BIT Numerical Mathematics
Submitted October 2005, pp. 001-017

0006-3835/03/4301-0001 $16.00

Polynomial Cost for Solving IVP for High-Index DAE

ROBERT M. CORLESS[1] * SILVANA ILIE[2] † ‡

[1] *Department of Applied Mathematics, University of Western Ontario, London, Ontario, N6A 5B7, Canada*
email: rcorless@uwo.ca
[2]*Computer Science Department, University of Toronto Toronto, Ontario, M5S 1A4, Canada*
e-mail: silvana@cs.toronto.edu

Abstract.

We show that the cost of solving initial value problems for high-index differential algebraic equations is polynomial in the number of digits of accuracy requested. The algorithm analyzed is built on a Taylor series method developed by Pryce for solving a general class of differential algebraic equations. The problem may be fully implicit, of arbitrarily high fixed index and contain derivatives of any order. We give estimates of the residual which are needed to design practical error control algorithms for differential algebraic equations. We shown that adaptive meshes are always more efficient than non-adaptive meshes. Finally, we construct sufficiently smooth interpolants of the discrete solution.

AMS subject classification (2000): 34A09, 65L80, 68Q25.

Key words: differential algebraic equations, initial value problem, adaptive step-size control, Taylor series, structural analysis, automatic differentiation, Hölder mean.

1 Introduction.

Many physical problems of high interest today are modeled as differential algebraic equations (DAE). Therefore, the need for efficient numerical methods of solving them has grown rapidly [2, 18]. Among these numerical methods, Taylor series methods have recently proved to be an attractive choice on many interesting problems arising in practice.

From the classical results of Riquier, Janet and Cartan [7, 25, 34] it follows that Taylor series methods apply in the more general context of partial differential algebraic equation (PDAE) solving, for generic systems. In this paper, we shall analyze the numerical solution of a certain class of DAE. For this class, Pryce developed a Taylor series method based on his structural analysis [30, 31] and on Pantelides' [29]. While

*Research partially supported by a grant from the Natural Science and Engineering Research Council of Canada.

†Corresponding author.

‡Research supported by a postdoctoral fellowship from the Natural Science and Engineering Research Council of Canada.

Pantelides' structural analysis relies on solving an algorithm for iteratively determining the hidden constraints, Pryce's analysis is based on solving a priori an assignment problem, allowing, in addition, higher order derivatives.

The Taylor series coefficients can be efficiently computed with automatic differentiation. The idea of automatically generating recursion formulae for efficient computation of Taylor series goes back at least to Moore [26]. Applications of automatic differentiation to solving differential equations were considered in Corliss and Chang [11].

In two previous papers, we studied the theoretical (asymptotic) complexity, as the accuracy of the computed residual, ε, gets better, of the numerical solution of IVP for ODE in [10] and for index-1 DAE in [22]. The main results of those papers are that the asymptotic cost of the numerical solution of such initial value problems is polynomial in the number of digits of accuracy, $\mathcal{B} = \lceil \log_2(1/\varepsilon) \rceil$. These results improved on the results of the standard theory of information based complexity (see, e.g., [36]), which predicted exponential cost of solving IVP for ODE. We believe that our model, assuming more smoothness of the problem and solution, is more realistic.

A second result is that adaptive step-size, under modest realistic assumptions, is never worse than fixed stepsizes and can be an unbounded factor better.

In the present paper, we extend the results to IVP for higher index DAE. We analyze an algorithm based on Pryce's method [30, 31] for solving numerically a general class of DAE that can be high-index, fully implicit and may contain derivatives of arbitrary order. We consider error control in terms of defect (or residual) as it is robust, of low cost, and it has a straightforward interpretation [13]. However, a similar analysis can be obtained for forward error. A study of complexity for solving initial value problems and boundary value problems for ordinary differential equations using local error control is given in [23].

The analysis of the Taylor series method for DAE, as discussed by Pryce [31] and Nedialkov & Pryce [28], is complex (see also [21]). Another difficulty is an appropriate (smooth enough) choice of interpolant to facilitate the analysis. In the end, the analysis of this paper requires more effort than the index-1 case, as might be expected.

The paper assumes that the index of the problem does not change throughout the integration, and that singularities and discontinuities are isolated and may be located at negligible cost and that solutions are otherwise smooth. We also assume that the DAE is suitable for automatic differentiation [33, 16]. Under these assumptions, we show that the numerical solution of IVP for index-n DAE is again of cost polynomial in the number of digits of accuracy. We analyze the asymptotic behavior of the residual and give error estimates which can be used together with the existing error controllers [35] for an efficient practical implementation of Pryce's method.

We have not investigated the numerical stability of this Taylor series method in floating point arithmetic as the order becomes large. Earlier heuristic work by Corliss & Chang [11] show the good behavior of Taylor series methods for ODE in practice, while Nedialkov & Pryce [28] report good experimental results for these methods for DAE. Pryce [32] gives a theoretical study of the growth of round-off error as the order of the Taylor method increases.

The study of asymptotic cost for solving differential equations sheds light on finite precision and finite order methods. Our theoretical analysis can guide the design and analysis of practical codes for highly accurate solutions of arbitrary index differential algebraic equations.

2 Pryce's structural analysis method for DAEs.

We consider an autonomous initial value problem for a differential algebraic equation in m dependent variables $x_j = x_j(t)$, with t, the scalar independent variable, in an interval $[a, b]$. The DAE has the general form

$$f_i(x_j, \text{ derivatives of } x_j) = 0 \quad \text{for } i = 1 \dots m \,. \tag{2.1}$$

As is well-known, non-autonomous problems can be put in an autonomous form by increasing the dimension by one. The functions f_i are assumed *piecewise analytic* in their variables. We allow higher order derivatives of the dependent variables and nonlinear expressions in them.

We assume that all standard functions used in defining the DAE are solutions of simple ODEs. Consequently, by adding them to the DAE the problem is converted to a larger DAE whose new functions involve only the four basic arithmetic operations. The functions f_i belong thus to the class of functions to which automatic differentiation applies. This class excludes, for example, the Gamma function.

A *consistent point* of the DAE (2.1) is defined as a scalar t^* together with a set of scalars η_{jl}, where (j, l) are all the elements of a finite set S_J, so that there exists a solution to the DAE (2.1) in a neighborhood of $t = t^*$ with $x_j^{(l)}(t^*) = \eta_{jl}$ for all $(j, l) \in S_J$ and the solution is unique. The set S_J need not be minimal.

In differential geometry, a consistent point is obtained as follows. First, the formally integrable form of the system is obtained after a number of differentiations and eliminations (which uncover all the hidden constraints). The formally integrable form defines a manifold in the appropriate jet space. A consistent point is a point of this manifold. As we shall see below, to obtain such a consistent point with Pryce's method, no elimination and generally *fewer differentiations* are needed. Therefore, for the class of problems on which the structural analysis is successful, Pryce's method is expected to be more efficient in unveiling the hidden constraints.

We assume, for simplicity, that the initial condition satisfies exactly a set of constraints given by Pryce's method for a consistent point for (2.1). The case when the initial condition satisfies with a good enough precision these constraints may add little to the overall cost.

