ALGEBRAIC-DELAY DIFFERENTIAL SYSTEMS: CONTINUOUSLY EXTENDABLE BANACH MANIFOLDS AND LINEARIZATION

NEMANJA KOSOVALIC

A DISSERTATION SUBMITTED TO THE FACULTY OF GRADUATE STUDIES
IN PARTIAL FULFILMENT OF THE REQUIREMENTS
FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

GRADUATE PROGRAM IN APPLIED MATHEMATICS
YORK UNIVERSITY
TORONTO, ONTARIO
JUNE 2014

Abstract

Consider a population of individuals occupying some habitat, and assume that the population is structured by age. Suppose that there are two distinct life stages, the immature stage and the mature stage. Suppose that the mature and immature population are not competing in the sense that they are consuming different resources. A natural question is "What determines the age of maturity?" A subsequent natural question is "How does the answer to the latter question affect the population dynamics?" In many biological contexts, including those from plant and insect populations, the age of maturity is not merely constant but is more accurately determined by whether or not the food concentration reaches a prescribed threshold.

We consider a model for such a population in terms of a nonlinear transport equation with nonlocal boundary conditions. The variable age of maturity gives rise to an implicit state-dependent delay in the system of first order partial differential equations. We explain the relevance of this problem and provide a mechanistic derivation of the model equations. We address the existence, positivity, and continuity of the solution semiflow arising from the model equations, and then we discuss the differentiability of the semiflow with respect to initial data, in a suitable weak sense. The problem of the differentiability of the solution semiflow arising from even ordinary differential equations containing state-dependent delays was a long standing open problem for some time. Prior to this work, there were no results which addressed the linearization of the solution semiflow corresponding to a partial differential equation having a state-dependent delay.

Acknowledgements

I would like to thank my doctoral advisor Dr. Jianhong Wu for recognizing my strengths, teaching me how to use them at full capacity, and for providing me with the opportunity to participate in conferences worldwide. I would like to thank my co-advisor Dr. Yuming Chen for his patience and support. I am grateful to Dr. Glenn Webb of Vanderbilt University for believing in me when I did not believe in myself. I would like to thank Dr. Bernhard Lani-Wayda of Justus-Liebig University for my visit to Giessen, and for visiting our group. I would like to thank Dr. Laurent Pujo-Menjouet of Lyon 1 University for his advice, my visit to Lyon, and his support. I am grateful to Dr. Yang Kuang, Dr. Hal Smith, and Dr. Horst Thieme of Arizona State University for my visit and their advice. I would like to thank Dr. Stan Kochman for his advice and utmost patience and support. I am grateful to Dr. Juris Steprans for improving my teaching. I would like to thank Dr. Peter Olofsson of Trinity University for teaching me about career options. Finally, I would like to thank Dr. Jacques Bélair and Dr. Peter Gibson for their service.

Table of Contents

Abstract **ii**

Acknowledgements **iv**

Table of Contents **v**

1 Introduction **1**

1.1 The Big Picture 1

1.2 Mechanistic Derivation 4

2 Existence and Continuity of the Abstract Solution Semiflow **9**

2.1 Comparison with Existing Results 9

2.2 Technical Preliminaries and Hypotheses 11

2.2.1 The Ambient Linear Space of Initial Data 12

2.2.2 Hypotheses 12

2.3 Local Solutions in M_0 15

2.4 Maximal Solutions and a Semiflow on M_0 28

2.5 The Model Equations: Part One 32

3 Differentiability of the Abstract Solution Semiflow with Respect to Initial Data 38

3.1 Background 38

3.1.1 Outline and Main Results 45

3.1.2 Morally Finite or Infinite Dimensional Problem? 47

3.2 Technical Preliminaries and Hypotheses 49

3.2.1 The Ambient Linear Space of Initial Data 50

3.2.2 Hypotheses 50

3.3 Differentiability of Solutions with respect to Time 54

3.4 The semiflow on M_0 and its Restriction to $\hat{M}_0$ 66

3.5 The Linear Variational System along Flowlines in $\hat{M}_0$ 69

3.6 Derivatives of Solution Operators $\hat{S}_t$ on $\hat{M}_0$ 79

3.7 The Model Equations: Part Two 93

4 C^0-Extendable Banach Manifolds 104

4.1 C^1 Banach Manifolds 104

4.2 The C^0-Extendable Smooth Structure 106

4.3 C^0-Extendable Submanifolds 108

5 Concluding Remarks **110**

5.1 Summary 110

5.2 Future Research Directions 111

Bibliography **111**

1 Introduction

We start off by discussing the big picture, that is, how the results presented in this thesis complement existing works in the field of functional differential equations, as well as their relevance for the broader research community. We then present a mechanistic derivation of the model equations.

1.1 The Big Picture

The work presented in this thesis is a coherent integration of two different themes. The two themes being integrated are ideas from the theory of differential equations containing *state-dependent delays* [35, 36], and the modeling of age structured populations using nonlinear transport equations [38]. The outcome is a general framework for the analysis of a class of age structured population models, having two distinct and non competing stages (the mature and immature stages), and with the special feature that the age of maturity of an individual at a given time is determined by whether or not the resource concentration, which depends on the *immature population*, reaches a prescribed threshold. An upshot of this thesis is that one can create new mathematics by letting oneself cross traditional interdisciplinary boundaries, instead of focusing on improving the technicalities of existing work. Below I briefly explain first the mathematical, and then the interdisciplinary significance of this work.

Differential equations containing state-dependent delays [ordinary differential equations (ODEs), or partial differential equations (PDEs)] are the focus of much of the current research on functional differential equations, and present the most challenging problems in this field [9, 31, 36, 14, 29, 28]. This is because the nonlinear term typically is not differentiable (or even Lipschitz!) on the initial history space of continuous functions, and the corresponding Cauchy problem is not well posed. Hence results from the well known monographs [7, 3, 8] do not apply to even ODEs containing state-dependent delays. An alternative for the case of ODEs is to work on a submanifold of the space of C^1 functions, or to work on the space of Lipschitz functions. As a result, this class of equations does not fit well into any classical framework for smooth semi-dynamical systems. In the case of PDEs containing state-dependent delays this problem is more serious since in general, solutions of PDEs are not locally Lipschitz in time. Consequently, there are very few works which deal with differential equations containing both state-dependent delays and partial differential operators.

Threshold phenomena in structured populations has been proposed in many biological contexts. This is usually manifested as a variable transition age between two distinct life stages. See [9] for examples in epidemiology, and [26, 6] for examples in fish and insect populations. Other examples of state-dependent delays appearing in population models, including one for hematopoiesis can be found in [1, 21]. The model described in the first paragraph of this introduction incorporates this idea for the case of a two stage population, in which there is no competition between juveniles and adults, and for which the age of transition from the juvenile to the adult stage depends implicitly on the history of itself and the history of the juvenile population, giving rise to a state-dependent delay. *Prior to this thesis there was no mathematical apparatus in which these models, in their natural PDE setting, could be embedded and not to mention the linearization of the corresponding solution semiflow. Furthermore, linearization of the solution semiflow for even ODEs with state-dependent delays was a long standing open problem, see e.g.* [35, 9]. Prior to the work [35] many researchers used to treat the linearization of a state-dependent delay equation in a purely formal way as in e.g. [16].

The situation above can be described in terms of a nonlinear transport equation coupled to an algebraic-delay term. This gives rise to an abstract algebraic-delay differential system. In Chapter 2, we establish sufficient conditions for the corresponding initial value problem to give rise to a continuous semiflow, on a subset of the ambient initial history space of continuous functions. In Chapter 3 we give sufficient conditions for the differentiability of the semiflow (in a weak sense), and in a certain phase space. *The main challenge here is to come up with the right notion of differentiability and the right phase space to recover the desired result. In particular, this leads to the notion of a "C^0-extendable submanifold", which is related to "almost Fréchet differentiable functions" introduced in [23].* We note for the convenience of the reader, that Chapter 2 and Chapter 3 although closely related, are entirely self contained and can be read in any order.

1.2 Mechanistic Derivation

The derivation below is adapted from [12] but morally goes back to [26].

Consider some abstract habitat and some population of individuals living in this habitat. Let $u(t, a)$ be the density of individuals of age a at time t. Let the immature population at time t be given by $I(t)$. Let $S(t)$ denote the concentration density of some resource per unit volume in the habitat at time t. To derive a deterministic model we need to make some assumptions.

First, we assume that $S(t)$ satisfies $S'(t) = S_0 - (\gamma_i I(t) + C)S(t)$. Here $S_0 > 0$ is a constant rate of food recruited in the habitat, $\gamma_i > 0$ is the rate of food consumption of the immature population per unit time, and $C > 0$ represents the resource consumption rate by anything else in the habitat. Since the resource consumption happens on a much faster time scale than that of life of the population, we can make a simplifying assumption. If we hold the immature population fixed, we get the equation, $S'(t) = S_0 - (\gamma_i I + C)S(t)$. The steady state is given by the formula $S = \frac{S_0}{\gamma_i I + C}$. Since this steady state is globally stable, the quasi steady state approximation gives

$$S(t) = \frac{S_0}{\gamma_i I(t) + C}. \tag{1.2.1}$$

For further details see [25].

Second we assume that the age of maturity at time t, $\tau(t)$, is defined by the condition

$$\int_{t-\tau(t)}^{t} S(\sigma)d\sigma = T > 0, \tag{1.2.2}$$

where $T > 0$ is a "size" threshold. This represents the difference between an individual's size at birth and their size τ units of time after birth. Combining (1.2.1) with (1.2.2) gives us

$$\int_{t-\tau(t)}^{t} \frac{S_0}{\gamma_i I(\sigma) + C} d\sigma = T \qquad \text{with} \qquad I(\sigma) = \int_0^{\tau(\sigma)} u(\sigma, a)da$$

or equivalently

$$\int_{t-\tau(t)}^{t} S_0 \left[\gamma_i \int_0^{\tau(\sigma)} u(\sigma, a)da + C\right]^{-1} d\sigma = T.$$

For convenience, let $S_0 = \gamma_i = 1$ and this can be achieved by rescaling the relevant parameters.

Finally we assume that the individuals have maximum age $0 < m \leq \infty$ and $u(t,a)$ satisfies the standard first order transport partial differential equations (PDEs),

$$\begin{aligned} &\partial_t u(t,a) + \partial_a u(t,a) = -d(a)u(t,a), \; t \geq 0 \text{ and } 0 \leq a < m; \\ &u(t,0) = b(\textstyle\int_{\tau(t)}^{m} \beta(\xi)u(t,\xi)d\xi), \end{aligned} \tag{1.2.3}$$

where $\tau(t)$ is given by

$$\int_{-\tau(t)}^{0} \left[\int_{0}^{\tau(t+\sigma)} u(t+\sigma,a)da + C \right]^{-1} d\sigma = T. \tag{1.2.4}$$

Note that we ignore technicalities concerning whether $\tau(t)$ is well defined by (1.2.4) at this stage. To have solutions for $t \geq 0$ we must specify the initial conditions,

$$\tau(t) = \varphi(t) \qquad \text{for } -a_m \leq t \leq 0$$

and

$$u(t,a) = \psi(t,a) \qquad \text{for } -a_m \leq t \leq 0 \text{ and } 0 \leq a < m.$$

Here $a_m \in (0,m)$ is the maximal age of maturity.

A look at (1.2.4) reveals that $\tau(t)$ depends on the history at time t of the population density, u_t, and the history of itself, τ_t. As usual, $u_t(\theta)(\cdot) = u(t+\theta)(\cdot)$ and $\tau_t(\theta) = \tau(t+\theta)$ for $\theta \in [-a_m, 0]$ and we abuse the notation that $u(t,a) =$

$u(t)(a)$. We assume naively that $\tau(t)$ is a function of u_t and τ_t, $\tau(t) = H(u_t, \tau_t)$ (see Section (2.2) and Section (2.5)).

To summarize, we have obtained the initial value problem,

$$\begin{aligned} &\partial_t u(t,a) + \partial_a u(t,a) = -d(a)u(t,a), \\ &u(t,0) = b(\textstyle\int_{\tau(t)}^m \beta(\xi)u(t,\xi)d\xi), \\ &\tau(t) = H(u_t, \tau_t) \end{aligned} \tag{1.2.5}$$

for $t \geq 0$ and $0 \leq a < m$ with initial conditions

$$\tau(t) = \varphi(t) \qquad \text{and} \qquad u(t,a) = \psi(t,a) \text{ for } -a_m \leq t \leq 0 \text{ and } 0 \leq a < m. \tag{1.2.6}$$

Note that for each $t \geq 0$, $\begin{pmatrix} u_t \\ \tau_t \end{pmatrix} \in M_0$, where

$$M_0 = \left\{ \begin{pmatrix} \psi \\ \varphi \end{pmatrix} \in \text{some subset of } C([-a_m, 0], L^1[0,m) \times \mathbf{R}) \mid \varphi(0) = H(\psi, \varphi) \right\}.$$

The precise definitions of H and M_0 are given in Sections (2.2) and (2.5). We can rewrite the initial value problem (1.2.5)–(1.2.6) abstractly as

$$\begin{aligned} &\tfrac{d}{dt} \begin{pmatrix} 0 \\ u(t,\cdot) \end{pmatrix} = \begin{pmatrix} -u(t,0) \\ -u_a(t,\cdot) \end{pmatrix} + \begin{pmatrix} b(\int_{\tau(t)}^m \beta(\xi)u(t,\xi)d\xi) \\ -d(\cdot)u(t,\cdot) \end{pmatrix}, \\ &\tau(t) = H(u_t, \tau_t), \\ &\begin{pmatrix} x_0 \\ \tau_0 \end{pmatrix} = \begin{pmatrix} \psi \\ \varphi \end{pmatrix} \in M_0. \end{aligned}$$

This mechanistic derivation shows how we are naturally led to considering the more general algebraic-delay differential system, the initial value problem (2.2.1) of Section 2.2, which is the object of study in this thesis.

2 Existence and Continuity of the Abstract Solution Semiflow

We make a comparison with related existing results in the literature and then proceed with our abstract framework. In Section 2.2, we state the relevant technical preliminaries and hypotheses, including the appropriate notion of mild solutions in the subset M_0 of the ambient linear space of continuous functions. In Section 2.3 we prove the existence and uniqueness of local mild solutions in M_0, (Theorem 1). In Section 2.4, we discuss the corresponding semiflow and show that it is continuous (Theorem 2). In Section 2.5, we give an application of the general theory (Proposition 3).

2.1 Comparison with Existing Results

Similar types of structured population models were already considered by Smith in [32]. The models considered by Smith were reduced to a single retarded functional differential equation whose nonlinear term is Lipschitz on the usual phase space of continuous functions. Due to the dependence of the age of maturity on the immature population, this is not possible for the case we are considering. This is because the second component of the system we are studying, which describes the age of maturity in the mechanistic derivation above, depends not only on the history of the population density but also on the history of itself. For a related work on threshold type delay differential equations, see [15].

More recently, Hbid et al. [11] considered a stage structured population model with the same feature determining the age of maturity that we have here. However, based on a simplifying assumption, they reduced the model to an integral equation containing a state-dependent delay, for which the immature population depends only on the history of the state variable, and consequently, does not need to be initialized.

The nonlinear semigroups approach we are using here was motivated by the works of Thieme [34], and Magal and Ruan [20, 19], specifically for the case of structured population models. For another semigroup approach for age structured models, see [38]. The basic reference for semigroup theory is [27].

A unification of various fundamental results for PDE with ordinary delay is given in [30], which uses a more general class of operators than we have here. It would be nice to see if the results presented here can find such generalizations. For a treatment of reaction diffusion systems with ordinary delay, see [39].

Also closely related is the recent work of Walther [36] on ODE algebraic-delay differential systems. Walther considered systems of the form

$$\begin{aligned} x'(t) &= f(x_t, r(t)), \\ 0 &= \Delta(r(t), x_t), \end{aligned}$$

where $x(t) \in \mathbf{R}^k$ and $r(t)$ is defined implicitly by the history of the state, x_t. As long as the derivative of Δ in the first component is nonsingular, such systems will be locally uniquely solvable thanks to the implicit function theorem. Unfortunately, we cannot apply the implicit function theorem for the case we are considering, so instead we impose a special Lipschitz condition on the function H given in the next section.

2.2 Technical Preliminaries and Hypotheses

In this section we state the relevant technical preliminaries and hypotheses. All Banach spaces are assumed to be over the real numbers. Whenever a product of Banach spaces is considered, we view it as a Banach space equipped with the corresponding product norm.

2.2.1 The Ambient Linear Space of Initial Data

Let $\delta > 0$ and $I = [-\delta, 0]$. For $F \subset E$, where E is a Banach space, $C(I, F)$ denotes the set of continuous functions mapping I into F. For $\psi \in C(I, F)$, we let $||\psi||$ be the supremum norm of ψ. Then $(C(I, E), || \cdot ||)$ is a Banach space.

Suppose that $0 < T < \infty$ and $y : I \cup [0, T] \to F$ is some map. As usual in the literature on delay equations, for each $t \in [0, T]$, we define $y_t : I \to F$ by $y_t(\theta) = y(t + \theta)$ for $\theta \in I$ and call y_t the history of y at time t. If $T = \infty$ then the same definition applies with $t \in [0, T]$ being replaced with $t \in [0, T)$.

2.2.2 Hypotheses

(H1) Let $(X, | \cdot |)$ denote a Banach space and suppose that $A : D(A) \to X$ with $D(A) \subset X$ is a linear operator satisfying the estimates of the Hille-Yosida theorem, that is, there is some $M \geq 1$ and some $\omega \in \mathbf{R}$ such that the ray $(\omega, \infty) \subset \rho(A)$ and $||(A - \lambda I)^{-n}|| \leq \frac{M}{(\lambda - \omega)^n}$ for $\lambda > \omega$ and for each positive integer n.

We let $X_0 = \overline{D(A)}$ and A_0 denote the part of A in X_0. Actually this class of operators falls under a more general class of well known operators as pointed out in [30]. Set $R_\lambda = (A - \lambda I)^{-1}$. Without loss of generality, assume that $\omega > 0$. It follows from (H1) that A_0 generates a C^0-semigroup of linear operators on X_0,

$\{T(t)\}_{t\geq 0}$, and that $||T(t)|| \leq Me^{\omega t}$.

(H2) Let $n > 0$ be given. Suppose that K is some compact subset of $\mathbf{R}^n$ such that K is contained in the closed ball of radius $h > 0$ centered at the origin. Set $I = [-h, 0] \subset \mathbf{R}$ and let C_0 be some closed and convex subset of X_0. Assume that $R_0 > 0$ and $f : \mathbf{R}^+ \to \mathbf{R}^+$ is a strictly increasing function with $f(R_0) = 1$. Let $D(H) = \left\{ \begin{pmatrix} \psi \\ \varphi \end{pmatrix} \in C(I, C_0 \times K) \mid ||\psi|| \leq R_0 \right\}$ and suppose $H : D(H) \to K$ is a function which satisfies the following Lipschitz condition: for each $Q > 0$ there is some $L_Q > 0$ such that, for $\begin{pmatrix} \psi_1 \\ \varphi_1 \end{pmatrix}, \begin{pmatrix} \psi_2 \\ \varphi_2 \end{pmatrix} \in D(H)$ with $||\psi_i|| \leq Q$ $(i = 1, 2)$, we have

$$|H(\psi_1, \varphi_1) - H(\psi_2, \varphi_2)| \leq \; f(Q)\; ||\varphi_1 - \varphi_2|| + L_Q||\psi_1 - \psi_2||.$$

For simplicity of notation, $|\cdot|$ has been used to denote the norm on X and also the norm on $\mathbf{R}^n$. This will not cause any confusion.

(H3) Let $M_0 = \left\{ \begin{pmatrix} \psi \\ \varphi \end{pmatrix} \in D(H) \mid \varphi(0) = H(\psi, \varphi) \text{ and } ||\psi|| < R_0 \right\}$. Assume $M_0 \neq \emptyset$.

(H4) Suppose $F : C_0 \times K \to X$ is a globally Lipschitz function, i.e., there is some $D > 0$ such that, for c_1, $c_2 \in C_0$ and k_1, $k_2 \in K$, we have $|F(c_1, k_1) - F(c_2, k_2)| \leq D(|c_1 - c_2| + |k_1 - k_2|)$.

(H5) (Subtangential Condition) We assume that, for each $(c,k) \in C_0 \times K$,

$$\lim_{h \downarrow 0} \frac{\operatorname{dist}\left(T(h)c + \lim\limits_{\mu \to \infty} \int_0^h T(s)\mu R_\mu F(c,k)ds, C_0 \right)}{h} = 0$$

holds. Here, $\operatorname{dist}(x, B) = \inf_{b \in B} |x - b|$ for $x \in X$ and $B \subset X$. (H5) is a well known condition which ensures the invariance of a closed and convex set, sometimes referred to as positivity. We refer readers to [24, 30, 34] for more detail.

Definition. Consider the following initial value problem,

$$\begin{cases} x'(t) = Ax(t) + F(x(t), a(t)), \\ a(t) = H(x_t, a_t), \\ \begin{pmatrix} x_0 \\ a_0 \end{pmatrix} = \begin{pmatrix} \psi \\ \varphi \end{pmatrix} \in M_0. \end{cases} \tag{2.2.1}$$

By a mild solution of (2.2.1) on $I \cup [0,T]$ in M_0 with $T < \infty$, we mean a pair of functions $\begin{pmatrix} x(t) \\ a(t) \end{pmatrix}$ with the following properties:

(i) $a : I \cup [0,T] \to K$ is continuous.

(ii) $x : I \cup [0,T] \to C_0$ is continuous such that, for each $t \in [0,T]$, $\int_0^t x(s)ds \in$

$D(A)$ and

$$x(t) = x(0) + A\int_0^t x(s)ds + \int_0^t F(x(s), a(s))ds.$$

(iii) For $0 \le t \le T$, $\begin{pmatrix} x_t \\ a_t \end{pmatrix} \in M_0$, i.e., $a(t) = H(x_t, a_t)$ and $||x_t|| < R_0$.

(iv) $\begin{pmatrix} x_0 \\ a_0 \end{pmatrix} = \begin{pmatrix} \psi \\ \varphi \end{pmatrix}$.

We similarly define mild solutions in M_0 on $I \cup [0, T)$ for $T = \infty$.

Note that (H1) implies that (ii) is equivalent to

$x(t) = T(t)\psi(0) + \lim\limits_{\mu\to\infty} \int_0^t T(t-s)\mu R_\mu F(x(s), a(s))ds$ for $t \in [0, T]$ (see [34]).

2.3 Local Solutions in M_0

In this section we establish the existence and uniqueness of local mild solutions for (2.2.1) in M_0.

Theorem 1 *Suppose $A : D(A) \to X$, $H : D(H) \to K$, $F : C_0 \times K \to X$, and M_0 are as in Section (2.2). Assume (H1)–(H5) hold. Then the initial value problem (2.2.1) has a unique mild solution $\begin{pmatrix} x(t) \\ a(t) \end{pmatrix}$ in M_0 on $I \cup [0, \tau]$ for some $0 < \tau < \infty$.*

Proof. We establish the existence and uniqueness of a local mild solution of (2.2.1) in M_0 by constructing a net of approximate solutions using a discrete approximation scheme. This is done in such a way that the histories of the approximate solutions lie in M_0. We show that the net constructed converges to a local mild solution of (2.2.1) in M_0. This method is a well known approach. See [34, 24, 30], for example. The difference between our version and others is that we must work on a "nonlinear submanifold" of the ambient space.

Step 1: Constructing an approximate solution of (2.2.1) in M_0.

We choose $R_1 > 0$ such that $||\psi|| < R_1 < R_0$. Set $R = R_1 - ||\psi||$. Then $0 < R < R_0$. Moreover, $f(R_1) < 1$. By (H2) we can find $J > 0$ such that if $||\gamma_1||$, $||\gamma_2|| \leq R_1$ then, for φ_1, $\varphi_2 \in C(I, K)$, we have

$$|H(\gamma_1, \varphi_1) - H(\gamma_2, \varphi_2)| \leq f(R_1)\, ||\varphi_1 - \varphi_2|| + J||\gamma_1 - \gamma_2||.$$

Let $\psi(0) = x^0$ and $\varphi(0) = a^0$. Fix some number $\epsilon \in (0, 1)$. Pick some $0 < \tau \leq R_1$ (another upper bound on τ independent of ϵ will be imposed later). Using (H5), the strong continuity of $T(t)$, and the uniform continuity of ψ and φ, we can find some $0 < t_1 \leq \min\{\epsilon, 2\tau\}$ such that

$$\frac{\text{dist}(G(t_1), C_0)}{t_1} < \frac{\epsilon}{2}, \tag{2.3.1}$$

if $s \in [0, t_1)$ then $|T(s)(x^0) - x^0| \leq \epsilon$,

if s_1, $s_2 \in I$ with $|s_1 - s_2| < t_1$ then $|\varphi(s_1) - \varphi(s_2)|$, $|\psi(s_1) - \psi(s_2)| \leq \epsilon$, where $G(t_1) = T(t_1)x^0 + \lim_{\mu\to\infty} \int_0^{t_1} T(t_1 - s)\mu R_\mu F(x^0, a^0)ds$. Choose $x^1 \in C_0$ such that

$$|x^1 - G(t_1)| \leq \frac{\epsilon t_1}{2} + \text{dist}(G(t_1), C_0) \leq \epsilon t_1.$$

It follows that

$$\begin{aligned} |x^1 - T(t_1)x^0| &\leq \epsilon\, t_1 + \int_0^{t_1} M^2 e^{\omega t_1} |F(x^0, a^0)| ds \\ &\leq 2\tau + M^2 2\tau e^{\omega 2\tau} |F(x^0, a^0)|. \end{aligned}$$

This, combined with $|x^1 - x^0| \leq |x^1 - T(t_1)(x^0)| + |T(t_1)(x^0) - x^0|$, tells us that we can choose τ independently of ϵ and t_1 so that $|x^1 - x^0| \leq R$.

We define a function $x^1 : I \cup [0, t_1] \to C_0$ by

$$x^1(t) = \begin{cases} \psi(t) & \text{if } t \in I, \\ \frac{t}{t_1} x^1 + \frac{t_1 - t}{t_1} x^0 & \text{if } t \in [0, t_1]. \end{cases}$$

Then, for $t \in [0, t_1]$, $x^1(t)$ is a parameterization of the straight line segment joining x^0 and x^1, meaning that $x^1(t) \in C_0 \cap B_R(x^0)$, where $B_R(x^0)$ denotes the closed ball of radius R in X_0 about x_0, which is convex. Consequently, for $t \in [0, t_1]$, $x^1_t \in C(I, C_0)$ and $||x^1_t|| \leq R_1 < R_0$.

To find a corresponding approximation for the second component of the system, we wish to solve the equation

$$a^1(t) = \begin{cases} \varphi(t) & \text{if } t \in I, \\ H(x^1_t, a^1_t) & \text{if } t \in [0, t_1]. \end{cases} \tag{2.3.2}$$

To show that (2.3.2) has a unique solution, we construct an appropriate contraction on $C(I \cup [0,t_1],K)$ which is a closed subset of the Banach space $C(I \cup [0,t_1],\mathbf{R}^n)$ since K is closed. Note that $\{x_t^1\} \times C(I,K) \subset D(H)$ for $t \in [0,t_1]$. So let $\mathcal{A} : C(I \cup [0,t_1],K) \to C(I \cup [0,t_1],\mathbf{R}^n)$ be given by the right hand side of (2.3.2). It follows from (H2) that

$$(\mathcal{A}a)(s) \in K \text{ for each } s \in I \cup [0,t_1] \text{ and that } \mathcal{A}a \text{ is continuous on } I \cup [0,t_1]$$

and

$$||\mathcal{A}a - \mathcal{A}b|| \leq W||a-b|| \text{ for some } W < 1.$$

Therefore, equation (2.3.2) has a unique solution, denoted by a^1.

