ISSN 1360-1725

THE UNIVERSITY of MANCHESTER

UMIST

Identification of the initial function for delay differential equations: Part I: The continuous problem & an integral equation analysis.

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. 431

Completed Aug 2003: Released: Jan 2004

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

Identification of the initial function for delay differential equations:
Part I: The continuous problem & an integral equation analysis.

Contents

Identification of the initial function for delay differential equations:
Part I: The continuous problem & an integral equation analysis.

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

narep431.tex

Abstract

In this report, which is the first of two related reports, we consider the "data assimilation problem" for delay differential equations. This problem consists of finding an initial function that gives rise to a solution which is a close fit to observed data. A rôle for adjoint equations and fundamental solutions is established, and related integral equations are obtained. In the second report a discrete analogue of the results obtained in the first part will be presented.

Keywords: Delay differential equations, initial function, adjoint equations, identification problem, data assimilation, fundamental matrices, regularization parameter.

*`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.

1 The nature of the problem

1.1 Modelling issues

Studies have been undertaken, in the context of the mathematical modelling of biological data (see §1.6), of the problem of determining a parametrized retarded differential equation, along with the corresponding initial function, such that the solution is a good fit to an observed function. The underlying question, "Given the solution, what was the problem?" has been addressed by others (*e.g.* [24]). Our approach to answering this question is computational, and relies upon the numerical solution of various retarded differential equations (differential equations with time lag) and the minimization of an objective function associated with the solutions. However, some interesting analysis arises in this exercise, and part of this is presented below.

1.2 Analysis

Let us consider an n-dimensional system of linear delay differential equations (DDEs) with *time-dependent* coefficients, of the form

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

subject to

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

Here

$$y(t),\ f(t),\ \varphi(t)\ \in \mathbb{R}^{n\times 1}, \quad A(t),\ B(t)\ \in \mathbb{R}^{n\times n},$$

τ is a prescribed positive constant (the "lag"), and these functions will be assumed to be continuous on $[0,T]$. We shall find it convenient to suppose that $T \geq \tau$. The solution of (1.1) is dependent on the function φ in (1.1b), and when $y(t) = \varphi(t)$ (for $t \in [-\tau, 0]$) we can therefore write

$$y(t) \equiv y(\varphi; t) \quad (\text{for } t \in [-\tau, T]). \tag{1.2}$$

Remark 1.1 *If $B(t)$ does not vanish for $t \in [0,T]$ then the functions $y(\varphi_1; t)$ and $y(\varphi_2; t)$ differ when the functions φ_1 and φ_2 differ, and φ is uniquely determined if $y(\varphi; t)$ is given on $[0, \tau]$.*

The problem addressed here involves the determination (given prescribed $\tau > 0$, $A(t)$, $B(t)$, $f(t)$) of an initial function $\varphi(t)$, chosen from some class $\mathcal{F}$ of functions defined on $[-\tau, 0]$, such that the solution $y(\varphi; t)$ of the given retarded equation is a "best" approximation to an observed function. Throughout, we suppose $\mathcal{F} \subseteq PC[-\tau, 0])$, so that the initial function is required to be piecewise continuous with finite jumps at points of discontinuity. We shall emphasize the case where

$$\varphi \in \mathcal{F}, \qquad \text{where } \mathcal{F} := \{\varphi \in C[-\tau, 0) \ \text{ with bounded } \varphi(0)\} \tag{1.3a}$$

$$A,\ B,\ f \in C[0,T]. \tag{1.3b}$$

The initial function $\varphi(t)$ is determined by minimization of a quadratic functional (given below) involving the deviation of $y(\varphi; t)$ from the observed function.

The problem as formulated is shown, using the adjoint equations with deviating arguments, to be equivalent to the solution of a Fredholm integral equation (see §1.4). This can be of a type which is generally recognized as ill-posed, and the effect of regularization parameters is established. Extensions to nonlinear equations are proposed, these once studied elsewere.

Our results here extend those for "data assimilation problems" in ordinary and partial differential equations (*cf.* [20] and related works). Since inverse problems are frequently ill-posed, the connection with integral equations of the first kind (and their regularization to integral equations of the second kind) is unsurprising. Moreover, the rôle of adjoint equations in variation-of-constants formulae can be found in the literature, but our results appear to be new and the theory provides insight into practical methods.

1.3 Details of the identification problem

We suppose that we are given functions $\widehat{\varphi}$, $\widehat{y}$, such that

$$\widehat{\varphi}(t) \text{ for } t \in [-\tau, 0], \quad \text{and} \quad \widehat{y}(t) \text{ for } t \in [0, T] \tag{1.4}$$

and we seek that initial function $\varphi_\star$ which, in a sense that we make precise,

- deviates little from $\widehat{\varphi}$ on $[-\tau, 0]$ yet
- gives rise to a solution $y(t) \equiv y(\varphi_\star; t)$ which is a good approximation to $\widehat{y}$ $\big(y(t) \approx \widehat{y}(t)\big)$ on $[0, T]$.

We introduce the functional

$$S_\alpha^{\beta,\gamma}(\varphi) := \frac{\alpha}{2}\int_{-\tau}^{0} \left\|\varphi(t)-\widehat{\varphi}(t)\right\|^2 dt + \frac{\beta}{2}\left\|\varphi(0)-\widehat{\varphi}(0)\right\|^2 + \frac{\gamma}{2}\left\|y(\varphi;0)-\widehat{y}(0)\right\|^2 + \frac{1}{2}\int_{0}^{T} \left\|y(\varphi;t)-\widehat{y}(t)\right\|^2 dt \tag{1.5}$$

(in which $\alpha, \beta, \gamma \geq 0$ and $y(\varphi; 0) = \varphi(0)$) and $\widehat{\varphi} = \widehat{\varphi}(t)$ and where $y(\varphi; t)$ satisfies (1.1). We intend the value γ (introduced to provide some flexibility in the formulation) to assume the value 0 or 1.

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

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.*

Find a function $\varphi_\star \in \mathcal{F}$, such that $y(\varphi_\star; t)$ satisfies (1.1) and

$$S_\alpha^{\beta,\gamma}(\varphi_\star) = \min_{\varphi \in \mathcal{F}} S_\alpha^{\beta,\gamma}(\varphi), \tag{1.6}$$

where $S_\alpha^{\beta,\gamma}(\varphi)$ is defined by (1.5).

This formulation embodies parameters $\alpha \geq$ and $\beta \geq 0$, $\gamma \geq 0$, which (when positive) are "regularization parameters" (see [11], for example). This applies, in particular, to α. Thus, if we introduce an abstract operator $\mathcal{A}$ such that $y(\varphi; t) = \mathcal{A}\varphi(t)$ we have

$$S_\alpha^{0,0}(\varphi) = \frac{\alpha}{2}\int_{-\tau}^{0} \left\|\varphi(t) - \widehat{\varphi}(t)\right\|_2^2 dt + \frac{1}{2}\int_{0}^{T} \left\|\mathcal{A}\varphi - \widehat{y}(t)\right\|_2^2 dt,$$

(which is of the form associated with Tikhonov regularization [26] for recovery of $\varphi(t)$ from $\mathcal{A}\varphi(t) = \widehat{y}(t)$). Clearly, $\varphi_\star(t)$ depends on these parameters α, β, γ.

We consider an idealized situation where the functions $\widehat{y}(t)$ and $\widehat{\varphi}(t)$ are supposed to be unambiguously defined, but in practice $\widehat{y}(t)$ is usually defined by *a priori* observational data which is subject to noise. The choice of $\widehat{\varphi}(t)$ is, roughly speaking, determined by modelling conditions.

Remark 1.2 *(a) The relative weight of $\widehat{\varphi}$ and $\widehat{y}$ in our criterion is governed by the choice of the parameters $\alpha, \beta, \gamma \geq 0$ in $S_\alpha^{\beta,\gamma}(\varphi)$. Referring ahead to Remark 1.3 the parameter α, if non-zero, will be seen to be a "regularization parameter".*

(b) The parameters α, β and γ should be non-negative. The first two parameters are factors of terms involving a function $\widehat{\varphi}(t)$ that may be taken as an expectation of the form of the true initial function. If α or β are chosen to be very large, the influence of $\widehat{\varphi}(t)$ will dominate, and one's expectation will be realized at the expense of fitting the data. However, if $\beta = \gamma = 0$ and $\widehat{\varphi}(t)$ vanishes, we shall be seeking a minimum norm solution familiar in classical methods of regularization: $S_\alpha^{\beta,\gamma}(\varphi)$ then becomes

$$\mathfrak{S}_\alpha(\varphi) := \frac{\alpha}{2}\int_{-\tau}^{0} \|\varphi(t)\|^2 dt + \frac{1}{2}\int_{0}^{T} \|y(\varphi;t) - \widehat{y}(t)\|^2 dt. \tag{1.7}$$

(c) In the expression

$$\frac{\alpha}{2}\int_{-\tau}^{0} \|\varphi(t) - \widehat{\varphi}(t)\|^2 dt + \frac{\beta}{2}\|\varphi(0) - \widehat{\varphi}(0)\|^2$$

the terms involving β may be regarded as a modification of the integral term $\int_{-\tau}^{0} \|\varphi(t) - \widehat{\varphi}(t)\|^2 dt$ to provide a Stieltjes integral. If $\alpha = 0$, $\beta > 0$ then β has the appearance of a regularization parameter. The term involving γ has a similar rôle to that involving β, but is included to take account of the fact that $\widehat{y}(0)$ may be an observed value, unrelated to $\varphi(0)$.

In practical applications, $\widehat{y}(t)$ may be given only for a discrete set of arguments $\{t_\ell\} \subseteq [0,T]$. A dense extension for $t \in [0,T]$ should then be considered; this could be recoverable numerically by a process that may itself involve some form of regularization. As a further complication, the data may be subject to observational errors or may embody the results of more than one observation.

1.4 Our results in brief

We show below that the optimal initial function $\varphi_\star$ satisfies a coupled set of delay equations (see (2.9)), involving "adjoint equations", and we give an iterative technique for obtaining successive approximations φ_n to $\varphi_\star$. We show that the function $\varphi_\star$ identified by our chosen formulation is associated with the solution of a Fredholm integral equation and the iteration we propose is related to an iterative solution of the integral equation. Our discussion establishes a connection with a regularization method due to Lavrent'ev [14].

