ACADEMIC PRESS

Available at WWW.MATHEMATICSWEB.ORG POWERED BY SCIENCE @ DIRECT•

J. Differential Equations 190 (2003) 364–406

Journal of Differential Equations

http://www.elsevier.com/locate/jde

Inertial and slow manifolds for delay equations with small delays ☆

Carmen Chicone*

Department of Mathematics, University of Missouri, Columbia, MO 65211, USA

Received December 20, 2001; revised June 11, 2002

Abstract

Yu.A. Ryabov and R.D. Driver proved that delay equations with small delays have Lipschitz inertial manifolds. We prove that these manifolds are smooth. In addition, we show that expansion in the small delay can be used to obtain the dynamical system on the inertial manifold. This justifies "post-Newtonian" approximation for delay equations.

MSC: 34K19

Keywords: Delay equation; Functional equation; Inertial manifold; Singular perturbation

1. Introduction

Ryabov [23–27] and Driver [7,8] studied retarded functional differential equations (RFDEs) with small retardations. For a globally Lipschitz RFDE with a sufficiently small retardation, they proved the existence of a finite-dimensional manifold of "special solutions" that attracts, exponentially fast, all solutions in the infinite-dimensional phase space of the RFDE. We will show that the manifold of special solutions is smooth. Thus, Ryabov's special solutions form an inertial manifold in the phase space of the RFDE. In particular, the long-term dynamics of the RFDE is determined by a (smooth) ordinary differential equation (ODE) on the inertial manifold. We will also show that the dynamical systems on the slow manifolds of the

☆ Research supported in part by a Grant from the University of Missouri Research Board.
*Fax: +1-573-882-1869.
E-mail address: carmen@chicone.math.missouri.edu.

0022-0396/03/$ - see front matter
doi:10.1016/S0022-0396(02)00148-1

singular perturbation problems obtained by expansion of the RFDE to some finite order in powers of the retardation agree with the dynamical system on the inertial manifold. Thus, the long-term dynamics of the original RFDE (the dynamics on its inertial manifold) can be obtained by reducing a "post-Newtonian" expansion to an appropriate slow manifold.

Let $C([a,b],\mathbb{R}^n)$ denote the Banach space of continuous functions from the interval $[a,b]$ to $\mathbb{R}^n$ with the supremum norm; and, in the special case where $a = -\tau < 0$ and $b = 0$, let $\mathscr{C} := C([-\tau,0],\mathbb{R}^n)$. Also, for each function g defined on the interval $[t-\tau,t]$, let $g_t \in \mathscr{C}$ denote the function given by $g_t(\theta) = g(t+\theta)$. A continuous function $F : \mathbb{R}\times\mathscr{C}\to\mathbb{R}^n$ determines the nonautonomous RFDE

$$\dot{x}(t) = F(t, x_t). \tag{1}$$

Similarly, a continuous function $F : \mathscr{C}\to\mathbb{R}^n$ determines the autonomous RFDE

$$\dot{x}(t) = F(x_t). \tag{2}$$

A solution of Eq. (1) is a continuous function $y : [\sigma-\tau, \sigma+c)\to\mathbb{R}^n$ defined for some $\sigma\in\mathbb{R}$ and $c>0$ such that $\dot{y}(t) = F(t, y_t)$ for $t\in[\sigma,\sigma+c)$, where $\dot{y}(t)$ denotes the right-hand derivative of y at t. In this case, for $\phi\in\mathscr{C}$ we will sometimes write $y(\phi)(t)$ (or $y(\sigma,\phi)(t)$) and say that $y(\phi)$ (or $y(\sigma,\phi)$) is the solution with initial condition ϕ at σ if $y(\phi)_\sigma = \phi$. The basic theorem of the subject is the following result on existence, uniqueness, and continuous dependence (see, for example, [6,10,15]).

Theorem 1.1. *If $F : \mathscr{C}\to\mathbb{R}^n$ is locally Lipschitz, $\phi\in\mathscr{C}$, and $\sigma\in\mathbb{R}$, then there is a unique continuous solution y of RFDE* (1) *such that $y_\sigma = \phi$. Moreover, $y(\phi)$ depends continuously on ϕ. If, in addition, F is C^1, then the solution $t\mapsto y(\phi)(t)$ is C^1 with respect to ϕ on every compact set in its domain of definition.*

Definition 1.2 (cf. Driver [7,8]). A solution y of RFDE (1) is called a *special solution* if y is defined on $\mathbb{R}$ and

$$\sup_{t\in\mathbb{R}} e^{-|t|/\tau}|y(t)| < \infty .$$

Suppose that $\eta : \mathbb{R}\times\mathbb{R}^n\to\mathbb{R}^n$ is a continuous function such that for each $\xi\in\mathbb{R}^n$ the function $t\mapsto\eta(t,\xi)$ is a special solution of RFDE (1). The function η is called a *special flow* for RFDE (1) if

$$\eta(t,\eta(s,\xi)) = \eta(t+s,\xi), \quad \eta(0,\xi) = \xi,$$

whenever $t,s\in\mathbb{R}$ and $\xi\in\mathbb{R}^n$.

We will often write $\eta(\xi)_t$ for the function in $\mathscr{C}$ given by

$$\eta(\xi)_t(\theta) = \eta(t+\theta,\xi).$$

In case $\tau > 0$, the initial conditions for an RFDE range over the infinite-dimensional space $\mathscr{C}$. Thus, $\mathscr{C}$ is the natural state-space for the associated dynamical system given by $T^t(\phi) = y(\phi)_t$ for $t \geqslant 0$ and $\phi \in \mathscr{C}$. The initial conditions for a special flow range over the finite-dimensional space $\mathbb{R}^n$. Thus, a special flow would seem to determine the dynamics of the RFDE only on a negligible subset of its phase space. On the other hand, a special flow would capture the long-term dynamics—the important dynamics in many physical models—if all solutions of the RFDE were attracted (in forward time) to the manifold $\mathscr{M} \subset \mathscr{C}$ given by the image of the map $\xi \to \eta(\xi)_0$ from $\mathbb{R}^n$ to $\mathscr{C}$. Even better, if $\mathscr{M}$ were smooth, then the special flow—the dynamical system that agrees with T^t on $\mathscr{M}$—would be generated by a smooth vector field, the infinitesimal generator of the special flow. In other words, the evolution of the RFDE on $\mathscr{M}$ would be uniquely determined by specifying an initial value in $\mathbb{R}^n$ for an ODE. In fact, under the assumption that there is a special flow, the vector field X with flow η is given (in local coordinates on $\mathscr{M}$) by

$$X(\xi) := \frac{d}{dt}\eta(t,\xi)\bigg|_{t=0} = F(\eta(\xi)_0). \tag{3}$$

Note that X is smooth whenever the function from $\mathbb{R}^n$ to $\mathscr{C}$ given by $\xi \mapsto \eta(\xi)_0$ is smooth. In particular, the smoothness of X depends only on η restricted to $[-\tau, 0] \times \mathbb{R}^n$.

Our investigation of special flows is motivated by the observation that the fundamental forces of nature—in either Maxwell's or Einstein's field theories—propagate at the speed of light, not at infinite speed as in Newtonian physics. In particular, there is no "action at a distance." As a result, the fundamental equations of motion are functional differential equations with space-dependent delays—a more general type of functional differential equation than the RFDEs considered here. Because these equations of motion are generally very complicated, it is common practice in physics to approximate them with ordinary (or partial) differential equations. We will consider one important approximation procedure, called post-Newtonian expansion, where differential equations are obtained by truncating at some finite order the Taylor expansion of the functional equations of motion in powers of some characteristic velocity divided by the speed of light. How can such an approximation be justified? In case the true functional differential equations of motion have an inertial manifold, the only viable finite-dimensional approximation is the dynamical system on the inertial manifold; it captures the long-term dynamics of the infinite-dimensional dynamical system. Therefore, the post-Newtonian approximation (or any other finite-dimensional approximation) would be justified if it agrees with the dynamical system on the inertial manifold.

As a first step in the direction of such a justification for post-Newtonian approximations of real physical systems, we will consider here exactly the same procedure for a special type of RFDE. To mimic the equations of motion, we will consider families of delay equations, where the delay is viewed as a parameter. More precisely, for a C^1-function $f : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$, the RFDE corresponding to the family of functionals $F_\tau : \mathscr{C} \to \mathbb{R}^n$ given by $F_\tau(\phi) = f(\phi(0), \phi(-\tau))$, where $\tau \in [0, b]$ for some

$b \geqslant 0$, is the delay equation

$$\dot{x}(t) = F_\tau(x_t) = f(x(t), x(t-\tau)). \tag{4}$$

Under the assumption that there is a special flow $\eta(t, \xi, \tau)$ for this delay equation which also depends smoothly on ξ and τ, the corresponding smooth vector field X with flow η is given by

$$X(\xi, \tau) = F_\tau(\eta(\xi, \tau)_0) = f(\xi, \eta(-\tau, \xi, \tau)) \tag{5}$$

in local coordinates on the manifold of special solutions $\mathscr{M}_\tau$. Note that the smoothness of the vector field X depends only on the smoothness of η restricted to $[-b, 0] \times \mathbb{R}^n \times [0, b]$.

As suggested previously, the vector field X (equivalently the ODE that it defines) on $\mathscr{M}_\tau$ will determine the long-term behavior of the original delay equation as long as $\mathscr{M}_\tau$ is an *inertial manifold*; that is, $\mathscr{M}_\tau$ is an invariant, finite dimensional, and smooth manifold that attracts all other solutions exponentially fast.

In this paper, "post-Newtonian expansion" means expansion of the vector field (5) to some *finite* order in powers of τ. This procedure produces a high-order ODE in the time-derivatives of the state x, where the highest-order time-derivative of x is multiplied by a power of τ. In other words, the resulting high-order ODE can be viewed as a singular perturbation problem with small parameter τ.

In the physics literature (see, for example, [18]), a famous example related to post-Newtonian expansion is the Lorentz–Dirac equation. For an electron confined to move on a line—the simplest example—and with radiation reaction taken into account, this model yields the third-order ODE

$$\ddot{x} = \tau \dddot{x} + \frac{q}{m} F(x)$$

for the position of the electron, where q is the charge, m is the mass, $\tau := 2q^2/(3mc^3)$, c is the speed of light, and F is the external force. As is well known, this type of equation does *not* give a satisfactory physical model. The fundamental difficulty is apparent, for example, with F given by Hooke's law (say $F(x) = -kx$ for $k > 0$). In this case, the resulting (linear) ODE has solutions—called runaway solutions in the physics literature—that are unbounded in forward time. Since the system is supposed to model the motion of an electron, which is supposed to radiate energy as its acceleration changes, the radiation reaction should cause damping, an effect that is incompatible with runaway solutions. Hence, this post-Newtonian system cannot be the correct (approximate) dynamical model.

A general resolution of the problem of runaway solutions and a rigorous foundation for post-Newtonian mechanics is proposed in [4,5]. There, the fundamental delay-type equations of motion in classical gravitation and electrodynamics are conjectured to have inertial manifolds in the low-velocity regime. As mentioned previously, in this case the long-term dynamics of the true equations of motion is given by the corresponding finite-dimensional differential equation on the inertial manifold. To be viable, a post-Newtonian model must produce a dynamical system that agrees with the dynamical system on the inertial manifold. By a second conjecture, an approximation to the dynamical system on the inertial manifold (up to the order of the truncation of the post-Newtonian expansion) can always be obtained by viewing the (truncation of a) post-Newtonian expansion as a singular perturbation problem and reducing the corresponding dynamical system to an appropriately chosen slow manifold. In this scenario, it is easy to see that the runaway solutions for the post-Newtonian system have no physical significance; they are merely artifacts of the expansion procedure that correspond to motions in the unstable manifold of the slow manifold. The correct post-Newtonian approximation—the ODE on the inertial manifold (a Newtonian equation with post-Newtonian corrections)—is thus approximated by reduction of the post-Newtonian high-order ODE to a slow manifold (cf. [2]).

We will justify the "post-Newtonian" approximation procedure for the delay equation (4). Three main results will be presented. Under the assumption that F is Lipschitz and τ is sufficiently small (an explicit bound will be given), we will show that RFDE (2) has a *smooth* inertial manifold. For the delay equation (4), we will also show that the inertial manifold depends smoothly on the parameter τ. Finally, for the delay equation, we will show that the singularly perturbed high-order ODE obtained by "post-Newtonian" expansion in the small parameter τ results in a vector field (the slow vector field) on an appropriate slow manifold that agrees with the vector field (the inertial vector field) given by the restriction of the infinite-dimensional dynamical system to its inertial manifold. More precisely, we will show that the slow manifold has the same dimension as the inertial manifold and the two vector fields agree to second order in τ—a result that is sufficient for most applications. We will also show these vector fields agree to all orders for the special case of the linear delay equation $\dot{x}(t) = Ax(t-\tau)$, where A is an invertible $n \times n$-matrix.

Note added in proof

In fact, these vector fields agree to all orders for the delay equation $\dot{x}(t) = f(x(t), x(t-\tau))$.

2. The existence and smoothness of special flows

The following theorem states some of the fundamental results of Driver and Ryabov (see [7,8]).

Theorem 2.1. (1) *Suppose that* $F : \mathbb{R} \times \mathscr{C} \to \mathbb{R}^n$ *is continuous and Lipschitz in its second argument* (*that is, there is some* $K > 0$ *such that*

$$|F(t,\phi) - F(t,\psi)| \leqslant K||\phi - \psi||,$$

whenever $t \in \mathbb{R}$ *and* $\phi, \psi \in \mathscr{C}$). *If*

$$\sup_{t \in \mathbb{R}} e^{-|t|/\tau} |F(t,0)| < \infty,$$

and $K\tau e < 1$, *then, for each* $s \in \mathbb{R}$ *and* $\xi \in \mathbb{R}^n$, *the nonautonomous RFDE* $\dot{x}(t) = F(t, x_t)$ *has a unique special solution* y *such that* $y(s) = \xi$. *Moreover, if the RFDE is autonomous, then it has a special flow*.

(2) *Suppose, in addition, that* p *is the least positive root of the equation* $p = Ke^{p\tau}$, $K\tau e^{1+K\tau+p\tau} < 1$, *and* $\bar{y}(t, t_0, \xi_0)$, *for* $t, t_0 \in \mathbb{R}$ *and* $\xi_0 \in \mathbb{R}^n$, *denotes the special solution such that* $\bar{y}(t_0, t_0, \xi_0) = \xi_0$. *Then, there is a number* $r > 0$ *with the following property*: *For each solution* $y(\sigma, \phi)(t)$ *of the RFDE such that* $y(\sigma, \phi)_\sigma = \phi$, *there is some* $\xi \in \mathbb{R}^n$ *such that* $\lim_{s \to \infty} \bar{y}(\sigma, s, y(\sigma, \phi)(s)) = \xi$ *and*

$$\sup_{t \geqslant \sigma + \tau} e^{rt} |y(\sigma, \phi)(t) - \bar{y}(t, \sigma, \xi)| < \infty.$$

The second part of Theorem 2.1 states that every solution is attracted exponentially fast to the manifold of special solutions (cf. [19]).

In this section we will prove two results about the smoothness of the special flow mentioned in Theorem 2.1. For simplicity, we will consider autonomous RFDEs.

Theorem 2.2. *If* $F : \mathscr{C} \to \mathbb{R}^n$ *is a continuously differentiable Lipschitz function with Lipschitz constant* $K > 0$, *and* $0 \leqslant 2K\tau\sqrt{e} < 1$, *then RFDE* (2) *has a continuously differentiable special flow* $\eta : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$.

