ISSN 1360-1725

THE UNIVERSITY of MANCHESTER

UMIST

Identification of the initial function for delay differential equations: Part III: Nonlinear DDEs & computational results.

Christopher T. H. Baker
Research Professor, The Victoria University of Manchester;
Visiting Professor, Chester College
&
Evgeniy I. Parmuzin
The Victoria University of Manchester.

Numerical Analysis Report No. 444
March 2004
For Part I and II see Report No. 431 and 443

Manchester Centre for Computational Mathematics
Numerical Analysis Reports

DEPARTMENTS OF MATHEMATICS

Reports available from:
Department of Mathematics
University of Manchester
Manchester M13 9PL
England

And over the World-Wide Web from URLs
http://www.ma.man.ac.uk/nareports
ftp://ftp.ma.man.ac.uk/pub/narep

Contents

Identification of the initial function for nonlinear delay differential equations.

Christopher T. H. Baker* and Evginiy I . Parmuzin†

narep444.tex

Abstract

In this report we consider the "data assimilation problem" for nonlinear delay differential equations. This problem consists of finding an initial function that gives rise to a solution of a given nonlinear delay differential equation, which is a close fit to observed data. A rôle for adjoint equations and fundamental solutions in the nonlinear case is established. The "Pseudo-Newton" method to solve nonlinear "data assimilation problem" will be presented. This is an extension of results obtain in [2] for linear delay differential equations.

Keywords: Nonlinear delay differential equations, initial function, linearized adjoint equations, identification problem, data assimilation, fundamental matrices, regularization parameter, Newton iterative method.

1 A nonlinear data assimilation problem

Consider a non-linear problem in the form of a delay differential equation

$$\frac{dy(t)}{dt} = f(t, y(t), y(t-\tau)), \quad t \in [0, T] \tag{1.1a}$$

with an initial condition

$$y(t) = \varphi(t), \quad t \in [-\tau, 0]. \tag{1.1b}$$

Here $y(t), \varphi(t), f(t), \in \mathbb{R}^{m\times 1}$, $\widehat{y}(t), \widehat{\varphi}(t) \in \mathbb{R}^{m\times 1}$. The solution $y(t)$ depends upon $\varphi(t)$, $y(t) \equiv y(\varphi; t)$. The question of interest is, roughly speaking, "given $y(t)$ what is $\varphi(t)$?" This question may be not well-posed; see Remark 2.4.

We therefore introduce the functional

$$S_\alpha^{\beta,\gamma}(\varphi) := \frac{\alpha}{2}\int_{-\tau}^{0} \left\|\varphi(t) - \widehat{\varphi}(t)\right\|^2 dt + \frac{\beta}{2}\left\|\varphi(0) - \widehat{\varphi}(0)\right\|^2 + \frac{\gamma}{2}\left\|y(\varphi; 0) - \widehat{y}(0)\right\|^2 + \frac{1}{2}\int_0 \left\|y(\varphi; t) - \widehat{y}(t)\right\|^2 dt \tag{1.2}$$

(in which $\alpha, \beta, \gamma \geq 0$ and $y(\varphi; 0) = \varphi(0)$) and $\widehat{\varphi} = \widehat{\varphi}(t)$ and $\widehat{y} = \widehat{y}(t)$ are given functions and where $y(\varphi; t)$ satisfies (1.1).

We can now formulate the *data assimilation problem* as follows:

*`cthbaker@na-net.ornl.gov` Research Professor, The University of Manchester; Visiting Professor, Chester College.

†`parmuzin@ma.man.ac.uk` Supported by an ORS award in The University of Manchester.

Definition 1.1 *Let $\mathcal{F} \subseteq PC[-\tau, 0]$ denote a smoothness class of bounded functions on $[-\tau, 0]$. Then the corresponding data assimilation problem for the identification of φ reads as follows.*

Define by $y(\varphi; t)$ the solution of (1.1) with initial function φ. Find $\varphi_\star \in \mathcal{F}$, such $y(\varphi_\star; t)$ minimizes $S_\alpha^{\beta,\gamma}(\varphi)$ over $\mathcal{F}$:

$$S_\alpha^{\beta,\gamma}(\varphi_\star) = \min_{\varphi \in \mathcal{F}} S_\alpha^{\beta,\gamma}(\varphi), \tag{1.3}$$

where $S_\alpha^{\beta,\gamma}(\varphi)$ is defined by (1.2) in terms of $y(\varphi; t)$.

as the problem in which $\alpha = \beta = \gamma = 0$ is potentially ill-posed, we may introduce positive regularization parameters α, β, γ.

2 The minimization problem

2.1 Derivatives of f

In this section we define some notation and state associated results, which will guide us later.

Definition 2.1 *For $u, v \in \mathbb{R}^{m\times 1}$, define*

$$\nabla_2 f(t,u,v) = \left\{\frac{\partial}{\partial u}\right\} f(t,u,v), \quad \nabla_3 f(t,u,v) = \left\{\frac{\partial}{\partial v}\right\} f(t,u,v)$$

as the Jacobians with elements $\frac{\partial f_k}{\partial u_l}$ and $\frac{\partial f_k}{\partial v_l}$ respectively. For (1.1), and a given solution $y(t) \equiv y(\varphi; t)$, define

$$A(t) = \nabla_2 f(t, y(t), y(t-\tau)), \quad B(t) = \nabla_3 f(t, y(t), y(t-\tau)) \ \in \mathbb{R}^{m\times m}. \tag{2.1}$$

We also introduce second derivatives of $f(t,u,v)$ as follows.

Definition 2.2 *For sufficiently smooth $f(t,u,v) \in \mathbb{R}^{m\times 1}$ we have*

$$\nabla_2 f(t,u,v) \equiv \left[\left\{\frac{\partial}{\partial u}\right\} f_1(t,u,v), \left\{\frac{\partial}{\partial u}\right\} f_2(t,u,v), \ldots, \left\{\frac{\partial}{\partial u}\right\} f_m(t,u,v)\right],$$
$$\text{where } f_i(t,u,v) \in \mathbb{R}^{m\times 1}, \quad i = 1, 2, \ldots, m.$$

From this we derive the second derivative $\mathfrak{C}(t)$ of the function $f(t,u,v)$ with respect to its first variable which can be written as

$$\mathfrak{C}(t)w \equiv \left[\left\{\frac{\partial}{\partial u}\right\} \nabla_2 f(t,u,v)\right] w =$$
$$\left[\left[\left\{\frac{\partial}{\partial u}\right\}\left\{\frac{\partial}{\partial u}\right\} f_1(t,u,v)\right] w, \left[\left\{\frac{\partial}{\partial u}\right\}\left\{\frac{\partial}{\partial u}\right\} f_2(t,u,v)\right] w, \ldots, \left[\left\{\frac{\partial}{\partial u}\right\}\left\{\frac{\partial}{\partial u}\right\} f_k(t,u,v)\right] w\right].$$

For the remaining second derivatives we have the expressions

$$\mathfrak{D}_1(t)w \equiv \left[\left\{\frac{\partial}{\partial v}\right\} \nabla_2 f(t,u,v)\right] w; \quad \mathfrak{D}_2(t)w \equiv \left[\left\{\frac{\partial}{\partial u}\right\} \nabla_3 f(t,u,v)\right] w;$$

$$\mathfrak{E}(t)w \equiv \left[\left\{\frac{\partial}{\partial v}\right\} \nabla_3 f(t,u,v)\right] w.$$

Remark 2.1 *In the case, where $u, v \in \mathbb{R}$, the derivatives of $f(t,u,v)$ can be written as*

$$\mathfrak{C}(t) = \left\{\frac{\partial^2}{\partial u^2}\right\} f(t,u,v)\Big]_{u=y(t),v=y(t-\tau)}, \quad \mathfrak{D}_1(t) = \left\{\frac{\partial^2}{\partial u \partial v}\right\} f(t,u,v)\Big]_{u=y(t),v=y(t-\tau)},$$

$$\mathfrak{D}_2(t) = \left\{\frac{\partial^2}{\partial v \partial u}\right\} f(t,u,v)\Big]_{u=y(t),v=y(t-\tau)}, \quad \mathfrak{E}(t) = \left\{\frac{\partial^2}{\partial v^2}\right\} f(t,u,v)\Big]_{u=y(t),v=y(t-\tau)}.$$

For sufficiently smooth $f(t,u,v)$ we have $\mathfrak{D}_1(t) = \mathfrak{D}_2(t)$.

The scalar form of Taylor's theorem gives us

$$f(t, u+\delta u, v+\delta v) = f(t,u,v) + \left[\left\{\frac{\partial}{\partial u}\right\} f(t,u,v)\right]\delta u + \left[\left\{\frac{\partial}{\partial v}\right\} f(t,u,v)\right]\delta v +$$

$$\frac{1}{2}\left\{\left[\left\{\frac{\partial^2}{\partial u^2}\right\} f(t,u,v)\right]\delta u^2 + 2\left[\left\{\frac{\partial^2}{\partial u \partial v}\right\} f(t,u,v)\right]\delta u \delta v + \left[\left\{\frac{\partial^2}{\partial v^2}\right\} f(t,u,v)\right]\delta v^2\right\} +$$

$$\mathcal{O}(\|\delta u^2\| + \|\delta v^2\|).$$

Remark 2.2 *Taylor's theorem for a vector-valued function can be regarded ([6, page 76-80]) as a special case of a result expressible in terms of Fréchet or Gateaux derivatives within the setting of operator on a Euclidean space: if $F: \mathbb{R}^m \to \mathbb{R}^m$ has a second Fréchet derivative*

$$\lim_{\|h\| \to 0} \frac{1}{\|h\|}\Big(F(x+h) - F(x) - F'(x)h - \frac{1}{2}F''(x)hh\Big) = 0.$$

Here $F'(x)$ corresponds to the Jacobian and

$$[F''(x)hk]^T = \Big(k^T H_1(x)h, k^T H_2(x)h, \ldots, k^T H_m(x)h\Big)$$

for $h, k \in \mathbb{R}^m$ and where $H_i(x) \equiv H_{f_i}(x)$ is the $m \times m$ Hessian matrix ($i = 1, 2, \ldots, m$):

$$H_i(x) = \begin{pmatrix} \partial_1 \partial_1 f_i(x) & \ldots & \partial_m \partial_1 f_i(x) \\ \vdots & \ddots & \vdots \\ \partial_1 \partial_m f_i(x) & \ldots & \partial_m \partial_m f_i(x) \end{pmatrix},$$

where ∂_i, ($i = 1, 2, \ldots, m$) is a partial derivative with respect of i-th component of the vector x.

Remark 2.3 *We use vector-valued functions where it is possible. But, in some stages, for simplicity, we obtain results only for a scalar-valued function. Notational complexities arise when we try to write Taylor's expansion up to second order terms for a vector-valued function.*

Remark 2.4 *We return briefly to the original mathematical problem of determining $\varphi(t)$ given $y(t)$ for $t \geq 0$. We suppose that, with $y(t)$ given, we also have $w(t) = y'(t)$. (However, the problem of determining $y'(t)$, given $y(t)$, is ill-posed.) Then $\varphi(t)$ satisfies the equation*

$$w(t) = f(t, y(t), \varphi(t-\tau)), \quad t \in [0, \tau].$$

By the implicit function theorem [4], this equation can be solved uniquely for $\varphi(t-\tau)$ ($t \in [0,\tau]$) provided that f is sufficiently smooth and $f_v(t, y(t), v) \neq 0$ for $t \in [0,\tau]$. If there are two different solutions φ_1 and φ_2 then $f(t, y(t), \varphi_1(t)) = f(t, y(t), \varphi_1(t))$, for $t \in [0,\tau]$ and there exist a function $\xi(t)$ such that $f_v(t, y(t), \xi(t)) = 0$, $t \in [0,\tau]$.

2.2 A necessary condition for a minimum

In order to discuss the minimum of the functional $S_\alpha^{\beta,\gamma}(\varphi)$ we need to analyze $S_\alpha^{\beta,\gamma}(\varphi+\varepsilon\psi)$. Let $\varphi_\star \in \mathcal{F}$ provide a minimum to the functional $S_\alpha^{\beta,\gamma}(\varphi)$, so that we have $S_\alpha^{\beta,\gamma}(\varphi_\star) \le S_\alpha^{\beta,\gamma}(\varphi_\star+\varepsilon\psi)$, where ε is a real parameter and ψ is an arbitrary function in $\mathcal{F}$. The term $\varepsilon\psi(t)$ is a perturbation in the initial function.

Provided that f is sufficiently smooth (f has continuous second-order derivatives), it can be shown that the solution of equation (1.1), when the initial condition is perturbed has the form

$$y(\varphi+\varepsilon\psi,t) = y(\varphi,t) + \varepsilon z(\psi,t) + \mathcal{O}(\varepsilon^2).$$

In the case of a linear DDE the term $\mathcal{O}(\varepsilon^2)$ are missing. For the nonlinear case, the function $z(\psi,t)$ satisfies (see Lemma 2.1 below) an equation that reduces to that obtained earlier for the linear case [2].