This concludes the first step of our recursion and we have obtained appropriate functions $x^1 : I \cup [0,t_1] \to C_0$ and $a^1 : I \cup [0,t_1] \to K$. By relabelling if necessary, we assume that t_1 is chosen maximally in the following way:

Let $S_1 = \sup\{s \in [0,2\tau] \mid 0 < s \leq \epsilon,\ \xi \in [0,s) \Rightarrow |T(\xi)x^0 - x^0| \leq \epsilon$, if $s_1, s_2 \in I$ and $|s_1 - s_2| < s$ then $|\varphi(s_1) - \varphi(s_2)|$ and $|\psi(s_1) - \psi(s_2)| \leq \epsilon$, $\text{dist}(T(s)x^0 + \lim_{\mu\to\infty} \int_0^s T(s-\xi)\mu R_\mu F(x(0),a(0))ds, C_0) \leq \epsilon s/2\}$. Clearly, $S_1 \neq \emptyset$. By a standard continuity argument, it is easy to see that $sup(S_1) \in S_1$ and we set $t_1 = max(S_1)$.

Let $t_0 = 0$. Suppose that $k \geq 1$ and that we are granted a sequence of mesh points $(t_j, x^j, a^j(t_j))$, and corresponding functions, $x^j \in C(I \cup [0,t_j], C_0)$ and $a^j \in$

$C(I \cup [0, t_j], K)$ such that, for each $1 \leq j < k$, the following properties hold:

If $t_{j-1} < \tau$ then (P1)–(P7) hold and if $t_{j-1} \geq \tau$ then $t_j = t_{j-1}$.

(P1) $t_j \leq 2\tau$ and $0 < t_j - t_{j-1} \leq \epsilon$.

(P2) If $s \in [0, t_j - t_{j-1})$ then $|T(s)x^{j-1} - x^{j-1}| \leq \epsilon$. Moreover, for $s_1,\ s_2 \in I \cup [0, t_{j-1}]$, if $|s_1 - s_2| < t_j - t_{j-1}$ then $|a^{j-1}(s_1) - a^{j-1}(s_2)|$, $|x^{j-1}(s_1) - x^{j-1}(s_2)| \leq \epsilon$.

(P3) $\text{dist}(T(t_j - t_{j-1})x^{j-1} + \lim_{\mu\to\infty} \int_{t_{j-1}}^{t_j} T(t_j - s)\mu R_\mu F(x^{j-1}, a^{j-1}(t_{j-1}))ds, C_0) \leq \epsilon(t_j - t_{j-1})/2$.

(P4) t_j is chosen maximally with respect to (P1)–(P3).

Namely, $t_j = \max_{\xi \in [0,2\tau]}\{$(P1)-(P3) hold with '$\xi$' in place of '$t_j$'$\}$.

(P5) $|x^j - T(t_j - t_{j-1})x^{j-1} - \lim_{\mu\to\infty} \int_{t_{j-1}}^{t_j} T(t_j - s)\mu R_\mu F(x^{j-1}, a^{j-1}(t_{j-1}))ds| \leq \epsilon(t_j - t_{j-1})$.

(P6) $x^j \in B_R(x^0)$.

(P7)

$$x^j(t) = \begin{cases} x^{j-1}(t) & \text{if } t \le t_{j-1}, \\ \\ \frac{t-t_{j-1}}{t_j-t_{j-1}} x^j + \frac{t_j-t}{t_j-t_{j-1}} x^{j-1} & \text{if } t \in [t_{j-1}, t_j] \end{cases}$$

and

$$a^j(t) = \begin{cases} a^{j-1}(t) & \text{if } t \le t_{j-1}, \\ \\ H(x_t^j, a_t^j) & \text{if } t \in [t_{j-1}, t_j]. \end{cases}$$

Note that we denote by 'x^j' and 'a^j' both members of C_0 and K, respectively, and the corresponding functions since this should not cause any confusion.

In order to complete the recursion, we show that (P1)–(P7) hold for $j = k$ whenever τ is small enough. It should be noted that τ has not yet been chosen.

If it happens that $t_{k-1} \ge \tau$ then we set $t_k = t_{k-1}$, and we are done. Otherwise, by the same procedure as in the first step of the recursion, we can find some $t_k \le 2\tau$ and $x_k \in C_0$ such that (P1)–(P5) hold. We need to verify (P6), then (P7) will follow exactly as in the first step of the recursion when (2.3.2) was solved using the contraction mapping principle. The purpose of the tedious estimates below is to show that τ can in fact be chosen *a priori* depending only on the initial data. These calculations are essentially those given in [34], but we repeat them here for completion. It should be noted that we use the hypothesis $\omega > 0$ from (H1) to establish (2.3.3) below.

For $j \leq k$, it follows from (P5) that $|x^j - T(t_j - t_{j-1})x^{j-1}| \leq \epsilon(t_j - t_{j-1}) + |\lim_{\mu\to\infty}\int_{t_{j-1}}^{t_j} T(t_j - s)\mu R_\mu F(x^{j-1}, a^{j-1}(t_{j-1}))ds|$. Then

$$\begin{aligned}
|F(x^{j-1}, a^{j-1}(t_{j-1}))| &\leq |F(x^{j-1}, a^{j-1}(t_{j-1})) - F(x^0, a^0)| + |F(x^0, a^0)| \\
&\leq D(|x^{j-1} - x^0| + |a^{j-1}(t_{j-1}) - a^0|) + |F(x^0, a^0)| \\
&\leq D(R + 2h) + |F(x^0, a^0)| := P.
\end{aligned}$$

Clearly P depends only on the initial data. Thus,

$$|x^j - T(t_j - t_{j-1})x^{j-1}| \leq Z(t_j - t_{j-1}), \qquad \text{where } Z = (1 + M^2 e^{\omega 2\tau} P).$$

Having this at our disposal, we next show that, for each $j \leq k$,

$$|x^j - T(t_j)x^0| \leq MZe^{\omega t_j} t_j. \tag{2.3.3}$$

In fact, we have

$$\begin{aligned}
& |x^j - T(t_j - t_{j-2})x^{j-2}| \\
\leq\ & |x^j - T(t_j - t_{j-1})x^{j-1}| + |T(t_j - t_{j-1})x^{j-1} - T(t_j - t_{j-2})x^{j-2}| \\
\leq\ & Z(t_j - t_{j-1}) + |T(t_j - t_{j-1})(x^{j-1} - T(t_{j-1} - t_{j-2})x^{j-2}| \\
\leq\ & Z(t_j - t_{j-1}) + Me^{\omega(t_j - t_{j-1})} Z(t_{j-1} - t_{j-2}) \\
\leq\ & MZe^{\omega(t_j - t_{j-2})}(t_j - t_{j-2}).
\end{aligned}$$

Continuing in this way, we can prove (2.3.3). It follows from (2.3.3) that

$$|x^k - x^0| \leq |x^k - T(t_k)x^0| + |T(t_k)x^0 - x^0| \leq MZe^{\omega 2\tau} 2\tau + |T(t_k)x^0 - x^0|. \tag{2.3.4}$$

Then we can choose $\tau > 0$ such that $x^k \in B_R(x^0)$ and note that, by virtue of (2.3.4) and the strong continuity of $T(t)$, this choice is independent of ϵ.

This completes the recursion and we conclude that for each positive integer j, we can find appropriate mesh points and functions such that (P1)–(P7) hold if $t_{j-1} < \tau$ and otherwise $t_j = t_{j-1}$.

To obtain an approximate solution in M_0, we need to show that this process ends after a finite number of steps. That is, we want to see that, for some positive integer j, $t_j \geq \tau$. We assume, by way of contradiction, that $t_j < \tau$ for each j. So there is some $0 < t \leq \tau$ such that $t_j \uparrow t$ and $t > t_j$. By the same calculations as those on pages 32-33 of [34], we deduce that $x^j \to x$ for some $x \in C_0$. Now we define the function $x : I \cup [0,t] \to C_0$ by

$$x(s) = \begin{cases} x^j(s) & \text{if } -h \leq s \leq t_j, \\ x & \text{if } s = t. \end{cases} \tag{2.3.5}$$

Clearly, x is continuous. Since for each $s \in [0,t]$, $||x_s|| \leq R_1$, the Lipschitz estimate for H with respect to R_1 and the contraction mapping principle give us a unique continuous solution to the equation

$$a(s) = \begin{cases} \varphi(s) & \text{if } s \in I, \\ H(x_s, a_s) & \text{if } s \in [0,t], \end{cases}$$

where x is given by (2.3.5). By uniqueness, it follows that $a(s) = a^j(s)$ for $s \in I \cup [0, t_j]$. By exploiting uniform continuity of x and a on $I \cup [0, t]$, and of the map $[0, t] \ni s \mapsto |T(s)x - x|$ we can find $0 < \delta < \epsilon$ such that $t + \delta \leq 2\tau$, and $|s_1 - s_2| < \delta \Rightarrow |x(s_1) - x(s_2)|, |a(s_1) - a(s_2)| < \epsilon$, and $0 \leq s < \delta \Rightarrow |T(s)x - x| < \epsilon/3$. Fix $\alpha \in (0, \delta)$. Since $t + \alpha > t_j$, by maximality, we see that for each j, one of (P1)–(P3) is not satisfied when 't_j' is replaced by '$t + \alpha$'. It is clear that (P1) is not satisfied for at most finitely many j when t_j is replaced with $t + \alpha$, and similarly for (P2). Therefore, there are infinitely many j such that

$$\text{dist}(T(t + \alpha - t_{j-1})x^{j-1} + \lim_{\mu \to \infty} \int_{t_{j-1}}^{t+\alpha} T(t + \alpha - s)\mu R_\mu F(x^{j-1}, a^{j-1}(t_{j-1}))ds, C_0) > \epsilon(t + \alpha - t_{j-1})/2.$$

Letting j tend to infinity and exploiting continuity shows that the subtangential condition, (H5), is violated, a contradiction.

Step 2: Estimates for the ϵ-approximate solution between mesh points.

The procedure in Step 1 granted us for each $0 < \epsilon < 1$ an approximate solution, which we denote by $\begin{pmatrix} x(t) \\ a(t) \end{pmatrix}$, where $x : I \cup [0, \tau] \to C_0$ and $a : I \cup [0, \tau] \to K$ satisfy

$$x(t) = x^j(t) \text{ and } a(t) = a^j(t) \qquad \text{if } s \leq t_j$$

and $[0, \tau] \subset \cup_{1 \leq j \leq k(\epsilon)}[t_{j-1}, t_j]$ for some $k(\epsilon) = k < \infty$ such that $t_{k-1} < \tau$ and $t_k \geq \tau$. Before moving on to Step 3, we obtain crucial estimates for the components of our approximate solutions.

First we note that by (P2), (P5), and (P7) if $s \in [t_{j-1}, t_j]$ then $|x(s) - x(t_{j-1})| \leq |x^j - x^{j-1}| \leq c_1\epsilon$ for some constant c_1 independent of ϵ. Similarly, we wish to show that there is some $c > 0$ independent of ϵ such that

$$|a(s) - a(t_{j-1})| \leq c\epsilon \qquad \text{for } s \in [t_{j-1}, t_j]. \tag{2.3.6}$$

We achieve this by showing that $||a_s - a_{t_{j-1}}|| \leq c\epsilon$ for each $s \in [t_{j-1}, t_j]$. Let $\theta \in [-h, 0]$ be given. If $s + \theta \leq t_{j-1}$ then $|a(s + \theta) - a(t_{j-1} + \theta)| \leq \epsilon$ by (P2) from Step 1. Otherwise, $s + \theta \geq t_{j-1}$. In this case, we have, by (P2), by the definition of a, and by the Lipschitz estimate for H with respect to R_1 in Step 1, that

$$\begin{aligned} |a(s+\theta) - a(t_{j-1}+\theta)| &\leq |a(s+\theta) - a(t_{j-1})| + |a(t_{j-1}) - a(t_{j-1}+\theta)| \\ &\leq |a(s+\theta) - a(t_{j-1})| + \epsilon \\ &\leq J||x_{s+\theta} - x_{t_{j-1}}|| + f(R_1)||a_{s+\theta} - a_{t_{j-1}}|| + \epsilon. \end{aligned}$$

Using (P2) and $\xi \in [t_{j-1}, t_j] \Rightarrow |x(\xi) - x(t_{j-1})| \leq c_1\epsilon$, it is easy to see that $||x_{s+\theta} - x_{t_{j-1}}|| \leq g\epsilon$ for some constant $g > 0$ independent of ϵ. Therefore, we have that, for each $s \in [t_{j-1}, t_j]$,

$$||a_s - a_{t_{j-1}}|| \leq Jg\epsilon + f(R_1) \sup_{\theta \in I \cap [t_{j-1} - s, 0]} ||a_{s+\theta} - a_{t_{j-1}}|| + \epsilon. \tag{2.3.7}$$

The function $(s, \theta) \mapsto ||a_{s+\theta} - a_{t_{j-1}}||$ defined on the compact set $K_0 := \{(s, \theta) \mid s \in [t_{j-1}, t_j],\ \theta \in I \cap [t_{j-1} - s, 0]\}$ is continuous and hence attains its maximum for some $(s^*, \theta^*) \in K_0$. By (2.3.7), we get

$$||a_{s^*+\theta^*} - a_{t_{j-1}}|| \leq Jg\epsilon + f(R_1)||a_{s^*+\theta^*} - a_{t_{j-1}}|| + \epsilon.$$

This, combined with the fact that $f(R_1) < 1$, gives us

$$||a_{s^*+\theta^*} - a_{t_{j-1}}|| \leq (Jg + 1)(1 - f(R_1))^{-1}\epsilon. \tag{2.3.8}$$

Then (2.3.8) and (2.3.7) together tell us that (2.3.6) holds with $c = (Jg + 1) + f(R_1)(Jg + 1)(1 - f(R_1))^{-1} > 0$. Clearly c depends only on the initial data.

Step 3: The net of approximate solutions converges to a solution as $\epsilon \downarrow 0$.

Using (P1), (P2), (P5), and the estimate (2.3.6) from Step 2, then proceeding exactly as on page 34 in [34], we obtain

$$\left| x^j - T(t_j)x^0 - \lim_{\mu\to\infty} \int_0^{t_j} T(t_j - s)\mu R_\mu F(x(s), a(s))ds \right| \leq d\, \epsilon e^{\omega t_j} t_j \tag{2.3.9}$$

for some constant $d > 0$ independent of ϵ. With the help of (2.3.9), we can argue in the same way as in [34] to get the critical estimate

$$\left| x(t) - T(t)x^0 - \lim_{\mu\to\infty} \int_0^{t} T(t - s)\mu R_\mu F(x(s), a(s))ds \right| \leq d\, \epsilon,$$

which holds for each $t \in [0, \tau]$. The constant d is larger than before (we relabeled) but still independent of ϵ. To complete this step, we must show that the net $\begin{pmatrix} x^\epsilon(t) \\ a^\epsilon(t) \end{pmatrix}$ for $\epsilon \in (0, 1)$ of approximate solutions converges to a solution of (2.2.1).

First we show that $\left\{ \begin{pmatrix} x^\epsilon \\ a^\epsilon \end{pmatrix} \right\}$ is Cauchy in the complete metric space $C(I \cup [0,\tau], C_0 \times K)$. If $\begin{pmatrix} x^\epsilon(t) \\ a^\epsilon(t) \end{pmatrix}$ and $\begin{pmatrix} y^\delta(t) \\ b^\delta(t) \end{pmatrix}$ for ϵ, $\delta \in (0,1)$ are approximate solutions, then (dropping the superscripts) we get

$$\begin{aligned} & |x(t) - y(t)| \\ \leq\ & (\epsilon+\delta)d + \left| \lim_{\mu\to\infty} \int_0^t T(t-s)\mu R_\mu (F(x(s),a(s)) - F(y(s),b(s))ds \right| \quad (2.3.10) \\ \leq\ & (\epsilon+\delta)d + \int_0^t M^2 e^{\omega(t-s)} D(|x(s)-y(s)| + |a(s)-b(s)|)ds. \end{aligned}$$

Since $||x_t||$, $||y_t|| \leq R_1$ for $t \in [0,\tau]$, we get

$$|a(t) - b(t)| \leq J||x_t - y_t|| + f(R_1)||a_t - b_t||$$

and hence

$$\sup_{-t-h\leq\theta\leq 0} |a(t+\theta) - b(t+\theta)| \leq (1-f(R_1))^{-1} J \sup_{-t-h\leq\theta\leq 0} |x(t+\theta) - y(t+\theta)|. \quad (2.3.11)$$

Then, by (2.3.10), (2.3.11), and an application of Gronwall's inequality, we have

$$\sup_{-t-h\leq\theta\leq 0} |x(t+\theta) - y(t+\theta)| \downarrow 0 \text{ uniformly with respect to } t \in [0,\tau] \text{ as } \epsilon,\ \delta \downarrow 0.$$

It follows that $||x-y||_\infty \downarrow 0$ and $||a-b||_\infty \downarrow 0$ as $\epsilon \downarrow 0$ and $\delta \downarrow 0$. Therefore, $\left\{ \begin{pmatrix} x^\epsilon(t) \\ a^\epsilon(t) \end{pmatrix} \right\}$ converges uniformly to a mild solution of (2.2.1) on $I \cup [0,\tau]$ in M_0 as $\epsilon \downarrow 0$.

The uniqueness deserves a few remarks. We suppose that $\begin{pmatrix} x(t) \\ a(t) \end{pmatrix}$ and $\begin{pmatrix} y(t) \\ b(t) \end{pmatrix}$ are two mild solutions of (2.2.1) respectively on $I \cup A_1$ and $I \cup A_2$ in M_0 with the same initial data. Here $A_i = [0, \tau_i]$ or $A_i = [0, \tau_i)$ for $0 < \tau_i \leq \infty$, for each $i = 1, 2$. Let $A = A_1 \cap A_2$. We will show that the two solutions agree on A. Assume first that $x \neq y$. Let $\alpha := \inf\{t \in A \mid x(t) \neq y(t)\}$. Then $x(t) = y(t)$ for $t \leq \alpha$. Choose $\delta > 0$ such that $(\alpha, \alpha + \delta] \subset A$ and $R_2 > 0$ such that for $t \in (\alpha, \alpha + \delta]$, $||x_t||$, $||y_t|| < R_2$ for some $R_2 < R_0$. By (H2) we have that

$$|a(t) - b(t)| \leq L_{R_2}||x_t - y_t|| + f(R_2)||a_t - b_t|| \qquad \text{for } t \in (\alpha, \alpha + \delta]. \qquad (2.3.12)$$

Now we are in a position to repeat the arguments for (2.3.11) and conclude by (2.3.12) and Gronwall's inequality, that $x(t) = y(t)$ for $t \in (\alpha, \alpha + \delta]$, violating the minimality of α. This shows that $x = y$ on A. Using (H2) it is easily seen that $a = b$ on A.

This completes the proof of Theorem 1. $\square$

2.4 Maximal Solutions and a Semiflow on M_0

Given $\Psi = \begin{pmatrix} \psi \\ \varphi \end{pmatrix} \in M_0$, the local solution granted in the previous section can be extended to a unique maximal solution $\begin{pmatrix} x^\Psi \\ a^\Psi \end{pmatrix}$ of (2.2.1) in M_0 defined for $t \in I \cup [0, t_e)$ for some $0 < t_e \leq \infty$ which depends on Ψ. Namely,

$t_e = \sup\{\tau \in (0, \infty) \mid$ (2.2.1) has a solution $\begin{pmatrix} x \\ a \end{pmatrix}$ on $I \cup [0, \tau]$ in M_0, with $\begin{pmatrix} x_0 \\ a_0 \end{pmatrix} = \begin{pmatrix} \psi \\ \varphi \end{pmatrix}\}$. In this section we discuss the semiflow on M_0 formed by these maximal solutions of (2.2.1) in M_0.

We first introduce some notations. Let $\Omega = \{(t, \Psi) \in [0, \infty) \times M_0 | t \in [0, t_e(\Psi))\}$. For $t \geq 0$, let $\Omega_t = \{\Psi \in M_0 | t < t_e(\Psi)\} \subset M_0$. Then $\Omega \subset \mathbf{R} \times C(I, X_0 \times \mathbf{R})$ and $\Omega_t \subset C(I, X_0 \times \mathbf{R})$. Both Ω and Ω_t are equipped with the relative topology. Define $S : \Omega \to M_0$ as

$$S(t, \Psi) = \begin{pmatrix} x_t^\Psi \\ a_t^\Psi \end{pmatrix} \qquad \text{for } (t, \Psi) \in \Omega.$$

Theorem 2 *The map S is a continuous semiflow on M_0. That is, S is continuous and satisfies the following two properties:*

(i) $S(0, \Psi) = \Psi$ *for* $\Psi \in M_0$.

(ii) *For each* s, $t \geq 0$ *with* $s < t_e(\Psi)$ *and* $t < t_e(S(s,\Psi))$, *we have* $t + s < t_e(\Psi)$ *and* $S(t, S(s,\Psi)) = S(t+s,\Psi) \in M_0$.

Proof. Properties (i) and (ii) are straightforward. It suffices to show that S is continuous. This is done in three steps, where Step 2 and Step 3 are merely adapting the corresponding proofs in [36] to our framework.

Step 1: Let $\Psi \in M_0$. We show that there is $\tau > 0$ and a neighborhood U of Ψ in M_0 such that $[0,\tau] \times U \subset \Omega$ and the restriction $S|_{[0,\tau]\times U}$ is continuous.

We take $0 < R_2 < R_1 < R_0$ such that $||\Psi|| < R_2$. Denote $R = (R_1 - R_2)/M$, where $M \geq 1$ is as in (H1). Let $\Phi = \begin{pmatrix} \phi^1 \\ \phi^2 \end{pmatrix} \in M_0$ such that $||\Phi - \Psi|| < R$. Denote the corresponding mild solution of Φ in M_0 as $\begin{pmatrix} x(t) \\ a(t) \end{pmatrix}$. Then, for $t \in [0, t_e(\Phi))$,

$$x(t) = T(t)\phi^1(0) + \lim_{\mu\to\infty} \int_0^t T(t-s)\mu R_\mu F(x(s), a(s))ds.$$

It follows that

$$|x(t)| \leq |T(t)(\phi^1(0) - \psi^1(0))| + |T(t)\psi^1(0)| + \int_0^t M^2 e^{\omega(t-s)} |F(x(s), a(s))|ds.$$

Observing that

$$\begin{aligned} |F(x(s), a(s))| &\leq |F(x(s), a(s)) - F(\psi^1(0), \psi^2(0))| + |F(\psi^1(0), \psi^2(0))| \\ &\leq 2D(R_0 + h) + |F(\psi^1(0), \psi^2(0))| \end{aligned}$$

and setting $C(\Psi) = 2D(R_0 + h) + |F(\psi^1(0), \psi^2(0))|$, we get

$$|x(t)| \leq (R_1 - R_2)e^{\omega t} + |T(t)\psi^1(0)| + tM^2 e^{\omega t} C(\Psi).$$

By continuity, there is some $\tau > 0$ such that, for any $\Phi \in M_0$ with $||\Phi - \Psi|| < R$, we have $|x(t)| < R_1$ for each $t \in [0, \tau] \cap [0, t_e(\Phi))$. Thus if $t \in [0, \tau] \cap [0, t_e(\Phi))$ then $||x_t|| < R_1$. In particular, this shows that $\tau < t_e(\Phi)$. Let U be the open ball of radius $R > 0$ about Ψ in M_0. We have shown $[0, \tau] \times U \subset \Omega$.

Now suppose we are given $(t_0, \Phi_0) \in [0, \tau] \times U$. Then for each $(t, \Phi) \in [0, \tau] \times U$, $||S(t, \Phi) - S(t_0, \Phi_0)|| \leq ||S(t, \Phi) - S(t, \Phi_0)|| + ||S(t, \Phi_0) - S(t_0, \Phi_0)||$. To complete the proof of Step 1, it is now clear that it suffices to show that the first term on the right hand side of the latter inequality is bounded by $c||\Phi - \Phi_0||$ for some constant $c > 0$ uniformly for $t \in [0, \tau]$.

Let $x(t), a(t)$ correspond to Φ_0 and $y(t), b(t)$ correspond to Φ. Then we have that for each $t \in [0, \tau]$

$$|x(t) - y(t)| \leq Me^{\omega\tau}||\Phi - \Phi_0|| + \int_0^t M^2 e^{\omega(t-s)} D(|x(s) - y(s)| + |a(s) - b(s)|)ds$$

and $|a(t) - b(t)| \leq L_{R_1}||x_t - y_t|| + f(R_1)||a_t - b_t||$. It is not difficult to see that the latter inequality implies

$$\sup_{-h \leq t+\theta \leq t}|a(t+\theta) - b(t+\theta)| \leq c(\sup_{-h \leq t+\theta \leq t}|x(t+\theta) - y(t+\theta)| + ||\Phi - \Phi_0||)$$

for some constant $c > 0$ depending on R_1. This information combined with a Gronwall inequality argument completes the proof of Step 1.

Step 2. Let $\Psi \in M_0$ and $t \in [0, t_e(\Psi))$. We show that $\Omega_t \subset M_0$ is open and the map $\Omega_t \ni \Phi \mapsto S(t, \Phi)$ is continuous at Ψ.

By continuity, we see that the set $K_1 = \{S(s, \Psi) \mid s \in [0, t]\} \subset M_0$ is compact. Therefore, applying Step 1, we find some $u > 0$ and some open subset N in M_0 containing K_1 such that $[0, u] \times N \subset \Omega$ and $S|_{[0,u]\times N}$ is continuous. Let J be the smallest positive integer such that $t/J < u$. Obviously, $(J-1)u \leq t < Ju$. Given $\epsilon > 0$, we find $\delta_1 > 0$ such that

$$\begin{aligned} &\text{if } ||\gamma - S((J-1)u, \Psi)|| < \delta_1 \text{ then } \gamma \in N \text{ and} \\ &||S(t-(J-1)u, S((J-1)u, \Psi)) - S(t-(J-1)u, \gamma)|| < \epsilon. \end{aligned} \tag{2.4.1}$$

Recursively we can find $\delta_j > 0$ for $j = 2, \ldots, J$ such that

$$\begin{aligned} &\text{if } ||\gamma - S((J-j)u, \Psi)|| < \delta_j \text{ then } \gamma \in N \text{ and} \\ &||S(u, \gamma) - S(u, S((J-j)u, \Psi))|| < \delta_{j-1}. \end{aligned} \tag{2.4.2}$$

Using (2.4.1), (2.4.2), the semigroup property, and induction, we see that if $\Phi \in M_0$ with $||\Phi - \Psi|| < \delta_J$ then $\Phi \in \Omega_t$ and $||S(t, \Phi) - S(t, \Psi)|| < \epsilon$. This completes the proof of Step 2.

Step 3. We prove that the map $S : \Omega \to M_0$ is continuous.

For $(t_0, \Psi_0) \in \Omega$, let U be a neighborhood of $S(t_0, \Psi_0)$ in M_0. We want to find a neighborhood $W \subset \Omega$ of (t_0, Ψ_0) such that $S(W) \subset U$. If $t_0 = 0$, by Step 1, we are done. Otherwise, $t_0 > 0$. By Step 1, we find some $0 < u < t_0$ and a neighborhood W_1 of Ψ_0 in M_0 such that $[0, u] \times W_1 \subset \Omega$ and $S|_{[0,u]\times W_1}$ is continuous. Let $0 < u_1 < u$. It follows from $S(t_0, \Psi_0) = S(t_0 - u_1, S(u_1, \Psi_0))$ that $S(u_1, \Psi_0) \in \Omega_{t_0 - u_1}$. By Step 2, we can find a neighborhood W_2 of $S(u_1, \Psi_0)$ in M_0 such that $S(t_0 - u_1, W_2) \subset U$. Take $0 < \delta < u_1$ such that $(u_1 - \delta, u_1 + \delta) \subset (0, u)$ and choose a neighborhood W_3 of Ψ_0 in M_0 with $S((u_1 - \delta, u_1 + \delta) \times W_3) \subset W_2$. If $s \in (t_0 - \delta, t_0 + \delta)$ then $s = (t_0 - u_1) + (s - t_0 + u_1)$ and therefore the semigroup property gives $S((t_0 - \delta, t_0 + \delta) \times W_3) \subset U$, which completes the proof. □

2.5 The Model Equations: Part One

In this section we present an application of the general theory. We will see that in practice, it is non-trivial to check that all of the relevant hypotheses are satisfied.

