Numerical Solutions of Index-1 Differential Algebraic Equations can be Computed in Polynomial Time

SILVANA ILIE* ROBERT M. CORLESS† GREG REID

Ontario Research Centre for Computer Algebra and
Department of Applied Mathematics, University of Western Ontario
London, ON, N6A 5B7, Canada
e-mails: silvana|rcorless|reid@uwo.ca

Abstract

The cost of solving an initial value problem for index-1 differential algebraic equations to accuracy ε is polynomial in $\ln(1/\varepsilon)$. This cost is obtained for an algorithm based on the Taylor series method for solving differential algebraic equations developed by Pryce. This result extends a recent result by Corless for solutions of ordinary differential equations. The results of the standard theory of information-based complexity give exponential cost for solving ordinary differential equations, being based on a different model.

AMS subject classification: 34A09, 65L80, 68Q25

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

1 Introduction

We show here that the cost to solve initial value problems for semi-explicit index-1 DAE is polynomial in the number of bits of accuracy. This result extends that of [5] which showed the cost of solving IVP for ODE is polynomial in the number of bits of accuracy. This result contradicted earlier results that modeled the numerical

*Research partially supported by an Ontario Graduate Scholarship.

†Research supported by a grant from the Natural Sciences and Engineering Research Council of Canada.

solution of IVP for ODE in such a way as to give exponential complexity [25]; of course that model is not realistic.

We use the residual error, or 'defect', because it makes the analysis simpler. This is equivalent to local error methods (see e.g. [22]) but is also easier to explain to users and is quite practical to implement, given that continuous extensions are needed also for other purposes [9]. The residual error is also, in the IVP for ODE case, directly connected to the global (forward) error by the Gröbner-Alekseev nonlinear variation-of-constants formula (see e.g. [13]). This separates the stability of the method (producing a small residual) from the conditioning of the problem [7, 9]. The Gröbner-Alekseev formula has been extended to IVP for certain classes of DAE [10, 19].

IVP for DAE are harder to solve in practice than IVP for ODE; but the results of this paper show that the difficulty is essentially a constant factor, and not a different exponent in the theoretical complexity.

The analysis assumes that the functions defining the DAE are piecewise analytic. This is the key assumption, allowing the use of arbitrary order methods (in fact the chosen order $p = \lceil (1/2)\ln(1/\varepsilon) \rceil$ depends on the tolerance ε for the residual). In practice, events and singularities of course occur; see e.g. [16]. We assume that singularities can be located accurately at negligible cost: we also assume that events are $\mathcal{O}(1)$ apart as $\varepsilon \to 0$ and that accurate location adds at most an extra factor of $\ln\ln(1/\varepsilon)$ to the cost. This assumption will be given detailed justification in a forthcoming paper [11].

We also use the fact that order-p accurate solutions may be computed on an interval of width $h \ll 1$ in $\mathcal{O}(p^2)$ operations using $\mathcal{O}(\ln(1/\varepsilon))$ bits and therefore at cost $\mathcal{O}(p^2 \left[\ln(1/\varepsilon)\right]^2)$. This can be done for IVP, for example, by Taylor series methods (see e.g. [8, 4, 13]) or Hermite-Obreshkov methods [4, 12] and recently it has been shown to be possible for DAE [17].

The analysis also relies on certain regularity assumptions that ensure that the error estimates are not fooled; without those assumptions, the problem is in fact undecidable [15].

2 The problem

Consider the following semi-explicit index-1 DAE:

$$\begin{aligned} y_e'(t) &= f(y_e(t), z_e(t)) \\ 0 &= g(y_e(t), z_e(t)) \end{aligned} \tag{1}$$

for $t \in I = (a, b)$, $y_e : I \to \mathbb{R}^k$ and $z_e : I \to \mathbb{R}^m$. We assume that the Jacobian $g_z(y_e, z_e)$ is invertible. The initial condition, given by $y_e(a) = y_0,\ z_e(a) = z_0$, is

assumed consistent with the constraints, so $g(y_0, z_0) = 0$.

We assume that f and g are analytic functions and that the solution $(y_e(t), z_e(t))$ exists and is unique on the interval of integration. We also assume that if we augment the DAE with extra ODEs that describe standard functions used in f and g, then the problem is converted to a larger DAE in which the new functions f and g involve only the four basic arithmetic operations. Thus f and g in (1) belong to the class of functions to which automatic differentiation applies. This includes most functions of practical interest, but excludes, for example, the Γ function.

