Journal of Computational and Applied Mathematics 279 (2015) 40–56

Contents lists available at ScienceDirect

Journal of Computational and Applied Mathematics

journal homepage: www.elsevier.com/locate/cam

Numerical solution of threshold problems in epidemics and population dynamics

Z. Bartoszewski[a,*], Z. Jackiewicz[b,c], Y. Kuang[b]

[a] Department of Applied Physics and Mathematics, Gdańsk University of Technology, 80-233 Gdańsk, Poland

[b] School of Mathematical and Statistical Sciences, Arizona State University, Tempe, AZ 85287-1804, United States

[c] Department of Applied Mathematics, AGH University of Science and Technology, 30-059 Kraków, Poland

ARTICLE INFO

Article history:
Received 17 April 2014
Received in revised form 15 September 2014

MSC:
65L03
65L05
65L06
65L20
65R20

Keywords:
Delay-differential equations
Threshold conditions
Continuous Runge–Kutta methods
Bisection method
Local error estimation
Convergence analysis

ABSTRACT

A new algorithm is proposed for the numerical solution of threshold problems in epidemics and population dynamics. These problems are modeled by the delay-differential equations, where the delay function is unknown and has to be determined from the threshold conditions. The new algorithm is based on embedded pair of continuous Runge–Kutta method of order $p = 4$ and discrete Runge–Kutta method of order $q = 3$ which is used for the estimation of local discretization errors, combined with the bisection method for the resolution of the threshold condition. Error bounds are derived for the algorithm based on continuous one-step methods for the delay-differential equations and arbitrary iteration process for the threshold conditions. Numerical examples are presented which illustrate the effectiveness of this algorithm.

1. Introduction

Denote by $C = C([\alpha, T], \mathbb{R}^m)$ the space of continuous functions from the interval $[\alpha, T]$ into $\mathbb{R}^m$ with the norm defined by

$$|||y|||_{[\alpha,T]} := \sup\left\{ e^{-a(t-t_0)} \|y\|_{[\alpha,t]} : \alpha \le t \le T \right\},$$

where $a > 0$ is a real parameter and $\|y\|_{[\alpha,t]}$ is the uniform norm on the interval $[\alpha, t]$. Let

$$f : [t_0, T] \times C\big([\alpha, T], \mathbb{R}^m\big) \times \mathbb{R}^m \to \mathbb{R}^m$$

and consider the initial-value problem for a functional-differential equation of the form

$$\begin{cases} y'(t) = f\Big(t, y(\cdot), y\Big(t - \tau\big(t, y(\cdot)\big)\Big)\Big), & t \in [t_0, T], \\ y(t) = g(t), & t \in [\alpha, t_0], \end{cases} \tag{1.1}$$

* Corresponding author.
E-mail addresses: zbart@pg.gda.pl (Z. Bartoszewski), jackiewicz@asu.edu (Z. Jackiewicz), kuang@asu.edu (Y. Kuang).

http://dx.doi.org/10.1016/j.cam.2014.10.020
0377-0427/

$\alpha \le t_0$. Problem (1.1) is a generalization of the problem elaborated in [1], where discrete variable methods for its numerical solution are investigated. The function $\tau(t, y(\cdot))$ appearing in this equation is determined from the so-called threshold condition

$$P\Big(t, y(\cdot), \tau\big(t, y(\cdot)\big)\Big) = m, \tag{1.2}$$

with given threshold $m > 0$. Here,

$$P : [t_0, T] \times C\big([\alpha, T], \mathbb{R}^m\big) \times \mathbb{R} \to \mathbb{R}$$

is a given operator. In applications P is usually an integral operator. Observe that (1.2) depends on the unknown function y. The solution to (1.1)–(1.2) will be denoted by $y(t)$ and $\tau(t, y(\cdot))$. Such equations find applications in modeling various problems in epidemics and population dynamics. Specific examples of such problems are presented in Section 2. The existence and uniqueness of the solution to (1.1)–(1.2) are discussed in Section 3.

Since in general, the operator P cannot be computed exactly (P may be an integral operator like in the applications introduced in the next section) we first replace (1.2) by the equation

$$\bar{P}\Big(t, y(\cdot), \tau\big(t, y(\cdot)\big)\Big) = m, \tag{1.3}$$

where $\bar{P}$ is a discrete approximation to P. Next, we describe the numerical approximation to the solution of (1.1), (1.3). Denote by $\bar{y}(t)$ and $\bar{\tau}(t, \bar{y}(\cdot))$ the solution to (1.1), (1.3). To compute numerical approximation $\bar{y}_h$ to $\bar{y}$ we consider the general class of continuous one-step methods of the form

$$\begin{cases} \bar{y}_h(t_n + \theta h_n) = \bar{y}_h(t_n) + h_n \Phi_h\Big(t_n, \theta, \bar{y}_h(\cdot), \bar{y}_h\Big(t_{n+\theta} - \bar{\tau}_h\big(t_{n+\theta}, \bar{y}_h(\cdot)\big)\Big)\Big), & \theta \in (0, 1], \\ \bar{y}_h(t) = g_h(t), & t \in [\alpha, t_0], \end{cases} \tag{1.4}$$

$n = 0, 1, \ldots, N$. Here, $t_{n+1} = t_n + h_n$, $t_{n+\theta} = t_n + \theta h_n$, $n = 0, \ldots, N$, $\theta \in (0, 1]$, with step-sizes h_n which satisfy

$$\sum_{n=0}^{N-1} h_n < T - t_0 \le \sum_{n=0}^{N} h_n. \tag{1.5}$$

Moreover, g_h is an approximation to the initial function g, and $\bar{\tau}_h(t_{n+\theta}, \bar{y}_h(\cdot))$ is an approximation to the solution $\bar{\tau}(t_{n+\theta}, \bar{y}_h(\cdot))$ to the operator equation

$$\bar{P}\Big(t_{n+\theta}, \bar{y}_h(\cdot), \bar{\tau}\big(t_{n+\theta}, \bar{y}_h(\cdot)\big)\Big) = m, \tag{1.6}$$

obtained from (1.3) by replacing t by $t_{n+\theta}$, $y(\cdot)$ by $\bar{y}_h(\cdot)$ and $\tau(t, y(\cdot))$ by $\bar{\tau}(t_{n+\theta}, \bar{y}_h(\cdot))$. In this formulation the increment function Φ_h and the operator equation (1.6) depend on $\bar{y}_h(\cdot)$ and $\bar{y}_h(t_{n+\theta} - \bar{\tau}_h(t_{n+\theta}, \bar{y}_h(\cdot)))$ although in practical applications this dependence is usually restricted to a discrete set of values such as, for example, $\bar{y}_h(t_{n+c_i})$ and $\bar{y}_h(t_{n+c_i} - \bar{\tau}_h(t_{n+c_i}, \bar{y}_h(\cdot)))$, $i = 1, 2, \ldots, s$, where c_i are given abscissas usually chosen from the interval $[0, 1]$. This is the case for continuous Runge–Kutta methods considered in Section 5.

Depending on the form of the increment function Φ_h the formulation (1.4) includes both the explicit and implicit formulas for (1.1), (1.3). Note that, although $\bar{y}_h(t_{n+\theta})$ is computed from (1.4) and the quantity $\bar{\tau}\big(t_{n+\theta}, \bar{y}_h(\cdot)\big)$ is computed from (1.6), these equations are not independent and (1.6) has to be resolved at each time step of numerical integration for the method (1.4).

We are interested to estimate the global error $y - \bar{y}_h$, where y is the solution to (1.1) with $\tau(t, y(\cdot))$ given by (1.2), and $\bar{y}_h$ is computed from (1.4) with the approximation $\bar{\tau}_h(t_n, \bar{y}_h(\cdot))$ to $\bar{\tau}(t_n, \bar{y}_h(\cdot))$ computed by some iterative procedure applied to Eq. (1.6). This error consists of two parts:

$$y - \bar{y}_h = \big(y - \bar{y}\big) + \big(\bar{y} - \bar{y}_h\big),$$

and we have

$$\big\|y - \bar{y}_h\big\|_{[\alpha,T]} \le \big\|y - \bar{y}\big\|_{[\alpha,T]} + \big\|\bar{y} - \bar{y}_h\big\|_{[\alpha,T]}. \tag{1.7}$$

Here, for $x \in C([\alpha, t], \mathbb{R}^m)$ and $t \in [\alpha, T]$ the norm $\|x\|_{[\alpha,t]}$ is defined by

$$\big\|x\big\|_{[\alpha,t]} := \sup\Big\{\big\|x(s)\big\| : \alpha \le s \le t\Big\},$$

where $\|\cdot\|$ is any norm on $\mathbb{R}^m$. The first term on the right hand side of the above inequality will be investigated in Section 4 using the theory of integral inequalities. The second term on the right hand side of (1.7) will be investigated in Section 5 using the generalization of the theory of one-step methods for functional differential equations. In Section 6 we describe the adaptation of continuous Runge–Kutta methods for ordinary differential equations to the problem (1.1), (1.3). In Section 7 we describe the numerical algorithm for the solution of (1.1), (1.3) based on embedded pair of continuous Runge–Kutta methods of order $p = 4$ and discrete method of order $q = p - 1 = 3$ which is used for error estimation. In this section we also describe the resolution of the threshold condition (1.6) for specific form of the operator $\bar{P}$. In Section 8 the results of numerical experiments will be presented using the examples of threshold problems presented in Section 2. Finally, in Section 9 some concluding remarks are given.

2. Examples of threshold problems

The first problem comes from the theory of epidemics. Assume that in a constant population we have $I(t)$ infectives and $S(t)$ susceptibles at time $t \geq 0$ with a known history of the number of infectives $I_0(t)$ prior to the time $t = 0$. Specifically it is assumed that the history $I_0(t)$ is described by a continuous increasing function on the interval $-\sigma \leq t < 0$ with $I_0(-\sigma) = 0$ and $I_0(0) = I_0$. Hoppensteadt and Waltman [2] considered the model of spread of infection under the following assumptions:

1. An individual who becomes infectious at time t fully recovers at time $t + \sigma$ and infectives I_0 added to the population at time $t = 0$ are also subject to this rule.
2. The rate of exposure of susceptibles to infectives is proportional to the number of infectives $I(t)$ and susceptibles $S(t)$ with a known proportionality function $r(t) > 0$.
3. The individual becomes infectious at time t after the accumulated dosage of infection $\int_{\widetilde{\tau}}^{t} \rho(s)I(s)ds$ reaches a known threshold $m > 0$. Here, $\rho(s) > 0$ is a known proportionality function and $\widetilde{\tau} = \widetilde{\tau}(t, I(\cdot))$ is an unknown function which depends on the history of $I(s)$ for $s \leq t$.