Consider the following class of scalar age structured models with threshold dependent age of maturity,

$$\begin{cases} \partial_t u(t,a) + \partial_a u(t,a) = -d(a)u(t,a), \\ u(t,0) = b(\int_{\tau(t)}^{m} \beta(\xi)u(t,\xi)d\xi), \\ \int_{t-\tau(t)}^{t} [\int_0^{\tau(\sigma)} u(\sigma,a)da + C]^{-1} d\sigma = T, \\ \begin{pmatrix} u_0 \\ \tau_0 \end{pmatrix} = \begin{pmatrix} \hat{\psi} \\ \hat{\varphi} \end{pmatrix} \in C([-a_m, 0], (L^1{}_+[0,m)) \times \mathbf{R}^+), \end{cases} \tag{2.5.1}$$

where $t \geq 0$, $0 \leq a < m$, and $a_m < m \leq \infty$. Here m represents the maximum age and a_m stands for the maximum juvenile age. We make the following assumptions:

(A1) $d : [0, m) \to \mathbf{R}^+$ and $\beta : [0, m) \to \mathbf{R}^+$ are essentially bounded.

(A2) $b : \mathbf{R}^+ \to \mathbf{R}^+$ is bounded, globally Lipschitz, and $0 < \max_{x\in\mathbf{R}^+} b(x) \leq \theta$ for some $\theta > 0$.

(A3) $a_m = (R_0 + C)T < m \leq \infty$, where $R_0 = C(\frac{1}{\sqrt{T\theta}} - 1) > 0$.

In order to apply Theorem 1, we rewrite (2.5.1) as follows. Let $X = \mathbf{R} \times L^1([0,m), \mathbf{R})$ and define $A : D(A) \to X$ by

$$A \begin{pmatrix} 0 \\ x \end{pmatrix} = \begin{pmatrix} -x(0) \\ -x' \end{pmatrix} \qquad \text{for} \begin{pmatrix} 0 \\ x \end{pmatrix} \in D(A) = \{0\} \times W^{1,1}([0,m), \mathbf{R}).$$

Note that $X_0 = \overline{D(A)} = \{0\} \times L^1[0,m)$. It is well known that A satisfies (H1) (see, for instance, [34, 19]). Denote

$$C_0 = \left\{ \begin{pmatrix} 0 \\ \gamma \end{pmatrix} \in 0 \times L^1[0,m) | 0 \leq \gamma(a) \leq \theta \; a.e. \; a \in [0,m) \right\}$$

and

$$D(H) = \left\{ \begin{pmatrix} \psi \\ \varphi \end{pmatrix} \in C([-a_m, 0], C_0 \times K) \mid ||\psi|| \leq R_0 \right\},$$

where $K = [0, a_m] \subset \mathbf{R}$. We prove that our "age of maturity function" is well defined in the following result.

Lemma 1 *The relation $H : D(H) \to K$, which is given by $(\psi, \varphi, \alpha) \in H$ if and only if $\int_{-\alpha}^{0} [\int_0^{\varphi(\sigma)} \psi(\sigma, \xi) d\xi + C]^{-1} d\sigma = T$, is a function.*

Proof. Given $\begin{pmatrix} \psi \\ \varphi \end{pmatrix} \in D(H)$, it suffices to show that there exists a unique $\alpha \in K$ such that $(\psi, \phi, \alpha) \in H$. In fact, note that the map $\alpha \mapsto \int_{-\alpha}^{0}[\int_{0}^{\varphi(\sigma)} \psi(\sigma,\xi)d\xi + C]^{-1}d\sigma$ defined for $\alpha \in [0, a_m]$ is strictly increasing and continuous. Moreover, $\int_{-a_m}^{0}[\int_{0}^{\varphi(\sigma)} \psi(\sigma,\xi)d\xi + C]^{-1}d\sigma \geq a_m/(R_0 + C) = T$. Now the result follows immediately. □

Define $f : \mathbf{R}^+ \to \mathbf{R}^+$ by $f(Q) = \frac{(Q+C)^2 T}{C^2}\theta$. The coming result tells us that H satisfies an appropriate Lipschitz condition.

Lemma 2 *For any $Q > 0$ there is some $L_Q > 0$ such that, for* $\begin{pmatrix} \psi_1 \\ \varphi_1 \end{pmatrix}, \begin{pmatrix} \psi_2 \\ \varphi_2 \end{pmatrix} \in D(H)$ *with $||\psi_i|| \leq Q$ $(i = 1\ , 2)$, we have*

$$|H(\psi_1, \varphi_1) - H(\psi_2, \varphi_2)| \leq\ f(Q)\ ||\varphi_1 - \varphi_2|| + L_Q||\psi_1 - \psi_2||.$$

Proof. Let $t_1 = H(\psi_1, \varphi_1)$ and $t_2 = H(\psi_2, \varphi_2)$. Without loss of generality, assume that $t_1 \leq t_2$. Then we have

$$\int_{-t_1}^{0}\left[\int_{0}^{\varphi_1(\sigma)} \psi_1(\sigma,\xi)d\xi + C\right]^{-1} d\sigma - \int_{-t_2}^{0}\left[\int_{0}^{\varphi_2(\sigma)} \psi_2(\sigma,\xi)d\xi + C\right]^{-1} d\sigma = 0$$

or

$$\int_{-t_2}^{-t_1}\left[\int_0^{\varphi_2(\sigma)}\psi_2(\sigma,\xi)d\xi + C\right]^{-1} d\sigma \;=\; \int_{-t_1}^{0}\left[\left(\int_0^{\varphi_1(\sigma)}\psi_1(\sigma,\xi)d\xi + C\right)^{-1} - \left(\int_0^{\varphi_2(\sigma)}\psi_2(\sigma,\xi)d\xi + C\right)^{-1}\right] d\sigma.$$

Using the fact that the function $u \mapsto 1/(u+C)$ is globally Lipschitz on $(0,\infty)$ with Lipschitz constant $1/C^2$, we get

$$\frac{|t_1-t_2|}{Q+C} \le \frac{1}{C^2}\int_{-t_1}^{0}\left|\int_0^{\varphi_1(\sigma)}\psi_1(\sigma,\xi)d\xi - \int_0^{\varphi_2(\sigma)}\psi_2(\sigma,\xi)d\xi\right| d\sigma.$$

It follows that $|t_1-t_2| \le (Q+C)t_1(\ ||\psi_1-\psi_2|| + \theta||\varphi_1-\varphi_2||\ )/C^2$. Since $t_1 \le (Q+C)T$, we obtain

$$|H(\psi_1,\varphi_1) - H(\psi_2,\varphi_2)| \le \frac{(Q+C)^2T}{C^2}||\psi_1-\psi_2|| + \frac{(Q+C)^2T}{C^2}\theta||\varphi_1-\varphi_2||.$$

This completes the proof. $\square$

Define $F : C_0 \times K \to X$ by $F(x,a) = \begin{pmatrix} b(\int_a^m \beta(\xi)x(\xi)d\xi) \\ -d(\cdot)x(\cdot) \end{pmatrix}$. By (A1) and (A2) it is clear that F is Lipschitz on $C_0 \times K$. The verification of the subtangential condition (H6) with respect to C_0, K, and F follows exactly as on pages 12–14 of the examples in [34]. Therefore, by Theorem 1, we have the following result.

Proposition 3 *In addition to (A1)–(A3), assume that* $\begin{pmatrix} \hat{\psi} \\ \hat{\varphi} \end{pmatrix} \in C([-a_m,0], L^1_+[0,m)\times \mathbf{R}^+)$ *satisfies the following two conditions:*

(i) *For each* $\sigma \in [-a_m, 0]$, $0 \leq \hat{\psi}(\sigma)(a) \leq \theta$ *a.e.* $a \in [0, m)$ *and* $\hat{\varphi}(\sigma) \in [0, a_m]$.

(ii) *For each* $\sigma \in [-a_m, 0]$, $\int_0^m \hat{\psi}(\sigma)(a)da < C(\frac{1}{\sqrt{T\theta}} - 1)$ *and* $\int_{-\hat{\varphi}(0)}^0 [\int_0^{\hat{\varphi}(\sigma)} \hat{\psi}(\sigma, \xi)d\xi + C]^{-1}d\sigma = T$.

Then the initial value problem (2.5.1) has a unique maximal solution $\begin{pmatrix} u \\ \tau \end{pmatrix} \in C([-a_m, t_e), L^1_+[0, m) \times [0, a_m])$ *on* $[-a_m, t_e)$ $(t_e > 0)$ *with* $\begin{pmatrix} u_0 \\ \tau_0 \end{pmatrix} = \begin{pmatrix} \hat{\psi} \\ \hat{\varphi} \end{pmatrix}$ *in the following sense:*

(i) *For* $0 \leq t < t_e$, $a \mapsto \int_0^t u(s, a)ds$ *is absolutely continuous, and for a.e.* $a \in [0, m)$,

$$\begin{aligned} u(t, a) &= u(0, a) - \partial_a \int_0^t u(s, a)ds - \int_0^t d(a)u(s, a)ds, \\ \int_0^t u(s, 0)ds &= \int_0^t b\left(\int_{\tau(s)}^m \beta(a)u(s, a)da\right) ds. \end{aligned}$$

(ii) *For* $0 \leq t < t_e$, $\int_{t-\tau(t)}^t [\int_0^{\tau(\sigma)} u(\sigma, a)da + C]^{-1}d\sigma = T$.

(iii) *For* $t \in [0, t_e)$ *the "total population" satisfies* $\int_0^m u(t, a)da < C(\frac{1}{\sqrt{T\theta}} - 1)$ *and* $0 \leq u(t, a) \leq \theta$ *for a.e.* $a \in [0, m)$.

Finally we note that, by Theorem 2, the corresponding semiflow is continuous. The reader should be warned that under the present hypothesis it is not necessarily true that the obtained semiflow is global, i.e. $t_e = \infty$. In case for each $t \in [0, t_e)$

$||u_t|| \leq R_1$ for some $R_1 < R_0$ then $t_e = \infty$. On one hand this is a limitation of the present framework, on the other hand it has the advantage that the delay is a priori bounded.

3 Differentiability of the Abstract Solution Semiflow with Respect to Initial Data

This chapter is devoted to finding general conditions under which the abstract semiflow from Chapter 2 is also differentiable with respect to initial data, in a suitable weak sense. We commence with Section 3.1 explaining the technical difficulties of the linearization problem for differential equations containing state-dependent delays, why our results are not contained in the existing literature, as well as why some well known approaches fail to resolve the linearization problem stemming from the model equations derived in the Introduction. Subsection 3.1.1 provides an outline for the main results of this chapter. This chapter is self contained and can be read without reading Chapter 2, if desired.

3.1 Background

A fundamental problem in the study of dynamical systems concerns the linearization of a flow or a semiflow along a trajectory. When the flow is induced by an ordinary differential equation (ODE) on $\mathbf{R}^n$ with a smooth nonlinearity, this problem is straightforward and the derivative of the semiflow with respect to initial data is given by the solution of the corresponding linearized system along flowlines. For semiflows on infinite dimensional spaces such as these given by solutions of certain nonlinear parabolic equations or solutions of delay differential equations with constant delays, this problem is merely an extrapolation of the finite dimensional ODE case with the help of an abstract variation of constants formula (see, for example, [33, 34]). This is possible because the nonlinearity appearing in the relevant equation is continuously differentiable on the appropriate function space and one can proceed to obtain the differentiability of the corresponding semiflow relying on Gronwall's inequality. It is well known from the works [9, 23, 31, 35, 36, 37, 14, 4, 22] that even ODEs containing a state dependent delay such as $x'(t) = x(t - x(t))$ do not fit into the standard frameworks for functional differential equations in [3, 7, 39]. The reason is that the nonlinear term is not differentiable (or even not Lipschitz!) on the commonly used phase space of continuous functions. In particular, the corresponding initial value problem is not well-posed on this phase space. A resolution for this problem is to restrict the phase space to a subset of the continuously differentiable functions so that the nonlinearity is continuously differentiable on it and to exploit the fact that its derivative has a bounded extension to the original space of continuous functions and this extension satisfies a componentwise continuity property. This weaker type of differentiability (with respect to the supremum norm from the space of continuous functions), sometimes called *almost Fréchet differentiability* as in [23] or more appropriately *extendable continuous differentiability* as in [31], is sufficient to obtain a continuously differentiable semiflow on a submanifold of the space of continuously differentiable functions for a class of equations including the one given above (see [37]).

Despite their emergence in the modeling of structured populations with developmental stages of variable length (see [11, 26, 32]), there are very few works dealing with differential equations containing both state-dependent delays and partial differential operators. The works [28, 29] deal with special classes of reaction diffusion systems containing state-dependent delays, but use special assumptions to circumvent the difficulties mentioned above.

Recall our model for a population structured by age with distinct juvenile and adult stages, and with a variable age of maturity. It is assumed that juveniles and adults are not competing for resources. As a result the model equations take the form:

$$\begin{cases} \partial_t u(t,a) + \partial_a u(t,a) = -d(a)u(t,a), \\ u(t,0) = b(\int_{\tau(t)}^{m} \beta(\xi)u(t,\xi)d\xi), \\ \int_{t-\tau(t)}^{t} [\int_0^{\tau(\sigma)} u(\sigma,a)da + C]^{-1} d\sigma = T, \\ \begin{pmatrix} u_0 \\ \tau_0 \end{pmatrix} = \begin{pmatrix} \psi \\ \varphi \end{pmatrix} \in C([-a_m,0], {L^1}_+[0,m) \times \mathbf{R}^+). \end{cases} \tag{3.1.1}$$

(See Section 1.2 for a detailed derivation.) Here $t \geq 0$, $0 \leq a < m$, $a_m < m \leq \infty$, and $0 < \tau(t) \leq a_m$ represents the variable age of maturity. The parameter $T > 0$ represents a resource concentration density threshold, m represents the maximum age, a_m the maximum juvenile age, and $C([-a_m,0], {L^1}_+[0,m) \times \mathbf{R}^+)$ denotes the space of continuous functions on $[-a_m,0]$ having values in ${L^1}_+[0,m) \times \mathbf{R}^+$. The natural setting for age structured population models is $L^1[0,m)$ since the total population at a given time is given by the L^1 norm of the population density. It was shown in Chapter 2 that under suitable hypotheses, the second component of system (3.1.1) can be written as an algebraic-delay equation and that system (3.1.1) can be written abstractly as:

$$\begin{aligned} \frac{d}{dt}\begin{pmatrix} 0 \\ u(t,\cdot) \end{pmatrix} &= \begin{pmatrix} -u(t,0) \\ -u_a(t,\cdot) \end{pmatrix} + \begin{pmatrix} b(\int_{\tau(t)}^{m} \beta(\xi)u(t,\xi)d\xi) \\ -d(\cdot)u(t,\cdot) \end{pmatrix}, \\ \tau(t) &= H(u_t,\tau_t), \\ \begin{pmatrix} x_0 \\ \tau_0 \end{pmatrix} &= \begin{pmatrix} \psi \\ \varphi \end{pmatrix} \in M_0. \end{aligned} \tag{3.1.2}$$

Here M_0 is a "nonlinear" subset of the ambient space of continuous functions induced by the algebraic component (For precise definitions of M_0 and H see Section 3.2 and Section 3.7). It was also shown that the abstract system gives rise to a continuous semiflow on M_0 via $S\left(t, \begin{pmatrix} \psi \\ \varphi \end{pmatrix}\right) = \begin{pmatrix} x_t \\ a_t \end{pmatrix}$, where $x_t = u_t(\cdot) \in C([-a_m, 0], L^1([0, m), \mathbf{R}^+))$ and $a_t = \tau_t \in C([-a_m, 0], \mathbf{R}^+)$. In this chapter we establish sufficient conditions which ensure that this semiflow is also continuously differentiable in a suitable weak sense.

Without going into too many technical details, we list reasons (R1)–(R3) below why the issue of differentiability of the semiflow induced by the above system is not addressed in existing works. For system (3.1.2), let $F : L^1([0, m), \mathbf{R}^+) \times [0, m) \to \mathbf{R}^+ \times L^1([0, m), \mathbf{R}^+)$ be given by $F(x(\cdot), a) = \begin{pmatrix} b(\int_a^m \beta(\xi)x(\xi)d\xi) \\ -d(\cdot)x(\cdot) \end{pmatrix}$. Then (3.1.2) has the form (3.2.1) (see Page 53). For the purpose of illustration, we take $b : \mathbf{R}^+ \to \mathbf{R}^+$ to be the identity mapping and assume that both β, $d : [0, m) \to \mathbf{R}^+$ are the constant function with value 1.

(R1) Poor smoothness properties of nonlinearities. Here $D_2F(x(\cdot), a) = \begin{pmatrix} -x(a) \\ 0 \end{pmatrix}$ and $D_1F(x(\cdot), a)\gamma = \begin{pmatrix} \int_a^m \gamma(\xi)d\xi \\ -\gamma(\cdot) \end{pmatrix}$. It is easy to see that $D_2F(x(\cdot), a)$ is not defined for general $x \in L^1[0, m)$ and that $D_1F(x(\cdot), a) \in \mathcal{L}(L^1[0, m), \mathbf{R} \times L^1[0, m))$ is not continuous. Even if x is continuous, although it can be shown that the partial derivatives of F exist at (x, a), F will *not* be differentiable with respect to the norm from $L^1[0, m) \times \mathbf{R}$. This means, in particular, that we cannot apply the results of e.g. [30] or [34] even indirectly since they require continuous differentiability of the nonlinear term, albeit on possibly thin subsets in [30]. Similarly, it will be seen in Section 3.7 that the other nonlinearity H has a similar lack of smoothness on the space $C(I, L^1[0, m) \times \mathbf{R})$.

(R2) Classical change of variables. In the work of Smith [32] on ODEs containing a threshold type state-dependent delay such as the one we have here, a change of variables is employed to reduce the system to one with a constant delay. Formally, employing such a transformation to system (3.1.1) amounts to setting $z(t) := \int_0^t [\int_0^{\tau(\sigma)} u(\sigma, a)da + C]^{-1} d\sigma$. Clearly, $z(t)$ is invertible. Denote $w(t, a) := u(z^{-1}(t), a)$ and $c(t) := \tau(z^{-1}(t))$. Then, after differentiating $c(t)$, the new system with a constant delay is given by

$$\begin{cases} w_t(t,a) + (\int_0^{c(t)} w(t,\xi)d\xi + C)w_a(t,a) = -d(a)(\int_0^{c(t)} w(t,\xi)d\xi + C)w(t,a), \\ w(t,0) = b(\int_{c(t)}^m \beta(\xi)w(t,\xi)d\xi), \\ c'(t) = \int_0^{c(t)} w(t,\xi)d\xi - \int_0^{c(t-T)} w(t-T,\xi)d\xi. \end{cases}$$

Although this (larger) system has a constant delay, it suffers from the same lack of smoothness given in (R1). Additionally, $w_a(t, a)$ is multiplied by an integral nonlinearity.

(R3) Monotonicity of $t \mapsto t - \tau(t)$. Other works including [12, 32] on differential or integral equations containing threshold type delays exploit monotonicity of the function $t \mapsto t - \tau(t)$, where $\tau(t)$ is the variable transition age in question. For instance, this property was used implicitly in (R2). We *do not* use monotonicity in our rendition for several reasons:

- For the problem at hand, it is not clear how the analysis can be simplified by using the monotonicity property even if an explicit representation formula is available such as for system (3.1.1) via the method of characteristics.
- One can construct systems which do not enjoy this property but can otherwise be included in the present framework.
- As we will show, the monotonicity property is not necessary to obtain the desired result on differentiability.

To obtain the desired result on differentiability, the problems caused by the poor smoothness of the nonlinearities are circumvented in an analogous fashion to existing works for ODEs with state-dependent delays. However, for the model equations (3.1.1), in contrast to the ODE case we will see that the appearance of the partial differential operator ∂_a also plays a key role.

3.1.1 Outline and Main Results

Although our results are of a more general nature, for clarity, we outline the structure of this chapter in terms of the model equations (3.1.1). The main goal of this chapter is to prove Theorem 11 in Section 3.6.

In Section 3.2, we cover the basic functional analytic preliminaries and the precise meaning of mild solutions of system (3.1.1). This functional analytic setup is captured in the way the first equation along with the nonlinear boundary condition for $u(t, 0)$ in system (3.1.1), is rewritten in system (3.1.2). This setup was motivated by the studies [19, 20, 34]. Moreover, the 'subtangential condition' (H5) adopted from [34] ensures that the population density remains non-negative for non-negative initial data.

In Section 3.3 we address the differentiability with respect to time of solutions of system (3.1.1). The existence and uniqueness of solutions of system (3.1.1) and continuity of the corresponding semiflow was established in [13]. Due to the poor smoothness properties of the nonlinearities discussed above, the methods used to obtain the differentiability of solutions must differ from the standard techniques from, e.g. [27]. This is where the assumption involving the Radon-Nikodym property in (H1) comes into play. We finally obtain a positively invariant set for the semiflow, denoted $\hat{M}_0$, on which every trajectory is C^1 in time and for which the population density $u(t, \cdot)$ is absolutely continuous. The set $\hat{M}_0$ is analogous to the infinitesimal generator for system (3.1.1).

Let $W^{1,1}[0, m)$ denote the space of absolutely continuous functions in $L^1[0, m)$ whose derivative lies in $L^1[0, m)$ (which, in the motivating example, contains the population density $u(t, \cdot)$). In Section 3.4 we show that the set $\hat{M}_0$ is a C^1 submanifold of the space $C([-a_m, 0], W^{1,1}[0, m) \times \mathbf{R})$. We show that $\hat{M}_0$ has an atlas of manifold charts whose derivatives have the special extension properties discussed above. In particular, for each $p \in \hat{M}_0$, we show that the tangent space $T_p\hat{M}$ which is a subspace of $C([-a_m, 0], W^{1,1}[0, m) \times \mathbf{R})$ has an extension to the larger function space $C([-a_m, 0], L^1[0, m) \times \mathbf{R})$.

In Section 3.5 we show that the (formal) linear variational system along flow lines in $\hat{M}_0$ can be solved uniquely for mild solutions for initial data belonging to the corresponding extended tangent space.

Section 3.6 develops the main results of this chapter. We show that the solution operators $\hat{S}_t$ at time t (whose domain is the set of initial data in $\hat{M}_0$ whose maximal interval of existence is bigger than t) are differentiable. Here the derivative at a point $p \in \hat{M}_0$ is a linear operator whose domain is the interpolation space $T_p\hat{M}^1 := T_p\hat{M} \cap C^1([-a_m, 0], L^1[0, m) \times \mathbf{R})$ and whose codomain is the space $C([-a_m, 0], L^1[0, m) \times \mathbf{R})$. Additionally, it is shown that the derivative map $d\hat{S}_t$ which is defined on an appropriate subset of the tangent bundle is continuous.

Finally, in Section 3.7, all of the relevant hypotheses are verified for the motivating example system (3.1.1) and some abstract results are used to infer about the regularity of its solutions.

3.1.2 Morally Finite or Infinite Dimensional Problem?

As mentioned above, in many cases the techniques used to obtain the differentiability of a semiflow with respect to initial data, which arises from some type of autonomous differential equation on an infinite dimensional phase space, having a smooth nonlinearity, are a glorification of the same techniques used in the case of an ODE. To bridge the gap, enough knowledge of functional analysis to manipulate an abstract variation of constants formula suffices. The same is *not* true for the model equations (3.1.1). We illustrate some reasons below.

The nonlinearity $F(x(\cdot), a) = \begin{pmatrix} \int_a^m x(\xi)d\xi \\ -x(\cdot) \end{pmatrix}$ can be written as a sum $F = F_1 + F_2$, where $F_1(x(\cdot), a) = \begin{pmatrix} \int_a^m x(\xi)d\xi \\ 0 \end{pmatrix}$ and $F_2(x(\cdot), a) = \begin{pmatrix} 0 \\ -x(\cdot) \end{pmatrix}$. The trouble maker is clearly F_1. Although $D_1F_1(x(\cdot), a) \in \mathcal{L}(L^1[0, m), \mathbf{R})$ exists, the map $L^1[0, m) \times [0, m) \ni (x(\cdot), a) \mapsto D_1F_1(x(\cdot), a) \in \mathcal{L}(L^1[0, m), \mathbf{R})$ is not continuous. However, it is easily checked that F_1 is C^1 on the smaller set $W^{1,1}[0, m) \times [0, m)$, where $D_1F_1(x(\cdot), a) \in \mathcal{L}(W^{1,1}[0, m), \mathbf{R})$, and $W^{1,1}[0, m)$ has the norm $|\gamma|_{W^{1,1}} =$

$|\gamma(0)| + |\gamma|_{L^1} + |\gamma'|_{L^1}$ for $\gamma \in W^{1,1}[0,m)$.

A key result used to obtain the differentiability of the corresponding semiflow with respect to initial data is

$$\begin{aligned} & |[D_1F_1(x^p(s), a^{p+\xi}(s)) - D_1F_1(x^p(s), a^p(s))](x^{p+\xi}(s) - x^p(s))| \\ = & \ |\int_{a^p(s)}^{a^{p+\xi}(s)} x^{p+\xi}(s)(\theta) - x^p(s)(\theta) d\theta| \\ \leq & \ |a^{p+\xi}(s) - a^p(s)| \ |x^{p+\xi}(s) - x^p(s)|_{W^{1,1}[0,m)}. \end{aligned}$$

Here $x^{p+\xi}(s)(\cdot)$, $x^p(s)(\cdot)$, $a^{p+\xi}(s)$, $a^p(s)$ denote the first and second components of solutions of (3.1.1) in $\hat{M}_0$ at time s corresponding to initial data $p+\xi$ and p, respectively (See Step 5 in the proof of Theorem 11).

In order to obtain the desired differentiability result, we need

$$|[D_1F_1(x^p(s), a^{p+\xi}(s)) - D_1F_1(x^p(s), a^p(s))](x^{p+\xi}(s) - x^p(s))| = o(\xi)$$

as $\xi \to 0$ for each s. Since it will turn out that $|a^{p+\xi}(s) - a^p(s)| = O(||\xi||)$, we require that, for each s, $|x^{p+\xi}(s) - x^p(s)|_{W^{1,1}[0,m)} \to 0$ as $\xi \to 0$. Here $||\cdot||$ denotes the supremum norm on the space $C([-a_m, 0], L^1[0,m) \times \mathbf{R})$. For even further illustration, we note that letting $s = 0$ gives us the requirement that

$$\begin{aligned} & |[D_1F_1(x^p(0), a^{p+\xi}(0)) - D_1F_1(x^p(0), a^p(0))](x^{p+\xi}(0) - x^p(0))| \\ \leq & \left| \int_{p_2(0)}^{p_2(0)+\xi_2(0)} \xi_1(0)(\theta) d\theta \right| = o(\xi) \qquad \text{as } \xi \to 0, \end{aligned}$$

where the subscripts 1 and 2 denote the first and second components of the initial data. Note that it is impossible for the latter to hold merely as $||\xi|| \to 0$. We can only expect this to hold as $\xi \to 0$ with respect to the supremum norm on the space $C([-a_m, 0], W^{1,1}[0, m) \times \mathbf{R})$. Namely, the supremum norm which includes a contribution from the partial differential operator ∂_a in system (3.1.1), since the right hand side is bounded by $||\xi||\ |\xi_1(0)|_{W^{1,1}[0,m)}$. That having been said, another key result is showing that $|x^{p+\xi}(s) - x^p(s)|_{W^{1,1}[0,m)} \to 0$ as $\xi \to 0$ with respect to this stronger norm. This part is given in Step 4 of the proof of Theorem 11 in Section 3.6, which is achieved by showing that $||\dot{x}_s^{p+\xi} - \dot{x}_s^p|| \to 0$ as $||\xi|| + ||\xi'|| \to 0$ for each s. Here the prime denotes differentiation with respect to time, not age. This leads us to the interpolation space $C^1([-a_m, 0], L^1[0, m) \times \mathbf{R}) \cap C([-a_m, 0], W^{1,1}[0, m) \times \mathbf{R})$. Due to the presence of the delay in system (3.1.1), we need to consider the latter interpolation space with the norm containing contributions from both the time derivative and the partial differential operator ∂_a in system (3.1.1) to obtain the desired differentiability of the semiflow with respect to initial data, which is Theorem 11 in Section 3.6. We will see in Section 3.7 that the partial derivative D_1H of the other nonlinearity H has similar properties as D_1F_1 above.