Lemma 2.1 *Suppose that f is sufficiently smooth and the nonlinear delay differential equation in the form*

$$\frac{dy(t)}{dt} = f(t, y(t), y(t-\tau)), \quad t \in [0,T] \tag{2.2}$$

with initial condition

$$y(t) = \varphi(t), \quad t \in [-\tau, 0],$$

has a unique solution $y(t) = y(\varphi;t)$. For the perturbed initial function $\varphi(t)+\varepsilon\psi(t)$ suppose there is a unique perturbed solution $y(\varphi+\varepsilon\psi;t)$ that is bounded on $[0,T]$. Suppose that $z(t) = z(\psi;t)$ satisfies

$$\frac{dz(t)}{dt} = A(t)z(t) + B(t)z(t-\tau), \quad t \in [0,T] \tag{2.3a}$$

(where $A(t)$ and $B(t)$ are given in Definition 2.1), with initial condition

$$z(t) = \psi(t), \; t \in [-\tau, 0]. \tag{2.3b}$$

Then

$$y(\varphi+\varepsilon\psi;t) = y(\varphi;t) + \varepsilon z(\psi;t) + \mathcal{O}(\varepsilon^2), \quad \text{as } \varepsilon \to 0. \tag{2.4}$$

Remark 2.5 *The term $\mathcal{O}(\varepsilon^2)$ in (2.4) contains (for the scalar case) higher order terms, of the type $\mathfrak{C}(t)\big(z(t)\big)z(t)$, $\mathfrak{E}(t)\big(z(t-\tau)\big)z(t-\tau)$, $\mathfrak{D}_1(t)\big(z(t-\tau)\big)z(t)$, $\mathfrak{D}_2(t)\big(z(t)\big)z(t-\tau)$, where $\mathfrak{C}(t)$, $\mathfrak{E}(t)$, $\mathfrak{D}_i(t)$ $(i=1,2)$ are certain second derivatives of $f(t,u,v)$.*

Proof. Let us define $w_\varepsilon(t) = y(\varphi+\varepsilon\psi;t) - y(\varphi;t)$, with $w_\varepsilon(t) = \varepsilon\psi(t)$ for $t \in [-\tau,0]$, and

$$w'_\varepsilon(t) = y'(\varphi+\varepsilon\psi;t) - y'(\varphi;t) =$$

$$f(t, y(\varphi+\varepsilon\psi;t), y(\varphi+\varepsilon\psi;t-\tau)) - f(t, y(\varphi;t), y(\varphi;t-\tau)). \tag{2.5}$$

By the mean value theorem, there exist bounded functions $\mathcal{A}_\varepsilon(t)$, $\mathcal{B}_\varepsilon(t)$ such that

$$w'_\varepsilon(t) = \mathcal{A}_\varepsilon(t)w_\varepsilon(t) + \mathcal{B}_\varepsilon(t)w_\varepsilon(t-\tau), \quad t \in [0,T];$$

$$w_\varepsilon(t) = \varepsilon\psi(t), \quad t \in [-\tau, 0].$$

By induction on n, and the method of steps, $|w_\varepsilon(t)| \le C_n\varepsilon$ for $t \in [(n-1)\tau, n\tau]$ and hence $w_\varepsilon(t) = \mathcal{O}(\varepsilon)$ for $t \in [-\tau, T]$. Returning to equation (2.5), Taylor's theorem yields

$$w'_\varepsilon(t) = A(t)w_\varepsilon(t) + B(t)w_\varepsilon(t-\tau) + r_\varepsilon(t), \quad t \in [0, T],$$

$$w_\varepsilon(t) = \varepsilon\psi(t), \quad t \in [-\tau, 0],$$

where $A(t)$ and $B(t)$ are given in Definition 2.1. Since the second derivatives of $f(t, u, v)$ with respect to u and v exist and all bounded, and $w_\varepsilon(t) = \mathcal{O}(\varepsilon)$, $\sup_{t\in[0,T]} |r_\varepsilon(t)| = \mathcal{O}(\varepsilon^2)$. Further, we can employ the fundamental soluton $Y(s, t)$ (see [1]) to write

$$w(t) = \varepsilon Y(t, 0)\psi(0) + \varepsilon \int_{-\tau}^{0} Y(t, s+\tau)B(s+\tau)\psi(s)ds + \int_{0}^{t} Y(t, s)r_\varepsilon(s)ds.$$

Thus,

$$w_\varepsilon(t) = \varepsilon z(\psi; t) + \mathcal{O}(\varepsilon^2),$$

as required and Lemma 2.1 is established.

We can write

$$S_\alpha^{\beta,\gamma}(\varphi + \varepsilon\psi) = S_\alpha^{\beta,\gamma}(\varphi) + \varepsilon P_\alpha^{\beta,\gamma}(\varphi, \psi) + \mathcal{O}(\varepsilon^2), \tag{2.6a}$$

where

$$P_\alpha^{\beta,\gamma}(\varphi, \psi) = \alpha \int_{-\tau}^{0} \{\varphi(t) - \widehat{\varphi}(t)\}\psi(t)dt + \tag{2.6b}$$

$$\int_{0}^{T} \{y(\varphi; t) - \widehat{y}(t)\}z(\psi; t)dt + \beta\{\varphi(0) - \widehat{\varphi}(0)\}\psi(0) + \gamma\{y(\varphi; 0) - \widehat{y}(0)\}z(\psi; 0).$$

Remark 2.6 *Given sufficiently differentiability Lemma 2.1 can be strengthened to establish a result of the form*

$$w_\varepsilon(t) = \varepsilon z(\psi; t) + \varepsilon^2\omega(\psi; t) + \mathcal{O}(\varepsilon^3)$$

and (2.6a) becomes

$$S_\alpha^{\beta,\gamma}(\varphi + \varepsilon\psi) = S_\alpha^{\beta,\gamma}(\varphi) + \varepsilon P_\alpha^{\beta,\gamma}(\varphi, \psi) + \varepsilon^2 Q_\alpha^{\beta,\gamma}(\varphi, \psi) + \mathcal{O}(\varepsilon^3), \tag{2.7}$$

where $Q_\alpha^{\beta,\gamma}(\varphi, \psi) = \dfrac{\alpha}{2} \int_{-\tau}^{0} |\psi(t)|^2 dt + \dfrac{1}{2} \int_{0}^{T} \Big(|z(\psi; t)|^2 + 2\omega(t)[y(\varphi; t) - \widehat{y}(t)]\Big)dt + \frac{\beta}{2}|\psi(0)|^2 + \frac{\gamma}{2}\Big(|z(\psi; 0)|^2 + 2\omega(0)[y(\varphi; 0) - \widehat{y}(0)]\Big)$.

$P_\alpha^{\beta,\gamma}(\varphi, \psi)$ is the Gateaux derivative of $S_\alpha^{\beta,\gamma}(\varphi)$ in the direction ψ:

$$\lim_{\varepsilon\to 0} \frac{S_\alpha^{\beta,\gamma}(\varphi + \varepsilon\psi) - S_\alpha^{\beta,\gamma}(\varphi)}{\varepsilon} = P_\alpha^{\beta,\gamma}(\varphi, \psi). \tag{2.8}$$

At the local minimum of the functional (1.2) this derivative must be zero for all $\psi(t)$, and hence (2.6b) must vanish for all solutions $z(\psi; t)$ of the equation (2.3) in which $z(\psi; t) = \psi(t)$ when $t \in [-\tau, 0]$.

We have the following result.

Theorem 2.1 *A necessary condition for a function $\varphi_\star(t)$ defined on $[-\tau, 0]$ to minimizes* $S_\alpha^{\beta,\gamma}(\varphi) = \frac{\alpha}{2}\int_{-\tau}^{0} \|\varphi(t) - \widehat{\varphi}(t)\|^2 dt + \frac{\beta}{2}\|\varphi(0) - \widehat{\varphi}(0)\|^2 + \frac{\gamma}{2}\|y(\varphi; 0) - \widehat{y}(0)\|^2 + \frac{1}{2}\int_0 \|y(\varphi; t) - \widehat{y}(t)\|^2 dt$ *for* $\varphi \in \mathcal{F}$ *is that* $P_\alpha^{\beta,\gamma}(\varphi, \psi)$ *as given in (2.6b) vanishes for all* $\psi \in \mathcal{F}$*, where* $z = z(\psi; t)$ *satisfies (2.3).*

2.3 Rôle of an adjoint equation in the non-linear case

In this section we shall obtain (in Theorem 2.2) an equivalent formulation of the problem (1.3) in the non-linear case, based upon the adjoint equations. Theorem 2.2 provides an alternative result to Theorem 2.1. We first need a lemma.

Lemma 2.2 *Let* $y = y(\varphi; t)$ *be a solution of the problem* (1.1) *and let* $z = z(\psi; t)$ *be a solution of the linearized problem* (2.3)*. Then the first variation of the functional* $S_\alpha^{\beta,\gamma}(\varphi)$ *can be represented in the form* $P_\alpha^{\beta,\gamma}(\varphi, \psi) :=$

$$\int_{-\tau}^{0} \Big\{\alpha[\varphi(t) - \widehat{\varphi}(t)] + x(t+\tau)B(t+\tau)\Big\}\psi(t)dt + \Big(x(0) + \beta[\varphi(0) - \widehat{\varphi}(0)] + \gamma[y(\varphi; 0) - \widehat{y}(0)]\Big)\psi(0),$$

where $x(t) \in \mathbb{R}^{1\times n}$ *is the solution* $(x(t) \equiv x(\varphi; t))$ *of the problem*

$$-\frac{dx(t)}{dt} - x(t)A(t) - x(t+\tau)B(t+\tau) = [y(\varphi; t) - \widehat{y}(t)], \quad \text{for } t \in [0, T], \tag{2.9a}$$

$$x(t) = 0, \quad \text{for } t \in [T, T+\tau] \tag{2.9b}$$

with coefficients defined by (2.1).

Equation (2.9) is a formal adjoint equation for (2.3), with a special forcing term $y(\varphi; t) - \widehat{y}(t)$.

Proof. For $P_\alpha^{\beta,\gamma}(\varphi, \psi)$ we have

$$P_\alpha^{\beta,\gamma}(\varphi, \psi) = \alpha\int_{-\tau}^{0} [\varphi(t) - \widehat{\varphi}(t)]\psi(t)dt + \beta[\varphi(0) - \widehat{\varphi}(0)]\psi(0) + \gamma[y(\varphi; 0) - \widehat{y}(0)]z(\psi; 0) +$$

$$\Big\{\int_0^T \big(-\frac{dx(t)}{dt} - x(t)A(t) - x(t+\tau)B(t+\tau)\big)z(\psi; t)dt\Big\}$$

$$= \alpha\int_{-\tau}^{0} [\varphi(t) - \widehat{\varphi}(t)]\psi(t)dt + \beta[\varphi(0) - \widehat{\varphi}(0)]\psi(0) + \gamma[y(\varphi; 0) - \widehat{y}(0)]z(\psi; 0) +$$

$$\underbrace{-\int_0^T \frac{dx(t)}{dt}z(\psi; t)dt}_{\text{term } (i)} \underbrace{-\int_0^T x(t)A(t)z(\psi; t)dt}_{\text{term } (ii)} \underbrace{-\int_0^T x(t+\tau)B(t+\tau)z(\psi; t)dt}_{\text{term } (iii)}. \tag{2.10}$$

Using integration by parts we can write the term (i) (involving the derivative) in (2.10) as

$$-x(t)z(\psi;t)\Big|_0^T + \int_0^T x(t)\frac{dz(\psi;t)}{dt}dt = x(0)\psi(0) + \int_0^T x(t)\frac{dz(\psi;t)}{dt}dt$$

since $x(T) = 0$ and $z(\psi, 0+) = \psi(0)$. If, in the term (iii) (involving $x(t+\tau)$), we substitute $t + \tau = s$ and take into account that $x(t) = 0$ for $t \in [T, T+\tau]$ we obtain

$$\int_0^T x(t+\tau)B(t+\tau)z(\psi;t)dt = \int_\tau^{T+\tau} x(s)B(s)z(\psi;s-\tau)ds$$

$$= \int_0^T x(s)B(s)z(\psi;s-\tau)ds - \int_{-\tau}^0 x(s+\tau)B(s+\tau)\psi(s)ds.$$

Finally, we can rewrite the first variation in the form

$$P_\alpha^{\beta,\gamma}(\varphi,\psi) = \alpha \int_{-\tau}^0 [\varphi(t) - \widehat{\varphi}(t)]\psi(t)dt + x(0)z(\psi;0) + \beta[\varphi(0) - \widehat{\varphi}(0)]\psi(0) + \gamma[y(\varphi;0) - \widehat{y}(0)]z(\psi;0)$$

$$+ \int_{-\tau}^0 x(t+\tau)B(t+\tau)\psi(t)dt + \int_0^T x(t)\left(\frac{dz(\psi;t)}{dt} - A(t)z(\psi;t) - B(t)z(\psi,t-\tau)\right)dt.$$

The function $z(\psi;t)$ satisfies the equation (2.3). Therefore, we can write

$$P_\alpha^{\beta,\gamma}(\varphi,\psi) = \int_{-\tau}^0 \big(\alpha[\varphi(t) - \widehat{\varphi}(t)] + x(t+\tau)B(t+\tau)\big)\psi(t)dt +$$

$$\Big(x(0) + \beta[\varphi(0) - \widehat{\varphi}(0)] + \gamma[y(\varphi;0) - \widehat{y}(0)]\Big)\psi(0),$$

and Lemma 2.2 is established.

Combining Lemma 2.2 and Theorem 2.1 we obtain the following result.

Theorem 2.2 *If a function $\varphi_\star(t)$ defined on $[-\tau, 0]$ minimizes $S_\alpha^{\beta,\gamma}(\varphi)$ for $\varphi \in \mathcal{F}$, where $\mathcal{F} = \{\varphi \in C[-\tau, 0)$ with bounded $\varphi(0)\}$ then*

$$\int_{-\tau}^0 \{\alpha[\varphi(t) - \widehat{\varphi}(t)] + x(t+\tau)B(t+\tau)\}\psi(t)dt = 0 \tag{2.11a}$$

for all $\psi \in C[-\tau, 0)$, where x satisfies (2.9), and

$$x(0) + \beta[\varphi(0) - \widehat{\varphi}(0)] + \gamma[\varphi(0) - \widehat{y}(0)] = 0. \tag{2.11b}$$

2.4 A system of equations that defines the optimal φ

Using the result of Lemma 2.2 and Theorem 2.2, we can obtain a system of equations that defines an initial function $\varphi_\star(t)$ which makes $P_\alpha^{\beta,\gamma}(\varphi,\psi)$ vanish for all ψ. We have

$$\frac{dy(t)}{dt} = f(t, y(t), y(t-\tau)), \quad \text{for } t \in [0, T], \tag{2.12a}$$

$$y(t) = \varphi_\star(t) \quad \text{for } t \in [-\tau, 0), \quad y(0) = \varphi_\star(0), \tag{2.12b}$$

$$-\frac{dx(t)}{dt} - x(t)A(t) - x(t+\tau)B(t+\tau) = [y(\varphi_\star; t) - \widehat{y}(t)], \quad \text{for } t \in [0, T], \tag{2.12c}$$

$$x(t) = 0, \quad \text{for } t \in [T, T+\tau], \tag{2.12d}$$

$$\alpha(\varphi_\star(t) - \widehat{\varphi}(t)) + [B(t+\tau)]x(t+\tau) = 0 \quad \text{for } t \in [-\tau, 0), \tag{2.12e}$$

$$x(0) + \beta\{\varphi_\star(0) - \widehat{\varphi}(0)\} + \gamma\{\varphi_\star(0) - \widehat{y}(0)\} = 0. \tag{2.12f}$$

Our formulation comprises the solution of the coupled equations (2.12a), (2.12c) and the additional equations for the initial function (2.12e), (2.12f).

Lemma 2.3 *If the solution $\varphi(t)$ of (2.12) is unique, it provide a minimum of $S_\alpha^{\beta,\gamma}(\varphi)$.*

Proof. $S_\alpha^{\beta,\gamma}(\varphi)$ in (1.2) is non-negative. If it has a unique stationary point, this must be a minimum.

3 A "Pseudo-Newton" iterative method

Our aim is to obtain an iterative method to solve the "data assimilation problem" for a nonlinear delay equation. In this section we present a "Pseudo-Newton" iterative process to solve the optimization problem (1.3). (Unfortunately, it is not quadratically convergent.)

