Verified Integration of Linear n-th Order ODEs Using Large Steps

Markus Neher
Institute for Applied and Numerical Mathematics
Universität Karlsruhe (TH), 76128 Karlsruhe, Germany
E-mail: markus.neher@math.uni-karlsruhe.de

June 26, 2006

Abstract

The solution $y(x)$ of an IVP for a linear ODE with analytic coefficient functions is represented as a power series. A high–order Taylor polynomial is used for an approximate numerical solution. The Taylor remainder series is rigorously estimated by some geometric series.

The method has been implemented and tested on a computer. Guaranteed enclosures are achieved by taking into account all roundoff errors.

AMS Subject Classification: 65G40, 65L05, 65L70.

Key Words: Ordinary differential equations; Initial value problems; Error bounds; Interval computations.

1 Introduction

The numerical solution of an initial value problem (IVP) usually yields an approximate solution, which is subject to discretization errors and which may also impaired by roundoff errors, if the computation is carried out on a computer with a floating point number system.

For rigorous integration, verified methods have been designed. Instead of approximate values of unknown quality, verified integration methods yield safe bounds for the solution of a given IVP. Interval computations as described in [2, 24, 38] have proved especially useful for verified integration. Started with the pioneering work of Moore [22, 23, 24], a rich variety of verified integration methods has been developed [3, 5, 6, 10, 14, 17, 18, 20, 21, 25, 26, 27, 28, 29, 41, 42]. Unfortunately, verified integration methods often suffer from overestimation of the local discretization errors and of the so–called wrapping effect, which comes with the propagation of the local errors.

Various methods have been suggested for reducing the wrapping effect [3, 21, 10, 14, 17, 20, 23, 25, 26, 39, 41]. While several of these methods are effective, they are often forced to use small step sizes in the integration. In this paper, we propose a new enclosure method that uses a novel approach to bound the truncation error in the Taylor series method. By analyzing the recurrence relation for the Taylor coefficients of the solution of the IVP, we determine some majorizing geometric series of the Taylor remainder series. The main advantage of our method is that it works with very large step sizes.

In [32], the method was presented for linear ODEs with polynomial coefficient functions. A preliminary extension to linear ODEs with analytic coefficients functions has been given in [33]. In this paper, an analytic transformation of the ODE that was required in [33] is removed. Also, the error bounds in this paper are sharper than those in [33], allowing for larger step sizes and making the computation faster. Numerical tests have shown that very large step sizes can be used in our method in practical calculations.

The paper is structured as follows: In Section 2, a new enclosure algorithm for linear IVPs is elaborated. Various aspects of the implementation of the algorithm on a computer are presented in Section 3. In the last section, numerical examples are given.

We assume that the reader is familiar with interval arithmetic and its implementation on a computer [15]. Our presentation is concentrated on the analytic foundations of the integration method rather than on customary modifications that are necessary when the calculations are performed in interval floating point arithmetic.

2 Enclosure Method for n–th Order Linear ODEs

In full generality, the verified integration method which is proposed in this paper can be applied to linear and to some nonlinear systems of ODEs with piecewise analytic coefficient functions. For simplicity, we restrict the presentation to the case of a single n-th order linear ODE. For $n \geq 2$, we consider the following IVP on some interval $[0, r]$:

$$\begin{aligned} y^{(n)} &= \sum_{i=0}^{n-1} p_i(x)\, y^{(i)} + p_{-1}(x), \quad x \in [0, r] \\ y^{(i)}(0) &= y_{io}, \quad i = 0, 1, \ldots, n-1, \end{aligned} \tag{1}$$

where $y_{io} \in \mathbb{R}$, $i = 0, 1, \ldots, n-1$. The functions $p_i(x)$, $i = -1, 0, \ldots, n-1$, are assumed to be analytic, with power series expansions

$$p_i(x) = \sum_{j=0}^{\infty} b_{ij}\, x^j, \quad x \in [0, r]. \tag{2}$$

It is well known (cf. [4], Ch. III) that the solution of (1) can be written as a power series

$$y(x) := \sum_{k=0}^{\infty} a_k\, x^k, \tag{3}$$

and that the series converges for all x with $|x| \leq r$.

The first n coefficients $a_0, \ldots, a_{n-1}$ are determined from the initial values in (1), namely

$$a_k = \frac{y_{k0}}{k!} \qquad \text{for} \quad k = 0, \ldots, n-1. \tag{4}$$

The recurrence relation for the coefficients with index greater than $n-1$ is derived by the usual way of differentiating (3), inserting the series and its derivatives into (1), and equating terms with equal powers of x. We obtain

$$a_{k+n} \;=\; \sum_{i=0}^{n-1}\sum_{j=0}^{k} \frac{(k-j+1)_i\, b_{ij}}{(k+1)_n}\, a_{k+i-j} \;+\; \frac{b_{-1,k}}{(k+1)_n} \quad \text{for} \quad k=0,1,\ldots, \tag{5}$$

where $(k)_i$ denotes Pochhammer's symbol [1, 6.1.22]

$$(k)_0 := 1, \quad (k)_i \;:=\; k\cdot(k+1)\cdots(k+i-1) \quad \text{for} \quad i,k\in\mathbb{N},$$

and $\mathbb{N}=\{1,2,3,\ldots\}$ denotes the set of natural numbers.

For an approximate solution to (1), we compute finitely many, say κ, coefficients a_k from (4) and (5) and then use the truncated series

$$s(\kappa) := \sum_{k=0}^{\kappa-1} a_k\, h^k \tag{6}$$

for an approximation to $y(h)$ for $h\in(0,r)$. The approximation error is then the truncation error given by the infinite series

$$t(\kappa) := \sum_{k=\kappa}^{\infty} a_k\, h^k. \tag{7}$$

Since the series (7) converges, for sufficiently large κ the remainder series $t(\kappa)$ becomes arbitrarily small.