In [31, 30] Pryce develops an algorithm for solving such a system by expanding the solution in Taylor series. The method is based on a pre-analysis stage which consists in solving an assignment problem. The assignment problem gives the *offsets*, that is a set of nonnegative integers $c_i \geq 0,\ d_j \geq 0$, with c_i being the number of times the equation f_i has to be differentiated to reduce the DAE to an ODE and d_j is the order of variable x_j in this ODE. The offsets also indicate which equations to solve for which unknowns, and give a systematic way of determining consistent initial conditions. The offsets are considered in the context of elliptic partial differential equation solving by Douglis & Nirenberg in [12].

The next step of the method consists in generating the system Jacobian. Provided the Jacobian is non-singular at each integration step, the method succeeds and the Taylor series coefficients can be computed to the desired order. This is based on the fact that, at each step of integration, after some initial stages, the Taylor coefficients can be obtained as solutions of some linear systems involving the system Jacobian.

2.1 Structural analysis.

We review below the main steps of Pryce's structural analysis and the corresponding algorithm, following [31, 30].

1. Form an $m \times m$ *signature matrix* $\Sigma = (\sigma_{ij})$ with

$$\sigma_{ij} = \begin{cases} \text{order of derivative to which the variable } x_j \\ \qquad \text{appears in equation } f_i \ , \\ \text{or } -\infty \text{ if the variable does not occur .} \end{cases}$$

2. Solve an assignment problem to determine a HVT (*highest value transversal*), which is a subset of indices (i, j) describing just one element in each row and each column, such that $\sum \sigma_{ij}$ is maximized and is finite. We make the assumption that such a HVT exists.

3. Determine the *offsets* of the problem, which are the vectors **c**, **d** with d_i, $c_j \geq 0$ the smallest such that

$$d_j - c_i \geq \sigma_{ij} \qquad \text{for all } 1 \leq i \leq m,\ 1 \leq j \leq m$$

with equality on the HVT.

The *structural index* is then defined as

$$\nu = \max_i c_i + \begin{cases} 0 & \text{if all } d_j > 0 \\ 1 & \text{if some } d_j = 0. \end{cases}$$

The structural index is no less that the differential index on first order DAE [31].

4. Form the *system Jacobian* **J**, given by

$$\mathbf{J}_{ij} = \begin{cases} \dfrac{\partial f_i}{\partial x_j^{(d_j - c_i)}} & \text{if } x_j^{(d_j-c_i)} \text{ appears in } f_i \ , \\ 0 & \text{otherwise.} \end{cases}$$

5. Choose a consistent point. If **J** is non-singular at that point, then the solution can be computed with Taylor series in a neighborhood of that point.

This is done by assuming an expansion

$$x_j(t) = \sum_{\ell \geq 0} \frac{1}{\ell!} x_{j,\ell} (t - t^*)^\ell \ .$$

Then substitute this in (2.1) and expand the equations in Taylor series

$$f_i(x_j(t), \text{ derivatives of } x_j(t)) = \sum_{q \geq 0} \frac{1}{q!} f_{i,q} (t - t^*)^q \ ,$$

and solve $f_{i,q} = 0$. Each $f_{i,q} = f_i^{(q)}(t^*)$ depends on a finite number of the $x_{j,\ell} = x_j^{(\ell)}(t^*)$.

2.2 Notation.

We define $k_d = -\max_j d_j$ and $k_c = -\max_i c_i$ (necessarily $k_d \leq k_c$). In this paper we consider $k_c \leq -1$, in which case the DAE (2.1) is at least of index 1. The case of index-1 DAEs with $k_c = 0$ is considered in [22].

We follow the notation from [31] and define for $k \geq k_d$:

$$\begin{aligned} J_k &= \{(j,\ell) : \ell = k + d_j \geq 0, 1 \leq j \leq m\} \\ I_k &= \{(i,q) : q = k + c_i \geq 0, 1 \leq i \leq m\}. \end{aligned}$$

We use the following notation: for a set I we define the vector

$$f_I = \left[f_i^{(q)} : \text{the } q\text{-th derivative of } f_i, (i,q) \in I \right] \tag{2.2}$$

and for a set J we define the vector

$$x_J = \left[x_j^{(\ell)} : \text{the } \ell\text{-th derivative of } x_j, (j,\ell) \in J \right] .$$

We denote by $I_{\leq k}$ ($I_{<k}$) the union of all I_r with $r \leq k$ ($r < k$ respectively) and by $J_{\leq k}$ ($J_{<k}$) the corresponding unions.

We shall also use the notation f_{I_k} (and x_{J_k}) for the values of the functions at a specific time and we shall omit to write the dependency of the time, if this is clear from the context.

2.3 Theoretical results.

The choice of the offsets induces a specific block triangular structure on the system of equations to be solved and this structure is exploited by the algorithm. At each stage k, f_{I_k} is a function only of the variables $x_{J_{\leq k}}$. The algorithm consists in solving, for each stage $k = k_d, k_d + 1, k_d + 2, \ldots$ the equations

$$f_{I_k} = f_{I_k}(x_{J_{<k}}, x_{J_k}) = 0 \quad \text{for variables } x_{J_k}. \tag{2.3}$$

This is equivalent to solving, for each k, the system $f_{i,k+c_i} = 0$ for those i with $k + c_i \geq 0$ in the unknowns $x_{j,k+d_j}$ for those j with $d_j + k \geq 0$.

If we consider

$$m_k = |I_k|, \qquad n_k = |J_k|,$$

where $|I|$ represents the cardinality of set I, then $m_k \leq n_k$ for all stages k, which means that at each stage we solve at most as many equations as unknowns (see [31]).

By permuting the variables $x_j(t)$ so that the d_j are in decreasing order and by permuting the equations f_i so that the c_i are in decreasing order, it can be shown ([31], Proposition 4.1) that the Jacobian

$$\mathbf{J}_k = \frac{\partial f_{I_k}}{\partial x_{J_k}} \tag{2.4}$$

is the leading $m_k \times n_k$ submatrix of $\mathbf{J}$. If $\mathbf{J}$ is non-singular, then $\mathbf{J}_k$ has full row-rank and thus there exists $\mathbf{J}_k^\dagger$, the Moore-Penrose inverse of $\mathbf{J}_k$ (see, e.g., [5]).

If a point $(t^*, x^*_{J_{\leq 0}}) \in \mathbb{R} \times \mathbb{R}^{J_{\leq 0}}$ satisfies (2.3) for all stages $k \leq 0$ then it is a consistent point of the DAE (2.1). If the system Jacobian $\mathbf{J}$ is non-singular at this

point and the functions f_i are analytic in a neighborhood, then the DAE is solvable in some neighborhood of this point and can be reduced to an ODE using this method ([31], Theorem 4.2). Alternatively, the DAE (2.1) can be solved by Taylor series in a neighborhood [28, 31].

We assume the initial conditions

$$\frac{d^\ell}{dt^\ell} x_j(a) = x^0_{j,\ell} \text{ for all } \ell \leq d_j$$

satisfy the stages $k \leq 0$ of Pryce's algorithm at $t = a$,

$$f_{I_k}(x^0_{J_{<k}}, x^0_{J_k}) = 0$$

for all $k \leq 0$ and thus is a consistent point for the DAE (2.1). We also assume that the system Jacobian is non-singular in a neighborhood of the exact solution on the interval $[a, b]$.

