ABSTRACT

ROBBINS, DANIELLE. Sensitivity Functions for Delay Differential Equation Models. (Under the direction of H.T. Banks.)

Delay differential equations are useful to model various biological, sociological, and physical processes in which there are hysteretic or memory effects. Nicholas Minorsky played a great role in establishing the use of these type of models for physical processes. From his work, physical processes like ship control systems are modeled using delay differential equations with delayed damping or delayed restoring force. G.E. Hutchinson also saw the importance of using delay systems to model ecological models. Hutchinson's equation, also known as the delay-logistic equation, is used to model population growth of a species.

For these biological and physical processes modeled using delay differential equations there are generally associated data sets. This data is used to estimate parameters within the model to gain the best predictive model for the process. When performing estimation procedures, parameter identifiability issues may occur resulting in unfavorable estimates. There also may not be enough data or repeated information in the data which will again produce unfavorable estimates. Sensitivity analysis improves the estimation process as traditional sensitivity functions can determine which parameters can be estimated and those that should be fixed. Generalized sensitivity functions will determine which regions in the data help estimate specific parameters. Thus using both type of sensitivity functions should lead to optimal parameter estimates.

We will derive and compute traditional sensitivity functions for the delay logistic model, delayed damping and restoring force harmonic oscillator models, as well as a sociological model for the behavior of alcoholics. We will also prove the existence of a solution for the derived sensitivity equation with respect to the delay. We will use the computed traditional sensitivity functions, to compute generalized sensitivity functions and illustrate the effect of a delay on generalized sensitivity functions (which provide insight on sensitivity of estimated parameters to data). We compare the numerical approximations of the generalized sensitivity functions for the delay-logistic equations to the equations without delay. From the traditional sensitivity functions and generalized sensitivity functions we simulate ideal data sets to obtain optimal estimates for the delay parameter τ via the inverse problem.

Sensitivity Functions for Delay Differential Equation Models

by
Danielle Robbins

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

Biomathematics

Raleigh, North Carolina

2011

APPROVED BY:

Ralph Smith	Hien Tran
Ronald Baynes	H.T. Banks Chair of Advisory Committee

DEDICATION

To Jesus, my parents, my grandparents, my brothers, my aunts and uncles, cousins (especially Shamel), Liz and Duby!!!

BIOGRAPHY

The author was born in Frederick, MD the only girl of 3 and the baby of Earl and Olivia Robbins. At a young age she loved all things science and mathematics, so it is only fitting she become a mathematician. She was also an avid athlete playing basketball, volleyball, and running track where she was state champion in the long jump in 2001. She attended the University of Maryland Baltimore County (UMBC) where she was on the track team and was a Meyerhoff Scholar, like her brother Darian, and the President was Dr. Freeman Hrabowski, an alum of Hampton University like her parents, grandfather, and oldest brother Dwaine. From UMBC she went on to attend Arizona State University where her Master's advisor was Dr. Carlos Castillo-Chavez. From here she went to NC State and finished here Doctorate under the direction of Dr. H.T. Banks. She will begin her career as a Research Analyst for the Southwest Interdisciplinary Research Center at Arizona State University.

ACKNOWLEDGEMENTS

I first would like to thank Jesus because without Him none of this would have been possible. My faith has sustained and kept me throughout this crazy experience called graduate school. I would like to thank my advisor Dr. Banks for his help because he has been more than an advisor but also a great mathematical father and I am appreciative. I would also like to thank Karyn Sutton who helped me get to NC State and has been supportive through everything. My friend, colleague, and fellow Biomathematican Kathleen Holm, you have been my greatest ally through out this whole ordeal and I truly love you! Thank you to Lesa Denning for all that you do and the CRSC for all of your support.

To my family, without you this would not of happened. Thank you to my parents who have followed me throughout my life and done whatever you could to make sure that I would be happy and successful. Thank you for loving me even when I was a brat, lol. Thank you for being great role models and instilling within me the importance of education. Thank you to my grandparents for all your love and support, for the summers of indulging my every want and need, and disciplining me when necessary, I love Jarratt, and Hampton, Va., they are a part of me. Thank you to my aunts and uncles and your listening to my mother's wishes every summer and giving me my workbooks to do. Thank you to Aunt Connie for making hands-on science experiments for me to do, and all the many history lessons. Thank you Aunt Rosie, Uncle T-man, and Aunt Miki, Ryan, Christian and Maya, when I moved back to the east coast your home was my home and it made school that much easier. Thank you for always listening and letting me be frank with my younger cousins. Thank you to Aunt BJ for your cards and loving emails and gifts, they were much appreciated! Thank you to Uncle Darrell and Aunt Karen for the many check-in calls, Aunt Sheila for your love and support over the years and for just being you. To Uncle Alfred and Aunt Gwen, and all of my great aunts and uncles, cousins etc. thank you for the love and support. Thank you to my brothers Dwaine, Darian, and Keith, I know I was annoying growing up but as we have gotten older you really have become great friends and great support. Thank you to my cousin, sister, bff, Shamel, girl what would I do without you????? God only knows we have been through many endeavors together and I am glad this one ends together too. To Mrs. Betty and the Green/Whiten family thanks for all your love and care throughout the years, you are my family.

To my friends, where would I be without such loving and caring individuals. Dr. Liz, you've been more than a friend a sister too, and thank you for everything. To Conswela, thank you for being my original mathematical support, as well as my bff in college and beyond, I love you, Ant and Gigi. To Shatema, thanks girl for never letting anyone get between our friendship I so appreciate it. To Nicole, Torey, and Chrisitie where would I be without my Mt. Peace girls!

To the Goodman's and Jenkins', thank you for your support and encouragement I grateful to have two other sets of parents! To my church families (Mt. Peace, Tanner Chapel, and Quinn Chapel), Thank You, Thank You, Thank You, you've been a home away from home for me, no matter where I go you are always with me.

To this guy named Chidozie Chukwumah, you rock, and are pretty awesome, glad I met you 6 months before I finished my Doctorate, it makes the perfect ending to my graduate school experience...Love you!

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

Introduction

1.1 History of Delay Differential Equations

Delay differential equations (DDEs) are used to model biological, physical, and sociological processes, as well as naturally occurring oscillatory systems. Minorsky in 1942 first introduces the idea of hysterodifferential equations in [38], using these type of equations to model self-excited oscillatory dynamical systems. He proposes the idea that some natural phenomena such as self-oscillations may be effected by the previous history of a motion or action, which describes a retarded dynamical system. A retarded dynamical system is a system that describes an action that has delayed time dependence [39]. Minorsky explains the importance of these systems due to their ability to model self-excitation within a control system. These physical systems are usually classified into systems with retarded damping given by

$$\ddot{x}(t) + K\dot{x}(t-\tau) + bx(t) = g(t), \tag{1.1}$$

and those with retarded restoring force described by

$$\ddot{x}(t) + K\dot{x}(t) + bx(t-\tau) = g(t), \tag{1.2}$$

where g is some external force. Minorksy used models such as (1.1) and (1.2) to study stabilization systems in ships [38]-[40]. He also defines the difference between ordinary differential equations (ODEs) and functional differential equations (FDEs) in the following manner. A ODE is an equation where the solution at the present time predicts the future, while a DDE/FDE is defined such that the present and future is dependent upon past in some form [40]. A DDE also has a infinite degree characteristic equation, while the degree of the characteristic equation for an ODE is finite [40]. The infinite degree of the characteristic equation of a DDE allows for infinite eigenvalues of the DDE, which promotes different solution behavior such as self-excited oscillations in the solution [40]. This property of the DDE makes it very advantageous in modeling physical control systems. Minorsky also gives insight to the necessary use a non-linear DDE to model a system with self-excited oscillations, as a linear DDE is unable to capture all of the properties of the self-excitation. This is a result of the unstable conditions on the harmonic root [40]. Minorsky lays a foundation for modeling oscillatory phenomena in physical systems. From his research we are introduced to delayed restoring force and delayed damping DDE models which prove useful for modeling many physical and biological processes.

In 1948 Hutchinson developed a delay differential equation model, known as the delay logistic equation, to describe the dynamics of a circular causal system [32]. A causal system is a system where the output depends on the current and/or past input. A circular causal system is any causal system where changes to one part of the system effects another part of the system at a different rate so that the system does not go extinct. Parasite host interaction is an example of an ecological circular causal system since if a parasite can complete its life cycle without killing the host or drastically altering the growth of the host population, the host population will continue to exist [34, 32]. The delay in this model can represent various naturally occurring attributes of the process being modeled like the gestation period in a growing population, or the life cycle of a parasite. Hutchinson's equation maybe used to model population growth, host population growth in the presence of a parasite, and various other biological and physical processes. From Minorsky and Hutchinson we learn early the importance of using DDE models.

1.2 Previous Works for Parameter Identification Problems and Sensitivity Analysis in DDE Models

Usually for biological and physical processes modeled there are associated data sets from corresponding experiments. From these data sets and using the mathematical model that describes the process, we can perform the inverse problem to estimate the parameters within the model. In order to perform the inverse problem, the problem must be well-posed. The model parameters must also be identifiable. In addition to these issues, a solution to the DDE model must exist.

When dealing with the parameter identifiability issues, traditional sensitivity functions (TSFs) are computed to give insight as to which parameters in the model solutions are most sensitive. As a result, a group of parameters may be fixed, which improves the parameter estimation process. Traditional sensitivity functions are usually derived from the model equation and will have the same form as the model equation, (i.e., if the mathematical model is a DDE, the derived sensitivity equation will also be a DDE). For these derived TSF equations, existence must also be determined. In our efforts, we perform a complete sensitivity analysis for all the parameters including the delay to help with parameter estimation for DDE models. This analysis will involve computing TSFs and generalized sensitivity functions (GSFs). GSFs will give insight as to how sensitive the estimate is to the data and determine regions in the data of high information content. Thus from sensitivity analysis the TSFs will determine which parameters maybe fixed and the GSFs will determine when and where data should be collected, therefore improving the parameter estimation process.

A main issue when performing sensitivity analysis for DDE models occurs when deriving the equation for sensitivity with respect to the delay. For this particular sensitivity equation proving existence is not straight forward as the solution is dependent on the derivative of the previous history of the solution to the model, which we will see later is $\dot{x}(t-\tau)$. One main goal is to reformulate this sensitivity equation such that we can provide the theory to establish a well-posed problem. We begin by discussing previous work done on DDE parameter identification problems, followed by a discussion of previous work on sensitivity analysis for differential equation models. We compare and contrast the ideas presented in these works with our approach to the problem. We present theoretical and numerical results for multiple examples.

In the summarized works on parameter identification problems for DDE models, we learn that issues within parameter estimation leads to use of sensitivity equations to improve the estimation process. Also the theory presented in these works helps to direct our efforts for numerical computation of sensitivity functions. Next we present the works chronologically and summarize the results.

Banks, Burns, and Cliff in 1981 compute parameter identification problems for delay systems [7]. Their research develops estimation algorithms for the parameters of the delay systems including estimation of the delay. They observed difficulty when estimating the delay since solutions of DDEs are not always differentiable with respect to the delays, which makes estimation methods such as least squares and maximum likelihood challenging. Banks, et al., also suggest the need for formal theory regarding the existence of sensitivity functions with respect to the delay. They formulate a class of estimation algorithms based on previous general approximation techniques for delay systems, and consider the following delay system identification problem

$$\begin{aligned} \dot{x}(t) &= L(q)x_t + B(\alpha)u(t),\ t \geq 0, &(1.3)\\ x(0) &= \eta,\ x_0 = \phi\ ,\ (\eta, \phi) \in \mathbb{R}^N \times L_2(-r, 0; \mathbb{R}^N) &(1.4)\end{aligned}$$

with output $y(t) = C(\alpha)x(t) + D(\alpha)u(t)$. Here the function x, x_t is defined such that $x_t(\theta) = x(t+\theta)$. Also given $r > 0$ exists and is fixed, and $\Omega \subset \mathbb{R}^\mu$ is a compact convex set, they define a compact convex parameter set $Q \subset \mathbb{R}^{\mu+\nu}$ by $Q \equiv \Omega \times \mathcal{H}$, where

$$\mathcal{H} = \{h = (r_1, r_2, \ldots, r_\nu) \in \mathbb{R}^\nu | 0 \leq r_i \leq r_{i+1}, i = 1, \ldots, \nu - 1\}.$$

Then for any $q = (\alpha, h) \in Q$, $L(q)$ is defined such that

$$L(q)\phi = \sum_{i=0}^{\nu} A_i(\alpha)\phi(-r_i) + \int_{-r_\nu}^{0} K(\alpha, \theta)\phi(\theta)d\theta.$$

Next, they introduce a more cohesive theoretical foundation for the identification problem of a delay system. Given the fundamental identification problem for the delay systems is linear, the problem can be reformulated in an abstract way such that its solution can be defined by a strongly continuous semigroup of linear operators. Thus equation (1.3)-(1.4) becomes

$$\dot{z}(t) = \mathcal{A}(q)z(t) + \mathcal{B}(\alpha)u(t), \ t \geq 0 \tag{1.5}$$

$$z(0) = (\eta, \phi), \tag{1.6}$$

$$y(t) = \mathcal{C}z(t) + D(\alpha)u(t). \tag{1.7}$$

where $(\eta, \phi) \in Z \equiv \mathbb{R}^N \times L_2(-r, 0; \mathbb{R}^N)$, $q \in Q$ and the infinitesimal generator $\mathcal{A}(q)$ is defined such that

$$\mathcal{A}(q)(\phi(0), \phi) = (L(q)\phi, \dot{\phi}).$$

Then given $t \geq 0$, $S(t; q) : Z \to Z$ is defined such that

$$S(t; q)(\eta, \phi) = (x(t; q), x_t(q)),$$

where $S(t; q)$ is a strongly continuous semigroup. By defining a strongly continuous semigroup, from which a solution maybe obtained, a well-posed identification problem can be formulated. Banks et al., then prove approximation theorems, using semigroup theory, for the original model. As a result of the approximation theory established in their paper, computationally efficient identification algorithms are established. We will use the abstract formulation with our example to prove existence of a solution for the derived TSF equations with respect to the delay. We note that the original DDE may not have an existing solution if the initial history is not continuous over the time interval. For our DDE problem, the derived sensitivity function with respect to the delay, not only must the history of the solution, $x(t - \tau)$ be continuous, but the derivative of that history, $\dot{x}(t - \tau)$, must also be defined over the time interval.

From Banks, Burns, and Cliff we learn the importance of reformulating the model in an abstract space. Next Brewer presents theory on *Frechet* derivatives for the solution of a linear abstract Cauchy problem [19]. The model observed is an autonomous linear abstract Cauchy problem

$$\begin{aligned} \dot{z}(t) &= A(q)z(t) + u(t), \\ z(0) &= z_0, \end{aligned} \tag{1.8}$$

where $z \in Z$, a Banach space with norm $||\cdot||$, $q \in Q$, a normed linear space with norm $|\cdot|$, and z_0 is the initial condition. Brewer sets up criteria such that differentiability of the solution with respect to a parameter may be established for both linear homogeneous and linear inhomogeneous equations. The operator in this paper, $A(q)$, that defines the linear abstract Cauchy problem must be such that the parameter q may appear in unbounded terms. In previous work by Gibson and Clark [31], the differentiability results for this type of linear abstract Cauchy problem were obtained when the operator $A(q)$ was represented as a linear combination of an operator independent of the parameter and a dependent bounded linear operator, $A(q) = A + B(q)$. Brewer expands the class results in [31] by considering the operator, $A(q)$, to represent the parameter in an unbounded format. The operator will generate a strongly continuous semigroup, and using semigroup theory Brewer is able to prove the existence of *Frechet* derivatives with respect to the parameters for the initial value problem (1.8). The solution to this problem via semigroups is $S(t;q)z_0$ when $u = 0$. As a result of the existence of the *Frechet* derivatives, he is able to carefully define sensitivity equations with respect to the parameters including the delay of the abstract system. Brewer applies his theory to a linear discrete delay system of the following format

$$\dot{x}(t) = a_0 x(t) + \sum_{k=1}^{n} a_k x(t - p_k) + u(t)$$

to show the application of his result. His theory is formulated on the Banach space $\mathbb{R} \times L^1(-p^*, 0)$, but maybe be extended to the Hilbert space $\mathbb{R}^m \times L^\nu(-p^*, 0)$ where $m, \nu \geq 1$. To use the results from this paper, there must be a general abstract *linear autonomous* system differentiable with respect to the parameter. Since our problem is non-linear and non-autonomous, Brewer's results are not readily extendable to our example.

Banks in a later work discusses spline methods for nonlinear delay systems which aides in performing the parameter identification problem [5]. Results in this paper are reintroduced in [13] by Banks and Lamm. In both of these papers, approximation theory for a general class of nonlinear functional differential equations is presented. This theory is proven without using the Trotter-Katto method and semigroup theory, but rather Grownwall's inequality and fixed point theory which are the methods we use to prove well-posedness for our problem. Banks considers the following system:

$$
\begin{aligned}
\dot{x}(t) &= f(x(t), x_t, x(t-\tau_1), \ldots, x(t-\tau_\nu)) + g(t),\ 0 \le t \le T \\
x_0 &= \phi,
\end{aligned}
\tag{1.9}
$$

where $f : Z \times \mathbb{R}^{n\nu} \to \mathbb{R}^n$, and $Z = \mathbb{R}^n \times L_2(-r, 0; \mathbb{R}^n)$. A nonlinear operator $\mathcal{A}$ is defined such that when reformulating (1.9) on Z the solution $z(t; \phi, g) = (x(t; \phi, g), x_t(\phi, g))$ will be a unique solution. Banks and Lamm [13] extend the definition of the operator $\mathcal{A}$ to be dependent on both the parameter and time, and are able to show existence and uniqueness of a solution $z(t; \phi, g) = (x(t; \phi, g), x_t(\phi, g))$ in Z. Although we employ the same techniques, Banks assumes a dissipative condition on the nonlinear operator for a class of functional differential equations to which our example does not fall. In this paper Banks' main results show convergence of the approximating solutions using piecewise linear splines and proves well-posedness for a class of FDEs; however, the theory presented is not applicable to our class of FDE, which includes sensitivity equations for the delay parameter. The theory is not readily extendable for our class of FDEs because the right hand side of the derived sensitivity equation for derivative of the solution with respect to the delay is driven by the derivative of the history of the original solution. Thus we need different continuity requirements for our initial function.

In 1989, Brewer, Burns, and Cliff [20] carried out the parameter identification problem for an abstract Cauchy problem using quasilinearization. The linear abstract Cauchy problem is defined in (1.8). Given a solution to an abstract Cauchy problem is dependent upon a parameter, and the Cauchy problem is defined by an unbounded parameter dependent evolution operator, $A(q)$, their goal was to establish convergence of the gradient-based parameter estimation algorithm. The use of quasilinearization with parameter estimation requires the derivative of the solution with respect to the parameter to be known (i.e., the gradient must exist). As a result Brewer, et al., show existence of *Frechet* derivatives with respect to the parameters, *including the delay*, using semigroup theory as applied to an autonomous linear delay differential equation. It is assumed that $A(q)$ generates a strongly continuous semigroup $S(t, q)$ on some Banach space with a norm, X. The *Frechet* derivative with respect to the delay will exist based on the theory in this paper if the right hand side of the linear abstract Cauchy problem is not dependent on the derivative of the previous history of the original solution. In our problem, the derived sensitivity equation with respect to the delay, is dependent on the derivative of the previous history of the original solution and is a nonlinear delay differential equation model. Thus we are not readily able to apply the theory from this paper to our example.

Banks, Banks, and Joyner [10] present a mathematical and statistical framework for performing the inverse problem on differential equations with history. They detail how to perform the inverse problem on a DDE system and determine which methods best estimate parameters in the model. The following system of functional differential equations is the example for which the mathematical and statistical framework is based. The system of equations models the changes in an insect population due to insecticide and is given by,

$$
\begin{aligned}
\frac{dA}{dt}(t) &= \int_{-7}^{5} N(t+\tau)m(\tau)d\tau - (d_A(t)+p_A(t)) \\
\frac{dN}{dt}(t) &= b(t)A(t) - (d_N(t)+p_N(t))N(t) - \int_{-7}^{-5} N(t+\tau)m(\tau)dt \\
A(\theta) &= \phi(\theta),\ N(\theta)=\psi(\theta),\ \theta\in[-7,0) \\
A(0) &= A^0,\ N(0)=N^0.
\end{aligned}
\tag{1.10}
$$

Here $A(t)$ is the number of adult insects, $N(t)$ is the number of neonate insects, and m is a probability density kernel with specific assumed properties. To approximate the solution to this model, they use Banks-Kappel splines since the model may be reformulated into an abstract evolution equation. We will discuss the Banks-Kappel method in a later section. In this paper Banks et al., discuss the use and formulation of sensitivity equations with respect to the parameters and density kernel m but do not present a formal proof on the existence of the *Frechet* derivatives that define the sensitivity equations. They do however reference a formal proof presented in [11], which uses a theoretical framework presented in [16]. From Banks, Banks, and Joyner we observe the necessity to formulate and compute sensitivity functions with respect to the delay to aide parameter estimation because in their example, like in many other DDE models, the delay parameter in particular has the least amount of information given from the data and research.

Banks, Rehm, and Sutton [18] study inverse problems for nonlinear delays systems. They give a theoretical framework for the convergence of approximations for nonautonomous nonlinear DDE models. To establish this convergence, they begin by determining if solutions to the nonautonomous nonlinear DDE exist. To do this they reformulate their model

$$\dot{x}(t) = f(t,x(t),x_t,x(t-\tau_1),\dots,x(t-\tau_m),q)+f_2(t),\ 0\le t\le T, \tag{1.11}$$

$$x_0 = \phi, \tag{1.12}$$

where $f:[0,T]\times X\times\mathbb{R}^{nm}\times\mathcal{Q}\to\mathbb{R}^n, X=\mathbb{R}^n\times L_2(-r,0;\mathbb{R}^n)$, and $\phi\in H^1(-r,0)$, into the abstract form

$$\begin{aligned} \dot{z}(t) &= \mathcal{A}(t;q)z(t) + (f_2(t), 0) \\ z(0) &= \xi = (\phi(0), \phi), \end{aligned} \tag{1.13}$$

where $\mathcal{A}(t;q) : D(\mathcal{A}) \subset X \to X$ is the nonlinear operator. They then determine if the abstract form has an existing solution. Banks et al.[18], suggests the existence of the solution can be established via fixed point theory and Picard iteration arguments. Existence and uniqueness for the solution of the approximation of (1.13) is established using the approach presented in [12]. Their main result ensures the convergence of solutions of the approximation to that of the original model. Banks, Rehm, and Sutton do not observe or formulate sensitivity equations; however, they do give insight to a theoretical framework for establishing existence and uniqueness for a DDE model. We use techniques presented here to establish well-posedness for our derived sensitivity equations with respect to the delay.

Next we describe previous work done on the computational aspects of sensitivity analysis for delay differential equations (DDEs). To be specific we discuss literature that uses sensitivity analysis on various DDE models [6, 17, 22] and the theory presented within these references. As we have previously mentioned, sensitivity analysis improves the parameter identification for process for DDE systems [5, 7, 13, 20], which use sensitivity functions in the identification process. From these works we discuss previous theoretical and computational tools and how there is a lack of proof of well-posedness for derived sensitivity equations with respect to the delay.

Baker and Rihan [17] show how to formally derive sensitivity equations for delay differential equation models, as well as the derivation of equations for the sensitivity of parameter estimates with respect to observations, what we know as GSFs. They consider a general system of delay differential equations such that $p \in \mathbb{R}^L$

$$\begin{aligned} \dot{x}(t,p) &= f(t, x(t), x(t-\tau), p),\ t \geq 0, \\ x(t,p) &= \psi(t,p),\ t \leq 0. \end{aligned} \tag{1.14}$$

From this general model local sensitivity functions, $\frac{\partial x}{\partial p_i}$, are obtained by solving

$$\begin{aligned} \frac{\partial \dot{x}(t,p)}{\partial p_i} &= \frac{\partial}{\partial p_i} f(t, x(t), x(t-\tau), p),\ t \geq 0, \\ \frac{\partial x(t,p)}{\partial p_i} &= 0. \end{aligned} \tag{1.15}$$

They derive the sensitivity of the optimum parameter estimate $\hat{p}$ to perturbations in the data

η_j by first defining the cost functional $\phi(p, \eta)$

$$\phi(p, \eta) = \sum_i \left[x(t_i, p) - \eta_i\right].^2 \tag{1.16}$$

Next they differentiate the cost functional and obtain

$$\frac{\partial}{\partial p_k}\phi(p, \eta) = 2\sum_i \left[x(t_i, p) - \eta_i\right] \frac{\partial x(t_i, p)}{\partial p_k}. \tag{1.17}$$

If the cost function (1.16) is minimized a $p = \hat{p}$ where $\hat{p} \equiv \hat{p}(\eta)$, then (1.17) is equal to zero. So (1.17) becomes

$$\sum_i \left[x(t_i, \hat{p}) - \eta_i\right] s_k(t_i, \hat{p}(\eta)) = 0, \tag{1.18}$$

where $s_k(t_i, p) = \frac{\partial x(t_i, p)}{\partial p_k}$. Both sides are then differentiated with respect to η_i so

$$\sum_{i=1}^{N}\sum_{l=1}^{L_p} \left[s_k(t_i, \hat{p}) s_l(t_i, \hat{p}) + \left[x(t_i, \hat{p}) - \eta\right] r_{lk}(t_i, \hat{p})\right] \frac{\partial \hat{p}_l}{\partial \eta_j} = s_k(t_j, \hat{p}). \tag{1.19}$$

Given that $x(t_i, \hat{p})$ is close to the observation η_i, $[x(t_i, \hat{p}) - \eta] = 0$, (1.19) becomes

$$\sum_{i=1}^{N}\sum_{l=1}^{L_p} \left[s_k(t_i, \hat{p}) s_l(t_i, \hat{p})\right] \frac{\partial \hat{p}_l}{\eta_j} \approx s_k(t_j, \hat{p}), \tag{1.20}$$

which can be rewritten as

$$\left[\sum_{i=1}^{N} s(t_i, \hat{p}) s^T(t_i, \hat{p})\right] \frac{\partial \hat{p}}{\partial \eta_j} \approx s(t_j, \hat{p}). \tag{1.21}$$

Thus

$$\frac{\partial \hat{p}}{\eta_j} \approx \left[\sum_{i=1}^{N} s(t_i, \hat{p}) s^T(t_i, \hat{p})\right]^{-1} s(t_j, \hat{p}).$$

Baker and Rihan also show that the sensitivity of the parameter estimates to observations (what we know as GSFs) maybe obtained by minimizing the previously defined objective function $\phi(p)$ in the following way

$$\frac{\partial}{\partial p}\phi(\hat{p}) = 2\sum_{j=1}^{N} \chi^T(t_j, \hat{p})[x(t_j, \hat{p}) - \eta_j] = 0, \tag{1.22}$$

where $\chi(t, \hat{p})$ is the sensitivity matrix.

Baker and Rihan also offer an outline of how to numerically compute both TSFs and GSFs for retarded delay differential equations and neutral delay differential equations. Baker and Rihan list issues that arise when doing parameter estimation in DDEs which includes difficulty in establishing existence of the derivative of the solution with respect to the parameters, as well as difficulty in establishing a well-posed problem for the derived sensitivity equations. The issues raised by Baker and Rihan are common when dealing with delay differential equations and we attempt to address and solve these issues within the current text.

Banks and Bortz [6] use sensitivity analysis to show how changes in parameters will effect the solutions of their delay differential equation model for HIV which has distributed delays. Their sensitivity equations are formulated from the following model:

$$\begin{aligned} \dot{x}(t) &= L(x(t), x_t) + f_1(x(t)) + f_2(t) \text{ for } 0 \leq t \leq t_f \\ (x(0), x_0) &= (\phi(0), \phi) \in \mathbb{R}^4 \times C(-r, 0; \mathbb{R}^4), \end{aligned} \tag{1.23}$$

where $x(t) = (V, A, C, T),^T$ which are the states used in the mathematical model for the HIV model, and $L(\eta, \phi)$ is a vector operator to account for the distributed delay. When deriving the sensitivity equations they obtain a system of DDEs, which are assumed to be well-posed. In their discussion of well-posedness for these sensitivity equations they assume the delay distributions are differentiable and parameterizable by a mean and standard deviation. In this paper they use theoretical steps presented in [5] to prove existence and uniqueness of the derived sensitivity equations (i.e., successive approximations, fixed point theory, Lipschitz continuity, etc.). While they are able to prove well-posedness for their particular class of sensitivity delay differential equations, their example does not have sufficient smoothness on the initial history functions, as it is assumed in $C(-r, 0; \mathbb{R}^4)$, such that the derivative of the history function will be defined over the time interval. Banks and Bortz also state that existence of the derived sensitivity function with respect to discrete delays can be established using the same manner of proof as their example; however, the arguments for this proof will be more tedious. From Banks and Bortz we gain more insight to the theoretical framework needed to prove a well-posed problem for derived sensitivity equations.

Banks and Nguyen [16] develop a theoretical framework for sensitivity functions of nonlinear dynamical systems in a Banach space where the parameters are dependent on a another Banach space. They observe the sensitivity of functional parameters in the following type of nonlinear ordinary differential equations

$$
\begin{aligned}
\dot{x}(t) &= f(t, x(t), \mu),\ t \geq t_0 \\
x(t_0) &= x_0,
\end{aligned}
\tag{1.24}
$$

where $f : \mathbb{R}_+ \times X \times \mathcal{M} \to X$ and $\mathcal{M}$ and X are complex Banach spaces. They establish well-posedness for (1.24), and existence of *Frechet* derivatives with respect to the solution $x(t)$ and parameter(s) μ. As a result, there is a solution to the respective sensitivity equation

$$
\begin{aligned}
\dot{y}(t) &= f_x(t, x(t, t_0, x_0, \mu), \mu)y(t) + f_\mu(t, x(t, t_0, x_0, \mu), \mu), t \geq t_0 \\
y(t_0) &= 0,
\end{aligned}
\tag{1.25}
$$

where $y(t) = \frac{\partial x(t)}{\partial \mu}$. To prove existence of (1.25), they use fixed point, successive approximations, and Gronwall's inequality theories. Banks and Nguyen apply their theoretical results to the DDE example for HIV dynamics, however they only display results for the the sensitivity with respect to the probability distribution for the delay and not a discrete delay which would require more details in the proof of existence. We follow closely the theoretical arguments within this paper and extend them for use in the proof of the existence of the *Frechet* derivative with respect to the delay for our examples.

Burns, Cliff, and Doughty [22] explain the use of continuous sensitivity equations for DDE models typical in biosciences, specifically for a model for Chlamydia Trachomatis. Their research focuses on parameter estimation techniques using sensitivity equations for DDE models specific to the biosciences. Parameter estimation for these typical models involves the inverse problem which uses sensitivity equations both theoretically and computationally. Burns et al., explore the parameter estimation problem for an ODE/PDE model of the cellular dynamics for Chlamydia Trachomatis which is then transformed into a simple DDE model to aid in the formulation of the continuous sensitivity functions. The following model describes the cellular changes between the retriculate body (RB) and extracellular elementary body (EB) in Chlamydia

$$
\frac{d}{dt}RB(t) = a(t)RB(t),\ t_1 < t < t_3, \tag{1.26}
$$

$$
a(t) = \begin{cases} \alpha,\ t_1 < t < t_2, \\ \alpha - \beta((t - t_2)/(t_3 - t)),\ t_2 < t < t_3, \end{cases} \tag{1.27}
$$

$$
\frac{\partial}{\partial t}\rho(t, s) = -\frac{\partial}{\partial s}\rho(t, s),\ t_2 < t < t_3,\, 0 < s < r \tag{1.28}
$$

$$
\frac{\partial}{\partial t}\rho(t, 0) = -\kappa\rho(t, 0) + \beta((t - t_2)/(t_3 - t))RB(t), \tag{1.29}
$$

where $\rho(t,s)$ is the number of RB cells that transform to EB cells. This ODE/PDE can then be transformed to the following DDE

$$\begin{aligned} \dot{x}(t) &= A_0(q)x(t) + A_1 x(t-r) + Bv(t;q),\ 0 < t < T, \\ x(0) &= \xi \in \mathbb{R}^n \text{ and } x(s) = \phi(s) \in L^2(-r,0;\mathbb{R}^N),\ -r < s < 0. \end{aligned} \tag{1.30}$$

Once the ODE/PDE model for Chlamydia Trachomatis is transformed into the simple DDE, it is then transformed to a Cauchy Problem on the state space $Z = \mathbb{R}^N \times L^2(-r,0;\mathbb{R}^N)$

$$\frac{dz(t;q,r)}{dt} = \mathcal{A}(q,r)z(t;q,r) + \mathcal{B}v(t;q), \tag{1.31}$$

$$z(0;q,r) = [\xi, \phi_0(\cdot)]^T \in Z \tag{1.32}$$

where

$$D(\mathcal{A}(q,r)) = \{[\xi,\phi(\cdot)]^T : \xi \in \mathbb{R}^N,\ \phi(\cdot) \in H^1(-r,0;\mathbb{R}^N),\ \phi(0) = \xi\} \tag{1.33}$$

$$\mathcal{A}(q,r)\begin{bmatrix} \xi \\ \phi(\cdot) \end{bmatrix} = \begin{bmatrix} A_0(q)\xi + A_1\phi(-r) \\ \phi'(\cdot) \end{bmatrix}, \tag{1.34}$$

and

$$\mathcal{B} = \begin{bmatrix} B \\ 0 \end{bmatrix}. \tag{1.35}$$

Transformation to the Cauchy problem allows for use of specific numerical schemes to simulate the solution as established in previous literature [7, 12]. From the simple DDE (1.30) for the model they formulate the following continuous sensitivity equations

$$\frac{\partial \dot{x}(t;q,r)}{\partial q} = A_0(q)\frac{\partial x(t;q,r)}{\partial q} + [\frac{A_0(q)}{\partial q}]x(t:q,r) + A_1\frac{\partial x(t-r;q,r)}{\partial q} + B\frac{\partial v(t;q)}{\partial q} \tag{1.36}$$

$$\frac{\partial \dot{x}(t;q,r)}{\partial r} = A_0(q)\frac{\partial x(t;q,r)}{\partial r} + A_1\frac{\partial x(t-r;q,r)}{\partial r} - A_1\frac{\partial x(t-r;q,r)}{\partial t} \tag{1.37}$$

which have a zero initial condition and initial function. Well-posedness is easily established for the sensitivity equations for the Chlamydia model with respect to parameters which are not the delay from semigroup theory for a linear autonomous operator $(A_0(q) + A_1)$ as discussed in [7]. They do not attempt to directly establish well-posedness for the sensitivity equation with respect to the delay. They only discuss ensuring that the initial past history of the solution to the original model is defined since it is a forcing term in the definition of this particular sensitivity equation, as shown in (1.37). Burns et al., give numerical results for estimated parameters of the Chlamydia model as well as sensitivities using SPLINE approximations coupled with MATLAB's ODE 23. Burns et al., also use the Gauss-Newton procedure along with a step-size selection scheme to minimize the least squares cost functional. From this paper the authors give insight to numerical schemes for computing sensitivity functions for DDE models as well as parameter estimation techniques. They also highlight well-posedness issues for the sensitivity equation with respect to the delay.