Theorem 1.1 *For appropriate $\mathfrak{g}(t)$, $K_{\beta,\gamma}(t,s)$ and $\gamma_0(\beta,\gamma)$, the function $\varphi_\star(t)$ solving the data assimilation problem satisfies equations of the form*

$$\alpha\varphi_\star(t) + \int_{-\tau}^{0} K_{\beta,\gamma}(t,s)\varphi_\star(s)ds = \mathfrak{g}(t), \quad \text{for } t \in [-\tau, 0), \tag{1.8}$$
$$\varphi_\star(0) = \gamma_0(\beta,\gamma).$$

In (1.8), the kernel $K_{\beta,\gamma}(t,s)$ is self-adjoint ($K_{\beta,\gamma}(t,s) = K^{\mathsf{T}}_{\beta,\gamma}(s,t)$) and positive-definite.

Remark 1.3 *If $\alpha > 0$ then equation (1.8) is a Fredholm equation of the second kind, and if $\alpha = 0$ it is a Fredholm equation of the first kind. The positive-definiteness of the kernel implies that the equation of the second kind ($\alpha > 0$) is uniquely solvable. However, Fredholm equations of the first kind ($\alpha = 0$) are ill-posed. For this reason, the introduction of $\alpha > 0$ is said to regularize the problem.*

The above theorem is derived through the use of an adjoint differential equation with a deviating argument. (The link with a variation of parameters formula based upon the fundamental solution is exploited.) We shall show properties of the kernel $K_{\beta,\gamma}(t,\sigma)$ using the fundamental matrix solution for a delay differential equation. The iteration that yields successive φ_n is related to the integral equations in (1.8). We shall prove the convergence of the described iteration algorithm, based upon properties of the kernel.

1.5 Continuity issues

In discussing a DDE of the form (1.1), where the condition (1.1b) is defined to the left of $[0,T]$, the term *derivative* is interpreted on $[0,\tau)$ as a right-hand derivative when a two-sided derivative does not exist, and 'differentiable' signifies the existence of a right-hand derivative.

In order to justify the following analysis (in particular, the application of integration by parts), we need to be aware of the continuity properties of solutions of DDEs of the form (1.1) when $\mathcal{F}$ is as defined in (1.3a). In what follows, we shall assume (1.3) unless otherwise stated, and in consequence integration by parts will be valid wherever we need it in the discussion that follows. The existence of a unique differentiable solution $y(t)$, for $t \in [0,T]$ is guaranteed if $\varphi(t)$ is continuous for $-\tau \le t < 0$ and, for arbitrary $T > 0$, $y(\cdot)$ has a continuous derivative (*cf.* Bellman & Cooke [9, pp. 52 & 73]) on the interval $[0,T]$ provided that $f(t)$ is continuous (as assumed).

A similar discussion applies to the propagation of discontinuities in an equation with deviating arguments of the form

$$-\frac{dw^{\mathsf{T}}(t)}{dt} - w^{\mathsf{T}}(t)A(t) - w^{\mathsf{T}}(t+\tau)B(t+\tau) = g^{\mathsf{T}}(t), \quad \text{for } t \in [0,T], \tag{1.9a}$$

or, equivalently, $-\dfrac{dw(t)}{dt} - A^{\mathsf{T}}(t)w(t) - B^{\mathsf{T}}(t+\tau)w(t+\tau) = g(t)$, for $t \in [0,T]$, with $w(t)$, $g(t) \in \mathbb{R}^{n\times 1}$, $A(t)$, $B(t) \in \mathbb{R}^{n\times n}$ and subject to

$$w(t) = \phi(t), \quad \text{for } t \in [T, T+\tau]. \tag{1.9b}$$

The final condition (1.9b) holds to the right of $[0,T]$ and, if required, the derivative in (1.9a) is interpreted on $(0,T]$ as a left-hand derivative. An example of (1.9) arises in (3.2), below. It is convenient at this point to note that by a change of variables ($s = t - \tau$), (1.9) can be re-written in the same form as (1.1), namely a linear DDE for a function ϖ related to w with an *initial* function related to ϕ and left-hand derivatives in (1.9a) become right-hand derivatives in the DDE formulation.

Remark 1.4 *The solutions of* neutral *delay differential equations (NDDEs) generally inherit stronger discontinuities, and extensions to NDDEs therefore require a refinement of the integration by parts.*

1.6 Additional remarks

Remark 1.5 *Results that relate to (1.1) will be pertinent when extending the discussion to systems of* non-linear *equations such as*

$$y'(t) = f\Big(t, y(t), y(t-\tau)\Big) \quad (\textit{for } t \in [0,T]) \tag{1.10a}$$

with

$$y(t) = \varphi(t) \quad (\textit{for } t \in [-\tau, 0]). \tag{1.10b}$$

Remark 1.6 *Note that we concentrate on the identification of the initial function $\varphi(t)$, it being assumed that the DDE is known. For discussions of parameter estimation for DDEs, in particular in the context of cell dynamics, see Baker, Bocharov & Paul [4], Baker, Bocharov, Paul, & Rihan [5], and [6], and the references therein. In addition to DDEs, both ordinary differential equations (ODEs) and parabolic partial differential equations (PPDEs) arise in the modeling of cell populations. Similar so-called — (see [1, 2, 19, 20, 23]) data assimilation problems*[1] *have been discussed for ODEs, see [19], and PPDEs, see [1, 2, 20]; the present work fills a gap in the discussion of DDEs.*

Remark 1.7 *It may be remarked that (1.1) can be transformed into a problem in which $\varphi(t)$ vanishes by changing $f(t)$ on the interval $[0,\tau]$ by adding $B(t)\varphi(t-\tau)$ when $t \in [0,\tau]$. With $x(t) = y(t) - \varphi(t)$, when $-\tau \le t \le T$, we can re-write system (1.1) in the form*

$$\frac{dx(t)}{dt} - A(t)x(t) - B(t)x(t-\tau) = f_0(t), \quad \textit{for } t \in [0,T],$$

subject to

$$x(t) = 0 \quad \textit{for } t \in [-\tau, 0],$$

where

$$f_0(t) = f(t) + B(t)\varphi(t-\tau), \ \textit{when } 0 \le t \le \tau, \ \textit{and } f_0(t) = f(t), \ \textit{when } t > \tau.$$

In consequence, the problem of identifying an initial function $\varphi(t)$ can be reformulated as the problem of identifying an inhomogeneous term $f_0(t)$.

[1] The term "data assimilation" appears to have originated in model identification in problems in meteorology [18, 21, 22].

2 The optimization problem

Let us consider the basic data assimilation problem (outlined in Section 1). Recall that the function in (1.5) is

$$S_\alpha^{\beta,\gamma}(\varphi) := \frac{\alpha}{2}\int_{-\tau}^{0} \|\varphi(t)-\widehat{\varphi}(t)\|^2 dt + \frac{\beta}{2}\|\varphi(0)-\widehat{\varphi}(0)\|^2 + \frac{\gamma}{2}\|y(\varphi;0)-\widehat{y}(0)\|^2 + \frac{1}{2}\int_{0}^{T} \|y(\varphi;t)-\widehat{y}(t)\|^2 dt. \tag{2.1}$$

2.1 Condition for a minimum

In order to consider the minimum of the functional (2.1), we analyze[2] $S_\alpha^{\beta,\gamma}(\varphi+\varepsilon\psi)$. If $\varphi_\star \in \mathcal{F}$ provides a minimum of the functional $S_\alpha^{\beta,\gamma}(\varphi)$, 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 the linear space $\mathcal{F}$. We note $S_\alpha^{\beta,\gamma}(\varphi) \ge 0$, and $S_\alpha^{\beta,\gamma}(\varphi+\varepsilon\psi)$ is a quadratic in ε. To write down $S_\alpha^{\beta,\gamma}(\varphi_\star+\varepsilon\psi)$ we need an expression for $y(\varphi_\star+\varepsilon\psi;t)$, and we have it in the following result.

Lemma 2.1 *Write*

$$Ly(t) := y'(t) - A(t)y(t) - B(t)y(t-\tau) \ (\textit{for } t\in[0,T]), \quad \textit{and} \quad My(t) = y(t) \ (\textit{for } t\in[-\tau,0]).$$

By virtue of the linearity of L and M,

$$y(\varphi_\star+\varepsilon\psi;t) = y(\varphi_\star;t) + \varepsilon z(\psi;t) \tag{2.2}$$

where $z(t)\equiv z(\psi;t)$ satisfies

$$Lz(t) = 0 \quad (\textit{for } t\in[0,T]) \quad \textit{and} \quad Mz(t) = \psi(t) \quad (\textit{for } t\in[-\tau,0]), \tag{2.3}$$

that is,

$$\frac{dz(\psi;t)}{dt} - A(t)z(\psi;t) - B(t)z(\psi;t-\tau) = 0, \quad \textit{for } t\in[0,T], \tag{2.4a}$$

$$z(\psi;t) = \psi(t), \quad \textit{for } t\in[-\tau,0), \ \textit{and } z(\psi;0) = \psi(0). \tag{2.4b}$$

The condition $z(\psi;t)=\psi(t)$, for $t\in[-\tau,0]$ is expressed as $z(\psi;t)=\psi(t)$, for $t\in[-\tau,0)$, and $z(\psi;0)=\psi(0)$ in order to *emphasize* the possibility that $\psi(0-)$ may not equal $\psi(0)$. The function $z(\psi;t)$ vanishes if and only if $\psi(t)$ vanishes.

We can write

$$\begin{aligned} &S_\alpha^{\beta,\gamma}(\varphi+\varepsilon\psi) = \\ &\frac{\alpha}{2}\int_{-\tau}^{0}\|\varphi(t)+\varepsilon\psi(t)-\widehat{\varphi}(t)\|^2 dt + \frac{\beta}{2}\|\varphi(0)+\varepsilon\psi(0)-\widehat{\varphi}(0)\|^2 + \\ &\frac{1}{2}\int_{0}^{T}\|y(\varphi;t)+\varepsilon z(\psi;t)-\widehat{y}(t)\|^2 dt + \frac{\gamma}{2}\|y(\varphi;0)+\varepsilon z(\psi;0)-\widehat{y}(0)\|^2. \end{aligned} \tag{2.5}$$

[2] We note the relationship with optimal control problems [3] when we regard the function φ as a control, but shall not pursue the details here.

We obtain

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

where

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

$$\int_0^T \{y(\varphi;t) - \widehat{y}(t)\}^{\mathsf{T}} z(\psi;t)dt + \beta\{\varphi(0) - \widehat{\varphi}(0)\}^{\mathsf{T}} \psi(0) + \gamma\{y(\varphi;0) - \widehat{y}(0)\}^{\mathsf{T}} z(\psi;0);$$

$$Q_\alpha^{\beta,\gamma}(\psi) = \frac{\alpha}{2}\int_{-\tau}^{0} \|\psi(t)\|dt + \frac{1}{2}\int_0^T \|z(\psi;t)\|dt + \frac{\beta}{2}\|\psi(0)\|^2 + \frac{\gamma}{2}\|z(\psi;0)\|^2. \tag{2.6c}$$

We observe that in the latter expressions we have $z(\psi;0) = \psi(0)$. We obtain the following result.

