Adaptive integration of delay differential equations

A. Bellen and M. Zennaro
Dipartimento di Scienze Matematiche
Università degli Studi di Trieste
E-mail: bellen@univ.trieste.it
E-mail: zennaro@univ.trieste.it

Abstract

We consider the numerical solution of nonstiff delay differential equations by means of a variable stepsize continuous Runge-Kutta method, the *advancing method*, of discrete order p and uniform order $q \leq p$. As in the well-known case of ordinary differential equations, the stepsize control mechanism is based on the use of a second method, the *error-estimating method*, of order $p' \neq p$, used to adapt the current stepsize in order that the local error fits a user-supplied tolerance TOL. We detect the minimal uniform order q needed for the correct performance of both the advancing and the error-estimating methods. After stating the relationships among global and local errors, we also discuss the effectiveness of the stepsize control mechanism in connection to the possible use of an additional continuous error-estimating method, which is used to monitor the uniform local error. A complete and detailed description of both the advancing method of uniform order 4 and the discrete error-estimating method of order 5 is given so as to enable the interested reader to implement his own code. A uniform error-estimating mechanism is also given which is recommendable for the reliability of the overall procedure. Finally, numerical experiments with a constructed equation are carried out aimed to test and check the results provided by the theory.

1 Introduction

We consider the numerical solution of nonstiff delay differential equations (DDEs) of the form

$$\begin{cases} y'(t) = f\Big(t, y(t), y(t - \tau(t, y(t)))\Big), & t_0 \leq t \leq t_f, \\ y(t) = \phi(t), \quad t \leq t_0, \end{cases}$$

by means of variable stepsize continuous Runge-Kutta (RK) methods of discrete (nodal) order p and uniform order $q \leq p$. Our aim is to ensure, as far as possible, the proportionality of the maximum global error to a given tolerance TOL and to minimize the computational cost.

It is well known that, for ordinary differential equations (ODEs), the stepsize control mechanism for a (discrete) *advancing method* of order p is mostly based on the use of a second method, the *error-estimating method*, used to adapt the current stepsize in order that the local error fits the tolerance TOL. The error-estimating method has order p', usually $p+1$ or $p-1$. When $p' = p+1$ the term TOL is used as a *tolerance per unit step* whereas, for $p' = p-1$, it is used as a *tolerance per step*. However, in both cases the global error turns out to be proportional to TOL, even though in the second case the proportionality can fail in presence of fortuitous and unpredictable superconvergence of the error-estimating method. Some imbedded pairs of advancing and error-estimating methods of order p and $p+1$ respectively has been proposed by Fehlberg for $p = 2$, $p = 4$

and $p = 7$ and are known as Runge-Kutta-Fehlberg pairs $RKF2(3)$, $RKF4(5)$ and $RKF7(8)$. The alternative approach with $p' = p - 1$ has been considered much later by Dormand and Prince [2] who provided, for $p = 5$ and $p' = 4$ the pair DoPri5(4). For a comprehensive analysis of this topic and presentation of more general imbedded pairs, we refer the interested reader to the book by Hairer, Nørsett and Wanner [4].

Since the step-by-step numerical integration of DDEs is based on the use of continuous approximations with possibly different orders of discrete and continuous local errors, the extension of the error control mechanism is not straighforward, and its mathematical justification is far from being obvious.

For DDEs, the local problem to be solved by the advancing method is

$$\begin{cases} w'_{n+1}(t) = f(t, w_{n+1}(t), x(t - \tau(t, w_{n+1}(t)))), & t_n \le t \le t_{n+1}, \\ w_{n+1}(t_n) = y_n, \end{cases} \tag{1}$$

$$x(s) = \begin{cases} \phi(s) & \text{for} \quad s \le t_0, \\ \eta(s) & \text{for} \quad t_0 \le s \le t_n, \quad \text{(known)} \\ w_{n+1}(s) & \text{for} \quad t_n \le s \le t_{n+1}, \quad \text{(to be computed)} \end{cases}$$

where, for $s < t_n$, the delayed approximate solution $\eta(s)$ has uniform global order p, that is

$$\max_{s \le t_n} \|y(s) - \eta(s)\| = O(h^p), \tag{2}$$

where $h = \max_n h_{n+1}$.

