From implicit to recursive equations*

Joris van der Hoeven

LIX, CNRS
École polytechnique
91128 Palaiseau Cedex
France

Email: `vdhoeven@lix.polytechnique.fr`
Web: `http://lix.polytechnique.fr/~vdhoeven`

August 2, 2011

The technique of relaxed power series expansion provides an efficient way to solve so called recursive equations of the form $F = \Phi(F)$, where the unknown F is a vector of power series, and where the solution can be obtained as the limit of the sequence 0, $\Phi(0)$, $\Phi(\Phi(0))$, ... With respect to other techniques, such as Newton's method, two major advantages are its generality and the fact that it takes advantage of possible sparseness of Φ. In this paper, we consider more general implicit equations of the form $\Phi(F) = 0$. Under mild assumptions on such an equation, we will show that it can be rewritten as a recursive equation. If we are actually computing with analytic functions, then recursive equations also provide a systematic device for the computation of verified error bounds. We will show how to apply our results in this context.

Keywords: Implicit equation, relaxed power series, algorithm

A.M.S. subject classification: 68W25, 42-04, 68W30, 65G20, 30B10

1. Introduction

Let $\mathbb{K}$ be an effective field of constants of characteristic zero. This means that elements in $\mathbb{K}$ can be encoded by data structures on a computer and that we have algorithms for performing the field operations of $\mathbb{K}$.

Let $F = (F^{[1]}, \ldots, F^{[r]})$ be a column vector of r indeterminate series in $\mathbb{K}[[z]]$. We may also consider F as a power series $F_0 + F_1 z + \cdots \in \mathbb{K}^r[[z]]$. Let $\Phi(F) = (\Phi(F)^{[1]}, \ldots, \Phi(F)^{[r]})$ be a column vector of expressions built up from F, z and constants in $\mathbb{K}$ using ring operations, differentiation and integration (with constant term zero). Finally, let $C_0, \ldots, C_{l-1} \in \mathbb{K}^r$ be a finite number of initial conditions. Assume that the system

$$\left\{ \begin{array}{rcl} \Phi(f) & = & 0 \\ f_0 & = & C_0 \\ & \vdots & \\ f_{l-1} & = & C_{l-1} \end{array} \right. \tag{1}$$

admits a unique solution $f \in \mathbb{K}[[z]]^r$. In this paper, we are interested in the efficient computation of this solution up to a given order n.

In the most favourable case, the equation $\Phi(f) = 0$ is of the form

$$f - \Psi(f) = 0, \tag{2}$$

*. This work has been supported by the ANR-09-JCJC-0098-01 MaGiX and ANR-10-BLAN-0109 projects, as well as a Digiteo 2009-36HD grant and Région Ile-de-France.

where the coefficient $\Psi(f)_n$ of z^n in $\Psi(f)$ only depends on earlier coefficients $f_0, \ldots, f_{n-1}$ of f, for each $n \in \mathbb{N}$. In that case,

$$f_n = \Psi(f)_n$$

actually provides us with a recurrence relation for the computation of the solution. Using the technique of relaxed power series expansions [vdH02, vdH07a], which will briefly be recalled in section 2.4, it is then possible to compute the expansion $f_{;n} = f_0 + \cdots + f_{n-1} z^{n-1}$ up till order n in time

$$\mathsf{T}(n) = s\,\mathsf{R}(n) + O(t\,n), \tag{3}$$

where s is the number of multiplications occurring in Ψ, t is the total size of Ψ as an expression, and $\mathsf{R}(n)$ denotes the complexity of relaxed multiplication of two power series up till order n. Here we assume that Ψ is represented by a directed acyclic graph, with possible common subexpressions. For large n, we have $\mathsf{R}(n) = O(\mathsf{M}(n) \log n)$, where $\mathsf{M}(n) = O(n \log n \log \log n)$ denotes the complexity [CT65, SS71, CK91] of multiplying two polynomials of degrees $< n$. If $\mathbb{K}$ admits sufficiently many 2^p-th roots of unity, then we even have $\mathsf{R}(n) = O(\mathsf{M}(n)\, \mathrm{e}^{2\sqrt{\log 2 \log \log n}})$ and $\mathsf{M}(n) = O(n \log n)$. For moderate n, when polynomial multiplication is done naively or using Karatsuba's method, relaxed multiplication is as efficient as the truncated multiplication of polynomials at order n.

One particularly important example of an equation of the above type is the integration of a dynamical system

$$f = f_0 + \textstyle\int \Psi(f), \tag{4}$$

where Ψ is algebraic (i.e. does not involve differentiation or integration). In that case, given the solution f up till order n, we may consider the linearized system

$$E' = \Psi(f) + J_{\Psi}(f)\, E + O(z^{2n})$$

up till order $2\,n$, where $J_{\Psi}(f)$ stands for the Jacobian matrix associated to Ψ at f. If we have a fundamental system of solutions of $E' = J_{\Psi}(f)\, E$ up till order n, then one step of Newton's method allows us to find the solution of (4) and a new fundamental system of solutions of the linearized equation up till order $2\,n$ [BK78, BCO+06]. A careful analysis shows that this leads to an algorithm of time complexity

$$\mathsf{T}(n) = \mathsf{M}(n)\,(2\,s\,r + 2\,s + 13/6\,r^2 + 4/3\,r + o(1)) + O(t\,r\,n). \tag{5}$$

In [vdH10], this bound has been further improved to

$$\mathsf{T}(n) = \mathsf{M}(n)\,(2\,s + 4/3\,r + o(1)) + O(t\,n), \tag{6}$$

under the assumptions that $\mathbb{K}$ admits sufficiently many 2^p-th roots of unity and that $r = O(\log n)$.

Although the complexity (5) is asymptotically better than (3) for very large n, the relaxed approach often turns out to be more efficient in practice. Indeed, Newton's method both suffers from a larger constant factor and the fact that we profit less from the potential sparsity of the system. In particular, if $r > \log n$, then the relaxed approach is generally faster. Moreover, as long as multiplications are done in the naive or Karatsuba model, the relaxed approach is optimal in the sense that the computation of the solution takes roughly the same time as its verification. Another advantage of the relaxed approach is that it generalizes to more general functional equations and partial differential equations.

Let us now return to our original implicit system (1). If Φ is a system of differentially algebraic equations, then we may also seek to apply Newton's method. For non degenerate systems and assuming that we have computed the solution f and a fundamental system of solutions for the linearized equation up till order n, one step of Newton's method yields an extension of the solutions up till order $2\,n-i$, for a fixed constant $i\in\mathbb{N}$. From an asymptotic point of view, this means that the complexities (5) and (6) remain valid, modulo multiplication of r by the differential order of the system in these bounds.

Another approach for the resolution of (1) is to keep differentiating the system with respect to f until it becomes equivalent to a system of the form (2). For instance, if Φ is algebraic, then differentiation of (1) yields

$$J_{\Phi}(f)\,f'+\frac{\partial\Phi}{\partial z}(f) \;=\; 0.$$

Consequently, if $J_{\Phi}(f)_0$ is invertible, then

$$f \;=\; f_0-\textstyle\int J_{\Phi}(f)^{-1}\,\frac{\partial\Phi}{\partial z}(f)$$

provides us with an equivalent system which can be solved by one of the previous methods. Unfortunately, this method requires the computation of the Jacobian, so we do not longer exploit the potential sparsity of the original system.

Yet another recent approach [vdH09b] is to consider not yet computed coefficients of f as formal unknowns, and solve the system of equations $\Phi(f)_0=\cdots=\Phi(f)_n=0$ for increasing values of n. For large n, the system $\Phi(f)_0=\cdots=\Phi(f)_n=0$ usually reduces to a linear system of equations. In particular, the coefficients of series with unknown coefficients are not polynomials but merely linear combinations. Using the so called "substitution product", the multiplication of series with unknown coefficients can be done while taking advantage of this linearity.

In this paper, we will present a variant of the approach of [vdH09b]. Roughly speaking, we reconsider the series with unknown coefficients as vectors of partially unknown series. Technically speaking, this is done *via* the concept of anticipators, which will be introduced in section 3. Using this technique, and under mild assumptions, we show in section 4.1 how to rewrite the original system of equations into a new system which is both recursive and not much larger in size. We may then apply a standard relaxed algorithm for its resolution. This leads to slightly sharper complexity bounds than those from [vdH09b], which will be presented in section 4.2.

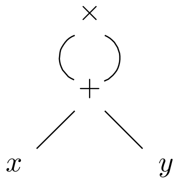
The main interest of the new technique though is the fact that it allows for the direct resolution of implicit equations by existing software for recursive equations. Moreover, recursive equations naturally occur in the area of reliable computing [Moo66, AH83, Neu90, MKC09, Rum10]. In this setting, our power series are replaced by so called Taylor models [MB96, MB04], which systematically require certified bounds for the remainders on polydisks. In section 5, we will show how our results apply in this case. In particular, the new algorithms are an asymptotically efficient device for the certified resolution of implicit equations and integration of dynamical systems on implicitly defined varieties.

Remark 1. Just before prepublication of this paper, it turned out that BERTHOMIEU and LEBRETON independently discovered the technique of section 3 in the index one case and in the setting of p-adic numbers instead of power series [BL11]. They also implemented their ideas in the MATHEMAGIX system and their first timings are very encouraging.

2. Preliminaries

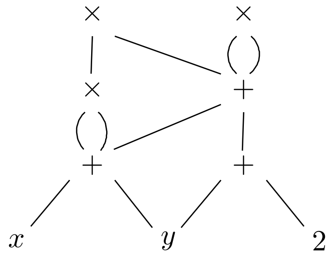
2.1. Dags

Assume that we have fixed a set Ω of function symbols, together with an *arity* n_ω for each $\omega \in \Omega$. Then a *dag* over Ω is a rooted finite directed acyclic Ω-labeled graph, such that each ω-labeled node has n_ω successors, together with an ordering on the successors. For instance,



is a typical dag for the expression $(x+y)^2$, with $\Omega = \{x, y, +, \times\}$, $n_x = n_y = 0$ and $n_+ = n_\times = 2$. We will denote by t_Φ the number of nodes of a dag Φ (also called its size) and by s_Φ its number of multiplications (also called its multiplicative size). For our example dag Φ, we thus have $t_\Phi = 4$ and $s_\Phi = 1$. We will denote by $\mathbb{D}_\Omega$ the set of dags over Ω.

More generally, we may consider *multivariate dags* with an arbitrary number of roots, which again come with ordering. For instance,



is a bivariate dag which represents a vector of two expressions $(x+y)^2((x+y)+(y+2))$ and $((x+y)+(y+2))^2$. We will denote by d_Φ the number of roots of a multivariate dag Φ, which we will also call its *dimension*. We will write $\Phi = (\Phi^{[1]}, \ldots, \Phi^{[d_\Phi]})$, where $\Phi^{[i]}$ stands for the subdag whose root is the i-th root of Φ. We will denote by $\mathbb{D}_\Omega^d$ the set of multivariate dags over Ω of dimension d and $\mathbb{D}_\Omega^* = \mathbb{D}_\Omega^1 \cup \mathbb{D}_\Omega^2 \cup \cdots$.