Lemma 2.2 *If*

$$\begin{aligned} J_\alpha^{\beta,\gamma}(\varphi,\psi) \equiv \tfrac{\alpha}{2}\int_{-\tau}^{0} \{\varphi(t) - \widehat{\varphi}(t)\}^{\mathsf{T}} \{\psi(t) - \widehat{\varphi}(t)\}dt + \tfrac{1}{2}\int_0^T \{y(\varphi;t) - \widehat{y}(t)\}^{\mathsf{T}} \{y(\psi;t) - \widehat{y}(t)\}dt + \\ + \tfrac{\beta}{2}\{\varphi(0) - \widehat{\varphi}(0)\}^{\mathsf{T}} \{\psi(0) - \widehat{\varphi}(0)\} + \tfrac{\gamma}{2}\{y(\varphi;0) - \widehat{y}(0)\}^{\mathsf{T}} \{y(\psi;0) - \widehat{y}(0)\}, \end{aligned} \tag{2.7}$$

then $S_\alpha^{\beta,\gamma}(\varphi) = J_\alpha^{\beta,\gamma}(\varphi,\varphi)$, *and* $P_\alpha^{\beta,\gamma}(\varphi,\psi) = J_\alpha^{\beta,\gamma}(\varphi,\psi) + J_\alpha^{\beta,\gamma}(\psi,\varphi) = 2J_\alpha^{\beta,\gamma}(\varphi,\psi)$. *Finally,* $Q_\alpha^{\beta,\gamma}(\psi) = \frac{1}{2}\frac{\partial^2}{\partial\varepsilon^2}\{J_\alpha^{\beta,\gamma}(\varphi+\varepsilon\psi,\varphi+\varepsilon\psi)\}$, *and* $Q_\alpha^{\beta,\gamma}(\psi) \geq 0$ *and* $Q_\alpha^{\beta,\gamma}(\psi) = 0$ *if and only if* $\psi = 0$ *on* $[-\tau,0]$.

Remark 2.1 $P_\alpha^{\beta,\gamma}(\varphi,\psi)$ *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 (2.1) this derivative must be zero for all solutions $z(w;t)$ *of the equation (2.4), where* $z(w;t) = w(t)$ *when* $t \in [-\tau,0]$. *In fact, the form of* $S_\alpha^{\beta,\gamma}(\varphi)$ *in (2.7) establishes that we obtain a* global *minimum for a function* $\varphi_\star$ *(which is unique given that* $Q_\alpha^{\beta,\gamma}(\psi) = 0$ *if and only if* $\psi = 0$*) such that* $P_\alpha^{\beta,\gamma}(\varphi_\star,\psi)$ *vanishes for all admissible functions* $\psi \in \mathcal{F}$.

We have the following result.

Theorem 2.1 *A function* $\varphi_\star(t)$ *defined on* $[-\tau,0]$ *minimizes* $S_\alpha^{\beta,\gamma}(\varphi)$ *in (2.1) for* $\varphi \in \mathcal{F}$ *if and only if* $P_\alpha^{\beta,\gamma}(\varphi,\psi)$ *in (2.6b) vanishes for all* $\psi \in \mathcal{F}$, *where* $z = z(\psi;t)$ *satisfies (2.2).*

2.2 A method for finding the optimal $\varphi_\star$

We propose a method to obtain the initial function $\varphi_\star \in \mathcal{F}$ which minimizes $S_\alpha^{\beta,\gamma}(\varphi)$ on $\mathcal{F}$.

This method comprises the solution of a set of coupled equations, written in the form

$$\frac{dy(t)}{dt} - A(t)y(t) - B(t)y(t-\tau) = f(t), \quad \text{for } t \in [0,T], \tag{2.9a}$$

$$y(t) = \varphi_\star(t) \quad \text{for } t \in [-\tau, 0), \quad y(0) = \varphi_\star(0), \tag{2.9b}$$

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

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

$$\alpha(\varphi_\star(t) - \widehat{\varphi}(t)) + [B(t+\tau)]^{\mathsf{T}} x(t+\tau) = 0 \quad \text{for } t \in [-\tau, 0), \tag{2.9e}$$

$$x(0) + \beta\{\varphi_\star(0) - \widehat{\varphi}(0)\} + \gamma\{\varphi_\star(0) - \widehat{y}(0)\} = 0. \tag{2.9f}$$

Remark 2.2 *If equation (2.9e) were to hold for $t \in [-\tau, 0]$ then the equation for $t = 0$ would have the form $\alpha\varphi_\star(0) + [x^{\mathsf{T}}(\tau)B(\tau)]^{\mathsf{T}} = \alpha\widehat{\varphi}(0)$. Equation (2.9f) defines $\varphi_\star(0)$ if $\beta + \gamma \neq 0$.*

In the process of determining $\varphi_\star(t)$ we also determine the corresponding $y(\varphi_\star; t)$ and $x^{\mathsf{T}}(\varphi_\star; t)$.

We shall show that the solution of (2.9) does provide a minimum. The method suggested for the solution of (2.9) is based upon the following iteration:

Definition 2.1 (An iteration for finding the optimal initial function)

$$\frac{dy_n(t)}{dt} - A(t)y_n(t) - B(t)y_n(t-\tau) = f(t), \quad \textit{for } t \in [0, T], \tag{2.10a}$$

$$y_n(t) = \varphi_n(t) \quad \textit{for } t \in [-\tau, 0), \quad y_n(0) = \varphi_n(0), \tag{2.10b}$$

$$-\frac{dx_n^{\mathsf{T}}(t)}{dt} - x_n^{\mathsf{T}}(t)A(t) - x_n^{\mathsf{T}}(t+\tau)B(t+\tau) = [y_n(\varphi_n; t) - \widehat{y}(t)]^{\mathsf{T}}, \quad \textit{for } t \in [0, T], \tag{2.10c}$$

$$x_n^{\mathsf{T}}(t) = 0, \quad \textit{for } t \in [T, T+\tau], \tag{2.10d}$$

$$\varphi_{n+1}(t) = \varphi_n(t) + \delta_n\big(\alpha(\varphi_n(t) - \widehat{\varphi}(t)) + [B(t+\tau)]^{\mathsf{T}} x_n(t+\tau)\big) \quad \textit{for } t \in [-\tau, 0), \tag{2.10e}$$

$$\varphi_{n+1}(0) = \varphi_n(0) + \delta_n'\big\{(\beta+\gamma)\varphi_n(0) + x_n(0) - \beta\widehat{\varphi}(0) - \gamma\widehat{y}(0)\big\}, \tag{2.10f}$$

for $n = 0, 1, 2, \ldots$ and $\{\delta_j\}$, $\{\delta_j'\}$ are appropriately chosen scalars. What is proposed is to find the solution of (2.9) using the following procedure:

- *Choose a starting approximation for the initial function ($\varphi_0(s)$, $s \in [-\tau, 0]$).*
- *For $n = 0, 1, \ldots, N$, where the choice of N yields appropriate accuracy:*
  1.) *Obtain the solution $y_n = y_n(t)$ of the original problem (2.10a) for $\varphi = \varphi_n(s)$.*
  2.) *Obtain the solution $x_n^{\mathsf{T}} = x_n^{\mathsf{T}}(t)$ of the adjoint problem (2.10c) with known right-hand side $y_n(t) - \widehat{y}(t)$ when $t \in [0, T]$ with the condition $x_n^{\mathsf{T}}(s) = 0$, $s \in [T, T+\tau]$.*
  3.) *Find the next iterate $\varphi_{n+1}(s)$ using equations (2.10e) and (2.10f).*

We shall establish that if we allow N to tend to infinity, this iteration converges to give $\varphi_\star$, for a feasible range of values of α and δ. Our convergence result (see Theorem 3.3) follows very simply from the fact that the solution of (2.9) satisfies an integral equation of the form

$$\alpha\varphi_\star(t) + \int_{-\tau}^{0} K_{\beta,\gamma}(t, s)\varphi_\star(s)ds = \mathfrak{g}(t), \tag{2.11}$$

for $t \in [-\tau, 0)$, in which $K_{\beta,\gamma}(t, s)$ and $\mathfrak{g}(t) \equiv \mathfrak{g}_\alpha^{\beta,\gamma}(t)$ are given in detail in (3.13). The proposed iteration is related to this integral equation (see (3.32), below).

3 The underlying theory

The theoretical discussion of our method depends upon aspects of the theory of DDEs and of integral equations.

3.1 A rôle for the adjoint equation

We shall obtain an equivalent formulation of the problem (1.6), based upon adjoint equations. This will provide an alternative characterization to that of Theorem 2.1. The purpose of this approach is to derive (2.9) and (2.10) in order to solve the "data assimilation problem" (1.6). The results presented here arise, in effect, due to the relation between the fundamental solution, certain adjoint problems, and variation of parameters formulae [7].

Lemma 3.1 *Let $y = y(\varphi; t)$ be a solution of the problem (1.1) and let $z = z(\psi; t)$ be a solution of the homogeneous problem (2.4). Then the first variation of the functional $S_\alpha^{\beta,\gamma}(\varphi)$ can be represented in the form* $P_\alpha^{\beta,\gamma}(\varphi,\psi) = \int_{-\tau}^{0} \Big\{\alpha[\varphi(t) - \widehat{\varphi}(t)] + x^{\mathsf{T}}(t+\tau)B(t+\tau)\Big\}^{\mathsf{T}} \psi(t)dt +$

$$\Big\{x(0) + \beta[\varphi(0) - \widehat{\varphi}(0)] + \gamma[y(\varphi;0) - \widehat{y}(0)]\Big\}^{\mathsf{T}} \psi(0), \tag{3.1}$$

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

$$-\frac{dx^{\mathsf{T}}(t)}{dt} - x^{\mathsf{T}}(t)A(t) - x^{\mathsf{T}}(t+\tau)B(t+\tau) = [y(\varphi;t) - \widehat{y}(t)]^{\mathsf{T}}, \quad \text{for } t \in [0,T], \tag{3.2a}$$

$$x^{\mathsf{T}}(t) = 0, \quad \text{for } t \in [T, T+\tau]. \tag{3.2b}$$

Equations (3.2) correspond to an adjoint problem, with a special forcing term $y(\varphi;t) - \widehat{y}(t)$.

