University of Chester

Author(s): Christopher T H Baker ; Gennady Bocharov ; F A R Rihan

Title: On some aspects of casual and neutral equations used in mathematical modelling

Date: 2007

Example citation: Baker, C. T. H., Bocharov, G., & Rihan, F. A. R. (2007). *On some aspects of casual and neutral equations used in mathematical modelling* (Applied Mathematics Group Research Report, 2007 : 2). Chester, United Kingdom: University of Chester.

Version of item: Published version

Available at: http://hdl.handle.net/10034/346437

University of Chester

Department of Mathematics

Applied Mathematics Group
Research Report

2007 Series

ON SOME ASPECTS OF CAUSAL & NEUTRAL EQUATIONS USED IN MATHEMATICAL MODELLING

Christopher T H Baker

Gennady A Bocharov

Eugene I Parmuzin

Fathalla A Rihan

Report 2007:2

ON SOME ASPECTS OF CAUSAL & NEUTRAL EQUATIONS USED IN MATHEMATICAL MODELLING

Christopher T H Baker[a,b] *,

[a] *Department of Mathematics, University of Chester, Chester CH1 4BJ;*
[b] *Emeritus professor, The University of Manchester, M13 9PL, UK.*

Gennady Bocharov[c 1], Eugene Parmuzin[c 2],

[c] *Institute of Numerical Mathematics, Russian Academy of Science, Moscow*

Fathalla Rihan [d]

[d] *Faculty of Science, Helwan University, Egypt*

Abstract

The problem that motivates the considerations here is the construction of mathematical models of natural phenomena that depend upon past states. The paper divides naturally into two parts: in the first, we expound the inter-connection between ordinary differential equations, delay-differential equations, neutral delay-differential equations and integral equations (with emphasis on certain linear cases). As we show, this leads to a natural hierarchy of model complexity when such equations are used in mathematical and computational modelling, and to the possibility of reformulating problems either to facilitate their numerical solution or to provide mathematical insight, or both. Volterra integral equations include as special cases the others we consider. In the second part, we develop some practical and theoretical consequences of results given in the first part. In particular, we consider various approaches to the definition of an adjoint, we establish (notably, in the context of sensitivity analysis for neutral delay-differential equations) rôles for well-defined adjoints and 'quasi-adjoints', and we explore relationships between sensitivity analysis, the variation of parameters formulae, the fundamental solution and adjoints.

Key words: Computational modelling, Analysis of models, Delay & neutral delay-differential equations, Volterra integral equations, Fundamental solutions, Variation of parameters, Adjoints, Resolvents, Sensitivity, Model selection

---

* *E-mail contact:* `cthbaker@na-net.ornl.gov` `c.baker@chester.ac.uk`.

[1] Supported in part by a Leverhulme visiting professorship at the University of Chester, UK and in part by the Russian Foundation for Basic Research; now honorary visiting professor at Chester.

[2] Supported by INTAS Fellowship YS04-83-2818 held at the University of Chester, UK and in part by the Russian Foundation for Basic Research.

1 INTRODUCTION

We have previously (severally, jointly, and with others – see, *e.g.* [1,4–6,8]) been involved in the determination of computational models that describe mathematically the evolution of phenomena that depend upon past states. However, adoption of a multidisciplinary and holistic approach has here been restrained by publication requirements, and we concentrate on the mathematical issues, with the motivation indicated but relegated to a lesser rôle.

In the context of this paper, a mathematical model comprises a set of equations, and possibly constraints, that purport to generate insight into some specified phenomena. The models we have employed have been, in a general sense, differential equations [DEs], ordinary differential equations [ODEs], delay-differential equations [DDEs], neutral delay differential equations [NDDEs] *cf.* [2], or related integral and integro-differential equations; see §2. We consider such equations here.

Our material divides into two parts, of which the first comprises a review of a variety of models, and explores interconnections between different types of models. This has a bearing on theoretical insight and the use of different methods for computational solution, and we hope this part will be of wide interest. The second part presents a search for a theory relating variation of parameters and adjoint theory to the issue of sensitivity, and an indication of its subsequent application. Here, we address a number of issues concerning the analysis and numerics of NDDEs that previously appeared uncertain, because of the complexity of NDDEs. For the second part we target, as a readership, a mathematical audience.

Indeed, our main motivation lies in the construction of *parametrized* models that provide a quantitatively consistent simulation of observed data. This involves a formulation of a family of putative models and the determination (computationally) of an actual model that is, in some sense, best of those available for selection. Frequently, this selection process can be effected using methods based on sensitivity and adjoint theory. The need to compute numerical solutions arises at every stage in computational modelling; The models can therefore be quite complex, which lessens the need to sacrifice realism for simplicity [3] .

For some modellers, parsimony [4] is important in the choice of a model. The material in §§2 – 3 indicates a natural hierarchy of model equations that can be used in an informal ranking of parsimony whereby equations at the lower end

[3] Models whose choice is motivated by an ability to treat them analytically can lead to equations inadequate for the purpose.
[4] Parsimony is a complex subject: a literature search for “Occams razor” will yield some further reading.

of the hierarchy can be rewritten as a special case of more general equations higher up the hierarchy. The different types of equations listed present differing challenges both numerically and theoretically. We give particular attention to NDDEs [5] (which are commonly termed either explicit or implicit but can sometimes be transformed from one type to another [§3.1]). We emphasize a rôle in the analysis played by Volterra integral equations [§3.2].

It is both of theoretical and of practical interest to know the sensitivity of a solution of a model [§5] to a change in the parameters. In particular, sensitivity with respect to changes in the model is an important topic in model selection; we therefore explore variation of parameters formulae, and their relation to resolvent and adjoint theory. To summarize: in §2 we introduce various classes of equation; we indicate in §3 how equations can be rewritten as a different type of equation subject to certain conditions; in §4 we concentrate on mathematical concepts such as adjoint operators and equations, fundamental solutions and solvent equations; finally, we are concerned with sensitivity theory in §5.

2 A VARIETY OF NONLINEAR HEREDITARY MODELS (A REVIEW)

Causal models can be classified as discrete, continuous, or hybrid. ODEs, DDEs, NDDEs and Volterra integral equations [VIEs] (and the discrete analogues – or similar partial differential equations, which we do not discuss here) provide classes of equations competing to be chosen as models, but there is a rich variety of possibilities going beyond these. Our emphasis in the paper on restricted classes of equations involves a simplification since, in practical modelling situations, parametrized models are actually based on combinations of such equations. For example, Marchuk's nonlinear equations in [30],

$$y_1'(t) = \beta y_1(t) - \gamma y_2(t) y_1(t), \quad y_2'(t) = \rho y_3(t) - \gamma\eta y_1(t) y_2(t) - \mu_2 y_2(t),$$
$$y_3'(t) = \alpha\xi(y_4(t)) y_2(t-\tau) y_1(t-\tau) - \mu_3\{y_3(t) - y_3^*\}, \quad y_4'(t) = \varsigma y_1(t) - \mu_4 y_4(t),$$

are of varying types, some ODEs and some DDEs. Some modellers emphasize the need for a model that is dimensionless and scaled (*cf.* [33]).

To focus attention, we shall be considering the formal, parametrized, model,

$$\frac{d}{dt}\{y(t)-cy(t-\tau)\} = ay(t)+by(t-\tau) \ (t \in [0,T]); \ y(t) = \phi(t) \ (t \in [-\tau,0], \tau \geq 0). \tag{2.1}$$

This is a linear scalar NDDE and selected parameters [6] $[a,b,c,\tau]^{\mathrm{T}}$ and ϕ define an actual (as opposed to a formal) model with solution $y(t) \equiv y([a,b,c,\tau]^{\mathrm{T}}, \phi; t)$.

[5] DDEs and ODEs are special cases of NDDEs. In Remark 3.1 we introduce a novel type of delay equation derived from an NDDE.

[6] $\mathbb{R}^N$ is the space of N-dimensional real column vectors and the notation $^{\mathrm{T}}$ denotes transpose: $y^{\mathrm{T}}(t)$ is $[y(t)]^{\mathrm{T}}$. The prime ($'$) denotes differentiation; see Remark 2.2 for the unrelated notation y_t.

More generally, supposing $\tau > 0$, the solution of the non-autonomous *system* of NDDEs

$$\frac{d}{dt}\{y(t) - C(t)y(t-\tau)\} - \{A(t)y(t) + B(t)y(t-\tau)\} = f(t) \ (t \in [0,T]) \quad (2.2)$$

with $y(t) = \phi(t)$ ($t \in [-\tau, 0]$), depends on the functions $A(\cdot)$, $B(\cdot)$, and $C(\cdot)$ with values in $\mathbb{R}^{N\times N}$, on $\tau \geq 0$, and on the vector-valued functions $f(\cdot)$ and $\phi(\cdot)$ (each potentially serving as a parameter). ODEs and DDEs result on setting $c = b = \tau = 0$ or $c = 0$ and $\tau > 0$ in (2.1) or alternatively $C(t) = B(t) = 0$ and $\tau = 0$ or $C(t) = 0$ and $\tau > 0$ in (2.2).

Remark 2.1 (*a*) The coefficients $A(t)$, $B(t)$, $C(t)$, and inhomogeneous term $f(t)$ in (2.2) often depend upon a vector parameter p, so that $A(t) = A(\mathrm{p}; t)$, $B(t) = B(\mathrm{p}; t)$, $C(t) = C(\mathrm{p}; t)$, $f(t) = f(\mathrm{p}; t)$. The number of parameter components is a poor measure of relative parsimony across models of different types. Regarding our motivation, it is important to note that most forms of modelling involve, at some stage, the determination of an actual parameter $\mathrm{p}_\star$ that minimize an objective function $\Phi(\mathrm{p})$ defined in terms of observation data [8]; see (5.12).

2.1 DDEs, Explicit NDDEs, Implicit NDDEs & related equations

Suppose $y : [0,T] \to \mathbb{R}^N$; equations on $[0,\infty)$ correspond to the absence of a bound on T. A classical system of ODEs with a specified initial value can be represented as

$$y'(t) = \mathfrak{f}\Big(t, y(t)\Big) \quad (t \in [0,T]) \text{ with initial condition } y(0) = y_0. \qquad (2.3\text{a})$$

Here, $\mathfrak{f} : [0,T] \times \mathbb{R}^N \to \mathbb{R}^N$; to write (2.3a) as $y' = \mathfrak{f}(t,y)$ or $y'(t) = \mathfrak{f}(t,y)$ would be a mild abuse of notation. On integration, the Cauchy problem (2.3a) assumes the form

$$y(t) = y_0 + \int_0^t \mathfrak{f}\Big(s, y(s)\Big)\mathrm{d}s \quad (t \in [0,T]) \qquad (2.3\text{b})$$

which is a *Volterra integral equation*. An absolutely continuous function y that satisfies (2.3b) satisfies the differential equation in (2.3a) almost everywhere on $[0,T]$ and $y(0) = y_0$. Closely related to (2.3b) are general Volterra integral equations, *e.g.*,

$$y(t) = g(t) + \int_0^t \mathsf{k}\Big(t, s, y(s)\Big)\mathrm{d}s \quad (t \geq 0) \text{ where, automatically, } y(0) = g(0). \qquad (2.4)$$

By integration (*cf.* (2.3b)) we can recast Volterra integro-differential equations such as

$$y'(t) = \gamma\Big(t, y(t), \int_0^t \mathsf{k}(t, s, y(s))\mathrm{d}s\Big) \quad (t \in [0,T]), \quad y(0) = y_0, \qquad (2.5)$$

as Volterra integral equations. Conversely, given sufficient differentiability, differentiation of a Volterra integral equation yields a Volterra integro-differential equation.