Our goal is to derive an explicit bound for $t(\kappa)$ in the form of a geometric series. For this purpose, we determine some index κ and some real number A such that

$$|a_k\, r^k| \;\le A \tag{8}$$

holds for all $k\ge\kappa$. Then the following estimate holds:

Theorem 1 *Let κ and A fulfill (8). Then for $h=\omega r$, $0\le\omega<1$, and all $k\ge\kappa$,*

$$|\,y(h)-\sum_{\nu=0}^{k-1} a_\nu\, h^\nu\,| \;\le\; \sum_{\nu=k}^{\infty} |a_\nu\, r^\nu|\,\omega^\nu \;\le\; A\sum_{\nu=k}^{\infty}\omega^\nu \;=\; \frac{A\,\omega^k}{1-\omega}.$$

The main result of this paper is a practical algorithm for determining suitable numbers κ and A for a given IVP (1). As a prerequisite, we recall a summation formula from [37].

Lemma 1 *For $i,k,m\in\mathbb{N}$ and $m>i$, let*

$$M(i,k,m) := \sum_{j=1}^{k} \frac{(k-j+1)_{i-1}}{(j)_m}. \tag{9}$$

Then the following identity holds:

$$\begin{aligned} M(i,k,m) &= \frac{(-1)^i (i-1)!}{(m-i)_i (k+i)_{m-i}} \\ &+ \frac{(i-1)!}{(m-1)!} \sum_{\nu=0}^{i-1} \frac{(-1)^{i-1-\nu}(i-1-\nu)!(k)_\nu}{\nu!(m-i+\nu)_{i-\nu}}. \end{aligned} \tag{10}$$

Lemma 1 contains a transformation rule between sums. It will be applied in the estimation of (7), when i and m are assigned fixed values, but k may be arbitrary. In this case, there are k summands in (9), but the number of summands in the representation (10) is bounded independently of k.

In the following, we consider functions $p_i(x)$ that are represented by a finite number of constants, variables, arithmetic operations, and standard functions like the exponential function, trigonometric functions, or their inverse functions, etc. For some h, $0 < h < r$, we seek lower and upper bounds of $y(h)$. A pair of such bounds is called an enclosure of $y(h)$. We assume that bounds for the Taylor coefficients of the functions $p_i(x)$ of the following form are available: For $i = -1, \ldots, n-1$ there are natural numbers $m_{-1} \geq 0$, $m_i \geq i+2$ for $i = 0, \ldots, i-1$, and real constants $B_i \geq 0$ so that

$$|b_{ij}| \leq \frac{B_i\, r^{m_i}}{(j-m_i+1)_{m_i} r^j} \quad \text{for all } j \geq m_i. \tag{11}$$

The computation of such bounds has been discussed in [9, 34, 35].

For $k > m$, we have from (5) and (11)

$$\begin{aligned} |a_{k+n}\, r^{k+n}| \leq & \sum_{i=0}^{n-1} \left(\sum_{j=0}^{m_i} \frac{r^{n-i+j}\,(k-j+1)_i\; |b_{ij}|}{(k+1)_n} \; |a_{k+i-j}|\; r^{k+i-j} \right. \\ & \left. + \sum_{j=m_i+1}^{k} \frac{r^{n-i+j}\,(k-j+1)_i\; |b_{ij}|}{(k+1)_n} \; |a_{k+i-j}|\; r^{k+i-j} \right) + \frac{|b_{-1,k}|\; r^{k+n}}{(k+1)_n} \\ \leq & \sum_{i=0}^{n-1} \left(\max_{\nu=k-m}^{k+n-1} |a_\nu\, r^\nu| \sum_{j=0}^{m_i} \frac{r^{n-i+j}\,(k-j+1)_i\; |b_{ij}|}{(k+1)_n} \right. \\ & \left. + \max_{\nu=0}^{k+n-1} |a_\nu\, r^\nu| \frac{B_i\, r^{n+m_i-i}}{(k+1)_n} \sum_{j=m_i+1}^{k} \frac{(k-j+1)_i}{(j-m_i+1)_{m_i}} \right) \\ & + \frac{B_{-1}\, r^{n+m_{-1}}}{(k-m_{-1}+1)_{m_{-1}}\,(k+1)_n} \end{aligned}$$

$$= \max_{\nu=0}^{k+n-1} |a_\nu\, r^\nu| \left(\frac{\max\limits_{\nu=k-m}^{k+n-1} |a_\nu\, r^\nu|}{\max\limits_{\nu=0}^{k+n-1} |a_\nu\, r^\nu|} \underbrace{\sum_{i=0}^{n-1} \sum_{j=0}^{m_i} \frac{r^{n-i+j}\,(k-j+1)_i\,|b_{ij}|}{(k+1)_n}}_{=:\mathcal{S}_1(k)} \right.$$

$$+ \underbrace{\sum_{i=0}^{n-1} \frac{B_i\, r^{n+m_i-i}}{(k+1)_n} \sum_{j=2}^{k+1-m_i} \frac{(k+1-m_i-j+1)_i}{(j)_{m_i}}}_{=:\mathcal{S}_2(k)}$$

$$\left. + \underbrace{\frac{B_{-1}\, r^{n+m_{-1}}}{(k-m_{-1}+1)_{m_{-1}+n} \max\limits_{\nu=0}^{k+n-1} |a_\nu\, r^\nu|}}_{=:\mathcal{S}_3(k)} \right),$$

that is

$$|a_{k+n}\, r^{k+n}| \;\leq\; \max_{\nu=0}^{k+n-1} |a_\nu\, r^\nu| \left(\frac{\max\limits_{\nu=k-m}^{k+n-1} |a_\nu\, r^\nu|}{\max\limits_{\nu=0}^{k+n-1} |a_\nu\, r^\nu|}\; \mathcal{S}_1(k) + \mathcal{S}_2(k) + \mathcal{S}_3(k) \right). \tag{12}$$