To keep the notation simple we consider the scalar case. We write $y(t) = \tilde{y}(t) + \delta\tilde{y}(t)$, $\varphi(t) = \tilde{\varphi}(t) + \delta\tilde{\varphi}(t)$, $x(t) = \tilde{x}(t) + \delta\tilde{x}(t)$ in (2.12). Then $y(t-\tau) = \tilde{y}(t-\tau) + \delta\tilde{y}(t-\tau)$ and $x(t+\tau) = \tilde{x}(t+\tau) + \delta\tilde{x}(t+\tau)$. We can rewrite (2.12) as

$$\frac{d}{dt}\Big\{\tilde{y}(t) + \delta\tilde{y}(t)\Big\} = f\big(t, \tilde{y}(t) + \delta\tilde{y}(t), \tilde{y}(t-\tau) + \delta\tilde{y}(t-\tau)\big), \quad \text{for } t \in [0, T], \tag{3.1a}$$

$$\tilde{y}(t) + \delta\tilde{y}(t) = \tilde{\varphi}(t) + \delta\varphi(t) \quad \text{for } t \in [-\tau, 0], \tag{3.1b}$$

$$-\frac{d}{dt}\Big\{\tilde{x}(t) + \delta\tilde{x}(t)\Big\} - \big(\tilde{x}(t) + \delta\tilde{x}(t)\big)\nabla_2 f(t, y(t) + \delta\tilde{y}(t), y(t-\tau) + \delta\tilde{y}(t-\tau))$$
$$-\big(\tilde{x}(t+\tau)+\delta\tilde{x}(t+\tau)\big)\nabla_3 f(t, y(t)+\delta\tilde{y}(t), y(t-\tau)+\delta\tilde{y}(t-\tau)) = [\tilde{y}(t)+\delta\tilde{y}(t)-\widehat{y}(t)], \quad \text{for } t \in [0, T], \tag{3.1c}$$

$$\tilde{x}(t) + \delta\tilde{x}(t) = 0, \quad \text{for } t \in [T, T+\tau]. \tag{3.1d}$$

(Here, $\nabla_2 f(t,u,v) = f_u(t,u,v)$ and $\nabla_3 f(t,u,v) = f_v(t,u,v)$.)

$$\alpha(\tilde{\varphi}(t)+\delta\varphi(t)-\widehat{\varphi}(t))+\Big[\nabla_3 f(t, y(t)+\delta\tilde{y}(t), y(t-\tau)+\delta\tilde{y}(t-\tau))\Big](\tilde{x}(t+\tau)+\delta\tilde{x}(t+\tau)) = 0, \quad \text{for } t \in [-\tau, 0), \tag{3.1e}$$

$$\tilde{x}(0) + \delta\tilde{x}(0) + \beta\{\tilde{\varphi}(0) + \delta\varphi(0) - \widehat{\varphi}(0)\} + \gamma\{\tilde{\varphi}(0) + \delta\varphi(0) - \widehat{y}(0)\} = 0. \tag{3.1f}$$

Using Taylor expansions, and assuming sufficient differentiability, we can write, for use in (3.1a),

$$f\big(t, \tilde{y}(t)+\delta\tilde{y}(t), \tilde{y}(t-\tau)+\delta\tilde{y}(t-\tau)\big) =$$

$$f\big(t, \tilde{y}(t), \tilde{y}(t-\tau)\big) + A(t)\delta\tilde{y}(t) + B(t)\delta\tilde{y}(t-\tau) + \mathcal{O}(\big|[\delta\tilde{y}(t), \delta\tilde{y}(t-\tau)]^T\big|^2), \tag{3.2a}$$

and for use in (3.1c),

$$\nabla_2 f(t, y(t)+\delta\tilde{y}(t), y(t-\tau)+\delta\tilde{y}(t-\tau)) = A(t) + \mathfrak{C}(t)\big(\delta\tilde{y}(t)\big) + \mathfrak{D}_1(t)\big(\delta\tilde{y}(t-\tau)\big) + \mathcal{O}(\big|[\delta\tilde{y}(t), \delta\tilde{y}(t-\tau)]^T\big|^2), \tag{3.2b}$$

while for use in (3.1c) and (3.1e) we have

$$\nabla_3 f(t+\tau, y(t+\tau)+\delta\tilde{y}(t+\tau), y(t)+\delta\tilde{y}(t)) = B(t+\tau) + \mathfrak{D}_2(t+\tau)\big(\delta\tilde{y}(t+\tau)\big) + \mathfrak{E}(t+\tau)\big(\delta\tilde{y}(t)\big) + \mathcal{O}(\big|[\delta\tilde{y}(t+\tau), \delta\tilde{y}(t)]^T\big|^2), \tag{3.2c}$$

where $\mathfrak{C}(t)$, $\mathfrak{D}_1(t)$, $\mathfrak{D}_2(t)$ and $\mathfrak{E}(t)$ were defined in Remark 2.1.

Substitute (3.2) into (3.1) and we can give an explicit expression for the first-order terms:

$$\frac{d}{dt}\Big\{\tilde{y}(t)+\delta\tilde{y}(t)\Big\} = f\big(t, \tilde{y}(t), \tilde{y}(t-\tau)\big) + A(t)\delta\tilde{y}(t) + B(t)\delta\tilde{y}(t-\tau) + \mathcal{O}(\big|\delta\tilde{y}(t)\big|^2 + \big|\delta\tilde{y}(t-\tau)\big|^2) \quad \text{for } t\in[0,T], \tag{3.3}$$

$$\tilde{y}(t)+\delta\tilde{y}(t) = \tilde{\varphi}(t)+\delta\varphi(t) \quad \text{for } t\in[-\tau,0], \tag{3.4}$$

so that

$$\frac{d}{dt}\Big\{\tilde{y}(t)+\delta\tilde{y}(t)\Big\} \approx f\big(t, \tilde{y}(t), \tilde{y}(t-\tau)\big) + A(t)\delta\tilde{y}(t) + B(t)\delta\tilde{y}(t-\tau), \quad \text{for } t\in[0,T], \tag{3.5a}$$

$$\tilde{y}(t)+\delta\tilde{y}(t) = \tilde{\varphi}(t)+\delta\varphi(t) \quad \text{for } t\in[-\tau,0]. \tag{3.5b}$$

The approximation symbol $\approx$ in (3.5a) results from ignoring the second order terms in (3.2). Likewise

$$-\frac{d}{dt}\Big\{\tilde{x}(t)+\delta\tilde{x}(t)\Big\} - \tilde{x}(t)\Big\{\ A(t) + \mathfrak{C}(t)\big(\delta\tilde{y}(t)\big) + \mathfrak{D}_1(t)\big(\delta\tilde{y}(t-\tau)\big)\Big\} - \delta\tilde{x}(t)A(t) + -\tilde{x}(t+\tau)\Big\{B(t+\tau) + \mathfrak{D}_2(t+\tau)\big(\delta\tilde{y}(t+\tau)\big) + \mathfrak{E}(t+\tau)\big(\delta\tilde{y}(t)\big)\Big\} - \delta\tilde{x}(t+\tau)B(t+\tau) \approx [\tilde{y}(t)+\delta\tilde{y}(t)-\widehat{y}(t)], \quad \text{for } t\in[0,T], \tag{3.5c}$$

$$\tilde{x}(t)+\delta\tilde{x}(t) = 0, \quad \text{for } t\in[T, T+\tau], \tag{3.5d}$$

$$\alpha(\tilde{\varphi}(t)+\delta\varphi(t)-\widehat{\varphi}(t)) + \Big\{B(t+\tau) + \mathfrak{D}_2(t+\tau)\big(\delta\tilde{y}(t+\tau)\big) + \mathfrak{E}(t+\tau)\big(\delta\varphi(t)\big)\Big\}\tilde{x}(t+\tau) + [B(t+\tau)]\delta\tilde{x}(t+\tau) \approx 0, \quad \text{for } t\in[-\tau,0), \tag{3.5e}$$

$$\tilde{x}(0)+\delta\tilde{x}(0) + \beta\{\tilde{\varphi}(0)+\delta\varphi(0)-\widehat{\varphi}(0)\} + \gamma\{\tilde{\varphi}(0)+\delta\varphi(0)-\widehat{y}(0)\} = 0. \tag{3.5f}$$

3.1 A "one stage" iterative process

To formulate a "Pseudo-Newton" iterative process for solving (2.12), we introduce

$$\tilde{y}^{[j+1]}(t) = \tilde{y}^{[j]}(t)+\delta\tilde{y}^{[j]}(t), \quad \tilde{\varphi}^{[j+1]}(t) = \tilde{\varphi}^{[j]}(t)+\delta\tilde{\varphi}^{[j]}(t), \quad \tilde{y}^{[j+1]}(t-\tau) = \tilde{y}^{[j]}(t-\tau)+\delta\tilde{y}^{[j]}(t-\tau),$$

$$\tilde{x}^{[j+1]}(t) = \tilde{x}^{[j]}(t)+\delta\tilde{x}^{[j]}(t), \ \text{ and } \ \tilde{x}^{[j+1]}(t+\tau) = \tilde{x}^{[j]}(t+\tau)+\delta\tilde{x}^{[j]}(t+\tau), \quad j\in\mathbb{N}.$$

Then, using (3.5) as motivation, we can define the following iterative process: given $\varphi^{[1]}(s)$ when $s \in [-\tau, 0]$ and $y^{[0]}(t)$, $x^{[0]}(t)$ when $t \in [0,T]$, define for $j = 0, 1, \ldots$

$$\frac{d}{dt}\Big\{\tilde{y}^{[j+1]}(t)\Big\} - A^{[j]}(t)\tilde{y}^{[j+1]}(t) - B^{[j]}(t)\tilde{y}^{[j+1]}(t-\tau) = f\big(t, \tilde{y}^{[j]}(t), \tilde{y}^{[j]}(t-\tau)\big) +$$

$$-A^{[j]}(t)\tilde{y}^{[j]}(t) - B^{[j]}(t)\tilde{y}^{[j]}(t-\tau), \quad \text{for } t \in [0,T], \tag{3.6a}$$

$$\tilde{y}^{[j+1]}(t) = \tilde{\varphi}^{[j+1]}(t) \quad \text{for } t \in [-\tau, 0], \tag{3.6b}$$

$$-\frac{d}{dt}\Big\{\tilde{x}^{[j+1]}(t)\Big\} - \tilde{x}^{[j+1]}(t)A^{[j]}(t) - \tilde{x}^{[j+1]}(t+\tau)B^{[j]}(t+\tau) =$$

$$\tilde{x}^{[j]}(t)\Big\{\ \mathfrak{C}^{[j]}(t)\big(\tilde{y}^{[j+1]}(t) - \tilde{y}^{[j]}(t)\big) + \mathfrak{D}_1^{[j]}(t)\big(\tilde{y}^{[j+1]}(t-\tau) - \tilde{y}^{[j]}(t-\tau)\big)\Big\} +$$

$$\tilde{x}^{[j]}(t+\tau)\Big\{\mathfrak{D}_2^{[j]}(t+\tau)\big(\tilde{y}^{[j+1]}(t+\tau) - \tilde{y}^{[j]}(t+\tau)\big) + \mathfrak{E}^{[j]}(t+\tau)\big(\tilde{y}^{[j+1]}(t) - \tilde{y}^{[j]}(t)\big)\Big\} +$$

$$[\tilde{y}^{[j+1]}(t) - \widehat{y}(t)], \quad \text{for } t \in [0,T], \tag{3.6c}$$

$$\tilde{x}^{[j+1]}(t) = 0, \quad \text{for } t \in [T, T+\tau], \tag{3.6d}$$

$$\frac{\tilde{\varphi}^{[j+2]}(t) - \tilde{\varphi}^{[j+1]}(t)}{\Delta_j} = \alpha(\tilde{\varphi}^{[j+1]}(t) - \widehat{\varphi}(t)) + [B^{[j]}(t+\tau)]\tilde{x}^{[j+1]}(t+\tau) +$$

$$\Big\{\mathfrak{D}_2^{[j]}(t+\tau)\big(\tilde{y}^{[j+1]}(t+\tau) - \tilde{y}^{[j]}(t+\tau)\big) + \mathfrak{E}^{[j]}(t+\tau)\big(\tilde{\varphi}^{[j+1]}(t) - \tilde{\varphi}^{[j]}(t)\big)\Big\}\tilde{x}^{[j]}(t+\tau), \tag{3.6e}$$

$$\frac{\tilde{\varphi}^{[j+2]}(0) - \tilde{\varphi}^{[j+1]}(0)}{\Delta'_j} = (\beta+\gamma)\tilde{\varphi}^{[j+1]}(0) + \tilde{x}^{[j+1]}(0) - \beta\widehat{\varphi}(0) - \gamma\widehat{y}(0). \tag{3.6f}$$

Here, j is the iteration index. Let us now, for simplicity, define

$$\mathcal{A}^{[j]}(t) = \mathfrak{C}^{[j]}(t)(\tilde{y}^{[j]}(t)) + \mathfrak{D}_1^{[j]}(t)(\tilde{y}^{[j]}(t-\tau)) \ \text{ and } \ \mathcal{B}^{[j]}(t) = \mathfrak{D}_2^{[j]}(t)(\tilde{y}^{[j]}(t)) + \mathfrak{E}^{[j]}(t)(\tilde{y}^{[j]}(t-\tau)). \tag{3.7}$$

Then, we can write (3.6) in brief as

$$\frac{d}{dt}\Big\{\tilde{y}^{[j+1]}(t)\Big\} - A^{[j]}(t)\tilde{y}^{[j+1]}(t) - B^{[j]}(t)\tilde{y}^{[j+1]}(t-\tau) = f\big(t, \tilde{y}^{[j]}(t), \tilde{y}^{[j]}(t-\tau)\big) +$$

$$-A^{[j]}(t)\tilde{y}^{[j]}(t) - B^{[j]}(t)\tilde{y}^{[j]}(t-\tau), \ \text{ for } t \in [0,T], \tag{3.8a}$$

$$\tilde{y}^{[j+1]}(t) = \tilde{\varphi}^{[j+1]}(t) \quad \text{for } t \in [-\tau, 0], \tag{3.8b}$$

$$-\frac{d}{dt}\Big\{\tilde{x}^{[j+1]}(t)\Big\} - \tilde{x}^{[j+1]}(t)A^{[j]}(t) - \tilde{x}^{[j+1]}(t+\tau)B^{[j]}(t+\tau) = \tilde{x}^{[j]}(t)\mathcal{A}^{[j+1]}(t) +$$

$$\tilde{x}^{[j]}(t+\tau)\mathcal{B}^{[j+1]}(t+\tau) - \tilde{x}^{[j]}(t)\mathcal{A}^{[j]}(t) - \tilde{x}^{[j]}(t+\tau)\mathcal{B}^{[j]}(t+\tau) + [\tilde{y}^{[j+1]}(t) - \widehat{y}(t)], \tag{3.8c}$$

$$\tilde{x}^{[j+1]}(t) = 0, \quad \text{for } t \in [T, T+\tau], \tag{3.8d}$$

$$\tilde{\varphi}^{[j+2]}(t) = \tilde{\varphi}^{[j+1]}(t) + \Delta_j\Big(\alpha(\tilde{\varphi}^{[j+1]}(t) - \widehat{\varphi}(t)) + B^{[j]}(t+\tau)\tilde{x}^{[j+1]}(t+\tau) +$$

$$\mathcal{B}^{[j+1]}(t+\tau)\tilde{x}^{[j]}(t+\tau) - \mathcal{B}^{[j]}(t+\tau)\tilde{x}^{[j]}(t+\tau)\Big) \quad \text{for } t \in [-\tau, 0), \tag{3.8e}$$

$$\tilde{\varphi}^{[j+2]}(0) = \tilde{\varphi}^{[j+1]}(0) + \Delta'_j\Big((\beta+\gamma)\tilde{\varphi}^{[j+1]}(0) + \tilde{x}^{[j+1]}(0) - \beta\widehat{\varphi}(0) - \gamma\widehat{y}(0)\Big). \tag{3.8f}$$

Thus, the system (3.8) is in some sense a Newton-like iterative method for solving the optimization problem (1.3) for the nonlinear DDE (1.1).

