JOURNAL OF INTEGRAL EQUATIONS
AND APPLICATIONS
Volume 2, Number 4, Fall 1990

AVERAGING IN INFINITE DIMENSIONS

J.K. HALE AND S.M. VERDUYN LUNEL

ABSTRACT. In this paper we study the method of averaging for functional differential equations and classes of partial differential equations when subjected to high frequency perturbations.

1. Introduction. It is a well-known fact that the motions described by systems modeled by ordinary differential equations (ODE) may change drastically when subjected to high frequency forcing functions. A classical example is the stabilization of a pendulum at the vertical position by rapidly oscillating the support. The amplitude of an undesirable limit cycle of an autonomous ODE may be decreased or the limit cycle even may be eliminated by such appropriately applied forces (see, for example, Andronov, Khaikin and Witt [**2**], Bogoliubov and Mitropolsky [**5**] or Hale [**14**] for these as well as other illustrations of similar phenomena).

The appropriate tool for understanding the effects of high frequency input functions is the method of averaging. It is the purpose of this paper to give an extension of this method to infinite dimensional systems which will include some types of partial differential equations (PDE) as well as functional differential equations (FDE). We give a few illustrations of the method. Further applications to the control of mechanical systems can be found in a forthcoming paper of Lehman and Verduyn Lunel.

To understand why new ideas are needed in our situation, it is worthwhile to present a brief summary of the method of averaging for ODE. Consider the system of ODE in $\mathbf{R}^n$,

$$\dot{x}(t) = f(t/\epsilon, x), \tag{1.1}$$

where ϵ is a small parameter and $f(s,x)$ is periodic in s of period one, the nonlinearity $f(s,x)$ is continuous together with its derivative with respect to x. Our theory below will hold for $f(s,x)$ which are almost periodic in s, but we present the ideas in the periodic case in order to avoid technical difficulties.

The averaged equation corresponding to (1.1) is

$$\dot{y} = f_0(y) \quad \text{with} \quad f_0(y) = \int_0^1 f(\tau, y)\, d\tau. \tag{1.2}$$

To show that there is a relationship between the solutions of (1.1) and (1.2) for ϵ small, we make a transformation of variables which is close to the identity which carries the vector field in (1.1) close to the one in (1.2). More specifically, we let

$$u(s, x) = \int_0^s [f(\tau, x) - f_0(x)]\, d\tau, \tag{1.3}$$

observe that $u(s, x)$ has period one in s and let

$$x = z + \epsilon\, u(t/\epsilon, z). \tag{1.4}$$

If we restrict x to a fixed bounded set B, then there is an $\epsilon_0 = \epsilon_0(B)$ such that this is a well-defined transformation for $0 \le \epsilon \le \epsilon_0$ and takes (1.1) into the equation

$$\dot{z} = f_0(z) + g(t/\epsilon, z, \epsilon), \tag{1.5}$$

where $g(s, z, \epsilon)$ has period one in s and $g(s, z, 0) = 0$. The vectorfield in (1.5) is, therefore, close to the one in (1.2), and one expects the dynamics of the two equations to be approximately the same.

The classical results on averaging compare the solution x of (1.1), where $x(0) = x_0$, with the approximate function $x^* = y^* + \epsilon u(t/\epsilon, y^*)$, where y^* is a solution of the averaged equation (1.2) and $y^*(0)$ the solution of the equation $x^*(0) = y^*(0) + \epsilon u(0, y^*(0))$. It is shown that, if $y^*(t)$ is bounded for $t \ge 0$, then, for any η and L, there is an ϵ_0 such that, for $0 \le \epsilon \le \epsilon_0$, the difference

$$|x(t) - x^*(t)| \le \eta$$

for $0 \le t \le L$. This is an easy consequence of the fact that (1.5) is equivalent to (1.1) through the transformation (1.4) (see Bogoliubov and Mitropolsky [**5**]).

It is also possible to deduce interesting information from the averaged equation on the infinite time interval $[0, \infty)$ provided that more is known about the solutions of (1.2). For example, to each hyperbolic equilibrium point (periodic orbit), there corresponds a hyperbolic periodic solution (invariant torus) for the original equation (1.1) if ϵ is small (see, for example, Bogoliubov and Mitropolsky [**5**] or Hale [**14**]). The examples mentioned in the first part of the introduction are discussed rather easily from this observation.

Let us now consider extensions of this method to abstract evolutionary equations in a Banach space X. More precisely, consider the equation

$$\dot{u} = Au + F(t/\epsilon, u, \epsilon), \tag{1.6}$$

where A is the generator of a $\mathcal{C}_0$-semigroup $T_A(t)$ on X and $F(t, u, \epsilon)$ is continuous in t, u, ϵ, continuously differentiable in u and almost periodic in t uniformly for u in compact subsets of X. Formally, we can consider the averaged equation

$$\dot{v} = Av + F_0(v) \tag{1.7}$$

and attempt the same type of transformation as above to relate the vectorfields of the two equations.

If (1.6) corresponds to a PDE, then A is an unbounded operator and X is infinite dimensional. Therefore, it becomes difficult to justify the transformation.

For the case in which A generates a group such that $e^{At}u_0$ is almost periodic for each $u_0 \in X$ and $F(t, u, \epsilon) = \epsilon G(u)$ is independent of t, previous work on averaging has been done (Mitropolsky [**23**], Dumas and Ellison [**10**], Dumas, Ellison and Sáenz [**11**], Lemlin and Ellison [**22**]). In this case, one can make an almost periodic transformation $u = e^{At}w$ to obtain the equation in "normal" form:

$$\dot{w} = \epsilon\, e^{-At} G(e^{At} w). \tag{1.8}$$

Equation (1.1) has this normal form under the transformation $t \mapsto \epsilon\, t$. Since (1.8) is almost periodic, it is reasonable to consider the equation averaged with respect to t. Applications of this procedure are contained in the references mentioned above.

In our situation, we are interested primarily in the effect of the forcing term. Now the operator $T_A(t)$ may contain dissipation so that it is not almost periodic and the above transformation is not natural.

Our approach is to use the variation of constants formula for (1.6); that is,

$$u(t) = T_A(t-s)u_0 + \int_s^t T_A(t-\tau)F(\tau/\epsilon, u(\tau), \epsilon)\, d\tau. \tag{1.9}$$

If $T^\epsilon(t,s)u_0 = u(t)$ with $u(s) = u_0$ is the solution operator, then we show in Section 2 that there is a nice function $G(t, v, \epsilon)$ such that

$$T^\epsilon(t,s)u_0 = S^\epsilon(t,s)u_0 + \epsilon\, G(t, S^\epsilon(t,s)u_0, \epsilon), \tag{1.10}$$

where $S^\epsilon(t,s)$ is the solution operator of the averaged equation (1.7) up to terms of order ϵ.

Of course, specific computations must be performed on the vector fields. If the initial data u_0 belongs to the domain $\mathcal{D}(A)$ of A in (1.7), then $S^\epsilon(t,s)u_0$ is a strong solution of the averaged equation up to terms of order ϵ and $T^\epsilon(t,s)u_0$ is a solution of (1.6). We show how the estimates between the solutions with initial data in $\mathcal{D}(A)$ yield estimates for any initial data. In Sections 4 and 5, these results are extended to (FDE), again working with the variation of constants formula rather than the differential equation. However, the presence of delays leads to additional complications. Let us explain. We consider the delay equation in $X = C([-r, 0]; \mathbf{R}^n)$:

$$\dot{x}(t) = f(t/\epsilon, x(t), x(t-r)), \tag{1.11}$$

where r is a nonnegative constant independent of ϵ and $f(t, x, y)$ is 1-periodic in t and smooth in x, y. We consider (1.11) as a perturbation of the delay equation $\dot{x}(t) = 0$. If $T_0(t)$ is the semigroup generated by $\dot{x}(t) = 0$, then the infinitesimal generator A_0 is given by $(A_0\varphi)(\theta) = d\varphi(\theta)/d\theta$ with the domain $\mathcal{D}(A_0)$ given by the C^1-functions which satisfy $\dot{\varphi}(0) = 0$. It is therefore tempting to write (1.11) as an abstract evolutionary equation,

$$\frac{dx_t}{dt} = A_0 x_t + F(t/\epsilon, x_t), \tag{1.12}$$

where

$$F(t/\epsilon, \varphi)(\theta) = \begin{cases} 0, & \text{for } \theta < 0 \\ f(t/\epsilon, \varphi), & \text{for } \theta = 0 \end{cases} \tag{1.13}$$

and $x_t(\theta) = x(t+\theta)$ for $-r \le \theta \le 0$. Of course, we must be careful about the space in which (1.1) is to be considered. If the operator $F(t/\epsilon, \cdot)$ can be considered as a bounded perturbation of the operator A_0, then (1.12) will generate a good evolutionary operator $T(t,s)$ which takes the initial data at s to the solution at t. Unfortunately, the operator $F(t/\epsilon, \cdot)$ is not bounded as a map from X into X! Consequently, we must find a larger space Y so that it is bounded from X to Y. A natural choice for Y is $\mathbf{R}^n \times L^\infty([-r, 0]; \mathbf{R}^n)$. Of course, the operator A_0 must be extended to an operator $A_0^{\odot *}$ from Y into Y. It is shown in Clément, Diekmann et al. [**6**] that this can be done, that $A_0^{\odot *}$ generates a weak$*$-continuous semigroup $\{T^{\odot *}(t)\}$ on Y and the variation of constants formula applied to elements $\varphi \in X$ gives solutions of (1.11); that is, the solution $x(t) = x(t, s, \varphi)$ through φ at s satisfies the integral equation

$$x_t = T_0(t-s)\varphi + \int_s^t T_0^{\odot *}(t-s) F(s/\epsilon, x_s)\, ds. \tag{1.14}$$

Thus, even though the map $F(t/\epsilon, \cdot)$ takes X into the larger space Y, the convolution integral is in X and one obtains solutions of (1.11).