Due to (9) and (10),

$$\mathcal{S}_2(k) \;=\; \sum_{i=0}^{n-1} \frac{B_i\, r^{n+m_i-i}}{(k+1)_n} \left(M(i+1, k+1-m_i, m_i) - \frac{(k+1-m_i)_i}{m_i!} \right)$$

$$= \sum_{i=0}^{n-1} \frac{B_i\, r^{n+m_i-i}}{(k+1)_n} \left(\frac{(-1)^{i+1} i!}{(m_i-1-i)_{i+1}\,(k-m_i+2+i)_{m_i-1-i}} \right.$$

$$\left. + \sum_{\nu=0}^{i} \frac{(-1)^{i-\nu} i!(i-\nu)!(k-m_i+1)_\nu}{\nu!(m_i-1-i+\nu)_{i-\nu+1}(m_i-1)!} - \frac{(k-m_i+1)_i}{m_i!} \right).$$

Now let

$$\mathcal{W}(k) := \frac{\max\limits_{\nu=k-m}^{k+n-1} |a_\nu\, r^\nu|}{\max\limits_{\nu=0}^{k+n-1} |a_\nu\, r^\nu|}\; \mathcal{S}_1(k) + \mathcal{S}_2(k) + \mathcal{S}_3(k). \tag{13}$$

The above calculations show that $\mathcal{W}(k)$ can be computed from a finite number of summands (independent of k) that each tend to zero for $k \to \infty$. Because the fraction of the maxima in (13) is at most 1, there is some number κ for which the *recess condition*

$$\mathcal{W}(k) \leq 1 \tag{14}$$

holds for all $k \geq \kappa$. Hence,

$$|a_{\kappa+n}\, r^{\kappa+n}| \;\leq\; \max_{\nu=0}^{\kappa+n-1} |a_\nu\, r^\nu|\,,$$

and, by induction,

$$|a_k\, r^k| \;\leq\; A \;:=\; \max_{\nu=0}^{\kappa+n-1} |a_\nu\, r^\nu| \quad \text{for all} \;\; k \in \mathbb{N}_0.$$

In practical calculations, the determination of some suitable value κ is not obvious. Some of the terms in the recess condition (14) are not *monotonically* decreasing to zero for $k \to \infty$. Hence, if (14) holds for some particular κ then it does not necessarily hold for all $k \geq \kappa$. Additional considerations including monotonicity properties are required for specifying κ. Monotonicity properties of some particular terms of the recess condition are stated in the following lemma. The proof is given in the appendix.

Lemma 2 *For $n \in \mathbb{N}$, $\mu \in \mathbb{N}_0$, $0 \leq \nu \leq n-1$, and $k > \dfrac{\mu(n+1)}{n-\nu}$,*

$$\frac{(k-\mu)_\nu}{(k+1)_n}$$

is a monotonically decreasing function of k.

Since $j \leq m$ and $i \leq n-1$ in (14), we have

Corollary 1 *For $k > mn-1$, all fractions of the form $\dfrac{(k-j+1)_i}{(k+1)_n}$ in the recess condition (14) are monotonically decreasing functions of k.*

In general, the denominator of the fraction of the maxima in (14),

$$\frac{\max\limits_{\nu=k-m}^{k+n-1} |a_\nu\, r^\nu|}{\max\limits_{\nu=0}^{k+n-1} |a_\nu\, r^\nu|}, \tag{15}$$

is not a monotonically decreasing function of k. An example for non-monotonic behavior is given in the appendix. We therefore modify the recess condition to obtain a monotonic convergence criterion. As a further benefit, this modification improves the error bound of the remainder series (7).

Theorem 2 *Suppose that for some $\kappa \geq mn$ and some $q \in [0,1]$ it holds that*

$$\frac{\max\limits_{\nu=\kappa-m}^{\kappa+n-1} |a_\nu r^\nu|}{\max\limits_{\nu=0}^{\kappa+n-1} |a_\nu r^\nu|} \;\leq\; q \tag{16}$$

and that

$$q\; \mathcal{S}_1(\kappa) + \mathcal{S}_2(\kappa) + \mathcal{S}_3(\kappa) \;\leq\; q. \tag{17}$$

Then, due to (12),

$$|a_{\kappa+n}r^{\kappa+n}| \;\leq\; q \max_{\nu=0}^{\kappa+n-1} |a_\nu r^\nu|\,,$$

and by induction,

$$|a_k r^k| \;\leq\; q \max_{\nu=0}^{\kappa+n-1} |a_\nu r^\nu| \quad \text{for all } k \geq \kappa + n.$$

qA can then be used instead of A in Theorem 1 for the estimation of the Taylor remainder series.

Proof: Because all summands on the left hand side of (17) tend to zero monotonically for $\kappa \to \infty$, $\kappa \geq mn$, (16) and (17) hold for arbitrary values $k \geq \kappa$, from which the assertions of the theorem follow. □

3 Implementation of the Enclosure Algorithm

We summarize the preceding results in Algorithm 1 for computing an enclosure of $y(h)$. Input data of the algorithm are r, h, the necessary Taylor coefficients b_{ij} of the functions p_i in the differential equation, the initial values y_{io}, and two parameters $\varepsilon_{\text{rel}}, \varepsilon_{\text{abs}} > 0$ for controlling the accuracy of the computed enclosure of $y(h)$. In real arithmetic, the iteration is stopped after finitely many steps when the relative or the absolute error is less than ε_{rel} or ε_{abs}, respectively.

3.1 Computation of Taylor Coefficients and Estimates

For the practical execution of Algorithm 1, Taylor coefficients of the functions $p_i(x)$ in (2) and suitable numbers B_i in (11) are required.