The continuous RK method for (1) takes the form

$$\eta(t_n + \theta h_{n+1}) = y_n + h_{n+1} \sum_{i=1}^{s} b_i(\theta) K_{n+1}^i, \quad 0 \le \theta \le 1, \tag{3}$$

where the system

$$\begin{aligned} K_{n+1}^i &= f\left(t_{n+1}^j, Y_{n+1}^j, x(t_{n+1}^j - \tau(t_{n+1}^j, Y_{n+1}^j))\right), \\ Y_{n+1}^i &= y_n + h_{n+1} \sum_{j=1}^{s} a_{ij} K_{n+1}^j, \end{aligned} \qquad i = 1, \ldots, s, \tag{4}$$

has to be solved for the *stage* values K_{n+1}^i (and, hence, for the Y_{n+1}^i's, that are called stage values too). Here $t_{n+1}^i = t_{n+1} + c_i h_{n+1}$ and the c_i are the *abscissae* of the RK method.

In particular, for $\theta = 1$ we get the nodal value

$$y_{n+1} = \eta(t_{n+1}) = y_n + h_{n+1} \sum_{i=1}^{s} b_i K_{n+1}^i.$$

If overlapping occurs, that is if, for some index i, the argument $t_{n+1}^i - \tau(t_{n+1}^i, Y_{n+1}^i)$ of $x(s)$ lies in the current interval $[t_n, t_{n+1}]$, then the *spurious stage* value $\tilde{Y}_{n+1}^i = x(t_{n+1}^i - \tau(t_{n+1}^i, Y_{n+1}^i))$ arises which is given by the formula (3) itself for

$$\theta = \theta_{n+1}^i = c_i - \frac{\tau(t_{n+1}^i, Y_{n+1}^i)}{h_{n+1}},$$

that is

$$\tilde{Y}_{n+1}^i = y_n + h_{n+1} \sum_{j=1}^{s} b_j(\theta_{n+1}^i) K_{n+1}^j. \tag{5}$$

It is worth remarking that, in the case of overlapping, the overall RK method becomes implicit even if the underlying RK method is explicit, that is the coefficient matrix (a_{ij}) is strictly lower triangular. Remark also that overlapping may occur in equations with state-dependent delay, where the value of the delay τ is impredictable, as well as in equations with constant delay for large values of the stepsize.

It is known that, as long as the solution is smooth, in order for the advancing DDE method to perform to the discrete and uniform global order p it is sufficient (and necessary) to use a discrete RK method of order p (local error $O(h^{p+1})$) and a continuous extension of order q (uniform local error $O(h^{q+1})$) with $q = p - 1$ only. Remark that, in this case, the local and global uniform errors in the DDE method are both $O(h^p)$. In other words, the error of the continuous extension does not propagate as t increases.

On the other hand, if the initial function $\phi(t)$ and the solution $y(t)$ do not link smoothly (or even continuously) at t_0, a number of discontinuity points (called *breaking points*) propagate along the integration interval, where the solution is not regular enough and the desired order of the error is no longer guaranteed by the numerical method. In this case the required conditions are recovered by simply including the breaking points as mesh points in order to proceed across steps where the solution is piecewise as smooth as necessary. Whereas locating the breaking points is trivial for constant and time-dependent delays, it is not an easy task for state-dependent delays where their location depends on the solution itself. For an insight on possible strategies for going around these difficulties see Bellen and Zennaro [1, Chapter 4]. In particular, the recent code RADAR5 developed by Guglielmi and Hairer [3], gives up locating the breaking points and try to control the local error by simply suitably reducing the stepsize in their proximity.

However, the choice of the uniform order q in the advancing method has an impact on the accuracy of the error-estimating method. More precisely:

- If the (discrete) error-estimating method has order $p' = p + 1$ (local error $O(h^{p+2})$), then the uniform order of the continuous advancing method should be raised to $q = p$ (local and global uniform error $O(h^{p+1})$). On the contrary, if $p' = p - 1$, then $q = p - 1$ is sufficient. For a rigorous proof of this statement, see [1, Chapter 7].