In this more general framework, it is possible to show that the solution operator $T^\epsilon(t,s)$ defined by (1.14) is represented in the form (1.10), where $S^\epsilon(t,s)$ is the solution operator of the averaged equation

$$\dot{y}(t) = f_0(y(t), y(t-r)) \quad \text{with} \quad f_0(y,z) = \int_0^1 f(\tau, y, z)\, d\tau \tag{1.15}$$

up to terms of order ϵ.

In Section 5, we show how these results give comparisons of the solutions of the averaged and original equation on an interval $[0, L]$. We also give results on an infinite time interval for special types of hyperbolic sets and show upper semicontinuity in ϵ of global attractors for the averaged and original equation.

Section 6 is devoted to some applications to parabolic PDE and FDE.

We end this introduction with the remark that the averaging method in delay equations has been justified when the equations are considered in normal form:

$$\dot{x}(t) = \epsilon f(t, x(t), x(t-\bar{r})) \tag{1.16}$$

(see Halanay [**12**], Hale [**13**], and Akmerov [**1**]). We remark that (1.11) is equivalent to (1.16) through the transformation $t \mapsto \epsilon t$ if and only if $r = \epsilon \bar{r}$; that is, the delay in (1.11) is small and approaches zero in ϵ. Therefore, results on averaging which can be justified only through the study of (1.16) are of limited application (see Lehman et al. [**4**] where control of a delayed mechanical system is obtained by rapidly oscillating forces). This remark was our primary motivation for extending the method of averaging in the manner mentioned above.

2. The method of averaging. Let $(X, \|\cdot\|)$ be a Banach space and consider the Cauchy problem

$$\dot{w} = Aw + F(t/\epsilon, w), \tag{2.1}$$

where $w(0) = w_0 \in X$, the operator A generates a strongly continuous semigroup $T_A(t)$ on X and the nonlinearity $F : \mathbf{R}_+ \times X \to X$ with $w \mapsto F(t,w)$ is Fréchet differentiable in w, strongly continuous and almost periodic in t uniformly with respect to w in compact subsets of X. Let M and ω be such that $\|T_A(t)\| \le Me^{\omega t}$.

Along with this Cauchy problem, we consider the averaged equation

$$\dot{v} = Av + F_0(v), \tag{2.2}$$

where

$$F_0(v) = \lim_{T\to 0} \frac{1}{T}\int_0^T F(\tau, v)\, d\tau. \tag{2.3}$$

The basic problem in the method of averaging is to determine in what sense the behavior of the solutions of the averaged Cauchy problem approximate the solutions of (2.1).

We shall assume the following smoothing property (H) of $T_A(t)$:

(H) If $h : [s,\infty) \to X$ is norm continuous, then

(i) $\displaystyle\int_s^t T_A(t-\tau)h(\tau)\, d\tau \in \mathcal{D}(A), \quad \text{for } s \le t;$

(ii) $\displaystyle\left\| A\int_s^t T_A(t-\tau)h(\tau)\, d\tau \right\| \le Me^{\omega t} \sup_{s\le\tau\le t} \|h(\tau)\|, \quad \text{for } s \le t.$

In this section, we prove that there exists a transformation of variables that takes a solution of (2.1) into a solution of the averaged problem (2.2) up to terms of order ϵ. Furthermore, this transformation of variables is almost periodic in t and is close to the identity if ϵ is sufficiently small.

This result, together with (H), enables us to prove the classical theorems on averaging for the Cauchy problem (2.1) which we will do in Section 5.

If $h : [s, \infty) \to X$ is continuously differentiable, then (H) holds and

$$\frac{d}{dt}\int_s^t T_A(t-\tau)h(\tau)\,d\tau = A\int_s^t T_A(t-\tau)h(\tau)\,d\tau + h(t). \tag{2.4}$$

Thus, the smoothing property (H) means that (2.4) is satisfied for norm continuous functions h and does allow to perturb A with certain classes of unbounded operators (compare Desch and Schappacher [**9**]). In concrete examples the hypothesis (H) can be replaced by weaker versions if we know more about the perturbation F (e.g., finite rank in the FDE case).

In this section we work with bounded perturbations from X into X; thus, all integrals are strong integrals. Without difficulties, we can handle the larger class of perturbations allowed in the "perturbed dual semigroup" framework (see Section 4). In this case, all integrals are still well defined (now in the weak$*$ topology) and the proofs go over immediately. In order to support this larger class of perturbations, we have made references to [**8**] for general results from the perturbation theory for nonlinear Lipschitz perturbations. In Section 4, when we discuss delay equations, we will explain this generalization.

Assume at first that the operator $U(t,\cdot) : X \to X$, given by

$$U(t,w) = \int_0^t (F(\tau,w) - F_0(w))\,d\tau, \tag{2.5}$$

is almost periodic. Since the operator $F(t,\cdot)$ is almost periodic, too, this means that the operator $U(t,\cdot)$ is uniformly bounded in t. Thus, in this case, $\epsilon\, U(t,w)$ and $\epsilon\, D_w U(t,w)$ approach 0 in norm as $\epsilon \to 0$, uniformly with respect to t and w in compact subsets of X.

In general, the operator $U(t,\cdot)$ is not almost periodic but we can always make an almost periodic approximation of $U(t,\cdot)$. This is the content of a classical result by Bogoliubov concerned with almost periodic approximation to the solution of the equation $\dot{y} - f(t,x) = 0$. The proof can be found in the appendix on almost periodic functions of the book by Hale [**14**]. Define

$$U(t,\cdot) = \int_{-\infty}^{t} e^{-\epsilon(t-\tau)}(F(\tau,\cdot) - F_0(\cdot))\, d\tau, \tag{2.6}$$

that is, $U(t,\cdot)$ satisfies the differential equation

$$\frac{\partial U}{\partial t}(t,\cdot) = -\epsilon\, U(t,\cdot) + F(t,\cdot) - F_0(\cdot).$$

For the approximation given by (2.6), we can prove that $\epsilon\, U(t,w)$ and $\epsilon\, D_w U(t,w)$ approach 0 in norm as $\epsilon \to 0$, uniformly, with respect to t and w in compact subsets of X. However, we are not able to conclude that these functions are of order ϵ as $\epsilon \to 0$. Since this modification only makes the results more difficult to state, we prefer to make the following hypothesis.

Hypothesis 2.1. *The operator function $U(t,\cdot)$, defined by* (2.5)*, as well as $D_w(t,\cdot)$ are almost periodic uniformly with respect to w in compact sets; that is, for any compact set W in X, there is a constant K_W such that*

$$\|U(t/\epsilon, w)\| \le K_W, \quad \|D_w U(t/\epsilon, w)\| \le K_W \tag{2.7}$$

uniformly with respect to t in $\mathbf{R}$ and w in W.

Next, we consider a mild solution w of (2.1); that is, a solution of the integral equation associated with (2.1),

$$w(t) = T_A(t-s)w_0 + \int_s^t T_A(t-\tau)F(\tau/\epsilon, w(\tau))\, d\tau. \tag{2.8}$$

To study the connection between the solutions of (2.1) and the solutions of the averaged equation (2.2), we introduce two function spaces. Define $BC([s,T];X)$ to be the Banach space of bounded continuous functions with values in X provided with the supremum norm

$$\|v\|_0 = \sup_{s \le \tau \le T} \|v(\tau)\|,$$

and define $BC^1([s,T];X)$ to be the Banach space of bounded continuous differentiable functions provided with the norm

$$\|v\|_1 = \|v\|_0 + \|\dot{v}\|_0.$$

LEMMA 2.2. *If* $v \in BC^1([s,T];S)$, *then*

$$\begin{aligned}
&\int_s^t T_A(t-\tau)(F(\tau/\epsilon, v(\tau)) - F_0(v(\tau)))\,d\tau \\
&\qquad = \epsilon\,A \int_s^t T_A(t-\tau)U(\tau/\epsilon, v(\tau))\,d\tau \\
&\qquad\quad - \epsilon \int_s^t T_A(t-\tau)H(\tau/\epsilon, v(\tau))\,d\tau \\
&\qquad\quad + \epsilon\,U(t/\epsilon, v(t)) - \epsilon\,T_A(t-s)U(s, v(s)),
\end{aligned}$$

where

$$H(t,v) = D_v U(t,v)\frac{dv}{dt}. \tag{2.9}$$

PROOF. We compute $\epsilon\,A\int_s^t T_A(t-\tau)U(\tau/\epsilon, v(\tau))\,d\tau$ explicitly. So,

$$\begin{aligned}
&\frac{1}{h}(T_A(h) - I)\int_s^t T_A(t-\tau)U(\tau/\epsilon, v(\tau))\,d\tau \\
&\qquad = \frac{1}{h}\int_{s-h}^s T_A(t-\tau)U((\tau+h)/\epsilon, v(\tau+h))\,d\tau \\
&\qquad\quad - \frac{1}{h}\int_{t-h}^t T_A(t-\tau)U((\tau+h)/\epsilon, v(\tau+h))\,d\tau \\
&\qquad\quad + \frac{1}{h}\int_s^t T_A(t-\tau)[U((\tau+h)/\epsilon, v(\tau+h)) \\
&\qquad\qquad\qquad\qquad - U((\tau/\epsilon, v(\tau+h))]\,d\tau \\
&\qquad\quad + \frac{1}{h}\int_s^t T_A(t-\tau)[U((\tau/\epsilon, v(\tau+h)) - U((\tau/\epsilon, v(\tau))]\,d\tau.
\end{aligned}$$