Theorem 2.3. *If* $f : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ *is a continuously differentiable Lipschitz function with Lipschitz constant* $K > 0$, *and* $0 \leqslant 8Kb < 1$, *then the delay equation* (4) *has a continuously differentiable family of special flows* $\eta : \mathbb{R} \times \mathbb{R}^n \times [0, b] \to \mathbb{R}^n$.

The proofs of these theorems are similar. We will obtain the special solutions for RFDE (2) as fixed points of a contraction. We will then use the fiber contraction principle (see [3,17]) to prove the smoothness of the family of special solutions with respect to their values in $\mathbb{R}^n$ at $t = 0$; and, for the case of delay equations, we will also prove their smoothness with respect to the delay τ.

Note that the smoothness of the infinitesimal generator of the special flow η, namely the vector field X given by $X(\xi, \tau) = f(\xi, \eta(-\tau, \xi, \tau))$, is determined by the smoothness of η restricted to $[-b, 0] \times \mathbb{R}^n \times [0, b]$, and therefore the smoothness of the forward extension follows from the usual results for ODEs.

The proofs given here are similar, but not the same as, the results in [14,16] where smoothness with respect to initial functions and delays is proved for the delay equation (4). Here, we will consider the smoothness of the special flows with respect to the finite-dimensional space of initial conditions $\mathbb{R}^n$, where we must consider backward-time solutions. Also, it turns out that, in our special situation, there is no loss of smoothness with respect to the delay, whereas in [14] the function f in the delay equation (4) is required to be C^2 in order for the solution to be C^1 with respect to the delay.

For another approach to the results presented here, note that the family of delay equations (4) is transformed, by the change of variables $t = s\tau$, to the family of delay equations

$$y'(t) = \tau f(y(s), y(s-1)), \tag{6}$$

where $y(s) := x(\tau s)$. This new family is equivalent to the original family (4) if $\tau \neq 0$. It is easy to see that the unperturbed delay equation, at $\tau = 0$, for family (6), namely $y'(s) = 0$, has a normally hyperbolic, finite dimensional, invariant manifold in $C([-1,0], \mathbb{R}^n)$. We could hope to obtain a smooth inertial manifold for family (6) (and therefore for the original family (4)) for $\tau \neq 0$ and sufficiently small, if we could apply an infinite-dimensional version of the usual finite-dimensional normal hyperbolicity theory (see, for example, the recent results of [1]). While this general approach might produce a useful result, the direct method seems to be a better choice in our special case: at least it produces an explicit estimate for the parameter interval containing $\tau = 0$ corresponding to the existence of inertial manifolds.

2.1. Proof of Theorem 2.2

We will split the proof of Theorem 2.2 into several propositions. The first part of the proof sets up an appropriate fiber contraction uses the fiber contraction principle is used in the second part of the proof to establish the smoothness of the special flow.

By the hypothesis, $2K\tau\sqrt{e} < 1$ where K is the Lipschitz constant for the functional F. Let $\lambda := (2\tau)^{-1}$ and note that

$$\frac{K}{\lambda} e^{\lambda\tau} < 1, \quad \tau K e < 1, \quad 2\lambda\tau = 1. \tag{7}$$

Let V be a compact subset of $\mathbb{R}^n$ and $\mathscr{B}$ the Banach space of continuous functions $\eta : \mathbb{R} \times V \to \mathbb{R}^n$ with

$$\|\eta\|_{\mathscr{B}} := \sup_{(t,\xi) \in \mathbb{R} \times V} |\eta(t,\xi)| e^{-\lambda |t|} < \infty.$$

Also, for an arbitrary Banach space E, a function $\alpha : \mathbb{R} \times V \to E$, and $(t,\xi) \in \mathbb{R} \times V$, let $\alpha(\xi)_t$ denote the function (defined on the interval $[-\tau, 0]$ with values in E) given by $\alpha(\xi)_t(\theta) = \alpha(t+\theta, \xi)$.

Proposition 2.4. *If $\alpha : \mathbb{R} \times V \to E$ is continuous, then $\alpha(\xi)_t \in C([-\tau,0], E)$ whenever $(t,\xi) \in \mathbb{R} \times V$. Moreover, the function $\tilde{\alpha} : \mathbb{R} \times V \to C([-\tau,0], E)$ given by $\tilde{\alpha}(t,\xi) = \alpha(\xi)_t$ is continuous.*

Proof. The first statement is obvious. The second statement follows from the local uniform continuity of α. $\square$

Using this proposition, note that if $\eta : \mathbb{R} \times V \to \mathbb{R}^n$ is continuous, then $\int_0^t F(\eta(\xi)_s)\, ds$ exists whenever $(t, \xi) \in \mathbb{R} \times V$. Hence, for every such η, there is a function $\Lambda(\eta) : \mathbb{R} \times V \to \mathbb{R}^n$ given by

$$\Lambda(\eta)(t, \xi) = \xi + \int_0^t F(\eta(\xi)_s)\, ds. \tag{8}$$

A fixed point of this operator in $\mathscr{B}$ is a special solution of RFDE (2).

Proposition 2.5. *If $\eta \in \mathscr{B}$, then $\Lambda(\eta) \in \mathscr{B}$. Moreover, $\Lambda : \mathscr{B} \to \mathscr{B}$ is a contraction.*

Proof. Using Proposition 2.4, it is easy to see that $\Lambda(\eta) : \mathbb{R} \times V \to \mathbb{R}^n$ is continuous; in fact, it is the composition of continuous functions. We will show first that $||\Lambda(\eta)||_{\mathscr{B}} < \infty$.

For $(s, \xi) \in \mathbb{R} \times V$, we have the key estimate

$$\begin{aligned} ||\eta(\xi)_s|| &= \sup_{\theta \in (-\tau, 0]} |\eta(s + \theta, \xi)| \\ &= \sup_{\theta \in (-\tau, 0]} e^{\lambda |s+\theta|} e^{-\lambda |s+\theta|} |\eta(s + \theta, \xi)| \\ &\leqslant e^{\lambda \tau} e^{\lambda |s|} ||\eta||_{\mathscr{B}}. \end{aligned} \tag{9}$$

Using this estimate and the Lipschitz constant K for F, we have that

$$\begin{aligned} |\Lambda(\eta)(t, \xi)| &\leqslant |\xi| + \left| \int_0^t |F(\eta(\xi)_s)|\, ds \right| \\ &\leqslant |\xi| + \left| \int_0^t |F(\eta(\xi)_s) - F(0)|\, ds + \int_0^t |F(0)|\, ds \right| \\ &\leqslant |\xi| + |t||F(0)| + K \left| \int_0^t ||\eta(\xi)_s||\, ds \right| \\ &\leqslant |\xi| + |t||F(0)| + \frac{K}{\lambda} e^{\lambda \tau} ||\eta||_{\mathscr{B}} (e^{\lambda |t|} - 1), \end{aligned}$$

and therefore

$$\sup_{(t, \xi) \in \mathbb{R} \times V} e^{-\lambda |t|} |\Lambda(\eta)(t, \xi)| \leqslant \sup_{\xi \in V} |\xi| + \frac{|F(0)|}{\lambda e} + \frac{K}{\lambda} e^{\lambda \tau} ||\eta||_{\mathscr{B}} < \infty,$$

as required.

To show that Λ is a contraction, suppose that $\eta, \gamma \in \mathscr{B}$, and use estimate (9) to obtain the inequalities

$$\begin{aligned} |\Lambda(\eta)(t,\xi) - \Lambda(\gamma)(t,\xi)| &\leqslant \left| \int_0^t |F(\eta(\xi)_s) - F(\gamma(\xi)_s)|\, ds \right| \\ &\leqslant K \left| \int_0^t ||\eta(\xi)_s - \gamma(\xi)_s||\, ds \right| \\ &\leqslant \frac{K}{\lambda} e^{\lambda\tau}(e^{\lambda|t|} - 1)||\eta - \gamma||_{\mathscr{B}}. \end{aligned}$$

Thus, we have the norm estimate

$$||\Lambda(\eta) - \Lambda(\gamma)||_{\mathscr{B}} \leqslant \frac{K}{\lambda} e^{\lambda\tau} ||\eta - \gamma||_{\mathscr{B}};$$

and, by the first inequality in display (7), Λ is a contraction. $\square$

For a Banach space E, let $L(\mathbb{R}^n, E)$ denote the linear transformations from $\mathbb{R}^n$ to E with the usual operator norm $|\ |$, and let $\mathscr{L}$ denote the Banach space of continuous functions $\Phi : \mathbb{R} \times V \to L(\mathbb{R}^n, \mathbb{R}^n)$ such that

$$||\Phi||_{\mathscr{L}} := \sup_{(t,\xi) \in \mathbb{R} \times V} |\Phi(t,\xi)| e^{-|t|/\tau} < \infty.$$

Also, let $\mathscr{F}$ denote the set of all continuous functions $\Phi : \mathbb{R} \times V \to L(\mathbb{R}^n, \mathbb{R}^n)$ such that

$$\sup_{(t,\xi) \in \mathbb{R} \times V} |\Phi(t,\xi)| e^{-\lambda|t|} \leqslant 1.$$

The Banach space $\mathscr{L}$ consists of the candidates for the derivatives with respect to ξ of the elements of $\mathscr{B}$.

Proposition 2.6. *The set $\mathscr{F}$ is a complete metric space relative to the metric d given by $d(\Phi, \Upsilon) = ||\Phi - \Upsilon||_{\mathscr{L}}$ for $\Phi, \Upsilon \in \mathscr{F}$.*

Proof. We will show that $\mathscr{F}$ is a closed subset of $\mathscr{L}$. If $\Phi \in \mathscr{F}$, then

$$|\Phi(t,\xi)| e^{-|t|/\tau} \leqslant e^{\lambda|t|} e^{-|t|/\tau}.$$

Hence, in view of the equality $2\lambda\tau = 1$ in display (7), $||\Phi||_{\mathscr{L}} < \infty$. Suppose that $\{\Phi_k\}_{k=1}^{\infty}$ is a sequence in $\mathscr{F}$ that converges to Φ in $\mathscr{L}$. Using estimate (9) and the definition of $\mathscr{F}$, we have that

$$\begin{aligned} |\Phi(t,\xi)| &\leqslant |\Phi(t,\xi) - \Phi_k(t,\xi)| + |\Phi_k(t,\xi)| \\ &\leqslant e^{|t|/\tau} ||\Phi - \Phi_k||_{\mathscr{L}} + e^{\lambda|t|}. \end{aligned}$$

By passing to the limit as $k \to \infty$, we obtain the desired estimate

$$|\Phi(t,\xi)| \leqslant e^{\lambda|t|}. \qquad \square$$

Using Proposition 2.4, if $(\eta, \Phi) \in \mathscr{B} \times \mathscr{L}$, then the function $\tilde{\Phi} : \mathbb{R} \times V \to L(\mathbb{R}^n, \mathscr{C})$ given by $\tilde{\Phi}(t,\xi) = \Phi(\xi)_t$ is continuous and

$$\Psi(\eta,\Phi)(t,\xi) := I + \int_0^t DF(\eta(\xi)_s)\Phi(\xi)_s \, ds \tag{10}$$

is an element of $L(\mathbb{R}^n, \mathbb{R}^n)$.

Proposition 2.7. *If $(\eta, \Phi) \in \mathscr{B} \times \mathscr{F}$, then $\Psi(\eta, \Phi) \in \mathscr{F}$. Moreover, the function $\Gamma : \mathscr{B} \times \mathscr{F} \to \mathscr{B} \times \mathscr{F}$ given by $(\eta, \Phi) \mapsto (\Lambda(\eta), \Psi(\eta, \Phi))$ is a continuous fiber contraction (that is, Γ is continuous, there is a number μ with $0 \leqslant \mu < 1$, and, for each fixed $\eta \in \mathscr{B}$, the function $\Phi \mapsto \Psi(\eta, \Phi)$ is a contraction with contraction constant μ).*

Proof. In view of Proposition 2.4, if $(\eta, \Phi) \in \mathscr{B} \times \mathscr{F}$, then function $(t, \xi) \mapsto \Psi(\eta, \Phi)(t, \xi)$ is continuous. We will show that $\Psi(\eta, \Phi) \in \mathscr{F}$.

Note that $||DF|| \leqslant K$. For v in the unit sphere of $\mathbb{R}^n$ and $\Phi \in F$, we have the following estimates:

$$\begin{aligned}
|\Psi(\eta,\Phi)(t,\xi)v| &\leqslant |v| + \left| \int_0^t |DF(\eta(\xi)_s)\Phi(\xi)_s v| \, ds \right| \\
&\leqslant 1 + K \left| \int_0^t ||\Phi(\xi)_s v|| \, ds \right| \\
&\leqslant 1 + K \left| \int_0^t \sup_{\theta \in (-\tau, 0]} e^{\lambda|s+\theta|} \, ds \right| \\
&\leqslant 1 + \frac{K}{\lambda} e^{\lambda\tau}(e^{\lambda|t|} - 1).
\end{aligned}$$

By using the first inequality in display (7), we have that

$$\begin{aligned}
|\Psi(\eta,\Phi)(t,\xi)v| &\leqslant \left(1 - \frac{K}{\lambda} e^{\lambda\tau}\right) + \frac{K}{\lambda} e^{\lambda\tau} e^{\lambda|t|} \\
&\leqslant 1 + \frac{K}{\lambda} e^{\lambda\tau}(e^{\lambda|t|} - 1) \\
&\leqslant e^{\lambda|t|},
\end{aligned}$$

and therefore $\Psi(\eta, \Phi) \in \mathscr{F}$.

We will show that Ψ is a uniform contraction. For $\eta \in \mathscr{B}$ and $\Phi, \Upsilon \in \mathscr{F}$, the analogue of the key inequality (9) for $\mathscr{L}$ is $||\Phi(\xi)_s|| \leqslant e^{1+|s|/\tau} ||\Phi||_{\mathscr{L}}$. It is used to obtain the estimate

$$\begin{aligned}|\Psi(\eta,\Phi)(t,\xi)v-\Psi(\eta,\Upsilon)(t,\xi)v| &\leqslant K\left|\int_0^t ||\Phi(\xi)_s v-\Upsilon(\xi)_s v||\,ds\right|\\ &\leqslant \tau Ke||\Phi-\Upsilon||_{\mathscr{L}}(e^{|t|/\tau}-1).\end{aligned}$$

Hence,

$$||\Psi(\eta,\Phi)-\Psi(\eta,\Upsilon)||_{\mathscr{L}}\leqslant K\tau e||\Phi-\Upsilon||_{\mathscr{L}},$$

and, by the second inequality in display (7), Ψ is a uniform contraction.

To complete the proof, we will show that Ψ is continuous. Because Ψ is a *uniform* contraction, it suffices to prove that the function $\eta\mapsto\Psi(\eta,\Phi)$ is a continuous map from $\mathscr{B}$ to $\mathscr{F}$ for each $\Phi\in\mathscr{F}$.

Remark. Although continuity with respect to the base point is an essential ingredient of the fiber contraction method, this nontrivial requirement is often ignored. There does not seem to be a general result that can be used to establish the required continuity; instead, the continuity must be checked in each case (see [9] for an approach to this difficulty in the setting of local contractions of Banach spaces).