3.1.1 Retrospective

We justify (heuristically) the application of the term "Pseudo-Newton" iteration for the system above.

The underlying problem associated with (2.12) can be put in the framework of

$$\Psi(y, x, \varphi) = 0$$

Then we have

$$\Psi(y_j + \delta y_j, x_j + \delta x_j, \varphi_j + \delta\varphi_j) = 0$$

and, if $\Psi_{j+1}(y_{j+1}, x_{j+1}, \varphi_{j+1}) = 0$ we have

$$\Psi_{j+1}((y_{j+1}, x_{j+1}, \varphi_{j+1})) \approx \Psi_j(y_j, x_j, \varphi_j) + \Psi'_j(y_j, x_j, \varphi_j) \times [\delta y_j, \delta x_j, \delta\varphi_j]^T.$$

Therefore, we solve

$$\Psi'_j(y_j, x_j, \varphi_j) \times [\delta y_j, \delta x_j, \delta\varphi_j]^T = -\Psi_j(y_j, x_j, \varphi_j)$$

3.2 A "two stage" iterative process

Residing in (3.8) is a linear problem. Thus, we can also consider a "two stage" iterative process, where at each step of the "Pseudo-Newton" method (for fixed j) we have to solve a linear problem to find $\bar{y}(t) = \tilde{y}^{[j+1]}(t)$, $\bar{x}(t) = \tilde{x}^{[j+1]}(t)$ and $\bar{\varphi}(t) = \tilde{\varphi}^{[j+1]}(t)$ in the form[1]

$$\frac{d\bar{y}(t)}{dt} - A^{[j]}(t)\bar{y}(t) - B^{[j]}(t)\bar{y}(t-\tau) = f\big(t, \tilde{y}^{[j]}(t), \tilde{y}^{[j]}(t-\tau)\big) + \\ -A^{[j]}(t)\tilde{y}^{[j]}(t) - B^{[j]}(t)\tilde{y}^{[j]}(t-\tau), \quad \text{for } t \in [0, T], \tag{3.9a}$$

$$\bar{y}(t) = \bar{\varphi}(t) \quad \text{for } t \in [-\tau, 0], \tag{3.9b}$$

$$-\frac{d\bar{x}(t)}{dt} - \bar{x}(t)A^{[j]}(t) - \bar{x}(t+\tau)B^{[j]}(t+\tau) = \tilde{x}^{[j]}(t)\mathcal{A}^{[j+1]}(t) + \\ \tilde{x}^{[j]}(t+\tau)\mathcal{B}^{[j+1]}(t+\tau) - \tilde{x}^{[j]}(t)\mathcal{A}^{[j]}(t) - \tilde{x}^{[j]}(t+\tau)\mathcal{B}^{[j]}(t+\tau) + [\bar{y}(t) - \widehat{y}(t)], \tag{3.9c}$$

$$\bar{x}(t) = 0, \quad \text{for } t \in [T, T+\tau], \tag{3.9d}$$

$$\alpha(\bar{\varphi}(t) - \widehat{\varphi}(t)) + B^{[j]}(t+\tau)\bar{x}(t+\tau) + \mathcal{B}^{[j+1]}(t+\tau)\tilde{x}^{[j]}(t+\tau) + \\ -\mathcal{B}^{[j]}(t+\tau)\tilde{x}^{[j]}(t+\tau) = 0 \quad \text{for } t \in [-\tau, 0), \tag{3.9e}$$

$$(\beta + \gamma)\bar{\varphi}(0) + \bar{x}(0) + \beta\widehat{\varphi}(0) + \gamma\widehat{y}(0) = 0. \tag{3.9f}$$

We shall use an iteration to determine $\bar{y}$, $\bar{x}$ and $\bar{\varphi}$ at the j-th step.

To solve (3.9), we can introduce an iteration index k and write

$$\frac{d}{dt}\Big\{\tilde{y}^{[j+1,k]}(t)\Big\} - A^{[j]}(t)\tilde{y}^{[j+1,k]}(t) - B_j(t)\tilde{y}^{[j+1,k]}(t-\tau) = f\big(t, \tilde{y}^{[j]}(t), \tilde{y}^{[j]}(t-\tau)\big) + \\ -A^{[j]}(t)\tilde{y}^{[j]}(t) - B^{[j]}(t)\tilde{y}^{[j]}(t-\tau), \quad \text{for } t \in [0, T], \tag{3.10a}$$

$$\tilde{y}^{[j+1,k]}(t) = \tilde{\varphi}^{[j+1,k]}(t) \quad \text{for } t \in [-\tau, 0], \tag{3.10b}$$

[1] A similar iterative process for ODEs have been considered, for example, in [3]

$$-\frac{d}{dt}\left\{\tilde{x}^{[j+1,k]}(t)\right\} - \tilde{x}^{[j+1,k]}(t)A^{[j]}(t) - \tilde{x}^{[j+1,k]}(t+\tau)B^{[j]}(t+\tau) = \tilde{x}^{[j]}(t)\mathcal{A}^{[j+1,k]}(t)+$$

$$\tilde{x}^{[j]}(t+\tau)\mathcal{B}^{[j+1,k]}(t+\tau) - \tilde{x}^{[j]}(t)\mathcal{A}^{[j]}(t) - \tilde{x}^{[j]}(t+\tau)\mathcal{B}^{[j]}(t+\tau) + [\tilde{y}^{[j+1,k]}(t) - \widehat{y}(t)], \quad (3.10c)$$

$$\tilde{x}^{[j+1,k]}(t) = 0, \quad \text{for } t \in [T, T+\tau], \quad (3.10d)$$

$$\tilde{\varphi}^{[j+1,k+1]}(t) = \tilde{\varphi}^{[j+1,k]}(t) + {}^1\Delta_{k+1}\Big(\alpha(\tilde{\varphi}^{[j+1,k]}(t) - \widehat{\varphi}(t)) + B^{[j]}(t+\tau)\tilde{x}^{[j+1,k]}(t+\tau)+$$

$$\mathcal{B}^{[j+1,k]}(t+\tau)\tilde{x}^{[j]}(t+\tau) - \mathcal{B}^{[j]}(t+\tau)\tilde{x}^{[j]}(t+\tau)\Big) + {}^2\Delta_{k+1}\big(\tilde{\varphi}^{[j+1,k]}(t) - \tilde{\varphi}^{[j+1,k-1]}(t)\big), \quad (3.10e)$$

$$\tilde{\varphi}^{[j+1,k+1]}(0) = \tilde{\varphi}^{[j+1,k]}(0) + \Delta'_{k+1}\Big((\beta+\gamma)\tilde{\varphi}^{[j+1,k]}(0) + \tilde{x}^{[j+1,k]}(0) + \beta\widehat{\varphi}(0) + \gamma\widehat{y}(0)\Big). \quad (3.10f)$$

We repeat the iterative process (3.10) until a convergence criterion is satisfied (for $k = 0, 1, \ldots, \mathcal{N}_j$) and then set

$$\tilde{y}^{[j+1]}(t) = \tilde{y}^{[j+1,\mathcal{N}_j]}(t), \quad \tilde{x}^{[j+1]}(t) = \tilde{x}^{[j+1,\mathcal{N}_j]}(t), \quad \text{and} \quad \tilde{\varphi}^{[j+1]}(t) = \tilde{\varphi}^{[j+1,\mathcal{N}_j]}(t).$$

Remark 3.1 *For different choices of* ${}^1\Delta_{k+1}$, ${}^2\Delta_{k+1}$ *and* Δ'_{k+1} *we obtain different iterative methods. For example, if* ${}^1\Delta_{k+1} = \Delta$ *and* $\Delta'_{k+1} = \Delta'$ *are constant, and* ${}^2\Delta_{k+1} = 0$ *we obtain the Picard iteration method. (It is not immediately obvious whether other standard iterations found in the literature can be obtained by simple choices of* ${}^1\Delta_{k+1}$, ${}^2\Delta_{k+1}$, Δ'_{k+1}.*)*

Remark 3.2 *According to [2] the system (3.9) can be written in the form (for appropriate* $\mathfrak{g}_\alpha^{\beta,\gamma}$ *and* $\mathfrak{L}_\alpha^{\beta,\gamma}$*)*

$$^{[j]}\mathfrak{L}_\alpha^{\beta,\gamma}\bar{\varphi}(t) = {}^{[j]}\mathfrak{g}_\alpha^{\beta,\gamma}(t), \quad \textit{for}, \ t \in [-\tau, 0). \quad (3.11)$$

We can apply various well-known iterative methods (see, for example, [5]) to equation (3.11). Our iterative process (3.10) associated with (3.11) gives us

$$\bar{\varphi}^{[k+1]}(t) = \bar{\varphi}^{[k]}(t) + {}^1\Delta_{k+1}\big({}^{[j]}\mathfrak{g}_\alpha^{\beta,\gamma}(t) - {}^{[j]}\mathfrak{L}_\alpha^{\beta,\gamma}\bar{\varphi}^{[k]}(t)\big) + {}^2\Delta_{k+1}\big(\bar{\varphi}^{[k]}(t) - \bar{\varphi}^{[k-1]}(t)\big). \quad (3.12)$$

3.3 A convergence result for the "Pseudo-Newton" method

In this section we shall state the convergence result for the "one stage" iterative process.

3.3.1 A system for the differences $y(t) - \tilde{y}^{[j]}(t)$ etc.

Let us consider the exact solutions of the system (2.12), namely the system for $y(t)$, $x(t)$ and $\varphi_\star(t)$. We have

$$\frac{dy(t)}{dt} = f(t, y(t), y(t-\tau)), \quad \text{for } t \in [0, T], \quad (3.13a)$$

$$y(t) = \varphi_\star(t) \quad \text{for } t \in [-\tau, 0), \quad y(0) = \varphi(0), \quad (3.13b)$$

$$-\frac{dx(t)}{dt} - x(t)A(t) - x(t+\tau)B(t+\tau) = [y(\varphi; t) - \widehat{y}(t)], \quad \text{for } t \in [0, T], \quad (3.13c)$$

$$x(t) = 0, \quad \text{for } t \in [T, T+\tau], \quad (3.13d)$$

$$\alpha(\varphi_\star(t) - \widehat{\varphi}(t)) + [B(t+\tau)]x(t+\tau) = 0 \quad \text{for } t \in [-\tau, 0), \quad (3.13e)$$

$$x(0) + \beta\{\varphi_\star(0) - \widehat{\varphi}(0)\} + \gamma\{\varphi_\star(0) - \widehat{y}(0)\} = 0. \quad (3.13f)$$

As an approximation to the exact solutions, let us consider the system arising after j iterations of the "one stage" process (3.6) (see page 10):

$$\frac{d\tilde{y}^{[j+1]}(t)}{dt} - A^{[j]}(t)\tilde{y}^{[j+1]}(t) - B^{[j]}(t)\tilde{y}^{[j+1]}(t-\tau) = f\big(t,\tilde{y}^{[j]}(t),\tilde{y}^{[j]}(t-\tau)\big)+$$

$$-A^{[j]}(t)\tilde{y}^{[j]}(t) - B^{[j]}(t)\tilde{y}^{[j]}(t-\tau), \quad \text{for } t\in[0,T], \tag{3.14a}$$

$$\tilde{y}^{[j+1]}(t) = \tilde{\varphi}^{[j+1]}(t) \quad \text{for } t\in[-\tau,0], \tag{3.14b}$$

$$-\frac{d\tilde{x}^{[j+1]}(t)}{dt} - \tilde{x}^{[j+1]}(t)A^{[j]}(t) - \tilde{x}^{[j+1]}(t+\tau)B^{[j]}(t+\tau) =$$

$$\tilde{x}^{[j]}(t)\Big\{\ \mathfrak{C}^{[j]}(t)\big(\tilde{y}^{[j+1]}(t)-\tilde{y}^{[j]}(t)\big) + \mathfrak{D}_1^{[j]}(t)\big(\tilde{y}^{[j+1]}(t-\tau)-\tilde{y}^{[j]}(t-\tau)\big)\Big\}+$$

$$\tilde{x}^{[j]}(t+\tau)\Big\{\mathfrak{D}_2^{[j]}(t+\tau)\big(\tilde{y}^{[j+1]}(t+\tau)-\tilde{y}^{[j]}(t+\tau)\big) + \mathfrak{C}^{[j]}(t+\tau)\big(\tilde{y}^{[j+1]}(t)-\tilde{y}^{[j]}(t)\big)\Big\}+$$

$$[\tilde{y}^{[j+1]}(t)-\widehat{y}(t)], \quad \text{for } t\in[0,T], \tag{3.14c}$$

$$\tilde{x}^{[j+1]}(t) = 0, \quad \text{for } t\in[T,T+\tau], \tag{3.14d}$$

$$\frac{\tilde{\varphi}^{[j+2]}(t)-\tilde{\varphi}^{[j+1]}(t)}{\Delta_j} = \alpha(\tilde{\varphi}^{[j+1]}(t)-\widehat{\varphi}(t)) + [B^{[j]}(t+\tau)]\tilde{x}^{[j+1]}(t+\tau)+$$

$$\Big\{\mathfrak{D}_2^{[j]}(t+\tau)\big(\tilde{y}^{[j+1]}(t+\tau)-\tilde{y}^{[j]}(t+\tau)\big) + \mathfrak{C}^{[j]}(t+\tau)\big(\tilde{\varphi}^{[j+1]}(t)-\tilde{\varphi}^{[j]}(t)\big)\Big\}\tilde{x}^{[j]}(t+\tau), \tag{3.14e}$$

$$\frac{\tilde{\varphi}^{[j+2]}(0)-\tilde{\varphi}^{[j+1]}(0)}{\Delta'_j} = (\beta+\gamma)\tilde{\varphi}^{[j+1]}(0) + \tilde{x}^{[j+1]}(0) - \beta\widehat{\varphi}(0) - \gamma\widehat{y}(0). \tag{3.14f}$$