Taylor coefficients are computed recursively via automatic differentiation [11, 24, 40]. The computation of the numbers B_i in (11) is performed with the ACETAF software package [9, 12], which is freely available at the author's website [8].

3.2 Continuing the Integration

Until now, only a single integration step has been described. In practical computations, however, the integration domain is usually split into subintervals on which the assumptions on the smoothness of the IVP (1) hold. In this case, initial values must be supplied for the second and all further integration steps. To maintain the validity of the computation, rigorous error bounds for the solution of (1) and its first $n-1$ derivatives at the grid points are used instead of approximate intermediate initial values.

Error bounds for the function value of the solution y of (1) are computed with Algorithm 1. Error bounds for derivatives are obtained in a similar way, by differentiating the power series (3) and estimating the truncation error as in the estimation of (7). From (3), we have

$$y^{(i)}(h) \;=\; \sum_{\nu=i}^{\infty} \;(\nu - i + 1)_i\, a_\nu\, h^{\nu-i}.$$

Algorithm 1
Single Integration Step of
Enclosure Method for Linear ODEs with Analytic Coefficient Functions

1. For $i = -1, \dots, n-1$:
   Choose $m_i \in \mathbb{N}_0$ ($m_{n-1} > 0$) and compute some B_i satisfying (11).
2. Let $m := \max_{i=-1}^{n-1} m_i$.
3. For $k = 0, \dots, n-1$:
   Compute a_k from (4).
4. For $k = 0, 1, \dots$:
   (a) If $k < mn$ or if (14) is not fulfilled then goto (d).
   (b) Let $q_1 := \dfrac{\max_{\nu=k-m}^{k+n-1} |a_\nu r^\nu|}{\max_{\nu=0}^{k+n-1} |a_\nu r^\nu|}, \quad q_2 := \dfrac{\mathcal{S}_2(k) + \mathcal{S}_3(k)}{1 - \mathcal{S}_1(k)}$
   (c) If $q_1, q_2 \in [0, 1]$:
      α) $q := \max\{q_1, q_2\}$.
      β) $A := \max_{\nu=0}^{k+n-1} |a_\nu r^\nu|$.
      γ) $s(k+n) := \sum_{\nu=0}^{k+n-1} a_\nu h^\nu, \quad T(k+n) := \dfrac{qA\,\omega^{k+n}}{1-\omega}$.
      δ) If $|\dfrac{T(k+n)}{s(k+n)}| \le \varepsilon_{\text{rel}}$ or $|T(k+n)| \le \varepsilon_{\text{abs}}$ then goto 5.
   (d) Compute a_{k+n} from (5).
5. Terminate with $y(h) \in s(k+n) + [-T(k+n), T(k+n)]$.

If $A = \max_{\nu=0}^{\infty} |a_\nu r^\nu|$ and if κ and q are such that (16) and (17) hold, then for $k \geq \kappa$,

$$|y^{(i)}(h) - \sum_{\nu=i}^{k-1} (\nu - i + 1)_i \, a_\nu \, h^{\nu-i}| \;\leq\; \sum_{\nu=k}^{\infty} (\nu - i + 1)_i \, |a_\nu| \, r^\nu \, r^{-i} \omega^{\nu-i}$$

$$\leq \frac{qA}{r^i} \sum_{\nu=k}^{\infty} (\nu - i + 1)_i \, \omega^{\nu-i} \;=\; \frac{qA}{r^i} \frac{d^i}{d\omega^i} \sum_{\nu=k}^{\infty} \omega^\nu \;=\; \frac{qA}{r^i} \frac{d^i}{d\omega^i} \frac{\omega^k}{1-\omega} \; .$$

An explicit formula for the latter derivative is presented in the following lemma, which is proved in the appendix.

Lemma 3 *For $\omega \in (0,1)$, $k \in \mathbb{N}$ and $0 \leq i \leq k-1$,*

$$\frac{d^i}{d\omega^i} \frac{\omega^k}{1-\omega} \;=\; \frac{(-1)^i (k-i)_{i+1} \omega^k}{(1-\omega)^{i+1}} \sum_{\nu=0}^{i} \binom{i}{\nu} \frac{(-\omega)^{-\nu}}{k-\nu}.$$

3.3 Interval Initial Values

An important aspect in the validated solution of our original IVP (1) is the treatment of the intermediate integration steps, when initial conditions are given by intervals instead of real numbers.

It is well known that inserting interval values into recurrence relations like (5) is unsuitable. Due to the dependency problem in interval arithmetic, the diameters of the computed enclosures are likely to grow with the increasing number of summands in (6). For the linear ODE (1), we use $n + 1$ IVPs with real initial conditions to compute one particular solution of the inhomogeneous ODE and a fundamental system of the corresponding homogeneous ODE. The enclosure of the flow of the interval initial set is then built from linear combinations of these solutions (see for example [33]).

3.4 Step Size Control

For many integration methods, the determination of a suitable step size for the integration is difficult. In our method, the modified recess condition (17) is a good indicator for the amount of work that is to be expected in a single integration step. To determine the local step size r_j, the modified recess condition is tested for appropriate values of κ, using approximations (instead of validated bounds) for the numbers B_i. Such approximations are computed with a few function evaluations of the functions p_i. Then q_2 in Step 4 of Algorithm 1 is computed approximately for suitable values of k. If q_2 is greater than 1, then r_j is reduced and the numbers B_i are recalculated. If $q_2 \ll 1$ holds, r_j is increased, to avoid unnecessarily small steps. This process is repeated until $q_2 < 1$ holds.

3.5 Multiple Precision Arithmetic

Using a large step size in the integration has the side-effect of delaying convergence of the Taylor series, so that several dozens or even a few hundred addends are involved in the summation. In floating point arithmetic, the accuracy of the result is then threatened by cancellation. Hence, any implementation of Algorithm 1 should be based on a multiple precision number format, such as the staggered correction format [19].