As explained in [2] these assumptions lead to the following model for the spread of infection

$$\begin{cases} S'(t) = -r(t)I(t)S(t), & t \geq 0, \\ S(0) = S_0, \end{cases} \tag{2.1}$$

$$I(t) = \begin{cases} I_0(t), & -\sigma \leq t \leq t_0, \\ I_0(t) + S_0 - S\Big(\widetilde{\tau}\big(t, I(\cdot)\big)\Big), & t_0 \leq t \leq t_0 + \sigma, \\ S\Big(\widetilde{\tau}\big(t-\sigma, I(\cdot)\big)\Big) - S\Big(\widetilde{\tau}\big(t, I(\cdot)\big)\Big), & t_0 + \sigma \leq t < \infty. \end{cases} \tag{2.2}$$

Here, $I_0(t)$ which is already defined on the interval $[-\sigma, 0)$ is extended on the interval $[0, \infty)$ by the formula

$$I_0(t) = \begin{cases} I_0(0) - I_0(t-\sigma), & 0 \leq t \leq \sigma, \\ 0, & \sigma < t < \infty, \end{cases}$$

and $0 < t_0 < \sigma$ is a unique time such that

$$\int_0^{t_0} \rho(s)I_0(s)ds = m. \tag{2.3}$$

It follows from [2] that in this model the function $\tau(t) := \widetilde{\tau}(t, I(\cdot))$ appearing in (2.2) is determined from the threshold condition

$$P\big(t, I(\cdot), \tau(t)\big) := \int_{\tau(t)}^{t} \rho(s)I(s)ds = m. \tag{2.4}$$

It was proved in [2] that the function $\tau(t)$ is continuously differentiable. As a consequence, we can reformulate the threshold condition (2.4) for $t \geq t_0$ in the differential form. This can be done as follows. Taking the derivative of (2.4) with respect to t we obtain

$$\frac{d}{dt}\int_{\tau(t)}^{t} \rho(s)I(s)ds = 0.$$

This is equivalent to

$$\frac{d}{dt}\left(\int_a^t \rho(s)I(s)ds - \int_a^{\tau(t)} \rho(s)I(s)ds\right) = 0,$$

where a is some constant. This leads to

$$\rho(t)I(t) - \tau'(t)\rho\big(\tau(t)\big)I\big(\tau(t)\big) = 0.$$

Observe also that

$$\int_{\tau(t_0)}^{t_0} \rho(s)I(s)ds = \int_{\tau(t_0)}^{t_0} \rho(s)I_0(s)ds = m,$$

and comparing this with (2.3) it follows that $\tau(t_0) = 0$. The above arguments lead to the initial value problem for $\tau(t)$ of the form

$$\begin{cases} \tau'(t) = \dfrac{\rho(t)I(t)}{\rho\big(\tau(t)\big)I\big(\tau(t)\big)}, & t \ge t_0, \\ \tau(t_0) = 0. \end{cases} \tag{2.5}$$

Here, $\tau'(t)$ is the total derivative of the function $\tau(t) = \widetilde{\tau}\big(t, I(\cdot)\big)$ with respect to t. The problem (2.1), (2.2) and (2.5) is a system of delay differential equations where the delay function $\tau(t)$ itself depends on the values of the unknown solution $I(s)$ for $s \le t$. This system was solved in [2] by imposing some rather restrictive conditions on the parameters of the model. These assumptions allowed them to reduce (2.1), (2.2), (2.5) to the system of difference–differential equations with constant delay which could then be solved numerically by the method of steps. A new more general variable stepsize variable order algorithm which is applicable to this system without any simplifying conditions on the parameters of the model was proposed recently in [3]. This algorithm is based on the Nordsieck representation of the family of diagonally implicit multistage integration methods of stage order equal to its order p for $1 \le p \le 4$. These methods were constructed in [4,5] and their Nordsieck representation developed in [6]. In this paper we propose a new algorithm for the numerical solution of (2.1)–(2.4) based on continuous Runge–Kutta method of order $p = 4$ constructed by Owren and Zennaro [7] for the integration of (2.1)–(2.2) and the resolution of the threshold condition (2.4) by the bisection method. This problem was also solved by Thompson and Shampine [8], where the threshold time was determined by using an event function of Matlab `dde23` solver for delay-differential equations [9].

The second example which comes from population dynamics is the predator–prey model with combined result of death and birth processes linked to the dynamic resources (population of prey). This model was recently proposed by Gourley and Kuang [10,11] and extends the previous work by Aiello and Freedman [12]. Denote by $x(t)$ the population of prey at time t and $y_j(t)$ and $y(t)$ the population of juvenile and adult predators. The model proposed in [11] takes the form

$$\begin{cases} x'(t) = \dfrac{r}{K}x(t)\big(1 - x(t)\big) - y(t)p\big(x(t)\big), \\ y'(t) = b\, e^{-d_j\tau\left(t,x(\cdot)\right)} y\Big(t - \tau\big(t, x(\cdot)\big)\Big) p\bigg(x\Big(t - \tau\big(t, x(\cdot)\big)\Big)\bigg) - d\, y(t), \\ y_j'(t) = b\, y(t)p\big(x(t)\big) - b\, e^{-d_j\tau\left(t,x(\cdot)\right)} y\Big(t - \tau\big(t, x(\cdot)\big)\Big) p\bigg(x\Big(t - \tau\big(t, x(\cdot)\big)\Big)\bigg) - d_j\, y_j(t), \\ x(t) = x_0(t), \quad t \in [\alpha, 0], \\ y(t) = y_0(t), \quad t \in [\alpha, 0], \end{cases} \tag{2.6}$$

$t \ge 0, \alpha \le 0$, where the given initial functions $x_0(t)$ and $y_0(t)$ are nonnegative and continuous on $\alpha \le t < 0$, and $x(0), y(0), y_j(0) > 0$. Observe that this model does not require the knowledge of the past history of y_j so this function does not need to be specified on the initial interval $[\alpha, 0]$. In the above system, r is the specific growth rate of the prey, K is its caring capacity, and the (given) function $p(x)$ is the adult predators' functional response. The parameters b and d are the adult predators' birth and death rates, respectively. In addition, it is assumed that juveniles suffer a mortality rate of d_j (the through-stage death rate) and take $\tau(t) = \tau(t, x(\cdot))$ units of time to mature. This delay function $\tau(t) = \tau(t, x(\cdot))$ which depends on the past history $x(s), s \le t$, of population of prey is determined from the threshold condition

$$P\Big(t, x(\cdot), \tau\big(t, x(\cdot)\big)\Big) = \int_{t-\tau\left(t,x(\cdot)\right)}^{t} p\big(x(s)\big)ds = m, \tag{2.7}$$

where $m > 0$ is a given threshold. The function $p(x)$ appearing in (2.6) and (2.7) is assumed to be differentiable, strictly increasing, and such that $p(0) = 0$ and $p(x)/x$ is bounded for all $x > 0$. Examples of such function relevant to this model are $p(x) = px$, where p is a constant, $p > 0$, and $p(x) = px/(1 + ax)$, where $a, p > 0$. The positivity preservation result for $x(t)$ and $y(t)$ was proved in [10,11] that if the initial functions $x_0(t), y_0(t)$ are nonnegative for $\alpha \le t < 0, x(0), y(0) > 0$, and the function $p(x)$ satisfy the conditions given above then the solutions $x(t)$ and $y(t)$ to (2.6) are positive for all $t > 0$. The results of numerical simulations for the problem (2.6) with constant delay τ are presented in [10].

Assuming that the function $\tau(t) = \tau(t, x(\cdot))$ is continuously differentiable we can reformulate (2.7) as

$$\tau'(t) = 1 - \frac{p\big(x(t)\big)}{p\big(x\big(t - \tau(t)\big)\big)}, \tag{2.8}$$

where $\tau(0)$ is determined from the condition

$$\int_{-\tau(0)}^{0} p\big(x(s)\big)ds = m.$$

Observe that $\tau'(t)$ is defined and equal to 0 if $\tau(t) = 0$. Then (2.5), (2.8) could be numerically solved using methods for the system of delay-differential equations [13,14]. However, this may lead to stiff differential equation (2.8) if $\tau(t)$ has sharp gradients. In this paper we pursue a different approach which avoids this problem, and we will develop an algorithm based on continuous Runge–Kutta method of order $p = 4$, where the function $\tau(t)$ will be determined directly from the threshold condition (2.7) by the bisection method. This new approach is applicable even if the function $\tau(t)$ is not differentiable.

3. Existence and uniqueness

Denote by $L([\alpha, T], \mathbb{R}^m)$ the space of Lipschitz continuous functions from $[\alpha, T]$ into $\mathbb{R}^m$ with the norm defined by

$$|||y|||_{[\alpha,T]} := \sup\Big\{e^{-a(t-t_0)}\,\|y\|_{[\alpha,t]} : \alpha \le t \le T\Big\},$$

where $a > 0$ is a real parameter which will be determined later. Assume that the function f in (1.1) is continuous, and satisfies the Lipschitz condition of the form

$$\Big\|f\Big(t, y_1(\cdot), u_1\Big) - f\Big(t, y_2(\cdot), u_2\Big)\Big\| \le L_1\Big\|y_1 - y_2\Big\|_{[\alpha,t]} + L_2\Big\|u_1 - u_2\Big\| \tag{3.1}$$

with constants $L_1, L_2 \ge 0$ for $t \in [\alpha, T]$, $y_1, y_2 \in C([\alpha, T], \mathbb{R}^m)$, and $u_1, u_2 \in \mathbb{R}^m$. Assume also that there exists a solution operator S for (1.2)

$$\tau\big(t, y(\cdot)\big) = S\big(t, y(\cdot), m\big), \tag{3.2}$$

which satisfies the Lipschitz condition

$$\Big|S\big(t, y_1(\cdot), m\big) - S\big(t, y_2(\cdot), m\big)\Big| \le L_S\Big\|y_1 - y_2\Big\|_{[\alpha,t]}, \tag{3.3}$$

with $L_S \ge 0$ for $t \in [\alpha, T]$, and $y_1, y_2 \in C([\alpha, T], \mathbb{R}^m)$, and the condition

$$0 \le \tau\big(t, y(\cdot)\big) \le t - \alpha, \quad t \in [t_0, T]. \tag{3.4}$$

Observe that the inequality $0 \le \tau(t, y(\cdot))$ in (3.4) implies that f is a Volterra operator, i.e., it depends only on the past history of the solution y, and the inequality $\tau(t, y(\cdot)) \le t - \alpha$ in (3.4) implies that (1.1) is well defined with initial function g specified on the interval $[\alpha, t_0]$.