For the differences between the iterated and the true solution we write

$$\varepsilon_y^{j+1}(t) = y(t) - \tilde{y}^{[j+1]}(t),\ \ \varepsilon_x^{j+1}(t) = x(t) - \tilde{x}^{[j+1]}(t) \ \text{ and } \ \varepsilon_\varphi^{j+1}(t) = \varphi_\star(t) - \tilde{\varphi}^{[j+1]}(t).$$

Differencing (3.13) and (3.14) (respectively (a),(b) and so on) and using Taylor's expansion for the functions involving the exact solutions, we can write the equations for the successive differences in the form

$$\frac{d\varepsilon_y^{j+1}(t)}{dt} - A^{[j]}(t)\varepsilon_y^{j+1}(t) - B^{[j]}(t)\varepsilon_y^{j+1}(t-\tau) = \Phi_1^j(t)\big[\varepsilon_y^j(t)\big]^2, \quad \text{for } t\in[0,T], \tag{3.15a}$$

$$\varepsilon_y^{j+1}(t) = \varepsilon_\varphi^{j+1}(t) \quad \text{for } t\in[-\tau,0], \tag{3.15b}$$

$$-\frac{d\varepsilon_x^{j+1}(t)}{dt} - A^{[j]}(t)\varepsilon_x^{j+1}(t) - B^{[j]}(t+\tau)\varepsilon_x^{j+1}(t+\tau) = \varepsilon_y^{j+1}(t)+$$

$$\Big[\varepsilon_x^j(t)\mathfrak{C}^{[j]}(t) + \varepsilon_x^j(t+\tau)\mathfrak{C}^{[j]}(t+\tau)\Big]\big(\varepsilon_y^j(t)-\varepsilon_y^{[j+1]}(t)\big) + \varepsilon_x^j(t)\mathfrak{D}_1^{[j]}(t)\big(\varepsilon_y^{[j]}(t-\tau)-\varepsilon_y^{[j+1]}(t-\tau)\big)+$$

$$\varepsilon_x^j(t+\tau)\mathfrak{D}_2^{[j]}(t+\tau)\big(\varepsilon_y^j(t+\tau)-\varepsilon_y^{[j+1]}(t+\tau)\big) + \Phi_2^j(t)[\varepsilon_y^j(t)]^2,\ \text{ for } t\in[0,T], \tag{3.15c}$$

$$\varepsilon_x^{j+1}(t) = 0, \quad \text{for } t\in[T,T+\tau], \tag{3.15d}$$

$$\frac{\varepsilon_\varphi^{j+1}(t)-\varepsilon_\varphi^{j+2}(t)}{\Delta_j} = \alpha\varepsilon_\varphi^{[j+1]}(t) + B^{[j]}(t+\tau)\varepsilon_x^{[j+1]}(t+\tau) + \varepsilon_x^j(t+\tau)\mathfrak{D}_2^{[j]}(t+\tau)\big(\varepsilon_y^j(t+\tau)-\varepsilon_y^{j+1}(t+\tau)\big)+$$

$$\varepsilon_x^j(t+\tau)\mathfrak{C}^{[j]}(t+\tau)\big(\varepsilon_\varphi^j(t)-\varepsilon_\varphi^{[j+1]}(t)\big) + \Phi_3^j(t)[\varepsilon_y^j(t)]^2,\ \text{ for } t\in[-\tau,0) \tag{3.15e}$$

$$\frac{\varepsilon_\varphi^{j+1}(0)-\varepsilon_\varphi^{j+2}(0)}{\Delta'_j} = (\beta+\gamma)\varepsilon_\varphi^{j+1}(0) + \varepsilon_x^{j+1}(0). \tag{3.15f}$$

Here the coefficients Φ_1^j, Φ_2^j and Φ_3^j involve higher-order derivatives of f and vary with j.

3.3.2 An integral equation for the $\varphi(t) - \tilde{\varphi}^{[j+1]}(t)$

Before we prove the main result we introduce some notation.

Definition 3.1 *Let us define*

$$\Psi^{[j]}(s,t) = Y(s,t) - Y(s,t)\Big[\varepsilon_x^j(s)\mathfrak{C}^{[j]}(s) + \varepsilon_x^j(s+\tau)\mathfrak{E}^{[j]}(s+\tau)\Big]$$
$$-Y(s+\tau,t)\varepsilon_x^j(s+\tau)\mathfrak{D}_1^{[j]}(s+\tau) - Y(s-\tau,t)\varepsilon_x^j(s)\mathfrak{D}_2^{[j]}(s).$$

Remark 3.3 *It is easy to see from the above expression, that* $\Psi^{[j]}(s,t) = Y(s,t) - \mathcal{O}(|\varepsilon_x^j|)$.

Lemma 3.1 *In the iteration (3.15),* $\varphi(t) - \tilde{\varphi}^{[j+1]}(t)$ *can be written for* $t \in [-\tau, 0)$, *in the form*

$$\varepsilon_\varphi^{j+2}(t) = \Big(1 - \delta_j\big(\alpha + \varepsilon_x^j(t+\tau)\mathfrak{E}^{[j]}(t+\tau)\big)\Big)\varepsilon_\varphi^{j+1}(t) +$$
$$-\delta_j \int_0^T B^{[j]}(t+\tau)\Psi^{[j]}(s,t+\tau)\Big[\big(\overset{\circ}{F}\varepsilon_\varphi^{j+1}\big)(s) - Y(s,0)\Upsilon(\varepsilon_\varphi^{j+1})\ \Omega^{[j]}\Big]ds +$$
$$-\delta_j \int_0^T B^{[j]}(t+\tau)\Psi^{[j]}(s,t+\tau)\ \Omega^{[j]} \int_{-\tau}^0 Y(s,0)Y(\mu+\tau,0)\varepsilon_x^j(\mu+\tau)\mathfrak{D}_1^{[j]}(\mu+\tau)\varepsilon_\varphi^{j+1}(\mu)d\mu ds +$$
$$-\delta_j\varepsilon_x^j(t+\tau)\mathfrak{D}_2^{[j]}(t+\tau)\Big[\big(\overset{\circ}{F}\varepsilon_\varphi^{j+1}\big)(t) - Y(t+\tau,0)\Upsilon(\varepsilon_\varphi^{j+1})\ \Omega^{[j]}\Big] +$$
$$\delta_j \int_t^0 B^{[j]}(t+\tau)Y(s+\tau,t+\tau)\varepsilon_x^j(s+\tau)\mathfrak{D}_1^{[j]}(s+\tau)\varepsilon_\varphi^{j+1}(s)ds + \mathcal{O}(\|\varepsilon_y^j(t)\|^2), \tag{3.16}$$

where $\Upsilon(\varepsilon_\varphi^{j+1}) = \int_0^T \Psi^{[j]}(\xi,0) \int_{-\tau}^0 Y(\xi,\mu+\tau)B^{[j]}(\mu+\tau)\varepsilon_\varphi^{j+1}(\mu)d\mu d\xi$ *and* $\Omega^{[j]} = \Big(\beta + \gamma + \int_0^T \Psi^{[j]}(\mu,0)Y(\mu,0)d\mu\Big)^{-1}$.

Proof. Equations (3.15a) and (3.15c) are linear equations for $\varepsilon_y^{j+1}(t)$ and $\varepsilon_x^{j+1}(t)$ respectively. According to [1, pp.7-8] we may write the solution of the equation (3.15a) in the form

$$\varepsilon_y^{j+1}(t) = Y(t,0)\varepsilon_\varphi^{j+1}(0) + \big(\overset{\circ}{F}\varepsilon_\varphi^{j+1}\big)(t) + \mathcal{O}(\|\varepsilon_y^j\|^2), \tag{3.17}$$

where $\big(\overset{\circ}{F}\varepsilon_\varphi^{j+1}\big)(t) = \int_{-\tau}^0 Y(t,s+\tau)B^{[j]}(s+\tau)\varepsilon_\varphi^{j+1}(s)ds$. The solution of (3.15c) can be written as

$$\varepsilon_x^{j+1}(t) = \int_t^T Y(s,t)\varepsilon_y^{j+1}(s)ds + \int_t^T Y(s,t)\Big[\varepsilon_x^j(s)\mathfrak{C}^{[j]}(s) + \varepsilon_x^j(s+\tau)\mathfrak{E}^{[j]}(s+\tau)\Big]\big(\varepsilon_y^j(s) - \varepsilon_y^{[j+1]}(s)\big)ds +$$
$$\underbrace{\int_t^T Y(s,t)\varepsilon_x^j(s)\mathfrak{D}_1^{[j]}(s)\big(\varepsilon_y^{[j]}(s-\tau) - \varepsilon_y^{[j+1]}(s-\tau)\big)ds}_{\mathfrak{I}_1} + \tag{3.18}$$

$$\underbrace{\int_t^T Y(s,t)\varepsilon_x^j(s+\tau)\mathfrak{D}_2^{[j]}(s+\tau)\big(\varepsilon_y^j(s+\tau)-\varepsilon_y^{[j+1]}(s+\tau)\big)\,ds}_{\mathfrak{J}_2} + \mathcal{O}(\|\varepsilon_y^j\|^2).$$

Let us change the variables in the terms $\mathfrak{J}_1$ and $\mathfrak{J}_2$. Since $\varepsilon_x^j(s) = 0$ when $s \in [T, T+\tau]$, and $Y(s,t) = 0$ when $s < t$, we can write, using the Definition 3.1, the solution of the equation (3.18) in the form

$$\varepsilon_x^{j+1}(t) = \int_t^T \Psi^{[j]}(s,t)\varepsilon_y^{j+1}(s)ds - \int_{t-\tau}^{t} Y(s+\tau,t)\varepsilon_x^j(s+\tau)\mathfrak{D}_1^{[j]}(s+\tau)\varepsilon_y^{[j+1]}(s) + \mathcal{O}(\|\varepsilon_y^j(t)\|^2). \quad (3.19)$$

Using (3.17) and (3.19) we have

$$\varepsilon_x^{j+1}(0) = \int_0^T \Psi^{[j]}(s,0)\Big[Y(s,0)\varepsilon_\varphi^{j+1}(0) + \big(\mathring{F}\varepsilon_\varphi^{j+1}\big)(s)\Big]ds +$$

$$- \int_{-\tau}^{0} Y(s+\tau,0)\varepsilon_x^j(s+\tau)\mathfrak{D}_1^{[j]}(s+\tau)\varepsilon_\varphi^{[j+1]}(s) + \mathcal{O}(\|\varepsilon_y^j(t)\|^2). \quad (3.20)$$

Then, for $\varepsilon_\varphi^{j+1}(0)$ we have $(\beta+\gamma)\varepsilon_\varphi^{j+1}(0) + \varepsilon_x^{j+1}(0) = 0$. We can therefore write

$$\Big[(\beta+\gamma) + \int_0^T \Psi^{[j]}(s,0)Y(s,0)ds\Big]\varepsilon_\varphi^{j+1}(0) = -\int_0^T \Psi^{[j]}(s,0)\big(\mathring{F}\varepsilon_\varphi^{j+1}\big)(s)ds +$$
$$\int_{-\tau}^{0} Y(s+\tau,0)\varepsilon_x^j(s+\tau)\mathfrak{D}_1^{[j]}(s+\tau)\varepsilon_\varphi^{[j+1]}(s) + \mathcal{O}(\|\varepsilon_y^j(t)\|^2). \quad (3.21)$$

Thus, we have

$$\varepsilon_\varphi^{j+1}(0) = -\Big(\int_0^T \Psi^{[j]}(s,0)\big(\mathring{F}\varepsilon_\varphi^{j+1}\big)(s)ds\Big)\Big/\Big((\beta+\gamma) + \int_0^T \Psi^{[j]}(s,0)Y(s,0)ds\Big) +$$
$$\Big(\int_{-\tau}^{0} Y(s+\tau,0)\varepsilon_x^j(s+\tau)\mathfrak{D}_1^{[j]}(s+\tau)\varepsilon_\varphi^{[j+1]}(s)\Big)\Big/\Big((\beta+\gamma) + \int_0^T \Psi^{[j]}(s,0)Y(s,0)ds\Big) + \mathcal{O}(\|\varepsilon_y^j(t)\|^2). \quad (3.22)$$

We also have following relation (see (3.15e))

$$\varepsilon_\varphi^{j+2}(t) = \Big(1 - \Delta_j\big(\alpha + \varepsilon_x^j(t+\tau)\mathfrak{C}^{[j]}(t+\tau)\big)\Big)\varepsilon_\varphi^{j+1}(t) - \Delta_j B^{[j]}(t+\tau)\varepsilon_x^{[j+1]}(t+\tau) +$$
$$-\Delta_j\varepsilon_x^j(t+\tau)\mathfrak{D}_2^{[j]}(t+\tau)\varepsilon_y^{j+1}(t+\tau) + \mathcal{O}(\|\varepsilon_y^j(t)\|^2). \quad (3.23)$$

By substituting (3.17) and (3.22) into (3.23) we obtain (3.16) and the Lemma is established.

Corollary 3.1 *According to Lemma 3.1 we can write the iteration (3.16) as*

$$\frac{\varepsilon_\varphi^{j+2}(t) - \varepsilon_\varphi^{j+1}(t)}{\delta_j} = \big(\alpha + \varepsilon_x^j(t+\tau)\mathfrak{C}^{[j]}(t+\tau)\big)\Big)\varepsilon_\varphi^{j+1}(t) - \Big\{\int_{-\tau}^{0} \widehat{K}^{\beta,\gamma}(t,s)\varepsilon_\varphi^{j+1}(s)ds +$$
$$\int_{-\tau}^{0} {}^1\Delta K^{\beta,\gamma}(t,\mu)\varepsilon_x^j(\mu+\tau)\varepsilon_\varphi^{j+1}(\mu)d\mu + \int_t^0 {}^2\Delta K^{0,0}(t,s)\varepsilon_x^j(s+\tau)\varepsilon_\varphi^{j+1}(s)ds\Big\}, \quad t \in [-\tau, 0) \quad (3.24)$$

Remark 3.4 *When $f(t, y(t), y(t-\tau))$ in (2.12a) is linear in $y(t)$ and $y(t-\tau)$, the quadratic terms (those involving $[\varepsilon_\psi^j]^2$, where $\psi \in \{y, x, \varphi\}$) vanish and the formula (3.24) becomes an iterative process via integral equation for "initial function differences", as in the linear case (see [2, pp. 13-14]).*

3.3.3 Convergence of the iterative method

We shall use induction to prove the convergence of the "one stage" Pseudo-Newton iterative method. We start with the formulation of two helpful Lemmas, which guide us later.