2 The stepsize control mechanism

In order to state the results on the stepsize control let us consider a mesh $\Delta = \{t_0, t_1, \ldots, t_N = t_f\}$ along the integration interval $[t_0, t_f]$ and denote the maximum discrete and uniform global errors by

$$e_n = \max_{i \le n} \|y(t_i) - y_i\| \quad \text{and} \quad E_n = \max_{t \le t_n} \|y(t) - \eta(t)\|.$$

Moreover, let

$$\sigma_{n+1} = \|w_{n+1}(t_{n+1}) - y_{n+1}\| / h_{n+1}$$

and

$$\Sigma_{n+1} = \max_{t_n \le t \le t_{n+1}} \|w_{n+1}(t) - \eta(t)\|$$

be the *discrete local error per unit step* and the *uniform local error* of the (local) problem (1) respectively. Finally, let us define

$$e_\Delta = \max_n e_n, \qquad \sigma_\Delta = \max_n \sigma_n \quad \text{and} \quad \Sigma_\Delta = \max_n \Sigma_n.$$

By a suitable modification of the standard propagation error analysis taking into account the error on the approximated delayed term x in (1), the conclusions we arrive at are synthetized in the following two points:

- The following relation among global and local errors holds independently of the presence of overlapping:
$$e_\Delta \leq K(\sigma_\Delta + R\Sigma_\Delta). \tag{6}$$
Therefore
$$\sigma_n \leq TOL \quad \text{and} \quad \Sigma_n \leq TOL \implies e_\Delta \leq K' \cdot TOL.$$
So we can say that both the discrete and the uniform local errors should be controlled, rather than the sole discrete local error as is done in ODE codes. More precisely, the given tolerance TOL should bound both the discrete local error per unit step and the uniform local error. For the latter, one can use a continuous error-estimating method of uniform order q' other than q, say $q' = q - 1$.

- Under the option $q = p$ in the uniform order of the advancing method, one might be satisfied with controlling the discrete local error only. In fact, we have
$$p = q \implies \Sigma_\Delta = hO(\sigma_\Delta).$$
Therefore, as TOL decreases the contribution of Σ_Δ in (6) becomes smaller and smaller and eventually it turns out to be negligible in comparison with σ_Δ.

A rigorous proof of the above statements is given in [1, Chapter 7] in the simplified case of variable non state-dependent delay.

The case $q = p-1$ leads to local errors σ_Δ and Σ_Δ of the same order so that it is recommendable to submit both of them to the test of tolerance.

3 A sample method

In this section we illustrate the foregoing theory by means of a sample method. Our presentation shows some degrees of freedom, so that some parameters are not fixed a priori. They have to be chosen by the reader who wants to actually implement the method we propose.

We illustrate the strategy based on an advancing method of uniform order $q = p$ endowed with a discrete error-estimating method of order $p' = p + 1$.

In this setting, we base our procedure on explicit RK methods of order $p = 4$, namely the classical four-stage RK method

$$\begin{array}{c|cccc}
0 & & & & \\
\frac{1}{2} & \frac{1}{2} & & & \\
\frac{1}{2} & 0 & \frac{1}{2} & & \\
1 & 0 & 0 & 1 & \\
\hline
 & \frac{1}{6} & \frac{1}{3} & \frac{1}{3} & \frac{1}{6}
\end{array}$$

In order to maintain the order $p = 4$ of the discrete RK method, when implementing formulae (3), (4), (5), it is sufficient to use an interpolant of order $p - 1 = 3$ in the current step $[t_n, t_{n+1}]$. It turns out that, as the interpolant of uniform order 3 in the current step $[t_n, t_{n+1}]$, it is convenient to consider the so-called *natural continuous extension* (NCE) of degree 3, introduced by Zennaro [5], defined as follows:

$$\hat{\eta}_{(0)}(t_n + \theta h_{n+1}) = y_n + h_{n+1} \sum_{i=1}^{4} b_i(\theta) K_{n+1}^i, \quad 0 \leq \theta \leq 1, \tag{7}$$

where

$$\begin{aligned} b_1(\theta) &= (\tfrac{2}{3}\theta^2 - \tfrac{3}{2}\theta + 1)\theta, \\ b_2(\theta) &= (-\tfrac{2}{3}\theta + 1)\theta^2, \\ b_3(\theta) &= (-\tfrac{2}{3}\theta + 1)\theta^2, \\ b_4(\theta) &= (\tfrac{2}{3}\theta - \tfrac{1}{2})\theta^2, \end{aligned}$$

