TRANSACTIONS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 369, Number 5, May 2017, Pages 3387–3419
http://dx.doi.org/10.1090/tran/6760
Article electronically published on January 6, 2017

ALGEBRAIC-DELAY DIFFERENTIAL SYSTEMS: C^0-EXTENDABLE SUBMANIFOLDS AND LINEARIZATION

N. KOSOVALIĆ, Y. CHEN, AND J. WU

ABSTRACT. Consider the abstract algebraic-delay differential system,

$$\begin{aligned} x'(t) &= Ax(t) + F(x(t), a(t)), \\ a(t) &= H(x_t, a_t). \end{aligned}$$

Here A is a linear operator on $D(A) \subseteq X$ satisfying the Hille-Yosida conditions, $x(t) \in \overline{D(A)} \subseteq X$, $a(t) \in \mathbf{R}^n$, and X is a real Banach space. Let $C_0 \subseteq \overline{D(A)}$ be closed and convex, and $K \subseteq \mathbf{R}^n$ be a compact set contained in the ball of radius $h > 0$ centered at 0. Under suitable Lipschitz conditions on the nonlinearities F and H and a subtangential condition, the system generates a continuous semiflow on a subset of the space of continuous functions $C([-h, 0], C_0 \times \mathbf{R}^n)$, which is induced by the algebraic constraint. The object of this paper is to find conditions under which this semiflow is also differentiable with respect to initial data. In the motivating example coming from modelling the dynamics of an age structured population, the nonlinearities F and H are *not* Fréchet differentiable on the sets $C_0 \times K$ and $C([-h, 0], C_0 \times K)$, respectively. The main challenge of obtaining the differentiability of the semiflow is to determine the right type of differentiability and the right phase space. We develop a novel approach to address this problem which also shows how the spaces on which the derivatives of solution operators act reflect the model structure.

1. Introduction

1.1. **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 those 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, [21, 22]). This is possible because the nonlinearity appearing in the relevant equation is continuously differentiable on the

Received by the editors December 9, 2013 and, in revised form, May 8, 2015.

2010 *Mathematics Subject Classification.* Primary 34K05; Secondary 34A09, 92D25.

Key words and phrases. Linearization, nonlinear transport equation, state dependent delay, structured population dynamics, Banach manifold, interpolation space, functional differential equation.

The research of the second author was partially supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) and by the Early Researchers Award Program of Ontario (ERA).

The research of the third author was partially supported by NSERC and by the Canada Research Chairs Program (CRC).

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 [4, 8, 13, 19, 23–25] 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 [1, 3, 26]. 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 joint 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 [13] or more appropriately *extendable continuous differentiability* as in [19], 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 simple scalar equation above (see [25]).

Despite their emergence in modelling structured populations with developmental stages of variable length (see [5, 14, 20]), there are very few works dealing with differential equations containing both state dependent delays and partial differential operators. The works [16, 17] deal with special classes of reaction diffusion systems containing state dependent delays, but use special assumptions to circumvent the difficulties mentioned above.

In [7] and [12] the authors considered a 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:

$$\tag{1.1} \begin{cases} \partial_t u(t,a) + \partial_a u(t,a) = -d(a)u(t,a), \\ u(t,0) = b(\int_{\tau(t)}^{m} \beta(\xi)u(t,\xi)d\xi), \\ \int_{t-\tau(t)}^{t} [\int_0^{\tau(\sigma)} u(\sigma,a)da + C]^{-1} d\sigma = T, \\ \begin{pmatrix} u_0 \\ \tau_0 \end{pmatrix} = \begin{pmatrix} \psi \\ \varphi \end{pmatrix} \in C([-a_m,0], {L^1}_+[0,m) \times \mathbf{R}^+). \end{cases}$$

Here $t \geq 0$, $0 \leq a \leq m$, $a_m < m \leq \infty$, 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 is 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 ${L^1}_+[0,m) \times \mathbf{R}^+$. See [7] or [12] for a detailed derivation. 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 [7] that under suitable hypotheses, the algebraic-delay system (1.1) can be written abstractly as

$$\tag{1.2} \begin{aligned} \frac{d}{dt}\begin{pmatrix} 0 \\ u(t,\cdot) \end{pmatrix} &= \begin{pmatrix} -u(t,0) \\ -u_a(t,\cdot) \end{pmatrix} + \begin{pmatrix} b(\int_{\tau(t)}^{m} \beta(\xi)u(t,\xi)d\xi) \\ -d(\cdot)u(t,\cdot) \end{pmatrix}, \\ \tau(t) &= H(u_t, \tau_t), \\ \begin{pmatrix} x_0 \\ \tau_0 \end{pmatrix} &= \begin{pmatrix} \psi \\ \varphi \end{pmatrix} \in M_0. \end{aligned}$$

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 2 and Section 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 paper we establish sufficient conditions which ensure that this semiflow is also continuously differentiable.

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 has not been addressed yet and cannot be addressed by existing works. For system (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 (1.2) has the form (2.1). 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's clear that $D_2F(x(\cdot), a)$ is not defined for general $x \in L^1[0, m)$ and it's easy to see 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. [18] or [22] even indirectly since they require continuous differentiability of the nonlinear term, albeit on possibly thin subsets in [18]. Similarly, it will be seen in Section 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 [20] 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 (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 [6, 20] 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 (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 (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.

