An efficient approach for solving stiff nonlinear boundary value problems

V. L. Makarov[*] D. V. Dragunov[†]

October 2015

Abstract

A new method for solving stiff boundary value problems is described and compared to other known approaches using the Troesch's problems as a test example. A c++ implementation of the proposed method is available at https://github.com/imathsoft/MathSoftDevelopment.

AMS 2010 Subject Classification: 65L04, 65L05, 65L10, 65L20, 65L50, 65Y15
Key Words: stiff boundary value problem, simple shooting method, multiple shooting method, the Troesch's problem

1 Introduction

In the current paper we consider the following nonlinear boundary value problem

$$\frac{d^2u(x)}{dx^2} = N\left(u(x), x\right) u(x), \quad x \in [a,b]\,,\ N(u,x) \in C^2(\mathbb{R} \times [a,b]), \tag{1}$$

$$u(a) = u_l \in \mathbb{R}, \quad u(b) = u_r \in \mathbb{R}, \tag{2}$$

which arises in many areas of physics and mathematics. Although, there is a huge variety of known methods for solving problems of type (1), (2) (see, for example [7], [6], [2] and the references therein), almost none of them fill comfortable when the problem turns out to be stiff.

As a quite famous example of stiff problems it is worth to mention the Troesch's problem:

$$\frac{d^2u(x)}{dx^2} = \lambda \sinh\left(\lambda u(x)\right), \quad x \in [0,1] \tag{3}$$

$$u(0) = 0, \quad u(1) = 1, \tag{4}$$

[*]Institute of Mathematics of NAS of Ukraine, e-mail: makarov@imath.kiev.ua
[†]Institute of Mathematics of NAS of Ukraine, e-mail: dragunovdenis@gmail.com

which is a partial case of problem (1), (2) with $N(u(x), x) \equiv \lambda \sinh(\lambda u(x))/u(x)$, $a = u_l = 0$, $b = u_r = 1$. It is well known, that the problem (3), (4) is inherently unstable and difficult (see [1], [3], [13], [8], [14] and the references therein), in the sense that

$$\lim_{\substack{\lambda\to+\infty\\x\to 0+}} u'_x(x,\lambda) = 0, \qquad \lim_{\substack{\lambda\to+\infty\\x\to 1-}} u'_x(x,\lambda) = +\infty.$$

Talking about the known methods for solving BVPs, it is impossible not to mention the *simple shooting method* (SSM) and the *multiple shooting method* (MSM) [11, Section 7.3] which are two the must simple and reliable techniques to deal with boundary value problems of type (1), (2). By calling them *techniques* and not just *methods* we would like to emphasize a fact that the basic idea behind them is very broad and can be used in many different modifications, which, in turn, might be called the *methods*. Since definitions of both SSM and MSM essentially relay on using methods for solving *initial value problems* (IVP), one of the ways to modify the methods consists in using different IVP solvers. Below we are going to adopt (or modify if you wish) the SSM and MSM for using a specific approach for numerical solution of IVP's called Straight-Inverse method (or, simply SI-method). The true meaning of the method's name will be explained in the coming sections, and for now we just note that the SI-method presented below has a genuine "immunity" towards problems whose solutions possess rapid variation (i.e. big values of the first derivative). In what follows we will use the Troesch's problem as a testing instrument to reveal strengths and weaknesses of the proposed modifications of the SSM and MSM via using SI-method.

2 Step functions

In this section we are going to introduce a pair of, so called, *step functions*, which play a crucial role throughout the rest of the paper.

Let us define the step function $U(h)$ to be the solution for the following initial value problem:

$$\frac{d^2U(s)}{ds^2} = (As+B)U(s), \quad U(0) = D; \quad U'(0) = C, \quad s \in \mathbb{R}. \tag{5}$$

It is easy to check that the function $U(s)$ can be expressed explicitly through the Airy functions (see [9, 283]) in the following way:

$$U(s) = C_1\,\mathbf{Ai}\left(\frac{As+B}{(-A)^{2/3}}\right) + C_2\,\mathbf{Bi}\left(\frac{As+B}{(-A)^{2/3}}\right), \tag{6}$$

where

$$C_1 = \left(-\mathbf{Bi}\left(\frac{B}{(-A)^{2/3}}\right) C\,(-A)^{2/3} + D\,A\,\mathbf{Bi}\left(1, \frac{B}{(-A)^{2/3}}\right)\right)\,E,$$

$$C_2 = \left(-C\,(-A)^{2/3}\,\mathbf{Ai}\left(\frac{B}{(-A)^{2/3}}\right) + D\,A\,\mathbf{Ai}\left(1, \frac{B}{(-A)^{2/3}}\right)\right)\ E,$$

$$E = A^{-1}\left(\mathbf{Bi}\left(\frac{B}{(-A)^{2/3}}\right)\mathbf{Ai}\left(1, \frac{B}{(-A)^{2/3}}\right) - \mathbf{Bi}\left(1, \frac{B}{(-A)^{2/3}}\right)\mathbf{Ai}\left(\frac{B}{(-A)^{2/3}}\right)\right)^{-1}.$$

Similarly to this, we define step function $V(s)$ as solution for the problem

$$\frac{d^2V(s)}{ds^2} = (As+B)\frac{dV(s)}{ds}, \quad V(0) = D, \quad V'(0) = C, \quad s \in \mathbb{R}, \tag{7}$$

possessing explicit general solution in the form of

$$V(s) = D + C\int\limits_0^s \exp\left(\frac{A}{2}\xi^2 + B\xi\right)d\xi. \tag{8}$$

Though explicit, formulas (6) and (8) are quite difficult to evaluate, especially when $\|A\|$ is quite small. However, in what follows, we are going to use functions $U(s)$ and $V(s)$ with $\|s\|$ quite small, which allows us to use more convenient approach for evaluating them instead of formulas (6), (8). The approach naturally follows from the theorems below.

Theorem 1.

$$U(s) = U(A,B,C,D,s) = \lim_{n\to+\infty} U_n(A,B,C,D,s), \tag{9}$$

where the successive approximation $U_n(A,B,C,D,s)$ can be found recursively

$$\begin{aligned} U_n(A,B,C,D,s) &= \int_0^s \left(\int_0^\xi (A\eta+B)U_{n-1}(A,B,C,D,\eta)d\eta\right)d\xi + Cs + D, \\ U_0(A,B,C,D,s) &= 0 \end{aligned} \tag{10}$$

and the following estimation holds true:

$$\|U_{n+1}(A,B,C,D,s) - U_n(A,B,C,D,s)\| \le \frac{(\|As\|+\|B\|)^n(\|Cs\|+\|D\|)\|s\|^{2n}}{(2n!)}. \tag{11}$$

Theorem 2.

$$V(s) = V(A,B,C,D,s) = \lim_{n\to+\infty} V_n(A,B,C,D,s), \tag{12}$$

where the successive approximations $V_n(A,B,C,D,s)$ can be found recursively

$$\begin{aligned} V_n'(A,B,C,D,s) =& \int_0^s (A\eta + B)V_{n-1}'(A,B,C,D,\eta)d\eta + C, \\ V_n(A,B,C,D,s) =& \int_0^s V_n'(A,B,C,D,\xi)d\xi + D, \\ V_0'(A,B,C,D,s) =& 0, \end{aligned} \tag{13}$$

and the following estimation holds true:

$$\|V_{n+1}(A,B,C,D,s) - V_n(A,B,C,D,s)\| \leq \frac{(\|As\ \ + \|B\||)^n \|C\| \|s\|^{n+1}}{n!}. \tag{14}$$

Equalities (9), (10) can be considered as another definition of the step function $U(s)$, as well as equalities (12), (13) can be used as another definition for $V(s)$. At the same time, the mentioned equalities make sense not only when A, B, C, D, s are real numbers but also when they are square matrices from the linear space $M_n(\mathbb{R})$[1]. The appearance of the norm notation $\|\cdot\|$ and not just absolute value $|\cdot|$ in the estimations (11), (14) is intended to cover the case of matrix arguments, which is important for us since using of matirx arquments provides us an easy and elegant way to compute partial derivatives of $U(A,B,C,D,s)$ and $V(A,B,C,D,s)$. The key to understanding what is meant lies in the following formula (see, for example [4, p. 98]):

$$f(J_2(a)) = f\left(\begin{bmatrix} a & 1 \\ 0 & a \end{bmatrix}\right) = \begin{bmatrix} f(a) & f'(a) \\ 0 & f(a) \end{bmatrix}, \tag{15}$$

which holds true for every function $f(x)$ defined in $x = a$ together with its first derivative. From (15) it follows that to calculate $U_{a_i}'(a_1, \ldots, a_5)$, for $i = 1, 2, \ldots, 5$, we can use the linear matrix equation

$$\begin{aligned} &U(a_1E_2 + \delta_{i1}J_2(0), a_2E + \delta_{i2}J_2(0), \ldots, a_5E_2 + \delta_{i5}J_2(0)) = \\ =&U(a_1, \ldots, a_5)E_2 + U_{a_i}'(a_1, \ldots, a_5)J_2(0), \end{aligned} \tag{16}$$

$$E_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \ \delta_{ij} = \begin{cases} 1, & i = j, \\ 0, & i \neq j. \end{cases}$$

