STABLE MANIFOLDS IN THE METHOD OF AVERAGING

STEPHEN SCHECTER

ABSTRACT. Consider the differential equation \( \dot{z} = \varepsilon f(z, t, \varepsilon) \), where \( f \) is \( T \)-periodic in \( t \) and \( \varepsilon > 0 \) is a small parameter, and the averaged equation \( \dot{z} = \bar{f}(z) := (1/T) \int_0^T f(z, t, 0) \, dt \). Suppose the averaged equation has a hyperbolic equilibrium at \( z = 0 \) with stable manifold \( \bar{W} \). Let \( \beta_\varepsilon(t) \) denote the hyperbolic \( T \)-periodic solution of \( \dot{z} = \varepsilon f(z, t, \varepsilon) \) near \( z \equiv 0 \). We prove a result about smooth convergence of the stable manifold of \( \beta_\varepsilon(t) \) to \( \bar{W} \times \mathbb{R} \) as \( \varepsilon \to 0 \). The proof uses ideas of Vanderbauwhede and van Gils about contractions on a scale of Banach spaces.

1. Introduction. Consider

\( \dot{z} = \varepsilon f(z, t, \varepsilon) \),

where \( f: \mathbb{R}^n \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}^n \) is \( C^1 \) and \( T \)-periodic in \( t \). Consider also the averaged equation

\( \dot{z} = \bar{f}(z) := \frac{1}{T} \int_0^T f(z, t, 0) \, dt \).

Suppose \( \bar{f}(0) = 0 \) and \( D\bar{f}(0) \) is hyperbolic, i.e., has no eigenvalues on the imaginary axis. Then it is well known that for each small \( \varepsilon > 0 \) there is a hyperbolic \( T \)-periodic solution \( \beta_\varepsilon(t) \) of \( 1 \) near \( z \equiv 0 \). Our goal is to study the dependence of the stable manifold of \( \beta_\varepsilon(t) \) on \( \varepsilon \) as \( \varepsilon \) approaches 0.

Let \( E_- \) (resp. \( E_+ \)) denote the stable (resp. unstable) subspace of \( D\bar{f}(0) \). Let \( B_\delta \) denote the ball of radius \( \delta \) about 0 in \( \mathbb{R}^n \). A result of Hale [3, pp. 166–167] states that there are constants \( k > 0, \delta > 0, \) and \( \varepsilon_0 > 0 \) such that the following is true: for each \( \varepsilon \) with \( 0 < \varepsilon < \varepsilon_0 \), there is a Lipschitz continuous function \( (E_- \cap B_\delta) \times \mathbb{R} \to E_+ \) with Lipschitz constant \( k \) whose graph is a local stable manifold of \( \beta_\varepsilon(t) \). (Note that the stable manifold of \( \beta_\varepsilon(t) \) is a subset of \( \mathbb{R}^n \times \mathbb{R} \), where the second variable is time.) This result, which is proved using only a Lipschitz assumption on \( f \), gives a “uniform” description of the stable manifolds of \( \beta_\varepsilon(t) \) for small \( \varepsilon > 0 \), but does not provide information about their limit as \( \varepsilon \) approaches 0.

We now state our result:

**Theorem 1.1.** Let \( r \geq 1 \) and assume \( f \) is \( C^s \) where \( s \geq 3(r + 1) \). Then there exists \( \delta > 0 \) and a \( C^s \) function \( \psi: (E_- \cap B_\delta) \times \mathbb{R} \times [0, \delta) \to E_+ \) such that (1) for \( \varepsilon > 0 \), the graph of \( \psi(\cdot, \cdot, \varepsilon) \) is a local stable manifold of \( \beta_\varepsilon(t) \); (2) there is a local stable manifold \( \bar{W} \subset \mathbb{R}^n \) of 0 for \( z = \bar{f}(z) \) such that the graph of \( \psi(\cdot, \cdot, 0) \) is \( \bar{W} \times \mathbb{R} \).

Received by the editors January 6, 1987. The contents of this paper were presented at the Canadian Mathematical Society Seminar on Oscillation, Bifurcation, and Chaos, Toronto, July 14–25, 1986.

1980 Mathematics Subject Classification (1985 Revision). Primary 34C29.
Of course, an analogous result holds for unstable manifolds. We remark that one knows from the usual stable manifold theorem that the restriction of $\psi$ to $(E_- \cap B_\delta) \times \mathbb{R} \times (0, \delta)$ is $C^s$. We make no claim that this result is the best possible; we hope that the differentiability assumption can eventually be reduced to $s \geq r + 1$.

Let us indicate an application of Theorem 1.1. Suppose $\dot{z} = f(z)$ has an orbit homoclinic to the hyperbolic equilibrium at $z = 0$. One wishes to use Melnikov’s method [3] to determine how this homoclinic orbit breaks as $\varepsilon$ increases from 0. Melnikov’s method involves the computation of an integral that is supposed to represent the derivative with respect to $\varepsilon$ at $\varepsilon = 0$ of the separation between the stable and unstable manifolds of $\beta_\varepsilon(t)$. One way to justify Melnikov’s method requires knowing that this derivative in fact exists; this is a consequence of Theorem 1.1. We remark that the fact that this derivative exists also follows from a version of Hartman’s Theorem for equations like (1) that is due to Murdock and Robinson [6]. A consequence of their theorem is that there is an asymptotic expansion in power of $\varepsilon$ for the position of the stable or unstable manifold of $\beta_\varepsilon(t)$.

2. Outline of proof. We consider equation (1), where $f: \mathbb{R}^n \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}^n$ is at least $C^1$ and is $T$-periodic in $t$. We consider also the averaged equation (2). We assume $\overline{f}(0) = 0$ and $A := Df(0)$ is hyperbolic. We first make a change of coordinates that takes the $T$-periodic solutions of (1) to zero.

For small $\varepsilon \neq 0$, let $\beta(t, \varepsilon) := \beta_\varepsilon(t)$ be the $T$-periodic solution of (1) near $z \equiv 0$, and let $\beta(t, 0) := 0$.

**Proposition 2.1.** In the above situation, assume $f$ is $C^s$, $s \geq 1$. Then there exists $\varepsilon_0 > 0$ such that $\beta(t, \varepsilon)$ is $C^s$ on $\varepsilon \in (-\varepsilon_0, \varepsilon_0)$. If we make the $T$-periodic coordinate change $z = y + \beta(t, \varepsilon)$, then (1) becomes

$$\dot{y} = \varepsilon g(y, t, \varepsilon),$$

where $g$ is $C^s$ on $\mathbb{R}^n \times \mathbb{R} \times (-\varepsilon_1, \varepsilon_1)$ and $T$-periodic in $t$. Moreover, $g(0, t, \varepsilon) \equiv 0$ and $\overline{g}(y) = \overline{f}(y)$, where

$$\overline{g}(y) := \frac{1}{T} \int_0^T g(y, t, 0) \, dt.$$

Thus (3) has the solution $y \equiv 0$ for each $\varepsilon$. Proposition 2.1 is proved in §3.