4 Numerical Examples

Algorithm 1 has been implemented in a Pascal-XSC [13] program called livp_ls (solving *l*inear *IVP*s with *l*arge *s*teps). The complete code of livp_ls is available at the author's website [36].

For the calculations of the numerical examples with livp_ls, besides the values of the computed enclosures, we list the values of L (denoting L-fold precision) in the staggered correction format and of κ in the partial sum $s(\kappa)$, for which the enclosures were obtained.

The numbers B_i and the truncation error bound $T(k)$ in Algorithm 1 depend on the choices of m_i and $h = \omega r_j$. There is no general rule for the optimal choice of these parameters. The bounds B_i are likely to grow with increasing m_i or decreasing ω. Analogously, if ω decreases then the smallest value of k, for which $0 \leq q_2 \leq 1$ holds, will probably become larger. On the other hand, for large values k, the truncation error bound $T(k)$ decreases with ω. Hence, for ω near 0, it may take a very large k to obtain $q_2 \in [0, 1]$, whereas for ω near 1, only for very large k will $T(k)$ be small enough to terminate Algorithm 1. In our numerical examples, we used $m_i = 10$ or $m_i = 20$, $\omega \in [0.5, 0.8]$. Each integration was carried out in a single integration step.

Example 1: $\quad y'' = \alpha e^x y + e^{-x} - \alpha,\ y(0) = 1,\ y'(0) = -1.$

Exact Solution: $y(x) = e^{-x}$.

This problem provides substantial insight into the behavior of our integration method. For all parameter values, the integration is performed on [0,1] using a single step.

The numerical solution of this problem gets more and more difficult with increasing α. For $\alpha \leq 1000$, double precision is sufficient for computing optimal bounds (with respect to single precision). The number of summands $k_{\max}$ in the approximating Taylor polynomial decreases with increasing m. For $\alpha \geq 2500$, optimal bounds are obtained with triple or even quadruple precision. The bounds computed for $\alpha = 2500$ show that not $k_{\max}$, but precision in computation plays the key role for the accuracy of the result (Table 1).

These results are compared with results obtained by Lohner's program AWA [16, 17, 18]. For the computations with AWA (Table 2), besides the enclosures, the respective degrees of the Taylor expansions and the number of integration steps are shown. In contrast to livp_ls, AWA aborts the integration prematurely for $\alpha > 500$. This is only in part due to the single precision used in AWA. The main source of error in this example is the wrapping effect, which in our method is eliminated by performing the integration in a single step.

α	$[y](1)$	Prec.	m	k_{max}
500	$3.678\ 794\ 411\ 714\ 42^{4}_{3}$E-01	2	0	209
500	$3.678\ 794\ 411\ 714\ 42^{4}_{3}$E-01	2	10	145
500	$3.678\ 794\ 411\ 714\ 42^{4}_{3}$E-01	2	20	53
1000	$3.678\ 794\ 411\ 714\ 42^{4}_{3}$E-01	2	0	148
1000	$3.678\ 794\ 411\ 714\ 42^{4}_{3}$E-01	2	10	145
1000	$3.678\ 794\ 411\ 714\ 42^{4}_{3}$E-01	2	20	74
2500	$3.678\ 794\ 4^{18}_{06}$E-01	2	0	216
2500	$3.678\ 794\ 4^{18}_{06}$E-01	2	10	180
2500	$3.678\ 794\ 4^{18}_{06}$E-01	2	20	185
2500	$3.678\ 794\ 411\ 714\ 42^{4}_{3}$E-01	3	20	117
5000	[-99.7, 122.9]	2	0	259
5000	$3.678\ 794\ 411\ 714\ ^{52}_{33}$E-01	3	0	220
5000	$3.678\ 794\ 411\ 714\ 42^{4}_{3}$E-01	4	0	220
5000	$3.678\ 794\ 411\ 714\ 42^{4}_{3}$E-01	4	20	165

Table 1: Integration of Problem 1 with livp_ls.

α	$[y](1)$/abort at	Order	Steps
500	$3.^{744}_{623}$E-01	20	84
1000	$x = 9.551$E-001	20	110
2500	$x = 6.519$E-001	20	113
5000	$x = 4.774$E-001	20	113
5000	$x = 4.650$E-001	30	879

Table 2: Integration of Problem 1 with AWA.

$y'(0)$	$[y](1)$	Prec.	m	k_{max}
$-19.999\ 999\ 917\ 553\ 86$	$-2.874\ 565\ 781\ 786\ {}^{799}_{809}\mathrm{E}-08$	2	10	102
$-19.999\ 999\ 917\ 553\ 85$	$1.434\ 566\ 783\ 738\ {}^{900}_{890}\mathrm{E}-08$	2	10	102

Table 3: Integration of Problem 2 with livp_ls.

Example 2: BVP

$$\begin{aligned} y'' &= 400y + (200 + 2\pi^2)\cos(2\pi x) + 200 \\ y(0) &= y(1) = 0. \end{aligned}$$

This boundary value problem from the book of Stoer and Bulirsch [43, Ch. 7.6] is first solved with a shooting method to compute approximate values for $y'(0)$. These approximations are then tested with livp_ls (Table 3). Thus optimal bounds (with respect to single precision) for $y'(0)$ are obtained:

$$y'(0) \in -19.999\ 999\ 917\ 553\ 8^{5}_{6}.$$

Example 3: Computation of eigenvalues of

$$\begin{aligned} -u'' + \cos(2x)\,u &= \lambda u \\ u(0) = u(pi) = 0. \end{aligned} \tag{18}$$

A real number λ is called an eigenvalue of the boundary value problem (18) if there is a nontrivial solution $u(x)$ of (18). As is well known, the eigenvalues of (18) form a sequence of real numbers $\{\lambda_\nu\}_{\nu=1}^{\infty}$ which is bounded from below and tends to infinity (cf. [7], Chap. V).