We are interested in obtaining an approximate solution, χ_j, with a tolerance ε in residual. We define each residual by

$$\delta_i(t) = f_i(t, \chi_j(t), \text{ derivatives of } \chi_j(t)) \tag{2.5}$$

and we denote the vector of residuals at each stage by δ_{I_k}, defined in a similar way as (2.2).

DEFINITION 2.1. *The approximate solution $\chi_j(t)$ satisfies the DAE (2.1) with a tolerance ε in residual if the residuals (2.5) satisfy*

$$\max_{t\in[a,b]} \|\delta_{I_k}(t)\| \leq \varepsilon \qquad \textit{for all } k_c \leq k \leq 0\,. \tag{2.6}$$

Here we used $\|u(t)\|$ to denote the vector norm at a fixed time t.

3 Numerical solution.

We are interested in finding a bound for the cost of the following algorithm for obtaining a numerical solution of the DAE (2.1) using the structural analysis of [31]: assume we have obtained at time t_n some values $x^n_{J_{\leq 0}}$ which satisfy the constraints $f_{I_{k\leq 0}} = 0$ more accurately than the desired tolerance, in a sense made precise later. We generate the Taylor coefficients at $t = t_n$ for the unknown functions up to the desired order and predict the values $\hat{x}^{n+1}_{J_{\leq 0}}$ at the next integration step by computing the Taylor series with a chosen stepsize h_n.

We correct these values by projecting the approximate solution back on the constraint manifold. We do this by applying only one Newton projection at each stage $k \leq 0$. As suggested by numerical experiments, one Newton iteration is sufficient for satisfying sufficiently accurately the constraint corresponding to that stage, see, e.g., [3]. Due to the use of a high order method, the approximate Taylor solution at one step stays very close to the exact solution at tight accuracies and the convergence of Newton iteration is assured. The advantage is that the cost of the algorithm is not increased unnecessarily.

The step is accepted if a given measure of the residuals satisfy the tolerance (and thus guaranteeing, as shown later, that a continuous extension of the discrete solution satisfies (2.6)).

3.1 Predictor step.

We consider the truncated Taylor solution at time $t = t_n + h$

$$\hat{x}_j(t) = \sum_{\ell=0}^{p+d_j-1} \frac{1}{\ell!} x_{j,\ell}^n (t-t_n)^\ell \quad \text{for } 1 \le j \le m\ . \tag{3.1}$$

Substitute it in the equations and expand in Taylor series:

$$f_i(\hat{x}_j, \text{ derivatives of } \hat{x}_j) = \sum_{q\ge 0} \frac{1}{q!} f_{i,q}(x_{j,\ell}^n)(t-t_n)^q \quad \text{for } 1 \le i \le m\ .$$

We solve the systems $f_{i,k+c_i}(x^n_{J_{\le k}}) = 0$ (for $k + c_i \ge 0$) in the unknowns $x^n_{j,k+d_j}$ (for $k + d_j \ge 0$) for each $k = k_d, k_d + 1, \ldots, p-1$.

For $k \le 0$, these equations are already satisfied with a much better accuracy than the tolerance. Indeed, at the first integration step, these equations are satisfied by the initial consistent point, and at the other integration steps, they are satisfied since the values are the ones projected on the constraint manifold at the previous step (see Remark 4.3 below).

The stages $k \ge 1$ are always linear in the corresponding unknowns and involve the Jacobian $\mathbf{J}(x^n_{J_{\le 0}})$, which is non-singular if the chosen tolerance is small enough.

For each $t \in [a, b]$ we define

$$\hat{\delta}_i(t) = f_i(\hat{x}_j(t), \text{ derivatives of } \hat{x}_j(t))\ , \qquad i = 1 \ldots m \tag{3.2}$$

to be the local residuals obtained for the approximate Taylor solution (3.1) on each $[t_n, t_n + h_n)$.

At time $t_{n+1} = t_n + h_n$ the predicted values are

$$\hat{x}^{n+1}_{j,k} = \hat{x}_j^{(k)}(t^-_{n+1}) \tag{3.3}$$

for $k \le d_j$, $1 \le j \le m$.

The predicted point, $\hat{x}^{n+1}_{J_{\le 0}}$ is close to a consistent point of the DAE, and its distance to the constraint manifold is given by the truncation error. So for a small ε the predicted point will be in a neighborhood of the exact solution where the system Jacobian, $\mathbf{J}$, is non-singular. In the same neighborhood, each $\mathbf{J}_k$ given by (2.4) is full row-rank and there exists $\mathbf{J}_k^\dagger$, the Moore-Penrose inverse of $\mathbf{J}_k$.

3.2 Corrector step.

For each $k \le 0$ we have at most as many equations as unknowns. We wish to satisfy these equations with a better accuracy than the one given by the truncation error. For each stage $k \le 0$ we compute a set of corrected values, by applying a single Newton iteration per stage. More would be superfluous.

For $k < k_c$, there are no constraints to satisfy and we preserve the values of the variables corresponding to that stage : $x_{J_k} = \hat{x}_{J_k}$ for all $k_d \le k < k_c$ (if $k_d < k_c$).

In what follows, the difference between the corrector at the end of the step, t_{n+1}, and the corresponding predictor defined by (3.3) will be denoted by

$$u^{n+1}_{J_k} = x^{n+1}_{J_k} - \hat{x}^{n+1}_{J_k} \tag{3.4}$$

for stages $k = k_c, \ldots, 0$.

Using the notation (3.4), we start with the Gauss-Newton projection

$$u_{J_k}^{n+1} = -\mathbf{J}_k^{\dagger}(\hat{x}_{J_k}^{n+1}) \cdot f_{I_{k_c}}(\hat{x}_{J_{k_c}}^{n+1}) \tag{3.5}$$

for stage $k = k_c$. We continue, at each stage $k = k_c + 1, \ldots, 0$, with the following Gauss-Newton projections

$$u_{J_k}^{n+1} = -\mathbf{J}_k^{\dagger}(\xi_k^{n+1}) \cdot f_{I_k}(\xi_k^{n+1}) \tag{3.6}$$

where

$$\xi_k^{n+1} = (x_{J_{<k}}^{n+1}, \hat{x}_{J_k}^{n+1}) \, . \tag{3.7}$$

The values $x_{J_{<k}}^{n+1}$ are already computed from the previous stages.

4 Error estimates.

We shall analyze the errors of the Newton projections and we shall compute the residuals over one integration step.

In this section we use the notation $f_i(t, x_{J_{\leq 0}}(t))$, although the functions f_i do not explicitly depend on t (system is autonomous), whenever we need to make use of the space in which the solution lies. Otherwise the dependence of t is dropped.