2.2. Dags as operators on power series

Consider a linear operator $\omega: \mathbb{K}[[z]] \to \mathbb{K}[[z]]$. We say that ω is a *coefficientwise* operator, if there exist fixed constants $\omega_0, \omega_1, \omega_2, \ldots \in \mathbb{K}$ such that

$$\omega(f_0 + f_1 z + f_2 z^2 + \cdots) = \omega_0 f_0 + \omega_1 f_1 z + \omega_2 f_2 z^2 + \cdots,$$

for all $f = \sum_i f_i z^i \in \mathbb{K}[[z]]$. For every $c \in \mathbb{K}$, the operator $\times_c$ defined by

$$\times_c(f) = c f$$

is an example of a coefficientwise operator. The truncation operators $\top_i: f \mapsto f^{\lfloor i \rfloor}$, which are defined by

$$f^{\lfloor i \rfloor} = f_i z^i + f_{i+1} z^{i+1} + \cdots$$

constitute another family of examples. We will denote by $\times_\mathbb{K}$ and $\top_\mathbb{N}$ the sets of all operators of the form $\times_c$ with $c \in \mathbb{K}$ resp. $\top_i$ with $i \in \mathbb{N}$. Finally, we define the coefficientwise operators $\delta = z\, {}^{\partial}\!/\!_{\partial z}$ and δ^{-1} by

$$\begin{aligned}\delta(f_0 + f_1 z + f_2 z^2 + \cdots) &= f_1 z + 2 f_2 z^2 + \cdots\\ \delta^{-1}(f_0 + f_1 z + f_2 z^2 + \cdots) &= f_1 z + \tfrac{1}{2} f_2 z^2 + \cdots,\end{aligned}$$

and we notice that δ^{-1} is the inverse of δ on $z\,\mathbb{K}[[z]]$.

Let $F=(F^{[1]},\ldots,F^{[r]})$ be r "indeterminate series" in $\mathbb{K}[[z]]$. We will sometimes consider F as a series with formal coefficients

$$\begin{aligned} F &= F_0+F_1 z^1+\cdots, \\ F_n &= (F_n^{[1]},\ldots,F_n^{[r]}). \end{aligned}$$

Let Λ be a set of coefficientwise linear operators. In what follows, we will take

$$\Omega_{F,\Lambda} \;=\; \mathbb{K}[z]\cup\left\{F^{[1]},\ldots,F^{[r]},+,-,\times\right\}\cup\times_{\mathbb{K}}\cup\top_{\mathbb{N}}\cup\Lambda$$

and denote by $\mathbb{D}_{F,\Lambda}$ the set of dags over $\Omega_{F,\Lambda}$. Similarly, we set $\mathbb{D}_{F,\Lambda}^d=\mathbb{D}_{\Omega_{F,\Lambda}}^d$ and $\mathbb{D}_{F,\Lambda}^*=\mathbb{D}_{\Omega_{F,\Lambda}}^*$. Dags in $\mathbb{D}_{F,\varnothing}^*$, $\mathbb{D}_{F,\{\delta\}}^*$ and $\mathbb{D}_{F,\{\delta^{-1}\}}^*$ will respectively be called *algebraic*, *differential* and *integral*. Notice that polynomials in $\mathbb{K}[z]$ are regarded as dags of size 1, independently of their degree; this is motivated by the fact that coefficient extraction is trivial for explicit polynomials.

Clearly, any dag $\Phi\in\mathbb{D}_{F,\Lambda}^d$ can be considered as a function $\mathbb{K}[[z]]^r\to\mathbb{K}[[z]]^d; f\mapsto\Phi(f)$. Given a small symbolic perturbation $E=\left(E^{[1]},\ldots,E^{[r]}\right)\in z^n\,\mathbb{K}[[z]]^r$, we may expand $\Phi(F+E)$ as a Taylor series in E

$$\Phi(F+E) \;=\; \Phi(F)+(D\Phi)(F)(E)+\tfrac{1}{2}(D^2\Phi)(F)(E)+\cdots$$

and truncation at order $2n$ yields

$$\Phi(F+E) \;=\; \Phi(F)+(D\Phi)(F)(E)+O(z^{2n}).$$

We claim that $(D\Phi)(F)(E)$ can be regarded as a dag in $\mathbb{D}_{(F,E),\Lambda}$. For instance, if

$$\Phi(F) \;=\; F\delta F+z\,\delta^{-1}(F\delta^{-1}F),$$

then

$$(D\Phi)(F)(E) \;=\; E\,\delta F+F\delta E+z\,\delta^{-1}(E\,\delta^{-1}F+F\delta^{-1}E).$$

In general, the claim is easily shown by induction over t_Φ.

2.3. Dags as series

We claim that any dag $\Phi\in\mathbb{D}_{F,\Lambda}^d$ can be regarded as a series in z

$$\Phi \;=\; \Phi_0+\Phi_1 z+\cdots,$$

such that each coefficient Φ_n is a dag over

$$\Omega_{F,\Lambda;n} \;=\; \mathbb{K}\cup\{F_0^{[1]},\ldots,F_0^{[r]},\ldots\ldots,F_n^{[1]},\ldots,F_n^{[r]},+,-,\times\}\cup\times_{\mathbb{K}}\cup\Lambda.$$

Indeed, by induction over the size of Φ, we first define the valuation v_Φ of Φ by

$$\begin{aligned} v_P &= \operatorname{val}P \qquad (P\in\mathbb{K}[z]) \\ v_{F^{[i]}} &= 0 \\ v_{\Psi\pm\Xi} &= \min(v_\Psi,v_\Xi) \\ v_{\Psi\Xi} &= v_\Psi+v_\Xi \\ v_{\omega(\Psi)} &= \max(v_\Psi,\max\{n:\omega_n\neq 0\}) \qquad (\omega\in\times_{\mathbb{K}}\cup\top_{\mathbb{N}}\cup\Lambda) \\ v_{(\Phi^{[1]},\ldots,\Phi^{[d]})} &= \min(v_{\Phi^{[1]}},\ldots,v_{\Phi^{[d]}}) \qquad (d\geqslant 2) \end{aligned}$$

We next define the coefficients Φ_n by another induction over the size of Φ. If $n < v_\Phi$, then we take $\Phi_n = 0$. Otherwise, we take

$$\begin{array}{rcl}
P_n & = & P_n \qquad (P \in \mathbb{K}[z]) \\
\top_i(\Psi)_n & = & \Psi_n \qquad (i \in \mathbb{N}, n \geqslant i) \\
(\Psi \pm \Xi)_n & = & \Psi_n \pm \Xi_n \\
(\Psi\,\Xi)_n & = & \Psi_{v_\Psi}\,\Xi_{n-v_\Psi} + \Psi_{v_\Psi+1}\,\Xi_{n-v_\Psi-1} + \cdots + \Psi_{n-v_\Xi}\,\Xi_{v_\Xi} \\
\omega(\Psi)_n & = & \omega_n\,\Psi_n \qquad (\omega \in \times_{\mathbb{K}} \cup \Lambda) \\
(\Phi^{[1]}, \ldots, \Phi^{[d]})_n & = & (\Phi_n^{[1]}, \ldots, \Phi_n^{[d]}) \qquad (d \geqslant 2)
\end{array}$$

As a result of the claim, we emphasize that Φ_n only depends on the coefficients $F_0, \ldots, F_n$ up till order z^n of F.

Remark 2. The formula for $(\Psi\ \Xi)_n$ can be replaced by any algebraically equivalent formula, as long as $(\Psi\,\Xi)_n$ only depends on $\Psi_0, \ldots, \Psi_{n-v_\Xi}$ and $\Xi_0, \ldots, \Xi_{n-v_\Psi}$. Assuming the concept of relaxed power series, to be introduced below, this means that we compute $\Psi\,\Xi$ using the formula

$$\Psi\,\Xi = [(\Psi \operatorname{div} z^{v_\Psi})\,(\Xi \operatorname{div} z^{v_\Xi})]\,z^{v_\Psi + v_\Xi},$$

where $(\Psi \operatorname{div} z^{v_\Psi})\,(\Xi \operatorname{div} z^{v_\Xi})$ is computed using a relaxed multiplication algorithm.

2.4. Relaxed power series

Let us briefly recall the technique of relaxed power series computations, which is explained in more detail in [vdH02]. In this computational model, a power series $f \in \mathbb{K}[[z]]$ is regarded as a stream of coefficients f_0, f_1, ... When performing an operation $g = \Phi(f_1, \ldots, f_k)$ on power series it is required that the coefficient g_n of the result is output as soon as sufficiently many coefficients of the inputs are known, so that the computation of g_n does not depend on the further coefficients. For instance, in the case of a multiplication $h = f\,g$, we require that h_n is output as soon as $f_0, \ldots, f_n$ and $g_0, \ldots, g_n$ are known. In particular, we may use the naive formula $h_n = \sum_{i=0}^{n} f_i\,g_{n-i}$ for the computation of h_n.

The additional constraint on the time when coefficients should be output admits the important advantage that the inputs may depend on the output, provided that we add a small delay. For instance, the exponential $g = \exp f$ of a power series $f \in z\,\mathbb{K}[[z]]$ may be computed in a relaxed way using the formula

$$g = \int f'\,g.$$

Indeed, when using the naive formula for products, the coefficient g_n is given by

$$g_n = \frac{1}{n}\,(f_1\,g_{n-1} + 2\,f_2\,g_{n-2} + \cdots + n\,f_n\,g_0),$$

and the right-hand side only depends on the previously computed coefficients $g_0, \ldots, g_{n-1}$.

The main drawback of the relaxed approach is that we cannot directly use fast algorithms on polynomials for computations with power series. For instance, assuming that $\mathbb{K}$ has sufficiently many 2^p-th roots of unity and that field operations in $\mathbb{K}$ can be done in time $O(1)$, two polynomials of degrees $< n$ can be multiplied in time $\mathsf{M}(n) = O(n \log n)$, using FFT multiplication [CT65]. Given the truncations $f_{;n} = f_0 + \cdots + f_{n-1}\,z^{n-1}$ and $g_{;n} = g_0 + \cdots + g_{n-1}\,z^{n-1}$ at order n of power series $f, g \in \mathbb{K}[[z]]$, we may thus compute the truncated product $(f\,g)_{;n}$ in time $\mathsf{M}(n)$ as well. This is much faster than the naive $O(n^2)$ relaxed multiplication algorithm for the computation of $(f\,g)_{;n}$. However, the formula for $(f\ g)_0$ when using FFT multiplication depends on all input coefficients f_0, ..., f_{n-1} and $g_0, \ldots, g_{n-1}$, so the fast algorithm is not relaxed. Fortunately, efficient relaxed multiplication algorithms do exist:

THEOREM 3. [VDH97, VDH02] *Let* $\mathsf{M}(n)$ *be the time complexity for the multiplication of polynomials of degrees* $< n$ *in* $\mathbb{K}[z]$. *Then there exists a relaxed multiplication algorithm for series in* $\mathbb{K}[[z]]$ *of time complexity* $\mathsf{R}(n) = O(\mathsf{M}(n) \log n)$.