1.2. **Outline and main results.** Although our results are of a more general nature, for clarity, we outline the structure of this paper in terms of the model equations (1.1) and (1.2). The main goal of this paper is to prove Theorem 8 in Section 6.

In Section 2, we cover the basic functional analytic preliminaries and the precise meaning of mild solutions of system (1.1). This functional analytic setup is captured in the way that the first equation along with the nonlinear boundary condition for $u(t, 0)$ in system (1.1) is rewritten in system (1.2). This setup was motivated by the studies [10,11,22]. Moreover, the "subtangential condition" (H5) adopted from [22] ensures that the population density remains non-negative for non-negative initial data.

In Section 3, we address the differentiability with respect to time of solutions of system (1.1). The existence and uniqueness of solutions of system (1.1) and continuity of the corresponding semiflow were established in [7]. We take this opportunity to make some remarks about the work [7]. The hypothesis (H2) (along with (H5)) enables one to find solutions of (1.1) in M_0 for initial data in M_0. Two consequences are that in the model equations (1.1), the age of maturity function is a priori bounded, and that solutions are not necessarily global in time, unless the total population is small enough. 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., [15]. 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 (1.1).

Let $W^{1,1}[0, m)$ denote the space of absolutely continuous functions whose a.e. derivative lies in $L^1[0, m)$. In the motivating example, this space corresponds to the population density $u(t, \cdot)$. In Section 4 we show that the set $\hat{M}_0$ is contained in a C^1 submanifold of the space $C([-a_m, 0], W^{1,1}[0, m) \times \mathbf{R})$, $\hat{M}$, which is induced by the algebraic constraint in (1.2). We show that $\hat{M}$ has an atlas of manifold charts whose derivatives have the special extension properties discussed above. In particular, for each $p \in \hat{M}$, 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 5, we show that the (formal) linear variational system along flowlines in $\hat{M}_0$ can be solved uniquely for mild solutions for initial data belonging to the corresponding extended tangent space.

Section 6 develops the main results of this paper. We show that the solution operators $\hat{S}_t$ at time t (whose domain is the set of initial data in $\hat{M}_0$ with maximal interval of existence 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 7, all of the relevant hypotheses are verified for the motivating example system (1.1), and some abstract results are used to infer the regularity of its solutions. It should be noted that the main result of this paper, Theorem 8 in Section 6, on differentiability, differs from the classical differentiability of a function defined on a Banach manifold.

1.3. **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 (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 the following observations:

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

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 (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 8).

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 (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 8 in Section 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 $C^1([-a_m,0],L^1[0,m)\times\mathbf{R})\cap C([-a_m,0],W^{1,1}[0,m)\times\mathbf{R})$. We will see in Section 7 that the partial derivative D_1H of the other nonlinearity H has similar properties as D_1F_1 above.

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.

2.1. **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 $t\in[0,T)$.

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 regularity 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 [18]. 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 $||T(t)|| \leq Me^{\omega t}$.

(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 $M_0 \neq \emptyset$.

Given $D(A)$ as the graph norm, 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

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

(H4): Suppose $F : C_0 \times K \to X$ has the form $F(c, k) = F_1(c, k) + F_2(c)$, where $F_1 : C_0 \times K \to X_1$ and $F_2 : C_0 \to X_2$. 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 such 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 $(c,k)\in C_0\times K$,

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

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

Definition. Consider the following initial value problem:

$$\begin{cases} x'(t)=Ax(t)+F(x(t),a(t)),\\ a(t)=H(x_t,a_t),\\ \begin{pmatrix}x_0\\ a_0\end{pmatrix}=\begin{pmatrix}\psi\\ \varphi\end{pmatrix}\in M_0.\end{cases} \tag{2.1}$$

By a mild solution of (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 [22].

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 (2.1) has a unique maximal mild solution on $I\cup[0,t_e)$ in M_0 (see [7]). In this section we discuss the differentiability of these mild solutions with respect to time. In Theorem 1 below, we give sufficient conditions under which mild solutions are locally Lipschitz in time. This result is used to derive Theorem 2, 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 3. We end this section 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 1. *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 $||x_t|| \le R_1$ 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, $\hat{a}(\xi) = \begin{cases}\varphi(-h) & \text{if } \xi \le -h,\\ a(\xi) & \text{if } \xi \in [-h, T]\end{cases}$ and $\hat{x}(\xi) = \begin{cases}\psi(-h) & \text{if } \xi \le -h,\\ x(\xi) & \text{if } \xi \in [-h, T].\end{cases}$ 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. For any $l \ge 0$, $A\int_0^l T(s)\psi(0)ds = \lim_{\mu\to\infty}\int_0^l T(l-s)\mu R_\mu A\psi(0)ds$.

This follows easily from Lemma 1.8 of [22].