and the stage values are

$$\begin{aligned} K_{n+1}^1 &= f(t_n, y_n, \eta(t_n - \tau(t_n, y_n))); \\ K_{n+1}^2 &= f(t_n + \tfrac{1}{2}h_{n+1}, y_n + \tfrac{1}{2}h_{n+1}K_{n+1}^1, \tilde{Y}_{n+1}^2), \end{aligned}$$

with

$$\begin{aligned} \tilde{Y}_{n+1}^2 &= \hat{\eta}_{(0)}(t_n + \tfrac{1}{2}h_{n+1} - \tau(t_n + \tfrac{1}{2}h_{n+1}, y_n + \tfrac{1}{2}h_{n+1}K_{n+1}^1)) \\ &= y_n + h_{n+1} \sum_{i=1}^{4} b_i\Big(\frac{1}{2} - \frac{\tau(t_n + \frac{1}{2}h_{n+1}, y_n + \frac{1}{2}h_{n+1}K_{n+1}^1)}{h_{n+1}}\Big) K_{n+1}^i \end{aligned}$$

if $t_n + \frac{1}{2}h_{n+1} - \tau(t_n + \frac{1}{2}h_{n+1}, y_n + \frac{1}{2}h_{n+1}K_{n+1}^1) > t_n$, and

$$\tilde{Y}_{n+1}^2 = \eta(t_n + \tfrac{1}{2}h_{n+1} - \tau(t_n + \tfrac{1}{2}h_{n+1}, y_n + \tfrac{1}{2}h_{n+1}K_{n+1}^1))$$

otherwise;

$$K_{n+1}^3 = f(t_n + \tfrac{1}{2}h_{n+1}, y_n + \tfrac{1}{2}h_{n+1}K_{n+1}^2, \tilde{Y}_{n+1}^3),$$

with

$$\begin{aligned} \tilde{Y}_{n+1}^3 &= \hat{\eta}_{(0)}(t_n + \tfrac{1}{2}h_{n+1} - \tau(t_n + \tfrac{1}{2}h_{n+1}, y_n + \tfrac{1}{2}h_{n+1}K_{n+1}^2)) \\ &= y_n + h_{n+1} \sum_{i=1}^{4} b_i\Big(\frac{1}{2} - \frac{\tau(t_n + \frac{1}{2}h_{n+1}, y_n + \frac{1}{2}h_{n+1}K_{n+1}^2)}{h_{n+1}}\Big) K_{n+1}^i \end{aligned}$$

if $t_n + \frac{1}{2}h_{n+1} - \tau(t_n + \frac{1}{2}h_{n+1}, y_n + \frac{1}{2}h_{n+1}K_{n+1}^2) > t_n$, and

$$\tilde{Y}_{n+1}^3 = \eta(t_n + \tfrac{1}{2}h_{n+1} - \tau(t_n + \tfrac{1}{2}h_{n+1}, y_n + \tfrac{1}{2}h_{n+1}K_{n+1}^2))$$

otherwise;

$$K_{n+1}^4 = f(t_n + h_{n+1}, y_n + h_{n+1}K_{n+1}^3, \tilde{Y}_{n+1}^4),$$

with

$$\begin{aligned} \tilde{Y}_{n+1}^4 &= \hat{\eta}_{(0)}(t_n + h_{n+1} - \tau(t_n + h_{n+1}, y_n + h_{n+1}K_{n+1}^3)) \\ &= y_n + h_{n+1} \sum_{i=1}^{4} b_i\Big(1 - \frac{\tau(t_n + h_{n+1}, y_n + h_{n+1}K_{n+1}^3)}{h_{n+1}}\Big) K_{n+1}^i \end{aligned}$$

if $t_n + h_{n+1} - \tau(t_n + h_{n+1}, y_n + h_{n+1}K^3_{n+1}) > t_n$, and

$$\tilde{Y}^4_{n+1} = \eta(t_n + h_{n+1} - \tau(t_n + h_{n+1}, y_n + h_{n+1}K^3_{n+1}))$$

otherwise.