From the various works summarized above we are able to obtain a theoretical foundation to establish existence and uniqueness for our derived sensitivity equations. We are also able to obtain insight for approximating these derived sensitivity equations. These works gives us the necessary background to accomplish our main task of determining existence and uniqueness for the derived sensitivity equation with respect to the delay parameter.

Chapter 2

Theoretical Framework

2.1 The Basic Model

We begin with the following equation model which may describe a physical, sociological or biological system:

$$\frac{dx(t)}{dt} = G(x(t), x(t-\tau), \theta), \quad t > 0 \tag{2.1}$$

$$x(\xi) = \begin{cases} \Phi(\xi), & -\tau \leq \xi < 0 \\ x_0, & \xi = 0 \end{cases} \tag{2.2}$$

$$\eta(t) = h(x(t), x(t-\tau), \theta), \quad t \in [0, T], \tag{2.3}$$

where $x(t) \in \mathbb{R}^n$ are vectors of state variables of the system, $\eta(t) \in \mathbb{R}^m$ is the vector of measurable or observable outputs, and $\theta \in \mathbb{R}^p$ is the vector of parameters. It is assumed that G, and h in (2.1) and (2.3) are sufficiently smooth, in order to carry out the construction of the generalized sensitivity functions (GSFs). GSFs determine how sensitive the parameter estimate is to the observations or data.

When solving (2.1), we obtain $x = x(t, \theta)$, where t $\in [0, T]$, and

$$\eta(t) = f(t, \theta), \quad t \in [0, T],$$

and f is defined as $f(t, \theta) = h(x(t, \theta), x(t-\tau, \theta), \theta)$, where $h = h(x, \tilde{x}, \theta)$ and $\tilde{x} = x(t-\tau)$. We then can define the traditional sensitivity functions (TSFs) by

$$s_k(t, \theta) = \frac{\partial \eta}{\partial \theta_k}(t, \theta) \in \mathbb{R}^M, k = 1, \ldots, p.$$

To compute the TSFs, $\eta(t)$ must be smooth with respect to θ which is assumed given the assumption that h is sufficiently smooth. We obtain the TSFs from the following equation:

$$\frac{d}{dt}\frac{\partial x(t)}{\partial \theta} = \frac{\partial G}{\partial x}\frac{\partial x}{\partial \theta}(t) + \frac{\partial G}{\partial \tilde{x}}\frac{\partial x}{\partial \theta}(t-\tau) + \frac{\partial G}{\partial \theta}(t), \tag{2.4}$$

where the $\frac{\partial}{\partial \theta}$ and $\frac{d}{dt}$ operators have been interchanged, due to the continuity assumptions made on G and x, such that we have a delay differential equation (DDE) for the sensitivity function $\frac{\partial x}{\partial \theta}$. This DDE can be solved using MATLAB function dde23. The routine dde23 is an extended ode23 solver using the method of steps to approximate the solution [43]. The TSFs are used in the definition for the GSF that will be discussed later.

2.2 Theoretical Foundations

We begin by establishing well-posedness for our model. Assuming that $G(x(t), x(t-\tau), \theta)$ in (2.1) is continuous for $t \geq 0$, a solution $x(t, x_0, \Phi, \tau, \theta)$ for (2.1) satisfies the following integral equation

$$x(t, x_0, \Phi, \tau, \theta) = \begin{cases} \Phi(t), \quad -\tau \leq t < 0; \\ x_0 \quad t = 0; \\ \int_0^t G(x(s, x_0, \Phi, \tau, \theta), x(s-\tau, x_0, \Phi, \tau, \theta), \theta) ds \ t > 0. \end{cases}$$

To determine if a solution for the sensitivity equations exist, we must first determine if the solution for the delay differential equation defined in (2.1), exists, and is unique, and depends continuously on data using theory described in [16, 33, 34]. To show existence and uniqueness, we use Lemma 1, where the idea of successive approximations normally used with ordinary differential equations arguments are applied [16]. We define *successive approximations* in the following way for (2.1). Let $k = 0, 1, 2, \ldots$, then

$$\begin{aligned} x^0(t, x_0, \Phi, \tau, \theta) &= \begin{cases} \Phi(t) \quad -\tau \leq t < 0 \\ x_0 \quad t \geq 0. \end{cases} \\ x^{k+1}(t, x_0, \Phi, \tau, \theta) &= \begin{cases} \Phi(t), \quad -\tau \leq t < 0 \\ x_0, \quad t = 0 \\ \int_0^t G(x^k(s, x_0, \Phi, \tau, \theta), x^k(s-\tau, x_0, \Phi, \tau, \theta), \theta) ds, \quad t > 0. \end{cases} \end{aligned} \tag{2.6}$$

Lemma 1 (*Existence and Uniqueness of Solutions*) Let $G : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^p \to \mathbb{R}^n$ be continuous and satisfy

$$|G(x_1, \tilde{x_1}, \theta) - G(x_2, \tilde{x_2}, \theta)| \leq C_1|x_1 - x_2| + C_2|\tilde{x_1} - \tilde{x_2}| \tag{2.7}$$

where $C_j > 0$ is a constant for $j = 1, 2$, and assume that $\Phi \in L^2(-\tau, 0; \mathbb{R}^n)$. Then the successive approximations x^k converge uniformly for $t \in [0, T]$ to a unique solution of the initial value problem (2.1) -(2.2) with the solution bounded on $[0, T]$ and in fact, $x \in H^1(0, T; \mathbb{R}^n)$. Moreover, if $\Phi \in H^{1,\infty}(-\tau, 0; \mathbb{R}^n)$, one has the solution satisfies $x \in H^{1,\infty}(-\tau, T; \mathbb{R}^n)$.

Proof: Let $I = [0, T]$ where $t \in I$ and τ and θ are fixed. We let $M > 0$ such that $|G(x, \tilde{x}, \theta)| \leq M$ for all $t \in I$. Thus

$$|x^1(t, x_0, \Phi, \tau, \theta) - x_0| \leq Mt \text{ for } t \geq 0.$$

To show that the successive approximations converge let

$$\Delta^k(t, x_0, \Phi, \tau, \theta) = |x^{k+1}(t, x_0, \Phi, \tau, \theta) - x^k(t, x_0, \Phi, \tau, \theta)|.$$

Then,

$$\begin{aligned}
\Delta^k(t, x_0, \Phi, \tau, \theta) &= |\int_0^t G(x^k(s), x^k(s-\tau), \theta) - G(x^{k-1}(s), x^{k-1}(s-\tau), \theta)ds| \\
&\leq \int_0^t |G(x^k(s), x^k(s-\tau), \theta) - G(x^{k-1}(s), x^{k-1}(s-\tau), \theta)|ds \\
&\leq \int_0^t \{C_1|x^k(s) - x^{k-1}(s)| + C_2|x^k(s-\tau) - x^{k-1}(s-\tau)|\}ds, \quad (2.8)
\end{aligned}$$

for $t > 0$. When $t < 0$, $\Delta^k(t, x_0, \Phi, \tau, \theta) = 0$, since $x^k(\xi) = x^{k-1}(\xi) = \Phi(\xi)$, for $\xi \leq 0$.

Let $\xi = s - \tau$; then the second term of (2.8) becomes

$$\int_0^t C_2|x^k(s-\tau) - x^{k-1}(s-\tau)|ds = \int_{-\tau}^{t-\tau} C_2|x^k(\xi) - x^{k-1}(\xi)|d\xi.$$

Then given $t \in [0, T]$, and $x^k(\xi) = x^{k-1}(\xi)$ for $\xi \leq 0$,

$$\int_{-\tau}^{t-\tau} C_2|x^k(\xi) - x^{k-1}(\xi)|d\xi \leq \int_0^t C_2|x^k(\xi) - x^{k-1}(\xi)|d\xi,$$

and thus

$$\begin{aligned}\int_0^t \{C_1|x^k(s)-x^{k-1}(s)| + C_2|x^k(s-\tau)-x^{k-1}(s-\tau)|\}ds \\ = \int_0^t \{C_1|x^k(s)-x^{k-1}(s)| + C_2|x^k(s)-x^{k-1}(s)|\}ds \\ = (C_1+C_2)\int_0^t |x^k(s)-x^{k-1}(s)|ds.\end{aligned}$$

Let $\tilde{C} = (C_1 + C_2)$, then

$$\Delta^k(t, x_0, \Phi, \tau, \theta) \le \tilde{C}\int_0^t \Delta^{k-1}(s, x_0, \Phi, \tau, \theta)ds.$$

We claim that

$$\Delta^k(t, x_0, \Phi, \tau, \theta) \le \frac{M\tilde{C}^k t^{k+1}}{(k+1)!}, \tag{2.9}$$

then for $k = 0$ and $t \in I$ we have

$$\Delta^0(t, x_0, \Phi, \tau, \theta) = |x^1(t, x_0, \Phi, \tau, \theta) - x_0| \le Mt.$$

By induction, we have

$$\begin{aligned}\Delta^{k+1}(t, x_0, \Phi, \tau, \theta) &\le \tilde{C}\int_0^t \Delta^k(s, x_0, \Phi, \tau, \theta)ds, \\ &\le \tilde{C}\int_0^t \frac{M\tilde{C}^k s^{k+1}}{(k+1)!}ds \\ &\le \frac{M\tilde{C}^{k+1}t^{k+2}}{(k+2)!}.\end{aligned}$$

Therefore, we have that the inequality in (2.9) holds for all k and the series for $\sum_{k=0}^{\infty}$ $\Delta^k(t, x_0, \Phi, \tau, \theta)$ is dominated by the power series for $\frac{Me^{\tilde{C}t}}{\tilde{C}}$. Thus using the comparison test, the series $\sum_{k=0}^{\infty} \Delta^k(t, x_0, \Phi, \tau, \theta)$ converges uniformly on I. This implies that the series

$$x_0 + \sum_{k=0}^{\infty}(x^{k+1}(t, x_0, \Phi, \tau, \theta) - x^k(t, x_0, \Phi, \tau, \theta))$$

converges uniformly and absolutely on I, and the partial sum

$$x^n(t, x_0, \Phi, \tau, \theta) = x_0 + \sum_{k=0}^{n-1}(x^{k+1}(t, x_0, \Phi, \tau, \theta) - x^k(t, x_0, \Phi, \tau, \theta))$$

converges uniformly to a continuous function x on interval I. Due to the existence of x on I, $G(x(t), x(t-\tau), \theta)$ exists for $t \in I$. Therefore given that $x^k(t, x_0, \Phi, \tau, \theta)$ converges uniformly to $x(t, x_0, \Phi, \tau, \theta)$, and $x, \tilde{x} \to G(x, \tilde{x}, \theta)$ is continuous,

$$x^{k+1}(t, x_0, \Phi, \tau, \theta) = x_0 + \int_0^t G(x^k(s, x_0, \Phi, \tau, \theta), x^k(s-\tau, x_0, \Phi, \tau, \theta), \theta)ds$$

becomes in the limit as $k \to \infty$

$$x(t, x_0, \Phi, \tau, \theta) = x_0 + \int_0^t G(x(s, x_0, \Phi, \tau, \theta), x(s-\tau, x_0, \Phi, \tau, \theta), \theta)ds.$$

Therefore $x(t, x_0, \Phi, \tau, \theta)$ exists and satisfies (2.1). Moreover, it is easily seen that $x \in H^{1,\infty}(-\tau, 0; \mathbb{R}^n)$ if $\Phi \in H^{1,\infty}(-\tau, 0; \mathbb{R}^n)$.

To show uniqueness of the solution, we assume there exists two solutions to (2.1), $x_1(t) = x_1(t, x_0, \Phi, \tau, \theta)$, and $x_2(t) = x_2(t, x_0, \Phi, \tau, \theta)$. We have

$$\begin{aligned}
|x_1(t) - x_2(t)| &= |\int_0^t G(x_1(s), x_1(s-\tau), \theta) - G(x_2(s), x_2(s-\tau), \theta)ds| \\
&\leq \int_0^t |G(x_1(s), x_1(s-\tau), \theta) - G(x_2(s), x_2(s-\tau), \theta)|ds \\
&\leq \tilde{C}\int_0^t |x_1(s) - x_2(s)|ds.
\end{aligned}$$

Then from Gronwall's Inequality, it follows that

$$|x_1(t, x_0, \Phi, \tau, \theta) - x_2(t, x_0, \Phi, \tau, \theta)| \leq 0e^{\tilde{C}t} = 0. \tag{2.10}$$

Thus $x_1(t, x_0, \Phi, \tau, \theta) = x_2(t, x_0, \Phi, \tau, \theta)$.

Remark 1 The above arguments can be readily extended to systems where G depends explicitly on t, i.e., $G = G(t, x, \tilde{x}, \theta)$. Then one obtains similar results for nonautonomous affine systems with bounded coefficients for the state and delay terms and bounded perturbations of nonautonomous linear systems.

Lemma 2 *(Continuous Dependence of Solutions on Parameters)* Let $G : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^p \to \mathbb{R}^n$, and given $\theta = \theta_0$ let $x(t, x_0, \Phi, \tau, \theta)$ be a solution of

$$\begin{aligned} \dot{x}(t) &= G(x(t), x(t-\tau), \theta_0), \\ x(\xi) &= \begin{cases} \Phi(\xi), & -\tau \leq \xi < 0 \\ x_0, & \xi = 0 \end{cases} \end{aligned} \tag{2.11}$$

for $t \in [0, T]$. Assume that

$$\lim_{\theta \to \theta_0} G(x, \tilde{x}, \theta) = G(x, \tilde{x}, \theta_0), \tag{2.12}$$

uniformly in $x, \tilde{x}$ and for $(x_1, \tilde{x}_1, \theta), (x_2, \tilde{x}_2, \theta) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^p$, and assume that (2.7) is satisfied. Then the Initial Value Problem (IVP) (2.11)-(2.12) has a unique solution $x(t, x_0, \Phi, \tau, \theta)$ that satisfies

$$\lim_{\theta \to \theta_0} x(t, x_0, \Phi, \tau, \theta) = x(t, x_0, \Phi, \tau, \theta_0), \quad t \in [0, T]. \tag{2.13}$$

Proof: Existence and uniqueness of a solution on any interval $[0, T]$ is provided by Lemma 1. To show continuous dependence of solutions on θ, let $t \in [0, T]$, and $x_0 \in \mathbb{R}^n$,$\Phi \in L^2(-\tau, 0; \mathbb{R}^n)$, $\tau \in \mathbb{R}$ be fixed, we define $z(t, \theta, \theta_0) = x(t, \theta) - x(t, \theta_0)$.

Then,

$$\begin{aligned} |z(t,\theta,\theta_0)| &= |x(t,\theta) - x(t,\theta_0)| \\ &\leq \int_0^t |G(x(s,\theta), x(s-\tau,\theta), \theta) - G(x(s,\theta_0), x(s-\tau,\theta_0), \theta_0)| ds \\ &= \int_0^t \{|G(x(s,\theta), x(s-\tau,\theta), \theta) + G(x(s,\theta_0), x(s-\tau,\theta_0), \theta) \\ &\quad -G(x(s,\theta_0), x(s-\tau,\theta_0), \theta) - G(x(s,\theta_0), x(s-\tau,\theta_0), \theta_0)|\} ds \\ &\leq \int_0^t \{|G(x(s,\theta), x(s-\tau,\theta), \theta) - G(x(s,\theta_0), x(s-\tau,\theta_0), \theta)| \\ &\quad +|G(x(s,\theta_0), x(s-\tau,\theta_0), \theta) - G(x(s,\theta_0), x(s-\tau,\theta_0), \theta_0)|\} ds \\ &\leq \int_0^t \tilde{C}|x(s,\theta) - x(s,\theta_0)| ds + \int_0^t \{|G(x(s,\theta_0), x(s-\tau,\theta_0), \theta) \\ &\quad -G(x(s,\theta_0), x(s-\tau,\theta_0), \theta_0)|\} ds. \end{aligned}$$

Let $h(s, \theta) = |G(x(t, \theta_0), x(s-\tau, \theta_0), \theta) - G(x(s, \theta_0), x(s-\tau, \theta_0), \theta_0)|$. Then $h(s, \theta) \to 0$ uniformly in s as $\theta \to \theta_0$ from the earlier described assumption in (2.13). Thus

$$|z(t, \theta, \theta_0)| \leq \int_0^T h(s, \theta) ds + \int_0^t \tilde{C}|x(t, \theta) - x(t, \theta_0)| ds, \tag{2.14}$$

and with application of Gronwall's inequality and the limit as $\theta \to \theta_0$ it follows that

$$\begin{aligned}\lim_{\theta \to \theta_0} |z(t, \theta, \theta_0)| &\leq \lim_{\theta \to \theta_0} (\int_0^T h(s, \theta) ds) e^{\tilde{C}t} \\ &= 0.\end{aligned}$$

Therefore,

$$\lim_{\theta \to \theta_0} x(t, x_0, \Phi, \tau, \theta) = x(t, x_0, \Phi, \tau, \theta_0).$$

Now that we have established existence, uniqueness, and continuous dependence on θ for the solution to the model described in (2.1), to perform sensitivity analysis and obtain GSFs (defined in a later section) we must determine if (2.4) has a solution. We want to prove that $\frac{\partial x(t)}{\partial \theta}$ exists and is given by $y(t)$ that satisfies the system (2.15), (2.16). Let $y(t)$ be the solution of

$$\dot{y}(t) = G_x(x(t), \tilde{x}(t), \theta) y(t) + G_{\tilde{x}}(x(t), \tilde{x}(t), \theta) y(t - \tau) + G_\theta(x(t), \tilde{x}(t), \theta), \quad t > 0 \tag{2.15}$$

$$y(t) = 0 \quad -\tau \leq t \leq 0. \tag{2.16}$$

To show that the above differential equation has a solution given by the appropriate derivative, we must prove that certain Frechet derivatives exists. Thus we must show

$$\lim_{|h| \to 0} \frac{1}{|h|} |G(x, \tilde{x}, \theta + h) - G(x, \tilde{x}, \theta) - A(h)| \to 0,$$

where A is the operator of the Frechet derivative of G at θ.

Lemma 3 *Mean Value Theorem* Let $G : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^p \to \mathbb{R}^n$ and $z = (x_0, \Phi)$ where Z is $\mathbb{R}^n \times L^2(-\tau, 0; \mathbb{R}^n)$.

(i) If $G_x(x, \tilde{x}, \theta)$ and $G_{\tilde{x}}(x, \tilde{x}, \theta)$ exists and are continuous for $x \in \mathbb{R}^n$, then for $x_1, x_2, \tilde{x}_1, \tilde{x}_2 \in \mathbb{R}^n$ and $\theta \in \mathbb{R}^p$, $t \geq 0$,

$$\begin{aligned}G(x_1, \tilde{x}_1, \theta) - G(x_2, \tilde{x}_2, \theta) &= \int_0^1 \{G_x(sx_1 + (1-s)x_2, s\tilde{x}_1 + (1-s)\tilde{x}_2, \theta)(x_1 - x_2) \\ &+ G_{\tilde{x}}(sx_1 + (1-s)x_2, s\tilde{x}_1 + (1-s)\tilde{x}_2, \theta)(\tilde{x}_1 - \tilde{x}_2)\} ds.\end{aligned}$$

(ii) If $G_\theta(x,\tilde{x},\theta)$ exists and is continuous for $x \in \mathbb{R}^n$, then for $x,\tilde{x} \in \mathbb{R}^n$ and $\theta_1,\theta_2 \in \mathbb{R}^p$, $t \geq 0$,

$$G(x,\tilde{x},\theta_1) - G(x,\tilde{x},\theta_2) = \int_0^1 G_\theta(x,\tilde{x},s\theta_1 + (1-s)\theta_2)(\theta_1 - \theta_2)ds.$$

(iii) Suppose $x(t,z,\tau,\theta)$ is a solution of (2.1), which is continuous in $\mathbb{R}^n$ and continuous and continuously differentiable for $z \in Z$, such that $D_z x(t;\cdot) \in \mathcal{L}(Z,\mathbb{R}^n)$, thus $D_z = \frac{\partial}{\partial z}$ is a bounded differential operator. Then for $z_1, z_2 \in Z$, and a fixed $\tau \in \mathbb{R}$, $\theta \in \mathbb{R}^p$, for $t \in [0,T]$,

$$x(t,z_1,\tau,\theta) - x(t,z_2,\tau,\theta) = \int_0^1 D_z x(t; sz_1 + (1-s)z_2)[z_1 - z_2]ds.$$

Proof: We consider (i). Let

$$H_1(s) = G(sx_1 + (1-s)x_2, s\tilde{x}_1 + (1-s)\tilde{x_2},\theta),\ \ 0 < s \leq 1$$

using the chain rule of Frechet derivative, we have

$$\begin{aligned} H_1^{'}(s) &= G_x(sx_1 + (1-s)x_2, s\tilde{x}_1 + (1-s)\tilde{x}_2,\theta)(x_1 - x_2) \\ &+ G_{\tilde{x}}(sx_1 + (1-s)x_2, s\tilde{x}_1 + (1-s)\tilde{x}_2,\theta)(\tilde{x}_1 - \tilde{x}_2). \end{aligned}$$

Then if we integrate $H'{}_1(s)$ for $s \in (0,1]$, we have $H_1(1) - H_1(0)$ which is $G(x_1,\tilde{x}_1,\theta) - G(x_2,\tilde{x}_2,\theta)$, thus we have (i).

The proof of (ii) and (iii) are similar to the previous proof, thus we omit it.

Theorem 1 Suppose that $G(x,\tilde{x},\theta)$ has continuous Frechet derivatives $G_\theta, G_x, G_{\tilde{x}}$ such that $|G_x| \leq M_0$,$|G_{\tilde{x}}| \leq M_1$, and $|G_\theta| \leq M_2$. Then the Frechet derivative $y_1(t) = \frac{\partial x(t)}{\partial \theta} \in \mathbb{R}^{n\times p}$ exists and is the unique solution for

$$\begin{aligned} \dot{y}_1(t) &= G_x(x(t),x(t-\tau),\theta)y_1(t) + G_{\tilde{x}}(x(t),x(t-\tau),\theta)y_1(t-\tau) + \\ & \qquad G_\theta(x(t),x(t-\tau),\theta), \quad t > 0 \qquad (2.17) \\ y_1(0) &= 0 \qquad -\tau \leq t < 0. \end{aligned}$$

Proof: Using a ready extension of Lemma 1 (see Remark 1), we can easily establish that (2.17) has a unique solution $y_1(t)$. For a fixed $\tau \in \mathbb{R}$, $t \in [0,T]$, $x,\tilde{x} \in \mathbb{R}^n$, let $h \in \mathbb{R}^p$, and

$m_1(t,\theta,h) = x(t,\theta+h) - x(t,\theta)$.

$$\begin{aligned}
m_1(t,\theta,h) &= \int_0^t \{G(x(s,\theta+h), x(s-\tau,\theta+h), \theta+h) - G(x(s,\theta), x(s-\tau,\theta),\theta)\}ds \\
&= \int_0^t \{G(x(s,\theta+h), x(s-\tau,\theta+h), \theta+h) - G(x(s,\theta+h), x(s-\tau,\theta+h),\theta) \\
&+ \ G(x(s,\theta+h), x(s-\tau,\theta+h),\theta) - G(x(s,\theta), x(s-\tau,\theta),\theta)\}ds. \qquad (2.18)
\end{aligned}$$

Using Lemma 3, with *Frechet* differentiability then we can write (2.18) as

$$\begin{aligned}
m_1(t,\theta,h) &= \int_0^t \Big\{G_\theta(x(s,\theta+h), x(s-\tau,\theta+h),\theta)[\theta+h-\theta] + w_1(s,h) \\
&+ \ G_x(x(s,\theta), x(s-\tau,\theta),\theta)[x(s,\theta+h) - x(s,\theta)] + w_2(s, m_1(s,\theta,h)) \\
&+ \ G_{\tilde{x}}(x(s,\theta), x(s-\tau,\theta),\theta)[x(s-\tau,\theta+h) - x(s-\tau,\theta)] \\
&\qquad + \ w_2(s, m_1(s-\tau,\theta,h))\Big\}ds,
\end{aligned}$$

where $w_1(t,h) \to 0$ as $h \to 0$, and $w_2(s, m_1(s,\theta,h)) \to 0$ as $m_1(s,\theta,h) \to 0$. We define $b_1(t,h) = \frac{w_1(t,h)}{|h|}$ and $b_2(t,h) = \frac{w_2(t,m_1(t,\theta,h))}{|m_1(t,\theta,h)|}$ so that $b_1(t,h) \to 0$, and $b_2(t,h) \to 0$ as $|h| \to 0$ and $|m_1(t,\theta,h)| \to 0$, respectively. Then for the Frechet derivative $\frac{\partial x(t)}{\partial \theta}$ to exist, as $|h| \to 0$ we must argue

$$\frac{|m_1(t,\theta,h) - y_1(t)h|}{|h|} \to 0,$$

where $y_1(t)$ is defined by (2.17). We argue that

$$\begin{aligned}
\frac{|m_1(t,\theta,h) - y_1(t)h|}{|h|} = \frac{1}{|h|}\int_0^t \{|G_\theta(x(s,\theta+h), x(s-\tau,\theta+h),\theta)[h] + w_1(s,h) \\
+G_x(x(s,\theta), x(s-\tau,\theta),\theta)[m_1(s,\theta,h)] + w_2(s, m_1(s,\theta,h)) \\
+G_{\tilde{x}}(x(s,\theta), x(s-\tau,\theta),\theta)[m_1(s-\tau,\theta,h)] + w_2(s-\tau, m_1(s-\tau,\theta,h)) \\
-h[G_x(x(s,\theta), x(s-\tau,\theta),\theta)y_1(s) + G_{\tilde{x}}(x(s,\theta), x(s-\tau,\theta),\theta)y_1(s-\tau) \\
+G_\theta(x(s,\theta), x(s-\tau,\theta),\theta)]\}ds.
\end{aligned}$$

Hence,

$$
\begin{aligned}
&\frac{|m_1(t,\theta,h)-y_1(t)h|}{|h|}\\
&\le \int_0^T \frac{|G_\theta(x(s,\theta+h),x(s-\tau,\theta+h),\theta)-G_\theta(x(s,\theta),x(s-\tau,\theta),\theta)||h|}{|h|}ds\\
&\quad+\int_0^T \frac{|G_x(x(s,\theta),x(s-\tau,\theta),\theta)||m_1(s,\theta,h)-y_1(s)h|}{|h|}ds\\
&\quad+\int_0^T \frac{|G_{\tilde{x}}(x(s,\theta),x(s-\tau,\theta),\theta)||m_1(s-\tau,\theta,h)-y(s-\tau)h|}{|h|}ds\\
&\quad+\int_0^T \frac{w_1(s,h)}{|h|}+\int_0^T \frac{w_2(s,m_1(s,\theta,h))}{|h|}ds\\
&\quad+\int_0^T \frac{w_2(s-\tau,m_1(s-\tau,\theta,h))}{|h|}ds.
\end{aligned}
$$

Then

$$
\begin{aligned}
&\frac{|m_1(t,\theta,h)-y_1(t)h|}{|h|}\\
&\le \int_0^T \frac{|G_\theta(x(s,\theta+h),x(s-\tau,\theta+h),\theta)-G_\theta(x(s,\theta),x(s-\tau,\theta),\theta)||h|}{|h|}ds\\
&\quad+\int_0^T \frac{M_0|m_1(s,\theta,h)-y_1(s)h|}{|h|}ds+\int_0^T \frac{M_1|m_1(s-\tau,\theta,h)-y_1(s-\tau)h|}{|h|}ds\\
&\quad+\int_0^T b_1(s,h)ds+\int_0^T \frac{w_2(s,m_1(s,\theta,h))}{|h|}ds+\int_0^T \frac{w_2(s-\tau,m_1(s-\tau,\theta,h))}{|h|}ds.
\end{aligned}
$$

We need to show that $\frac{w_2(t,m_1(t,\theta,h))}{|h|} \le K\frac{w_2(t,m_1(t,\theta,h))}{|m_1(t,\theta,h)|}$, so we consider

$$
\begin{aligned}
|m_1(t,\theta,h)| &= \Big|\int_0^T \big\{G_\theta(x(s,\theta+h),x(s-\tau,\theta+h),\theta)[h]\\
&\quad +G_x(x(s,\theta),x(s-\tau,\theta),\theta)m_1(s,\theta,h)\\
&\quad + G_{\tilde{x}}(x(s,\theta),x(s-\tau,\theta),\theta)m_1(s-\tau,\theta,h)+w_1(s,h)\\
&\quad + w_2(s,m_1(s,\theta,h))+w_2(s-\tau,m_1(s-\tau,\theta,h))ds\big\}\Big|\\
&\le \int_0^T M_2|h|ds+\int_0^T M_0|m_1(s,\theta,h)|ds+\int_0^T M_1|m_1(s-\tau,\theta,h)|ds\\
&+ \int_0^T |h|b_1(s,h)ds+\int_0^T b_2(s,h)|m_1(s,\theta,h)|ds\\
&+ \int_0^T b_2(s-\tau,h)|m_1(s-\tau,\theta,h)|ds.
\end{aligned}
$$

Let $\xi = s - \tau$, then

$$\int_0^T (b_2(s-\tau) + M_1)|m_1(s-\tau,\theta,h)|ds = \int_{-\tau}^{T-\tau} (b_2(\xi,h) + M_1)|m_1(\xi,\theta,h)|d\xi$$
$$\leq \int_{-\tau}^{0} (b_2(\xi,h) + M_1)|m_1(\xi,\theta,h)|d\xi + \int_{0}^{T-\tau} (b_2(\xi,h) + M_1)|m_1(\xi,\theta,h)|d\xi$$
$$\leq 0 + \int_{0}^{T-\tau} (b_2(\xi,h) + M_1)|m_1(\xi,\theta,h)|d\xi \leq \int_{0}^{T} (b_2(\xi,h) + M_1)|m_1(\xi,\theta,h)|d\xi$$

since for $t \in [-\tau, 0]$, $x(t,\theta+h) = x(t,\theta)$, so $m_1(t,\theta,h) = 0$. Using the change of variables, we obtain the following:

$$\begin{aligned}
|m_1(t,\theta,h)| &\leq \int_0^T M_2|h|ds + \int_0^T M_0|m_1(s,\theta,h)|ds + \int_0^T M_1|m_1(s,\theta,h)|ds \\
&\quad + \int_0^T b_1(s,h)|h|ds + \int_0^T b_2(s,h)|m_1(s,\theta,h)|ds + \int_0^T b_2(s,h)|m_1(s,\theta,h)|ds \\
&\leq \int_0^T (M_2 + b_1(s,h))|h|ds + \int_0^T (M_0 + M_1 + 2b_2(s,h))|m_1(s,\theta,h)|ds \\
&\leq |h|K,
\end{aligned}$$

where $K = (M_2 + b_1(s,h))e^{\int_0^t (M_0+M_1+2b_2(s,h))ds}$ when applying Gronwall's inequality. Thus

$$\frac{w_2(t, m_1(t,\theta,h))}{|h|} \leq K \frac{w_2(t, m_1(t,\theta,h))}{|m_1(t,\theta,h)|}.$$

Then

$$\begin{aligned}
&\frac{|m_1(t,\theta,h) - y_1(t)h|}{|h|} \\
&\quad \leq \int_0^T \frac{|G_\theta(x(s,\theta+h), x(s-\tau,\theta+h),\theta) - G_\theta(x(s,\theta), x(s-\tau,\theta),\theta)||h|}{|h|} ds \\
&\quad + \int_0^T \frac{M_0|m_1(s,\theta,h) - y_1(s)h|}{|h|} ds + \int_0^T \frac{M_1|m_1(s-\tau,\theta,h) - y_1(s-\tau)h|}{|h|} ds \\
&\quad + \int_0^T b_1(s,h)ds + \int_0^T K\frac{w_2(s, m_1(s,\theta,h))}{|m_1(s,\theta,h)|} ds + \int_0^T K\frac{w_2(s-\tau, m_1(s-\tau,\theta,h))}{|m_1(s-\tau,\theta,h)|} ds.
\end{aligned}$$

Then using the change of variables as shown earlier, we have that

$$\begin{aligned}
&\frac{|m_1(t,\theta,h)-y_1(t)h|}{|h|} \\
&\qquad \le \int_0^T \frac{|G_\theta(x(s,\theta+h),x(s-\tau,\theta+h),\theta)-G_\theta(x(s,\theta),x(s-\tau,\theta),\theta)||h|}{|h|}ds \\
&\qquad\quad + \int_0^T \frac{M_0|m_1(s,\theta,h)-y_1(s)h|}{|h|}ds + \int_0^T \frac{M_1|m_1(s,\theta,h)-y_1(s)h|}{|h|}ds \\
&\qquad\quad + \int_0^T b_1(s,h)ds + \int_0^T Kb_2(s,h)ds + \int_0^T Kb_2(s,h)ds \\
&\qquad \le \int_0^T \frac{|G_\theta(x(s,\theta+h),x(s-\tau,\theta+h),\theta)-G_\theta(x(s,\theta),x(s-\tau,\theta),\theta)||h|}{|h|}ds \\
&\qquad\quad + \int_0^T (M_0+M_1)\frac{|m_1(s,\theta,h)-y_1(s)h|}{|h|}ds + \int_0^T \{b_1(s,h)+2Kb_2(s,h)\}ds
\end{aligned} \tag{2.19}$$

Since $x(t,\theta)$ is continuously dependent on θ for $t \in [0,T]$, and G_θ is continuous, using Lemma 2, we have

$$\lim_{|h|\to 0} |G_\theta(x(s,\theta+h),x(s-\tau,\theta+h)) - G_\theta(x(s,\theta),x(s-\tau,\theta),\theta)| = 0.$$

Thus

$$|G_\theta(x(s,\theta+h),x(s-\tau,\theta+h,\theta),\theta) - G_\theta(x(s,\theta),x(s-\tau,\theta),\theta)| \le b_3(t,h),$$

where $b_3(t,h) \to 0$ as $|h| \to 0$. Then

$$\begin{aligned}
\frac{|m_1(t,\theta,h)-y_1(t)h|}{|h|} &\le \int_0^T \{b_1(s,h)+2Kb_2(s,h)+b_3(s,h)\}ds \\
&\quad + \int_0^T (M_0+M_1)\frac{|m_1(s,\theta,h)-y_1(s)h|}{|h|}ds.
\end{aligned}$$

Since as $|h| \to 0$, $b_1(t,h), b_2(t,h)$, and $b_3(t,h) \to 0$, and all functions are bounded, due to dominated convergence we find $K(h) \to 0$ as $h \to 0$ where

$$K(h) = \int_0^T \{b_1(s,h)+2Kb_2(s,h)+b_3(s,h)\}ds.$$

Thus with an application of Gronwall's inequality we have

$$\begin{aligned}
\lim_{|h|\to 0} \frac{|m_1(t,\theta,h)-y_1(t)h|}{|h|} &\le \lim_{|h|\to 0} \{K(h)\} e^{\int_0^t (M_0+M_1)ds} \\
&= 0.
\end{aligned}$$

Theorem 2 Suppose the function $G(x, \tilde{x}, \theta)$ of (2.1) has continuous Frechet derivatives $G_x(x, \tilde{x}, \theta)$, $G_{\tilde{x}}(x, \tilde{x}, \theta)$, with respect to x and $\tilde{x}$, with

$$|G_x(x, \tilde{x}, \theta)| \leq M_0, \quad |G_{\tilde{x}}(x, \tilde{x}, \theta)| \leq M_1$$

for some constants $M_j > 0$ for $j = 0, 1$. Then the Frechet derivative $y_2(t) = \frac{\partial}{\partial z} x(t, z, \theta)$ exists with $y_2(t) \in \mathcal{L}(Z, \mathbb{R}^n)$ (recall $z = (x_0, \Phi), Z = \mathbb{R}^n \times L^2(-\tau, 0; \mathbb{R}^n)$), and satisfies the equation

$$\begin{aligned} \dot{y}_2(t)[h] &= G_x(x(t), x(t-\tau), \theta) y_2(t)[h] + G_{\tilde{x}}(x(t), x(t-\tau), \theta) y_2(t-\tau)[h], t > 0 \quad (2.20) \\ y_2(\xi) &= \mathcal{I} \quad -\tau \leq \xi \leq 0, \end{aligned}$$

where $\mathcal{I} \in \mathcal{L}(Z, \mathbb{R}^n)$ is the identity.