Other assumptions (detailed later) ensure that the minimum stepsize is bounded away from zero, and thus the integration does not 'grind to a halt'.

We are interested in computing a numerical solution with a tolerance ε for both the residual error of the differential equation $\|\delta_1(t)\|$ and the residual error of the algebraic equation $\|\delta_2(t)\|$, with

$$\begin{array}{rcl} \delta_1(t) & = & y'(t) - f(y(t), z(t)) \\ \delta_2(t) & = & g(y(t), z(t)) \; . \end{array} \tag{2}$$

By controlling the magnitude of the residual we also control the forward error. A relationship between errors can be obtained either by using the Gröbner-Alekseev formula or by using Gronwall's lemma for the underlying ODE [14]. We review below the second approach.

If the functions f and g are analytic in a neighborhood of the exact solution path then there exist some positive constants L_1 and L_2 so that

$$\|f(y, z) - f(y_e, z_e)\| \leq L_1 \|y - y_e\| + L_2 \|z - z_e\| \; .$$

We can choose $L_1 = \sup \|f_y\|$ and $L_2 = \sup \|f_z\|$. Since g_z is invertible, the implicit function theorem will give

$$\|z(t) - z_e(t)\| \leq \ell_1 \|y(t) - y_e(t)\| + \ell_2 \|\delta_2(t)\|$$

where $\ell_1 = \sup \|g_z^{-1} g_y\|$, $\ell_2 = \sup \|g_z^{-1}\|$ and the suprema exist in an appropriate neighborhood of the exact solution. By applying Gronwall's lemma [13] on the compact interval $[a, b]$, we obtain that there exists a positive constant K so that

$$\|y(t) - y_e(t)\| \leq K(\|y(a) - y_0\| + \max_{a \leq s \leq t} \|\delta_2(s)\| + \max_{a \leq s \leq t} \|\delta_1(s)\|) \; . \tag{3}$$

If the functions f and g are piecewise analytic, we assume that, in a neighborhood of the solution, f is Lipschitz continuous in its arguments and g_y, g_z^{-1} are bounded in norm on each region where the functions have one analytic formula. We also assume that the boundaries of these regions are continuous and piecewise analytic. Under these assumptions a similar result to (3) holds.

In this paper, we consider the Taylor series method developed by Pryce in [20] and [21], and we analyse it for the fixed order case.

3 Numerical solution

We shall analyse the cost of the following algorithm: assume we have obtained at time t_n the values (y_n, z_n) which satisfy the algebraic constraints more accurately than the desired tolerance. After generating the Taylor coefficients for the unknowns, we predict the values $(y_{n+1}, \hat{z}_{n+1})$ by computing the Taylor series with a chosen stepsize h_n. Then we correct $\hat{z}_{n+1}$ by applying one Newton iteration for the algebraic variables which ensures that the algebraic constraints are satisfied more accurately. If y_{n+1} and the new value for z, namely z_{n+1}, satisfy the differential equation with the residual below the tolerance ε then the step is accepted.

3.1 Solution by Taylor series

The method proposed by Pryce in [20] and [21] for solving a general class of DAEs consists of generating the Taylor coefficients for the differential and algebraic equations and equating them to zero to solve for the Taylor coefficients of the variables y and z.

Pryce's method starts with a pre-processing stage which reveals a certain structure of the problem. The structure may be used for analysing the DAE: the structural index and the degrees of freedom for the solution manifold can be computed based on it. Also, once the structure is known, the system of the Taylor coefficients can be solved automatically.

The first step of the analysis consists in determining a set of integers, called *offsets* of the problem, that indicate which equations to solve for which unknowns. The next step is to generate the system Jacobian. If at each integration step, the system Jacobian is non-singular, the method succeeds, and the Taylor coefficients can be determined up to the desired order.

A critical observation is that at each integration step the current Jacobian J, once computed, generates after a few initial stages the Taylor coefficients for the unknowns iteratively, as solutions of some linear systems which have the same matrix J.

For the particular case of semi-explicit index-1 DAE, the offsets of the problem and the system Jacobian can be easily determined. The offsets are given by the following vectors with $k + m$ components: $d = (1, \cdots 1, 0, \cdots, 0)$ (with the first k components 1) and $c = (0, \cdots, 0)$.

The system Jacobian is given by:

$$J \quad = \quad \begin{bmatrix} I & -f_z \\ 0 & g_z \end{bmatrix}$$

and it is non-singular in a neighborhood of the solution since the DAE (1) is of index 1. This implies that the Taylor series method is successful for our problem.