3.2 Technical Preliminaries and Hypotheses

In this section we state the relevant technical preliminaries and hypotheses. All Banach spaces are assumed to be over the real numbers. Whenever a product of Banach spaces is considered, we view it as a Banach space equipped with the corresponding product norm.

3.2.1 The Ambient Linear Space of Initial Data

Let $0 < \delta < \infty$ and $I = [-\delta, 0]$. For $F \subset E$, where E is a Banach space, $C(I, F)$ denotes the set of continuous functions mapping I into F. For $\psi \in C(I, F)$, we let $||\psi||$ be the supremum norm of ψ. Then $(C(I, E), ||\cdot||)$ is a Banach space. Similarly, we let $C^1(I, F)$ be the set of continuously differentiable functions mapping I into F. If $\delta = \infty$ and $I = (-\infty, 0]$, we let $BUC(I, F)$ denote the set of bounded uniformly continuous functions mapping I into F and similarly $BUC(I, E)$ is a Banach space when equipped with the supremum norm.

Suppose that $0 < T < \infty$ and $y : I \cup [0, T] \to F$ is a map. As usual in the literature on delay equations, for each $t \in [0, T]$, we define $y_t : I \to F$ by $y_t(\theta) = y(t + \theta)$ for $\theta \in I$ and call y_t the history of y at time t. If $T = \infty$ then the same definition applies with $t \in [0, T]$ being replaced with $t \in [0, T)$.

3.2.2 Hypotheses

(H1) Let $(X, |\cdot|)$ denote a Banach space. Suppose that $A : D(A) \to X$ with $D(A) \subset X$ is a linear operator satisfying the estimates of the Hille-Yosida theorem, that is, there is some $M \geq 1$ and some $\omega \in \mathbf{R}$ such that the ray $(\omega, \infty) \subset \rho(A)$ and $||(A - \lambda I)^{-n}|| \leq \frac{M}{(\lambda-\omega)^n}$ for $\lambda > \omega$ and for each positive integer n. In order to derive reasonable regualirty properties of solutions of our system, we further assume that X has the direct sum decomposition, $X = X_1 \oplus X_2$, where X_1 and X_2 are closed subspaces and X_1 has the Radon-Nikodym property, that is, given an open subset $O \subset \mathbf{R}$, every Lipschitz map $g : O \to X_1$ is a.e. differentiable.

Let $X_0 = \overline{D(A)}$ and A_0 denote the part of A in X_0. Actually this class of operators falls under a more general class of well known operators as pointed out in [30]. Set $R_\lambda = (A - \lambda I)^{-1}$. Without loss of generality, assume that $\omega > 0$. It follows from (H1) that A_0 generates a C^0-semigroup of linear operators $\{T(t)\}_{t \geq 0}$ on X_0 and satisfies $||T(t)|| \leq M e^{\omega t}$.

(H2) Let n be a given positive integer. Suppose that K is some compact subset of $\mathbf{R}^n$ such that K is contained in the closed ball of radius $h > 0$ centered at the origin. Set $I = [-h, 0] \subset \mathbf{R}$. Let C_0 be some closed and convex subset of X_0. Assume that there is some $R_0 > 0$, a strictly increasing function $f : \mathbf{R}^+ \to \mathbf{R}^+$ with $f(R_0) = 1$, and a function $H : D(H) \to K$ satisfying the following Lipschitz condition: for each $Q > 0$, there is some $L_Q > 0$ such that, for $\begin{pmatrix} \psi_1 \\ \varphi_1 \end{pmatrix}, \begin{pmatrix} \psi_2 \\ \varphi_2 \end{pmatrix} \in D(H)$ with $||\psi_i|| \leq Q$ $(i = 1, 2)$, we have

$$|H(\psi_1, \varphi_1) - H(\psi_2, \varphi_2)| \leq \; f(Q)\, ||\varphi_1 - \varphi_2|| + L_Q ||\psi_1 - \psi_2||,$$

where $D(H) = \left\{ \begin{pmatrix} \psi \\ \varphi \end{pmatrix} \in C(I, C_0 \times K) \mid ||\psi|| \leq R_0 \right\}$ (Both norms of the spaces X and $\mathbf{R}^n$ will be denoted by $|\cdot|$ since this should not cause any confusion).

(H3) Let $M_0 = \left\{ \begin{pmatrix} \psi \\ \varphi \end{pmatrix} \in D(H) \mid \varphi(0) = H(\psi, \varphi) \text{ and } ||\psi|| < R_0 \right\}$. Assume $M_0 \neq \emptyset$.

Give $D(A)$ the graph norm and view $C(I, D(A) \times \mathbf{R}^n)$ as a Banach space. We assume that $C_0 \cap D(A) \neq \emptyset$. Let $D(\hat{H}) := D(H) \cap C(I, D(A) \times \mathbf{R}^n)$. We let the function $\hat{H}$ with domain $D(\hat{H})$ be the restriction of H to $D(\hat{H})$.

Remark. When $D(H)$ and $D(\hat{H})$ are respectively given the relative topology from $C(I, X_0 \times \mathbf{R}^n)$ and $C(I, D(A) \times \mathbf{R}^n)$, we have the continuous inclusions,

$$\begin{array}{ccc} D(H) & \rightarrow & C(I, X_0 \times \mathbf{R}^n) \\ \uparrow & & \uparrow \\ D(\hat{H}) & \rightarrow & C(I, D(A) \times \mathbf{R}^n) \end{array}$$

(H4) Suppose $F : C_0 \times K \rightarrow X$ has the form $F(c, k) = F_1(c, k) + F_2(c)$, where $F_1 : C_0 \times K \rightarrow X_1$ and $F_2 : C_0 \rightarrow X_2$. We assume that F_1 is globally Lipschitz (there is some $D > 0$ such that, for c_1, $c_2 \in C_0$ and k_1, $k_2 \in K$, we have $|F_1(c_1, k_1) - F_1(c_2, k_2)| \leq D(|c_1 - c_2| + |k_1 - k_2|)$) and that F_2 is continuously differentiable on C_0 (for each $c \in C_0$, there is a bounded linear operator $DF_2(c) : X_0 \to X_2$ which satisfies $\lim_{\xi \to 0,\, c+\xi \in C_0,\, \xi \in X_0} \frac{|F_2(c+\xi) - F_2(c) - DF_2(c)\xi|}{|\xi|} = 0$ and the map $C_0 \ni c \mapsto DF_2(c)$ is continuous with respect to the uniform operator topology). We also assume that $\sup_{c \in C_0} ||DF_2(c)|| < \infty$ so that F_2 is globally Lipschitz on C_0 since C_0 is convex. Note that it follows that F is also globally Lipschitz.

(H5) (Subtangential Condition) We assume that, for each $(c, k) \in C_0 \times K$,

$$\lim_{h \downarrow 0} \frac{\text{dist}\left(T(h)c + \lim\limits_{\mu \to \infty} \int_0^h T(s)\mu R_\mu F(c,k)ds, C_0 \right)}{h} = 0$$

holds. Here $\text{dist}(x, B) = \inf_{b \in B} |x - b|$ for $x \in X$ and $B \subset X$.

Definition. Consider the following initial value problem,

$$\begin{cases} x'(t) = Ax(t) + F(x(t), a(t)), \\ a(t) = H(x_t, a_t), \\ \begin{pmatrix} x_0 \\ a_0 \end{pmatrix} = \begin{pmatrix} \psi \\ \varphi \end{pmatrix} \in M_0. \end{cases} \tag{3.2.1}$$

By a mild solution of (3.2.1) on $I \cup [0, T]$ in M_0 with $0 < T < \infty$, we mean a pair of functions $\begin{pmatrix} x(t) \\ a(t) \end{pmatrix}$ with the following properties:

(i) $a : I \cup [0, T] \to K$ is continuous.

(ii) $x : I \cup [0,T] \to C_0$ is continuous such that, for each $t \in [0,T]$, $\int_0^t x(s)ds \in D(A)$ and

$$x(t) = x(0) + A\int_0^t x(s)ds + \int_0^t F(x(s), a(s))ds.$$

(iii) For $0 \le t \le T$, $\begin{pmatrix} x_t \\ a_t \end{pmatrix} \in M_0$, *i.e.*, $a(t) = H(x_t, a_t)$ and $||x_t|| < R_0$.

(iv) $\begin{pmatrix} x_0 \\ a_0 \end{pmatrix} = \begin{pmatrix} \psi \\ \varphi \end{pmatrix}$.

We similarly define mild solutions in M_0 on $I \cup [0,T)$ for $T = \infty$.

Note that (H1) implies that (ii) is equivalent to

$x(t) = T(t)\psi(0) + \lim_{\mu\to\infty} \int_0^t T(t-s)\mu R_\mu F(x(s), a(s))ds$ for $t \in [0,T]$. See [34].

3.3 Differentiability of Solutions with respect to Time

Under the assumptions (H1)–(H5), given $\begin{pmatrix} \psi \\ \varphi \end{pmatrix} \in M_0$, there is some $t_e > 0$ such that (3.2.1) has a unique maximal mild solution on $I \cup [0, t_e)$ in M_0 (see Chapter 2 and [13]). In this section we discuss the differentiability of these mild solutions with respect to time. In Theorem 4 below, we give sufficient conditions under which mild solutions are locally Lipschitz in time. This result is used to derive Theorem 5, which gives sufficient conditions for the C^1 smoothness of the x component. Finally, with the aid of an additional hypothesis, we derive sufficient conditions for the C^1 smoothness of the a component in Theorem 6. We end Section 3.2 by identifying a positively invariant set for the corresponding solution semiflow $S(\cdot,\cdot)$, called $\hat{M}_0$, on which every trajectory is C^1 in time.

Theorem 4 *Suppose that (H1)–(H5) hold. Given* $\begin{pmatrix} \psi \\ \varphi \end{pmatrix} \in M_0$, *let* $\begin{pmatrix} x(t) \\ a(t) \end{pmatrix}$ *denote the corresponding (maximal) mild solution on* $I \cup [0, t_e)$ *in* M_0. *If* $\psi(0) \in D(A)$ *and* $\begin{pmatrix} \psi \\ \varphi \end{pmatrix}$ *is Lipschitz on* I, *then* $\begin{pmatrix} x(t) \\ a(t) \end{pmatrix}$ *is locally Lipschitz on* $I \cup [0, t_e)$ *and* $a : I \cup [0, t_e) \to K$ *is differentiable almost everywhere.*

Proof. Fix $T \in [0, t_e)$. Choose $0 < R_1 < R_0$ such that $||x_t|| \leq R_1$ for $t \in [0, T]$. Denote the trivial extensions of x and a to $(-\infty, T]$ respectively by $\hat{x}$ and $\hat{a}$, that is, $\hat{a}(\xi) = \begin{cases} \varphi(-h) & \text{if } \xi \leq -h, \\ a(\xi) & \text{if } \xi \in [-h, T] \end{cases}$ and $\hat{x}(\xi) = \begin{cases} \psi(-h) & \text{if } \xi \leq -h, \\ x(\xi) & \text{if } \xi \in [-h, T]. \end{cases}$ Note that, for each $t \in [0, T]$, $\hat{x}_t$ and $\hat{a}_t$ are members of the Banach spaces $BUC((-\infty, 0], X_0)$ and $BUC((-\infty, 0], \mathbf{R}^n)$, respectively. Moreover, $Lip(\hat{x}_0) = Lip(\psi)$, $Lip(\hat{a}_0) = Lip(\varphi)$. The proof is done in the following four steps.

Step 1. $A \int_0^h T(s)\psi(0)ds = \lim_{\mu\to\infty} \int_0^h T(h-s)\mu R_\mu A\psi(0)ds.$

This follows easily from Lemma 1.8 of [34].

Step 2. There is $L > 0$ (depending possibly on T) such that $||\hat{x}_s - \hat{x}_0|| \leq Ls$ and $||\hat{a}_s - \hat{a}_0|| \leq Ls$ for each $s \in [0, T]$.

Let $s \in [0, T]$ be given. The result in Step 1 combined with $T(s)\psi(0) - \psi(0) = A \int_0^s T(\xi)\psi(0)d\xi$ implies that $|x(s) - x(0)| \leq Cs$ for some $C > 0$. Note that C may depend on T. Then, for each $\theta \leq 0$, we have $|\hat{x}(s+\theta) - \hat{x}(\theta)| \leq |x(s+\theta) - x(0)| + |x(0) - \hat{x}(\theta)| \leq Cs + Lip(\psi)(-\theta) \leq Cs + Lip(\psi)s$ if $s + \theta \geq 0$ and $|\hat{x}(s+\theta) - \hat{x}(\theta)| \leq Lip(\psi)s$ if $s+\theta < 0$. These observations imply that $||\hat{x}_s - \hat{x}_0|| \leq Ls$ for some $L > 0$ depending on T. Now we turn to $\hat{a}$. If $s + \theta < 0$ then $|\hat{a}(s+\theta) - \hat{a}(\theta)| \leq Lip(\varphi)s$; if $s+\theta \geq 0$, then by (H2) there is some constant $J > 0$ depending on R_1 such that $|\hat{a}(s+\theta) - \hat{a}(\theta)| \leq |a(s+\theta) - a(0)| + |a(0) - \hat{a}(\theta)| \leq J||x_{s+\theta} - x_0|| + f(R_1)||a_{s+\theta} - a_0|| + Lip(\varphi)(-\theta)$. With $||x_{s+\theta} - x_0|| \leq ||\hat{x}_{s+\theta} - \hat{x}_0|| \leq Ls$ and $||a_{s+\theta} - a_0|| \leq ||\hat{a}_s - \hat{a}_0|| + Lip(\varphi)s$, we get $|\hat{a}(s+\theta) - \hat{a}(\theta)| \leq Qs + f(R_1)||\hat{a}_s - \hat{a}_0||$ for some $Q > 0$ depending on T. By virtue of $f(R_1) < 1$ we can conclude that $||\hat{a}_s - \hat{a}_0|| \leq Ls$ for a possibly larger constant L than the one found before.

Step 3. There is an $L > 0$ such that, for each t, $h \in [0, T]$ with $t + h \leq T$, we have $||\hat{a}_{t+h} - \hat{a}_t|| \leq L(||\hat{x}_{t+h} - \hat{x}_t|| + h)$.

Given $\theta \leq 0$, if $t+h+\theta \leq 0$ then $|\hat{a}(t+h+\theta) - \hat{a}(t+\theta)| \leq Lip(\varphi)h$; if $t+h+\theta \geq 0$ and $t + \theta < 0$ then using the result in Step 2 we have $|\hat{a}(t+h+\theta) - \hat{a}(t+\theta)| \leq |a(t+h+\theta) - a(0)| + |a(0) - \hat{a}(t+\theta)| \leq L|t+h+\theta| + Lip(\varphi)h \leq Lh + Lip(\varphi)h =$

$(L + Lip(\varphi))h$; if $t + h + \theta \geq 0$ and $t + \theta \geq 0$ then $|\hat{a}(t + h + \theta) - \hat{a}(t + \theta)| \leq J||x_{t+h+\theta} - x_{t+\theta}|| + f(R_1)||a_{t+h+\theta} - a_{t+\theta}|| \leq J||\hat{x}_{t+h} - \hat{x}_t|| + f(R_1)||\hat{a}_{t+h} - \hat{a}_t||$. The required result is now obvious since $f(R_1) < 1$.

Step 4. There is an $L > 0$ such that, for each t, $h \in [0, T]$ with $t + h \leq T$, we have $||\hat{x}_{t+h} - \hat{x}_t|| \leq Lh$ and $||\hat{a}_{t+h} - \hat{a}_t|| \leq Lh$.

For each $t \in [0, T]$, we have

$$\begin{aligned}
& x(t+h) - x(t) \\
= \; & (T(t+h) - T(t))\psi(0) + \lim_{\mu\to\infty} \int_0^h T(t+h-s)\mu R_\mu F(x(s), a(s))ds \\
& + \lim_{\mu\to\infty} \int_0^t T(t-s)\mu R_\mu (F(x(s+h), a(s+h)) - F(x(s), a(s)))ds \\
= \; & T(t)(T(h)\psi(0) - \psi(0) + \lim_{\mu\to\infty} \int_0^h T(h-s)\mu R_\mu F(x(s), a(s))ds) \\
& + \lim_{\mu\to\infty} \int_0^t T(t-s)\mu R_\mu (F(x(s+h), a(s+h)) - F(x(s), a(s)))ds.
\end{aligned}$$

Using the result in Step 1 and the fact that $T(h)\psi(0) - \psi(0) = A\int_0^h T(s)\psi(0)ds$, we obtain

$$\begin{aligned}
& x(t+h) - x(t) \\
= \; & T(t)\left(\lim_{\mu\to\infty} \int_0^h T(h-s)\mu R_\mu (F(x(s), a(s)) + A\psi(0))ds\right) \\
& + \lim_{\mu\to\infty} \int_0^t T(t-s)\mu R_\mu (F(x(s+h), a(s+h)) - F(x(s), a(s)))ds.
\end{aligned}$$

Then there exists $C > 0$, depending on T, such that, for each $t \in [0, T]$,

$$\begin{aligned} & |x(t+h) - x(t)| \\ \leq\ & C\left(h + \int_0^t e^{\omega(t-s)}(|x(s+h) - x(s)| + |a(s+h) - a(s)|)ds\right) \\ \leq\ & C(h + \int_0^t e^{\omega(t-s)}(||\hat{x}_{s+h} - \hat{x}_s|| + ||\hat{a}_{s+h} - \hat{a}_s||)ds). \end{aligned}$$

Now the result in this step follows from that in Step 3 and an application of Gronwall's inequality. The statement of the Theorem follows from Step 4 and Rademacher's theorem (see [5]) applied to the function a.

Theorem 5 *Suppose (H1)–(H5) hold. Let* $\begin{pmatrix} \psi \\ \varphi \end{pmatrix} \in M_0$ *and let* $\begin{pmatrix} x(t) \\ a(t) \end{pmatrix}$ *denote the corresponding (maximal) mild solution on* $I \cup [0, t_e)$ *in* M_0*. If* $\psi(0) \in D(A)$*,* $A\psi(0) + F(\psi(0), \varphi(0)) \in X_0$*, and* $\begin{pmatrix} \psi \\ \varphi \end{pmatrix}$ *is Lipschitz on* I*, then, for each* $t \in [0, t_e)$*,* $x(t)$ *is continuously differentiable,* $x(t) \in D(A)$*,* $x'(t) = Ax(t) + F(x(t), a(t)) \in X_0$*, and* $a : I \cup [0, t_e) \to K$ *is differentiable almost everywhere.*

Proof. Fix $T \in (0, t_e)$. By Theorem 4, we know that both $x(t)$ and $a(t)$ are Lipschitz on $[0, T]$. Therefore, the function $[0, T] \ni t \mapsto F_1(x(t), a(t)) \in X_1$ is also Lipschitz and hence almost everywhere differentiable since X_1 has the Radon-Nikodym property. Let $g(t) = d/dtF_1(x(t), a(t))$. Consider the non-autonomous initial value problem

$$\begin{cases} w'(t) = Aw(t) + g(t) + DF_2(x(t))w(t), & t \in [0,T], \\ w(0) = Ax(0) + F(x(0), a(0)) \in X_0, \end{cases}$$

which has a unique (continuous) mild solution $w(t)$ on $[0,T]$. By Theorem 1.9 of [34], we know that x is right differentiable at zero since $x(0) \in D(A)$ and $Ax(0) + F(x(0), a(0)) \in X_0$. With standard arguments, we can finish the proof.

In order to derive C^1-smoothness of a, we make the following hypothesis, which is also crucial for the main theorem of this chapter in Section 3.6.

(H6) Equip $D(A)$ with the graph norm. Assume that there is an open subset U of the Banach space $C(I, D(A) \times \mathbf{R}^n)$ such that $D(\hat{H}) \subset U$ and $\hat{H} : D(\hat{H}) \to K$ has a continuously differentiable extension (in the Fréchet sense) to a map $H_e : U \to K$ with $D_1 H_e(\psi, \varphi) \in \mathcal{L}(C(I, D(A)), \mathbf{R}^n)$ having rank n. We further assume that, for each $(\psi, \varphi) \in U$, the partial derivative $D_1 H_e(\psi, \varphi) \in \mathcal{L}(C(I, X_0), \mathbf{R}^n)$ exists as a *relative Fréchet derivative* on U (note the larger space) and that the map $U \times C(I, X_0) \ni (\psi, \varphi, \gamma) \mapsto D_1 H_e(\psi, \varphi)\gamma \in \mathbf{R}^n$ is continuous, where U inherits the topology from $C(I, D(A) \times \mathbf{R}^n)$.

Remarks. Hypothesis (H6) deserves some remarks.

(i) By "relative Fréchet derivative on U", we mean that

$$\lim_{\xi \to 0,\ \xi \in C(I, X_0),\ (\psi+\xi, \varphi) \in U} \frac{|H_e(\psi + \xi, \varphi) - H_e(\psi, \varphi) - D_1 H_e(\psi, \varphi)\xi|}{|\xi|_{C(I, X_0)}} = 0$$

for $(\psi, \varphi) \in U$.

(ii) For each $(\psi, \varphi) \in U$, we can extend $DH_e(\psi, \varphi) \in \mathcal{L}(C(I, D(A) \times \mathbf{R}^n), \mathbf{R}^n)$ to a linear operator $DH_e^1 \in \mathcal{L}(C(I, X_0 \times \mathbf{R}^n), \mathbf{R}^n)$ using the fact that $D_1 H_e(\psi, \varphi)$ has such an extension. Moreover, the map $U \times C(I, X_0 \times \mathbf{R}^n) \ni (\psi, \varphi, \gamma) \mapsto DH_e^1(\psi, \varphi)\gamma \in \mathbf{R}^n$ is continuous.

(iii) We will drop the subscript 'e' and the superscript '1' from now on.

(iv) The second extension property of the derivative of the function H appearing above is analogous to those appearing in [35, 36] and in (H7) below.

Theorem 6 *Suppose (H1)–(H6) hold. Let* $\begin{pmatrix} \psi \\ \varphi \end{pmatrix} \in D(\hat{H}) \cap M_0 \cap C^1(I, C_0 \times K)$ *with* $\psi'(0) = A\psi(0) + F(\psi(0), \varphi(0))$ *and* $\varphi'(0) = D_1 H(\psi, \varphi)\psi' + D_2 H(\psi, \varphi)\varphi'$. *Then the corresponding maximal mild solution on* $I \cup [0, t_e)$ *in* M_0, $\begin{pmatrix} x(t) \\ a(t) \end{pmatrix}$, *satisfies* $\begin{pmatrix} x_t \\ a_t \end{pmatrix} \in D(\hat{H}) \cap M_0 \cap C^1(I, C_0 \times K)$, $x'(t) = Ax(t) + F(x(t), a(t))$, *and* $a'(t) = D_1 H(x_t, a_t)x_t' + D_2 H(x_t, a_t)a_t'$ *for each* $t \in [0, t_e)$.

Proof. The fact that $x_t \in C(I, D(A))$ for $t \in [0, t_e)$ follows from Theorem 5 since $[0, t_e) \ni t \mapsto x'(t) \in X_0$ is continuous and $Ax(t) = x'(t) - F(x(t), a(t))$ implies that $x \in C(I \cup [0, t_e), D(A))$. It remains to prove that a is C^1 on $I \cup [0, t_e)$ and that its derivative is in fact given by the formula above. To this end, fix $0 < T < t_e$. Let $R_1 = \max_{s\in[0,T]}||x_s|| < R_0$. The equation

$$b(s) = \begin{cases} \varphi'(s) & \text{if } s \in I \\ D_1H(x_s, a_s)x'_s + D_2H(x_s, a_s)b_s & \text{if } s \in [0, T] \end{cases}$$

has a unique continuous solution, $b : I \cup [0, T] \to \mathbf{R}^n$, thanks to the contraction mapping principle since $||D_2H(x_s, a_s)|| \leq f(R_1) < 1$ for each $s \in [0, T]$. We will show that $a'(t) = b(t)$ for each $t \in [0, T]$. Firstly, for $t \in [0, T]$ and $r \in [0, h]$ such that $t + r \leq T$, we have

$$\begin{aligned} & a(t + r) - a(t) - rb(t) \\ = \;& H(x_{t+r}, a_{t+r}) - H(x_t, a_{t+r}) + H(x_t, a_{t+r}) - H(x_t, a_t) - rb(t) \\ = \;& D_1H(x_t, a_{t+r})(x_{t+r} - x_t) - D_1H(x_t, a_t)rx'_t + \omega_1(x_{t+r} - x_t, x_t, a_{t+r}) \\ & +D_2H(x_t, a_t)(a_{t+r} - a_t - rb_t) + \omega_2(a_{t+r} - a_t, x_t, a_t), \end{aligned} \tag{3.3.1}$$

where $\omega_1 : \Omega_1 \to \mathbf{R}^n$ and $\omega_2 : \Omega_2 \to \mathbf{R}^n$ are the remainder terms. Here $\Omega_1 = \{(\xi, \beta, \chi) \in C(I, X_0) \times C(I, D(A)) \times C(I, \mathbf{R}^n) \mid (\beta, \chi) \in U \text{ and } (\beta + \xi, \chi) \in U\}$ and Ω_2 is given similarly. It follows from (H6) that ω_2 is continuous on Ω_2 where Ω_2 inherits the relative topology from $C(I, \mathbf{R}^n) \times C(I, D(A)) \times C(I, \mathbf{R}^n)$ (Note carefully how the assumption concerning the relative partial Fréchet derivative from (H6) is used). By (H6) the function $g : [0, T] \to \mathbf{R}^n$ given by $g(s) = H(x_s, a_{t+r})$ is C^1 with

$g'(s) = D_1H(x_s, a_{t+r})x'_s$. Moreover,

$$
\begin{aligned}
& \omega_1(x_{t+r} - x_t, x_t, a_{t+r}) \\
= \;& g(t+r) - g(t) - D_1H(x_t, a_{t+r})(x_{t+r} - x_t) \\
= \;& g'(t)r + \left(\int_0^1 g'(t+sr) - g'(t)ds)r - D_1H(x_t, a_{t+r})(x_{t+r} - x_t \right) \\
= \;& D_1H(x_t, a_{t+r})(x'_t r - x_{t+r} + x_t) \\
& + \left(\int_0^1 D_1H(x_{t+sr}, a_{t+r})x'_{t+sr} - D_1H(x_t, a_{t+r})x'_t ds \right) r.
\end{aligned}
$$

Note that $\lim_{r\to 0} \frac{\omega_1(x_{t+r}-x_t, x_t, a_{t+r})}{r} = 0$. Secondly, it follows from Theorem 4 that a is Lipschitz on $I \cup [0, T]$ and hence it is clear that $\lim_{r\to 0} \frac{\omega_2(a_{t+r}-a_t, x_t, a_t)}{r} = 0$. The proof is finished in the coming two steps, where we will use the notation $j = o(k)$ for functions j and k to mean $\lim_{x\to 0} \frac{|j(x)|}{|k(x)|} = 0$.

Step 1. $\lim_{r\to 0,\ 0<r\le \min\{T,h\}} \frac{||a_r - a_0 - rb_0||}{r} = 0$.