By definition, taking the limit $h \downarrow 0$ yields

$$\begin{aligned}\epsilon A \int_s^t & T_A(t-\tau)U(\tau/\epsilon, v(\tau))\,d\tau \\ &= \epsilon\, T_A(t-s)U(s/\epsilon, v(s)) - \epsilon\, U(t/\epsilon, v(t)) \\ &\quad + \int_s^t T_A(t-\tau)(F(\tau/\epsilon, v(\tau)) - F_0(v(\tau)))\,d\tau \\ &\quad + \epsilon \int_s^t T_A(t-\tau)D_v U(\tau/\epsilon, v(\tau))\dot{v}(\tau)\,d\tau,\end{aligned}$$

where we used that $U(t,\cdot)$ satisfies the differential equation

$$\frac{\partial U}{\partial t}(t,\cdot) = +F(t,\cdot) - F_0(\cdot). \ \square$$

We define the following transformation of variables $w(t) = \mathcal{F}v(t)$ with $\mathcal{F} : BC([s,T];X) \to BC([s,T];X)$ and

$$\text{(2.10)}\quad \mathcal{F}v(t) = v(t) - \epsilon\, A \int_s^t T_A(t-\tau)U(\tau/\epsilon, v(\tau))\,d\tau + \epsilon\, U(t/\epsilon, v(t)).$$

LEMMA 2.3. *The transformation $w = \mathcal{F}v$ given by* (2.10) *is well defined, almost periodic and close to the identity; that is, there is a constant $K = K(v)$ and an $\epsilon_0 > 0$ such that, for $0 \le \epsilon \le \epsilon_0$, the difference*

$$\sup_{s \le t \le T} \|w(t) - v(t)\| < \epsilon K.$$

PROOF. From (H), we find

$$\left\| \epsilon\, A \int_s^t T_A(t-\tau)U(\tau/\epsilon, v(\tau))\,d\tau \right\| \le \epsilon\, M e^{\omega t} \sup_{s \le \tau \le t} \|U(\tau/\epsilon, v(\tau))\|.$$

So, from (2.4) and the conditions on the nonlinearity F, we derive that $\mathcal{F}$ is well defined. The function v remains in a compact subset of X, thus Hypothesis 2.1 yields that $\mathcal{F}$ is almost the identity. Finally, the existence of $v \in BC([s,T];X)$ such that $w = \mathcal{F}v$ now follows from an easy contraction argument. $\square$

Next we derive the integral equation for v when w satisfies (2.8). If we substitute (2.10) into (2.8) and use Lemma 2.2 to rewrite the expression, we obtain

$$\begin{aligned} v(t) = T_A(t-s)v_0 &+ \int_s^t T_A(t-\tau)F_0(v(\tau))\,d\tau \\ &+ \int_s^t T_A(t-\tau)N(\tau/\epsilon, v(\tau), \epsilon)\,d\tau, \end{aligned} \tag{2.11}$$

where $w_0 = v_0 + \epsilon\, U(s/\epsilon, v_0)$ and

$$N(t, v, \epsilon) = -\,\epsilon\, H(t, v) + F(t, \mathcal{F}v) - F(t, v). \tag{2.12}$$

To prove that v is a solution of the averaged equation up to terms of order ϵ, it remains to analyze the nonlinearity $N(t, v, \epsilon)$.

LEMMA 2.4. *For $v \in BC^1([s,T]; X)$, there is a constant $C > 0$ such that*

$$\left\| \int_s^t T_A(t-\tau)N(\tau/\epsilon, v(\tau), \epsilon)\,d\tau \right\| \le \epsilon\,\frac{M}{\omega}(e^{\omega t} - 1)C\|v\|_1. \tag{2.13}$$

PROOF. For $v \in BC^1([s,T]; X)$ the function $N(\tau/\epsilon, v(\tau), \epsilon)$ is norm continuous in τ. So, we can estimate

$$\left\| \int_s^t T_A(t-\tau)N(\tau/\epsilon, v(\tau), \epsilon)\,d\tau \right\| \le \frac{M}{\omega}(e^{\omega t} - 1) \sup_{s \le \tau \le t} \|N(\tau/\epsilon, v(\tau), \epsilon)\|. \tag{2.14}$$

The nonlinearity F is continuously differentiable and almost periodic, and $v(\tau)$ remains in a compact subset of X for all $s \le \tau \le t$. Therefore, the derivatives D_vU and D_vF are uniformly bounded, say by K. This shows

$$\sup_{s \le \tau \le t} \|N(\tau/\epsilon, v(\tau), \epsilon)\| \le \epsilon\, K\|v\|_1 + K\|\mathcal{F}v - v\|.$$

and (2.13) follows from Lemma 2.3. ◻

Thus, if $w = \mathcal{F}v$, then v is a solution of the averaged equation up to terms of order ϵ. As a first application of Lemmas 2.3 and 2.4, we compare the solution w of (2.1), with $w(s) = w_0 \in X$, with the approximate solution $w^* = \mathcal{F}v^*$ where v^* is a solution of the averaged equation with $v^*(s) = v_0$ and $w_0 = v_0 + \epsilon\, U(s/\epsilon, v_0)$.

THEOREM 2.5. *If, for $w_0 \in X$, the solution v^* of the averaged equation* (2.2) *with $v(s) = v_0$ and $w_0 = v_0 + \epsilon\, U(s/\epsilon, v_0)$ is uniformly bounded for $t \geq s$, then, for any η and L, there is an ϵ_0 such that, for $0 < \epsilon < \epsilon_0$, the difference*

$$\|w(t) - w^*(t)\| \leq \eta \tag{2.15}$$

for $s \leq t \leq L$.

PROOF. The transformation $\mathcal{F}$ is close to the identity for ϵ sufficiently small. So,

$$\|w - w^*\|_0 \leq \|v - v^*\|_0 + K_1\, \epsilon,$$

where $K_1 = K_1(\|v\|)$. Next we approximate v by $\tilde{v} \in BC^1([s,T]; X)$ and derive from Lemma 2.4,

$$\|\tilde{v} - v^*\| \leq \int_s^t M e^{w(t-\tau)} K_2 \|\tilde{v} - v^*\|\, d\tau + \epsilon\, K_3 \|\tilde{v}\|_1.$$

The Gronwall inequality yields

$$\|\tilde{v} - v^*\| \leq \epsilon\, K_4 \|\tilde{v}\|_1.$$

Thus

$$\|w - w^*\|_0 \leq \epsilon\, K_1 + \epsilon\, K_4 \|\tilde{v}\|_1 + \|v - \tilde{v}\|_0$$

and we can choose $\tilde{v}$ such that $\|v - \tilde{v}\|_0 < \eta 2$ and ϵ so small such that

$$\epsilon\, K_1 + \epsilon\, K_4 \|\tilde{v}\|_1 < \eta/2. \quad \square$$

Since v^* depends on t through t rather than $\epsilon\, t$, we have (2.15) for $s \leq t \leq L$ rather than $s \leq t \leq L/\epsilon$.

3. An invariant manifold theorem. To compare the solution w with the approximate solution w^* for all time we need more structure in the equations. We start with some more precise estimates for the integral equation (2.11) for v. The nonlinearity $F_0 : X \to X$ is a Fréchet differentiable perturbation. Thus, the operator $A + F_0(\cdot)$ generates a Fréchet differentiable nonlinear semigroup $S^0(t)$. If we assume (without loss of generality) that $F_0(0) = 0$, then the derivative of this nonlinear semigroup is just the linear semigroup $S_0(t)$ with generator $A+D_wF_0(0)$ associated with the linearization around zero. Therefore, to study the variation of constants formula (2.11), we also can view $N(t, v, \epsilon)$ as a nonlinear perturbation of the nonlinear semigroup $S^0(t)$ associated with the averaged equation (2.2). This yields the following integral equation for v (see Proposition 2.5 of Clément et al. [**8**]):

$$v(t) = S_0(t-s)v_0 + \int_s^t S_0(t-\tau)R(\tau/\epsilon, v(\tau), \epsilon)\, d\tau, \tag{3.1}$$

where

$$R(t, v, \epsilon) = N(t, v, \epsilon) = F_0(v) - D_vF_0(0)v.$$

LEMMA 3.1. *The mapping $R(t, \epsilon, \cdot) : BC^1([s, T]; X) \to BC([s, T]; X)$ is bounded and Lipschitzian. More precisely, for $\|v_1\|_1, \|v_2\|_1 < \rho$ and $|\epsilon| < \sigma$,*

$$\|R(t, 0, \epsilon)\| \leq M(|\epsilon|),$$
$$\|R(t, v_1, \epsilon) - R(t, v_2, \epsilon)\| \leq \eta(\rho, \sigma)(\|v - w\|_1),$$

where $\eta(0, 0) = 0$, $M(0) = 0$ and the functions η and M are nondecreasing.