Next we use averaging to make (3) autonomous to order $2r + 1$:

**Proposition 2.2.** Let $r \geq 1$ and assume $f$ is $C^s$, $s \geq 3(r + 1)$. Let $g$ be as in Proposition 2.1. Then there is a $C^{r+3}$ change of coordinates $y = \varphi(x, t, \varepsilon)$, with $\varphi(x, t, 0) = x$ and $\varphi(0, t, \varepsilon) \equiv 0$, that transforms (3) to

$$\dot{x} = \varepsilon Ax + \sum_{i=0}^{2r} \varepsilon^{i+1} g_i(x) + \varepsilon^{2r+2} h(x, t, \varepsilon).$$

The coordinate change is defined on a set of the form $B_\rho \times \mathbb{R} \times (-\rho, \rho)$ in $x t \varepsilon$-space. Moreover, $g_i$ and $h$ are at least $C^{r+2}$, $h$ is $T$-periodic in $t$, $g_i(0) = 0$, $h(0, t, \varepsilon) \equiv 0$, $D_g(0) = 0$, and $Az + g_0(x) = \overline{g}(x) = \overline{f}(x)$.

The proof of Proposition 2.2 is in §4.

We now scale time by $\tau := \varepsilon t$, which converts (4) to

$$\frac{dx}{d\tau} = Ax + \sum_{i=0}^{2r} \varepsilon^{i+1} g_i(x) + \varepsilon^{2r+1} h \left( x, \frac{\tau}{\varepsilon}, \varepsilon \right).$$

License or copyright restrictions may apply to redistribution; see http://www.ams.org/journal-terms-of-use
When $\varepsilon = 0$, we take
\begin{equation}
(5b) \quad \frac{dx}{d\tau} := Ax + g_0(x) = \bar{g}(x) = \bar{f}(x).
\end{equation}

In the remainder of the proof, we shall assume for simplicity that the $g_i$ are defined and $C^{r+2}$ on all of $\mathbb{R}^n$, that $h$ is defined and $C^{r+2}$ on all of $\mathbb{R}^n \times \mathbb{R} \times \mathbb{R}$, and that $h$ is $T$-periodic in $t$. Also, the term $\varepsilon^{2r+1} h(x, \tau/\varepsilon, \varepsilon)$ will always be taken to be 0 for $\varepsilon = 0$.

Recall that $E_-$ and $E_+$ denote the stable and unstable subspaces of $A$. Let $\pi_-$ and $\pi_+$ denote the corresponding projections.

Equation (5) has the bounded solution $x \equiv 0$ for each $\varepsilon$. In order to find the stable manifolds of these solutions, we shall look for solutions of (5) that are bounded for $\tau \geq 0$. Such a solution satisfies, for some $x_0 \in E_-$, the equation
\begin{equation}
x(\tau) = e^{A\tau}x_0 + \int_0^\tau e^{A(\tau-\sigma)}\pi_- \left( \sum_{i=0}^{2r} \varepsilon^i g_i(x(\sigma)) + \varepsilon^{2r+1} h\left(x(\sigma), \frac{\sigma}{\varepsilon}, \varepsilon\right) \right) d\sigma
\end{equation}
\begin{equation}
\quad + \int_\tau^\infty e^{A(\tau-\sigma)}\pi_+ \left( \sum_{i=0}^{2r} \varepsilon^i g_i(x(\sigma)) + \varepsilon^{2r+1} h\left(x(\sigma), \frac{\sigma}{\varepsilon}, \varepsilon\right) \right) d\sigma.
\end{equation}


We now formulate equation (6) in an abstract manner.

Let $C^0(\mathbb{R}_+, \mathbb{R}^n)$ denote the Banach space of bounded continuous functions from $\mathbb{R}_+ := [0, \infty)$ to $\mathbb{R}^n$ with the sup norm. Define the bounded linear operators $L: E_- \to C^0(\mathbb{R}_+, \mathbb{R}^n)$ and $K: C^0(\mathbb{R}_+, \mathbb{R}^n) \to C^0(\mathbb{R}_+, \mathbb{R}^n)$ by
\begin{align*}
Lx_0(\tau) &:= e^{A\tau}x_0; \\
Kx(\tau) &:= \int_0^\tau e^{A(\tau-\sigma)}\pi_- x(\sigma) d\sigma + \int_\tau^\infty e^{A(\tau-\sigma)}\pi_+ x(\sigma) d\sigma.
\end{align*}
Define the nonlinear operator $N: C^0(\mathbb{R}_+, \mathbb{R}^n) \times \mathbb{R} \to C^0(\mathbb{R}_+, \mathbb{R}^n)$ by
\begin{align*}
N(x, \varepsilon)(\tau) &:= \left\{ \begin{array}{ll}
\sum_{i=0}^{2r} \varepsilon^i g_i(x(\tau)) + \varepsilon^{2r+1} h(x(\tau), \frac{\tau}{\varepsilon}, \varepsilon), & \varepsilon \neq 0, \\
g_0(x(\tau)), & \varepsilon = 0.
\end{array} \right.
\end{align*}
Define $F: C^0(\mathbb{R}_+, \mathbb{R}^n) \times E_- \times \mathbb{R} \to C^0(\mathbb{R}_+, \mathbb{R}^n)$ by
\begin{align*}
F(x, x_0, \varepsilon) &:= Lx_0 + KN(x, \varepsilon).
\end{align*}

Then (6) is equivalent to the fixed point problem $x = F(x, x_0, \varepsilon)$.

If $F$ were $C^r$ and a contraction in $x$ for each $(x_0, \varepsilon)$ near $(0, 0)$, then the fixed point $x(x_0, \varepsilon)$ would be a $C^r$ function of $(x_0, \varepsilon)$ near $(0, 0)$, say for $(x_0, \varepsilon) \in (E_- \cap B_\delta) \times (0, \delta)$. The proof of Theorem 1.1 would then be completed as follows.

By standard arguments it suffices to find a $C^r$ function $\tilde{\psi}: (E_- \cap B_\delta) \times [0, \delta) \to E_+$ such that: (1) for $\varepsilon > 0$, the graph of $\tilde{\psi}(\cdot, \varepsilon)$ is the intersection of a local stable manifold of $\beta_\delta(t)$ with the plane $t = 0$; (2) the graph of $\tilde{\psi}(\cdot, 0)$ is a local stable manifold of 0 for $\dot{z} = \tilde{f}(z)$. Define $\chi: (E_- \cap B_\delta) \times (-\delta, \delta) \to E_+$ by $\chi(x_0, \varepsilon) = \pi_+ x(x_0, \varepsilon)(0)$. If $x(x_0, \varepsilon)$ were $C^r$, then $\chi$ would be $C^r$. For each $\varepsilon$, the graph of $\chi(\cdot, \varepsilon)$ would be the intersection of a local stable manifold of $x \equiv 0$ for (5) with the plane $t = 0$.

The map $\tilde{\psi}$ is then constructed by tracing through the coordinate changes (for a smaller $\delta$). The restriction $\varepsilon \geq 0$ enters because for $\varepsilon < 0$ the stable manifold of $x \equiv 0$ for (5) corresponds to the unstable manifold of $x \equiv 0$ for (4).
Unfortunately, $F$ is not differentiable because $N$ is not differentiable with respect to $\varepsilon$, even when $\varepsilon \neq 0$. The reason is that

$$\frac{\partial}{\partial \varepsilon} h \left(x(\tau), \frac{\tau}{\varepsilon}, \varepsilon \right) = D_2 h \left(x(\tau), \frac{\tau}{\varepsilon}, \varepsilon \right) \cdot -\frac{\tau}{\varepsilon^2} + D_3 h \left(x(\tau), \frac{\tau}{\varepsilon}, \varepsilon \right),$$

which is not bounded in $\tau$ for all $x \in C^0(\mathbb{R}_+, \mathbb{R}^n)$.