For $\eta,\gamma\in\mathscr{B}$ and $\Phi\in\mathscr{F}$, we have the estimates

$$\begin{aligned}|\Psi(\eta,\Phi)(t,\xi)v-\Psi(\gamma,\Phi)(t,\xi)v| &\leqslant \left|\int_0^t |DF(\eta(\xi)_s)-DF(\gamma(\xi)_s)|||\Phi(\xi)_s v||\,ds\right|\\ &\leqslant e^{\lambda\tau}\left|\int_0^t |DF(\eta(\xi)_s)-DF(\gamma(\xi)_s)|e^{\lambda|s|}\,ds\right|. \end{aligned}\tag{11}$$

Claim 2.8. *Fix $\gamma\in\mathscr{B}$. For each $\varepsilon>0$, there is a $\delta>0$ such that*

$$|DF(\eta(\xi)_s)-DF(\gamma(\xi)_s)|e^{-\lambda|s|}<\varepsilon,$$

whenever $(s,\xi)\in\mathbb{R}\times V$ and $||\eta-\gamma||_{\mathscr{B}}<\delta$.

Remark. Define the space $\mathscr{N}$ of continuous functions $G:\mathbb{R}\times V\to L(\mathscr{C},\mathbb{R}^n)$ such that $\sup_{(t,\xi)\in\mathbb{R}\times V}|G(t,\xi)|e^{-\lambda|t|}<\infty$. Claim 2.8 states that the map $P:\mathscr{B}\to\mathscr{N}$ given by $P(\eta)(t,\xi)=DF(\eta(\xi)_t)$ is continuous.

To begin the proof of the claim, recall that $||DF||\leqslant K$ and choose a number $\sigma>0$ such that $2K\exp(-\lambda\sigma)<\varepsilon$. For all $\eta,\gamma\in\mathscr{B}$, $\xi\in V$, and $|s|>\sigma$,

$$|DF(\eta(\xi)_s)-DF(\gamma(\xi)_s)|e^{-\lambda|s|}\leqslant 2Ke^{-\lambda|\sigma|}<\varepsilon. \tag{12}$$

On the other hand, using the uniform continuity of the function DF on the compact subset $S \coloneqq \{\gamma(\xi)_s \in \mathscr{C} : s \in [-\sigma, \sigma]\}$ of $\mathscr{C}$ (and a compactness argument), there is some $\delta_1 > 0$ such that $||DF(\phi) - DF(\psi)|| < \varepsilon$ whenever $\phi \in \mathscr{C}$, $\psi \in S$ and $||\phi - \psi|| < \delta_1$. Also, using the definition of the $\mathscr{B}$-norm, there is a $\delta > 0$ such that $||\eta(\xi)_s - \gamma(\xi)_s|| < \delta_1$ for all $(s, \xi) \in [-\sigma, \sigma] \times V$ whenever $||\eta - \gamma||_{\mathscr{B}} < \delta$. Hence, for such an η and for all $(s, \xi) \in [-\sigma, \sigma] \times V$, we have that

$$|DF(\eta(\xi)_s) - DF(\gamma(\xi)_s)|e^{-\lambda|s|} \leqslant |DF(\eta(\xi)_s) - DF(\gamma(\xi)_s)| < \varepsilon.$$

This result, combined with inequality (12), proves the claim.

Fix $\varepsilon > 0$. By Claim 2.8, there is a $\delta > 0$ such that

$$|DF(\eta(\xi)_s) - DF(\gamma(\xi)_s)|e^{-\lambda|s|} < 2\lambda\varepsilon e^{-\lambda\tau},$$

whenever $||\eta - \gamma||_{\mathscr{B}} < \delta$. Using this result and estimate (11), we have that

$$\begin{aligned} |\Psi(\eta, \Phi)(t, \xi)v - \Psi(\gamma, \Phi)(t, \xi)v| &< 2\lambda\varepsilon \left| \int_0^t e^{2\lambda|s|}\, ds \right| \\ &\leqslant \varepsilon(e^{2\lambda|t|} - 1), \end{aligned}$$

whenever $||\eta - \gamma||_{\mathscr{B}} < \delta$. Finally, by using the equation $1/\tau - 2\lambda = 0$ and the last equality in display (7), it follows that

$$|\Psi(\eta, \Phi)(t, \xi)v - \Psi(\gamma, \Phi)(t, \xi)v|e^{-|t|/\tau} < \varepsilon,$$

whenever $||\eta - \lambda|| < \delta$; that is, the function $\eta \mapsto \Psi(\eta, \Phi)$ is continuous.

Proof of Theorem 2.2. Choose a point $(\sigma, \zeta) \in \mathbb{R} \times \mathbb{R}^n$, an open subset U of $\mathbb{R}^n$ with compact closure V such that $\zeta \in U$, and let the Banach spaces $\mathscr{B}$ and $\mathscr{L}$ be defined relative to the compact set V.

Because Λ is a contraction, it has a unique fixed point $\eta \in \mathscr{B}$. Because F is a complete metric space and (by Proposition 2.7) the map $\Phi \mapsto \Psi(\eta, \Phi)$ is a contraction, it has a unique fixed point $\Phi \in \mathscr{F}$. By the fiber contraction theorem [17], the point $(\eta, \Phi) \in \mathscr{B} \times \mathscr{F}$ is a globally attracting fixed point for the fiber contraction Γ. Let $\eta_1(t, \xi) \equiv 0$ and $\Phi_1(t, \xi) \equiv 0$, and note that $(\eta_1, \Phi_1) \in \mathscr{B} \times \mathscr{F}$. Also, for each integer $k > 1$, let $(\eta_k, \Phi_k) = \Gamma(\eta_{k-1}, \Phi_{k-1})$. Proceeding by induction, we have that $D_\xi \eta_1(t,\xi) = \Phi_1(t,\xi)$, and if $D_\xi \eta_k(t,\xi) = \Phi_k(t,\xi)$, then

$$\begin{aligned} D_\xi \eta_{k+1}(t,\xi) &= D_\xi\left(\xi + \int_0^t F(\eta_k(\xi)_s)\,ds\right) \\ &= I + \int_0^t DF(\eta_k(\xi)_s) D_\xi \eta_k(\xi)_s\,ds \\ &= I + \int_0^t DF(\eta_k(\xi)_s)\Phi_k(\xi)_s\,ds \\ &= \Phi_{k+1}(t,\xi), \end{aligned}$$

where the differentiation under the integral sign is justified because the interval of integration is finite and the integrand is continuously differentiable with respect to ξ. Thus, we have that $\{\eta_k\}_{k=1}^\infty$ converges to η in $\mathscr{B}$ and $\{D_\xi \eta_k\}_{k=1}^\infty$ converges to Φ in $\mathscr{F}$. Finally, because continuity and differentiability are local properties, to prove that $D_\xi \eta$ exists and is continuous at (σ,ζ), it suffices to restrict the functions in these sequences to the domain $[\sigma-1,\sigma+1]\times U$ where

$$\sup_{[\sigma-1,\sigma+1]\times U} |\eta_k(t,\xi) - \eta(t,\xi)| \leqslant e^{\lambda(1+|\sigma|)} ||\eta_k - \eta||_{\mathscr{B}},$$

$$\sup_{[\sigma-1,\sigma+1]\times U} |\Phi_k(t,\xi) - \Phi(t,\xi)| \leqslant e^{\lambda(1+|\sigma|)} ||\Phi_k - \Phi||_{\mathscr{L}}.$$

Hence, on this domain, the sequences $\{\eta_k\}_{k=1}^\infty$ and $\{D_\xi \eta_k\}_{k=1}^\infty$ are uniformly convergent, and therefore $D_\xi \eta = \Phi$ is a continuous function. $\square$

2.2. *Proof of Theorem 2.3*

While the proof of Theorem 2.3 is similar in structure to the proof of Theorem 2.2, it is more difficult. To see why, recall the delay equation (4), and note that its variational equation with respect to τ along the solution $t \mapsto x(t)$ is given by

$$\dot{w}(t) = D_1 f(x(t), x(t-\tau))w(t) + D_2 f(x(t), x(t-\tau))(w(t-\tau) - \dot{x}(t-\tau)),$$

where $D_1 f$ (respectively, $D_2 f$) denotes the partial derivative with respect to the first (respectively, the second) argument of f. A problem arises because the factor $\dot{x}(t-\tau)$ appears in the integrand of the basic integral equation that will be used to define the principal part of a required fiber map. In order for this fiber map to be defined, the function $\dot{x}(t-\tau)$ must, at least, be integrable. This requirement must be taken into account in the definition of the function space where we will seek a special solution of the original delay equation as the fixed point of a contraction. The natural candidate for this space is a certain weighted Sobolev-type space that we will define (cf. [14,16]) after the statement and proof of a technical lemma.

Lemma 2.9. *The function* $p:[0,\infty)\to\mathbb{R}$ *given by*

$$p(x) = (2(2+x))^{-1}\ln(1+x)$$

achieves its maximum value pmax *at a unique point* $\sigma\in[0,\infty)$. *The value of* pmax *gives* $1/\text{pmax}\approx 7.18$. *If* b *is such that* $Kb/\text{pmax}<1$ *and* $\lambda := 2K(2+\sigma)$, *then there is a number* $v>\lambda$ *such that, for* $\mu := \lambda + v$,

$$\frac{2K}{\lambda}(1+e^{b\lambda})<1, \qquad \frac{K}{v}(1+e^{bv})<1, \qquad \frac{K}{\mu}(1+e^{b\mu})<1.$$

Proof. The first statement of the proposition is an exercise in calculus. It is also easy to show that

$$\text{pmax} < (4(2+\sigma))^{-1}\ln(7+4\sigma). \tag{13}$$

By the hypothesis,

$$b\lambda = 2Kb(2+\sigma) < 2(2+\sigma)\,\text{pmax} = \ln(1+\sigma).$$

Hence, $e^{b\lambda}<1+\sigma$, and therefore

$$\frac{2K}{\lambda}(1+e^{b\lambda})<1.$$

By continuity, there is some number $v_0>\lambda$ such that

$$\frac{K}{v}(1+e^{bv})<1,$$

whenever $v_0>v>\lambda$. Let $\mu=\mu(v)=\lambda+v$ and note that

$$\begin{aligned}\frac{K}{\mu}(1+e^{b\mu}) &< \frac{K}{2\lambda}(1+e^{b(\lambda+v)})\\ &= \frac{1}{4(2+\sigma)}(1+e^{b(\lambda+v)}).\end{aligned}$$

Also, note that

$$\lim_{v\to\lambda^+}\frac{1}{4(2+\sigma)}(1+e^{b(\lambda+v)}) = \frac{1}{4(2+\sigma)}(1+e^{2b\lambda}).$$

Using inequality (13), we have that

$$2b\lambda < \ln(7+4\sigma).$$

A rearrangement of this estimate gives

$$1 + e^{2b\lambda} < 4(2 + \sigma),$$

and therefore

$$\frac{1}{4(2+\sigma)}(1 + e^{2b\lambda}) < 1.$$

By continuity, there is some $v > \lambda$ such that

$$\frac{K}{\mu}(1 + e^{b\mu}) < 1,$$

as required. $\square$

By the hypothesis of Theorem 2.3, we have $0 \leqslant 8Kb < 1$. Hence, there are numbers λ, v, and μ satisfying the relations stated in Lemma 2.9. Let us fix these numbers for the remainder of this section. Also, let $V \subset \mathbb{R}^n$ be a compact set.

If $\eta : \mathbb{R} \times V \times [0, b] \to \mathbb{R}^n$ is continuous and $J \subset \mathbb{R}$ is compact, then $\int_J |\eta(t, \xi, \tau)|\, dt < \infty$ whenever $(\xi, \tau) \in V \times [0, b]$. Thus, there is an associated function $\eta_* : V \times [0, b] \to L^1_{\mathrm{loc}}(\mathbb{R}, \mathbb{R}^n)$ given by $\eta_*(\xi, \tau)(t) = \eta(t, \xi, \tau)$. By the usual theory of Schwartz distributions (see, for example, [21]), the L^1_{loc}-function $\eta_*(\xi, \tau)$ has a distributional derivative $D_\tau(\eta_*(\xi, \tau))$ for each $(\xi, \tau) \in V \times [0, b]$. If, in addition, $D_\tau(\eta_*(\xi, \tau)) \in L^1_{\mathrm{loc}}(\mathbb{R}, \mathbb{R}^n)$ for every $(\xi, \tau) \in V \times [0, b]$, then we have a function $\dot\eta : V \times [0, b] \to L^1_{\mathrm{loc}}(\mathbb{R}, \mathbb{R}^n)$ given by $\dot\eta(\xi, \tau) = D_\tau(\eta_*(\xi, \tau))$. By the definition of convergence in L^1_{loc}, the function $\dot\eta$ is continuous, if for each compact set $J \subset \mathbb{R}$, the function $\dot\eta : V \times [0, b] \to L^1(J, \mathbb{R}^n)$ is continuous. Note that if $\dot\eta$ is continuous, then for each such J, the map from $V \times [0, b]$ to $\mathbb{R}^n$ given by $(\xi, \tau) \mapsto \int_J \dot\eta(\xi, \tau)\, ds$ is continuous. In fact, we have that

$$\left| \int_J \dot\eta(\xi, \tau)\, ds - \int_J \dot\eta(\zeta, \sigma)\, ds \right| \leqslant \|\eta(\xi, \tau) - \eta(\zeta, \sigma)\|_1,$$

where the norm is the L^1-norm for $L^1(J, \mathbb{R}^n)$. Let $\mathscr{S}$ denote the set of all continuous functions $\eta : \mathbb{R} \times V \times [0, b] \to \mathbb{R}^n$ such that, for each $(\xi, \tau) \in V \times [0, b]$, the function $\eta_*(\xi, \tau)$ has a distributional derivative $\dot\eta(\xi, \tau)$ in $L^1_{\mathrm{loc}}(\mathbb{R}, \mathbb{R}^n)$, the function $\dot\eta : V \times [0, b] \to L^1_{\mathrm{loc}}(\mathbb{R}, \mathbb{R}^n)$ is continuous, and

$$\|\eta\|_{\mathscr{S}} := \sup_{(t,\xi,\tau) \in \mathbb{R} \times V \times [0,b]} e^{-\lambda|t|} \left(|\eta(t, \xi, \tau)| + \left| \int_0^t |\dot\eta(\xi, \tau)|\, ds \right| \right) < \infty.$$

The following proposition is proved in the appendix.

Proposition 2.10. *The set $\mathscr{S}$ endowed with the $\mathscr{S}$-norm is a Banach space.*

For $\eta : \mathbb{R} \times V \times [0, b] \to \mathbb{R}^n$, we have the operator

$$\Lambda(\eta)(t, \xi, \tau) = \xi + \int_0^t f(\eta(s, \xi, \tau), \eta(s - \tau, \xi, \tau))\, ds. \tag{14}$$

Note that a fixed point of Λ in $\mathscr{S}$ is a solution of the delay equation (4).

Let $\rho > 0$ be a number such that

$$\rho \geqslant \left(\sup_{\xi \in V} |\xi| + \frac{2|f(0,0)|}{\lambda e} \right) \left(1 - \frac{2K}{e} (1 + e^{\lambda b}) \right)^{-1}, \tag{15}$$

and let $\mathscr{B}$ denote the closed ball with radius ρ at the origin in $\mathscr{S}$.

Proposition 2.11. *If $\eta \in \mathscr{B}$, then $\Lambda(\eta) \in \mathscr{B}$. Also, $\Lambda : \mathscr{B} \to \mathscr{B}$ is a contraction.*

Proof. Because η and f are continuous, the function $\Lambda(\eta)$ is continuous with its range in $\mathbb{R}^n$. Similarly, since

$$D_t \Lambda(\eta)(t, \xi, \tau) = f(\eta(t, \xi, \tau), \eta(t - \tau, \xi, \tau)),$$

the function $D_t \Lambda(\eta)$ is continuous. Moreover, the function $D_t \Lambda(\eta)(\xi, \tau) : \mathbb{R} \to \mathbb{R}^n$ is in $L^1_{\text{loc}}(\mathbb{R}, \mathbb{R}^n)$.