Step 2. There is $L > 0$ (depending possibly on T) such that $||\hat{x}_s - \hat{x}_0|| \le Ls$ and $||\hat{a}_s - \hat{a}_0|| \le Ls$ 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 $|x(s) - x(0)| \le Cs$ for some $C > 0$. Note that C may depend on T. Then, for each $\theta \le 0$, we have $|\hat{x}(s+\theta) - \hat{x}(\theta)| \le |x(s+\theta) - x(0)| + |x(0) - \hat{x}(\theta)| \le Cs + Lip(\psi)(-\theta) \le Cs + Lip(\psi)s$ if $s + \theta \ge 0$ and $|\hat{x}(s+\theta) - \hat{x}(\theta)| \le Lip(\psi)s$ if $s + \theta < 0$. These observations imply that $||\hat{x}_s - \hat{x}_0|| \le Ls$ for some $L > 0$ depending on T. Now we turn to $\hat{a}$. If $s + \theta < 0$, then $|\hat{a}(s+\theta) - \hat{a}(\theta)| \le Lip(\varphi)s$; if $s + \theta \ge 0$, then by (H2) there is some constant $J > 0$ depending on R_1 such that $|\hat{a}(s+\theta) - \hat{a}(\theta)| \le |a(s+\theta) - a(0)| + |a(0) - \hat{a}(\theta)| \le J||x_{s+\theta} - x_0|| + f(R_1)||a_{s+\theta} - a_0|| + Lip(\varphi)(-\theta)$. With $||x_{s+\theta} - x_0|| \le ||\hat{x}_{s+\theta} - \hat{x}_0|| \le Ls$ and $||a_{s+\theta} - a_0|| \le ||\hat{a}_s - \hat{a}_0|| + Lip(\varphi)s$, we get $|\hat{a}(s+\theta) - \hat{a}(\theta)| \le Qs + f(R_1)||\hat{a}_s - \hat{a}_0||$ for some $Q > 0$ depending on T. By virtue of $f(R_1) < 1$ we can conclude that $||\hat{a}_s - \hat{a}_0|| \le Ls$ for a possibly larger constant L than the one found before.

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

Given $\theta \le 0$, if $t + l + \theta \le 0$, then $|\hat{a}(t+l+\theta) - \hat{a}(t+\theta)| \le Lip(\varphi)l$; if $t + l + \theta \ge 0$ and $t + \theta < 0$, then using the result in Step 2 we have $|\hat{a}(t+l+\theta) - \hat{a}(t+\theta)| \le |a(t+l+\theta) - a(0)| + |a(0) - \hat{a}(t+\theta)| \le L|t+l+\theta| + Lip(\varphi)l \le Ll + Lip(\varphi)l = (L + Lip(\varphi))l$; if $t + l + \theta \ge 0$ and $t + \theta \ge 0$, then $|\hat{a}(t+l+\theta) - \hat{a}(t+\theta)| \le J||x_{t+l+\theta} - x_{t+\theta}|| + f(R_1)||a_{t+l+\theta} - a_{t+\theta}|| \le J||\hat{x}_{t+l} - \hat{x}_t|| + f(R_1)||\hat{a}_{t+l} - \hat{a}_t||$. The required result is now obvious since $f(R_1) < 1$.

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

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

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

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

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

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

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

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 [2]) applied to the function a. ☐

Theorem 2. *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 1, 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 $g(t) = d/dt F_1(x(t), a(t))$. 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 [22], 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 paper in Section 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 [23, 24] and in (H7) below.

Theorem 3. *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 2 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

$$b(s) = \begin{cases} \varphi'(s) & \text{if } s \in I, \\ D_1 H(x_s, a_s)x_s' + D_2 H(x_s, a_s)b_s & \text{if } s \in [0, T] \end{cases}$$

has a unique continuous solution, $b : I \cup [0, T] \to \mathbf{R}^n$, thanks to the contraction mapping principle since $||D_2 H(x_s, a_s)|| \le f(R_1) < 1$ 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 $t + r \le T$, we have

$$
\begin{aligned}
& a(t+r) - a(t) - rb(t) \\
= \; & H(x_{t+r}, a_{t+r}) - H(x_t, a_{t+r}) + H(x_t, a_{t+r}) - H(x_t, a_t) - rb(t) \\
= \; & D_1 H(x_t, a_{t+r})(x_{t+r} - x_t) - D_1 H(x_t, a_t) r x_t' + \omega_1(x_{t+r} - x_t, x_t, a_{t+r}) \\
& + D_2 H(x_t, a_t)(a_{t+r} - a_t - rb_t) + \omega_2(a_{t+r} - a_t, x_t, a_t),
\end{aligned} \tag{3.1}
$$

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 $g'(s) = D_1 H(x_s, a_{t+r}) x_s'$. Moreover,

$$
\begin{aligned}
& \omega_1(x_{t+r} - x_t, x_t, a_{t+r}) \\
= \; & g(t+r) - g(t) - D_1 H(x_t, a_{t+r})(x_{t+r} - x_t) \\
= \; & g'(t) r + \int_0^1 r(g'(t+sr) - g'(t)) ds - D_1 H(x_t, a_{t+r})(x_{t+r} - x_t) \\
= \; & D_1 H(x_t, a_{t+r})(x_t' r - x_{t+r} + x_t) \\
& + \int_0^1 \big(D_1 H(x_{t+sr}, a_{t+r}) x_{t+sr}' - D_1 H(x_t, a_{t+r}) x_t' ds \big) \, r.
\end{aligned}
$$

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 1 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 next 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

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

For I_1, using (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

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

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

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

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

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

Now, if $t+\theta>0$, then it follows from (3.1) with $t+\theta$ replacing t that

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

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 $s+r\le T$ and $r\in(0,h]$, we have

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

(note that $o(r)$ does not depend on s). Applying this to $s=t+\theta^*\ge 0$ 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 7, this will turn out to be merely an ordinary differential equation with a state dependent delay.