Assume that the Taylor series for the equations and for the unknowns calculated at step n of integration and at time $t = t_n + h$ are, respectively,

$$\begin{aligned} y(t) &= \sum_{j\geq 0} y_{n,j}(t-t_n)^j/j! \\ z(t) &= \sum_{j\geq 0} z_{n,j}(t-t_n)^j/j! \\ F(t) &= y'(t) - f(y(t),z(t)) = \sum_{j\geq 0} F_{n,j}(t-t_n)^j/j! \\ G(t) &= g(y(t),z(t)) = \sum_{j\geq 0} G_{n,j}(t-t_n)^j/j! \ . \end{aligned}$$

According to Pryce's algorithm we have to solve, for each stage $j = 0, 1, \cdots, p-1$, the systems $(F_{n,j} = 0, G_{n,j} = 0)$ in the unknowns $(y_{n,j+1}, z_{n,j})$.

The value $y_{n,0}$ is given either from the initial condition for the first integration step or from the previous step for the other steps.

The stage $j = 0$ is atypical. The equations $(F_{n,0} = 0,\ G_{n,0} = 0)$ may be nonlinear in the corresponding unknown $z_{n,0}$. The value $z_{n,0}$ should be also given from the initial condition at the beginning or from the value at the previous step. Since one Newton correction is applied at the end of the previous step, we shall see that the condition $g(y_{n,0}, z_{n,0}) \approx 0$ will be satisfied accurately enough so that its residual does not essentially perturb the algebraic residual for the current integration step. The stage is completed by taking $y_{n,1} = f(y_{n,0}, z_{n,0})$.

The stages $j \geq 1$ are all linear in the corresponding unknowns and involve the same Jacobian $J(y_{n,0}, z_{n,0})$.

The approximate solution, given by

$$\begin{aligned} \hat{y}(t) &= \sum_{j=0}^{p} y_{n,j}(t-t_n)^j/j! \\ \hat{z}(t) &= \sum_{j=0}^{p-1} z_{n,j}(t-t_n)^j/j! \ , \end{aligned}$$

satisfies the problem

$$\begin{aligned} \hat{y}'(t) &= f(\hat{y}(t),\hat{z}(t)) + \hat{\delta}_1(t) \\ 0 &= g(\hat{y}(t),\hat{z}(t)) + \hat{\delta}_2(t) \ . \end{aligned} \tag{4}$$

The predicted values at time $t_{n+1} = t_n + h_n$ are $y_{n+1} = \hat{y}(t_{n+1})$ and $\hat{z}_{n+1} = \hat{z}(t_{n+1})$. As proved in the next section, the algebraic residual after the Newton

correction at the step $(n-1)$ (which equals the algebraic residual for the stage $j=0$ at the nth integration step) is $\mathcal{O}(h_{n-1}^{2p})$. This error is assumed negligible with respect to h_n^p (a reasonable approximation for small enough tolerance) so the residuals are $\hat{\delta}_1 = \phi_{1,n} h_n^p$ and $\hat{\delta}_2 = \phi_{2,n} h_n^p$.

An asymptotic evaluation for small h_n gives

$$\phi_{1,n} = \frac{1}{p!}\frac{d^p}{dt^p} f(y,z)(t_n) + \mathcal{O}(h_n) \; .$$

Thus the local error coefficients can be represented asymptotically as the value at (y_n, z_n) of some fixed function involving multiplication, addition or subtraction of f, g, their derivatives with respect to y and z up to order p and g_z^{-1}, depending only on p. A similar expression can be obtained for $\phi_{2,n}$.

3.2 Newton projection

We wish to satisfy the algebraic equations more accurately than the differential equations, but we do not want to increase the cost unnecessarily. It has been observed in practice [18] that after one Newton iteration the algebraic constraints become sufficiently accurate and this agreement improves as the tolerance $\varepsilon \to 0$.

Since g is an analytic function in its variables and g_z is non-singular for the exact solution then g_z is non-singular in a neighbouhood of the solution, and we can consider a simplified Newton iteration. If the predictor step is given by $(y_n, \hat{z}_n)$, we keep y_n constant and we correct only the algebraic variable, by considering a new value, z_n, which satisfies

$$z_n = \hat{z}_n - g_z^{-1}(y_n, \hat{z}_n) g(y_n, \hat{z}_n) \; . \tag{5}$$

From (5) we obtain that

$$\|z_n - \hat{z}_n\| \leq L \|g(y_n, \hat{z}_n)\| \tag{6}$$

where $L = \sup_U \|g_z^{-1}\|$ and U is a convex open neighborhood of the exact solution on which g_z is invertible and g_z^{-1} and g_{zz} are bounded.