The statement and the proof of the following existence and uniqueness theorem are more or less a direct extension of the basic existence and uniqueness results for functional differential equations [15–17]. Observe that finding a solution of Eq. (1.1) with $\tau(t, y(\cdot))$ given by (3.2) is equivalent to solving the integral equation

$$\begin{cases} y(t) = y(t_0) + \displaystyle\int_{t_0}^{t} f\Big(s, y(\cdot), y\Big(s - \tau\big(s, y(\cdot)\big)\Big)\Big)ds, & t \in [t_0, T], \\ y(t) = g(t), & t \in [\alpha, t_0]. \end{cases} \tag{3.5}$$

To prove the existence of a solution through a point (t_0, g) from $[t_0, t_0+A] \times C([\alpha, t_0+A], \mathbb{R}^m)$ for some $A > 0$, we consider the set of all functions y on $[\alpha, t_0 + A]$ that are continuous and coincide with g on $[\alpha, t_0]$. The values of these functions on $[t_0, t_0+A]$ are required to satisfy $\|y(t)-g(t_0)\| \le B$ for some $B > 0$. The solution mapping F obtained from the corresponding integral equation can be defined, and A and B can be chosen so that F maps this class into itself and is *completely continuous* (that is, the mapping F is continuous and takes closed bounded sets into compact sets). Thus, *Schauder's fixed point theorem* (which asserts that if K is a convex subset of a Banach space V and F is a continuous mapping of K into itself so that $F(K)$ is contained in a compact subset of K, then F has a fixed point) implies the existence of a solution. We have the following theorem.

Theorem 3.1. *Assume that the function f is continuous, it satisfies (3.1) and there exists a solution operator S given by (3.2) satisfying (3.3) and (3.4). Then the problem (1.1) with $\tau(t, y(\cdot))$ given by (3.2) has a unique solution $y \in L([\alpha, T], \mathbb{R}^m)$ for any $T \ge t_0$.*

Proof. We present below a proof for the case of $m = 1$. The case for $m > 1$ is the same conceptually with only notational differences.

We consider first the existence. Let $C = C([\alpha, t_0 + A], \mathbb{R})$ and define the set

$$K = \Big\{\phi : \phi \in C,\ \phi(t) = g(t),\ t \in [\alpha, t_0],\ \text{and } |\phi(t) - g(t_0)| \le B \text{ for } t \in [t_0, t_0 + A]\Big\}.$$

Observe that C is a Banach space with the norm defined in Section 1 and K is closed and convex. Based on (3.5), we define a continuous mapping of K into a set in C as follows

$$F(\phi)(t) = \begin{cases} g(t), & t \in [\alpha, t_0], \\ g(t_0) + \displaystyle\int_{t_0}^{t} f\Big(s, y(\cdot), y\Big(s - \tau\big(s, y(\cdot)\big)\Big)\Big)ds, & t \in [t_0, t_0 + A]. \end{cases}$$

Since f satisfies the Lipschitz condition on C and K is bounded, it follows that the set $F(K)$ is compact. We will show next that if A is small enough, then F maps K into itself. Let

$$M = \max\left\{\left|f\left(s, \phi(\cdot), \phi\left(s - \tau\left(s, \phi(\cdot)\right)\right)\right)\right| : \phi \in K\right\},$$

and

$$G = \max\{|g(t)| : t \in [\alpha, t_0]\}.$$

Clearly both G and M are finite. Observe that

$$\left|F(\phi)(t) - g(t_0)\right| = \left|\int_{t_0}^{t} f\left(s, y(\cdot), y\left(s - \tau\left(s, y(\cdot)\right)\right)\right)ds\right| \le MA$$

for $t \in [t_0, t_0 + A]$. It is easy to see that if we select A so that $MA < B$, then $F(K)$ is a subset of K. The Lipschitz condition on f (3.1) implies that

$$M \le (L_1 + L_2)(G + B) + \left|f\left(t_0, g(\cdot), g\left(s - \tau\left(s, g(\cdot)\right)\right)\right)\right|.$$

Let $B = G$ and

$$D = 2(L_1 + L_2)B + \left|f\left(t_0, g(\cdot), g\left(s - \tau\left(s, g(\cdot)\right)\right)\right)\right|.$$

We see that $MA < B$ is true if $A < G/D$. Choosing $A = G/(2D)$ we can see that the initial value problem (1.1) with $\tau(t, y(\cdot))$ given by (3.2) has a solution $y \in L([\alpha, t_0 + A], \mathbb{R})$. By simple induction argument, the solution exists on $[\alpha, t_0 + nA]$ for all positive integers n. This shows that the solution exists for all $t > t_0$.

We now consider the uniqueness. Assume that the initial value problem (1.1) with $\tau(t, y(\cdot))$ given by (3.2) has two solutions y_1 and y_2 in $L([\alpha, t_0 + A], \mathbb{R})$. Put

$$F_2 = \max\left\{y_2'(t) : t \in [t_0, t_0 + A]\right\}.$$

Let $v = y_1 - y_2$ and

$$V(t) = \|v\|_{[\alpha,t]} := \sup\left\{\|v\|_{[\alpha,s]} :\ \alpha \le s \le T\right\}.$$

Clearly $v(t) = 0$ for $t \in [\alpha, t_0]$. For $t \ge t_0$, we have

$$v(t) = \int_{t_0}^{t} \left[f\left(s, y_1(\cdot), y_1\left(s - \tau\left(s, y_1(\cdot)\right)\right)\right) - f\left(s, y_2(\cdot), y_2\left(s - \tau\left(s, y_2(\cdot)\right)\right)\right)\right]ds.$$

The Lipschitz condition on f implies that

$$\begin{aligned}&\left|f\left(s, y_1(\cdot), y_1\left(s - \tau\left(s, y_1(\cdot)\right)\right)\right) - f\left(s, y_2(\cdot), y_2\left(s - \tau\left(s, y_2(\cdot)\right)\right)\right)\right|\\ &\quad \le L_1\left|y_1(\cdot) - y_2(\cdot)\right| + L_2\left|y_1\left(s - \tau\left(s, y_1(\cdot)\right)\right) - y_2\left(s - \tau\left(s, y_2(\cdot)\right)\right)\right|.\end{aligned}$$

Let $s_1 = s - \tau\left(s, y_1(\cdot)\right)$ and $s_2 = s - \tau\left(s, y_2(\cdot)\right)$. We have

$$\left|y_1(s_1) - y_2(s_2)\right| \le \left|y_1(s_1) - y_2(s_1)\right| + \left|y_2(s_1) - y_2(s_2)\right|.$$

Similarly, the Lipschitz condition on the delay function τ implies that

$$|y_2(s_1) - y_2(s_2)| \le F_2|s_1 - s_2| \le F_2 L_S V(s).$$

Hence for $t \in [t_0, t_0 + A]$, since $V(t)$ is a nondecreasing function we have

$$\left|v(t)\right| \le \int_{t_0}^{t} \left(L_1 + L_2(1 + F_2 L_S)\right)V(t)ds \le \left(L_1 + L_2(1 + F_2 L_S)\right)AV(t). \tag{3.6}$$

Observe that if we select $A < 1/(L_1 + L_2(1 + F_2 L_S))$ the inequality (3.6) implies that

$$V(t_0 + A) \le [L_1 + L_2(1 + F_2 L_S)]AV(t_0 + A) < V(t_0 + A)$$

if $V(t_0 + A) > 0$. This contradiction implies that $V(t_0 + A) = 0$ and therefore $y_1(t) = y_2(t)$ for $t \in [t_0, t_0 + A]$. This local uniqueness result can be extended to the whole interval $[t_0, \infty)$ by a simple contradiction argument. □

4. A bound on the term $\|y-\bar{y}\|_{[\alpha,T]}$

Assume that the function f in (1.1) is continuous, satisfies the Lipschitz condition of the form

$$\Big\|f\Big(t,y_1(\cdot),u_1\Big)-f\Big(t,y_2(\cdot),u_2\Big)\Big\|\le L_1\Big\|y_1-y_2\Big\|_{[\alpha,t]}+L_2\Big\|u_1-u_2\Big\| \tag{4.1}$$

with constants $L_1, L_2\ge 0$ for $t\in[\alpha,T]$, $y_1,y_2\in C([\alpha,T],\mathbb{R}^m)$, and $u_1,u_2\in\mathbb{R}^m$. Assume also that there exists a solution operator S for (1.2)

$$\tau\big(t,y(\cdot)\big)=S\big(t,y(\cdot),m\big), \tag{4.2}$$

which satisfies the Lipschitz condition

$$\Big|S\big(t,y_1(\cdot),m\big)-S\big(t,y_2(\cdot),m\big)\Big|\le L_S\Big\|y_1-y_2\Big\|_{[\alpha,t]}, \tag{4.3}$$

with $L_S\ge 0$ for $t\in[\alpha,T]$, and $y_1,y_2\in C([\alpha,T],\mathbb{R}^m)$, and the condition

$$0\le\tau\big(t,y(\cdot)\big)\le t-\alpha,\quad t\in[t_0,T]. \tag{4.4}$$

Observe that the inequality $0\le\tau(t,y(\cdot))$ in (4.4) and the Lipschitz condition (4.1) imply that f is a Volterra operator, i.e., it depends only on the past history of the solution y, and the inequality $\tau(t,y(\cdot))\le t-\alpha$ in (4.4) implies that (1.1) is well defined with initial function g specified on the interval $[\alpha,t_0]$. It can be verified that the condition (4.4) is satisfied for the examples considered in Section 2.

In this section we derive a bound on the quantity $\|y-\bar{y}\|_{[\alpha,T]}$, where y is the solution to (1.1)–(1.2) and $\bar{y}$ is the solution to (1.1), (1.3). Let $\bar{S}$ be the solution operator for (1.3), i.e.,

$$\bar{\tau}\big(t,\bar{y}(\cdot)\big)=\bar{S}\big(t,\bar{y}(\cdot),m\big).$$

The exact expressions for the solution operator S defined in Section 3 and $\bar{S}$ defined above are usually not known and in applications an approximation to $\bar{\tau}(t,\bar{y}(\cdot))$ is usually computed by some iterative procedure applied to (1.3), for example the bisection method. We have the following theorem.

Theorem 4.1. *Assume that the function f is continuous and satisfies the Lipschitz condition (4.1) and that the solution operator S given by (4.2) satisfies the condition (4.3). Assume also that*

$$\Big\|S\big(t,\bar{y}(\cdot),m\big)-\bar{S}\big(t,\bar{y}(\cdot),m\big)\Big\|\le\varepsilon,$$

for some $\varepsilon>0$ and $t\in[\alpha,T]$. Then

$$\Big\|y-\bar{y}\Big\|_{[\alpha,t]}\le\frac{\widetilde{\varepsilon}}{L}\Big(e^{L(t-t_0)}-1\Big),\quad t\in[\alpha,T], \tag{4.5}$$

with $\widetilde{\varepsilon}=D\varepsilon$, where D and L are some nonnegative constants.