Lemma 3.2 *There exist ε_1^* and C_y such that, for all $\varepsilon \le \varepsilon_1^*$ when $\|\varepsilon_\varphi^{j+1}(t)\|_\infty < \varepsilon$ and $\|\varepsilon_y^j(t)\|_\infty \le C_y\varepsilon$, then $\|\varepsilon_y^{j+1}(t)\|_\infty \le C_y\varepsilon$.*

Proof. Using (3.17) there exist N_0 and N_1 such that $\|\varepsilon_y^{j+1}(t)\|_\infty \le N_0\|\varepsilon_\varphi^{j+1}(t)\|_\infty + N_1\|\varepsilon_y^j(t)\|_\infty^2$ and, therefore, $\|\varepsilon_y^{j+1}(t)\|_\infty \le N_0\varepsilon + N_1(C_y\varepsilon)^2$. Let $C_y = 2N_0$. Then $\|\varepsilon_y^{j+1}(t)\|_\infty = \frac{C_y}{2}\varepsilon + N_1(C_y\varepsilon)^2 = \frac{C_y}{2}\big(1 + 2N_1C_y\varepsilon\big)\varepsilon$. Thus we need $2N_1C_y\varepsilon < 1$ and therefore $\varepsilon \le \varepsilon_1^* = 1\big/4N_1N_0$. Lemma 3.2 is established.

Lemma 3.3 *Under the condition of Lemma 3.2 there exist ε_2^* and C_x such that, for all $\varepsilon \le \varepsilon_2^*$ when $\|\varepsilon_x^j(t)\|_\infty \le C_x\varepsilon$, then $\|\varepsilon_x^{j+1}(t)\|_\infty \le C_x\varepsilon$.*

Proof. Using the same approach as in Lemma 3.2 we can write, from (3.19),

$$\|\varepsilon_x^{j+1}(t)\|_\infty \le M_0\|\varepsilon_y^{j+1}(t)\|_\infty + M_1\|\varepsilon_x^j(t)\|_\infty\|\varepsilon_y^{j+1}(t)\|_\infty + M_2\|\varepsilon_y^j\|_\infty^2$$

and, further

$$\|\varepsilon_x^{j+1}(t)\|_\infty \le M_0C_y\varepsilon + M_1C_xC_y\varepsilon^2 + M_2(C_y\varepsilon)^2 = \frac{C_x}{2}\big(1 + (2M_1 + M_2/M_0)C_y\varepsilon\big)\varepsilon,$$

where we set $C_x = 2M_0C_y$. We need $(2M_1 + M_2/M_0)C_y\varepsilon < 1$. Therefore $\varepsilon \le \varepsilon_2^* = M_0\big/2(M_0M_1 + N_0M_2)$ and the Lemma is established.

Using the results of Lemma 3.2 and Lemma 3.3, we have the following theorem.

Theorem 3.1 *Sufficient conditions for the iteration (3.6) to converge (when the initial approximation are sufficiently close) are*

$$0 < \Delta_j < \frac{2}{\alpha + K_1}, \quad \text{and } 0 < \Delta_j' < \frac{2}{\beta + \gamma + \widehat{K}}, \quad \text{uniformly for all } j, \tag{3.25}$$

where K_1 and $\widehat{K}$ are appropriate finite positive constants.

The values K_1 and $\widehat{K}$ are given in our proof below.

Proof. Suppose that $\sup\limits_{k \le j+1} \|\varepsilon_\varphi^k(t)\|_\infty < \varepsilon^*$, where $\varepsilon^* = \min(\varepsilon_1^*, \varepsilon_2^*)$. Define $e_j := \sup\limits_{\varphi} \max\limits_{k<j+1} \|\varepsilon_\varphi^k(t)\|_\infty$. Take as induction hypothesis the assumption $e_j < \varepsilon^*$, where we assume this true for $j = 1$. Then, for sufficiently small ε^*, from (3.16), we can write

$$\sup_{t\in[-\tau,0)} |\varepsilon_\varphi^{j+2}(t)| \le |1 - \Delta_j(\alpha + K_1)|\varepsilon^* + \mathcal{O}([\varepsilon^*]^2) \tag{3.26}$$

where $K_1 = \sup_t \int_{-\tau}^{0} |\mathcal{M}(s,t)| ds$ and $\mathcal{M}(s,t) = \int_0^T B^{[j]}(t+\tau)Y(\xi,t+\tau)Y(\xi,s+\tau)B^{[j]}(s+\tau)d\xi +$

$$-\int_0^T \int_0^T B^{[j]}(t+\tau)Y(\xi,t+\tau)\Omega Y(\xi,0)Y(\mu,0)Y(\mu,s+\tau)B(s+\tau)d\mu d\xi$$

Thus a condition for this iteration to converge is $|1-\Delta_j(\alpha+K_1)| < 1$, uniformly for all j and we have the condition for Δ_j:

$$0 < \Delta_j < \frac{2}{\alpha + K_1}, \quad \text{uniformly for all } j. \tag{3.27a}$$

Using (3.21) and (3.15f) we can write

$$|\varepsilon_\varphi^{j+2}(0)| \le \left|1 - \Delta'_j(\beta+\gamma+K_2)\right| \varepsilon^* + \Delta'_j K_3 \|\varepsilon_\varphi^{j+1}(t)\|_\infty + \mathcal{O}([\varepsilon^*]^2),$$

where $K_2 = \left|\int_0^T Y(s,0)Y(s,0)ds\right|$ and $K_3 = \sup_t \int_0^T \left|Y(s,0)Y(t,s+\tau)B^{[j]}(s+\tau)\right| ds$. We therefore seek $\left|1-\Delta'_j(\beta+\gamma+K_2+K_3)\right| < 1$, uniformly for all j. Under the condition (3.27a), and provided also that

$$0 < \Delta'_j < \frac{2}{\beta+\gamma+K_2+K_3}, \quad \text{uniformly for all } j, \tag{3.27b}$$

$\|\varepsilon_\varphi^{j+2}(t)\|_\infty \le \varrho\varepsilon^*$ for some $\varrho \in (0,1)$. Hence $\|\varepsilon_\varphi^n(t)\|_\infty \to 0$ as $n \to \infty$. Thus, Theorem 3.1 is established.

Remark 3.5 *According to Remark 3.3, if ε^* is sufficiently small $\Psi^{[j]}(s,0) \approx Y(s,0)$ and $\beta+\gamma+\int_0^T \Psi^{[j]}(s,0)Y(s,0)ds > 0$, therefore, $\Omega^{[j]}$ and K_1 are finite.*

4 Computational results

In this section we shall present results of numerical experiments of identification of the initial function $\varphi_\star(t)$ for nonlinear DDE.

4.1 Posing the problem

Let us consider the following nonlinear delay differential equation

$$\frac{dy(t)}{dt} = \lambda y(t)\big(y(t-\tau)-2\big), \quad t \in [0,T] \tag{4.1a}$$

with an initial condition

$$y(t) = \varphi(t), \quad t \in [-\tau, 0]. \tag{4.1b}$$

Here λ is a real number, a parameter of the equation. This problem corresponds to

$$f\big(t, y(t), y(t-\tau)\big) = \lambda y(t)\big(y(t-\tau)-2\big)$$

in (1.1).

Remark 4.1 *According to Definition 2.1 and Remark 2.1 we have:*

$$A(t) = \lambda\Big(y(t-\tau)-2\Big);\;\; B(t) = \lambda y(t);\quad \mathfrak{C}(t) = 0;\;\; \mathfrak{E}(t) = 0;\;\; \mathfrak{D}_i(t) = \lambda,\; i = 1,2.$$

The identification problem of finding the optimal initial function for nonlinear DDE (4.1) has, according to §2.4, the form

$$\frac{dy(t)}{dt} = \lambda y(t)\big(y(t-\tau)-2\big), \quad \text{for } t \in [0,T], \tag{4.2a}$$

$$y(t) = \varphi(t) \quad \text{for } t \in [-\tau, 0), \quad y(0) = \varphi(0), \tag{4.2b}$$

$$-\frac{dx(t)}{dt} - \lambda x(t)\big(y(t-\tau)-2\big) - \lambda x(t+\tau)y(t+\tau) = y(\varphi;t) - \widehat{y}(t), \quad \text{for } t \in [0,T], \tag{4.2c}$$

$$x(t) = 0, \quad \text{for } t \in [T, T+\tau], \tag{4.2d}$$

$$\alpha(\varphi(t) - \widehat{\varphi}(t)) + \lambda y(t+\tau)x(t+\tau) = 0 \quad \text{for } t \in [-\tau, 0), \tag{4.2e}$$

$$x(0) + \beta\{\varphi(0) - \widehat{\varphi}(0)\} + \gamma\{\varphi(0) - \widehat{y}(0)\} = 0. \tag{4.2f}$$

4.2 Numerical method for determine initial function

We apply to the system (4.2) the iterative methods described in section 3. Thus, using the "one-step" method we obtain the following iterative process

$$\frac{d\tilde{y}^{[j+1]}(t)}{dt} - \lambda\big(\tilde{y}^{[j]}(t-\tau)-2\big)\tilde{y}^{[j+1]}(t) - \lambda\tilde{y}^{[j]}(t)\tilde{y}^{[j+1]}(t-\tau) = -\lambda\tilde{y}^{[j]}(t)\tilde{y}^{[j]}(t-\tau),\; t \in [0,T], \tag{4.3a}$$

$$\tilde{y}^{[j+1]}(t) = \tilde{\varphi}^{[j+1]}(t), \;\; t \in [-\tau, 0], \tag{4.3b}$$

$$\begin{aligned} -\frac{d\tilde{x}^{[j+1]}(t)}{dt} - \lambda\tilde{x}^{[j+1]}(t)\big(\tilde{y}^{[j]}(t-\tau)-2\big) - \lambda\tilde{x}^{[j+1]}(t+\tau)\tilde{y}^{[j]}(t+\tau) = \\ \lambda\tilde{x}^{[j]}(t)\Big\{\tilde{y}^{[j+1]}(t-\tau) - \tilde{y}^{[j]}(t-\tau)\Big\} + \lambda\tilde{x}^{[j]}(t+\tau)\Big\{\tilde{y}^{[j+1]}(t+\tau) - \tilde{y}^{[j]}(t+\tau)\Big\} + \\ (\tilde{y}^{[j+1]}(t) - \widehat{y}(t)), \;\; t \in [0,T], \end{aligned} \tag{4.4a}$$

$$\tilde{x}^{[j+1]}(t) = 0, \;\; t \in [T, T+\tau], \tag{4.4b}$$

$$\begin{aligned} \frac{\tilde{\varphi}^{[j+2]}(t) - \tilde{\varphi}^{[j+1]}(t)}{\Delta_j} = \alpha(\tilde{\varphi}^{[j+1]}(t) - \widehat{\varphi}(t)) + \lambda[\tilde{y}^{[j]}(t+\tau)]\tilde{x}^{[j+1]}(t+\tau) + \\ \lambda\Big\{\tilde{y}^{[j+1]}(t+\tau) - \tilde{y}^{[j]}(t+\tau)))\Big\}\tilde{x}^{[j]}(t+\tau), \;\; t \in [-\tau, 0), \end{aligned} \tag{4.5a}$$

$$\frac{\tilde{\varphi}^{[j+2]}(0) - \tilde{\varphi}^{[j+1]}(0)}{\Delta_j'} = (\beta+\gamma)\tilde{\varphi}^{[j+1]}(0) + \tilde{x}^{[j+1]}(0) + \beta\widehat{\varphi}(0) + \gamma\widehat{y}(0). \tag{4.5b}$$

To solve the system (4.3a) and (4.4) numerically we can use an Euler-type discretization. For example, on the uniform grid with step h (here $\tau = hN$ and $T = hK$) the discrete equation (4.3a) has the form

$$\frac{\tilde{y}^{[j+1]}_{n+1} - \tilde{y}^{[j+1]}_n}{h} - \lambda\big(\tilde{y}^{[j]}_{n-N} - 2\big)\tilde{y}^{[j+1]}_n - \lambda\tilde{y}^{[j]}_n\tilde{y}^{[j+1]}_{n-N} = -\lambda\tilde{y}^{[j]}_n\tilde{y}^{[j]}_{n-N}, \quad \text{for } n = 0,1,\ldots,K-1, \tag{4.6a}$$

$$\tilde{y}^{[j+1]}_n = \tilde{\varphi}^{[j+1]}_n \quad \text{for } n = -N,\ldots,-1,0. \tag{4.6b}$$

(The remaining equations are obvious.)

4.3 Numerical results

4.3.1 A "one stage" iterative process

Here we present some numerical experiments using the "one stage" "Pseudo-Newton" method. In all experiments presented in this section we choose time-lag $\tau = 1$ and an "observation data" $\widehat{y}(t)$ was given on a uniform mesh.We start with the simplest experiment to illustrate the behaviour of the algorithm.

Experiment 1

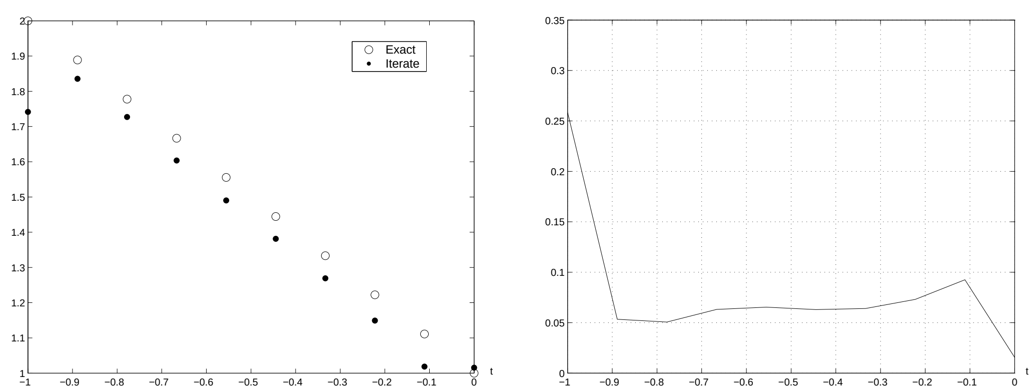
In the first set of experiments we consider system (4.3a) with the "exact" initial function in the form $\varphi_\star(t) = 1 - t$ and we start with initial guess $\varphi^{[0]} = 0$. The number of points in the initial interval is $N = 10$, parameters are $\beta = 0$, $\gamma = 1$, $\lambda = 1$ and the termination condition is $\varepsilon := \|\varphi^{[j+1]} - \varphi^{[j]}\| \Big/ \|\varphi^{[j]}\| \leq 10^{-6}$. A summary of the first experiments, for different values of the parameter α, is given in Table 1. (Timings are broadly indicative only; in a time-sharing

α	1	0.2	0.1	0.05	0.001	0.0005
The number of iterations	93	397	820	1657	10551	13863
The relative error $\mathcal{R}$	0.9712	0.9399	0.8996	0.8154	0.0661	0.0492
cpu time min:sec	0:0.28	0:1.49	0:3.84	0:5.82	0:45.88	1:23.57
The value of the functional $S_\alpha^{\beta,\gamma}(\varphi)$	4.0e-02	3.6e-02	3.4e-02	2.9e-02	1.9e-04	8.1e-05

Table 1: A summary of the experiment 1 for the nonlinear DDE

environment, they depend on a number of factors.)