THEOREM 4. [VDH07A] *If* $\mathbb{K}$ *admits a primitive* 2^p*-th root of unity for all* p*, then there exists a relaxed multiplication algorithm of time complexity* $\mathsf{R}(n) = O(n \log n \, \mathrm{e}^{2\sqrt{\log 2 \log \log n}})$. *In practice, the existence of a* 2^{p+1}*-th root of unity with* $2^p \geqslant n$ *suffices for multiplication up to order* n.

In what follows, we will denote by $\mathsf{R}(n)$ the complexity of relaxed multiplication up till order n. Let us now consider a general equation of the form

$$f = \Phi(f), \tag{7}$$

where $\Phi \in \mathbb{D}_{F,\Lambda}^r$ an r-dimensional dag. We say that (7) is a *recursive equation*, if each coefficient $\Phi(F)_n$ only depends on earlier coefficients $F_0, \ldots, F_{n-1}$ of F. That is, $\Phi(F)_n \in \mathbb{D}_{(F_0,\ldots,F_{n-1}),\Lambda}^r$ for all n. In order to solve (7) up till order n, we then need to perform s_Φ relaxed multiplications at order n and t_Φ coefficientwise operations $+, -$ or $\omega \in \times_{\mathbb{K}} \cup \mathbb{C}$ at order n. This yields the following complexity bound:

PROPOSITION 5. *Any recursive equation* (7) *can be solved up till order* n *in time*

$$T(n) = s_\Phi \mathsf{R}(n) + O(t_\Phi n).$$

3. ANTICIPATORS

When solving an implicit equation in f using a relaxed algorithm, the coefficients f_n are computed only gradually. During the resolution process, it might happen that we wish to evaluate dags at higher orders than the number of known coefficients of f. That is, given $\Phi \in \mathbb{D}_{F,\Lambda}^*$ and $i \geqslant 1$, we might need $\Phi(f)_{n+i}$, even though only $f_0, \ldots, f_n$ are known. In that case, we have a problem, but we may still do the best we can, and compute $\Phi(f_0 + \cdots + f_n z^n)_{n+i}$ instead of $\Phi(f)_{n+i}$.

This motivates the introduction of the i-th order *anticipator* $\Phi^{\langle i \rangle}$ of Φ by

$$\begin{aligned} \Phi^{\langle i \rangle}(F)_{0;i-1} &= 0 \\ \Phi^{\langle i \rangle}(F)_{n+i} &= \Phi(F_{0;n})_{n+i}, \end{aligned}$$

where

$$F_{i;j} = F_i z^i + \cdots + F_j z^j.$$

On the one hand, we will show in this section that $\Phi, \ldots, \Phi^{\langle i \rangle}$ can be computed simultaneously by a dag Ψ of multiplicative size $s_\Psi = s_\Phi$ and total size $t_\Psi = O(i \, s_\Phi)$. On the other hand, we will show that $\Phi^{\langle i \rangle}$ is essentially a linear perturbation of Φ, which can be computed explicitly.

3.1. Computation of $\Phi^{\langle i \rangle}$ as a dag

Let us show how to compute a dag for $\Phi^{\langle i \rangle}$. The following rules are straightforward:

$$\begin{aligned} P^{\langle i \rangle} &= P \qquad (P \in \mathbb{K}[z]) \\ \left(F^{[k]}\right)^{\langle i \rangle} &= 0 \\ (\Phi \pm \Psi)^{\langle i \rangle} &= \Phi^{\langle i \rangle} \pm \Psi^{\langle i \rangle} \\ \omega(\Phi)^{\langle i \rangle} &= \omega\left(\Phi^{\langle i \rangle}\right) \qquad (\omega \in \times_{\mathbb{K}} \cup \top_{\mathbb{N}} \cup \Lambda). \end{aligned}$$

As to multiplication, for $n \geqslant i-1$, we have

$$\begin{aligned}
(\Phi\,\Psi)^{\langle i\rangle}(F)_{n+i} &= \sum_{k=0}^{n+i} \Phi(F_{0;n})_k\,\Psi(F_{0;n})_{n+i-k}\\
&= \sum_{k=i}^{n} \Phi(F_{0;n})_k\,\Psi(F_{0;n})_{n+i-k}+\\
&\quad \sum_{k=0}^{i-1} \Phi(F_{0;n})_k\,\Psi(F_{0;n})_{n+i-k}+\Phi(F_{0;n})_{n+i-k}\,\Psi(F_{0;n})_k\\
&= \sum_{k=i}^{n} \Phi(F)_k\,\Psi(F)_{n+i-k}+\\
&\quad \sum_{k=0}^{i-1} \Phi(F)_k\,\Psi^{\langle i-k\rangle}(F)_{n+i-k}+\Phi^{\langle i-k\rangle}(F)_{n+i-k}\,\Psi(F)_k\\
&= \left[\left(\Phi^{\lfloor i\rfloor}\,\Psi^{\lfloor i\rfloor}\right)(F)+\sum_{k=0}^{i-1}\left(\Phi(F)_k\,\Psi^{\langle i-k\rangle}(F)+\Psi(F)_k\,\Phi^{\langle i-k\rangle}(F)\right)z^k\right]_{n+i}
\end{aligned}$$

Consequently,

$$(\Phi\,\Psi)^{\langle i\rangle} = P_{\Phi,\Psi}^{\langle i\rangle}+\Phi^{\lfloor i\rfloor}\,\Psi^{\lfloor i\rfloor}+\sum_{k=0}^{i-1}\left(\Phi(F)_k\,\Psi^{\langle i-k\rangle}+\Psi(F)_k\,\Phi^{\langle i-k\rangle}\right)z^k, \tag{8}$$

for some polynomial $P_{\Phi,\Psi}^{\langle i\rangle}\in\mathbb{K}[z]$ with $\operatorname{val} P_{\Phi,\Psi}^{\langle i\rangle}\geqslant i$ and $\deg P_{\Phi,\Psi}^{\langle i\rangle}\leqslant 2\,i-2$ (in particular, $P_{\Phi,\Psi}^{\langle 1\rangle}=0$). Notice also that

$$\begin{aligned}
\Phi^{\lfloor i-1\rfloor}\,\Psi^{\lfloor i-1\rfloor} &= \Phi^{\lfloor i\rfloor}\,\Psi^{\lfloor i\rfloor}+\left(\Phi(F)_{i-1}\,\Psi^{\lfloor i\rfloor}+\Psi(F)_{i-1}\,\Phi^{\lfloor i\rfloor}\right)z^{i-1}+\\
&\quad \Phi(F)_{i-1}\,\Psi(F)_{i-1}\,z^{2i-2}.
\end{aligned}\tag{9}$$

Now assume that $\Phi^{\langle 0\rangle},\ldots,\Phi^{\langle i\rangle}$ and $\Psi^{\langle 0\rangle},\ldots,\Psi^{\langle i\rangle}$ are known. Then we may simultaneously compute $(\Phi\,\Psi)^{\langle 0\rangle},\ldots,(\Phi\,\Psi)^{\langle i\rangle}$ in the following way:

$$\begin{aligned}
\Phi^{\lfloor 1\rfloor} &= \Phi-\Phi_0\\
&\vdots\\
\Phi^{\lfloor i\rfloor} &= \Phi^{\lfloor i-1\rfloor}-\Phi_{i-1}\,z^{i-1}\\
\Psi^{\lfloor 1\rfloor},\ldots,\Psi^{\lfloor i\rfloor} &\quad \text{similarly}\\
\Phi^{\lfloor i\rfloor}\,\Psi^{\lfloor i\rfloor} &\\
\Phi^{\lfloor i-1\rfloor}\,\Psi^{\lfloor i-1\rfloor},\ldots,\Phi^{\lfloor 0\rfloor}\,\Psi^{\lfloor 0\rfloor} &\quad \text{using formula (9)}\\
(\Phi\,\Psi)^{\langle 1\rangle},\ldots,(\Phi\,\Psi)^{\langle i\rangle} &\quad \text{using formula (8)}
\end{aligned}$$

This computation involves one series product and $O(i^2)$ additions and scalar multiplications. For large i, we may further reduce the cost to $O(\mathsf{M}(i))$ since the computation of $(\Phi\,\Psi)^{\langle 1\rangle},\ldots,(\Phi\,\Psi)^{\langle i\rangle}$ really comes down to the computation of two truncated power series products $\Phi(F)\,(\Psi^{\langle 0\rangle}+\Psi^{\langle 1\rangle}\,z+\cdots)$ and $\Psi(F)\,(\Phi^{\langle 0\rangle}+\Phi^{\langle 1\rangle}\,z+\cdots)$ at order i. In summary, we obtain

Lemma 6. *Given $\Phi\in\mathbb{D}_{F,\Lambda}^{*}$, there exists a simultaneous dag Ψ for $\Phi^{\langle 0\rangle}$, ..., $\Phi^{\langle i\rangle}$ of multiplicative size $s_\Psi=s_\Phi$ and total size $t_\Psi=O(i\,t_\Phi+\mathsf{M}(i)\,s_\Phi)$.*

3.2. Computation of $\Phi^{\langle i\rangle}$ as a perturbation of Φ

Since $\Phi(F)_n$ only depends on $F_0,\ldots,F_n$, we notice that

$$\Phi^{\langle 0\rangle} = \Phi.$$

In general, for $n\geqslant i$ and $E=F_{n+1;n+i}$, we may expand

$$\begin{aligned}\Phi(F_{0;n}+E) &= \Phi(F_{0;n})+(D\Phi)(F_{0;n})(E)+O(z^{2n+2})\\ &= \Phi(F_{0;n})+(D\Phi)(F_{0;i-1})(E)+O(z^{n+i+1}).\end{aligned}$$

Let $\mathbb{e}_{[k]}$ denote the k-th basis element of $\mathbb{K}^r$, so that $F_j=F_j^{[1]}\,\mathbb{e}_{[1]}+\cdots+F_j^{[r]}\,\mathbb{e}_{[r]}$ for all j. When considering F as a column vector, it follows by linearity that

$$\Phi(F)_{n+i} = \Phi^{\langle i\rangle}(F)_{n+i}+\Phi^{\{i,1\}}(n)\,F_{n+1}+\cdots+\Phi^{\{i,i\}}(n)\,F_{n+i}, \tag{10}$$

where $\Phi^{\{i,j\}}$ is a row matrix with entries

$$\Phi^{\{i,j\}}_{[k]}(n) = [D(\Phi)(F_{0;i-1})(z^{n+j}\,\mathbb{e}_{[k]})]_{n+i}.$$

Notice that $\Phi^{\{i,j\}}(n)$ depends on n, but $D(\Phi)(F_{0;i-1})$ does not.