For $\eta \in \mathscr{B}$, we have that

$$\begin{aligned} e^{-\lambda|t|} |\Lambda(\eta)(t, \xi, \tau)| &\leqslant e^{-\lambda|t|} |\xi| + e^{-\lambda|t|} \left| \int_0^t |f(\eta(s, \xi, \tau), \eta(s - \tau, \xi, \tau))|\, ds \right| \\ &\leqslant |\xi| + e^{-\lambda|t|} \left| \int_0^t (K(|\eta(s, \xi, \tau)| + |\eta(s - \tau, \xi, \tau))|) \right. \\ &\qquad \left. + \ |f(0,0)|)\, ds \right| \\ &\leqslant |\xi| + e^{-\lambda|t|} \left| \int_0^t (K(e^{\lambda|s|} ||\eta||_{\mathscr{S}} + e^{\lambda\tau} e^{\lambda|s|} ||\eta||_{\mathscr{S}}) \right. \\ &\qquad \left. + |f(0,0)|)\, ds \right| \\ &\leqslant \sup_{\xi \in V} |\xi| + \frac{K}{\lambda} (1 + e^{\lambda b}) ||\eta||_{\mathscr{S}} + \frac{|f(0,0)|}{\lambda e} \end{aligned}$$

and

$$\begin{aligned} e^{-\lambda|t|} \left| \int_0^t |D_t \Lambda(\eta)(t, \xi, \tau)|\, ds \right| &= e^{-\lambda|t|} \left| \int_0^t |f(\eta(s, \xi, \tau), \eta(s - \tau, \xi, \tau))|\, ds \right| \\ &\leqslant \frac{K}{\lambda} (1 + e^{\lambda b}) ||\eta||_{\mathscr{S}} + \frac{|f(0,0)|}{\lambda e}, \end{aligned}$$

that is,

$$||\Lambda(\eta)||_{\mathscr{S}} \leqslant \sup_{\xi \in V} |\xi| + \frac{2|f(0,0)|}{\lambda e} + \frac{2K}{\lambda}(1+e^{\lambda b})\rho.$$

Note that this inequality has the form $||\Lambda(\eta)||_{\mathscr{S}} \leqslant A + B\rho$. By the choice of ρ as in display (15), we have that $\rho \geqslant A/(1-B)$. Hence, $A + B\rho \leqslant \rho$; that is, $\Lambda : \mathscr{B} \to \mathscr{B}$.

By similar estimates, for $\eta, \gamma \in \mathscr{S}$ we have that

$$e^{-\lambda|t|}|\Lambda(\eta)(t,\xi,\tau) - \Lambda(\gamma)(t,\xi,\tau)| \leqslant \frac{K}{\lambda}(1+e^{\lambda b})||\eta - \gamma||_{\mathscr{S}}$$

and

$$e^{-\lambda|t|}\left|\int_0^t |D_t\Lambda(\eta)(t,\xi,\tau) - D_t\Lambda(\gamma)(t,\xi,\tau)|\,ds\right| \leqslant \frac{K}{\lambda}(1+e^{\lambda b})||\eta-\gamma||_{\mathscr{S}}.$$

Hence,

$$||\Lambda(\eta) - \Lambda(\gamma)||_{\mathscr{S}} \leqslant \frac{2K}{\lambda}(1+e^{\lambda b})||\eta-\gamma||_{\mathscr{S}},$$

and, by the choice of λ as in Lemma 2.9, the function Λ is a contraction. $\square$

For λ, μ, v, and b as in Lemma 2.9, let $\mathscr{L}$ denote the Banach space of all continuous functions $\Phi : \mathbb{R} \times V \times [0,b] \to L(\mathbb{R}, \mathbb{R}^n)$ that are bounded with respect the norm given by

$$||\Phi||_{\mathscr{L}} := \sup_{(t,\xi,\tau)\in \mathbb{R}\times V\times[0,b]} e^{-\mu|t|}|\Phi(t,\xi,\tau)|.$$

We will use the natural identification of $L(\mathbb{R}, \mathbb{R}^n)$ with $\mathbb{R}^n$ to identify $\Phi(t,\xi,\tau)$ with an element of $\mathbb{R}^n$. The Banach space $\mathscr{L}$ consists of the candidates for the derivatives with respect to τ of the elements of $\mathscr{S}$. For $(\eta, \Phi) \in \mathscr{S} \times \mathscr{L}$, let Ψ denote the operator given by

$$\begin{aligned}
&\Psi(\eta,\Phi)(t,\xi,\tau)\\
&\quad = \int_0^t (D_1 f(\eta(s,\xi,\tau), \eta(s-\tau,\xi,\tau))\Phi(s,\xi,\tau)\\
&\qquad + D_2 f(\eta(s,\xi,\tau), \eta(s-\tau,\xi,\tau))(\Phi(s,\xi,\tau) - \dot{\eta}(\xi,\tau)(s-\tau))\,ds.
\end{aligned}$$

Choose $r > 0$ such that

$$\frac{r}{2Ke^{\lambda b}}\left(1 - \frac{K}{v}(1+e^{vb})\right) \geqslant \rho, \tag{16}$$

where ρ is as in display (15), and let $\mathscr{F}$ denote the set of all continuous functions $\Phi : \mathbb{R} \times V \times [0, b] \to L(\mathbb{R}, \mathbb{R}^n)$ such that

$$\sup_{(t,\xi,\tau)\in\mathbb{R}\times V\times[0,b]} e^{-\nu|t|}|\Phi(t,\xi,\tau)| \leqslant r.$$

Proposition 2.12. *The set $\mathscr{F}$ is a closed subset of $\mathscr{L}$. If $(\eta, \Phi) \in \mathscr{B} \times \mathscr{F}$, then $\Psi(\eta, \Phi) \in \mathscr{F}$. Moreover, the function $\Gamma : \mathscr{B} \times \mathscr{F} \to \mathscr{B} \times \mathscr{F}$ given by $\Gamma(\eta, \Phi) = (\Lambda(\eta), \Psi(\eta, \Phi))$ is a continuous fiber contraction.*

Proof. To show that $\mathscr{F} \subset \mathscr{L}$, let $\Phi \in \mathscr{F}$ and use the inequality $\mu > \nu$ to obtain the estimate

$$\sup_{(t,\xi,\tau)\in\mathbb{R}\times V\times[0,b]} e^{-\mu|t|}|\Phi(t,\xi,\tau)| \leqslant \sup_{t\in\mathbb{R}} re^{\nu|t|}e^{-\mu|t|} < \infty$$

required for Φ to be in $\mathscr{L}$. The proof that $\mathscr{F}$ is closed in $\mathscr{L}$ is similar to the proof of Proposition 2.6.

To show that $\Psi(\eta, \Phi) \in \mathscr{F}$ whenever $(\eta, \Phi) \in \mathscr{B} \times \mathscr{F}$, let us first recall that if $g \in L^1_{\text{loc}}(\mathbb{R})$, then $t \mapsto \int_0^t g\, ds$ is continuous. In fact, this map is absolutely continuous (see [22, p. 50]). Thus, because the image of $\dot{\eta}$ is in $L^1_{\text{loc}}(\mathbb{R}, \mathbb{R}^n)$ and f is continuously differentiable, it follows that $\Psi(\eta, \Phi)$ is continuous. Also, because f is globally Lipschitz, we have that $||D_1 f|| + ||D_2 f|| \leqslant K$, where the norm on the cross product $\mathbb{R}^n \times \mathbb{R}^n$ is the sum of the $\mathbb{R}^n$-norms.

Using the estimates

$$\begin{aligned}
|\Psi(\eta,\Phi)(t,\xi,\tau)| &\leqslant K\left|\int_0^t (|\Phi(s,\xi,\tau)| + |\Phi(s-\tau,\xi,\tau)| + |\dot{\eta}(\xi,\tau)(s-\tau)|)\, ds\right| \\
&\leqslant \frac{rK}{\nu}(1+e^{\nu\tau})(e^{\nu|t|}-1) + K\left|\int_{-\tau}^{t-\tau} |\dot{\eta}(\xi,\tau)|\, ds\right| \\
&\leqslant \frac{rK}{\nu}(1+e^{\nu b})(e^{\nu|t|}-1) + 2Ke^{\lambda b}\rho e^{\lambda|t|}
\end{aligned}$$

(where, in case $t > \tau$, the last integral is split into integrals over $[-\tau, 0]$ and $[0, t-\tau]$ before the norm estimate is made), we have that

$$\begin{aligned}
|\Psi(\eta,\Phi)(t,\xi,\tau)|e^{-\nu|t|} &\leqslant \frac{rK}{\nu}(1+e^{\nu b}) + 2\rho Ke^{\lambda b} \\
&\leqslant \frac{rK}{\nu}(1+e^{\nu b}) + r\left(1 - \frac{K}{\nu}(1+e^{\nu b})\right) \\
&\leqslant r,
\end{aligned}$$

that is, $\Psi(\eta, \Phi) \in \mathscr{F}$.

We will show that $\Psi : \mathscr{B} \times \mathscr{F} \to \mathscr{F}$ is a uniform contraction. In fact, for $\eta \in \mathscr{S}$ and Φ, $\Upsilon \in \mathscr{L}$, we have the inequalities

$$\begin{aligned}|\Psi(\eta,\Phi)(t,\xi,\tau) - \Psi(\eta,\Upsilon)(t,\xi,\tau)| &\leqslant K\left|\int_0^t (|\Phi(s,\xi,\tau) - \Upsilon(s,\xi,\tau)| \right.\\ &\quad \left. + |\Phi(s-\tau,\xi,\tau) - \Upsilon(s-\tau,\xi,\tau)|)\, ds\right|\\ &\leqslant \frac{K}{\mu}(1+e^{\mu\tau})||\Phi - \Upsilon||_{\mathscr{L}}(e^{\mu|t|}-1),\end{aligned}$$

and therefore

$$||\Psi(\eta,\Phi) - \Phi(\eta,\Upsilon)||_{\mathscr{L}} \leqslant \frac{K}{\mu}(1+e^{\mu\tau})||\Phi - \Upsilon||_{\mathscr{L}},$$

as required.

We will show that Γ is a continuous fiber contraction. As in the proof of Proposition 2.7, it suffices to show that, for each $\Phi \in \mathscr{F}$, the function from $\mathscr{B}$ to $\mathscr{F}$ given by $\eta \mapsto \Psi(\eta,\Phi)$ is continuous.

Claim 2.13. *Fix $\gamma \in \mathscr{S}$ and let $i \in \{1,2\}$. For each $\varepsilon > 0$, there is a $\delta > 0$ such that*

$$e^{-\lambda|s|}|D_i f(\eta(s,\xi,\tau), \eta(s-\tau,\xi,\tau)) - D_i f(\gamma(s,\xi,\tau), \gamma(s-\tau,\xi,\tau))| < \varepsilon,$$

whenever $(s,\xi,\tau) \in \mathbb{R} \times V \times [0,b]$ and $||\eta - \gamma||_{\mathscr{S}} < \delta$.

The proof of Claim 2.13 is similar to the proof of Claim 2.8.

Fix $\gamma \in \mathscr{B}$, $\Phi \in \mathscr{F}$, and $\varepsilon > 0$. Let

$$M := \frac{r}{\lambda + v}(1 + e^{vb}) + 2Ke^{\lambda b} + 2\rho e^{2\lambda b},$$

and apply Claim 2.13 to obtain a $\delta > 0$ such that $\delta < \varepsilon/M$ and, for $i \in \{1,2\}$,

$$e^{-\lambda|s|}|D_i f(\eta(s,\xi,\tau), \eta(s-\tau,\xi,\tau)) - D_i f(\gamma(s,\xi,\tau), \gamma(s-\tau,\xi,\tau))| < \frac{\varepsilon}{M}, \tag{17}$$

whenever $(s,\xi,\tau) \in \mathbb{R} \times V \times [0,b]$ and $||\eta - \gamma||_{\mathscr{S}} < \delta$. Also, note that

$$|\Psi(\eta,\Phi)(t,\xi,\tau) - \Psi(\gamma,\Phi)(t,\xi,\tau)| \leqslant I_1 + I_2 + I_3,$$

where

$$\begin{aligned}I_1 := \left|\int_0^t \right. &|D_1 f(\eta(s,\xi,\tau), \eta(s-\tau,\xi,\tau))\\ &\left. - D_1 f(\gamma(s,\xi,\tau), \gamma(s-\tau,\xi,\tau))||\Phi(s,\xi,\tau)|\, ds\right|,\end{aligned}$$

$$I_2 := \left| \int_0^t |D_2 f(\eta(s,\xi,\tau),\eta(s-\tau,\xi,\tau)) \right.$$
$$\left. - D_2 f(\gamma(s,\xi,\tau),\gamma(s-\tau,\xi,\tau))||\Phi(s-\tau,\xi,\tau)|\, ds \right|,$$

$$I_3 := \left| \int_0^t |D_2 f(\eta(s,\xi,\tau),\eta(s-\tau,\xi,\tau))\dot{\eta}(\xi,\tau)(s-\tau) \right.$$
$$\left. - D_2 f(\gamma(s,\xi,\tau),\gamma(s-\tau,\xi,\tau))\dot{\gamma}(\xi,\tau)(s-\tau)|\, ds \right|.$$

Using inequality (17) and the definition of $\mathscr{F}$, we have that

$$I_1 + I_2 < \varepsilon \frac{r}{M(\lambda+v)}(1+e^{vb})(e^{(\lambda+v)|t|}-1). \tag{18}$$

To bound I_3, add and subtract the quantity

$$D_2 f(\eta(s,\xi,\tau),\eta(s-\tau,\xi,\tau))\dot{\gamma}(\xi,\tau)(s-\tau),$$

apply the triangle inequality, and make a change of variables to obtain the estimates

$$\begin{aligned} I_3 &< K\left|\int_{-\tau}^{t-\tau} |\dot{\eta}(\xi,\tau)-\dot{\gamma}(\xi,\tau)|\, ds\right| + \frac{\varepsilon}{M}\left|\int_{-\tau}^{t-\tau} e^{\lambda|s|}|\dot{\gamma}(\xi,\tau)|\, ds\right| \\ &\leqslant K\left|\int_{-\tau}^{t-\tau} |\dot{\eta}(\xi,\tau)-\dot{\gamma}(\xi,\tau)|\, ds\right| + \frac{\varepsilon}{M}e^{\lambda|t|}e^{\lambda\tau}\left|\int_{-\tau}^{t-\tau} |\dot{\gamma}(\xi,\tau)|\, ds\right|. \end{aligned}$$

The integrals are bounded above (using norm estimates in $\mathscr{S}$ and $\mathscr{B}$) by considering separately the cases $t-\tau\leqslant 0$ and $t-\tau>0$, and in the latter case by splitting the integral into integrals over $[-\tau,0]$ and $[0,t-\tau]$. Using the resulting estimates, the choice of δ, and the definition of $\mathscr{B}$, we find that

$$I_3 < 2Ke^{\lambda b}e^{\lambda|t|}\delta + \frac{2\varepsilon}{M}e^{2\lambda|t|}e^{2\lambda b}\rho.$$

By combining this inequality with inequality (18), using the hypotheses $v>\lambda$ and $\mu=\lambda+v$, the definition of δ (recall that $\delta<\varepsilon/M$), and the definition of M, we have that

$$e^{-\mu|t|}|\Psi(\eta,\Phi)(t,\xi,\tau)-\Psi(\gamma,\Phi)(t,\xi,\tau)|<\varepsilon,$$

whenever $(t,\xi,\tau)\in\mathbb{R}\times V\times[0,b]$ and $||\eta-\gamma||_{\mathscr{S}}<\delta$. Hence, $\eta\mapsto\Psi(\eta,\Phi)$ is continuous. $\square$

Proof Theorem 2.3. The proof is similar to the proof of Theorem 2.2 except for a modification of the induction argument that is used to show the equality $D_\tau\eta_k=\Phi_k$ for the elements of the sequences $\{\eta_k\}_{k=1}^{\infty}$ in $\mathscr{B}$ and $\{\Phi_k\}_{k=1}^{\infty}$ in $\mathscr{F}$ which are defined recursively by the iterates of the point $(0,0)\in\mathscr{B}\times\mathscr{F}$ with respect to the fiber map Γ.