Of course, the same remains true as related to the function V. Here we are not going to discuss the efficiency of the proposed approach for finding partial derivatives of the step functions (and this might be a good topic for further investigations), however, we would like to mention that on practice it proved to be a quite efficient, reliable and easy to use

[1]Linear space of $n \times n$ matrices over the field of real numbers

algorithm[2]. Another thing worth mentioning is that equation (16) should not be used for calculating partial derivatives of $U(A,B,C,D,s)$ and $V(A,B,C,D,s)$ with respect to s. The thing is that no matter which matrix norm $\|\cdot\|$ one would use, the right hand sides of inequalities (11), (14) tend to 0 as $s\to 0$, and they do that much faster when s is substituted with sE_2 than when it is substituted with $J_2(s)$ (especially when $\|A\|$ is quite big, which is the case for stiff problems). At the same time the most natural way to calculate $V'_s(s)$ follows directly from the Theorem 2, and to calculate $U'_s(s)$ it is easier to use the formula

$$U'_s(s)=\int_0^s(A\eta+B)U(\eta)d\eta+C. \tag{17}$$

The importance of the *step functions* introduced above can be explained by the approximation properties they possess. The properties are described in the theorems below.

Theorem 3. *Let $u(x)$ be the solution to equation* (1) *supplemented by the following initial conditions:*

$$u(a)=u_l\in\mathbb{R},\quad u'(a)=u'_l\in\mathbb{R}. \tag{18}$$

Then, for the sufficiently small $h>0$, the inequalities

$$\|u(a+x)-U(A,B,C,D,x)\|_{x\in[-h,h]}\le h^4K\|u(x)\|_{x\in[-h,h]}\exp(hM), \tag{19}$$

$$\|u'(a+x)-U'_x(A,B,C,D,x)\|_{x\in[-h,h]}\le h^3K\|u(x)\|_{x\in[-h,h]}\exp(hM) \tag{20}$$

hold true, where

$$\begin{aligned}A=&N'_u(u_l,a)u'_l+N'_x(u_l,a),\\ B=&N(u_l,a),\\ C=&u'_l,\\ D=&u_l,\end{aligned} \tag{21}$$

$$\begin{aligned}K&=\frac{1}{2}\left\|(N(u(x),x))''_{u^2}\right\|_{x\in[-h,h]}=\\ &=\frac{1}{2}\|N''_{u^2}(u(x),x)\left(u'(x)\right)^2+2N''_{u\,x}(u(x),x)u'(x)+\\ &+N''_{x^2}(u(x),x)+N'_u(u(x),x)N(u(x),x)u(x)\|_{x\in[-h,h]},\\ M&=\max\{1,|B|+|A|h\}\end{aligned} \tag{22}$$

and

$$\|f(x)\|_{x\in[a,b]}\overset{def}{=}\max_{\forall x\in[a,b]}|f(x)|,\ f(x)\in C([a,b]),$$

[2]The algorithm is used in the implementation available at https://github.com/imathsoft/MathSoftDevelopment.

$$\|f(x)\|_{1,x\in[a,b]} \stackrel{def}{=} \max\{\|f(x)\|_{x\in[a,b]}, \|f'(x)\|_{x\in[a,b]}\}.$$

Proof. From the assumption (1) about smoothness of the function $N(u,x)$ and the Picard-Lindelöf theorem (see, for example [12, p. 38]) it follows that the solution $u(x)$ of the IVP (1), (18) exists at least in some closed neighborhood $\overline{B_\delta(a)} = [a-\delta, a+\delta], \delta > 0$ of the point $x = a$. From now one assume that $h = \delta$ and the solution $u(x) \in C^2\left(\overline{B_\delta(a)}\right)$ is known.

Combining equations (1) and (5) we come to a linear system of the first order ordinary differential equations with respect to unknown vector-function $Z(x) = [z(x), z'(x)]^T$, $z(x) = u(a+x) - U(A,B,C,D,x)$:

$$\dot{Z}(x) = \begin{bmatrix} 0 & 1 \\ B(a) + A(a)(x-a) & 0 \end{bmatrix} Z(x) + \begin{bmatrix} 0 \\ F(x)u(x) \end{bmatrix}, \tag{23}$$

supplemented with zero initial conditions

$$Z(a) = 0, \tag{24}$$

where

$$\begin{aligned} A(x) &= (N(u(x),x))'_x, \\ B(x) &= N(u(x),x), \\ F(x) &= N(u(x),x) - B(a) - A(a)(x-a). \end{aligned}$$

From (23), (24) it follows that

$$\|Z(x)\| \stackrel{def}{=} \max\{|z(x)|, |z'(x)|\} \le M \begin{cases} \int\limits_0^x \|Z(\xi)\| d\xi, & x \in [0,h] \\ \int\limits_x^0 \|Z(\xi)\| d\xi & x \in [-h,0) \end{cases} \tag{25}$$
$$+ h^3 K \|u(x)\|_{x\in[-h,h]},\ x \in [-h,h].$$

The Gronwall's inequality (see, for example, [12, 42]), being applied to (25), leads us to the inequality

$$\|Z(x)\|_{x\in[-h,h]} \le h^3 K \|u(x)\|_{x\in[-h,h]} \exp(hM)$$

which immediately implies estimation (20). Estimation (19) follows from the one (20) and the fact that

$$\|z(x)\|_{x\in[-h,h]} = \| \int_0^x z'(\xi) d\xi \|_{x\in[-h,h]}.$$

□

Lemma 1. *Let $u'_l \neq 0$. Then in some neighborhood $B(u_l, h) = (u_l - h; u_l + h)$ of the point*

u_l there exists the unique solution $x(u)$ to the IVP

$$\frac{d^2x(u)}{du^2} = -N\left(u, x(u)\right) u \left(\frac{dx(u)}{du}\right)^3, \tag{26}$$

$$x(u_l) = a, \quad x'(u_l) = x'_l = 1/u'_l \tag{27}$$

and $x(u)$ is the inverse function to $u(x)$ satisfying (1), (18), i.e. $\forall u_1 \in B(u_l, h)$ we have that $x(u_i) \in \bar{B}(a, h_1)$ for some $h_1 > 0$ and $u_1 \equiv u(x(u_1))$.

Theorem 4. *Let $x(u)$ be a function satisfying IVP (26), (27) in some neighborhood $B(u_l, h)$. Then*

$$\|x(u_l + u) - V(A, B, C, D, u)\|_{u\in[-h,h]} = h^4 K \|x(u)\|_{u\in[-h,h]} \exp(hM), \tag{28}$$

$$\|x'(u_l + u) - V'_u(A, B, C, D, u)\|_{u\in[-h,h]} = h^3 K \|x(u)\|_{u\in[-h,h]} \exp(hM), \tag{29}$$

where

$$\begin{aligned}
A =& -\left((N'_u(u_l, a) + N'_x(u_l, a)x'_l)u_l + N(u_l, a)\right)(x'_l)^2 + 2\left(N(u_l, a)u_l\right)^2 (x'_l)^4,\\
B =& - N(u_l, a)u_l\,(x'_l)^2,\\
C =& x'_l,\\
D =& a,\\
K =& \frac{1}{2}\left\|\left(N\left(u, x(u)\right) u \left(x'(u)\right)^2\right)''_{u^2}\right\|_{u\in[-h,h]},\\
M =& |B| + |A|h.
\end{aligned} \tag{30}$$

Proof. The proof is similar to Theorem 3. So we skip it. □

3 Straight-Inverse method for solving IVPs for the second order differential equations

3.1 Preliminary comments

In this section we describe the SI-method for solving equation (1) subjected to the initial conditions (18). However, before doing that, let us first introduce a list of requirements which we expect the SI-method to fulfill, and which, historically, led us to the SI-method as such:

1. the method should approximate the exact solution $u(x)$ for IVP (1), (18) on a discrete mesh $\varnothing \neq \omega \in [a, b]$, which should depend on the problem itself and on the desired accuracy of approximation;

2. for a given positive $h \in \mathbb{R}$, the method should provide an algorithm to construct a mesh $\omega(h) \in [a, b]$, containing $N_\omega \leq h^{-1} \int_a^b \sqrt{1 + (u'(x))^2} dx$ points, such that, the SI-method's approximation $u_\omega(x)$ of the solution $u(x)$ on the mesh $\omega(h)$ satisfies the following asymptotic equalities:

$$\|u(x) - u_\omega(x)\|_\omega = \mathcal{O}(h^2), \tag{31}$$

$$\|u'(x) - u'_\omega(x)\|_\omega = \mathcal{O}(h^2); \tag{32}$$

3. from the method's point of view, there should be no qualitative difference between solving IVP (1), (18) with respect to $u(x)$ or with respect to its inverse $x(u)$, i.e. the method being applied to the IVP (26), (27), which is the "inverse" equivalent of IVP (1), (18) (in the sense that the graph of x(u) (wherever it exists exists) coincide with that of u(x)), gives the same result (or almost the same result) as when it is applied to the "straight" IVP[3].