The following result follows immediately from Theorem 3.

Corollary 4. *Suppose* (H1)–(H6) *hold. The set* $\hat{M}_0 := \left\{ \begin{pmatrix} \psi \\ \varphi \end{pmatrix} \in D(\hat{H}) \cap M_0 \cap C^1(I, C_0 \times K) \,|\, \psi'(0) = A\psi(0) + F(\psi(0), \varphi(0)) \text{ and } \varphi'(0) = D_1H(\psi,\varphi)\psi' + D_2H(\psi,\varphi)\varphi' \right\}$ *is a positively invariant subset of* M_0 *for the semiflow* S.

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 (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 [7]. 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 next lemma is immediate from the fact that S is a semiflow on M_0 (see [7, Theorem 2]).

Lemma 1. *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* $0 \le s,\, t$ *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 notation 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).

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 the reader to the book by Lang [9].

Proposition 5. *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 5, 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 $C(I, D(A) \times \mathbf{R}^n)$.

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 (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 $\hat{S}(t, \Psi_0)$,

$$(5.1)\qquad \begin{cases} y'(t) &= Ay(t) + D_1F_1(x(t), a(t))y(t) \\ &\quad +D_2F_1(x(t), a(t))b(t) + DF_2(x(t))y(t), \quad t \in [0, t_e), \\ b(t) &= D_1H(x_t, a_t)y_t + D_2H(x_t, a_t)b_t, \\ \begin{pmatrix}y_0\\ b_0\end{pmatrix} &= \begin{pmatrix}\psi_1\\ \varphi_1\end{pmatrix}. \end{cases}$$

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 the one given in [4, 23].

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 5.

Definition. Suppose (H1)–(H7) hold. By a mild solution of (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

$$\begin{aligned} y(t) &= y(0) + A\int_0^t y(s)ds + \int_0^t \Big[D_1F_1(x(s), a(s))y(s) \\ &\quad + D_2F_1(x(s), a(s))b(s) + DF_2(x(s))y(s) \Big] ds. \end{aligned}$$

(iii) For $0 \leq t \leq T$, $\begin{pmatrix} y_t \\ b_t \end{pmatrix} \in TM_0^t$, *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 (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 $t_0 < T \leq t_e$.

It does *not* follow from [7] or other related works on partial functional differential equations such as [18, 22, 26] that (5.1) has a mild solution. We address this issue with the following lemma and proposition.

Lemma 2. *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 $||D_1F_1(c,k)||_{\mathcal{L}(X_0,X_1)} \le B$ 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 6. *Suppose* (H1)–(H7) *hold. If the initial data* $\begin{pmatrix}\psi_1\\ \varphi_1\end{pmatrix} \in TM_0^0$*, then* (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 $0 \le t_0 < t_e$ 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 (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

$$(5.2)\qquad \begin{cases} y'(t) = Ay(t) + D_1F_1(x(t),a(t))y(t) \\ \qquad\quad +D_2F_1(x(t),a(t))b(t) + DF_2(x(t))y(t), & t \in [t_0,T], \\ y_{t_0} = \psi \end{cases}$$

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 [22]. 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 2, there is some $B > 0$ such that $||D_1F_1(x(s),a(s))||_{\mathcal{L}(X_0,X_1)} \le B$ for each $s \in [t_0,T]$. By (H4) there is some $B' > 0$ such that $||DF_2(x(s))|| \le B'$ 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 (5.1), we let $\mathcal{A} : \mathcal{C} \to \mathcal{C}$ be given by

$$(5.3)\qquad (\mathcal{A}b)(t) = \begin{cases} \varphi(t-t_0) & \text{if } t \in [t_0-h,t_0], \\ D_1H(x_t,a_t)y(b)_t + D_2H(x_t,a_t)b_t & \text{if } t \in [t_0,T], \end{cases}$$

where $y(b)$ denotes the solution to (5.2). That $(\mathcal{A}b) \in \mathcal{C}$ follows from the continuity of the maps $[t_0, T] \ni t \mapsto y(b)_t \in C([t_0-h, t_0], X_0)$, $[t_0, T] \ni t \mapsto D_1H(x_t, a_t)y(b)_t \in \mathbf{R}^n$, and $[t_0, T] \ni t \mapsto D_2H(x_t, a_t) \in \mathcal{L}(C(I, \mathbf{R}^n), \mathbf{R}^n)$. 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 (5.3), (H6), (H2), and Lemma 2 we have for $t_0 < T_0 < t_e$ and $t \in [t_0, T_0]$ that

$$\begin{aligned} |(\mathcal{A}b_1)(t) - (\mathcal{A}b_2)(t)| &\leq |D_1H(x_t, a_t)(y(b_1)_t - y(b_2)_t)| \\ &\quad +|D_2H(x_t, a_t)((b_1)_t - (b_2)_t)| \\ &\leq C||y(b_1)_t - y(b_2)_t|| + f(R_1)||(b_1)_t - (b_2)_t||, \end{aligned} \tag{5.4}$$

where $C = \sup_{s\in[0,T_0]} ||D_1H(x_s, a_s)||$. Moreover, using the abstract variation of constants formula (see the Remark following (2.1)), Lemma 2, (H7)(iii), (H4), and (H1), we have