Now suppose that $\hat{\tau}(t) \geq 0$ denotes a (time-)lag; if $\hat{\tau}(t)$ is constant, we write its value as "τ". *Explicit NDDEs* with one time-dependent lag $\hat{\tau}(t)$ commonly have the form

$$y'(t) = f\Big(t, y(t), y(t-\hat{\tau}(t)), y'(t-\hat{\tau}(t))\Big) \; (t \in [0,T]; \; f : [0,T] \times \mathbb{R}^N \times \mathbb{R}^N \times \mathbb{R}^N \to \mathbb{R}^N). \tag{2.6}$$

The corresponding (so-called) *implicit NDDEs* commonly have the form

$$\frac{d}{dt}\Big\{y(t) - g_*\Big(t, y(t), y(t-\hat{\tau}(t))\Big)\Big\} = f_*\Big(t, y(t), y(t-\hat{\tau}(t))\Big). \tag{2.7}$$

To determine a solution of a DDE or NDDE we require an initial condition of the form

$$y(t) = \phi(t), \quad \text{for } t \in [-\tau_{\min}, 0] \text{ (for suitable } \tau_{\min} \geq 0). \tag{2.8}$$

(For equations with a single lag $\hat{\tau}(t)$, $\tau_{\min} = \inf_{t\in[0,T]} \hat{\tau}(t)$.) Dependence of $y(\cdot)$ on ϕ can be denoted by writing $y(\phi;\cdot)$. The smoothness of $y(\phi;\cdot)$ is an important issue in the numerics [2,11] of DDEs and NDDEs (the solution or its derivatives may inherit jumps from the behaviour of the initial function, through the lag). In consequence, a distinction between two-sided and right-hand derivatives may be necessary. We may suppose that $\phi(t)$ is, on $[-\tau_{\min}, 0]$, (a) bounded (b) piecewise-continuous, (c) continuous, (d) differentiable, or (e) continuously differentiable. For (2.6) we need differentiability; for (2.7) we assume condition (b) for simplicity.

Implicit NDDEs (2.7) can be regarded as a system of constrained DDEs:

$$u'(t) = f_*\Big(t, y(t), y(t-\hat{\tau}(t))\Big), \quad u(t) = y(t) - g_*\Big(t, y(t), y(t-\hat{\tau}(t))\Big).$$

The expression $y(t) - g_*\big(t, y(t), y(t-\hat{\tau}(t))\big)$ is called *the difference part* (*cf.* the difference equation (3.9) below). Solvability of (2.7) was discussed in [7,9] with the condition that equations $u = g(t,u,v) + w$ can be solved in the form $u = \gamma(t,v,w)$ where γ is continuous and satisfies a certain Lipschitz condition.

For cases with multiple lags, refer to, *e.g.*, [24,25]. A DDE with a single lag, say

$$y'(t) = f_\circ\Big(t, y(t), y(t-\hat{\tau}(t))\Big) \tag{2.9}$$

is a special case of (2.6) and of (2.7). In contrast, the integro-differential equation in (2.5) generalizes, on replacing integrals by Riemman-Stieltjes integrals, to a more general form

$$y'(t) = \gamma\Big(t, y(t), \int_0^t \mathsf{k}_\sharp\big(t, s, y(s)\big)\mathrm{d}s + \sum_{\hat{\tau}_i(t)\in[0,t]} \mathsf{k}_\sharp\big(t, \hat{\tau}_i(t), y(\hat{\tau}_i(t))\big)\Big) \ (t \in [0,T]), \tag{2.10}$$

and when only the sum is present we obtain a novel form of DDE; see Remark 3.1.

Remark 2.2 In past decades (starting with Krasovskii's Russian school and Hale and his co-workers), the analysis of hereditary problems developed from the perspective of functional DEs. In common usage v_t denotes (given a function v with a suitable domain, and $\tau > 0$) the function such that $v_t(s) = v(t+s)$ for $s \in [-\tau, 0]$, for appropriate t. An implicit NDDE (2.7) with a *fixed* lag $\tau > 0$,

$$\frac{d}{dt}\Big\{y(t) - g_*\Big(t, y(t), y(t-\tau)\Big)\Big\} = f_*\Big(t, y(t), y(t-\tau)\Big), \tag{2.11}$$

can be seen as defining the evolution of $\{y_t \big| t \geq 0\}$ with $y_0 = \phi$. Hale and contemporaries (see [20,21,23] and the original papers) considered functional DEs $\frac{d}{dt}D_\sharp(t, y_t) = f_\sharp(t, y_t)$, where $D_\sharp$ and $f_\sharp$ are continuous mappings of $\mathbb{R}_+ \times C([-\tau, 0] \to \mathbb{R}^N)$ into $\mathbb{R}^N$. If the *difference operator* $D_\sharp$ is atomic at zero (see [20,23–25]), the initial value problem for the above equation is well-posed, and existence and uniqueness theorems have been given. (The terminology 'atomic', 'non-atomic', appears to be related to measure theory [21, p.69].) The class includes *functional differential equations of retarded and neutral type*

$$\frac{d}{dt}\{y(t) - g_\sharp(t, y_t)\Big\} = f_\sharp(t, y_t) \ (t \geq 0), \text{ with } y_0 = \phi. \tag{2.12}$$

This equation can be rewritten [7] as $\Big\{y(t) - g_\sharp(t, y_t)\Big\} = \int_0^t f_\sharp(s, y_s)\mathrm{d}s + \{\phi(0) - g_\sharp(0, \phi)\}$.

Remark 2.3 Discrete analogues of the foregoing equations also have a direct rôle in modelling, and an indirect rôle in numerical simulation of the models in this section. One can construct discrete parallels for many of the results given for DEs and VIEs in the remainder of the paper.

2.2 Causal, or Volterra, equations

PROPOSITION 2.1 *ODEs are special cases of DDEs, which are special cases of NDDEs. All the equations displayed above can be recast as Volterra integral equations.*

It follows from Proposition 2.1 that a facility in the theory and numerics of Volterra integral (and integro-differential) equations is an advantage. The close relationship between various types of models, and the possibility of regarding many of them as special cases of Volterra integral equations (or, in the discrete case, Volterra summation equations) opens the study of DDEs, NDDEs, etc., to the application of numerical and theoretical techniques used to study Volterra integral (or, summation) equations. This route may lose advantages associated with the special structure of DDEs and NDDEs. However, there are areas in the study of NDDEs in particular – notably variation of parameters and adjoint theory – where one is uncertain how to proceed (the uncertainty appears to be widely shared, though largely unexpressed); the integral equation theory may then provide guidance on possible routes to pursue. The relationships also indicate a simple hierarchy of complexity that can be used, in the quest for parsimony, to rank various types of model.

Remark 2.4 From an abstract perspective, our models are expressible in terms of *causal* operators (also called Volterra operators): An operator $\mathcal{V}$ acting on a space of functions $\mathbb{S}$ each defined on $\mathbb{D} \subseteq \mathbb{R}$ is said to be causal if $\mathcal{V}y_1(t) = \mathcal{V}y_2(t)$ whenever $y_1(s) = y_2(s)$ for all $s \leq t$ with $s, t \in \mathbb{D}$. Classical causal equations can generally be expressed in the form $\mathcal{V}y(t) = f(t)$ (with $t \in \mathbb{D}$, and for the unknown $y \in \mathbb{S}$) and a neutral functional differential equation counterpart (*cf.* [14, p.123 onwards]) has the form $\frac{d}{dt}\mathcal{V}_1 y(t) = \mathcal{V}_2 y(t) + f(t)$ $(t \in \mathbb{D})$ where $\mathcal{V}_{1,2}$ are both causal.

3 RELATION OF ONE TYPE OF MODEL TO ANOTHER

We now concentrate on linear equations, since (*a*) nonlinear equations are often solved by iterative techniques involving linearization (*e.g.*, Newton iteration), (*b*) qualitative analysis of nonlinear equations often proceeds via a study of an approximating linearized equation, and (*c*) linear equations are simpler to investigate. To discuss the relation of each type of model equation to the other types we study evolutionary equations having the form

$$\mathcal{L}_\star\{y\}(t) = f(t) \quad (t \in [0, T]). \tag{3.1}$$

This equation is to be satisfied, given f, by a solution y that (in general) satisfies a supplementary *initial condition* on an *initial interval* $[-\tau, 0]$.

One may regard $\mathcal{L}_\star$ as defining an integral or differential expression, or one may introduce a corresponding operator, also denoted, without loss of clarity, by $\mathcal{L}_\star$ (or by $\mathcal{L}_\star : \mathbb{F} \to \mathbb{F}_1$), on a suitable function space $\mathbb{F}$ that includes the solutions of the problems under consideration. As examples of $\mathcal{L}_\star$ we introduce, below, $\mathcal{L}_{ODE}$, $\mathcal{L}_{DDE}$, $\mathcal{L}_{NDDE}$, $\mathcal{L}_{VIE}$, *etc.* corresponding to ODEs, DDEs, NDDEs, and VIEs. In general, the models incorporate parameters $\{p_\ell\}$ which correspond to features of the phenomenon under investigation. If appropriate, we write $f(\cdot) = f(\mathrm{p}; \cdot)$ ($\mathrm{p} \in \mathbb{R}^N$ has components p_ℓ, $\ell = 1, 2, \cdots, L$) and $\mathcal{L}_\star \equiv \mathcal{L}_\star^{\mathrm{p}}$. (For $\mathcal{L}_\star$, we place a parameter in a superscript, rather than an additional argument, for clarity in expressions such as $\mathcal{L}_\star^{\mathrm{p}}\{y(\mathrm{p}; t)\}$; $\mathbb{F} = \mathrm{dom}(\mathcal{L}_\star^{\mathrm{p}})$ may depend on p (then, $\mathbb{F} \equiv \mathbb{F}^{\mathrm{p}}$)).

3.1 From implicit to explicit NDDEs & vice versa

In the choice of mathematical models, one may ask which form (explicit or implicit) of NDDE is to be preferred. A *general* explicit NDDE cannot be rewritten (see [24, p. 54], [25, p. 119]) as an implicit NDDE. However, setting aside modelling or algorithmic considerations, we can show that subsets of NDDEs of explicit and of implicit types can (with sufficient differentiability conditions) be transformed to the alternative type.

Consider, as example, the linear explicit equation

$$\overset{\star}{\mathcal{L}}_{EXP}\{y\}(t) = f(t), \tag{3.2a}$$

with

$$\overset{\star}{\mathcal{L}}_{EXP}\{y\}(t) := y'(t) - \{A_*(t)y(t) + B_*(t)y(t-\tau) + C_*(t)y'(t-\tau)\}. \tag{3.2b}$$

Our equations are considered on $[0, T]$, with initial conditions that determine the solution and its smoothness. By implication, a solution of (3.2) has a derivative on $[-\tau, T]$ so we here suppose ϕ differentiable. Eqn (3.2) yields, if ${C'}_*$ exists,

$$\overset{\star}{\mathcal{L}}_{IMP}\{y\}(t) = f(t), \text{ for } t \in [0, T], \tag{3.3a}$$

with

$$\overset{\star}{\mathcal{L}}_{IMP}\{y\}(t) = \frac{d}{dt}\{y(t) - C_*(t)y(t-\tau)\} - \Big\{A_*(t)y(t) + \{B_*(t) - {C'}_*(t)\}y(t-\tau)\Big\} \tag{3.3b}$$

($t \in [0, T]$). This is an implicit neutral equation of the form

$$\mathcal{L}_{IMP}\{y\}(t) = f(t), \text{ for } t \in [0, T], \tag{3.4a}$$

with

$$\mathcal{L}_{IMP}\{y\}(t) := \frac{d}{dt}\{y(t) - C(t)y(t-\tau)\} - \{A(t)y(t) + B(t)y(t-\tau)\}. \tag{3.4b}$$

If C' and ϕ' exist and the solution of (3.4) has a derivative, then $\mathcal{L}_{EXP}\{y\}(t) = f(t)$, where

$$\mathcal{L}_{EXP}\{y\}(t) := y'(t) - \Big\{A(t)y(t) + \{B(t) + C'(t)\}y(t-\tau) + C(t)y'(t-\tau)\Big\}. \tag{3.5}$$

Note that (3.4) (3.5) reduce, if $C(t) \equiv 0$, to the simpler $\mathcal{L}_{DDE}\{y\}(t) = f(t)$, with

$$\mathcal{L}_{DDE}\{y\}(t) := y'(t) - \{A(t)y(t) + B(t)y(t-\tau)\}. \tag{3.6}$$

LEMMA 3.1 *Eqns (3.2), (3.4) and (3.5) are equivalent if their solutions are differentiable, ϕ', ${C'}_*$ and C' exist, and $A_*(t) = A(t)$, $B_*(t) - {C'}_*(t) = B(t)$, and $C_*(t) = C(t)$.*

Notation 3.1 *If we concentrate on simple linear NDDEs, we may base models either on the explicit form (3.2) or the implicit (3.4). When (assuming adequate differentiability) the discussion refers to both (3.4) and the equivalent (3.5),*

$$\mathcal{L}_{NDDE} \text{ represents either } \mathcal{L}_{IMP} \text{ or } \mathcal{L}_{EXP}. \tag{3.7}$$

For definiteness, the notation $\mathcal{L}_{NDDE}$ is taken as synonymous with $\mathcal{L}_{IMP}$ in the manipulation.

