NUMERICAL SOLUTION TO INITIAL VALUE PROBLEM FOR ONE CLASS OF DIFFERENTIAL EQUATION WITH MAXIMUM

Sergey N. Dashkovskiy\textsuperscript{1}, Olga D. Kichmarenko\textsuperscript{2,5}, Kateryna Yu. Sapozhnikova\textsuperscript{3}, Alexander N. Vityuk\textsuperscript{4}

\textsuperscript{1}University of Würzburg
Sanderring 2, 97070 Würzburg, GERMANY
\textsuperscript{2,4}Odesa National I.I. Mechnikov University
2, Dvoriaska, Odessa, 65000, UKRAINE
\textsuperscript{3}University of Applied Sciences Erfurt
Altonaer Str. 25, 99085 Erfurt, GERMANY

Abstract: This paper presents a numerical method for solution computation for one type of initial value problems given by differential equations with maximum of the unknown function over a prehistory. We prove the convergence of the method and illustrate it by several examples.

AMS Subject Classification: 34K, 37M99
Key Words: differential equations with maximum, numerical methods, nonlinear systems

1. Introduction

In many applications one has to deal with infinite dimensional systems which are hardly possible to solve analytically, especially in the nonlinear case. Here we consider a special subclass of functional-differential equations, and look for their numerical solutions. Computation of a solution requires studying functional-differential equations. Among the vast majority of functional-differential prob-
lems one of the most interesting is differential equation with maximum which contains a maximum of the studied function on a prehistory interval. This kind of problems have a great deal of applications in economics, ecology, biology, oscillation theory, automatic regulation etc. ([10],[11]). Such problems are often nonlinear, which makes them very challenging in view of control, stabilization, optimization and numerical solution. For example, in 1966 E.P.Popov considered a system for voltage regulation of a generator of constant current. The equation describing the work of the regulator involves the maximum of the unknown function and it has the form

\[ T_0u'(t) + u(t) + q \max_{s \in [t-h,t]} u(s) = f(t), \]

where \( T_0, q \) are constants characterizing the object, \( u(t) \) is the regulated voltage and \( f(t) \) is a perturbation. Similar problems were studied in [1],[2],[8]. For instance [2] provides a qualitative analysis of solutions to differential equations with maximum, see also [4].

Among recent developments for systems with maximum we recall stability results [12], averaging methods [9], where the dynamics can depend also on the maximum of the input function (see [6]). However many problems like stabilization or optimal control are not studied yet. In many applications such systems can be nonlinear. Due to nonlinearity analytic solution is in many cases not possible and a numerical method is needed. Standard methods cannot be applied directly in case of non-monotone solutions because of the maximum of the unknown function in the equation.

In this paper we consider the following problem

\[
\dot{x}(t) = f(t, x(t), \max_{\tau \in [t-h,t]} x(\tau)), \quad t \in [0, T],
\]

\[
x(t) = 0, \quad t \in [-h, 0],
\]

where \( h > 0 \) is a constant and \( T > 0 \) is finite. We are interested in numerical methods for solution calculation. Sufficient conditions of existence and uniqueness of solutions for this problem were obtained in [13]. Recall that in [5] an algorithm for constructions of two sequences of successive approximations of the solution to the initial value problem for nonlinear differential equations with maximum was developed and is based on the monotone iterative techniques. Here we propose a numerical method based on a simpler approach and illustrate it by several examples.
2. Notation and Preliminaries

Let $f$ from (1) satisfy the following

**Assumption 1.** $f : [0, T] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is

(i) continuous in all variables, and

(ii) satisfy the Lipschitz condition

$$|f(t, x, y) - f(t, \bar{x}, \bar{y})| \leq K|x - \bar{x}| + L|y - \bar{y}| \quad \forall t, x, y$$  \hspace{1cm} (3)

where $K > 0, L > 0$ are Lipschitz constants.

Under a solution to the problem (1),(2) we understand a function $x \in C^1([-h, T])$ satisfying (1) and the initial condition (2). Assumption 1 implies its existence and uniqueness.

For the reader’s convenience and to make our paper self-contained we present the following theorem with proof.