Proof. Subtracting the integral equations for y and $\bar{y}$

$$\begin{cases} y(t)=y(t_0)+\displaystyle\int_{t_0}^{t}f\Big(s,y(\cdot),y\Big(s-\tau\big(s,y(\cdot)\big)\Big)\Big)ds, & t\in[t_0,T],\\ y(t)=g(t), & t\in[\alpha,t_0],\end{cases}$$

and

$$\begin{cases} \bar{y}(t)=\bar{y}(t_0)+\displaystyle\int_{t_0}^{t}f\Big(s,\bar{y}(\cdot),\bar{y}\Big(s-\bar{\tau}\big(s,\bar{y}(\cdot)\big)\Big)\Big)ds, & t\in[t_0,T],\\ \bar{y}(t)=g(t), & t\in[\alpha,t_0],\end{cases}$$

then taking norms on both sides of the resulting equation, and using (4.1) and the triangle inequality we obtain

$$\begin{aligned}\Big\|y(t)-\bar{y}(t)\Big\|\le&\int_{t_0}^{t}\Big(L_1\Big\|y-\bar{y}\Big\|_{[\alpha,s]}+L_2\Big\|y\Big(s-\tau\big(s,y(\cdot)\big)\Big)-y\Big(s-\bar{\tau}\big(s,\bar{y}(\cdot)\big)\Big)\Big\|\\ &+L_2\Big\|y\Big(s-\bar{\tau}\big(s,\bar{y}(\cdot)\big)\Big)-\bar{y}\Big(s-\bar{\tau}\big(s,\bar{y}(\cdot)\big)\Big)\Big\|\Big)ds,\end{aligned} \tag{4.6}$$

$t\in[t_0,T]$. Since the function y is Lipschitz continuous we get

$$\Big\|y\Big(s-\tau\big(s,y(\cdot)\big)\Big)-\bar{y}\Big(s-\bar{\tau}\big(s,\bar{y}(\cdot)\big)\Big)\Big\|\le L_y\Big|\tau\big(s,y(\cdot)\big)-\bar{\tau}\big(s,\bar{y}(\cdot)\big)\Big|,$$

where L_y is Lipschitz constant of the solution y. We have also

$$\begin{aligned}\Big|\tau\big(s, y(\cdot)\big) - \bar{\tau}\big(s, \bar{y}(\cdot)\big)\Big| &\le \Big\|S\big(s, y(\cdot), m\big) - \bar{S}\big(s, \bar{y}(\cdot), m\big)\Big\| \\ &\le \Big\|S\big(s, y(\cdot), m\big) - S\big(s, \bar{y}(\cdot), m\big)\Big\| + \Big\|S\big(s, \bar{y}(\cdot), m\big) - \bar{S}\big(s, \bar{y}(\cdot), m\big)\Big\| \\ &\le L_S \Big\|y - \bar{y}\Big\|_{[\alpha,s]} + \varepsilon,\end{aligned}$$

and substituting the above inequalities into (4.6) we obtain

$$\Big\|y(t) - \bar{y}(t)\Big\| \le \int_{t_0}^{t} \Big(L\Big\|y - \bar{y}\Big\|_{[\alpha,s]} + \widetilde{\varepsilon}\Big)ds,$$

where

$$L = L_1 + L_2\big(1 + L_y L_S\big), \qquad \widetilde{\varepsilon} = D\varepsilon, \quad D = L_2 L_y.$$

Since the right hand side of the above inequality is nondecreasing with respect to t we have also

$$\Big\|y - \bar{y}\Big\|_{[\alpha,t]} \le \int_{t_0}^{t} \Big(L\Big\|y - \bar{y}\Big\|_{[\alpha,s]} + \widetilde{\varepsilon}\Big)ds. \tag{4.7}$$

Consider the integral equation

$$u(t) = \int_{t_0}^{t} \Big(Lu(s) + \widetilde{\varepsilon}\Big)ds.$$

Then it follows from the theory of integral inequalities, compare [18,19], that

$$\Big\|y - \bar{y}\Big\|_{[\alpha,t]} \le u(t) = \frac{\widetilde{\varepsilon}}{L}\Big(e^{L(t-t_0)} - 1\Big), \quad t \in [\alpha, T],$$

which is our claim. □

In applications the parameter $\widetilde{\epsilon}$ will be related to the discretization parameter h,

$$h = \max\{h_n : \; n = 0, \ldots, N\}.$$

To be more precise, the following relation will hold

$$\widetilde{\epsilon} = O(h^r) \quad \text{as } h \to 0,$$

where r is the order of the quadrature formula used to approximate the threshold operator P. This is further discussed in Section 6.

5. A bound on the term $\|\bar{y} - \bar{y}_h\|_{[\alpha,T]}$

In this section we derive a bound on the quantity $\|\bar{y} - \bar{y}_h\|_{[\alpha,T]}$ using the generalization of the theory of one-step methods for functional differential equations which was developed in [14].

It will be always assumed that the increment function Φ_h appearing in (1.4) satisfies the Lipschitz condition of the form

$$\Big\|\Phi_h\Big(t_n, \theta, y(\cdot), u\Big) - \Phi_h\Big(t_n, \theta, z(\cdot), v\Big)\Big\| \le M_1\Big\|y - z\Big\|_{[\alpha, t_{n+1}]} + M_2\Big\|u - v\Big\| \tag{5.1}$$

with $M_1, M_2 \ge 0$, for $y, z \in C([\alpha, t_{n+1}], \mathbb{R}^m)$ and $u, v \in \mathbb{R}^m$. Observe that the norm of the difference $y - z$ in the condition (5.1) is taken over the interval $[\alpha, t_{n+1}]$. This allows us to include in our discussion the implicit methods. Taking this norm over the interval $[\alpha, t_n]$ only would correspond to explicit methods.

Define the local discretization error $h_n\bar{\xi}(t_n, \theta, h)$ of the method (1.4) at the point $t_n + \theta h_n$ as the residuum obtained by replacing $\bar{y}_h$ by $\bar{y}$ and $\bar{\tau}_h$ by $\bar{\tau}$ in (1.4), i.e.,

$$\bar{y}(t_n + \theta h_n) = \bar{y}(t_n) + h_n\Phi_h\bigg(t_n, \theta, \bar{y}(\cdot), \bar{y}\Big(t_n - \bar{\tau}\big(t_n, \bar{y}(\cdot)\big)\Big)\bigg) + h_n\bar{\xi}(t_n, \theta, h), \tag{5.2}$$

$n = 0, 1, \ldots, N-1$, $\theta \in (0, 1]$, and put

$$\bar{\xi}(h) = \sup\Big\{\big\|\bar{\xi}(t_n, \theta, h)\big\| : 0 \le n \le N, \; \theta \in (0, 1]\Big\}.$$

The method (1.4) is said to be consistent if

$$\lim_{h\to 0}\bar{\xi}(h)=0.$$

Denote also by $\bar{\eta}(t_n,\theta,h)$ the error of the approximation $\bar{\tau}_h(t_{n+\theta},\bar{y}_h(\cdot))$ to $\bar{\tau}(t_{n+\theta},\bar{y}_h(\cdot))$, i.e.,

$$\bar{\eta}(t_n,\theta,h)=\bar{\tau}\big(t_{n+\theta},\bar{y}_h(\cdot)\big)-\bar{\tau}_h\big(t_{n+\theta},\bar{y}_h(\cdot)\big),$$

$n=0,1,\ldots,N$, $\theta\in(0,1]$, and put

$$\bar{\eta}(h)=\sup\Big\{\big|\bar{\eta}(t_n,\theta,h)\big|:0\le n\le N,\ \theta\in(0,1]\Big\}.$$

Observe that this quantity is the error of the numerical procedure used to resolve the approximation (1.6) to the threshold condition (1.3).

Assume that the solution $\bar{y}$ to (1.1) and the solution operator $\bar{S}(t,\bar{y}(\cdot),m)$ to (1.3) satisfy Lipschitz conditions with constants $L_{\bar{y}}\ge 0$ and $L_{\bar{S}}\ge 0$, respectively, of the form

$$\big\|\bar{y}(t_1)-\bar{y}(t_2)\big\|\le L_{\bar{y}}\big|t_1-t_2\big|, \tag{5.3}$$

$t_1,t_2\in[t_0,T]$, and

$$\Big\|\bar{S}\big(t,y(\cdot),m\big)-\bar{S}\big(t,z(\cdot),m\big)\Big\|\le L_{\bar{S}}\Big\|y-z\Big\|_{[\alpha,t]}, \tag{5.4}$$

$t\in[\alpha,T]$, $y,z\in C([\alpha,T],\mathbb{R}^m)$. Denote the global error of the method (1.4) by $\bar{e}_h=\bar{y}-\bar{y}_h$. This method is said to be convergent if

$$\lim_{h\to 0}\big\|\bar{e}_h\big\|_{[\alpha,T]}=0.$$

For a fixed approximation $\bar{P}$ to the operator P we have the following convergence theorem.

Theorem 5.1. *Assume that the problem (1.1) has a unique solution on the interval $[\alpha,T]$. Assume also that the increment function Φ_h satisfies the Lipschitz condition (5.1) and the function $\bar{y}$ and the operator $\bar{S}$ satisfy the Lipschitz conditions (5.3) and (5.4), respectively. Assume also that the method (1.4) is consistent, i.e., $\lim_{h\to 0}\bar{\xi}(h)=0$, the starting error $g-g_h$ satisfies $\lim_{h\to 0}\|g-g_h\|_{[\alpha,t_0]}=0$, and that the error $\bar{\eta}(h)$ of the approximation $\bar{\tau}_h(t_n,\bar{y}_h(\cdot))$ to $\bar{\tau}(t_n,\bar{y}_h(\cdot))$ satisfies $\lim_{h\to 0}\bar{\eta}(h)=0$. Then the method (1.4) is convergent and we have the following error bound*

$$\big\|\bar{e}_h\big\|_{[\alpha,T]}\le\Big(\big\|\bar{e}_h\big\|_{[\alpha,t_0]}+\sigma(h)\Big)e^{Q(T-t_0)}, \tag{5.5}$$

where Q is some nonnegative constant and $\lim_{h\to 0}\sigma(h)=0$. Moreover, if $\|\bar{e}_h\|_{[\alpha,t_0]}=O(h^p)$, $\bar{\xi}(h)=O(h^p)$ and $\bar{\eta}(h)=O(h^p)$ as $h\to 0$ then the method is convergent with order p, i.e., $\|\bar{e}_h\|_{[\alpha,T]}=O(h^p)$ as $h\to 0$.