3.2 Definition

Denoting by $u_\omega(x)$ the SI-method's approximation of the solution to IVP (1), (18), we define it by the following recurrence chain of equalities:

$$x_0 = a, \quad u_0 = u_l, \quad u'_0 = u'_l; \tag{33}$$

if $|u'_{i-1}| \leq 1$ then

$$\begin{aligned}
x_i =& x_{i-1} + h, \\
u_i =& U(A_{i-1}, B_{i-1}, C_{i-1}, D_{i-1}, h), \\
u'_i =& U'_h(A_{i-1}, B_{i-1}, C_{i-1}, D_{i-1}, h), \\
A_{i-1} =& N'_u(u_{i-1}, x_{i-1})u'_{i-1} + N'_x(u_{i-1}, x_{i-1}), \\
B_{i-1} =& N(u_{i-1}, x_{i-1}), \\
C_{i-1} =& u'_{i-1}, \\
D_{i-1} =& u_{i-1},
\end{aligned} \tag{34}$$

[3]Apparently, the statement, as it is now, has a lack of rigor and it might even seem contradictory. However, its meaning will become more clear as we get more familiar with the SI-method and its properties.

otherwise, if $|u'_{i-1}| > 1$

$$\begin{aligned}
x_i =& V(A_{i-1}, B_{i-1}, C_{i-1}, D_{i-1}, h_i^*),\\
u_i =& u_{i-1} + h_i^*,\\
u'_i =& 1/V'_h(A_{i-1}, B_{i-1}, C_{i-1}, D_{i-1}, h_i^*),\\
A_{i-1} =& -\left((N'_u(u_{i-1}, x_{i-1}) + N'_x(u_{i-1}, x_{i-1})x'_{i-1})u_{i-1} + \right.\\
& \left. + N(u_{i-1}, x_{i-1})\right)\left(x'_{i-1}\right)^2 + \\
& + 2\left(N(u_{i-1}, x_{i-1})u_{i-1}\right)^2\left(x'_{i-1}\right)^4,\\
B_{i-1} =& -N(u_{i-1}, x_{i-1})u_{i-1}\left(x'_{i-1}\right)^2,\\
C_{i-1} =& 1/u'_{i-1},\\
D_{i-1} =& x_{i-1},\\
h_i^* =& \operatorname{sign}(u'_{i-1})h,
\end{aligned} \tag{35}$$

where

$$u_i \overset{def}{=} u_\omega(x_i), \quad u'_i \overset{def}{=} u'_\omega(x_i)$$

and h — some fixed positive real number hereinafter referenced to as a *step size* of the SI-method.

As a matter of the fact, the recurrence equalities (33), (34), (35) provide us a way to construct an ordered set of two dimensional points

$$\omega(h) = \{(u'_i, u_i, x_i), i = 0, 1, 2 \ldots\} \tag{36}$$

which will be referenced to as a *mesh* of the SI-method. From the recurrence formulas (33), (34), (35) it follows that as long as function $N(u, x)$ belongs to $C^1(\mathbb{R} \times [0, +\infty))$ the mesh $\omega(h)$ contains infinite number of points, i.e. the recurrence process of calculation (u'_i, u_i, x_i) can be continued infinitely long. In the light of this a reasonable question arises: whether the mesh $\omega(h)$ (which is infinite) have something to do with the exact solution $u(x)$ of the Cauchy problem (1), (18) (which can exist only on some finite subset of $[0, +\infty)$) and, if yes, which approximation properties does the mesh possess with respect to the exact solution? To some extent the question is addressed in the paragraph below.

3.3 Error analysis.

In the present section we investigate approximation properties of the SI-method introduced above. The main result can be stated as the following theorem.

Theorem 5. *Let the nonlinear function $N(u,x)$ be independent on x, i.e.*

$$\frac{\partial N(u,x)}{\partial x} \equiv 0, \tag{37}$$

and

$$N(u) \equiv N(u,x) \in C^3([u_l,+\infty)), \tag{38}$$

$$N(u) > 0,\ \forall u \in (u_l,+\infty), \tag{39}$$

$$\lim_{u\to+\infty} \frac{1}{u^{2+\lambda}} \int\limits_{u_l}^{u} N(\xi)\xi d\xi > 0, \tag{40}$$

for some $\lambda > 0$.

If

$$0 < h < \max\left\{1, \sqrt{\frac{\varepsilon}{P^*}}, \frac{\varepsilon}{L_0^* M^*}\right\} \tag{41}$$

then there exists an integer $i^* \geq 0$, such that

$$u_i' < 1, \forall i \in \overline{0, i^*-1},\ u_{i^*}' \geq 1 \tag{42}$$

and the estimation holds true

$$\max\{|u(x_i) - u_i|, |u'(x_i) - u_i'|\} \leq h^2 P^*,\ \forall i \in \overline{0, i^*}, \tag{43}$$

where

$$P^* = \frac{M^*}{2}\left(L_2^* + L_1^* L_0^* M^*\right) \frac{\exp\left((S^*+1)(1+L_0^*+L_1^*) + S^*(2L_1^*+L_2^*)\right)}{(1+L_0^*+L_2^*)(2L_1^*+L_2^*)}, \tag{44}$$

$$L_i^* = \max_{|u|<M^*+\varepsilon} \left|N^{(i)}(u)\right|,\ i = 0,1,2, \tag{45}$$

$$M^* = \frac{1}{2} S^* (1 + 3\varepsilon - u_l'), \tag{46}$$

$$S^* = \lim_{u\to+\infty} \int\limits_{u_l}^{u} \frac{d\eta}{\sqrt{(u_l')^2 + 2\int\limits_{u_l}^{\eta} N(\xi)\xi d\xi}}.[4] \tag{47}$$

and ε denotes an arbitrary positive parameter.

If, additionally,

$$N^{(i)}(u) \geq 0,\ \forall u \in [u_l,+\infty),\ i = 1,2,3, \tag{48}$$

[4]The existance of the limit follows from condition (40),

$$\lim_{u\to+\infty}\frac{N(u)u}{\int\limits_0^u N(\xi)\xi d\xi}<+\infty, \tag{49}$$

$$h<\max\left\{\frac{1-2\varepsilon}{P^*},\frac{1}{3\mu}\right\}, \tag{50}$$

where

$$\mu=\sup_{u\in[u_{i^*},+\infty)}\frac{\mathcal{N}(u)}{1+\int\limits_{u_{i^*}}^{u}\mathcal{N}(\xi)d\xi}, \tag{51}$$

then $u_i'\geq 1$, $\forall i>i^*$ *and the estimations hold true*

$$|x_i-x(x_i)|\leq\left(\frac{L_0^*M^*}{(1-2\varepsilon)^2}+1\right)\frac{P^*h^2}{1-2\varepsilon}+ \tag{52}$$

$$+\frac{h^2}{2}\int\limits_{u_{i^*}}^{u_i}\left(1+\frac{\mathcal{N}''(\zeta)+\mathcal{N}'(\zeta)}{1+2\int\limits_{u_{i^*}}^{\zeta-h}\mathcal{N}(\xi)d\xi}\right)\frac{d\zeta}{\sqrt{(1-2\varepsilon)^2+2\int\limits_{u_{i^*}}^{\zeta}\mathcal{N}(\xi)d\xi}},$$

$$|x_i'-x'(x_i)|\leq\frac{h^2P^*}{1-P^*h^2-L_0^*M^*h}+\left(\frac{L_0^*M^*}{(1-2\varepsilon)^2}+1\right)\frac{P^*h^2(u_i-u_{i^*})}{1-2\varepsilon}+ \tag{53}$$

$$+\frac{h^2}{2}\int\limits_{u_{i^*}}^{u_i}\left(1+\frac{\mathcal{N}''(\zeta)+\mathcal{N}'(\zeta)}{1+2\int\limits_{u_{i^*}}^{\zeta-h}\mathcal{N}(\xi)d\xi}\right)\frac{(u_{i+1}-\zeta)d\zeta}{\sqrt{(1-2\varepsilon)^2+2\int\limits_{u_{i^*}}^{\zeta}\mathcal{N}(\xi)d\xi}}.$$

where $x(\cdot)\stackrel{def}{=}u^{-1}(\cdot)$.