**Theorem 2** ([13]). Let $f$ be continuous with

$$f(0, 0, 0) = 0.$$  \hspace{1cm} (4)

Then the problem (1),(2) is equivalent to the equation

$$v(t) = \begin{cases} 
0, & t \in [-h, 0], \\
\frac{d}{dt}x(t) + \max_{\tau \in [t-h, t]} \int_0^\tau v(s)ds, & t \in [0, T],
\end{cases}$$  \hspace{1cm} (5)

and if $v \in C([-h, T])$ is a solution to (5), then

$$x(t) = \begin{cases} 
0, & t \in [-h, 0], \\
\int_0^t v(s)ds, & t \in [0, T]
\end{cases}$$

solves the problem (1),(2).

**Proof.** Note that (4) guarantees the continuity of $v$ at $t = 0$. Let $x \in C^1([-h, T])$ be a solution to (1),(2). Define $v = \dot{x} \in C([-h, T])$, then

$$\max_{\tau \in [t-h, t]} x(\tau) = \max_{\tau \in [t-h, t]} \int_0^\tau v(s)ds,$$
since
\[
x(t) = \begin{cases} 
0, & t \in [-h, 0], \\
\int_0^t v(s)ds, & t \in [0, T]. 
\end{cases}
\]

Since \(x\) satisfies (1), (2) it follows that \(v\) solves (5).

Conversely, let \(v \in C([-h, T])\) be a solution to (5). Define
\[
x(t) = \begin{cases} 
0, & t \in [-h, 0], \\
\int_0^t v(s)ds, & t \in [0, T]. 
\end{cases}
\]

Taking the time derivative of \(x\) we obtain a continuous function
\[
\dot{x}(t) = \begin{cases} 
0, & t \in [-h, 0], \\
v(t), & t \in [0, T], 
\end{cases}
\]
that satisfies (1), (2) due to (5).

**Definition 3** ([7]). The modulus of continuity of a function \(f : [0, T] \to \mathbb{R}\) is a function defined by:
\[
\omega_f(\delta) = \sup \{|f(t_1) - f(t_2)| : t_1, t_2, \in [0, T]; |t_1 - t_2| \leq \delta\}.
\]

It is known that \(f \in C[0, T]\) if and only if \(\lim_{\delta \to 0} \omega_f(\delta) = 0\) holds.

Our numerical approach to the problem (1), (2) is based on left rectangle method of integration, which requires only the \(C^1\) property for \(f\) to derive the error estimation. Recall that for the trapezoid or Simpson method more regularity is required.

Consider
\[
\int_a^b f(x)dx = l \sum_{k=0}^{n-1} f_k + r_n(f), \quad \text{where} \quad f(x) \in C([a, b]), \quad b - a = nl, \quad x_k = kl, \quad f_k = f(x_k), \quad l \text{ is integration step.}
\]

Then
\[
|r_n(f)| = \left| \sum_{k=0}^{n-1} \int_{x_k}^{x_{k+1}} (f(x) - f_k) \, dx \right| \leq \sum_{k=0}^{n-1} \int_{x_k}^{x_{k+1}} |f(x) - f_k| \, dx.
\]

For \(x \in [x_k, x_{k+1}]\) we have \(|f(x) - f_k| \leq |f(x) - f(x_k)| + |f(x_k) - f_k| \leq \delta_k + \omega_f(l)\) with \(\delta_k = |f(x_k) - f_k|\). Hence
\[
|r_n(f)| \leq l \sum_{k=0}^{n-1} \delta_k + l \omega(f; l). \tag{6}
\]

This proves the following
Lemma 4. If $f(x) \in C([a, b])$, then for the error of the left rectangle method the estimation (6) is valid.