Consider $H(s) = g(y_n, \hat{z}_n + s(z_n - \hat{z}_n))$. By integrating the function $H''(s)(1-s)$ with respect to s on the interval $[0,1]$, we derive the following Taylor expansion:

$$g(y_n, z_n) = g(y_n, \hat{z}_n) + g_z(y_n, \hat{z}_n)(z_n - \hat{z}_n) + \frac{1}{2}(z_n - \hat{z}_n)^T G(y_n, z_n, \hat{z}_n)(z_n - \hat{z}_n)$$

where

$$G(y_n, z_n, \hat{z}_n) = \int_0^1 2(1-s) g_{zz}(y_n, \hat{z}_n + s(z_n - \hat{z}_n)) ds \; .$$

The first two terms in the Taylor expansion cancel each other due to (5), and by using (6) we obtain

$$\|g(y_n, z_n)\| \leq \frac{1}{2} M L^2 \|g(y_n, \hat{z}_n)\|^2 \ , \tag{7}$$

where $M = \sup_U \|g_{zz}\|$.

We conclude the analysis of the Newton projection by observing that if the tolerance ε is sufficiently small then the residual is small enough to guarantee that $(y_n, \hat{z}_n)$ is in the basin of attraction of the Newton method.

3.3 Error Analysis

We give below estimates for the residual errors (2). The residual in the algebraic equation at time t_{n+1}, after the corrector step, is given by $\|\delta_2\| = \|g(y_{n+1}, z_{n+1})\|$. Using (4) and (7), we obtain

$$\|\delta_2\| \leq \frac{1}{2} M L^2 \|\phi_{1,n}\|^2 h_n^{2p} \ .$$

The residual in the differential equation is

$$\begin{aligned} \|\delta_1\| &= \|\hat{y}'(t_{n+1}) - f(y_{n+1}, z_{n+1})\| \\ &= \|\hat{\delta}_1 + f(y_{n+1}, \hat{z}_{n+1}) - f(y_{n+1}, \hat{z}_{n+1} + g_z^{-1}(y_{n+1}, \hat{z}_{n+1})\hat{\delta}_2)\| \\ &= \|\phi_{1,n} - f_z(y_{n+1}, \zeta_{n+1}) g_z^{-1}(y_{n+1}, \hat{z}_{n+1})\phi_{2,n}\| h_n^p \\ &= \psi_n h_n^p \end{aligned}$$

where ζ_{n+1} is close to z_{n+1}. Since the residual $\|\delta_2\|$ at the end of the step is $\mathcal{O}(h_n^{2p})$, for small enough tolerance it is enough to control just the differential residual, which is $\mathcal{O}(h_n^p)$. We require that $\psi_n h_n^p \leq \varepsilon$ for all $n \leq N$.

4 The minimum cost of the algorithm

We claim that the minimum cost to find an approximate solution which satisfies the desired accuracy with the algorithm described above is achieved if the residual error is equidistributed. In order to see this, we need a result from [5], which we reproduce 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

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

Note that the s-norm is not, in fact, a true norm for $s < 0$, because then 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. The first assumption is 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. The second assumption is that the same property holds for the maximum norm of the local error coefficients, $\|\Phi_N\|_\infty$. What we are really assuming is that the implementation can do this in practice.

Minimax theorem 1 *Given $p, N \in \mathbb{N}$ and a vector with positive coefficients $[\psi_i]_{1 \leq i \leq N}$, the following inequality is true:*

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

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

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

We give a proof by linearization below. A proof based on Hölder's inequality can be found in [6].

Proof. The following inequality is valid for all $a_j > 0$ (see e.g. [3])

$$\max_{1 \leq j \leq N} b_j \geq \frac{\sum_{j=1}^{N} b_j a_j}{\sum_{j=1}^{N} a_j}$$

and the equality holds iff the b_j are all equal. By choosing $a_j = \psi_j^{-1/p}$ and $b_j = \psi_j^{1/p} h_j$, we derive

$$(\max_{1 \leq j \leq N} \psi_j h_j^p)^{1/p} = \max_{1 \leq j \leq N} \psi_j^{1/p} h_j \geq \frac{b-a}{\sum_{j=1}^{N} \psi_j^{-1/p}} \tag{8}$$

with equality iff $\psi_j h_j^p$ are all equal. We conclude by applying the power p in (8).