In the present case, it is convenient to use an induction hypothesis with two parts: (1) $D_t\eta_k$ (the classical partial derivative) is continuous. (2) $D_\tau\eta_k = \Phi_k$. Note that because η_k is continuous and

$$D_t\eta_{k+1}(t,\xi,\tau) = f(\eta_k(t,\xi,\tau), \eta_k(t-\tau,\xi,\tau)),$$

the partial derivative $D_t\eta_{k+1}$ is automatically continuous. Two main ingredients are used to prove that $D_\tau\eta_{k+1} = \Phi_{k+1}$. The first induction hypothesis is used to justify the interchange of the partial derivative operator D_τ and the integral in the expression for the operator Λ. The second ingredient is an easy result from the theory of distributions. It states that the function $t \mapsto D_t\eta_k(t,\xi,\tau)$ corresponding to the classical partial derivative $D_t\eta_k$, which is assumed to exist and be continuous by the induction hypothesis, is in the equivalence class of the corresponding distributional derivative $\dot\eta(\xi,\tau)$ for each $(\xi,\tau) \in V \times [0,b]$. $\square$

3. Expansion in the small parameter

By Theorems 2.1–2.3, we know that under appropriate restrictions on the size of the delay, the delay equation (4), given by $\dot x(t) = f(x(t), x(t-\tau))$, has an n-dimensional C^1 inertial manifold consisting of special solutions. Theorem 2.1 implies that the inertial manifold is exponentially attracting; Theorems 2.2 and 2.3 imply that the inertial manifold is smooth. In this case, the inertial vector field (5), given by $X(\xi,\tau) = f(\xi, \eta(-\tau,\xi,\tau))$, is the generator of the corresponding special flow. Moreover, X is a C^1 function. In this section, we will assume that the function f, the solutions of the delay equation, and the inertial vector field are sufficiently smooth so that their Taylor expansions are defined. As mentioned in the Introduction, we will show that the inertial vector field agrees with the slow vector field on an appropriately chosen slow manifold for the singular perturbation problem obtained by expanding the delay equation to second order in powers of the delay τ. We will also show agreement to all orders for the linear delay equation $\dot x(t) = Ax(t-\tau)$, where A is an invertible $n \times n$-matrix.

3.1. Inertial manifold reduction

The expansion of the family of inertial vector fields $X(\xi,\tau) = f(\xi, \eta(-\tau,\xi,\tau))$ with respect to τ at $\tau = 0$ is

$$\begin{aligned} X(\xi,\tau) = {} & f(\xi,\xi) - \tau D_2 f(\xi,\xi) f(\xi,\xi) + \frac{\tau^2}{2!}\{D_2^2 f(\xi,\xi)(f(\xi,\xi), f(\xi,\xi)) \\ & + D_2 f(\xi,\xi)(D_1 f(\xi,\xi) + 3D_2 f(\xi,\xi)) f(\xi,\xi)\} + O(\tau^3). \end{aligned} \tag{19}$$

This result is obtained by using the invariance of the special flow η; in fact, we have that

$$\dot{\eta}(t,\xi,\tau) = f(\eta(t,\xi,\tau),\eta(t-\tau,\xi,\tau)), \quad \eta(0,\xi,\tau) = \xi.$$

Clearly, $X(\xi,0) = f(\xi,\xi)$ and

$$X_\tau(\xi,\tau) = D_2 f(\xi,\eta(-\tau,\xi,\tau))(\eta_\tau(-\tau,\xi,\tau) - f(\eta(-\tau,\xi,\tau),\eta(-2\tau,\xi,\tau))).$$

Since $\eta(0,\xi,\tau) = \xi$, all derivatives of the function $\tau \mapsto \eta(0,\xi,\tau)$ vanish. Hence, it follows that

$$X_\tau(\xi,0) = -D_2 f(\xi,\xi) f(\xi,\xi)$$

and

$$\begin{aligned} X_{\tau\tau}(\xi,0) = {} & D_2^2 f(\xi,\xi)(f(\xi,\xi),f(\xi,\xi)) + D_2 f(\xi,\xi)(-\dot{\eta}_\tau(0,\xi,0) \\ & + D_1 f(\xi,\xi) f(\xi,\xi) + 2D_2 f(\xi,\xi) f(\xi,\xi)). \end{aligned}$$

Finally, note that

$$\begin{aligned} \dot{\eta}_\tau(0,\xi,0) &= \frac{\partial}{\partial\tau}\frac{\partial}{\partial t}\eta(t,\xi,\tau)\bigg|_{t=0,\tau=0} \\ &= \frac{\partial}{\partial\tau} f(\xi,\eta(-\tau,\xi,\tau))\bigg|_{\tau=0} \\ &= -D_2 f(\xi,\xi) f(\xi,\xi). \end{aligned}$$

In view of the representation of X in display (19), it is clear that sufficiently small delays do not matter if the vector field given by $\xi \mapsto f(\xi,\xi)$ is structurally stable (cf. [20]). On the other hand, using the theorems in Section 2, the range of delays for which the inertial manifold exists can be estimated. Once this is done, the parameter τ is rescued from the "realm of the sufficiently small," and the effect of the perturbation caused by the delay remains to be determined. In case the vector field $\xi \mapsto f(\xi,\xi)$ is not structurally stable (for instance, if the vector field is Hamiltonian), then even sufficiently small delays do matter. As a simple illustration, consider the delay (Duffing) equation

$$\ddot{x} + \omega^2 x = -ax(t-\tau) + bx^3(t-\tau). \tag{20}$$

An equivalent first-order system of delay equations has the form of Eq. (4), but it is not Lipschitz on $\mathbb{R}^2$. This difficulty is easily remedied by using a cut-off function, defined on $\mathbb{R}^2$, to create a new system that agrees with the original system on some open ball at the origin and is constant in the complement of a larger open ball. The modified system has an inertial manifold for small τ and the first-order approximation (computed using Eq. (19)) of the reduced system on the corresponding inertial manifold is given by

$$\ddot{x} + \tau(3bx^2 - a)\dot{x} + (a + \omega^2)x - bx^3 = 0 \tag{21}$$

on the open ball at the origin where the modified system agrees with the original system. While the delay equation (20) with $\tau = 0$ is conservative, the second-order differential equation (21) is a form of van der Pol's oscillator; it has a stable limit cycle for appropriate choices of its parameters. Thus, small delays certainly do matter in this case.

3.2. *Post-Newtonian expansion*

As discussed in Section 1 (see also [4,5]), we will mimic post-Newtonian expansion in classical field theory using the delay equation (4),

$$\dot{x}(t) = f(x(t), x(t - \tau)).$$

Here, post-Newtonian expansion means Taylor expansion of the function $\tau \mapsto f(x(t), x(t - \tau))$ to some *finite* order in powers of τ. We will show how to obtain the inertial vector field—the vector field which gives the correct long-term dynamics of the delay equation—from the post-Newtonian expansion. It turns out that the slow vector field on a slow manifold of an appropriately chosen singularly perturbed system, naturally derived from the post-Newtonian expansion, agrees with the inertial vector field.

The post-Newtonian expansion of the function $\tau \mapsto f(x(t), x(t - \tau))$ at $\tau = 0$ is given by

$$\begin{aligned} f(x(t), x(t - \tau)) = {} & f(x(t), x(t)) - \tau D_2 f(x(t), x(t))\dot{x}(t) \\ & + \frac{\tau^2}{2!}(D_2 f(x(t), x(t))\ddot{x}(t) + D_2^2 f(x(t), x(t))(\dot{x}(t), \dot{x}(t))) \\ & + O(\tau^3). \end{aligned} \tag{22}$$

Note that the truncation of this expansion at order N in τ, when set equal to $\dot{x}(t)$, produces an Nth-order ODE of the form

$$(-1)^N \frac{\tau^N}{N!} D_2 f(x, x) x^{(N)} = F(x, \dot{x}, \ldots, x^{(N-1)}, \tau), \tag{23}$$

where

$$\begin{aligned} F(x, \dot{x}, \ldots, x^{(N-1)}, \tau) := {} & \dot{x} - f(x, x) + \tau D_2 f(x, x) \\ & - \frac{\tau^2}{2!} D_2^2 f(x, x)(\dot{x}, \dot{x})) + O(\tau^3). \end{aligned}$$

Because Eq. (23) is singular in the limit as $\tau \to 0$, many examples of such systems will contain "runaway" solutions (that is, solutions that are unbounded in forward time).

For this reason, the post-Newtonian approximation is not satisfactory as a physical model as mentioned in Section 1.

Since, for small delays, the long-term dynamics of the delay equation (4) is obtained by reduction to its inertial manifold, it is clear that the utility of post-Newtonian expansion can be justified only if there is a natural way to extract the dynamical system on the inertial manifold from the post-Newtonian expansion. Fortunately, there is a well-established method for approaching this problem: reduction to a slow manifold.

We will illustrate the method for the singularly perturbed Nth-order ODE (23). Suppose that $N>1$ and let $\mu := \tau^{1/(N-1)}$. The differential equation (23) is equivalent to the singularly perturbed first-order system

$$\begin{aligned}
\dot{x} &= y_1,\\
\mu^N \dot{y}_1 &= y_2,\\
&\vdots\\
\mu^N \dot{y}_{N-2} &= y_{N-1},\\
\mu^N (-1)^N \frac{1}{N!} D_2 f(x,x)\dot{y}_{N-1} &= F(x, y_1, \ldots, y_{N-1}, \mu^{N-1}).
\end{aligned} \tag{24}$$

Under the assumption that $D_2 f(x,x)$ has no eigenvalue with zero real part and τ is sufficiently small, this system has an n-dimensional slow manifold, an invariant manifold with the same dimension as the inertial manifold for the underlying delay equation (4) (see Proposition 3.1). Moreover, the reduction of the dynamical system (19) to this slow manifold agrees with the dynamical system (19) on the inertial manifold.

While the inertial manifold attracts nearby solutions, the slow manifold generally has both stable and unstable manifolds, that is, some solutions are attracted to the slow manifold and some solutions are repelled. The unstable directions correspond to the runaway modes. Also, it should be clear that *only* the solutions on the slow manifold of system (24) have physical significance; all other solutions are merely artifacts of the truncation of the post-Newtonian expansion.

As a convenient terminology, let us call the dynamical system on the slow manifold of system (24) the post-Newtonian approximation. We will show that this approximation is useful by proving that it agrees with the dynamical system on the corresponding inertial manifold.

The geometric theory for singular perturbation problems—initiated by the pioneering work of Fenichel [11,12]—is by now well developed. We will explain the basic idea and then apply the result to the truncation of system (24).

The basic singular perturbation problem is given by a system of the form

$$\dot{x} = f(x, y), \qquad \varepsilon\dot{y} = g(x, y),$$

where ε is a small parameter and f and g are smooth functions. Note that the reduced system obtained by setting $\varepsilon = 0$ is not a differential equation. To overcome this difficulty, let us introduce the "fast time" $s := t/\varepsilon$, and thereby recast the system into the regular perturbation form

$$x' = \varepsilon f(x, y), \qquad y' = g(x, y),$$

where "′" denotes differentiation with respect to s. This fast-time system is equivalent to the original system for $\varepsilon \neq 0$, the important values of this parameter.

The unperturbed fast-time system

$$x' = 0, \qquad y' = g(x, y)$$

has the invariant set $\Gamma_0 := \{(x, y) : g(x, y) = 0\}$ consisting entirely of rest points. Under the generic assumption that the partial derivative $g_y(x, y)$ is an invertible linear map whenever $(x, y) \in \Gamma_0$, an application of the implicit function theorem can be used to show that Γ_0 is a smooth manifold given by the graph of a function $y = \alpha(x)$; that is, $\Gamma_0 := \{(x, y) : y = \alpha(x)\}$. Under the stronger assumption that $g_y(x, y)$ has no eigenvalue with real part zero, the solutions of the unperturbed system starting near Γ_0 are all either attracted to, or repelled from Γ_0. This fact is easily seen by linearization of the system at a rest point on Γ_0. At each such point (x, y), the system matrix of the linearization is given by

$$\begin{pmatrix} 0 & 0 \\ g_x(x, y) & g_y(x, y) \end{pmatrix}.$$

The block of zero eigenvalues corresponds to the motion along the invariant set. Eigenvalues of $g_y(x, y)$ with positive real parts correspond to exponentially fast expansion from Γ_0 (runaway modes); eigenvalues with negative real parts correspond to exponentially fast contraction to Γ_0. In other words, the rate of contraction in the normal direction dominates the fastest rate of contraction—in this case zero—on the invariant manifold and, likewise, the normal rate of expansion dominates the fastest expansion on the invariant manifold. An invariant manifold consisting entirely of rest points with these properties is called *normally hyperbolic* (see [11] for the general definition).

By a fundamental result of Fenichel, a normally hyperbolic invariant manifold persists (for sufficiently small values of the small parameter) as a normally hyperbolic invariant manifold—again given as a graph—in the full nonlinear fast-time system. Because the flow on the normally hyperbolic invariant manifold for the unperturbed system is stationary, it is (infinitely) slow relative to the ambient flow. Under a small perturbation, the flow on the new invariant manifold likewise is slow relative to the perturbed ambient flow. For this reason, these invariant manifolds are called *slow manifolds*, and the corresponding flows on these manifolds are also called slow. It is important to realize that the slow manifolds for $\varepsilon \neq 0$ remain invariant sets for the original singularly perturbed system. In fact, the qualitative features of the dynamical behavior of the fast- and slow-time systems for $\varepsilon \neq 0$ are identical; only the speed at which points move along trajectories is different.

The existence of a family of *invariant* manifolds for the family of fast-time systems, given as a family of graphs

$$\Gamma_\varepsilon := \{(x, y) : y = \alpha(x, \varepsilon)\},$$

ensures that

$$y' = \alpha_x(x, \varepsilon)x',$$

whenever $y = \alpha(x, \varepsilon)$. It follows that α is defined implicitly by the relation

$$g(x, \alpha(x, \varepsilon)) = \varepsilon\alpha_x(x, \varepsilon)f(x, \alpha(x, \varepsilon)),$$

and therefore α can be approximated in the usual manner by equating coefficients after power series expansion in the small parameter. The reduction of the dynamical system to the slow manifold is given by the ODE

$$x' = \varepsilon f(x, \alpha(x, \varepsilon));$$

the corresponding slow-time ODE is

$$\dot{x} = f(x, \alpha(x, \varepsilon)).$$

For many problems, this last equation determines the essential dynamical behavior of the original nonlinear singular perturbation problem.