Proof. As it was pointed out above, the function $x(u)$, which is (by definition) inverse of the exact solution $u(x)$ should be solution to the IVP (26), (27). Under the assumptions of the theorem, equation (26) becomes a partial case of the well known Bernoulli equation, which allows us to express the function $x(u)$ in the closed form (see, for example, [15]):

$$x(u)=a+\int\limits_{u_l}^{u}\frac{d\eta}{\sqrt{(u_l')^2+2\int\limits_{u_l}^{\eta}N(\xi)\xi d\xi}}. \tag{54}$$

From (37), (39), (38) and the Picard-Lindelöf theorem (see [12, p. 38]) it follows that function $x(u)$ belongs to $C^3([0,+\infty))$ and is the unique solution to the IVP (26), (27) on $[0,+\infty)$.

Using inequalities (39), (40), assumption $u_l'>0$ and the *Limit Comparison Theorem for*

Improper Integrals, from (54) we can easily derive that $x(u)$ is a monotonically increasing function on $[0, +\infty)$ with bounded range :

$$[0, +\infty) \overset{x(\cdot)}{\Rightarrow} [a, S),\ S = \lim_{u \to +\infty} x(u) < +\infty.$$

The letter fact means that its inverse $u(x)$, exists on $\big[a, S\big)$ and is the unique solution to (26), (27) on the segment. As it follows from equation (54), condition (40) also means that $x'(u)$ is a monotonically decreasing towards zero function

$$\lim_{u \to +\infty} x'(u) = 0, \tag{55}$$

and, consequently, $u'(x)$ is a function which monotonically increase towards infinity as x tends to S :

$$u'(\xi_1) < u'(\xi_2),\ \forall \xi_1, \xi_2 \in [a, S) : \xi_1 < \xi_2,\ \lim_{x \to +S} u'(x) = +\infty. \tag{56}$$

From (56) it follows that for each $\delta \geq u'_l$, there exists a unique $x_\delta \in [a, S)$ such that

$$u'(x_\delta) = \delta.$$

This in conjunction with the fact that function $u'(x)$ is convex on $[a, S)$, allows us to establish the following inequality (see Figure 1)

$$\max_{x \in [a,S): u'(x) \leq \delta} |u(x)| = \int\limits_a^{x_\delta} u'(\xi) d\xi \leq \frac{1}{2}(x_\delta - a)(\delta - u'_l) < \frac{1}{2}(S - a)(\delta - u'_l), \tag{57}$$

which is of crucial importance for the rest of proof.

Without loss of generality we confine ourselves to consider a case when $u'_l < 1$.

Using notation

$$E_i = \begin{bmatrix} u(x_i) - u_i \\ u'(x_i) - u'_i \end{bmatrix},\ \ e_i = \|E_i\|$$

we can find that

$$e_0 = 0.$$

From Theorem 3 it follows that

$$e_1 \leq h^3 K \|u(x)\|_{x \in [0,h]} \exp(hM),$$

where

$$K = \frac{1}{2}\|N''(u(x))(u'(x))^2 + N'(u(x))N(u(x))u(x)\|_{[a,a+h]},$$

$$M = \max\{1, |N(u_0)| + |N'(u_0)u'_0|h\}.$$

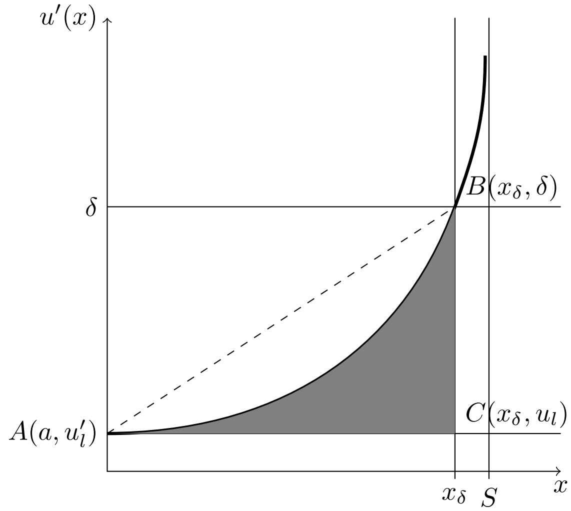

Figure 1: Area of the shaded region is equal to $\int\limits_a^{x_\delta} u'(\xi)d\xi$, which is, apparently, less or equal to the area of $\triangle ABC$, which, in turn, is equal to $\frac{1}{2}(x_\delta - a)(\delta - u_l')$.

In general case e_i can be estimated from the system of differential equation

$$\dot{Z}_i(x) = \begin{bmatrix} 0 & 1 \\ N(u_{i-1}) + N'(u_{i-1})u'_{i-1}(x - x_{i-1}) & 0 \end{bmatrix} Z_i(x) + \begin{bmatrix} 0 \\ F_i(x)u(x) \end{bmatrix}, \tag{58}$$

$$x \in [x_{i-1}, x_i],\, ,\ Z_i(x_i) = Z_{i-1}(x_i),$$

where

$$F_i(x) = N(u(x)) - N(u_{i-1}) - N'(u_{i-1})u'_{i-1}(x - x_{i-1}),$$

$$Z_i(x) = \begin{bmatrix} z_i(x) \\ z_i'(x) \end{bmatrix},\ \ z_i(x) = u(x) - U(N'(u_{i-1})u'_{i-1}, N(u_{i-1}), u'_{i-1}, u_{i-1}, x - x_{i-1}),\ i = 1, 2, \ldots,$$

$$Z_0(x) \equiv 0. \tag{59}$$

We are not going to estimate e_i for all integer i, but only for all $i \leq i^*$, where i^* is defined in (42). However, at this point, the very existence of such an integer value i^* is yet to be proved. Let us fix some arbitrary $\varepsilon > 0$ and assume that

$$e_j \leq \|Z_i(x)\|_{x \in [x_{j-1}, x_j]} < \varepsilon,\ \forall j : u'(x_j) \leq 1 + 3\varepsilon,\ u_j < 1, \tag{60}$$

for sufficiently small values of h.

We are also free to consider the constants L_i^*, $i = 0, 1, 2$ defined in (45), where

$$M^* = \frac{1}{2}(S - a)(1 + 3\varepsilon - u_l') > \max_{x \in [a;S): u'(x) \leq 1 + 3\varepsilon} u(x). \tag{61}$$

Here, in the latter inequality of (61), we have used (57).

Now, requiring that

$$h < \frac{\varepsilon}{L_0^* M^*}. \tag{62}$$

we can easily prove that i^* exists and belongs to $(a, u'^{-1}(1+2\varepsilon))$. Indeed, if there exists $x_j \in (a, u'^{-1}(1+2\varepsilon)]$, such that $u'_j \geq 1$, then we can put

$$i^* = \min_{u'(x_j) \geq 1+2\varepsilon} \{j\}.$$

If this is not the case at least for a single h satisfying (62), then, from (62) it follows that

$$h < u'^{-1}(1+2\varepsilon) - u'^{-1}(1+\varepsilon),$$

which, in turn, means that there exists at least one x_j belonging to the interval

$$(u'^{-1}(1+\varepsilon), u'^{-1}(1+2\varepsilon)].$$

Taking into account (60), the latter fact yields us $u'_j \geq 1$ and, consequently, we have a contradiction.

Note, that constants defined in (45), (61), do not depend on h and, technically, have nothing to do with the assumption (60). Using the constants and assumption (60), we can derive estimation for $\|Z_i(x)\|_{x\in[x_{i-1},x_i]}$, for $i = 1, 2, \ldots, i^*$ in the following way:

$$\|Z_i(x)\| \leq (1 + hL_1^* + h^2(L_1^* + L_2^*))\|Z_{i-1}(x)\|_{x\in[x_{i-2},x_{i-1}]} + \tag{63}$$

$$+ \max\{1, L_0^* + L_1^* h\} \int_{x_{i-1}}^{x} \|Z_i(\xi)\| d\xi + h^3 \frac{M^*}{2} (L_2^* + L_1^* L_0^* M^*),$$

$$\forall x \in [x_{i-1}, x_i],\ i = 1, 2, \ldots, i^*,\ \|Z_0(x)\|_{x\in[x_{-1},x_0]} \equiv 0.$$

Applying the Gronwall's inequality (see, for example, [12, 42]) to (63) we get

$$\|Z_i(x)\|_{x\in[x_{i-1},x_i]} \leq$$

$$\leq \left((1 + hL_1^* + h^2(L_1^* + L_2^*))\|Z_{i-1}(x)\|_{x\in[x_{i-2},x_{i-1}]} + h^3 \frac{M^*}{2} (L_2^* + L_1^* L_0^* M^*) \right) \times \tag{64}$$

$$\times \exp\left(h\left(\max\{1, L_0^* + L_1^* h\}\right)\right),\ i = 1, 2, \ldots, i^*.$$