$$\begin{aligned} &|y(b_1)(t) - y(b_2)(t)| \\ = &\left| \lim_{\mu\to\infty} \int_{t_0}^{t} T(t-s)\mu R_\mu[D_1F_1(x(s), a(s))(y(b_1)(s) - y(b_2)(s)) \right. \\ &+D_2F_1(x(s), a(s))(b_1(s) - b_2(s)) \\ &\left. +DF_2(x(s))(y(b_1)(s) - y(b_2)(s))]ds \right| \\ \leq &\int_{t_0}^{t} M^2e^{\omega(t-s)}(C_1|y(b_1)(s) - y(b_2)(s)| + C_2||b_1 - b_2||)ds \\ \leq &M^2C_2e^{\omega T_0}||b_1 - b_2||t + \int_{t_0}^{t} M^2e^{\omega(t-s)}C_1|y(b_1)(s) - y(b_2)(s)|ds, \end{aligned} \tag{5.5}$$

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 (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 (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 (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 (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 (5.5), we have

$$|y_1(t) - y_2(t)| \leq \int_0^t M^2 e^{\omega(t-s)}(C_1|y_1(s) - y_2(s)| + C_2|b_1(s) - b_2(s)|)ds \tag{5.6}$$

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

$$||(\hat{b}_1)_t - (\hat{b}_2)_t|| \leq C(1 - f(R_1))^{-1}||(\hat{y}_1)_t - (\hat{y}_2)_t||. \tag{5.7}$$

Combining (5.6) with (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 (5.7) that $b_1 = b_2$ on $I \cup [0,T]$. As T is arbitrary, this completes Step 2.

Step 3. Let

$$t'_e(\psi_1,\varphi_1) := \sup\{\rho \in (0,t_e) \mid (5.1) \text{ has a mild solution } (y,b) \text{ on } I \cup [0,\rho]\}.$$

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 2, (H7)(iii), and (H4) that

$$\max \left\{ \begin{array}{l} \sup_{s\in[0,\rho_0]} ||D_1F_1(x(s),a(s))||, \\ \max_{s\in[0,\rho_0]} ||DF_2(x(s))||, \\ \max_{s\in[0,\rho_0]} ||D_2F_1(x(s),a(s))|| \end{array} \right\} < \infty.$$

By Lemma 2, $C' := \sup_{s\in[0,\rho_0]} ||D_1H(x_s,a_s)|| < \infty$. Let $(y,b) : I \cup [0,\rho_0) \to X_0 \times \mathbf{R}^n$ be the mild solution of (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 $||\hat{b}_t|| \leq (1-f(R_1))^{-1}||\varphi_1|| + (1 - f(R_1))^{-1}C'||\hat{y}_t||$ for each $t \in [0,\rho_0)$. Therefore, there is some $C'' > 0$ (which depends on ρ_0) such that $|y(t)| \leq Me^{\omega t}|\psi_1(0)| + C'' \int_0^t (|y(s)| + |b(s)|)ds \leq Me^{\omega t}|\psi_1(0)| + C'' \int_0^t (||\hat{y}_s|| + ||\hat{b}_s||)ds$ 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 6, we can obtain the following result.

Corollary 7. *Suppose* (H1)–(H7) *hold. For* $\Psi \in TM_0^0$*, let* $\begin{pmatrix} y \\ b \end{pmatrix}$ *be the corresponding mild solution to* (5.1) *on* $I \cup [0,t_e)$*. Then, for each* $T \in [0,t_e)$ *and* $t \in [0,T]$*,* $\left\| \begin{pmatrix} y_t \\ b_t \end{pmatrix} \right\| \le C||\Psi||$*, where* C *depends on* Ψ_0 *and* T*. (Recall that* Ψ_0 *is fixed throughout this section.)*

6. Derivatives of solution operators $\hat{S}_t$ on $\hat{M}_0$

Recall from Section 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 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]

[1]Throughout this section we use the following notation for a function g defined on an appropriate subset of the product of two normed spaces and whose image is contained in another normed space: $|g(\xi,w)| \le o(|\xi|)$ means that $\lim_{\xi\to 0} \frac{|g(\xi,w)|}{|\xi|} = 0$ pointwise. Similarly, $|g(\xi,w)| \le O(|\xi|)$

(H6)*: In addition to (H6), we further assume that, for $(\psi_0, \varphi_0), (\psi, \varphi) \in \hat{M}_0$,

$$|DH(\psi,\varphi)(\psi',\varphi') - DH(\psi_0,\varphi_0)(\psi_0',\varphi_0')| \\ \le O(||\psi-\psi_0|| + ||\varphi-\varphi_0||) + O(||\varphi-\varphi_0||)||\psi_0'|| + O(||\varphi-\varphi_0||)||\varphi_0'||$$

and

$$||D_1H(\psi,\varphi) - D_1H(\psi_0,\varphi_0)||_{\mathcal{L}(C(I,D(A)),\mathbf{R}^n)} \le O(||\psi-\psi_0|| + ||\varphi-\varphi_0||)$$

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

$$||D_2F_1(c,k) - D_2F_1(c_0,k_0)||_{\mathcal{L}(\mathbf{R}^n,X_1)} \\ \le O(|c-c_0|_{D(A)}) + Z_{(c_0,k_0)}(|c-c_0| + |k-k_0|).$$

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:

$$||D_1F_1(c,k) - D_1F_1(c_0,k_0)||_{\mathcal{L}(D(A),X_1)} \le O(|c-c_0| + |k-k_0|)$$

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 8. *Assume* (H1)–(H5), (H6)**, and* (H7)* *hold. Then the function $\hat{S}_t : \hat{\Omega}_t \to \hat{M}_0$ 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 (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 (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 7 that $\left\|\begin{pmatrix}y_t\\ b_t\end{pmatrix}\right\| \le C\left\|\begin{pmatrix}\xi_1\\ \xi_2\end{pmatrix}\right\|$ for a constant $C > 0$ depending on p and t.

means $|g(\xi,w)| \le C|\xi|$ 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.