Remark 3.1 *If required, the derivative in (3.2a) is interpreted as the left-hand derivative. The derivative of the function $x^{\mathsf{T}}(t)$ satisfying (3.2) inherits from $y(\varphi;t) - \widehat{y}(t)$ any jump discontinuities at points in $[0,T]$; if $y(\varphi;t) - \widehat{y}(t)$ is continuous (in particular if $y(t)$ satisfies (1.1) where $f \in C[0,T]$ and $u \in C[-\tau,0]$) then $\frac{d}{dt}x^{\mathsf{T}}(t)$ is continuous on $[0,T]$.*

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

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

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

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

$$\underbrace{-\int_0^T \frac{dx^{\mathsf{T}}(t)}{dt} z(\psi;t)dt}_{\text{term } (i)} \underbrace{-\int_0^T x^{\mathsf{T}}(t)A(t)z(\psi;t)dt}_{\text{term } (ii)} \underbrace{-\int_0^T x^{\mathsf{T}}(t+\tau)B(t+\tau)z(\psi;t)dt}_{\text{term } (iii)}. \tag{3.3}$$

Using integration by parts, we can write the term (i) in (3.3) (involving the derivative) as

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

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

$$\begin{aligned}\int_0^T x^{\mathsf{T}}(t+\tau)B(t+\tau)z(\psi;t)dt &= \int_\tau^{T+\tau} x^{\mathsf{T}}(s)B(s)z(\psi;s-\tau)ds \\ &= \int_0^T x^{\mathsf{T}}(s)B(s)z(\psi;s-\tau)ds - \int_{-\tau}^0 x^{\mathsf{T}}(s+\tau)B(s+\tau)\psi(s)ds.\end{aligned}$$

Finally, we can rewrite the first variation in the form

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

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

The function $z(\psi;t)$ satisfies the homogeneous equation (2.4). For $P_\alpha^{\beta,\gamma}(\varphi,\psi)$ we therefore obtain the expression in (3.1) and our lemma follows.

$$\begin{aligned}S_{\alpha,\beta}^{\gamma}(\varphi+\varepsilon\psi) = S_\alpha^{\beta,\gamma}(\varphi) + \varepsilon\Bigg(&\int_{-\tau}^0 \{\alpha[\varphi(t)-\widehat{\varphi}(t)]^{\mathsf{T}} + x^{\mathsf{T}}(t+\tau)B(t+\tau)\}\psi(t)dt + \\ &(x^{\mathsf{T}}(0) + \beta[\varphi(0)-\widehat{\varphi}(0)]^{\mathsf{T}} + \gamma[y(\varphi;0)-\widehat{y}(0)]^{\mathsf{T}})\psi(0)\Bigg) + \varepsilon^2 Q_\alpha^{\beta,\gamma}(\psi).\end{aligned} \tag{3.4}$$

and since $y(\varphi;0) = \varphi(0)$, we have established the following theorem.

Theorem 3.1 *A function $\varphi_\star(t)$ defined on $[-\tau,0]$ minimizes $S_\alpha^{\beta,\gamma}(\varphi)$ for $\varphi \in \mathcal{F}$ where $\mathcal{F} = C[-\tau,0) \cap \{\varphi | \varphi(0)$ is bounded$\}$ if, where x^{T} satisfies (2.2),*

$$\int_{-\tau}^0 \{\alpha[\varphi(t)-\widehat{\varphi}(t)]^{\mathsf{T}} + x^{\mathsf{T}}(t+\tau)B(t+\tau)\}\psi(t)dt = 0 \tag{3.5a}$$

for all $\psi \in C[-\tau,0)$, and

$$\{x^{\mathsf{T}}(0) + \beta[\varphi(0)-\widehat{\varphi}(0)]^{\mathsf{T}} + \gamma[\varphi(0)-\widehat{y}(0)]^{\mathsf{T}}\}\psi(0) = 0. \tag{3.5b}$$

If we consider the problem of minimizing over a subset $\mathcal{F}_n$ of $\mathcal{F}$, we obtain the analogous results with $\mathcal{F}$ replaced in $\mathcal{F}_n$.

3.2 A rôle for fundamental solutions

Consider a linear delay differential equation of the form

$$\frac{dy(t)}{dt} - A(t)y(t) - B(t)y(t-\tau) = f(t), \quad \text{for } t \in [0,T], \tag{3.6a}$$

with initial condition

$$y(t) = \varphi(t) \quad \text{for } t \in [-\tau, 0]. \tag{3.6b}$$

We introduce (see [12, pp. 359-363], [13, pp. 148-150]) the fundamental solution for (3.6).

Definition 3.1 *Let $Y(s,t)$ be a solution of the equation*

$$\frac{\partial Y(s,t)}{\partial s} + Y(s,t)A(s) + Y(s+\tau,t)B(s+\tau) = 0, \quad \textit{for } s < t, \tag{3.7a}$$

which satisfies

$$Y(s,t) = \begin{cases} 0, & \textit{for } t < s, \\ I, & \textit{for } s = t. \end{cases} \tag{3.7b}$$

The function $U(t,s) \equiv Y(s,t)$ is called the fundamental matrix *(or* fundamental solution.

Lemma 3.2 *The solution of the system (3.6) is given by*

$$y(t) = Y(0,t)\varphi(0) + \int_{-\tau}^{0} Y(s+\tau,t)B(s+\tau)\varphi(s)ds + \int_{0}^{t} Y(s,t)f(s)ds. \tag{3.8}$$

For related results, see [7]. One may see that the essential structure in (3.8) is that

$$y(\varphi, f; t) \equiv y(t) = Y(0,t)\varphi(0) + \mathrm{F}\varphi(t) + \mathrm{V}f(t) \tag{3.9a}$$

where F is a Fredholm integral operator defined on $\mathcal{F}$ and V is a Volterra integral operator defined on $C[0,T]$. Hence, also, the solution of the homogeneous problem with initial function ψ is

$$z(\psi; t) \equiv z(t) = Y(0,t)\psi(0) + \mathrm{F}\psi(\mathrm{t}). \tag{3.9b}$$

Remark 3.2 *The function $U(t,s)$ in Definition 3.1 satisfies, for $t > s$,*

$$\frac{\partial U(t,s)}{\partial t} - A(t)U(t,s) - B(t)U(t-\tau,s) = 0, \tag{3.10}$$

with the initial conditions $U(t,s) = 0$ for $s - \tau \le t < s$, $U(t,s) = I$ for $t = s$.
(For a proof, see Hale[13, pp. 148-151].) We note that $U(t,s)$ is the fundamental solution (see [12, 13, 17]) for a problem of the form

$$-\frac{d}{dt}x^{\mathsf{T}}(t) - x^{\mathsf{T}}(t)A(t) - x^{\mathsf{T}}(t+\tau)B(t+\tau) = g(t) \textit{ for } t \in [0,T], \tag{3.11a}$$

$$x^{\mathsf{T}}(t) = v^{\mathsf{T}}(t) \textit{ for } t \in [T, T+\tau]. \tag{3.11b}$$

Lemma 3.3 *The solution of (3.11) is expressible as*

$$x^{\mathsf{T}}(t) = v^{\mathsf{T}}(T)U(T,t) + \int_{T}^{T+\tau} v^{\mathsf{T}}(s)B(s)U(s-\tau,t)ds + \int_{t}^{T} g^{\mathsf{T}}(s)U(s,t)ds.$$

It follows that for the adjoint equation (3.2) we have an expression for the solution $x^{\mathsf{T}}(t)$ in terms of $U(s,t)$. Since $x^{\mathsf{T}}(T)$ and $\int_{T}^{T+\tau} x^{\mathsf{T}}(s)B(s)U(s-\tau,t)ds$ vanish, we find:

Corollary 3.1 *The solution $x^{\mathsf{T}}(t)$ of (3.2) is given by*

$$x^{\mathsf{T}}(t) = \int_{t}^{T} \left[y(\varphi;s) - \widehat{y}(s)\right]^{\mathsf{T}} U(s,t)ds, \tag{3.12}$$

where $U(s,t)$ is the fundamental solution.

3.3 An integral equation for the optimal initial function $\varphi_\star$

We shall establish the following theorem.

Theorem 3.2 *For $t \in [-\tau, 0)$,*

$$\alpha\varphi_\star(t) + \int_{-\tau}^{0} K_{\beta,\gamma}(t,s)\varphi_\star(s)ds = \mathfrak{g}(t), \tag{3.13}$$

in which

$$\begin{aligned} K_{\beta,\gamma}(t,s) = &\int_{\xi=t+\tau}^{T} [B(t+\tau)]^{\mathsf{T}} [Y(t+\tau,\xi)]^{\mathsf{T}} Y(s+\tau,\xi)B(s+\tau)d\xi + \\ &- \int_{\xi=t+\tau}^{T}\int_{\mu=s+\tau}^{T} [B(t+\tau)]^{\mathsf{T}} [Y(t+\tau,\xi)]^{\mathsf{T}} Y(0,\xi)\mathcal{D}_{\beta,\gamma}^{-1}[Y(0,\mu)]^{\mathsf{T}} Y(s+\tau,\mu)B(s+\tau)d\mu d\xi \end{aligned} \tag{3.14a}$$

and

$$\begin{aligned} \mathfrak{g}_{\alpha}^{\beta,\gamma}(t) = \alpha\widehat{\varphi}(t) - &\int_{t+\tau}^{T} [B(t+\tau)]^{\mathsf{T}} [Y(t+\tau,s)]^{\mathsf{T}} Y(0,s)\mathcal{D}_{\beta,\gamma}^{-1} F_0(\widehat{\varphi}(0),\widehat{y},f)ds + \\ &- \int_{t+\tau}^{T} [B(t+\tau)]^{\mathsf{T}} [Y(t+\tau,s)]^{\mathsf{T}} \left(\int_{0}^{s} Y(\xi,s)f(\xi)d\xi - \widehat{y}(s)\right) ds \end{aligned} \tag{3.14b}$$

with symmetric and positive-definite constant matrix

$$\mathcal{D}_{\beta,\gamma} := (\beta+\gamma)I + \int_{0}^{T} [Y(0,s)]^{\mathsf{T}} Y(0,s)ds. \tag{3.14c}$$