From (64), taking into account (59), we can get the estimation

$$\|Z_i(x)\|_{x\in[x_{i-1},x_i]} \leq h^3 \frac{M^*}{2} (L_2^* + L_1^* L_0^* M^*) \sum_{j=1}^{j=i} (1 + hL_1^* + h^2(L_1^* + L_2^*))^{j-1} \exp(jhE^*) = \tag{65}$$

$$= h^3 \frac{M^*}{2} \left(L_2^* + L_1^* L_0^* M^*\right) \exp\left(hE^*\right) \frac{(1 + hL_1^* + h^2(L_1^* + L_2^*))^i \exp\left(ihE^*\right) - 1}{(1 + hL_1^* + h^2(L_1^* + L_2^*)) \exp\left(hE^*\right) - 1} <$$

$$< h^3 \frac{M^*}{2} \left(L_2^* + L_1^* L_0^* M^*\right) \exp\left(hE^*\right) \frac{\exp\left((S-a)\left(E^* + L_1^* + h(L_1^* + L_2^*)\right)\right) - 1}{(1 + hL_1^* + h^2(L_1^* + L_2^*)) \exp\left(hE^*\right) - 1} \leq$$

$$\leq h^2 \frac{M^*}{2} \left(L_2^* + L_1^* L_0^* M^*\right) \frac{\exp\left((S^*+1)(1 + L_0^* + L_1^*) + S^*(2L_1^* + L_2^*)\right)}{(1 + L_0^* + L_2^*)(2L_1^* + L_2^*)} = h^2 P^*,$$

$$i = 1, 2, \ldots, i^*,$$

where

$$E^* = \max\left\{1, L_0^* + L_1^* h\right\} \leq 1 + L_0^* + L_1^*, \; S^* = S - a.$$

The last inequality in (65) holds true under assumption that

$$0 \leq h \leq 1, \tag{66}$$

which we accept from now on.

Going back to inequality (63), it is important to mention that to derive it for each particular $i = 1, 2, \ldots, i^*$, we have used assumption (60) only for $j = i - 1$, but not for $j = i$. Besides that, the inequality (63) for $i = 1$ does not rely upon (60) at all. With this in mind and taking into account (65), we can easily prove (by means of mathematical induction) that for h satisfying

$$h < \sqrt{\frac{\varepsilon}{P^*}}, \tag{67}$$

assumption (60) holds true. This concludes proof of the first part of the theorem, which states the existence of i^* (42) and the fulfillment of approximation estimates (43), provided that h satisfies (41).

Above we showed that $x_{i^*} \in \left[\alpha_0, \alpha_1\right]$, where $\alpha_0 = u'^{-1}(1 - P^* h^2)$, $\alpha_1 = u'^{-1}(1 + 2\varepsilon)$. Using similar reasoning it is easy to verify that

$$1 - 2\varepsilon \leq u'(x_{i^*} - h) \leq u'(x_{i^*} + h) \leq 1 + 3\varepsilon. \tag{68}$$

In what follows we will need the estimate

$$\min_{x \in [x_{i^*} - h, x_{i^*} + h]} u'(x) = u'\left(x_{i^*} - h\right) = \tag{69}$$

$$= u'\left(x_{i^*} - h\right) - u'\left(x_{i^*}\right) + u'\left(x_{i^*}\right) \geq u'\left(x_{i^*} - h\right) - u'\left(x_{i^*}\right) + u'\left(\alpha_0\right) \geq$$

$$\geq -u''(u'^{-1}(1 + 3\varepsilon))h + 1 - P^* h^2 \geq 1 - P^* h^2 - L_0^* M^* h = \tau(h).$$

Using (69) we can easily find that

$$\max_{u\in[u(x_{i^*}-h),u(x_{i^*}+h)]} x'(u) \le (\tau(h))^{-1}. \tag{70}$$

Now if we require that

$$h \le \frac{\tau(h)}{P^*}, \tag{71}$$

then

$$u_{i^*} \in [u(x_{i^*}-h), u(x_{i^*}+h)], \tag{72}$$

indeed

$$u_{i^*} - u(x_{i^*}-h) = u_{i^*} - u(x_{i^*}) + u(x_{i^*}) - u(x_{i^*}-h) \ge h\tau(h) - P^*h^2 \ge 0,$$

$$u(x_{i^*}+h) - u_{i^*} = u(x_{i^*}+h) - u(x_{i^*}) + u(x_{i^*}) - u_{i^*} \ge h\tau(h) - P^*h^2 \ge 0.$$

From (72) it follows that $u(x_{i^*}) + \theta(u_{i^*} - u(x_{i^*})) \in [u(x_{i^*}-h), u(x_{i^*}+h)]$, $\forall\theta \in [0,1]$. With this in mind, and taking into account (70), we can easily derive the inequality

$$|x(u_{i^*}) - x(u(x_{i^*}))| = x'(u(x_{i^*}) + \theta(u_{i^*} - u(x_{i^*})))|u_{i^*} - u(x_{i^*})| \le \tag{73}$$

$$\le (\tau(h))^{-1} P^*h^2.$$

where $0 \le \theta \le 1$.

From (68) it follows that

$$\max_{u\in[u(x_{i^*}-h),u(x_{i^*}+h)]} |x''(u)| = \max_{u\in[u(x_{i^*}-h),u(x_{i^*}+h)]} N(u)u\left(x'(u)\right)^3 \le \frac{L_0^*M^*}{(1-2\varepsilon)^3}. \tag{74}$$

Now using (72) and inequality (74) we can get the estimate

$$|x'(u_{i^*}) - x'_{i^*}| \le |x'(u_{i^*}) - x'(u(x_{i^*}))| + |x'(u(x_{i^*})) - x'_{i^*}| \le \tag{75}$$

$$\le \frac{L_0^*M^*P^*}{(1-2\varepsilon)^3}h^2 + \left|\frac{1}{u(x_{i^*})} - \frac{1}{u'_{i^*}}\right| \le \left(\frac{L_0^*M^*}{(1-2\varepsilon)^2} + 1\right)\frac{P^*}{1-2\varepsilon}h^2.$$

At this point we have proved that after the first i^* iterations the algorithm should switch from formulas (34) to formulas (35). The lemma below, among other things, states that starting from i^* the algorithm will never switch back to formulas (34), i.e. u'_i remains greater or equal to 1, for all $i > i^*$, provided h is small enough. For the sake of convenience, hereinafter we will use notation $\mathcal{N}(u) = N(u)u$.

Lemma 2. *Let $h > 0$ satisfy condition*

$$h < \frac{1}{3\mu}, \tag{76}$$

where constant μ is defined in (51). *Then*

$$A_i \frac{u^2}{2} + B_i u \le 0 \; \forall u \in [0, h], \tag{77}$$

$$0 \le x_i' \le \left(\frac{1}{(x_{i^*}')^2} + \int\limits_{u_{i^*}}^{u_{i-1}} \mathcal{N}(u) du \right)^{-\frac{1}{2}}, \; i = i^* + 1, i^* + 2, \dots. \tag{78}$$

Proof. Let us consider an auxiliary sequence

$$\bar{x}_i' = \bar{x}_{i-1}' \exp\left(s_{1,i} \mathcal{N}(u_{i-1}) \left(\bar{x}_{i-1}'\right)^2 h + s_{2,i} \right), \; s_{1,i} \le -\frac{1}{2}, \; s_{2,i} \le 0, \tag{79}$$

$$i = i^* + 1, i^* + 2, \dots, \; \bar{x}_{i^*}' = x_{i^*}' = \frac{1}{u_{i^*}'}.$$

It is easy to see that $\bar{x}_{i^*}' > 0$ and from

$$(\bar{x}_i')^2 - \left(\frac{1}{\left(\bar{x}_{i-1}'\right)^2} + \mathcal{N}(u_{i-1}) h \right)^{-1} =$$

$$= \left(\bar{x}_{i-1}'\right)^2 \exp\left(2 s_{1,i} \mathcal{N}(u_{i-1}) \left(\bar{x}_{i-1}'\right)^2 h + 2 s_{2,i} \right) - \frac{\left(\bar{x}_{i-1}'\right)^2}{1 + \left(\bar{x}_{i-1}'\right)^2 \mathcal{N}(u_i) h} =$$

$$= \frac{\left(\bar{x}_{i-1}'\right)^2 \left(\exp\left(2 s_{1,i} \mathcal{N}(u_{i-1}) \left(\bar{x}_{i-1}'\right)^2 h + 2 s_{2,i} \right) \left(1 + \left(\bar{x}_{i-1}'\right)^2 \mathcal{N}(u_{i-1}) h \right) - 1 \right)}{1 + \left(\bar{x}_{i-1}'\right)^2 \mathcal{N}(u_i)} \le$$

$$\le \frac{\left(\bar{x}_{i-1}'\right)^2 \left(\exp\left(2 s_{1,i} \mathcal{N}(u_{i-1}) \left(\bar{x}_{i-1}'\right)^2 h + 2 s_{2,i} \right) \exp\left(\mathcal{N}(u_{i-1}) \left(\bar{x}_{i-1}'\right)^2 h \right) - 1 \right)}{1 + \left(\bar{x}_{i-1}'\right)^2 \mathcal{N}(u_i)} \le 0$$

it follows that

$$\bar{x}_i' \le \left(\frac{1}{\left(\bar{x}_{i-1}'\right)^2} + \mathcal{N}(u_{i-1}) h \right)^{-\frac{1}{2}}. \tag{80}$$