In Figure 1 we present the results of our experiment with parameter $\alpha = 0.001$.



(a) The exact function $\varphi_\star(t)$ and the iterated function $\varphi^{[\mathcal{N}]}(t)$, $\mathcal{N} = 12697$

(b) The error curve $|\varphi_\star(t) - \varphi^{[\mathcal{N}]}(t)|$

Figure 1: An experiment with $\alpha = 0.0001$

As we see from Figure 1, for a nonlinear equation we have the similar problem near the boundaries as in the linear case. In the nonlinear case, the regularization parameter α plays a similar rôle as in the linear one: for $\alpha = 0$ we effectively solve a nonlinear integral equation of the first kind.

Experiment 2

An experiment with another initial function shown in Figure 2. Here the "exact" initial function is $\varphi_\star(t) = 1 - \sin(\pi t)/2$, the initial guess is $\varphi^{[0]}(t) = 0$ and the number of points in the initial interval $N = 10$. The parameters in this experiment are $\alpha = 0.0001$, $\beta = 0$, $\gamma = 1$ and $\lambda = 1$. The termination condition is $\varepsilon := \|\varphi^{[j+1]} - \varphi^{[j]}\| \Big/ \|\varphi^{[j]}\| \leq 10^{-6}$.

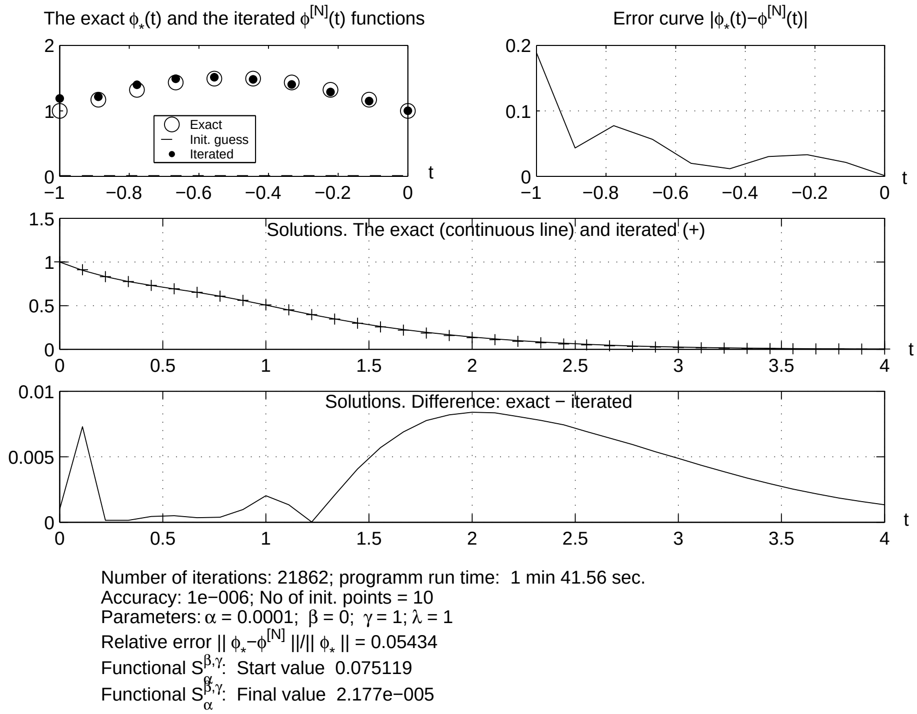


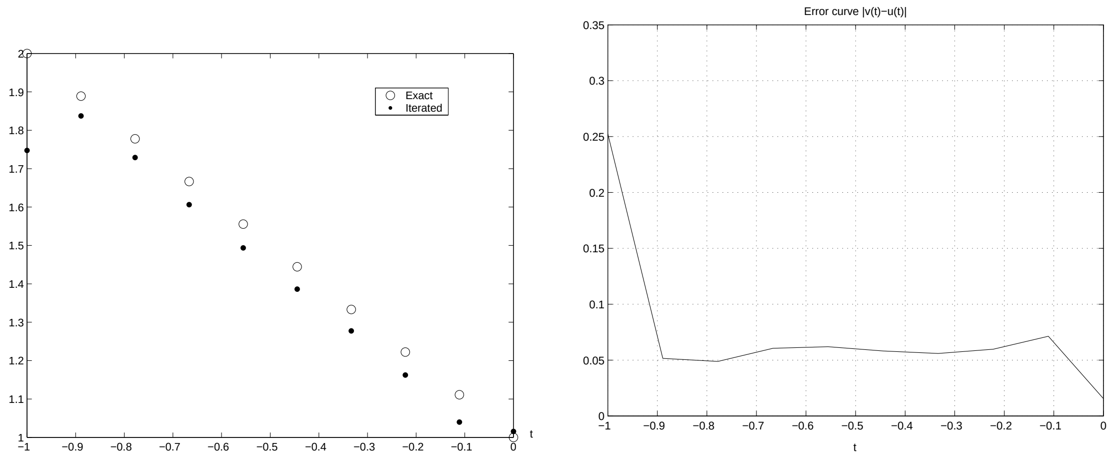
Figure 2: A summary of the experiment with $\varphi_\star(t) = 1 - \sin(\pi t)/2$

As we see from this experiment the behaviour of the iterative process for nonlinear DDEs is similar to that in linear case (see [1, 16-28]). In both the linear and nonlinear cases, the increasing parameter α speeds up the convergence of the iterative process, but at the same time the relative error increases.

As a conclusion we can state that the iterative process described in §3 converges and the solution obtained by this method provides a near-minimum of the functional $S_\alpha^{\beta,\gamma}$. As we can see in Figure 2 the solution of the problem (3.8) is sufficiently close to the solution of the nonlinear problem (2.12a), which we used as a "observation" data.

Experiment 3

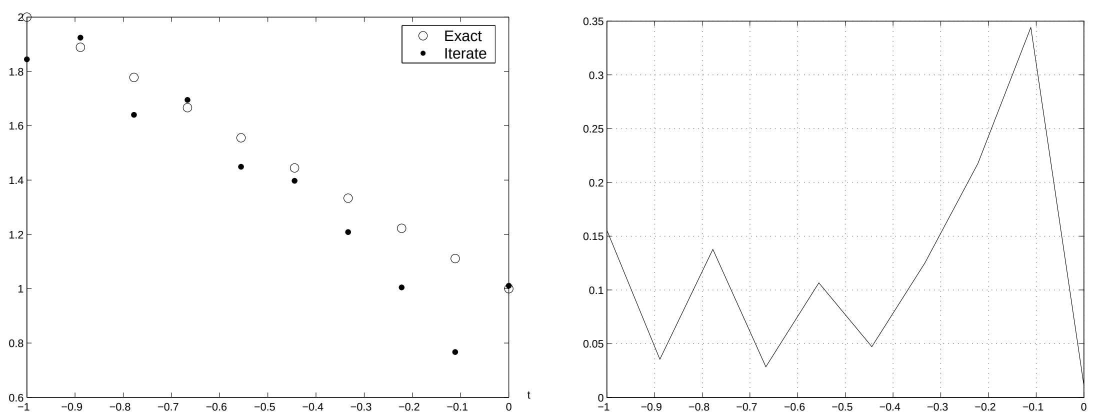
In the next series of experiments we investigate the dependence of the algorithm on the parameter λ. One might expect difficulties to arise for certain values of λ and this proves to be the case. The "true" initial function is $\varphi_\star(t) = 1 - t$. The initial guess is $\varphi^{[0]} = 0$. Parameters are $\alpha = 0.001$, $\beta = 0$, $\gamma = 1$. The number of point in the initial interval is $N = 10$ and the termination condition is $\varepsilon := \|\varphi^{[j+1]} - \varphi^{[j]}\| \Big/ \|\varphi^{[j]}\| \leq 10^{-6}$. In Figures 3 and 4 we present the exact and iterated function (Figure (a)) and the difference between exact and iterated functions (Figure (b)) with parameter $\lambda = 1$ and $\lambda = 3$ respectively. A summary of this experiment is shown in Table 2.



(a) The exact function $\varphi_\star(t)$ and the iterated function $\varphi^{[\mathcal{N}]}(t)$, $\mathcal{N} = 10563$

(b) The error curve $|\varphi_\star(t) - \varphi^{[\mathcal{N}]}(t)|$

Figure 3: An experiment with the parameter $\lambda = 1$



(a) The exact function $\varphi_\star(t)$ and the iterated function $\varphi^{[\mathcal{N}]}(t)$, $\mathcal{N} = 42015$

(b) The error curve $|\varphi_\star(t) - \varphi^{[\mathcal{N}]}(t)|$

Figure 4: An experiment with the parameter $\lambda = 3$

When we the increase the parameter λ in equation (4.1), the iterative process slows down and the error (the difference between the "exact" function and iterated function) increases. Increasing λ changes the behaviour of the solution of the nonlinear problem. Here we mention some difficulties which arise, when we solve the identification problem for a nonlinear equation with increasing λ.

	Number of iterations	cpu time min:sec	The relative error $\mathcal{R}$	the value of functional $\times 10^{-4}$
$\lambda = 1$	10563	0:27.08	0.0661	1.89
$\lambda = 3$	42015	2:43.09	0.1749	3.71

Table 2: A summary of the experiment with different λ

In our approach we used a linearized system to solve a nonlinear equation. Asymptotic stability of the linear equation is assumed when $|B(t)| < -A(t)$. Unfortunately, the coefficient in the linearized problem, which depend on solution of the original problem, changes not only throughout time, but also in each Newton iteration, and this affects both the original and adjoint equations. For small α, the number of iterations may be in the thousands and if, in some stage, one of the linearized equations is highly unstable, this may cause divergence of the process.

The other problem, which may cause the divergence of the algorithm, is the qualitative behaviour of the solution of the original problem. It is known that for some λ chaotic behaviour in the solution of the nonlinear equation can arise (for example the solution of the nonlinear equation $y'(t) = \lambda y(t)(1 - y(t-1))$ for $\lambda = 3$ has narrow peaks). Since, in our algorithm, we use the solution of the original equation as a forcing term in the adjoint equation, such chaotic behaviour of the original equation leads to an exceptionally large solution of the adjoint equation.

Experiment 4

Even when the algorithms converge we may have a problem with the identification of the initial function. Let us consider the following example. Suppose that we solve the identification problem for the nonlinear equation (4.1) with the "true" initial function $\varphi(t) > 0$ on $[-\tau, 0]$. Then, using known properties of the solution of the nonlinear equation (4.1) we can state the following results. Given $\delta > 0$ and $\varepsilon > 0$ there exist $\lambda(\varepsilon.\delta)$ such that for $\lambda > \lambda(\varepsilon.\delta)$ we have $|y(t)| < \varepsilon$ for $t \in [\delta, \tau]$.

Therefore we have $|\widehat{y}(t)| < \varepsilon$ (as a solution of the nonlinear problem (4.1)) and $|y(\lambda, t)| < \varepsilon$ (as a solution of the linearized equation (4.3a)). Thus this implies $|\widehat{y}(t) - y| < \varepsilon$ on $[\delta, \tau]$. Since the initial condition for the adjoint equation is zero, for the solution of the adjoint equation we have $|x(t)| < \varepsilon \mathcal{M}$ on $[\delta, \tau]$, where $\mathcal{M}$ is a positive finite constant. Therefore, from equation (4.5), we obtain $|\varphi^{[j]}| \leq \varepsilon^2 \mathcal{M}$ on $[\delta - \tau, 0]$[2]. Thus, for sufficiently large λ, we can recover the initial function only in the interval $[-\tau, \delta - \tau]$ (for very large λ we may have $\delta \ll \tau$). One of the experiments for nonlinear equation (4.1) with "true" initial function $\varphi_*(t) = 1 - t$ is shown in Figure 5.

[2]For the zero initial guess the result is immediate.

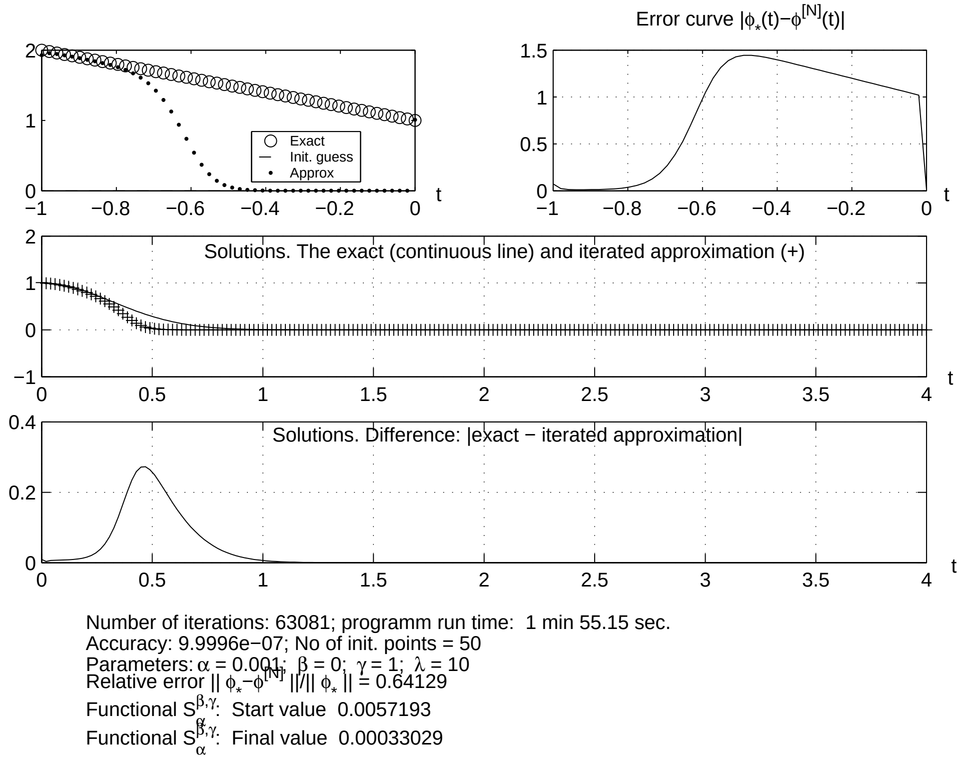


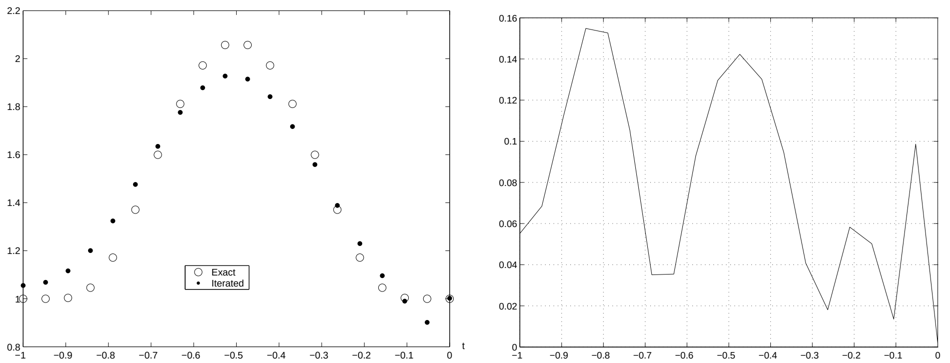
Figure 5: A summary of the experiment with $\varphi_\star(t) = 1 - \sin(\pi t)/2$

Experiment 5

In the next experiments we examine the behaviour of the algorithm when the length of the interval T is increased. We may address here the following question: how increasing T affected the solution of the identification problem for some particular "true" initial function. To answer this question we compare results that we obtain for different values of m (here $m\tau = T$, where τ is time-lag). We carry out all the experiments with parameters: $\alpha = 0.0001$, $\beta = 0$, $\gamma = 1$, $\lambda = 1$. The number of points in the initial interval is $N = 20$. The "exact" initial function is $\varphi_\star(t) = 1 + 50\exp(1/(t(1+t))$ and the initial guess is $\varphi^{[0]} = 0$. Results of this experiments are shown in Figures 6, 7 and summary information is presented in Table 3.