PROOF. For $v = 0$, we find $R(t, 0, \epsilon) = \epsilon H(t, 0) = \epsilon\, U(t, 0)$. So, the first part follows form Lemma 2.1. Recall that

$$H(t, v) = D_vU(t, v)\frac{dv}{dt} + U(t, v)$$

and write

$$\begin{aligned} R(t, v_1, \epsilon) - R(t, v_2, \epsilon) = {} & \epsilon\,(H(t, v_1) - H(t, v_2)) \\ & + F(t, \mathcal{F}v_1) - F(t, v_1) \\ & + F(t, v_2) - F(t, \mathcal{F}v_2) \\ & + F_0(v_1) - F_0(v_2) \\ & + D_vF_0(0)(v_2 - v_1). \end{aligned}$$

The operators $F(t,\cdot)$ and $U(t,\cdot)$ are Fréchet differentiable and almost periodic, uniformly on compact subsets of X. So, we can estimate, for $v \in BC^1([s,T];X)$,

$$\|R(t,v_1,\epsilon)-R(t,v_2,\epsilon)\| \le \epsilon\, C\|v_1-v_2\|_1+\tilde{\eta}(\rho,\sigma)\sup_{s\le\tau\le T}\|v_1(\tau)-v_2(\tau)\|,$$

where $\tilde{\eta}(0,0)=0$ and $\tilde{\eta}$ is nondecreasing. □

LEMMA 3.2. *The mapping*

$$v \mapsto \int_s^t S_0(t-\tau)R(\tau/\epsilon, v(\tau),\epsilon)\,d\tau$$

is bounded from $BC^1([s,T];X)$ into $BC^1([s,T;X)$.

PROOF. Since $v \in BC^1([s,T];X)$ we have that $R(\tau, v(\tau),\epsilon)$ is continuous in τ. Thus,

(3.2)
$$\left\|\int_s^t S_0(t-\tau)R(\tau/\epsilon, v(\tau),\epsilon)\,d\tau\right\| \le \frac{M}{\omega}(e^{\omega t}-1)\sup_{s\le\tau\le t}\|R(\tau/\epsilon,v(\tau),\epsilon)\|,$$

where $\|S_0(t)\| \le Me^{\omega t}$. Furthermore, $D_vF_0(0)$ is a bounded perturbation of A. So, the semigroup $S_0(t)$ satisfies (H) as well and

(3.3)
$$\left\|A\int_s^t S_0(t-\tau)R(\tau/\epsilon, v(\tau),\epsilon)\,d\tau\right\| \le Me^{\omega t}\sup_{s\le\tau\le t}\|R(\tau/\epsilon,v(\tau),\epsilon)\|.$$

If we realize that

$$\frac{d}{dt}\int_s^t S_0(t-\tau)R(\tau/\epsilon, v(\tau),\epsilon)\,d\tau = A\int_s^t S_0(t-\tau)R(\tau/\epsilon, v(\tau),\epsilon)\,d\tau + R(t/\epsilon, v(t),\epsilon),$$

the proof follows from Lemma 3.1. □

Introduce the space $X_A = (\mathcal{D}(A), \|\cdot\|_A)$, where $\|\cdot\|_A$ denotes the graph norm and is defined by

$$\|w\|_A := \|w\| + \|Aw\|.$$

We will identify X_A with its embedding into X. The closed graph theorem shows that the linear operator $\Omega_\lambda(A) = \lambda I - A$ is a bounded invertible operator from X_A into X. We therefore conclude that X_A is a Banach space.

Define the solution operator $T^\epsilon(t,s) : X \to X$ for (2.8) by $T^\epsilon(t,s)w_0 := w(t)$ with $w(s) = w_0$ and w given by (2.8). The transformation $\mathcal{F}$ maps the solution operator $T^\epsilon(t,s)$ for (2.8) into another solution operator $S^\epsilon(t,s) : X \to X$, where $S^\epsilon(t,s)v_0 := v(t)$ with $v(s) = v_0$ and $w_0 = v_0 + \epsilon U(s/\epsilon, v_0)$. For the solution operators, Lemma 3.2 states that there is a well-defined variation of constants formula for $S^\epsilon(t,s)$ on the Banach space X_A given by

$$S^\epsilon(t,s)v_0 = S_0(t-s)v_0 + \int_s^t S_0(t-s)R(\tau/\epsilon, S^\epsilon(\tau,s)v_0,\epsilon)\,d\tau. \tag{3.4}$$

Thus, for $v_0 \in X_A$, we have $S^\epsilon(t,s)v_0 \in X_A$, and $v(t) = S^\epsilon(t,s)v_0$ is a strong solution of the averaged equation (2.2) up to terms of order ϵ. For further references, we state the following theorem.

THEOREM 3.3. *The transformation $\mathcal{F} : X \to X$ given by* (2.10) *maps the solution operator $T^\epsilon(t,s)$ associated with* (2.1) *into a new solution operator $S^\epsilon(t,s) : X \to X$. Furthermore, for $v_0 \in X_A$, the variation of constants formula* (3.4) *holds, the solution $v(t) = S^\epsilon(t,s)v_0$ is in $BC^1([s,T];X)$ and is a strong solution of the averaged equation up to terms of order ϵ.*

Thus, because of the hypothesis (H), we can handle the relatively bounded perturbation $R(t,\cdot,\epsilon)$ of A, and there is a variation of constants formula on X_A of the resulting solution operator. The idea now is to analyze $S^\epsilon(t,s)$ on X_A through the variation of constants formula.

To prove the existence of a solution of (3.1) that remains in a small neighborhood of 0 (the equilibrium of the averaged equation) for all time, we need more structure. More precisely, we need hyperbolicity of the semigroup $S_0(t)$ and the existence of an exponential dichotomy:

(i) There is a decomposition $X = X_+ \oplus X_-$ with continuous projections $P_\pm : X \to X_\pm$ such that

$$P_\pm S_0(t) = S_0(t) P_\pm, \qquad \text{for } t \geq 0;$$

(ii) $P_\pm(X_A) \subset X_A$;

(iii) $S_0(t)$ can be extended to a group on X_+;

(iv) There are positive constants M, α, β so that

$$\|S_0(t)P_+x\|_A \leq Me^{\alpha t}\|x\|_A, \quad t \leq 0,\ x \in X_A,$$

and

$$\|S_0(t)P_-x\|_A \leq Me^{-\beta t}\|x\|_A, \quad t \geq 0,\ x \in X_A.$$

A basic lemma is that the dichotomy interacts nicely with the variation of constants formula.

LEMMA 3.4. (i) *If $v(t, v_0)$ is a bounded solution of* (3.1) *for $t \leq 0$, then $v(t, v_0)$ satisfies the integral equation*

$$\begin{aligned} v(t) = {} & S_0(t-s)P_+v_0 + \int_0^t S_0(t-\tau)P_+R(\tau/\epsilon, v(\tau), \epsilon)\,d\tau \\ & + \int_{-\infty}^t S_0(t-\tau)P_-R(\tau/\epsilon, v(\tau), \epsilon)\,d\tau; \end{aligned}$$

(ii) *If $v(t, v_0)$ is a bounded solution of* (3.1) *for $t \geq 0$, then $v(t, v_0)$ satisfies the integral equation*

$$\begin{aligned} v(t) = {} & S_0(t-s)P_-v_0 + \int_0^t S_0(t-\tau)P_-R(\tau/\epsilon, v(\tau), \epsilon)\,d\tau \\ & - \int_t^\infty S_0(t-\tau)P_+R(\tau/\epsilon, v(\tau), \epsilon)\,d\tau; \end{aligned}$$

(iii) *If* $v(t, v_0)$ *is a bounded solution of* (3.1) *on* $\mathbf{R}$*, then* $v(t, v_0)$ *satisfies the integral equation*

$$
\begin{aligned}
v(t) = (\mathcal{T}v)(t) = &+\int_{\infty}^{t} S_0(t-\tau)P_+R(\tau/\epsilon, v(\tau), \epsilon)\,d\tau \\
&-\int_{-\infty}^{t} S_0(t-\tau)P_-R(\tau/\epsilon, v(\tau), \epsilon)\,d\tau.
\end{aligned}
\tag{3.5}
$$

Let $\mathcal{AP} \subset BC(\mathbf{R}; X)$ be the class of almost periodic functions and $\mathcal{P}_p \subset BC(\mathbf{R}; X)$ the class of periodic functions of period p (in case $t \mapsto F(t, \cdot)$ is periodic of period p).

We can now prove the existence of a continuous bounded, almost periodic (or periodic) solution of (2.1). The proof is similar to the ODE case [**14**, Theorem IV.2.1], but, in this case, we make the transformation through $\mathcal{F}$ and work with respect to the $BC^1(\mathbf{R}; X)$ topology.