Once we have computed the discrete approximation $y_{n+1} = \hat{\eta}_{(0)}(t_{n+1})$ obtained from (7) for $\theta = 1$, we do not consider the interpolant $\hat{\eta}_{(0)}(t)$ any more, but we switch to the *cubic Hermite interpolant* $\eta_{(0)}(t)$ at the end-points t_n and t_{n+1}, which also makes use of the extra stage

$$K^5_{n+1} = f(t_{n+1}, y_{n+1}, \tilde{Y}^5_{n+1}),$$

where

$$\begin{aligned} \tilde{Y}^5_{n+1} &= \hat{\eta}_{(0)}(t_{n+1} - \tau(t_{n+1}, y_{n+1})) \\ &= y_n + h_{n+1} \sum_{i=1}^{4} b_i \Big(1 - \frac{\tau(t_{n+1}, y_{n+1})}{h_{n+1}}\Big) K^i_{n+1} \end{aligned}$$

if $t_{n+1} - \tau(t_{n+1}, y_{n+1}) > t_n$, and

$$\tilde{Y}^5_{n+1} = \eta(t_{n+1} - \tau(t_{n+1}, y_{n+1}))$$

otherwise.

In other words, we consider the interpolant

$$\eta_{(0)}(t_n + \theta h_{n+1}) = d_1(\theta) y_n + d_2(\theta) y_{n+1} + d_3(\theta) h_{n+1} K^1_{n+1} + d_4(\theta) h_{n+1} K^5_{n+1},$$

where

$$\begin{aligned} d_1(\theta) &= (\theta - 1)^2 (2\theta + 1), \\ d_2(\theta) &= \theta^2 (3 - 2\theta), \\ d_3(\theta) &= \theta(\theta - 1)^2, \\ d_4(\theta) &= \theta^2 (\theta - 1). \end{aligned}$$

Note that, if overlapping does not occur in the current step, then K^5_{n+1} equals $K^1_{n+2} = f(t_{n+1}, y_{n+1}, \eta(t_{n+1} - \tau(t_{n+1}, y_{n+1})))$, which is available for free from the next step $[t_{n+1}, t_{n+2}]$.

In order to construct the continuous approximation $\eta(t)$ of order $q = 4$ for the advancing method in the current step $[t_n, t_{n+1}]$, we perform one uniform correction of $\eta_{(0)}(t)$. Such a procedure is a particular case of a more general approach for raising the uniform order, based on successive uniform corrections starting from a lower order interpolant (see [1, Chapter 5]).

To this aim, consider the extra abscissa $\theta_1 \in (0, 1)$, $\theta_1 \neq \frac{1}{2}$, and the extra stage

$$K^6_{n+1} = f(t_n + \theta_1 h_{n+1}, \eta_{(0)}(t_n + \theta_1 h_{n+1}), \tilde{Y}^6_{n+1}),$$

where

$$\tilde{Y}^6_{n+1} = \eta_{(0)}(t_n + \theta_1 h_{n+1} - \tau(t_n + \theta_1 h_{n+1}, \eta_{(0)}(t_n + \theta_1 h_{n+1})))$$

if $t_n + \theta_1 h_{n+1} - \tau(t_n + \theta_1 h_{n+1}, \eta_{(0)}(t_n + \theta_1 h_{n+1})) > t_n$, and

$$\tilde{Y}^6_{n+1} = \eta(t_n + \theta_1 h_{n+1} - \tau(t_n + \theta_1 h_{n+1}, \eta_{(0)}(t_n + \theta_1 h_{n+1})))$$

otherwise.