To deal with this difficulty, we shall work in a "scale" of Banach spaces. The idea of doing this comes from Vanderbauwhede and van Gils [8], who use this idea to give a nice proof of the $C^r$ center manifold theorem. The same idea was used earlier by Diekmann and van Gils [2] and by van Gils [9]. There is also some resemblance to the Nash-Moser implicit function theorem [5]. For $\gamma > 0$ we define

$$X_\gamma := \{x \in C^0(\mathbb{R}_+, \mathbb{R}^n) : e^\gamma \|x(\tau)\| \text{ is bounded}\}.$$ 

Note that $\|\|$ denotes a norm in $\mathbb{R}^n$. In $X_\gamma$ we use the norm

$$\|x\|_{\gamma} = \sup_{\tau \geq 0} e^{\gamma \tau} \|x(\tau)\|,$$

which makes $X_\gamma$ a Banach space.

There exist constants $M > 0$ and $\alpha > 0$ such that

$$\|\pi_- e^{\alpha \tau} x\| \leq Me^{-\alpha \tau} \|x\| \quad \text{for all } x \in \mathbb{R}^n \text{ and for all } \tau \geq 0;$$

$$\|\pi_+ e^{\alpha \tau} x\| \leq Me^{\alpha \tau} \|x\| \quad \text{for all } x \in \mathbb{R}^n \text{ and for all } \tau \leq 0.$$

**Proposition 2.3.** If $\alpha > \gamma \geq \eta > 0$, then $F$ maps $X_\gamma \times E_- \times \mathbb{R}$ into $X_\eta$.

Thus if $\alpha > \gamma \geq \eta > 0$, we may define $F^\eta_\gamma : X_\gamma \times E_- \times \mathbb{R} \to X_\eta$ by the same formula used to define $F$.

**Proposition 2.4.** If $\gamma \in (0, \alpha)$ and $\varepsilon$ is fixed, then $F^\eta_\gamma (\cdot , \cdot , \varepsilon)$ is $C^{r+2}$. If $\alpha > \gamma > \eta > 0$, then $F^\eta_\gamma$ is $C^r$.

**Proposition 2.5.** Let $\alpha_1 \in (0, \alpha)$ and $\kappa \in (0, 1)$. There exist $\delta > 0$ and $\delta_1 > 0$ such that if $0 < \gamma < \alpha_1$, $\|x_0\| < \delta$, $|\varepsilon| < \delta$, and $\|x\|_\gamma \leq \delta_1$, then $\|F(x, x_0, \varepsilon)\|_\gamma \leq \delta_1$ and $\|D_1 F^\gamma_\gamma (x, x_0, \varepsilon)\| \leq \kappa$.

Thus, by Proposition 2.5, $F^\gamma_\gamma (\cdot , x_0, \varepsilon)$ is a contraction of the ball of radius $\delta_1$ in $X_\gamma$ for $0 < \gamma < \alpha_1$ and $(x_0, \varepsilon) \in B_\delta \times (-\delta, \delta)$. However, $F^\gamma_\gamma$ is still not differentiable with respect to $\varepsilon$: the difficulty is the same as before. On the other hand, by Proposition 2.4, if $\gamma > \eta$ then $F^\eta_\gamma$ is $C^r$; but $F^\eta_\gamma$ maps $X_\eta$ into the larger space $X_\eta$. Nevertheless, with Propositions 2.4 and 2.5 in hand, one may appeal to an abstract theorem, due essentially to Vanderbauwhede and van Gils, to conclude that for each $\gamma \in (0, \alpha_1)$, the fixed point of $F^\gamma_\gamma (\cdot , x_0, \varepsilon)$ is a $C^r$ function of $(x_0, \varepsilon)$. One can then define the map $x$ as above and complete the proof of Theorem 1.1.

We remark that the reason we assume $f$ is $C^{3(r+1)}$ is as follows: To prove that $F^\eta_\gamma$, $\eta > \gamma$, is $C^r$ away from $\varepsilon = 0$, we need that $h$ is $C^{r+2}$; to prove this at $\varepsilon = 0$, we need that the power of $\varepsilon$ multiplying $h$ in (6) is at least $\varepsilon^{2r+1}$.

In §5 and §6 we study the differentiability of the nonlinear map $N$, viewed as a map from $X_\gamma$ to $X_\eta$ with $\gamma \geq \eta$. In §7 we use the results of this study and some standard arguments to prove Propositions 2.3, 2.4 and 2.5. The abstract theorem just mentioned is stated and proved in §8.

The remaining sections of the paper can be read independently, except that §5 should be read before §6.
3. Sending periodic solutions to zero. In this section we prove Proposition 2.1. Only the assertion that $0(t, e)$ is $C^s$ requires proof.

Let $\Phi(z, t, e)$ be the solution of (1) with $\Phi(z, 0, e) = z$. Then $\Phi$ is $C^s$ and $D_2\Phi(z, t, e) = \varepsilon f(\Phi(z, t, e), t, e)$. Let $\Psi(z, t, e) := (1/\varepsilon)[\Phi(z, t, e) - z]$, with $\Psi(z, 0, e)$ defined by continuity. Now $\Psi(z, 0, e) \equiv 0$, and $D_2\Psi(z, t, e) = (1/\varepsilon)D_2\Phi(z, t, e) = f(\Phi(z, t, e), t, e)$. Therefore

$$\Psi(z, t, e) = \int_0^t f(\Phi(z, u, e), u, e)\,du.$$  

Since $f$ and $\Phi$ are $C^s$, $\Psi$ is $C^s$.

For $e \neq 0$, $\Phi(z, T, e) = z$ if and only if $\Psi(z, T, e) = 0$. We compute from (7) that $\Psi(0, T, 0) = 0$ and $D_1\Psi(0, T, 0) = TDf(0)$, which is invertible. By the implicit function theorem we can solve the equation $\Psi(z, T, e) = 0$ near $(z, e) = (0, 0)$. The solution $z = \alpha(e)$ is a $C^s$ function with $\alpha(0) = 0$. Then we define $\beta(t, e) := \Phi(\alpha(e), t, e)$.

4. Averaging. Proposition 2.2 follows from the following more general formulation.

**Theorem 4.1.** Let $y := \varepsilon g(y, t, e)$, where $g: \mathbb{R}^n \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}^n$ is $C^s$ and $T$-periodic in $t$. Let $1 \leq k < s$. Then there is a $T$-periodic change of coordinates

$$y = x + \sum_{i=1}^k \varepsilon^i \varphi_i(x, t),$$

valid on a set of the form $B_\rho \times \mathbb{R} \times (-\rho, \rho)$ in $xte$-space, such that in the new coordinates

$$\dot{x} = \sum_{i=1}^k \varepsilon^i h_i(x) + \varepsilon^{k+1} h(x, t, e).$$

Here $\varphi_i$ and $h_i$ are $C^{s-1+1}$, $h$ is $C^{s-k}$ and $T$-periodic in $t$, and $h_1(x) = \bar{g}(x)$. Moreover, if $g(0, t, e) \equiv 0$, then $h_i(0) = 0$ and $h(0, t, e) \equiv 0$.