The definition of an operator requires specification of the space on which it acts, and both of the operators $\mathcal{L}_{DDE}$ and $\mathcal{L}_{NDDE}$ can be regarded as mapping certain functions defined on $[-\tau, T]$ to functions defined on $[0, T]$; later, we define quasi-adjoint operators that map functions defined on $[0, T + \tau]$ to functions defined on $[0, T]$ (for $T < \infty$). The difference in the domains is the source of some difficulties for which differing remedies can be proposed.

3.2 From linear implicit NDDEs to Volterra integral equations

We consider the implicit form (3.4) under the assumption $\phi \in C[-\tau, 0]$, and we show that the problem can be rewritten as the Volterra integral equation (3.15). This is primarily of theoretical significance; the equivalent result for DDEs or ODEs is a special case.

Definition 3.1 *We denote by $\lfloor w \rfloor$ (to be read as "floor w") the greatest integer not exceeding $w \in \mathbb{R}$; if $n = \lfloor \frac{t}{\tau} \rfloor$, $t \in [n\tau, (n+1)\tau)$. The identity matrix is I and we define*

$$\mathfrak{C}_{-1} = I,\ \mathfrak{C}_0(t) = C(t), \quad \mathfrak{C}_r(t) = \mathfrak{C}_{r-1}(t)C(t - r\tau) \ \ \text{for } r \in \{1, 2, 3, \cdots, n\}. \tag{3.8}$$

Lemma 3.2 *Suppose that $y(t) = \phi(t)$ for $t \in [-\tau, 0]$, and consider the relation*

$$y(t) - C(t)y(t - \tau) = u(t) \ \ \text{for } t \in [0, T]. \tag{3.9}$$

Then

$$y(t) = \sum_{r=-1}^{n-1} \mathfrak{C}_r(t)u(t - [r+1]\tau) + \mathfrak{C}_n(t)\phi(t - [n+1]\tau). \tag{3.10}$$

Proof: With $\mathfrak{C}_r(t) = C(t)C(t-\tau)\cdots C(t - r\tau)$, multiply $y(t - r\tau) - C(t - r\tau)y(t - [r+1]\tau) = u(t - r\tau)$ by $\mathfrak{C}_r(t)$, and sum over r. □

Now (3.4) with (3.9) gives $u(t) = \int_0^t \{A(s)y(s) + B(s)y(s-\tau) + f(s)\}\mathrm{d}s + u(0)$ and this yields the next lemma, on substituting $u(0) = \phi(0) - C(0)\phi(-\tau)$. Here,

$$x_+^0 := 1 \text{ if } x \geq 0, \quad x_+^0 := 0 \text{ if } x < 0. \tag{3.11}$$

Lemma 3.3 *The functions $u \equiv u(\phi;\cdot)$ and $y \equiv y(\phi;\cdot)$ in (3.9) satisfy*

$$u(t) = \int_0^t K(t,s)y(s)\mathrm{d}s + g(t), \ \textit{for } t \in [0,T], \tag{3.12}$$

where $g(t) = \int_{-\tau}^0 B(s+\tau)\phi(s)\mathrm{d}s + \{\phi(0) - C(0)\phi(-\tau)\} + \int_0^t f(s)\mathrm{d}s$ *and*

$$K(t,s) = A(s)(t-s)_+^0 + B(s+\tau)(t-\tau-s)_+^0. \tag{3.13}$$

Let $\mathcal{I}$ be the identity operator, and $\mathcal{K}$ the integral operator on $C[0,T]$ with kernel $K(t,s)$. If $C(t) \equiv 0$ (the DDE case), $u = y$ and (3.12) is an integral equation $\{\mathcal{I} - \mathcal{K}\}y = g$ for y.

Definition 3.2 *With $n = \lfloor t/\tau \rfloor$, we define*

$$\mathfrak{C}(t) = \sum_{r=-1}^{n-1} \mathfrak{C}_r(t), \ \textit{and } \mathfrak{C}_{\lfloor t/\tau \rfloor}(t) \equiv \mathfrak{C}_n(t) = C(t)C(t-\tau)\cdots C(t-n\tau), \tag{3.14a}$$

$$\mathrm{f}(t,s) = \sum_{r=-1}^{n-1} \mathfrak{C}_r(t) f(s)(t-s-[r+1]\tau)_+^0, \tag{3.14b}$$

$$\gamma(\phi;t) = \mathfrak{C}(t)\Big\{ \int_{-\tau}^0 B(s+\tau)\phi(s)\mathrm{d}s + \{\phi(0) - C(0)\phi(-\tau)\}\Big\} + \mathfrak{C}_{\lfloor t/\tau \rfloor}(t)\phi(t - \lfloor t/\tau \rfloor \tau) + \\ + \int_0^t \mathrm{f}(t,s)\mathrm{d}s, \tag{3.14c}$$

$$\mathfrak{K}(t,s) = \sum_{r=-1}^{n-1} \mathfrak{C}_r(t) K(t-[r+1]\tau, s) \tag{3.14d}$$

Thus, $\mathfrak{K}(t,s) = \sum_{r=-1}^{\lfloor t/\tau \rfloor - 1} \mathfrak{C}_r(t)\{A(s)(t-[r+1]\tau-s)_+^0 + B(s+\tau)(t-[r+2]\tau-s)_+^0\}$.

Theorem 3.4 *The solution $y \equiv y(\phi,\cdot)$ of (3.4), with $y(t) = \phi(t)$ for $t \in [-\tau, 0]$, is the unique solution of the Volterra integral equation*

$$y(t) = \int_0^t \mathfrak{K}(t,s)y(s)\mathrm{d}s + \gamma(\phi;t), \tag{3.15}$$

which can be written

$$\mathcal{L}_{INT}\{y\}(t) = \gamma(\phi;t) \ \textit{where } \mathcal{L}_{INT}\{y\}(t) := y(t) - \int_0^t \mathfrak{K}(t,s)y(s)\mathrm{d}s. \tag{3.16}$$

Proof: Apply Lemma 3.2 to (3.12). The theorem results after manipulation. □

Remark 3.1 As an alternative to the approach above, one may proceed from (3.2) to obtain a delay-differential equation for y of a more general type than (2.9): it involves lags τ, $2\tau, \ldots (n-1)\tau$ where $n = \lfloor t/\tau \rfloor$. Indeed, eqn (3.2) yields

$$y'(t-r\tau) = A_*(t-r\tau)y(t-r\tau) + B_*(t-r\tau)y((t-[r+1]\tau) + \\ + C_*(t-r\tau)y'((t-[r+1]\tau) + f(t-r\tau), \tag{3.17}$$

for $r \in \{1, 2, 3, \cdots, n\}$. If we multiply (3.17) by $\widehat{\mathfrak{C}}_{r-1}(t)$, where (*cf.* (3.8)),

$$\widehat{\mathfrak{C}}_{-1} = I,\ \widehat{\mathfrak{C}}_0(t) = C_*(t), \quad \widehat{\mathfrak{C}}_j(t) = \widehat{\mathfrak{C}}_{j-1}(t)C_*(t-j\tau) \ \text{ for } j = 1, 2, 3, \cdots, n, \tag{3.18}$$

($n = \lfloor t/\tau \rfloor$) we obtain, on summing, an equation of the form

$$\hat{\mathcal{L}}\{y\}(t) = \widehat{f}(t) \quad (t \in [0, T]) \tag{3.19a}$$

where

$$\hat{\mathcal{L}}\{y\}(t) := y'(t) - \sum_{r=0}^{\lfloor t/\tau \rfloor} \widehat{G}_r(t)y(t-r\tau), \tag{3.19b}$$

$$y(t) = \phi(t) \quad (t \in [-\tau, 0]). \tag{3.20}$$

Here, $\widehat{f}(\cdot) = \widehat{f}(\phi, \phi'; \cdot)$. We regard this as the basis for plausible algorithms; we leave the reader to verify expressions for $\widehat{f}(\cdot)$ and the coefficient functions $\{\widehat{G}_r(t)\}$.

4 ROUTES TO VARIATION OF PARAMETERS FORMULAE

Variation of parameters (VoP) formulae – or 'variation of constants' formulae – are a regular feature in the mathematics associated with modelling. They arise in methods of identifying actual parameters or initial conditions, in the correct analysis of perturbations in initial conditions and in studies of stability and of periodicity or bifurcation. As is well-known, VoP formulae [26] can be used [3] to show how solutions of differential and related equations vary with the parameters of the problem. Such formulae are often associated with fundamental solutions, or resolvent or solvent equations, or adjoint equations (see [3] and its references).

4.1 Solvent equations for NDDEs

Here, we apply integral equation theory to NDDEs. Let us return to the study of the solution $y = y(\phi; \cdot)$ (which we now suppose to be continuous) of the NDDE (3.4), *viz.*

$$\frac{d}{dt}\{y(t) - C(t)y(t-\tau)\} = A(t)y(t) + B(t)y(t-\tau) + f(t) \quad (t \in [0, T]), \\ y(t) = \phi(t) \quad (t \in [-\tau, 0]); \quad \phi \in C[-\tau, 0]. \tag{4.1}$$

We concentrate on linear equations, to gain insight into more general NDDEs. The solution of (4.1) satisfies (3.15), so the theory of Volterra integral equations of the type (3.15) allows us to deduce the form of y from Theorem 3.4. (Likewise, if $w \in C[0, T]$, $w - \mathcal{K}w = v$, and this equation is of the type in (3.15), w satisfies an NDDE of the type (4.1).)

We recall the rule of integration by parts; the usual case of vanishing Δ_ℓ often suffices.

Lemma 4.1 *Suppose, for* $1 \le m \in \mathbb{Z}$, $\ell \in \{0, 1, \ldots, m-1\}$, *the functions* u, v *have continuous first derivatives on* $[t_\ell, t_{\ell+1}]$ $(0 = t_0 < t_1 \ldots \le t_m = T)$, *and* $\Delta_\ell = \lim_{t\searrow t_\ell} u^{\mathrm{T}}(t)v(t) - \lim_{t\nearrow t_\ell} u^{\mathrm{T}}(t)v(t)$ *denotes the jump in* $u(t)v(t)$ *at* t_ℓ *(if* $u, v \in C[0,T]$, *all* Δ_ℓ *vanish). Then,* $\int_0^t [u(s)]^{\mathrm{T}} v'(s)\mathrm{d}s = \Big\{u^{\mathrm{T}}(t)v(t) + \sum_{t_\ell<t} \Delta_\ell - u^{\mathrm{T}}(0)v(0)\Big\} - \int_0^t [u'(s)]^{\mathrm{T}} v(s)\mathrm{d}s$, *for* $t \in [0,T]$.

Definition 4.1 *(The solvent kernels, cf [13, p.56 & pp.59-60].) The* resolvent kernel $\mathfrak{R}(t,s)$ *for* $\mathfrak{K}(t,s)$ *satisfies*

$$\mathfrak{R}(t,s) - \int_s^t \mathfrak{K}(t,\varsigma)\mathfrak{R}(\varsigma,s)d\varsigma = \mathfrak{K}(t,s) \;(\text{ for } 0 \le s \le t \le T) \tag{4.2}$$

and $\mathfrak{R}(t,s) = 0$ *for* $0 \le t < s \le T$. *The kernel* $\mathfrak{U}(t,s)$ *of the* differential resolvent *is* $I + \displaystyle\int_s^t \mathfrak{R}(t,\varsigma)d\varsigma$, *if* $0 \le s \le t \le T$, $\mathfrak{U}(t,s) = 0$ *if* $s > t$, *where* I *is the identity matrix and* 0 *is the zero matrix.*

Lemma 4.2 *Denote by* $u(\cdot) \equiv u[g;\cdot]$ *the solution of* $u(t) - \displaystyle\int_0^t \mathfrak{K}(t,s)u(s)\mathrm{d}s = g(t)$ $(t \in [0,T])$. *(a) When* $g \in C[0,T]$, $u(t) = g(t) + \displaystyle\int_0^t \mathfrak{R}(t,s)g(s)\mathrm{d}s$. *(b) When, in addition,* g' *is continuous except for jumps at points* $\{t_\ell\} \subset [0,T]$, $u(t) = \mathfrak{U}(t,0)g(t) + \displaystyle\int_0^t \mathfrak{U}(t,s)g'(s)\mathrm{d}s$.