Proof: Using Lemma 1, we know that the differential equation (2.20) has a unique solution, $y_2(t)[h]$. For a fixed $\tau \in \mathbb{R}$, $\theta \in \mathbb{R}^p$, and $t \in [0, T]$, let $h \in Z$, and $m_2(t, z, h) = x(t, z+h) - x(t, z)$.

$$m_2(t, z, h) = \int_0^t \{G(x(s, z+h), x(s-\tau, z+h), \theta) - G(x(s, z), x(s-\tau, z), \theta)\} ds.$$

With the *Frechet* differentiability of G with respect to $x \in \mathbb{R}^n$ and for $z \in \mathcal{L}(Z, \mathbb{R}^n)$, we have

$$\begin{aligned} m_2(t, z, h) = \int_0^t \Big\{ & G_x(x(s, z), x(s-\tau, z), \theta)[x(s, z+h) - x(s, z)] + w_3(s, m_2(s, z, h)) \\ & + G_{\tilde{x}}(x(s, z), x(s-\tau, z), \theta)[x(s-\tau, z+h) - x(s-\tau, z)] + w_3(s, m_2(s-\tau, z, h)) \Big\} ds \end{aligned}$$

where

$$\frac{|w_3(t, m_2(t, z, h))|}{|m_2(t, z, h)|} \to 0$$

as $|m_2(t, z, h)|$ approaches zero. We define $b_4(t, h)$ as follows:

$$b_4(t, h) = \frac{|w_3(t, m_2(t, z, h))|}{|m_2(t, z, h)|},$$

then $b_4(t, h) \to 0$ uniformly in t as $|h| \to 0$. To show that $y_2(t)$ is a solution for (2.20), we must show that

$$\frac{|m_2(t, z, h) - y_2(t)[h]|}{|h|} \to 0$$

as $|h| \to 0$, where $y_2(t)$ is $\in \mathcal{L}(Z, \mathbb{R}^n)$.

By definition,

$$
\begin{aligned}
&\frac{|m_2(t,z,h)-y_2(t)[h]|}{|h|}\\
&\qquad = |\int_0^t \Big\{\frac{G_x(x(s,z),x(s-\tau,z),\theta)[x(s,z+h)-x(s,z)]+w_3(s,m_2(s,z,h))}{|h|}\\
&\qquad +\frac{G_{\tilde{x}}(x(s,z),x(s-\tau,z),\theta)[x(s-\tau,z+h)-x(s-\tau,z)]+w_3(m_2(s-\tau,z,h))}{|h|}\\
&\qquad -\frac{\{G_x(x(s,z),x(s-\tau,z),\theta)y_2(s)[h]+G_{\tilde{x}}(x(s-\tau,z),x(s-\tau,z),\theta)y_2(s-\tau)[h]\}}{|h|}\Big\}ds|,
\end{aligned}
$$

then

$$
\begin{aligned}
&\frac{|m_2(t,z,h)-y_2(t)[h]|}{|h|}\\
&\qquad \leq \int_0^t \Big\{|\frac{G_x(x(s,z),x(s-\tau,z),\theta)[x(s,z+h)-x(s,z)]+w_3(s,m_2(s,z,h))}{|h|}|\\
&\qquad +|\frac{G_{\tilde{x}}(x(s,z),x(s-\tau,z),\theta)[x(s-\tau,z+h)-x(s-\tau,z)]+w_3(s-\tau,m_2(s-\tau,z,h))}{|h|}|\\
&\qquad -|\frac{[G_x(x(s,z),x(s-\tau,z),\theta)y_2(s)[h]+G_{\tilde{x}}(x(s-\tau,z),x(s-\tau,z),\theta)y_2(s-\tau)[h]]}{|h|}|\Big\}ds.
\end{aligned}
$$

We want to show that

$$
\frac{|w_3(t,m_2(t,z,h))|}{|h|} \leq K\frac{|w_3(t,m_2(t,z,h))|}{|m_2(t,z,h)|}
$$

for some constant $K>0$. Thus we consider

$$
\begin{aligned}
|m_2(t,z,h)| &= \frac{1}{|h|}\int_0^T |G_x(x(s,z),x(s-\tau,z),\theta)[x(s,z+h-x(s,z)]+w_3(s,m_2(s,z,h))\\
&\quad +G_{\tilde{x}}(x(s,z),x(s-\tau,z),\theta)[x(s-\tau,z+h)-x(s-\tau,z)]+w_3(s,m_2(s-\tau,z,h))|ds\\
&= \frac{1}{|h|}\int_0^T \Big\{|G_x(x(s,z),x(s-\tau,z),\theta)[x(s,z+h)-x(s,z)]+w_3(s-\tau,m_2(s,z,h))\\
&\qquad +G_{\tilde{x}}(x(s,z),x(s-\tau,z),\theta)\{\int_0^1 D_z x(s-\tau,r(z+h)+(1-r)z)[h]dr\}\\
&\qquad +w_3(s-\tau,m_2(s-\tau,z,h))|\Big\}ds.
\end{aligned}
$$

Then given the earlier definition of $b_4(t,h)$ we know that $|w_3(t,m_2(t,z,h)|=b_4(t,h)|m_2(t,z,h)|$.

Thus with the assumptions that $|G_x| \leq M_0$, and $|G_{\tilde{x}}| \leq M_1$, the function $|m_2(t,z,h)|$ is bounded by

$$\int_0^T M_0 |m_2(s,z,h)| ds + \int_0^T b_4(s,h)|m_2(s,z,h)| ds$$
$$+ \int_0^T M_1 \int_0^1 \{D_z x(s, r(z+h) + (1-r)z)[h] dr\} ds + \int_0^T b_4(s-\tau,h)|m_2(s-\tau,z,h)| ds.$$

We let $\xi = s - \tau$ and then we can write the previous as follows:

$$\begin{aligned} &\int_0^T M_0 |m_2(s,z,h)| ds + \int_0^T b_4(s,h)|m_2(s,z,h)| ds \\ &+ \int_0^T M_1 \int_0^1 \{D_z x(s, r(z+h) + (1-r)z)[h] dr\} ds \\ + \int_{-\tau}^0 b_4(\xi,h)|m_2(\xi,z,h)| d\xi &+ \int_0^{T-\tau} b_4(\xi,h)|m_2(\xi,z,h)| d\xi. \end{aligned}$$

When $\xi \in [-\tau, 0]$, $m_2(\xi,z,h) = I - I = 0$, thus the function $|m_2(t,z,h)|$ is bounded by:

$$\int_0^T M_0 |m_2(s,z,h)| ds + \int_0^T b_4(s,h)|m_2(s,z,h)| ds$$
$$+ \int_0^T M_1 \int_0^1 \{D_z x(s, r(z+h) + (1-r)z)[h] dr\} ds + \int_0^{T-\tau} b_4(\xi,h)|m_2(\xi,z,h)| d\xi.$$

Then since D_z is a bounded operator, $\int_0^1 D_z\ x(s, r(z+h) + (1-r)z)[h] dr < h$ and

$$\begin{aligned} |m_2(t,z,h)| \quad \leq \quad & \int_0^T M_0 |m_2(s,z,h)| ds + \int_0^T 2 b_4(s,h)|m_2(s,z,h)| ds \\ & + \int_0^T M_1\ h\ ds. \end{aligned}$$

We apply Gronwall's inequality to obtain

$$|m_2(t,z,h)| \leq K h$$

where $K = M_1 T e^{\int_0^T (M_0 + 2 b_4(s,h)) ds}$ and $b_4(t,h) \to 0$ uniformly in t as $|h| \to 0$. Thus it follows that

$$\frac{|w_3(t, m_2(t,z,h))|}{|h|} \leq K \frac{|w_3(t, m_2(t,z,h))|}{|m_2(t,z,h)|},$$

therefore

$$\frac{|w_3(t,m_2(t-\tau,z,h))|}{|h|} \le K\frac{|w_3(t,m_2(t-\tau,z,h))|}{|m_2(t,z,h)|}.$$

Thus we find

$$\begin{aligned}\frac{|m_2(t,z,h)-y_2(t)[h]|}{|h|} \le & \int_0^T M_0\frac{|m_2(s,z,h)-y_2(s)[h]|}{|h|}ds + \int_0^T K\frac{|w_3(s,m_2(s,z,h))|}{|m_2(s,z,h)|}ds \\ &+ \int_0^T M_1\frac{|m_2(s-\tau,z,h)-y_2(s-\tau)[h]|}{|h|}ds + \int_0^T K\frac{|w_3(s-\tau,m_2(s-\tau,z,h))|}{|m_2(s-\tau,z,h)|}ds \\ &\le \int_0^T M_0\frac{|m_2(s,z,h)-y_2(s)[h]|}{|h|}ds \\ &+ \int_0^T \{K(b_4(s,h)+b_4(s-\tau,h)+M_1\frac{|m_2(s-\tau,z,h)-y_2(s-\tau)[h]|}{|h|}\}ds.\end{aligned}$$

Letting $\xi = s-\tau$ in the last two integral terms, we have

$$\begin{aligned}\frac{|m(t,z,h)-y(t)[h]|}{|h|} \le & \int_0^T M_0\frac{|m_2(s,z,h)-y_2(s)[h]|}{|h|}ds + \int_0^T K(b_4(s,h)+b_4(\xi,h)ds \\ &+ \int_{-\tau}^0 M_1\frac{|m_2(\xi,z,h)-y_2(\xi)[h]|}{|h|}d\xi + \int_0^{T-\tau} M_1\frac{|m_2(\xi,z,h)-y_2(\xi)[h]|}{|h|}d\xi.\end{aligned}$$

When $\xi \in [-\tau, 0]$

$$\frac{|m_2(\xi,z,h)-y_2(\xi)[h]|}{|h|} = \frac{I-I-0\cdot h}{|h|},$$

then

$$\frac{|m_2(t,z,h)-y_2(t)[h]|}{|h|} \le \int_0^T (M_0+M_1)\frac{|m_2(s,z,h)-y_2(s)[h]|}{|h|}ds + \int_0^T 2K(b_4(s,h))ds.$$

We apply Gronwall's inequality and dominated convergence, and taking the limit as $|h| \to 0$, we conclude

$$\begin{aligned}\lim_{|h|\to 0}\frac{|m_2(t,z,h)-y_2(t)[h]|}{|h|} &\le \lim_{|h|\to 0}\{(\int_0^T 2K(b_4(s,h)ds)e^{\int_0^t M_0+M_1 ds}\} \\ &= 0\end{aligned}$$

since $b_4(s,h) \to 0$ as $h \to 0$.

Theorem 3 Suppose that $G(x, \tilde{x}, \theta)$ has continuous Frechet derivatives $G_x, G_{\tilde{x}}$ such that $|G_x| \le M_0$, and $|G_{\tilde{x}}| \le M_1$ and suppose that the solution x of (2.1)-(2.2) satisfies $x \in H^{1,\infty}(-\tau, T; \mathbb{R}^n)$, for $0 < \tau < r$ for fixed $r > 0$. Then the Frechet derivative $y_3(t) = \frac{\partial x(t)}{\partial \tau} \in \mathbb{R}^n$ exists and is the unique solution for

$$
\begin{aligned}
\dot{y}_3(t) &= G_x(x(t), x(t-\tau), \theta) y_3(t) + G_{\tilde{x}}(x(t), x(t-\tau), \theta)[y_3(t-\tau) - \dot{x}(t-\tau)] \\
y_3(\nu) &= 0, \quad -\tau \le \nu \le 0.
\end{aligned} \tag{2.21}
$$

Moreover, $\frac{\partial x(t)}{\partial \tau}$ is continuous in θ and, if $x \in C^1(-\tau, T; \mathbb{R}^n)$ it is also continuous in τ.

Proof: We first reformulate (2.21) as a Cauchy problem on the state space $Z_1 = \mathbb{R}^n \times L^2(-r, 0; \mathbb{R}^n)$ with the norm $|(\xi, \phi)|^2 = |\xi|^2 + \int_{-r}^0 |\phi(s)|^2 ds$. We may then consider solutions of the system for $\tau' s$ satisfying $-r < -\tau < 0$.

Let $y_3(t, \tau)$ be a solution to (2.21) (we temporally suppress notation indicating the dependence of solutions on θ). Then for $t > 0$ we define $y_{3t}(\cdot) \in L^2(-\tau, 0; \mathbb{R}^n)$ by the past history $y_{3t}(s, \tau) = y_3(t+s, \tau), \quad -\tau < s < 0$. If $z_1(t, \tau) = (y_3(t, \tau), y_{3t}(\cdot, \tau))^T$, then $z_1(t, \tau)$ is a solution to the abstract Cauchy problem

$$
\begin{aligned}
\frac{dz_1(t)}{dt} &= A(t, \tau) z_1(t, \tau) \\
z_1(0, \tau) &= (0, 0)^T \in Z_1,
\end{aligned} \tag{2.22}
$$

where $D(A(t, \tau)) = \{(\xi, \phi(\cdot))^T : \xi \in \mathbb{R}^n, \phi(\cdot) \in H^1(-\tau, 0; \mathbb{R}^n), \phi(0) = \xi\}$, and

$$
A(t, \tau) \begin{bmatrix} \xi \\ \phi(\cdot) \end{bmatrix} = \begin{bmatrix} G_x(x(t), x(t-\tau), \theta)\xi + G_{\tilde{x}}(x(t), x(t-\tau), \theta)[\phi(-\tau) - \dot{x}(t-\tau)] \\ \phi'(\cdot) \end{bmatrix}.
$$

Note that $A(t, \tau)$ is a vector affine operator on $z_1(t) = (y_3(t), y_{3_t}(\cdot))^T$. Moreover, we note that for $x \in H^{1,\infty}(-\tau, T; \mathbb{R}^N)$, we have existence of a unique solution to (2.21) or equivalently, (2.22). Thus we must argue that $\frac{\partial x}{\partial \tau}$ exists and also satisfies (2.21) (or (2.22)).

The proof follows the arguments for Theorem 1 with the G_θ term replaced by $G_{\tilde{x}}\{-\dot{x}(t-\tau)\}$. That is, one defines the differences $\tilde{m}_1(t, \tau, h) = x(t, \tau+h) - x(t, \tau)$ and the difference quotients

$$
\frac{|\tilde{m}_1(t, \tau, h) - y_3(t, \tau)h|}{|h|}
$$

corresponding to solutions z_1 of (2.22). Then arguments exactly like those in the proof of

Theorem 1 provide the desired results for existence of $\frac{\partial x(t)}{\partial \tau}$ and that it satisfies (2.21).

We begin by defining the solution to (2.22) with successive approximations for $k = 0, 1, 2, \ldots,$

$$z_1^0(t;\tau,\theta) = \begin{cases} \Phi(t) & -\tau \le t < 0 \\ z_{10} & t \ge 0. \end{cases} \tag{2.23}$$

$$z_1^{k+1}(t, x_0, \Phi, \tau, \theta) = \begin{cases} \Phi(t), & -\tau \le t < 0 \\ z_{10}, & t = 0 \\ \int_0^t A(s;\tau,\theta) z^k(s;\tau,\theta) ds, & t > 0. \end{cases} \tag{2.24}$$

Existence of the solution occurs when the successive approximations for $z_1(t;\tau,\theta)$ converge as $k \to \infty$.

For $t \in I = [0, T]$, $\tau \in [-r, 0]$, and a fixed θ,

$$\begin{aligned} |z^1(t;\tau,\theta) - z_{10}| &= |\int_0^t A(s;\tau,\theta) z_1^0(s;\tau,\theta) ds| \\ &\le \int_0^t | \begin{bmatrix} M_0 + M_1 \\ \phi'(\cdot) \end{bmatrix} (z_1^0(s;\tau,\theta))| ds \\ &\le Q \int_0^t |z_1^0(s;\tau,\theta)| ds \\ &\le Q \int_0^t |N_z| ds \\ &\le Q N_z t \end{aligned}$$

for $t \ge 0$, where $Q \ge \begin{bmatrix} M_0 + M_1 \\ \phi'(\cdot) \end{bmatrix}$, and z_1^0 is bounded by N_z. Thus the first two approximations for $z_1(t;\tau,\theta)$ converges.

Let $\Delta_z^k(t;\tau,\theta) = |z_1^{k+1}(t;\tau,\theta) - z_1^k(t;\tau,\theta)|$, then

$$\begin{aligned} \Delta_z^k(t;\tau,\theta) &= |\int_0^t A(s;\tau,\theta) z_1^{k+1}(s;\tau,\theta) - A(s;\tau,\theta) z_1^k(s;\tau,\theta) ds| \\ &\le \int_0^t |A(s;\tau,\theta)(z_1^k(s;\tau,\theta) - z_1^{k-1}(s;\tau,\theta))| ds \\ &\le Q \int_0^t |(z_1^k(s;\tau,\theta) - z_1^{k-1}(s;\tau,\theta))| ds \\ &\le Q \int_0^t \Delta_z^{k-1}(s;\tau,\theta) ds. \end{aligned}$$

We claim that

$$\Delta_z^k(t;\tau,\theta) \leq \frac{M_z Q^k t^{k+1}}{(k+1)!}, \tag{2.25}$$

where $M_z = QN_z$. Then for $k=0$ and $t \in I$

$$\Delta_z^0(t;\tau,\theta) = |z^1(t;\tau,\theta) - z_{10}| \leq M_z t$$

thus by induction we have

$$\begin{aligned} \Delta_z^k(t;\tau,\theta) &\leq Q\int_0^t \Delta_z^{k-1}(s;\tau,\theta)ds & (2.26)\\ &\leq Q\int_0^t \frac{M_z Q^k s^{k+1}}{(k+1)!}ds & (2.27)\\ &\leq \frac{M_z Q^{k+1} t^{k+2}}{(k+2)!}. & (2.28) \end{aligned}$$

Therefore (2.25) holds for all k, and the series for $\sum_{k=0}^{\infty} \Delta_z^k(t;\tau,\theta)$ is dominated by the power series for $\frac{M_z e^{Qt}}{Q}$. Using the comparison test, the series $\sum_{k=0}^{\infty} \Delta_z^k(t;\tau,\theta)$ converges uniformly on I. As a result, the series

$$z_{10} + \sum_{k=0}^{\infty}(z_1^{k+1}(t;\tau,\theta) - z_1^k(t;\tau,\theta))$$

converges uniformly and absolutely on I, and the partial sum

$$z_{10} + \sum_{k=0}^{n-1}(z_1^{k+1}(t;\tau,\theta) - z_1^k(t;\tau,\theta))$$

converges uniformly to a continuous function z_1 on interval I. z_1 exists on I, and $A(t;\tau,\theta)$ exists for $t \in I$, then given that $z_1^k(t;\tau,\theta)$ converges uniformly to $z_1(t;\tau,\theta)$, and $t \to A(t;\tau,\theta)$ is continuous

$$z_1^{k+1}(t;\tau,\theta) = z_{10} + \int_0^t A(s;\tau,\theta) z_1^k(s;\tau,\theta)ds$$

becomes

$$z_1(t;\tau,\theta) = z_{10} + \int_0^t A(s;\tau,\theta) z_1(s;\tau,\theta)ds$$

as $k \to \infty$. Thus $z_1(t;\tau,\theta)$ exists and is a solution to (2.22).

Next, we assume that $z_1^1(t,\tau)$ and $z_1^2(t,\tau)$ are both solutions to (2.22). Then

$$\begin{aligned} |z_1^1(t,\tau) - z_1^2(t,\tau)| &= |\int_0^t A(s,\tau)z_1^1(s,\tau) - A(s,\tau)z_1^2(s,\tau)ds| \\ &\leq Q\int_0^t |z_1^1(s,\tau) - z_1^2(s,\tau)|ds \end{aligned}$$

for a constant Q.

We apply Grownwall's inequality:

$$|z_1^1(s,\tau) - z_1^2(s,\tau)| \leq 0e^{Qt} = 0.$$

Thus $z_1^1(s,\tau) = z_1^2(s,\tau)$, so $z_1(s,\tau)$ is the unique solution of (2.22).

To argue that the solution to (2.22) depends continuously on θ, let $h(t;\tau,\theta,\theta_0) = z_1(t,\tau,\theta) - z_1(t,\tau,\theta_0)$, where now we need to express explicitly the dependence of solutions on θ. Then

$$\begin{aligned} |h(t;\tau,\theta,\theta_0)| &= |z_1(t,\tau,\theta) - z_1(t,\tau,\theta_0)| \\ &\leq \int_0^t |A(s,\tau,\theta)z_1(s,\tau,\theta) - A(s,\tau,\theta_0)z_1(s,\tau,\theta_0)|ds \\ &\leq \int_0^t \Big\{|A(s,\tau,\theta)z_1(s,\tau,\theta) - A(s,\tau,\theta)z_1(s,\tau,\theta_0)| \\ &\qquad +|A(s,\tau,\theta)z_1(s,\tau,\theta_0) - A(s,\tau,\theta_0)z_1(s,\tau,\theta_0)|\Big\}ds \\ &\leq \int_0^t |A(s,\tau,\theta)(z_1(s,\tau,\theta) - z_1(s,\tau,\theta_0))|ds \\ &+ \int_0^t |(A(s,\tau,\theta) - A(s,\tau,\theta_0))z_1(s,\tau,\theta_0)|ds \\ &\leq Q\int_0^t |z_1(s,\tau,\theta) - z_1(s,\tau,\theta_0)|ds + r(T;\tau,\theta), \end{aligned} \tag{2.29}$$

where

$$r(T;\tau,\theta) = \int_0^t |A(s,\tau,\theta) - A(s,\tau,\theta_0)||z_1(s,\tau,\theta_0)|ds.$$

Since $A(t;\tau,\theta)$ is continuous due to assumptions on $G(x(t),\tilde{x}(t),\theta)$, so that as $\theta \to \theta_0$

$$\lim_{\theta\to\theta_0} |r(T;\tau,\theta)| = \lim_{\theta\to\theta_0} \int_0^T |A(s,\tau,\theta) - A(s,\tau,\theta_0)||z_1(s,\tau,\theta_0)|ds = 0.$$

If we apply Gronwall's inequality and take the limit as $\theta \to \theta_0$ in (2.29), then

$$\lim_{\theta \to \theta_0} |h(t; \tau, \theta, \theta_0)| \leq \lim_{\theta \to \theta_0} |r(T; \tau, \theta)| e^{Qt} = 0.$$

Thus, we have

$$\lim_{\theta \to \theta_0} z_1(t, \tau, \theta) = z_1(t, \tau, \theta_0).$$

Next we argue that the solution to (2.22), depends continuously on the delay τ whenever $\dot{x}$ is continuous. Let $h^\tau(t; \tau, \tau^*, \theta) = z_1(t, \tau, \theta) - z_1(t, \tau^*, \theta)$ for a fixed $\theta \in \mathbb{R}^p$ and fixed $\tau^* \in [-r, 0]$. We have

$$\begin{aligned}
|h^\tau(t; \tau, \tau^*, \theta)| &= |z_1(t, \tau, \theta) - z_1(t, \tau^*, \theta)| \\
&\leq \int_0^t |A(s, \tau, \theta) z_1(s, \tau, \theta) - A(s, \tau^*, \theta_0) z_1(s, \tau^*, \theta)| ds \\
&\leq \int_0^t \Big\{ |A(s, \tau, \theta) z_1(s, \tau, \theta) - A(s, \tau, \theta) z_1(s, \tau^*, \theta)| \\
&\quad + |A(s, \tau, \theta) z_1(s, \tau^*, \theta) - A(s, \tau^*, \theta) z_1(s, \tau^*, \theta)| \Big\} ds \\
&\leq \int_0^t |A(s, \tau, \theta)(z_1(s, \tau, \theta) - z_1(s, \tau^*, \theta))| ds \\
&\quad + \int_0^t |(A(s, \tau, \theta) - A(s, \tau^*, \theta)) z_1(s, \tau^*, \theta)| ds \\
&\leq Q \int_0^t |(z_1(s, \tau, \theta) - z_1(s, \tau^*, \theta))| ds \\
&\quad + \int_0^t |[A(s, \tau, \theta) - A(s, \tau^*, \theta)] z_1(s, \tau^*, \theta)| ds \\
&\leq Q \int_0^t |(z_1(s, \tau, \theta) - z_1(s, \tau^*, \theta))| ds + r^\tau(T; \tau, \theta), \qquad (2.30)
\end{aligned}$$

where

$$r^\tau(T; \tau, \theta) = \int_0^T |A(s, \tau, \theta) - A(s, \tau^*, \theta)| |z_1(s, \tau^*, \theta)| ds.$$

Since $x \in C^1(-\tau, T; \mathbb{R}^N)$, then as $\tau \to \tau^*$, $|A(t, \tau, \theta) - A(t, \tau^*, \theta)| \to 0$. Moreover, $A(t, \tau, \theta)$ is bounded in t, τ. As a result, when $\tau \to \tau^*$, the limit of $|r^\tau(T; \tau, \theta)|$ is 0. Then if we apply Gronwall's inequality and take the limit as $\tau \to \tau^*$ in (2.30), we find

$$\lim_{\tau \to \tau^*} |h^\tau(t; \tau, \tau^*, \theta)| \leq \lim_{\tau \to \tau^*} |r^\tau(T; \tau, \theta)| e^{Qt} = 0,$$

and thus

$$\lim_{\tau \to \tau^*} z_1(t, \tau, \theta) = z_1(t, \tau^*, \theta).$$

Chapter 3

Parameter Estimation and Sensitivity Functions

3.1 Parameter Estimation

3.1.1 Mathematical and Statistical Models

Before we determine the generalized sensitivity functions of a physical, sociological, or biological situation we first introduce the details that surround parameter estimation or inverse problems. Given the dynamical system model we have previously discussed with the observation process described in (2.1), we assume the observations are discrete such that

$$\eta(t_j) = h(t_j, x(t_j), x(t-\tau), \theta), j = 1, \ldots, n_d.$$

In general, the data $\{\eta_j\}$ is not exactly $\eta(t_j)$, due to uncertainty within the observations. Thus we use a statistical model to better model the uncertainty in the data. We consider a statistical model of the form

$$Y_j = f(t_j, q_0) + \mathcal{E}_j, j = 1, \ldots, n_d,$$

where $f(t_j, q_0) = h(t_j, x(t_j), x(t-\tau), \theta)$, $q_0 = (\theta_0, \tau_0, x_{00})$, and $j = 1, \ldots, n_d$, corresponds to the solution of the mathematical model described in (2.1), at the j^{th} time of the solution vector $x(t) \in \mathbb{R}^n$, $\eta(t) \in \mathbb{R}^m$ and $q \in \mathbb{R}^{p+1+n}$. Note that since $x(t)$ depends on the intial condition x_0 this n vector may also be a "parameter" to be estimated from data. Hence we have $q = (\theta, \tau, x_0) \in \mathbb{R}^{p+1+n}$. Here q_0 represents the true value of the parameters that generates the observations $\{Y_j\}_{j=1}^{n_d}$[9]. The existence of $q_0 = (\theta_0, \tau_0, x_{00})$ is standard in statistical formulation, and assuming that $E[\mathcal{E}_j] = 0$, it is implied that (2.1) describes the biological, sociological, or physical process correctly. $\mathcal{E}_j$ are random variables that account for measurement error, and or random phenomena that happens such that $\{\eta_j\}$ is not exactly $\eta(t_j)$ [9]. We assume that $\mathcal{E}_j$ has an unknown probability distribution with mean zero. The previously described assumptions on the model and measurement error will help us determine the correct way to estimate q to represent the true parameter value q_0.

As discussed earlier we assume that we have the following observation process

$$Y_j = f(t_j, q_0) + \mathcal{E}_j, \tag{3.1}$$

where $j = 1, \ldots, n_d$, and the $\mathcal{E}_j$ are independent identically distributed, such that $E(\epsilon_j) = 0$, and the constant variance is σ_0^2. Realizations of Y_j are given by

$$y_j = f(t_j, q_0) + \epsilon_j. \tag{3.2}$$

These realizations are used to obtain $\hat{q}$, and $\hat{\sigma}^2$ using ordinary least squares estimation method since the statistical model assumes constant variance. We define

$$q_{OLS}(Y) = q_{OLS}^{n_d}(Y) = \arg\min_{q \in Q_{ad}} \sum_{j=1}^{n_d} \left[Y_j - f(t_j, q)\right]^2, \tag{3.3}$$

where $Q_{ad} \in \mathbb{R}^m$ and $m = p + 1 + n$, then

$$\hat{q} = \arg\min_{\theta \in Q_{ad}} \sum_{j=1}^{n_d} \left[y_j - f(t_j, q)\right]^2, \tag{3.4}$$

since $\mathcal{E}_j = Y_j - f(t_j, q)$ is a random variable and y_j is a realization of Y_j. Thus $\hat{q}$ will be a realization of q_{OLS}.