Proposition 2.2 follows from Theorem 4.1 by setting $k := 2r + 1$, $Ax + g_0(x) := h_1(x)$, $g_i(x) := h_{i+1}(x)$ for $1 \leq i \leq 2r$.

**Proof of Theorem 4.1.** The proof is based on Perko’s analysis of high order averaging [7].

We begin by computing with equation (8), in which the $\varphi_i$ are to be chosen later. Differentiation with respect to $t$ gives

$$\dot{y} = \dot{x} + \sum_{i=1}^k \varepsilon^i(D_1\varphi_i(x, t)\dot{x} + D_2\varphi_i(x, t)).$$
We set \( y = \varepsilon g(y, t, \varepsilon) \) and solve for \( x \):

\[
\dot{x} = \left( I + \sum_{i=1}^{k} \varepsilon^{i} D_{1} \varphi_{i} \right)^{-1} \left( \varepsilon g \left( x + \sum_{i=1}^{k} \varepsilon^{i} \varphi_{i}, t, \varepsilon \right) - \sum_{i=1}^{k} \varepsilon^{i} D_{2} \varphi_{i} \right)
\]

\[
= \left( I + \sum_{i=1}^{k-1} \varepsilon^{i} A_{i}(x, t) + \varepsilon^{k} A(x, t, \varepsilon) \right)
\]

\[
\left( \sum_{i=1}^{k} \varepsilon^{i} (b_{i}(x, t) - D_{2} \varphi_{i}) + \varepsilon^{k+1} h(x, t, \varepsilon) \right).
\]

For each \( i = 1, \ldots, k - 1 \),

\[
A_{i}(x, t) = \sum_{i_{1} + \cdots + i_{j} = i} a_{i_{1} \ldots i_{j}} (D_{1} \varphi_{i_{1}}) \cdots (D_{1} \varphi_{i_{j}}),
\]

where the coefficients \( a_{i_{1} \ldots i_{j}} \) are not important for our purposes. The \( b_{i} \) are determined by expanding \( \varepsilon g(x + \sum_{i=1}^{k} \varepsilon^{i} \varphi_{i}, t, \varepsilon) \) in powers of \( \varepsilon \). Thus \( b_{1}(x, t) = g(x, t, 0) \).

For \( i = 2, \ldots, k \),

\[
b_{i}(x, t) = \sum_{j=1}^{i-1} \sum_{m=j}^{i-1} \frac{1}{j!(i-1-m)!} D_{1}^{j} D_{3}^{i-1-m} g(x, t, 0) \sum_{i_{1} + \cdots + i_{j} = m} \varphi_{i_{1}} \cdots \varphi_{i_{j}} + D_{3}^{i-1} g(x, t, 0).
\]

Grouping (10) by powers of \( \varepsilon \), we obtain

\[
\dot{x} = \sum_{i=1}^{k} \varepsilon^{i} (c_{i}(x, t) - D_{2} \varphi_{i}) + \varepsilon^{k+1} h(x, t, \varepsilon).
\]

Here

\[
c_{1}(x, t) = b_{1}(x, t) = g(x, t, 0),
\]

and, for \( 2 \leq i \leq k \),

\[
c_{i}(x, t) = \sum_{j=1}^{i-1} A_{i-j}(x, t)(b_{j}(x, t) - D_{2} \varphi_{j}) + b_{i}(x, t).
\]

Thus, for \( 2 \leq i \leq k \), \( c_{i} \) is defined in terms of \( A_{1}, \ldots, A_{i-1}, b_{1}, \ldots, b_{i} \), which are all defined in terms of \( \varphi_{1}, \ldots, \varphi_{i-1} \).

We must now choose \( T \)-periodic \( \varphi_{i} \) so that in (11), the coefficient of \( \varepsilon^{i} \) for \( 1 \leq i \leq k \) is independent of \( t \). For \( i = 1, \ldots, k \), define inductively

\[
h_{i}(x) := \tilde{c}_{i}(x, t) := \frac{1}{T} \int_{0}^{T} c_{i}(x, t) \, dt;
\]

\[
D_{2} \varphi_{i}(x, t) := c_{i}(x, t) - h_{i}(x) \quad \varphi_{i}(x, 0) := 0.
\]

Thus each \( \varphi_{i} \) is \( T \)-periodic in \( t \).

Clearly \( c_{1}(x, t) = g(x, t, 0) \) is \( C^{s} \). The following implications are immediate from the definitions:

\[
c_{i} \text{ is } C^{s-i+1} \Rightarrow h_{i} \text{ is } C^{s-i+1} \Rightarrow D_{2} \varphi_{i} \text{ is } C^{s-i+1}
\]

\[
\Rightarrow \varphi_{i} \text{ is } C^{s-i+1} \Rightarrow A_{i} \text{ is } C^{s-i} \text{ and } b_{i+1} \text{ is } C^{s-i} \Rightarrow c_{i+1} \text{ is } C^{s-i}.
\]
(The last two implications hold if 1 ≤ i ≤ k - 1, the others if 1 ≤ i ≤ k.) We conclude that for 1 ≤ i ≤ k, b_i, c_i and φ_i are C_{s-i+1}; and for 1 ≤ i ≤ k - 1, A_i is C_{s-i}.

We must show that h is C_{s-k}. For this it suffices, by (10) and (11), to show that A and b are C_{s-k}.

We shall say that a function p(x, t, e) is C_{q}, 0 ≤ m ≤ q, if D_1^i D_2^j D_3^l p(x, t, e) exists and is continuous whenever i + j + l ≤ q and i + j ≤ m. Note that if p is C_{q}, then p is C_{m}. We shall say that p is C_{m} if p is C_{q} for every q ≥ m.

Now
\[ φ(x, t, e) := x + \sum_{i=1}^{k} ε^i φ_i(x, t) \]
is C_{s-k+1}, and q is C_{s}. Therefore (φ(x, t, e), t, e) is C_{s-k+1}. (To see this, let \( g(x, t, e) := φ(x, t, e), t, e) \). Note that if 0 ≤ l ≤ k - 1, D_1^i g(x, t, e) is at least C_{s-k+1}. Thus, since each b_i(x, t) and D_2^i φ_i(x, t) is at least C_{s-k+1},

\[ ε^k b = g - \sum_{i=1}^{k} ε^{i-1}(b_i - D_2^i φ_i) \]
is C_{s-k+1}. Therefore b is C_{s-k}. (To see this, use the following fact and induction: If ε^j c(x, t, e) is C_{q}, then ε^{j-1} c(x, t, e) is C_{q-1} if q > m, and is C_{m-1} if q = m. To prove this, let c(x, t, e) := ε^j c(x, t, e). Then

\[ ε^j c(x, t, e) = c(x, t, e) = ε \int_{0}^{1} D_3 c(x, t, ue) du. \]

Therefore
\[ ε^{j-1} c(x, t, e) = \int_{0}^{1} D_3 c(x, t, ue) du, \]

which is clearly C_{m-1} if q > m and C_{m-1} if q = m.)

Moreover, (I + \[ \sum_{i=1}^{k} ε^i D_1 φ_i(x, t) \])^{-1} is C_{s-k}, and I + \[ \sum_{i=1}^{k-1} ε^i A_i(x, t) \] is C_{s-k+1}. Therefore ε^k A is C_{s-k}, so A is C_{s-k}. It follows that h is C_{s-k}.