Let us now return to a truncation of the post-Newtonian expansion of the delay equation (4). The next proposition states a sufficient condition for the singular system (24) to have a normally hyperbolic invariant manifold. More importantly, even if this condition is not satisfied, the formal slow vector field (that is, the hypothetical restriction of the vector field corresponding to system (24) to a hypothetical slow manifold) agrees with the inertial vector field of the delay equation (4). Indeed, if we wish to determine the long-term dynamics of the delay equation, our objective is to obtain the dynamical system on its inertial manifold. From this point of view, the existence of the slow-manifold is not important; it is simply a construct that gives an alternative way to obtain the inertial vector field. On the other hand, it might be possible to prove the existence of the desired inertial manifold under the assumption that an infinite sequence of post-Newtonian truncations have slow-manifolds. Also, there are cases where post-Newtonian approximations are obtained without reference to a specific delay-type equation. Thus, the conditions for the existence of slow manifolds for the post-Newtonian truncations has some independent interest.

Proposition 3.1. *Suppose that N is a positive even integer and, for all* $x \in \mathbb{R}^n$, *the matrix* $D_2 f(x,x)$ *has no eigenvalue with zero real part. If* $\tau > 0$ *is sufficiently small, then system* (24) *has a normally hyperbolic n-dimensional slow manifold. Moreover, if* $N \geqslant 2$ *is a positive integer, then the formal slow vector field agrees to order two in* τ *with the corresponding inertial vector field on the* (*n-dimensional*) *inertial manifold of the delay equation* (4).

Remark. The simple condition for normal hyperbolicity given in Proposition 3.1 for truncations at even orders is due (in part) to the formulation used here of a corresponding first-order singular perturbation problem (see system (26)) where the small parameter is taken to be $\mu := \tau^{1/(N-1)}$. This relation is invertible in a neighborhood of the origin only if N is even.

Proof. We will consider two cases: $N = 2$ and $N > 2$. Note that because zero is not an eigenvalue of $D_2 f(x,x)$, this linear transformation is invertible. For $N = 2$, we have $\mu = \tau$ and system (24) has the form

$$\begin{aligned} \dot{x} &= y_1, \\ \tau^2 \dot{y}_1 &= 2(D_2 f(x,x))^{-1}(y_1 - f(x,x) + \tau D_2 f(x,x) y_1 \\ &\quad - \frac{\tau^2}{2} D_2^2 f(x,x)(y_1, y_1)). \end{aligned} \tag{25}$$

In case $N > 2$ system (24) can be recast as

$$\begin{aligned} \dot{x} &= y_1, \\ \mu^N \dot{y}_1 &= y_2, \\ &\vdots \\ \mu^N \dot{y}_{N-2} &= y_{N-1}, \\ \mu^N \dot{y}_{N-1} &= (-1)^N N! (D_2 f(x,x))^{-1} \big(y_1 - f(x,x) + \mu^{N-1} D_2 f(x,x) y_1 \\ &\quad - \frac{\mu^{2(N-1)}}{2!} \left(\frac{1}{\mu^N} D_2 f(x,x) y_2 + D_2^2 f(x,x)(y_1, y_1) \right) + \mathscr{F} \bigg), \end{aligned} \tag{26}$$

where $\mathscr{F}$ is a sum of terms obtained from the terms of order three through N in the Taylor expansion (22). The essential observation is that in these terms—and in every other term—the jth time-derivative of x appearing in expansion (22) is replaced by $(1/\mu^{jN}) y_j$.

By changing to the fast-time $s := t/\mu^N$, we obtain the system

$$
\begin{aligned}
x' &= \mu^N y_1,\\
y_1' &= y_2,\\
y_2' &= y_3,\\
&\vdots\\
y_{N-2}' &= y_{N-1},\\
y_{N-1}' &= (-1)^N N!(D_2 f(x,x))^{-1}\Big(y_1 - f(x,x) + \mu^{N-1} D_2 f(x,x) y_1\\
&\quad - \frac{\mu^{2(N-1)}}{2!}\Big(\frac{1}{\mu^N} D_2 f(x,x) y_2 + D_2^2 f(x,x)(y_1,y_1)\Big) + \mathscr{F}\Big),
\end{aligned}
\tag{27}
$$

which is equivalent to system (26) for $\mu > 0$. Note that

$$\Gamma := \{(x, y_1, \ldots, y_{N-1}) : y_1 = f(x,x), y_2 = y_3 = \cdots = y_{N-1} = 0\}$$

is an invariant manifold, consisting entirely of rest points, for system (27) with $\mu = 0$. Because Γ is defined by $n(N-1)$ equations in $\mathbb{R}^{nN}$, Γ is an n-dimensional manifold. Equivalently, Γ is the graph of the function $G : \mathbb{R}^n \to \mathbb{R}^{n(N-1)}$ given by

$$x \mapsto (f(x,x), 0, 0, \ldots, 0).$$

We will show that Γ is normally hyperbolic whenever N is even.

The linearized system at a point $(x, y_1, y_2, \ldots, y_{N-1})$ on Γ is given by the system matrix

$$
\begin{pmatrix}
0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0\\
0 & 0 & I & 0 & 0 & 0 & \cdots & 0 & 0\\
0 & 0 & 0 & I & 0 & 0 & \cdots & 0 & 0\\
\vdots & & & & & & & & \\
0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & I\\
\mathscr{A}(x,N) & \mathscr{B}(x,N) & 0 & 0 & 0 & 0 & \cdots & 0 & 0
\end{pmatrix},
$$

where

$$
\begin{aligned}
\mathscr{A}(x,N) &:= (-1)^{N+1} N!(D_2 f(x,x))^{-1}(D_1 f(x,x) + D_2 f(x,x)),\\
\mathscr{B}(x,N) &:= (-1)^N N!(D_2 f(x,x))^{-1}.
\end{aligned}
$$

The vector $v \coloneqq (\xi, \eta_1, \eta_2, \ldots, \eta_{N-1})$ is in the kernel of the system matrix if and only if $\eta_2 = \eta_3 = \cdots \eta_{N-1} = 0$ and $\eta_1 - (D_1 f(x,x) + D_2 f(x,x))\xi = 0$, exactly the set of conditions required for v to be tangent to the manifold Γ. Hence, we have proved that the system has exactly n zero eigenvalues and these correspond to the "eigendirections" tangent to Γ. To prove that Γ is normally hyperbolic, it suffices to show that the system matrix has no nonzero pure imaginary eigenvalue.

If $i\beta$, with $\beta \neq 0$, is an eigenvalue with (complex) eigenvector

$$v \coloneqq (\xi, \eta_1, \eta_2, \ldots, \eta_{N-1}),$$

then $0 = (i\beta)\xi,\ \eta_j = i\beta\eta_{j-1}$ for $j \in \{2, 3, \ldots, N-1\}$, and

$$\mathscr{A}(x,N)\xi + B(x,N)\eta_1 = i\beta\eta_{N-1}.$$

It follows that $\xi = 0$ and

$$\mathscr{B}(x,N)\eta_1 = (i\beta)^{N-1}\eta_1.$$

Because $N-1$ is an odd integer, $(i\beta)^{N-1}$ is pure imaginary. Hence, $D_2 f(x,x)$ would have an eigenvalue with zero real part, contrary to the hypothesis. By an application of Fenichel's theory, the normally hyperbolic manifold Γ persists for sufficiently small $\mu \neq 0$. Moreover, because $N-1$ is odd, the relation $\mu = \tau^{1/(N-1)}$ is invertible and system (24) also has a normally hyperbolic slow manifold for sufficiently small $\tau \neq 0$.

It remains to show that the slow vector field (that is, the restriction of the vector field corresponding to system (24) to the slow manifold) agrees to order two in τ with the inertial vector field (19) for system (24). The differential equation on the slow manifold is given by

$$x' = \mu^N y_1(x,\mu),$$

which, in the original time-scale, is

$$\dot{x} = y_1(x,\mu).$$

Thus, the slow vector field is given by $x \mapsto y_1(x,\mu)$.

Note that the last equation in system (27) has the form

$$y'_{N-1} = \mathscr{B}(x,N)\left(y_1 - \sum_{j=0}^{N-1} a_j(x, \mu^{N-1}y_1, \mu^{N-2}y_2, \ldots, \mu^{N-j}y_j)\right), \tag{28}$$

where the functions $a_j,\ j \in \{1, 2, \ldots, N-1\}$, are polynomials in their last j variables and $a_0(x) = f(x,x)$.

Claim 3.2. *The expansion of the slow vector field y_1 in powers of μ has the form*

$$y_1(x,\mu) = \sum_{i=0}^{N-1} \mu^{i(N-1)} y_{1,i(N-1)}(x) + O(\mu^{(N-1)(N-1)+1}). \tag{29}$$

The formal series expansion of y_1 has the form

$$y_1(x,\mu) = \sum_{i=0}^{\infty} \mu^i y_{1,i}(x). \tag{30}$$

Using the invariance of the slow manifold, each y_j is obtained from this expression, in turn, by differentiating with respect to x and multiplication by $\mu^N y_1(x,\mu)$. To see this, consider a solution of system (27) where the first component is $t \mapsto x(t,\mu)$. For example, we have that

$$y_2(x(t,\mu),\mu) = \mu^N \frac{d}{dt} y_1(x(t,\mu),\mu) = \mu^N D y_1(x(t,\mu),\mu) y_1(x(t,\mu),\mu),$$

and therefore

$$y_2(x,\mu) = \mu^N D y_1(x,\mu) y_1(x,\mu).$$

By this procedure, it is clear that the leading term of the series expansion for y_j has order $\mu^{(j-1)N}$.

To determine the form of the series expansion (30), substitute it into system (27) and note that the left-hand side of the last equation of the resulting system has leading-order $\mu^{(N-1)N}$. Hence, all terms of lower order in the series expansion for y_1 are determined by the equation

$$y_1 = \sum_{j=0}^{N-1} a_j(x, \mu^{N-1} y_1, \mu^{N-2} y_2, \dots, \mu^{N-j} y_j). \tag{31}$$

Recall that the leading term of the expansion of y_j has order $\mu^{(j-1)N}$ and substitute these series into the right-hand side of Eq. (31). After these substitutions, the leading term of the series expansion of the resulting right-hand side of Eq. (31) has order zero, and its next term has order $N-1$.

Thus, the leading term of the series expansion of y_1, the left-hand side of Eq. (31), has order zero, and its next term has order $N-1$. Using this fact, recompute the series for the y_j as indicated above and note that

$$y_j = \mu^{(j-1)N}(y_{j,0} + \mu^{N-1} y_{j,N-1} + O(\mu^N)).$$

To recompute the series expansion for the right-hand side of Eq. (31), we now substitute

$$\mu^{N-j}y_j = \mu^{j(N-1)}y_{j,0} + \mu^{(j+1)(N-1)}y_{j,N-1} + O(\mu^{(j+1)(N-1)+1}).$$

Note that the first three terms of the series expansion of the right-hand side of Eq. (31) must now have orders 0, $N-1$, and $2(N-1)$. Hence, the first three terms on the left-hand side of Eq. (31) have the same orders, and it follows that

$$y_1 = y_{1,0} + \mu^{N-1}y_{1,N-1} + \mu^{2(N-1)}y_{1,2(N-1)} + O(\mu^{2(N-1)+1}).$$

Proceeding by induction, let us suppose that

$$\begin{aligned} y_1 = {} & y_{1,0} + \mu^{N-1}y_{1,N-1} + \mu^{2(N-1)}y_{1,2(N-1)} \\ & + \cdots + \mu^{i(N-1)}y_{1,i(N-1)} + O(\mu^{i(N-1)+1}). \end{aligned} \tag{32a}$$

In this case,

$$y_j = \mu^{(j-1)N}(y_{j,0} + \mu^{N-1}y_{j,N-1} + \cdots + \mu^{i(N-1)}y_{j,i(N-1)} + O(\mu^{i(N-1)+1}))$$

and

$$\begin{aligned} \mu^{N-j}y_j = {} & \mu^{j(N-1)}y_{j,0} + \mu^{(j+1)(N-1)}y_{j,N-1} \\ & + \cdots + \mu^{(j+i)(N-1)}y_{j,i(N-1)} + O(\mu^{(j+i)(N-1)+1}). \end{aligned}$$

The essential feature of these series is that the order of every term, whose order is less than or equal to $(i+1)(N+1)$, has the form $k(N-1)$ for some integer $k \in \{0, 1, 2, \ldots, i+1\}$. After substitution of these series, the series expansion of the resulting right-hand side of Eq. (31) has the same property. Hence, so does the left-hand side of Eq. (31), as required.

Consider system (25). Using Claim 3.2 and retaining only terms with order less than three in τ, the slow vector field is given by

$$y_1(x) = f(x,x) + \tau y_{1,1}(x) + \tau^2 y_{1,2}(x).$$

By substitution of this expression into the last equation in system (25) and by retaining only the appropriate low-order terms, the functions $y_{1,1}$ and $y_{1,2}$ are determined by equating terms of the same order from the left- and right-hand sides of the equation

$$\begin{aligned} \tau^2(D_1 f(x,x)y_1(x) + D_2 f(x,x)y_1(x)) = 2(D_2 f(x,x))^{-1}\Big(& \tau y_{1,1}(x) + \tau^2 y_{1,2}(x) \\ & + \tau D_2 f(x,x)(f(x,x) + \tau y_{1,1}(x)) \\ & - \frac{\tau^2}{2} D_2^2 f(x,x)(f(x,x), f(x,x))\Big). \end{aligned}$$

In fact, $y_{1,1}$ and $y_{1,2}$ agree with the first- and second-order terms in expansion (19) of the inertial vector field.

In case $N > 2$, we again use Claim 3.2 and the ansatz

$$y_1(x) = f(x,x) + \mu^{N-1} y_{1,N-1}(x) + \mu^{2(N-1)} y_{1,2(N-1)}(x). \tag{32b}$$

Also, we note that after substitution into equations two through $N-2$ in system (26) it follows that the leading term in the expansion of y_i has order $\mu^{(i-1)N}$. Thus, after substitution in the last equation of this system, the leading term on its left-hand side has order $\mu^{(N-1)N}$; therefore, unlike in the case $N = 2$, these terms do not enter into the determination of the coefficients of the slow vector field to order $\mu^{2(N-1)}$. In fact, the coefficients of Eq. (32) are determined by equating to zero the terms of order $\mu^{i(N-1)}$, for $i = 0, 1, 2$, in the expression

$$\begin{aligned} & y_1(x) - f(x,x) + \mu^{N-1} D_2 f(x,x) y_1(x) \\ & - \frac{\mu^{2(N-1)}}{2!} \left(\frac{1}{\mu^N} D_2 f(x,x) y_2 + D_2 f(x,x)(y_1, y_1) \right) \end{aligned}$$

after substitution using Eq. (32). Since the expansion of y_2 has leading-order μ^N, it "cancels" the factor $1/\mu^N$. The coefficients $y_{1,N-1}$ and $y_{1,2(N-1)}$ determined in this manner agree with the first- and second-order terms in expansion (19) of the inertial vector field. □

Conjecture 3.3. *The formal expansion of the slow vector field corresponding to system* (26) *agrees to order N in τ with expansion* (19) *of the inertial vector field.*

We will prove a special case of Conjecture 3.3.

Theorem 3.4. *Suppose that A is an $n \times n$ matrix. If $|\tau| \|A\| e < 1$, then the delay equation $\dot{x}(t) = Ax(t-\tau)$ has an inertial manifold, and its inertial vector field is given by*

$$X(x,\tau) = \sum_{j=0}^{\infty} (-1)^j \frac{(1+j)^j}{(1+j)!} \tau^j A^{1+j} x.$$

Moreover, if $N \geqslant 1$ and A is invertible, then the expansion in powers of τ of the slow vector field corresponding to system (26) *agrees to order N with the inertial vector field X.*

Proof. By the ratio test, if $|\tau| \|A\| e < 1$, then the series in the statement of the theorem converges.

By Theorem 2.1, the linear delay equation has a special flow $y = y(t,x,\tau)$. We will show that its generator is given by $X = X(x,\tau)$.