Proof. Subtracting (1.4) from (5.2), then taking norms on both sides of the resulting equation and using the condition (5.1) we obtain

$$\begin{aligned}\Big\|\bar{e}_h(t_n+\theta h_n)\Big\| &\le \Big\|\bar{e}_h(t_n)\Big\|+h_nM_1\Big\|\bar{e}_h\Big\|_{[\alpha,t_{n+1}]}\\ &\quad+h_nM_2\Big\|\bar{y}\Big(t_{n+\theta}-\bar{\tau}\big(t_{n+\theta},\bar{y}(\cdot)\big)\Big)-\bar{y}_h\Big(t_{n+\theta}-\bar{\tau}_h\big(t_{n+\theta},\bar{y}_h(\cdot)\big)\Big)\Big\|+h_n\bar{\xi}(h),\end{aligned}$$

$n=0,1,\ldots,N$, $\theta\in(0,1]$. This leads to

$$\begin{aligned}\Big\|\bar{e}_h(t_n+\theta h_n)\Big\| &\le \Big\|\bar{e}_h\Big\|_{[\alpha,t_n]}+h_nM_1\Big\|\bar{e}_h\Big\|_{[\alpha,t_{n+1}]}\\ &\quad+h_nM_2\Big\|\bar{y}\Big(t_{n+\theta}-\bar{\tau}\big(t_{n+\theta},\bar{y}(\cdot)\big)\Big)-\bar{y}\Big(t_{n+\theta}-\bar{\tau}_h\big(t_{n+\theta},\bar{y}_h(\cdot)\big)\Big)\Big\|\\ &\quad+h_nM_2\Big\|\bar{y}\Big(t_{n+\theta}-\bar{\tau}_h\big(t_{n+\theta},\bar{y}_h(\cdot)\big)\Big)-\bar{y}_h\Big(t_{n+\theta}-\bar{\tau}_h\big(t_{n+\theta},\bar{y}_h(\cdot)\big)\Big)\Big\|+h_n\bar{\xi}(h),\end{aligned}$$

$n=0,1,\ldots,N$, $\theta\in(0,1]$. Using (5.3) and the definition of $\bar{e}_h$ we get

$$\begin{aligned}\Big\|\bar{e}_h(t_n+\theta h_n)\Big\| &\le \Big\|\bar{e}_h\Big\|_{[\alpha,t_n]}+h_nM_1\Big\|\bar{e}_h\Big\|_{[\alpha,t_{n+1}]}\\ &\quad+h_nM_2L_{\bar{y}}\Big\|\bar{\tau}\big(t_{n+\theta},\bar{y}(\cdot)\big)-\bar{\tau}_h\big(t_{n+\theta},\bar{y}_h(\cdot)\big)\Big\|+h_nM_2\Big\|\bar{e}_h\Big\|_{[\alpha,t_{n+1}]}+h_n\bar{\xi}(h).\end{aligned} \tag{5.6}$$

It follows from (5.4), the definition of the solution operator $\bar{S}$, and the definition of $\bar{\eta}(h)$ that

$$\left\|\bar{\tau}\big(t_{n+\theta},\bar{y}(\cdot)\big)-\bar{\tau}_h\big(t_{n+\theta},\bar{y}_h(\cdot)\big)\right\| \le \left\|\bar{S}\big(t_{n+\theta},\bar{y}(\cdot)\big)-\bar{S}\big(t_{n+\theta},\bar{y}_h(\cdot)\big)\right\| + \left\|\bar{\tau}\big(t_{n+\theta},\bar{y}_h(\cdot)\big)-\bar{\tau}_h\big(t_{n+\theta},\bar{y}_h(\cdot)\big)\right\| \le L_{\bar{S}}\left\|\bar{e}_h\right\|_{[\alpha,t_{n+1}]}+\bar{\eta}(h).$$

Substituting the above inequality into (5.6) leads to

$$\left\|\bar{e}_h\right\|_{[t_n,t_{n+1}]} \le \left\|\bar{e}_h\right\|_{[\alpha,t_n]} + h_nM_1\left\|\bar{e}_h\right\|_{[\alpha,t_{n+1}]} + h_nM_2(1+L_{\bar{y}}L_{\bar{S}})\left\|\bar{e}_h\right\|_{[\alpha,t_{n+1}]} + h_n\big(\bar{\xi}(h)+M_2L_{\bar{y}}\bar{\eta}(h)\big).$$

Since the sequence $\|\bar{e}_h\|_{[\alpha,t_n]}$ is nondecreasing with respect to n the last inequality leads to

$$\left\|\bar{e}_h\right\|_{[\alpha,t_{n+1}]} \le \left\|\bar{e}_h\right\|_{[\alpha,t_n]} + h_nM\left\|\bar{e}_h\right\|_{[\alpha,t_{n+1}]} + h_n\big(\bar{\xi}(h)+M_2L_{\bar{y}}\bar{\eta}(h)\big),$$

or

$$\Big(1-h_nM\Big)\left\|\bar{e}_h\right\|_{[\alpha,t_{n+1}]} \le \left\|\bar{e}_h\right\|_{[\alpha,t_n]} + h_n\big(\bar{\xi}(h)+M_2L_{\bar{y}}\bar{\eta}(h)\big),$$

where the constant M is defined by $M = M_1 + M_2(1 + L_{\bar{y}}L_{\bar{S}})$. If $M > 0$ then

$$0 < \frac{1}{1-h_nM} \le 1 + h_nCM \quad \text{for } h_n < h^* := \min\left\{\frac{1}{M}, \frac{C-1}{CM}\right\},$$

where $C > 1$ is an arbitrary constant. Hence, it follows that

$$\left\|\bar{e}_h\right\|_{[\alpha,t_{n+1}]} \le \Big(1+h_nQ\Big)\left\|\bar{e}_h\right\|_{[\alpha,t_n]} + h_n\Big(\widetilde{\xi}(h)+\widetilde{\eta}(h)\Big), \tag{5.7}$$

$n = 0, 1, \ldots, N$, $0 < h_n < h^*$, where the constant Q and the quantities $\widetilde{\xi}(h)$ and $\widetilde{\eta}(h)$ are defined by

$$Q = CM, \qquad \widetilde{\xi}(h) = (1+h^*Q)\bar{\xi}(h), \qquad \widetilde{\eta}(h) = M_2L_{(1+h^*Q)}\bar{y}\bar{\eta}(h).$$

If $M = 0$ then (5.7) is true with

$$Q = 0, \qquad \widetilde{\xi}(h) = \bar{\xi}(h), \qquad \widetilde{\eta}(h) = M_2L_{\bar{y}}\bar{\eta}(h),$$

and without any restriction imposed on h_n. Using the standard induction arguments the inequality (5.7) leads to

$$\left\|\bar{e}_h\right\|_{[\alpha,t_n]} \le \left\|\bar{e}_h\right\|_{[\alpha,t_0]}\prod_{j=0}^{n-1}\Big(1+h_jQ\Big) + \Big(\widetilde{\xi}(h)+\widetilde{\eta}(h)\Big)\sum_{j=0}^{n-1}h_j\prod_{i=j+1}^{n-1}\Big(1+h_iQ\Big),$$

$n = 0, 1, \ldots, N$. It follows from (1.5) that

$$\sum_{j=0}^{n-1}h_j\prod_{i=j+1}^{n-1}\Big(1+h_iQ\Big) \le \sum_{j=0}^{n-1}h_j\prod_{i=0}^{n-1}\Big(1+h_iQ\Big) \le (T-t_0)\prod_{j=0}^{n-1}\Big(1+h_jQ\Big)$$

and we obtain

$$\left\|\bar{e}_h\right\|_{[\alpha,t_n]} \le \left(\left\|\bar{e}_h\right\|_{[\alpha,t_0]} + (T-t_0)\big(\widetilde{\xi}(h)+\widetilde{\eta}(h)\big)\right)\prod_{j=0}^{n-1}\Big(1+h_jQ\Big).$$

The last inequality and the bound

$$\prod_{j=0}^{n-1}\Big(1+h_jQ\Big) < e^{(h_0+\cdots+h_{n-1})Q} \le e^{(h_0+\cdots+h_{N-1})Q} < e^{(T-t_0)Q}$$

lead to

$$\left\|\bar{e}_h\right\|_{[\alpha,T]} < \left(\left\|\bar{e}_h\right\|_{[\alpha,t_0]} + (T-t_0)\big(\widetilde{\xi}(h)+\widetilde{\eta}(h)\big)\right)e^{Q(T-t_0)}$$

which is equivalent to (5.5) with $\sigma(h) = (T-t_0)(\widetilde{\xi}(h)+\widetilde{\eta}(h))$. It is also clear that $\|\bar{e}_h\|_{[\alpha,t_0]} = O(h^p)$, $\|\bar{\xi}(h)\| = O(h^p)$ and $\|\bar{\eta}(h)\| = O(h^p)$ imply the convergence of order p. This completes the proof. □

6. Continuous Runge–Kutta methods

Continuous Runge–Kutta methods for ordinary differential equations have been investigated by Zennaro [20,21], and Owren and Zennaro [22,7]. The adaptations of these methods to various forms of functional differential equations is discussed in the monograph [23]. We formulate this extension for a somewhat less general class of problems than that discussed in Section 1, namely the class of delay-differential equations of the form

$$\begin{cases} y'(t) = f\Big(t, y(t), y\Big(t-\tau\big(t, y(\cdot)\big)\Big)\Big), & t \in [t_0, T], \\ y(t) = g(t), & t \in [\alpha, t_0], \end{cases} \tag{6.1}$$

$\alpha \le t_0$, where $f : [t_0, T] \times \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}^m$ is continuous and satisfies the Lipschitz condition

$$\big\|f\big(t, y_1, u_1\big) - f\big(t, y_2, u_2\big)\big\| \le L_1\big\|y_1 - y_2\big\| + L_2\big\|u_1 - u_2\big\| \tag{6.2}$$

with constants $L_1, L_2 \ge 0$ for $t \in [\alpha, T], y_1, y_2, u_1, u_2 \in \mathbb{R}^m$, and the function $\tau(t, y(\cdot))$ satisfies (1.2). Observe that the specific examples discussed in Section 2 are of the form (6.1), (1.2). As in Section 1 we also consider the discrete approximation (1.3) to the threshold condition (1.2). Following the approach of [23] explicit continuous Runge–Kutta methods adapted to (6.1), (1.3) take the following form

$$\begin{cases} \bar{Y}_i = \bar{y}_h(t_n) + h_n \displaystyle\sum_{j=1}^{i-1} a_{ij}\bar{F}_j, \\ \bar{F}_i = f\Big(t_n + c_ih_n, \bar{Y}_i, \bar{y}_h\Big(t_n + c_ih_n - \bar{\tau}_h\big(t_n + c_ih_n, \bar{y}_h(\cdot)\big)\Big)\Big), \\ \bar{y}_h(t_n + \theta h_n) = \bar{y}_h(t_n) + h_n \displaystyle\sum_{j=1}^{s} b_j(\theta)\bar{F}_j, \end{cases} \tag{6.3}$$

$i = 1, 2, \dots, s, n = 0, 1, \dots, N-1, \theta \in (0, 1]$. Here, $\bar{\tau}_h(t_n+c_ih_n, \bar{y}_h(\cdot))$ are approximations to the solutions $\bar{\tau}(t_n+c_ih_n, \bar{y}_h(\cdot))$ of the discrete threshold conditions

$$\bar{P}\Big(t_n + c_ih_n, \bar{y}_h(\cdot), \bar{\tau}\big(t_n + c_ih_n, \bar{y}_h(\cdot)\big)\Big) = m.$$

Observe that this equation corresponds to (1.6) with $\theta = c_i$, $i = 1, 2, \dots, s$. The increment function Φ_h for this method is given by

$$\Phi_h(t, \theta, y, u) = \sum_{j=1}^{s} b_j(\theta)\bar{F}_j(t + c_jh_n, y_j, u_j),$$

where

$$y = \begin{bmatrix} y_1 \\ \vdots \\ y_s \end{bmatrix} \in \mathbb{R}^{ms}, \qquad u = \begin{bmatrix} u_1 \\ \vdots \\ u_s \end{bmatrix} \in \mathbb{R}^{ms}.$$

Observe that (6.3) is a nonlinear system for $\bar{Y}_i$ or $\bar{F}_i$ if the delay is smaller than the stepsize of integration. This is the case for example (2.1)–(2.4) at the beginning of integration. Such systems will be solved by functional iterations.