3.1.2 Generalized Sensitivity Functions

To compute the generalized sensitivity, based on the mathematical and statistical model defined in equation (3.1) the following must be known: TSFs (∇f), the Fisher Information Matrix (FIM) F_G, and the variance σ_0^2.

The TSFs gives insight to the relationship between the model and the parameter. When computing this function, we are able to determine the time intervals where the model is most affected by the parameter in question. A general definition for the TSF of a model is

$$\frac{d}{dt}\frac{\partial x}{\partial \theta} = \frac{\partial G}{\partial x}\frac{\partial x}{\partial \theta}(t) + \frac{\partial G}{\partial \theta}(t),$$

given

$$\begin{aligned} \frac{dx(t)}{dt} &= G(x(t),\theta), t>0 \\ x(0) &= x_0, \\ \eta(t) &= f(x(t),\theta), t\in[0,T], \end{aligned}$$

where $x(t) \in \mathbb{R}^n$ are vectors of state variables of the system, $\eta(t) \in \mathbb{R}^m$ is the vector of measurable outputs, and $\theta \in \mathbb{R}^p$ is the vector of parameters. G, and f must be differentiable in order to construct the TSFs, and also must be sufficiently smooth to construct the GSFs that will be defined later.

The continuous FIM is defined as

$$F_G(T) = \int_0^T \frac{1}{\sigma(t)} \nabla_q f(t,q_0) \nabla_q f(t,q_0)^T dP(t), \tag{3.5}$$

where q_0 is the true parameter value for $f(t,q)$, T is the final time, and P is some measure P defined on the time interval [8]. We observe that the definition of the measure P can change the definition of the FIM, and should be chosen such that optimal information can be gained from the data used to estimate the parameters.

The variance is defined as

$$\sigma_0^2 = \frac{1}{n-(p+1)} E\left[\sum_{j=1}^{n_d} [Y_j - f(t_j,q_0),]\right] \tag{3.6}$$

where n is the number of time points, and Y_j represents the observation process described in equation (3.1). These functions are approximated using the estimate $\hat{q}$ for q_0, which is unknown, and they must be smooth with respect to some general measure P over the time interval $[0,T]$.

We use the traditional sensitivity functions, variance and the FIM, to compute the generalized sensitivity functions which has the following definition:

$$gs(t) = \int_0^t \left[F_G(T)^{-1} \frac{1}{\sigma^2(s)} \nabla_q f(s,q_0)\right] \bullet \nabla_q f(s,q_0) dP(s), \;\; t\in[0,T].$$

When using a discrete measure P (i.e., $P = \sum_{i=1}^n \Delta_{t_i}$), the generalized sensitivity functions are then approximated with $\hat{\theta}$ as shown below:

$$gs(t_l) = \sum_{k=1}^{l} \left(\left[\sum_{j=1}^{n_d} \frac{1}{\sigma^2(t_j)} \nabla_q f(t_j,\hat{q}) \nabla_q f(t_j,\hat{q})'\right]^{-1} \times \frac{\nabla_q f(t_k,\hat{q})}{\sigma^2(t_k)} \right) \bullet \nabla_q f(t_k,\hat{q}) \;,$$

where $l = 1 \ldots n_d$. A discrete measure is usually appropriate when modeling biological, physiological, and sociological processes because generally data is collected on a specific set of time points.

3.2 Example: Delay-Logistic Equation

The delay-logistic equation is appropriate when modeling a biological process such as tumor growth, a single species growth model with time delay, or a biological situation where a parasite's life cycle is shorter than the hosts life cycle, thus having a delayed effect on the population size of the hosts. This model is popularly known in population ecology as Hutchinson's equation (1948) [32], or Wright's equation. Wright used the following formulation of the delay logistic equation to describe the distribution of prime numbers in 1955 [45],

$$\frac{dx(t)}{dt} = rx(t)\left[1 - \frac{x(t-\tau)}{K}\right]. \tag{3.7}$$

The delay τ represents biological factors such as pregnancy time, hatching period, and renewal of food, any factors that influence changes in the population size [26]. The delay-logistic equation is appropriate for circular causal systems, which are systems where changes in one part of the system alter other parts in the system but at a different rate [32]. The following is an ecological example of a circular causal system, cheetahs hunt gazelles, as cheetahs evolve to be faster, gazelles evolve to be faster to avoid extinction [32]. The delay logistic equation is also appropriate to model tumor growth since tumor growth occurs when tumor cells proliferate. Proliferation will occur after the cell cycle, which is a circular causal system, is complete; thus the delay in equation (3.7) represents the time it takes the cell cycle to finish [29]. Hutchinson's equation is also used to model population growth and is appropriate for modeling of population growth with any mammal species, since gestation must occur to increase the population size. Thus the τ in equation (3.7) could represent the gestation period for the population modeled [32]. The delay logistic equation is also ideal when modeling a parasite-host interaction if the parasite finishes its life cycle in the host without killing the host [34]. In this case τ represents the lifespan of the parasite. As result of the many uses of the delay logistic equation, we use this model to gain insight about generalized sensitivities as they pertain to DDEs.

The follow equations represent the traditional sensitivity functions of the delay logistic equation, and are used to compute the generalized sensitivity functions as previously discussed.

$$
\begin{aligned}
\frac{\partial}{\partial r}\frac{dx(t)}{dt} &= r\left[1-\frac{x(t-\tau)}{K}\right]\frac{\partial x(t)}{\partial r}-\frac{rx(t)}{K}\frac{\partial x(t-\tau)}{\partial r}+x(t)\left[1-\frac{x(t-\tau)}{K}\right]\\
\frac{\partial}{\partial K}\frac{dx(t)}{dt} &= r\left[1-\frac{x(t-\tau)}{K}\right]\frac{\partial x(t)}{\partial K}-\frac{rx(t)}{K}\frac{\partial x(t-\tau)}{\partial K}+rx(t)\left[\frac{x(t-\tau)}{K^2}\right]\\
\frac{\partial}{\partial x_0}\frac{dx(t)}{dt} &= r\left[1-\frac{x(t-\tau)}{K}\right]\frac{\partial x(t)}{\partial x_0}-\frac{rx(t)}{K}\frac{\partial x(t-\tau)}{\partial x_0}
\end{aligned}
$$

Let $s_1(t)=\frac{\partial x(t)}{\partial r}$, $s_2(t)=\frac{\partial x(t)}{\partial K}$, $s_3=\frac{\partial x(t)}{\partial x_0}$ then

$$
\begin{aligned}
\frac{\partial s_1(t)}{\partial t} &= r\left[1-\frac{x(t-\tau)}{K}\right]s_1(t)-\frac{rx(t)}{K}s_1(t-\tau)+x(t)\left[1-\frac{x(t-\tau)}{K}\right]\\
\frac{\partial s_2(t)}{\partial t} &= r\left[1-\frac{x(t-\tau)}{K}\right]s_2(t)-\frac{rx(t)}{K}s_2(t-\tau)+rx(t)\left[\frac{x(t-\tau)}{K^2}\right].\\
\frac{\partial s_3(t)}{\partial t} &= r\left[1-\frac{x(t-\tau)}{K}\right]s_3(t)-\frac{rx(t)}{K}s_3(t-\tau).
\end{aligned}
$$

Then we can solve the previous system of DDEs, after solving for the solution $x(t)$ to the original model, to obtain the following sensitivity functions:

$$
\begin{pmatrix} \frac{\partial x}{\partial r}\\ \frac{\partial x}{\partial K}\\ \frac{\partial x}{\partial x_0} \end{pmatrix}
$$

and eventually their corresponding GSFs.

3.2.1 Traditional Sensitivity function for the Delay Equation

To obtain the GSF for τ, we must solve the following retarded DDE for the TSF with respect to τ.

$$
\frac{d}{dt}\frac{\partial x}{\partial \tau}=\frac{\partial}{\partial \tau}\left[rx(t)\left[1-\frac{x(t-\tau)}{K}\right]\right] \tag{3.8}
$$

The above equation is a neutral since when applying the chain rule the right hand side of the equation will be defined both by state variables and their derivatives on distinct time levels [21].

Note that the derivatives of the state variables may also be state dependent [21]. A neutral equation will tend to have more continuity issues in comparison to a retarded DDE. A more general form of the neutral DDE is as follows:

$$\dot{x}(t) = f(t, x(t,\theta), x(t-\tau,\theta), \dot{x}(t-\tau,\theta)), \quad t \geq 0 \tag{3.9}$$

By decoupling equation (3.8) from the model equation (3.7), and solving the equation by itself, (3.8) is not a neutral equation because the derivative of $\frac{\partial x(t-\tau)}{\partial \tau}$ with respect to time is not needed to determine the solution, and $\dot{x}(t-\tau)$ is defined.

To determine the solution numerically, we simplify equation (3.8) by following the rules of differentiation. Using the product rule on equation (3.8) we obtain the following:

$$r\frac{\partial x(t)}{\partial \tau}\left[1 - \frac{x(t-\tau)}{K}\right] + rx(t)\frac{\partial}{\partial \tau}\left[1 - \frac{x(t-\tau)}{K}\right].$$

The second part of the previous sum

$$rx(t)\frac{\partial}{\partial \tau}\left[1 - \frac{x(t-\tau)}{K}\right],$$

becomes

$$rx(t)\left[\frac{\partial}{\partial \tau}(1) - \frac{\partial}{\partial \tau}(\frac{x(t-\tau)}{K})\right].$$

This simplifies to

$$-\frac{rx(t)}{K}\left[\frac{\partial}{\partial \tau}(x(t-\tau))\right].$$

Next we apply the chain rule to

$$\frac{\partial}{\partial \tau}x(t-\tau),$$

and the simplified sum becomes

$$-\frac{rx(t)}{K}\left[-\dot{x}(t-\tau) + \frac{\partial x(t-\tau)}{\partial \tau}\right].$$

The entire equation can then be simplified and rewritten using $s_4(t) = \frac{\partial x(t)}{\partial \tau}$ as follows:

$$\frac{\partial s_4(t)}{\partial t} = r\left[1 - \frac{x(t-\tau)}{K}\right]s_4(t) - \frac{rx(t)}{K}\left[s_4(t-\tau) - \dot{x}(t-\tau)\right]. \tag{3.10}$$

We cannot directly solve $\frac{\partial x(t)}{\partial \tau}$ shown in equation (4.1), since existence is not automatically guaranteed by the initial past history of the solution $x(t)$. If the initial past history is not sufficiently smooth then $\dot{x}(t-\tau)$ may not exist on $t \in [0, \tau]$. We apply Theorem 3 to equation

(4.1) to ensure existence of a solution for $\frac{\partial x(t)}{\partial \tau}$.

We define $z(t) = (s_4(t), s_{4t}(t))$ and transform (4.1) to be

$$\frac{dz(t)}{dt} = A(t, \tau, \theta)z(t) + F(t) \tag{3.11}$$

where

$$A(t, \tau, \theta) \begin{bmatrix} \xi \\ \phi(\cdot) \end{bmatrix} = \begin{bmatrix} r[1 - \frac{x(t-\tau)}{K}]\xi + \frac{rx(t)}{K}[\phi(-\tau)] \\ \phi'(\cdot) \end{bmatrix},$$

$$F(t) = \begin{bmatrix} -\frac{rx(t)}{K}\dot{x}(t-\tau) \\ 0 \end{bmatrix},$$

and $z(t)$ is the assumed solution of the Cauchy problem (4.2), with a zero condition.

Chapter 4

Approximation of Delay Equations

4.1 Banks-Kappel Spline Approximation

To numerically solve our delay differential equations we must use a numerical method that has numerical convergence for a DDE model existing in an Hilbert Space, like $\mathbb{R}^n \times L^2(-\tau, 0)$. The Banks-Kappel spline approximation [12] can handle both constant and non-constant initial functions, for these reasons we use this spline approximation, which we will call the Banks-Kappel Method. The Banks-Kappel (BK) method is an approximation technique for functional differential equations (FDEs), and may be applied to linear and non-linear models. For the BK spline approximation the history of the function (i.e., initial condition ϕ, or $x(t-\tau)$) is approximated using splines and then discretized over time. This method is most advantageous for numerically approximating our dde example models and their corresponding traditional sensitivity functions since the method is adaptable to both changing initial functions, and the space to which the solution exists [12].

4.2 Banks-Kappel Splines for Hutchinson's Equation

Given our first example, the delay logistic equation $\dot{x} = rx[1 - \frac{x(t-\tau)}{K}]$, the traditional sensitivity function with respect to the delay is defined as follows:

$$\frac{\partial s_4(t)}{\partial t} = a(t)s_4(t) + b(t)s_4(t-\tau) + f(t) \tag{4.1}$$

where $a(t) = r\left[1 - \frac{x(t-\tau)}{K}\right]$, $b(t) = -\frac{rx(t)}{K}$, $f(t) = \frac{rx(t)\dot{x}(t-\tau)}{K}$, and initially $s_4(0) = 0$, and $s_4(t+\theta) = s_{4t}(\theta) = 0$ for $t \in [-\tau, 0]$. The solution to (4.1) may not be defined for all t, given certain initial functions, so to ensure a solution exists we define $z(t) = (s_4(t), s_{4t}(\theta))$ and transform (4.1) to be

$$\begin{aligned} \frac{dz(t)}{dt} &= A(t)z(t) + F(t) \\ z(0) &= (s_4(0), s_{40}(\theta)) = (0,0) \end{aligned} \tag{4.2}$$

where $F(t) = (f(t), 0)$. $A(t)$ is defined

$$A(t)(\phi(0), \phi) = (G(t, \phi), D\phi)$$

where $G(t, \phi) = a(t)\phi(0) + b(t)\phi(-\tau)$, and $D(A(t)) = \{(\phi(0), \phi) | \phi(0) \in \mathbb{R}, \phi \in H^1(-\tau, 0; \mathbb{R})\}$. We let $z(t) \in Z_1$ be the unique solution (known to exist uniquely by Theorem 3, of the Cauchy problem (4.2), where $Z_1 = \mathbb{R} \times L_2(-\tau, 0; \mathbb{R})$, and $D(A(t)) \subset Z_1 \to Z_1$.

To approximate (4.2) we use the extension of Banks-Kappel splines [12] with nonautonomous theory as described in the Banks and Rosen paper [17], since we have time dependent coefficients. We define Z_1^N to be a piecewise linear spline subspace of Z_1, and approximate (4.2) using the following differential equation

$$\begin{aligned} \frac{dz^N(t)}{dt} &= A^N(t)z^N(t) + F^N(t) \\ z^N(0) &= P^N(z(0), z_0) = 0 \end{aligned} \tag{4.3}$$

where $A^N(t) = P^N A(t) P^N$, $F^N(t) = P^N F(t)$, and $P^N : Z_1 \to Z_1^N$ is the orthogonal projection. We fix a basis $\hat{\beta}_1^N, \ldots, \hat{\beta}_k^N$ that spans Z_1^N such that

$$\beta^N = (\beta_1^N, \ldots, \beta_N^N)$$

and

$$\hat{\beta}^N = (\beta^N(0), \beta^N),$$

so the dimension of Z_1^N is $N+1$. Then $z^N(t) = \hat{\beta}^N w^N(t)$ where $w^N(t)$ is the coordinate vector of z^N with respect to the chosen basis. Our fixed basis for Z_1^N corresponds to the partition $t_i = -i(\tau/N)$ for $i = 0, \ldots, N$ and we define this basis by

$$\hat{\beta}^N = (\beta^N(0), \beta^N) \text{ where } \beta^N = (e_0^N, e_1^N, \ldots, e_N^N).$$

The basis elements e_j^N's are piecewise linear splines defined by the Kronecker symbol δ_{ij}, so

$$e_j^N(t_i) = \delta_{ij} \text{ for } i, j = 0, 1, \ldots, N.$$

An equivalent system for (4.2) is represented in the following equation

$$\begin{aligned} \frac{dw^N(t)}{dt} &= A^N(t)w^N(t) + f^N(t) \\ w^N(0) &= w_0^N, \end{aligned} \tag{4.4}$$

where $w^N(t)$ and $f^N(t)$ are coordinate vectors of length $N+1$ for $z^N(t)$ and $F^N(t)$ respectively. Moreover, $z^N(t) = \hat{\beta}^N w^N(t)$ and $F^N(t) = \hat{\beta}^N f^N(t)$. To solve (4.4) we must compute the initial condition $P^N(\xi, \phi)$, $A^N(t)$ a $(N+1) \times (N+1)$ matrix representation of $A^N(t)$, and $f^N(t)$ a $(N+1) \times 1$ vector representation of $F^N(t)$.

Using the orthogonal projection, $P^N : Z_1 \to Z_1^N$, we can uniquely determine an element, $\hat{\beta}^N \alpha^N = P^N(\xi, \phi)$, in Z_1^N for any element $(\xi, \phi) \in Z_1$ with the following orthogonality relationship,

$$P^N(\xi, \phi) - (\xi, \phi) \perp Z_1^N$$

which is equivalent to

$$<\hat{\beta}^N \alpha^N - (\xi, \phi), \hat{\beta}^N>_{Z_1} = 0.$$

To compute $P^N(\xi, \phi)$, we solve

$$Q^N \alpha^N = h^N(\xi, \phi), \tag{4.5}$$

where

$$Q^N =< \hat{\beta}^N, \hat{\beta}^N >_{Z_1} = \beta^N(0)^T \beta^N(0) + \int_{-\tau}^{0} \beta^N(p)^T \beta^N(p) dp$$

is a $(N+1) \times (N+1)$ matrix defined in the following way

$$Q^N = \frac{\tau}{N} \begin{bmatrix} N/\tau + \frac{1}{3} & \frac{1}{6} & 0 & \dots & 0 \\ \frac{1}{6} & \frac{2}{3} & \frac{1}{6} & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & & \frac{1}{6} & \frac{2}{3} & \frac{1}{6} \\ 0 & \dots & 0 & \frac{1}{6} & \frac{1}{3} \end{bmatrix}$$

and

$$h^N(\xi, \phi) =< \hat{\beta}^N, (\xi, \phi) >_{Z_1} = \beta^N(0)^T \xi + \int_{-\tau}^{0} \beta^N(p)^T \phi(p) dp$$

is a vector of length $(N+1)$. When we solve (4.5), we get the coordinate vector α^N of length $(N+1)$.

Next to compute $F^N(t) = \hat{\beta}^N f^N(t)$ where $f^N(t)$ is a coordinate vector of length $(N+1)$, we solve

$$Q^N f^N(t) = \beta^N(0) f(t)$$

where $f(t) = (\frac{rx(t)\dot{x}(t-\tau)}{K})$.

Finally to determine $A^N(t)$ for $\hat{\phi}^N = (\phi^N(0), \phi^N) \in Z_1^N$, we define $\alpha^N(t) \in \mathbb{R}^(N+1)$, and $\gamma^N(t) \in \mathbb{R}^(N+1)$ such that

$$\hat{\phi}^N = \hat{\beta}^N \alpha^N(t) \text{ and } A^N(t)\hat{\phi}^N(t) = \hat{\beta}^N \gamma^N(t).$$

Then $A^N(t)\hat{\phi}^N = P^N A(t)\hat{\phi}^N = P^N(G(t,\phi^N), D\phi^N)$, and $Q^N\gamma^N(t) = H^N(t)\alpha^N(t)$ where

$$\begin{aligned}
H^N(t) &= h(G(t,\phi^N), D\phi^N) \\
&= < \hat{\beta}^N, (G(t,\phi^N), D\phi^N) >_{Z_1} \\
&= \beta^N(0)^T G(t,\phi^N) + \int_{-\tau}^{0} \beta^N(p)^T D[\phi^N(p)]dp \\
&= \beta^N(0)^T G(t,\beta^N\alpha^N(t)) + \int_{-\tau}^{0} \beta^N(p)^T D[\beta^N(p)]\alpha^N(t)dp \\
&= \left(\beta^N(0)^T G(t,\beta^N)\right)\alpha^N(t) + \left(\int_{-\tau}^{0} \beta^N(p)^T D[\beta^N(p)]dp\right)\alpha^N(t) \\
&= \left(\beta^N(0)^T[a(t)\beta^N(0) + b(t)\beta^N(-\tau)] + \int_{-\tau}^{0} \beta^N(p)^T D[\beta^N(p)]dp\right)\alpha^N(t) \\
&= (H_1(t) + H_2)\alpha^N(t)
\end{aligned}$$

is a $(N+1)\times(N+1)$ matrix at time t. $H_1(t)$ and H_2 are explicitly defined as follows:

$$H_1(t) = \begin{bmatrix} a(t) & 0 & \dots & 0 & b(t) \\ 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ \vdots & \vdots & \dots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & 0 \end{bmatrix}$$

and

$$H_2 = \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} & 0 & \dots & 0 \\ \frac{1}{2} & 0 & -\frac{1}{2} & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & & \frac{1}{2} & 0 & -\frac{1}{2} \\ 0 & \dots & 0 & \frac{1}{2} & -\frac{1}{2} \end{bmatrix}.$$

Then

$$A^N(t) = (Q^N)^{-1} H^N(t),$$

is a $(N+1) \times (N+1)$ matrix, and $\gamma^N(t)$ is a solution of

$$Q^N \gamma^N(t) = H^N(t)\alpha^N(t)$$

where $\gamma^N(t)$ is a vector of length $(N+1)$ at time t. We use $\gamma^N(t)$ to obtain

$$\gamma^N(t) = A(t)\alpha^N.$$

4.3 Pseudocode for Implementation of Method

We want to solve the following differential equation

$$\begin{aligned} \frac{d\alpha^N(t)}{dt} &= A^N(t)\alpha^N(t) + f^N(t) \\ \alpha^N(0) &= \alpha_0^N, \end{aligned} \tag{4.6}$$

where $\gamma^N(t) = A^N(t)\alpha^N(t)$, $f^N(t)$ is the coordinate vector for $F^N(t)$, and $\alpha^N(t)$ is the coordinate vector for $z^N(t)$ where

$$\begin{aligned} \frac{dz^N(t)}{dt} &= A^N(t)z^N(t) + F^N(t) \\ z^N(0) &= P^N(z(0), z_0). \end{aligned}$$

Implementing the following algorithm will compute the traditional sensitivity functions (4.6), the subroutines will be defined in the Appendix.

1. Set time vector

2. Calculate the solution for the model using the Banks-Kappel spline, our example will be the delay logistic equation:

$$\frac{dx(t)}{dt} = rx[1 - \frac{x(t-\tau)}{K}].$$

3. Set the number of nodes, $(N+1)$, and then define our nodes such that $\{\theta_i\} = \{-\tau, \ldots, \frac{-\tau}{N}, 0\}$, so $\theta_0 = 0$, and $\theta_N = -\tau$.

4. Compute Q^N in the subroutine **makeQ(N)** which returns the matrix Q^N in echelon form and is now called QN. This subroutine is discussed in the Appendix.

5. Set initial condition (ξ, ϕ), and find coordinate vector $\alpha_0^N = \alpha^N(t_0)$. To find the coordinate vector we solve $Q^N \alpha^N = h^N(\xi, \phi)$ for α^N where $\xi = z(0)$, and $\phi = z_0$. Next $h^N(\xi, \phi)$ will be computed using subroutine **makeh** where you pass the initial conditions and return the vector h. Then in the subroutine **backSub** we pass QN computed in step 3, and h and return the coordinate vector α_0^N. We will discuss both subroutines in later sections.

6a. We call

$$solution = ODE45(@RHS, tspan, \alpha_0^N),$$

where we pass the initial coordinate vector we computed in step four and

$$RHS = \gamma^N(t) + f^N(t).$$

The coordinate vectors $\gamma^N(t)$, $f^N(t)$, and $\alpha^N(t)$ are computed during the use of MATLAB'S ODE 45 (steps 5b & 5c), a variable step Runge-Kutta Method.

6b. Compute $f^N(t)$ by solving $Q^N f^N(t) = (e_0(0), e_1(0), \ldots, e_N(0))^T(\frac{rx(t)\dot{x}(t-\tau)}{K})$.

$$\begin{aligned} Q^N f^N(t) &= (e_0(0), e_1(0), \ldots, e_N(0))^T(\frac{rx(t)\dot{x}(t-\tau)}{K}) \\ &= \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}(\frac{rx(t)\dot{x}(t-\tau)}{K}) \\ Q^N f^N(t) &= \begin{pmatrix} \frac{rx(t)\dot{x}(t-\tau)}{K} \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}. \end{aligned}$$

We would normally use our subroutine backSub to solve for $f^N(t)$, however since the vector

$$(e_0(0), e_1(0), \ldots, e_N(0))^T f(t)$$

has only one non-zero entry as shown above, we can directly compute

$$f^N(t) = \begin{pmatrix} \frac{rx(t)\dot{x}(t-\tau)}{K}/a_1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

since a_1 is the first entry from QN, the row reduced version of Q^N. Note that $QN(1,1) = Q^N(1,1)$. Also to compute $\dot{x}(t-\tau)$ we evaluate the following equation $\dot{x}(t-\tau) = rx(t-\tau)\left[1 - \frac{x(t-2\tau)}{K}\right.$.

6c. We next compute $H^N(t) = H_1(t) + H_2$ in the subroutine **makeH** where we pass $x(t), x(t-\tau), r, K$ and return H(t). Note in the subroutine H_2 will already be defined as it never changes in time, and

$$
H_1(t) = \begin{pmatrix} r[1-\frac{x(t-\tau)}{K}] & 0 & \dots & 0 & -\frac{rx(t)}{K} \\ 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ \vdots & \vdots & \dots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & 0 \end{pmatrix}
$$

and

$$
H_2 = \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} & 0 & \dots & 0 \\ \frac{1}{2} & 0 & -\frac{1}{2} & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & & \frac{1}{2} & 0 & -\frac{1}{2} \\ 0 & \dots & 0 & \frac{1}{2} & -\frac{1}{2} \end{pmatrix}.
$$

Once we compute $H(t)$, we solve $Q^N\gamma^N(t) = H^N(t)\alpha^N(t)$ by passing QN, the echelon form of Q^N, and $H(t)\alpha^N(t)$ in the subroutine **backSub** and we will return the coordinate vector $\gamma^N(t)$, where $\gamma^N(t) = A^N(t)\alpha^N(t)$.

4.4 Numerical Comparison of BK-Splines, DDE23, and Method of Steps

We compare the Banks-Kappel (BK) spline approximation [12] for $N = 8, 16, 32, 64$ with MATLAB's DDE23 [43], using the delay logistic example at delay values of $\tau = .1, 1$, and $\frac{\pi}{2r}$. These values of τ will exhibit different solution behaviors. MATLAB's DDE23 is an extended ode23 solver using the method of steps to approximate the solution [43]. The Banks-Kappel spline approximation scheme is based on classical least squares approximations [12]. We later compare the BK method with the analytical solution obtained using the Method of Steps [34]. The parameter values used for the delay logistic equation are as follows: $x_0 = .1, r = .7, K = 17.5$ for $t \in [0, 50]$, where the initial function is $\phi(t) = .1$ for the constant case, and $\phi(t) = sin(\frac{2\pi t}{\tau})$ for the non-constant initial function. We will also use as an initial non-constant function $\phi(t) = y(t)$ which will be defined later.

The delay logistic equation

$$
\frac{dx(t)}{dt} = rx(t)\left[1 - \frac{x(t-\tau)}{K}\right]
$$

exhibits different behavior based on the value of the delay. For smaller values of τ the solution tends to behave similar to that of the non-delay logistic equation; however, if $r\tau > \pi/2$ the solution oscillates around the carrying capacity, K [34]. Also if the delay τ is sufficiently large it will cause the solution to die out rather quickly. Given our parameter values and the values for τ, the solution will go slightly past the carrying capacity, and then return to the steady state K for $\tau = 1$, and a oscillatory solution around K will occur at $\tau = \frac{\pi}{2r}$. Since we do not have the exact solution over the entire interval $t \in [0, 50]$ we compare the BK spline approximation with MATLAB's DDE23 method by looking at the convergence of the solutions for different N, where the mesh for t is chosen by DDE23, thus the mesh varies for different values of τ. We also compare BK spline approximation with the exact solution obtained using the Method of Steps, over the interval $t \in [0, 2\tau]$ for a constant initial function, and $t \in [0, \tau]$ for a non-constant initial function.

BK Spline Approximation vs. DDE23 for $\phi(t) = .1$

We compare the solution for the delay logistic equation using the Banks-Kappel spline approximation and MATLAB's DDE23 for $\tau = .1, 1$, and $\frac{\pi}{2r}$, and $N = 8, 16, 32, 64$ for the constant initial function $\phi(t) = .1$.

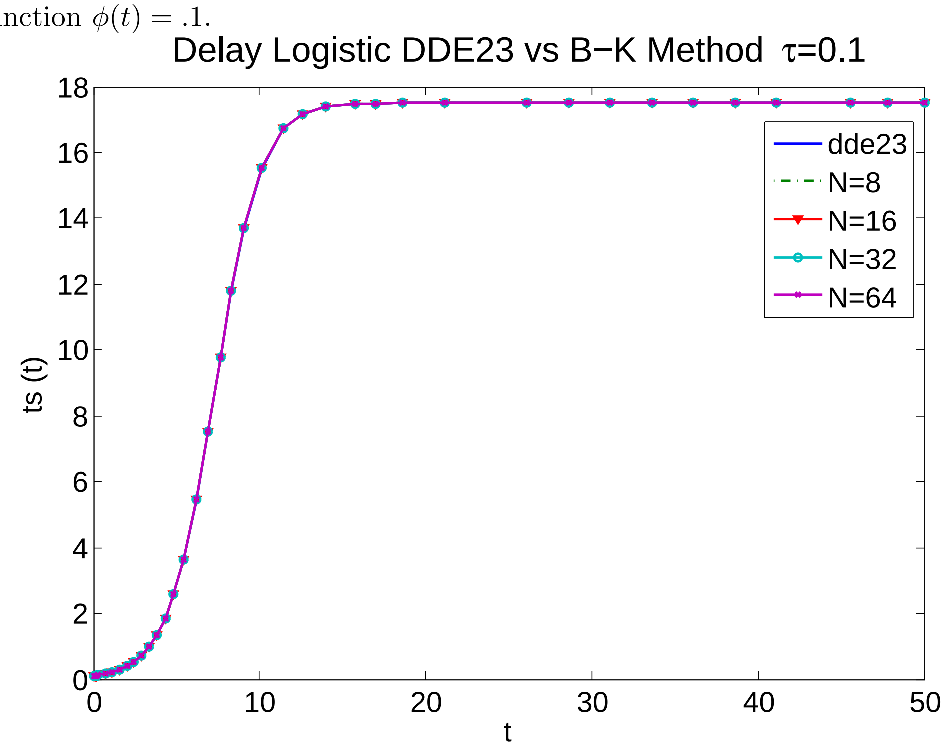


Figure 4.1: The numerical solution using the constant function $\phi(t) = .1$ for $\tau = .1$.

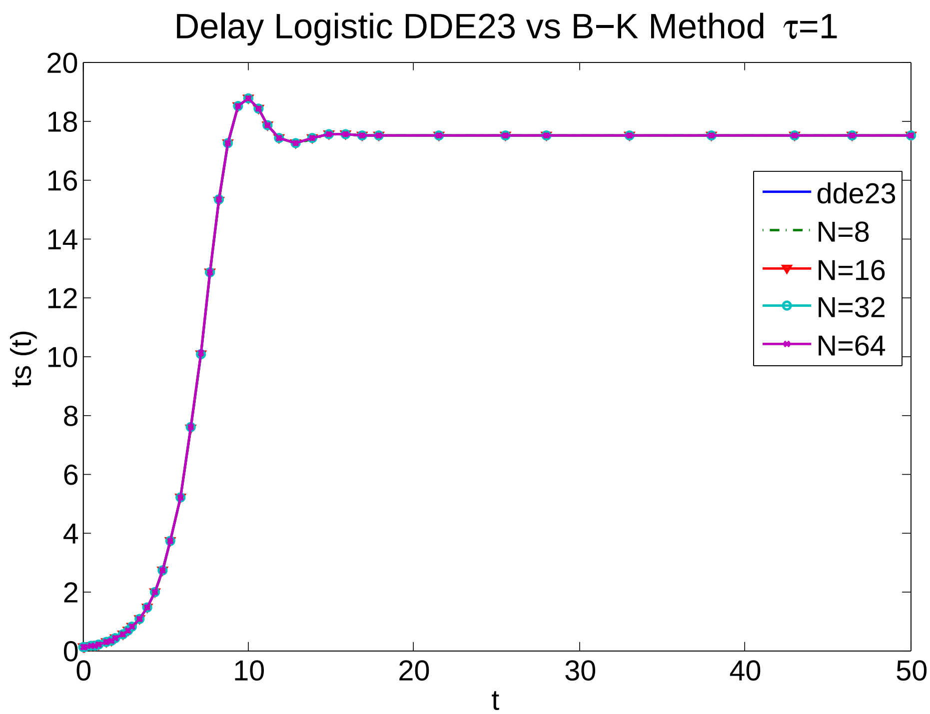


Figure 4.2: The numerical solution using the constant function $\phi(t) = .1$ for $\tau = 1$.