Let $\theta \in [-h, 0]$. First, if $r + \theta \le 0$, then

$$
\begin{aligned}
|a(r+\theta) - a(\theta) - rb(\theta)| &= |\varphi(r+\theta) - \varphi(\theta) - r\varphi'(\theta)| \\
&\le \int_0^1 |\varphi'(\theta + sr) - \varphi'(\theta)| ds\ r \\
&\le \max_{-h\le\xi\le -r} \int_0^1 |\varphi'(\xi + sr) - \varphi'(\xi)| ds\ r \\
&= o(r).
\end{aligned}
$$

Next, if $r + \theta > 0$, then

$$
|a(r+\theta) - a(\theta) - rb(\theta)| \le I_1 + I_2,
$$

where

$$I_1 = |a(r+\theta) - \varphi(0) - (r+\theta)\varphi'(0)|$$

and

$$I_2 = |\varphi(0) + (r+\theta)\varphi'(0) - \varphi(\theta) - r\varphi'(\theta)|.$$

We have

$$\begin{aligned} I_2 &= |\varphi(0) + (r+\theta)\varphi'(0) - \varphi(\theta) - r\varphi'(\theta)| \\ &= |\varphi(\theta) - \varphi(0) - \varphi'(0)\theta + r(\varphi'(\theta) - \varphi'(0))| \\ &\leq \max_{-r\leq\xi\leq 0} |\varphi(\xi) - \varphi(0) - \varphi'(0)\xi| + r \max_{-r\leq\xi\leq 0} |\varphi'(\xi) - \varphi'(0)| \\ &= o(r). \end{aligned}$$

For I_1, using (3.3.1) for $t = 0$ with r being replaced by $r+\theta$ and the continuity of $D_1H(\psi,\varphi)\gamma$ in $(\psi,\varphi,\gamma) \in U \times C(I, X_0)$ from (H6), we obtain

$$\begin{aligned} I_1 &= |a(r+\theta) - \varphi(0) - (r+\theta)\varphi'(0)| \\ &\leq o(r+\theta) + |\omega_1(x_{r+\theta} - x_0, x_0, a_{r+\theta})| \\ &\quad + f(R_1)||a_{r+\theta} - a_0 - (r+\theta)b_0|| + |\omega_2(a_{r+\theta} - a_0, x_0, a_0)|. \end{aligned}$$

Note that $\sup_{-r<\xi\leq 0} \frac{|o(r+\xi)|}{r} \to 0$ as $r \to 0$. Since $|\omega_1(x_{r+\theta} - x_0, x_0, a_{r+\theta})| = o(r+\theta)$ and $|\omega_2(a_{r+\theta} - a_0, x_0, a_0)| = o(r+\theta)$, it follows that

$$I_1 = f(R_1)||a_{r+\theta} - a_0 - (r+\theta)b_0|| + o(r).$$

Let $K_0 := \{(r, \theta) \in \mathbf{R}^2 \mid r + \theta \geq 0, r \in [0, T] \cap [0, h],\ \ \theta \in [-h, 0]\}$ and note that the compactness of K_0 and continuity of the function $K_0 \ni (r, \theta) \mapsto ||a_{r+\theta} - a_0 - (r + \theta)b_0|| \in \mathbf{R}$ implies that we can find $(r^*, \theta^*) \in K_0$ which maximizes this function. Hence, collectively, we can conclude that, for each $r \in (0, h] \cap (0, T]$, $||a_r - a_0 - rb_0|| \leq o(r) + f(R_1)||a_{r^*+\theta^*} - a_0 - (r^* + \theta^*)b_0||$. As $f(R_1) < 1$, it is clear that $||a_r - a_0 - rb_0|| = o(r)$ as desired.

Step 2. For each $t \in [0, T)$, $\lim_{r\to 0,\, t+r\leq T,\, 0<r\leq h} \frac{||a_{t+r}-a_t-rb_t||}{r} = 0$.

Let $\theta \in [-h, 0]$. If either $t + r + \theta \leq 0$ or $t + r + \theta > 0$ with $t + \theta \leq 0$, then by Step 1 we have

$$\begin{aligned} |a(t + r + \theta) - a(t + \theta) - rb(t + \theta)| &= |a_r(t + \theta) - a_0(t + \theta) - rb_0(t + \theta)| \\ &\leq ||a_r - a_0 - rb_0|| \\ &= o(r). \end{aligned}$$

Now, if $t + \theta > 0$, then it follows from (3.3.1) with $t + \theta$ replacing t that

$$\begin{aligned} & |a(t + r + \theta) - a(t + \theta) - rb(t + \theta)| \\ \leq\ & |D_1H(x_{t+\theta}, a_{t+\theta+r})(x_{t+\theta+r} - x_{t+\theta}) - D_1H(x_{t+\theta}, a_{t+\theta})rx'_{t+\theta}| \\ & +|\omega_1(x_{t+\theta+r} - x_{t+\theta}, x_{t+\theta}, a_{t+\theta+r})| + f(R_1)||a_{t+\theta+r} - a_{t+\theta} - rb_{t+\theta}|| \\ & +|\omega_2(a_{t+\theta+r} - a_{t+\theta}, x_{t+\theta}, a_{t+\theta})|. \end{aligned}$$

It follows from continuity that there is $(t^*, \theta^*) \in \{(s, \xi) \mid s \in [0, T]$ and $\xi \in [-s, 0] \cap [-h, 0]\}$ such that the maximum in (t, θ) of the right hand side of the above inequality is achieved at (t^*, θ^*). Then, for $s \in [0, T)$ with $s + r \leq T$ and $r \in (0, h]$, we have

$$||a_{s+r} - a_s - rb_s|| \leq o(r) + f(R_1)||a_{t^*+\theta^*+r} - a_{t^*+\theta^*} - rb_{t^*+\theta^*}||$$

(Note that $o(r)$ does not depend on s). Applying this to $s = t + \theta^* \geq 0$ and using the fact that $f(R_1) < 1$, we obtain $||a_{t+r} - a_t - rb_t|| = o(r)$ as desired. This completes the proof.

We remark that although in general, a satisfies what is called a neutral differential equation, for the concrete example given in the introduction and in Section 3.7, this will turn out to be merely an ordinary differential equation with a state-dependent delay.

The following result follows immediately from Theorem 6.

Corollary 7 *Suppose (H1)–(H6) hold. The set* $\hat{M}_0 := \left\{ \begin{pmatrix} \psi \\ \varphi \end{pmatrix} \in D(\hat{H}) \cap M_0 \cap C^1(I, C_0 \times K) \,\middle|\, \psi'(0) = A\psi(0) + F(\psi(0), \varphi(0)) \text{ and } \varphi'(0) = D_1 H(\psi, \varphi)\psi' + D_2 H(\psi, \varphi)\varphi' \right\}$ *is a positively invariant subset of* M_0 *for the semiflow* S.

3.4 The semiflow on M_0 and its Restriction to $\hat{M}_0$

In this section we briefly discuss the semiflow on M_0 and the smaller positively invariant set $\hat{M}_0$.

Denote the semiflow induced by maximal mild solutions of (3.1.1) in M_0 by $S : \Omega \to M_0$, where $\Omega := \{(t, \Psi) \in [0, \infty) \times M_0 \mid t < t_e(\Psi)\}$. The fact that S is a semiflow and is continuous with respect to the relative topologies from $\mathbf{R} \times C(I, X_0 \times \mathbf{R}^n)$ and $C(I, X_0 \times \mathbf{R}^n)$, respectively, is established in [13]. Let $\hat{\Omega} := \Omega \cap ([0, \infty) \times \hat{M}_0) = \{(t, \Psi) \in [0, \infty) \times \hat{M}_0 \mid t < t_e(\Psi) \text{ and } \Psi \in \hat{M}_0\}$. Define $\hat{S} := S|_{\hat{\Omega}}$. The coming lemma is immediate from the fact that S is a semiflow on M_0 (see Section (2.4)).

Lemma 3 *The map $\hat{S} : \hat{\Omega} \to \hat{M}_0$ has the semigroup property, that is,*

(i) $\hat{S}(0, \Psi) = \Psi$ *for each* $\Psi \in \hat{M}_0$*;*

(ii) *If* $\Psi \in \hat{M}_0$ *and* $0 \leq s, t$ *with* $s < t_e(\Psi)$ *and* $t < t_e(\hat{S}(s, \Psi))$*, then* $t + s < t_e(\Psi)$ *and* $\hat{S}(t, \hat{S}(s, \Psi)) = \hat{S}(t + s, \Psi)$.

Next we introduce notations for the solution operators. Let $\Omega_t := \{\Psi \in M_0 \mid t < t_e(\Psi)\}$ and $S_t : \Omega_t \to M_0$ be given by $S_t(\Psi) := S(t, \Psi)$. Similarly, $\hat{\Omega}_t := \Omega_t \cap \hat{M}_0$ and $\hat{S}_t : \hat{\Omega}_t \to \hat{M}_0$ is given by $\hat{S}_t := S_t|_{\hat{\Omega}_t}$.

Let $\hat{M} := \left\{ \begin{pmatrix} \psi \\ \varphi \end{pmatrix} \in U \mid \varphi(0) = H(\psi, \varphi) \right\}$, where U is given in (H6) (see Page 59).

Let $D(A)$ be given the graph norm. We can turn $C(I, D(A) \times \mathbf{R}^n)$ into a C^1-Banach manifold by assigning it the standard C^1-smooth structure. For Banach manifolds, we refer readers to the book by Lang [17].

Proposition 8 *Suppose (H6) holds. The set $\hat{M}$ is a C^1-Banach submanifold of $C(I, D(A) \times \mathbf{R}^n)$ of codimension n. For each $p \in \hat{M}$, the tangent space at p, $T_p\hat{M}$, is given by the kernel of the derivative of the map $U \ni (\psi, \varphi) \mapsto \varphi(0) - H(\psi, \varphi)$ at the point p.*

Proof. Let $J(\psi, \varphi) = \varphi(0) - H(\psi, \varphi)$ for $(\psi, \varphi) \in U$. Then J is C^1 on U. Fix $(\psi_0, \varphi_0) \in U$. We have

$$DJ(\psi_0, \varphi_0)(\gamma_1, \gamma_2) = \gamma_2(0) - D_1H(\psi_0, \varphi_0)\gamma_1 - D_2H(\psi_0, \varphi_0)\gamma_2$$

for $(\gamma_1, \gamma_2) \in C(I, D(A) \times \mathbf{R}^n)$. Let $e_1, \ldots, e_n$ form a basis of $\mathbf{R}^n$. For each j, we set $\gamma_2 = 0 \in \mathbf{R}^n$ and (by (H6)) choose γ_1 such that $D_1H(\psi_0, \varphi_0)\gamma_1 = -e_j$. Then $DJ(\psi_0, \varphi_0)(\gamma_1, \gamma_2) = e_j$. This shows that $DJ(\psi_0, \varphi_0)$ is surjective. Therefore, we have the decomposition $C(I, D(A) \times \mathbf{R}^n) = \ker(DJ(\psi_0, \varphi_0)) \oplus N$ for some n-dimensional subspace N such that $DJ(\psi_0, \varphi_0)|N$ is an isomorphism. Hence we can write $(\psi_0, \varphi_0) = k_0 + n_0$ for $k_0 \in \ker DJ(\psi_0, \varphi_0)$ and $n_0 \in N$ and $J(k_0 + n_0) = 0$. We can find relatively open neighborhoods U_1 of k_0 in the subspace $\ker(DJ(\psi_0, \varphi_0))$ and V_1 of n_0 in the subspace N such that $U_1 + V_1 \subset U$. Define $\tilde{J} : U_1 \times V_1 \to \mathbf{R}^n$ by $\tilde{J}(k', n') = J(k' + n')$. Since $D_2\tilde{J}(k_0, n_0)$ is an isomorphism, the implicit function theorem gives relatively open sets U_0 in the subspace $\ker DJ(\psi_0, \varphi_0)$ and V_0 in the subspace N with $(k_0, v_0) \in U_0 \times V_0$, and a C^1 map $h : U_0 \to V_0$ satisfying $J(k' + n') = 0$ for $(k', n') \in U_0 \times V_0$ if and only if $n' = h(k')$. It follows that $U_0 + V_0$ is an open neighborhood of (ψ_0, φ_0). Let $\beta : U_0 + V_0 \to K \times N$ be given by $\beta(k' + n') = (k', h(k') - n')$. Observe that β is a C^1 homeomorphism and satisfies $\beta((U_0 + V_0) \cap \hat{M}) = U_0 \times \{0\}$. It is not difficult to verify the statement concerning the tangent space at (ψ_0, φ_0).

Let us make some comments about the special manifold charts above and the tangent spaces. In light of Proposition 8, we have $T_{(\psi,\varphi)}\hat{M} = \{(\gamma_1, \gamma_2) \in C(I, D(A) \times \mathbf{R}^n) \mid \gamma_2(0) = D_1H(\psi, \varphi)\gamma_1 + D_2H(\psi, \varphi)\gamma_2\}$ for each $(\psi, \varphi) \in \hat{M}$. Note that, by (H6), $T_{(\psi,\varphi)}\hat{M}$ has an extension to the larger space $C(I, X_0 \times \mathbf{R}^n)$, which we call $T_{(\psi,\varphi)}M$ and is given by the same formula. For each $p \in \hat{M}$, we can find ambient-open sets $U_0 \subset T_p\hat{M}$ and V_0 such that $(U_0 + V_0) \cap \hat{M}$ is a neighborhood of p in $\hat{M}$, and a chart whose inverse is a map $g : U_0 \to U_0 + V_0$ given by $g(k) = k + h(k)$, where $h : U_0 \to V_0$ is C^1. Since $J(k + h(k)) = 0$ for $k \in U_0$, differentiating with respect to 'k' yields $DJ(k + h(k))(1_{T_p\hat{M}} + Dh(k)) = 0$. Thus $Dh(k) = -(DJ(k+h(k))|N_p)^{-1}DJ(k+h(k))1_{T_p\hat{M}}$, where N_p is the complementary n-dimensional subspace. Notice that the right hand side of the expression for $Dh(k)$ is defined on the larger space T_pM by (H6) and that this induces a bounded linear operator in $\mathcal{L}(T_pM, N_p)$, where T_pM is a Banach space with the weaker supremum norm. We denote this extension by $Dh_e(k) \in \mathcal{L}(T_pM, N_p)$ and lastly we note that, by (H6), the map $U_0 \times T_pM \ni (k, \gamma) \mapsto Dh_e(k)\gamma$ varies continuously.

Remark. It is natural to call $\hat{M}$ a C^0-*extendable submanifold* of $C(I, D(A) \times \mathbf{R}^n)$.

3.5 The Linear Variational System along Flowlines in $\hat{M}_0$

Throughout this section, let $\Psi_0 = \begin{pmatrix} \psi_0 \\ \varphi_0 \end{pmatrix} \in \hat{M}_0$ and let $\begin{pmatrix} x(t) \\ a(t) \end{pmatrix}$ be the corresponding (maximal) (classical) solution of (3.2.1) on $I \cup [0, t_e)$ which lies in $\hat{M}_0$. We consider (for now formally) the linear variational system along the trajectory $\hat{S}(t, \Psi_0)$,

$$\begin{cases} y'(t) = Ay(t) + D_1F_1(x(t), a(t))y(t) \\ \qquad +D_2F_1(x(t), a(t))b(t) + DF_2(x(t))y(t), \quad t \in [0, t_e), \\ b(t) = D_1H(x_t, a_t)y_t + D_2H(x_t, a_t)b_t, \\ \begin{pmatrix} y_0 \\ b_0 \end{pmatrix} = \begin{pmatrix} \psi_1 \\ \varphi_1 \end{pmatrix}. \end{cases} \tag{3.5.1}$$

We make the following hypothesis concerning the partial derivatives of $F_1 : C_0 \times K \to X_1$.

(H7) (i) For each $(c, k) \in C_0 \times K$ there is a bounded linear map $D_1F_1(c, k) \in \mathcal{L}(X_0, X_1)$ with $\lim_{\xi \to 0, c+\xi \in C_0} \frac{|F_1(c+\xi,k)-F_1(c,k)-D_1F_1(c,k)(\xi)|}{|\xi|} = 0$.

(ii) For each $(c,k) \in (D(A) \cap C_0) \times K$ there is a bounded linear map $D_2F_1(c,k) \in \mathcal{L}(\mathbf{R}^n, X_1)$ with

$$\lim_{\xi \to 0, k+\xi \in K} \frac{|F_1(c,k+\xi) - F_1(c,k) - D_2F_1(c,k)(\xi)|}{|\xi|} = 0.$$

(iii) The maps $(C_0 \times K \times X_0)_{X_0 \times \mathbf{R}^n \times X_0} \ni (c,k,\gamma) \mapsto D_1F_1(c,k)\gamma \in X_1$ and $[(D(A) \cap C_0) \times K]_{D(A) \times \mathbf{R}^n} \ni (c,k) \mapsto D_2F_1(c,k) \in \mathcal{L}(\mathbf{R}^n, X_1)$ are continuous (The subscripts attached to the domains indicate the choices of topology on the domains).

Remark. The weaker form of continuity of the partial derivative given in (H7) is reminiscent of one given in [9, 35].

We start with the following definitions.

Definition. Suppose (H6) holds. Let $\Psi = \begin{pmatrix} \psi \\ \varphi \end{pmatrix} \in \hat{M}_0$. For $t \in [0, t_e)$, let

$$TM_0^t := \left\{ \begin{pmatrix} \rho \\ \chi \end{pmatrix} \in C(I, X_0 \times \mathbf{R}^n) \mid \chi(0) = D_1H(x_t, a_t)\rho + D_2H(x_t, a_t)\chi \right\}.$$

Remark. $TM_0^t = T_{\hat{S}_t(\Psi)}M$ is simply an extension of the tangent space $T_{\hat{S}_t(\Psi)}\hat{M}$ introduced in Section 3.5.

Definition. Suppose (H1)–(H7) hold. By a mild solution of (3.5.1) on $I \cup [0,T]$ for $0 < T < t_e \leq \infty$ we mean a pair of functions $\begin{pmatrix} y \\ b \end{pmatrix}$ such that

(i) $b : I \cup [0,T] \to K$ is continuous;

(ii) $y : I \cup [0,T] \to X_0$ is continuous and, for each $t \in [0,T]$, $\int_0^t y(s)ds \in D(A)$ and

$$y(t) = y(0) + A\int_0^t y(s)ds + \int_0^t \Big[D_1F_1(x(s),a(s))y(s) + D_2F_1(x(s),a(s))b(s) + DF_2(x(s))y(s)\Big]ds;$$

(iii) For $0 \le t \le T$, $\begin{pmatrix} y_t \\ b_t \end{pmatrix} \in TM_0^t$, *i.e.*, $b(t) = D_1H(x_t,a_t)y_t + D_2H(x_t,a_t)b_t$;

(iv) $\begin{pmatrix} y_0 \\ b_0 \end{pmatrix} = \begin{pmatrix} \psi_1 \\ \varphi_1 \end{pmatrix}$.

In case $T = t_e$, we make appropriate modifications to the above definition.

Remark. Given $t_0 \in (0,t_e)$, we can also consider (3.5.1) for $t \in (t_0,t_e)$ with initial data $\begin{pmatrix} \psi_1 \\ \varphi_1 \end{pmatrix} \in TM_0^{t_0}$ and similarly define a mild solution on $[t_0-h,t_0]\cup[t_0,T]$ for $t_0 < T < t_e$ or $[t_0-h,t_0]\cup[t_0,T)$ for $t_0 < T \le t_e$.

It does *not* follow from [13] or other related works on partial functional differential equations such as [30, 34, 39] that (3.5.1) has a mild solution. We address this issue with the following lemma and proposition.

Lemma 4 *Suppose that (H6) and (H7) hold. The map $C_0\times K \ni (c,k) \mapsto D_1F_1(c,k) \in \mathcal{L}(X_0,X_1)$ is locally bounded in the following sense: Each $X_0\times\mathbf{R}^n$-compact set $J \subset$*

$C_0 \times K$ has an $X_0 \times \mathbf{R}^n$-open neighborhood N such that $D_1F_1 : C_0 \times K \to \mathcal{L}(X_0, X_1)$ is bounded on $N \cap (C_0 \times K)$. Similarly, the map $U \ni (\psi^1, \varphi^1) \mapsto D_1H(\psi^1, \varphi^1) \in \mathcal{L}(C(I, X_0), \mathbf{R}^n)$ is also locally bounded, where U has the relative topology from $C(I, D(A) \times \mathbf{R}^n)$.

Proof. We only give the proof of the first part since that of the second part is similar. Let $J \subset C_0 \times K$ be compact. For each $(c_0, k_0) \in J$, we show that there is a relative neighborhood $N_{(c_0,k_0)}$ of (c_0, k_0) and $B > 0$ such that $||D_1F_1(c, k)||_{\mathcal{L}(X_0,X_1)} \leq B$ for each $(c, k) \in N_{(c_0,k_0)}$. By way of contradiction, there is a $(c_0, k_0) \in J$ and a sequence $(c_n, k_n) \to (c_0, k_0)$ in $C_0 \times K$ such that $||D_1F_1(c_n, k_n)||_{\mathcal{L}(X_0,X_1)} \to \infty$ as $n \to \infty$. It follows from (H7)(iii) that $\{D_1F_1(c_n, k_n)\gamma\}$ is bounded for each $\gamma \in X_0$. Then the uniform boundedness principle implies that $\{||D_1F_1(c_n, k_n)||_{\mathcal{L}(X_0,X_1)}\}$ is also bounded, which is a contradiction. Now the result follows since J is compact.

Proposition 9 *Suppose (H1)–(H7) hold. If the initial data $\begin{pmatrix} \psi_1 \\ \varphi_1 \end{pmatrix} \in TM_0^0$, then* (3.5.1) *has a unique mild solution $\begin{pmatrix} y(t) \\ b(t) \end{pmatrix}$ on $I \cup [0, t_e)$.*

Proof. The proof is completed in three steps.

Step 1. Let $0 \leq t_0 < t_e$ *and* $\begin{pmatrix} \psi \\ \varphi \end{pmatrix} \in TM_0^{t_0}$ *be given. Then there is* $\tau \in (t_0, t_e)$ *such that* (3.5.1) *has a mild solution on* $[t_0 - h, \tau]$ *with initial data* $\begin{pmatrix} \psi \\ \varphi \end{pmatrix}$.

Let $T \in (t_0, t_e)$ and $\mathcal{C} := \{b : [t_0 - h, T] \to \mathbf{R}^n \mid b \text{ is continuous and } b_{t_0} = \varphi\}$. Note that $\mathcal{C}$ is a closed subset of $C([t_0 - h, T], \mathbf{R}^n)$. Furthermore, for each $b \in \mathcal{C}$ the non-autonomous equation

$$\begin{cases} y'(t) &= Ay(t) + D_1F_1(x(t), a(t))y(t) \\ & \quad +D_2F_1(x(t), a(t))b(t) + DF_2(x(t))y(t), \qquad t \in [t_0, T], \\ y_{t_0} &= \psi \end{cases} \tag{3.5.2}$$

can be solved for a unique mild solution $y = y(b) : [t_0 - h, T] \to X_0$. To justify the latter statement, we note by (H7)(iii) that the map $[0, T] \ni s \mapsto D_2F_1(x(s), a(s))b(s) \in X$ is continuous. Therefore, it suffices to show that for each fixed $t \in [t_0, T]$ the term $G : [t_0, T] \times X_0 \to X$ given by $G(t, y) := D_1F_1(x(t), a(t))y + D_2F_1(x(t), a(t))b(t) + DF_2(x(t))y$ is Lipschitz on X_0 uniformly in t as we then can apply Proposition 2.10 of [34]. Given y_1, $y_2 \in X_0$, we have

$$|G(t, y_1) - G(t, y_2)| = |D_1F_1(x(t), a(t))(y_1 - y_2) + DF_2(x(t))(y_1 - y_2)|.$$

By Lemma 4, there is some $B > 0$ such that $||D_1F_1(x(s), a(s))||_{\mathcal{L}(X_0, X_1)} \leq B$ for each $s \in [t_0, T]$. By (H4) there is some $B' > 0$ such that $||DF_2(x(s))|| \leq B'$ for each $s \in [t_0, T]$. It follows that $G(t, \cdot)$ is Lipschitz on X_0 uniformly in t.

To obtain a solution for the second component of (3.5.1), we let $\mathcal{A} : \mathcal{C} \to \mathcal{C}$ be given by

$$(\mathcal{A}b)(t) = \begin{cases} \varphi(t - t_0) & \text{if } t \in [t_0 - h, t_0], \\ D_1 H(x_t, a_t) y(b)_t + D_2 H(x_t, a_t) b_t & \text{if } t \in [t_0, T], \end{cases} \tag{3.5.3}$$

where $y(b)$ denotes the solution to (3.5.2). That $(\mathcal{A}b) \in \mathcal{C}$ follows from the continuity of the maps $[t_0, T] \ni t \mapsto y(b)_t \in C([t_0 - h, t_0], X_0)$, $[t_0, T] \ni t \mapsto D_1 H(x_t, a_t) y(b)_t \in \mathbf{R}^n$, and $[t_0, T] \ni t \mapsto D_2 H(x_t, a_t) \in \mathcal{L}(C(I, \mathbf{R}^n), \mathbf{R}^n)$. The continuity of the latter two maps is a consequence of (H6) while that of the former is a consequence of the continuity of $y(b)$ on $[t_0 - h, T]$. In the following, we show that $\mathcal{A}$ is a contraction provided T is small enough.

Let $T_0 \in (t_0, t_e)$. It follows that $\max_{s \in [0, T_0]} ||x_s|| = R_1$ for some $R_1 \in [0, R_0]$. Using (3.5.3), (H6), (H2), and Lemma 4 we have for $t_0 < T_0 < t_e$ and $t \in [t_0, T_0]$ that

$$\begin{aligned} |(\mathcal{A}b_1)(t) - (\mathcal{A}b_2)(t)| &\leq |D_1 H(x_t, a_t)(y(b_1)_t - y(b_2)_t)| \\ &\quad + |D_2 H(x_t, a_t)((b_1)_t - (b_2)_t)| \\ &\leq C ||y(b_1)_t - y(b_2)_t|| + f(R_1) ||(b_1)_t - (b_2)_t||, \end{aligned} \tag{3.5.4}$$

where $C = \sup_{s \in [0, T_0]} ||D_1 H(x_s, a_s)||$. Moreover, using the abstract variation of constants formula (see the Remark following (3.2.1)), Lemma 4, (H7)(iii), (H4), and (H1), we have

$$
\begin{aligned}
& |y(b_1)(t) - y(b_2)(t)| \\
= & \left| \lim_{\mu\to\infty} \int_{t_0}^{t} T(t-s)\mu R_\mu [D_1F_1(x(s),a(s))(y(b_1)(s)-y(b_2)(s)) \right. \\
& +D_2F_1(x(s),a(s))(b_1(s)-b_2(s)) \\
& \left. +DF_2(x(s))(y(b_1)(s)-y(b_2)(s))]ds \right| \\
\leq & \int_{t_0}^{t} M^2 e^{\omega(t-s)} (C_1|y(b_1)(s)-y(b_2)(s)| + C_2||b_1-b_2||)ds \\
\leq & M^2 C_2 e^{\omega T_0} ||b_1-b_2||t + \int_{t_0}^{t} M^2 e^{\omega(t-s)} C_1 |y(b_1)(s)-y(b_2)(s)|ds, \quad (3.5.5)
\end{aligned}
$$

where

$$
\begin{aligned}
C_1 &= \max\{\sup_{s\in[0,T_0]} ||D_1F_1(x(s),a(s))||, \max_{s\in[0,T_0]} ||DF_2(x(s))||\}, \\
C_2 &= \max_{s\in[0,T_0]} ||D_2F_1(x(s),a(s))||.
\end{aligned}
$$

An application of Gronwall's inequality to (3.5.5) yields that, for each $t_0 \leq t \leq T_0$, $|y(b_1)(t)-y(b_2)(t)| \leq JC_2t||b_1-b_2||$ for some $J>0$ which depends on T_0. It follows from $||y(b_1)_t - y(b_2)_t|| \leq ||y(b_1)-y(b_2)|| \leq JC_2T||b_1-b_2||$ and (3.5.4) that

$$
|(\mathcal{A}b_1)(t) - (\mathcal{A}b_2)(t)| \leq CJC_2T||b_1-b_2|| + f(R_1)||b_1-b_2||
$$

for each $t_0 \leq t \leq T \leq T_0$. Since $f(R_1) < 1$, it is clear that $\mathcal{A}$ is a contraction provided $T = \tau$ is chosen small enough.