THEOREM 3.5. *Suppose* $\mathcal{D}$ *is one of the classes* $BC(\mathbf{R}; X)$*,* $\mathcal{AP}$ *or* $\mathcal{P}_p$*. There are constants* $\rho > 0$*,* $\epsilon_1 > 0$ *and a function* $w^* : \mathbf{R} \times [0, \epsilon_0] \to X$ *with*

(i) $(t, \epsilon) \mapsto w^*(t, \epsilon)$ *is continuous and* $w^*(t, 0) = 0$*;*

(ii) $w^*(\cdot, \epsilon) \in \mathcal{D}$ *and* $\|w^*(\cdot, \epsilon)\|_0 \le \rho$ *for* $0 \le \epsilon \le \epsilon_1$*;*

and such that $t \mapsto w^*(t, \epsilon)$ *is a unique solution of* (3.1) *in* $\mathcal{D}$ *which has norm* $\le \rho$*.*

PROOF. We first make the transformation $w = \mathcal{F}v$ given by (2.10). Set $\mathcal{D}^1$ to be one of the classes

$$\mathcal{P}_p^1 \subset \mathcal{AP}^1 \subset BC^1(\mathbf{R}; X).$$

The proof is a standard application of the uniform contraction principle. Because of (H) the range of the mapping given by (3.5) is contained in $BC^1([s, \mathcal{T}]; X)$. Thus consider the mapping $\mathcal{T} : \mathcal{D}^1 \to \mathcal{D}^1$ and we have to verify that the contraction mapping is well defined. From Lemma 3.2, we conclude that $\mathcal{T} : BC^1(\mathbf{R}; X) \to BC^1(\mathbf{R}; X)$. Furthermore, from the representation (3.5), it follows that the subspace $\mathcal{AP}^1$ or $\mathcal{P}_T^1$

is invariant under $\mathcal{T}$. Thus $\mathcal{T}$ is a well-defined mapping from $\mathcal{D}^1$ into $\mathcal{D}^1$.

Define $\mathcal{D}^1_\rho = \{v \in \mathcal{D}^1 : \|v\|_1 < \rho\}$. To prove the theorem we have to show that $\mathcal{T} : \mathcal{D}^1_\rho \to \mathcal{D}^1_\rho$ is a contraction, but this now follows from the estimates for $R(t, v, \epsilon)$ derived in Lemma 3.2. This shows the existence of the solution $v^*(\cdot, \epsilon) \in \mathcal{D}^1_\rho$ of (3.1) and, hence, the theorem follows. □

To discuss the stability properties of $w^*(t, \epsilon)$ or $v^*(t, \epsilon)$, we linearize around $v^*(t, \epsilon)$ and use the general saddle point property for bounded perturbations on X_A. Similar to the proof above, using Lemma 3.4, one can prove the existence of stable and unstable manifolds in $BC^1(\mathbf{R}_\pm; X)$. Thus, in this space, the solution $v^*(t, \epsilon)$ has the same stability properties as the zero solution of the averaged equation. A similar approximation as in the proof of Theorem 2.5 yields that $w^*(\cdot, \epsilon)$ has the same stability properties as the zero solution of the averaged equation in $BC(\mathbf{R}\pm; X)$.

THEOREM 3.6. *The unique solution* $w^*(\cdot, \epsilon)$ *of* (2.1) *in* $\mathcal{D}$ *with norm less than* ρ *has the same stability properties as the zero solution of the averaged equation* (2.2).

4. Functional differential equations. Let Ω be a neighborhood of 0 in $X = C([-r, 0]; \mathbf{R}^n)$, the supremum normed Banach space of continuous functions from $[-r, 0]$ to $\mathbf{R}^n$. Suppose $f : \mathbf{R} \times \Omega \to \mathbf{R}^n$ is continuous. For $\varphi \in \Omega$, we assume that $f(t, \varphi)$ is almost periodic in t uniformly with respect to φ in compact subsets of Ω and f has a continuous Fréchet derivative $\partial f(t, \varphi)/\partial\varphi$ in φ on $\mathbf{R} \times \Omega$. Let ϵ be a real parameter and $x_t(\theta) = x(t + \theta)$. Along with the system of delay equations

$$\begin{aligned} \dot{x}(t) &= f(t/\epsilon, x_t), \qquad \text{for } t > 0, \\ x_0 &= \varphi, \end{aligned} \tag{4.1}$$

we consider the averaged system

$$\begin{aligned} \dot{x}(t) &= f_0(x_t), \qquad \text{for } t > 0, \\ x_0 &= \varphi, \end{aligned} \tag{4.2}$$

where

$$f_0(\varphi) = \lim_{T\to\infty} \frac{1}{T} \int_0^T f(s,\varphi)\, ds.$$

In this section, we shall apply the result from Section 2 and show that there exists an almost periodic transformation of variables. This transformation is close to the identity, for ϵ small, and takes system (4.1) into the averaged system (4.2) up to terms of order ϵ:

$$\dot{z}(t) = f_0(z_t) + O(\epsilon) \quad \text{as } \epsilon \to 0. \tag{4.3}$$

This result enables us to prove the classical theorems on averaging for delay equations of the type (4.1) which we will do in the next section.

It is tempting to try an almost periodic transformation of variables of the form

$$\begin{cases} x(t) = z(t) + \epsilon u(t/\epsilon, y_t), & \text{for } \theta = 0, \\ x(t+\theta) = z(t+\theta), & \text{for } -r \le \theta < 0, \end{cases}$$

and then study the behavior of the transformed equation through a variation of constants formula. However, a rigorous proof is not that easy. First of all, because this is not a transformation from the state space $\mathcal{C}$ into itself, and, secondly, how does the variation of constants formula appear in this particular case?

To overcome these problems, we shall work within the abstract framework developed by Clément, Diekmann et al. [**6**], and then we use the abstract result from Section 2. Let us start to recall this so-called perturbed dual semigroup approach in the case of nonlinear delay equations.

As the unperturbed equation, we take

$$\dot{x}(t) = 0 \tag{4.4}$$

considered as a delay equation. This gives rise to a $\mathcal{C}_0$-semigroup $\{T_0(t)\}$ on X defined by

$$(T_0(t)\varphi)(\theta) = \begin{cases} \varphi(t+\theta), & \text{for } t+\theta \le 0, \\ \varphi(0), & \text{for } t+\theta > 0. \end{cases}$$

Next we develop the framework for the unperturbed equation with the result in mind that it should be invariant under the class of perturbations we consider shortly.

The dual space X^* is represented by $NBV[0,r]$, the space of functions of bounded variation f, normalized such that $f(0) = 0$, f is left continuous and $f(t) = f(r)$ for $t \geq r$. The duality pairing is given by

$$\langle \varphi, f \rangle = \int_0^r df(\sigma)\varphi(-\sigma), \quad \text{for } \varphi \in X \text{ and } f \in X^*.$$

The dual semigroup $\{T^*(t)\}$ is a translation in the other direction and is a weak$*$ continuous semigroup. The subspace $X^{\odot}$ of X^* will be defined by the largest closed invariant subspace on which $\{T^*(t)\}$ is actually strongly continuous. From the general theory, it follows that $X^{\odot}$ is the norm closure of $\mathcal{D}(A^*)$ in X^* and given by the space of all functions f of the form

$$f(\theta) = \begin{cases} 0, & \text{for } \theta \leq 0, \\ c + \int_0^\theta g(\tau)\,d\tau, & \text{for } \theta > 0, \end{cases}$$

for some row vector $c \in \mathbf{R}^n$ and $g \in L^1(\mathbf{R}_+)$ with $\operatorname{supp}(g) \subset [0,r]$. Thus, $X^{\odot}$ can be represented by $\mathbf{R}^n \times L^1[0,r]$, and $\{T_0^{\odot}(t)\}$, the restriction of $\{T_0^*(t)\}$ to $X^{\odot}$, is given by

$$T_0^{\odot}(t)(c,g) = \left(c + \int_0^\theta g(\tau)\,d\tau, g(t+\cdot) \right),$$

for $(c,g) \in X^{\odot}$. The dual space $X^{\odot *}$ can be represented by $\mathbf{R}^n \times L^\infty[-r,0]$ with pairing

$$\langle (\alpha,\varphi),(c,g) \rangle = c\alpha + \int_0^r g(\tau)\varphi(-\tau)\,d\tau,$$

for $(\alpha,\varphi) \in X^{\odot *}$ and $(c,g) \in X^{\odot}$.

Thus, we arrive at a weak$*$ continuous semigroup $\{T_0^{\odot *}(t)\}$ given by

$$(T_0^{\odot *}(t)\varphi)(\theta) = \begin{cases} \varphi(t+\theta), & \text{for } t+\theta \leq 0, \\ \alpha, & \text{for } t+\theta > 0, \end{cases}$$

and we can repeat the above procedure. The space X is called $\odot$-reflexive with respect to $T_0(t)$ if and only if $X^{\odot\odot} \equiv X$. Clearly, the infinitesimal generator $A_0^{\odot *}$ is given by

$$A_0^{\odot *}(\alpha, \varphi) = (0, \dot{\varphi})$$

with domain the equivalence classes that contain a Lipschitz φ with $\varphi(0) = \alpha$. Therefore, the norm closure of $\mathcal{D}(A_0^{\odot *})$ is isomorphic to X, and X is called sun-reflexive with respect to $T_0(t)$.