In case $\xi \in T_p\hat{M}^1$, we note that $\left\| \begin{pmatrix} y_t \\ b_t \end{pmatrix} \right\| \le C||\xi||_{T_p\hat{M}^1}$. 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, $p+\xi \in \hat{\Omega}_t$, $||x_s^{p+\xi} - x_s^p|| + ||a_s^{p+\xi} - a_s^p|| \le O(||\xi||)$ 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 6 or Step 1 in the proof of Theorem 2 in [7]).

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 (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 2 and by Theorem 3, 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 3 (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 2 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 2 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

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

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 $|\omega_3(x^{p+\xi}(s) - x^p(s), x^p(s))| \leq o(||\xi||)$ uniformly in $s \in [0,t]$. Similarly, by Step 1 and (H7), we know that $|\omega_2(a^{p+\xi}(s) - a^p(s), x^p(s), a^p(s))| \leq o(||\xi||)$ 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

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

Note carefully how the second to last inequality follows from (H7)* and Step 1, and that the constant coming from the latter big O depends only on p. Then $|x^{p+\xi}(s) - x^p(s)|_{D(A)} \leq O(||\xi||) + |\dot{x}^{p+\xi}(s) - \dot{x}^p(s)| \leq Z_p(||\xi|| + ||\xi'||)$, where the first inequality follows from the argument starting with "Note that by Theorem 2 ..." in Step 4 and the second inequality follows from Step 4. This shows that $|\omega_1(x^{p+\xi}(s) - x^p(s), x^p(s), a^{p+\xi}(s))| \leq o(||\xi|| + ||\xi'||)$ 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 2, 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 the 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

$$\begin{aligned} ||x_s^{p+\xi} - x_s^p||_{C(I, D(A))} &= \max_{\theta \in I} |x^{p+\xi}(s+\theta) - x^p(s+\theta)| + |Ax^{p+\xi}(s+\theta) - Ax^p(s+\theta)| \\ &\leq ||\xi_1|| + ||A\xi_1|| + Z_p(||\xi|| + ||\xi'||). \end{aligned}$$

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 that

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

holds pointwise in p and uniformly in s, where $C = \sup_{s\in[0,t]} ||D_1H(x_s^p, a_s^p)|| < \infty$ is granted by Lemma 2. Hence $||\hat{a}_s^{p+\xi} - \hat{a}_s^p - \hat{b}_s|| \leq o(||\xi|| + ||\xi'|| + ||A\xi_1||) + O(||\hat{x}_s^{p+\xi} - \hat{x}_s^p||) + f(R_1)||\hat{a}_s^{p+\xi} - \hat{a}_s^p||$, which implies that $||\hat{a}_s^{p+\xi} - \hat{a}_s^p - \hat{b}_s|| \leq o(||\xi|| + ||\xi'|| + ||A\xi_1||) + O(||\hat{x}_s^{p+\xi} - \hat{x}_s^p||)$ 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 $||x_t^{p+\xi} - x_t^p - y_t|| \leq ||\hat{x}_t^{p+\xi} - \hat{x}_t^p - \hat{y}_t|| \leq o(||\xi|| + ||\xi'|| + ||A\xi_1||)$ holds pointwise in p. By Step 6 it follows that $||\hat{a}_t^{p+\xi} - \hat{a}_t^p - \hat{b}_t|| \leq o(||\xi|| + ||\xi'|| + ||A\xi_1||)$ 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 [7] (concerning the continuity of $\hat{S}_t$), (H4), (H6), (H7), Lemma 2, 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. □

7. An application

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

Consider the following class of scalar age structured model with threshold dependent age of maturity:

$$(7.1)\qquad \begin{cases} \partial_t u(t,a) + \partial_a u(t,a) = -d(a)u(t,a), \\ u(t,0) = b(\int_{\tau(t)}^m \beta(\xi)u(t,\xi)d\xi), \\ \int_{t-\tau(t)}^t [\int_0^{\tau(\sigma)} u(\sigma,a)da + C]^{-1}d\sigma = T, \\ \begin{pmatrix} u_0 \\ \tau_0 \end{pmatrix} = \begin{pmatrix} \psi \\ \varphi \end{pmatrix} \in C^1([-a_m,0], {L^1}_+[0,m) \times \mathbf{R}^+), \end{cases}$$

where $t \geq 0$, $0 \leq a \leq m$, 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 (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, [10, 22]). 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 [7], 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 3. *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 $| \int_0^{\varphi(\sigma)} \psi(\sigma,\xi)d\xi - \int_0^{\gamma_1(\sigma)} \gamma_2(\sigma,\xi)d\xi| \le ||\psi - \gamma_2|| + | \int_{\varphi(\sigma)}^{\gamma_1(\sigma)} \psi(\sigma,\xi)d\xi|$. 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 $| \int_0^{\varphi(\sigma)} \psi(\sigma,\xi)d\xi - \int_0^{\gamma_1(\sigma)} \gamma_2(\sigma,\xi)d\xi| \le ||\psi - \gamma_2|| + |\theta(\sigma, \gamma_1(\sigma))|$. 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 4. *We have:*