Step 2. Local solutions of (3.5.1) *are unique.*

Suppose that $\begin{pmatrix} y_1 \\ b_1 \end{pmatrix}$ and $\begin{pmatrix} y_2 \\ b_2 \end{pmatrix}$ are two mild solutions of (3.5.1) respectively on $I \cup A_1$ and $I \cup A_2$ having the same initial data $\begin{pmatrix} \psi_1 \\ \varphi_1 \end{pmatrix} \in TM_0^0$, where $A_i = [0, \tau_i] \subset [0, t_e)$ or $A_i = [0, t_e)$, $i = 1, 2$. We show that the two solutions agree on $A = A_1 \cap A_2$. For each $T \geq 0$ such that $[0, T] \subset A$, by the same argument as in Step 1 which established (3.5.5), we have

$$|y_1(t) - y_2(t)| \leq \int_0^t M^2 e^{\omega(t-s)}(C_1|y_1(s) - y_2(s)| + C_2|b_1(s) - b_2(s)|)ds \qquad (3.5.6)$$

for $t \in [0, T]$. Let

$$\hat{b}_i(\xi) = \begin{cases} \varphi_1(-h) & \text{if } \xi \leq -h \\ b_i(\xi) & \text{if } \xi \in [-h, T] \end{cases}$$

and

$$\hat{y}_i(\xi) = \begin{cases} \psi_1(-h) & \text{if } \xi \leq -h \\ y_i(\xi) & \text{if } \xi \in [-h, T] \end{cases}$$

be the trivial extensions of b_i and y_i to $(-\infty, T]$, respectively, $i = 1, 2$. Denote $R_1 = \max_{s \in [0,T]} ||x_s||$. Clearly, $\begin{pmatrix} \hat{y}_t \\ \hat{b}_t \end{pmatrix} \in BUC((-\infty, 0], X_0 \times \mathbf{R}^n)$. Arguing as in Step 1, we have that $|b_1(t) - b_2(t)| \leq C||(y_1)_t - (y_2)_t|| + f(R_1)||(b_1)_t - (b_2)_t||$ for $t \in [0, T]$. It is not difficult to see that

$$||(\hat{b}_1)_t - (\hat{b}_2)_t|| \leq C(1 - f(R_1))^{-1}||(\hat{y}_1)_t - (\hat{y}_2)_t||. \qquad (3.5.7)$$

Combining (3.5.6) with (3.5.7) and using an application of Gronwall's inequality, we get $y_1(t) = y_2(t)$ for $t \in I \cup [0, T]$. It now follows from (3.5.7) that $b_1 = b_2$ on $I \cup [0, T]$. As T is arbitrary, this completes Step 2.

Step 3. Let

$$t_e'(\psi_1, \varphi_1) := \sup\{\rho \in (0, t_e) \mid (3.5.1) \textit{ has a mild solution } (y, b) \textit{ on } I \cup [0, \rho]\}.$$

Then $t_e'(\psi_1, \varphi_1) = t_e$.

By Step 1, $t_e' > 0$. Suppose $t_e' = \rho_0 < t_e$. It follows from Lemma 4, (H7)(iii), and (H4) that

$$\max \left\{ \begin{array}{l} \sup_{s \in [0, \rho_0]} ||D_1 F_1(x(s), a(s))||, \\ \\ \max_{s \in [0, \rho_0]} ||DF_2(x(s))||, \\ \\ \max_{s \in [0, \rho_0]} ||D_2 F_1(x(s), a(s))|| \end{array} \right\} < \infty.$$

By Lemma 4, $C' := \sup_{s \in [0, \rho_0]} ||D_1 H(x_s, a_s)|| < \infty$. Let $(y, b) : I \cup [0, \rho_0) \to X_0 \times \mathbf{R}^n$ be the mild solution of (3.5.1) having initial data $\begin{pmatrix} \psi_1 \\ \\ \varphi_1 \end{pmatrix}$. Let $R_1 = \max_{s \in [0, \rho_0]} ||x_s||$. As in Step 2, let $(\hat{y}, \hat{b}) : (-\infty, 0] \cup [0, \rho_0)$ be the trivial extension of (y, b) to $[-\infty, \rho_0)$. It is not difficult to see that $||\hat{b}_t|| \le (1 - f(R_1))^{-1} ||\varphi_1|| + (1 - f(R_1))^{-1} C' ||\hat{y}_t||$ for each $t \in [0, \rho_0)$. Therefore, there is some $C'' > 0$ (which depends on ρ_0) such that $|y(t)| \le M e^{\omega t} |\psi_1(0)| + C'' \int_0^t (|y(s)| + |b(s)|) ds \le M e^{\omega t} |\psi_1(0)| + C'' \int_0^t (||\hat{y}_s|| + ||\hat{b}_s||) ds$

for $t \in [0, \rho_0)$. It follows that there is $C''' > 0$ such that

$$||\hat{y}_t|| \leq ||\psi_1|| + Me^{\omega t}||\psi_1|| + C''' \int_0^t ||\hat{y}_s|| + ||\varphi_1||ds$$
$$\leq (Me^{\omega t} + 1)||\psi_1|| + C'''||\varphi_1||t + C''' \int_0^t ||\hat{y}_s||ds$$

for $t \in [0, \rho_0)$. Then the continuity of the map $[0, \rho_0) \ni t \mapsto \hat{y}_t \in BUC((-\infty, 0], X_0)$ and Gronwall's inequality imply that y is bounded on $I \cup [0, \rho_0)$. By setting $\tilde{y}(\rho_0) := T(\rho_0)\psi_1(0) + \lim_{\mu\to\infty} \int_0^{\rho_0} T(\rho_0 - s)\mu R_\mu[D_1F_1(x(s), a(s))y(s) + D_2F_1(x(s), a(s))b(s) + DF_2(x(s))y(s)]ds$, it is not difficult to see that y can be extended to a continuous map $\tilde{y} : I \cup [0, \rho_0] \to X_0$. Then, the equation

$$\tilde{b}(t) = \begin{cases} \varphi_0(t) & \text{if } t \in I \\ D_1H(x_t, a_t)\tilde{y}_t + D_2H(x_t, a_t)\tilde{b}_t & \text{if } t \in [0, \rho_0] \end{cases}$$

can be solved for a unique continuous map $\tilde{b} : I \cup [0, \rho_0] \to \mathbf{R}^n$ by using the fact that $f(R_1) < 1$, where $R_1 = \max_{s\in[0,\rho_0]} ||x_s||$, and the contraction mapping principle. Note that it is obvious that $\tilde{b}(t) = b(t)$ for $t < \rho_0$. Then applying Step 1 for $t_0 = \rho_0$ and $(\psi, \varphi) = (\tilde{y}_{\rho_0}, \tilde{b}_{\rho_0}) \in TM_0^{\rho_0}$, we can extend (y, b) beyond ρ_0, which is a contradiction.

By applying similar arguments as those in Step 3 of the proof of Proposition 9, we can obtain the following result.

Corollary 10 *Suppose (H1)–(H7) hold. For* $\Psi \in TM_0^0$*, let* $\begin{pmatrix} y \\ b \end{pmatrix}$ *be the corresponding mild solution to* (3.5.1) *on* $I \cup [0, t_e)$*. Then, for each* $T \in [0, t_e)$ *and*

$t \in [0,T]$, $\left\| \begin{pmatrix} y_t \\ b_t \end{pmatrix} \right\| \leq C||\Psi||$, *where* C *depends on* Ψ_0 *and* T. *(Recall that* Ψ_0 *is fixed throughout this section).*

3.6 Derivatives of Solution Operators $\hat{S}_t$ on $\hat{M}_0$

Recall from Section 3.4 that $\hat{M}_0$ is a positively invariant subset for the semiflow S of the C^1-submanifold $\hat{M}$ of $C(I, D(A) \times \mathbf{R}^n)$. At each point $p \in \hat{M}_0$ the tangent space at p, denoted by $T_p\hat{M}$, is contained in a larger set T_pM which is a Banach space with the weaker supremum norm (*i.e.*, the supremum norm which does not include the contribution from the operator A). Moreover, we let $T\hat{M}_0 = \{(p, \gamma) \mid p \in \hat{M}_0 \text{ and } \gamma \in T_p\hat{M}\}$ denote the tangent bundle of $\hat{M}$ restricted to $\hat{M}_0$ and point out that it has an obvious extension which we call $TM_0 = \{(p, \gamma) \mid p \in \hat{M}_0 \text{ and } \gamma \in T_pM\}$. In order to derive the desired differentiability of $\hat{S}_t$ on $\hat{\Omega}_t$, we consider the interpolation space $(C^1(I, X_0 \times \mathbf{R}^n) \cap C(I, D(A) \times \mathbf{R}^n), ||\cdot||_1)$, where $\left\| \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} \right\|_1 = ||\xi_1'|| + ||\xi_2'|| + \left\| \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} \right\|_{C(I, D(A) \times \mathbf{R}^n)}$. From now on, we let $T_p\hat{M}^1 = T_p\hat{M} \cap C^1(I, X_0 \times \mathbf{R}^n)$ and view it as a Banach space with the $||\cdot||_1$ norm. We note that the norm $||\cdot||_1$ is equivalent to the norm given by $||\xi||_1' = ||\xi|| + ||\xi'|| + ||A\xi_1||$, where $\xi =$

$\begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} \in T_p\hat{M}^1$. Before stating the main theorem of this section, we strengthen hypotheses (H6) and (H7) as follows[1]:

(H6)* In addition to (H6), we further assume that, for (ψ_0, φ_0), $(\psi, \varphi) \in \hat{M}_0$, $|DH(\psi,\varphi)(\psi',\varphi') - DH(\psi_0,\varphi_0)(\psi_0',\varphi_0')| \leq O(||\psi-\psi_0|| + ||\varphi-\varphi_0||) + O(||\varphi - \varphi_0||)||\psi_0'|| + O(||\varphi-\varphi_0||)||\varphi_0'||$ and $||D_1H(\psi,\varphi) - D_1H(\psi_0,\varphi_0)||_{\mathcal{L}(C(I,D(A)),\mathbf{R}^n)} \leq O(||\psi - \psi_0|| + ||\varphi - \varphi_0||)$ both hold uniformly.

(H7)* In addition to (H7), we assume that for each (c,k), $(c_0,k_0) \in (C_0 \cap D(A)) \times K$ we have $||D_2F_1(c,k) - D_2F_1(c_0,k_0)||_{\mathcal{L}(\mathbf{R}^n, X_1)} \leq O(|c-c_0|_{D(A)}) + Z_{(c_0,k_0)}(|c - c_0| + |k - k_0|)$. Moreover, for $(c,k) \in (C_0 \cap D(A)) \times K$, $D_1F_1(c,k) \in \mathcal{L}(D(A), X_1)$ exists and satisfies the special Lipschitz condition: $||D_1F_1(c,k) - D_1F_1(c_0,k_0)||_{\mathcal{L}(D(A),X_1)} \leq O(|c-c_0| + |k-k_0|)$ uniformly.

Remark. Note that the Lipschitz conditions in each of (H6)* and (H7)* involve a weaker norm on the right hand side.

The following is the main result of this section.

Theorem 11 *Assume (H1)–(H5), (H6)*, and (H7)* hold. Then the function* $\hat{S}_t$:

[1]Throughout this section we use the following notation for a function g defined on a neighborhood of zero, in the product of two normed spaces, and whose image is contained in another normed space. $|g(\xi,w)| \leq o(|\xi|)$ means that $\lim_{\xi\to 0}\frac{|g(\xi,w)|}{|\xi|} = 0$ pointwise. Similarly, $|g(\xi,w)| \leq O(|\xi|)$ means $|g(\xi,w)| \leq C|\xi|$ for some $C > 0$ which depends on w. Lastly, $|g(\xi,w)| = Z_w(|\xi|)$ means $|g(\xi,w)| \to 0$ as $\xi \to 0$ pointwise.

$\hat{\Omega}_t \to \hat{M}_0$ is differentiable in the following sense: For each $p \in \hat{\Omega}_t$, $D\hat{S}_t(p) \in \mathcal{L}(T_p\hat{M}^1, T_{\hat{S}_t(p)}M)$ and satisfies

$$\lim_{\xi\to 0, p+\xi\in\hat{\Omega}_t, \xi\in T_p\hat{M}^1} \frac{||\hat{S}_t(p+\xi) - \hat{S}_t(p) - D\hat{S}_t(p)\xi||_{C(I,X_0\times\mathbf{R}^n)}}{||\xi||_{T_p\hat{M}^1}} = 0.$$

In fact, the mapping $z : I \cup [0, t_e(p)) \to X_0 \times \mathbf{R}^n$ given by $z(t) = D\hat{S}_t(p)(\xi)(0)$ for $t \in [0, t_e(p))$ is a solution of the linear variational system (3.5.1) *along $\hat{S}(t,p)$ with initial data $z_0 = \xi$. Furthermore, the map $d\hat{S}_t : T\hat{M}_0 \cap (\hat{\Omega}_t \times C^1(I, X_0 \times \mathbf{R}^n)) \to TM_0$ given by $d\hat{S}_t(p,\gamma) = (\hat{S}_t(p), D\hat{S}_t(p)\gamma)$ is continuous when the domain inherits the relative product topology induced from the $||\cdot||_1$ norm on $C^1(I, X_0 \times \mathbf{R}^n) \cap C(I, D(A) \times \mathbf{R}^n)$ and TM_0 has the relative product topology from $C(I, X_0 \times \mathbf{R}^n)$.*

Proof. Given $t > 0$ and $p = \begin{pmatrix} \psi \\ \varphi \end{pmatrix} \in \hat{\Omega}_t$, write $\hat{S}(t,p) = \begin{pmatrix} x_t^\psi \\ a_t^\varphi \end{pmatrix} = \begin{pmatrix} x_t^p \\ a_t^p \end{pmatrix} \in \hat{M}_0$. Let $\begin{pmatrix} y \\ b \end{pmatrix} : I \cup [0, t_e(p)) \to X_0 \times \mathbf{R}^n$ be the mild solution of the linear variational system (3.5.1) along $\hat{S}_t(p)$ having initial data $\xi = \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} \in T_p\hat{M}$. It follows from Corollary 10 that $\left\| \begin{pmatrix} y_t \\ b_t \end{pmatrix} \right\| \leq C \left\| \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} \right\|$ for a constant $C > 0$ depending on p and t. In case $\xi \in T_p\hat{M}^1$, we note that $\left\| \begin{pmatrix} y_t \\ b_t \end{pmatrix} \right\| \leq C \left\|\xi\right\|_{T_p\hat{M}^1}$. We can find $0 < R_1 < R_0$ such that $\max_{s\in[0,t]} ||x_s^p|| < R_1$. The proof is achieved in the following eight steps.

Step 1. For $s \in [0,t]$, p, $p+\xi \in \hat{\Omega}_t$, $||x_s^{p+\xi} - x_s^p|| + ||a_s^{p+\xi} - a_s^p|| \leq O(||\xi||)$ uniformly in s and pointwise in p.

This follows from a standard argument using Gronwall's inequality and (H2) (See Step 2 of the proof of Proposition 9 or Step 1 in the proof of Theorem 2 in [13]).

Step 2. For $\mu \in \hat{\Omega}_t$, the pair $\begin{pmatrix} \dot{x^\mu} \\ \dot{a^\mu} \end{pmatrix}$ is a mild solution of the linear variational system (3.5.1) on $I \cup [0,t]$ along $\begin{pmatrix} x^\mu \\ a^\mu \end{pmatrix}$. In particular,

$$\begin{aligned} \ddot{x^\mu}(s) &= A\dot{x^\mu}(s) + D_1F_1(x^\mu(s), a^\mu(s))\dot{x^\mu}(s) \\ &\quad + D_2F_1(x^\mu(s), a^p(s))\dot{a^\mu}(s) + DF_2(x^\mu(s))\dot{x^\mu}(s) \end{aligned}$$

in the mild sense and

$$\dot{a^\mu}(s) = D_1H(x^\mu(s), a^\mu(s))\dot{x^\mu}_s + D_2H(x^\mu(s), a^\mu(s))\dot{a^\mu}_s.$$

By the proof of Theorem 5 and by Theorem 6, it suffices to check that

$$\frac{d}{ds}F_1(x^\mu(s), a^\mu(s)) = D_1F_1(x^\mu(s), a^\mu(s))\dot{x^\mu}(s) + D_2F_1(x^\mu(s), a^p(s))\dot{a^\mu}(s).$$

Note that this is not an immediate consequence of the chain rule since in general F_1 is not differentiable. However, the same arguments as those before Step 1 in the proof of Theorem 6 (with the use of (H6) being replaced by (H7)) can be used to obtain the desired result here.

Step 3. For $s \in [0,t]$ and p, $p+\xi \in \hat{\Omega}_t$, $||\dot{a}_s^{p+\xi} - \dot{a}_s^p|| \le O(||\xi|| + ||\xi'||)$ holds pointwise in p and uniformly in s.

Let $\theta \in [-h, 0]$. If $s \in [0,t]$ and $s+\theta \le 0$ then $|\dot{a}^{p+\xi}(s+\theta) - \dot{a}^p(s+\theta)| \le ||\dot{\xi}||$. If $s+\theta \ge 0$, then by Step 2, (H6)*, and Step 1, we have

$$
\begin{aligned}
& |\dot{a}^{p+\xi}(s+\theta) - \dot{a}^p(s+\theta)| \\
\le\ & |DH(x_{s+\theta}^{p+\xi}, a_{s+\theta}^{p+\xi})(\dot{x}_{s+\theta}^{p+\xi}, \dot{a}_{s+\theta}^{p+\xi}) - DH(x_{s+\theta}^{p}, a_{s+\theta}^{p})(\dot{x}_{s+\theta}^{p}, \dot{a}_{s+\theta}^{p})| \\
\le\ & O(||\xi||) + O(||\xi||)||\dot{x}_{s+\theta}^p|| + O(||\xi||)||\dot{a}_{s+\theta}^p|| \\
\le\ & O(||\xi||),
\end{aligned}
$$

where the constant coming from the latter big O depends on $\max_{\mu \in I \cup [0,t]} |\dot{x}^p(\mu)|, \max_{\mu \in I \cup [0,t]} |\dot{a}^p(\mu)|$ and clearly depends on p.

Therefore, $||\dot{a}_s^{p+\xi} - \dot{a}_s^p|| \le O(||\xi|| + ||\xi'||)$ pointwise in p and uniformly in s.

Step 4. For $s \in [0,t]$ and p, $p+\xi \in \hat{\Omega}_t$, $||\dot{x}_s^{p+\xi} - \dot{x}_s^p|| \to 0$ as $||\xi|| + ||\xi'|| \to 0$ uniformly in s and pointwise in p.

Let $w(s) := \dot{x^{p+\xi}}(s) - \dot{x^p}(s)$. From Step 2 we have

$$\begin{aligned}|w(s)| &\leq |T(s)(\dot{\xi}(0))| + \int_0^s M^2 e^{\omega(s-\theta)} |D_1F_1(x^{p+\xi}(\theta), a^{p+\xi}(\theta))\dot{x^{p+\xi}}(\theta) \\ &\quad -D_1F_1(x^p(\theta), a^p(\theta))\dot{x^p}(\theta) + D_2F_1(x^{p+\xi}(\theta), a^{p+\xi}(\theta))\dot{a^{p+\xi}}(\theta) \\ &\quad -D_2F_1(x^p(\theta), a^p(\theta))\dot{a^p}(\theta) + DF_2(x^{p+\xi}(\theta))\dot{x^{p+\xi}}(\theta) \\ &\quad -DF_2(x^p(\theta))\dot{x^p}(\theta)|d\theta \\ &= |T(s)(\dot{\xi}(0))| + \int_0^s M^2 e^{\omega(s-\theta)} |I(\theta)| d\theta\end{aligned}$$

for $s \in [0,t]$. Since the set $\{(x^p(s), a^p(s)) \mid s \in [0,t]\} \subset C_0 \times K$ is $X_0 \times \mathbf{R}^n$ compact, by Lemma 4 we can find an open neighborhood N of it in $X_0 \times \mathbf{R}^n$ such that $C := \sup_{(c,k)\in N\cap(C_0\times K)} ||D_1F_1(c,k)|| < \infty$. By Step 1, we can choose $||\xi||$ small enough such that $(x^{p+\xi}(s), a^{p+\xi}(s)) \in N$ for each $s \in [0,t]$. Therefore, it follows from (H7), (H7)*, and Step 3 that

$$\begin{aligned}
|I(\theta)| &\leq C|\dot{x^{p+\xi}}(\theta) - \dot{x^p}(\theta)| \\
&\quad + |D_1F_1(x^{p+\xi}(\theta), a^{p+\xi}(\theta))\dot{x^p}(\theta) - D_1F_1(x^p(\theta), a^p(\theta))\dot{x^p}(\theta)| \\
&\quad + |[D_2F_1(x^{p+\xi}(\theta), a^{p+\xi}(\theta)) - D_2F_1(x^p(\theta), a^p(\theta))]\dot{a^{p+\xi}}(\theta)| \\
&\quad + ||D_2F_1(x^p(\theta), a^p(\theta))||\ |\dot{a^{p+\xi}}(\theta) - \dot{a^p}(\theta)| \\
&\quad + |DF_2(x^{p+\xi}(\theta))\dot{x}^{p+\xi}(\theta) - DF_2(x^p(\theta))\dot{x}^p(\theta)| \\
&\leq C|\dot{x^{p+\xi}}(\theta) - \dot{x^p}(\theta)| + Z_{(x^p(\theta), a^p(\theta))}(||\xi||) \\
&\quad + (|x^{p+\xi}(\theta) - x^p(\theta)|_{D(A)} + Z_{(x^p(\theta), a^p(\theta))}(||\xi||))|\dot{a}^{p+\xi}(\theta)| \\
&\quad + \max_{\mu\in[0,t]} ||D_2F(x^p(\mu), a^p(\mu))||\ |O(||\xi|| + ||\xi'||)| \\
&\quad + ||DF_2(x^{p+\xi}(\theta))||\ |\dot{x}^{p+\xi}(\theta) - \dot{x}^p(\theta)| \\
&\quad + ||DF_2(x^{p+\xi}(\theta)) - DF_2(x^p(\theta))||\ |\dot{x}^p(\theta)|.
\end{aligned}$$

Note that by Theorem 5 and Step 1,

$$\begin{aligned}
|x^{p+\xi}(\theta) - x^p(\theta)|_{D(A)} &= |x^{p+\xi}(\theta) - x^p(\theta)| + |Ax^{p+\xi}(\theta) - Ax^p(\theta)| \\
&\leq |x^{p+\xi}(\theta) - x^p(\theta)| + |\dot{x}^{p+\xi}(\theta) - \dot{x}^p(\theta)| \\
&\quad + |F(x^{p+\xi}(\theta), a^{p+\xi}(\theta)) - F(x^p(\theta), a^p(\theta))| \\
&\leq O(||\xi||) + |\dot{x}^{p+\xi}(\theta) - \dot{x}^p(\theta)|,
\end{aligned}$$

where the constant coming from the big O depends only on p. Furthermore, by Step 3 we have that

$$|\dot{a}^{p+\xi}(\theta)| \leq |\dot{a}^{p+\xi}(\theta) - \dot{a}^{p}(\theta)| + |\dot{a}^{p}(\theta)| \leq O(||\xi|| + ||\xi'||) + \max_{\mu\in[0,t]} |\dot{a}^{p}(\mu)|,$$

where the constant coming from the big O depends only on p. Hence, choosing $||\xi|| + ||\xi'||$ small enough, we have $|\dot{a}^{p+\xi}(\theta)| \leq 1 + \max_{\mu\in[0,t]} |\dot{a}^{p}(\mu)|$. This gives $(|x^{p+\xi}(\theta) - x^{p}(\theta)|_{D(A)} + Z_{(x^p(\theta),a^p(\theta))}(||\xi||))|\dot{a}^{p+\xi}(\theta)| \leq O(||\xi|| + |\dot{x}^{p+\xi}(\theta) - \dot{x}^{p}(\theta)|) + Z_{(x^p(\theta),a^p(\theta))}(||\xi||)$. Then by (H4), the continuity of DF_2 on C_0 implies that we can find an X_0-open neighborhood, N_1, of the X_0-compact set $\{x^p(\mu) \mid \mu \in [0,t]\}$ and some $C_1 > 0$ such that, for each $c \in N_1 \cap C_0$, $||DF_2(c)||_{\mathcal{L}(X_0,X_2)} \leq C_1$. By Step 1, we can choose $||\xi||$ small enough such that $x^{p+\xi}(\mu) \in N_1$ for each $\mu \in [0,t]$. Hence $||DF_2(x^{p+\xi}(\theta))|| < C_1$. Finally, we can conclude that, for $||\xi|| + ||\xi'||$ small enough (depending on only p),

$$\begin{aligned}
|I(\theta)| &\leq C|\dot{x^{p+\xi}}(\theta) - \dot{x^{p}}(\theta)| + Z_{(x^p(\theta),a^p(\theta))}(||\xi||) \\
&\quad +O(\ ||\xi|| + |\dot{x}^{p+\xi}(\theta) - \dot{x}^{p}(\theta)|\ ) + Z_{(x^p(\theta),a^p(\theta))}(||\xi||) \\
&\quad +O(||\xi|| + ||\xi'||) + O(|\dot{x}^{p+\xi}(\theta) - \dot{x}^{p}(\theta)|) + Z_{(x^p(\theta))}(||\xi||) \\
&\leq O(||\xi|| + ||\xi'||) + O(|\dot{x^{p+\xi}}(\theta) - \dot{x^{p}}(\theta)|) + Z_{(x^p(\theta),a^p(\theta))}(||\xi||).
\end{aligned}$$

Therefore, for each $s \in [0, t]$,

$$\begin{aligned} |w(s)| &\leq Me^{\omega t}||\xi'|| + \int_0^s M^2 e^{\omega(s-\theta)}[O(||\xi|| + ||\xi'||) + O(|w(\theta)|) \\ &\quad + Z_{(x^p(\theta), a^p(\theta))}(||\xi||)]d\theta \\ &\leq O(||\xi|| + ||\xi'||) + \int_0^s O(|w(\theta)|)d\theta + \int_0^t Z_{(x^p(\theta), a^p(\theta))}(||\xi||)d\theta \\ &\leq O(||\xi|| + ||\xi'||) + \int_0^s O(|w(\theta)|)d\theta + Z_p(||\xi||). \end{aligned}$$

Here we have used the dominated convergence theorem to obtain the last line above. Step 4 now follows from Gronwall's inequality.

Step 5. If p, $p + \xi \in \hat{\Omega}_t$ *and* $\xi \in T_p\hat{M}$, *then* $||\hat{x}_t^{p+\xi} - \hat{x}_t^p - \hat{y}_t|| \leq o(||\xi|| + ||\xi'||) + \int_0^t O(||\hat{x}_s^{p+\xi} - \hat{x}_s^p - \hat{y}_s|| + ||\hat{a}_s^{p+\xi} - \hat{a}_s^p - \hat{b}_s||)ds$ *pointwise in* p, *where* $\hat{\cdot}$ *indicates the trivial extension of the corresponding function to* $(-\infty, 0]$ *by its value at* $-h$ *for* $\theta \leq -h$.