Thus $\eta(t)$ is defined as the fourth-degree Hermite–Birkhoff interpolant relevant, for the generic polynomial $P(\theta)$ of degree four, to the interpolation conditions on $P(0)$, $P(1)$, $P'(0)$, $P'(1)$ and

$P'(\theta_1)$. Note that $\theta_1 \neq \frac{1}{2}$ because the interpolation condition on $P'(\frac{1}{2})$ is linearly dependent on the other four conditions. We obtain

$$\begin{aligned}\eta(t_n + \theta h_{n+1}) &= d_1(\theta)y_n + d_2(\theta)y_{n+1} + d_3(\theta)h_{n+1}K^1_{n+1} \\ &\quad + d_4(\theta)h_{n+1}K^5_{n+1} + d_5(\theta)h_{n+1}K^6_{n+1},\end{aligned}$$

where

$$\begin{aligned}
d_1(\theta) &= \tfrac{1}{2\theta_1 - 1}(\theta - 1)^2(-3\theta^2 + 2(2\theta_1 - 1)\theta + 2\theta_1 - 1),\\
d_2(\theta) &= \tfrac{1}{2\theta_1 - 1}\theta^2(3\theta^2 - 4(\theta_1 + 1)\theta + 6\theta_1),\\
d_3(\theta) &= \tfrac{1}{2\theta_1(2\theta_1 - 1)}\theta(\theta - 1)^2((1 - 3\theta_1)\theta + 2\theta_1(2\theta_1 - 1)),\\
d_4(\theta) &= \tfrac{1}{2(\theta_1 - 1)(2\theta_1 - 1)}\theta^2(\theta - 1)((2 - 3\theta_1)\theta + \theta_1(4\theta_1 - 3)),\\
d_5(\theta) &= \tfrac{1}{2\theta_1(2\theta_1 - 1)(\theta_1 - 1)}\theta^2(\theta - 1)^2.
\end{aligned}$$

In order to estimate the discrete local error by means of an error-estimating method of higher order $p' = p + 1 = 5$, we construct an approximation $\tilde{y}_{n+1}$ by using a suitable quadrature formula. More precisely, we consider the Lobatto abscissae $\pi_1 = \frac{5-\sqrt{5}}{10}$ and $\pi_2 = \frac{5+\sqrt{5}}{10}$ and the additional stages

$$K^7_{n+1} = f(t_n + \pi_1 h_{n+1}, \eta(t_n + \pi_1 h_{n+1}), \tilde{Y}^7_{n+1}),$$

where

$$\tilde{Y}^7_{n+1} = \eta(t_n + \pi_1 h_{n+1} - \tau(t_n + \pi_1 h_{n+1}, \eta(t_n + \pi_1 h_{n+1})))$$

if $t_n + \pi_1 h_{n+1} - \tau(t_n + \pi_1 h_{n+1}, \eta(t_n + \pi_1 h_{n+1})) > t_n$, and

$$\tilde{Y}^7_{n+1} = \eta(t_n + \pi_1 h_{n+1} - \tau(t_n + \pi_1 h_{n+1}, \eta(t_n + \pi_1 h_{n+1})))$$

otherwise;

$$K^8_{n+1} = f(t_n + \pi_2 h_{n+1}, \eta(t_n + \pi_2 h_{n+1}), \tilde{Y}^8_{n+1}),$$

where

$$\tilde{Y}^8_{n+1} = \eta(t_n + \pi_2 h_{n+1} - \tau(t_n + \pi_2 h_{n+1}, \eta(t_n + \pi_2 h_{n+1})))$$

if $t_n + \pi_2 h_{n+1} - \tau(t_n + \pi_2 h_{n+1}, \eta(t_n + \pi_2 h_{n+1})) > t_n$, and

$$\tilde{Y}^8_{n+1} = \eta(t_n + \pi_2 h_{n+1} - \tau(t_n + \pi_2 h_{n+1}, \eta(t_n + \pi_2 h_{n+1})))$$

otherwise.

Then the quadrature formula is

$$\tilde{y}_{n+1} = y_n + h_{n+1}(\tfrac{1}{12}K^1_{n+1} + \tfrac{5}{12}K^7_{n+1} + \tfrac{5}{12}K^8_{n+1} + \tfrac{1}{12}\bar{K}^5_{n+1}),$$

where, as we have already observed,

$$\bar{K}^5_{n+1} = K^1_{n+2} = f(t_{n+1}, y_{n+1}, \eta(t_{n+1} - \tau(t_{n+1}, y_{n+1})))$$

may or may not equal the already available stage K^5_{n+1}.

As is usually done, we estimate the discrete local error per unit step by

$$\tilde{\sigma}_{n+1} = \|\tilde{y}_{n+1} - y_{n+1}\| / h_{n+1}.$$

The current step is accepted if and only if

$$\tilde{\sigma}_{n+1} \leq TOL, \tag{8}$$

where TOL is the user-supplied tolerance.