We also recall the following related result:

Lemma 5. ([14]) Let on the uniform grid $T_N = \{t_k = kl, k = 0, N, Nl = a\}$ the grid functions $q : T_N \to \mathbb{R}, w : T_N \to [0, \infty)$ be such that

$$
\begin{cases}
|q_{n+1}| \leq Ql \sum_{k=0}^{n-1} |q_k| + S, \\
w_{n+1} \geq Ql \sum_{k=0}^{n-1} w_k + S, n = 0, N - 1, \\
q_0 = 0, w_0 = S, S > 0, Q > 0.
\end{cases}
$$

Then $|q_n| \leq w_n, n = 0, N,$ and $w_n = z(t_n),$ satisfies (7), where $z(t) = Se^{Qt}$ is a solution to integral equation

$$z(t) = Q \int_0^t z(\tau) d\tau + S.$$

3. Numerical Method

Let $l$ be the step of integration, $T = Nl, t_k = kl, k = 0, N$. By $x_k$ we denote the approximation of solution $x(t_k)$ at $t = t_k$ to the problem (1),(2), also by $v_k$ we denote the approximation of solution $v(t_k)$ of (5). Assume, that $x_k$ and $v_k, k = 0, N$ are already found. Let us show how $x_{n+1}, v_{n+1}$ can be calculated. Note, that for $t \in [0, T], \ x(t) = \int_0^t v(s) ds$. By the left rectangle formula we obtain

$$x(t_{n+1}) = \int_0^{t_{n+1}} v(s) ds = \sum_{k=0}^n x_{k+1} \int_{t_k}^{t_{k+1}} v(s) ds \approx l \sum_{k=0}^n v_k.$$

Hence $x_{n+1} = l \sum_{k=0}^n v_k$. Let $\delta_k = |v(t_k) - v_k|$, then

$$|x(t_{n+1}) - x_{n+1}| \leq \sum_{n=0}^k \int_{t_k}^{t_{k+1}} |v(s) - v_k| ds.$$
For \( t \in [t_k, t_{k+1}] \) we can write

\[ |v(t) - v_k| \leq |v(t) - v(t_k)| + |v(t_k) - v_k| \leq \omega(v; l) + \delta_k. \tag{9} \]

From (8), (9) and Lemma 4 follows that

\[ |x(t_{n+1}) - x_{n+1}| \leq l \sum_{n=0}^{k} \delta_k + T\omega(v; l). \tag{10} \]

According to (5) if \( t \geq 0 \), then \( v(t) = f(t, \int_0^t v(s)ds, \max_{\tau \in [t-h,t]} x(\tau)) \). Then \( v(t_{n+1}) = f(t_{n+1}, x(t_{n+1}), \max_{\tau \in [t_{n+1}-h,t_{n+1}]} x(\tau)) \). Let \( \tau \) be such that \( x(\tau) = \max_{\tau \in [t_{n+1}-h,t_{n+1}]} x(\tau) \). Let us calculate an approximation \( x_{n+1} \) of \( x(\tau) \). If \( \tau \in [-h,0] \cap [t_{n+1}-h, t_{n+1}] \), then \( x(\tau) = 0 \), since \( v(t) = 0 \) for \( t \in [-h,0] \). Let \( \tau \) belong to one of the intervals \( [t_{n+1}-k, t_{n+1}-2-k] \), \( 1 \leq k \leq N \), and \( t_{n+1}-k \geq 0 \). If \( \tau = t_{n+1}-k \), then using left rectangle formula we get

\[ x(\tau) = \int_0^{t_{n+1}-k} v(s)ds = \sum_{i=0}^{n-k} \int_{t_i}^{t_{i+1}} v(s)ds \approx l \sum_{i=0}^{n-k} v(t_i) \approx l \sum_{i=0}^{n-k} v_i = \bar{x}_{n+1}. \]

From (10) we obtain

\[ |x(\tau) - \bar{x}_{n+1}| \leq l \sum_{i=0}^{n-k} \delta_i + a\omega_v(l). \tag{11} \]

Now consider the case \( t_{n+1}-k < \tau < t_{n+1}-2-k \). Then

\[ x(\tau) = \int_0^{t_{n+1}-k} v(s)ds + \int_{t_{n+1}-k}^{\tau} v(s)ds \approx l \sum_{i=0}^{n-k} v(t_i) + v(t_{n+1}-k)(\tau - t_{n+1}-k) \approx l \sum_{i=0}^{n-k} v(t_i) \approx l \sum_{i=0}^{n-k} v_i. \]

In this case \( \bar{x}_{n+1} = l \sum_{i=0}^{n-k} v_i \) and the following estimation holds

\[ |x(\tau) - \bar{x}_{n+1}| \leq l \sum_{i=0}^{n-k} \delta_i + l\omega_v(l). \tag{12} \]
If \( \tau = t_{n+2-k} \), then analogically follows that \( \overline{x}_{n+1} = \sum_{i=0}^{n+1-k} v_i \) and estimation (12) is valid.