For each $s \in [0,t]$, it follows from (H4) and (H7) that

$$\begin{aligned}
& F_1(x^{p+\xi}(s), a^{p+\xi}(s)) - F_1(x^p(s), a^p(s)) - D_1F_1(x^p(s), a^p(s))y(s) \\
& -D_2F_1(x^p(s), a^p(s))b(s) + F_2(x^{p+\xi}(s)) - F_2(x^p(s)) - DF_2(x(s))y(s) \\
= \; & F_1(x^{p+\xi}(s), a^{p+\xi}(s)) - F_1(x^p(s), a^{p+\xi}(s)) + F_1(x^p(s), a^{p+\xi}(s)) \\
& -F_1(x^p(s), a^p(s)) - D_1F_1(x^p(s), a^p(s))y(s) - D_2F_1(x^p(s), a^p(s))b(s) \\
& +DF_2(x^p(s))(x^{p+\xi}(s) - x^p(s) - y(s)) + \omega_3(x^{p+\xi}(s) - x^p(s), x^p(s)) \\
= \; & D_1F_1(x^p(s), a^{p+\xi}(s))(x^{p+\xi}(s) - x^p(s)) - D_1F_1(x^p(s), a^p(s))y(s) \\
& +\omega_1(x^{p+\xi}(s) - x^p(s), x^p(s), a^{p+\xi}(s)) \\
& +D_2F_1(x^p(s), a^p(s))(a^{p+\xi}(s) - a^p(s) - b(s)) \\
& +\omega_2(a^{p+\xi}(s) - a^p(s), x^p(s), a^p(s)) \\
& +DF_2(x^p(s))(x^{p+\xi}(s) - x^p(s) - y(s)) + \omega_3(x^{p+\xi}(s) - x^p(s), x^p(s)) \\
= \; & [D_1F_1(x^p(s), a^{p+\xi}(s)) - D_1F_1(x^p(s), a^p(s))](x^{p+\xi}(s) - x^p(s)) \\
& +D_1F_1(x^p(s), a^p(s))(x^{p+\xi}(s) - x^p(s) - y(s)) \\
& +\omega_1(x^{p+\xi}(s) - x^p(s), x^p(s), a^{p+\xi}(s)) \\
& +D_2F_1(x^p(s), a^p(s))(a^{p+\xi}(s) - a^p(s) - b(s)) \\
& +\omega_2(a^{p+\xi}(s) - a^p(s), x^p(s), a^p(s)) \\
& +DF_2(x^p(s))(x^{p+\xi}(s) - x^p(s) - y(s)) + \omega_3(x^{p+\xi}(s) - x^p(s), x^p(s)),
\end{aligned}$$

where ω_1, ω_2, and ω_3 denote the error terms associated with D_1F_1, D_2F_1, and DF_2, respectively. By Step 1 and (H4), we know that $|\omega_3(x^{p+\xi}(s) - x^p(s), x^p(s))| \le o(||\xi||)$ uniformly in $s \in [0,t]$. Similarly, by Step 1 and (H7), we know that $|\omega_2(a^{p+\xi}(s) - a^p(s), x^p(s), a^p(s))| \le o(||\xi||)$ uniformly in s. The meat of the matter lies in ω_1. To this end, let $g_s : [0,1] \to X_1$ be given by $g_s(\mu) = F_1(x^p(s) + \mu(x^{p+\xi}(s) - x^p(s)), a^{p+\xi}(s))$. By (H7), g_s is C^1 in μ and

$$
\begin{aligned}
& |\omega_1(x^{p+\xi}(s) - x^p(s), x^p(s), a^{p+\xi}(s))| \\
= \; & |g(3.1.1) - g(0) - g'(0)| \\
= \; & \left| \int_0^1 g'(\mu) - g'(0) d\mu \right| \\
\le \; & \int_0^1 |D_1F_1(x^p(s) + \mu(x^{p+\xi}(s) - x^p(s)), a^{p+\xi}(s))(x^{p+\xi}(s) - x^p(s)) \\
& - D_1F_1(x^p(s), a^{p+\xi}(s))(x^{p+\xi}(s) - x^p(s))| d\mu \\
\le \; & \int_0^1 O(||\xi||)|x^{p+\xi}(s) - x^p(s)|_{D(A)} d\mu \\
\le \; & O(||\xi||)|x^{p+\xi}(s) - x^p(s)|_{D(A)}.
\end{aligned}
$$

Note carefully how the second last inequality follows from (H7)* and Step 1, and that the constant coming from the latter big O depends only on p. Then $|x^{p+\xi}(s) - x^p(s)|_{D(A)} \le O(||\xi||) + |\dot{x}^{p+\xi}(s) - \dot{x}^p(s)| \le Z_p(||\xi|| + ||\xi'||)$, where the first inequality follows from the argument starting with "Note that by Theorem 5 ..." in Step 4 and the second inequality follows from Step 4. This shows that $|\omega_1(x^{p+\xi}(s) - x^p(s), x^p(s), a^{p+\xi}(s))| \le o(||\xi|| + ||\xi'||)$ pointwise in p and uniformly in s. Similarly, it follows from (H7)*, Step 1, and Step 4 that

$$
\begin{aligned}
& |[D_1F_1(x^p(s), a^{p+\xi}(s)) - D_1F_1(x^p(s), a^p(s))](x^{p+\xi}(s) - x^p(s))| \\
\leq\ & O(||\xi||)|x^{p+\xi}(s) - x^p(s))|_{D(A)} \\
\leq\ & o(||\xi|| + ||\xi'||)
\end{aligned}
$$

uniformly in s and pointwise in p. Therefore, it follows from the abstract variation of constants formula that, for $\theta \in (-\infty, 0]$ and $t + \theta \geq 0$,

$$
\begin{aligned}
& |x^{p+\xi}(t+\theta) - x^p(t+\theta) - y(t+\theta)| \\
\leq\ & \int_0^{t+\theta} M^2 e^{\omega(t+\theta-s)}[o(||\xi|| + ||\xi'||) \\
& +|D_1F_1(x^p(s), a^p(s))(x^{p+\xi}(s) - x^p(s) - y(s))| + o(||\xi|| + ||\xi'||) \\
& +|D_2F_1(x^p(s), a^p(s))(a^{p+\xi}(s) - a^p(s) - b(s))| + o(||\xi||) \\
& +|DF_2(x^p(s))(x^{p+\xi}(s) - x^p(s) - y(s))| + o(||\xi||)]ds.
\end{aligned}
$$

Since $\sup_{s\in[0,t]} ||D_1F_1(x^p(s), a^p(s))||$, $\max_{s\in[0,t]} ||D_2F_1(x^p(s), a^p(s))||$, $\max_{s\in[0,t]} ||DF_2(x^p(s))|| < \infty$ (see (H7), Lemma 4, and (H4)), Step 5 follows.

Step 6. If $s \in [0, t]$, p, $p + \xi \in \hat{\Omega}_t$, *and* $\xi \in T_p\hat{M}$, *then* $||\hat{a}_s^{p+\xi} - \hat{a}_s^p - \hat{b}_s|| \leq o(||\xi|| + ||\xi'|| + ||A\xi_1||) + O(||\hat{x}_s^{p+\xi} - \hat{x}_s^p - \hat{y}_s||)$ *uniformly in* s *and pointwise in* p, *where the meaning of* $\hat{\cdot}$ *is same as in Step 5.*

Proceeding analogously as in Step 5 and using (H6), we have that

$$
\begin{aligned}
& a^{p+\xi}(s) - a^p(s) - b(s) \\
= \; & H(x_s^{p+\xi}, a_s^{p+\xi}) - H(x_s^p, a_s^p) - D_1H(x_s^p, a_s^p)y_s - D_2H(x_s^p, a_s^p)b_s \\
= \; & (D_1H(x_s^p, a_s^{p+\xi}) - D_1H(x_s^p, a_s^p))(x_s^{p+\xi} - x_s^p) \\
& + D_1H(x_s^p, a_s^p)(x_s^{p+\xi} - x_s^p - y_s) + \omega_1(x_s^{p+\xi} - x_s^p, x_s^p, a_s^{p+\xi}) \\
& + D_2H(x_s^p, a_s^p)(a_s^{p+\xi} - a_s^p - b_s) + \omega_2(a_s^{p+\xi} - a_s^p, x_s^p, a_s^p)
\end{aligned}
$$

for all $s \in [0,t]$. It follows from (H6) that the error term $|\omega_2(a_s^{p+\xi} - a_s^p, x_s^p, a_s^p)| \leq o(||\xi||)$ uniformly in $s \in [0,t]$ and pointwise in p. Arguing as in Step 5 and using (H6)*, we can obtain

$$
\begin{aligned}
& |\omega_1(x_s^{p+\xi} - x_s^p, x_s^p, a_s^{p+\xi})| \\
\leq \; & \int_0^1 |D_1H_1(x_s^p + \mu(x_s^{p+\xi} - x_s^p), a_s^{p+\xi})(x_s^{p+\xi} - x_s^p) \\
& - D_1H_1(x_s^p, a_s^{p+\xi})(x_s^{p+\xi} - x_s^p)| d\mu \\
\leq \; & \int_0^1 O(||\xi||)||x_s^{p+\xi} - x_s^p||_{C(I,D(A))} d\mu \\
\leq \; & O(||\xi||)||x_s^{p+\xi} - x_s^p||_{C(I,D(A))}.
\end{aligned}
$$

Now $||x_s^{p+\xi} - x_s^p||_{C(I,D(A))} = \max_{\theta \in I} |x^{p+\xi}(s+\theta) - x^p(s+\theta)| + |Ax^{p+\xi}(s+\theta) - Ax^p(s+\theta)| \leq ||\xi_1|| + ||A\xi_1|| + Z_p(||\xi|| + ||\xi'||)$. Hence $|\omega_1(x_s^{p+\xi} - x_s^p, x_s^p, a_s^{p+\xi})| \leq o(||A\xi_1|| + ||\xi|| + ||\xi'||)$ pointwise in p and uniformly in s. Similarly, using (H6)* and Step 1 gives $|(D_1H(x_s^p, a_s^{p+\xi}) - D_1H(x_s^p, a_s^p))(x_s^{p+\xi} - x_s^p)| \leq o(||A\xi_1|| + ||\xi|| + ||\xi'||)$

pointwise in p and uniformly in s. Then, for $\theta \in (-\infty, 0]$ with $s+\theta \geq 0$, we have

$$\begin{aligned} & |a^{p+\xi}(s+\theta) - a^p(s+\theta) - b(s+\theta)| \\ \leq \;& o(||\xi|| + ||A\xi_1|| + ||\xi'||) + C||x^{p+\xi}_{s+\theta} - x^p_{s+\theta}|| + f(R_1)||a^{p+\xi}_{s+\theta} - a^p_{s+\theta}|| \end{aligned}$$

holds pointwise in p and uniformly in s, where $C = \sup_{s\in[0,t]} ||D_1 H(x^p_s, a^p_s)|| < \infty$ is granted by Lemma 4. Hence $||\hat{a}^{p+\xi}_s - \hat{a}^p_s - \hat{b}_s|| \leq o(||\xi|| + ||\xi'|| + ||A\xi_1||) + O(||\hat{x}^{p+\xi}_s - \hat{x}^p_s||) + f(R_1)||\hat{a}^{p+\xi}_s - \hat{a}^p_s||$, which implies that $||\hat{a}^{p+\xi}_s - \hat{a}^p_s - \hat{b}_s|| \leq o(||\xi|| + ||\xi'|| + ||A\xi_1||) + O(||\hat{x}^{p+\xi}_s - \hat{x}^p_s||)$ holds uniformly in s and pointwise in p since $f(R_1) < 1$.

Step 7. If $t \geq 0$ and $p \in \hat{\Omega}_t$, then $D\hat{S}_t(p) \in \mathcal{L}(T_p\hat{M}^1, T_{\hat{S}_t(p)}M)$ exists and is given by $D\hat{S}_t(p)(\xi) = \begin{pmatrix} y_t \\ b_t \end{pmatrix}$ *for $\xi \in T_p\hat{M}^1$.*

From Steps 5 and 6, and Gronwall's inequality we see that if $p+\xi \in \hat{\Omega}_t$ and $\xi \in T\hat{M}_p$ then $||x^{p+\xi}_t - x^p_t - y_t|| \leq ||\hat{x}^{p+\xi}_t - \hat{x}^p_t - \hat{y}_t|| \leq o(||\xi|| + ||\xi'|| + ||A\xi_1||)$ holds pointwise in p. By Step 6 it follows that $||\hat{a}^{p+\xi}_t - \hat{a}^p_t - \hat{b}_t|| \leq o(||\xi|| + ||\xi'|| + ||A\xi_1||)$ also holds pointwise in p. This completes the proof of Step 7.

Step 8. $d\hat{S}_t(p,\gamma) = (\hat{S}_t(p), D\hat{S}_t(p)\gamma)$ is continuous with respect to the topologies stated in the hypothesis of this theorem.

Step 8 follows from Theorem 4.2 of [13] (concerning the continuity of $\hat{S}_t$), (H4), (H6), (H7), Lemma 4, Step 1, Step 4, and arguments similar to those used in Step 6, which involve trivial extensions of the relevant functions to $(-\infty, 0]$ by their values at $-h$, the fact that $f(R_1) < 1$, and Gronwall's inequality.

3.7 The Model Equations: Part Two

In this section we present an application of the general theory.

Consider the following class of scalar age structured models with threshold dependent age of maturity,

$$\begin{cases} \partial_t u(t,a) + \partial_a u(t,a) = -d(a)u(t,a), \\ u(t,0) = b(\int_{\tau(t)}^{m} \beta(\xi)u(t,\xi)d\xi), \\ \int_{t-\tau(t)}^{t} [\int_0^{\tau(\sigma)} u(\sigma,a)da + C]^{-1} d\sigma = T, \\ \begin{pmatrix} u_0 \\ \tau_0 \end{pmatrix} = \begin{pmatrix} \psi \\ \varphi \end{pmatrix} \in C^1([-a_m,0], {L^1}_+[0,m) \times \mathbf{R}^+), \end{cases} \tag{3.7.1}$$

where $t \geq 0$, $0 \leq a < m$, and $0 < a_m < m \leq \infty$. Here m represents the maximum age and a_m stands for the maximum juvenile age. We make the following assumptions.

(A1) $d : [0,m) \to \mathbf{R}^+$ and $\beta : [0,m) \to \mathbf{R}^+$ are bounded and continuous.

(A2) $b : \mathbf{R}^+ \to \mathbf{R}^+$ is C^2, b, b', and b'' are bounded, and $0 < \max_{x \in \mathbf{R}^+} b(x) \leq \theta$ for some $\theta > 0$.

(A3) $a_m = (R_0 + C)T < m \leq \infty$, where $R_0 = C(\frac{1}{\sqrt{T\theta}} - 1) > 0$.

Next we rewrite (3.7.1) as follows. Let $X = \mathbf{R} \times L^1([0,m),\mathbf{R})$ and define $A : D(A) \to X$ by

$$A \begin{pmatrix} 0 \\ x \end{pmatrix} = \begin{pmatrix} -x(0) \\ -x' \end{pmatrix} \qquad \text{for} \begin{pmatrix} 0 \\ x \end{pmatrix} \in D(A) = \{0\} \times W^{1,1}([0,m),\mathbf{R}).$$

Note that $X_0 = \overline{D(A)} = \{0\} \times L^1[0,m)$. It is well known that A satisfies (H1) (see, for instance, [19, 34]). Denote

$$C_0 = \left\{ \begin{pmatrix} 0 \\ \gamma \end{pmatrix} \in \{0\} \times L^1[0,m) \mid 0 \leq \gamma(a) \leq \theta \ a.e. \ a \in [0,m) \right\}$$

and

$$D(H) = \left\{ \begin{pmatrix} \psi \\ \varphi \end{pmatrix} \in C(I, C_0 \times K) \mid ||\psi|| \leq R_0 \right\},$$

where $K = [\frac{TC}{2}, a_m] \subset \mathbf{R}$ and $I = [-a_m, 0]$ for simplicity of notation.

As in Lemma 5.1 of [13], it follows that the relation $H : D(H) \to K$, which is given by $(\psi,\varphi,\alpha) \in H$ if and only if $\int_{-\alpha}^{0}[\int_0^{\varphi(\sigma)} \psi(\sigma,\xi)d\xi + C]^{-1}d\sigma = T$, is a function which satisfies the appropriate Lipschitz condition from (H2) with $f(Q) = \frac{(Q+C)^2T}{C^2}\theta$.

Let M_0 be as in (H3). We give $D(A) = \{0\} \times W^{1,1}[0,m)$ the graph norm, namely, $|\gamma| = |\gamma(0)| + |\gamma|_{L^1} + |\gamma'|_{L^1}$ for $\gamma \in D(A)$.

We would like to study the differentiability of the function H. To this end, let

$$\Gamma = \left\{ \begin{pmatrix} \psi \\ \varphi \end{pmatrix} \in C(I, L^1[0,m) \times [0,m)) \;\middle|\; \begin{array}{l} \|\psi\| < R_0 \text{ and, for } \sigma \in I, \\ \int_0^{\varphi(\sigma)} \psi(\sigma,\xi)d\xi > -C/2 \\ \text{and } 0 < \varphi(\sigma) < m \end{array} \right\}.$$

Define $G : (0, a_m) \times \Gamma \to \mathbf{R}$ by $G(\alpha, \psi, \varphi) = \int_{-\alpha}^{0}[\int_0^{\varphi(\sigma)} \psi(\sigma, \xi)d\xi + C]^{-1}d\sigma - T$. We first study the differentiability of G. We commence with some lemmas.

Lemma 5 *The set Γ is open in $C(I, L^1[0, m) \times \mathbf{R})$. In particular, $\hat{\Gamma} := \Gamma \cap C(I, D(A) \times \mathbf{R})$ is open in $C(I, D(A) \times \mathbf{R})$, where $D(A)$ is given the graph norm.*

Proof. Let $\begin{pmatrix} \psi \\ \varphi \end{pmatrix} \in \Gamma$. Fix a_1, $a_2 \in (0, m)$ such that $a_1 < \varphi(\sigma) < a_2$. We find some $r_1 > 0$ such that if $\gamma_1 \in C(I, \mathbf{R})$ with $||\gamma_1 - \varphi|| < r_1$ then $a_1 < \gamma_1(\sigma) < a_2$ for each $\sigma \in I$. The continuity of the map $I \ni \sigma \mapsto \int_0^{\varphi(\sigma)} \psi(\sigma, \xi)d\xi \in \mathbf{R}$ implies that we can find $r_2 > 0$ such that, for any $\sigma \in I$, if $|x - \int_0^{\varphi(\sigma)} \psi(\sigma, \xi)d\xi| < r_2$ then $x > -C/2$. We note that, for each $\begin{pmatrix} \gamma_1 \\ \gamma_2 \end{pmatrix} \in C(I, L^1[0, m) \times \mathbf{R})$ with $||\gamma_1 - \varphi|| < r_1$, we have $|\int_0^{\varphi(\sigma)} \psi(\sigma, \xi)d\xi - \int_0^{\gamma_1(\sigma)} \gamma_2(\sigma, \xi)d\xi| \leq ||\psi - \gamma_2|| + |\int_{\varphi(\sigma)}^{\gamma_1(\sigma)} \psi(\sigma, \xi)d\xi|$. Next observe that the map $\theta : I \times [a_1, a_2] \ni (\sigma, s) \mapsto |\int_{\varphi(\sigma)}^{s} \psi(\sigma, \xi)d\xi|$ is uniformly continuous. Then $|\int_0^{\varphi(\sigma)} \psi(\sigma, \xi)d\xi - \int_0^{\gamma_1(\sigma)} \gamma_2(\sigma, \xi)d\xi| \leq ||\psi - \gamma_2|| + |\theta(\sigma, \gamma_1(\sigma))|$. Note that $|\theta(\sigma, \gamma_1(\sigma)|$ converges to zero uniformly as $||\varphi - \gamma_1|| \to 0$. It follows that we can choose $||\varphi - \gamma_1|| + ||\psi - \gamma_2||$ small enough such that $|\int_0^{\varphi(\sigma)} \psi(\sigma, \xi)d\xi - \int_0^{\gamma_1(\sigma)} \gamma_2(\sigma, \xi)d\xi| < r_2$, which gives the desired result.

Lemma 6 (i) *The partial derivatives $D_1G(\alpha, \psi, \varphi)$ and*

$D_2G(\alpha, \psi, \varphi) \in \mathcal{L}(C(I, L^1[0, m)), \mathbf{R})$ exist in the Fréchet sense and are given respectively by $D_1G(\alpha, \psi, \varphi)1 = [\int_0^{\varphi(-\alpha)} \psi(-\alpha, \xi)d\xi + C]^{-1}$ and

$D_2G(\alpha,\psi,\varphi)\gamma = \int_{-\alpha}^{0} \frac{-\int_0^{\varphi(\sigma)} \gamma(\sigma,\xi)d\xi}{(\int_0^{\varphi(\sigma)} \psi(\sigma,\xi)d\xi + C)^2} d\sigma$ for $(\alpha,\psi,\varphi) \in (0,a_m) \times \Gamma$.

(ii) *The map $(0,a_m) \times \Gamma_\varphi \ni (\alpha,\psi) \mapsto D_{1,2}G(\alpha,\psi,\varphi) \in \mathcal{L}(\mathbf{R} \times C(I,L^1[0,m)),\mathbf{R})$ is continuous, where $\Gamma_\varphi = \{\psi \in C(I,L^1[0,m)) \mid (\psi,\varphi) \in \Gamma\}$.*

(iii) *The map G is continuously differentiable in the Fréchet sense on $(0,a_m) \times \hat{\Gamma}$ and $D_3G(\alpha,\psi,\varphi)\gamma = \int_{-\alpha}^{0} \frac{-\psi(\sigma,\varphi(\sigma))\gamma(\sigma)}{(\int_0^{\varphi(\sigma)} \psi(\sigma,\xi)d\xi + C)^2} d\sigma$, where $(0,a_m) \times \hat{\Gamma}$ inherits the norm from $\mathbf{R} \times C(I,D(A) \times \mathbf{R})$.*

(iv) *For $(\alpha,\psi,\varphi) \in \hat{\Gamma}$, the partial derivative $D_2G(\alpha,\psi,\varphi) \in \mathcal{L}(C(I,D(A)),\mathbf{R})$ has a bounded extension to $\mathcal{L}(C(I,L^1[0,m)),\mathbf{R})$, $L(\alpha,\psi,\varphi)$, and the map $(0,a_m) \times \hat{\Gamma} \times C(I,L^1[0,m)) \ni (\alpha,\psi,\varphi,\gamma) \mapsto L(\alpha,\psi,\varphi)\gamma \in \mathbf{R}$ is continuous, where $\hat{\Gamma}$ has the relative topology induced from $C(I,D(A) \times \mathbf{R})$.*

Proof. Let $(\alpha,\psi,\varphi) \in (0,a_m) \times \Gamma$. It follows from the continuity of $I \ni \sigma \mapsto \int_0^{\varphi(\sigma)} \psi(\sigma,\xi)d\xi \in \mathbf{R}$ that $D_1G(\alpha,\psi,\varphi)1 = [\int_0^{\varphi(-\alpha)} \psi(-\alpha,\xi)d\xi + C]^{-1}$ and it is easy to check that D_1G is continuous when Γ is given the relative topology from $\mathbf{R} \times C(I,L^1[0,m) \times \mathbf{R})$. We turn our attention to $D_2G(\alpha,\psi,\varphi)$. Define $y : (-\frac{C}{2},\infty) \to \mathbf{R}$ by $y(\sigma) = \frac{1}{\sigma + C}$, $l_\alpha : C(I,\mathbf{R}) \to \mathbf{R}$ by $l_\alpha(\gamma) = \int_{-\alpha}^{0} \gamma(\sigma)d\sigma$, $g_\varphi : C(I,L^1[0,m)) \to C(I,\mathbf{R})$ by $g_\varphi(\gamma)(\sigma) = \int_0^{\varphi(\sigma)} \gamma(\sigma,\xi)d\xi$, and $h : C(I,(-\frac{C}{2},\infty)) \to C(I,\mathbf{R})$ by $h(\gamma)(\sigma) = y(\gamma(\sigma))$. Then $G(\alpha,\psi,\varphi) = l_\alpha(h(g_\varphi(\psi))) - T$ and the chain rule gives us $D_2G(\alpha,\psi,\varphi)\gamma = Dl_\alpha(h(g_\varphi(\psi))Dh(g_\varphi(\psi))Dg_\varphi(\psi)\gamma = \int_{-\alpha}^{0} \frac{-\int_0^{\varphi(\sigma)} \gamma(\sigma,\xi)d\xi}{(\int_0^{\varphi(\sigma)} \psi(\sigma,\xi)d\xi + C)^2} d\sigma$. This completes the proof of item (i).

Next we verify item (ii). First, it is easy to check that D_1G is continuous on $(0, a_m) \times \Gamma$. Second, it is easily checked that when φ is fixed and (α, ψ) are allowed to vary, each of the linear operators in the latter composition vary continuously, which verifies item (ii).

To show (iii), given $(\alpha, \psi, \varphi) \in (0, a_m) \times \hat{\Gamma}$, define $g_\psi : C(I, (0, m)) \to C(I, \mathbf{R})$ by $g_\psi(\gamma)(\sigma) = \int_0^{\gamma(\sigma)} \psi(\sigma, \xi)d\xi$. Then $G(\alpha, \psi, \varphi) = l_\alpha(h(g_\psi(\varphi))) - T$ and hence $D_3G(\alpha, \psi, \varphi)\gamma = Dl_\alpha(h(g_\psi(\varphi))Dh(g_\psi(\varphi))Dg_\psi(\varphi)\gamma = \int_{-\alpha}^0 \frac{-\psi(\sigma,\varphi(\sigma))\gamma(\sigma)}{(\int_0^{\varphi(\sigma)} \psi(\sigma,\xi)d\xi + C)^2} d\sigma$. It is clear that the other two partial derivatives of G on $(0, a_m) \times \hat{\Gamma}$ (with respect to the stronger norm) are given by the same formulas as in (i). Since D_1G is continuous on $(0, a_m) \times \Gamma$, it suffices to check the continuity of D_2G and D_3G on $(0, a_m) \times \hat{\Gamma}$. We have $Dl_\alpha(h(g_\varphi(\psi))\gamma = \int_{-\alpha}^0 \gamma(\sigma)d\sigma$, which is clearly continuous. Furthermore, $(Dh(g_\varphi(\psi))\gamma)(\sigma) = \frac{-\gamma(\sigma)}{(\int_0^{\varphi(\sigma)} \psi(\sigma,\xi)d\xi + C)^2}$ is easily checked to be continuous in (ψ, φ) (even with respect to the weaker norm) and $Dg_\varphi(\psi)\gamma(\sigma) = \int_0^{\varphi(\sigma)} \gamma(\sigma, \xi)d\xi$ for $\gamma \in C(I, D(A))$. So, for $(\psi^i, \varphi^i) \in \hat{\Gamma}$ ($i = 1, 2$), we have

$$\begin{aligned}
|Dg_{\varphi^1}(\psi^1)\gamma(\sigma) - Dg_{\varphi^2}(\psi^2)\gamma(\sigma)| &= \left| \int_{\varphi^1(\sigma)}^{\varphi^2(\sigma)} \gamma(\sigma, \xi)d\xi \right| \\
&= \left| \int_{\varphi^1(\sigma)}^{\varphi^2(\sigma)} \gamma(\sigma, 0) + \int_0^\xi d_2\gamma(\sigma, \theta)d\theta \, d\xi \right| \\
&\leq ||\varphi^1 - \varphi^2|| \, ||\gamma||_{C(I,D(A))},
\end{aligned}$$

which shows that D_2G is continuous on $(0, a_m) \times \hat{\Gamma}$. Turning our attention to D_3G, it suffices to check that $Dg_\psi(\varphi)$ varies continuously in (ψ, φ). If $(\psi^i, \varphi^i) \in \hat{\Gamma}$ ($i = 1$, 2), then

$$
\begin{aligned}
& |Dg_{\psi^1}(\varphi^1)\gamma(\sigma) - Dg_{\psi^2}(\varphi^2)\gamma(\sigma)| \\
= \; & \left|(\psi^1(\sigma,\varphi^1(\sigma)) - \psi^2(\sigma,\varphi^2(\sigma)))\gamma(\sigma)\right| \\
\leq \; & \left(||\psi^1 - \psi^2||_{C(I,D(A))} + \left| \int_{\varphi^1(\sigma)}^{\varphi^2(\sigma)} d_2\psi^2(\sigma,\xi)d\xi \right| \right) ||\gamma||
\end{aligned}
$$

and it is now obvious that $Dg_\psi(\varphi)$ is continuous. This proves (iii).

The first part of (iv) follows from (i). In light of the above discussion, to complete the proof of (iv), it suffices to check that the map $\hat{\Gamma} \times C(I, L^1[0,m]) \ni (\psi, \varphi, \gamma) \mapsto Dg_\varphi(\psi)\gamma \in C(I, \mathbf{R})$ is continuous. For $(\psi^i, \varphi^i, \gamma^i) \in \hat{\Gamma} \times C(I, L^1[0,m])$ $(i = 1,\, 2)$, we have

$$
|Dg_{\varphi^1}(\psi^1)\gamma^1(\sigma) - Dg_{\varphi^2}(\psi^2)\gamma^2(\sigma)| \leq ||\gamma^1 - \gamma^2|| + \left| \int_{\varphi^1(\sigma)}^{\varphi^2(\sigma)} \gamma^2(\sigma,\xi)d\xi \right|
$$

and the desired result is now obvious.