The convergence and order of convergence of the method (6.3) can be analyzed using Theorem 5.1 in Section 4 and Theorem 6.1.2 in [23]. First of all, it follows from (6.2) that the increment function Φ_h of the method (6.3) satisfies the inequality

$$\big\|\Phi_h\big(t, \theta, y, u\big) - \Phi_h\big(t, \theta, \bar{y}, \bar{u}\big)\big\| \le \sum_{j=1}^{s}\big|b_j(\theta)\big|\Big\|f(t + c_jh_n, y_j, u_j) - f(t + c_jh_n, \bar{y}_j, \bar{u}_j)\Big\|.$$

This inequality implies that Φ_h satisfies the Lipschitz condition of the form

$$\big\|\Phi_h\big(t, \theta, y, u\big) - \Phi_h\big(t, \theta, \bar{y}, \bar{u}\big)\big\| \le M_1|||y - \bar{y}||| + M_2|||u - \bar{u}|||, \tag{6.4}$$

where $M_1 = BL_1, M_2 = BL_2, B = \sup\big\{\sum_{j=1}^{s}\big|b_j(\theta)\big| : \theta \in (0, 1]\big\}$, and for $x = [x_1^T, x_2^T, \dots, x_s^T]^T \in \mathbb{R}^{ms}$ the norm $|||\cdot|||$ is defined by $|||x||| = \max\{\|x_i\| : 1 \le i \le s\}$. Since $\bar{\tau}_h(t_n + c_ih_n, \bar{y}_h(\cdot))$ depends only on the discrete number of the values of the function $\bar{y}_h$ it follows from Theorem 6.1.2 in [23] that the method (6.3) preserves the order of the underlying continuous Runge–Kutta method for ordinary differential equations

$$\begin{cases} y'(t) = f\big(t, y(t)\big), & t \in [t_0, T], \\ y(t_0) = y_0. \end{cases} \tag{6.5}$$

Hence, if the local discretization error of the underlying Runge–Kutta method for (6.5) tends to zero as $h \to 0$, $h = \max\{h_n : n = 0, \dots, N\}$, then the local discretization error $\bar{\xi}(h)$ of the method (6.3) for (6.1), (1.3) has also this property. If we assume, in addition, that $\lim_{h\to 0} \|g - g_h\|_{[\alpha,t_0]} = 0$ and the error $\bar{\eta}(h)$ of the approximations $\bar{\tau}_h(t_n + c_i h_n, \bar{y}_h(\cdot))$ to $\bar{\tau}(t_n + c_i h_n, \bar{y}_h(\cdot))$ satisfies $\lim_{h\to 0} \bar{\eta}(h) = 0$ then it follows from (6.4) and Theorem 5.1 that the method (6.3) is convergent to the solution $\bar{y}$ of (6.1), (1.3). Moreover, if the underlying Runge–Kutta method for (6.5) has order p, $\|g - g_h\|_{[\alpha,t_0]} = O(h^p)$ and $\bar{\eta}(h) = O(h^p)$ then the method (6.3) is convergent with order p.

Our algorithm for (1.1)–(1.2) which will be described in the next section is based on continuous Runge–Kutta method of order p and discrete method of order $q = p - 1$. Such embedded pairs were constructed in [7] using the strategy based on the minimization of the error constant of continuous Runge–Kutta method. The coefficients of embedded pair of order $p = 4$ and $q = p - 1 = 3$ derived in [7] are given by

$$\begin{array}{c|cccccc}
0 & & & & & & \\
\frac{1}{6} & \frac{1}{6} & & & & & \\
\frac{11}{37} & \frac{44}{1369} & \frac{363}{1369} & & & & \\
\frac{11}{17} & \frac{3388}{4913} & -\frac{8349}{4913} & \frac{8140}{4913} & & & \\
\frac{13}{15} & -\frac{36764}{408375} & \frac{767}{1125} & -\frac{32708}{136125} & \frac{210392}{408375} & & \\
1 & \frac{1697}{18876} & 0 & \frac{50653}{116160} & \frac{299693}{1626240} & \frac{3375}{11648} & \\
\hline
\bar{y}_h(t_n + \theta h_n) & b_1(\theta) & b_2(\theta) & b_3(\theta) & b_4(\theta) & b_5(\theta) & b_6(\theta) \\
\hline
\widehat{y}_{n+1} & \frac{101}{363} & 0 & -\frac{1369}{14520} & \frac{11849}{14520} & 0 & 0
\end{array}, \tag{6.6}$$

with

$$\begin{aligned}
b_1(\theta) &= -\frac{866577}{824252}\,\theta^4 + \frac{1806901}{618189}\,\theta^3 - \frac{104217}{37466}\,\theta^2 + \theta, \\
b_2(\theta) &= 0, \\
b_3(\theta) &= \frac{12308679}{5072320}\,\theta^4 - \frac{2178079}{380424}\,\theta^3 + \frac{861101}{230560}\,\theta^2, \\
b_4(\theta) &= -\frac{7816583}{10144640}\,\theta^4 + \frac{6244423}{5325936}\,\theta^3 - \frac{63869}{293440}\,\theta^2, \\
b_5(\theta) &= -\frac{624375}{217984}\,\theta^4 + \frac{982125}{190736}\,\theta^3 - \frac{1522125}{762944}\,\theta^2, \\
b_6(\theta) &= \frac{296}{131}\,\theta^4 - \frac{461}{131}\,\theta^3 + \frac{165}{131}\,\theta^2.
\end{aligned} \tag{6.7}$$

In (6.6) $\bar{y}_h(t_n + \theta h_n)$ corresponds to continuous Runge–Kutta method with continuous weights $b_i(\theta)$ given in (6.7), and $\widehat{y}_{n+1}$ corresponds to discrete method which is used for error estimation.

7. Description of the algorithm

To integrate (1.1) we use the continuous Runge–Kutta method with coefficients listed in (6.6) and (6.7). After computing $\bar{y}_h(t_{n+1})$ and $\widehat{y}_{n+1}$ we compute the estimate of the local discretization error at the point t_{n+1} according to the formula

$$\mathrm{est}(t_{n+1}) = \left\| \bar{y}_h(t_{n+1}) - \widehat{y}_{n+1} \right\|.$$

Once this estimate is computed the stepsize h_n from t_n to t_{n+1} is accepted if

$$\mathrm{est}(t_{n+1}) \le \mathrm{tol},$$

where tol is a prescribed accuracy tolerance. A new stepsize h_{n+1} from t_{n+1} to t_{n+2} is then computed according to the standard formula

$$h_{n+1} = \delta\, h_n \left(\frac{\mathrm{tol}}{\mathrm{est}(t_{n+1})} \right)^{1/5},$$

where δ is a safety coefficient chosen as $\delta = 0.8$. If

$$\mathrm{est}(t_{n+1}) > \mathrm{tol}$$

the step is rejected and we restart the integration at the point t_n with the new stepsize $\bar{h}_n = 0.5 h_n$. Following the approach of [24] the initial stepsize h_0 is computed from the formula

$$h_0 = \min\left\{ 0.1|T - t_0|, \frac{\mathrm{tol}^{1/5}}{\left\| f\Big(t_0, y(\cdot), y\big(t_0 - \tau(t_0, g(\cdot))\big)\Big) \right\|} \right\}.$$

In addition to continuous Runge–Kutta method to resolve (1.1) we also need to implement a solver to compute approximations $\bar{\tau}_h(t,\bar{y}_h(\cdot))$ to the solution $\bar{\tau}(t,\bar{y}_h(\cdot))$ to the nonlinear equation

$$\bar{P}\Big(t,\bar{y}_h(\cdot),\bar{\tau}\big(t,\bar{y}_h(\cdot)\big)\Big)=m, \tag{7.1}$$

where $\bar{P}$ is a discrete approximation to P. For this purpose we will use the bisection method. The implementation of this algorithm for specific examples discussed in Section 2 requires the computation of approximations to the integral of the form

$$\int_{\tau}^{t}\rho(s)I(s)ds \tag{7.2}$$

in the first example described in Section 2 or

$$\int_{t-\tau}^{t}p\big(x(s)\big)ds \tag{7.3}$$

in the second example described in Section 2. Here, τ stands for the approximation to $\tau(t,I(\cdot))$ in the first integral and approximation to $\tau(t,x(\cdot))$ in the second integral. If $\tau\le t_0$ or $t-\tau\le t_0$ these integrals can be written in the form

$$\int_{\tau}^{t_0}\rho(s)I(s)ds+\int_{t_0}^{t}\rho(s)I(s)ds$$

or

$$\int_{t-\tau}^{t_0}p\big(x(s)\big)ds+\int_{t_0}^{t}p\big(x(s)\big)ds.$$

The approximations to the first integrals in the above two relations are computed using the subroutine `quad` in Matlab with the same accuracy tolerance tol used by continuous Runge–Kutta method. The approximations to the second integrals in the above two relations are computed by composite Simpson rule using the grid generated by continuous Runge–Kutta method, with the exception that the approximations on the intervals $[t_n,t_n+c_ih_n]$ are computed using approximations to $I(t_n)$ or $x(t_n)$. If $\tau>t_0$ in (7.2) or $t-\tau>t_0$ in (7.3) these integrals can be written in the form

$$\int_{\tau}^{t}\rho(s)I(s)ds=\int_{\tau}^{t_q}\rho(s)I(s)ds+\int_{t_q}^{t_n}\rho(s)I(s)ds+\int_{t_n}^{t}\rho(s)I(s)ds$$

or

$$\int_{t-\tau}^{t}p\big(x(s)\big)ds=\int_{t-\tau}^{t_q}p\big(x(s)\big)ds+\int_{t_q}^{t_n}p\big(x(s)\big)ds+\int_{t_n}^{t}p\big(x(s)\big)ds.$$

The first integrals on the right hand side of the above two relations are then approximated by the Simpson rule

$$\int_{\tau}^{t_q}\rho(s)I(s)ds\approx\frac{t_q-\tau}{6}\Big(\rho(t_{q-1}+\theta_1h_{q-1})I(t_{q-1}+\theta_1h_{q-1})+4\,\rho(t_{q-1}+\theta_2h_{q-1})I(t_{q-1}+\theta_2h_{q-1})+\rho(t_q)I(t_q)\Big)$$

or

$$\int_{t-\tau}^{t_q}\rho(s)I(s)ds\approx\frac{t_q-t+\tau}{6}\Big(p\big(x(t_{q-1}+\theta_1h_{q-1})\big)+4\,p\big(x(t_{q-1}+\theta_2h_{q-1})\big)+p\big(x(t_q)\big)\Big).$$