Let now \( k = 1 \), \( \tau \in [t_n, t_{n+1}] \). In this case \( \overline{x}_{n+1} = \sum_{i=0}^{n} v_i \),

\[
|x(\tau) - \overline{x}_{n+1}| \leq l \sum_{i=0}^{n} \delta_i + l\omega_v(l). \tag{13}
\]

Therefore, in all considered cases we get \( \overline{x}_{n+1} \) which is the approximate value of \( \overline{x} = \max_{\tau \in [t_n - h, t_{n+1}]} x(\tau) \). For \( \overline{x}_{n+1} \) the estimation (13) is valid. Hence we set

\[
v_{n+1} = f(t_{n+1}, x_{n+1}, \overline{x}_{n+1}). \tag{14}
\]

The next theorem shows that the presented numerical method converges if

\[
\lim_{N \to \infty} \max_k |v(t_k) - v_k| = 0, \tag{15}
\]

\[
\lim_{N \to \infty} \max_k |x(t_k) - x_k| = 0, \tag{16}
\]

**Theorem 6.** Let \( f \) satisfy Assump. 1 and \( f(0,0,0) = 0 \). Then (15),(16) hold for the approximated solutions to (1),(2) and (5).

**Proof.** Using our previous notation and from (10),(13) we have

\[
\delta_{n+1} = |v(t_{n+1}) - v_{n+1}| =
\]

\[
= |f(t_{n+1}, x_{n+1}, \max_{\tau \in \sigma_{(t_{n+1})}} x(\tau)) - f(t_{n+1}, x_{n+1}, \overline{x}_{n+1})| \leq
\]

\[
\leq K |x(t_{n+1}) - x_{n+1}| + L |x(\tau) - \overline{x}_{n+1}| \leq
\]

\[
\leq K \left( \sum_{k=0}^{n} \delta_k + l\omega_v(l) \right) + L \left( \sum_{k=0}^{n} \delta_k + T\omega_v(l) \right) = Al \sum_{k=0}^{n} \delta_k + \gamma
\]

where \( A = K + L, \quad \gamma = AT\omega_v(l), \quad n = 0, N - 1 \). Applying Lemma 5 in case \( Q = A, S = \gamma \) we get

\[
\delta_n \leq z(t_n) = Se^{Qt_n} \leq Se^{QT}.
\]

Therefore, \( \lim_{N \to \infty} \max_{n=0,N} \delta_n = 0 \), since \( \gamma \to 0 \). From (10) follows that

\[
|x(t_{n+1}) - x_{n+1}| \leq l \sum_{k=0}^{n} Se^{QT} + l\omega_v(l) = l (Se^{QT} + \omega_v(l)),
\]

and hence \( \lim_{N \to \infty} \max_{n=0,N} |x(t_n) - x_n| = 0 \). \( \square \)
Remark 7. This method can be applied for problems with nonzero initial condition:
\[
\dot{x}(t) = f(t, x(t), \max_{\tau \in [t-h,t]} x(\tau)), \ t \in [0, T],
\]
\[
x(t) = \alpha, \ t \in [-h, 0], \ \alpha = \text{const.}
\] (17)
\[
\dot{x}(t) = f(t, x(t), x(t)), \ t \in [0, T], \ x(0) = 0,
\]
\[
x(t) = 0, \ t \in [-h, 0],
\] (18)
Let \( f(0, x(0), \max_{\tau \in [-h,0]} x(\tau)) = f(0, \alpha, \alpha) = 0 \). Assume \( u(t) = \alpha, \ t \in [-h, T] \). Let \( y \) be defined by \( x(t) = y(t) + u(t) \). It can be shown (see [13]), that \( y(t) \) is the solution to the problem
\[
\dot{y}(t) = f(t, y(t) + \alpha, \max_{\tau \in [t-h,t]} y(\tau) + \alpha), \ t \in [0, T],
\]
\[
y(t) = 0, \ t \in [-h, 0],
\]
and \( f(0, y(0), \max_{\tau \in [-h,0]} y(\tau)) = f(0, \alpha, \alpha) = 0 \). Hence there exists a unique solution to (17),(18).