Additionally, $\varphi_\star(0)$ satisfies a relation

$$\mathcal{D}_{\beta,\gamma}\varphi_\star(0) = \beta\widehat{\varphi}(0) + \gamma\widehat{\mathfrak{y}}(0) + \int_{-\tau}^{0}\int_{s+\tau}^{T} [Y(0,\xi)]^{\mathsf{T}} Y(s+\tau,\xi)B(s+\tau)\varphi(s)d\xi ds - \int_0^T [Y(0,s)]^{\mathsf{T}} \left(\int_0^s Y(\xi,s)f(\xi)d\xi - \widehat{\mathfrak{y}}(s)\right) ds, \tag{3.15}$$

Remark 3.3 *It is easy to show that the constant matrix $\mathcal{D}_{\beta,\gamma}$ is symmetric, positive-definite (and therefore has an inverse). Consider $\mathcal{J}(t,s) = \int_0^T [Y(t,\xi)]^{\mathsf{T}} Y(s,\xi)d\xi$ we have $\mathcal{J}(t,s) = \mathcal{J}^{\mathsf{T}}(s,t)$, and therefore $D_{\beta,\gamma} \equiv (\beta+\gamma)I + \mathcal{J}(0,0) = D_{\beta,\gamma}^{\mathsf{T}}$. By definition*

$$\int_0^T u^{\mathsf{T}}(t)\int_0^T [Y(0,s)]^{\mathsf{T}} Y(0,s)ds\ u(t)dt = \int_0^T\int_0^T u^{\mathsf{T}}(t)[Y(0,s)]^{\mathsf{T}} Y(0,s)u(t)dsdt \geq 0 \tag{3.16}$$

for some $u(t)$, and $\beta \geq 0$, $\gamma \geq 0$ so we have $\left(D_{\beta,\gamma}u, u\right) \geq 0$. Thus the positive semi-definiteness of $D_{\beta,\gamma}$ is obtained. Since $Y(0,s)$ is a solution of the equation (3.7a) we can conclude from (3.16) that $(D_{\beta,\gamma}u,u) = 0$ implies $u \equiv 0$. The existence of $\mathcal{D}_{\beta,\gamma}^{-1}$ follows.

Remark 3.4 *If we define $\mathfrak{R}(t,\xi) \equiv [Y(0,\xi)]^{\mathsf{T}} Y(t,\xi)B(t)$ and $\mathfrak{S}(t,\xi) \equiv Y(t,\xi)B(t)$, we may write (3.14) in brief*

$$K_{\beta,\gamma}(t,s) = \int_{t+\tau}^{T} [\mathfrak{S}(t+\tau,\xi)]^{\mathsf{T}}\mathfrak{S}(s+\tau,\xi)d\xi - \int_{\xi=t+\tau}^{T} [\mathfrak{R}(t+\tau,\xi)]^{\mathsf{T}} d\xi\ \mathcal{D}_{\beta,\gamma}^{-1}\int_{\mu=s+\tau}^{T} \mathfrak{R}(s+\tau,\mu)d\mu; \tag{3.17a}$$

$$\mathfrak{g}_\alpha^{\beta,\gamma}(t) = \alpha\widehat{\varphi}(t) - \int_{t+\tau}^{T} [\mathfrak{R}(t+\tau,s)]^{\mathsf{T}}\mathcal{D}_{\beta,\gamma}^{-1}F_0(\widehat{\varphi}(0),\widehat{\mathfrak{y}},f)ds + \int_{t+\tau}^{T} [\mathfrak{S}(t+\tau,s)]^{\mathsf{T}} \left(\int_0^s Y(\xi,s)f(\xi)d\xi - \widehat{\mathfrak{y}}(s)\right) ds, \tag{3.17b}$$

where

$$F_0(\widehat{\varphi}(0),\widehat{\mathfrak{y}},f) = \beta\widehat{\varphi}(0) + \beta\widehat{\mathfrak{y}}(0) - \int_0^T [Y(0,s)]^{\mathsf{T}} \left(\int_0^s Y(\xi,s)f(\xi)d\xi - \widehat{\mathfrak{y}}(s)\right) ds,$$

Corollary 3.2 *When $\alpha = 0$, we have*

$$\int_{-\tau}^{0} K_{\beta,\gamma}(t,s)\varphi_\star(s)ds = \mathfrak{g}_\alpha^{\beta,\gamma}(t) \ \ \textit{for} \ \ t \in [-\tau,0). \tag{3.18}$$

where $K_{\beta,\gamma}(t,s)$ and $\mathfrak{g}_\alpha^{\beta,\gamma}(t)$ defined by (3.14).

The equation (3.13) is obtainable from (3.18) by applying Lavrent'ev's method ([14, 25], [27, p. 89]) – which is sometimes called the "method of singular perturbation" [15, 25] – to (3.18).

3.3.1 Properties of the kernel $K_{\beta,\gamma}(t,s)$.

Since the function $\varphi = \varphi_\star(t)$ minimizes $S_\alpha^{\beta,\gamma}(\varphi)$ over the space $\mathcal{F}$, we can write

$$P_\alpha^{\beta,\gamma}(\varphi,\psi) = 0, \quad \text{for all } \psi \in \mathcal{F}. \tag{3.19}$$

From (2.6b) we have

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

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

Clearly,

$$P_\alpha^{0,0}(\varphi,\psi) = \alpha \int_{-\tau}^{0} \{\varphi(t) - \widehat{\varphi}(t)\}^{\mathsf{T}} \psi(t)dt + \int_{0}^{T} \{y(\varphi,f;t) - \widehat{\mathrm{y}}(t)\}^{\mathsf{T}} z(\psi;t)dt =$$

$$= \alpha \int_{-\tau}^{0} \varphi^{\mathsf{T}}(t)\psi(t)dt + \int_{0}^{T} y^{\mathsf{T}}(\varphi,f;t)z(\psi;t)dt - \Big\{\alpha \int_{-\tau}^{0} \widehat{\varphi}^{\mathsf{T}}(t)\psi(t)dt + \int_{0}^{T} \widehat{\mathrm{y}}^{\mathsf{T}}(t)z(\psi;t)dt\Big\}.$$

Using the expressions for $y(t)$ and $z(t)$ associated with (3.9),

$$\int_{0}^{T} \{y(\varphi,f;t) - \widehat{\mathrm{y}}(t)\}^{\mathsf{T}} z(\psi;t)dt =$$

$$\Big\{\int_{0}^{T} \{Y(0,t)\varphi(0) + \mathrm{F}\varphi(t) + \mathrm{V}f(t) - \widehat{\mathrm{y}}(t)\}Y(0,t)dt\Big\}^{\mathsf{T}} \psi(0) +$$

$$\int_{0}^{T} \Big\{Y(0,t)\varphi(0) + \mathrm{F}\varphi(t) + \mathrm{V}f(t) - \widehat{\mathrm{y}}(t)\Big\}^{\mathsf{T}} \mathrm{F}\psi(t)dt.$$

Let us write $P_\alpha^{\beta,\gamma}(\varphi,\psi)$ in the form

$$P_\alpha^{\beta,\gamma}(\varphi,\psi) = \overbrace{\underbrace{{}^0P_\alpha^{0,0}(\varphi,\psi)}_{\text{terms in } \varphi} + \underbrace{{}^1P_\alpha^{0,0}(\widehat{\varphi},\psi)}_{\text{terms in } \widehat{\varphi}} + \underbrace{{}^2P_0^{0,0}(\{\mathrm{V}f - \widehat{y}\},\psi)}_{\text{terms in } f \ \& \ \widehat{y}}}^{\text{integrals with } \psi(t) \text{ in the integrand}} + \overbrace{\underbrace{{}^0\nabla P_0^{0,0}(\varphi) + {}^1\nabla P_0^{0,0}(\{\mathrm{V}f - \widehat{y}\})}_{\text{involving } \psi(0)}}^{\text{integrals}}$$

$$+ \overbrace{\underbrace{{}^0\Delta P_0^{0,0}}_{\text{independent of } \beta,\ \gamma} + \underbrace{{}^1\Delta P_0^{\beta,\gamma} + \overbrace{{}^2\Delta P_0^{\beta,\gamma}(\widehat{\varphi},\widehat{y})}^{\text{involving only } \psi(0)}}_{\text{involving } \beta \text{ and } \gamma}}^{\text{discrete terms: involving } \varphi(0) \text{ and } \psi(0);} . \tag{3.20}$$

Let us now consider the bilinear form

$$\mathcal{P}_\alpha^{\beta,\gamma}(\varphi,\psi) = {}^0P_\alpha^{0,0}(\varphi,\psi) + \Big({}^0\nabla P_0^{0,0}(\varphi) + {}^0\Delta P_0^{0,0} + {}^1\Delta P_0^{\beta,\gamma}\Big)\psi(0) \tag{3.21}$$

Lemma 3.4 *The bilinear form $\mathcal{P}_\alpha^{\beta,\gamma}(\varphi,\psi)$ is symmetric and positive semidefinite (definite when $\alpha>0$) on $\mathcal{F}$.*

Proof. If we return to the detailed expression for ${}^0P_\alpha^{0,0}(\varphi,\psi)$, ${}^0\nabla P_0^{0,0}(\varphi)$, ${}^0\Delta P_0^{0,0}$ and ${}^1\Delta P_0^{\beta,\gamma}$ we obtain the following expression for the bilinear form

$$\mathcal{P}_\alpha^{\beta,\gamma}(\varphi,\psi)=\alpha\int\limits_{-\tau}^{0}\varphi^{\mathsf{T}}(s)\psi(s)ds+\int\limits_{0}^{T}\int\limits_{-\tau}^{0}\int\limits_{-\tau}^{0}\varphi^{\mathsf{T}}(s)[B(s+\tau)]^{\mathsf{T}}[Y(s+\tau),t]^{\mathsf{T}}Y(\mu+\tau,t)B(\mu+\tau)\psi(\mu)d\mu dsdt+$$

$$\begin{aligned}&\int\limits_{0}^{T}\int\limits_{-\tau}^{0}\varphi^{\mathsf{T}}(s)[B(s+\tau)]^{\mathsf{T}}[Y(s+\tau,t)]^{\mathsf{T}}Y(0,t)\psi(0)dsdt+\\&\int\limits_{0}^{T}\int\limits_{-\tau}^{0}\varphi^{\mathsf{T}}(0)[Y(0,t)]^{\mathsf{T}}Y(s+\tau,t)B(s+\tau)\psi(s)dsdt+\\&\varphi^{\mathsf{T}}(0)(\beta+\gamma)\psi(0)+\int\limits_{0}^{T}\varphi^{\mathsf{T}}(0)[Y(0,t)]^{\mathsf{T}}Y(0,t)\psi(0)dt.\end{aligned}\tag{3.22}$$

It is easy to see from (3.22) that $\mathcal{P}_\alpha^{\beta,\gamma}(\varphi,\psi)=\mathcal{P}_\alpha^{\beta,\gamma}(\psi,\varphi)$. Now consider $\mathcal{P}_\alpha^{\beta,\gamma}(\varphi,\varphi)$. We can write

$$\mathcal{P}_\alpha^{\beta,\gamma}(\varphi,\varphi)=\alpha\int\limits_{-\tau}^{0}\varphi^{\mathsf{T}}(s)\varphi(s)ds\ +(\beta+\gamma)\varphi^{\mathsf{T}}(0)\varphi(0)$$

$$\int\limits_{0}^{T}\Big[Y(0,t)\varphi(0)+\int\limits_{-\tau}^{0}Y(s+\tau,t)B(s+\tau)\varphi(s)ds\Big]^{\mathsf{T}}\Big[Y(0,t)\varphi(0)+\int\limits_{-\tau}^{0}Y(s+\tau,t)B(s+\tau)\varphi(s)ds\Big]dt\geq 0.$$

Thus, positive definiteness of the bilinear form $\mathcal{P}_\alpha^{\beta,\gamma}(\varphi,\varphi)$ is obtained.