From Clément, Diekmann et al. [**6**], we have the following perturbation result. If the perturbation $F(t, \cdot) : X \to X^{\odot *}$ is defined by

$$F(t, \varphi) = (f(t, \varphi), 0), \tag{4.5}$$

then there is a one-to-one correspondence between the solutions of (4.1) and the solutions of the integral equation

$$u(t) = T_0(t-s)u(s) + \int_s^t T_0^{\odot *}(t-\tau)F(\tau/\epsilon, u(\tau))\, d\tau, \quad \text{for } t > s, \tag{4.6}$$

where the integral should be interpreted as a weak$*$ integral. Thus, the solution operator $T^\epsilon(t, s)\varphi := u(t)$ with $u(s) = \varphi$ of (4.6) is an evolutionary system for (4.1), and the variation of constants formula

$$T^\epsilon(t, s)\varphi = T_0(t-s)\varphi + \int_s^t T_0^{\odot *}(t-\tau)F(\tau/\epsilon, T^\epsilon(\tau, s)\varphi)\, d\tau \tag{4.7}$$

holds for $t \geq s$ on X. So, although the perturbation $F(t, \cdot)$ maps out of the space X into the larger space $X^{\odot *}$, the convolution in (4.7) maps the result back into X.

In this special case even more is true (see Example 4.9 of [**7**]). Let $h : [s, \infty) \to \mathbf{R}^n \times L^\infty[-r, 0]$ be norm continuous so that the range of h is contained in $\mathbf{R}^n$, i.e., $h(t) = (\alpha(t), 0)$. From the representations for $T_0^{\odot *}(t)$ derived earlier, we see that

$$\int_s^t T_0^{\odot *}(t-\tau)h(\tau)\, d\tau = (\psi(0), \psi), \tag{4.8}$$

where

$$\psi(\theta) = \begin{cases} \int_s^{t+\theta} \alpha(\tau)\, d\tau, & \text{for } -\min\{r, t-s\} < \theta \leq 0, \\ 0, & \text{for } -r \leq \theta \leq -\min\{r, t-s\}. \end{cases} \tag{4.9}$$

Since ψ is Lipschitz continuous, we have

$$\int_s^t T_0^{\odot *}(t-\tau)h(\tau)\,d\tau \in \mathcal{D}(A_0^{\odot *})$$

and

$$A_0^{\odot *}\int_s^t T_0^{\odot *}(t-\tau)h(\tau)\,d\tau = (0,\dot{\psi}). \tag{4.10}$$

Thus, we have proved

LEMMA 4.1. *Let $h:[s,\infty)\to \mathbf{R}^n\times L^\infty[-r,0]$ be norm continuous. If the range of h is contained in $\mathbf{R}^n$, then*

(i) $\int_s^t T_0^{\odot *}(t-\tau)h(\tau)\,d\tau \in \mathcal{D}(A_0^{\odot *})$, *for $s\le t$.*

(ii) $\|A_0^{\odot *}\int_s^t T_0^{\odot *}(t-\tau)h(\tau)\,d\tau\| \le M\sup_{s\le\tau\le t}\|h(\tau)\|$, *for $s\le t$.*

This is not exactly the hypothesis (H) since we restrict the class of continuous functions h to those with finite dimensional range. But this is all that we need to obtain a variation of constants formula for the perturbation R, given the special nature of the perturbations we consider (i.e., finite rank). Since we work with the weak$*$ topology, not only the integrals but also the derivatives have to be considered with respect to this topology. Thus, the function space $BC^1([s,T];X)$ denotes the bounded weakly$*$ continuously differentiable functions.

The transformation $\mathcal{F}: BC^1([s,T];X)\to BC([s,T];X)$ is well defined [**7**, 2.2] and given by

$$\begin{aligned}\mathcal{F}S^\epsilon(t,s)\varphi = S^\epsilon(t,s)\varphi - \epsilon A_0^{\odot *}\int_s^t T_0^{\odot *}(t-\tau)U(\tau/\epsilon, S^\epsilon(\tau,s)\varphi)\,d\tau\\ + \epsilon\, U(t/\epsilon, S^\epsilon(t,s)\varphi),\end{aligned} \tag{4.11}$$

where $U(t,\varphi)=(u(t,\varphi),0)$ and $u(t,\varphi)=\int_0^t (f(\tau,\varphi)-f_0(\varphi))\,d\tau$. In coordinates, the transformation reads

$$x(t+\theta)=\begin{cases} z(t+\theta)+\epsilon u((t+\theta)/\epsilon, z(t+\theta)), & \text{for } -\min\{r,t-s\}<\theta\le 0,\\ z(t+\theta), & \text{for } -r\le\theta\le-\min\{r,t-s\},\end{cases}$$

where $x(t+\theta) = (T^\epsilon(t,s)\varphi)(\theta)$ and $z(t+\theta) = (S^\epsilon(t,s)\varphi)(\theta)$.

Now we will analyze the operators H and R in the FDE case. Let

$$X_{A_0^{\odot *}} = (\mathcal{D}(A_0^{\odot *}), \|\cdot\|_{A_0^{\odot *}}).$$

If we substitute $T^\epsilon(t,s)\varphi = \mathcal{F}S^\epsilon(t,s)\varphi$ into (4.7), we obtain a variation of constants formula for the perturbation R on $X_{A_0^{\odot *}}$ (see (2.11)):

$$\begin{aligned} S^\epsilon(t,s)\varphi = T_0(t-s)\varphi + \int_s^t T_0^{\odot *}(t-\tau)\frac{\partial F_0}{\partial \varphi}(0) S^\epsilon(t,s)\varphi\, d\tau \\ + \int_s^t T_0^{\odot *}(t-\tau) R(\tau/\epsilon, S^\epsilon(t,s)\varphi, \epsilon)\, d\tau, \end{aligned} \tag{4.12}$$

where

$$R(t,\varphi,\epsilon) = -\,\epsilon\, H(t,\varphi) + F(t,\mathcal{F}\varphi) - F(t,\varphi) + F_0(\varphi) - \frac{\partial F_0}{\partial \varphi}(0)\varphi$$

and

$$H(t,\varphi) = (h(t,\varphi), 0),$$

with

$$h(t,\varphi) = \frac{\partial u}{\partial \varphi}(t,\varphi)(0)\frac{d\varphi}{dt} + u(t,\varphi)$$

and

$$u(t,\varphi) = \int_0^t (f(\tau,\varphi) - f_0(\varphi))\, d\tau.$$

Thus, we conclude from Lemma 4.1 that the variation of constants formula for $S^\epsilon(t,s)$ holds on $X_{A_0^{\odot *}}$, and we can apply the abstract results from Sections 2 and 3 to the FDE (4.1).

5. The classical theorems on averaging for FDE. In this section we present the classical result on averaging for FDE. We first recall the conditions on the nonlinearity from Section 4.

Let Ω be a neighborhood of 0 in $X = C([-r,0];\mathbf{R}^n)$, the supremum normed Banach space of continuous functions from $[-r,0]$ to $\mathbf{R}^n$. Suppose $f : \mathbf{R}\times\Omega \to \mathbf{R}^n$ is continuous. For $\varphi \in \Omega$ we assume that

$f(t,\varphi)$ is almost periodic in t uniformly with respect to φ in compact subsets of Ω and f has a continuous Fréchet derivative $(\partial f(t,\varphi)/\partial\varphi)$ in φ on $\mathbf{R}\times\Omega$. Let ϵ be a real parameter and $x_t(\theta) = x(t+\theta)$. Along with the system of delay equations

$$\begin{aligned} \dot{x}(t) &= f(t/\epsilon, x_t), \quad \text{for } t > 0, \\ x_0 &= \varphi \end{aligned} \tag{5.1}$$

we consider the averaged system

$$\begin{aligned} \dot{y}(t) &= f_0(y_t), \quad \text{for } t > 0, \\ y_0 &= \varphi, \end{aligned} \tag{5.2}$$

where

$$f_0(\varphi) = \lim_{T\to\infty} \frac{1}{T}\int_0^T f(s,\varphi)\, ds.$$

As a first result, we compare the solution $x(t;\varphi) = (T^\epsilon(t,s)\varphi)(0)$ of (5.1) with the approximate solution $x^*(t;\varphi^*) = (\mathcal{F}S^0(t-s)\varphi^*)(0)$, with

$$\varphi = \varphi^* + \epsilon\, U(s/\epsilon, \varphi^*),$$

where $S^0(t)$ is the nonlinear semigroup associated with the averaged equation (5.2).

THEOREM 5.1. *If the solution $y(\cdot;\varphi^*)$ of the averaged equation is bounded, then, for any η and L, there is an ϵ_0 such that, for $0 \le \epsilon \le \epsilon_0$, we have*

$$|x(t) - x^*(t)| \le \eta, \tag{5.3}$$

for $s \le t \le L$.

PROOF. Set $w(t) = T^\epsilon(t,s)\varphi$ with $w(s) = \varphi$, $v^*(t) = S^0(t-s)\varphi^*$ with $v^*(s) = \varphi^*$. The result follows from Theorem 2.5. □

Since the transformation $\mathcal{F}$ is close to the identity, we can also compare the solution $x(\cdot;\varphi)$ with the corresponding solution of the averaged equation directly.

COROLLARY 5.2. *If the solution $y(\cdot;\varphi)$ of the averaged equation is bounded, then, for any η and L, there is an ϵ_0 such that, for $0 \le \epsilon \le \epsilon_0$, the difference*

$$|x(t;\varphi) - y(t;\varphi)| \le \eta,$$

for $s \le t \le L$.