Let $D(\hat{H}) := D(H) \cap C(I, D(A) \times \mathbf{R})$ and $\hat{H} = H|_{D(H)}$.

Lemma 7 (i) *The function $\hat{H} : D(\hat{H}) \to K$ can be extended to a continuously differentiable function $H_e : U \to K$, where U is an open subset of the Banach space $C(I, D(A) \times \mathbf{R})$.*

(ii) *For each $(\psi, \varphi) \in U$, $D_1 H_e(\psi, \varphi) \in \mathcal{L}(C(I, L^1[0,m)), \mathbf{R})$ exists as a relative Fréchet derivative on U and the map $U \times C(I, L^1[0,m)) \ni (\psi, \varphi, \gamma) \mapsto$*

$D_1H_e(\psi,\varphi)\gamma$ is continuous when U has the relative topology induced from $C(I, D(A)\times \mathbf{R})$.

Proof. For $\begin{pmatrix}\psi\\ \varphi\end{pmatrix} \in D(\hat{H})$, we have $(\hat{H}(\psi,\varphi),\psi,\varphi)\in(0,a_m)\times\hat{\Gamma}$, $G(H(\psi,\varphi),\psi,\varphi) = 0$, and $D_1G(H(\psi,\varphi),\psi,\varphi)\neq 0$. By Lemma 6(iii) and an application of the implicit function theorem, we can find an open set $U\subset\hat{\Gamma}$ in $C(I, D(A)\times\mathbf{R})$ and a C^1 extension $H_e : U\to(0,a_m)$. The image of H_e is actually contained in $K = [\frac{TC}{2}, a_m]$ by the definition of Γ and G. To verify (ii), fix $\varphi\in C(I,(0,m))$ and let $G_\varphi : (0,a_m)\times\Gamma_\varphi\to\mathbf{R}$ be given by $G_\varphi(\alpha,\psi) = G(\alpha,\psi,\varphi)$. Note that Γ_φ is defined in Lemma 6(ii) and it is open in $C(I, L^1[0,m))$. By Lemma 6(ii), we know that G_φ is C^1 in the Fréchet sense on $(0,a_m)\times\Gamma_\varphi$. Therefore, if $(\psi,\varphi)\in U$ then $\psi\in\Gamma_\varphi$, $G_\varphi(H_e(\psi,\varphi),\psi) = 0$, and $D_1G_\varphi(H_e(\psi,\varphi),\psi)\neq 0$. The implicit function theorem gives us an open set $U(\varphi)\subset\Gamma_\varphi$ of ψ and a C^1-function $H(\varphi) : U(\varphi)\to(0,a_m)$ satisfying $H(\varphi) = H_{e\varphi}$ on $U_\varphi\cap U(\varphi)$, where U_φ and $H_{e\varphi}$ are defined in the obvious way. Then for each $\xi\in C(I, L^1[0,m))$ such that $\psi+\xi\in U\cap U(\varphi)$ we have $H_e(\psi+\xi,\varphi) - H_e(\psi,\varphi) - DH(\varphi)(\xi) = H(\varphi)(\psi+\xi) - H(\varphi)(\psi) - DH(\varphi)(\xi) = o(\xi)$. This proves the first part of (ii). The continuity property stated in (ii) follows from the formula $D_1H_e(\psi,\varphi)\gamma = -D_1G(H_e(\psi,\varphi),\psi,\varphi)^{-1}D_2G(H_e(\psi,\varphi),\psi,\varphi))\gamma$, the continuity of D_1G (see proof of Lemma 6), and Lemma 6(iv).

It follows from Lemma 7 that

$$D_1H_e(\psi,\varphi)\gamma = \int_{-H_e(\psi,\varphi)}^{0} \frac{\int_0^{\varphi(\sigma)}\gamma(\sigma,\xi)d\xi}{(\int_0^{\varphi(\sigma)}\psi(\sigma,\xi)d\xi+C)^2} d\sigma (\int_0^{\varphi(-H_e(\psi,\varphi))}\psi(-H_e(\psi,\varphi),\xi)d\xi + C)$$

and

$$D_2H_e(\psi,\varphi)\gamma = \int_{-H_e(\psi,\varphi)}^{0} \frac{\psi(\sigma,\varphi(\sigma))\gamma(\sigma)}{(\int_0^{\varphi(\sigma)}\psi(\sigma,\xi)d\xi+C)^2} d\sigma (\int_0^{\varphi(-H_e(\psi,\varphi))}\psi(-H_e(\psi,\varphi),\xi)d\xi + C).$$

Then $\mathrm{rank}(D_1H_e(\psi,\varphi)) = 1$ and hence (H6) is verified. It is not difficult to check that, for $(\psi,\varphi) \in \hat{M}_0$, $DH(\psi,\varphi)(\psi',\varphi') = 1 - \frac{\int_0^{\varphi(-\varphi(0))}\psi(-\varphi(0),\xi)d\xi+C}{\int_0^{\varphi(0)}\psi(0,\xi)d\xi+C}$ and hence it is easy to verify the first statement in (H6)* by using this formula. The second statement of (H6)* can be checked using the above formula for D_1H.

Remark. This is the 'special property of the derivative of H' mentioned in the Future Work section of [13].

Define $F : C_0 \times K \to X$ by $F(x,a) = \begin{pmatrix} b(\int_a^m \beta(\xi)x(\xi)d\xi) \\ -d(\cdot)x(\cdot) \end{pmatrix}$. Then the verification of the subtangential condition (H5) with respect to C_0, K, and F follows exactly as in [34]. We write $F(x,a) = F_1(x,a) + F_2(x)$, where $F_1(x,a) = \begin{pmatrix} b(\int_a^m \beta(\xi)x(\xi)d\xi) \\ 0 \end{pmatrix}$ and $F_2(x) = \begin{pmatrix} 0 \\ -d(\cdot)x(\cdot) \end{pmatrix}$. Taking $X_1 = \mathbf{R} \times \{0\}$ and $X_2 = \{0\} \times L^1[0,m)$ gives $X = X_1 \oplus X_2$ and hypotheses (H7), (H7)*, and (H4) are easily verified.

Therefore, by Theorem 5, we have the following result.

Proposition 12 *In addition to (A1)–(A3), assume that* $\begin{pmatrix} \psi \\ \varphi \end{pmatrix} \in C^1(I, L^1_+ \times \mathbf{R}^+)$ *satisfies the following three conditions.*

(i) $\psi(0,0) = b(\int_{\varphi(0)}^m \beta(\xi)\psi(0,\xi)d\xi)$ *and* $\psi(0)(\cdot) \in W^{1,1}[0,m)$.

(ii) *For each* $\sigma \in I$, $0 \le \psi(\sigma)(a) \le \theta$ *for all* $a \in [0,m)$ *and* $\varphi(\sigma) \in [\frac{TC}{2}, a_m]$.

(iii) *For each* $\sigma \in I$, $\int_0^m \psi(\sigma)(a)da < C(\frac{1}{\sqrt{T\theta}} - 1)$ *and* $\int_{-\varphi(0)}^0 [\int_0^{\varphi(\sigma)} \psi(\sigma,\xi)d\xi + C]^{-1}d\sigma = T$.

Then the initial value problem (3.7.1) has a unique maximal solution $\begin{pmatrix} u(t,\cdot) \\ \tau(t) \end{pmatrix} \in C([-a_m, t_e), L^1[0,m] \times \mathbf{R})$ $(t_e > 0)$ *in* M_0*, such that* $t \mapsto u(t,\cdot) \in C^1([0,t_e), L^1[0,m])$, $\tau(t)$ *is locally Lipschitz on* $[0,t_e)$*, and* $\begin{pmatrix} u_0 \\ \tau_0 \end{pmatrix} = \begin{pmatrix} \psi \\ \varphi \end{pmatrix}$. *Moreover,*

(i) *For* $0 \le t < t_e$, $[0,m) \ni a \mapsto u(t,a)$ *is absolutely continuous, and for a.e.* $a \in [0,m)$,

$$\partial_t u(t,a) + \partial_a u(t,a) = -d(a)u(t,a) \qquad \text{for } 0 \le t < t_e,$$
$$u(t,0) = b\left(\int_{\tau(t)}^m \beta(\xi)u(t,\xi)d\xi\right).$$

(ii) *For* $0 \le t < t_e$, $\int_{t-\tau(t)}^t [\int_0^{\tau(\sigma)} u(\sigma,a)da + C]^{-1}d\sigma = T$.

(iii) *For* $t \in [0,t_e)$*, the "total population" satisfies* $\int_0^m u(t,a)da < C(\frac{1}{\sqrt{T\theta}} - 1)$ *and* $0 \le u(t,a) \le \theta$ *for each* $a \in [0,m)$.

Actually, we can say more about the differentiability of $\tau(t)$. From Proposition 12(ii), we note that, for a.e. $t \in [0, t_e)$,

$$\tau'(t) = 1 - \frac{\int_0^{\tau(t-\tau(t))} u(t-\tau(t), a)da + C}{\int_0^{\tau(t)} u(t,a)da + C}.$$

Since a Lipschitz function with continuous a.e. derivative is continuously differentiable, it follows that $\tau(t)$ is C^1 on $[0, t_e)$. Note that Proposition 12 also implies that the map $[0, t_e) \times [0, m) \ni (t, \theta) \mapsto \int_0^\theta u(t,a)da \in \mathbf{R}$ is C^1. Therefore, we obtain the following result.

Corollary 13 *Under the hypothesis of Proposition 12,* $\tau(t)$ *is* C^2 *on* $[0, t_e)$ *and* $\tau'(t) = 1 - \frac{\int_0^{\tau(t-\tau(t))} u(t-\tau(t),a)da+C}{\int_0^{\tau(t)} u(t,a)da+C}$.

This smoothing in time effect for the age of maturity function is caused by the fact that it satisfies an ODE with a state-dependent delay. The same is *not* true for the population density.

In order to derive the "integration along the characteristics" formula we make the following observations. Define $q : [-a_m, t_e) \times [0, m) \to \mathbf{R}^2$ by

$$q(t,a) = \begin{cases} \psi(0, a-t)\exp(-\int_{a-t}^{a} d(\theta)d\theta) & \text{if } 0 \le t \le a, \\ b(\int_{\tau(t-a)}^{m} \beta(\theta)u(t-a,\theta)d\theta)\exp(-\int_0^a d(\theta)d\theta) & \text{if } t \ge a, \\ \psi(t,a) & \text{if } t \in I. \end{cases}.$$

It is not difficult to check that $(q(t,\cdot), \tau(t))^t$ is a mild solution of (3.7.1) in M_0 on $[-a_m, t_e)$. By uniqueness, it follows that $q(t,\cdot) = u(t,\cdot)$ for $t \in [-a_m, t_e)$.

We conclude this discussion by noting that classical solutions to (3.7.1) in M_0, that is, solutions corresponding to initial conditions given in the hypothesis of Proposition 12, will be even more regular than the abstract semigroup theory tells us if we assume that the initialization $\psi(t, a)$ and the model paramemters in (A1)-(A3) are more regular. However, the population density can never become smoother than the initialization ψ, which is clear from the integration along the characterstics formula.

4 C^0-Extendable Banach Manifolds

The purpose of this section is to give an *intrinsic* definition of a new class of Banach manifolds and their submanifolds, which has appeared in this thesis concerning the phase space for our state-dependent delay system. The definition is general enough to include the "C^0-extendable submanifold" $\hat{M}$ introduced in Section 3.4. We begin by recalling the definition of a C^1 Banach manifold in the way it is given in [17].

4.1 C^1 Banach Manifolds

Definition 14 *Let X be a set. A C^1 atlas on X is a collection of pairs $(U_i, \varphi_i), i \in I$ (for some index set I) satisfying the following conditions:*

- ***(A1)*** *Each U_i is a subset of X and the collection $\{U_i\}_{i\in I}$ covers X.*
- ***(A2)*** *Each map φ_i is a bijection of U_i onto an open subset $\varphi_i(U_i)$ of some Banach space $\hat{E}_i$. For any i, j, $\varphi_i(U_i \cap U_j)$ is open in $\hat{E}_i$.*
- ***(A3)*** *For any i, j the map $\varphi_j \circ {\varphi_i}^{-1} : \varphi_i(U_i \cap U_j) \to \varphi_j(U_i \cap U_j)$ is a C^1-*

diffeomorphism.

Then the collection $\{U_i\}_{i\in I} \cup \{\varphi_j^{-1}(W) \mid W$ is an open subset of $\varphi(U_j)$ and $j \in I\}$ is a subbasis for a topology on X in which each U_i is open and for which each map $\varphi_i : U_i \to \varphi(U_i)$ is a homeomorphism. For each $i \in I$ the pair (U_i, φ_i) is called a chart of the atlas. Suppose that $U \subset X$ is open and that $\varphi : U \to U'$ is a topological isomorphism, where U' is an open subset of some Banach space. The pair (U, φ) is said to be compatible with the atlas (U_i, φ_i) if the conditions **(A2), (A3)** are satisfied when 'φ_j' is replaced with 'φ' and 'U_j' is replaced with 'U' for each $i \in I$. Two atlases are compatible if their charts are pairwise compatible. The compatibility relation on the collection of all C^1 atlases on X is an equivalence relation. Given a C^1 atlas on X, its equivalence class is said to define a C^1 smooth structure on X.

Definition 15 *A set X together with a C^1 smooth structure on X is called a C^1 Banach manifold.*

4.2 The C^0-Extendable Smooth Structure

We will define a new smooth structure useful for our purposes below.

Definition 16 *Suppose the collection (U_i, φ_i) is a C^1 atlas for X. If in addition to* **(A1)-(A3)** *from Definition 14, we assume the following:*

- ***(A4)*** *For each i the Banach space $(\hat{E}_i, ||\cdot||_{i,1})$ is contained in some larger Banach space $(E_i, ||\cdot||_{i,0})$ such that $||\cdot||_{i,0} \leq ||\cdot||_{i,1}$.*

- ***(A5)*** *For any i, j and any $p \in \varphi_i(U_i \cap U_j)$, $D(\varphi_j \circ {\varphi_i}^{-1})(p) \in \mathcal{L}(\hat{E}_i, \hat{E}_j)$ has a bounded extension $D_e(\varphi_j \circ {\varphi_i}^{-1})(p) \in \mathcal{L}(E_i, E_j)$ which is invertible.*

- ***(A6)*** *For any i, j the map $\varphi_i(U_i \cap U_j) \times E_i \ni (p, \gamma) \mapsto D_e(\varphi_j \circ {\varphi_i}^{-1})(p)\gamma \in E_j$ is continuous, where $\varphi_i(U_i \cap U_j)$ inherits the relative topology from $\hat{E}_i$. Furthermore, $\varphi_i(U_i \cap U_j) \times E_j \ni (p, \gamma) \mapsto [D_e(\varphi_j \circ {\varphi_i}^{-1})(p)]^{-1}\gamma \in E_i$ is continuous.*

We call (U_i, φ_i) a C^0-extendable atlas for X.

Proposition 17 *The relation* $(U_i, \varphi_i)_{i\in I} \simeq (V_j, \psi_j)_{j\in J}$ *if and only if for each* i, j *conditions* ***(A2),(A3),(A5),(A6)*** *hold (with* ψ_j *in place of* φ_j *and* V_j *in place of* U_j*), is an equivalence relation on the collection of* C^0*- extendable atlases for* X*.*

Proof. It is clear that $(U_i, \varphi_i) \simeq (U_i, \varphi_i)$. Now suppose $(U_i, \varphi_i)_{i\in I} \simeq (V_j, \psi_j)_{j\in J}$. It suffices to only check conditions **(A5)** and **(A6)**. We have that for each $i \in I$ and each $j \in J$, and any $p \in \varphi_i(U_i \cap V_j)$, $D(\psi_j \circ {\varphi_i}^{-1})(p) \in \mathcal{L}(\hat{E}_i, \hat{F}_j)$ has a bounded extension $D_e(\psi_j \circ {\varphi_i}^{-1})(p) \in \mathcal{L}(E_i, F_j)$ which is invertible. (Here $\psi_j : V_j \to \psi_j(V_j) \subset \hat{F}_j \subset F_j$). Let $q \in \psi_j(U_i \cap V_j)$ then $D(\varphi_i \circ {\psi_j}^{-1})(q) = [D(\psi_j \circ {\varphi_i}^{-1})(p)]^{-1}$ where $\psi_j \circ {\varphi_i}^{-1}(p) = q$. Let $L_e = D_e(\psi_j \circ {\varphi_i}^{-1})(p) \in \mathcal{L}(E_i, F_j)$. Then $(L_e)^{-1} \in \mathcal{L}(F_j, E_j)$ is an extension of $D(\varphi_i \circ {\psi_j}^{-1})(q)$ having the desired property, which verifies **(A5)**. The continuity property **(A6)** is trivial. This shows that $(V_j, \psi_j)_{j\in J} \simeq (U_i, \varphi_i)_{i\in I}$. Now suppose that $(U_i, \varphi_i)_{i\in I} \simeq (V_j, \psi_j)_{j\in J} \simeq (W_k, \beta_k)_{k\in K}$, where $\beta_k : W_k \to \beta_k(W_k) \subset \hat{G}_k \subset G_k$. Let $p \in U_i \cap W_k$ and choose some $V_j \ni p$. Then we have that $(\beta_k \circ \varphi_i^{-1})(p) = ((\beta_k \circ \psi_j^{-1}) \circ (\psi_j \circ \varphi_i^{-1}))(p)$ and we are granted bounded invertible extensions $D_e(\psi_j \circ \varphi_i^{-1})(p) \in \mathcal{L}(E_i, F_j)$ and $D_e(\beta_k \circ \psi_j^{-1})((\psi_j \circ \varphi_i^{-1})(p)) \in \mathcal{L}(F_j, G_k)$. We take $D_e(\beta_k \circ \varphi_i^{-1})(p)$ to be the composition of the latter two in the appropriate order. The continuity property from **(A6)** follows from the easily checked fact that the set $\varphi_i(U_i \cap V_j \cap W_j)$ is open in $\hat{E}_i$ and $\psi_j(U_i \cap V_j \cap W_j)$ is open in $\hat{F}_j$. (Recall that V_j depends on p).

Definition 18 *A set X together with a C^0-extendable smooth structure is called a C^0-extendable Banach manifold.*

4.3 C^0-Extendable Submanifolds

In this section we give the definition of a submanifold of a C^0-extendable Banach manifold, we then briefly point out why the examples encountered in this thesis are an instance of these.

Definition 19 *By a submanifold of a C^0-extendable Banach manifold X we mean a subset Y having the property that for each $y \in Y$ there is some chart (U, φ) of X at y such that φ is a homeomorphism of U onto a product $V_1 \times V_2$ where V_1, V_2 are open in some Banach spaces E_1, E_2, respectively, and such that $\varphi(Y \cap U) = V_1 \times \{0\}$.*

The remarks following Proposition 8 in Section 3.4 shows that the set $\hat{M}$ is a C^0-extendable submanifold of $C(I, D(A) \times \mathbf{R}^n)$ when the latter is viewed as a C^0-extendable Banach manifold whose smooth structure is generated by the trivial atlas, consisting only of the identity map on $C(I, D(A) \times \mathbf{R}^n)$. It should be pointed out that the solution manifolds appearing in other literature on state-dependent delay equations such as [35] are *not* an instance of our definition. The reason for this is that our examples are induced by a purely 'algebraic-delay' condition, whereas in [35] a derivative term is involved. Despite this difference among the latter examples, both share the property that the tangent spaces of the solution manifolds are embedded in a larger space with a weaker norm, and that this embedding is continuous.

5 Concluding Remarks

5.1 Summary

We have motivated an abstract algebraic-delay differential system in Chapter 1.2 via a two stage age structured population model, in which it is assumed that there is no competition between juvenilles and adults for resources. The key feature of the model is that the age of maturity at a given time depends on the history of the population density, as well as the history of itself. With the help of semigroup theory, we addressed the existence, positivity and continuity of the solution semiflow corresponding to the model equations. Finally, we showed that the solution semiflow is differentiable with respect to initial data, in a suitable weak sense, in Chapter 3.6. The latter was carried out by proving the well posedness of a formal linear variational system in Chapter 3.5, and then finding the right interpolation of spaces, for which the solution of the linear variational system is actually the derivative of the solution operator. In order to state the main differentiability result, Theorem 11 of 3.6, we used the notion of a C^0-extendable submanifold of a Banach space. In Chapter 4 we gave an intrinsic definition of a new class of Banach manifolds, and their submanifolds to facilitate the latter and briefly indicated how they differ with similar examples in other places in the literature on state-depdendent delay equations.

5.2 Future Research Directions

There are various research directions which can be pursued related to the work given in the present thesis. Since this work grants a linearization of an abstract solution semiflow, one can try to derive a principle of local linearized stability. See [34, 12, 3] for renditions of the linearized stability principle for semiflows generated by delay differential equations and structured population models, and [9] in case of ODEs with state-dependent delays. One can try to adpat a proof of the Hopf bifurcation theorem in [7] and the corrected version in [18], which exploits only C^1 smoothness of the solution semiflow, to the present case for the weak type of differentiability introduced here. The problem of the higher order differentiability (in the classical sense) for semiflows arising from even ODEs having state depdendent delays remains open, see [9]. The question of higher order differentiability (in a weak sense) with respect to parameters for a system of state-dependent delay equations is investigated in the work [2] (see also [10]). Perhaps a similar approach may be looked into here.

Bibliography

[1] J. Bélair. Population models with state-dependent delays. In *Second International Conference on Mathematical Population Dynamics*, pages 165–176. Marcel Dekker, 1991.

[2] Y. Chen, Q. Hu, and J. Wu. Second-order differentiability with respect to parameters for differential equations with adaptive delays. *Front. Math. China*, 5(2):221–286, 2010.

[3] O. Diekmann, Van Gils S.A., S.M.V. Lunel, and Walther H.O. *Delay Equations: Functional-,Complex-, and Nonlinear Analysis.* Applied Mathematical Sciences 110. Springer, 1995.

[4] R.D. Driver. A two body problem of classical electrodynamics: The one dimensional case. *Ann. Phys.*, 21:122–142, 1963.

[5] L. Evans. *Partial Differential Equations.* American Mathematical Society, 1998.

[6] K.J. Frogner. Variable developmental period: Intraspecific competition models with conditional age-specific maturity and mortality schedules. *Ecology*, 61:1099–1106, 1980.

[7] J.K. Hale. *Theory of Functional Differential Equations.* Springer-Verlag, 1977.

[8] J.K. Hale and S.M.V. Lunel. *Introduction to Functional Differential Equations.* Springer-Verlag, 1993.

[9] F. Hartung, T. Krisztin, H.O. Walther, and J. Wu. Functional differential equations with state-dependent delays: Theory and applications. In *HANDBOOK OF DIFFERENTIAL EQUATIONS Ordinary Differential Equations*, volume 3, pages 435–545. Elsevier, 2006.

[10] F. Hartung and J. Turi. On differentiability of solutions with respect to parameters for state-dependent delay equations. *J. Diff. Eq.*, 135:192–237, 1997.

[11] M.E. Hbid, M. Louihi, and E. Sanchez. A threshold state-dependent delayed functional equation arising from marine population dynamics: modelling and analysis. *J. Evol. Equ.*, 10:905–928, 2010.

[12] K. Korvasova. Linearized stability in case of state-dependent delay: a simple test example. Master's thesis, Universiteit Utrecht, 2011.

[13] N. Kosovalić, F.M.G. Magpantay, Y. Chen, and J. Wu. Abstract algebraic-delay differential systems and age structured population dynamics. *J. Diff. Eq.*, 255:593–609, 2013.

[14] T. Krisztin. An unstable manifold near a hyperbolic equilibrium for a class of differential equations with state-dependent delay. *Disc. and Cont. Dyn. Systems*, 9:993–1028, 2003.

[15] Y. Kuang and H.L. Smith. Periodic solutions of differential delay equations with threshold type delay. In *Oscillation & Dynamics in Delay Equations. Contemporary Mathematics*, volume 120, pages 153–176. Amer Mathematical Society, 1992.

[16] Y. Kuang and H.L. Smith. Slowly oscillating periodic solutions of autonomous state-dependent delay equations. *Nonlin. An. T.M.A.*, 19(9):855–872, 1992.

[17] S. Lang. *Differential Manifolds.* Springer-Verlag, 1985.

[18] B. Lani-Wayda. Hopf bifurcation for retarded functional differential equations and for semiflows in banach spaces. *J. Dyn. Diff. Eq.*, 25(4):1159–1199, 2013.

[19] P. Magal and S. Ruan. Center manifolds for semilinear equations with non dense domain and applications to Hopf bifurcation in age structured models. *Mem. Amer. Math. Soc.*, 202(951), 2009.

[20] P. Magal and S. Ruan. On semilinear cauchy problems with non-dense domain. *Adv. Diff. Equ.*, 14:1041–1084, 2009.

[21] J. Mahaffy, J. Bélair, and M. Mackey. Hematopoietic model with moving boundary condition and state dependent delay. *J. of Theor. Biology*, 190:135–146, 1998.

[22] J. Mallet-Paret and R.D. Nussbaum. Stability of periodic solutions of state-dependent delay-differential equations. *J. Diff. Eq.*, 250:4085–4103, 2011.

[23] J. Mallet-Paret, R.D. Nussbaum, and P. Paraskevopoulos. Periodic solutions for functional differential equations with multiple state-dependent time lags. *Top. Meth. Nonlin. Anal.*, 3, 1994.

[24] R.H. Jr. Martin and H. L. Smith. Abstract functional differential equations and reaction-diffusion systems. *Trans. Amer. Math. Soc.*, 321(1):1–44, 1990.

[25] J.A.J. Metz and O. Diekmann. *The Dynamics of Physiologically Structured Populations*. Lecture notes in Biomathematics 68. Springer-Verlag, New York, 1986.

[26] R. M. Nisbet and W.S.C. Gurney. The systematic formulation of population models for insects with dynamically varying instar duration. *Theor. Pop. Biology*, 23:114–135, 1983.

[27] A. Pazy. *Semigroups of linear operators and applications to partial differential equations*. Applied Mathematical Sciences 44. Springer, 1983.

[28] A.V. Rezounenko. Differential equations with discrete state-dependent delay: uniqueness and well-posedness in the space of continuous functions. *Nonlin. Anal.*, 70:3978–3986, 2009.

[29] A.V. Rezounenko. Non-linear partial differential equations with discrete state-dependent delays in a metric space. *Nonlin. Anal.*, 73:1707–1714, 2010.

[30] W. Ruess. Flow invariance for nonlinear partial differential delay equations. *Trans. Amer. Math. Soc.*, 361(8):4367–4403, 2009.

[31] J. Sieber. Finding periodic orbits in state-dependent delay differential equations as roots of algebraic equations. *Disc. and Cont. Dyn. Sys.*, 32(8):2607–2651, 2012.

[32] H. L. Smith. Reduction of structured population models to threshold-type delay equations and functional differential equations: A case study. *Math. Biosci.*, 113:1–23, 1993.

[33] J. Smoller. *Shock Waves and Reaction-Diffusion Equations*. A Series of Comprehensive Studies in Mathematics 258. Springer-Verlag, 1994.

[34] H.R. Thieme. Semiflows generated by Lipschitz perturbations of non-densely defined operators. *Diff. Int. Equ.*, 3:1035–1066, 1990.

[35] H.O. Walther. The solution manifold and C^1-smoothness of solution operators for differential equations with state dependent delay. *J. Diff. Equ.*, 195:46–65, 2003.

[36] H.O. Walther. Algebraic-delay differential systems, state dependent delay, and temporal order of reactions. *J. Dyn. Diff. Equ.*, 21(1):195–232, 2009.

[37] H.O. Walther. Differential equations with locally bounded delay. *J. Diff. Equ.*, 252:3001–3039, 2011.

[38] G.F. Webb. *Theory of Nonlinear Age-Dependent Population Dynamics*. Monographs and textbooks in pure and applied mathematics. Dekker, 1985.

[39] J. Wu. *Theory and Applications of Partial Functional Differential Equations*. Applied Mathematical Sciences 119. Springer, 1996.