Denote by $\mathcal{K}$ the integral operator on $C[0,T]$ with kernel $\mathfrak{K}(t,s)$, and by $\mathcal{R}$ the integral operator on $C[0,T]$ with kernel $\mathfrak{R}(t,s)$, then $(\mathcal{I}-\mathcal{K})(\mathcal{I}+\mathcal{R}) = \mathcal{I}$ (the identity); part (a) follows. $\mathfrak{U}(t,s)$ is continuous for $s \in [0,t]$ though $\{\partial/\partial s\}\mathfrak{U}(t,s) = -\mathfrak{R}(t,s)$ for $s < t$ is piecewise continuous, and if g is continuous and g' is piecewise continuous, Lemma 4.1 applied to the components of $\int_0^t \mathfrak{R}(t,s)g(s)\mathrm{d}s$ yields part (b). The obvious result that if $g(t) = g_1(t) + g_2(t)$ then $u[g;t] = u[g_1;t] + u[g_2;t]$ can be useful.

Theorem 4.3 *In the notation (3.14), the solution* $y(\cdot) \equiv y(\phi;\cdot)$ *of (4.1) satisfies*

$$y(\phi;t) = \gamma(\phi;t) + \int_0^t \mathfrak{R}(t,s)\gamma(\phi;s)\mathrm{d}s \quad (t \in [0,T]), \tag{4.3a}$$

$$y(\phi;t) = \mathfrak{U}(t,0)\gamma(0) + \int_0^t \mathfrak{U}(t,s)\gamma'(s)\mathrm{d}s \;(t \in [0,T]). \tag{4.3b}$$

Note that $\gamma(\phi;t)$, defined in (3.14c), acquires its differentiability from that of $\mathfrak{C}(t)$, $\mathfrak{C}_{\lfloor t/\tau\rfloor}t)$, $\phi(t)$, and $\int_0^t \mathrm{f}(t,s)\mathrm{d}s$. If we substitute $\gamma(\phi;t)$ into (4.3a) we obtain an equation for $\phi(\cdot)$ given $y(\phi;\cdot)$; solution of this solves the *inverse problem* to (4.1); see also (4.23). We define $f_\star(t) \equiv f_\star(C;t)$ by setting

$$f_\star(t) := \int_0^t \mathrm{f}(t,s)\mathrm{d}s \text{ where } \mathrm{f}(t,s) := \sum_{r=-1}^{n-1} \mathfrak{C}_r(t)f(s)(t-s-[r+1]\tau)_+^0. \tag{4.4}$$

Corollary 4.4 *(a) If f vanishes, $\gamma(\phi;t)$ reduces to $\gamma_0(\phi;t) = \mathfrak{C}(t)\Big\{\int_{-\tau}^{0} B(s+\tau)\phi(s)\mathrm{d}s + \{\phi(0) - C(0)\phi(-\tau)\}\Big\} + \mathfrak{C}_{\lfloor t/\tau \rfloor}(t)\phi(t - \lfloor t/\tau \rfloor \tau)$ and $y(\phi;t)$ reduces to $y_0(\phi;t) = \gamma_0(\phi;t) + \int_0^t \mathfrak{R}(t,s)\gamma_0(s)\mathrm{d}s$. (b) If ϕ vanishes, $y(\phi;t)$ reduces to $y(0;t) = f_\star(t) + \displaystyle\int_0^t \mathfrak{R}(t,s) f_\star(s)\mathrm{d}s$. (c) The solution $y(\phi;t)$ of the general inhomogeneous problem is $y(\phi;t) = y_0(\phi;t) + y(0;t)$.*

4.2 Adjoint operators, adjoint equations, and related concepts

The term "adjoint" has a context-dependent meaning, in the literature. The discussion of adjoints in the case of integral equations can be regarded as an extension of that for systems of algebraic equations. The expressions *adjoint function*, *adjoint equation*, *adjoint operator* and *adjoint expression* or *Lagrange adjoint* are to be found in use [12,19,20] in the context of DDEs and explicit NDDEs. Applied to DEs, the concepts date to Lagrange in the eighteenth century. The discussion in [15] relates to DEs and differential expressions and the results required for DDEs and NDDEs may be viewed as generalizations of Green's formula and the bilinear concomitant [7]. What one seeks in the literature is a clear statement of the definition and significance of adjoints, but the older texts introduce "adjoints" in an opportunistic (often informal) fashion while some others write only for the *cognoscenti.* Especially for NDDEs, clear formulations are difficult to locate; our discussion below, in particular the definition of a quasi-adjoint and of an adjoint function, is intended to overcome this problem.

Remark 4.1 Conventionally, one may study adjoints from a functional analytic approach; this permits a unified framework. With that in mind, first let $\mathcal{L}$ be a bounded linear operator mapping a normed linear space $\mathbb{F}$ onto a normed linear space $\mathbb{F}_1$, and let λ_1 be a bounded linear functional on $\mathbb{F}_1$. If $x_1 = \mathcal{L}x_0$ then $\lambda_1(x_1) = \lambda_1(\mathcal{L}x_0)$ defines a linear functional λ_0 on $\mathbb{F}$ such that $\lambda_0(x_0) = \lambda_1(\mathcal{L}x_0)$. The map $\mathcal{L}^*$ assigning λ_0 to λ_1 is the functional analytic adjoint of $\mathcal{L}$. The above remarks sketch concepts discussed rigorously in [31], in the context where $\mathbb{F}$ is a Banach space (complete normed linear space); *cf.* [18, p.172], [27, p. 114], and [34] for further reading.

To simplify we now consider a special case where $\mathcal{L}$ acts on a linear space $\mathbb{F}$ is equipped with an inner-product $\langle\cdot,\cdot\rangle$ over $\mathbb{R}$. Occasionally, we consider a pseudo-inner-product: the notation $\prec \cdot \succ$ in place of $\langle\cdot,\cdot\rangle$ includes this possibility. When equipped with this structure, we denote the resulting space by $\mathcal{F} := [\mathbb{F}, \langle\cdot,\cdot\rangle]$ or by $\mathcal{F} := [\mathbb{F}, \prec\cdot,\cdot\succ]$; $\mathcal{L}$ is supposed to be a linear operator defined on $\mathcal{F}$. Now, $\prec\ell, x_0\succ$ defines an association between $\ell \in \mathcal{F}$ and the functional λ (in the conjugate space) with $\lambda(x_0) = \prec\ell, x_0\succ$.

[7] Integration by parts (Lemma 4.1) provides the simplest example. For an introduction to the concepts, and related material, see [15, p. 277 *et seq.*], [29,31,32].

Definition 4.2 *A* "true adjoint" *(or "functional analytic adjoint") of the operator $\mathcal{L}$ on $\mathcal{F}$ is an operator $\mathcal{L}^*$ such that $\prec \mathcal{L}w_0, w_1 \succ = \prec w_0, \mathcal{L}^* w_1 \succ$ for all $w_{0,1} \in \mathcal{F}$. An* adjoint equation *corresponding to an equation $\mathcal{L}y = f$ then has the form $\mathcal{L}^* x = g$.*

The definition of a true adjoint here depends on the definition of $\mathcal{F}$ (that is, $\mathbb{F}$ and $\prec \cdot, \cdot \succ$); for an inner-product space the true adjoint is unique. Definitions of f, g are needed to specify solutions $y \in \mathcal{F}$, $x \in \mathcal{F}$ in Definition 4.2. A (real-valued) inner-product generates a norm with which $[\mathbb{F}, \langle \cdot, \cdot \rangle]$ may be a Hilbert space or a pre-Hilbert space; a pseudo-inner-product generates a semi-norm. Throughout, we let (w_1, w_2) denote $\int_0^T [w_1(s)]^{\mathrm{T}} w_2(s) \mathrm{d}s$, the conventional inner-product, whenever this is defined.

Functions in $\mathbb{F}$ may be required to satisfy certain homogeneous end conditions. In work involving adjoints found in the literature, some part of the structure (such as the prescription of end conditions that contribute to the definition of a function space, or the definition of $\langle \cdot, \cdot \rangle$ or $\prec \cdot, \cdot \succ$) is, often, not clearly specified or is assumed, and the term "formal" is applied to the term adjoint or quotation marks are placed around the word adjoint; we use *formal adjoint* where detail or rigour seems, initially, to be lacking.

Example 4.1 *For the DDE case (3.6),*

$$\mathcal{L}_{DDE}\{y\}(t) = f(t),$$

(for some f) where

$$\mathcal{L}_{DDE}\{y\}(t) \equiv y'(t) - \{A(t)y(t) + B(t)y(t-\tau)\} \quad (t \in [0,T]), \tag{4.5a}$$

let us define a "formal adjoint" equation *as an equation of the form*

$$\mathcal{L}^{\dagger}_{DDE}\{x\}(t) = g(t) \tag{4.5b}$$

(for some g) where[8]

$$\mathcal{L}^{\dagger}_{DDE}\{x\}(t) \equiv -x'(t) - \{A^{\mathrm{T}}(t)x(t) + B^{\mathrm{T}}(t+\tau)x(t+\tau)\} \quad (t \in [0,T]). \tag{4.5c}$$

The "adjoint" in (4.5c) was motivated in [12] by considering an invariant bilinear function.

One can show (using integration by parts) a correspondence between (4.5c) and a true adjoint. Let $\mathbb{F}$ denote the space consisting of functions defined on $[-\tau, T+\tau]$, that have support $[0,T]$ and are differentiable on $[0,T]$; for operators assumed to act on $\mathbb{F}$, this restricts the problems considered. For

[8] Note that $\mathcal{L}^{\dagger}_{DDE}$ should be thought of as $[\mathcal{L}_{DDE}]^{\dagger}$; the equation (4.5b) is an *advanced* equation, not a DDE. Similar remarks apply to $\mathcal{L}^{\dagger}_{EXP}$, $\mathcal{L}^{\dagger}_{IMP}$, $\mathcal{L}^{\dagger}_{NDDE}$, $\mathcal{L}^{\natural}_{NDDE}$ *etc.*, below, and $\mathcal{L}^{\dagger}_{VIE}$.

$\mathcal{F} = [\mathbb{F}, (\cdot,\cdot)]$, with $(u, v) := \int_0^T [u(s)]^{\mathrm{T}} v(s)\mathrm{d}s$

$$\int_0^T [w_1(s)]^{\mathrm{T}} \mathcal{L}_{DDE}\{w_2\}(s)\mathrm{d}s = \int_0^T [\mathcal{L}^{\dagger}_{DDE}\{w_1\}(s)]^{\mathrm{T}} w_2(s)\mathrm{d}s \quad (w_{1,2} \in \mathcal{F}); \quad (4.6)$$

thus, $\mathcal{L}^{\dagger}_{DDE}$ is the true adjoint of $\mathcal{L}_{DDE}$ on $\mathcal{F}$. However, if $v_{1,2} \in C^1[-\tau, T+\tau]$ do not have support $[0, T]$ *then the inner-product and $\mathcal{L}^{\dagger}_{DDE}\{v_1\}$, $\mathcal{L}_{DDE}\{v_2\}$ defined by (4.5) have a meaning but, in general, $(v_1, \mathcal{L}_{DDE}\{v_2\}) \neq (\mathcal{L}^{\dagger}_{DDE}\{v_1\}, v_2)$. For such functions (cf. Lemma 4.1),*

$$(v_1, \mathcal{L}_{DDE}\{v_2\}) = \kappa_{DDE}(v_1, v_2) + (\mathcal{L}^{\dagger}_{DDE}\{v_1\}, v_2) \quad (4.7)$$

where $\kappa_{DDE}(v_1, v_2)$ is an extension of the bilinear concomitant in Green's formula.