We can state the following result

Lemma 3.5 *The kernel $K_{\beta,\gamma}(t,s)$ defined by (3.17) is self-adjoint and positive semi-definite.*

Proof. Since the bilinear form (3.21) is positive semidefinite for all $\varphi(0)$ we can pick some particular value of the function φ at $t=0$, namely, let

$$\varphi(0)=-\mathcal{D}_{\beta,\gamma}^{-1}\int\limits_{-\tau}^{0}\int\limits_{s+\tau}^{T}[Y(0,\xi)]^{\mathsf{T}}Y(s+\tau,\xi)B(s+\tau)\varphi(s)d\xi ds,$$

and we define the bilinear form (3.22) after substitution of $\varphi(0)$ as $\widehat{\mathcal{P}}_\alpha^{\beta,\gamma}(\varphi,\varphi)$. We have

$$\widehat{\mathcal{P}}_\alpha^{\beta,\gamma}(\varphi,\varphi)=\alpha\int\limits_{-\tau}^{0}\varphi^{\mathsf{T}}(s)\varphi(s)ds+\int\limits_{0}^{T}\int\limits_{-\tau}^{0}\int\limits_{-\tau}^{0}\varphi^{\mathsf{T}}(s)[B(s+\tau)]^{\mathsf{T}}[Y(s+\tau),t]^{\mathsf{T}}Y(\mu+\tau,t)B(\mu+\tau)\varphi(\mu)d\mu dsdt-$$

$$\int\limits_{0}^{T}\int\limits_{-\tau}^{0}\varphi^{\mathsf{T}}(s)[B(s+\tau)]^{\mathsf{T}}[Y(s+\tau,t)]^{\mathsf{T}}Y(0,t)\mathcal{D}_{\beta,\gamma}^{-1}\int\limits_{-\tau}^{0}\int\limits_{\mu+\tau}^{T}[Y(0,\xi)]^{\mathsf{T}}Y(\mu+\tau,\xi)B(\mu+\tau)\varphi(\mu)d\xi d\mu dsdt+$$

$$\Bigg\{ \int_0^T \int_{-\tau}^0 \Big[\mathcal{D}_{\beta,\gamma}^{-1} \int_{-\tau}^0 \int_{\mu+\tau}^T [Y(0,\xi)]^{\mathsf{T}} Y(\mu+\tau,\xi) B(\mu+\tau)\varphi(\mu) d\xi d\mu\Big]^{\mathsf{T}} [Y(0,t)]^{\mathsf{T}} Y(s+\tau,t) B(s+\tau)\varphi(s) ds dt +$$

$$\Big[\mathcal{D}_{\beta,\gamma}^{-1} \int_{-\tau}^0 \int_{t+\tau}^T [Y(0,\mu)]^{\mathsf{T}} Y(t+\tau,\mu) B(t+\tau)\varphi(t) d\mu dt\Big]^{\mathsf{T}} \Big((\beta+\gamma)I + \int_0^T [Y(0,t)]^{\mathsf{T}} Y(0,t) dt\Big) \times$$

$$\times \mathcal{D}_{\beta,\gamma}^{-1} \int_{-\tau}^0 \int_{s+\tau}^T [Y(0,\xi)]^{\mathsf{T}} Y(s+\tau,\xi) B(s+\tau)\varphi(s) d\xi ds \Bigg\}$$

Let us consider last two terms in the expression for $\widehat{\mathcal{P}}_\alpha^{\beta,\gamma}(\varphi,\varphi)$ (within the figure parenthesis). Taking into account that $\mathcal{D}_{\beta,\gamma} = (\beta+\gamma)I + \int_0^T [Y(0,t)]^{\mathsf{T}} Y(0,t) dt$ we can write

$$-\int_0^T \int_{-\tau}^0 \Big\{\mathcal{D}_{\beta,\gamma}^{-1} \int_{-\tau}^0 \int_{\mu+\tau}^T [Y(0,\xi)]^{\mathsf{T}} Y(\mu+\tau,\xi) B(\mu+\tau)\varphi(\mu) d\xi d\mu\Big\}^{\mathsf{T}} [Y(0,t)]^{\mathsf{T}} Y(s+\tau,t) B(s+\tau)\varphi(s) ds dt +$$

$$\Big\{\mathcal{D}_{\beta,\gamma}^{-1} \int_{-\tau}^0 \int_{t+\tau}^T [Y(0,\mu)]^{\mathsf{T}} Y(t+\tau,\mu) B(t+\tau)\varphi(t) d\mu dt\Big\}^{\mathsf{T}} \int_{-\tau}^0 \int_{s+\tau}^T [Y(0,\xi)]^{\mathsf{T}} Y(s+\tau,\xi) B(s+\tau)\varphi(s) d\xi ds = 0.$$

Hence, we have

$$0 \le \widehat{\mathcal{P}}_\alpha^{\beta,\gamma}(\varphi,\varphi) = \alpha(\varphi,\varphi) + (K_{\beta,\gamma}\varphi,\varphi), \tag{3.23}$$

where

$$\begin{aligned} K_{\beta,\gamma} \equiv K_{\beta,\gamma}(t,s) = & \int_{\xi=t+\tau}^T [B(t+\tau)]^{\mathsf{T}} [Y(t+\tau,\xi)]^{\mathsf{T}} Y(s+\tau,\xi) B(s+\tau) d\xi + \\ & - \int_{\xi=t+\tau}^T \int_{\mu=s+\tau}^T [B(t+\tau)]^{\mathsf{T}} [Y(t+\tau,\xi)]^{\mathsf{T}} Y(0,\xi) \mathcal{D}_{\beta,\gamma}^{-1} [Y(0,\mu)]^{\mathsf{T}} Y(s+\tau,\mu) B(s+\tau) d\mu d\xi \end{aligned} \tag{3.24}$$

Since $\mathcal{D}_{\beta,\gamma}^{-1}$ is symmetric we can show, from (3.23) and (3.24), that $\widehat{\mathcal{P}}_\alpha^{\beta,\gamma}(\varphi,\psi)$ is symmetric and Lemma 3.5 is established.

3.3.2 Equivalence to an integral equation for φ

In this section we shall establish that the initial function which satisfies equations (3.5) also satisfies the integral equation (3.13).

According to Corollary 3.1 we may write the solution of the adjoint problem (3.2) in the form

$$x^{\mathsf{T}}(t) = \int_t^T [y(\varphi;s) - \widehat{y}(s)]^{\mathsf{T}} U(s,t) ds. \tag{3.25}$$

Using (3.8), we can write (3.25) as

$$x^{\mathsf{T}}(t) = \int_t^T \Big\{ \underbrace{\Big[Y(0,s)\varphi(0) + \int_{-\tau}^0 Y(\xi+\tau,s) B(\xi+\tau)\varphi(\xi) d\xi + \int_0^s Y(\xi,s) f(\xi) d\xi\Big]^{\mathsf{T}}}_{y(s)} - [\widehat{y}(s)]^{\mathsf{T}} \Big\} U(s,t) ds$$

or

$$x^{\mathsf{T}}(t) = \int_t^T \left[Y(0,s)\varphi(0) + \int_{-\tau}^0 Y(\xi+\tau,s)B(\xi+\tau)\varphi(\xi)d\xi \right]^{\mathsf{T}} U(s,t)ds +$$

$$\int_t^T \left[\int_0^s Y(\xi,s)f(\xi)d\xi - \widehat{y}(s) \right]^{\mathsf{T}} U(s,t)ds$$

(for $t \le T$), and therefore

$$x^{\mathsf{T}}(t+\tau) = \int_{t+\tau}^T \left[\int_0^s Y(\xi,s)f(\xi)d\xi - \widehat{y}(s) \right]^{\mathsf{T}} U(s,t+\tau)ds +$$

$$+ \int_{t+\tau}^T \left[Y(0,s)\varphi(0) + \int_{-\tau}^0 Y(\xi+\tau,s)B(\xi+\tau)\varphi(\xi)d\xi \right]^{\mathsf{T}} U(s,t+\tau)ds. \tag{3.26}$$

By virtue of (2.9e), $\varphi(t)$ satisfies

$$\alpha\varphi(t) + \left[x^{\mathsf{T}}(t+\tau)B(t+\tau)\right]^{\mathsf{T}} = \alpha\widehat{\varphi}(t) \quad \text{for } t \in [-\tau, 0).$$

Thus, using (3.26), we have

$$\alpha\varphi(t) + \left[\int_{t+\tau}^T \left[Y(0,s)\varphi(0) + \int_{-\tau}^0 Y(\xi+\tau,s)B(\xi+\tau)\varphi(\xi)d\xi \right]^{\mathsf{T}} U(s,t+\tau)ds B(t+\tau) \right]^{\mathsf{T}} =$$

$$\alpha\widehat{\varphi}(t) - \left[\int_T^{t+\tau} \left[\int_0^s Y(\xi,s)f(\xi)d\xi - \widehat{y}(s) \right]^{\mathsf{T}} U(s,t+\tau)ds B(t+\tau) \right]^{\mathsf{T}}, \quad t \in [-\tau, 0). \tag{3.27}$$

From the expression for $x^{\mathsf{T}}(t)$, we have

$$x^{\mathsf{T}}(0) = \int_0^T \left[\int_0^s Y(\xi,s)f(\xi)d\xi - \widehat{y}(s) \right]^{\mathsf{T}} U(s,0)ds +$$

$$+ \int_0^T \left[Y(0,s)\varphi(0) + \int_{-\tau}^0 Y(\xi+\tau,s)B(\xi+\tau)\varphi(\xi)d\xi \right]^{\mathsf{T}} U(s,0)ds. \tag{3.28}$$