The functions f_i are analytic in their variables, the interval of integration is compact, and the system Jacobian for the exact solution is non-singular. Therefore there exists a neighborhood, U, in the space $[a, b] \times \mathbb{R}^{J_{\leq 0}}$, of the exact solution

$$\{(t, x^e_{J_{\leq 0}}(t)), t \in [a, b]\}$$

such that the Jacobian $\mathbf{J}(z)$ is non-singular and all the Moore-Penrose inverses $\mathbf{J}_k^{\dagger}(z)$, with $k \leq 0$ are uniformly bounded for $z \in U$. More precisely, there exists $L > 0$ so that

$$\|\mathbf{J}_k^{\dagger}(z)\| \leq L \quad \text{for all } k \leq 0 \text{ and for all } z \in U \, . \tag{4.1}$$

For $k_c + 1 \leq k$, $f_{I_k}(t, y_{J_{\leq k}})$ is *Lipschitz continuous* with respect to the variables $\{y_{J_l}\}_{l=k_c,\ldots,k-1}$ in some neighborhood U_0 in $[a, b] \times \mathbb{R}^{J_{\leq 0}}$ of the exact solution, if there exists $M_k > 0$ so that

$$\begin{aligned} &\|f_{I_k}(t, y_{J_{<k_c}}, y^1_{J_{k_c}}, \ldots, y^1_{J_{k-1}}, y_{J_k}) - f_{I_k}(t, y_{J_{<k_c}}, y^2_{J_{k_c}}, \ldots, y^2_{J_{k-1}}, y_{J_k})\| \\ &\leq M_k \sum_{\ell=k_c}^{k-1} \|y^1_{J_\ell} - y^2_{J_\ell}\| \end{aligned} \tag{4.2}$$

for all $(t, y^q_{J_{\leq k}}) \in U_0$ with $y^q_{J_{\leq k}} = (y_{J_{<k_c}}, y^q_{J_{k_c}}, \ldots, y^q_{J_{k-1}}, y_{J_k})$, and $q = 1, 2$.

Let $t = t_n + h$ and $h \in [0, h_n]$.

PROPOSITION 4.1. *Assume that for all $k_c + 1 \leq k \leq 0$, $f_{I_k}(t, y_{J_{\leq k}})$ is Lipschitz continuous with respect to the variables $\{y_{J_l}\}_{l=k_c,\ldots,k-1}$ in some neighborhood $U_0 \subset U$ in $[a, b] \times \mathbb{R}^{J_{\leq 0}}$ of the exact solution. Then asymptotically, as for each n, h_n is small*

enough such that the points of the corresponding predictor step and all the points obtained at the corrector step are within the neighborhood U_0 and the Taylor expansions are valid, the following holds:

$$u_{J_k}^{n+1} = \mathcal{O}(h_n^{p-k}) , \quad k_c \le k \le 0 \tag{4.3}$$

and

$$f_{I_k}(t, \hat{x}_{J_{\le k}}(t)) = \Phi_{I_k}^n h^{(p-k)} + \mathcal{O}(h^{p-k+1}) , \quad k_c \le k \le 0,\ t \in [t_n, t_n + h_n] , \tag{4.4}$$

$$f_{I_k}(\xi_k^{n+1}) = \Phi_{I_k}^n h_n^{(p-k)} + \mathcal{O}(h_n^{p-k+1}) , \quad k_c + 1 \le k \le 0 \tag{4.5}$$

where the vectors $\Phi_{I_k}^n$ do not depend on h_n.

The symbols $\mathcal{O}$ are considered uniformly for $t \in [a, b]$. We note that we are only interested in meshes on which the residual is below a given tolerance ε. Asymptotically, as the tolerance becomes small, all the predicted points and the points resulting at the projection stages for such meshes will be within the neighborhood U_0.

PROOF. Consider

$$\phi_i^n = \frac{1}{(p + c_i)!} \frac{d^{p+c_i}}{dt^{p+c_i}} f_i(x_{J_{\le 0}})(t_n) .$$

These coefficients are independent of h and h_n, depending only on the value of some fixed function at $(x_{J_{\le 0}}^n)$ (the beginning of the step). Then

$$f_i(\hat{x}_j(t), \text{ derivatives of } \hat{x}_j(t)) = \phi_i^n (t - t_n)^{p+c_i} + \mathcal{O}((t - t_n)^{p+c_i+1})$$

where $\hat{x}$ is given by (3.1). Here the symbol $\mathcal{O}$ hides the bound on the derivative $\dfrac{d^{p+c_i+1}}{dt^{p+c_i+1}} f$ across $[a, b]$.

By differentiating $(k + c_i) \ge 0$ times with respect to t, we obtain (4.4), where

$$\Phi_{i,k+c_i}^n = \frac{1}{(p - k)!} \frac{d^{p+c_i}}{dt^{p+c_i}} f_i(x_{J_{\le 0}})(t_n) , \tag{4.6}$$

provided $k \le 0$. This is the coefficient of $(t - t_n)^{p-k}$ in the Taylor series expansion of $\dfrac{d^{c_i+k}}{dt^{c_i+k}} f$, This observation is important for interpreting the principal residual error coefficients (5.8). In particular, we see that the coefficients (4.6) are bounded above *independent of the order of the method and the particular point in the mesh.*

We also note that, by using (4.2) for $k_c + 1 \le k$, we derive

$$f_{I_k}(\xi_k^{n+1}) = f_{I_k}(\hat{x}_{J_{\le k}}^{n+1}) + \sum_{\ell=k_c}^{k-1} \mathcal{O}(\|u_{J_\ell}^{n+1}\|) \tag{4.7}$$

We now prove (4.3) for k and (4.5) for $k + 1$ by induction over $k \ge k_c$.

- Let $k = k_c$. By using (3.5) and the uniform-boundedness condition (4.1), we obtain
$$\|u_{J_{k_c}}^{n+1}\| \leq L\|f_{I_{k_c}}(\hat{x}_{J_{k_c}}^{n+1})\| .$$
If we take $k = k_c$ in (4.4), we derive
$$f_{I_{k_c}}(\hat{x}_{J_{k_c}}^{n+1}) = \mathcal{O}(h_n^{p-k_c})$$
and thus we obtained (4.3) for $k = k_c$.
If we consider $k = k_c + 1$ in (4.7) and we apply (4.3) for $k = k_c$ (obtained above), then we find :
$$f_{I_{k_c+1}}(\xi_k^{n+1}) = f_{I_{k_c+1}}(\hat{x}_{J_{\leq k_c+1}}^{n+1}) + \mathcal{O}(h_n^{p-k_c}) .$$

- We now assume the hypothesis is true for all $\ell \leq k-1$ and show that it is true for k. By using (3.6) and (4.1), we obtain
$$\|u_{J_k}^{n+1}\| \leq L\|f(\xi_k^{n+1})\| .$$
But from the induction step at $k-1$ we know that
$$\|f(\xi_k^{n+1})\| = \mathcal{O}(h_n^{p-k})$$
and thus (4.3) is also true for k.
We now show that (4.5) holds for $k+1$. By applying the induction hypothesis (4.3) for $\ell \leq k$ in (4.7) for $k+1$, we obtain
$$f_{I_{k+1}}(\xi_k^{n+1}) = f_{I_{k+1}}(\hat{x}_{J_{\leq k+1}}^{n+1}) + \mathcal{O}(h_n^{p-k}) .$$

□

REMARK 4.1. In Proposition 4.1, we only need the Lipschitz condition (4.2) to hold for small right hand sides in (4.2). This is satisfied when each derivative $\dfrac{\partial f_{I_k}}{\partial y_{J_\ell}}$ is bounded, for all $\ell = k_c, \ldots, k-1$, on some neighborhood of the exact solution. In particular, if f_i, $i = 1, \ldots, m$ and the exact solution contains only analytic functions on the interval of integration, then Proposition 4.1 applies.

REMARK 4.2. According to Proposition 4.1, for each stage $k = k_c, \ldots, 0$, the Taylor residuals together with the errors committed in the projections, all corresponding to the stage k, are of *the same order*, $\mathcal{O}(h_n^{p-k})$.