Finally we assume that g(0, t, e) = 0. Clearly c_1(0, t) = 0. We have the following implications:
\[ c_1(0, t) = 0 \Rightarrow h_1(0) = 0 \Rightarrow D_2 φ_1(0, t) = 0 \Rightarrow b_{i+1}(0, t) = 0 \Rightarrow c_{i+1}(0, t) = 0. \]
The last two implications hold for 1 ≤ i ≤ k - 1, the others for 1 ≤ k. Setting x = 0 in (9) shows that \( \dot{x} = 0 \) if x = 0. Then (11) implies that h(0, t, e) = 0. □

5. Differentiability of N for \( ε \neq 0 \). Our goal in this section is the following result. Let the spaces \( X_γ \) be as in §2.

**Theorem 5.1.** Let \( r ≥ 1 \).

1. Suppose g: \( \mathbb{R}^n \to \mathbb{R}^n \) is C^r and g(0) = 0. Let \( γ ≥ η > 0 \). Then the mapping \( x(τ) \to g(x(τ)) \) is C^r from \( X_γ \) to \( X_η \). Its ith derivative, \( i ≤ i ≤ r, \) at \( x \in X_γ \) is the mapping \( G_{γ}^{(i)}(x) \) defined below.

2. Suppose h: \( \mathbb{R}^n \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}^n \) is C^r, T-periodic in t, and h(0, t, e) = 0. Let \( γ ≥ η > 0 \). Then the mapping \( x(τ) \to h(x(τ), τ/ε, e) \) is C^r from \( X_γ \) to \( X_η \) for each
fixed $\varepsilon \neq 0$. Its $i$th derivative, $1 \leq i \leq r$, at $x \in X_\gamma$, is the mapping $H^{(i,0)}_{\gamma}(x, \varepsilon)$ defined below.

(3) Suppose $h: \mathbb{R}^n \times \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}^n$ is $C^{r+2}$, $T$-periodic in $t$, and $h(0, t, \varepsilon) \equiv 0$. Let $\gamma > \eta > 0$. Then the mapping

$$(x(\tau), \varepsilon) \mapsto h(x(\tau), \tau/\varepsilon, \varepsilon)$$

is $C^r$ from $X_\gamma \times (\mathbb{R} \setminus \{0\})$ to $X_\eta$. Its $(i,j)$th partial derivative, $1 \leq i + j \leq r$, at $(x, \varepsilon)$, is the mapping $H^{(i,j)}_{\eta}(x, \varepsilon)$ defined below.

We shall first give the proof of (3). Then we shall discuss the proofs of (1) and (2) more briefly.

To prove (3) we shall successively prove the following. Let $\gamma > \eta > 0$.

1. The mapping (12) takes $X_\gamma \times (\mathbb{R} \setminus \{0\})$ to $X_\eta$, and the mappings $H^{(i,j)}_{\eta}(x, \varepsilon)$ are bounded $(i + j)$-multilinear mappings from $X^3 \times \mathbb{R}^3$ to $X_\eta$.

2. Each $H^{(i,j)}_{\eta}(x, \varepsilon)$ depends continuously on $(x, \varepsilon)$.

3. $H^{(i,j)}_{\eta}(x, \varepsilon)$ is the $(i,j)$th partial derivative of the mapping (12).

Throughout this section and the next two, we use the symbol $\| \cdot \|$ to denote a norm in $\mathbb{R}^n$ or the operator norm of a linear or multilinear operator between specific Banach spaces. We use $\| \cdot \|_\gamma$ to denote the norm in $X_\gamma$.

**Lemma 5.2.** Suppose $h: \mathbb{R}^n \times \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}^n$ is $C^1$, $T$-periodic in $t$, and $h(0, t, \varepsilon) \equiv 0$. Let $x \in X_\gamma$ and $\varepsilon \neq 0$. If $\gamma > \eta > 0$, then $h(x(\tau), \tau/\varepsilon, \varepsilon) \in X_\eta$.

**Proof.** Since $\gamma > 0$ and $x \in X_\gamma$, $\{x(t): r > 0\}$ lies in $B_\rho$ for some $\rho > 0$. Since $h$ is $T$-periodic in $t$, $\|D_1 h(x(t, \varepsilon))\|$ is bounded on $B_\rho \times \mathbb{R} \times \{\varepsilon\}$. Let $b$ be a bound. Then for $t > 0$,

$$e^{nT} \| h \left( x(\tau), \frac{T}{\varepsilon}, \varepsilon \right) \| = e^{nT} \left\| h \left( x(\tau), \frac{T}{\varepsilon}, \varepsilon \right) - h \left( 0, \frac{T}{\varepsilon}, \varepsilon \right) \right\| \leq e^{nT} b \| x(\tau) \| \leq b \| x \|_\gamma. \quad \square$$

Define $\tilde{h}: \mathbb{R}^n \times \mathbb{R} \times (\mathbb{R} \setminus \{0\}) \rightarrow \mathbb{R}^n$ by

$$\tilde{h}(x, \tau, \varepsilon) := h(x(\tau), \tau/\varepsilon, \varepsilon).$$

**Lemma 5.3.** Suppose $h: \mathbb{R}^n \times \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}^n$ is $C^{r+2}$ and $T$-periodic in $t$. Let $1 \leq i + j \leq r + 2$. Then for $\tau > 0$,

$$\| D_1^i D_3^j \tilde{h}(x, \tau, \varepsilon) \| \leq b_{ij}(x, \tau, \varepsilon),$$

where $b_{ij}: \mathbb{R}^n \times \mathbb{R} \times (\mathbb{R} \setminus \{0\}) \rightarrow \mathbb{R}$ is a polynomial in $\tau$ of degree $j$ whose coefficients are continuous functions of $(x, \varepsilon)$.

**Proof.** $D_1^i D_3^j \tilde{h}(x, \tau, \varepsilon)$ is a linear combination of terms of the form

$$D_1^i D_2^j D_3^k h \left( x, \frac{\tau}{\varepsilon}, \varepsilon \right) \frac{\partial m_1}{\partial x_m} \left( \frac{\tau}{\varepsilon} \right) \cdots \frac{\partial m_k}{\partial x_m} \left( \frac{\tau}{\varepsilon} \right),$$

where $k + l \leq j$ and $m_1 + \cdots + m_k + l = j$. Note that $\| D_1^i D_3^j \tilde{h}(x, \tau, \varepsilon) \|$ is bounded by a continuous function of $(x, \varepsilon)$, and

$$\left| \frac{\partial^m}{\partial \varepsilon^m} \left( \frac{\tau}{\varepsilon} \right) \right| \leq m! \tau \varepsilon^{-(m+1)}. \quad \square$$

**Proposition 5.4.** Suppose $h: \mathbb{R}^n \times \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}^n$ is $C^{r+2}$ and $T$-periodic in $t$, and $\tilde{h}$ is defined as above. Let $x \in X_\gamma$ and $\varepsilon \neq 0$. License or copyright restrictions may apply to redistribution; see http://www.ams.org/journal-terms-of-use
If \( \gamma \geq \eta > 0 \) and \( 1 \leq i \leq r + 2 \), there is a bounded \( i \)-multilinear map from \( X_{\gamma}^i \) to \( X_{\eta} \) defined by

\[
(x_1, \ldots, x_i) \rightarrow D_1^i \tilde{h}(x(\tau), \tau, \epsilon)x_1(\tau) \cdots x_i(\tau).
\]