Let us investigate the functions $\Phi^{\{i,j\}}(n)$ more closely. If Φ is algebraic, then we have

$$D(\Phi)(F_{0;i-1})(z^{n+j}\,\mathbb{e}_{[k]}) = \frac{\partial\Phi}{\partial F^{[k]}}(F_{0;i-1})\,z^{n+j},$$

whence

$$\Phi^{\{i,j\}}_{[k]}(n) = \frac{\partial\Phi}{\partial F^{[k]}}(F_{0;i-1})_{i-j}. \tag{11}$$

In particular, $\Phi^{\{i,j\}}(n)$ is actually constant. If Φ is differential, of differential order d (this means that d is the maximal number of δ-nodes on a path from the root of Φ to a leaf), then, considering Φ as a differential polynomial in $F,\delta F,\ldots,\delta^{d)}F$, we have

$$D(\Phi)(F_{0;i-1})(z^{n+j}\,\mathbb{e}_{[k]}) = \frac{\partial\Phi}{\partial F^{[k]}}(F_{0;i-1})\,z^{n+j}+\cdots+\frac{\partial\Phi}{\partial(\delta^d F^{[k]})}(F_{0;i-1})\,(\delta^d z^{n+j}),$$

whence

$$\Phi^{\{i,j\}}_{[k]}(n) = \left[\frac{\partial\Phi}{\partial F^{[k]}}(F_{0;i-1})\right]_{i-j}+\cdots+\left[\frac{\partial\Phi}{\partial(\delta^d F^{[k]})}(F_{0;i-1})\right]_{i-j}(n+j)^d \tag{12}$$

is a polynomial of degree at most d. Similarly, if Φ is algebraic in $F,\int F,\ldots,\int^d F$, where $\int=\delta^{-1}z$, then

$$D(\Phi)(F_{0;i-1})(z^{n+j}\,\mathbb{e}_{[k]}) = \frac{\partial\Phi}{\partial F^{[k]}}(F_{0;i-1})\,z^{n+j}+\cdots+\frac{\partial\Phi}{\partial(\int^d F^{[k]})}(F_{0;i-1})\,(\textstyle\int^d z^{n+j}),$$

whence

$$\Phi^{\{i,j\}}_{[k]}(n) = \left[\frac{\partial\Phi}{\partial F^{[k]}}(F_{0;i-1})\right]_{i-j}+\cdots+\left[\frac{\partial\Phi}{\partial(\int^d F^{[k]})}(F_{0;i-1})\right]_{i-j-d}\frac{1}{(n+j)\cdots(n+j+d-1)}. \tag{13}$$

Consequently, there exists a polynomial $A_{i,j,k}$ of degree $\leqslant d$ with

$$\Phi^{\{i,j\}}_{[k]}(n) = \frac{A_{i,j,k}(n)}{(n+j)\,(n+j+1)\cdots(n+j+d-1)},$$

for all $n\geqslant i$.

For more general integral dags Φ, it can be checked by induction over the size of Φ that $\Phi^{\{i,j\}}_{[k]}(n)$ is still a rational function in n, which remains bounded for $n\to\infty$, and whose denominator has integer coefficients. Similarly, if $\Lambda\subseteq\{\delta,\delta^{-1}\}$, then $\Phi^{\{i,j\}}_{[k]}(n)$ is a rational function in n, whose denominator has integer coefficients.

4. Relaxed resolution of implicit equations

Assume now that we want to solve a system of power series equations

$$\begin{cases} \Phi(f) & = & 0 \\ f_0 & = & C_0 \\ & \vdots & \\ f_{l-1} & = & C_{l-1} \end{cases} \tag{14}$$

where $\Phi = (\Phi^{[1]}, \ldots, \Phi^{[r]}) \in \mathbb{D}^r_{F,\Lambda}$ is a vector of dags and $C_0, \ldots, C_{l-1} \in \mathbb{K}^r$ a finite number of initial conditions. For definiteness, it is also important that (14) admits a unique solution f. This will be guaranteed by an even stronger technical assumption to be detailed below. Roughly speaking, for a given *index* $i \leqslant \frac{1}{2}(l+1)$, we will assume that each coefficient Φ_n with $n \geqslant l$ can be determined as a function of the previous coefficients $\Phi_0, \ldots, \Phi_{n-1}$ using only the equations $\Phi(f)_0 = \cdots = \Phi(f)_{n+i-1} = 0$.

4.1. Construction of a recursive equation

Let $i \leqslant \frac{1}{2}(l+1)$. For each n and $j \in \{1, \ldots, i\}$, we introduce the $r \times r$ matrix

$$M_n^{\{j\}} = \begin{pmatrix} (\Phi^{[1]})^{\{i,j\}}_{[1]}(n) & \cdots & (\Phi^{[1]})^{\{i,j\}}_{[r]}(n) \\ \vdots & & \vdots \\ (\Phi^{[r]})^{\{i,j\}}_{[1]}(n) & \cdots & (\Phi^{[r]})^{\{i,j\}}_{[r]}(n) \end{pmatrix},$$

the $i\,r \times (2\,i-1)\,r$ block matrix

$$\boldsymbol{M}_n = \begin{pmatrix} M^{\{1\}}_{n-i} & \cdots & M^{\{i\}}_{n-i} & & \\ & \ddots & & \ddots & \\ & & M^{\{1\}}_{n-1} & \cdots & M^{\{i\}}_{n-1} \end{pmatrix}, \tag{15}$$

the $(i-1)\,r \times 1$, $i\,r \times 1$ and $(2\,i-1) \times r$ block column vectors

$$\check{\boldsymbol{f}}_n = \begin{pmatrix} f_{n-(i-1)} \\ \vdots \\ f_{n-1} \end{pmatrix}, \qquad \boldsymbol{f}_n = \begin{pmatrix} f_{n-(i-1)} \\ \vdots \\ f_n \end{pmatrix}, \qquad \hat{\boldsymbol{f}}_n = \begin{pmatrix} f_{n-(i-1)} \\ \vdots \\ f_{n+i-1} \end{pmatrix},$$

and the $i\,r \times 1$ column vector

$$\boldsymbol{g}_n = \begin{pmatrix} -\Phi^{\langle i \rangle}(f)_n \\ \vdots \\ -\Phi^{\langle i \rangle}(f)_{n+i-1} \end{pmatrix}.$$

In view of (10), the equations $\Phi(f)_n = \cdots = \Phi(f)_{n+i-1} = 0$ then translate into

$$\boldsymbol{M}_n \hat{\boldsymbol{f}}_n = \boldsymbol{g}_n.$$

We will say that (14) is *i-predictive* or *predictive* of *index* i, if, for all $n \geqslant l$, there exist $r \times i\,r$ and $r \times (i-1)\,r$ matrices $\boldsymbol{P}_n$ and $\boldsymbol{Q}_n$, such that

$$\boldsymbol{P}_n \boldsymbol{M}_n = (\ \boldsymbol{Q}_n \ \ \mathrm{Id}_r \ \ 0\ ).$$

In that case, we have

$$\begin{aligned}\boldsymbol{P}_n \boldsymbol{M}_n \hat{\boldsymbol{f}}_n &= \boldsymbol{Q}_n \check{\boldsymbol{f}}_n + f_n\\ &= \boldsymbol{P}_n \boldsymbol{g}_n,\end{aligned}$$

whence

$$f_n = \boldsymbol{P}_n \boldsymbol{g}_n - \boldsymbol{Q}_n \check{\boldsymbol{f}}_n \tag{16}$$

provides us with an explicit formula for f_n. Now let $\boldsymbol{P}$ and $\boldsymbol{Q}$ be the operators on vectors of power series V with the property that $(\boldsymbol{P} V)_n = \boldsymbol{P}_n V_n$ and $(\boldsymbol{Q} V)_n = \boldsymbol{Q}_n V_n$. Then we may rewrite (16) into

$$f = \boldsymbol{P} \begin{pmatrix} -z^i \Phi^{\langle i \rangle}(f) \\ \vdots \\ -z \Phi^{\langle i \rangle}(f) \end{pmatrix} z^{-i} - \boldsymbol{Q} \begin{pmatrix} z^i f \\ \vdots \\ z f \end{pmatrix} \tag{17}$$

This is the desired recursive equation for f.

Remark 7. The main sources of unpredictivity are an insufficient number of initial conditions and the existence of multiple solutions. In the latter case, we may usually restore predictivity by differentiating the equation a finite number of times. It seems unlikely that there exist equations with a unique solution and which are unpredictive for any number of initial conditions. However, we do not have a formal proof for this intuition yet.

4.2. Complexity analysis

Let us first consider the case of an algebraic dag Φ. In that case, the matrix $\boldsymbol{M} = \boldsymbol{M}_n$ does not depend on n and its coefficients are explicitly given by (11). We may now determine $r \times i r$ and $r \times (i-1) r$ matrices $\boldsymbol{P}$ and $\boldsymbol{Q}$ with

$$\boldsymbol{P} \boldsymbol{M} = (\ \boldsymbol{Q} \ \ \mathrm{Id}_r \ \ 0\ ), \tag{18}$$

using Gaussian elimination in order, and whenever such matrices exist. The equation (14) is i-predictive, if and only if this is indeed possible.

Theorem 8. *Consider an i-predictive equation* (14)*, such that Φ is algebraic. Then we may compute n terms of its unique solution f in time*

$$T_{\mathrm{alg}}(n) = s_\Phi \, \mathsf{R}(n) + O(\mathsf{M}(i) \, s_\Phi \, n + i \, t_\Phi \, n) + O(i \, r^2 \, n).$$

Proof. By what precedes, the operators $\boldsymbol{P}$ and $\boldsymbol{Q}$ in (17) are really the constant matrices from (18). By lemma 6, the size of the righthand side of (17) as a dag is therefore bounded by $O(\mathsf{M}(i) \, s_\Phi + i \, t_\Phi + i \, r^2)$ and its multiplicative size is exactly s_Φ. The result thus follows from proposition 5. □

Assume now that $\Lambda \subseteq \{\delta, \delta^{-1}\}$. Then we claim that there exists an algorithms for checking i-predictivity and constructing a general formula for the corresponding matrices $\boldsymbol{P}_n$ and $\boldsymbol{Q}_n$. Indeed, we recall from section 3.2 that $\boldsymbol{M}_n$ is the evaluation at $N = n$ of a matrix $\boldsymbol{M}^*$ with entries in $\mathbb{K}(N)$ and denominators in $\mathbb{Z}[N]$. We may thus use Gaussian elimination in order to compute $r \times i r$ and $r \times (i-1) r$ matrices $\boldsymbol{P}^*$ and $\boldsymbol{Q}^*$ with entries in $\mathbb{K}(N)$ and

$$\boldsymbol{P}^* \boldsymbol{M}^* = (\ \boldsymbol{Q}^* \ \ \mathrm{Id}_r \ \ 0\ ),$$

whenever such matrices exist. For those $n \geqslant l$ which are not positive integer roots of one of the denominators of the entries of $\boldsymbol{P}^*$, we now have $\boldsymbol{P}_n = \boldsymbol{P}^*(n)$ and $\boldsymbol{Q}_n = \boldsymbol{Q}^*(n)$. For each of the finite number of integer roots $n \geqslant l$, we may directly compute the matrices $\boldsymbol{P}_n$ and $\boldsymbol{Q}_n$ by Gaussian elimination over $\mathbb{K}$, whenever such matrices exist.