Following the well–known Sturm–Liouville theory, we compute eigenvalues of (18) with a shooting method with shooting parameter λ, using the initial value problem

$$\begin{aligned} y'' &= (\cos x - \lambda)\,y \\ y(0) &= 0,\ y'(0) = 1. \end{aligned} \tag{19}$$

For $\lambda \in \mathbb{R}$ fixed, let $y(\cdot;\lambda)$ denotes the solution of (19). If $y(\cdot;\lambda)$ has $\nu-1$ zeros within $(0,\pi)$ then λ is a lower bound of λ_ν an upper bound of $\lambda_{\nu-1}$. If additionally $y(\pi;\lambda) = 0$ then $\lambda = \lambda_\nu$. Due to the symmetry of the differential equation and the boundary conditions, eigenvalues can also be computed by testing $y(\frac{\pi}{2};\lambda) = 0$ (for eigenvalues with even indexes) or $y'(\frac{\pi}{2};\lambda) = 0$ (for eigenvalues with odd indexes).

Lower and upper eigenvalue bounds for specific eigenvalues are computed by solving (19) for different values of λ and $x \in [0,\pi]$, and by determining the number $N(\lambda)$ of zeros of the respective solutions $y(\cdot;\lambda)$ within $(0,\pi)$ (see [30] for the computation of suitable starting values λ and for the reliable determination of $N(\lambda)$). Eigenvalue bounds are refined by bisection with respect to λ (faster algorithms were given in [30]).

In Table 4, the enclosures of $y(\pi;\lambda)$ from (19) are listed for four different values of λ. From these enclosures, the eigenvalues λ_4 and λ_{11} of (18) are determined with optimal

λ	$N(\lambda)$	$[y(\frac{\pi}{2})]$	Prec.	m	k_{max}
16.008 310 459 709 47	3	< -3.8E-16	2	20	171
16.008 310 459 709 48	4	> 1.3E-16	2	20	173
λ	$N(\lambda)$	$[y'(\frac{\pi}{2})]$	Prec.	m	k_{max}
121.001 041 672 579 0	10	< -1.0E-16	2	20	233
121.001 041 672 579 1	11	> 7.0E-15	2	20	221

Table 4: Integration of Problem 3 with livp_ls.

accuracy (with respect to single precision):

$$\begin{aligned} \lambda_4 &\in \quad 16.008\ 310\ 459\ 709\ 4_7^8, \\ \lambda_{11} &\in 121.001\ 041\ 672\ 57_{90}^{91}. \end{aligned}$$

5 Conclusion

We have presented a new enclosure method for linear ODEs with analytic coefficient functions. Our numerical examples demonstrate the successful implementation of the method on a computer.

The estimation procedure can be extended to linear ODEs with piecewise analytic coefficient functions (e.g. splines), provided that the breakpoints of analyticity are known. With some modifications, it can also be applied to systems of linear ODEs, and to a class of nonlinear ODEs (cf. [31]). These extensions will be the subject of future research.

Appendix

Proof of Lemma 2: For $n \in \mathbb{N}$, $\mu \in \mathbb{N}_0$, and $0 \leq \nu \leq n-1$,

$$\begin{aligned} \frac{(k+1-\mu)_\nu}{(k+2)_n} &= \frac{(k-\mu)_\nu}{(k+1)_n} \cdot \frac{(k-\mu+\nu)(k+1)}{(k-\mu)(k+1+n)} \\ &\overset{!}{<} \frac{(k-\mu)_\nu}{(k+1)_n} \quad \Longleftrightarrow \quad (k-\mu+\nu)(k+1) \ < \ (k-\mu)(k+1+n) \\ &\Longleftrightarrow \quad k \ > \ \frac{\mu(n+1)}{n-\nu}. \end{aligned}$$

□

Proof of Lemma 3: The assertion obviously holds for $i = 0$. By induction, we have

$$\frac{d^{i+1}}{d\omega^{i+1}}\frac{\omega^k}{1-\omega} = \frac{d}{d\omega}\left[\frac{(-1)^i(k-i)_{i+1}\omega^k}{(1-\omega)^{i+1}}\sum_{\nu=0}^{i}\binom{i}{\nu}\frac{(-\omega)^{-\nu}}{k-\nu}\right]$$

$$= \frac{(-1)^i(k-i)_{i+1}\left(k\omega^{k-1}(1-\omega)^{i+1}+\omega^k(1-\omega)^i(i+1)\right)}{(1-\omega)^{2i+2}}\sum_{\nu=0}^{i}\binom{i}{\nu}\frac{(-\omega)^{-\nu}}{k-\nu}$$

$$+ \frac{(-1)^i(k-i)_{i+1}\omega^k}{(1-\omega)^{i+1}}\sum_{\nu=0}^{i}\binom{i}{\nu}\frac{\nu(-\omega)^{-\nu-1}}{k-\nu}$$

$$= \frac{(-1)^i(k-i)_{i+1}\omega^{k-1}}{(1-\omega)^{i+2}}\left\{\Big(k(1-\omega)+\omega(i+1)\Big)\sum_{\nu=0}^{i}\binom{i}{\nu}\frac{(-\omega)^{-\nu}}{k-\nu}\right.$$

$$\left. + \ \omega(1-\omega)\sum_{\nu=0}^{i}\binom{i}{\nu}\frac{\nu(-\omega)^{-\nu-1}}{k-\nu}\right\} .$$