Therefore, we can write $\varphi(0)$ as (see (2.9f))

$$(\gamma+\beta)\varphi(0) + \left[x^{\mathsf{T}}(0)\right]^{\mathsf{T}} = \beta\widehat{\varphi}(0) + \gamma\widehat{y}(0).$$

Here, using (3.28), we can write

$$(\beta+\gamma)\varphi(0) + \left[\int_0^T \left[Y(0,s)\varphi(0) + \int_{-\tau}^0 Y(\xi+\tau,s)B(\xi+\tau)\varphi(\xi)d\xi \right]^{\mathsf{T}} U(s,0)ds \right]^{\mathsf{T}} =$$

$$\beta\widehat{\varphi}(0) + \gamma\widehat{y}(0) - \left[\int_0^T \left[\int_0^s Y(\xi,s)f(\xi)d\xi - \widehat{y}(s) \right]^{\mathsf{T}} U(s,0)ds \right]^{\mathsf{T}}. \tag{3.29}$$

The latter equation allows us to eliminate $\varphi_\star(0)$ from (3.27) to obtain an integral equation for $\varphi_\star$. Taking the transposes we can write (3.29) in the form

$$(\beta+\gamma)\varphi(0)+\int_0^T [U(s,0)]^{\mathsf{T}}\left(Y(0,s)\varphi(0)+\int_{-\tau}^0 Y(\xi+\tau,s)B(\xi+\tau)\varphi(\xi)d\xi\right)ds=$$

$$\beta\widehat{\varphi}(0)+\gamma\widehat{y}(0)-\int_0^T \Big[U(s,0)\Big]^{\mathsf{T}}\left(\int_0^s Y(\xi,s)f(\xi)d\xi-\widehat{y}(s)\right)ds.$$

Using the equivalence $U(t,s)=Y(s,t)$ we obtain

$$(\gamma+\beta)I\varphi(0)+\int_0^T[Y(0,s)]^{\mathsf{T}}Y(0,s)\varphi(0)ds=-\int_0^T\int_{-\tau}^0[Y(0,s)]^{\mathsf{T}}Y(\xi+\tau,s)B(\xi+\tau)\varphi(\xi)d\xi ds+F_0(\widehat{\varphi}(0),\widehat{y},f),$$

where $F_0(\widehat{\varphi}(0),\widehat{y},f)=\beta\widehat{\varphi}(0)+\beta\widehat{y}(0)-\int_0^T[Y(0,s)]^{\mathsf{T}}\left(\int_0^s Y(\xi,s)f(\xi)d\xi-\widehat{y}(s)\right)ds$. We may therefore write

$$\mathcal{D}_{\beta,\gamma}\varphi(0)=-\int_0^T\int_{-\tau}^0[Y(0,s)]^{\mathsf{T}}Y(\xi+\tau,s)B(\xi+\tau)\varphi(\xi)d\xi ds+F_0(\widehat{\varphi}(0),\widehat{y},f). \tag{3.30}$$

Using the equation (3.30), we may write (3.27) as

$$\alpha\varphi(t)-\int_{t+\tau}^T[B(t+\tau)]^{\mathsf{T}}[Y(t+\tau,s)]^{\mathsf{T}}Y(0,s)\left(\mathcal{D}^{-1}\int_0^T\int_{-\tau}^0[Y(0,\mu)]^{\mathsf{T}}Y(\xi+\tau,\mu)B(\xi+\tau)\varphi(\xi)d\xi d\mu\right)ds+$$

$$\int_{t+\tau}^T[B(t+\tau)]^{\mathsf{T}}[Y(t+\tau,s)]^{\mathsf{T}}Y(0,s)\mathcal{D}_{\beta,\gamma}^{-1}F_0(\widehat{\varphi}(0),\widehat{y},f)ds+$$

$$\int_{t+\tau}^T\int_{-\tau}^0[B(t+\tau)]^{\mathsf{T}}[Y(t+\tau,s)]^{\mathsf{T}}Y(\xi+\tau,s)B(\xi+\tau)\varphi(\xi)d\xi ds=$$

$$\alpha\widehat{\varphi}(t)-\int_{t+\tau}^T[B(t+\tau)]^{\mathsf{T}}[Y(t+\tau,s)]^{\mathsf{T}}\left(\int_0^s Y(\xi,s)f(\xi)d\xi-\widehat{y}(s)\right)ds,$$

or

$$\alpha\varphi(t) + \int_{t+\tau}^{T}\int_{-\tau}^{0} [B(t+\tau)]^{\mathsf{T}}\,[Y(t+\tau,s)]^{\mathsf{T}}\,Y(\xi+\tau,s)B(\xi+\tau)\varphi(\xi)d\xi ds+$$

$$-\int_{s=t+\tau}^{s=T}\int_{\mu=0}^{\mu=T}\Big\{\int_{-\tau}^{0}[B(t+\tau)]^{\mathsf{T}}\,[Y(t+\tau,s)]^{\mathsf{T}}\,Y(0,s)\mathcal{D}_{\beta,\gamma}^{-1}[Y(0,\mu)]^{\mathsf{T}}\,Y(\xi+\tau,\mu)B(\xi+\tau)\varphi(\xi)d\xi\Big\}d\mu ds =$$

$$\alpha\widehat{\varphi}(t) - \int_{t+\tau}^{T}[B(t+\tau)]^{\mathsf{T}}\,[Y(t+\tau,s)]^{\mathsf{T}}\,Y(0,s)\mathcal{D}_{\beta,\gamma}^{-1}F_0(\widehat{\varphi}(0),\widehat{y},f)ds+$$

$$-\int_{t+\tau}^{T}[B(t+\tau)]^{\mathsf{T}}\,[Y(t+\tau,s)]^{\mathsf{T}}\left(\int_{0}^{s}Y(\xi,s)f(\xi)d\xi - \widehat{y}(s)\right)ds. \tag{3.31}$$

Remark 3.5 *From the Definition 3.1 we have $Y(\xi+\tau,\mu)=0$ for $\mu<\xi+\tau$. Therefore we may change the lower limit in the third term in the equation (3.31) to $\mu=\xi+\tau$.*

We have derived (3.31) which is the required integral equation formulation and Theorem 3.2 is now established.

Remark 3.6 *The solution $\varphi_\star(t)$ of (3.13) is unique in $L_2[-\tau,0]$ and hence in $\mathcal{F}_n$. Using a similar method to the one described above we can show that in finite dimensional subspace $\mathcal{F}_n$ of $\mathcal{F}$ the integral equation has the form (3.13), where all functions are treated as functions from $\mathcal{F}_n$.*

3.4 An iteration defining φ_n expressed via an integral equation

We shall consider the convergence of the iteration described in §2.2 by studying the iteration

$$\frac{\varphi_{n+1}(t)-\varphi_n(t)}{\delta_n} = \mathfrak{g}_\alpha^{\beta,\gamma}(t) - \Big(\alpha\varphi_n(t) + \int_{-\tau}^{0}K_{\beta,\gamma}(t,s)\varphi_n(s)ds\Big). \tag{3.32}$$

This iteration is based upon the integral equation (3.13).It is, in a certain sense, linked to the "method of sequential approximation" [14].

In (3.32), $K_{\beta,\gamma}(t,s)$ has been shown to be symmetric and positive-definite; the corresponding integral operator on $L_2[-\tau,0]$ is bounded, self-adjoint, and positive-definite. We state the following result.

Lemma 3.6 *The iteration (3.32) is equivalent to the iteration (2.10) described in §2.2. (For a given φ_0, the two sequences $\{\varphi_n\}$ are identical.)*

Proof. From (2.10e), the functions defined by the iteration (2.1) satisfy the relation

$$\frac{\varphi_{n+1}(t)-\varphi_n(t)}{\delta_n} = \alpha(\varphi_n(t)-\widehat{\varphi}(t)) + [B(t+\tau)]^{\mathsf{T}}\,x_n(t+\tau) \quad \text{for } t\in[-\tau,0)$$

and we have shown in §3.3.2 that

$$\alpha(\varphi_n(t) - \widehat{\varphi}(t)) + [B(t+\tau)]^{\mathsf{T}} x_n(t+\tau) = \alpha\varphi_n(t) + \int_{-\tau}^{0} K_{\beta,\gamma}(t,s)\varphi_n(s)ds - \mathfrak{g}_\alpha^{\beta,\gamma}(t),$$

so the result is immediate.

Theorem 3.3 (Convergence) *Suppose $\rho(K_{\beta,\gamma})$ is the spectral radius of the integral operator $K_{\beta,\gamma}$ on $L_2[-\tau,0]$ having the kernel $K_{\beta,\gamma}(t,s)$. Then, a sufficient condition for the iteration (2.10) in Definition 2.1 to converge in the mean-square norm is*

$$\delta_n \leq \frac{2}{\max\left(\alpha, \rho(K_{\beta,\gamma})\right)}, \quad \textit{for all } n. \tag{3.33}$$

Proof. We shall write $\mathfrak{L}_\alpha^{\beta,\gamma}\varphi(t) = \alpha\varphi(t) + \int_{-\tau}^{0} K_{\beta,\gamma}(t,s)\varphi(s)ds$ and the operator $\mathfrak{L}_\alpha^{\beta,\gamma}$ on $L_2[-\tau,0]$ inherits self-adjointness and (with $\alpha > 0$) positive-definiteness from the corresponding properties of the integral operator $K_{\beta,\gamma}$. For a sequence $\{\delta_n\}$ with $\delta_n > 0$ for all n, we can write the iteration process (3.32) in the form

$$\frac{\varphi_{n+1}(t) - \varphi_n(t)}{\delta_n} = \mathfrak{g}_\alpha^{\beta,\gamma}(t) - \mathfrak{L}_\alpha^{\beta,\gamma}\varphi_n(t). \tag{3.34}$$

Let $\varphi_\star$ be the solution of the equation $\mathfrak{L}_\alpha^{\beta,\gamma}\varphi_\star(t) = \mathfrak{g}(t)$ and let us define $\varepsilon_{n+1} = \varphi_{n+1} - \varphi_\star$. Then, according to (3.34), we have the relation

$$\varepsilon_{n+1} = (I - \delta_n \mathfrak{L}_\alpha^{\beta,\gamma})\varepsilon_n,$$

and

$$\varepsilon_{n+1} = \prod_{i=0}^{n}(I - \delta_n \mathfrak{L}_\alpha^{\beta,\gamma})\varepsilon_0. \tag{3.35}$$