Theorem 9. *Consider an i-predictive equation* (14) *and let d be the maximal degree of an entry of $\boldsymbol{M}^*$. Then we may compute n terms of the solution f to* (14) *in time*

$$T(n) \;=\; s_\Phi \mathsf{R}(n) + O(\mathsf{M}(i)\, s_\Phi\, n + i\, t_\Phi\, n) + O((i^2\, r^3 + i\, r^2\, d)\, n).$$

Proof. The computation of $\boldsymbol{M}^*$ (and the finite number of exceptional $\boldsymbol{M}_n$ for which n is a root of one of the denominators) is a precomputation. The determination of every next $\boldsymbol{M}_n$ can be done in time $O(i^2\, r^2 + i\, r^2\, d)$, via a diagonal translation of $\boldsymbol{M}_{n-1}$ and evaluation of the $O(i\, r^2)$ rational functions which are the entries of the bottom $r \times (2\, i - 1)\, r$ submatrix. Now assume that we maintain upper and lower triangular matrices $\boldsymbol{U}_n$, $\boldsymbol{L}_n$ and a permutation matrix $\boldsymbol{\Pi}_n$ at each stage such that $\boldsymbol{L}_n = \boldsymbol{U}_n\, \boldsymbol{\Pi}_n\, \boldsymbol{M}_n$. Then the determination of $\boldsymbol{U}_n$, $\boldsymbol{L}_n$ and $\boldsymbol{\Pi}_n$ as a function of $\boldsymbol{U}_{n-1}$, $\boldsymbol{L}_{n-1}$ and $\boldsymbol{\Pi}_{n-1}$ can be done in time $O(i^2\, r^3)$ using naive linear algebra. The determination of $\boldsymbol{P}_n$ and $\boldsymbol{Q}_n$ from $\boldsymbol{U}_n$, $\boldsymbol{L}_n$ and $\boldsymbol{\Pi}_n$ can again be done in time $O(i^2\, r^3)$. Consequently, the cost of applying the operators $\boldsymbol{P}$ and $\boldsymbol{Q}$ during the relaxed resolution of (17) at order n is bounded by $O((i^2\, r^3 + i\, r^2\, d)\, n)$. The cost of the evaluation of the remaining dag is bounded by $s_\Phi \mathsf{R}(n) + O(\mathsf{M}(i)\, s_\Phi\, n + i\, t_\Phi\, n)$, as in the algebraic case. □

5. Reliable resolution of numeric implicit equations

If $\mathbb{K} \subseteq \mathbb{C}$, then power series solutions to recursive or implicit equations often converge in a small disk around the origin. For instance, if $\Phi \in \mathbb{D}_{F,\Lambda}^r$ is an algebraic dag and $C \in \mathbb{K}^r$, then the recursive equation

$$f \;=\; C + \textstyle\int \Phi(f)$$

admits a convergent solution, by Cauchy-Kovalevskaya's theorem. Given a point $z \in \mathbb{K}$ in a sufficiently small disk where f converges, we might like to compute an approximation $\tilde{f}(z) \in \mathbb{K}^r$ of $f(z)$, together with a bound $\delta \in (\mathbb{K}^>)^r$ for the error: $|\tilde{f}^{[k]}(z) - f^{[k]}(z)| < \delta^{[k]}$. Interval arithmetic [Moo66, AH83, Neu90, MKC09, Rum10] provides a classical framework for deriving such enclosures in a systematic way. We will rather rely on ball arithmetic [vdH09a], which is more suitable for complex and multiple precision arithmetic.

5.1. Ball arithmetic

Let us briefly recall the principles behind ball arithmetic. Given a normed vector space $\mathbb{K}$, we will denote by $\boldsymbol{\mathbb{K}}$ or $\mathcal{B}(\mathbb{K}, \mathbb{R})$ the set of closed balls with centers in $\mathbb{K}$ and radii in $\mathbb{R}^{\geqslant} = \{x \in \mathbb{R} : x \geqslant 0\}$. Given such a ball $\boldsymbol{z} \in \mathcal{B}(\mathbb{K}, \mathbb{R})$, we will denote its center by $\operatorname{cen}(\boldsymbol{z})$ and its radius by $\operatorname{rad}(\boldsymbol{z})$. Conversely, given $z \in \mathbb{K}$ and $r \in \mathbb{R}$, we will denote by $z + \mathcal{B}(r)$ the closed ball with center z and radius r.

A continuous operation $f : \mathbb{K}^d \to \mathbb{K}$ is said to *lift* into an operation $\boldsymbol{f} : \boldsymbol{\mathbb{K}}^d \to \boldsymbol{\mathbb{K}}$ on balls, which is usually also denoted by f, if the *inclusion property*

$$f(x^{[1]}, \ldots, x^{[d]}) \;\in\; f(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_d) \tag{19}$$

is satisfied for any $\boldsymbol{x}^{[1]},\ldots,\boldsymbol{x}^{[d]}\in\mathbb{IK}$ and $x^{[1]}\in\boldsymbol{x}^{[1]},\ldots,x^{[d]}\in\boldsymbol{x}^{[d]}$. For instance, if $\mathbb{K}$ is a Banach algebra, then we may take

$$\begin{aligned}\boldsymbol{x}+\boldsymbol{y} &= \operatorname{cen}(\boldsymbol{x})+\operatorname{cen}(\boldsymbol{y})+\mathcal{B}(\operatorname{rad}(\boldsymbol{x})+\operatorname{rad}(\boldsymbol{y}))\\ \boldsymbol{x}-\boldsymbol{y} &= \operatorname{cen}(\boldsymbol{x})-\operatorname{cen}(\boldsymbol{y})+\mathcal{B}(\operatorname{rad}(\boldsymbol{x})+\operatorname{rad}(\boldsymbol{y}))\\ \boldsymbol{x}\,\boldsymbol{y} &= \operatorname{cen}(\boldsymbol{x})\operatorname{cen}(\boldsymbol{y})+\mathcal{B}(\operatorname{rad}(\boldsymbol{x})\,(|\operatorname{cen}(\boldsymbol{y})|+\operatorname{rad}(\boldsymbol{y}))+|\operatorname{cen}(\boldsymbol{y})|\operatorname{rad}(\boldsymbol{x})).\end{aligned}$$

Similar formulas can be given for division and elementary functions.

It is convenient to extend the notion of a ball to more general radius types, which only carry a partial ordering. This allows us for instance to regard a vector $\boldsymbol{x}=(\boldsymbol{x}^{[1]},\ldots,\boldsymbol{x}^{[d]})\in\mathcal{B}(\mathbb{K},\mathbb{R})^d$ of balls as a "vectorial ball" with center $\operatorname{cen}(\boldsymbol{x})=(\operatorname{cen}(\boldsymbol{x}_1),\ldots,\operatorname{cen}(\boldsymbol{x}_d))\in\mathbb{K}^d$ and radius $\operatorname{rad}(\boldsymbol{x})=(\operatorname{rad}(\boldsymbol{x}^{[1]}),\ldots,\operatorname{rad}(\boldsymbol{x}^{[d]}))\in\mathbb{R}^d$. If $x=(x^{[1]},\ldots,x^{[d]})\in\mathbb{K}^d$, then we write $x\in\boldsymbol{x}$ if and only if $x^{[i]}\in\boldsymbol{x}^{[i]}$ for all $i\in\{1,\ldots,d\}$. A similar remark holds for matrices and power series with ball coefficients.

In concrete machine computations, numbers are usually approximated by floating point numbers with a finite precision. Let $\tilde{\mathbb{R}}$ be the set of floating point numbers at a given working precision, which we will assume fixed. It is customary to include the infinities $\pm\infty$ in $\tilde{\mathbb{R}}$ as well. The IEEE754 standard [ANS08] specifies how to perform basic arithmetic with floating point numbers in a predictable way, by specifying a rounding mode $\mathsf{R}\in\{\downarrow,\uparrow,\updownarrow\}$ (down, up and nearest). A multiple precision implementation of this standard is available in the Mpfr library [HLRZ00]. Given an operation $f:\mathbb{R}^d\to\mathbb{R}$, we will denote by $f^{\mathsf{R}}:\tilde{\mathbb{R}}^d\to\tilde{\mathbb{R}}$ its approximation using floating pointing arithmetic with rounding mode R. This notation extends to the case when $\mathbb{R}$ and $\tilde{\mathbb{R}}$ are replaced by their complexifications $\mathbb{C}$ and $\tilde{\mathbb{C}}=\tilde{\mathbb{R}}[\mathrm{i}]$.

Let $\mathbb{K}=\mathbb{R}$ and $\tilde{\mathbb{K}}=\tilde{\mathbb{R}}$ or $\mathbb{K}=\mathbb{C}$ and $\tilde{\mathbb{K}}=\tilde{\mathbb{C}}$. We will denote by $\tilde{\mathbb{IK}}$ or $\mathcal{B}(\tilde{\mathbb{K}},\tilde{\mathbb{R}})$ the set of closed balls in $\mathbb{K}$ with centers in $\tilde{\mathbb{K}}$ and radii in $\tilde{\mathbb{R}}^{\geqslant}$. In this case, we will also allow for balls with an infinite radius. A continuous operation $f:\mathbb{K}^d\to\mathbb{K}$ is again said to *lift* to an operation f: $\tilde{\mathbb{IK}}^d\to\tilde{\mathbb{IK}}$ on balls if (19) holds for any $\boldsymbol{x}^{[1]},\ldots,\boldsymbol{x}^{[d]}\in\tilde{\mathbb{IK}}$ and $x^{[1]}\in\boldsymbol{x}^{[1]},\ldots,x^{[d]}\in\boldsymbol{x}^{[d]}$. The formulas for the ring operations may now be adapted to

$$\begin{aligned}\boldsymbol{x}+\boldsymbol{y} &= \operatorname{cen}(\boldsymbol{x})+^{\updownarrow}\operatorname{cen}(\boldsymbol{y})+\mathcal{B}(\operatorname{rad}(\boldsymbol{x})+^{\uparrow}\operatorname{rad}(\boldsymbol{y})+^{\uparrow}\epsilon_{+,\boldsymbol{x},\boldsymbol{y}})\\ \boldsymbol{x}-\boldsymbol{y} &= \operatorname{cen}(\boldsymbol{x})-^{\updownarrow}\operatorname{cen}(\boldsymbol{y})+\mathcal{B}(\operatorname{rad}(\boldsymbol{x})+^{\uparrow}\operatorname{rad}(\boldsymbol{y})+^{\uparrow}\epsilon_{-,\boldsymbol{x},\boldsymbol{y}})\\ \boldsymbol{x}\,\boldsymbol{y} &= \operatorname{cen}(\boldsymbol{x})\times^{\updownarrow}\operatorname{cen}(\boldsymbol{y})+\\ &\quad\ \mathcal{B}(\operatorname{rad}(\boldsymbol{x})\times^{\uparrow}(|\operatorname{cen}(\boldsymbol{y})|+^{\uparrow}\operatorname{rad}(\boldsymbol{y}))+^{\uparrow}|\operatorname{cen}(\boldsymbol{y})|\times^{\uparrow}\operatorname{rad}(\boldsymbol{x})+^{\uparrow}\epsilon_{\times,\boldsymbol{x},\boldsymbol{y}}),\end{aligned}$$

where $\epsilon_{+,\boldsymbol{x},\boldsymbol{y}}$, $\epsilon_{-,\boldsymbol{x},\boldsymbol{y}}$ and $\epsilon_{\times,\boldsymbol{x},\boldsymbol{y}}$ are reliable bounds for the rounding errors induced by the corresponding floating point operations on the centers; see [vdH09a] for more details.