If \( \gamma > \eta > 0 \), \( 1 \leq i + j \leq r + 2 \), and \( i \geq 1 \), there is a bounded \((i+j)\)-multilinear map from \( X_{\gamma}^i \times \mathbb{R}^j \) to \( X_{\eta} \) defined by

\[
(x_1, \ldots, x_i, \epsilon_1, \ldots, \epsilon_j) \rightarrow D_1^i D_3^j \tilde{h}(x(\tau), \tau, \epsilon)x_1(\tau) \cdots x_i(\tau)\epsilon_1 \cdots \epsilon_j.
\]

Assume in addition that \( h(0, t, \epsilon) \equiv 0 \). If \( \gamma > \eta > 0 \) and \( 1 \leq j \leq r + 1 \), there is a bounded \( j \)-multilinear map from \( \mathbb{R}^j \) to \( X_{\eta} \) defined by

\[
(\epsilon_1, \ldots, \epsilon_j) \rightarrow D_3^j \tilde{h}(x(\tau), \tau, \epsilon)\epsilon_1 \cdots \epsilon_j.
\]

**Proof.** Let \( \{x(\tau) : \tau \geq 0\} \subset B_\rho \).

(1) Let \( b \) be an upper bound for \( b_{i0}(x, \tau, \epsilon) \) on \( B_\rho \times \mathbb{R} \times \{\epsilon\} \). Then for \( \tau \geq 0 \),

\[
e^\eta \tau \||D_1^i \tilde{h}(x(\tau), \tau, \epsilon)x_1(\tau) \cdots x_i(\tau)|| \leq e^\eta b \||x_1(\tau)|| \cdots \||x_i(\tau)||
\]

\[
\leq e^{(\eta - \gamma) \tau} b \||x_1||_\gamma \cdots \||x_i||_\gamma \leq b \||x_1||_\gamma \cdots \||x_i||_\gamma.
\]

(2) Let \( b(\tau) \) be a polynomial that bounds \( b_{ij}(x, \tau, \epsilon) \) on \( B_\rho \times \mathbb{R}^+ \times \{\epsilon\} \). Then for \( \tau \geq 0 \),

\[
e^\eta \tau \||D_1^i D_3^j \tilde{h}(x(\tau), \tau, \epsilon)x_1(\tau) \cdots x_i(\tau)\epsilon_1 \cdots \epsilon_j||
\]

\[
\leq e^{(\eta - \gamma) \tau} \||x_1||_\gamma \cdots \||x_i||_\gamma \||\epsilon_1||_1 \cdots \||\epsilon_j||_1
\]

for some constant \( b \), because \( i \geq 1 \) and \( \gamma > \eta \).

(3) Notice that \( \tilde{h}(0, \tau, \epsilon) \equiv 0 \), so \( D_3^j \tilde{h}(0, \tau, \epsilon) \equiv 0 \) for any \( j \). Let \( b(\tau) \) be a polynomial that bounds \( b_{ij}(x, \tau, \epsilon) \) on \( B_\rho \times \mathbb{R}^+ \times \{\epsilon\} \). Then for \( \tau \geq 0 \),

\[
e^\eta \tau \||D_3^j \tilde{h}(x(\tau), \tau, \epsilon)\epsilon_1 \cdots \epsilon_j||
\]

\[
\leq e^{(\eta - \gamma) \tau} \||u(\tau)|| \leq e^{(\eta - \gamma) \tau} \||x(\tau)|| \leq b \||x||_\gamma
\]

for some constant \( b \) because \( \gamma > \eta \).

**Proposition 5.5.** Assume the hypotheses of Proposition 5.4.

(1) If \( \gamma > \eta > 0 \) and \( 1 \leq i \leq r + 1 \), the \( i \)-multilinear map from \( X_{\gamma}^i \) to \( X_{\eta} \) defined by (13) depends continuously on \((x, \epsilon) \in X_{\gamma} \times (\mathbb{R} \setminus \{0\})\).

(2) If \( \gamma > \eta > 0 \), \( 1 \leq i + j \leq r + 1 \), and \( i \geq 1 \), the \((i+j)\)-multilinear map from \( X_{\gamma}^i \times \mathbb{R}^j \) to \( X_{\eta} \) defined by (14) depends continuously on \((x, \epsilon) \in X_{\gamma} \times (\mathbb{R} \setminus \{0\})\).

(3) Assume in addition that \( h(0, t, \epsilon) \equiv 0 \). If \( \gamma > \eta > 0 \) and \( 1 \leq j \leq r \), the \( j \)-multilinear map from \( \mathbb{R}^j \) to \( X_{\eta} \) defined by (15) depends continuously on \((x, \epsilon) \in X_{\gamma} \times (\mathbb{R} \setminus \{0\})\).

**Proof.** Let \((x, \epsilon) \in X_{\gamma} \times (\mathbb{R} \setminus \{0\})\). Choose \( \rho > 0 \) so that if \( 0 \leq \tau < \infty \) then \( x(\tau) + B_{\epsilon/2} \subset B_\rho \). Let \( b \) be an upper bound for \( b_{i+1,0}(x, \tau, \epsilon) \) on \( B_\rho \times \mathbb{R} \times [\epsilon/2, 3\epsilon/2] \).
Choose a polynomial $c(r)$ so that $b_{i_1}(\tilde{x}, r, \tilde{e}) \leq c(r)$ on $B_\rho \times \mathbb{R}_+ \times [\varepsilon/2, 3\varepsilon/2]$. Let $\lambda \in (0, \varepsilon/2)$. If $\|y - x\|_\gamma < \lambda$ and $|\delta - \varepsilon| < \lambda$, then for $r \geq 0$:

$$
\begin{align*}
& e^{\eta r}\|D^1_1 h(y(r), r, \delta) - D^1_1 h(x(r), r, \varepsilon)\| x_1(\tau) \cdots x_t(\tau) \\
\leq & e^{\eta r} \left\{ \int_0^1 \|D^1_1 h(uy(\tau) + (1 - u)x(\tau), \tau, u\delta + (1 - u)\varepsilon)\| du \|y(\tau) - x(\tau)\| \\
& \quad + \int_0^1 \|D^1_1 D^1_2 h(uy(\tau) + (1 - u)x(\tau), \tau, u\delta + (1 - u)\varepsilon)\| du |\delta - \varepsilon| \right\} \\
\times & e^{-ir\gamma} \|x_1\|_\gamma \cdots \|x_t\|_\gamma \\
\leq & e^{\eta r} \{b(r)e^{-r\gamma} \|y - x\|_\gamma + c(\tau)|\delta - \varepsilon|\} e^{-i\gamma r} \|x_1\|_\gamma \cdots \|x_t\|_\gamma \\
\leq & C\lambda \|x_1\|_\gamma \cdots \|x_t\|_\gamma
\end{align*}
$$

for some constant $C$ because $\gamma > \eta$.

(2) The proof is similar to (1).