The iteration (3.34) converges in the mean-square norm if $\|\varepsilon_n\|_2 \to 0$ as $n \to \infty$. From (3.35) we have

$$\|\varepsilon_{n+1}\|_2 \leq \left\| \prod_{i=0}^{n}(I - \delta_n \mathfrak{L}_\alpha^{\beta,\gamma}) \right\|_2 \|\varepsilon_0\|_2 \leq \prod_{i=0}^{n} \left\| (I - \delta_n \mathfrak{L}_\alpha^{\beta,\gamma}) \right\|_2 \|\varepsilon_0\|_2.$$

Thus, a sufficient condition for convergence of this iteration is

$$\|I - \delta_n \mathfrak{L}_\alpha^{\beta,\gamma}\|_2 \leq \vartheta < 1 \quad \text{for all } n. \tag{3.36}$$

Given the properties of $\mathfrak{L}_\alpha^{\beta,\gamma}$ on $L_2[-\tau,0]$, we have $\|\mathfrak{L}_\alpha^{\beta,\gamma}\|_2 = \max_r \kappa_r$ (the spectral radius $\rho(\mathfrak{L}_\alpha^{\beta,\gamma})$), where $\{\kappa_r\}_{r\geq 0}$ are the positive eigenvalues of $\mathfrak{L}_\alpha^{\beta,\gamma}$. Indeed, $\kappa_r = \alpha + \varkappa_r$, where $\{\varkappa_r\}_{r\geq 0}$ are the positive eigenvalues of $K_{\beta,\gamma}$. Then condition (3.36) becomes

$$\max_r |1 - \delta_n\alpha - \delta_n\varkappa_r| < 1.$$

We have $1-\delta_n\alpha-\delta_n\varkappa_r \in [1-\delta_n\alpha-\delta_n\rho(K_{\beta,\gamma}), 1-\delta_n\alpha) \subseteq (-1,1)$ provided $1-\delta_n\alpha-\delta_n\rho(K_{\beta,\gamma}) > -1$ and Theorem 3.3 established.

Remark 3.7 *In general, the explicit form of the kernel $K_{\beta,\gamma}(t,s)$ is unknown and we cannot implement the iteration process for the integral equation itself; we use the iterative process (2.10).*

4 Extensions to nonlinear problems

In a later report, we consider a non-linear problem in the form

$$\frac{\partial y(t)}{\partial t} = f(t, y(t), y(t-\tau)), \quad t \in [0, T] \tag{4.1a}$$

with an initial condition

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

Observe that for a general nonlinear problem, various uniqueness conditions cannot be assumed. In our later report, we shall extend our result from §2 and formulate a "data assimilation problem" (1.6) for a nonlinear system (4.1). (The work is also to be reported in the PhD thesis of E. Parmuzin.)

References

[1] Agoshkov, V. I. and Marchuk, G. I., On the solvability and numerical solution of data assimilation problems, Russian J. Numer. Anal. Math. Modelling, **8**, (1993), pp. 1–16,

[2] Agoshkov, V. I., Parmuzin, E. I., Shutyaev, V. P. and Bardos, C., Numerical analysis of iterative algorithms for an inverse boundary transport problem, Math. Models Methods Appl. Sci., Mathematical Models & Methods in Applied Sciences, **10**, (2000), pp. 11–29,

[3] Agoshkov, V. I. Optimal control methods in inverse problems and computational processes, J. Inverse Ill-Posed Probl., **9**, (2001), pp. 205–218.

[4] Baker C.T.H, Bocharov G.A. and Paul C.A.H. Mathematical Modelling of the Interleukin-2 T-Cell System: A Comparative Study of Approaches based on Ordinary and Delay Differential Equations, J. of Theoretical Medicine, **2**, (1997), pp. 117-128.

[5] Baker C.T.H, Bocharov G.A. and Paul C.A.H. and Rihan F. Modelling and Analysis of Time-Lags in Some Basic Patterns of Cell Proliferation, J. of Mathematical Biology, **37**, (1998), pp. 341-371.

[6] Baker C.T.H, Bocharov G.A., Paul C.A.H. and Rihan F. Computational modelling with functional differential equations: identification, selection, and sensitivity. *Submitted for publication.*

[7] Baker C.T.H., Parmuzin E.I, A guide tour of VPF for continuous and discretized DDE, NA Report 430, Manchester, 2003.

[8] H. T. Banks, J. A. Burns and E. M. Cliff, *Parameter estimation and identification for systems with delays*, SIAM J. Control Optim. , **19**, No. 6, (1981), 791–828.

[9] Bellman, R. and Cooke, K.L., Differential-Difference Equation, Academic Press, New York, (1963). Lib. Congress No. 61-18904.

[10] H. W. Engl, and C. W. Groetsch (eds.), *Inverse and ill-posed problems*, Academic, Boston, 1987

[11] H. W. Engl, M. Hanke and A. Neubauer, *Regularization of inverse problems*, Kluwer Academic, Dordrecht, 1996. ISBN 0-7923-4157-0

[12] Halanay, A., Differential Equations: stability, oscillation, time lag, Academic Press, New York, (1966), Lib. Congress No. 65-25005.

[13] Hale, J., Theory of Functional Differential Equations, Springer-Verlag, New York, (1977), ISBN 0-387-90203-1.

[14] M. M. Lavrent'ev and L. Ya. Savel'e, *Linear operators and ill-posed problems*, Consultants Bureau, New York, 1995.

[15] F. Liu and M. Z. Nashed, *Convergence of regularized solutions of nonlinear ill-posed problems with monotone operators*, [16], (1996), 353–361.

[16] P. Marcellini, G. G. Talenti and E. Vesentini (eds.), *Partial differential equations and applications* Lecture Notes in Pure and Applied Mathematics, 177, Marcel Dekker, New York, 1996.

[17] Lakshmikantham, V. and Deo S.G., Method of variation of parameters for dynamic system. Volume I., Gordon and Breach Science Publishers, New Delhi, 1998, ISBN 90-5699-160-4.

[18] Le Dimet, F.-X. and Talagrand, O., Variational algorithms for analysis and assimilation of meteorological observations, Tellus, **38A**, (1986), pp. 97-110.

[19] Parmuzin, E. I. and Shutyaev, V. P., Numerical algorithms for solving a problem of data assimilation, Zh. Vychisl. Mat. Mat. Fiz., **37**, (1997), pp. 816–827.

[20] Parmuzin, E. I. and Shutyaev, V. P., Numerical analysis of iterative methods for solving evolution data assimilation problems, Russian J. Numer. Anal. Math. Modelling, **14**, (1999), pp. 275–289.

[21] Pokrovski, O. M., An optimal 4-D assimilation of conventional and satellite data for meteorological fields analysis, Sov.Meteorology and Hydrology, Allerton Press Inc., New York, **8**, (1974), pp. 29-36.

[22] Sasaki, Y., Some basic formalisms on numerical variational analysis, Mon. Wea. Review, **98**, (1970), pp. 875-883.

[23] Shutyaev, V., Control operators and iterative algorithms in variational data assimilation problems, J. Inverse Ill-Posed Probl., **9**, (2001), pp. 177–188.

[24] Verduyn Lunel, Sjoerd M. Parameter identifiability of differential delay equations. *Int. J. Adapt. Control Signal Process* 15, No.6, 655-678 (2001).

[25] U. Tautenhahn, *On the method of Lavrentiev regularization for nonlinear ill-posed problems*, Inverse Problems, **18**, (2002), 191–207.

[26] A. N. Tikhonov and V. Y. Arsenin, *Solution of ill-posed problems*, Wiley, New York, 1977.

[27] G. Vainikko, *On the optimality of regularization methods*, [10], (1987), 77–95.

A Appendix

A.1 An another proof of Lemma 3.1.

We can write the solution $z(\psi,t)$ of this system (2.4) as in (3.8), where the last term vanishes because $z(\psi,t)$ is a solution of the homogeneous problem. We have

$$z(t) = Y(0,t)\psi(0) + \int_{-\tau}^{0} Y(s+\tau,t)B(s+\tau)\psi(s)ds. \tag{A.1}$$

Let us consider the first variation of the functional $S_\alpha^{\beta,\gamma}(\varphi)$ (2.6b). When using (A.1) we obtain

$$\begin{aligned} P_\alpha^{\beta,\gamma}(\varphi,\psi) = \alpha \int_{-\tau}^{0} [\varphi(t) - \widehat{\varphi}(t)]^{\mathsf{T}} \psi(t)dt + \beta[\varphi(0) - \widehat{\varphi}(0)]^{\mathsf{T}} \psi(0) + \gamma[y(\varphi;0) - \widehat{y}(0)]^{\mathsf{T}} \psi(0) + \\ + \int_{0}^{T} [y(\varphi;t) - \widehat{y}(t)]^{\mathsf{T}} \left(Y(0,t)\psi(0) + \int_{-\tau}^{0} Y(s+\tau,t)B(s+\tau)\psi(s)ds \right) dt. \end{aligned} \tag{A.2}$$

Using (3.12) we can write the solution of the adjoint equation (3.2) in the form

$$x^{\mathsf{T}}(t) = \int_{t}^{T} [y(\varphi;s) - \widehat{y}(s)]^{\mathsf{T}} U(s,t)ds. \tag{A.3a}$$

Therefore, for $t=0$ we have

$$x^{\mathsf{T}}(0) = \int_{0}^{T} [y(\varphi;s) - \widehat{y}(s)]^{\mathsf{T}} U(s,0)ds, \tag{A.3b}$$

and for $t = t+\tau$

$$x^{\mathsf{T}}(t+\tau) = \int_{t+\tau}^{T} [y(\varphi;s) - \widehat{y}(s)]^{\mathsf{T}} U(s,t+\tau)ds.$$

We may write $x^{\mathsf{T}}(t+\tau)$ in the form

$$x^{\mathsf{T}}(t+\tau) = \int_{0}^{T} [y(\varphi;s) - \widehat{y}(s)]^{\mathsf{T}} U(s,t+\tau)ds - \int_{0}^{t+\tau} [y(\varphi;s) - \widehat{y}(s)]^{\mathsf{T}} U(s,t+\tau)ds.$$

Since we have $U(s,t) = 0$ for $s < t$, the last term vanishes and we have

$$x^{\mathsf{T}}(t+\tau) = \int_{0}^{T} [y(\varphi;s) - \widehat{y}(s)]^{\mathsf{T}} U(s,t+\tau)ds. \tag{A.3c}$$

Thus, using (A.3b) and (A.3c), we may write (A.2) as

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