	Number of iterations	cpu time min:sec	The relative error $\mathcal{R}$	the value of functional $\times 10^{-5}$
$m = 2$	33235	3:31.03	0.0633	1.098
$m = 4$	33372	4:57.42	0.0630	1.154
$m = 8$	33372	8:50.96	0.0630	1.155

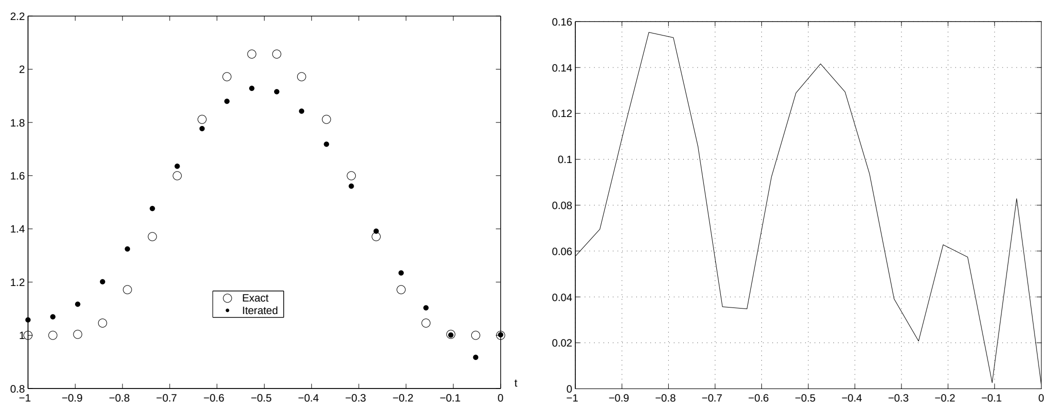
Table 3: A summary of the experiment for different m



(a) The exact function $\varphi_\star(t)$ and the iterate function $\varphi^{[\mathcal{N}]}(t)$, $\mathcal{N} = 33235$

(b) The error curve $|\varphi_\star(t) - \varphi^{[\mathcal{N}]}(t)|$

Figure 6: An experiment for $m = 2$



(a) The exact function $\varphi_\star(t)$ and the iterated function $\varphi^{[\mathcal{N}]}(t)$, $\mathcal{N} = 33372$

(b) The error curve $|\varphi_\star(t) - \varphi^{[\mathcal{N}]}(t)|$

Figure 7: An experiment for $m = 4$

By analyzing Table 3 we surmise that the iterated solutions are similar and the convergence and accuracy of the method is apparently independent of m.

Experiment 6

It is known that equation (4.1) has two steady states $y(t) \equiv 2$ and $y(t) \equiv 0$. The null solution $y(t) \equiv 0$ is an attractor for other solution if $\lambda > 0$, so $y(t) \to 0$ as $t \to \infty$. Therefore the following problem can be investigated. Suppose that we have two solutions of the equation (4.1) $y(\lambda, \varphi_1)$ and $y(\lambda, \varphi_2)$ on $[0, T]$, where $\varphi_1(t) > 0$ and $\varphi_2(t) > 0$ when $t \in [-\tau, 0]$. Then, for sufficiently large λ, the solutions $y(\lambda, \varphi_1)$ and $y(\lambda, \varphi_2)$ will be practically identical almost everywhere, except on an interval $[0, \delta]$, where $\delta < T$. Therefore, the "observation data", which is generated by solving the problem (4.1) with the initial functions $\varphi_1(t)$ and $\varphi_2(t)$, will be also identical, except on the interval $[0, \delta]$. Thus we can address here the following question: does our method distinguish the function $\varphi_1(t)$ from $\varphi_2(t)$.

In order to answer this question, let us consider the same experiments with two "exact" initial functions $\varphi^1_\star(t) = 1 + 50\exp(1/(t(1+t))$ and $\varphi^2_\star(t) = 1 - 50\exp(1/(t(1+t))$. The results of this experiment are shown on Figures 8 and 9. If we compare this Figures we can conclude that the difference between the solutions in $[0, 3\tau]$ is enough to find the correct initial function.

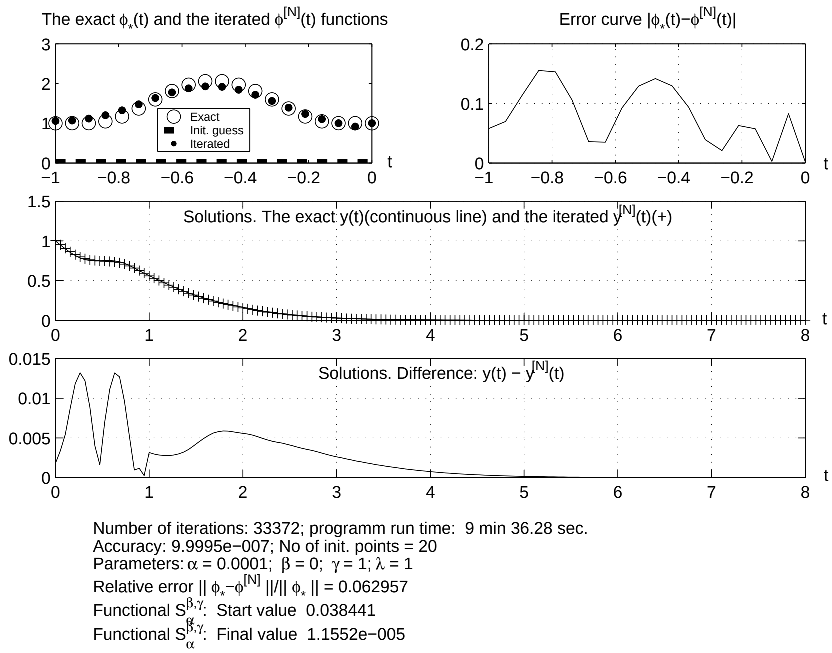


Figure 8: A summary of the experiment for $m = 8$ with $\varphi_\star(t) = 1 + 50\exp(1/(t(1+t))$

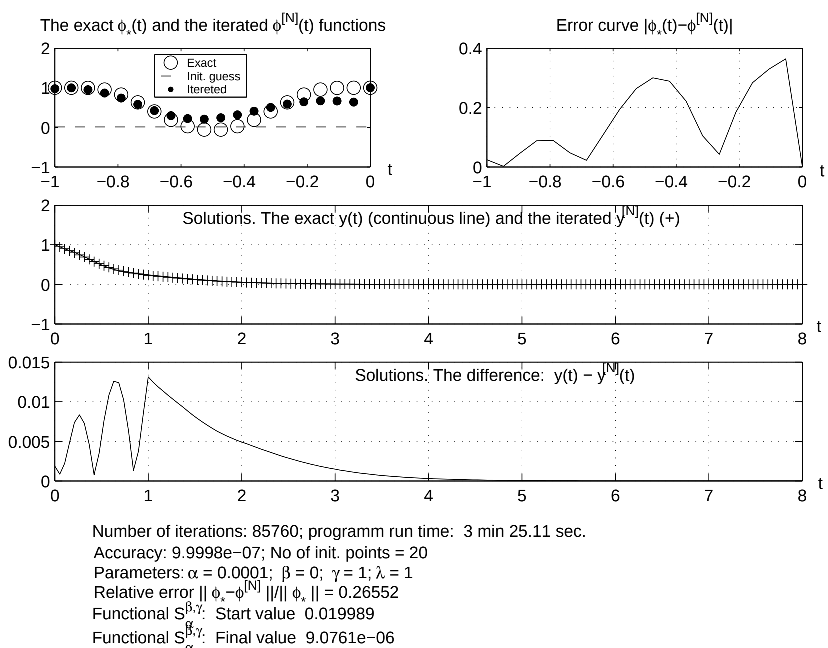


Figure 9: A summary of the experiment for $m = 8$ with $\varphi_\star(t) = 1 - 50\exp(1/(t(1+t))$

4.3.2 A "two stage" iterative process

In this section we illustrate briefly the behaviour of the "two stage" iterative method, described in §3. We consider a numerical experiment with the "exact" initial function

$$\varphi_\star(t) = 1 + 50\exp\left(1/(t(1+t)\right).$$

The initial guess is $\varphi^{[0]} = 0$ and $\alpha = 0.0001$, $\beta = 0$, $\gamma = 1$, $\lambda = 1$. The number of points on the initial interval is $N = 20$ and $m = 8$ ($T = 8\tau$). In Figure 10 we present the result of this experiment.

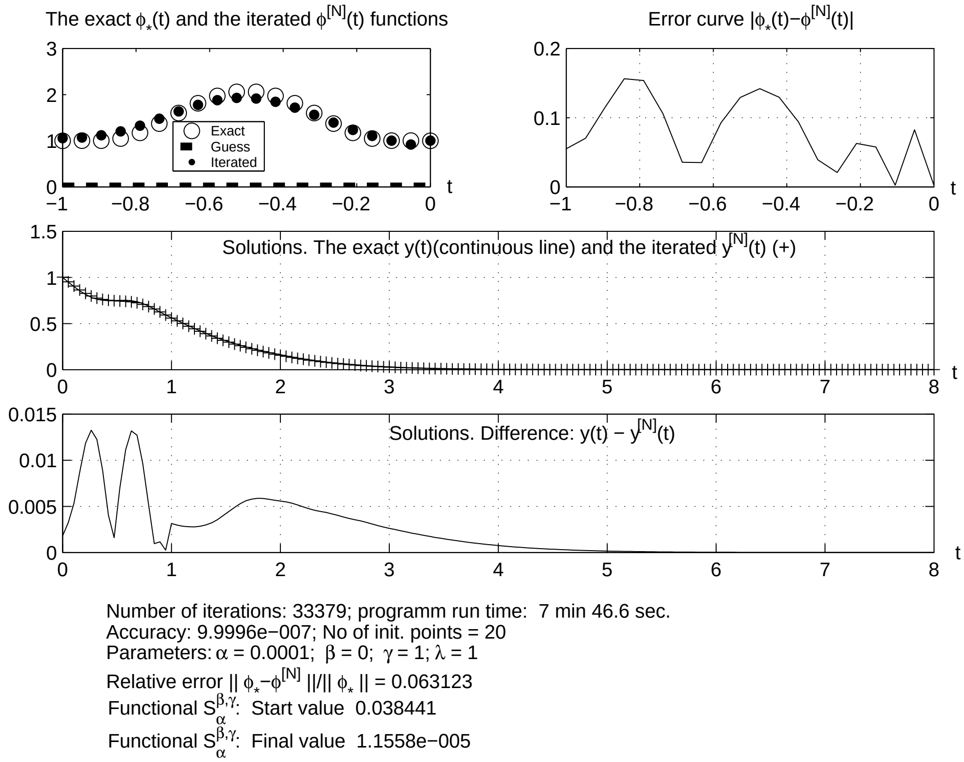


Figure 10: A "two stage" iterative process. A summary.

Since the inner process in the "two stage" iterative process is a "Picard-type" iteration, the behaviour of the "one stage" iterative process and "two stage" iterative process are similar. But the advantage of using the "two stage" process lies in the fact that we can control the accuracy of the iterated solution in each stage of the Newton method and this leads us to a more accurate solution of the Newton method.

4.4 Concluding remarks

In this section we presented numerical experiments for solving the identification problem (IP) in order to find the initial function for a given nonlinear DDE. Here we summarize the main experimental results.

The method proposed can usually be used to solve the IP of finding the initial function for various kinds of nonlinear equations, which formally can be written $y'(t) = f(t, y(t), y(t-\tau))$, when $t > 0$, subject to some initial condition $y(t) = \varphi(t)$, when $t \in [-\tau, 0]$. However, we have seen that cases can arise where the method fails in practice.

The rôle of the regularization parameter in the IP for nonlinear DDE is similar to its rôle in the linear case. The number of iteration increases, but at the same time the relative error (of the "true" and iterated functions) decreases when $\alpha \to 0$.

To solve the IP for nonlinear equation we linearize our equation and write down the linearized system of equations. The coefficients, both in the original and the adjoint equation are now dependent on the solution of the problem and change not only throughout the interval, but also in each iteration of the Newton method. For small α, the number of iterations may be in the thousands and if, in some stage, one of the linearized equations is highly unstable, this may cause divergence of the process.

We defined two possible iterative method to solve the IP for nonlinear DDEs: a "one stage" method and a "two stage" method. In our numerical experiments we found no difference between the two methods. This is probably because in both methods we used the same iterative method (the Picard iteration) to compute $\varphi^{[j]}$.

References

[1] Baker C.T.H., Parmuzin E.I., A guided tour of VPF for continuous and discretized DDEs, NA Report 430, MCCM, Manchester, England, January 2004.

[2] Baker C.T.H., Parmuzin E.I, Identification of the initial function for delay differential equation: Part I: The continuous problem & an integral equation analysis, NA Report 431, MCCM, Manchester, England, January 2004.

[3] Le Dimet, F.-X. and Shutyaev, V.P, On Newton methods in data assimilation, Russ. J. Numer. Anal. Math. Modelling, **15**, No. 5,(2000), pp. 419–434.

[4] Liusternik L.A. and Sobolev V.J., Elements of functional analysis, Frederick Ungar Publishing Company, New York, 1961.

[5] Marchuk G.I. and Lebedev V.I., Numerical Methods in the Theory of Neutron Transport, Harwood Academic Publishers, New York, 1986.

[6] Ortega J. M. and Rheinboldt W. C., Iterative solution of nonlinear equations in several variables, Academic Press, New York, 1970.