In order to ease the remainder of our exposition, we will avoid technicalities related to rounding problems, and compute with "idealized" balls with centers in $\mathbb{K}\in\{\mathbb{R},\mathbb{C}\}$ and radii in $\mathbb{R}^{\geqslant}$. For those who are familiar with rounding errors, it should not be difficult though to adapt our results to more realistic machine computations.

5.2. Taylor models

If we are computing with analytic functions on a disk, or multivariate analytic functions on a polydisk, then Taylor models [MB96, MB04] provide a suitable functional analogue for ball arithmetic. We will use a multivariate setup with $z=(z_1,\ldots,z_d)$ as our coordinates and a polydisk $D=\mathcal{B}(\rho)=\{z,|z|\leqslant|\rho|\}$ for a fixed $\rho=(\rho_1,\ldots,\rho_d)\in(\mathbb{R}^{>})^d$, where we use vector notation. Taylor models come in different blends, depending on whether we use a global error bound on D or individual bounds for the coefficients of the polynomial approximation. Individual bounds are sharper (especially if we truncate up to an small order such that the remainder is not that small), but more expensive to compute. Our general setup will cover all possible blends of Taylor models.

We first need some more definitions and notations. Assume that $\mathbb{N}^d$ is given the natural partial ordering. Let $\mathbb{e}_k$ denote the k-th canonical basis vector of $\mathbb{N}^d$, so that $(\mathbb{e}_k)_k = 1$ and $(\mathbb{e}_k)_l = 0$ for $l \neq k$. For every $i \in \mathbb{N}^d$, we will write $\|i\| = |i_1| + \cdots + |i_d|$. A subset $\mathcal{I} \subseteq \mathbb{N}^d$ is called an *initial segment*, if for any $i \in \mathcal{I}$ and $j \in \mathbb{N}^d$ with $j \leqslant i$, we have $j \in \mathcal{I}$. In that case, we write $\mathring{\mathcal{I}} = \{i \in \mathcal{I} \colon i + \{\mathbb{e}_1, \ldots, \mathbb{e}_d\} \subseteq \mathcal{I}\}$ and $\partial \mathcal{I} = \mathcal{I} \setminus \mathring{\mathcal{I}}$. In what follows, we assume that $\mathcal{I}$ and $\mathcal{J}$ are fixed initial segments of $\mathbb{N}^d$ with $\mathring{\mathcal{J}} \subseteq \mathcal{I}$. For instance, we may take $\mathcal{I} = \mathcal{T}_n = \{i \in \mathbb{N}^d \colon \|i\| \leqslant n\}$ and $\mathcal{J} = \mathcal{T}_{n+1}$ or $\mathcal{J} = \mathcal{T}_n$ or $\mathcal{J} = \{0\}$.

Recall that $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$. Given a series $f = \sum_{i \in \mathbb{N}^d} f_i z^i \in \mathbb{K}[[z]]$, we will write $\operatorname{supp} f = \{i \in \mathbb{N}^d \colon f_i \neq 0\}$ for its *support*. Given a subset $\mathbb{S} \subseteq \mathbb{K}[[z]]$ and a subset $\mathcal{S} \subseteq \mathbb{N}^d$, we write $f_{\mathcal{S}} = \sum_{i \in \mathcal{S}} f_i z^i$ and $\mathbb{S}_{\mathcal{S}} = \{g \in \mathbb{S} \colon \operatorname{supp} g \subseteq \mathcal{S}\}$. If f is analytic on D, then we denote its sup norm by

$$\|f\|_D = \sup_{z \in D} |f(z)|.$$

A *Taylor model* is a tuple $\boldsymbol{P} = (\rho, \mathcal{I}, \mathcal{J}, \operatorname{cen}(\boldsymbol{P}), \operatorname{rad}(\boldsymbol{P}))$, where ρ, $\mathcal{I}$ and $\mathcal{J}$ are as above, $\operatorname{cen}(\boldsymbol{P}) \in \mathbb{K}[z]_{\mathcal{I}}$ and $\operatorname{rad}(\boldsymbol{P}) \in \mathbb{R}[z]_{\mathcal{J}}$. We will write $\mathbb{T} = \mathbb{T}_{D,\mathcal{I},\mathcal{J}} = \mathcal{B}_D(\mathbb{K}[z]_{\mathcal{I}}, \mathbb{R}[z]_{\mathcal{J}})$ for the set of such Taylor models. Given $\boldsymbol{P} \in \mathbb{T}$ and $i \in \mathbb{N}^d$, we will also denote $\boldsymbol{P} = \operatorname{cen}(\boldsymbol{P}) + \mathcal{B}_D(\operatorname{rad}(\boldsymbol{P}))$ and $\boldsymbol{P}_i = \operatorname{cen}(\boldsymbol{P})_i + \mathcal{B}(\operatorname{rad}(\boldsymbol{P})_i)$. Given an analytic function f on D, we write $f \in \boldsymbol{P}$, if there exists a decomposition

$$f = \operatorname{cen}(\boldsymbol{P}) + \sum_{i \in \mathcal{J}} \varepsilon_i z^i$$

with $\varepsilon_i \in \mathbb{C}[[z]]$ and $\|\varepsilon_i\|_D \leqslant \operatorname{rad}(\boldsymbol{P})_i$ for all i. Optionally, we may also require that $f_i \in \boldsymbol{P}_i$ for all $i \in \mathcal{I}$. Given two Taylor models $\boldsymbol{P}, \boldsymbol{Q} \in \mathbb{T}$, we will say that $\boldsymbol{P}$ is *included* in $\boldsymbol{Q}$, and we write $\boldsymbol{P} \subseteq \boldsymbol{Q}$ if $f \in \boldsymbol{Q}$ for any $f \in \boldsymbol{P}$. This holds in particular if $\boldsymbol{P}_i \subseteq \boldsymbol{Q}_i$, in which case we say that $\boldsymbol{P}$ is *strongly included* in $\boldsymbol{Q}$ and write $P \sqsubseteq Q$.

Addition, subtraction and scalar multiplication are defined in a natural way on Taylor models. For multiplication, we need a projection $\pi = \pi_{\mathcal{J}} \colon \mathbb{N}^d \to \mathcal{J}$ with $\pi(i) \leqslant i$ for all i and $\pi(i) = i$ if $i \in \mathcal{J}$. One way to construct such a mapping is as follows. For $i \in \mathcal{J}$, we must take $\pi(i) = i$. For $i \notin \mathcal{J}$, let k be largest such that $i_k \neq 0$. Then we recursively define $\pi(i) = \pi(i - \mathbb{e}_k)$. Given $\boldsymbol{P}, \boldsymbol{Q} \in \mathbb{T}$, we now define their product by

$$\boldsymbol{P}\boldsymbol{Q} = \sum_{i,j \in \mathcal{I}} \boldsymbol{P}_i \boldsymbol{Q}_j \mathcal{B}_D(\rho)^{i+j-\pi(i+j)} z^{\pi(i+j)}.$$

Using the observation that $z^{i+j} \in \mathcal{B}_D(\rho)^{i+j-\pi(i+j)} z^{\pi(i+j)}$, this product satisfies the inclusion property that $f g \in \boldsymbol{P}\boldsymbol{Q}$ for any analytic functions $f \in \boldsymbol{P}$ and $Q \in \boldsymbol{Q}$ on D. Finally, let ω be a coefficientwise operator on $\mathbb{K}[[z]]$ such that there exist constants ω_i^* with

$$\|\omega f\|_D \leqslant \omega_i^* \|f\|_D, \tag{20}$$

for any $f \in z^i \mathbb{K}[[z]]$ which converges on D. Then we say that ω is *bounded* and we may lift ω into an operator on Taylor models by taking

$$\omega \boldsymbol{P} = \sum_i \omega_i^* \boldsymbol{P}_i z^i.$$

Indeed, the formula (20) implies the inclusion property $\omega f \in \omega \boldsymbol{P}$ for any $f \in \boldsymbol{P}$.

Remark 10. If $\mathcal{I} = \mathcal{J}$, then there exists a more algebraic alternative for the definition of the product $\boldsymbol{P}\boldsymbol{Q}$. Let $\mathcal{F}$ be the finite set of minimal elements of $\mathbb{N}^d \setminus \mathcal{I}$ and consider the ideal $\mathfrak{I}$ in $\mathbb{K}[z]$ generated by the relations $z^i - \mathcal{B}(\rho)^{i-\pi(i)} z^{\pi(i)}$, for $i \in \mathcal{F}$. Then $\mathbb{T}_{D,\mathcal{I},\mathcal{I}}$ is naturally isomorphic to $\mathbb{K}[z]/\mathfrak{I}$ as a vector space and we may transport the product of $\mathbb{K}[z]/\mathfrak{I}$ to $\mathbb{T}_{D,\mathcal{I},\mathcal{I}}$. However, this more algebraic definition does not seem to be more convenient from the computational point of view.

Remark 11. The strong inclusion relation $\sqsubseteq$ has the advantage that it is easy to check, contrary to general inclusion of Taylor models. An intermediate relation $\trianglelefteqslant$ which can still be checked can be defined by recursion over $\mathcal{I}$ and $\mathcal{J}$: if $\mathcal{I}=\mathcal{J}=\{0\}$, then we take $\boldsymbol{P}\trianglelefteqslant\boldsymbol{Q}$ if and only if $\boldsymbol{P}_0\subseteq\boldsymbol{Q}_0$. Otherwise, let $i\in\mathcal{I}\cup\mathcal{J}$ be largest for a total ordering on the monoid $\mathbb{N}^d$ (e.g. the lexicographical ordering). Let $\mathcal{I}'=\mathcal{I}\setminus\{i\}$, $\mathcal{J}'=\mathcal{J}\setminus\{i\}$, $\boldsymbol{P}'=\boldsymbol{P}_{\mathcal{I}'\cup\mathcal{J}'}$ and $\boldsymbol{Q}'=\boldsymbol{Q}_{\mathcal{I}'\cup\mathcal{J}'}$. If $\boldsymbol{P}_i\subseteq\boldsymbol{Q}_i$, then we take $\boldsymbol{P}\trianglelefteqslant\boldsymbol{Q}$ if and only if $\boldsymbol{P}'\trianglelefteqslant\boldsymbol{Q}'$. Otherwise, we first compute the smallest $c\in\mathbb{R}$ with $\boldsymbol{P}_i\subseteq\boldsymbol{Q}_i+\mathcal{B}(c)$. Setting $\boldsymbol{P}''=\boldsymbol{P}'+\mathcal{B}(\rho)^{i-\pi_{\mathcal{J}'}(i)}\,z^{\pi_{\mathcal{J}}(i)}$, we now take $\boldsymbol{P}\trianglelefteqslant\boldsymbol{Q}$ if and only if $\boldsymbol{P}''\trianglelefteqslant\boldsymbol{Q}'$.