Applying inequality (80) recursively we get the estimation

$$\bar{x}_i' \le \left(\frac{1}{(\bar{x}_{i^*}')^2} + \sum_{j=i^*}^{i-1} \mathcal{N}(u_j) h \right)^{-\frac{1}{2}} \le \left(1 + \int\limits_{u_{i^*}}^{u_i} \mathcal{N}(u) du - \mathcal{N}(u_i) h \right)^{-\frac{1}{2}}. \tag{81}$$

To derive the last inequality in (81) we exploit the fact that function $\mathcal{N}(u)$ is non-decreasing (see (48)). From (81), using (76) we can easily get

$$0 \le \mathcal{N}(u_i)\left(\bar{x}_i'\right)^2 h \le \mathcal{N}(u_i)h\left(1+\int\limits_{u_{i^*}}^{u_i}\mathcal{N}(u)du-\mathcal{N}(u_i)h\right)^{-1} \le \frac{\mu h}{1-\mu h} \le \frac{1}{2}, \tag{82}$$

$$i = i^*+1, i^*+2, \ldots .$$

Inequality (82) together with (39) and (48) imply that estimation (81) remains valid if

$$s_{1,i} = \mathcal{N}(u_{i-1})\left(\bar{x}_{i-1}'\right)^2 h - 1 \le -\frac{1}{2},\ s_{2,i} = -\frac{h^2}{2}\mathcal{N}'(u)\left(\bar{x}_{i-1}'\right)^2. \tag{83}$$

On the other hand, sequence $\{\bar{x}_i'\}$ (79) together with substitution (83) totally coincide with sequence $\{x_i'\}$ (35) :

$$\begin{aligned} x_i' = x_{i-1}' \exp\left(-h\mathcal{N}(u_{i-1})\left(x_{i-1}'\right)^2 + \frac{h^2}{2}\left(-\mathcal{N}'(u_{i-1})\left(x_{i-1}'\right)^2 + 2\left(\mathcal{N}(u_{i-1})\left(x_{i-1}'\right)^2\right)^2\right)\right) = \\ = x_{i-1}'\exp\left(A_i\frac{h^2}{2}+Bh\right),\ i = i^*+1, i^*+2, \ldots . \end{aligned} \tag{84}$$

In the light of the latter observation, estimates (78), immediately follow from (81), whereas inequalities (77) follow from (82). $\square$

From now on we assume that h satisfies requirements (76), (71), which is equivalent to (50).

Let us now consider a sequence of functions $\{y_i(u)\}$, defined as follows

$$\begin{aligned} y_i(u) &= x(u) - V_i(u), \\ V_i(u) &= V\left(A_i, B_i, x_{i-1}', x_{i-1}, u-u_{i-1}\right), \\ A_i &= -\mathcal{N}'(u_{i-1})\left(x_{i-1}'\right)^2 + 2\left(\mathcal{N}(u_{i-1})\right)^2\left(x_{i-1}'\right)^4, \\ B_i &= -\mathcal{N}(u_{i-1})\left(x_{i-1}'\right)^2, i = i^*+1, i^*+2, \ldots . \end{aligned} \tag{85}$$

It is easy to see that $y_i(u)$ should satisfy the recurrence system of Cauchy problems

$$y_i''(u) = G_i(u)y_i'(u) + F_i(u)x'(u), \tag{86}$$

$$F_i(u) = -\mathcal{N}(u)\left(V_i'(u)\right)^2 - A_i(u-u_{i-1}) - B_i,\ u \in [u_{i-1}, u_i],$$

$$G_i(u) = \left(B_i + A_i(u-u_{i-1}) - \mathcal{N}(u)x'(u)\left(x'(u)+V_i'(u)\right)\right),$$

$$Y_i(u_i) = Y_{i-1}(u_i),\ u_i = u_{i-1}+h,\ i = i^*+1, i^*+2, \ldots . \tag{87}$$

Inequalities (73), (75) allow us to estimate $|y_{i^*}^{(k)}(u_{i^*})|$, $k=0,1$ in the following way:

$$|y_{i^*}(u_{i^*})| \leq \frac{P^*h^2}{1-P^*h^2-L_0^*M^*h},\ |y'_{i^*}(u_{i^*})| \leq \left(\frac{L_0^*M^*}{(1-2\varepsilon)^2}+1\right)\frac{P^*h^2}{1-2\varepsilon}. \tag{88}$$

Using mean value theorem, we can easily find that

$$F_i(u) = \frac{(u-u_{i-1})^2}{2}\Big(-\mathcal{N}''(u_{i-1}+\theta h)\left(V_i'(u_{i-1}+\theta h)\right)^2 + \tag{89}$$

$$+2\mathcal{N}(u_{i-1}+\theta h)\mathcal{N}'(u_{i-1}+\theta h)\left(V_i'(u_{i-1}+\theta h)\right)^4 -$$

$$-8\left(\mathcal{N}(u_{i-1}+\theta h)\right)^3\left(V_i'(u_{i-1}+\theta h)\right)^6\Big),\ \forall u\in[u_{i-1},u_i],\ 0\leq\theta\leq 1,$$

which, together with (78) and (82), yields us an estimate[5]

$$\|F_i(u)\|_{[u_{i-1},u_i]} \leq \frac{(u-u_{i-1})^2}{2}\left(\frac{\mathcal{N}''(u)+\mathcal{N}'(u)}{1+\int\limits_{u_{i^*}}^{u_{i-1}}\mathcal{N}(\xi)d\xi}+1\right),\ i=i^*+1,i^*+2,\ldots. \tag{90}$$

Solution to (86) can be expressed in the form

$$y_i'(u) = \int\limits_{u_{i-1}}^{u}\exp\left(\int\limits_{\xi}^{u}G_i(\zeta)d\zeta\right)F_i(\xi)x'(\xi)d\xi + y'_{i-1}(u_{i-1})\exp\left(\int\limits_{u_{i-1}}^{u}G_i(\zeta)d\zeta\right), \tag{91}$$

$$y_i(u) = \int\limits_{u_{i-1}}^{u}y_i'(\xi)d\xi + y_{i-1}(u_{i-1}), \tag{92}$$

$$u\in[u_{i-1},u_i],\ i=i^*+1,i^*+2,\ldots.$$

Using initial estimates (88), from (91) we can easily get inequalities

$$\|y_i'(u)\|_{[u_{i-1},u_i]} \leq \left(\frac{L_0^*M^*}{(1-2\varepsilon)^2}+1\right)\frac{P^*h^2}{1-2\varepsilon}+ \tag{93}$$

$$+\frac{h^2}{2}\int\limits_{u_{i^*}}^{u_i}\left(1+\frac{\mathcal{N}''(\zeta)+\mathcal{N}'(\zeta)}{1+\int\limits_{u_{i^*}}^{\zeta-h}\mathcal{N}(\xi)d\xi}\right)\frac{d\zeta}{\sqrt{(1-2\varepsilon)^2+2\int\limits_{u_{i^*}}^{\zeta}\mathcal{N}(\xi)d\xi}},\ i=i^*+1,i^*+2,\ldots,$$

[5]Here we use conditions (48), implying that functions $\mathcal{N}'(u)$ and $\mathcal{N}''(u)$ are nondecreasing.

which, being combined with (92), yield us

$$\|y_i(u)\|_{[u_{i-1},u_i]} \leq \frac{h^2 P^*}{1 - P^* h^2 - L_0^* M^* h} + \left(\frac{L_0^* M^*}{(1-2\varepsilon)^2} + 1\right) \frac{P^* h^2 (u_i - u_{i^*})}{1 - 2\varepsilon} + \tag{94}$$

$$+\frac{h^2}{2} \int\limits_{u_{i^*}}^{u_i} \left(1 + \frac{\mathcal{N}''(\zeta) + \mathcal{N}'(\zeta)}{1 + \int\limits_{u_{i^*}}^{\zeta - h} \mathcal{N}(\xi) d\xi} \right) \frac{(u_{i+1} - \zeta) d\zeta}{\sqrt{(1-2\varepsilon)^2 + 2 \int\limits_{u_{i^*}}^{\zeta} \mathcal{N}(\xi) d\xi}} \quad i = i^* + 1, i^* + 2, \ldots.$$

Estimates (52), (53) follows immediately from (93) and (94) respectively, which completes the proof. □

As it can be easily verified, the estimations given in Theorem 5 are quite rough and become quite useless when, for instance, applied to the Troesch's equation (3) with λ sufficiently large. The main reason for this is the roughness of estimate (61). The latter can be improved, and this is reflected in the remark below.