The terms within the curly brackets are equal to

$$\sum_{\nu=0}^{i}\binom{i}{\nu}\frac{\Big((i+1)\omega+(k-\nu)(1-\omega)\Big)(-\omega)^{-\nu}}{k-\nu}$$

$$= \omega\sum_{\nu=0}^{i}(i+1)\binom{i}{\nu}\frac{(-\omega)^{-\nu}}{k-\nu}+(1-\omega)\sum_{\nu=0}^{i}\binom{i}{\nu}(-\omega)^{-\nu}$$

$$= \omega\sum_{\nu=0}^{i}(i+1-\nu)\binom{i+1}{\nu}\frac{(-\omega)^{-\nu}}{k-\nu}+(1-\omega)(1-\frac{1}{\omega})^i$$

$$= \omega\sum_{\nu=0}^{i}(i+1-k)\binom{i+1}{\nu}\frac{(-\omega)^{-\nu}}{k-\nu}+\omega\sum_{\nu=0}^{i}\binom{i+1}{\nu}(-\omega)^{-\nu}-\omega(1-\frac{1}{\omega})^{i+1}$$

$$= \omega(i+1-k)\sum_{\nu=0}^{i+1}\binom{i+1}{\nu}\frac{(-\omega)^{-\nu}}{k-\nu}-\omega(i+1-k)\frac{(-\omega)^{-i-1}}{k-i-1}$$

$$+ \ \omega\sum_{\nu=0}^{i+1}\binom{i+1}{\nu}(-\omega)^{-\nu}-\omega(-\omega)^{-i-1}-\omega(1-\frac{1}{\omega})^{i+1}$$

$$= \omega(i+1-k)\sum_{\nu=0}^{i+1}\binom{i+1}{\nu}\frac{(-\omega)^{-\nu}}{k-\nu}.$$

From

$$(i+1-k)\,(k-i)_{i+1} \;=\; -\,(k-i-1)_{i+2}$$

we finally have

$$\frac{d^{i+1}}{d\omega^{i+1}}\frac{\omega^k}{1-\omega} = \frac{(-1)^{i+1}(k-i-1)_{i+2}\omega^k}{(1-\omega)^{i+2}}\sum_{\nu=0}^{i+1}\binom{i+1}{\nu}\frac{(-\omega)^{-\nu}}{k-\nu}.$$

□

Example of Non–Monotonicity in (14)

Consider the IVP

$$y'' + 100\,y = 250 + 90\,x^8 + 100\,x^{10}$$
$$y(0) = 2.5, \quad y'(0) = 10^{-3}$$

which is solved by $y(x) = 2.5 + 10^{-4}\sin(10x) + x^{10}$.

The first eleven Taylor coefficients a_k of y are

$$a_0 = 2.5, \quad a_1 = 10^{-3}, \quad a_2 = 0, \quad a_3 = -\tfrac{1}{60}, \quad a_4 = 0, \quad a_5 = \tfrac{1}{12},$$

$$a_6 = 0, \quad a_7 = \tfrac{25}{126}, \quad a_8 = 0, \quad a_9 = \tfrac{625}{2268}, \quad a_{10} = 1.$$

For $r = 1$, $m_1 = m_0 = 2$, $m_{-1} = 0$, $B_1 = B_0 = 0$, $B_{-1} = 190$, $b_{00} = 100$, $A := \max_{\nu=0}^{\infty} |a_\nu| = 2.5$, we have the following multipliers in the recess condition (14):

$$k = 8: \quad \frac{\max_{\nu=6}^{9} |a_\nu|}{A}\frac{|b_{00}|}{9\cdot 10} + \frac{B_{-1}}{9\cdot 10A} = \frac{1}{225}(100a_9 + 190) \lesssim 0.9670,$$

$$k = 9: \quad \frac{\max_{\nu=7}^{10} |a_\nu|}{A}\frac{|b_{00}|}{10\cdot 11} + \frac{B_{-1}}{10\cdot 11A} = \frac{1}{275}(100a_{10} + 190) \lesssim 1.055.$$

For $k = 8$, (14) gives $|a_{10}| < 0.9670\,A$, but only $|a_{11}| < 1.055\,A$ can be guaranteed for $k = 9$, even though all fractions in (14) except (15) are monotonically decreasing functions of k for $k \geq 4$ according to Corollary 1.

References

[1] M. Abramowitz and I. Stegun. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. National Bureau of Standards, Washington, 1964.

[2] G. Alefeld and J. Herzberger. *Introduction to Interval Computations*. Academic Press, New York, 1983.

[3] M. Berz and K. Makino. Verified integration of ODEs and flows using differential algebraic methods on high-order Taylor models. *Reliable Computing*, 4:361–369, 1998.

[4] G. Birkhoff and G.-C. Rota. *Ordinary Differential Equations*. Ginn and Company, Boston, 1962.

[5] G. F. Corliss. Survey of interval algorithms for ordinary differential equations. *Appl. Math. Comput.*, 31:112–120, 1989.

[6] G. F. Corliss. Guaranteed error bounds for ordinary differential equations. In M. Ainsworth, J. Levesley, W. A. Light, and M. Marletta, editors, *Theory and Numerics of Ordinary and Partial Differential Equations*. Clarendon Press, Oxford, 1995.

[7] R. Courant and D. Hilbert. *Methods of Mathematical Physics, Vol. 1.* Interscience Publishers, New York, 1953.

[8] I. Eble and M. Neher. ACETAF 2.8: A c++ program for the automatic computation of estimates for Taylor coefficients of analytic functions. `http://www.uni-karlsruhe.de/~Markus.Neher/acetaf.html`, March 2003.

[9] I. Eble and M. Neher. ACETAF: A software package for computing validated bounds for Taylor coefficients of analytic functions. *ACM TOMS*, 29:263–286, 2003.

[10] P. Eijgenraam. *The Solution of Initial Value Problems Using Interval Arithmetic.* Mathematical Centre Tracts No. 144, Stichting Mathematisch Centrum, 1981.

[11] A. Griewank. *Evaluating Derivatives: Principles and Techniques of Algorithmic Differentiation.* SIAM, Philadelphia, 2000.