Define $\phi(t,x,\tau) := y(\tau t, x, \tau)$ so that

$$\begin{aligned}\dot{\phi}(t,x,\tau) &= \tau \dot{y}(\tau t, x, \tau)\\ &= \tau A y(\tau t - \tau, x, \tau)\\ &= \tau A y(\tau(t-1), x, \tau)\\ &= \tau A \phi(t-1, x, \tau)\end{aligned}$$

and

$$\phi(t,x,0) \equiv x.$$

Also, in this case, the inertial vector field is given by

$$Ay(-\tau, x, \tau) = A\phi(-1, x, \tau).$$

We will show that

$$X(\tau, x) := A\phi(-1, x, \tau)$$

has the series expansion in the statement of the theorem.

Note that $X(0,x) = A\phi(-1,x,0) = Ay(0,x,0) = Ax$. The Taylor series at $\tau = 0$ is determined from the partial derivatives of ϕ with respect to τ. We will compute these partial derivatives from appropriate variational equations. We have that

$$\dot{\phi}(t,x,\tau) = \tau A \phi(t-1,x,\tau),$$

$$\frac{d}{dt}\frac{\partial \phi}{\partial \tau}(t,x,\tau) = A\phi(t-1,x,\tau) + \tau A \frac{\partial \phi}{\partial \tau}(t-1,x,\tau),$$

and, by induction for $j \geqslant 1$,

$$\frac{d}{dt}\frac{\partial^j \phi}{\partial \tau^j}(t,x,\tau) = jA\frac{\partial^{j-1}\phi}{\partial \tau^{j-1}}(t-1,x,\tau) + \tau A \frac{\partial^j \phi}{\partial \tau^j}(t-1,x,\tau).$$

After evaluation at $\tau = 0$,

$$\frac{d}{dt}\frac{\partial^j \phi}{\partial \tau^j}(t,x,0) = jA\frac{\partial^{j-1}\phi}{\partial \tau^{j-1}}(t-1,x,0),$$

and, by integration,

$$\frac{\partial^j \phi}{\partial \tau^j}(t,x,0) = jA\int_0^t \frac{\partial^{j-1}\phi}{\partial \tau^{j-1}}(s-1,x,0)\,ds.$$

For $j \geqslant 1$, the jth Taylor coefficient $X_j(x)$ of $X(\tau, x)$ is given by

$$X_j(x) = \frac{1}{j!} A \frac{\partial^j \phi}{\partial \tau^j}(-1, x, 0).$$

It is now clear that the Taylor coefficients of X are determined by the following algorithm:

Input j;
  $\Phi_0(t, x) := x$;
  $X_0(x) := Ax$;
If $j = 0$ Go To Output;
For k From 1 To j Do
  $\Phi_k(t, x) := kA \int_0^t \Phi_{k-1}(s-1, x)\, ds$,
  $X_k(x) := \frac{1}{k!} \Phi_k(-1, x)$,
End For Loop;
Output $X_j(x)$.

By induction, it is easy to see that

$$\Phi_j(t, x) = t(t-j)^{j-1} A^j x,$$

$$X_j(x) = (-1)^j \frac{(1+j)^j}{(1+j)!} A^{1+j} x, \tag{33}$$

as required.

Let us now consider the slow vector field. By replacing the right-hand side of the delay equation $\dot{x}(t) = Ax(t-\tau)$ with its Taylor polynomial of degree N at $\tau = 0$ and rearranging the terms in the resulting equation, we obtain the Nth-order ODE

$$\tau^N x^{(N)} = (-1)^N N! A^{-1} \left(x^{(1)} - \sum_{j=0}^{N-1} (-1)^j \frac{\tau^j}{j!} A x^{(j)} \right), \tag{34}$$

where $x^{(j)}$ denotes the jth derivative with respect to the slow-time t.

For convenience of notation, let us consider the series expansion of the slow vector field, given by y_1 (as in Claim 3.2), in the form

$$y_1(x, \tau) = \sum_{k=0}^{N} \mu^{k(N-1)} \rho_{1,k}(x), \tag{35}$$

where $\mu = 1/\tau^{N-1}$. The corresponding expansions of the elements of the sequence $\{y_j\}_{j=2}^{N-1}$ are determined in turn using the invariance of the slow manifold for system (27) and the recursive definition $y_j = y'_{j-1}$ where the differentiation is with respect to the fast-time s. More precisely,

$$y_j(x, \tau) = Dy_{j-1}(x, \tau)x' = Dy_{j-1}(x, \tau)\mu^N y_1(x, \tau),$$

where D denotes the derivative with respect to the space variable x. It follows that

$$\begin{aligned} y_2(x,\tau) &= \mu^N \left(\sum_{k=0}^{N} \mu^{k(N-1)} D\rho_{1,k}(x) \right) \left(\sum_{i=0}^{N} \mu^{i(N-1)} \rho_{1,i}(x) \right) \\ &= \mu^N \sum_{k=0}^{N} \mu^{k(N-1)} \left(\sum_{i=0}^{k} D\rho_{1,i}(x)\rho_{1,k-i}(x) \right) + O(\mu^N \mu^{N^2-1}) \\ &= \mu^N \sum_{k=0}^{N} \mu^{k(N-1)} \rho_{2,k}(x) + O(\mu^N \mu^{N^2-1}), \end{aligned}$$

where

$$\rho_{2,k}(x) \coloneqq \sum_{i=0}^{k} D\rho_{1,i}(x)\rho_{1,k-i}(x),$$

and, by induction for $j = 2, 3, \ldots, N-1$,

$$y_j(x,\tau) = \mu^{(j-1)N} \sum_{k=0}^{N-1} \mu^{k(N-1)} \rho_{j,k}(x) + O(\mu^{(j-1)N} \mu^{N^2-1}), \tag{36}$$

where

$$\rho_{j,k}(x) \coloneqq \sum_{i=0}^{k} D\rho_{j-1,i}(x)\rho_{1,k-i}(x).$$

Let us determine the coefficients $\{\rho_{1,k}\}_{k=0}^{N}$ by substitution into the fast-time system (27). Since the leading-order term of the expansion for y'_{N-1} has order $\mu^{(N-1)N}$, the coefficients $\{\rho_{1,k}\}_{k=0}^{N-1}$ are determined by equating to zero the right-hand side of the last equation in system (27). An easy computation shows that $y'_{N-1} = \tau^N x^{(N)}$. Hence, we can instead determine these coefficients by equating to zero the right-hand side Eq. (34). After substituting into this equation for the time-derivatives $x^{(j)}$ according to the definitions of the y_i in system (26), it follows immediately that the determining equation for the coefficients $\{\rho_{1,k}\}_{k=0}^{N-1}$ is given by

$$y_1 = Ax + \sum_{j=1}^{N-1} (-1)^j \frac{\mu^{N-j}}{j!} Ay_j. \tag{37}$$

Using the expansions of the y_i from Eq. (36), we have that

$$\begin{aligned} Ax + \sum_{j=1}^{N-1} \frac{(-1)^j}{j!} \mu^{N-j} A y_j &= Ax + A \sum_{j=1}^{N-1} \frac{(-1)^j}{j!} \mu^{N-j} \sum_{k=0}^{N-1} \mu^{k(N-1)} \rho_{j,k}(x) \\ &= Ax + A \sum_{j=1}^{N-1} \sum_{k=0}^{N-1} \frac{(-1)^j}{j!} \mu^{(j+k)(N-1)} \rho_{j,k}(x). \end{aligned}$$

By summing along "negative slope" diagonals in the (j,k)-index space, the last double sum can be rearranged so that

$$\begin{aligned} & Ax + \sum_{j=1}^{N-1} (-1)^j \frac{\mu^{N-j}}{j!} A y_j \\ &= Ax + A \sum_{k=1}^{N-1} \mu^{k(N-1)} \sum_{\ell=1}^{k} \frac{(-1)^\ell}{\ell!} \rho_{\ell,k-\ell}(x) + O(\mu^{(N-1)(N-1)+1}). \end{aligned} \tag{38}$$

Using Eq. (37) and comparing coefficients in expansions (35) and (38), it follows that the $\rho_{1,j}(x)$, for $j = 0, 1, \ldots, N-1$, are given by the following algorithm:

Input j;
  $\rho_{1,0}(x) := Ax$;
If $j = 0$ Go To Output;
  $\rho_{1,1}(x) := -A^2 x$;
If $j = 1$ Go To Output;
For k From 2 To j Do
  For ℓ From 2 To k Do
    $\rho_{\ell,k-\ell}(x) := \sum_{i=0}^{k-\ell} D\rho_{\ell-1,i}(x)\rho_{1,k-\ell-i}(x)$,
  End For Loop;
  $\rho_{1,k}(x) := A \sum_{i=1}^{k} \frac{(-1)^i}{i!} \rho_{i,k-i}(x)$,
End For Loop;
Output $\rho_{1,j}(x)$.

For $i \in \{1, 2, \ldots, N-1\}$ and $j \in \{0, 1, \ldots, N-1\}$, we will show that

$$\rho_{i,j}(x) = (-1)^j \frac{i(i+j)^{j-1}}{j!} A^{i+j} x. \tag{39}$$

In particular, if this representation is valid, then $X_j(x) = \rho_{1,j}(x)$ for $j \in \{0, 1, \ldots, N-1\}$ (see Eq. (33)). We will also use formula (39) to prove that $X_N(x) = \rho_{1,N}(x)$.

By inspection, $\rho_{1,0}$ and $\rho_{1,1}$, as defined by the algorithm, are given by the representation in display (39).

Suppose that $\rho_{\alpha,\beta}$, as defined by the algorithm for $1 \leqslant \alpha + \beta < k$, are given by the representation in display (39). We will show that the $\rho_{\ell,k-\ell}$, defined by the algorithm for $\ell \in \{2, 3, \dots, k\}$, are also given by the representation in display (39).

Using the induction hypothesis, we have

$$\begin{aligned}\rho_{\ell,k-\ell}(x) &:= \sum_{i=0}^{k-\ell} D\rho_{\ell-1,i}(x)\rho_{1,k-\ell-i}(x)\\ &= (-1)^{k-\ell} \sum_{i=0}^{k-\ell} \frac{(\ell-1)(\ell-1+i)^{i-1}(1+k-\ell-i)^{k-\ell-i-1}}{i!(k-\ell-i)!} A^k x.\end{aligned}$$

Thus, it suffices to show that

$$(\ell-1) \sum_{i=0}^{k-\ell} \frac{(\ell-1+i)^{i-1}(1+k-\ell-i)^{k-\ell-i-1}}{i!(k-\ell-i)!} = \frac{\ell k^{k-\ell-1}}{(k-\ell)!},$$

or equivalently,

$$\sum_{i=0}^{k-\ell} \binom{k-\ell}{i} (\ell-1+i)^{i-1}(1+k-\ell-i)^{k-\ell-i-1} = \frac{\ell}{\ell-1} k^{k-\ell-1}.$$

With $m := k - \ell$, the required identity is

$$\sum_{i=0}^{m} \binom{m}{i} (\ell-1+i)^{i-1}(m+1-i)^{m-i-1} = \frac{\ell}{\ell-1} (m+\ell)^{m-1}. \tag{40}$$

This nontrivial combinatorial identity is a special case of Abel's generalization of the binomial theorem; namely,

$$\alpha\beta \sum_{i=0}^{m} \binom{m}{i} (\alpha+i)^{i-1}(\beta+m-i)^{m-i-1} = (\alpha+\beta)(\alpha+\beta+m)^{m-1}$$

(see for example [13, p. 19]). In fact, identity (40) is obtained from Abel's identity with the replacements $\alpha = \ell - 1$ and $\beta = 1$.

To complete this part of the proof, we will show that $\rho_{1,k}$, as defined in the algorithm, is given by formula (39); or, in other words,

$$\begin{aligned}(-1)^k \frac{(1+k)^{k-1}}{k!} A^{k+1} x &= A \sum_{i=1}^{k} \frac{(-1)^i}{i!} \left((-1)^{k-i} \frac{ik^{k-i-1}}{(k-i)!} A^k x \right)\\ &= (-1)^k k^{k-1} \sum_{i=1}^{k} \frac{k^{-i}}{(k-i)!(i-1)!} A^{k+1} x.\end{aligned}$$

Clearly, it suffices to show that

$$\left(\frac{1+k}{k}\right)^{k-1} = \sum_{i=1}^{k} \frac{k!}{k^i (k-i)!(i-1)!}. \tag{41}$$

But, since

$$\begin{aligned}\left(1+\frac{1}{k}\right)^{k-1} &= \sum_{i=0}^{k-1} \binom{k-1}{i} \frac{1}{k^i} \\ &= \sum_{i=0}^{k-1} \binom{k}{i+1} \frac{i+1}{k^{i+1}} \\ &= \sum_{i=1}^{k} \binom{k}{i} \frac{i}{k^i} \\ &= \sum_{i=1}^{k} \frac{k!}{k^i (k-i)!(i-1)!},\end{aligned}$$

the required identity is a corollary of the binomial theorem.

To prove that $X_N = \rho_{1,N}$, let us equate the terms of order $\mu^{N(N-1)}$ in the last equation of system (27). In the present case, this equation is obtained from Eq. (34). After substitution of the series expansions for the y_i, the left-hand side of the equation has one term of the required order, namely $\mu^{N(N-1)} A^n x$. After multiplication of both sides of the equation by the inverse of the factor $(-1)^N N! A^{-1}$ and some algebraic manipulation, it follows that

$$\rho_{1,N}(x) = \frac{(-1)^N}{N!} A^{N+1} x + A \sum_{j=1}^{N-1} \frac{(-1)^j}{j!} [y_j]_{N(N-1)+j-N},$$

where $[y_j]_{N(N-1)+j-N}$ denotes the coefficient of order $\mu^{N(N-1)+j-N}$ in the series expansion of y_j. Using formula (36), this coefficient is $\rho_{j,N-j}$, which is given explicitly in display (39). After some simplification, it follows that

$$\rho_{1,N}(x) = (-1)^N \sum_{j=1}^{N} \frac{N^{N-j-1}}{(j-1)!(N-j)!} A^{N+1} x.$$

By inspection, this coefficient is equal to the coefficient of τ^N in the expansion of $X(x,\tau)$ if

$$\left(\frac{1+N}{N}\right)^{N-1} = \sum_{j=1}^{N} \frac{N!}{N^j (N-j)!(j-1)!};$$

therefore, the desired result follows from identity (41). $\square$

Acknowledgments

The author thanks the anonymous referee for suggesting many improvements of the first draft of this paper.

Appendix

Proof of Proposition 2.10. We will prove that $\mathcal{S}$ is complete.

Suppose that $\{\eta_k\}_{k=1}^{\infty}$ is a Cauchy sequence in $\mathcal{S}$; that is, for every $\varepsilon > 0$, there is some $N \geqslant 1$ such that

$$e^{-\lambda|t|}\left(|\eta_k(t,\xi,\tau) - \eta_\ell(t,\xi,\tau)| + \left|\int_0^t |\dot{\eta}_k(\xi,\tau) - \dot{\eta}_\ell(\xi,\tau)|ds\right|\right) < \varepsilon,$$

whenever $(t,\xi,\tau) \in \mathbb{R} \times V \times [0,b]$ and $k,\ell \geqslant N$. The space $\mathcal{A}$ of continuous functions $\eta : \mathbb{R} \times V \times [0,b] \to \mathbb{R}^n$ that are bounded with respect to the norm

$$||\eta||_{\mathcal{S}} := \sup_{(t,\xi,\tau) \in \mathbb{R} \times V \times [0,b]} e^{-\lambda|t|}|\eta(t,\xi,\tau)|$$

is a Banach space. Thus, the sequence $\{\eta_k\}_{k=1}^{\infty}$ has a limit in $\mathcal{A}$.