Remark 1. *The estimates of Theorem 5 remain valid and can be essentially improved if the constant M^* defined by formula* (46) *is substituted by the one defined as*

$$M^* = F^{-1}\left(\frac{1}{2}\left(\frac{1}{(1+3\varepsilon)^2} - (u_l')^2\right)\right), \; F(u) = \int\limits_{u_l}^{u} N(\xi)\xi d\xi. \tag{95}$$

4 Straight-Inverse method for solving BVPs for second order differential equations.

In this section we introduce the SI-method for solving boundary value problems. As a matter of the fact we are going to talk about the *simple* and *multiple shooting techniques* supplemented by the SI-method for solving IVPs described above.

Since any *simple shooting technique* is nothing but a bisection algorithm supplemented by an IVP solver, the meaning of the *SI simple shooting* method is evident and self-explanatory. At the same time the *SI multiple shooting method* requires considerably deeper introduction and the rest of the current section is devoted strictly to it.

4.1 SI multiple shooting method.

Talking about the *multiple shooting technique* for solving boundary value problems we (as a rule) mean a way how the given BVP can be transformed to a system of nonlinear algebraic equations together with an algorithm how to solve the system.

Assume that we have fixed some positive parameter $\mathbf{h}$, hereinafter referenced to as a *step size*. In addition to that we need to have some initial guess Ω_k which is some discrete approximation[6] of the exact solution $u(x)$ of the BVP (1), (2):

$$\begin{aligned}\Omega_k = \{\omega_{k,i} = (u'_{k,i}, u_{k,i}, x_{k,i})| \\ & x_{k,i} \in [a,b],\ u(x_{k,i}) \approx u_{k,i},\ u'(x_{k,i}) \approx u'_{k,i},\ i \in \overline{0, N_k},\ x_{k,i} < x_{k,j} \Leftrightarrow i < j < N_k \\ & u_{k,0} = u_l,\ u_{k,N_0} = u_r,\ x_{k,0} = 0,\ x_{k,N_0} = 1\},\end{aligned} \tag{96}$$

From now on, we assume that

$$\begin{aligned} h_{k,i} &\stackrel{def}{=} x_{k,i+1} - x_{k,i},\ i \in \overline{0, N_k - 1}, \\ \bar{h}_{k,i} &\stackrel{def}{=} u_{k,i+1} - u_{k,i},\ i \in \overline{0, N_k - 1}, \\ \max\{h_{k,i}, \bar{h}_{k,i}\} &\le \mathbf{h},\ \forall i \in \overline{0, N_k - 1}. \end{aligned} \tag{97}$$

Given that, we transform the set Ω_0 into an ordered set of nonlinear equations

$$\Sigma_0 = \{\sigma_{0,i,j},\ i \in \overline{0, N_0 - 1},\ j = 0, 1\}$$

using the rule described below.

For the first two equations we have

$$\begin{aligned}&\sigma_{k,0,0}(\omega_{k,0},\ \omega_{k,1}) \stackrel{def}{=} \\ &\begin{cases} U(A_U(x_{k,0}, u_{k,0}, \mathbf{u}'_0),\ B_U(x_{k,0}, u_{k,0}, \mathbf{u}'_0),\ \mathbf{u}'_0, u_{k,0}, h_{k,0}) = \mathbf{u_1},\ |u'_{k,0}| \le 1, \\ V(A_V(u_{k,0}, x_{k,0}, \mathbf{x}'_0),\ B_V(u_{k,0}, x_{k,0}, \mathbf{x}'_0),\ \mathbf{x}'_0, x_{k,0}, \bar{h}_{k,0}) = \mathbf{x_1},\ |u'_{k,0}| > 1, \end{cases}\end{aligned} \tag{98}$$

$$\begin{aligned}&\sigma_{k,0,1}(\omega_{k,0},\ \omega_{k,1}) \stackrel{def}{=} \\ &\begin{cases} U'_h(A_U(x_{k,0}, u_{k,0}, \mathbf{u}'_0),\ B_U(x_{k,0}, u_{k,0}, \mathbf{u}'_0),\ \mathbf{u}'_0, u_{k,0}, h)|_{h=h_{k,0}}, & |u'_{k,0}| \le 1, \\ V'_h(A_V(u_{k,0}, x_{k,0}, \mathbf{x}'_0),\ B_V(u_{k,0}, x_{k,0}, \mathbf{x}'_0),\ \mathbf{x}'_0, x_{k,0}, h)|_{h=\bar{h}_{k,0}}, & |u'_{k,0}| > 1, \end{cases} = \\ &= \begin{cases} \mathbf{u}'_1, & |u'_{k,0}| \le 1,\ |u'_{k,1}| \le 1, \\ 1/\mathbf{x}'_1, & |u'_{k,0}| \le 1,\ |u'_{k,1}| > 1, \\ \mathbf{x}'_1, & |u'_{k,0}| > 1,\ |u'_{k,1}| > 1, \\ 1/\mathbf{u}'_1, & |u'_{k,0}| > 1,\ |u'_{k,1}| \le 1. \end{cases}\end{aligned} \tag{99}$$

As you can see, the equations are dependent on the absolute values of $|u'_{k,0}|$ and $|u'_{k,1}|$, which (in general) represent values of the first derivative of the unknown function $u(x)$. The same is true for rest of the equations, although their descriptions seem even more complicated.

[6]Here we avoid talking about how close the approximation should be, however, in practice, if the approximation is too bed the method described below can do not work at all. At the same time, using the simple shooting approach described above it is always possible to get the desired approximation.

For $i \in \overline{1, N_k - 1}$. we have:

$$\sigma_{k,i,0}(\omega_{k,i-1},\ \omega_{k,i},\ \omega_{k,i+1}) \stackrel{def}{=}$$

$$\begin{cases} U(A_U(x_{k,i}, \mathbf{u}_i, \mathbf{u}'_i),\ B_U(x_{k,i}, \mathbf{u}_i, \mathbf{u}'_i),\ \mathbf{u}'_i, \mathbf{u}_i, h_{k,i}), & |u'_{k,i}| \le 1,\ |u'_{k,i-1}| \le 1, \\ U(A_U(\mathbf{x}_i, u_{k,i}, \mathbf{u}'_i),\ B_U(\mathbf{x}_i, u_{k,i}, \mathbf{u}'_i),\ \mathbf{u}'_i, u_{k,i}, x_{k,i+1} - \mathbf{x}_i), & |u'_{k,i}| \le 1,\ |u'_{k,i-1}| > 1, \\ V(A_V(u_{k,i}, \mathbf{x}_i, \mathbf{x}'_i),\ B_V(u_{k,i}, \mathbf{x}_i, \mathbf{x}'_i),\ \mathbf{x}'_i, \mathbf{x}_i, \bar{h}_{k,i}), & |u'_{k,i}| > 1,\ |u'_{k,i-1}| > 1, \\ V(A_V(\mathbf{u}_i, x_{k,i}, \mathbf{x}'_i),\ B_V(\mathbf{u}_i, x_{k,i}, \mathbf{x}'_i),\ \mathbf{x}'_i, x_{k,i}, u_{k,i+1} - \mathbf{u}_i), & |u'_{k,i}| > 1,\ |u'_{k,i-1}| \le 1, \end{cases} =$$

$$= \begin{cases} \begin{cases} \mathbf{u}_{i+1}, & i \in \overline{1, N_k - 2} \\ u_{k,N_k}, & \text{otherwise} \end{cases}, & |u'_{k,i}| \le 1, \\ \begin{cases} \mathbf{x}_{i+1}, & i \in \overline{1, N_k - 2} \\ x_{k,N_k}, & \text{otherwise} \end{cases} & |u'_{k,i}| > 1, \end{cases} \tag{100}$$

$$\sigma_{k,i,1}(\omega_{k,i-1},\ \omega_{k,i},\ \omega_{k,i+1}) \stackrel{def}{=}$$

$$\begin{cases} U'_h(A_U(x_{k,i}, \mathbf{u}_i, \mathbf{u}'_i),\ B_U(x_{k,i}, \mathbf{u}_i, \mathbf{u}'_i),\ \mathbf{u}'_i, \mathbf{u}_i, h)|_{h=h_{k,i}}, & |u'_{k,i-1}| \le 1,\ |u'_{k,i}| \le 1, \\ U'_h(A_U(\mathbf{x}_i, u_{k,i}, \mathbf{u}'_i),\ B_U(\mathbf{x}_i, u_{k,i}, \mathbf{u}'_i),\ \mathbf{u}'_i, u_{k,i}, h)|_{h=x_{k,i+1}-\mathbf{x}_i}, & |u'_{k,i-1}| > 1, |u'_{k,i}| \le 1, \\ V'_h(A_V(u_{k,i}, \mathbf{x}_i, \mathbf{x}'_i),\ B_V(u_{k,i}, \mathbf{x}_i, \mathbf{x}'_i),\ \mathbf{x}'_i, \mathbf{x}_i, h)|_{h=\bar{h}_{k,i}}, & |u'_{k,i-1}| > 1,\ |u'_{k,i}| > 1, \\ V'_h(A_V(\mathbf{u}_i, x_{k,i}, \mathbf{x}'_i),\ B_V(\mathbf{u}_i, x_{k,i}, \mathbf{x}'_i),\ \mathbf{x}'_i, x_{k,i}, h)|_{h=u_{k,i+1}-\mathbf{u}_i}, & |u'_{k,i-1}| \le 1,\ |u'_{k,i}| > 1, \end{cases} =$$