[12] R. B. Kearfott, M. Neher, S. Oishi, and F. Rico. Libraries, tools, and interactive systems for verified computations: Four case studies. In R. Alt, A. Frommer, R. B. Kearfott, and W. Luther, editors, *Numerical Software with Result Verification, Springer Lecture Notes in Computer Science 2991*, pages 36–63. Springer, Berlin, 2004.

[13] R. Klatte, U. Kulisch, M. Neaga, D. Ratz, and Ch. Ullrich. *Pascal-XSC – Language Reference with Examples.* Springer, Berlin, 1992.

[14] W. Kühn. Rigorously computed orbits of dynamical systems without the wrapping effect. *Computing*, 61:47–67, 1998.

[15] U. Kulisch and W. L. Miranker. *Computer Arithmetic in Theory and Practice.* Academic Press, New York, 1981.

[16] R. Lohner. Enclosing the solutions of ordinary initial- and boundary-value problems. In E. Kaucher, U. Kulisch, and Ch. Ullrich, editors, *Computerarithmetic: Scientific Computation and Programming Languages*, pages 255–286. Teubner, Stuttgart, 1987.

[17] R. Lohner. *Einschließung der Lösung gewöhnlicher Anfangs- und Randwertaufgaben und Anwendungen.* PhD thesis, Universität Karlsruhe, 1988.

[18] R. Lohner. Computation of guaranteed solutions of ordinary initial and boundary value problems. In J. R. Cash and I. Gladwell, editors, *Computational Ordinary Differential Equations*, pages 425–435. Clarendon Press, Oxford, 1992.

[19] R. Lohner. Interval arithmetic in staggered correction format. In E. Adams and U. Kulisch, editors, *Scientific Computing with Automatic Result Verification*, pages 301–321. Academic Press, Boston, 1993.

[20] K. Makino. *Rigorous analysis of nonlinear motion in particle accelerators.* PhD thesis, Michigan State University, 1998.

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

[22] R. E. Moore. The automatic analysis and control of error in digital computation based on the use of interval numbers. In L. B. Rall, editor, *Error in Digital Computation, Vol. I*, pages 61–130. John Wiley and Sons, New York, 1965.

[23] R. E. Moore. Automatic local coordinate transformations to reduce the growth of error bounds in interval computation of solutions of ordinary differential equations. In L. B. Rall, editor, *Error in Digital Computation, Vol. II*, pages 103–140. John Wiley and Sons, New York, 1965.

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

[25] N. S. Nedialkov. *Computing rigorous bounds on the solution of an IVP for an ODE.* PhD thesis, University of Toronto, 1999.

[26] N. S. Nedialkov and K. R. Jackson. A new perspective on the wrapping effect in interval methods for initial value problems for ordinary differential equations. In U. Kulisch, R. Lohner, and A. Facius, editors, *Perspectives of Enclosure Methods*, pages 219–264. Springer, Wien, 2001.

[27] N. S. Nedialkov and K. R. Jackson. Some recent advances in validated methods for ivps for odes. *Appl. Numer. Math.*, 42:269–284, 2003.

[28] N. S. Nedialkov, K. R. Jackson, and G. F. Corliss. Validated solutions of initial value problems for ordinary differential equations. *Appl. Math. Comput.*, 105:21–68, 1999.

[29] N. S. Nedialkov, K. R. Jackson, and J. Pryce. An effective high-order interval method for validating existence and uniqueness of the solution of an IVP for an ODE. *Reliable Computing*, 7:449–465, 2001.

[30] M. Neher. Inclusion of eigenvalues and eigenfunctions of the Sturm-Liouville problem. In L. Atanassova and J. Herzberger, editors, *Computer Arithmetic and Enclosure Methods*, pages 401–408. Elsevier, Amsterdam, 1992.

[31] M. Neher. Enclosing power series solutions of ODEs. *ZAMM*, 77(S2):S635–S636, 1997.

[32] M. Neher. An enclosure method for the solution of linear ODEs with polynomial coefficients. *Numer. Funct. Anal. and Optimiz.*, 20:779–803, 1999.

[33] M. Neher. Geometric series bounds for the local errors of Taylor methods for linear n-th order ODEs. In G. Alefeld, J. Rohn, S. Rump, and T. Yamamoto, editors, *Symbolic Algebraic Methods and Verification Methods*, pages 183–193. Springer, Wien, 2001.

[34] M. Neher. Validated bounds for Taylor coefficients of analytic functions. *Reliable Computing*, 7:307–319, 2001.

[35] M. Neher. Improved validated bounds for Taylor coefficients and for Taylor remainder series. *J. Comput. Appl. Math.*, 152:393–404, 2003.

[36] M. Neher. livp_ls: A Pascal-XSC program for the verified integration of linear n-th order ODEs using large steps. http://www.uni-karlsruhe.de/~Markus.Neher/livp_ls.html, June 2006.

[37] M. Neher. *A Note on a Sum Associated with the Generalized Hypergeometric Function.* Preprint 06/19, Fakultät für Mathematik, Universität Karlsruhe, Germany, 2006.

[38] A. Neumaier. *Interval Methods for Systems of Equations.* Cambridge University Press, Cambridge, 1990.

[39] K. Petras. Validated solution of ODEs via Runge-Kutta methods. Presented at the SCAN 98 International Conference, 1998.

[40] L. B. Rall. *Automatic Differentiation: Techniques and Applications, Lecture Notes in Computer Science, Vol. 120.* Springer, Berlin, 1981.

[41] R. Rihm. *Über Einschließungsverfahren für gewöhnliche Anfangswertprobleme und ihre Anwendung auf Differentialgleichungen mit unstetiger rechter Seite.* PhD thesis, Universität Karlsruhe, Germany, 1993.

[42] R. Rihm. Implicit methods for enclosing solutions of ODEs. *J. UCS*, 4:202–209, 1998.

[43] J. Stoer and R. Bulirsch. *Introduction to Numerical Analysis.* Springer, Berlin, 1980.