REMARK 4.3. For small enough ε, the Newton projections double the number of digits of accuracy in the corresponding residuals, $f_{I_k}(x_{J_{\leq k}}^{n+1}) = \mathcal{O}(h_n^{2(p-k)})$, for $k \leq 0$. Indeed, by a Taylor expansion we obtain
$$\begin{aligned} f_{I_k}(x_{J_{\leq k}}^{n+1}) &= f_{I_k}(\xi_k^{n+1}) + \mathbf{J}_k(\xi_k^{n+1})u_{J_k}^{n+1} \\ &+ \frac{1}{2}u_{J_k}^{n+1\,T} \cdot G_k^{n+1}u_{J_k}^{n+1} \end{aligned}$$
where
$$G_k^{n+1} = \int_0^1 2(1-s)\frac{\partial^2 f_{I_k}}{\partial x_{J_k}\partial x_{J_k}}(x_{J_{<k}}^{n+1}, \hat{x}_{J_k}^{n+1} + su_{J_k}^{n+1})ds .$$

The first two terms in the expansion cancel each other due to (3.5)–(3.6), and then by using (4.1) we derive

$$\|f_{I_k}(x^{n+1}_{J_{\leq k}})\| \leq \frac{1}{2} N_k L^2 \|f_{I_k}(\xi^{n+1}_k)\|^2$$

where $N_k = \sup \|\dfrac{\partial^2 f_{I_k}}{\partial x_{J_k} \partial x_{J_k}}\|$ on some neighborhood. We end the proof by applying Proposition 4.1.

EXAMPLE 4.1. We consider the example of the simple nonlinear pendulum, which is described by the following equations

$$\left\{ \begin{array}{rcrcl} 0 & = & f_1 & = & x'' + x\lambda \\ 0 & = & f_2 & = & y'' + y\lambda - g \\ 0 & = & f_3 & = & x^2 + y^2 - L^2 \end{array} \right. \tag{4.8}$$

where x, y are the ordinary orthogonal coordinates and $L\lambda$ is the tension in the string. The length of the pendulum, L, and the acceleration due to gravity, g, are positive constants.

The offsets of the variables (x, y, λ) are $(2, 2, 0)$, and those of the equations (f_1, f_2, f_3) are $(0, 0, 2)$. Since the system Jacobian is non-singular [31], then the system is solvable by Pryce's method. This is an index-3 DAE [31].

According to the algorithm, the equations to be solved are grouped as follows:

(f_3) in unknowns (x, y) at stage $k = -2$,

(f_3') in unknowns (x', y') at stage $k = -1$,

(f_1, f_2, f_3'') in unknowns (x'', y'', λ) at stage $k = 0$.

According to Remark 4.1, in order for Proposition 4.1 to apply, it suffices to check that the following are bounded:

(I). Stage $k = -1$:

$$\left[\begin{array}{cc} \dfrac{\partial f_3'}{\partial x} & \dfrac{\partial f_3'}{\partial y} \end{array} \right] = \left[\begin{array}{cc} 2x' & 2y' \end{array} \right]$$

(II). Stage $k = 0$:

$$\left[\begin{array}{cccc} \dfrac{\partial f_1}{\partial x} & \dfrac{\partial f_1}{\partial y} & \dfrac{\partial f_1}{\partial x'} & \dfrac{\partial f_1}{\partial y'} \\ \dfrac{\partial f_2}{\partial x} & \dfrac{\partial f_2}{\partial y} & \dfrac{\partial f_2}{\partial x'} & \dfrac{\partial f_2}{\partial y'} \\ \dfrac{\partial f_3''}{\partial x} & \dfrac{\partial f_3''}{\partial y} & \dfrac{\partial f_3''}{\partial x'} & \dfrac{\partial f_3''}{\partial y'} \end{array} \right] = \left[\begin{array}{cccc} \lambda & 0 & 0 & 0 \\ 0 & \lambda & 0 & 0 \\ 2x'' & 2y'' & 4x' & 4y' \end{array} \right]$$

Indeed, with a compactness and smoothness argument, we see that all of these variables are bounded on some neighborhood of the exact solution.

Hence Proposition 4.1 applies and we can derive the order of the residuals and of the errors between the corrector and predictor for the variables at each stage. If, for a given order p, the approximate Taylor solution is computed according to formula (3.1), then these errors, at the initial stages, are those listed in Table 4.1.

Table 4.1: The order of the Taylor residuals and of the errors of projections, for the initial stages in the simple pendulum problem.

Stage	equations	order of residual	unknowns	order of error in projection
$k=-2$	$f_3=0$	h^{p+2}	x, y	h^{p+2}
$k=-1$	$f_3'=0$	h^{p+1}	x', y'	h^{p+1}
$k=0$	$f_1=0$ $f_2=0$ $f_3''=0$	h^p	x'', y'', λ	h^p

5 Dense Output.

In the present paper we are not concerned with finding the optimal interpolant (the one that minimizes some norm of the residual). We want to measure the size of the residual associated with a sufficiently accurate continuous approximation.

We wish to find some interpolants which are at least as smooth as required by the problem (2.1). More precisely, we require each interpolant $\chi_j(t)$, corresponding to the dependent variable $x_j(t)$, to have continuous derivatives up to order d_j on the interval $[a,b]$.

The following construction satisfies this requirement: let $[t_n, t_{n+1}]$ be the current interval of integration and t arbitrary in this interval. Consider χ_j to be a piecewise-polynomial approximation of the variable x_j. On each interval $[t_n, t_{n+1}]$, define χ_j to be the Hermite interpolant which satisfies the following conditions:

$$\begin{aligned} \chi_j^{(k)}(t_n) &= x_{j,k}^n \quad \text{for all } k=0,\dots,p+d_j-1 \\ \chi_j^{(k)}(t_{n+1}) &= x_{j,k}^{n+1} \quad \text{for all } k=0,\dots,d_j\ . \end{aligned} \tag{5.1}$$

This interpolant has continuous derivatives up to order d_j on $[a,b]$.

If for each χ_j we consider the basis

$$\begin{aligned} &(t-t_n)^k && \text{for} \quad 0\le k\le p+d_j-1 \\ &(t-t_n)^{p+d_j}(t-t_{n+1})^k && \text{for} \quad 0\le k\le d_j \end{aligned}$$

then we find the following representation of the interpolant:

$$\chi_j(t) = \hat{x}_j(t) + \sum_{k=0}^{d_j} a^j_{p+d_j+k}(t-t_{n+1})^k(t-t_n)^{p+d_j} \tag{5.2}$$

where

$$a^j_{p+d_j+k} = h_n^{-(p+d_j+k)} \sum_{\ell=0}^{k} c^j_{k,\ell} h_n^\ell u^{n+1}_{j,\ell}$$

for all $1\le j\le m$ and $0\le k\le d_j$ and $\hat{x}_j(t)$ is the truncated Taylor solution (3.1). The coefficients $c_{k,\ell}$ are some constants depending only on p and the offsets.

