Let us finally consider the case when φ is represented by a general element $P \in \check{\mathcal{S}}$. Then we may rewrite P as a fraction Φ/Ψ with $\Phi \in \mathcal{R}[F]$ and $\Psi = B_1^{\nu(P)} \cdots B_k^{\nu(P)}$, and we have

$$\nu(\Phi) \leqslant \nu(P) + k\nu(P)\max\{\nu(B_1), \ldots, \nu(B_k)\} \leqslant kC\nu(P).$$

Since $\boldsymbol{v}(\rho(\Psi)) = \mathbf{0}$, we thus get

$$\begin{aligned} |\boldsymbol{v}(\rho(P))| &\leqslant \xi_{\mathcal{R}}(2\nu(\Phi)(3\nu(\Phi)^3 + C + 1)^2) \\ &\leqslant \xi_{\mathcal{R}}(2kC\nu(P)(3(kC\nu(P))^3 + C + 1)^2) \\ &\leqslant \xi_{\mathcal{R}}(50(kC\nu(P)^7)), \end{aligned}$$

since $\nu(P) = 0$ or $C + 1 \leqslant 2(kC\nu(P))^3$. □

Remark 4 It is plausible that the second and third part in the above proof may be further optimized, so as to reduce the exponent from 7 to 2. In the second part, one might for instance consider the factorization of P instead of $\mathtt{gcd}(P, \partial P/\partial F)$ as in the zero-test algorithm.

As a consequence of the above bound for the valuations of non-zero series $\varphi \in \mathcal{S}$, we have a straightforward zero-test algorithm for series $\varphi \in \mathcal{S}$ which consists of testing whether all coefficients of φ up to the bound vanish using relaxed evaluation [van der Hoeven (2002b)]. This algorithm satisfies the following complexity bound:

Theorem 4 *Let $P \in \check{\mathcal{S}}$. With the notation from Theorem 3, we may test whether P represents zero in time $O(\xi_{\mathcal{R}}((2kC\nu(\varphi))^7)^k \log^2 \xi_{\mathcal{R}}((2kC\nu(\varphi))^7)k^3)$.*

Remark 5 Of course, the complexity bound from Theorem 4 is very pessimistic, since it reflects the theoretical worst case bounds for the valuations. In practice, we recommend using `zero_test`, which we expect to be much faster in average.

Consider a tower of regular D-algebraic ring extensions $\mathcal{R}_0 \subseteq \mathcal{R}_1 \subseteq \cdots \subseteq \mathcal{R}_h$ starting with $\mathcal{R}_0 = \mathcal{C}[z_1, \ldots, z_k]$. We have natural representations

$$\rho : \check{\mathcal{R}}_i = \check{\mathcal{R}}_{i-1}\left[F_i, \frac{1}{B_{i,1}(F_i)}, \ldots, \frac{1}{B_{i,n_i}(F_i)}\right] \to \mathcal{R}_i$$

for all $i \in \{1, \ldots, k\}$. The repeated application of Theorem 3 yields

Corollary 1 *There exists a constant K, such that for all $P \in \check{\mathcal{R}}_h$, we have either $\rho(P) = 0$ or $|\boldsymbol{v}(\rho(P))| \leqslant K\nu(P)^{7^h}$.*

Remark 6 In other words, for fixed h, we have a polynomial time algorithm zero-test in $\mathcal{R}_h$. Theoretically speaking, we already knew this, because $\mathcal{R}_h \cong \check{\mathcal{R}}_h/I$ for a certain ideal I of $\check{\mathcal{R}}_h$. Hence, it would suffice to reduce a polynomial in $\check{\mathcal{R}}_h$ with respect to a Groebner basis for I in order to know whether it represents zero. Unfortunately, we do not know of any algorithm to compute such a Groebner basis for I. Nevertheless, even without such a Groebner basis the above corollary tells us that we still have a polynomial time zero-test.

Let us now return to exp-log series in the ring $\mathcal{E}_k$ considered in the introduction. Recall that the size of an element in $\mathcal{E}_k$ is the number of nodes in the corresponding expression tree. Repeated application of Theorem 3 yields

Corollary 2 *Consider an exp-log series $f \in \mathcal{E}_k$, which can be represented by an expression of size σ. Then either $f = 0$ or $|\boldsymbol{v}(f)| \leqslant (4k\sigma)^{7^\sigma}$.*

Proof Let $\check{f}$ be an expression which represents f and let $\check{f}_1, \ldots, \check{f}_\sigma$ be its subexpressions listed in the order of a postfix traversal. We construct a tower $\mathcal{R}_0 \subseteq \cdots \subseteq \mathcal{R}_h$ with representations $\check{\mathcal{R}}_0 \subseteq \cdots \subseteq \check{\mathcal{R}}_h$, such that $\check{f}_i \in \mathcal{R}_{p_i}$ for all i and $0 = p_1 \leqslant \cdots \leqslant p_\sigma = h$. We construct the tower by induction over i. For $i = 0$ we have nothing to show, so suppose $i > 0$ and that we have performed the construction up to stage $i - 1$.