(i) *The partial derivatives $D_1G(\alpha, \psi, \varphi)$, $D_2G(\alpha, \psi, \varphi) \in \mathcal{L}(C(I, L^1[0,m)), \mathbf{R})$ 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

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

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

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

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

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

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 5. (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* $U \times C(I, L^1[0, m)) \ni (\psi, \varphi, \gamma) \mapsto D_1 H_e(\psi, \varphi)\gamma$ *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 $D_1 G(H(\psi, \varphi), \psi, \varphi) \neq 0$. By Lemma 4(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 4(ii) and it is open in $C(I, L^1[0, m))$. By Lemma 4(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 $D_1 G_\varphi(H_e(\psi, \varphi), \psi) \neq 0$. 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_φ} 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 $D_1 H_e(\psi, \varphi)\gamma = -D_1 G(H_e(\psi, \varphi), \psi, \varphi)^{-1} D_2 G(H_e(\psi, \varphi), \psi, \varphi))\gamma$, the continuity of $D_1 G$ (see proof of Lemma 4), and Lemma 4(iv). □

It follows from Lemma 5 that

$$D_1 H_e(\psi, \varphi)\gamma = \int_{-H_e(\psi,\varphi)}^{0} \frac{\int_0^{\varphi(\sigma)} \gamma(\sigma, \xi) d\xi}{(\int_0^{\varphi(\sigma)} \psi(\sigma, \xi) d\xi + C)^2} d\sigma \times \left(\int_0^{\varphi(-H_e(\psi,\varphi))} \psi(-H_e(\psi, \varphi), \xi) d\xi + C \right)$$

and

$$D_2 H_e(\psi, \varphi)\gamma = \int_{-H_e(\psi,\varphi)}^{0} \frac{\psi(\sigma, \varphi(\sigma))\gamma(\sigma)}{(\int_0^{\varphi(\sigma)} \psi(\sigma, \xi) d\xi + C)^2} d\sigma \times \left(\int_0^{\varphi(-H_e(\psi,\varphi))} \psi(-H_e(\psi, \varphi), \xi) d\xi + C \right).$$

Then $\operatorname{rank}(D_1 H_e(\psi, \varphi)) = 1$, 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 $D_1 H$.

Remark. This is the "special property of the derivative of H" mentioned in the Future Work section of [7].

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 [22]. 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 2, we have the following result.

Proposition 9. *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* $\sigma \in I$,

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

Then the initial value problem (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 9(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 9 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 10. *Under the hypothesis of Proposition* 9, $\tau(t)$ *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 (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 (7.1) in M_0, that is, solutions corresponding to initial conditions given in the hypothesis of Proposition 9, will be even more regular than what the abstract semigroup theory tells us if we assume that the initialization $\psi(t,a)$ and the model parameters 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.

Acknowledgement

The authors would like to thank the anonymous referee for helpful comments which improved the presentation of this article.

References

[1] Odo Diekmann, Stephan A. van Gils, Sjoerd M. Verduyn Lunel, and Hans-Otto Walther, *Delay equations: Functional, complex, and nonlinear analysis*, Applied Mathematical Sciences, vol. 110, Springer-Verlag, New York, 1995. MR1345150 (97a:34001)

[2] Lawrence C. Evans, *Partial differential equations*, Graduate Studies in Mathematics, vol. 19, American Mathematical Society, Providence, RI, 1998. MR1625845 (99e:35001)

[3] Jack Hale, *Theory of functional differential equations*, 2nd ed., Applied Mathematical Sciences, vol. 3, Springer-Verlag, New York-Heidelberg, 1977. MR0508721 (58 #22904)

[4] Ferenc Hartung, Tibor Krisztin, Hans-Otto Walther, and Jianhong Wu, *Functional differential equations with state-dependent delays: theory and applications*, Handbook of differential equations: ordinary differential equations. Vol. III, Handb. Differ. Equ., Elsevier/North-Holland, Amsterdam, 2006, pp. 435–545, DOI 10.1016/S1874-5725(06)80009-X. MR2457636 (2010b:34165)

[5] M. L. Hbid, M. Louihi, and E. Sánchez, *A threshold state-dependent delayed functional equation arising from marine population dynamics: modelling and analysis*, J. Evol. Equ. **10** (2010), no. 4, 905–928, DOI 10.1007/s00028-010-0075-x. MR2737164 (2012d:35382)

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

[7] N. Kosovalić, F. M. G. Magpantay, Y. Chen, and J. Wu, *Abstract algebraic-delay differential systems and age structured population dynamics*, J. Differential Equations **255** (2013), no. 3, 593–609, DOI 10.1016/j.jde.2013.04.025. MR3053479