Let $R_j(t) = \chi_j(t) - \hat{x}_j(t)$, $1 \leq j \leq m$, be the difference between the interpolant and the truncated Taylor solution (3.1). We may write

$$R_j(t) = \sum_{k=0}^{d_j} q_{j,k}(\tau) h_n^k u_{j,k}^{n+1} \tag{5.3}$$

where $\tau = \frac{1}{h_n}(t - t_n)$, $\tau \in [0, 1]$ and the polynomials $q_{j,k}$ have coefficients depending only on p and the offsets (by convention, the polynomials corresponding to those variables which are not projected are identically zero). In fact, we can show that

$$\|R_{J_0}(t)\| \leq \alpha \sum_{k=k_c}^{0} h_n^k \|u_{J_k}^{n+1}\| \tag{5.4}$$

where we can choose, for example,

$$\alpha = \max |q_{j,k}^{(d_j)}(\tau)|$$

and the maximum is taken over all $\tau \in [0, 1]$, all $1 \leq j \leq m,\ k \leq d_j$ and depends only on p and the offsets.

By using equations (3.5)–(3.6) and Proposition 4.1, we derive that

$$\|R_{J_0}(t)\| \leq \alpha \sum_{k=k_c}^{0} h_n^k \|\mathbf{J}_k^{\dagger}(\xi_k^{n+1}) \cdot \hat{\delta}_{I_k}(t_{n+1}^-)\|) + \mathcal{O}(h_n^{p+1}) \tag{5.5}$$

for each $t \in [t_n, t_{n+1}]$ and is of order $\mathcal{O}(h_n^p)$. The residuals $\hat{\delta}_{I_k}$, given by (3.2), correspond to the Taylor truncated solution (3.1).

REMARK 5.1. We note that the bounds (5.5) for the residual of interest for each step of integration *depend on derivatives of the residual functions* δ_i at the corresponding mesh points.

5.1 Residual control.

We are interested in measuring the magnitude of the residuals associated with the continuous approximation (5.2). Using Proposition 4.1, we obtain that at each stage $k \leq 0$ on the interval $[t_n, t_{n+1}]$, the corresponding residual is of order $\mathcal{O}(h_n^{p-k})$, thus the dominating residual is obtained at stage $k = 0$ and is of order $\mathcal{O}(h_n^p)$. This residual is given by

$$\delta_{I_0}(t) = f_{I_0}(\chi_{J_{\leq 0}}(t)) \,. \tag{5.6}$$

By using (5.3) and a first-order Taylor expansion of (5.6) around the predicted solution (3.1), we derive

$$\delta_{I_0}(t) = \hat{\delta}_{I_0}(t) + \mathbf{J}_0(\hat{x}_{J_{\leq 0}}(t)) \cdot R_{J_0}(t) + \mathcal{O}(h_n^{p+1}) \,. \tag{5.7}$$

on the interval $[t_n, t_{n+1})$ where the truncated Taylor approximations $\hat{x}(t)$ have the desired smoothness.

Therefore, by using (5.4), (5.7) and Proposition 4.1, we find that the principal error term is of the form $\rho_n h_n^p$ where

$$\max_{t\in[t_n,t_{n+1})} \|\delta_{I_0}(t)\| = \rho_n h_n^p + \mathcal{O}(h_n^{p+1}) \, .$$

The principal error coefficient is bounded as follows: $\rho_n \leq (\alpha+1)\psi_n$ with

$$\psi_n h_n^p = \|\mathbf{J}_0(\hat{x}_{J_{\leq 0}}^{n+1})\| \sum_{k=0}^{k_c} h_n^k \|\mathbf{J}_k^{\dagger}(\xi_k^{n+1})\hat{\delta}_{I_k}(t_{n+1}^-)\| \, . \tag{5.8}$$

It is sufficient to require that $\psi_n h_n^p \leq \dfrac{1}{\alpha+1}\varepsilon$ to guarantee that the residual is below the tolerance (higher order terms are neglected).

REMARK 5.2. The maximum $\alpha = \alpha(p)$ depends on p and there exists $\mathcal{K} > 0$ so that $[\alpha(p)+1]$ is $\mathcal{O}(p^{\mathcal{K}})$ ($\mathcal{K}$ may be chosen, e.g., the highest among all degrees of the polynomials $q_{j,k}^{(d_j)}$ with $k = 0, \ldots, d_j$).

EXAMPLE 5.1. We return now to the nonlinear simple pendulum problem governed by the equations (4.8). We want to construct the corresponding interpolants, as described above, and estimate the residual.

The Hermite interpolants corresponding to (x, y, z) and the step n of integration are respectively:

$$\left\{ \begin{array}{rcl} \chi_1(t) &=& \hat{x}(t) + q_{1,0}(\tau)u_{1,0}^{n+1} + h_n q_{1,1}(\tau)u_{1,1}^{n+1} + h_n^2 q_{1,2}(\tau)u_{1,2}^{n+1} \\ \chi_2(t) &=& \hat{y}(t) + q_{2,0}(\tau)u_{2,0}^{n+1} + h_n q_{2,1}(\tau)u_{2,1}^{n+1} + h_n^2 q_{2,2}(\tau)u_{2,2}^{n+1} \\ \chi_3(t) &=& \hat{\lambda}(t) + q_{3,0}(\tau)u_{3,0}^{n+1} \end{array} \right. .$$

Here,

$$\left\{ \begin{array}{rcl} q_{1,0}(\tau) = q_{2,0}(\tau) &=& \tau^{p+2}[1-(p+2)(\tau-1)+\frac{1}{2}(p+2)(p+3)(\tau-1)^2] \\ q_{1,1}(\tau) = q_{2,1}(\tau) &=& \tau^{p+2}(\tau-1)[1-(p+2)(\tau-1)] \\ q_{1,2}(\tau) = q_{2,2}(\tau) &=& \frac{1}{2}\tau^{p+2}(\tau-1)^2 \\ q_{3,0}(\tau) &=& \tau^p \, . \end{array} \right.$$

and

$$\left\{ \begin{array}{ll} u_{1,j}^{n+1} & = x_j^{n+1} - \hat{x}_j^{n+1} \\ u_{2,j}^{n+1} & = y_j^{n+1} - \hat{y}_j^{n+1} \\ u_{3,\ell}^{n+1} & = \lambda_j^{n+1} - \hat{\lambda}_j^{n+1} \end{array} \right.$$

for $j = 0, 1, 2$ and $\ell = 0$.

Using (5.3), we derive the errors in the interpolation at stage $k = 0$:

$$R_{J_0}(t) = \left(\begin{array}{c} \frac{1}{h_n^2} q_{1,0}^{(2)}(\tau)u_{1,0}^{n+1} + \frac{1}{h_n} q_{1,1}^{(2)}(\tau)u_{1,1}^{n+1} + q_{1,2}^{(2)}(\tau)u_{1,2}^{n+1} \\ \frac{1}{h_n^2} q_{2,0}^{(2)}(\tau)u_{2,0}^{n+1} + \frac{1}{h_n} q_{2,1}^{(2)}(\tau)u_{2,1}^{n+1} + q_{2,2}^{(2)}(\tau)u_{2,2}^{n+1} \\ q_{3,0}(\tau)u_{3,0}^{n+1} \end{array} \right)$$

where

$$\left\{ \begin{array}{rcl} q_{1,0}^{(2)}(\tau) &=& q_{2,0}^{(2)}(\tau) = \frac{1}{2}(p+2)(p+3)(p+4)\tau^p(\tau-1)[(p+3)(\tau-1)+2] \\ q_{1,1}^{(2)}(\tau) &=& q_{2,1}^{(2)}(\tau) = -(p+2)(p+3)(\tau-1)[(p+4)(\tau-1)+3] \\ q_{1,2}^{(2)}(\tau) &=& q_{2,2}^{(2)}(\tau) = \frac{1}{2}(p+3)(p+4)(\tau-1)^2 + 2(p+3)(\tau-1) + 1 \, . \end{array} \right.$$

Thus, we derive $\alpha(p) = \mathcal{O}(p^4)$.