If $\check{f}_i \in \mathcal{C}$ or $\check{f}_i \in \{z_1, \ldots, z_k\}$, then we clearly have $\check{f}_i \in \check{\mathcal{R}}_0 = \mathcal{C}[z_1, \ldots, z_k]$. If $\check{f}_i = \check{f}_{j_1} + \check{f}_{j_2}$, $\check{f}_i = \check{f}_{j_1} - \check{f}_{j_2}$, $\check{f}_i = \check{f}_{j_2} - \check{f}_{j_1}$ or $\check{f}_i = \check{f}_{j_1}\check{f}_{j_2}$ with $j_1 < j_2 < i$, then $\check{f}_i \in \check{\mathcal{R}}_{j_2}$. Assume finally that $\check{f}_i = \varphi \circ \check{f}_j$ with $j < i$, where $\varphi \in \{1/(1+z), \exp z, \log(1+z)\}$. Then we take $\check{\mathcal{R}}_{p_i} = \check{\mathcal{R}}_{p_{i-1}}[\check{f}_i]$ if $\varphi = \exp z$ or $\check{\mathcal{R}}_{p_i} = \check{\mathcal{R}}_{p_{i-1}}[\check{f}_i, 1/(1+\check{f}_i)]$ otherwise, and one the relations

$$\begin{aligned}
\partial f_i &= -(\partial f_j) f_i^2;\\
\partial f_i &= (\partial f_j) f_i;\\
\partial f_i &= \frac{\partial f_j}{1 + f_j}
\end{aligned}$$

holds for all $\partial \in \{\partial_1, \ldots, \partial_k\}$. Notice that the pseudo-norm of $\check{f}_i$ is bounded by i (whence by σ) for all i. Consequently, the C from Theorem 3 is bounded by 2σ at each stage. By induction over i, it therefore follows that $\xi_{\mathcal{R}_i}(s) \leqslant (4k\sigma)^{\frac{7^{i-1}-1}{6}} s^{7^{i-1}}$ for all $i \geqslant 1$. If $f \neq 0$, we conclude that $|\boldsymbol{v}(f)| \leqslant \xi_{\mathcal{R}_\sigma}(\sigma) \leqslant (4k\sigma)^{7^\sigma} = k^{O(1)^\sigma}$. □

References

Ax, J., 1971. On Schanuel's conjecture. Ann. of Math. 93, 252–268.

Bareiss, E., 1968. Sylvester's identity and multistep integer-preserving Gaussian elimination. Math. Comp. 22 22, 565–578.

Caviness, B., Prelle, M., 1978. A note on algebraic independence of logarithmic and exponential constants. SIGSAM Bull. 12/2, 18–20.

Gabrielov, A., Vorobjov, N., April 2004. Complexity of computations with Pfaffian and Noetherian functions. In: et al, Y. I. (Ed.), Normal forms, bifurcations and finiteness problems in differential equations. Vol. 137 of NATO Science series II. Kluwer, to appear.

Khovanskii, A., 1991. Fewnomials. Vol. 88 of Translations of Mathematical Monographs. A.M.S., Providence RI.

Kollar, J., 1999. Effective nullstellensatzn for arbitrary ideals. Vol. 1. pp. 313–337.

Lang, S., 1971. Transcendental numbers and diophantine approximation. Bull. Amer. Math. Soc. 77/5, 635–677.

Loos, R., 1983. Generalized polynomial remainder sequences. In: Buchberger, B., Collins, G., Loos, R. (Eds.), Computer Algebra: Symbolic and Algebraic Computation. Springer-Verlag, Wien/New York, pp. 115–137.

Macintyre, A., Wilkie, A., 1995. On the decidability of the real exponential field. In: Odifreddi, P. (Ed.), Kreisel 70th birthday volume. CLSI. A.K. Peters.

Péladan-Germa, A., 1995. Testing identities of series defined by algebraic partial differential equations. In: Cohen, G., Giusti, M., Mora, T. (Eds.), Proc. of AAECC-11. Vol. 948 of Lect. Notes in Comp. Science. Springer, Paris, pp. 393–407.

Richardson, D., 1997. How to recognise zero. JSC 24, 627–645.

Richardson, D., 2001. The uniformity conjecture. In: Lecture Notes in Computer Science. Vol. 2064. Springer Verlag, pp. 253–272.

Shackell, J., 1989. A differential-equations approach to functional equivalence. In: Proc. ISSAC '89. ACM Press, Portland, Oregon, A.C.M., New York, pp. 7–10.

Shackell, J., 1993. Zero equivalence in function fields defined by differential equations. Proc. of the A.M.S. 336 (1), 151–172.

Shackell, J., 2004. Symbolic asymptotics. Vol. 12 of Algorithms and computation in Mathematics. Springer-Verlag.

van der Hoeven, J., 1997. Automatic asymptotics. Ph.D. thesis, École polytechnique, France.

van der Hoeven, J., 2001a. Fast evaluation of holonomic functions near and in singularities. JSC 31, 717–743.

van der Hoeven, J., 2001b. Zero-testing, witness conjectures and differential diophantine approximation. Tech. Rep. 2001-62, Prépublications d'Orsay.

van der Hoeven, J., July 2002a. A new zero-test for formal power series. In: Mora, T. (Ed.), Proc. ISSAC '02. Lille, France, pp. 117–122.

van der Hoeven, J., 2002b. Relax, but don't be too lazy. JSC 34, 479–542.

van der Hoeven, J., 2003. Counterexamples to witness conjectures. Tech. Rep. 2003-43, Université Paris-Sud, Orsay, France, to appear in JSC.

van der Hoeven, J., 2004. Transseries and real differential algebra. Tech. Rep. 2004-47, Université Paris-Sud, Orsay, France, to appear in Lecture Notes in Mathematics, Springer-Verlag.