[8] Tibor Krisztin, *A local unstable manifold for differential equations with state-dependent delay*, Discrete Contin. Dyn. Syst. **9** (2003), no. 4, 993–1028, DOI 10.3934/dcds.2003.9.993. MR1975366 (2004d:34155)

[9] Serge Lang, *Differential manifolds*, 2nd ed., Springer-Verlag, New York, 1985. MR772023 (85m:58001)

[10] Pierre Magal and Shigui Ruan, *Center manifolds for semilinear equations with non-dense domain and applications to Hopf bifurcation in age structured models*, Mem. Amer. Math. Soc. **202** (2009), no. 951, vi+71, DOI 10.1090/S0065-9266-09-00568-7. MR2559965 (2012c:37163)

[11] Pierre Magal and Shigui Ruan, *On semilinear Cauchy problems with non-dense domain*, Adv. Differential Equations **14** (2009), no. 11-12, 1041–1084. MR2560868 (2010h:35407)

[12] F. M. G. Magpantay, N. Kosovalić, and J. Wu, *An age-structured population model with state-dependent delay: derivation and numerical integration*, SIAM J. Numer. Anal. **52** (2014), no. 2, 735–756, DOI 10.1137/120903622. MR3187671

[13] John Mallet-Paret, Roger D. Nussbaum, and Panagiotis Paraskevopoulos, *Periodic solutions for functional-differential equations with multiple state-dependent time lags*, Topol. Methods Nonlinear Anal. **3** (1994), no. 1, 101–162. MR1272890 (96e:34119)

[14] R. M. Nisbet and W. S. C. Gurney, *The systematic formulation of population models for insects with dynamically varying instar duration*, Theoret. Population Biol. **23** (1983), no. 1, 114–135, DOI 10.1016/0040-5809(83)90008-4. MR700823 (84f:92046)

[15] A. Pazy, *Semigroups of linear operators and applications to partial differential equations*, Applied Mathematical Sciences, vol. 44, Springer-Verlag, New York, 1983. MR710486 (85g:47061)

[16] Alexander V. Rezounenko, *Differential equations with discrete state-dependent delay: uniqueness and well-posedness in the space of continuous functions*, Nonlinear Anal. **70** (2009), no. 11, 3978–3986, DOI 10.1016/j.na.2008.08.006. MR2515314 (2010c:35017)

[17] Alexander V. Rezounenko, *Non-linear partial differential equations with discrete state-dependent delays in a metric space*, Nonlinear Anal. **73** (2010), no. 6, 1707–1714, DOI 10.1016/j.na.2010.05.005. MR2661353 (2011j:35245)

[18] Wolfgang M. Ruess, *Flow invariance for nonlinear partial differential delay equations*, Trans. Amer. Math. Soc. **361** (2009), no. 8, 4367–4403, DOI 10.1090/S0002-9947-09-04833-8. MR2500891 (2010b:35475)

[19] Jan Sieber, *Finding periodic orbits in state-dependent delay differential equations as roots of algebraic equations*, Discrete Contin. Dyn. Syst. **32** (2012), no. 8, 2607–2651, DOI 10.3934/dcds.2012.32.2607. MR2903983

[20] Hal L. Smith, *Reduction of structured population models to threshold-type delay equations and functional-differential equations: a case study*, Math. Biosci. **113** (1993), no. 1, 1–23, DOI 10.1016/0025-5564(93)90006-V. MR1201765 (93i:92016)

[21] Joel Smoller, *Shock waves and reaction-diffusion equations*, 2nd ed., Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 258, Springer-Verlag, New York, 1994. MR1301779 (95g:35002)

[22] Horst R. Thieme, *Semiflows generated by Lipschitz perturbations of non-densely defined operators*, Differential Integral Equations **3** (1990), no. 6, 1035–1066. MR1073056 (92e:47121)

[23] Hans-Otto Walther, *The solution manifold and C^1-smoothness for differential equations with state-dependent delay*, J. Differential Equations **195** (2003), no. 1, 46–65, DOI 10.1016/j.jde.2003.07.001. MR2019242 (2006b:34169)

[24] Hans-Otto Walther, *Algebraic-delay differential systems, state-dependent delay, and temporal order of reactions*, J. Dynam. Differential Equations **21** (2009), no. 1, 195–232, DOI 10.1007/s10884-009-9129-6. MR2482014 (2009m:34143)

[25] Hans-Otto Walther, *Differential equations with locally bounded delay*, J. Differential Equations **252** (2012), no. 4, 3001–3039, DOI 10.1016/j.jde.2011.11.004. MR2871791

[26] Jianhong Wu, *Theory and applications of partial functional-differential equations*, Applied Mathematical Sciences, vol. 119, Springer-Verlag, New York, 1996. MR1415838 (98a:35135)

Department of Mathematics and Statistics, York University, 4700 Keele Street, Toronto, Ontario, Canada M3J 1P3

Current address: Department of Mathematics and Statistics, University of South Alabama, 411 University Boulevard N, Mobile, Alabama 36688-0002

E-mail address: kosovalic@southalabama.edu

Department of Mathematics, Wilfrid Laurier University, 75 University Avenue West, Waterloo, Ontario, Canada N2L 3C5

E-mail address: ychen@wlu.ca

Department of Mathematics and Statistics, York University, 4700 Keele Street, Toronto, Ontario, Canada M3J 1P3

E-mail address: wujh@mathstat.yorku.ca