Note, that the proposed numerical method to compute a solution to (1), (2) coincides with Euler method for the problem
\[
\dot{x}(t) = f(t, x(t), x(t)), \ t \in [0, T], \ x(0) = 0,
\]
in case the solution for (1), (2) is not decreasing.

4. Examples

Example 1. Consider the following differential equation
\[
\begin{cases}
\dot{x}(t) = t + \max_{\tau \in [t-h,t]} x(\tau), & t \in [0, T] \\
x(t) = 0, & t \in [-h, 0],
\end{cases}
\] (19)
with \( h = 0.5, T = 1 \). For this example we can compare the numerical solution which we get by the proposed method with the analytic solution. Using theorem about differential inequalities it easy to prove that the solution \( x \) is a monotonically increasing function, that is why \( \max_{\tau \in [t-h,t]} x(\tau) = x(t) \). So the Cauchy problem (19) is equivalent to:
\[
\begin{cases}
\dot{x}(t) = t + x(t), & t \in [0, T] \\
x(0) = 0, & t \in [-h, 0].
\end{cases}
\] (20)
Figure 1: Solutions to (19) and (20) for different values of $l$

Table 1: Estimation of the difference between analytic and numerical solutions for the Example 1

<table>
<thead>
<tr>
<th>$t_k$</th>
<th>$0.1$</th>
<th>$0.3$</th>
<th>$0.5$</th>
<th>$0.7$</th>
<th>$1.0$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x(t_k)$</td>
<td>0.0052</td>
<td>0.0499</td>
<td>0.1487</td>
<td>0.3138</td>
<td>0.7183</td>
</tr>
<tr>
<td>$l = 0.1$, $x_k$</td>
<td>0</td>
<td>0.0310</td>
<td>0.1105</td>
<td>0.2487</td>
<td>0.5937</td>
</tr>
<tr>
<td>$</td>
<td>x(t_k) - x_k</td>
<td>$</td>
<td>0.0052</td>
<td>0.0189</td>
<td>0.0382</td>
</tr>
<tr>
<td>$l = 0.05$, $x_k$</td>
<td>0.0025</td>
<td>0.0401</td>
<td>0.1289</td>
<td>0.2799</td>
<td>0.6533</td>
</tr>
<tr>
<td>$</td>
<td>x(t_k) - x_k</td>
<td>$</td>
<td>0.0027</td>
<td>0.0098</td>
<td>0.0198</td>
</tr>
<tr>
<td>$l = 0.0025$, $x_k$</td>
<td>0.0050</td>
<td>0.0494</td>
<td>0.1477</td>
<td>0.3120</td>
<td>0.7149</td>
</tr>
<tr>
<td>$</td>
<td>x(t_k) - x_k</td>
<td>$</td>
<td>0.0002</td>
<td>0.0005</td>
<td>0.0010</td>
</tr>
</tbody>
</table>

The solution to the problem (20) is $x(t) = e^t - t - 1$.

Figure 1 illustrates the approximations of the solution to the problem with maximum (19) by the proposed numerical method and the analytic solution to the problem (20). Table 1 shows the values of $x_k$ – numerical solution by the proposed method and $x(t_k)$ – the analytic solution calculated at points $t_k$. In the same table the estimation between analytic and numerical solutions for steps $l = 0.1$, $l = 0.05$, $l = 0.0025$ are presented. In particular, for the step $l = 0.1$ we get maximum value of $|x(t_k) - x_k|$ on an interval $[0, 1]$ which is 0.1245, for the step $l = 0.05$ the maximum value is 0.0650, and for $l = 0.0025$ is 0.0034. Thus for step $l$ we can point numbers $\varepsilon_l$ such that $|x(t_k) - x_k| \leq \varepsilon_l$, and $\varepsilon_{0.1} = 0.1245, \varepsilon_{0.05} = 0.0650, \varepsilon_{0.0025} = 0.0034$.