To give estimates for all time, we assume that $y = y_0$ is an equilibrium solution for the averaged equation. Thus, $f_0(y_0) = 0$, and, from the general theory [**15**], it follows that the semigroup $S_0(t)$ associated with the variational equation has a hyperbolic structure if and only if the generator of $S_0(t)$ has no spectrum on the imaginary axis. Therefore, we can apply the results from Sections 3 and 4 and compare the solution x with the approximate solution x^* for all time.

THEOREM 5.3. *If $y = y_0$ is a hyperbolic equilibrium for the averaged equation* (5.2)*, then, for some $\epsilon_0 > 0$ and $0 < \epsilon \le \epsilon_0$, there is a unique almost periodic solution $t \mapsto x^*(t,\epsilon)$ of* (5.1) *continuous in t and ϵ with $x^*(t,0) = y_0$ and $|x^*(t,\epsilon) - y_0| \le \rho$ for $t \in \mathbf{R}$, which has the same stability properties as the equilibrium y_0 of* (5.2)*. If, in addition, the nonlinearity f is periodic $f(s+p,\varphi) = f(s,\varphi)$, then the solution $s \mapsto x^*(t,\epsilon)$ is periodic in t of period p.*

COROLLARY 5.4. *If $y = y_0$ is hyperbolic and uniformly asymptotically stable, then the unique almost periodic solution $t \mapsto x^*(t,\epsilon)$ is hyperbolic and uniformly asymptotically stable, and there is a $\rho > 0$ such that*

(i) *If $x(\cdot;\varphi)$ is a solution of* (5.1) *with $x(s) = \varphi$ and $\|\varphi - y_0\| < \rho$, then*

$$|x(t) - x^*(t,\epsilon)| \le Ce^{-\gamma(t-s)}. \tag{5.4}$$

(ii) *If $y(\cdot;\varphi)$ is the solution of the averaged equation, then*

$$|x(t;\varphi) - y(t;\varphi)| < \eta, \tag{5.5}$$

for $t \ge s$.

The proofs of these theorems are immediate consequences of the previous results. In Section 2, we proved the existence of a transformation

$\mathcal{F}$ that is close to the identity and maps a solution $x(\cdot;\varphi)$ of (5.1) with $\varphi \in X$ into a solution $z(\cdot;\varphi^*)$ of the averaged equation (5.2) up to terms of the order ϵ. In Section 3, we used the transformation $\mathcal{F}$ to prove the existence of an almost periodic solution $x^*(\cdot,\epsilon)$ of (5.1) with the properties listed in Theorem 5.3. The exponential estimate in (5.4) follows from the standard stable manifold theorem, and a combination of this result with the finite time result from Corollary 5.2 yields (5.5).

As an illustration of these results, we can study the effect of rapid oscillations in the linear system

$$\dot{x}(t) = \sum_{j=0}^{m} F_j(t/\epsilon)x(t-r_j), \tag{5.6}$$

where $0 = r_0 < r_1 < \cdots < r_m = r$ and the coefficients F_j are almost periodic $n \times n$-matrices. Along with equation (5.6), consider the autonomous equation

$$\dot{x}(t) = A_0 x(t) + \sum_{j=1}^{m} A_j x(t-r_j), \tag{5.7}$$

where

$$A_j = \lim_{T\to\infty} \frac{1}{T}\int_0^T F_j(s)\,ds$$

and the characteristic equation

$$\det\left[\lambda I - \sum_{j=0}^{m} A_j e^{-\lambda r_j}\right] = 0. \tag{5.8}$$

Thus, if the characteristic equation (5.8) has no roots on the imaginary axis, then the zero solution of (5.6) has the same stability properties as the zero solution of (5.7). In particular, the zero solution of (5.7) is uniformly asymptotically stable if all the roots of the characteristic equation (5.8) have negative real part and unstable if the characteristic equation has a root with positive real part.

Next we will assume that the nonlinearity is periodic in time of period p and study the relation between the dynamics of (5.1) and (5.2) more closely. Let $S^0(t)$ be the semigroup generated by the averaged equation

(5.2), and let $T^\epsilon(t,s)$ be the solution operator for (5.1). Since we are assuming that $f(t,\varphi)$ is periodic in t of period p, the operator $T^\epsilon(t,s)$ satisfies

$$T^\epsilon(t+\epsilon\, p, s+\epsilon\, p) = T^\epsilon(t,s). \tag{5.9}$$

The Poincaré map Π^ϵ for (5.1) is defined by be $\Pi^\epsilon \equiv T^\epsilon(\epsilon\, p, 0)$.

A subset $\mathcal{A}_0 \subset \mathcal{C}$ is said to be a local attractor for (5.2) if $\mathcal{A}_0$ is compact, invariant and there is an open neighborhood U of $\mathcal{A}_0$ such that

$$\operatorname{dist}(S^0(t)\, U, \mathcal{A}_0) \to 0 \quad \text{as } t \to \infty$$

($\mathcal{A}_0$ is attracting under time evolution of the solution operator).

A set $\mathcal{A}_\epsilon \subset \mathcal{C}$ is said to be a local attractor for the Poincaré map Π^ϵ if $\mathcal{A}_\epsilon$ is compact, invariant ($\Pi^\epsilon \mathcal{A}_\epsilon = \mathcal{A}_\epsilon$) and there is an open neighborhood V of $\mathcal{A}_\epsilon$ such that

$$\operatorname{dist}(\Pi_\epsilon^n\, V, \mathcal{A}_\epsilon) \to 0 \quad \text{as } n \to \infty$$

($\mathcal{A}_\epsilon$ is attracting under iteration of the Poincaré map).

We then have

Theorem 5.5. *If the averaged equation* (5.2) *has a local attractor* $\mathcal{A}_0$*, then there is an* $\epsilon_0 > 0$ *such that, for* $0 < \epsilon \le \epsilon_0$*, the Poincaré map of* (5.1) *has a local attractor* $\mathcal{A}_\epsilon$ *and* $\operatorname{dist}(\mathcal{A}_\epsilon, \mathcal{A}_0) \to 0$ *as* $\epsilon \to 0$.

Proof. From the assumption that the averaged equation has a local attractor, we derive that there exists a δ_2 neighborhood

$$B_{\delta_2}(\mathcal{A}_0) = \{\varphi \in X : \operatorname{dist}(\varphi, \mathcal{A}_0) < \delta_2\}$$

with the property that, for any $0 < \delta_1 < \delta_2$, there is a $t_0 > 0$ such that

$$S^0(t) B_{\delta_2}(\mathcal{A}_0) \subset \mathcal{B}_{\delta_1}(\mathcal{A}_0)$$

for $t \ge t_0$. Thus, from Theorem 3.3 and the Gronwall inequality, we find that, for any $\eta > 0$, there is an N_0 and an $\epsilon_0 > 0$ such that

$$T^\epsilon(n\epsilon\, p, s) B_{\delta_1/2}(\mathcal{A}_0) \subset B_\eta(\mathcal{A}_0) \tag{5.10}$$

for $n \geq N_0$ and $0 \leq \epsilon \leq \epsilon_0$. The variation of constants formula for $T^\epsilon(t,s)$ shows that the map $T^\epsilon(\epsilon p, s)$ is an α-contraction, and from (5.10), we conclude that the orbits $T^\epsilon(n\epsilon p, s)B_{\delta_1/2}(\mathcal{A}_0)$ are bounded. Thus, Lemma 2.2.3 of Hale [**16**] implies that the ω-limit set is compact and attracts $B_{\delta_1/2}(\mathcal{A}_0)$. So, the ω-limit set is a local attractor $\mathcal{A}_\epsilon$ for the Poincaré map Π^ϵ in X. Finally, equation (5.10) yields $\mathcal{A}_\epsilon \subset B_\eta(\mathcal{A}_0)$, and since η is arbitrary, this implies the theorem. □

If, in addition, $y = 0$ is a hyperbolic equilibrium for the averaged equation and $\mathcal{A}_0$ equals this equilibrium, then, from the upper semicontinuity of $\mathcal{A}_\epsilon$, we find

Corollary 5.6. *If $\mathcal{A}_0$ is the hyperbolic equilibrium $y = 0$, then $\mathcal{A}_\epsilon = \{\varphi_\epsilon\}$ is a singleton and $\varphi_\epsilon \to 0$ as $\epsilon \to 0$.*

Further investigations are needed to study other applications of these results. For example, what happens if the attractor of the averaged equation consists of a single hyperbolic periodic orbit? From the upper semicontinuity we expect $\mathcal{A}_\epsilon$ to be a torus for ϵ sufficiently small.

6. Applications. In this last section we present some applications to FDE and PDE.

Example 6.1. Consider the equation

$$\dot{x}(t) = (A_1 + f(t/\epsilon))x(t) + (A_2 + f(t/\epsilon))x(t-r), \tag{6.1}$$

where f is almost periodic with average zero. There is a result by Hale [**13**] which states: If $r = O(\epsilon)$ and the matrix $(A_1 + A_2)$ has no eigenvalues on the imaginary axis, then there exists an ϵ_0 such that, for $0 \leq \epsilon \leq \epsilon_0$, the stability properties of (6.1) and the ODE

$$\dot{x} = (A_1 + A_2)x \tag{6.2}$$

are the same. Let us show how this result is related to the results from Section 5. The averaged equation for (6.1) is given by

$$\dot{x}(t) = A_1 x(t) + A_2 x(t-r). \tag{6.3}$$

Thus, the set of eigenvalues is given by the solutions of the transcendental equation

$$\det \Delta(z) = \det \left[zI - A_1 - A_2 e^{-zr}\right] = 0. \tag{6.4}$$

It is an easy estimate to show that the characteristic equation (6.4) has finitely many solutions in any right-half plane $\Re(z) > 0$, we derive that, for ϵ sufficiently small, the equations (6.2) and (6.3) have the same stability properties.