Here, $q\ge 1$ is the integer such that $\tau\in(t_{q-1},t_q]$ or $t-\tau\in(t_{q-1},t_q]$,

$$\theta_1=\frac{\tau-t_{q-1}}{t_q-t_{q-1}}\quad\text{or}\quad\theta_1=\frac{t_q-t+\tau}{t_q-t_{q-1}},$$

and $\theta_2=(1+\theta_1)/2$. The approximations to $I(t_{q-1}+\theta_1h_{q-1})$, $I(t_{q-1}+\theta_2h_{q-1})$, $x(t_{q-1}+\theta_1h_{q-1})$ and $x(t_{q-1}+\theta_2h_{q-1})$ are computed by continuous Runge–Kutta method (6.3). The approximations to the integrals over the interval $[t_q,t_n]$ are computed by composite Simpson rules

$$\int_{t_q}^{t_n}\rho(s)I(s)ds\approx\sum_{s=q}^{n-1}\frac{t_{s+1}-t_s}{6}\Big(\rho(t_s)I(t_s)+4\,\rho\left(t_{s+\frac{1}{2}}\right)I\left(t_{s+\frac{1}{2}}\right)+\rho(t_{s+1})I(t_{s+1})\Big)$$

or

$$\int_{t_q}^{t_n}p\big(x(s)\big)ds\approx\sum_{s=q}^{n-1}\frac{t_{s+1}-t_s}{6}\Big(p\big(x(t_s)\big)+4\,p\Big(x\left(t_{s+\frac{1}{2}}\right)\Big)+p\big(x(t_{s+1})\big)\Big).$$

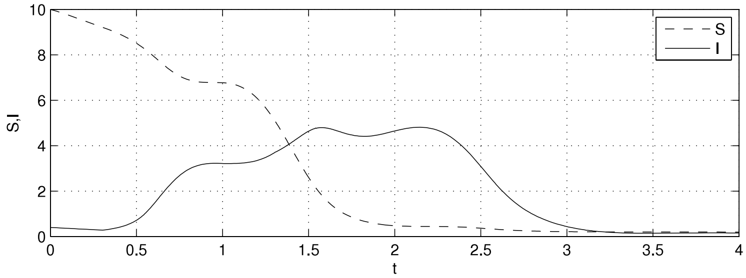

Fig. 1. Solution $S(t)$ and $I(t)$ versus t to the problem (2.1)–(2.4) computed with $\mathrm{tol} = 10^{-6}$.

The approximations to the integrals over the interval $[t_n, t]$ are computed by

$$\int_{t_n}^{t} \rho(s)I(s)ds \approx (t - t_n)\rho(t_n)I(t_n) \quad \text{or} \quad \int_{t_n}^{t} p\big(x(s)\big)ds \approx (t - t_n)p\big(x(t_n)\big).$$

If τ or $t - \tau$ is greater than t_n then we use the approximations

$$\int_{\tau}^{t} \rho(s)I(s)ds \approx (t - \tau)\rho(t_n)I(t_n) \quad \text{or} \quad \int_{t-\tau}^{t} p\big(x(s)\big)ds \approx \tau p\big(x(t_n)\big).$$

In the implementation of our algorithm we have to compute approximations

$$\bar{\tau}_h(t_n + c_i h_n, \bar{y}_h(\cdot))$$

to the solution $\bar{\tau}(t_n + c_i h_n, \bar{y}_h(\cdot))$ of (7.1) for $n = 0, 1, \dots, i = 2, 3, \dots, s$. We use the bisection method on the interval

$$\Big[\bar{\tau}_h\big(t_n + c_{i-1}h_n, \bar{y}_h(\cdot)\big) - \Delta\bar{\tau}, \bar{\tau}_h\big(t_n + c_{i-1}h_n, \bar{y}_h(\cdot)\big) + \Delta\bar{\tau}\Big],$$

where we have chosen $\Delta\bar{\tau} = \mathrm{tol}^{1/5}$. Since $c_s = 1$ we also define

$$\bar{\tau}_h\big(t_{n+1}, \bar{y}_h(\cdot)\big) = \bar{\tau}_h\big(t_n + c_s h_n, \bar{y}_h(\cdot)\big)$$

to start computations on the interval

$$\Big[\bar{\tau}_h\big(t_{n+1} + c_{i-1}h_{n+1}, \bar{y}_h(\cdot)\big) - \Delta\bar{\tau}, \bar{\tau}_h\big(t_{n+1} + c_{i-1}h_{n+1}, \bar{y}_h(\cdot)\big) + \Delta\bar{\tau}\Big],$$

where h_{n+1} is a new stepsize. The initial approximation $\bar{\tau}_h(t_0, \bar{y}_h(\cdot))$ to $\bar{\tau}(t_0, \bar{y}_h(\cdot))$ is computed by solving Eq. (7.1) for $t = t_0$. This equation takes the form

$$\int_{\bar{\tau}(t_0, I_0(\cdot))}^{t_0} \rho(s)I_0(s)ds = m$$

for Example 1 and

$$\int_{t_0 - \bar{\tau}(t_0, x_0(\cdot))}^{t_0} p\big(x_0(s)\big)ds = m$$

for Example 2. Here, $I_0(s)$ and $x_0(s)$ are given initial functions, compare Section 2.

8. Numerical experiments

We have applied the algorithm described in Section 6 to the examples of threshold problems (2.1)–(2.4) and (2.6)–(2.7) discussed in Section 2. The selection of numerical results for problem (2.1)–(2.4) is presented in Figs. 1–3.

These figures correspond to $m = 0.1$, $\sigma = 1$, $S_0 = 10$, $\rho(t) = \exp(-t^2)$, $r(t) = 0.5(1 + \sin(5t))$, and $I_0(t)$ defined by

$$I_0(t) = \begin{cases} 0.4(1 + t), & -1 \le t \le 0, \\ 0.4(1 - t), & 0 \le t \le 1, \\ 0, & \text{otherwise}. \end{cases}$$

It can be verified that the condition (2.3) implies $t_0 = 0.302817$. In Fig. 1 we have plotted the approximations to the solution $S(t)$ and $I(t)$ to the problem (2.1)–(2.2) computed with the tolerance $\mathrm{tol} = 10^{-6}$. In Fig. 2 we have plotted the approximation to the delay function $\tau(t) = \tau(t, I(\cdot))$ for $t \ge t_0$. The stepsize pattern of our algorithms is presented in Fig. 3, where the rejected steps are marked by the symbol '$\times$'. There were $n_r = 105$ rejected steps out of $n = 383$ steps for this error tolerance.

The selection of numerical results for problem (2.6)–(2.7) is presented in Figs. 4–6. These figures correspond to the parameter values $r = 1, K = 1, b = 10, d = 0.5, d_j = 1, m = 1$, the function $p(x) = x$, and the initial functions

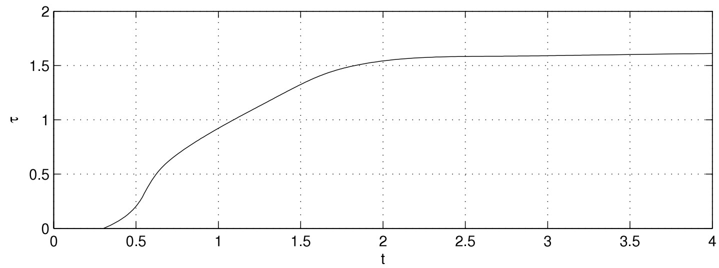

Fig. 2. The function $\tau(t) = \tau(t, I(\cdot))$ versus t for $t \geq t_0$ for the problem (2.1)–(2.4) computed with tol $= 10^{-6}$.

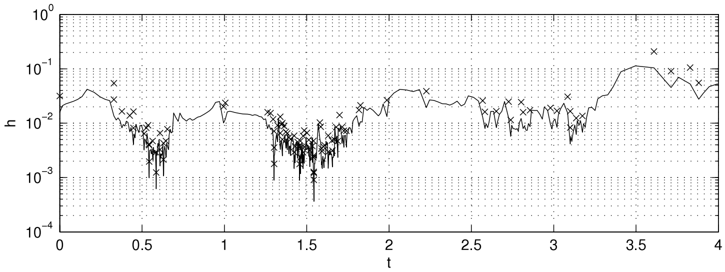

Fig. 3. Stepsize pattern of the algorithm applied to the problem (2.1)–(2.4) with tol $= 10^{-6}$.

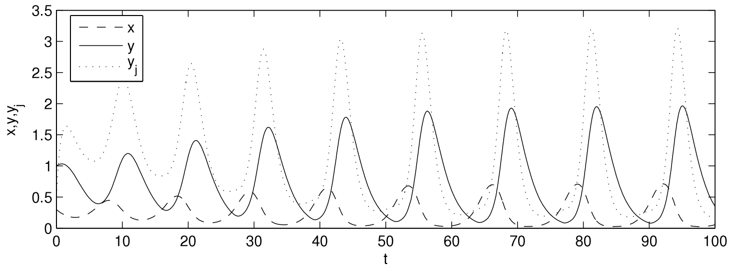

Fig. 4. Solution $x(t)$, $y(t)$ and $y_j(t)$ versus t to the problem (2.6)–(2.7) computed with tol $= 10^{-4}$.

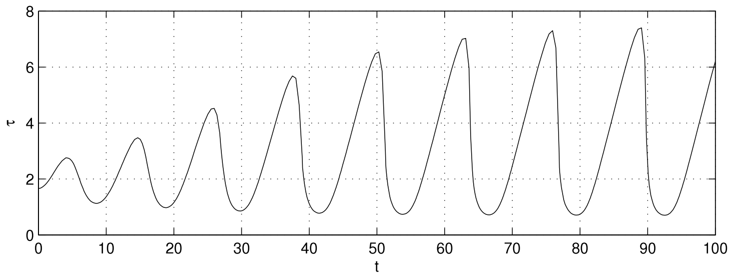

Fig. 5. The function $\tau(t) = \tau(t, x(\cdot))$ versus t for the problem (2.6)–(2.7) computed with tol $= 10^{-4}$.

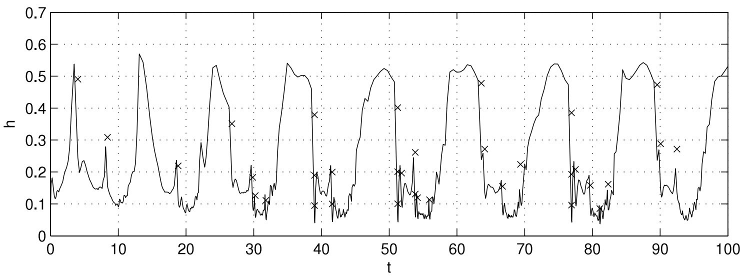

Fig. 6. Stepsize pattern of the algorithm applied to the problem (2.6)–(2.7) with tol $= 10^{-4}$.