4.3 Towards an adjoint theory for NDDEs

We turn to NDDEs. A definition of a formal adjoint equation for the explicit NDDE

$$\mathcal{L}_{EXP}\{y\}(t) = f(t) \quad (4.8a)$$

in (3.5) was proposed in [12, p.320] and it reduces, if $C(t) \equiv 0$, to $\mathcal{L}^{\dagger}_{DDE}\{x\}(t) = g(t)$ in (4.5c). Our question, here, is what form we should consider as suitable for the description "formal adjoint equation" in the case of implicit NDDEs, $\mathcal{L}_{IMP}\{y\}(t) = f(t)$ in (3.4), where

$$\mathcal{L}_{IMP}\{y\}(t) := \frac{d}{dt}\{y(t) - C(t)y(t-\tau)\} - \{A(t)y(t) + B(t)y(t-\tau)\}. \quad (4.9)$$

Remark 4.2 Reviewing the literature, a candidate for the description "formal adjoint equation" is the equation

$$\mathcal{L}^{\dagger}_{IMP}\{x\}(t) = g(t) \ (t \in [0, T]). \quad (4.10a)$$

where

$$\mathcal{L}^{\dagger}_{IMP}\{x\}(t) := -\frac{d}{dt}x(t) + C^{\mathrm{T}}(t+\tau)\frac{d}{dt}\{x(t+\tau)\} - \{A^{\mathrm{T}}(t)x(t) + B^{\mathrm{T}}(t+\tau)x(t+\tau)\}. \quad (4.10b)$$

While $y(\phi;\cdot)$ is specified by $y(t) = \phi(t)$ for $t \in [-\tau, 0]$, a solution $x(\psi;\cdot)$ of $\mathcal{L}^{\ddagger}_{IMP}\{x\}(t) = g(t)$ is specified by defining $x(\psi;t) = \psi(t)$ for $t \in [T, T+\tau]$. and there are thus issues centered on the existence of x', given ψ; indeed, (4.10) is not in an advanced form that is analogous to the implicit NDDE, so much as one analogous to the explicit NDDE. Furthermore, *motivation* for writing down (4.10) for the NDDE case is not immediate from a scan of the literature. (In [20,23], "adjoints" and "true adjoints" are discussed using the notation referred to in §2.2.)

Let us assume that $C(t)$ is differentiable and consider the form

$$\mathcal{L}^{\ddagger}_{IMP}\{x\}(t) = g(t) \text{ for } t \in [0, T], \quad (4.11)$$

where

$$\mathcal{L}^{\ddagger}_{IMP}\{x\}(t) := -\frac{d}{dt}\{x(t) - C^{\mathrm{T}}(t+\tau)x(t+\tau)\} - \{A^{\mathrm{T}}(t)x(t) + [B(t+\tau) + C'(t+\tau)]^{\mathrm{T}}x(t+\tau)\}, \tag{4.12}$$

for $t \in [0, T]$, with $x(t) = \psi(t)$ for $t \in [T, T+\tau]$ (with $\psi \in C[T, T+\tau]$).

Recall that we take $\mathcal{L}_{NDDE}$ as $\mathcal{L}_{IMP}$ in (4.9) and, assuming sufficient differentiability, it also represents $\mathcal{L}_{EXP}$ in (3.5); then $\mathcal{L}^{\ddagger}_{NDDE}$, taken as $\mathcal{L}^{\ddagger}_{IMP}$ in (4.12), also represents $\mathcal{L}^{\ddagger}_{EXP}$.

4.4 Adjoints and quasi-adjoints for NDDEs

What we seek, for application in manipulation in this work, is a generalization of (4.7) to NDDEs (the case of DDEs is subsumed in our discussion), of the general form (say)

$$\prec v_1, \mathcal{L}_{NDDE}\{v_2\}\succ = \kappa^{\natural}_{NDDE}(v_1, v_2) + \prec \mathcal{L}^{\natural}_{NDDE}\{v_1\}, v_2 \succ \tag{4.13}$$

for all $v_{1,2} \in \mathcal{V} \equiv [\mathbb{V}; \prec\cdot,\cdot\succ]$. Here, $\mathcal{V}$ is some (pseudo-)inner-product space. But a relation of the type $\prec v_1, \mathcal{L}_{NDDE}\{v_2\}\succ = \kappa^{\sharp}(v_1, v_2) + \prec \mathcal{L}^{\sharp}_{NDDE}\{v_1\}, v_2\succ$ may be satisfied by more than one operator $\mathcal{L}^{\sharp}_{NDDE}$ with differing $\kappa^{\sharp}$, if the properties of $\kappa^{\sharp}(v_1, v_2)$ are not pre-defined. We can term $\mathcal{L}^{\natural}$ a *quasi-adjoint* of $\mathcal{L}$ if, for some pseudo-inner-product space $\mathcal{V} := [\mathbb{V}; \prec, \cdot, \cdot, \succ]$, there exist $\mathbb{V}_{1,2} \subset \mathbb{V}$ such that, whenever $v_1 \in \mathbb{V}_1,\ v_2 \in \mathbb{V}_2$, (4.13) holds with $\kappa^{\natural}(v_1(\cdot), v_2(\cdot)) = 0$. If $\kappa^{\natural}(v_1, v_2) = 0$ for arbitrary $v_{1,2} \in \mathbb{V}$ then we term $\mathcal{L}^{\natural}$ a *true adjoint* of $\mathcal{L}$ on $[\mathbb{V}; \prec\cdot,\cdot\succ]$, and we write $\mathcal{L}^{\natural}$ as $\mathcal{L}^{\star}$. In some respects, it is simpler (in our applications to NDDEs) to implement a related definition of *(quasi-) adjoint functions.*

Definition 4.3 *Regard $\mathcal{L}_{NDDE}\{u\}(t) = f(t)$, and $\mathcal{L}^{\ddagger}_{NDDE}\{w\}(t) = g(t)$ as equations for $t \in [0, T]$ with solutions $u(\cdot) = u(f, \phi; \cdot)$, $w(\cdot) = w(g, \psi, \cdot)$ satisfying $u(t) = \phi(t)$ for $t \in [-\tau, 0]$, and $w(t) = \psi(t)$ for $t \in [T, T+\tau]$. Then w is a* quasi-adjoint function *for u if*

$$\prec w, f\succ = \kappa_{NDDE}(w, u) + \prec g, u\succ \quad (f = \mathcal{L}_{NDDE}\{u\},\ g = \mathcal{L}^{\ddagger}_{NDDE}\{w\}); \tag{4.14}$$

further, w is called an adjoint function *for u if ϕ, ψ and f, g are such that $\kappa_{NDDE}(w, u) = 0$.*

Lemma 4.5 has a bearing on our discussion of adjoint functions. $\mathcal{L}_{NDDE}$ represents both $\mathcal{L}_{IMP}$ and $\mathcal{L}_{EXP}$, assuming differentiability, and $\mathcal{L}^{\ddagger}_{NDDE}\{\ \}$ thus represents $\mathcal{L}^{\ddagger}_{IMP}$ and $\mathcal{L}^{\ddagger}_{EXP}$. Part ($b$) based on the conventional inner-product, is open to some ready generalizations.

Lemma 4.5 *Suppose $\mathcal{L}_{\ NDDE}\{u\}(t) = f(t)$ for $t \in [0, T]$, $u(t) = 0$ for $t \in [-\tau, 0]$, and suppose $\mathcal{L}^{\ddagger}_{NDDE}\{v\}(t) = g(t)$ for $t \in [0, T]$, $v(t) = 0$ for $t \in [T, T+\tau]$, where $f, g \in C[0, T]$. Then (a) $u, v \in C[0, T]$ have derivatives that are* *continuous on* $[0,T]$ *except possibly at points* $t \in \{\ell\tau\}_{\ell\in\mathbb{N}} \subseteq [0,T]$*, and (b)* $\int_0^T [v(s)]^{\mathrm{T}} \mathcal{L}_{NDDE}\{u\}(s)\mathrm{d}s = \int_0^T [\mathcal{L}^{\ddagger}_{NDDE}\{v\}(s)]^{\mathrm{T}} u(s)\mathrm{d}s$.

Proof: Part(a) can be established using a method of steps. We focus on (b) and consider terms contributing to $\int_0^T [\mathcal{L}^{\ddagger}_{NDDE}\{v\}(s)]^{\mathrm{T}} u(s)\mathrm{d}s$, *i.e.*, $\int_0^T [\mathcal{L}^{\ddagger}_{IMP}\{v\}(s)]^{\mathrm{T}} u(s)\mathrm{d}s$. Our proof is based on manipulation using integration by parts, selective changes of variable, and the conditions $u(t) = u'(t) = 0$ for $t \in [-\tau, 0]$, and $v(t) = v'(t) = 0$ for $t \in [T, T+\tau]$. Clearly, $\int_0^T [A^{\mathrm{T}}(s)v(s)]^{\mathrm{T}} u(s)\mathrm{d}s = \int_0^T [v(s)]^{\mathrm{T}} A(s)u(s)\mathrm{d}s$; $\int_0^T [B^{\mathrm{T}}(s+\tau)v(s+\tau)]^{\mathrm{T}} u(s)\mathrm{d}s = \int_0^T [v(s)]^{\mathrm{T}} B(s)u(s-\tau)\mathrm{d}s$. Apply integration by parts to $\int_0^T [\{v(s) - C^{\mathrm{T}}(s+\tau)v(s+\tau)\}' - [C'(t)]^{\mathrm{T}} v(s+\tau)]^{\mathrm{T}} u(s)\mathrm{d}s$, change variables of integration as appropriate, and exploit end conditions, to obtain $-\int_0^T [v(s)]^{\mathrm{T}} \{u(s) - C(s)u(s-\tau)\}'\mathrm{d}s$. Assembling the component parts, the result (b) follows.

4.5 The resolvent, and true adjoint of an integral operator

We now examine a true adjoint in the context of the integral equation formulation in §4.1. Our matrix-valued kernel functions $\mathfrak{K}$, $\mathfrak{R}$, and $\mathfrak{U}$ define operators $\mathcal{K}$, $\mathcal{R}$, $\mathcal{U}$, on appropriate subspaces of the space $L_2[0,T]$ of real-vector-valued functions, *e.g.*: $\mathcal{K}w(t) := \int_0^T \mathfrak{K}(t,s)w(s)\mathrm{d}s = \int_0^t [\mathfrak{K}(t,s)]w(s)\mathrm{d}s$ for $w \in L_2[0,T]$, $t \in [0,T]$. With (w_1, w_2) denoting $\int_0^T [w_1(s)]^{\mathrm{T}} w_2(s)\mathrm{d}s$, we have $(w_1, \mathcal{K}w_2) = (\mathcal{K}^* w_1, w_2)$ where $\mathcal{K}^* w(t) = \int_t^T [\mathfrak{K}(s,t)]^{\mathrm{T}} w(s)\mathrm{d}s$, $(w \in L_2[0,T]$, $t \in [0,T])$. Likewise, the integral operator $\mathcal{R}$ on $L_2[0,T]$ has as true adjoint $\mathcal{R}^*$, the integral operator with $\mathcal{R}^* w(t) = \int_t^T [\mathfrak{R}(s,t)]^{\mathrm{T}} w(s)\mathrm{d}s$; $\mathcal{U}$, the integral operator with kernel $\mathfrak{U}(t,s)$, has as its adjoint the integral operator with kernel $[\mathfrak{U}(s,t)]^{\mathrm{T}}$.

Example 4.2 *If, cf.(3.14d),* $\mathfrak{K}(t,s) = \sum_{r=-1}^{\lfloor t/\tau \rfloor - 1} \mathfrak{C}_r(t) K(t - [r+1]\tau, s)$*, then* $[\mathfrak{K}(s,t)]^{\mathrm{T}}$ *can be written* $\sum_{r=-1}^{\lfloor s/\tau \rfloor - 1} K^{\mathrm{T}}(s, t - [r+1]\tau)\mathfrak{C}_r^{\mathrm{T}}(s)$*; the true adjoint of* $\mathcal{L}_{INT}$ *on* $[L_2[0,T]; (\cdot,\cdot)]$ *follows.*

4.5.1 Further directions

Definition 4.4 *The equation* $\hat{\mathcal{L}}^{\dagger}\{x\}(t) = \breve{g}(t)$ $(t \in [0,T])$ *where* $\hat{\mathcal{L}}^{\dagger}\{x\}(t) := -x'(t) + \sum_{r=0}^{\lfloor t/\tau \rfloor} \widehat{G}_r^{\mathrm{T}}(t + r\tau)x(t + r\tau)$*, is a formal adjoint equation for (3.19), and a solution is specified by defining* $\breve{g}(t)$*, and requiring* $x(t) = \breve{\psi}(t)$ *for* $t \in [T, T+\tau]$.