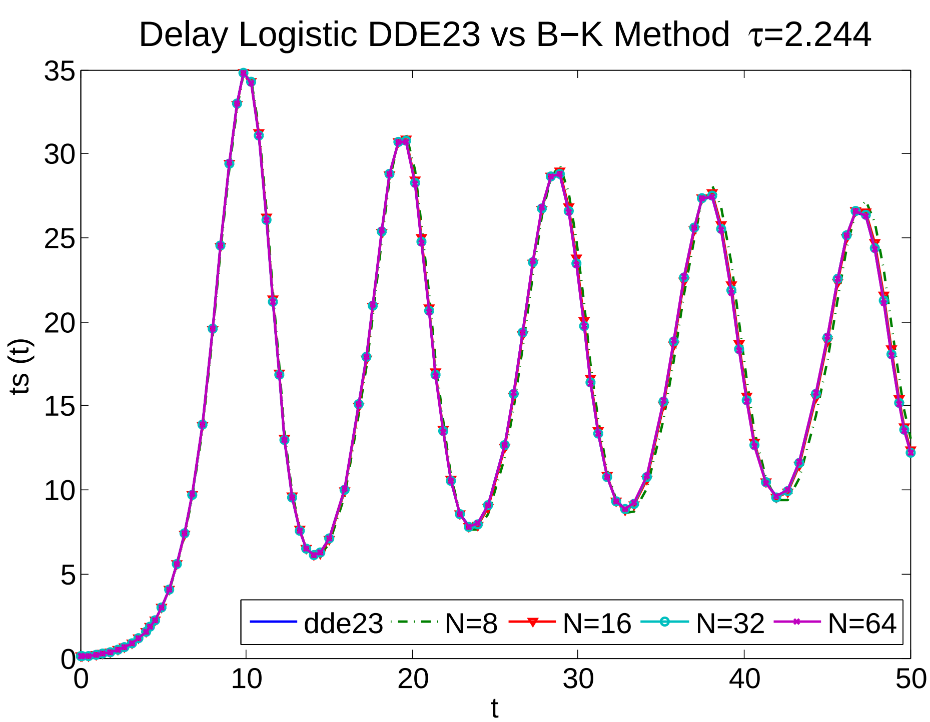


Figure 4.3: The numerical solution using the constant function $\phi(t) = .1$ for $\tau = \frac{\pi}{2r}$.

From the figures above we can see that both methods perform well in approximating the solution to the delay logistic equation with a constant initial function. For the BK method the solution for $N = 8$ is comparable to the solution for $N = 64$ with very little difference. There is also little difference between the dde23 and the BK method at any value of N.

BK Spline Approximation vs. DDE23 for $\phi(t) = sin(\frac{2\pi t}{\tau})$

We compare both methods in the same manner as in the last section, except we change the initial function, $\phi(t) = sin(\frac{2\pi t}{\tau})$.

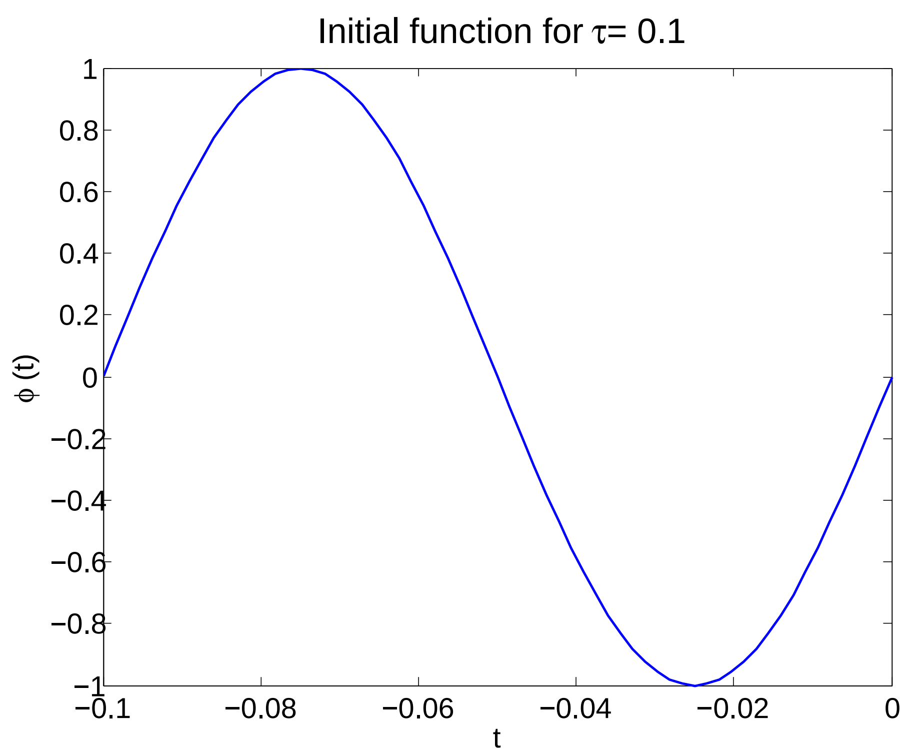


Figure 4.4: The initial condition $\phi(t) = sin(\frac{2\pi t}{\tau})$ for $t \in [-\tau, 0]$,

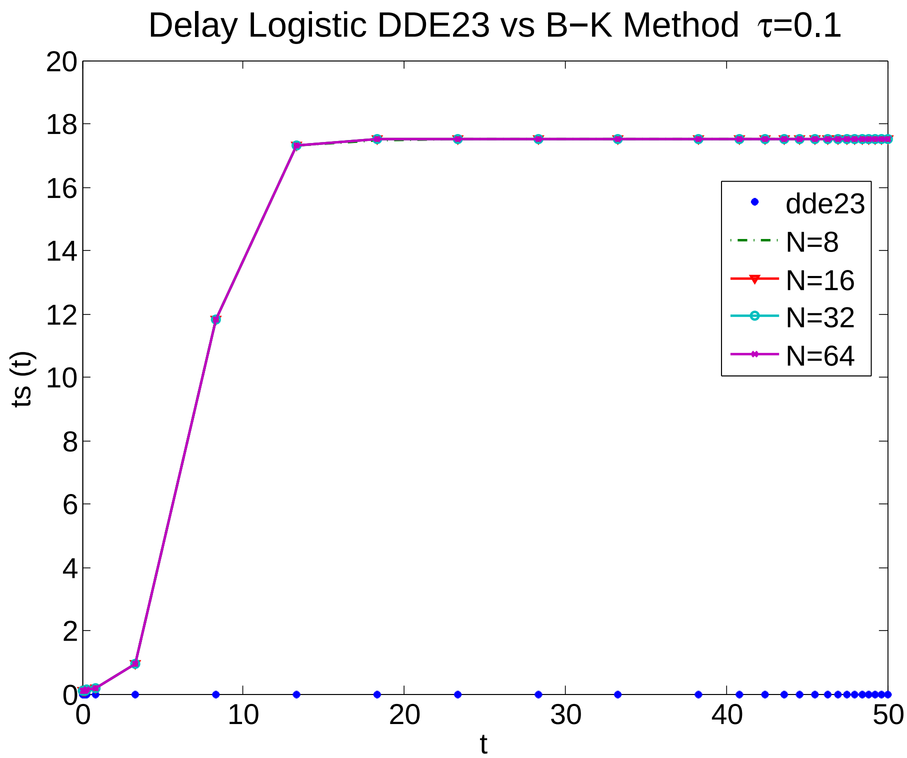


Figure 4.5: The numerical solution at $\tau = .1$.

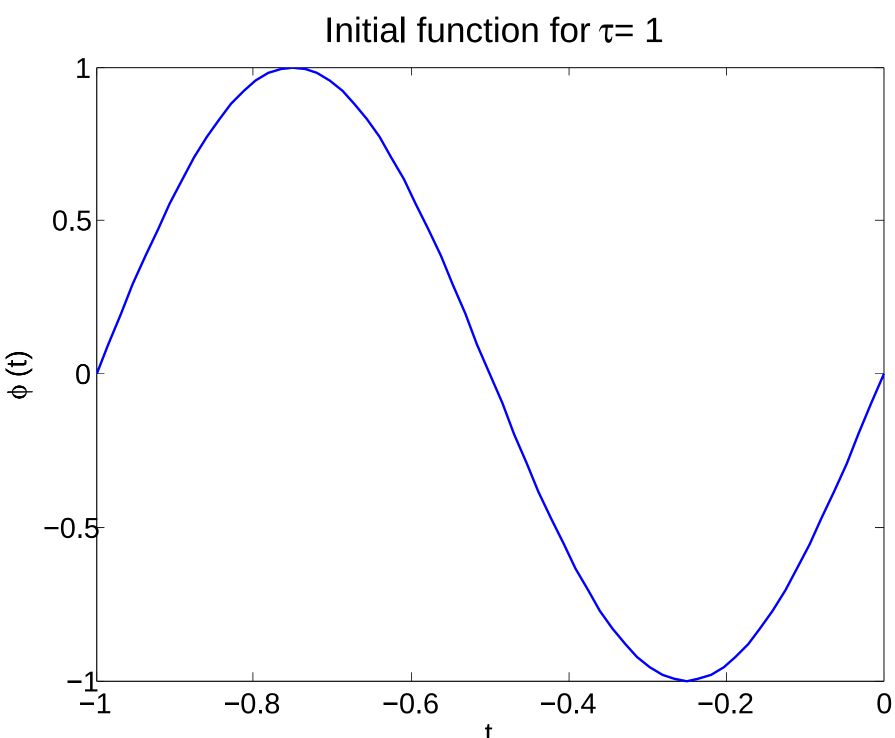


Figure 4.6: The initial condition $\phi(t) = sin(\frac{2\pi t}{\tau})$ for $t \in [-\tau, 0]$.

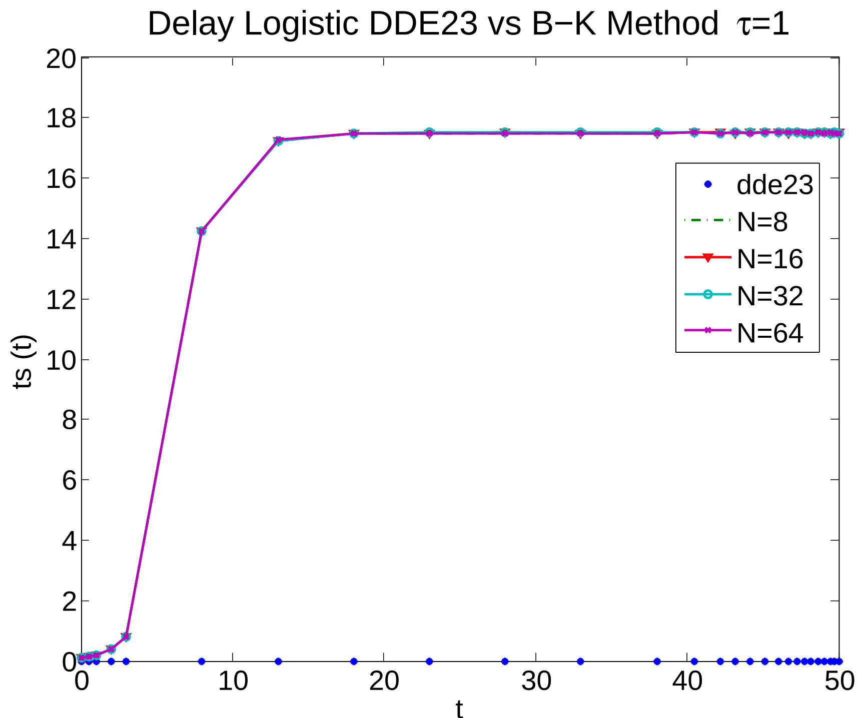


Figure 4.7: The numerical solution at $\tau = 1$.

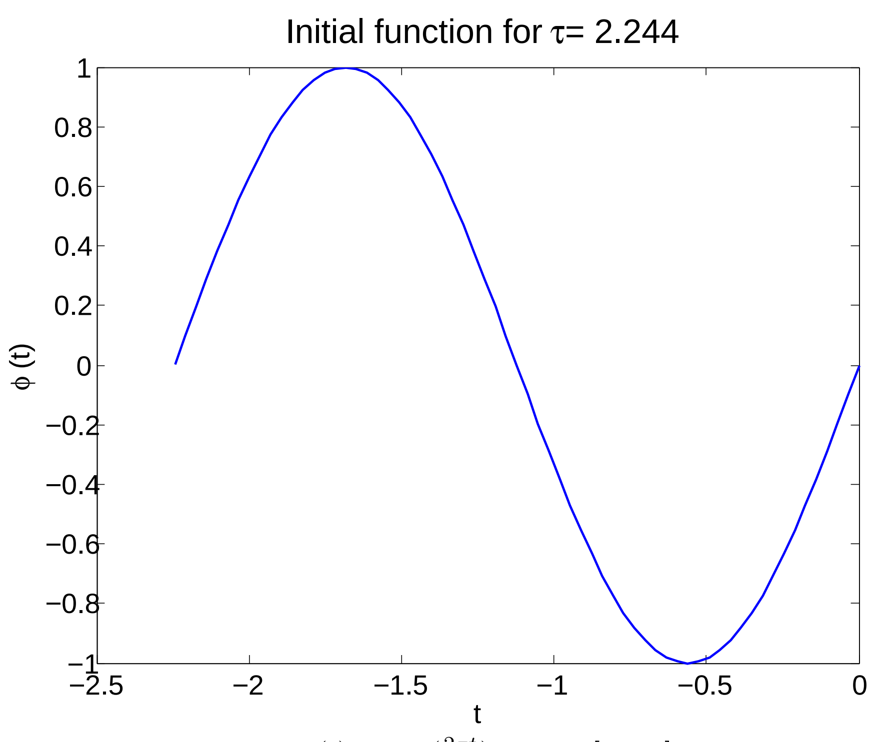


Figure 4.8: The initial condition $\phi(t) = sin(\frac{2\pi t}{\tau})$ for $t \in [-\tau, 0]$, and the numerical solution at $\tau = \frac{\pi}{2r}$.

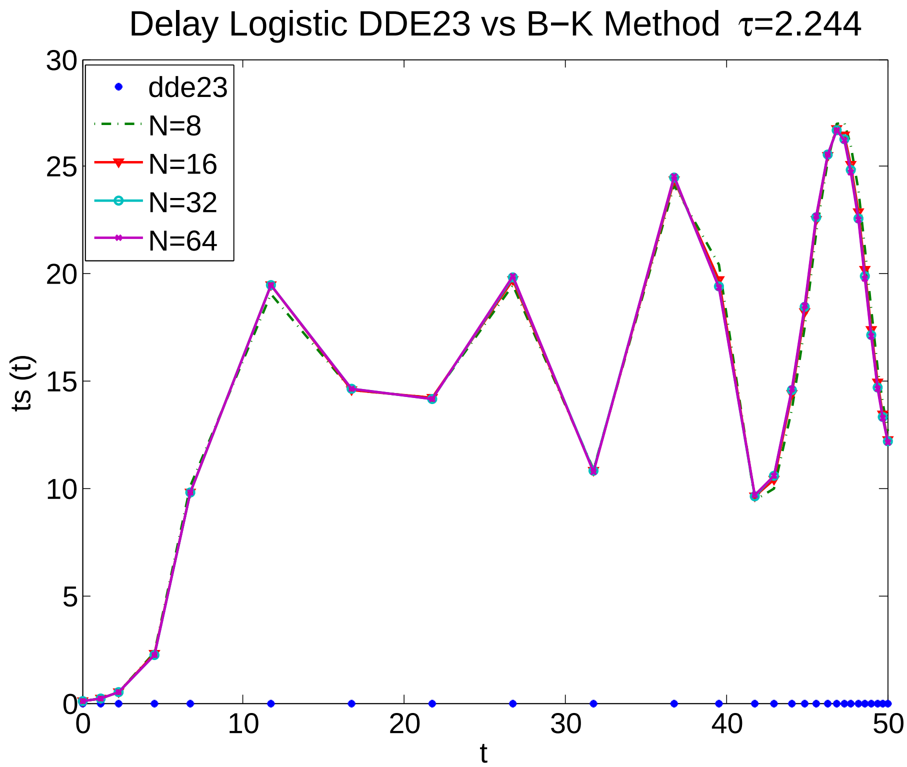


Figure 4.9: The initial condition $\phi(t) = sin(\frac{2\pi t}{\tau})$ for $t \in [-\tau, 0]$, and the numerical solution at $\tau = \frac{\pi}{2r}$.

As shown above MATLAB's dde23 does not perform as well when approximating the solution for the delay logistic equation for a non-constant initial function. For the BK spline approximation the solution for $N = 16$ is comparable to $N = 64$, and the BK method handles a varying initial condition better than dde23. The matlab routine cannot handle the initial function of $\phi(t) = sin(\frac{2\pi t}{\tau})$ due to the oscillatory nature of the solution and stiffness of the system. This behavior is expected and explained further in [43].

BK Spline Approximation vs. DDE23 for $\phi(t) = y(t)$

We compare both methods in the same manner as in the previous cases, now $\phi(t) = y(t)$, where

$$y(t) = \begin{cases} \tau + t, & -\tau \leq t \leq \frac{-\tau}{2} \\ \frac{\tau}{2}, & -\frac{-\tau}{2} < t < 0,. \end{cases}$$

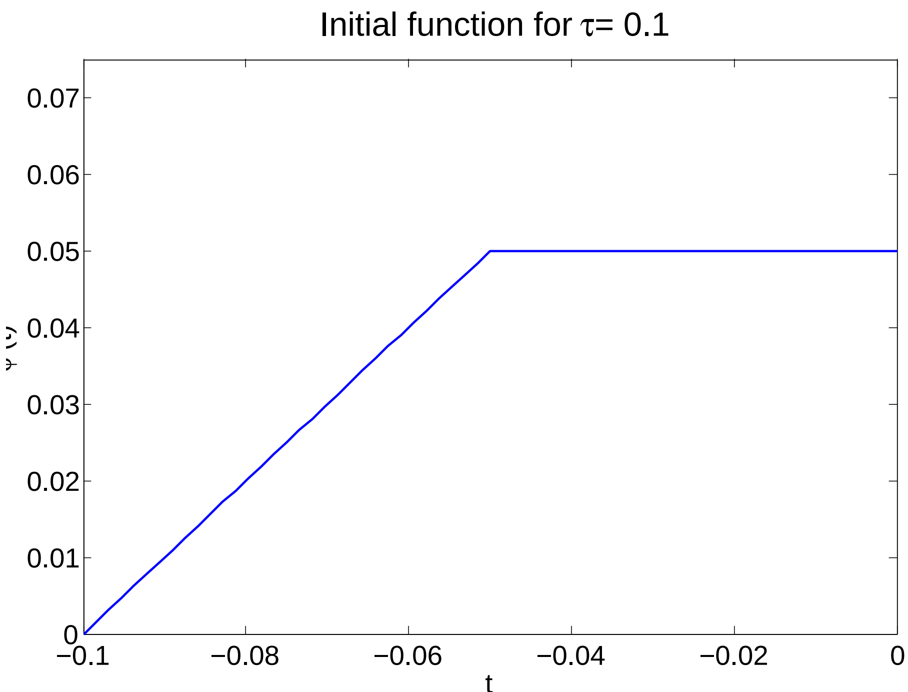


Figure 4.10: The initial condition $\phi(t) = y(t)$ for $t \in [-\tau, 0]$.

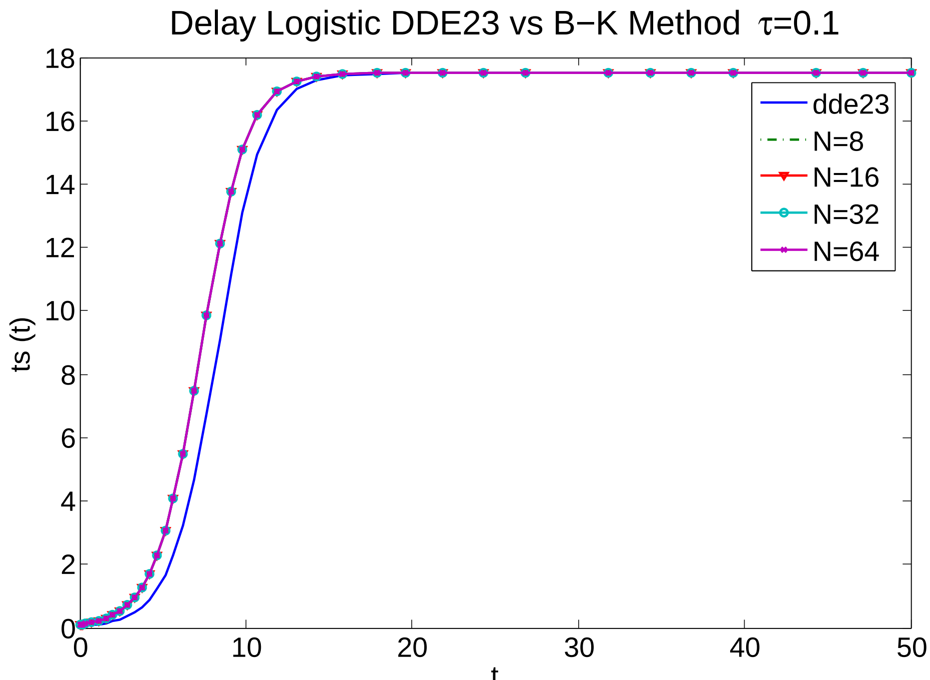


Figure 4.11: The numerical solution at $\tau = .1$.

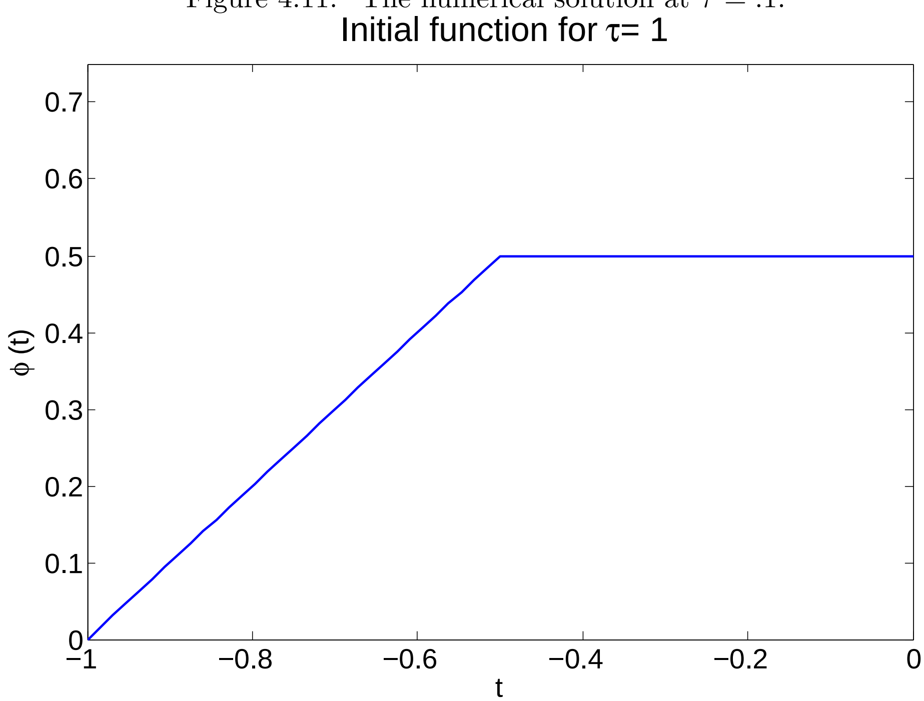


Figure 4.12: The initial condition $\phi(t) = y(t)$ for $t \in [-\tau, 0]$.

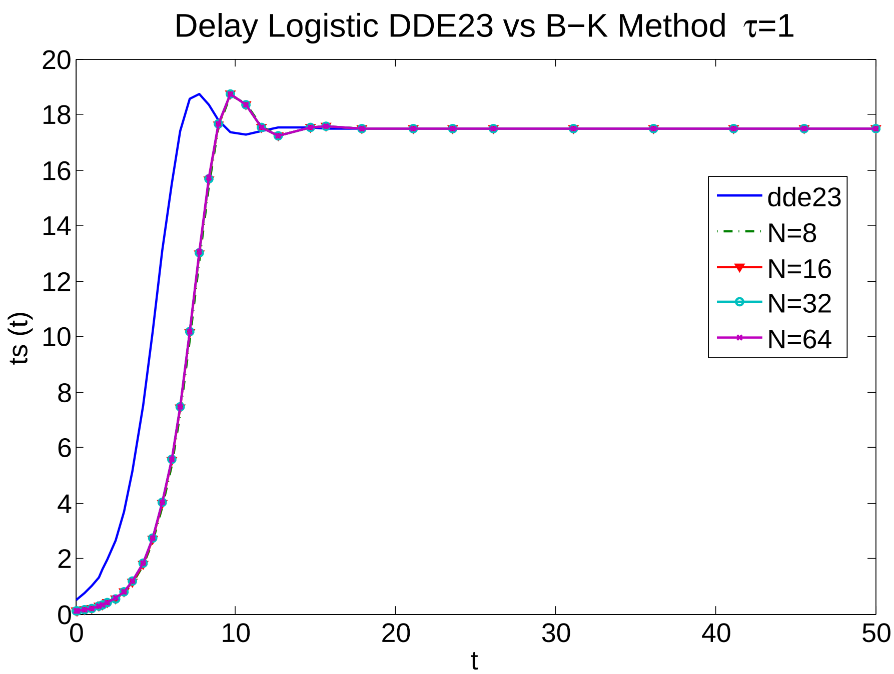


Figure 4.13: The numerical solution at $\tau = 1$.

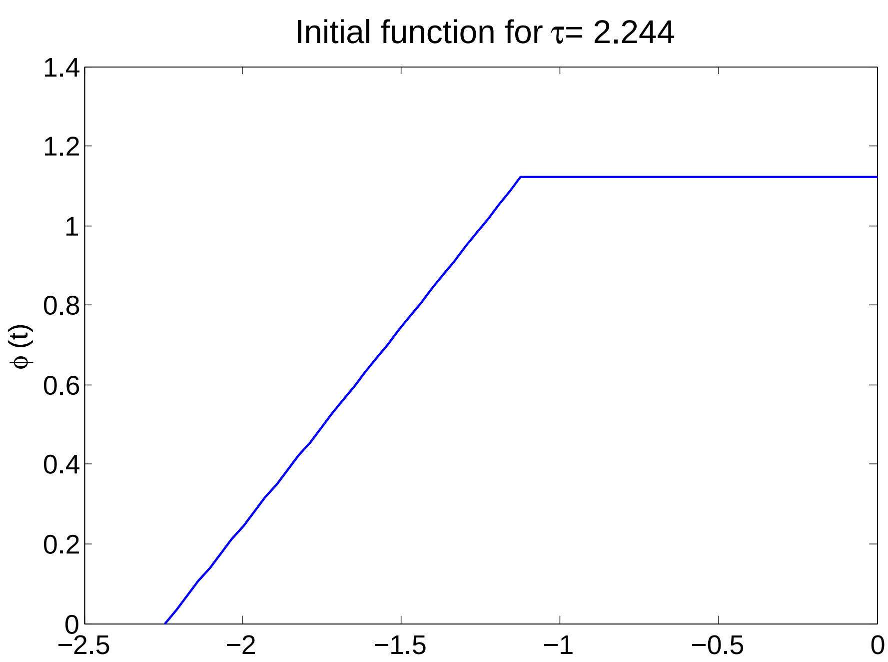


Figure 4.14: The initial condition $\phi(t) = y(t)$ for $t \in [-\tau, 0]$.

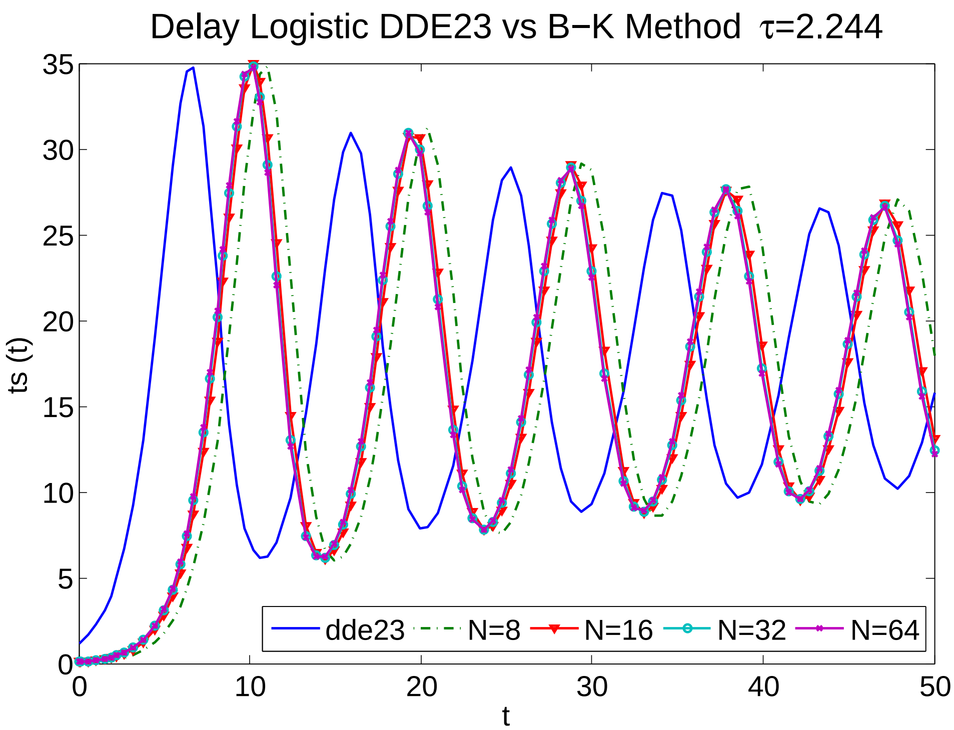


Figure 4.15: The numerical solution at $\tau = \frac{\pi}{2r}$.

Similar behavior occurs for both methods like with the previous non-constant initial condition. For the BK spline approximation $N = 16$ and $N = 64$ have similar solutions, while dde23 handles the initial function differently in comparison.

BK Method vs. Method of Steps for $\phi(t) = .1$

Using the Method of Steps given a constant initial condition, $\phi(t) = .1$ for $t \in [-\tau, 0]$, we obtain the following solution for the first two steps:

for $t \in [0, \tau]$

$$
\begin{aligned}
\frac{dx(t)}{dt} &= rx(t)\left[1 - \frac{\phi(t)}{K}\right] \\
\int \frac{dx}{x} &= \int r\left[1 - \frac{.1}{K}\right] dt \\
lnx &= r\left[1 - \frac{.1}{K}\right] t + C \\
x(t) &= .1e^{r(1-\frac{.1}{K})t},
\end{aligned}
$$

for $t \in [\tau, 2\tau]$, $B = r(1 - \frac{.1}{K})$

$$
\begin{aligned}
\frac{dx(t)}{dt} &= rx(t)\left[1 - \frac{.1e^{B}(t-\tau)}{K}\right] \\
\int \frac{dx}{x} &= \int r\left[1 - \frac{.1e^{B}(t-\tau)}{K}\right] dt \\
lnx &= rt - \frac{.1}{BK}e^{B(t-\tau)} + C \\
x(t) &= .1e^{rt-\frac{.1}{BK}e^{B(t-\tau)}}.
\end{aligned}
$$

We compare the Method of Steps to the BK method for $\tau = 2, 4, 8$. We choose larger values values for the delay so that the step intervals in the method of steps are larger.

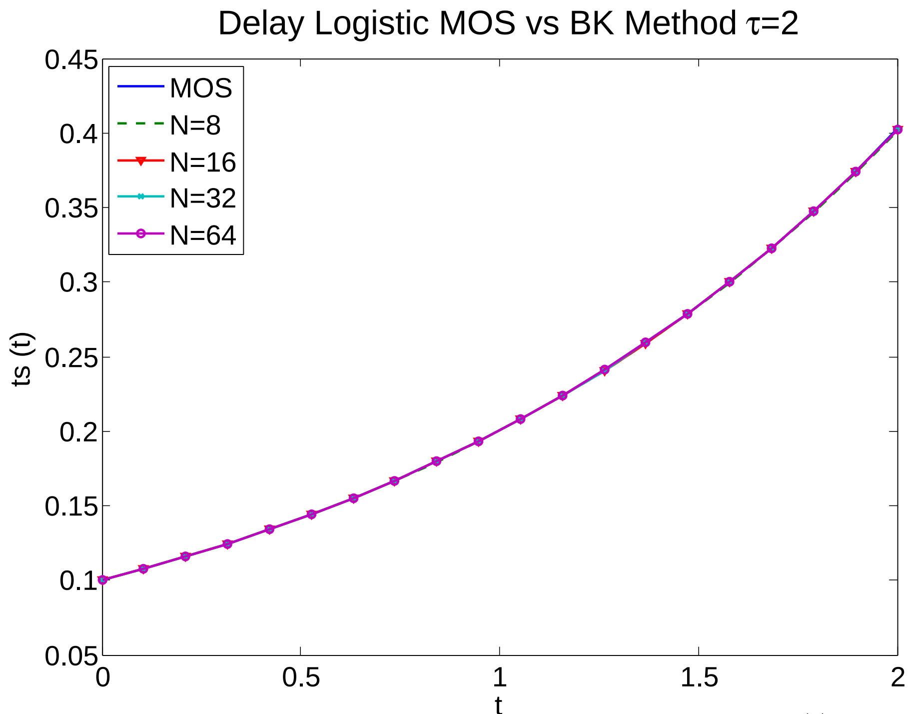


Figure 4.16: The numerical solution using the constant function $\phi(t) = .1$ for $\tau = 2$.