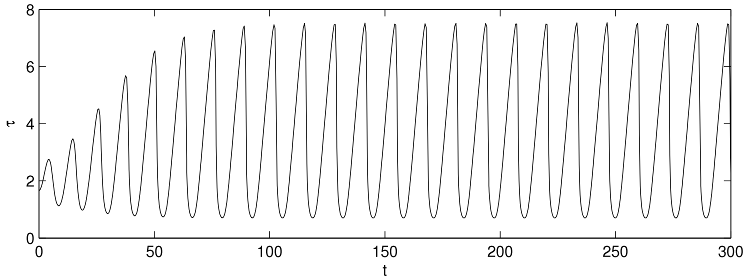

Fig. 7. Function $\tau(t)$ versus t for $p(x) = x$ and $t \in [0, 300]$ obtained with tol $= 10^{-4}$ by direct resolution of threshold condition.

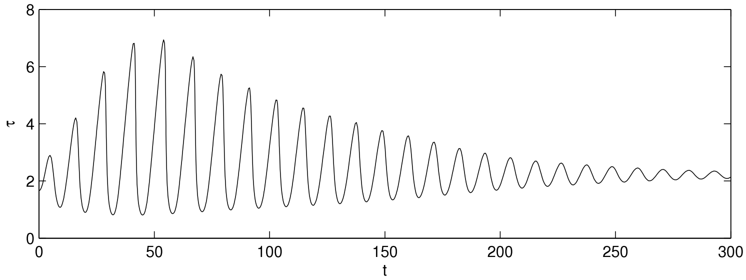

Fig. 8. Function $\tau(t)$ versus t for $p(x) = x$ and $t \in [0, 300]$ obtained with tol $= 10^{-4}$ by differentiating threshold condition.

$x_0(t) = 0.3$ and $y_0(t) = 1$. In Fig. 4 we have plotted the approximations to the solution $x(t)$, $y(t)$ and $y_j(t)$ to the problem (2.6)–(2.7) computed with the tolerance tol $= 10^{-4}$. In Fig. 5 we have plotted the approximation to the delay function $\tau(t) = \tau(t, x(\cdot))$. The stepsize pattern of our algorithms is presented in Fig. 6, where the rejected steps are marked by the symbol '$\times$'. There were $n_r = 34$ rejected steps out of $n = 610$ steps for this error tolerance.

We have also implemented the algorithm, which is based on differentiating the threshold condition, and then integrating the resulting system of delay differential equations by continuous Runge–Kutta method. However, the disadvantage of this approach, as compared with the approach based on the direct resolution of threshold condition, is that now the computation of $\tau(t)$ is subject to propagated errors and is, in general, less accurate for the same tolerances. In contrast, in our approach, where we resolve directly the threshold condition, we can compute the delay function $\tau(t)$ to any accuracy at any point t. To illustrate this point we have integrated problem (2.6)–(2.7) for $t \in [0, 300]$ with the function $p(x) = x$ using the approach based on the direct resolution of the threshold condition (2.7) with the tolerance tol $= 10^{-4}$ and this resulted in the function $\tau(t)$ plotted in Fig. 7. This requires $n = 1801$ steps and there were $nr = 108$ rejected steps. This is a correct behavior of the function $\tau(t)$. Then we differentiated the threshold condition (2.7) to obtain (2.8) and integrated the resulting system of delay differential equations by the same continuous Runge–Kutta method of order four with the same accuracy tolerance tol $= 10^{-4}$ and this resulted in the function $\tau(t)$ plotted in Fig. 8. This requires $n = 912$ step and there were $nr = 97$ rejected steps. So the integration by numerical method for delay differential equations is more efficient but much less accurate. We can recover the behavior similar to that in Fig. 7 using the continuous Runge–Kutta method with a much smaller tolerance. For example, using tol $= 10^{-8}$ leads to the function $\tau(t)$ plotted in Fig. 9. But now the integration requires $n = 21761$ steps and there were $nr = 7649$ rejected steps. There are also other problems with the approach based on differentiating the threshold condition and integrating the resulting system by numerical methods for delay differential equations. For some tolerances and functions $p(x)$ the numerical method for delay differential systems was not able to integrate the resulting system until the end of the interval of integration. For example for tol $= 10^{-4}$ and $p(x) = 1.5x$ the integration was terminated at $t_{\text{end}} = 32.4165$ and for tol $= 10^{-4}$ and $p(x) = 2x$ the integration was terminated at $t_{\text{end}} = 19.6658$. In contrast, using our approach we were still able to integrate the problem on the whole interval $[0, 300]$. The resulting function $\tau(t)$ corresponding to tol $= 10^{-4}$ and $p(x) = 2x$ is plotted in Fig. 10. This integration requires $n = 1360$ steps and there were $nr = 144$ rejected steps. Observe very sharp gradients which suggest that this function may be discontinuous. This does not prevent our algorithm from computing it with a high accuracy.

9. Concluding remarks

We described a new variable stepsize algorithm for the numerical solution of threshold problems in epidemics and population dynamics. This algorithm is based on embedded pair of continuous Runge–Kutta method of order $p = 4$ and discrete Runge–Kutta method of order $q = p - 1 = 3$ which is used for the estimation of local discretization errors which form a basis for adaptive selection of stepsizes. The integral threshold conditions are approximated by the composite Simpson rule and resolved using the bisection method. In contrast to previous approaches to the numerical solution of

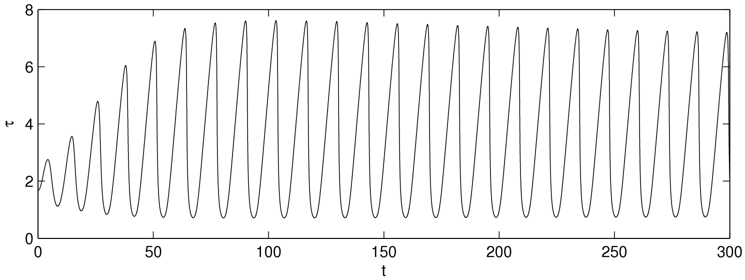

Fig. 9. Function $\tau(t)$ versus t for $p(x) = x$ and $t \in [0, 300]$ obtained with tol $= 10^{-8}$ by differentiating threshold condition.

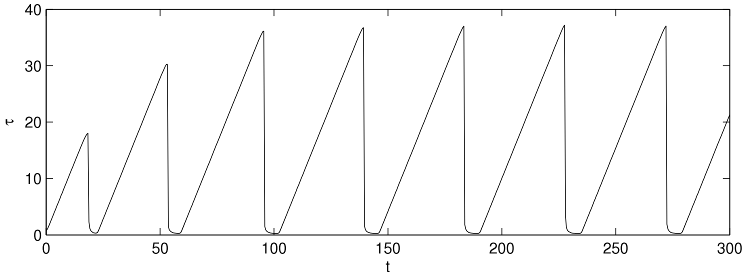

Fig. 10. Function $\tau(t)$ versus t for $p(x) = 2x$ and $t \in [0, 300]$ obtained with tol $= 10^{-4}$ by direct resolution of threshold condition.

this problem, our algorithm is applicable if the solution $\tau(t)$ to the threshold condition has sharp gradients which leads to stiffness, or even if $\tau(t)$ is not differentiable. The results of numerical experiments on examples of threshold problems from epidemics and population dynamics are presented which illustrate the accuracy, reliability and robustness of the new algorithm.

Acknowledgments

The authors wish to express their gratitude to Frank Hoppensteadt and Stephen Gourley for providing examples of threshold problems in epidemics and population dynamics which motivated this research.

References

[1] Z. Jackiewicz, B. Zubik-Kowal, Discrete variable methods for delay-differential equations with threshold-type delays, J. Comput. Appl. Math. 228 (2009) 514–523.
[2] F.C. Hoppensteadt, P. Waltman, A problem in the theory of epidemics, Math. Biosci. 9 (1970) 71–91.
[3] F.C. Hoppensteadt, Z. Jackiewicz, Numerical solution of a problem in the theory of epidemics, Appl. Numer. Math. 56 (2006) 533–543.
[4] J.C. Butcher, Z. Jackiewicz, Construction of diagonally implicit general linear methods of type 1 and 2 for ordinary differential equations, Appl. Numer. Math. 21 (1996) 385–415.
[5] W.M. Wright, The construction of order 4 DIMSIMs for ordinary differential equations, Numer. Algorithms 26 (2001) 123–130.
[6] J.C. Butcher, P. Chartier, Z. Jackiewicz, Nordsieck representation of DIMSIMs, Numer. Algorithms 16 (1997) 209–230.
[7] B. Owren, M. Zennaro, Derivation of efficient, continuous, explicit Runge–Kutta methods, SIAM J. Sci. Stat. Comput. 13 (1992) 1488–1501.
[8] S. Thompson, L.F. Shampine, A note on Zdzislaw's epidemic model, http://www.radford.edu/thompson/ffddes/Infection/infection.pdf.
[9] L.F. Shampine, I. Gladwell, S. Thompson, Solving ODEs with MATLAB, Cambridge University Press, Cambridge, 2003.
[10] S.A. Gourley, Y. Kuang, A stage structured predator–prey model and its dependence on maturation delay and death rate, J. Math. Biol. 49 (2004) 188–200.
[11] S.A. Gourley, Y. Kuang, A stage structured predator–prey model: importance of maturation delay, in: Talk at the Third International Conference on the Numerical Solution of Volterra and Delay Equations, Arizona State University, 2004, May 18–21.
[12] W.G. Aiello, H.I. Freedman, A time-delay model of single species growth with stage structure, Math. Biosci. 101 (1990) 139–153.
[13] Z. Jackiewicz, Convergence of multistep methods for Volterra functional differential equations, Numer. Math. 32 (1979) 307–332.
[14] Z. Jackiewicz, One-step methods of any order for neutral functional differential equations, SIAM J. Numer. Anal. 21 (1984) 486–510.
[15] J.K. Hale, S.M. Verduyn Lunel, Introduction to Functional Differential Equations, Springer-Verlag, 1993.
[16] Y. Kuang, Delay Differential Equations with Applications in Population Dynamics, Academic Press, Boston, 1993.
[17] H.L. Smith, An Introduction to Delay Differential Equations with Applications to the Life Sciences, in: Texts in Applied Mathematics, Springer, Berlin, 2011.
[18] H. Brunner, P.J. van der Houwen, The Numerical Solution of Volterra Equations, North-Holland, Amsterdam, New York, 1986.
[19] E. Hairer, S.P. Nørsett, G. Wanner, Solving Ordinary Differential Equations I. Nonstiff Problems, Springer-Verlag, Berlin, Heidelberg, New York, 1993.
[20] M. Zennaro, Natural continuous extensions of Runge–Kutta methods, Math. Comp. 46 (1986) 119–133.
[21] M. Zennaro, Natural Runge–Kutta methods and projection methods, Numer. Math. 53 (1988) 423–438.
[22] B. Owren, M. Zennaro, Order barriers for continuous explicit Runge–Kutta methods, Math. Comp. 56 (1991) 645–661.
[23] A. Bellen, M. Zennaro, Numerical Methods for Delay Differential Equations, in: Oxford Science Publications, Clarendon Press, Oxford, 2003.
[24] I. Gladwell, L.F. Shampine, R.W. Brankin, Automatic selection of the initial stepsize for ODE solver, J. Comput. Appl. Math. 18 (1987) 175–192.