The following, which in some sense generalizes Example 4.1, justifies Definition 4.4.

Theorem 4.6 *Let* $\mathbb{F}$ *denote the functions defined on* $\mathbb{R}$*, that have support, and are differentiable on* $[0,T]$ *and form an inner-product space* $\mathcal{F} := [\mathbb{F}; (\cdot,\cdot)]$ *with the conventional inner-product* $(v_1, v_2) := \int_0^T v_1^{\mathrm{T}}(s)v_2(s)\mathrm{d}s$*. Suppose that* $v_{1,2} \in \mathcal{F}$ *and for* $t \in [0,T]$ *define operators* $\hat{\mathcal{L}}$ *and* $\hat{\mathcal{L}}^{\dagger}$ *on* $\mathcal{F}$ *where* $\hat{\mathcal{L}}\{v\}(t) :=$

$\sum_{r=0}^{\lfloor t/\tau \rfloor} \widehat{G}_r(t)v(t-r\tau)$, *as in (3.19), and* $\widehat{\mathcal{L}}^{\dagger}\{v\}(t) := \sum_{r=0}^{\lfloor t/\tau \rfloor} \widehat{G}_r^{\mathrm{T}}(t+r\tau)v(t+r\tau)$ *as in Definition 4.4. Then* $\int_0^T v_2^{\mathrm{T}}(s)\widehat{\mathcal{L}}\{v_1\}(s)\mathrm{d}s = \int_0^T [\mathcal{L}^{\dagger}\{v_2\}(s)]^{\mathrm{T}} v_1(s)\mathrm{d}s$*, and* $\widehat{\mathcal{L}}^{\dagger}$ *is the true adjoint* $\widehat{\mathcal{L}}^{\star}$ *of* $\widehat{\mathcal{L}}$ *on* $\mathcal{F} = [\mathbb{F}; (\cdot,\cdot)]$.

4.6 Linear functionals expressed using adjoints or quasi-adjoints

We now recall a result of Marchuk *et al.* (for further reading, see [29,31]). Following [32], we here consider, from a formal perspective, a basic result in the use of operators on a Hilbert space, and their adjoints. We quote a result which states that a certain linear functional can be evaluated in terms of an adjoint. Suppose (i) $\mathcal{H} := [\mathbb{H}; \langle\cdot,\cdot\rangle]$ is a Hilbert space (over $\mathbb{R}$), where the norm in $\mathcal{H}$ is defined as $\|\cdot\| = \sqrt{\langle\cdot,\cdot\rangle}$; (ii) $\mathcal{L}$ is a linear operator with (for simplicity) domain $\mathcal{H}$. By definition, the true adjoint $\mathcal{L}^{\star}$ satisfies D

$$\langle v_1, \mathcal{L}v_2\rangle = \langle \mathcal{L}^{\star}v_1, v_2\rangle, \ \forall v_{1,2} \in \mathcal{H} \tag{4.15}$$

(equally, $\langle \mathcal{L}v_1, v_2\rangle = \langle v_1, \mathcal{L}^{\star}v_2\rangle$, for all $v_{1,2} \in \mathcal{H}$) and, again for simplicity, $\mathcal{L}^{\star}$ has domain $\mathcal{H}$. Consider the "basic equation",

$$\mathcal{L}\sigma = \delta, \text{ where } \delta \in \mathcal{H}, \tag{4.16}$$

and the problem of evaluating a bounded linear functional, $M_\omega[\sigma]$, of the solution σ, where

$$M_\omega[\sigma] := \langle \omega, \sigma\rangle \equiv \langle \sigma, \omega\rangle \ \text{ (wherein } \omega \in \mathcal{H} \text{ is given).} \tag{4.17}$$

Consider, with (4.16), the 'adjoint equation' with the non-homogeneous term ω in (4.17):

$$\mathcal{L}^{\star}x_\omega = \omega. \tag{4.18}$$

Theorem 4.7 *$M_\omega[\sigma]$ in (4.17) can be written $\langle x_\omega, \delta\rangle$ where x_ω satisfies (4.18).*

Proof: From (4.16) and (4.18), $\langle x_\omega, \mathcal{L}\sigma\rangle - \langle \mathcal{L}^{\star}x_\omega, \sigma\rangle = \langle x_\omega, \delta\rangle - \langle \omega, \sigma\rangle$, and using (4.15) we deduce $\langle x_\omega, \delta\rangle = \langle \omega, \sigma\rangle$, the required result. □

To find the value $M_\omega[\sigma]$ we may use either its definition (4.17), or $\langle x_\omega, \delta\rangle = \langle \delta, x_\omega\rangle$. The preceding material is largely a paraphrase of material in [32], somewhat simplified. In Theorem 4.8 we indicate a natural extension of Theorem 4.7. Our assumptions presents no difficulty in the case of integral operators. DDEs and NDDEs are more complex: in a sense, some of the difficulties can be associated with the fact that the definition of $\mathcal{L}_{NDDE}\{u\}(t)$ for $t \in [0,T]$ requires knowledge of $u(t)$ for $t \in [-\tau, T]$ whereas for the natural definitions of an adjoint or quasi-adjoint $\mathcal{L}^{\natural}_{NDDE}$ the definition of $\mathcal{L}^{\natural}_{NDDE}\{v\}(t)$ for $t \in [0,T]$ requires $v(t)$ for $t \in [0, T+\tau]$. However, the following result is straightforward.

Theorem 4.8 *Suppose* $\mathcal{L}_{NDDE}\{\sigma\} = \delta$, $\mathcal{L}^{\ddagger}_{NDDE}\{x_\omega\} = \omega$, $\widehat{M}_\omega[\sigma] := \prec\omega, \sigma\succ$. *Then* $\widehat{M}_\omega[\sigma] = \kappa_{NDDE}(x_\omega, \delta) + \prec x_\omega, \delta\succ$ *when* $\prec x_\omega, \mathcal{L}_{NDDE}\{\sigma\}\succ = \kappa_{NDDE}(x_\omega, \sigma) + \prec\mathcal{L}^{\ddagger}_{NDDE}\{x_\omega\}, \sigma\succ$.

We apply Theorem 4.8 below but note that it can be modified (*cf.* Theorem 4.7) for the case where $\mathcal{L}^\natural$ is a quasi-adjoint for a bounded linear operator $\mathcal{L}$, both operators acting on a pseudo-inner-product space $\mathcal{V} := [V; \prec\cdot,\cdot\succ]$ and satisfying $\prec v_1, \mathcal{L}\{v_2\}\succ = \kappa^\natural(v_1, v_2) + \prec\mathcal{L}^\natural\{v_1\}, v_2\succ$ for all $v_{1,2} \in \mathcal{V}$.

4.7 The fundamental solution and adjoint equations

In the theory of linear DDEs, and NDDEs, the "adjoint" equation is associated, in a modified form, with the definition of a function Cauchy matrix or 'fundamental solution'. It provides a variation of parameters formula. We again make the differentiability assumptions used in Lemma 3.1.

To motivate our discussion, we consider the case $\phi(t) = 0$ for $t \in [-\tau, 0]$. Suppose that

$$\mathcal{L}_{NDDE}\{z_0\}(t) = f_\natural(t) \ (t \in [0,T]) \text{ where } z_0(t) = 0 \text{ for } t \in [-\tau, 0]. \tag{4.19}$$

As is clear from Theorem 4.3, $y_0(t) \equiv y_0(f_\natural; t)$, can be expressed in the form $\int_0^t \mathfrak{X}(t,s) f_\natural(s) \mathrm{d}s$ for some function $\mathfrak{X}$. (For the integral equation theory based on the resolvent $\mathfrak{R}$, it suffices that $f_\natural$ is piecewise continuous.) Rather than use the integral equation theory to investigate $\mathfrak{X}$, we use an alternative approach based, in part, on adjoint theory (see Remark 4.3).

Definition 4.5 *(a) Denote by $Y(s,t)$ the solution (for $0 \le s \le t \le T$) of*

$$-\frac{\partial}{\partial s}\Big\{Y(s,t) - Y(s+\tau,t)C(s+\tau)\Big\} = Y(s,t)A(s) + Y(s+\tau,t)[B(s+\tau) + C'(s+\tau)], \tag{4.20}$$

which satisfies $Y(s,t) = 0$, for $t < s$, $Y(t,t) = I$. (b) Let $\mathfrak{C}_i(t)$, $\mathfrak{C}(t)$ be defined as in (3.8). We define the generalized fundamental solution *for the implicit NDDE (4.1)*

$$\mathfrak{Y}(s,t) := \sum_{i=-1}^{\lfloor t/\tau \rfloor - 1} \mathfrak{C}_i(t) Y(s, t-(i+1)\tau). \tag{4.21}$$

Remark 4.3 Here, $\mathfrak{C}_{-1}(t) = I$ and $\mathfrak{C}_i(t) := \prod_{j=0}^{i} C(t - j\tau)$ for $i \in \mathbb{N}$ (multiply left-to-right with increasing j) and $\mathfrak{C}(t) := \mathfrak{C}_n(t)$, $n = \lfloor t/\tau \rfloor$. Note the simplification in (4.20) if $C'(t) \equiv 0$. If $C(t)$ vanishes then $\mathfrak{Y}(s,t) = Y(s,t)$. Note the *discontinuity* in $Y(s,t)$ at $s = t$. Denote by ρ^{T} any chosen row of the identity matrix (and fix $t_* \in [0,T]$); then (4.20) yields

$$-\frac{\partial}{\partial s}\Big\{\rho^{\mathrm{T}} Y(s,t_*) - \rho^{\mathrm{T}} Y(s+\tau,t_*)C(s+\tau)\Big\} = \rho^{\mathrm{T}} Y(s,t_*)A(s) + \rho^{\mathrm{T}} Y(s+\tau,t_*)D(s+\tau); \tag{4.22a}$$

where

$$D(t) := B(t) + C'(t). \tag{4.22b}$$

Every row $\rho^{\mathrm{T}} Y(\cdot, t_*)$ satisfies the equation obtained on taking transposes in the "adjoint" (4.12).

THEOREM 4.9 *(Variation of Parameters) The solution $y(t) \equiv y(\phi; t)$ of the NDDE (4.1) is expressible as*

$$y(\phi; t) = y_0(\phi; t) + \int_0^t \mathfrak{Y}(s,t) f(s) ds \tag{4.23a}$$

where

$$y_0(\phi; t) = \mathfrak{Y}(0,t)\Big(\phi(0) - C(0)\phi(-\tau)\Big) + \int_{-\tau}^0 \mathfrak{Y}(s+\tau, t) B(s+\tau)\phi(s) ds + \\ - \int_{-\tau}^0 \frac{\partial}{\partial s}\Big\{\mathfrak{Y}(s+\tau,t)C(s+\tau)\Big\}\phi(s)ds + \mathfrak{C}_{\lfloor t/\tau \rfloor}(t)\phi(t - \lfloor \frac{t}{\tau} \rfloor \tau). \tag{4.23b}$$

Proof: Use (4.20) and (4.21); replace t in the NDDE (4.1) by s, multiply by $\mathfrak{Y}(s,t)$, and integrate for $s \in [0,t]$. The result follows after manipulation.

EXAMPLE 4.3 *For the solution of the DDE (3.6) ($C(t) = 0$ in the preceding theorem) $y(\phi; t) = Y(0,t)\phi(0) + \int_{-\tau}^0 Y(s+\tau, t)B(s+\tau)\phi(s)ds + \int_0^t Y(s,t)f(s)ds$.*

5 SENSITIVITY THEORY FOR DDES AND NDDES

The preceding results supply tools for an analysis of DDEs or NDDEs, respectively

$$y'(t) = A(t)y(t) + B(t)y(t-\tau) + f(t) \ (t \geq 0), \tag{5.1}$$

$$\frac{d}{dt}\{y(t) - C(t)y(t-\tau)\} = A(t)y(t) + B(t)y(t-\tau) + f(t) \ (t \geq 0), \tag{5.2}$$

in each case with $y(t) = \phi(t)$ ($t \in [-\tau, 0]$). The variation of parameter formulae for DDEs and for NDDEs, the related theory for Volterra integral equation theory, and the direct use of adjoint equations, all provide alternative means of deriving perturbation results.