6 Polynomial cost of Pryce's algorithm.

We show below that the cost of computing an approximate solution of the DAE (2.1) which satisfies the tolerance with the algorithm under investigation is polynomial in the number of digits of accuracy requested. Moreover, provided the local error coefficients given by (5.8) satisfy some regularity conditions described below, the cost of computing the solution with the same accuracy is minimized on the equidistributed mesh. In order to obtain this, we make use of a result on equidistribution from [10], which is given below.

For a vector $\Psi = [\psi_1, \cdots, \psi_N]$ we define the *s-norm* as

$$\|\Psi\|_s = \left(\sum_{i=1}^{N} \psi_i^s \right)^{1/s}$$

and the *Hölder s-mean* as

$$\mathcal{M}_s(\Psi) = \left(\frac{1}{N} \sum_{i=1}^{N} \psi_i^s \right)^{1/s}.$$

The s-norm is not a true norm for $s < 1$, when the triangle inequality is not always satisfied. We shall not need this property.

We assume that the vector of the error coefficients $\Psi_N = [\psi_1, \cdots, \psi_N]$ satisfies some regularity conditions. First, we assume that there exists a positive integer N_0 such that $\mathcal{M}_s(\Psi_{N_2}) \leq \mathcal{M}_s(\Psi_{N_1})$ for $N_2 > N_1 \geq N_0$. This means that if the mesh is sufficiently fine, the Hölder mean does not increase with further refinement. Second, we assume that the same property holds for the maximum norm of the local error coefficients, $\|\Phi_N\|_\infty$.

In fact, we assume that the practical implementation satisfies the regularity conditions, that is the estimates are not fooled. Theoretically, these realistic assumptions regarding the regularity conditions may be lifted by using a continuous representation of the local residual. The meshes are generated by introducing a grid deformation map which maps a uniform mesh in a new, auxiliary, independent variable into a non-uniform mesh in the initial independent variable. The discrete representations of both the Hölder s-mean and the infinity norm of the principal error coefficients can then be interpreted in terms of the corresponding integral representations (across the interval $[a, b]$) of the principal error function. This continuous approach was introduced in [23] for ODE, and for DAE is subject of our future work.

We evaluate the cost of an algorithm by counting the number of arithmetic operations, which is a more appropriate measure of efficiency for series methods than counting number of function evaluations [11]. Following the standard theory of computational complexity [36], we ignore memory hierarchy, overheads and interpolation costs.

The cost per step for the method investigated in this paper, for a fixed order, is the same for all steps and consists of the cost of obtaining the Taylor series plus the cost of $[1 - k_c]$ Newton projections. Thus, the cost of the method for fixed order is directly proportional to the number of steps taken.

A mesh is optimal if it allows us to take the minimum number of steps to compute an approximate solution on the interval $[a, b]$, satisfying the tolerance.

THEOREM 6.1 (MINIMAX). *Given $p, N \in \mathbb{N}$ and a vector with positive coefficients*

$[\psi_i]_{1\le i\le N}$, *the following inequality is true*

$$\max\{\psi_i h_i^p : \sum_1^N h_i = b-a\} \ge (b-a)^p \|\Psi\|_{-1/p} = \bar{h}^p \mathcal{M}_{-1/p}(\Psi)$$

where $\bar{h} = (b-a)/N$ *is the average stepsize. Equality holds iff*

$$\psi_i h_i^p = \bar{h}^p \mathcal{M}_{-1/p}(\Psi) \text{ for all } 1 \le i \le N.$$

A proof of this theorem based on Hölder's inequality is available in [10], and one by linearization in [22].

A first consequence of the Minimax Theorem (and the regularity condition on the Hölder mean of the local error coefficients) is that the equidistributing mesh ($\psi_i h_i^p = \bar{h}^p \mathcal{M}_{-1/p}(\Psi) = \varepsilon$ for all $1 \le i \le N$) is an optimal mesh for the algorithm. It is straightforward to derive that the number of steps corresponding to the equidistributing mesh is

$$N = \frac{b-a}{\bar{h}} = (b-a)\cdot(\mathcal{M}_{-1/p}(\Psi_N))^{1/p}(\frac{\varepsilon}{\alpha+1})^{-1/p}. \tag{6.1}$$

A second consequence of the Minimax Theorem (and regularity conditions on the Hölder mean and infinity norm) is that the cost of computing an approximate solution satisfying a given tolerance is never less on a fixed step mesh than on an equidistributed mesh for the above algorithm. In fact, the ratio of the two costs (the ratio of the number of steps corresponding to each mesh) is given by $\|\Psi_{fix}\|_\infty/(\mathcal{M}_{-1/p}(\Psi))^{1/p}$.

Proofs of the above results are similar to those in [10].

THEOREM 6.2. *The minimal cost of computing the solution of (2.1) with Pryce's method is polynomial in the number of digits of accuracy requested and is bounded above by*

$$\mathcal{C}(b-a)e^{2+(\mathcal{K}/e)}(\mathcal{M}_{-1/p}(\Psi_N))^{1/p}\mathcal{B}^4$$

where $\mathcal{C}, \mathcal{K}$ *are some constants. The minimum cost is reached on the equidistributing mesh.*

PROOF. Order-p accurate solutions computed with Taylor series using automatic differentiation can be obtained in $\mathcal{O}(p^2)$ operations (see [11]) and, if naive multiplication is used, the cost of doing arithmetics with $\mathcal{B}$ bits of accuracy is $\mathcal{O}(\mathcal{B}^2)$. Thus the cost of obtaining the Taylor series at one step costs $cp^2\mathcal{B}^2$. The constant c depends on the dimension of the problem.

The cost of $[1-k_c]$ Newton projection is independent of p and costs $\mathcal{O}(\mathcal{B}^2)$. Therefore the cost of one step is $C_1 p^2 \mathcal{B}^2$.

Using (6.1), we obtain the total cost corresponding to the equidistributed mesh:

$$C_1(b-a)(\mathcal{M}_{-1/p}(\Psi_N))^{1/p}\mathcal{B}^2p^2(\frac{\varepsilon}{\alpha+1})^{-1/p}\ . \tag{6.2}$$

Since there exist positive constants C_2 and $\mathcal{K}$ so that $\alpha+1 \le C_2 p^{\mathcal{K}}$, then

$$(\alpha+1)^{1/p} \le C_2 e^{\mathcal{K}/e}\ .$$

Choosing $p = \lceil (\mathcal{B} \ln 2)/2 \rceil$ in (6.2) and $\mathcal{C} = C_1 C_2 (\ln 2)^2/4$, we obtain an upper bound of the minimum cost given by

$$\mathcal{C}(b-a) e^{2+(\mathcal{K}/e)} (\mathcal{M}_{-1/p}(\Psi_N))^{1/p} \mathcal{B}^4 \ .$$

□

REMARK 6.1. The cost of computing the solution on a fixed-step mesh is also polynomial in the number of digits of accuracy. However, we note that the dimension of the problem is hidden in the constant factor in the expression of the cost.