5.3. The fixed point theorem

Most of the definitions and results of sections 2.1 and 2.4 generalize in a straightforward way to the multivariate case. Only the complexity results need additional attention, since $\mathsf{M}(n)$ needs to be replaced by a suitable multivariate analogue; see [vdH02, LS03, vdH05, vdHL10, vdHS10] for various results in this direction. Let us take

$$\Omega \;=\; \mathbb{K}[z_1,\ldots,z_d]\cup\{F^{[1]},\ldots,F^{[r]},+,-,\times\}\cup\times_{\mathbb{K}}\cup\top_{1,\mathbb{N}}\cup\cdots\cup\mathbb{T}_{r,\mathbb{N}}\cup\Lambda,$$

with $\top_{k,j}(f)=\sum_{i_k\geqslant j} f_i\,z^i$ for all k and j, and where Λ is set of bounded coefficientwise operators on $\mathbb{K}[[z]]$. Given $\Phi\in\mathbb{D}_F^r$, consider the equation

$$f \;=\; \Phi(f). \tag{21}$$

This equation is said to be *recursive* if $\Phi(F)_i$ only depends on F_j with $j<i$, for every i. In that case, the equation admits a unique solution, whose coefficients may be computed using a relaxed algorithm.

Let $\mathbb{K}\in\{\mathbb{R},\mathbb{C}\}$, $\boldsymbol{\mathbb{K}}=\mathcal{B}(\mathbb{K},\mathbb{R})$, $\rho\in(\mathbb{R}^>)^d$, $D=\mathcal{B}(\rho)$, $\mathcal{I}$ and $\mathcal{J}$ be as in the previous section. Consider a polydisk $D'=\mathcal{B}(\rho')$ with $D'\subseteq D$. Then any Taylor model $\boldsymbol{P}\in\mathbb{T}_{D,\mathcal{I},\mathcal{J}}$ can naturally be reinterpreted as a Taylor model in $\mathbb{T}_{D',\mathcal{I},\mathcal{J}}$, which we will denote by $\boldsymbol{P}_{D'}$.

THEOREM 12. *Assume that* (21) *is a recursive equation. Let* $0<\tau<1$ *and let* $\boldsymbol{P}\in\mathbb{T}^r$ *be a vector of Taylor models with the property that* $\Phi(\boldsymbol{P})\subseteq\boldsymbol{P}$. *Then the unique solution* f *to* (21) *is analytic on* τD *and* $f\in\boldsymbol{P}_{\tau D}$.

Proof. Given power series $f\in\mathbb{K}[[z]]$ and $g\in\mathbb{R}^{\geqslant}[[z]]$, we will say that g is a *majorant* for f, and write $f\trianglelefteqslant g$, if $|f_i|\leqslant g_i$ for all $i\in\mathbb{N}^d$. Let $f^{\langle 0\rangle}$ be an arbitrary element of $\boldsymbol{P}$ and consider the sequence $f^{\langle n\rangle}=\Phi^n(f^{\langle 0\rangle})$. By induction, we have $f^{\langle n\rangle}\subseteq\boldsymbol{P}$. In particular, there exists a bound M with $\|f^{\langle n\rangle}\|_D\leqslant M/2$ for all n. For any $n\leqslant m$, it follows that

$$f^{\langle m\rangle}-f^{\langle n\rangle} \;\trianglelefteqslant\; \frac{M}{(1-z_1/\rho_1)\cdots(1-z_d/\rho_d)}.$$

Since (21) is recursive, the sequence $f^{\langle n\rangle}$ tends coefficientwise to the unique power series solution f of (21). More precisely, for any n and $i\in\mathbb{N}^d$ such that $\|i\|<n$, we have $f_i^{\langle n\rangle}=f_i$. Therefore, given $m\geqslant n$, the valuation of $f^{\langle m\rangle}-f^{\langle n\rangle}$ is at least n. Consequently,

$$f^{\langle m\rangle}-f^{\langle n\rangle} \;\trianglelefteqslant\; M\,\frac{(z_1/\rho_1)^{\lfloor n/d\rfloor}\cdots(z_d/\rho_d)^{\lfloor n/d\rfloor}}{(1-z_1/\rho_1)\cdots(1-z_d/\rho_d)}.$$

In particular,

$$\|f^{\langle m\rangle}-f^{\langle n\rangle}\|_{\tau D} \;\leqslant\; M\,\frac{\tau^{d\lfloor n/d\rfloor}}{(1-\tau)^d}.$$

We conclude that $\|f^{\langle m\rangle}-f^{\langle n\rangle}\|_{\tau D}\to 0$ for $n\to\infty$, whence $f^{\langle n\rangle}$ converges to f on the polydisk τD. Since $f^{\langle n\rangle}\in\boldsymbol{P}_{\tau D}$ for all n, we also get $f\in\boldsymbol{P}_{\tau D}$, by continuity. □

Remark 13. It is somewhat unsatisfactory that the final bound $f \in \boldsymbol{P}_{\tau D}$ only holds on slightly smaller disks. In would be interesting to investigate under which conditions we have $f \in \boldsymbol{P}_D$. This is in particular the case if $\Phi(\boldsymbol{P}) \sqsubseteq \boldsymbol{P}$ and $\Phi(\boldsymbol{P})_i \subseteq \mathring{\boldsymbol{P}}_i$ for all $i \in \partial \mathcal{J}$. Indeed, by continuity with respect to ε, this stronger condition implies that $\Phi(\boldsymbol{P}_{(1+\varepsilon)D}) \subseteq \boldsymbol{P}_{(1+\varepsilon)D}$ for some small $\varepsilon > 0$. We also notice that ball arithmetic can be slightly "inflated" by adding a tiny constant to the radius at every operation. Now assume that $\Phi(\boldsymbol{P})_i \subseteq \boldsymbol{P}_i$ for this kind of inflated ball arithmetic. For the usual arithmetic and each j, we then either have $\Phi(\boldsymbol{P})_i^{[j]} \subseteq \mathring{\boldsymbol{P}}_i^{[j]}$, or the expression which computes $\Phi^{[j]}(\boldsymbol{P})_i$ is identically equal to $\boldsymbol{P}_i^{[j]}$.

5.4. Contraction of Φ on sufficiently small disks

With the notations from the previous section, assume in addition that $\mathcal{J} \subseteq \mathcal{I}$. Let $0 < \varepsilon < 1$ and let $\boldsymbol{f} \in \mathbb{K}[[z]]$ be a power series with ball coefficients. Then the restriction $\boldsymbol{f}_{\mathcal{I}} \in \mathbb{K}[z]_{\mathcal{I}}$ can naturally be reinterpreted as a Taylor model in $\mathbb{T}_{D,\mathcal{I},\mathcal{J}}$, which we will denote by $\boldsymbol{f}_{D,\mathcal{I},\mathcal{J}}$. For what follows, we recall that $\Phi(\boldsymbol{f})$ is computed in $\mathbb{K}[[z]]$, whereas $\Phi(\boldsymbol{f}_{D,\mathcal{I},\mathcal{J}})$ is computed in $\mathbb{T}_{D,\mathcal{I},\mathcal{J}}$.

Lemma 14. *Let Φ be a dag, $\boldsymbol{f} \in \mathbb{K}[[z]]$ and $\boldsymbol{P} = \boldsymbol{f}_{\varepsilon D,\mathcal{I},\mathcal{J}} \in \mathbb{T}_{\varepsilon D,\mathcal{I},\mathcal{J}}$. For $\varepsilon \to 0$, we have*

$$\Phi(\boldsymbol{P}) = \Phi(\boldsymbol{f})_{\varepsilon D,\mathcal{I},\mathcal{J}} + O(\varepsilon).$$

Proof. The lemma follows by an easy induction over the size of Φ. Let us for instance treat the case when $\Phi = \Psi\,\Xi$. Then

$$\begin{aligned}
(\Psi\,\Xi)(\boldsymbol{P})_i &= \sum_{j+k=i} \Psi(\boldsymbol{P})_j\,\Xi(\boldsymbol{P})_k + \sum_{\substack{j+k>i \\ \pi_{\mathcal{J}}(j+k)=i}} \Psi(\boldsymbol{P})_j\,\Xi(\boldsymbol{P})_k \mathcal{B}(\varepsilon\,\rho)^{\pi_{\mathcal{J}}(j+k)-j-k} \\
&= \sum_{j+k=i} (\Psi(\boldsymbol{f})_j + O(\varepsilon))\,(\Xi(\boldsymbol{f})_k + O(\varepsilon)) + O(\varepsilon) \\
&= \sum_{j+k=i} \Psi(\boldsymbol{f})_j\,\Xi(\boldsymbol{f})_k + O(\varepsilon) \\
&= (\Psi\,\Xi)(\boldsymbol{f})_i + O(\varepsilon),
\end{aligned}$$

for all $i \in \mathcal{I}$. □

Theorem 15. *Assume that* (21) *is a recursive equation. Then there exists an $\varepsilon > 0$ and a Taylor model $\boldsymbol{P} \in \mathbb{T}_{\varepsilon D,\mathcal{I},\mathcal{J}}$, such that $\Phi(\boldsymbol{P}) \sqsubseteq \boldsymbol{P}$.*

Proof. Let $f \in \mathbb{C}[[z]]^r$ be the exact solution to (21). Let $\nu = \max_{i \in \mathcal{J}} \|i\|$, $\delta = \varepsilon^{1/(\nu+1)}$ and

$$E = \sum_{i \in \mathcal{J}} \delta^{\nu - \|i\|} z^i.$$

Consider the Taylor model

$$\boldsymbol{P} = f_{\varepsilon D,\mathcal{I},\mathcal{J}} + \mathcal{B}_{\varepsilon D}(E).$$

By lemma 14, we have

$$\begin{aligned}
\Phi(\boldsymbol{P})_i &= \Phi(f_{\mathcal{I}} + \mathcal{B}_{\varepsilon D}(E))_i + O(\delta^{\nu+1}) \\
&= \Phi(f_{\mathcal{I}})_i + O(\delta^{\nu - \|i\| - 1}) \\
&= f_i + O(\delta^{\nu - \|i\| + 1}) \\
&\sqsubseteq f_i + \mathcal{B}_{\varepsilon D}\big(\delta^{\nu - \|i\|}\big),
\end{aligned}$$

for sufficiently small ε. □

Remark 16. In practice, in order to find a suitable $\boldsymbol{P}$ with $\Phi(\boldsymbol{P}) \sqsubseteq \boldsymbol{P}$, we first compute an enclosure $\boldsymbol{f}_{\mathcal{I}}$ for $f_{\mathcal{I}}$ by solving (21) in $\mathbb{K}[[z]]$ using a relaxed algorithm. Then the problem reduces to the computation of a bound $R \in \mathbb{R}^{\geqslant}[z]_{\partial\mathcal{J}}$ with $\Phi(\boldsymbol{P}) \sqsubseteq \boldsymbol{P}$ for $\boldsymbol{P} = \boldsymbol{f}_{\mathcal{I}} + \mathcal{B}(R)$. This really comes down to determining a fixed point for the mapping $R \mapsto \operatorname{rad}(\Phi(\boldsymbol{f}_{\mathcal{I}} + \mathcal{B}_D(R))_{\partial\mathcal{J}})$. We refer to [vdH07b, vdH09a] for details on how to do this.