Remark 5.1 While there is an extensive theory for integral equations, some care should be exercised when employing integral equation results in the case of DDEs or NDDEs. *General* perturbation theory will here permit changes that destroy the connection with DDEs or NDDEs. For precision one requires careful specification of the permitted admissible perturbations and $\mathbb{F}$.

5.1 Sensitivity of $y(\cdot)$ with respect to the coefficient functions

We can now indicate how the solutions $y(\cdot)$ of (5.1) or of (5.2) respond to perturbations of the coefficients $A(\cdot)$, $B(\cdot)$, $C(\cdot)$. With $\|\delta A(\cdot)\| = \|\delta B(\cdot)\| = \|\delta C(\cdot)\| = 1$ we change

$$A(\cdot) \text{ to } A(\cdot) + \varepsilon\delta A(\cdot), B(\cdot) \text{ to } B(\cdot) + \varepsilon\,\delta B(\cdot), \text{ and } C(\cdot) \text{ to } C(\cdot) + \varepsilon\,\delta C(\cdot); \tag{5.3a}$$

$$D(t) := B(t) + C'(t) \text{ changes to } D(t) + \varepsilon\delta D(t), \text{ with } \delta D(t) = \delta B(t) + \delta C'(t). \tag{5.3b}$$

These changes invoke changes of $Y(t,s)$, $\mathfrak{Y}(t,s)$, $\mathfrak{R}(t,s)$ from which one can deduce the change in $y(\cdot)$ to within $\mathcal{O}(\varepsilon^2)$.

THEOREM 5.1 *Given (5.3), $Y(s,t)$ changes to $Y(t,s) + \varepsilon\delta Y(t,s)\ +\ \mathcal{O}(\varepsilon^2)$ where*

$$\frac{\partial}{\partial s}\Big\{\delta\,Y(s,t)\ -\ \delta\,Y(s+\tau,t)C(s+\tau)\Big\}\ +\ \delta\,Y(s,t)A(s)\ +\ \delta\,Y(s+\tau,t)D(s+\tau)\ =$$

$$= -\frac{\partial}{\partial s}\{Y(s+\tau,t)\delta C(s+\tau)\} + Y(s,t)\delta A(s) + Y(s+\tau,t)\delta D(s+\tau). \quad (5.4)$$

and $\delta Y(t,s) = 0$ for $t \le s$. Further, $\mathfrak{C}_i(t)$ changes to $\mathfrak{C}_i(t) + \varepsilon\delta\mathfrak{C}_i(t) + \mathcal{O}(\varepsilon^2)$ where

$$\delta\mathfrak{C}_{-1} = 0,\ \delta\mathfrak{C}_0(t) = \delta C(t), \quad \delta\mathfrak{C}_r(t) = \delta\mathfrak{C}_{r-1}(t)C(t-r\tau) + \mathfrak{C}_{r-1}(t)\delta C(t-r\tau), \quad (5.5)$$

and the consequent change in $\mathfrak{Y}(t,s)$ is $\varepsilon\delta\mathfrak{Y}(t,s)\ +\ \mathcal{O}(\varepsilon^2)$ where

$$\delta\mathfrak{Y}(t,s) = \sum_{i=0}^{\lfloor t/\tau\rfloor - 1}\{\mathfrak{C}_i(t)\delta Y(s,t-i\tau) + \delta\mathfrak{C}_i(t)Y(s,t-i\tau)\}.$$

COROLLARY 5.2 *(a) Let $y(\cdot) \equiv y(A,B,C;\cdot)$ be the solution of (4.1) with a given (fixed) ϕ and τ, then $y(A+\varepsilon\,\delta A, B+\varepsilon\,\delta B, C+\varepsilon\,\delta C;t) - y(A,B,C;t) = \varepsilon\delta y(t) + \mathcal{O}(\varepsilon^2)$ where $\delta y(t) = \delta y_1(t) + \delta y_2(t)$ with*

$$\begin{aligned}
\delta y_1(t) &= \delta\mathfrak{Y}(0,t)\Big(\phi(0) - C(0)\phi(-\tau)\Big) + \int_{-\tau}^{0}\delta\mathfrak{Y}(s+\tau,t)B(s+\tau)\phi(s)ds+\\
&\quad -\int_{-\tau}^{0}\frac{\partial}{\partial s}\Big\{\delta\mathfrak{Y}(s+\tau,t)C(s+\tau)\Big\}\phi(s)ds + \int_{0}^{t}\delta\mathfrak{Y}(s,t)f(s)ds,\\
\delta y_2(t) &= \mathfrak{Y}(0,t)\Big(\phi(0) - \delta C(0)\phi(-\tau)\Big) + \int_{-\tau}^{0}\mathfrak{Y}(s+\tau,t)\delta B(s+\tau)\phi(s)ds+\\
&\quad -\int_{-\tau}^{0}\frac{\partial}{\partial s}\Big\{\mathfrak{Y}(s+\tau,t)\delta C(s+\tau)\Big\}\phi(s)ds + \delta\mathcal{C}(t)\phi(t-\lfloor\frac{t}{\tau}\rfloor\tau).
\end{aligned}$$

Remark 5.2 There are parallel theories in which one considers the first-order perturbations from an integral equation perspective. Theorem 5.4 provides an indication of this approach. *Another alternative is to base the development on first-order perturbations $\{\varepsilon\delta\hat{G}_r\}$ in (3.19).*

5.2 Sensitivity with respect to the initial function

A number of difficulties are encountered when τ, the time-lag, is regarded as a parameter defining a model NDDE and one considers the response of the solution to changes in τ (and hence, by inference, in ϕ). For reasons of space, we content ourselves with a single result for this case.

THEOREM 5.3 *The sensitivity of $y(\phi;t)$ with respect to the lag τ may not be a continuous function of t.*

We have enough theory to indicate how $y(\phi;\cdot)$ responds to changes in ϕ. For convenience we suppose that ϕ and $\delta\phi$ (with $\|\delta\phi\| = 1$) are both defined on a *fixed* interval $[-\tau, 0]$ and consider the solution $y(\phi + \varepsilon\delta\phi;\cdot)$ of the NDDE (5.2) – or, in particular, the DDE (4.1). We use the notation $\mathcal{D}$, with suitable embellishments, to denote a Gateaux derivative. In the case of the NDDE (5.2), Theorem 4.9 yields $y(\phi + \varepsilon\delta\phi; t) - y(\phi; t)$ and we define:

$$s(\phi, \delta\phi; t) := \lim_{\varepsilon\to 0} \frac{y(\phi + \varepsilon\delta\phi; t) - y(\phi; t)}{\varepsilon} \quad \text{where } \|\delta\phi\| = 1. \tag{5.6}$$

The limit in (5.6) is the *first-order sensitivity of* $y(\phi; t)$ *with respect to changes in* ϕ. It is the Gateaux (or directional) derivative $\mathcal{D}^{\delta\phi}y(\phi; t)$ of $y(\phi;\cdot)$ in the direction $\delta\phi$. (A reassessment of the notation $\{\partial/\partial\phi\}y$ in [16] is suggested.) Our definition (5.6) can also be applied when $y(\phi; t)$ is a solution of (2.6) or (2.7); however, for linear equations (5.2) $s(\phi, \delta\phi; t)$ is independent of ϕ.

Let $y_0(\phi; t)$ be the solution (with initial function ϕ), corresponding to vanishing f, defined in Corollary 4.4. Then $y_0(\delta\phi; t)$ is the corresponding solution with initial function $\delta\phi$ where $\|\delta\phi\| = 1$. The following theorem results from Theorems 4.3 and 4.9.

Theorem 5.4 *When $y(\phi;\cdot)$ is the solution of (5.2), we have $s(\phi, \delta\phi; t) = y_0(\delta\phi; t)$ (which is independent of ϕ), and in consequence*

$$s(\phi, \delta\phi; t) = \mathfrak{Y}(0, t)\Big(\delta\phi(0) - C(0)\delta\phi(-\tau)\Big) + \mathcal{C}(t)\delta\phi(t - \lfloor\frac{t}{\tau}\rfloor\tau) +$$
$$\int_{-\tau}^{0} \Big[\mathfrak{Y}(s + \tau, t)B(s + \tau) - \frac{\partial}{\partial s}\Big\{\mathfrak{Y}(s + \tau, t)C(s + \tau)\Big\}\Big]\delta\phi(s)ds. \tag{5.7}$$

Further, $s(\phi, \delta\phi; t)$ can also be written as $\gamma(\delta\phi; t) + \int_0^t \mathfrak{R}(t, s)\gamma(\delta\phi; t)\mathrm{d}s$.

5.3 Sensitivity based upon adjoint functions; parameter selection

Suppose $y(\mathrm{p};\cdot)$ satisfies (5.2), or the equivalent explicit form assuming C', ϕ' exist:

$$\mathcal{L}_{NDDE}\{y(\mathrm{p};\cdot)\}(t) = f(\mathrm{p}; t),\ t \in [0, T]; \quad y(t) = \phi(t),\ t \in [-\tau, 0] \tag{5.8}$$

and (based on data relating to observed phenomena) $\widehat{y} \in \mathbb{F}$ approximates y. The coefficients of the NDDE depend on p, $\mathcal{L}_{NDDE} = \mathcal{L}^{\mathrm{p}}_{NDDE}$, and we discuss sensitivity with respect to p; we assume τ, ϕ fixed. The analogue of (5.6) is the first order sensitivity of $y(\mathrm{p}; t)$ with respect to changes $\delta\mathrm{p}$,

$$\sigma(\mathrm{p}, \delta\mathrm{p}; t) := \lim_{\varepsilon\to 0} \frac{y(\mathrm{p} + \varepsilon\delta\mathrm{p}; t) - y(\mathrm{p}; t)}{\varepsilon} \quad (\text{where } \|\delta\mathrm{p}\| = 1). \tag{5.9}$$

Denoting Gateaux derivatives in the direction $\delta\mathrm{p}$ by the use of $\mathcal{D}^{\delta\mathrm{p}}$, (5.9) is $\mathcal{D}^{\delta\mathrm{p}}y(\mathrm{p};t)$. We suppose

$$y(\mathrm{p}+\varepsilon\delta\mathrm{p};t) = y(\mathrm{p};t) + \varepsilon\sigma(\mathrm{p},\delta\mathrm{p};t) + o(\varepsilon) \;\text{ as } \varepsilon \to 0 \tag{5.10}$$

is valid for $t \in [0,T]$. (Such a relationship is not always valid for every t (*e.g.*, if τ is a perturbed component of p), but it may be sufficient that it holds almost everywhere.)

Now suppose that $\mathcal{L}^{\mathrm{p}+\varepsilon\delta\mathrm{p}}_{NDDE}\{u\} = \mathcal{L}^{\mathrm{p}}_{NDDE}\{u\} + \varepsilon\{\delta\mathcal{L}^{\mathrm{p},\delta\mathrm{p}}_{NDDE}\}\{u\} + o(\varepsilon)$ and that $f(\mathrm{p}+\varepsilon\delta\mathrm{p};t) = f(\mathrm{p};t)+\varepsilon\delta f(\mathrm{p},\delta\mathrm{p};t)+o(\varepsilon)$ (uniformly for all u with sufficiently small $u(\cdot) - y(\mathrm{p};\cdot)$ and for $t \in [0,T]$) as $\varepsilon \to 0$. Thus, we write the Gateaux derivatives $\mathcal{D}^{\delta\mathrm{p}}\mathcal{L}^{\mathrm{p}}\{u\}$ as $\{\delta\mathcal{L}^{\mathrm{p},\delta\mathrm{p}}\}\{u\}$ and $\mathcal{D}^{\delta\mathrm{p}}f(\mathrm{p};t)$ as $\delta f(\mathrm{p},\delta\mathrm{p};t)$.