(3) Let $b(r)$ be a polynomial upper bound for $b_{i_1}(\tilde{x}, r, \tilde{e})$ on $B_\rho \times \mathbb{R}_+ \times [\varepsilon/2, 3\varepsilon/2]$. Let $c(r)$ be a polynomial upper bound for $b_{i_1, j_1}(\tilde{x}, \tilde{r}, \varepsilon)$ on $B_\rho \times \mathbb{R}_+ \times [\varepsilon/2, 3\varepsilon/2]$. Let $\lambda \in (0, \varepsilon/2)$. If $\|y - x\|_\gamma < \lambda$ and $|\delta - \varepsilon| < \lambda$, then for $r \geq 0$:

$$
\begin{align*}
& e^{\eta r}\|D^1_1 h(y(r), r, \delta) - D^1_1 h(x(r), r, \varepsilon)\| \\
\leq & e^{\eta r} \left\{ \int_0^1 \|D^1_1 h(uy(\tau) + (1 - u)x(\tau), \tau, u\delta + (1 - u)\varepsilon)\| du \|y(\tau) - x(\tau)\| \\
& \quad + \int_0^1 \|D^1_1 D^1_2 h(uy(\tau) + (1 - u)x(\tau), \tau, u\delta + (1 - u)\varepsilon)\| du |\delta - \varepsilon| \right\} \\
& \quad \cdot \|uy(\tau) + (1 - u)x(\tau)\| dv du |\delta - \varepsilon| \\
\leq & e^{\eta r} \{b(r)e^{-r\gamma} + c(\tau)e^{-r\gamma}\} \lambda \leq C\lambda
\end{align*}
$$

for some constant $C$. □

Let $\gamma > \eta > 0$. Suppose $h: \mathbb{R}^n \times \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}^n$ is $C^{r+2}$, $T$-periodic in $t$, and $h(0, t, \varepsilon) \equiv 0$. Then by Lemma 5.2 we may define $H^n_\gamma: X_\gamma \times (\mathbb{R} \setminus \{0\}) \rightarrow X_\eta$ by

$$
H^n_\gamma(x, \varepsilon)(\tau) := h(x(\tau), \tau/\varepsilon, \varepsilon) = \tilde{h}(x(\tau), \tau, \varepsilon).
$$

If $1 \leq i + j \leq r$, $x \in X_\gamma$, and $\varepsilon \neq 0$, we define $H^{n(i, j)}_\gamma(x, \varepsilon)$ to be the bounded $(i + j)$-multilinear map from $X^i_\gamma \times \mathbb{R}^j$ to $X_\eta$ given by

$$
H^{n(i, j)}_\gamma(x, \varepsilon)(x_1, \ldots, x_i, \epsilon_1, \ldots, \epsilon_j)(\tau) := D^i_1 D^j_2 \tilde{h}(x(\tau), \tau, \varepsilon)(x_1(\tau), \ldots, x_i(\tau), \epsilon_1, \ldots, \epsilon_j).
$$

This makes sense by Proposition 5.4.

We shall now complete the proof of Theorem 5.1(3). For uniformity of exposition we define $H^{n(0, 0)}_\gamma := H$.
Consider a pair \((i, j)\) with \(0 \leq i + j \leq r - 1\). Fix \(x \in X_\gamma\) and let \(y \in X_\gamma\) be close to \(x\). We must show
\[
H^\eta_{i,j}(y, \varepsilon) - H^\eta_{i,j}(x, \varepsilon) - H^\eta_{i+1,j}(x, \varepsilon)(y - x) = o(\|y - x\|_\gamma);
\]
\[
H^\eta_{i,j}(x, \delta) - H^\eta_{i,j}(x, \varepsilon) - H^\eta_{i,j+1}(x, \varepsilon)(\delta - \varepsilon) = o(\|\delta - \varepsilon\|).
\]
For then \(D_1H^\eta_{i,j}(x, \varepsilon) = H^\eta_{i+1,j}(x, \varepsilon)\), and \(D_2H^\eta_{i,j}(x, \varepsilon) = H^\eta_{i,j+1}(x, \varepsilon)\).
Since \(H^\eta_{i+1,j}\) and \(H^\eta_{i,j+1}\) are continuous by Proposition 5.5, \(H^\eta_{i,j}\) is continuously differentiable for \(0 \leq i + j \leq r - 1\), and its partial derivatives are as stated.
Thus Theorem 5.1(3) is proved.

Let \(S = \{x_1, \ldots, x_i, \varepsilon_1, \ldots, \varepsilon_j\} \in X_\gamma^i \times \mathbb{R}^j:\|x_1\|_\gamma = \cdots = \|x_i\|_\gamma = \varepsilon_1 = \cdots = \varepsilon_j = 1\). To show (16), we compute
\[
\|H^\eta_{i,j}(y, \varepsilon) - H^\eta_{i,j}(x, \varepsilon) - H^\eta_{i+1,j}(x, \varepsilon)(y - x)\|
\leq \sup_{S} \sup_{\tau \geq 0} e^{\eta \tau} \left\| \int_0^1 [D^2_1H^\eta_{i,j}(u(y(\tau) + (1-u)x(\tau), \tau, \varepsilon) - D^2_1H^\eta_{i+1,j}(u(y(\tau), \tau, \varepsilon)](x(\tau), \tau, \varepsilon) du \right\| \|y - x\|_\gamma.
\]
Since \(H^\eta_{i+1,j}\) depends continuously on \((x, \varepsilon)\), this expression can be made an arbitrarily small multiple of \(\|y - x\|_\gamma\) by taking \(\|y - x\|_\gamma\) sufficiently small.

The proof of (17) is similar.

To prove Theorem 5.1(2), assume the hypotheses and let \(\gamma \geq \eta > 0\). We define \(H^\eta_\gamma\) by the same formula used previously; we may do this for \(\gamma \geq \eta\) by Lemma 5.2.
We also define \(H^\eta_{i,0}(x, \varepsilon), 1 \leq i \leq r\), by the formula used previously; Proposition 5.4(1) shows we may do this for \(\gamma \geq \eta\) and \(1 \leq i \leq r\). To see that \(H^\eta_{i,0}(x, \varepsilon)\) depends continuously on \(x\) for fixed \(\varepsilon\), we appeal to:

**Proposition 5.6.** Suppose \(h: \mathbb{R}^n \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}^n\) is \(C^\tau\) and \(T\)-periodic in \(t\).
If \(\gamma \geq \eta > 0\) and \(1 \leq i \leq r\), then \(H^\eta_{i,0}(x, \varepsilon)\) depends continuously on \(x \in X_\gamma\) for fixed \(\varepsilon\).