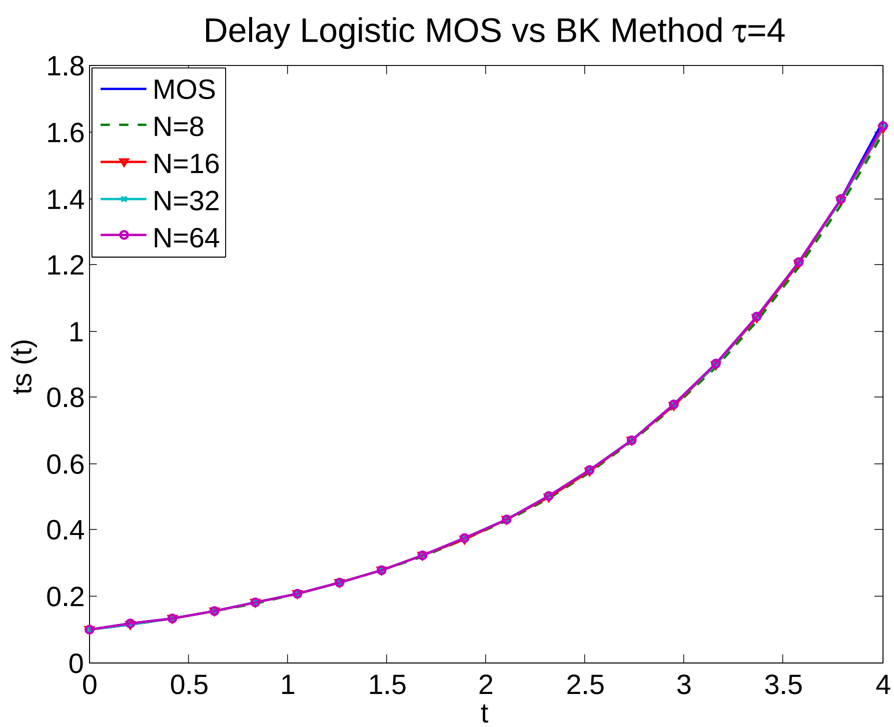


Figure 4.17: The numerical solution using the constant function $\phi(t) = .1$ for $\tau = 4$.

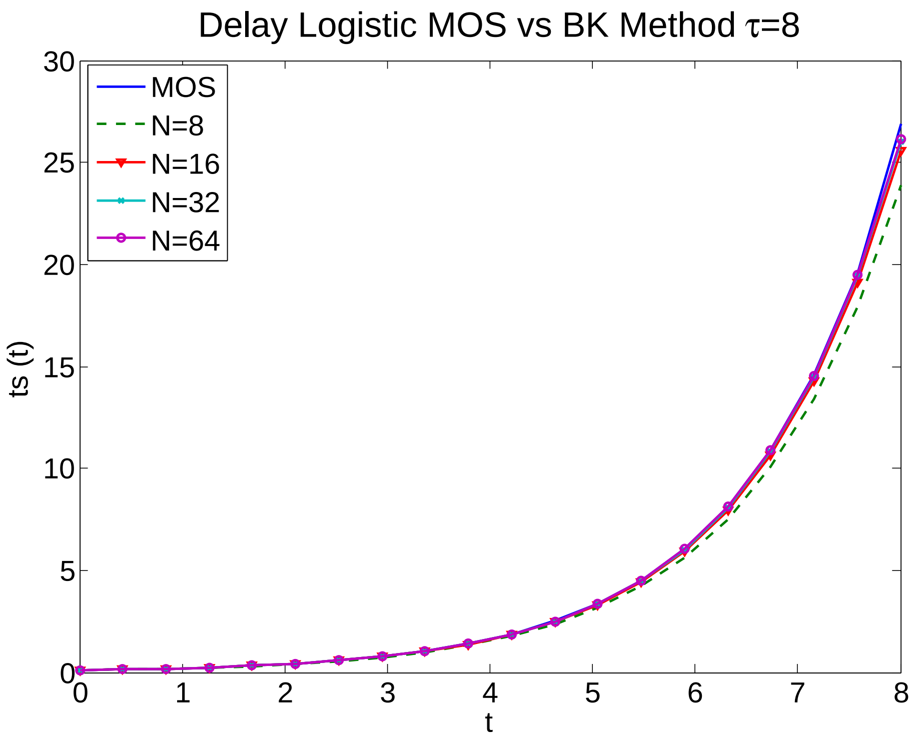


Figure 4.18: The numerical solution using the constant function $\phi(t) = .1$ for $\tau = 8$.

In figures $(4.16 - 4.18)$, we observe that as the value of τ increase, N must increase to obtain numerical convergence.

BK Method vs. Method of Steps for $\phi(t) = sin(\frac{2\pi t}{\tau})$

Using the Method of Steps given a initial function, $\phi(t) = sin(\frac{2\pi t}{\tau})$ for $t \in [-\tau, 0]$, we get the following solution for the first step:

$t \in [0, .2]$

$$
\begin{aligned}
\frac{dx(t)}{dt} &= rx(t)\left[1 - \frac{\phi(t)}{K}\right] \\
\int \frac{dx}{x} &= \int r\left[1 - \frac{\frac{2\pi(t-tau)}{\tau}}{K}\right] dt \\
lnx &= rt + \frac{r\tau}{2K}cos(t-\tau) + C \\
x(t) &= .1e^{-rt}e^{rt+\frac{r\tau}{2K}cos(\frac{2\pi(t-\tau)}{\tau})}.
\end{aligned}
$$

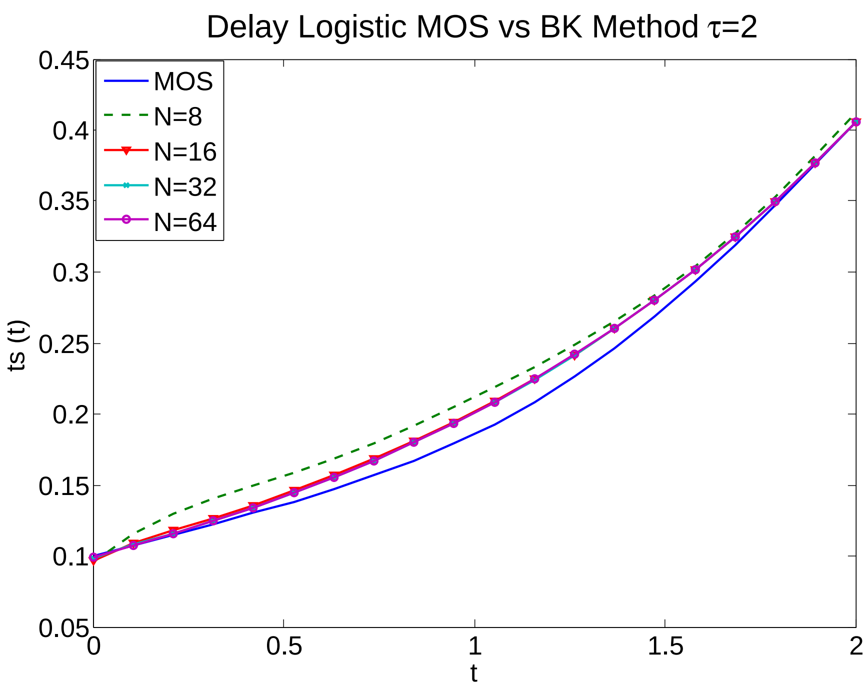


Figure 4.19: The numerical solution using the constant function $\phi(t) = sin(\frac{2\pi t}{\tau})$ for $t \in [-\tau, 0]$ for $\tau = 2$.

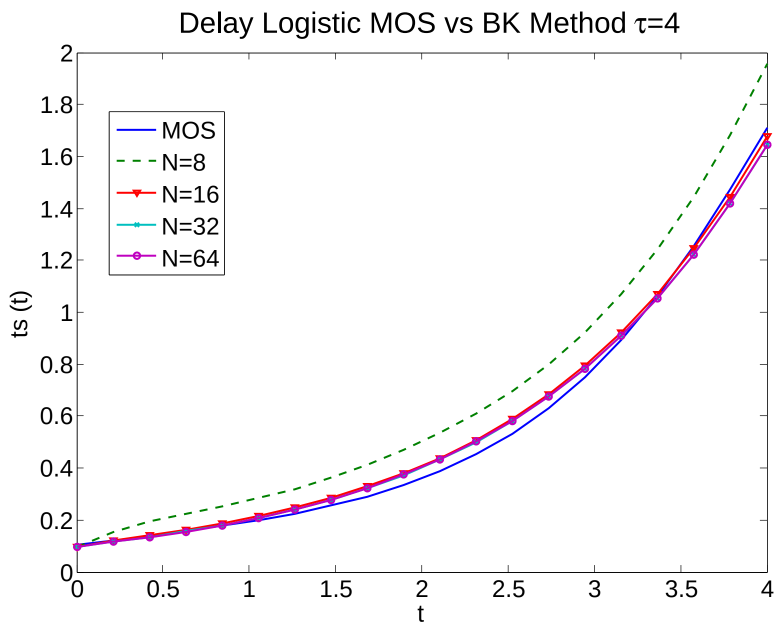


Figure 4.20: The numerical solution using the constant function $\phi(t) = sin(\frac{2\pi t}{\tau})$ for $t \in [-\tau, 0]$ for $\tau = 4$.

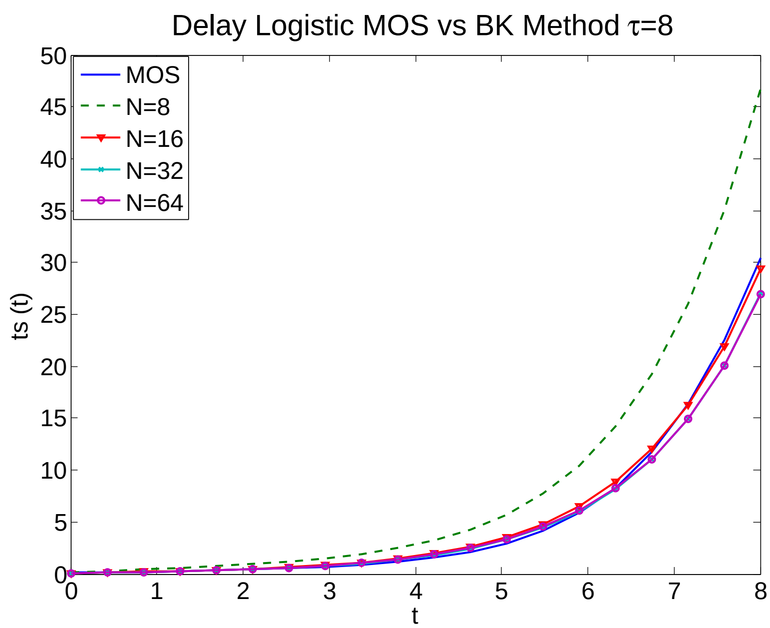


Figure 4.21: The numerical solution using the constant function $\phi(t) = sin(\frac{2\pi t}{\tau})$ for $t \in [-\tau, 0]$ for $\tau = 8$.

We observe again that as τ increases N must be increased to obtain numerical convergence.

BK Method vs. Method of Steps for $\phi(t) = y(t)$

The function $y(t)$ is

$$y(t) = \begin{cases} \tau + t, & -\tau \le t \le \frac{-\tau}{2} \\ \frac{\tau}{2}, & -\frac{-\tau}{2} < t < 0, \end{cases}$$

then the solution to the delay logistic equation using the Method of Steps on the interval $t \in [-\tau, \frac{-\tau}{2}]$ is

$$x(t) = .1e^{\frac{r\tau^2}{K}} e^{rt - \frac{r(t-\tau)^2}{K}}$$

and

$$x(t) = .1e^{r(1-\frac{\tau/2}{K})t}$$

on $t \in (\frac{-\tau}{2}, 0]$.

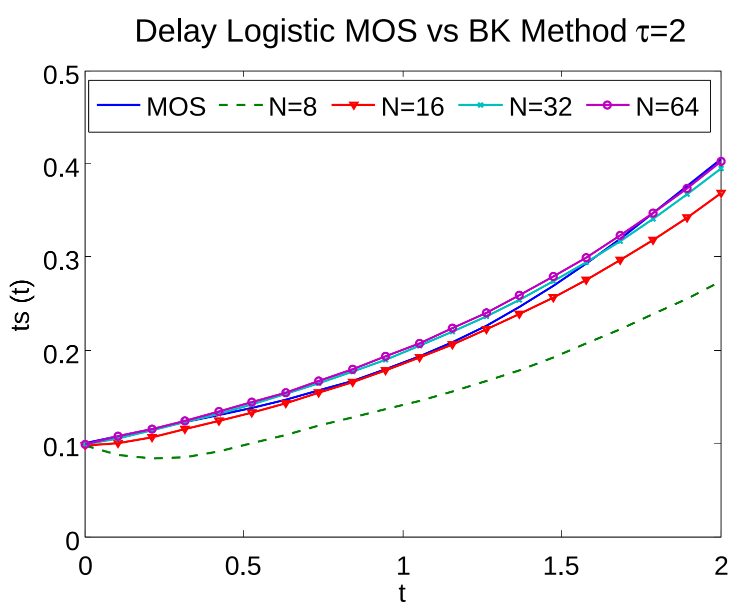


Figure 4.22: The numerical solution using the constant function $\phi(t) = y(t)$ for $\tau = 2$.

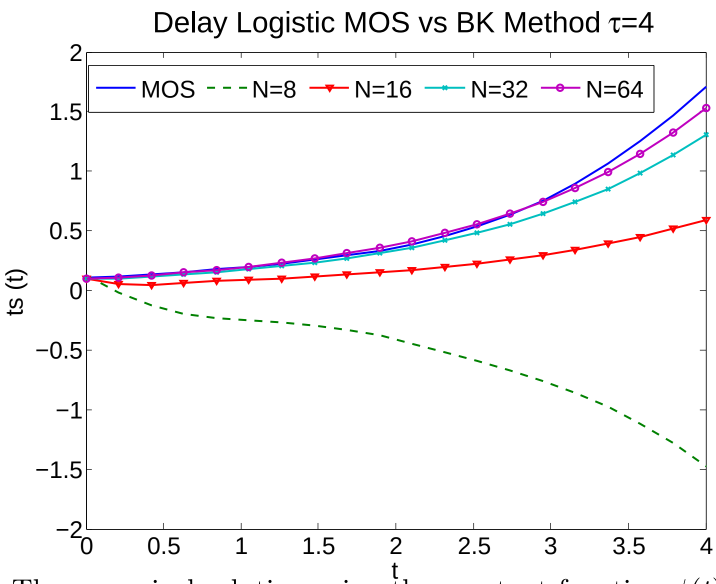


Figure 4.23: The numerical solution using the constant function $\phi(t) = y(t)$ for $\tau = 4$.

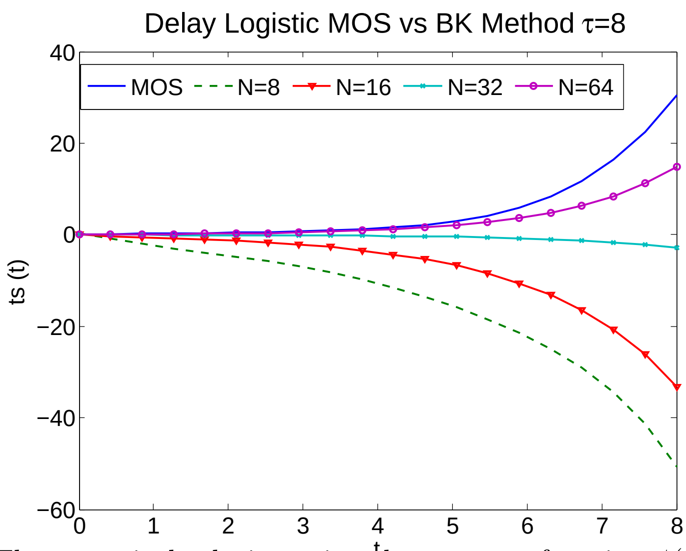


Figure 4.24: The numerical solution using the constant function $\phi(t) = y(t)$ for $\tau = 8$.

For this example we observe similar behavior for numerical convergence when compared to the previous two cases of the Method of Steps versus the BK spline approximation. However for the BK spline approximation to converge numerically when $y(t)$ defined previously is the initial function, N must be larger in comparison to the constant initial function. This happens because at every $\frac{\tau}{2}$ distance in the interval a jump discontinuity has been propagated through the solution and should occur again at τ where the interval ends, thus making it harder for the numerical method to converge.

Summary

For the Delay Logistic example, if the τ is in a range such that the solution does not oscillate and its solution reaches the steady state of the carrying capacity K, then a minimum of 8 nodes are needed to approximate the solution for a constant initial condition, and $N = 16$ is necessary for a non-constant initial condition such as $y(t)$. For oscillatory solutions that occur when $\tau = \frac{\pi}{2r}$, a minimum of 32 nodes are necessary to have numerical convergence. We also observe that MATLAB's dde23 cannot handle non-constant initial functions, especially initial functions with periodic behavior. However, the BK spline approximation can handle such types of initial functions; it just requires more nodes to obtain numerical convergence.

Chapter 5

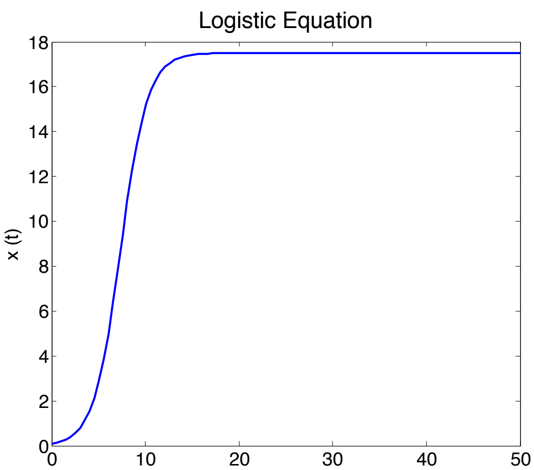
Numerical Examples

We numerically compute the TSFs and GSFs for three different examples: Hutchinson's equation, the delayed harmonic oscillator, and the behavior alcohol model. We assume we know the true parameter values θ_0, and the delay τ_0 for each example. We also assume we have constant variance, $\sigma_0^2 = .1$. We compute the TSFs and GSFs for all examples at the nominal values for the parameters and delay, θ_0 and τ_0, respectively.

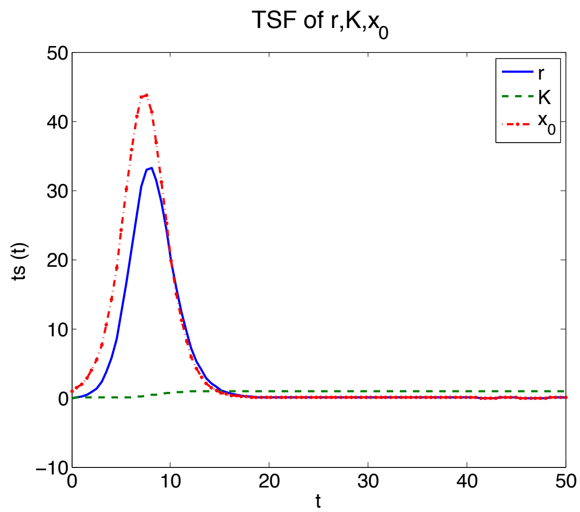
5.1 Delay Logistic Equation

For Hutchinson's equation we compute TSFs and GSFs for $\tau = .1, 1, \frac{\pi}{2r}$ at the nominal parameter values $\theta_0 = K_0 = 17.5, r_0 = .7$, and $x_{00} = .1$. We compute Hutchinson's equation for a constant initial function $\phi = .1$. We numerically compute the TSFs, GSFs using Banks-Kappel splines [12] and compare results with sensitivity functions computed in [8] for the non-delay logistic equation. The results obtained from the non-delay logistic equation as reported in [8] in Figures 5.1 (a)-(c).

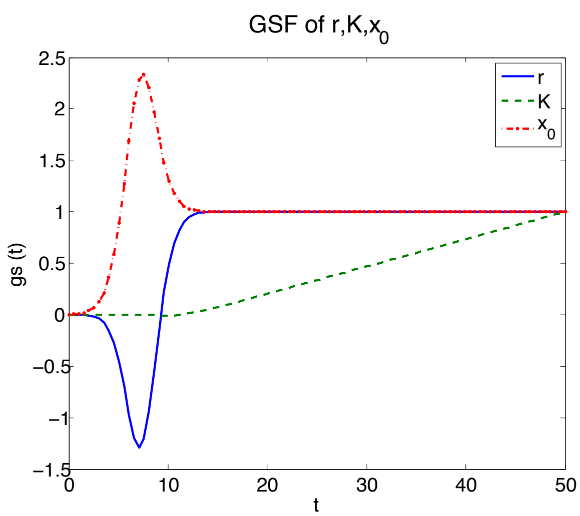
For the non-delay system it is observed in [8] that the model is very sensitive to the initial condition x_0 and r when the solution is growing the quickest as shown in Figure 5.1 (b). Once the solution reaches the carrying capacity the model becomes sensitive to K. The GSFs show regions of high information content where the function is increasing and decreasing. The parameters x_0 and r have correlated regions of high information, shown in Figure 5.1 (c), which again corresponds to the region where the solution is increasing to the carrying capacity (Figure 5.1 (a)). The GSF for K exhibits the "force-to-one" nature that is described in [8].



(a): Solution

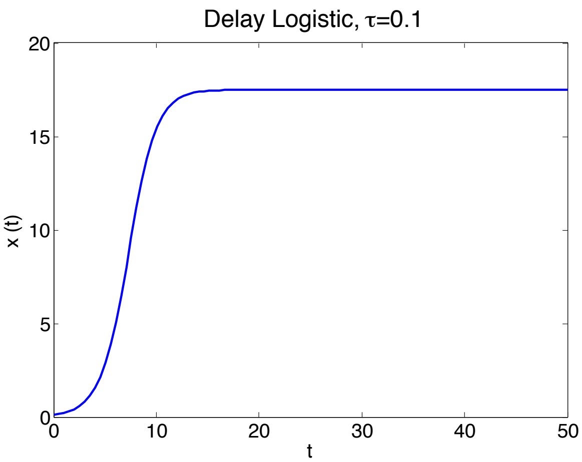


(b): TSF

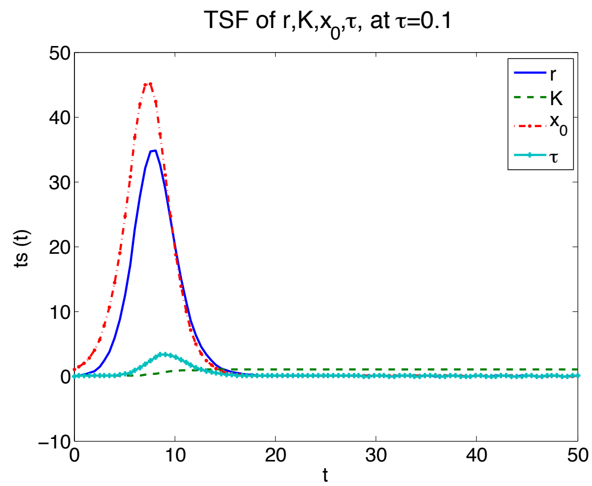


(c): GSF

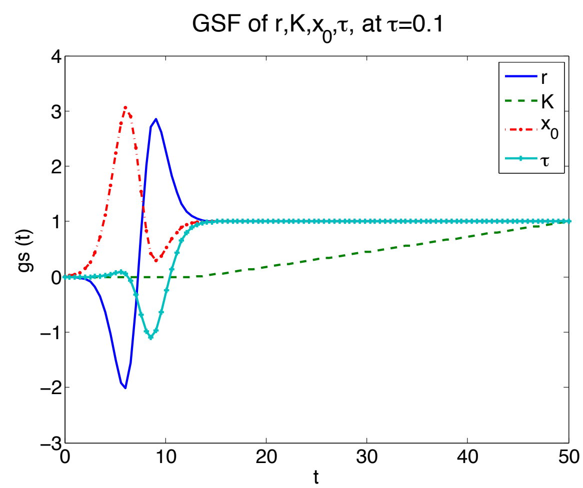
Figure 5.1: The numerical solution to the non-delay solution for $x_0 = .1, r = .7, K = 17.5$ in (a). The numerical solution to the TSFs with respect to r, K, x_0 for $x_0 = .1, r = .7, K = 17.5$ for the non-delay solution in (b). The numerical solution to the GSFs with respect to r, K, x_0 for $x_{00} = .1, r_0 = .7, K_0 = 17.5$ for the non-delay solution in (c).



(a): Solution

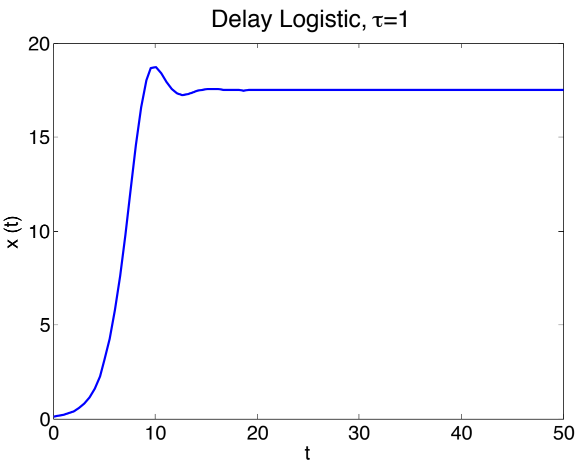


(b): TSF

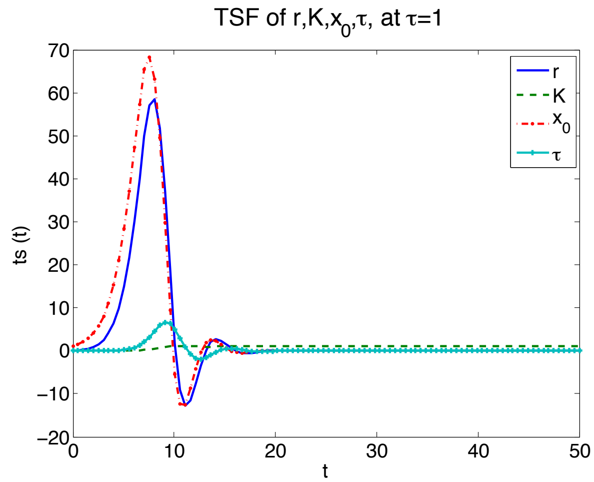


(c): GSF

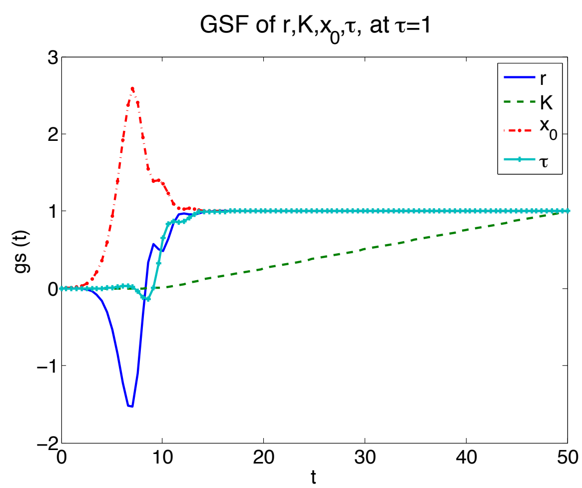
Figure 5.2: The numerical solution to the solution for $\phi = .1, x_0 = .1, r = .7, K = 17.5$ and $\tau = .1$ in (a). The numerical solution to the TSFs with respect to r, K, x_0, τ for $x_{00} = .1, r_0 = .7, K_0 = 17.5$ and $\tau_0 = .1$ in (b). The numerical solution to the GSFs with respect to r, K, x_0, τ for $x_{00} = .1, r_0 = .7, K_0 = 17.5$ and $\tau_0 = .1$ in (c).



(a): Solution

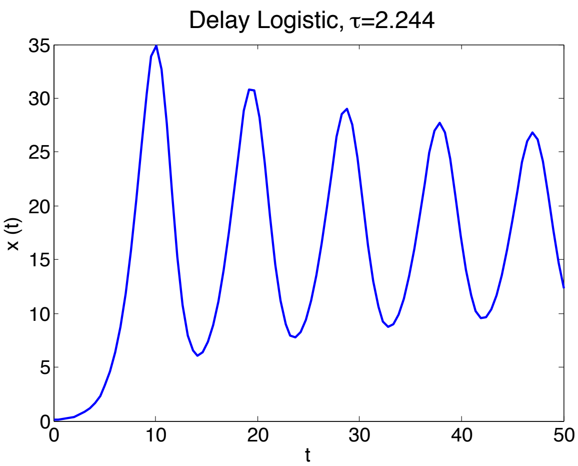


(b): TSF

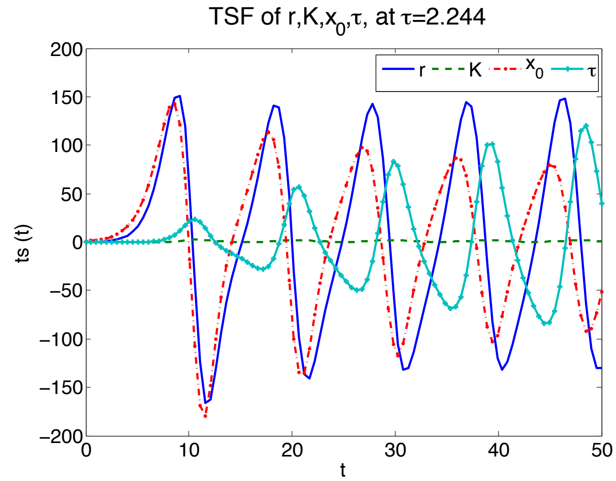


(c): GSF

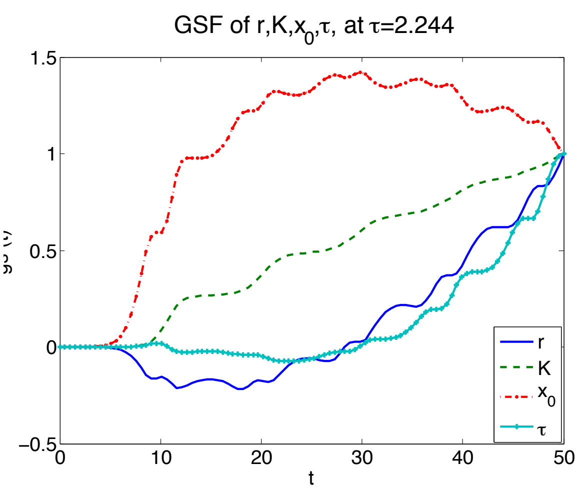
Figure 5.3: The numerical solution to the solution for $\phi = .1, x_0 = .1, r = .7, K = 17.5$ and $\tau = 1$ in (a). The numerical solution to the TSFs with respect to r, K, x_0, τ for $x_{00} = .1, r_0 = .7, K_0 = 17.5$ and $\tau_0 = 1$ in (b). The numerical solution to the GSFs with respect to r, K, x_0, τ for $x_{00} = .1, r_0 = .7, K_0 = 17.5$ and $\tau_0 = 1$ in (c).



(a): Solution



(b): TSF



(c): GSF

Figure 5.4: The numerical solution to the solution for $\phi = .1, x_0 = .1, r = .7, K = 17.5$ and $\tau = \frac{\pi}{2r}$ in (a). The numerical solution to the TSFs with respect to r, K, x_0, τ for $x_{00} = .1, r_0 = .7, K_0 = 17.5$ and $\tau_0 = \frac{\pi}{2r}$ in (b). The numerical solution to the GSFs with respect to r, K, x_0, τ for $x_{00} = .1, r_0 = .7, K_0 = 17.5$ and $\tau_0 = \frac{\pi}{2r}$ in (c).

When comparing Figures 5.1 (b) and 5.2 (b), there is similar behavior between the traditional sensitivity functions for r, K, and x_0. We observe that the model is most sensitive to parameters r and x_0 in the region where the solution is changing quickly. The model is sensitive to the parameter K once the solution approaches the carrying capacity. The TSF for τ displayed in Figure 5.2 (b) shows that the model is sensitive to the delay when the solution history is large enough to effect the solution at the current time. Although the behavior of the TSFs when $\tau = .1$ for the delay logistic system is similar to the behavior of the TSFs for the non-delay system, the same is not true when comparing the GSFs. We observe with the addition of the GSF with respect to the delay, the GSFs for r and x_0 are still correlated in the same way as the non-delay case, as shown in Figures 5.1 (c) and 5.2 (c); however the shape of those functions have changed. This is a reasonable outcome because with the addition of another parameter the FIM becomes larger. Also, the GSF for r is also inversely correlated with the GSF for τ. We also observe there are no dynamical changes in the GSF of τ from $t \in [0, 5]$ because the history of the solution $x(t-\tau)$ is not sufficiently large to affect the solution $x(t)$ in this region. In both cases the behavior for the GSF of K remains the same as it does not start to increase until the solution to the logistic/delay logistic equations nears the carrying capacity K. It is also observed that for the delay logistical model the dynamical regions for the TSFs and GSFs are the same.