EXAMPLE 6.2. In the theory of vibrational control for nonlinear systems, one studies stabilizability of nonlinear systems using zero mean parametric excitation (see [**3**] and [**4**]). To explain the idea, we consider the nonlinear system of FDE,

$$\dot{x}(t) = f(x(t), x(t-r)), \tag{6.5}$$

where $x_0 = \varphi \in X$ and f is continuously differentiable. Introduce linear multiplicative vibrations into the system

$$\dot{x}(t) = \epsilon^{-1} B(t/\epsilon) x + f(x(t), x(t-r)), \tag{6.6}$$

where $s \mapsto B(s)$ is periodic with average zero such that the fundamental solution $Y(t,0)$ of $\dot{x} = B(t)x$ is almost periodic. Now make the transformation of variables $x(t) = Y(t/\epsilon, -r)y(t)$ for $t \geq -r$ to obtain

$$\dot{y}(t) = Y^{-1}(t/\epsilon, -r) f(Y(t/\epsilon, -r)y(t), Y((t-r)/\epsilon, -r)y(t-r)). \tag{6.7}$$

Thus, we obtain an almost periodic system of FDE that satisfies the conditions from Section 5. Therefore, we can average (6.7) and study the averaged equation. For example, if $x = 0$ is an unstable hyperbolic equilibrium of (6.5), then the goal is to find a parametric excitation B such that the solution $x = 0$ of (6.6) is stable. Thus, from the averaging results it is sufficient to analyze the autonomous equation derived from averaging (6.7).

EXAMPLE 6.3. As an illustration of Theorem 5.5, consider,

$$\dot{x}(t) = -x(t) + b\frac{x(t-r)}{1 + x(t-r)^n}, \tag{6.8}$$

where n is a fixed even integer and $b > 0$ is a parameter.

The solution map is point dissipative; that is, there is a bounded set $B \subset \mathcal{C}$ such that, for any state φ, the solution $x(\cdot;\varphi)$ of (5.10) satisfies $T(t)\varphi \in B$ for $t \geq t_0 = t_0(\varphi)$. Thus, there exists a global attractor $\mathcal{A}$ for (6.8).

This equation was introduced and studied by Mackey as a model to describe different periodic diseases. It is known that, for n even with $n \geq 8$, there exists a b_0 such that, for $b \geq b_0$ (at least numerically, see Hale and Sternberg [**17**], Glass and Mackey [**20**]), there is some chaotic motion on $\mathcal{A}$. Consider the class of rapidly oscillating disturbances of (6.8),

$$\dot{x}(t) = -x(t) + b\frac{\alpha\cos(t/\epsilon) + x(t-r)}{1 + (\alpha\cos(t/\epsilon) + x(t-r))^n}, \tag{6.9}$$

where α measures the energy of the disturbance. We can show the following quenching chaos result [**18**]: For ϵ sufficiently small, the attractor $\mathcal{A}_\epsilon$ of the Poincaré map for (5.11) is just a singleton provided that $\alpha > \max\{2b, 3\}$. So, the result states that high frequency perturbations through the feedback mechanism can eliminate complicated motion on the attractor with relatively low energy. To prove the result we average the equation (6.9) and estimate the nonlinearity in the averaged equation. A Razumikhin-type theorem yields the result for the averaged equation and, hence, Corollary 5.6 completes the proof.

EXAMPLE 6.4. Let A be a sectorial operator and consider the parabolic PDE

$$\dot{x} + Ax = F(t/\epsilon, x) \tag{6.10}$$

on a Banach space X (see Henry [**19**] for the general theory). Suppose F is almost periodic and satisfies the smoothness conditions from Section 2. Since A is sectorial, the semigroup $T_A(t)$ generated by A is analytic and hypothesis (H) is trivially satisfied. Thus, the results from Section 3 hold. In particular, equation (6.10) has an almost periodic solution $x^*(t,\epsilon)$ on $\mathbf{R}$ with $\|x^*(t,\epsilon)\| \leq \rho$, that has the same stability properties as the equilibrium solution $y = 0$ of the averaged equation

$$\dot{y} + Ay = F_0(y) \tag{6.11}$$

of (6.10). Similarly, we can apply the results from Sections 2 and 3 to the interpolation spaces X^α and study relatively bounded perturbations of A, where the results about existence and stability of the almost periodic solution $x^*(t, \epsilon)$ are now in the X^α space and topology (compare Henry [**19**]).

Applications to hyperbolic problems are currently under investigation. Here the condition (H) is not trivially satisfied, but, like in the FDE case, we can analyze the perturbation R from Section 3 and restrict the class of continuous functions h for which (H) must be satisfied.

REFERENCES

1. R.R. Akmerov, *On the principle of averaging for functional differential equations of neutral type*, Ukrain Mat. Zh. **25** (1973), 579–588 (Russian).

2. A.A. Andronov, A.A. Witt, and S.E. Khaikin, *Theory of oscillations*, Pergamon Press, New York, 1966.

3. R.E. Bellman, J. Bentsman, and S.M. Meerkov, *Vibrational control of nonlinear systems*: *Vibrational stabilizability*, IEEE Trans. Automat. Control AC-**31** (1986), 710–716.

4. J. Bentsman, J. Fakhfakh, H. Hvostov, and B. Lehman, *Vibrational stabilization of linear systems with time delay*, in *Analysis and optimization of systems* (A. Bensoussan and J.L. Lions, eds.), Springer Lecture Notes in Control and Inform. Sci. **111** (1988), 72–79.

5. N.N. Bogoliubov and Y.A. Mitropolsky, *Asymptotic methods in the theory of nonlinear oscillations*, Gordon and Breach, New York, 1961.

6. P. Clément, O. Diekmann, M. Gyllenberg, H.J.A.M. Heijmans and H.R. Thieme, *Perturbation theory for dual semigroups*; *The sun-reflexive case*, Math. Ann. **277** (1987), 709–725.

7. ———, *Time-dependent perturbations in the sun-reflexive case*, Proc. Roy. Soc. Edinburgh Sect. A **109** (1989), 145–172.

8. ———, *Nonlinear Lipschitz continuous perturbations in the sun-reflexive case*, in *Volterra integro-differential equations in Banach spaces and applications* (G. DaPrato and M. Iannelli, eds.), Pitman Research Notes in Math. **190** (1989), 67–89.

9. W. Desch and W. Schappacher, *On relatively bounded perturbations of linear* $\mathcal{C}_0$*-semigroups*, Ann. Sci. Norm. Super. Pisa **11** (1984), 327–341.

10. H.S. Dumas and J.A. Ellison, *Particle channeling and the method of averaging, local and global methods of nonlinear dynamics* (R. Cawley, A.W. Sáenz and W.W. Zachary, eds.), Springer Lecture Notes in Physics **252**, 1986, 200–230.

11. H.S. Dumas, J.A. Ellison, and A.W. Sáenz, *Improved* N^{th} *order averaging for periodic systems*, J. Differential Equations, to appear.

12. A. Halanay, *On the method of averaging for differential equations with retarded argument*, J. Math. Anal. Appl. **14** (1966), 70–76.

13. J.K. Hale, *Averaging methods for differential equations with retarded arguments*, J. Differential Equations **2** (1966), 57–73.

14. ———, *Ordinary differential equations*, Kreiger, Malabar, 1980.

15. ———, *Theory of functional differential equations*, Springer-Verlag, New York, 1977.

16. ———, *Asymptotic behavior of dissipative systems*, Amer. Math. Soc., Providence, 1988.

17. J.K. Hale and N. Sternberg, *Onset of chaos in differential delay equations*, J. Comput. Phys. **77** (1988), 221–239.

18. J.K. Hale and S.M. Verduyn Lunel, *The effect of rapid oscillations in the dynamics of delay equations*, to appear in Proc. Mathematical Population Dynamics, Rutgers University, May 17–20, 1989.

19. D. Henry, *Geometric theory of semilinear parabolic equations*, Springer-Verlag, New York, 1981.

20. L. Glass and M.C. Mackey, *Pathological conditions resulting from instabilities in physiological control systems*, Ann. N.Y. Acad. Sci. **316** (1979), 214–235.

21. A.B. Lemlih, *An extension of the method of averaging to partial differential equations*, Ph.D. Dissertation, Univ. of New Mexico, Albuquerque, 1986.

22. A.B. Lemlih and J.A. Ellison, *The method of averaging and the quantum anharmonic oscillator*, Phys. Rev. Lett. **55** (1985), 1950.

23. Y.A. Mitropolsky, *Problèmes de la théorie asymptotique des oscillations non stationaires*, Gauthier-Villars, Paris, 1966.

CENTER FOR DYNAMICAL SYSTEMS AND NONLINEAR STUDIES, GEORGIA INSTITUTE OF TECHNOLOGY, ATLANTA, GA 30332