The new stepsize h_{new}, to be used either as h_{n+2} in the next step $[t_{n+1}, t_{n+2}]$ if (8) is satisfied or, again, as h_{n+1} in the current step $[t_n, t_{n+1}]$ if (8) is violated, is computed by using the common formula

$$h_{\text{new}} = \max\Big\{\omega_{\min}, \min\Big\{\omega_{\max}, \rho\Big(\frac{TOL}{\tilde{\sigma}_{n+1}}\Big)^{1/4}\Big\}\Big\} \cdot h_{n+1}, \tag{9}$$

where $\omega_{\min} \in (0,1)$, $\omega_{\max} > 1$ and $\rho \in (0,1)$ are suitable *safety factors.*

A maximum number of rejections per step may be fixed by the user as well.

Since the uniform order is $q = p$, controlling the discrete local error is sufficient for the stepsize selection. Nevertheless, for more reliability, we also propose an additional optional control of the uniform local error Σ_{n+1} by a continuous error-estimating method of lower order $q' = q - 1 = 3$, by employing local extrapolation. More precisely, we consider the computable quantity

$$\hat{\Sigma}_{n+1} = \max_{0 \le \theta \le 1} \|\eta_{(0)}(t_n + \theta h_{n+1}) - \eta(t_n + \theta h_{n+1})\|. \tag{10}$$

It can be seen that

$$\Sigma_\Delta = O(TOL) \tag{11}$$

is assured if we impose

$$h_{n+1}\hat{\Sigma}_{n+1} \le TOL. \tag{12}$$

Note that, by (6), relation (11) is necessary to ensure error-tolerance proportionality.

The test (12) is made whenever (8) holds. If (12) is not satisfied, then we reduce the stepsize by using the formula

$$h_{\text{new}2} = \max\Big\{\omega_{\min}, \rho\Big(\frac{TOL}{h_{n+1}\hat{\Sigma}_{n+1}}\Big)^{1/5}\Big\} \cdot h_{n+1} \tag{13}$$

and repeat the computations in the current step $[t_n, t_{n+1}]$ with $h_{n+1} = h_{\text{new}2}$.

In any case, even if (8) is satisfied, the stepsize $h_{\text{new}2}$ suggested by (13) is used to bound the next step h_{n+2}, so that (9) is substituted by

$$h_{n+2} = \max\Big\{\omega_{\min}, \min\Big\{\omega_{\max}, \rho\Big(\frac{TOL}{\tilde{\sigma}_{n+1}}\Big)^{1/4}, \rho\Big(\frac{TOL}{h_{n+1}\hat{\Sigma}_{n+1}}\Big)^{1/5}\Big\}\Big\} \cdot h_{n+1}.$$

In order to compute $\hat{\Sigma}_{n+1}$ from (10), we set

$$\delta(\theta) = \eta_{(0)}(t_n + \theta h_{n+1}) - \eta(t_n + \theta h_{n+1}).$$

It turns out that

$$\begin{aligned}\delta(\theta) \quad = \quad & h_{n+1}\frac{\theta^2(\theta-1)^2}{2(2\theta_1 - 1)}\Big(\frac{2\theta_1 - 1}{\theta_1}K^1_{n+1} - (2K^2_{n+1} + 2K^3_{n+1} + K^4_{n+1}) \\ & + \frac{3\theta_1 - 2}{\theta_1 - 1}K^5_{n+1} + \frac{1}{\theta_1(\theta_1 - 1)}K^6_{n+1}\Big),\end{aligned}$$

so that

$$\begin{aligned}\hat{\Sigma}_{n+1} \quad & = \quad \|\delta(\tfrac{1}{2})\| \\ & = \quad \frac{h_{n+1}}{32|2\theta_1 - 1|}\Big\|\frac{2\theta_1 - 1}{\theta_1}K^1_{n+1} - (2K^2_{n+1} + 2K^3_{n+1} + K^4_{n+1}) \\ & \quad + \frac{3\theta_1 - 2}{\theta_1 - 1}K^5_{n+1} + \frac{1}{\theta_1(\theta_1 - 1)}K^6_{n+1}\Big\|.\end{aligned}$$