When $\tau = 1$ we observe that the solution goes slightly past the carrying capacity but nears the carrying capacity in the same time frame as that for the non-delay and $\tau = .1$ solutions. The TSF dynamics for r and x_0 are similar to those for the non-delay system, the model is sensitive to these parameters where the solution is changing as shown in Figure 5.3 (b). The only difference in comparing the TSFs is the shape of the curves for x_0 and r after $t = 10$ which is reasonable since this is where changes in the solution occur. The TSF for K exhibits the same behavior as observed in the previous example. The GSFs for x_0 and r are again inversely correlated but exhibit comparable behavior to the non-delay system GSFs, as shown in Figure 5.3 (c). For the GSF of τ the region of steepest increase, as shown in Figure 5.3 (c), corresponds to region where the solution goes slightly past the carrying capacity, and then returns to the steady state solution. The dynamical regions for the sensitivity functions again are the same.

When $\tau = \frac{\pi}{2r}$ the solution oscillates around the carrying capacity K, Figure 5.4 (a), as expected from Hutchinson's previous results [32]. The model is always sensitive to r and x_0 when $\tau = \frac{\pi}{2r}$ because the solution oscillates, and each of those parameters effect how the solution changes, see Figure 5.4 (b). The sensitivity to K is still minimal in comparison to the other parameters, but the TSF for K still exhibits oscillatory behavior. The model becomes sensitive to the delay τ once the history of the solution is sufficiently large to affect the solution at the current time t. The model then remains sensitive to τ over the rest of the time interval, as shown in Figure 5.4 (b). The GSFs for all parameters are always changing making it difficult to decipher the meaning of the mainly monotone functions seen in Figure 5.4 (c). The GSFs also have multiple regions where steep increases and decreases occur, leading us to infer that the region of high information content for all parameters would be nearly the entire time interval.

From the three different solution examples for the delay logistic equation we learn that for $\tau < \tau_c$, where τ_c is some critical delay value, the TSF and GSF behavior is comparable to that of the non-delay model first reported in [8]. We also can determine clear regions of high information content, which can help with parameter estimation. However with oscillatory solutions it becomes more difficult to distinguish clear regions of high information from the GSFs and the TSFs which have oscillatory behavior. Overall, the delay is a catalyst in changing solution dynamics, thus being able to estimate the delay properly will aid proper analysis of any DDE model.

5.2 Harmonic Oscillators with Delayed Damping

One of earliest practical examples where it was discovered that small delays could have a profound influence on solution behavior were shown in the efforts modeling ship stabilization systems and nonlinear oscillations by Nicholas Minorsky [38, 39, 40]. As we earlier summarized in Chapter 1, Minorsky introduces and uses this idea of hysterodifferential equations to correctly describe delayed oscillations within a ship's control system. He also uses these type of models since the associated eigenvalues with the solutions can give a range of behavior, allowing the same model to both describe non-oscillatory and oscillatory behavior. For our next example we compute TSFs and GSFs for the harmonic oscillator with delayed damping and delayed restoring force.

The following model is the harmonic oscillator with retarded damping

$$\begin{aligned} \frac{d^2y(t)}{dt^2} + K\frac{dy(t-\tau)}{dt} + by(t) = g(t), \\ y(0) = 10, \frac{dy}{dt}(0) = 0. \end{aligned} \tag{5.1}$$

Let $x_1(t) = y(t)$, and $x_2(t) = \frac{dy(t)}{dt}$, then we can rewrite equation (5.1) to be

$$\frac{dx_1(t)}{dt} = x_2(t) \tag{5.2}$$

$$\frac{dx_2(t)}{dt} = g(t) - bx_1(t) - Kx_2(t-\tau). \tag{5.3}$$

5.2.1 Traditional Sensitivity

Let $s_1(t) = \frac{\partial x_1(t)}{\partial K}, s_2(t) = \frac{\partial x_1(t)}{\partial b}, s_3(t) = \frac{\partial x_1(t)}{\partial \tau}, s_4(t) = \frac{\partial x_2(t)}{\partial K}, s_5(t) = \frac{\partial x_2(t)}{\partial b}$, and $s_6(t) = \frac{\partial x_2(t)}{\partial \tau}$, then the following system of delay differential equations may be solved at (K_0, b_0, τ_0) to obtain the traditional sensitivity functions:

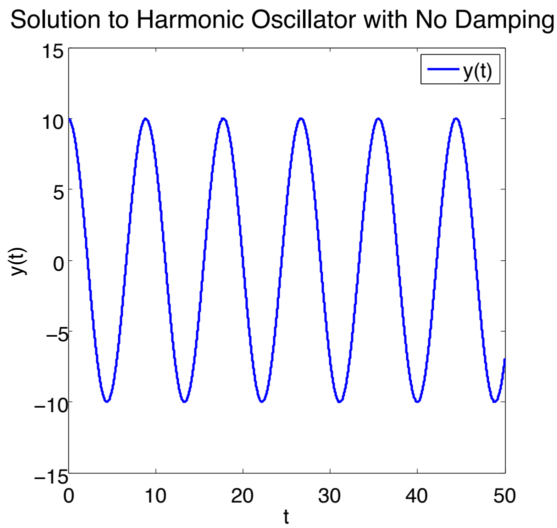
$$\begin{aligned}
\frac{ds_1(t)}{dt} &= s_4(t) \\
\frac{ds_2(t)}{dt} &= s_5(t) \\
\frac{ds_3(t)}{dt} &= s_6(t) \\
\frac{ds_4(t)}{dt} &= -bs_1(t) - Ks_4(t-\tau) - x_2(t-\tau) \\
\frac{ds_5(t)}{dt} &= -bs_2(t) - Ks_5(t-\tau) - x_1(t) \\
\frac{ds_6(t)}{dt} &= -bs_3(t) - Ks_6(t-\tau) + K\dot{x}_2(t-\tau).
\end{aligned}$$

We use Banks Kappel splines to numerically compute the TSFs and GSFs for the model at the parameter point $q_0 = (K_0, b_0, \tau_0)$ where the delay $\tau_0 = 1$ for $t \in [0, 50]$ or $t \in [0, 25]$. We choose a nominal value for the variance $\sigma^2 = .1$. To check our numerical solution we observe the oscillatory solution by setting $b = .5$, and $K = 0$ and $g(t) = 0$. As a result of $K = 0$ we observe an oscillatory solution with oscillatory sensitivity functions. We then further observe the behavior of the sensitivity functions for the following parameter sets, $\{K = .5,\ b = 2\}$ with $g(t) = 10$, and $\{K = .5, b = 2\}$ with $g(t) = g_1(t)$ where

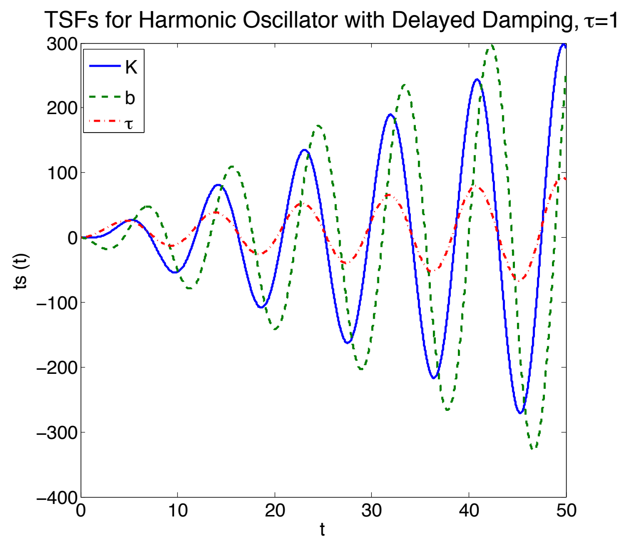
$$g_1(t) = \begin{cases} 10, & 0 \le t < 15 \\ 0, & \text{otherwise.} \end{cases}$$

We observe in Figure 5.5 (a) the undamped solution to the harmonic oscillator. As a result of the oscillatory solution, we observe oscillatory TSFs for the parameters K, b and τ as show in Figure 5.5 (b). Since the TSFs tell how sensitive the model is to the parameter we would infer that the model is always sensitive to all of the parameters for an oscillatory solution. In Figure 5.5 (c) there are multiple regions where the GSFs are vastly increasing and decreasing for K and b, and strictly increasing for τ, thus it would be wise to use all data collected over the entire time period to estimate the parameters for this type of model solution. We use this example to compare and contrast the behavior of both the TSFs and GSFs of the Harmonic Oscillator with delayed damping at different initial parameter points $(K_0, b_0,\ \tau_0)$.

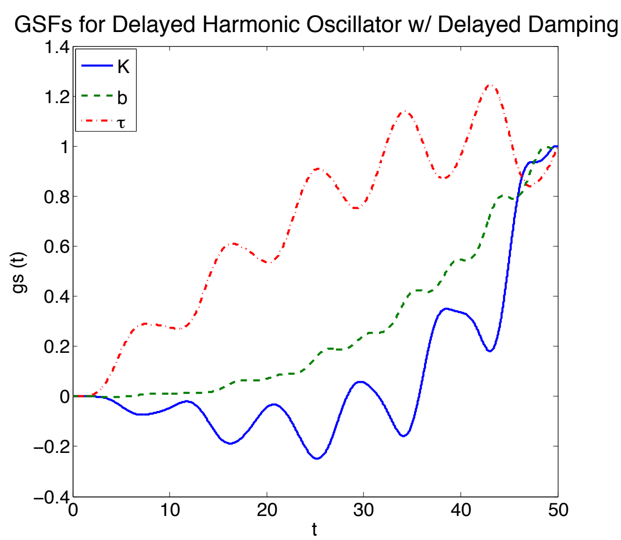
In Figure 5.6 (a) we change the values for K, b, and $g(t)$ such that after some time the



(a): Solution

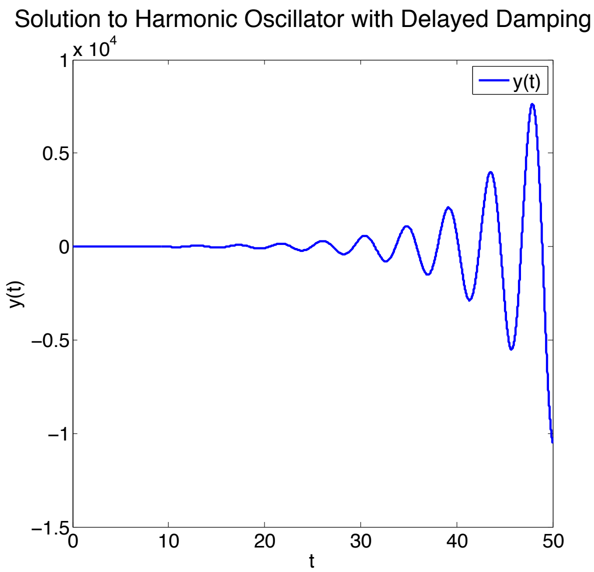


(b): TSF

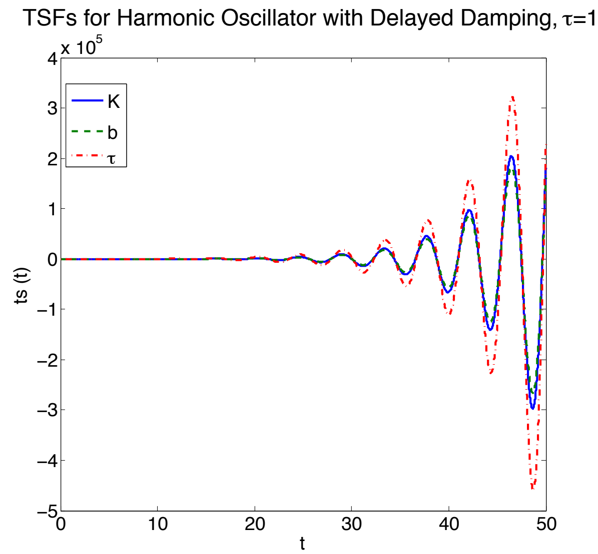


(c): GSF

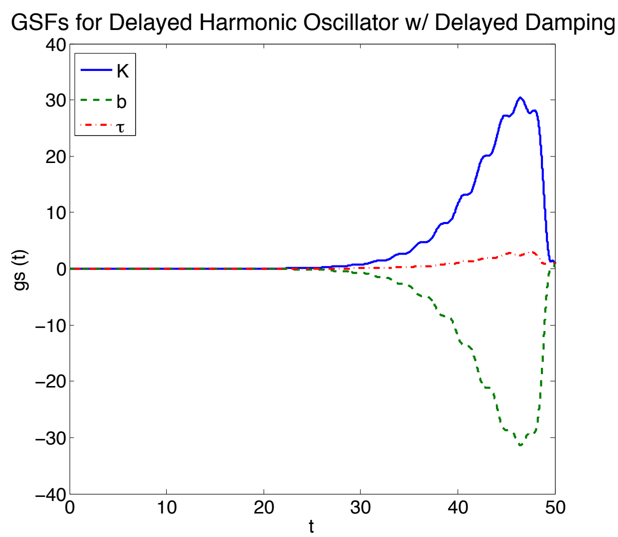
Figure 5.5: The numerical solution to the Harmonic Oscillator $K = 0, b = .5$, $\tau = 1$, and $g(t) = 0$ in (a). The numerical solution to the TSFs for the model with respect to K, b, τ at $K_0 = 0, b_0 = .5$, $\tau_0 = 1$, and $g(t) = 0$ in (b). The numerical solution to the GSFs for the model with respect to K, b, τ at $K_0 = 0, b_0 = .5$, $\tau_0 = 1$, and $g(t) = 0$ in (c).



(a): Solution

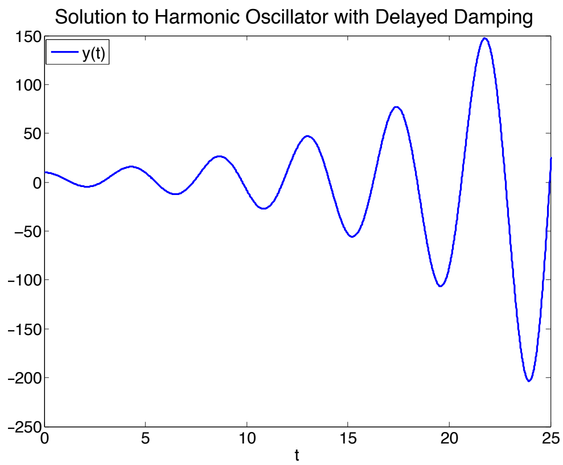


(b): TSF

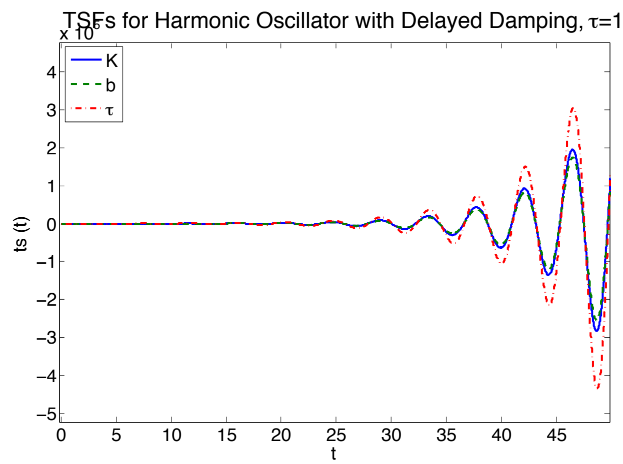


(c): GSF

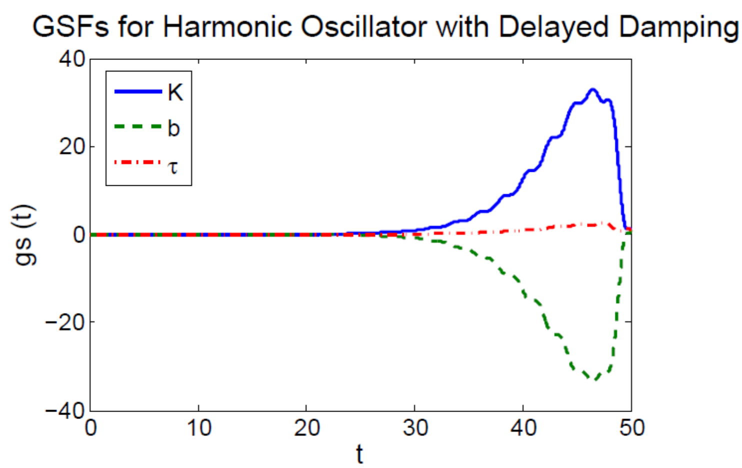
Figure 5.6: The numerical solution to the Harmonic Oscillator $K = .5, b = 2$, $\tau = 1$, and $g(t) = 10$ in (a). The numerical solution to the TSFs for the model with respect to K, b, τ at $K_0 = .5, b_0 = 2$, $\tau_0 = 1$, and $g(t) = 10$ in (b). The numerical solution to the GSFs for the model with respect to K, b, τ at $K_0 = .5, b_0 = 2$, $\tau_0 = 1$, and $g(t) = 10$ in (c).



(a): Solution



(b): TSF



(c): GSF

Figure 5.7: The numerical solution to the Harmonic Oscillator $K = .5, b = 2$, $\tau = 1$, and $g = g_1$ in (a). The numerical solution to the TSFs for the model with respect to K, b, τ at $K_0 = .5, b_0 = 2$, $\tau_0 = 1$, and $g = g_1$ in (b). The numerical solution to the GSFs for the model with respect to K, b, τ at $K_0 = .5, b_0 = 2$, $\tau_0 = 1$, and $g(t) = g_1$ in (c).

restoring force will overtake the delayed damping on the solution. As a result of the solution we observe the model is sensitive to the parameters where the solution begins to oscillate. Thus the TSFs oscillate on the same interval of the solution as shown in Figure 5.6 (b). The GSFs for K and b are inversely correlated and all the parameters have regions of high information content occurring where the solution oscillates. In comparison with the undamped solution, we observe that inclusion of the delayed damping reduces the sensitivity regions due to changes in the solution behavior.

When $g(t) = g_1(t)$, we observe that the solution oscillates over the entire time only changing in amplitude after $t \approx 15$. The TSFs are now oscillatory and imply the model is sensitive to all of the parameters. The GSFs have multiple regions of high information content after $t \approx 5$. Thus increasing data collection after $t = 5$ could possibly aid in the parameter estimation of K, b and τ. We observe that for oscillatory solutions the sensitivity functions are best used together to improve data collection and estimation processes for the model.

5.3 Harmonic Oscillator with Delayed Restoring Force

The following model is the delayed harmonic oscillator with retarded damping

$$\frac{d^2y(t)}{dt^2} + K\frac{dy(t)}{dt} + by(t-\tau) = g(t), \tag{5.4}$$
$$y(0) = 10, \frac{dy}{dt}(0) = 0.$$

Let $x_1(t) = y(t)$, and $x_2(t) = \frac{dy(t)}{dt}$, then we can rewrite equation (5.4) to be

$$\frac{dx_1(t)}{dt} = x_2(t) \tag{5.5}$$
$$\frac{dx_2(t)}{dt} = g(t) - bx_1(t-\tau) - Kx_2(t). \tag{5.6}$$

5.3.1 Traditional Sensitivity

Let $s_1(t) = \frac{\partial x_1(t)}{\partial k}, s_2(t) = \frac{\partial x_1(t)}{\partial \lambda}, s_3(t) = \frac{\partial x_1(t)}{\partial \tau}, s_4(t) = \frac{\partial x_2(t)}{\partial k}, s_5(t) = \frac{\partial x_2(t)}{\partial \lambda}$, and $s_6(t) = \frac{\partial x_2(t)}{\partial \tau}$, then the following system of delay differential equations may be solved at (K_0, b_0, τ_0) to obtain the traditional sensitivity functions:

$$\begin{aligned}
\frac{ds_1(t)}{dt} &= s_4(t) \\
\frac{ds_2(t)}{dt} &= s_5(t) \\
\frac{ds_3(t)}{dt} &= s_6(t) \\
\frac{ds_4(t)}{dt} &= -bs_1(t-\tau) - Ks_4(t) - x_2(t) \\
\frac{ds_5(t)}{dt} &= -bs_2(t-\tau) - Ks_5(t) - x_1(t-\tau) \\
\frac{ds_6(t)}{dt} &= -bs_3(t-\tau) - Ks_6(t) + \dot{x}_1(t-\tau)
\end{aligned}$$

We use Banks-Kappel splines we numerically compute the TSFs and GSFs for the model at some initial parameter point $q_0 = (K_0, b_0, \tau_0)$ with the delay $\tau_0 = 1$ for $t \in [0, 50]$. We choose nominal values for the variance $\sigma^2 = .1$. To check our numerical solution we use $K = 1$, $b = 0$, and $g(t) = 0$, this means the solution will have no restoring force and the solution to $x_2(t) = 0$, making the solution to $x_1(t)$ constant. In Figure 5.8 we observe a constant solution. We then use the parameter values $k = 5$ and $b = .5$ with $g(t) = 10$ and $g(t) = g_1(t)$ to observe different behavior in the sensitivity functions.

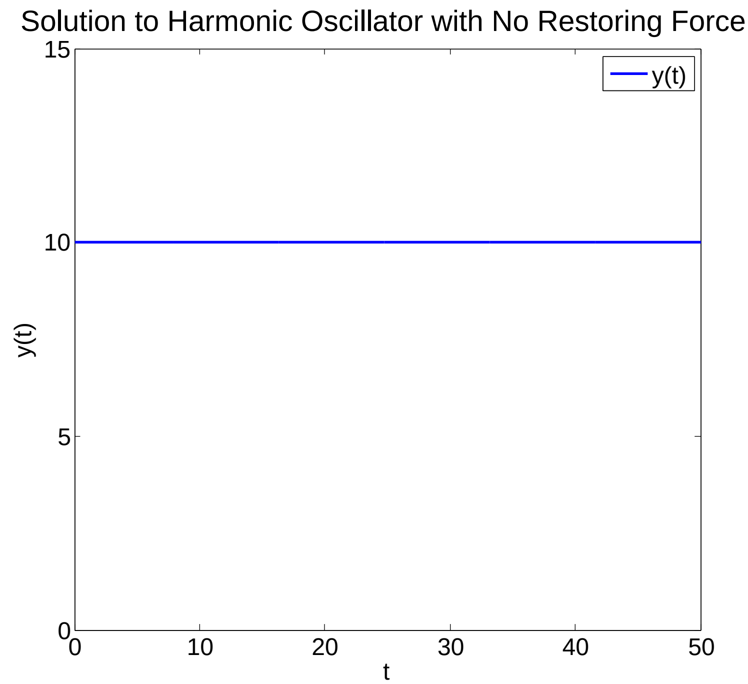
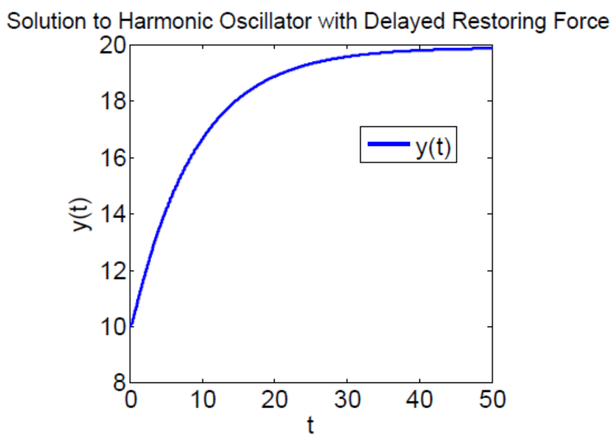


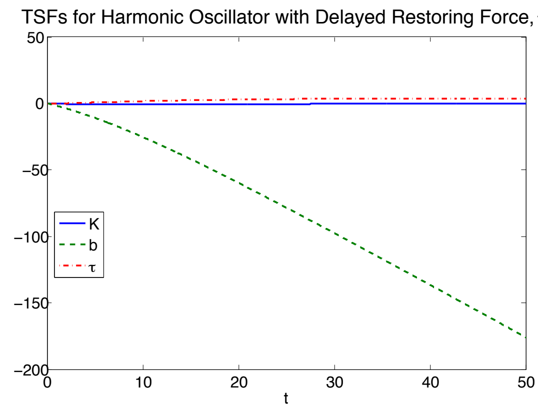
Figure 5.8: The numerical solution to the Harmonic Oscillator with no Restoring Force for $K = 1$, $b = 0$, and $g(t) = 0$.

When $K = 5, b = .5$ and $g(t) = 10$ we observe a increasing non-oscillatory solution. The corresponding TSFs show that the model is most sensitive to the restoring force parameter b, shown in Figure 5.9 (b). The model is also sensitive to the other parameters as the TSFs for these parameters are also changing over the entire interval. The GSFs in Figure 5.9 (c) show that data collected over the entire interval will help in estimating b and τ. While for the parameter K collecting more data from $t = 0$ to $t \approx 20$ should improve the parameter estimate.

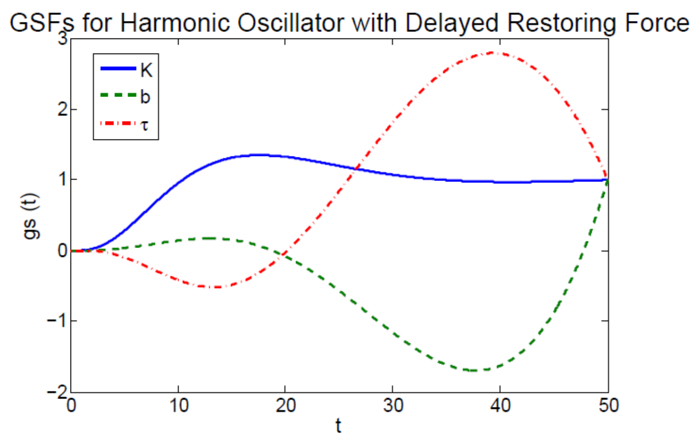
If $g(t)$ is changed to from 10 to $g_1(t)$ the solution changes as shown in Figure 5.10 (a). As a result, changes in the shape of the TSFs occurs, however the overall sensitivity behavior is still the same, i.e. the model is most sensitive to b, shown in Figure 5.10 (b). In Figure 5.10 (c) the GSFs show that the estimates for all parameters are sensitive to the observations for most of the entire time interval. The GSFs for b, τ are strictly increasing wile K decreases slightly at the end of the time interval. This outcome suggests that data collected over the entire time interval will aide in the estimation of each parameter.



(a): Solution

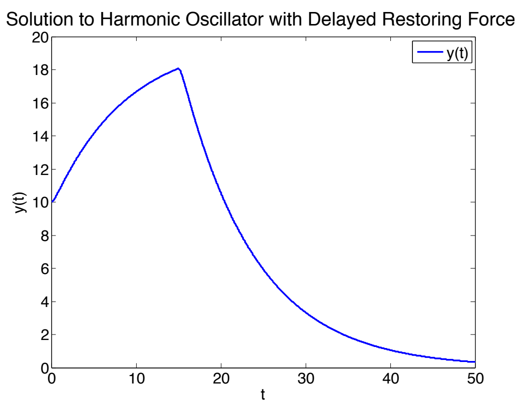


(b): TSF

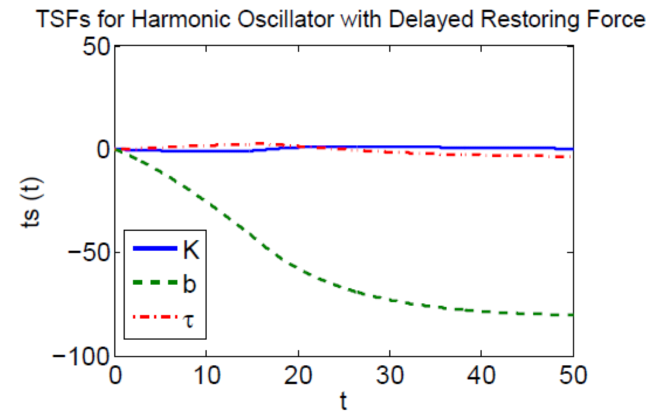


(c): GSF

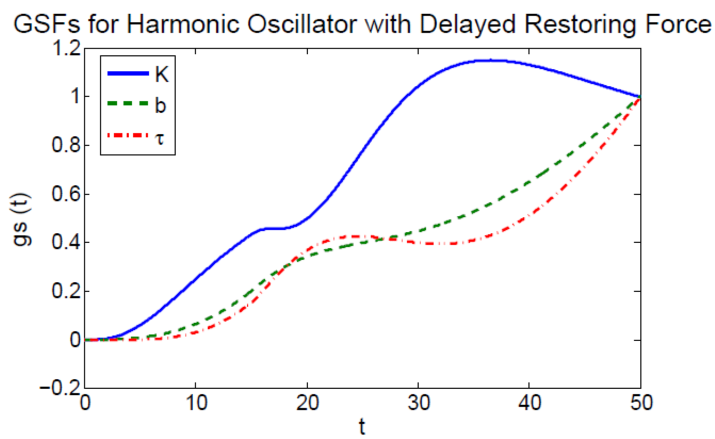
Figure 5.9: The numerical solution to the Harmonic Oscillator with Delayed Restoring Force when $K = 5, b = .5,\ \tau = 1$, and $g(t) = 10$ in (a). The numerical solution to the TSFs for the model with respect to K, b, τ for $K_0 = 5, b_0 = .5,\ \tau_0 = 1$, and $g(t) = 10$ in (b). The numerical solution to the GSFs for the model with respect to K, b, τ for $K_0 = 5, b_0 = .5,\ \tau_0 = 1$, and $g(t) = 10$ in (c).



(a): Solution



(b): TSF



(c): GSF

Figure 5.10: The numerical solution to the Harmonic Oscillator with Delayed Restoring Force when $K = 5, b = .5$, $\tau = 1$, and $g = g_1$ in (a). The numerical solution to the TSFs for the model with respect to K, b, τ for $K_0 = 5, b_0 = .5$, $\tau_0 = 1$, and $g = g_1$ in (b). The numerical solution to the GSFs for the model with respect to K, b, τ for $K_0 = 5, b_0 = .5$, $\tau_0 = 1$, and $g = g1$ in (c).

5.4 A Behavior Change Model

For the last example, we will compute TSFs and GSFs with respect to the delay for the behavior change model for problem drinking in individuals. We also determine proper regions to collect data to obtain the best estimates when performing the inverse problem.

Model

The following model is a simplified version of a model of behavior change in problem drinkers given in [18]

$$\frac{d}{dt}A(t) = -a_{12}\chi_{G>G^*}G(t) + a_{13}D(t), \tag{5.7}$$

$$\frac{d}{dt}G(t) = a_{21}[A(t-\tau) - (1 + c_1\chi_{W(t)})A^*], \tag{5.8}$$

$$\frac{d}{dt}D(t) = -a_{32}\chi_{G>G^*}(G(t) - G^*) + ch(t). \tag{5.9}$$

Here $A(t)$ represents an individual's drinking rate, $G(t)$ is his or her guilt as a result of drinking the previous day, and $D(t)$ is an individual's desire to drink.

Using the alcohol behavior model, the TSFs with respect to the delay τ are

$$\begin{aligned}
\frac{d}{dt}\frac{\partial A(t)}{\partial \tau} &= g_A(A(t),G(t),D(t))\frac{\partial A(t)}{\partial \tau} + g_G(A(t),G(t),D(t))\frac{\partial G(t)}{\partial \tau} \\
&\quad + g_D(A(t),G(t),D(t))\frac{\partial D(t)}{\partial \tau} + g_\tau(A(t),G(t),D(t)), &(5.10)\\
\frac{d}{dt}\frac{\partial G(t)}{\partial \tau} &= k_A(A(t-\tau),G(t),D(t),\tau)\frac{\partial A(t-\tau)}{\partial \tau} + k_G(A(t-\tau),G(t),D(t),\tau)\frac{\partial G(t)}{\partial \tau} \\
&\quad + k_D(A(t-\tau),G(t),D(t),\tau)\frac{\partial D(t)}{\partial \tau} + k_\tau(A(t-\tau),G(t),D(t)), &(5.11)\\
\frac{d}{dt}\frac{\partial D(t)}{\partial \tau} &= m_A(A(t),G(t),D(t))\frac{\partial A(t)}{\partial \tau} + m_G(A(t),G(t),D(t))\frac{\partial G(t)}{\partial \tau} \\
&\quad + m_D(A(t),G(t),D(t))\frac{\partial D(t)}{\partial \tau} + m_\tau(A(t),G(t),D(t)), &(5.12)
\end{aligned}$$

where $g(A(t),G(t),D(t)) = -a_{12}\chi_{G>G^*}G(t) + a_{13}D(t)$, $k(A(t-\tau),G(t),D(t)) = a_{21}[A(t-\tau) - (1+c_1\chi_{W(t)})A^*]$, and $m(A(t),G(t),D(t)) = -a_{32}\chi_{G>G^*}(G(t)-G^*) + c\,h(t)$.

We define $s_1(t) = \frac{\partial A(t)}{\partial \tau}$, $s_2(t) = \frac{\partial G(t)}{\partial \tau}$, $s_3(t) = \frac{\partial D(t)}{\partial \tau}$, then equations (5.10)-(5.12) become

$$\frac{ds_1(t)}{dt} = -a_{12}\chi_{G>G^*}s_2(t) + a_{13}s_3(t), \tag{5.13}$$

$$\frac{ds_2(t)}{dt} = a_{21}[s_1(t-\tau) - \dot{A}(t-\tau)], \tag{5.14}$$

$$\frac{ds_3(t)}{dt} = -a_{32}\chi_{G>G^*}s_2(t). \tag{5.15}$$

We use the BK spline approximation to numerically compute the solution for the model, and the TSFs and GSFs for the delay $\tau = 1$ as it relates to $A(t)$, $G(t)$, and $D(t)$ for $t \in [0, 28]$. We choose a nominal value for the variance $\sigma^2 = .1$, and the parameter values are $a_{12} = .1, a_{13} = 2, a_{21} = .2, c_1 = .01, A^* = 2, a_{32} = 2, G^* = .5, c_2 = 3$.