As in the standard theory of computational complexity [25], we ignore memory hierarchy, overheads and interpolation costs. We measure the cost of the algorithm by the number of arithmetic operations [8].

The cost per step for this method with a fixed order is the same for all steps and consists of the cost to obtain the Taylor series plus the cost of one Newton projection. The cost of computing a solution increases then with the number of steps to go from $t = a$ to $t = b$.

A mesh is called optimal if it takes the minimum number of steps to go from a to b while satisfying the tolerance.

By analogy with [5], we obtain:

1. The solution of the minimax theorem gives also the optimal mesh for the algorithm considered in this paper

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

   and the corresponding total number of steps

$$N = \frac{b-a}{\bar{h}} = (b-a) \cdot (\mathcal{M}_{-1/p}(\Psi_N))^{1/p} \varepsilon^{-1/p} \ . \tag{9}$$

2. Computing the solution of (1) on a fixed step size mesh with an error tolerance ε is more expensive than computing the solution with the same tolerance on an equidistributed mesh for the algorithm analysed. The cost for a fixed step size mesh is greater than the cost on an equidistributed mesh by a factor of $\|\Psi_{fix}\|_\infty / (\mathcal{M}_{-1/p}(\Psi_{ad}))^{1/p}$. This factor is asymptotic to 1 as $p \to \infty$, but for realistic values of p may be quite substantial.

Remark. Equidistribution ensures that the mean stepsize controls the integration (the minimum stepsize is related to the maximum ψ_n); assumed bounded here, this ensures that the integration does not grind to a halt.

Theorem 1 *The minimal cost of obtaining the solution with error ε of the IVP for (1) using Pryce's method is polynomial in the number of digits of accuracy. It is reached on the equidistributing mesh and is bounded above by*

$$C(b-a)\frac{e^2}{4} \cdot (\mathcal{M}_{-1/p}(\Psi_N))^{1/p} \Big[\ln(1/\varepsilon)\Big]^4. \tag{10}$$

If the Hölder p-mean of the local error coefficients depends weakly on the order p of the method, then the minimum is given by (10) and is reached for $p = \lceil (1/2)\ln(1/\varepsilon) \rceil$.

Proof. Computing the solution with Taylor series to order p and precision ε for one step takes $C_1 p^2 \left[\ln(1/\varepsilon)\right]^2$ operations. Indeed, the cost of generating and computing the Taylor series with p terms using automatic differentiation is $\mathcal{O}(p^2)$ (see [8]). Also, doing arithmetic with precision ε, that is with $\left[\ln(1/\varepsilon)\right]$ digits of accuracy, costs $\mathcal{O}(\left[\ln(1/\varepsilon)\right]^2)$ if the naive algorithm for multiplication is used. The constant C_1 depends on the dimension of the problem (k and m).

The cost of the Newton projection is independent of p and is $\mathcal{O}(\left[\ln(1/\varepsilon)\right]^2)$, with the constant depending on the dimension of the problem.

Thus, overall, the cost of one step is $Cp^2\left[\ln(1/\varepsilon)\right]^2$. By using (9), we obtain the total cost (the number of steps times the cost for one step) on the equidistributed mesh

$$C(b-a)\cdot(\mathcal{M}_{-1/p}(\Psi_N))^{1/p}\left[\ln(1/\varepsilon)\right]^2 p^2 \varepsilon^{-1/p}\ . \tag{11}$$

Under the hypothesis on the local error coefficients, we find that the total cost (11) reaches a minimum for the value $p = (1/2)\ln(1/\varepsilon)$, given by

$$C(b-a)\frac{e^2}{4}\cdot(\mathcal{M}_{-1/p}(\Psi_N))^{1/p}\left[\ln(1/\varepsilon)\right]^4$$

Since the Taylor series method allows a variable order, we can choose $p = \lceil (1/2)\ln(1/\varepsilon)\rceil$.

If the order is $p = \lceil (1/2)\ln(1/\varepsilon)\rceil$ and if the dependence of the Hölder mean of the local error coefficients on p is not weak, then (10) gives an upper bound for the minimal cost.

Remarks. The cost of computing the solution on a fixed step size mesh is also polynomial in number of digits of accuracy (it is obtained by replacing the Hölder mean norm by the infinity norm in (10)).