4 A numerical experiment

We conclude by applying the method introduced in the previous section to the constructed equation

$$\begin{cases} y'(t) = -y(t) + y\left(\alpha(t)\right) + \frac{t}{20}\cos\left(\frac{t}{20}\right) + \sin\left(\frac{t}{20}\right) - \sin\left(\frac{\alpha(t)}{20}\right), & 0 \le t \le 1000, \\ y(t) = \sin\left(\frac{t}{20}\right), \quad -20 \le t \le 0, \end{cases} \tag{14}$$

with variable (non state-dependent) deviated argument $\alpha(t) = t-1+\sin(t)$ corresponding to a delay $\tau(t) = 1 - \sin(t)$ which periodically vanishes at the points $\frac{\pi}{2} + 2k\pi$, k positive integer. The exact solution $y(t) = \sin\left(\frac{t}{20}\right)$ is smooth and hence no breaking points have to be included as mesh points for preserving the piecewise regularity of the solution and hence the accuracy of the methods. Remark that overlapping occurs in the proximity of every point where the delay vanishes, and possibly somewhere else. Being the exact solution available, the example is suitable to illustrate that, according to the theory, the local error estimation as well as the expected proportionality with the given tolerance are succesfully obtained also in presence of overlapping.

For the actual implementation of the method described in Section 3, we have chosen $\theta_1 = \frac{1}{3}$ for the computation of K^6_{n+1}, and $\omega_{\min} = 0.5$, $\omega_{\max} = 1.5$, $\rho = 0.9$ in (9) and (13).

Since $q = p$, the error-tolerance proportionality is guaranteed under the sole discrete error control $\sigma_n \le TOL$ (see Table 1 where $\sigma-$rejections stands for the number of steps rejected by the test on the discrete error). In Table 2 we provide the results obtained by performing also the continuous error control $\Sigma_n \le TOL$. The term Σ-rejections stands for the number of steps rejected by the test on the continuous error control. As expected from the theory, for sufficiently small tolerance TOL, the continuous error control does not work any more and the global error is essentially determined by the discrete error controller. However, the estimation of the uniform error Σ_n is taken into account in the prediction of the next stepsize. This is why the results in the last three rows of Table 1 and Table 2 are not just identical.

TOL	$\sigma-$rejections	steps	overlappings	$\frac{e_N}{TOL}$
10^{-1}	5	17	17	32.17
10^{-2}	108	200	200	25.34
10^{-3}	210	429	429	26.72
10^{-4}	208	687	430	11.03
10^{-5}	253	1211	523	14.50
10^{-6}	280	2120	649	20.28
10^{-7}	265	3741	770	20.12

Table 1. Numerical results for equation (14) with the sole discrete test of tolerance.

TOL	$\sigma-$rejections	$\Sigma-$rejections	steps	overlappings	$\frac{e_N}{TOL}$
10^{-1}	2	158	276	276	1.47
10^{-2}	1	238	415	415	2.86
10^{-3}	16	238	573	462	3.61
10^{-4}	128	112	780	446	10.38
10^{-5}	209	0	1196	509	16.40
10^{-6}	235	0	2121	650	15.64
10^{-7}	253	0	3740	771	17.19

Table 2. Numerical results for equation (14) with the discrete and uniform test of tolerance.

References

[1] A. Bellen, M. Zennaro, *Numerical Methods for Delay Differential Equations*, Numerical Mathematics and Scientific Computation Series, Oxford University Press, 2003.

[2] J.R. Dormand and P.J. Prince, A family of embedded Runge–Kutta formulae, *J. Comput. Appl. Math.* **6** (1980), 19–26

[3] N. Guglielmi, E. Hairer, Implementing Radau IIA methods for stiff delay differential equations, *Computing* **67** (2001), 1–12

[4] E. Hairer, S.P. Nørsett and G. Wanner, *Solving Ordinary Differential Equations I, Nonstiff Problems*, Springer-Verlag, Berlin, 1993.

[5] M. Zennaro, Natural Continuous Extensions of Runge-Kutta methods, *Math. Comput.* **46** (1986), 119–133