$$= \begin{cases} \mathbf{u}'_{i+1}, & |u'_{k,i}| \le 1,\ |u'_{k,i+1}| \le 1, \\ 1/\mathbf{x}'_{i+1}, & |u'_{k,i}| \le 1,\ |u'_{k,i+1}| > 1, \\ \mathbf{x}'_{i+1}, & |u'_{k,i}| > 1,\ |u'_{k,i+1}| > 1, \\ 1/\mathbf{u}'_{i+1}, & |u'_{k,i}| > 1,\ |u'_{k,i+1}| \le 1, \end{cases} \tag{101}$$

where

$$\begin{aligned} A_U(x, u, u') =& N'_u(u, x)u' + N'_x(u, x), \\ B_U(x, u, u') =& N(u, x), \\ A_V(u, x, x') =& -\left((N'_u(u, x) + N'_x(u, x)x')u + N(u, x)\right)(x')^2 + 2\,(N(u, x)u)^2\,(x')^4\,, \\ B_V(u, x, x') =& -N(u, x)u\,(x')^2\,, \end{aligned} \tag{102}$$

bold variables describe unknowns and

$$\mathbf{u}'_i \stackrel{def}{=} 1/\mathbf{x}'_i,\ \forall i \in \overline{0, N_k}.$$

Using Ω_k as an initial guess and applying a single iteration of the generalized Newton's method (see, for example, [11, p. 293]) to the system Σ_k (98), (99), (100), (101) we will get a new set Ω_{k+1} as a combination of Ω_k and the results brought by the Newton's method's iteration. In practice it (quite often) happens that the set Ω_{k+1} obtained as described above needs to be sorted out (to fulfill the requirement $x_{k+1,i} < x_{k+1,j} \Leftrightarrow i < j < N_{k+1}$) and then refined (by linear interpolation to satisfy the inequality $\max\{h_{k+1,i}, \bar{h}_{k+1,i}\} \leq \mathbf{h}$, $\forall i \in \overline{0, N_{k+1}-1}$). Once it is done, we can use Ω_{k+1} to construct a new system Σ_{k+1} and, applying a single Newton's method's iteration to it, get Ω_{k+2} and so on and so forth, until the difference between to subsequent Ω's is not small enough[7].

5 Numerical examples

In this section we present numerical results of the SI-method when applied to the Troesch's problem (3), (4), [10].

λ	Value	[14]	[1]	[5]	SI-method, $h = 10^{-4}$	SI-method, $h = 10^{-5}$
20	$u'(0)$	1.648773182e-08	—	1.6487734e-8	1.648773647e-008	1.648773188e-008
20	$u'(1)$	22026.29966	22026.4657	—	22026.4657494062	22026.4657494068
30	$u'(0)$	7.486093793e-13	—	7.4861194e-13	7.486098431e-013	7.486093844e-013
50	$u'(0)$	1.542999878e-21	—	1.5430022e-21	1.543002448e-021	1.542999906e-021
61	$u'(0)$	—	—	2.5770722e-26	2.577078525e-026	2.577072299e-026
100	$u'(0)$	2.976060781e-43	—	—	2.976075557e-043	2.976060927e-043

Table 1: Troesch's problem solved by different approaches.

Value	[3]	[14]	SI-method, $h = 10^{-4}$	SI-method, $h = 10^{-5}$
$u(0.1)$	4.211183679705e-05	4.211189927237e-05	4.21119023173e-05	4.21118993037e-05
$u(0.2)$	1.299639238293e-04	1.299641158237e-04	1.29964125220e-04	1.29964115920e-04
$u(0.3)$	3.589778855481e-04	3.589784013896e-04	3.58978427345e-04	3.58978401657e-04
$u(0.4)$	9.779014227050e-04	9.779027718029e-04	9.77902842508e-04	9.77902772532e-04
$u(0.5)$	2.659017178062e-03	2.659020490351e-03	2.659020682593-03	2.65902049234e-03
$u(0.999)$	8.889931171768e-01	8.889931181558e-01	8.89035025083e-01[8]	8.88994612232e-01[9]

Table 2: Solution to the Troesch's problem with $\lambda = 10$ obtained by different approaches.

Source	$u'(0)$	$\|\Omega(h)\|$[10]	CPU time, sec.[11]	Rel. diff. to [14][12]
SIM, $h = 10^{-2}$	3.141990565e-43	240	0.022	5.6e-2
SIM, $h = 10^{-3}$	2.977378936e-43	2208	0.054	4.4e-4
SIM, $h = 10^{-4}$	2.976075557e-43	21753	0.275	5.0e-6
SIM, $h = 10^{-5}$	2.976060927e-43	203143	2.135	4.9e-8
SIM, $h = 10^{-6}$	2.976060782e-43	2081478	16.05	3.4e-10
[14]	2.976060781e-43	—	—	0.0

Table 3: Solution to the Troesch's problem with $\lambda = 100$.

[7] Of course, it might not always be the case and requires the iteration process to be convergent.
[8] For $x = 0.999000491899$
[9] For $x = 0.999000017539$
[10] Number of knots in the final mesh.
[11] System capabilities : Intel(R) Core(TM) i3-3120M, 2.5 GHz, 8 Gb RAM. Single thread implementation.
[12] Relative difference as compared to $u'(0)$ calculated in [14].

References

[1] Snyman J. A. Continuous and discontinuous numerical solutions to the troesch problem. *Journal of Computational and Applied Mathematics*, 5(3):171–175, 1979.

[2] E. L. Allgower and S. F. McCormick. Newton's method with mesh refinements for numerical solution of nonlinear two-point boundary value problems. *Numerische Mathematik*, 29(3):237–260, 1978.

[3] Shih-Hsiang Chang. A variational iteration method for solving troesch's problem. *Journal of Computational and Applied Mathematics*, 234(10):3043 – 3047, 2010.

[4] F. R. Gantmacher. *The theory of matrices. Vol. 1.* AMS Chelsea Publishing, Providence, RI, 1998. Translated from the Russian by K. A. Hirsch, Reprint of the 1959 translation.

[5] I. P. Gavrilyuk, M. Hermann, M. V. Kutniv, and V. L. Makarov. Difference schemes for nonlinear bvps using runge-kutta ivp-solvers. *Advances in Difference Equations*, 2006:Article ID 12167, 29 pages, 2006.

[6] Sung N. Ha. A nonlinear shooting method for two-point boundary value problems. *Computers and Mathematics with Applications*, 42(1011):1411 – 1420, 2001.

[7] M. Hermann and D. Kaiser. Rwpm: a software package of shooting methods for nonlinear two-point boundary value problems. *Applied Numerical Mathematics*, 13(1):103 – 108, 1993.

[8] D.J Jones. Solution of troesch's, and other, two point boundary value problems by shooting techniques. *Journal of Computational Physics*, 12(3):429 – 434, 1973.

[9] William H. Press, Saul A. Teukolsky, William T. Vetterling, and Brian P. Flannery. *Numerical Recipes 3rd Edition: The Art of Scientific Computing*. Cambridge University Press, New York, NY, USA, 3 edition, 2007.

[10] S.M Roberts and J.S Shipman. On the closed form solution of troesch's problem. *Journal of Computational Physics*, 21(3):291 – 304, 1976.

[11] J. Stoer and R. Bulirsch. *Introduction to numerical analysis*. Springer-Verlag, 2002.

[12] Gerald Teschl. *Ordinary differential equations and dynamical systems*, volume 140 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2012.

[13] B.A Troesch. A simple approach to a sensitive two-point boundary value problem. *Journal of Computational Physics*, 21(3):279 – 290, 1976.

[14] Hector Vazquez-Leal, Yasir Khan, Guillermo Fernández-Anaya, Agustín Herrera-May, Arturo Sarmiento-Reyes, Uriel Filobello-Nino, Víctor-M. Jimenez-Fernández, and Domitilo Pereyra-Díaz. A general solution for Troesch's problem. *Mathematical Problems in Engineering*, 2012.

[15] V.F. Zaitsev and A.D. Polyanin. *Handbook of Exact Solutions for Ordinary Differential Equations*. CRC Press, 2002.