By using asymptotically faster multiplication (e.g. FFT methods, [2]) we may reduce the bound on the minimum cost to $\mathcal{O}(\ln[\ln(1/\varepsilon)]\left[\ln(1/\varepsilon)\right]^3)$, ignoring factors of $\ln[\ln(1/\varepsilon)]^2$ from event location.

5 Conclusions

We have obtained that the cost of approximating the solution of the semi-explicit index-1 DAE (1) with the Taylor series method developed by Pryce is a polynomial in the number of digits of accuracy. We have shown that nonadaption is more expensive than adaption depending on the ratio of the maximum norm of the vector of local error coefficients to their Hölder mean.

Future work will consider polynomial cost in the number of digits of accuracy for computing the solution for high index DAE.

We would like to thank the anonymous reviewers for their detailed comments which helped greatly to improve the paper. We would also like to thank Stephen Campbell, Wayne Enright, Ned Nedialkov and Larry Shampine for helpful discussions regarding aspects of DAE solving.

References

[1] E.L. Allgower, K. Georg, *Introduction to numerical continuation methods*, SIAM Classics in Applied Mathematics, Philadelphia, 2003.

[2] R.P. Brent, Fast multiple-precision evaluation of elementary functions, J. Assoc. Comput. Mach 23(2) (1976), 242-251.

[3] P.S. Bullen, D.S. Mitrinovic, P.M. Vasic, *Means and their inequalities*, Mathematics and its Applications, East European Series, D. Reidel Publishing Company, Dordrecht, 1988.

[4] J.C. Butcher, *The numerical analysis of ordinary differential equations*, John Wiley & Sons Ltd., Chichester, 1987.

[5] R.M. Corless, A new view of the computational complexity of IVP for ODE, Numerical Algorithms 31 (2002), 115-124.

[6] R.M. Corless, An elementary solution of a minimax problem arising in algorithms for automatic mesh selection, SIGSAM Bulletin: Commun. Comput. Algebra 34(4) (2001), 7-15.

[7] R.M. Corless, Error backward, Chaotic Numerics, AMS Contemporary Mathematics Series 172 (1994), 31-62.

[8] G. Corliss, Y.F. Chang, Solving ordinary differential equations using Taylor series, ACM Tansactions on Mathematical Software 8(2) (1982), 114-144.

[9] W.H. Enright, A new error control for initial value solvers, Appl. Math. Comput. 31 (1989), 288-301.

[10] W.H. Enright, Continuous numerical methods for ODEs with defect control, J. Comp. Appl. Math 125(1-2) (2000), 159-170.

[11] S. Ilie, R.M. Corless, The cost of locating singularities and rank changes, in preparation.

[12] K.R. Jackson, N. Nedialkov, Some recent advances in validated methods for IVPs for ODEs, Appl. Numer. Math. 42(1) (2002), 269-284.

[13] E. Hairer, S.P. Norsett, G. Wanner, *Solving Ordinary Differential Equations I*, Computational Mathematics Vol. 8, Springer, Berlin, 1987.

[14] E. Hairer, G. Wanner, *Solving Ordinary Differential Equations II*, Computational Mathematics, Vol. 14, Springer, Berlin, 1991.

[15] W. Kahan, Handheld calculator evaluates integrals, Hewlett-Packard J. 31(8) (1980), 23–32.

[16] C. Moler, Are we there yet? Zero crossing and event handling for differential equations, Matlab News & Notes (1997), 16-17.

[17] N.S. Nedialkov, J.D. Pryce, Solving differential-algebraic equations by Taylor series (I): computing Taylor coefficients, submitted.

[18] N.S.Nedialkov, *private communication*.

[19] H. Nguyen, Interpolation and error control schemes for DAE using continuous implicit Runge-Kutta methods, Ph.D. thesis, Univ. of Toronto, 1995.

[20] J.D. Pryce, A simple structural analysis method for DAEs, BIT 41(2) (2001), 364-394.

[21] J.D. Pryce, Solving high-index DAEs by Taylor series, Numerical Algorithms 19 (1998), 195-211.

[22] L. Shampine, Solving ODEs and DDEs with residual control, J. Comp. Appl. Math., accepted.

[23] L.F. Shampine, M.W. Reichelt, J.A. Kierzenka, Solving Index-1 DAEs in Matlab and Simulink, SIAM Review 41(3) (1999), 538-552.

[24] G. Söderlind, Automatic control and adaptive time-stepping, Numerical Algorithms 31 (2002), 281-310.

[25] A.G. Werschulz, *The Computational Complexity of Differential and Integral Equations*, Oxford Science, Oxford, 1991.