We also notice that the coefficients of $\Phi(\boldsymbol{f}_{\mathcal{I}} + \mathcal{B}_D(R))$ which are not in $\partial\mathcal{J}$ do not depend on R. For large expansion orders, it may be wise to implement Taylor models in such a way that these coefficients need not to be recomputed for different values of R. In [vdH07b, vdH09a], we describe how to do this in the univariate case.

5.5. Implicit equations

Let us now return to the system (14) of implicit equations and assume that Φ is algebraic in $F, \int F, \int\int F, \ldots$. We will use a domain $\mathbb{T}_{D,\mathcal{I},\mathcal{J}}$ of Taylor models with $D = \mathcal{B}(\rho)$, $\rho \in \mathbb{R}^{>}$ and $\mathcal{J} \subseteq \mathcal{I} = \{0, \ldots, \nu\}$. Using the theory of sections 3 and 4.1, we may construct a recursive equation for the solution f of (14). By (17), this equation has the form

$$f \;=\; \Psi(f) = \Pi \begin{pmatrix} -z^i\,\Phi^{\langle i\rangle}(f) \\ \vdots \\ -z\,\Phi^{\langle i\rangle}(f) \end{pmatrix} z^{-i} - \mathrm{X} \begin{pmatrix} z^i f \\ \vdots \\ z f \end{pmatrix},$$

for certain matrices Π and X whose entries are coefficientwise operators. In order to apply theorems 12 and 15 we need to show that these operators and δ^{-1} are bounded.

Let us first show that δ^{-1} is bounded. Let $h \in \mathbb{C}[[z]]$ be a convergent series on $\mathcal{B}(\rho)$. For any $n > 0$ and $z \in \mathcal{B}(\rho)$, we have

$$\begin{aligned} |(\delta^{-1}(z^n h))(z)| &\leqslant \int_0^\rho \left| t^{n-1} h\!\left(\tfrac{z}{|z|}\,t\right)\right| \mathrm{d}t \\ &\leqslant \|h\|_\rho \int_0^\rho t^{n-1}\,\mathrm{d}t \\ &= \frac{1}{n}\,\|h\|_\rho\,\rho^n \\ &= \frac{1}{n}\,\|z^n h\|_\rho. \end{aligned}$$

We may therefore take $(\delta^{-1})^*_n = \frac{1}{n}$ for $n > 0$ and $(\delta^{-1})^*_0 = 2$.

Now consider one of the entries ω of Π or X. By choosing ν sufficiently large, we may assume without loss of generality that ω_n is given by the evaluation of a rational function in $\mathbb{K}[N]$ at $N = n$ for all $n \geqslant \nu$. Moreover, the coefficients of the matrix (15) are given by (13), whence they are bounded for $n \to \infty$. Consequently, ω_n also remains bounded for $n \to \infty$, and there exists an absolutely convergent expansion

$$\omega_n \;=\; c_0 + \frac{c_1}{n} + \frac{c_2}{n^2} + \cdots. \tag{22}$$

Hence, we may regard ω as an infinite sum

$$\omega \;=\; c_0 + c_1\,\delta^{-1} + c_2\,\delta^{-2} + \cdots.$$

Taking ν sufficiently large, we may also ensure that (22) converges for $n \geqslant \nu$. Consequently, it suffices to take

$$\omega_n^* \;\geqslant\; |c_0| + \frac{|c_1|}{n} + \frac{|c_2|}{n^2} + \cdots$$

for all $n \geqslant \nu$. We may use crude bounds for the remaining ω_n^* with $n < \nu$. Combining theorems 9, 12 and 15, we now obtain the following theorem.

Theorem 17. *Consider an i-predictive system of equations* (14). *Then we may compute $\varepsilon>0$ and a Taylor model $\boldsymbol{P}\in\mathbb{T}_{\mathcal{B}(\varepsilon),\mathcal{I},\mathcal{J}}$, such that its unique solution f is convergent on $\mathcal{B}(\varepsilon)$ and $f\in\boldsymbol{P}$.*

In the Taylor model setting, it is interesting to study the solution f under small perturbations of the initial conditions. This can be done by switching to a multivariate context as in the previous subsections. In the system (14), we now replace the time z by z_d. We also replace the initial conditions $C_0, \ldots, C_l$ by Taylor models $\tilde{C}_0, \ldots, \tilde{C}_l$ in $\mathbb{T}_{D,\mathcal{I},\mathcal{J}}$ which only depend on $z_1, \ldots, z_{d-1}$, and such that $(\tilde{C}_k)_0 = C_k$ for all k. Although the above theory mostly generalizes in a straightforward way, there is a small quirk: when a matrix of the type (15) does not have full rank, then the rank of a small perturbation of it is generally strictly higher. Consequently, i-predictivity is not preserved under small perturbations.

Nevertheless, if we restrict our attention to 1-predictive systems, for which the matrix (15) does have full rank, then small perturbations of the system remain 1-predictive and the generalization of theorem 17 provides us with the complete flow of the implicit equation at any order. Such 1-predictive systems are very common in practice and a well known example is the pendulum, whose equations are given by

$$\begin{aligned}\ddot{x} &= -\lambda x\\ \ddot{y} &= g-\lambda y\\ x^2+y^2 &= \mu.\end{aligned}$$

More generally, dynamical systems on implicitly defined varieties are 1-predictive systems. Notice that the classical implicit function theorem also falls into this category.

Of course, if the solution f is known in a certain disk D (as well as some of its derivatives, if necessary), then we may use the value of f at a given point $z\in D$ as a new initial condition at z and compute an analytic continuation of the solution around z with the same method. In general 1-predictivity is preserved throughout this process, whence our method yields an efficient method for high precision integration of dynamical systems on implicitly defined varieties. Predictive systems of a higher index usually degrade into 1-predictive systems after one analytic continuation. If not, then there is usually an algebraic reason for that, in which case the system can be rewritten into a "simpler system" using differential algebra [Rit50, Kol73, Bou94]. Unfortunately, although the order (or ranking) of the new system is lower, its size as a dag usually explodes.

Bibliography

[AH83] G. Alefeld and J. Herzberger. *Introduction to interval analysis*. Academic Press, New York, 1983.

[ANS08] ANSI/IEEE. IEEE standard for binary floating-point arithmetic. Technical report, ANSI/IEEE, New York, 2008. ANSI-IEEE Standard 754-2008. Revision of IEEE 754-1985, approved on June 12, 2008 by IEEE Standards Board.

[BCO+06] A. Bostan, F. Chyzak, F. Ollivier, B. Salvy, É. Schost, and A. Sedoglavic. Fast computation of power series solutions of systems of differential equation. preprint, april 2006. submitted, 13 pages.

[BK78] R.P. Brent and H.T. Kung. Fast algorithms for manipulating formal power series. *Journal of the ACM*, 25:581–595, 1978.

[BL11] J. Berthomieu and R. Lebreton. Relaxed p-adic hensel lifting for algebraic systems. Work in preparation, 2011.

[Bou94] F. Boulier. *Étude et implantation de quelques algorithmes en algèbre différentielle*. PhD thesis, University of Lille I, 1994.

[CK91] D.G. Cantor and E. Kaltofen. On fast multiplication of polynomials over arbitrary algebras. *Acta Informatica*, 28:693–701, 1991.

[CT65] J.W. Cooley and J.W. Tukey. An algorithm for the machine calculation of complex Fourier series. *Math. Computat.*, 19:297–301, 1965.

[HLRZ00] G. Hanrot, V. Lefèvre, K. Ryde, and P. Zimmermann. MPFR, a C library for multiple-precision floating-point computations with exact rounding. `http://www.mpfr.org`, 2000.

[Kol73] E.R. Kolchin. *Differential algebra and algebraic groups*. Academic Press, New York, 1973.

[LS03] G. Lecerf and É. Schost. Fast multivariate power series multiplication in characteristic zero. *SADIO Electronic Journal on Informatics and Operations Research*, 5(1):1–10, September 2003.

[MB96] K. Makino and M. Berz. Remainder differential algebras and their applications. In M. Berz, C. Bischof, G. Corliss, and A. Griewank, editors, *Computational differentiation: techniques, applications and tools*, pages 63–74, SIAM, Philadelphia, 1996.

[MB04] K. Makino and M. Berz. Suppression of the wrapping effect by Taylor model-based validated integrators. Technical Report MSU Report MSUHEP 40910, Michigan State University, 2004.

[MKC09] R.E. Moore, R.B. Kearfott, and M.J. Cloud. *Introduction to Interval Analysis*. SIAM Press, 2009.

[Moo66] R.E. Moore. *Interval Analysis*. Prentice Hall, Englewood Cliffs, N.J., 1966.

[Neu90] A. Neumaier. *Interval methods for systems of equations*. Cambridge university press, Cambridge, 1990.

[Rit50] J.F. Ritt. *Differential algebra*. Amer. Math. Soc., New York, 1950.

[Rum10] S.M. Rump. Verification methods: Rigorous results using floating-point arithmetic. *Acta Numerica*, 19:287–449, 2010.

[SS71] A. Schönhage and V. Strassen. Schnelle Multiplikation grosser Zahlen. *Computing*, 7:281–292, 1971.

[vdH97] J. van der Hoeven. Lazy multiplication of formal power series. In W. W. Küchlin, editor, *Proc. ISSAC '97*, pages 17–20, Maui, Hawaii, July 1997.

[vdH02] J. van der Hoeven. Relax, but don't be too lazy. *JSC*, 34:479–542, 2002.

[vdH05] J. van der Hoeven. Notes on the Truncated Fourier Transform. Technical Report 2005-5, Université Paris-Sud, Orsay, France, 2005.

[vdH07a] J. van der Hoeven. New algorithms for relaxed multiplication. *JSC*, 42(8):792–802, 2007.

[vdH07b] J. van der Hoeven. On effective analytic continuation. *MCS*, 1(1):111–175, 2007.

[vdH09a] J. van der Hoeven. Ball arithmetic. Technical report, HAL, 2009. `http://hal.archives-ouvertes.fr/hal-00432152/fr/`.

[vdH09b] J. van der Hoeven. Relaxed resolution of implicit equations. Technical report, HAL, 2009. `http://hal.archives-ouvertes.fr/hal-00441977/fr/`.

[vdH10] J. van der Hoeven. Newton's method and FFT trading. *JSC*, 45(8):857–878, 2010.

[vdHL10] J. van der Hoeven and G. Lecerf. On the bit-complexity of sparse polynomial multiplication. Technical report, HAL, 2010. `http://hal.archives-ouvertes.fr/hal-00476223/fr/`.

[vdHS10] J. van der Hoeven and É. Schost. Multi-point evaluation in higher dimensions. Technical report, HAL, 2010. `http://hal.archives-ouvertes.fr/hal-00477658/fr/`.