Claim A.1. *For each integer $p \geqslant 1$, there is a continuous function $g_p : V \times [0,b] \to L^1([-p,p],\mathbb{R}^n)$ such that for each $\varepsilon > 0$, there is an integer $N \geqslant 1$ and*

$$e^{-\lambda|t|}\left|\int_0^t g_p(\xi,\tau) - \dot{\eta}_k(\xi,\tau)\, ds\right| < \varepsilon$$

whenever $(t,\xi,\tau) \in [-p,p] \times V \times [0,b]$ and $k \geqslant N$. Moreover, there is a number $r > 0$ such that

$$e^{-\lambda|t|}\left|\int_0^t |g_p(\xi,\tau)|\, ds\right| \leqslant r.$$

To prove the claim, note first that the Cauchy sequence $\{\eta_k\}_{k=1}^{\infty} \subset \mathcal{S}$ is bounded. Thus, there is some $r > 0$ such that

$$e^{-\lambda|t|}\left|\int_0^t |\dot{\eta}_k(\xi,\tau)|\, ds\right| \leqslant r, \tag{A.1}$$

whenever $(t,\xi,\tau) \in [-p,p] \times V \times [0,b]$ and $k \geqslant 1$; and therefore, for each $k \geqslant 1$, the continuous function $\dot{\eta}_k : V \times [0,b] \to L^1((-p,p),\mathbb{R}^n)$ is bounded. Moreover, the sequence $\{\dot{\eta}_k\}_{k=1}^{\infty}$ is a Cauchy sequence in the Banach space $\mathscr{E}$ of bounded continuous functions from $V \times [0,b]$ to $L^1((-p,p),\mathbb{R}^n)$. This fact is an immediate consequence of the inequality

$$\begin{aligned}\sup_{\xi,\tau} ||\dot{\eta}_k(\xi,\tau) - \dot{\eta}_\ell(\xi,\tau)|| &= \sup_{\xi,\tau} \int_{-p}^{p} |\dot{\eta}_k(\xi,\tau) - \dot{\eta}_\ell(\xi,\tau)|\, ds \\ &= \sup_{\xi,\tau} e^{\lambda|p|}e^{-\lambda|p|} \int_{-p}^{p} |\dot{\eta}_k(\xi,\tau) - \dot{\eta}_\ell(\xi,\tau)|\, ds \\ &\leqslant 2e^{\lambda|p|}||\eta_k - \eta_\ell||_{\mathscr{S}}.\end{aligned}$$

Because $\mathscr{E}$ is complete, the sequence $\{\dot{\eta}_k\}_{k=1}^{\infty}$ converges to some $g_p \in \mathscr{E}$. Hence, for every $\varepsilon > 0$, there is an $N \geqslant 1$ such that

$$e^{-\lambda|t|}\left|\int_0^t |g_p(\xi,\tau) - \dot{\eta}_k(\xi,\tau)|\, ds\right| < \varepsilon, \tag{A.2}$$

whenever $(t,\xi,\tau) \in [-p,p] \times V \times [0,b]$ and $k \geqslant N$. Using inequalities (A.1) and (A.2) and a triangle law estimate, it follows that

$$e^{-\lambda|t|}\left|\int_0^t |g_p(\xi,\tau)|\, ds\right| \leqslant r.$$

This completes the proof of the claim.

For each $(\xi,\tau) \in V \times [0,b]$ and $p \geqslant 1$, choose a function $g_p(\xi,\tau)$ in the equivalence class of $g_p(\xi,\tau) \in L^1((-p,p),\mathbb{R}^n)$. (We are using the same name for two different objects.) For each $t \in \mathbb{R}$, define $g(\xi,\tau)(t) = g_p(\xi,\tau)(t)$, where p is the smallest integer such that $t \in (-p,p)$. The function $g(\xi,\tau) : \mathbb{R} \to \mathbb{R}^n$ is measurable. In fact, for an open set U in $\mathbb{R}^n$, the set $U_p := g(\xi,\tau)^{-1}(U) \cap (-p,p)$ is measurable because

$$g(\xi,\tau)^{-1}(U) \cap (-p,p) = g_p(\xi,\tau)^{-1}(U) \cap (-p,p),$$

and therefore $g(\xi,\tau)^{-1}(U)$ is the countable union of measurable sets. We will show that $g(\xi,\tau) \in L^1_{\mathrm{loc}}(\mathbb{R},\mathbb{R}^n)$. For this, choose a compact set $J \subset \mathbb{R}$. There is some $p \geqslant 1$ such that $J \subset [-p,p]$. Hence, we have that

$$\begin{aligned}\int_J |g(\xi,\tau)|\, ds &\leqslant \int_{-p}^{p} |g(\xi,\tau)|\, ds \\ &= \sum_{\ell=1}^{p} \int_{-\ell}^{\ell} |g_\ell(\xi,\tau)|\, ds \\ &\leqslant \sum_{\ell=1}^{p} ||g_\ell(\xi,\tau)||_1 < \infty.\end{aligned}$$

To show that $g : V \times [0, b] \to L^1_{\mathrm{loc}}(\mathbb{R}, \mathbb{R}^n)$ is continuous, we fix $(\zeta, \sigma) \in V \times [0, b]$ and use essentially the same estimate to obtain the inequality

$$\int_J |g(\xi, \tau) - g(\zeta, \sigma)|\, ds \leqslant \sum_{\ell=1}^{p} ||g_\ell(\xi, \tau) - g_\ell(\zeta, \sigma)||_1.$$

Since each g_ℓ is continuous, so is g.

We will show that

$$\lim_{k\to\infty} \sup_{(t,\xi,\tau)\in\mathbb{R}\times V\times[0,b]} e^{-\lambda|t|}\left|\int_0^t |g(\xi,\tau) - \dot{\eta}_k(\xi,\tau)|\, ds\right| = 0.$$

Choose $\varepsilon > 0$. By using a triangle-law estimate, we have the inequality

$$e^{-\lambda|t|}\left|\int_0^t |g(\xi,\tau) - \dot{\eta}_k(\xi,\tau)|\, ds\right| \leqslant e^{-\lambda|t|}\left|\int_0^t |g(\xi,\tau) - \dot{\eta}_\ell(\xi,\tau)|\, ds\right| + ||\eta_\ell - \eta_k||_{\mathscr{S}}.$$

Because $\{\eta_k\}_{k=1}^{\infty}$ is a Cauchy sequence in $\mathscr{S}$, there is some $N \geqslant 1$ such that

$$e^{-\lambda|t|}\left|\int_0^t |g(\xi,\tau) - \dot{\eta}_k(\xi,\tau)|\, ds\right| < e^{-\lambda|t|}\left|\int_0^t |g(\xi,\tau) - \dot{\eta}_\ell(\xi,\tau)|\, ds\right| + \varepsilon,$$

whenever $(t, \xi, \tau) \in \mathbb{R} \times V \times [0, b]$, $\ell \geqslant N$, and $k \geqslant N$. Also, for $t \in (-p, p)$, we have the inequality

$$e^{-\lambda|t|}\left|\int_0^t |g(\xi,\tau) - \dot{\eta}_\ell(\xi,\tau)|\, ds\right| \leqslant e^{-\lambda|t|} \sum_{j=1}^{p} ||g_j(\xi,\tau) - \dot{\eta}_\ell(\xi,\tau)||_1.$$

Hence, for each (t, ξ, τ),

$$\lim_{\ell\to\infty} e^{-\lambda|t|}\left|\int_0^t |g(\xi,\tau) - \dot{\eta}_\ell(\xi,\tau)|\, ds\right| = 0,$$

and therefore

$$e^{-\lambda|t|}\left|\int_0^t |g(\xi,\tau) - \dot{\eta}_k(\xi,\tau)|\, ds\right| < \varepsilon,$$

whenever $(t, \xi, \tau) \in \mathbb{R} \times V \times [0, b]$ and $k \geqslant N$, as required.

To complete the proof, we will show that $\dot{\eta} = g$. The function $\eta_*(\xi, \tau) \in L^1_{\mathrm{loc}}(\mathbb{R}, \mathbb{R}^n)$ defines the distribution (a linear functional on the space of test functions $\mathscr{D}(\mathbb{R}, \mathbb{R}^n)$) given by $\phi \mapsto \int_{-\infty}^{\infty} \eta_*(\xi, \tau)\phi\, ds$, where the product in the integrand is the inner product in $\mathbb{R}^n$. By definition, the distributional derivative of this distribution is the distribution $\phi \mapsto -\int_{-\infty}^{\infty} \eta_*(\xi,\tau)D\phi\, ds$. We must show that

$$-\int_{-\infty}^{\infty} \eta_*(\xi,\tau)D\phi\, ds = \int_{-\infty}^{\infty} g(\xi,\tau)\phi\, ds$$

for each test function ϕ.

Choose $\phi \in \mathscr{D}(\mathbb{R},\mathbb{R}^n)$. Since the support of ϕ is compact, it is contained in some interval $[-p,p]$. Thus, it suffices to show that the quantities

$$-\int_{-p}^{p} \eta_*(\xi,\tau)D\phi\, ds, \qquad \int_{-p}^{p} g(\xi,\tau)\phi\, ds$$

are equal. We will show that each of these quantities is the limit of the same sequence. In fact, since the distributional derivative of $(\eta_k)_*(\xi,\eta)$ is $\dot{\eta}_k(\xi,\eta)$, we have that

$$-\int_{-p}^{p} (\eta_k)_*(\xi,\tau)D\phi\, ds = \int_{-p}^{p} \dot{\eta}_k(\xi,\tau)\phi\, ds, \tag{A.3}$$

and the required sequence is one of the sequences corresponding to this equality.

Because

$$\begin{aligned}\left|\int_{-p}^{p} g_p(\xi,\tau)\phi\, ds - \int_{-p}^{p} \dot{\eta}_k(\xi,\tau)\phi\, ds\right| &\leqslant \int_{-p}^{p} |g_p(\xi,\tau) - \dot{\eta}_k(\xi,\tau)||\phi|\, ds \\ &\leqslant ||\phi|| ||g_p - \dot{\eta}_k||_1,\end{aligned}$$

it follows that

$$\lim_{k\to\infty} \int_{-p}^{p} \dot{\eta}_k(\xi,\tau)\phi\, ds = \int_{-p}^{p} g(\xi,\tau)\phi\, ds.$$

Similarly,

$$\begin{aligned}&\left|\int_{-p}^{p} \eta_*(\xi,\tau)D\phi\, ds - \int_{-p}^{p} (\eta_k)_*(\xi,\tau)D\phi\, ds\right| \\ &\quad\leqslant ||D\phi|| \int_{-p}^{p} |\eta(s,\xi,\tau) - \eta_k(s,\xi,\tau)|\, ds \\ &\quad\leqslant 2||\phi||e^{\lambda p}||\eta - \eta_k||_{\mathscr{S}},\end{aligned}$$

and therefore

$$\lim_{k\to\infty} -\int_{-p}^{p} (\eta_k)_*(\xi,\tau)D\phi\, ds = -\int_{-p}^{p} \eta_*(\xi,\tau)D\phi\, ds. \qquad \square$$

References

[1] P.W. Bates, K. Lu, C. Zeng, Persistence of overflowing manifolds for semiflows, Comm. Pure. Appl. Math. LII (1999) 983–1046.
[2] L. Bel, Spontaneous predictivisation, in: J. Llosa (Ed.), Relativistic Action at a Distance: Classical and Quantum Aspects, Lecture Notes in Physics, Vol. 162, 1982, pp. 21–49.
[3] C. Chicone, Ordinary Differential Equations with Applications, Springer, New York, 1999.
[4] C. Chicone, S. Kopeikin, B. Mashhoon, D.G. Retzloff, Delay equations and radiation damping, Phys. Lett. A 285 (2001) 17–26.
[5] C. Chicone, What are the equations of motion in classical mechanics? Canad. Appl. Math. Quart., to appear.
[6] O. Diekmann, S.A. van Gils, S.M. Verduyn Lunel, H.O. Walther, Delay Equations: Functional-, Complex-, and Nonlinear Analysis, Springer, New York, 1995.
[7] R.D. Driver, On Ryabov's asymptotic characterization of the solutions of quasi-linear differential equations with small delays, SIAM Rev. 10 (3) (1968) 329–341.
[8] R.D. Driver, Linear differential systems with small delays, J. Differential Equations 21 (1976) 148–166.
[9] M.S. Elbialy, Local contractions of Banach spaces and spectral gap conditions, J. Funct. Anal. 182 (1) (2001) 108–150.
[10] L.E. El'sgol'ts, Introduction to the Theory of Differential Equations with Deviating Arguments, R.J. Mclaughlin, Translator, Holden-Day, San Francisco, CA, 1966.
[11] N. Fenichel, Persistence and smoothness of invariant manifolds for flows, Indiana Univ. Math. J. 21 (1971) 193–226.
[12] N. Fenichel, Geometric singular perturbation theory for ordinary differential equation, J. Differential Equations 31 (1979) 53–98.
[13] I.P. Gouldent, D.M. Jackson, Combinatorial Enumeration, Wiley, New York, 1983.
[14] J.K. Hale, L.A.C. Ladeira, Differentiability with respect to delays, J. Differential Equations 92 (1991) 14–26.
[15] J.H. Hale, S.M. Verduyn Lunel, Introduction to Functional Differential Equations, Springer, New York, 1993.
[16] F. Hartung, J. Turi, On differentiability of solutions with respect to parameters in state-dependent delay equations, J. Differential Equations 135 (1997) 192–237.
[17] M. Hirsch, C. Pugh, Stable manifolds and hyperbolic sets, in: (Proc. Sympos. Pure Math., Vol. XIV, Berkeley, CA, 1968), S.S. Chern, S. Smale (Eds.), Global Analysis Amer. Math. Soc., Providence, RI, 1970, pp. 133–168.
[18] J.D. Jackson, Classical Electrodynamics, Wiley, New York, 1999.
[19] J. Jarnik, J. Kurzweil, Ryabov's special solutions of functional differential equations, Boll. Un. Mat. Ital. 11 (4 Suppl. 3) (1975) 198–208.
[20] J. Kurzweil, Small delays don't matter, in: Lecture Notes in Mathematics, Vol. 206, Springer, New York, 1971.
[21] E.H. Lieb, M. Loss, Analysis, 2nd Edition, Amer. Math. Soc, Providence, RI, 2001.
[22] F. Riesz, B. Sz.-Nagy, Functional Analysis, 2nd Edition, L.F. Boron, Translator, Frederick Ungar Publishing Company, New York, 1955.
[23] Yu.A. Ryabov, Application of the method of small parameters to the investigations of automatic control systems with delay, Automat. Remote Control 21 (1960) 507–514.
[24] Yu.A. Ryabov, Application of the method of small parameters for the construction of solutions of differential equations with retarded argument, Soviet Math. Dokl. 1 (1960) 852–855.
[25] Yu.A. Ryabov, Application of the method of small parameters of Lyapunov-Poincaré in the theory of systems with delay, Inzh. Zh. 1 (1961) 3–15.
[26] Yu.A. Ryabov, Certain asymptotic properties of linear systems with small time delay, Soviet Math. Dokl. 4 (1963) 928–930.
[27] Yu.A. Ryabov, Certain asymptotic properties of linear systems with small time lag, Trudy Sem. Teor. Differential. Uravnenii s Otklon. Argumentom Univ. Druzhby Narodov Patrisa Lumumby 1 (1965) 153–164.