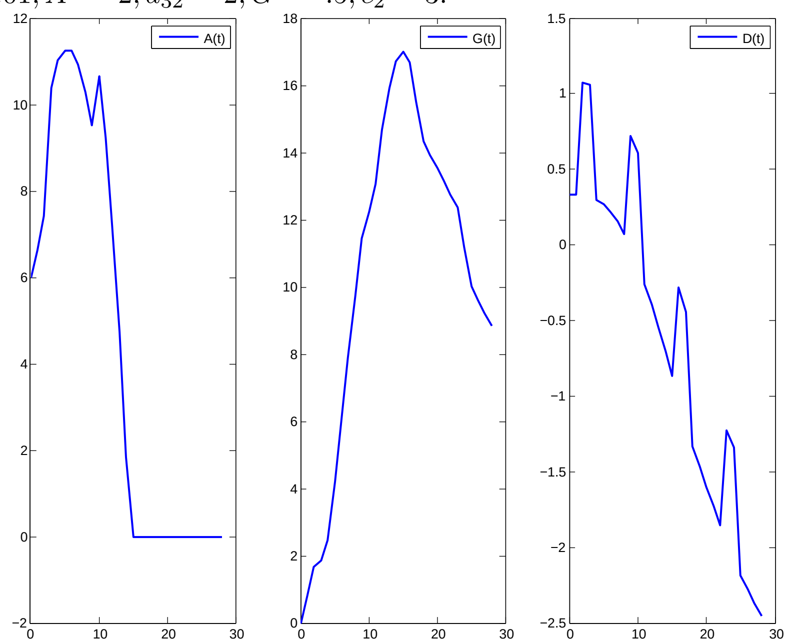
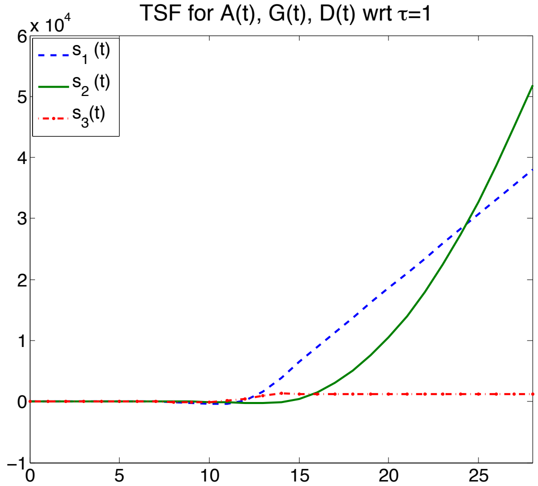
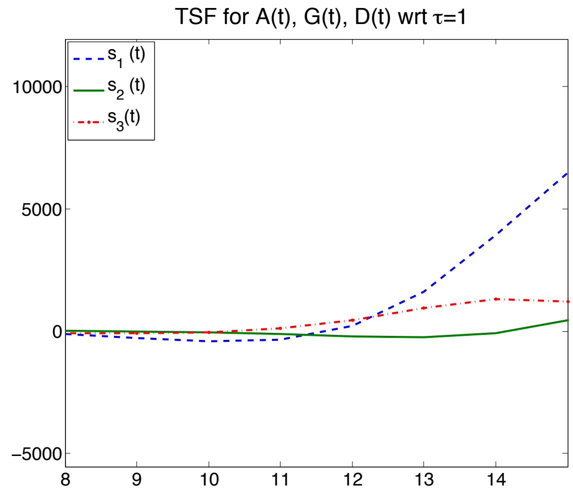


Figure 5.11: The numerical solution to the alcohol behavior model when $\tau = 1$.

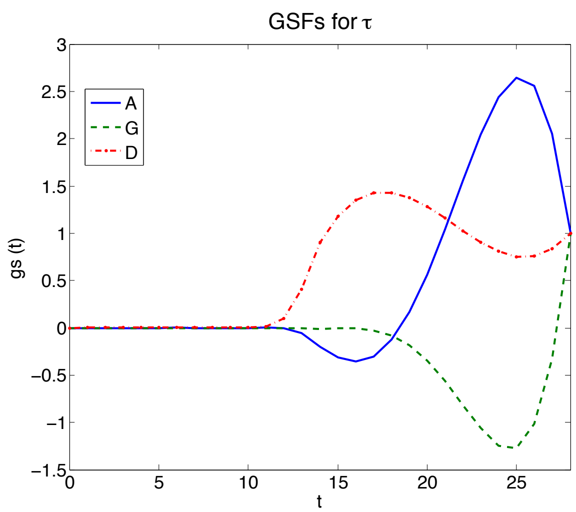
In Figure 5.12 (a) we observe that an individual's desire to drink $D(t)$ is initially not as sensitive to the delay as the TSF remains close to zero from $t \in [0, 10]$, thus it would be difficult to estimate τ using data on this interval for equation (5.9). Although when we observe that same TSF in Figure 5.12 (b) after $t = 10$ the TSF is not so close to zero, and hits a steady state after $t = 14$, this means that an estimate for τ maybe reasonably obtained for data collected after $t = 10$. An individual's guilt, $G(t)$, as represented by equation (5.8), and his or her drinking rate $A(t)$ are also sensitive to the delay τ. As a person's guilt is dependent upon his or her previous drinking rate, and a change in person's drinking rate with respect to the delay is



(a): TSF-1



(b): TSF-2



(c): GSF

Figure 5.12: The numerical solution to the TSFs for $\frac{dA(t)}{d\tau}$, $\frac{dG(t)}{d\tau}$, and $\frac{dD(t)}{d\tau}$ when $\tau = 1$ in (a). The numerical solution to the TSFs for $\frac{dA(t)}{d\tau}$, $\frac{dG(t)}{d\tau}$, and $\frac{dD(t)}{d\tau}$ for $t \in [8, 14]$ when $\tau = 1$ in (b). The numerical solution to the GSF with respect to τ for $A(t), G(t), D(t)$ when $\tau = 1$ in (c).

more dependent upon the change in his or her guilt with respect to the delay τ. When looking at the graph of the TSF $\frac{dG(t)}{d\tau}$ in Figures 5.12 (a) and (b), we can deduce that if there is data collected after $t = 10$ for the model an estimate for τ can be obtained. After $t = 10$ the TSF $\frac{dG(t)}{d\tau}$ is no longer close to zero and is changing for the remaining time interval. When observing the graph of the TSF $\frac{dA(t)}{d\tau}$, we observe that the TSF begins to move away from zero after $t = 8$, thus if we use the equation for an individual's drinking rate to estimate τ, the data should be collected a little earlier when compared to data collected for an individual's guilt.

We compute the GSF with respect to τ using the TSFs, and observe that $D(t)$'s region of high information content occurs when $t \in [10, 15]$. As described by Thomaseth and Cobelli [44], high information content occurs in a GSF where the function has a steep increase. A steep decrease usually occurs due to some correlation and has valuable information in the GSF also as shown in [8]. Also in Figure 5.12 (c) we observe that between $t = 15$ and $t = 24$ the model is sensitive to the data for $A(t)$, and then $G(t)$'s region of importance is the remaining time interval after $t = 24$. The GSF for $A(t)$ has a region of steep increase that is inversely correlated to the region of steep decrease of the GSF for $G(t)$. If we combine the information gained from both the TSFs and GSFs we may obtain time intervals where data can be collected to improve the estimate for the delay in each of model equations. For example, when doing parameter estimation for the delay τ using equation (5.9) its best to have data collected on the region $t \in [10, 14]$.

Sensitivity functions are useful in identifying regions of high information content which can aide in parameter estimation. These functions also give insight to how the model is related to the parameter (TSFs), and how the model is related to the data (GSFs). The sensitivity functions with respect to the delay are insured to have dynamical changes if the formulation of the equation is dependent on the previous history of the sensitivity function, as shown in our example $\frac{dG(t)}{d\tau}$. TSFs and GSFs when used together improve experimental design giving the experimentalist the ability to determine where to collect optimal information.

Chapter 6

Improving the Inverse Problem

6.1 A Generic Inverse Problem

Suppose we have a given data set $\hat{\eta} = \{\eta_i\}$ that are corresponding observations to the solution $x(t_i;\theta)$ of the following equation

$$\begin{aligned} \frac{dx(t)}{dt} &= G(x(t), x(t-\tau), \theta), \quad t > 0 \\ x(\xi) &= \begin{cases} \Phi(\xi), & -\tau \le \xi < 0 \\ x_0, & \xi = 0. \end{cases} \end{aligned} \tag{6.1}$$

The model dynamics in (6.1) are parameter, θ, dependent where G in our case represents a functional differential equation. To estimate θ, we use a least squares approach where the problem is to minimize the following cost functional

$$J(\theta, \hat{\eta}) = \sum_i |x(t_i;\theta) - \eta_j|^2 \tag{6.2}$$

over $\theta \in \Theta$, where $\Theta \subset \mathbb{R}^p$ [9].

We compute the inverse problem for Hutchinson's Equation to estimate the delay parameter. We improve this computation by observing the TSFs and GSFs computed in the previous chapter for this example. Based on the observations of these sensitivity functions at various values for the delay, we determine regions within the solution data that will aid in the estimation of the delay τ. We also obtain an appropriate final time T to end data collection for this particular parameter. Use of these functions allows us to simulate data sets that will give the most optimal estimate for the delay.

6.2 Example: Hutchinson's Equation

Constant variance data sets are simulated using $1\%, 5\%$, and 10% noise, for the delay parameter when $\tau = .1, 1, \frac{\pi}{2r}$. We create noisy data sets with 30 evenly spaced data points for $t \in [0, 25]$. This type of data sets will be identified as "even" data ets. We also create noisy data sets with 30 points where the mesh is not evenly space, and most of the data points are concentrated in the region of high information content. An example of this type of data set is as follows: for $t \in [0, 5]$ there will be 5 data points, for $t \in (5, 15]$ there will be 20 data points, and for $t \in (15, 25]$ there will 5 data points, where the region of high information content is $t \in [5, 15]$. We identify these data sets as "enhanced" data sets. We also create "enhanced+" noisy data sets where an additional 5 data points are added to the region of high information content. When the delay is $\tau = .1, 1$, the region of high information content is $t \in [5, 15]$. While for $\tau = \frac{\pi}{2r}$, the region of high information content corresponds with $t \in [6, T]$, where T is the final time. The parameter values r, K, and x_0 are $.7, 17.5$, and $.1$ respectively. We perform the inverse problem to estimate only the delay parameter, τ, with an initial guess of $\tau^* = .25$ when the true value is $\tau_0 = .1$, and $\tau^* = .5$ when the true value is $\tau = 1$ or $\frac{\pi}{2r}$. We also compute the associated 95% confidence interval.

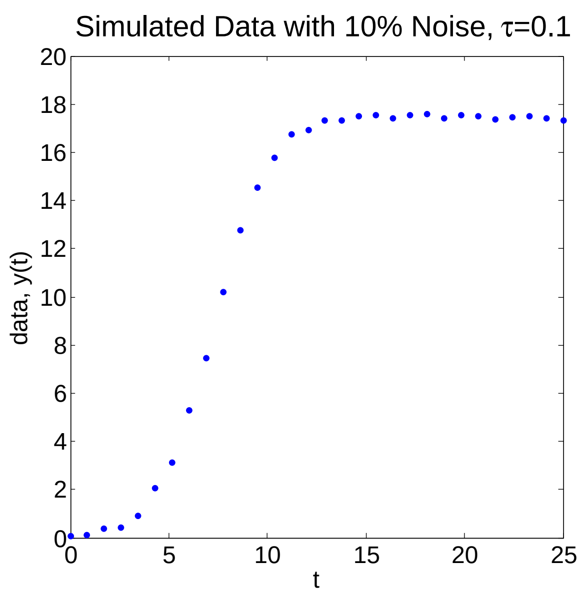


Figure 6.1: An example of a simulated data set with 10% noise and evenly spaced data points when $\tau = .1$.

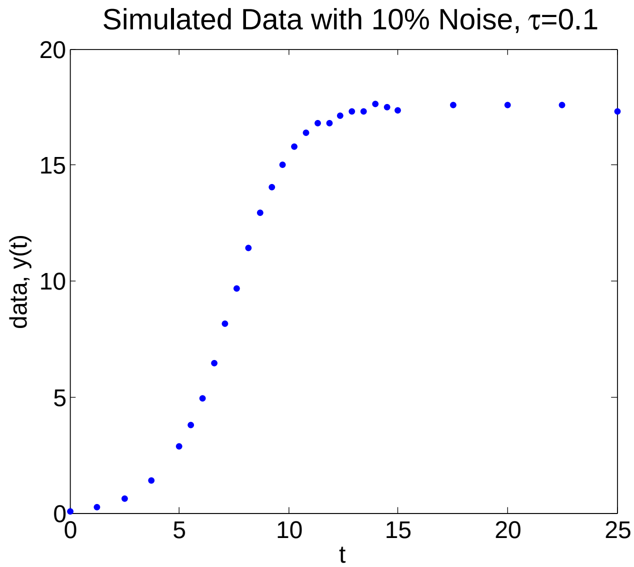


Figure 6.2: An example of a simulated data set with 10% noise and data points concentrated in the region of high information content for the delay parameter τ when $\tau = .1$.

Table 6.1: Parameter estimates for the delay from the 3 types of simulated data sets (even or enhanced or enhanced+) at noise levels nl of 1%, 5%, and 10%, Standard Error (SE) estimates, and 95% Confidence Intervals (CI), when $\tau = .1$.

Data Type	nl	$\hat{\tau}$	SE	CI
even	.01	.1023	.0071	(.0884, .1163)
even	.05	.1000	.0334	(.0345, .1655)
even	.10	.1129	.0861	(-.0558, .2816)
enhanced	.01	.1008	.0078	(.0855, .1161)
enhanced	.05	.1062	.0380	(.0316, .1807)
enhanced	.10	.1103	.0743	(-.0353, .2560)
enhanced+	.10	.1070	.0683	(-.0268, .2409)

In Table (6.1), we observe comparable estimates of the delay from the simulated data sets with lower amount amounts of noise (i.e., $nl = 1\%, 5\%$). The estimate does not appear optimal in the presence of 10% noise for the evenly space simulated noisy data set. We then simulate data for the same parameter values using the enhanced mesh by concentrating the data in the region of high information content as described earlier. We observe a slight improvement in the estimate of the delay. When the delay is estimated from the simulated enhanced+ data set and we observe an improvement in the estimate of the delay parameter from $\hat{\tau} = .1103$ to $\hat{\tau} = .1070$.

Table 6.2: Parameter estimates for the delay from the types of simulated data sets (even or enhanced or enhanced+) at noise levels, nl, of $1\%, 5\%$, and 10%, Standard Error (SE) estimates, and 95% Confidence Intervals (CI), when $\tau = 1$.

Data Type	nl	$\hat{\tau}$	SE	CI
even	.01	1.0019	.0027	(.9966,1.0071)
even	.05	1.0013	.0104	(.9808, 1.0217)
even	.10	1.0097	.0195	(.9715, 1.0479)
enhanced	.01	1.0003	.0011	(.9981, 1.0025)
enhanced	.05	1.0003	.0082	(.9841, 1.0164)
enhanced	.10	1.0006	.0167	(.9679, 1.0332)
enhanced+	.10	1.0010	.0174	(.9669, 1.0351)

Table 6.3: Parameter estimates for the delay from the types of simulated data sets (even or enhanced or enhanced+) at noise levels, nl, of $1\%, 5\%$, and 10%, Standard Error (SE) estimates, and 95% Confidence Intervals (CI), when $\tau = \frac{pi}{2r} \approx 2.244$.

Data Type	nl	$\hat{\tau}$	SE	CI
even	.01	2.2439	.0001	(2.2437, 2.2442)
even	.05	2.2438	.0008	(2.2421, 2.2454)
even	.10	2.2435	.0015	(2.2406, 2.2464)
enhanced	.01	2.2440	.0002	(2.2436, 2.2445)
enhanced	.05	2.2432	.0007	(2.2418, 2.2445)
enhanced	.10	2.2446	.0016	(2.2415, 2.2478)
enhanced+	.10	2.2445	.0013	(2.2419, 2.2471)

In Table (6.2), we observe that the enhanced data sets improve the estimate of the delay for all levels of noise. In Table (6.3), we observe that the enhanced data sets improve the estimate of the delay when $nl = 5\%, 10\%$.

6.3 Summary

Based on the results we see an overall improvement in the estimate for the delay when using enhanced data based on the region of high information content. We also observe better estimates when we increase the amount of data within the region of high information content and use enhanced+ data sets to estimate the delay. As a result, it is advantageous to use sensitivity functions to help obtain optimal estimates for parameters specifically in our case the delay.

Chapter 7

Final Remarks

7.1 Research Conclusions

From this research we learn the importance of sensitivity functions for parameter estimation of delay differential equation (dde) models. We formulate traditional sensitivity equations (TSFs) of the dde form and prove well-poseness for different classes of ddes. The TSFs for the various models (delay-logistic, harmonic oscillator, and alcohol behavior model) show that each model is always sensitive to the delay parameter τ for both oscillatory and non-oscillatory solutions. From here the TSFs, along with the variance, and Fisher Information Matrix, are used to compute the generalized sensitivity functions (GSFs). The GSFs give us insight to the regions within the data that can improve the estimate of the parameters within the model. We have learned that for different types of solution behaviors this region of importance changes. For example, oscillatory solutions tend to have larger regions of importance and would encourage us to suggest that data collection be increased over the entire time interval to improve the estimate of the model parameters. Non-oscillatory solutions have more distinct regions of high information content within the data and we can reasonably suggest specific regions in that data that may improve the parameter estimates for the models.

To show the ability of the sensitvity functions to improve the parameter estimate of the delay, we simulate three types of noisy data sets (even, enhanced, and enhanced+). Enhanced data sets has most of its points located in the region of high information content that corresponds to the delay, (this region is first presented in Chapter 5). We estimate the delay parameter, τ, and determine that focusing the data in the region of high information content corresponding to the delay improves the estimate in the presence of higher amounts of noise. Overall we learn that sensitivity function behavior is solution dependent and is useful in improving parameter estimates for delay differential equation models.

7.2 Future Work

In the future it would be advantageous to use this idea of generalized sensitvity and apply it to real data sets to possibly improve the experimental process. These functions may be used to change the observation process and/or observation times of the experiment. For other future work it would reasonable to explore the correlation of GSF functions for the model parameters and determine if this correlation affects the region of high information content for the model parameters. Futher complete sensitvity analysis (traditional and generalized) should be performed on different types of dde models, such as those with functional delays, in order to further understand the information that may be gained from generalized sensitivity.

REFERENCES

[1] J. Arino, L. Wang and G. Wolkowicz, An alternative formulation for a delayed logistic equation, *J. Theo. Bio.*, **241** (2006), 109–118.

[2] C. Baker and F. Rihan, Sensitivity analysis of parameters in modelling with delay-differential equations, *MCCM Tec. Rep.*, **349** (1999), Manchester, ISSN 1360-1725.

[3] H.T. Banks, *Modeling and Control in the Biomedical Sciences*, Lecture Notes in Biomath., Vol. 6, Springer-Verlag, Berlin, Heidelberg, New York, 1975.

[4] H.T. Banks, Delay systems in biological models: approximation techniques, *Nonlinear Systems and Applications* (V. Lakshmikantham, ed.), Academic Press, New York, 1977, 21–38.

[5] H.T. Banks. Identification of nonlinear delay systems using spline methods. In *Nonlinear Phenomena in Mathematical Sciences,* V. Lakshmikanitham, Ed. Academic Press, Inc., New York, NY, 1982, pp. 47-45.

[6] H.T. Banks, and D. M. Bortz, A parameter sensitivity methodology in the context of HIV delay equation models, *Journal of Mathematical Biology*, **50(6)** (2005), 607-625.

[7] H.T. Banks, J.A. Burns and E.M. Cliff, Parameter estimation and identification for systems with delays, *SIAM J. Control and Optimization*, **19** (1981), 791–828.

[8] H.T. Banks, S. Dediu, S. Ernstberger and F. Kappel, Generalized sensitivities and optimal experimental design, CRSC-TR08-12 (Revised), November, 2009; *J. Inverse and Ill-posed Problems*, **18** (2010), 25–83.

[9] H.T. Banks, M. Davidian, J.R. Samuels, Jr. and Karyn L. Sutton, An inverse problem statistical methodology summary, CRSC-TR08-01, January, 2008; Chapter 11 in *Mathematical and Statistical Estimation Approaches in Epidemiology*, (edited by Gerardo Chowell, et al.), Springer, Berlin Heidelberg New York, 2009, pp. 249–302.

[10] H.T. Banks, J.E. Banks and S.L. Joyner, Estimation in time-delay modeling of insecticide-induced mortality, CRSC-TR08-15, October, 2008; *J. Inverse and Ill-posed Problems*, **17** (2009), 101–125.

[11] H.T. Banks, S. Dediu, H. K. Nguyen, Sensitivity of dynamical systems to parameters in a convex susbset of a topological vector space, CRSC-TR06-25, November, 2006;*Mathematical biosciences and Engineering*, **4** (2007), 403–430.

[12] H.T. Banks and F. Kappel, Spline approximations for functional differential equations, *J. Differential Equations*, **34** (1979), 496–522.

[13] H.T. Banks and P.K. Daniel Lamm, Estimation of delays and other parameters in nonlinear functional differential equations, *SIAM J. Control and Optimization*, **21** (1983), 895–915.

[14] H.T. Banks and J.M. Mahaffy, Global asymptotic stability of certain models for protein synthesis and repression, *Quart. Applied Math.*, **36** (1978), 209–221.

[15] H.T. Banks and J.M. Mahaffy, Stability of cyclic gene models for systems involving repression, *J. Theoretical Biology*, **74** (1978), 323–334.

[16] H.T. Banks and H. Nguyen, Sensitivity of dynamical systems to Banach space parameters, CRSC-TR05-13, February, 2005; *J. Math. Analysis and Applications*, **323** (2006), 146–161.

[17] H.T. Banks and I.G. Rosen, Spline approximations for linear nonautonomous delay systems, ICASE Rep. No. 81-33, NASA Langley Res. Center, Oct., 1981; *J. Math. Anal. Appl.,* **96** (1983), 226–268.

[18] H.T. Banks, K. Rehm and K. Sutton, Inverse problems for nonlinear delay systems, CRSC-TR10-17, N.C. State University, November, 2010; *Methods and Applications of Analysis,* **17** (2010), 331-356.

[19] D. Brewer, The differentiability with respect to a parameter of the solution of a linear abstract Cauchy problem, *SIAM J. Math. Anal.*, **13** (1982), 607–620.

[20] D. Brewer, J.A. Burns, and E. M. Cliff, Parameter Identification for An Abstract Cauchy Problem by Quazilinearization, ICASE Rep. No. 89-75, NASA Langley Res. Center, Oct., 1989.

[21] M.D. Buhmann and A. Iserles, On the dynamics of a discretized neutral equation, *IMA J. of Numerical Analysis*, **12** (1992), 339–363.

[22] J.A. Burns, E.M. Cliff and S.E. Doughty, Sensitivity analysis and parameter estimation for a model of Chlamydia Trachomatis infection, *J. Inverse Ill-Posed Problems*, **15** (2007), 19–32.

[23] S.N. Busenberg and K.L. Cooke, eds., *Differential Equations and Applications in Ecology, Epidemics, and Population Problems*, Academic Press, New York, 1981.

[24] V. Capasso, E. Grosso and S.L. Paveri-Fontana, eds., *Mathematics in Biology and Medicine*, Lecture Notes in Biomath., Vol. 57, Springer-Verlag, Berlin, Heidelberg, New York, 1985.

[25] J. Caperon, Time lag in population growth response of *Isochrysis Galbana* to a variable nitrate environment, *Ecology*, **50** (1969), 188–192.

[26] S. Choi and N. Koo, Oscillation theory for delay and neutral differential equations, *Tends in Mathematics*, **2** (1999), 170–176.

[27] K.L. Cooke, Functional differential equations: Some models and perturbation problems, in *Differential Equations and Dynamical Systems* (J.K. Hale and J.P. LaSalle, eds.), Academic Press, New York, 1967, 167–183.

[28] J.M. Cushing, *Integrodifferential Equations and Delay Models in Population Dynamics*, Lec. Notes in Biomath., Vol. 20, Springer-Verlag, Berlin and New York, 1977.

[29] U. Forys and A. Marciniak-Czochra, Logistic equations in tumor growth modelling, *Int.J. Appl. Math. Comput. Sci.*, **3** (2003), 317–325.

[30] F.C. Hoppensteadt, ed., *Mathematical Aspects of Physiology*, Lectures in Applied Math, Vol. 19, American Mathematical Society, Providence, 1981.

[31] Gibson, Clark, Sensitivity analysis for a class of evolution equations, *J. Mathematical Analysis and Applications*, **58** (1977), 22-31.

[32] G.E. Hutchinson, Circular causal systems in ecology, *Ann. N.Y. Acad. Sci.,* **50** (1948) 221–246.

[33] J. Hale, *Theory of Functional Differential Equations*, Springer-Verlag, New York, 1977.

[34] Y. Kuang, *Delay Differential Equations: with Applications in Population Dynamics*, Academic Press, Inc., San Diego, 1993.

[35] N. MacDonald, Time lag in a model of a biochemical reaction sequence with end-product inhibition, *J. Theor. Biol.*, **67** (1977), 727–734.

[36] M. Martelli, K.L. Cooke, E. Cumberbatch, B. Tang and H. Thieme, eds., *Differential Equations and Applications to Biology and to Industry*, World Scientific, Singapore. 1996.

[37] J.A.J. Metz and O. Diekmann, eds., *The Dynamics of Physiologically Structured Populations*, Lecture Notes in Biomath., Vol. 68, Springer-Verlag, Berlin, Heidelberg, New York, 1986.

[38] N. Minorsky, Self-excited oscillations in dynamical systems possessing retarded actions, *J. Applied Mechanics*, **9** (1942), A65–A71.

[39] N. Minorsky, On non-linear phenomenon of self-rolling, *Proc. National Academy of Sciences*, **31** (1945), 346–349.

[40] N. Minorsky, *Nonlinear Oscillations*, Van Nostrand, New York, 1962.

[41] A. Pazy, *Semigroups of Linear Operators and Applications to Partial Differential Equations,* Springer-Verlag, New York, 1983.

[42] K. Schmitt, ed., *Delay and Functional Differental Equations and Their Applications*, Academic Press, New York, 1972.

[43] L.F. Shampine and S. Thompson, *Delay Differential Equations With dde23,* (2000), SMU/RU.

[44] K. Thomseth and C. Cobelli, Generalized sensitivity functions in physiological system identification, *Ann. Biomed. Engr.*, **27** (1999), 606–616.

[45] E.M. Wright, A non-linear difference-differential equation, *J. Reine Angew. Math*, **494** (1955), 66–87.

APPENDIX

Appendix A

Numerical Implementation

A.1 Pseudocode Subroutines

makeQ

In this subroutine we will pass N and return QN, where QN is in row echelon form of Q, i.e. it is in an upper triangular matrix. The function syntax is $QN = makeQ(N)$.

$$
\begin{aligned}
Q^N &= (e_0(0), e_1(0), \ldots, e_N(0))^T (e_0(0), e_1(0), \ldots, e_N(0)) \\
&\quad + \int_{-\tau}^{0} (e_0(p), e_1(p), \ldots, e_N(p))^T (e_0(p), e_1(p), \ldots, e_N(p)) dp \\
&= \begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \ldots & 0 \end{pmatrix} \\
&\quad + \begin{pmatrix} \int_{-\tau}^{0} e_0(p)e_0(p)dp & \int_{-\tau}^{0} e_0(p)e_1(p)dp & \int_{-\tau}^{0} e_0(p)e_2(p)dp & \ldots & \int_{-\tau}^{0} e_0(p)e_N(p)dp \\ \int_{-\tau}^{0} e_1(p)e_0(p)dp & \int_{-\tau}^{0} e_1(p)e_1(p)dp & \int_{-\tau}^{0} e_1(p)e_2(p)dp & \ldots & \int_{-\tau}^{0} e_1(p)e_n(p)dp \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \int_{-\tau}^{0} e_N(p)e_0(p)dp & \int_{-\tau}^{0} e_N(p)e_1(p)dp & \ldots & \int_{-\tau}^{0} e_N(p)e_{N-1}(p)dp & \int_{-\tau}^{0} e_N(p)e_N(p)dp \end{pmatrix}.
\end{aligned}
$$

Thus

$$
Q^N = \begin{pmatrix} 1 & 0 & \ldots & 0 & 0 \\ 0 & 0 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \ldots & 0 & 0 \end{pmatrix} + \frac{\tau}{N} \begin{pmatrix} \frac{1}{3} & \frac{1}{6} & 0 & \ldots & 0 \\ \frac{1}{6} & \frac{2}{3} & \frac{1}{6} & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & & \frac{1}{6} & \frac{2}{3} & \frac{1}{6} \\ 0 & \ldots & 0 & \frac{1}{6} & \frac{1}{3} \end{pmatrix}.
$$

So to compute the above matrix we define $Q1 = zeros(N+1, N+1)$, and $Q2 = zeros(N+1, N+1)$. Then we set $Q1(1,1) = 1$. Then $Q2 = \frac{\tau}{N}[diag((1/6) * ones(1, N-1), -1) + (diag((2/3) * ones(1, N))) + (diag((1/6) * ones(1, N-1), 1)]$. We then set $Q2(1,1) = \frac{\tau}{N}\frac{1}{3}$, and $Q2(N+1, N+1) = \frac{\tau}{N}\frac{1}{3}$, now $Q^N = Q1 + Q2$. We row reduce to put Q^N in upper triangular form in the following way:

$$
\begin{aligned}
QN = zeros(N+1, N+1) \\
QN(1,:) = Q^N(1,:) \\
\text{for i = 2: N+1} \\
QN(i,:) = QN(i-1,:)(\frac{-Q(i,i-1)}{QN(i-1,i-1)}) + Q(i,:). \\
\text{end}
\end{aligned}
$$

The subroutine will return the matrix QN.

makeh

For this subroutine the syntax will be $h = makeh(\xi, \phi)$, then

$$
\begin{aligned}
h(\xi, \phi) &= (e_0(0), e_1(0), \ldots, e_N(0))^T \xi + \int_{-\tau}^{0} (e_0(p), e_1(p), \ldots, e_N(p))^T \phi(p) dp \\
&= \begin{pmatrix} \xi \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \begin{pmatrix} \int_{-\tau}^{0} e_0(p)\phi(p)dp \\ \int_{-\tau}^{0} e_1(p)\phi(p)dp \\ \vdots \\ \int_{-\tau}^{0} e_N(p)\phi(p)dp \end{pmatrix}.
\end{aligned}
$$

backSub

For the back substitution subroutine we use the following syntax

$$x = backSub(A, b)$$

where $Ax = b$ and the matrix A is in echelon form making it an upper triangular matrix.

From here the following steps take place:

$$x = zeros(N, 1)$$
$$x(N) = b(N)/A(N, N)$$
$$\text{for j=N-1:-1:1}$$
$$x(j) = (b(j) - A(j, j+1:N) * x(j+1:N))/A(j, j)$$
$$\text{end}$$

makeH

This subroutine syntax is as follows:

$$H = makeH(x(t), x(t-\tau), r, K, N).$$

For the subroutine we will only need to define $a = r[1 - \frac{x(t-\tau)}{K}]$, the first entry of the first column, and $b = -\frac{rx(t)}{K}$, the first entry of the last column. By passing N we can then define a matrix H1= zeros(N+1,N+1), then set the $H(1,1) = a$, and $H(1, end) = b$. From here we define H2=zeros(N+1,N+1), and create the known unchanging tridiagonal matrix by $H2 = [(diag((1/2) * ones(1, N-1), -1)) + (diag(zeroes(1, N))) + (diag((-1/2) * ones(1, N-1), 1)]$. We then set $H2(1,1) = 1/2$, and $H2(N+1, N+1) = -1/2$, now we return $H = H1 + H2$.

RHS

This function computes the right hand side of (4.6). In this function the coordinate vector $\gamma(t)$ is determined at each *value*, so we use other functions backSub and makeH to determine this coordinate vector. We also compute $f^N(t)$ within the function, as well as $x(t)$ and $x(t-\tau)$, which is needed to compute matrix $H(t)$ and $f^N(t)$ for the delay logistic example. The syntax for RHS is as follows:

$$\begin{aligned}
dzdt = rhs(t, z, r, k, N, QN, tau, x(t), x(t-\tau))\\
gamma = zeros(N+1, 1)\\
H = zeros(N+1, N+1)\\
H = makeH(x(t), x(t-tau), r, k, N)\\
f = zeros(N+1, 1)\\
f(1,1) = \frac{(rx(t)x(t-\tau))/k}{QN(1,1)}\\
gamma = backSub(QN, H * z)\\
dzdt = gamma + f
\end{aligned}$$