Example 2. Now we consider differential equation with maximum and with constant initial function where the solution is not monotone, so that the ”max”
sign cannot be omitted, but such that an analytic solution can be derived with some more effort:

\[
\begin{aligned}
\dot{x}(t) &= t - \max_{\tau \in [t-h, t]} x(\tau), \quad t \in [0, T] \\
x(t) &= 1, \quad t \in [-h, 0].
\end{aligned}
\]

(21)

where \( h = 0.5, \ T = 5 \). The analytic solution to the problem (21) can be written (after simplifications of numbers given by large expressions with fractions with roots) as follows

\[
x(t) = \begin{cases} 
\frac{t^2}{2} - t + 1, & t \in [0, 0.5], \\
-\frac{t^3}{6} + \frac{5}{4}t^2 - \frac{13}{8}t + \frac{55}{48}, & t \in [0.5, 1.02], \\
t - 1 + 1.64e^{-t}, & t \in [1.02, \infty].
\end{cases}
\]

According to Remark 7 in order to apply our numerical method to the problem (21) we first have to solve the following problem

\[
\begin{aligned}
\dot{y}(t) &= t - \max_{\tau \in [t-h, t]} y(\tau) + 1, \quad t \in [0, T] \\
y(t) &= 0, \quad t \in [-h, 0],
\end{aligned}
\]

(22)

where \( h = 0.5, \ T = 5 \). Then we compute the solution to (21) as \( x(t) = y(t) + 1 \), where \( y(t) \) is the solution to the problem (22) by the proposed numerical method with step \( l = 0.05 \). Note that the graph of the true solution coincides with the graph of the numerical solution. Here \( |x(t_k) - x_k| \leq \varepsilon_{0.05} = 0.0125 \).

Figure 2: Plot of numerical solution to the problem (21)
Example 3. In the last example we consider a system with maximum described by the following differential equation

\[
\begin{align*}
\dot{x}(t) &= e^{-t} \sin t - t \max_{\tau \in [t-h,t]} x(\tau), \quad t \in [0, T] \\
x(t) &= 1, \quad t \in [-h, 0],
\end{align*}
\]

where \( h = 1, T = 4 \). In this case one can verify that \( \dot{x}(t) < 0, \quad t \in [0, T] \) so that the problem (23) can be reduced to the following problem with time delay because of the monotonicity of the solution

\[
\begin{align*}
\dot{x}(t) &= e^{-t} \sin t - tx(t-h), \quad t \in [0, T] \\
x(t) &= 1, \quad t \in [-h, 0].
\end{align*}
\] (24)

On the following Figure 3 we present the graphs of numerical solutions calculated by the proposed method to the problem (23) (note that we apply Remark 7) with the steps \( l = 0.1, \quad l = 0.05, \quad l = 0.0025 \). There exist \( \varepsilon_l \) such that \( |x(t_k) - x_k| \leq \varepsilon_l \) and for particular steps we can specify \( \varepsilon_{0.1} = 0.0983, \quad \varepsilon_{0.05} = 0.0407, \quad \varepsilon_{0.0025} = 0.0018 \).

![Figure 3: Plot of solutions to problem (23) with different steps.](image-url)
5. Conclusion

In this work we have proposed a simple but efficient numerical method to calculate solutions for initial value problems of one type of differential equations with maximum and proved its convergence. From examples it can be seen that the estimation in (10) is confirmed. As a continuation of this research our numerical method can be extended to a problem of a more general type such as

$$\dot{x}(t) = f(t, x(t), \max_{\tau \in [g(t), \gamma(t)]} x(\tau)), \quad t \in [t_0, T],$$

$$x(t) = \varphi(t), \quad t \in [\inf_{t \geq 0} g(t), t_0].$$

where $g(t) \leq \gamma(t) \leq t$, $t_0 \geq \inf_{t \geq 0} g(t)$. Also a numerical method for computation of solutions to initial value problem for differential equations with maximum in case when the right hand side is a measurable function will be a matter of our future investigation.

Acknowledgments

Kateryna Yu. Sapozhnikova thank Ernst-Abbe Foundation for financial support.

References