We also remark that *very high accuracies* (of order $\mathbf{10^{-128}}$) have been very recently reported for solving DAEs with Pryce's method [3, 4]. Currently, existing software using this method may solve differential algebraic systems of hundred equations [27]. Detailed description on implementation and results of such a solver are given in [28]. These results support our theory which predicts polynomial cost in the number of bits of accuracy.

We mention that an interesting open problem is to extend this complexity analysis to the numerical solution of more general systems of partial differential algebraic equations for which Pryce's structural analysis can be applied [37, 38].

7 Conclusions.

In this paper, we have investigated the cost of solving initial value problems for high-index differential algebraic equations depending on the number of digits of accuracy requested.

We analyzed an algorithm based on a Taylor series method developed by Pryce for a general class of differential algebraic equations. We showed that the cost of computing a solution with this algorithm is polynomial in the number of digits of accuracy. We also showed that adaptation performs better than non-adaptation and we included a sufficiently smooth dense output.

The cost of locating singularities and rank changes will be addressed in a future work.

Acknowledgement.

We thank Stephen Campbell, George Corliss, Wayne Enright, Ned Nedialkov and John Pryce for helpful discussions on numerical DAE solving and Greg Reid for insight on geometrical PDAE solving. We would like to thank the reviewers for their detailed comments which helped improve the paper.

REFERENCES

1. Allgower, E.L., Georg, K., *Introduction to numerical continuation methods*, SIAM Classics in Applied Mathematics, Philadelphia, 2003.
2. Ascher, U., Petzold, L., *Computer methods for ODE and DAE*, SIAM, Philadelphia, 1998.
3. Barrio, R., Performance of the Taylor series methods for ODEs/DAEs, *Applied Mathematics and Computation*, **163**(2) (2005) 525 – 545.
4. Barrio, R., Blesa, F., Lara, M., VSVO formulation of the Taylor method for the numerical solution of ODEs, *Computers & Mathematics with Applications*, **50**(1-2) (2005) 93 – 111.

5. Ben-Israel, A., Greville, T.N.E., *Generalized inverses, theory and applications*, Wiley-Interscience, New York, 1974.
6. Butcher, J.C., *The numerical analysis of ordinary differential equations*, John Wiley & Sons Ltd., Chichester, 1987.
7. Cartan, É., *Les systémes différentiels extérieurs et leur applications géométriques*, Hermann, 1945.
8. Corless, R.M., Error backward, *Chaotic Numerics*, *AMS Contemporary Mathematics Series* **172** (1994) 31 – 62.
9. Corless, R.M., An elementary solution of a minimax problem arising in algorithms for automatic mesh selection, *SIGSAM Bulletin: Commun. Comput. Algebra* **34**(4) (2001) 7 – 15.
10. Corless, R.M., A new view of the computational complexity of IVP for ODE, *Numerical Algorithms* **31** (2002) 115 – 124.
11. Corliss, G., Chang, Y.F., Solving ordinary differential equations using Taylor series, *ACM Tansactions on Mathematical Software* **8**(2) (1982) 114 – 144.
12. Douglis, A., Nirenberg, L., Interior estimates for elliptic systems of partial differential equations, *Comm. Pure Appl. Math.*, **8** (1955), 503 – 538.
13. Enright, W.H., A new error control for initial value solvers, *Appl. Math. Comput.* **31** (1989) 288 – 301.
14. Enright, W.H., Jackson, K.R., Norsett, S.P., Thomsen, P.G., Interpolants for Runge-Kutta formulas, *ACM Transactions on Mathematical Software* **12**(3) (1986) 193 – 218.
15. Gladwell, I., Shampine, L.F., Baca, L.S., Brankin, R.W., Practical aspects of interpolation in Runge-Kutta codes, *SIAM J. Sci. Stat. Comput* **8**(3) (1987) 322 – 341.
16. Griewank, A., *Evaluating Derivatives: Principles and Techniques of Algorithmic Differentiation*, Frontiers in Applied Mathematics, SIAM, Philadelphia, PA, 2000.
17. Hairer, E., Nørsett, S.P., Wanner, G., *Solving Ordinary Differential Equations I*, Computational Mathematics, Vol. 8, Springer, Berlin, 1987.
18. Hairer, E., Wanner, G., *Solving Ordinary Differential Equations II*, Computational Mathematics, Vol. 14, Springer, Berlin, 1991.
19. Hoefkens, J., Berz, M., and Makino, K., Efficient high-order methods for ODEs and DAEs, in: G.F. Corliss et al., eds., *Automatic Differentiation: From Simulation to Optimization*, Springer-Verlag, New York, 2001, 343 – 350.
20. van der Hoeven, J., Fast evaluation of holonomic functions, *Theoret. Comput. Sci.* **210** (1999) 199 – 215.
21. Ilie, S., Computational complexity of numerical solutions of initial value problems for differential algebraic equations, PhD thesis, University of Western Ontario, 2005.
22. Ilie, S., Corless, R.M., Reid, G., Numerical solutions of index-1 differential algebraic equations can be computed in polynomial time, *Numerical Algorithms*, **41**(2) (2006) 161 –171.
23. Ilie, S., Söderlind, G., Corless, R.M., Adaptivity and computational complexity in the numerical solution of ODEs, *Journal of Complexity*, to appear.

24. Jackson, K.R., Nedialkov, N., Some recent advances in validated methods for IVPs for ODEs, *Appl. Numer. Math.* **42**(1) (2002) 269 – 284.
25. Janet, M., *Leçons sur les systèmes d'équations aux dérivées partielles*, Gauthier-Villars, 1929.
26. Moore, R.E., *Interval Analysis*, Prentice Hall, Englewood Cliffs, New York, 1966.
27. Nedialkov, N.S., *private communication.*
28. Nedialkov, N.S., Pryce, J.D. , Solving differential-algebraic equations by Taylor series (I): computing Taylor coefficients, *BIT*, **45**(3) (2005) 561 – 591.
29. Pantelides, C., The consistent initialization of differential-algebraic systems, *SIAM J. Sci. Stat. Comput.* **9**(2) (1988) 213 – 231.
30. Pryce, J.D., Solving high-index DAEs by Taylor series, *Numerical Algorithms* **19** (1998) 195 – 211.
31. Pryce, J.D., A simple structural analysis method for DAEs, *BIT* **41**(2) (2001) 364 – 394.
32. Pryce, J.D., *private communication.*
33. Rall, L.B., *Automatic Differentiation: Techniques and Applications*, Springer-Verlag, Berlin, 1981.
34. Riquier, C., *Les systémes d'équations aux dérivées partielles*, Gautier-Villars, 1910.
35. Söderlind, G., Automatic control and adaptive time-stepping, *Numerical Algorithms* **31** (2002) 281 – 310.
36. Werschulz, A.G., *The Computational Complexity of Differential and Integral Equations*, Oxford Science, Oxford, 1991.
37. Wu, W., Reid, G., Symbolic-numeric computation of implicit Riquier bases for PDE, Proc. of ISSAC 2007, *ACM* (2007) 377 – 386.
38. Wu, W., Reid, G., Ilie, S., Implicit Riquier bases for PDAE and their discretizations, *submitted*.