Lemma 5.5 *With the preceding notation,*

$$\mathcal{L}^{\mathrm{p}}_{NDDE}\sigma(\mathrm{p},\delta\mathrm{p};t) = \delta f(\mathrm{p},\delta\mathrm{p};t) - \{\delta\mathcal{L}^{\mathrm{p},\delta\mathrm{p}}_{NDDE}\}\{y(\mathrm{p};t)\} \text{ for } t \in [0,T], \tag{5.11a}$$

$$\sigma(\mathrm{p},\delta\mathrm{p};t) = 0 \text{ for } t \in [-\tau,0]. \tag{5.11b}$$

Adjoint theory can be used to discuss (5.9) but here we emphasize a different but related aspect: parameter selection; (*cf.* the treatment for DDEs in [32, §6.2]) A common issue is determination of $\mathrm{p}_\star \in \mathbb{P}$ to minimize

$$\Phi(\mathrm{p}) := \mathcal{J}\{y(\mathrm{p};\cdot) - \widetilde{y}(\cdot)\} \text{ where } \mathcal{J}\{u\} := \prec u, u \succ \tag{5.12}$$

and $\mathcal{J}\{u\}$ is defined on $[\mathbb{F}, \prec\cdot,\cdot\succ]$. Thus, $\Phi(\mathrm{p}) = \sum_{j=1}^{M}[y(\mathrm{p};t_j)-\widetilde{y}(t_j)]^{\mathrm{T}}\mathrm{W}_j[y(\mathrm{p};t_j)-\widetilde{y}(t_j)]$ with $\{t_j\}_1^M \in [0,T]$ gives $\Phi_{WLS}(\mathrm{p})$ in [8] (here, $\sum_{j=1}^{M} u_j^{\mathrm{T}}\mathrm{W}_j u_j$ is a discrete analogue of (u,u)). If the objective function Φ achieves a minimum at $\mathrm{p}_\star$ in the interior of $\mathbb{P}$, then $\mathcal{D}^{\delta\mathrm{p}}\Phi(\mathrm{p}) = \{\partial/\partial\varepsilon\}\{\Phi(\mathrm{p}+\varepsilon\delta\mathrm{p})\}_{\varepsilon=0}$, vanishes for $\mathrm{p} = \mathrm{p}_\star$ and all $\delta\mathrm{p}$ with $\|\delta\mathrm{p}\| = 1$.

Now, $\mathcal{D}^{\delta\mathrm{p}}\Phi(\mathrm{p}) = \mathcal{D}^{\delta\mathrm{p}}\mathcal{J}\{y(\mathrm{p};\cdot) - \widetilde{y}(\cdot)\}$ and, from (5.10),

$$\mathcal{J}\{y(\mathrm{p}+\varepsilon\delta\mathrm{p};\cdot)-\widetilde{y}(\cdot)\} = \mathcal{J}\{y(\mathrm{p};\cdot)-\widetilde{y}(\cdot)\}+2\varepsilon\prec y(\mathrm{p};\cdot) - D\widetilde{y}(\cdot), \sigma(\mathrm{p},\delta\mathrm{p};\cdot)\succ + o(\varepsilon), \tag{5.13}$$

as $\varepsilon \to 0$. The first-order sensitivity of $\Phi(\mathrm{p};\cdot)$ with respect to perturbations $\delta\mathrm{p}$ is the coefficient of ε in (5.13), namely $\mathcal{D}^{\delta\mathrm{p}}\Phi(\mathrm{p}) = 2\prec y(\mathrm{p};\cdot) - \widetilde{y}(\cdot), \sigma(\mathrm{p},\delta\mathrm{p};\cdot)\succ$. The evaluation of this term, which must vanish if p coincides with $\mathrm{p}_\star$ in the interior of $\mathbb{P}$, provides an opportunity to apply Theorem 4.8. Recall that $\mathcal{L}^{\ddagger}_{NDDE}$ is defined in (4.12).

Lemma 5.6 *Suppose that* $x(\cdot) \equiv x(\mathrm{p};\cdot)$ *satisfies*

$$\mathcal{L}^{\ddagger}_{NDDE}\{x\}(t) = \omega(t) \text{ with } \omega(t) := 2\{y(\mathrm{p};t) - \widetilde{y}(t)\},\; t \in [0,T], \tag{5.14a}$$

$$x(t) = 0,\; t \in [T, T+\tau], \tag{5.14b}$$

and $\sigma(\cdot) \equiv \sigma(\mathrm{p}, \delta\mathrm{p}; \cdot)$ *satisfies (5.11). Then, in terms of the classical inner product* $(\cdot,\cdot)$, x *is an adjoint function to* σ*; that is* $(x, \delta f(\mathrm{p}, \delta\mathrm{p}; \cdot) - \delta\mathcal{L}^{\mathrm{p},\delta\mathrm{p}} y(\mathrm{p}; \cdot)) = (\omega, \sigma)$.

The adjoint property of x (Definition 4.3) arises from the properties $\sigma(\mathrm{p}, \delta\mathrm{p}; t) = 0$, $t \in [-\tau, 0]$, $x(t) = 0$, $t \in [T, T+\tau]$ and integration by parts (see Lemma 4.5).

Theorem 5.7 *Suppose that* x *in Lemma 5.6 is a quasi-adjoint function to* σ*, so that* $\prec x, \delta f(\mathrm{p}, \delta\mathrm{p}; \cdot) - \delta\mathcal{L}^{\mathrm{p},\delta\mathrm{p}} y(\mathrm{p}; \cdot) \succ = \prec \omega, \sigma \succ + \kappa(x, \sigma)$*. Then,*

$$2 \prec y(\mathrm{p}; \cdot) - \widetilde{y}(\cdot), \sigma(\mathrm{p}, \delta\mathrm{p}; \cdot) \succ,$$

or $\mathcal{D}^{\delta\mathrm{p}}\Phi(\mathrm{p})$*, is expressible as*

$$\prec x, \delta f(\mathrm{p}, \delta\mathrm{p}; t) - \delta\mathcal{L}^{\mathrm{p},\delta\mathrm{p}} y(\mathrm{p}; t) \succ - \kappa(x, \sigma(\mathrm{p}, \delta\mathrm{p}; \cdot)).$$

Lemma 5.6 verifies the applicability of Theorem 5.7 in the case $\prec \cdot, \cdot \succ$ is the classical inner-product $(\cdot,\cdot)$ but to apply the theorem in the case of a pseudo-inner product requires the verification of the assumptions of the theorem and the evaluation of the term in κ.

Standard optimization algorithms for locating a minimum value $\Phi(\mathrm{p}_\star) = \mathcal{J}\{y(\mathrm{p}_\star; \cdot) - \widetilde{y}(\cdot)\}$ can encounter practical difficulties. Now, at a stationary value not on the boundary of $\mathbb{P}$, $\mathcal{D}^{\delta\mathrm{p}}\mathcal{J}\{y(\mathrm{p}; \cdot) - \widetilde{y}(\cdot)\}$ is to vanish, and Theorem 5.7 leads to a computational alternative to standard algorithms for minimizing Φ. The details are omitted for reasons of space.

References

[1] S.M. Andrew, C.T.H. Baker, G.A. Bocharov, Rival approaches to mathematical modelling in immunology, J. Comput. Appl. Math. 205 (2007) 669-686; published online 25 July 2006.

[2] C.T.H. Baker, Retarded differential equations, J. Comput. Appl. Math. 125 (2000), 309-335.

[3] C.T.H. Baker, E.O. Agyingi, E.I. Parmuzin, F.A. Rihan, Yihong Song, Sense from sensitivity and variation of parameters. Appl. Numer. Math. 56 (2006), no. 3-4, 397–412.

[4] C.T.H. Baker, G.A. Bocharov, Computational aspects of time-lag models of Marchuk type that arise in immunology. Russian J. Numer. Anal. Math. Modelling 20 (2005), 247–262.

[5] C.T.H. Baker, G.A. Bocharov, J.M. Ford, P.M. Lumb, S.J. Norton, C.A.H. Paul, T. Junt, P. Krebs, B. Ludewig, Computational approaches to parameter estimation and model selection in immunology, J. Comput. Appl. Math. 184 (2005) 50-76.

[6] C.T.H. Baker, G.A. Bocharov, C.A.H. Paul, F.A. Rihan, Modelling and analysis of time-lags in some basic patterns of cell proliferation. J. Math. Biol. 37, No.4, 341-371 (1998).

[7] C.T.H. Baker, P.M. Lumb, On integral equation formulations of a class of evolutionary equations with time-lag, J. Integral Equations Appl. 18 (2006) 227–247.

[8] C.T.H. Baker, G.A. Bocharov, C.A.H. Paul, F.A. Rihan, Computational modelling with functional differential equations: identification, selection, and sensitivity. Appl. Numer. Math. 53 (2005), no. 2-4, 107–129.

[9] C.T.H. Baker, C.A.H. Paul, Discontinuous solutions of neutral delay differential equations. Appl. Numer. Math. 56 (2006), 284–304.

[10] C.T.H. Baker, E.I. Parmuzin, Analysis via integral equations of an identification problem for delay-differential equations, J. Integral Equations Appl. 16, 111-135 (2004).

[11] A. Bellen, M. Zennaro, *Numerical methods for delay differential equations*, Oxford University Press, New York, 2003.

[12] R. Bellman, K.L. Cooke, *Differential-difference equations*, Academic Press, New York, 1963.

[13] H. Brunner, *Collocation methods for Volterra integral and related functional differential equations*, Cambridge University Press, Cambridge, 2004.

[14] C. Corduneanu, *Functional equations with causal operators*, Taylor & Francis, London, 2002.

[15] R. Courant, D. Hilbert, *Methods of mathematical physics.* Vol. I. Interscience Publishers, Inc. (John Wiley & Sons), New York, N.Y., 1953

[16] P.C. Das, S.S. Prabhu, Determination of the optimal initial function and the parameters in a system described by delay-differential equations. Journal of Optimization Theory and Applications, 40 (4), (1983) 583-593.

[17] O. Diekmann, S.A. van Gils, S.M. Verduyn Lunel, H-O. Walther, *Delay equations. Functional, complex, and nonlinear analysis*, Springer-Verlag, New York, 1995.

[18] A. Friedman *Foundations of Modern Analysis*, Dover. New York (1982).

[19] A. Halanay, *Differential equations: Stability, oscillations, time lags*, Academic Press, New York, 1966.

[20] J.K. Hale, *Theory of functional differential equations*, First edition, 1971; second edition, Springer, New York, 1977.

[21] J.K. Hale, M.A. Cruz, Existence, uniqueness, and continuous dependence for hereditary systems, Ann. Mat. Pura Appl. (4) 85, 1970, 63–81.

[22] J.K. Hale, K.R. Meyer, A class of functional equations of neutral type. Memoirs of the American Mathematical Society, No. 76. AMS, Providence, R.I. 1967.

[23] J.K. Hale, S.M. Verduyn Lunel, *Introduction to functional-differential equations*, Springer, New York, 1993.

[24] V.B. Kolmanovskiĭ, A.D. Myshkis, *Applied theory of functional-differential equations*, Kluwer Acad. Publ., Dordrecht, 1992.

[25] V.B. Kolmanovskĭ, A.D. Myshkis, *Introduction to the theory and applications of functional-differential equations*, Kluwer Acad. Publ., Dordrecht, 1999.

[26] V. Lakshmikantham, S.G. Deo, *Method of variation of parameters for dynamic systems*, Gordon and Breach, Amsterdam, 1998.

[27] L.A. Liusternik, V.J. Sobolev, *Elements of functional analysis* (transl. from the Russian by A.E. Labarre, H. Izbicki, H.W. Crowley), Frederick Ungar, New York, 1961.

[28] N. MacDonald, *Time lags in biological models*, Springer, Berlin-New York, 1978.

[29] G.I. Marchuk, *Adjoint equations and analysis of complex systems.* Revised by the author, and transl. from the 1992 Russian edition by G. Kontarev, Kluwer Acad., Dordrecht, 1995.

[30] G.I. Marchuk, *Mathematical models of immune response in infectious diseases.* Transl. from Russian by G. Kontarev & I. Sidorov. Kluwer, Dordrecht, 1997.

[31] G.I. Marchuk, V.I. Agoshkov, V.P. Shutyaev, *Adjoint equations and perturbation algorithms in nonlinear problems*, CRC Press, New York, 1996.

[32] G.I. Marchuk, V. Shutyaev, G.A. Bocharov, Adjoint equations and analysis of complex systems: application to virus infection modelling. J.Comput. Appl. Math. 184 (2005), 177–204.

[33] L.A. Segel, Simplification and scaling, SIAM Rev., 14 (1972) 547–551.

[34] K. Yosida, Functional Analysis, 6th edition, Springer, Berlin (1995).