**Proof.** Fix \(x \in X_\gamma, \varepsilon \neq 0,\) and \(i\) with \(1 \leq i \leq r\). Choose \(\rho > 0\) so that \(B_\rho\) contains \(\{x(\tau): \tau \geq 0\}\) in its interior. Since \(h\) is \(T\)-periodic in \(t, D^1_1h\) is uniformly continuous on \(B_\rho \times \mathbb{R} \times \{\varepsilon\}\). Choose \(\lambda > 0\). Then there exists \(\delta > 0\) such that if \(\tau \geq 0\) and \(\|z - x(\tau)\| < \delta\) then \(z \in B_\rho\) and \(\|D^1_1h(z, t, \varepsilon) - D^1_1h(x(\tau), t, \varepsilon)\| < \lambda\). Now let \(y \in X_\gamma\) with \(\|y - x\| < \delta\). Then \(\|y(\tau) - x(\tau)\| < \delta\) for all \(\tau \geq 0\). If \(\gamma > \eta > 0\), we have
\[
e^{\|y(\tau), \varepsilon\|} - e^{\|x(\tau), \varepsilon\|} \|x(\tau), \varepsilon\| \leq \lambda \|x\|_\gamma \cdots \|x_i\|_\gamma. \]

The proof of Theorem 5.1(2) is now similar to the proof of Theorem 5.1(3) above.
To prove Theorem 5.1(1), we define \( G^n_\gamma : X^\gamma \to X_\eta, \gamma \geq \eta, \) by \( G^n_\gamma(x)(\tau) := g(x(\tau)) \), and we define \( G^{n(i)}_\gamma(x) : X^i_\gamma \to X_\eta \) by

\[
G^{n(i)}_\gamma(x)(x_1, \ldots, x_i)(\tau) = D^i g(x(\tau))(x_1(\tau), \ldots, x_i(\tau)).
\]

The proof is similar to that of Theorem 5.1(2).

6. Differentiability of \( N \) for \( \varepsilon = 0 \). Our goal in this section is the following result. Assume \( h : \mathbb{R}^n \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}^n \) is \( C^{r+2} \), \( T \)-periodic in \( t \), and \( h(0, t, \varepsilon) \equiv 0 \).

For \( \gamma > \eta > 0 \) define \( \tilde{H}_\eta^n : X^\gamma \times \mathbb{R} \to X_\eta \) by

\[
\tilde{H}_\eta^n(x, \tau) := \begin{cases} 
\varepsilon^{2r+1} h(x(\tau), \tau/\varepsilon, \varepsilon), & \varepsilon \neq 0; \\
0, & \varepsilon = 0.
\end{cases}
\]

**Theorem 6.1.** Under the above assumptions, \( \tilde{H}_\eta^n \) is \( C^r \).

**Proof.** Theorem 5.1(3) implies that \( \tilde{H}_\eta^n \) is \( C^r \) on \( X^\gamma \times (\mathbb{R} \setminus \{0\}) \). Thus we must study \( \tilde{H}_\eta^n \) at points \((x, 0)\).

Recall that \( h(x, \tau, \varepsilon) := h(x, \tau/\varepsilon, \varepsilon) \). For \( 1 \leq i + j \leq r \) and \((x, \varepsilon) \in X^\gamma \times (\mathbb{R} \setminus \{0\})\), define 

\( \tilde{H}^{n(i,j)}_\eta(x, \varepsilon) \) to be the bounded \((i+j)\)-multilinear map from \( X^i_\gamma \times \mathbb{R}^j \) to \( X_\eta \) given by

\[
(x_1(\tau), \ldots, x_i(\tau), \varepsilon_1, \ldots, \varepsilon_j) \to D^i_1 D^j_3 [\varepsilon^{2r+1} h(x(\tau), \tau, \varepsilon)](x_1(\tau), \ldots, x_i(\tau), \varepsilon_1, \ldots, \varepsilon_j).
\]

Define \( \tilde{H}^{n(0,0)}_\eta(x, 0) \) to be the zero map. For uniformity of exposition we define \( \tilde{H}^{n(0,0)}_\eta := \tilde{H}_\eta^n \).

To prove Theorem 6.2, we shall show:

1. \( \tilde{D}_1 \tilde{H}^{n(i,j)}_\eta(x, \varepsilon) \) depends continuously on \((x, \varepsilon)\) for \( 1 \leq i + j \leq r \).
2. \( D_1 \tilde{H}^{n(i,j)}_\eta(x, \varepsilon) = \tilde{H}^{n(i+1,j)}_\eta(x, \varepsilon) \) and \( D_2 \tilde{H}^{n(i,j)}_\eta(x, \varepsilon) = \tilde{H}^{n(i,j+1)}_\eta(x, \varepsilon) \) for \( 0 \leq i + j \leq r - 1 \).

These assertions need only be proved at points \((x, 0)\).

**Lemma 6.2.** Suppose \( h : \mathbb{R}^n \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}^n \) is \( C^{r+2} \) and \( T \)-periodic in \( t \). Let \( 1 \leq i + j \leq r + 2 \). Let \( \rho > 0 \). Then there is a polynomial \( b_{ij}(\tau) \) of degree \( j \) such that for all \( x \in B_\rho \) and for all \( \varepsilon \neq 0 \) sufficiently small, \( \|D^i_1 D^j_3 [\varepsilon^{2r+1} h(x, \tau, \varepsilon)]\| \) is bounded by \( \varepsilon^{2r+1-2j} b_{ij}(\tau) \) for \( \tau \geq 0 \).

The proof is an exercise in differentiation; see the proof of Lemma 5.2. Notice that if \( j \leq r \), then \( 2r + 1 - 2j \geq 1 \).

We shall show that \( \tilde{H}^{n(i,j)}_\eta(x, \varepsilon) \) depends continuously on \((x, \varepsilon)\) at \( \varepsilon = 0 \) by showing that \( \tilde{H}^{n(i,j)}_\eta(x, \varepsilon) \to 0 \) as \( \varepsilon \to 0 \) uniformly on bounded subsets of \( X^\gamma \). Let \( x \in B_\rho \) and let \( \varepsilon \neq 0 \) be small. If \( i \geq 1 \), then

\[
\|\tilde{H}^{n(i,j)}_\eta(x, \varepsilon)(x_1, \ldots, x_i, \varepsilon_1, \ldots, \varepsilon_j)\|_\eta \\
= \sup_{\tau \geq 0} \varepsilon^{\eta r} \|D^i_1 D^j_3 [\varepsilon^{2r+1} h(x(\tau), \tau, \varepsilon)](x_1(\tau), \ldots, x_i(\tau), \varepsilon_1, \ldots, \varepsilon_j)\| \\
\leq \sup_{\tau \geq 0} \varepsilon^{\eta r} e^{2r+1-2j} b_{ij}(\tau) e^{-\eta r} \|x_i\|_\gamma \cdots \|x_1\|_\gamma \|\varepsilon_1\| \cdots \|\varepsilon_j\|,
\]

License or copyright restrictions may apply to redistribution; see http://www.ams.org/journal-terms-of-use
and \( \sup_{\tau \geq 0} e^{\eta \tau} e^{2r+1-2j} b_{ij}(\tau) e^{-\gamma \tau} \to 0 \) as \( \varepsilon \to 0 \). If \( i = 0 \), we use the fact that \( \tilde{h}(0,\tau,\varepsilon) \equiv 0 \):

\[
\|\tilde{H}_0^{(0,j)}(x,\varepsilon)\|_\eta = \sup_{\tau \geq 0} e^{\eta \tau} \| D^2_{\tilde{h}} [e^{2r+1} \tilde{h}(x(\tau),\tau,\varepsilon)] \| \\
= \sup_{\tau \geq 0} e^{\eta \tau} \left[ \int_0^1 D_1 \tilde{h}(ux(\tau),\tau,\varepsilon) \, du \right] \| x(\tau) \| \\
\leq \sup_{\tau \geq 0} \sup_{0 \leq u \leq 1} e^{\eta \tau} \| D_1 D^2_{\tilde{h}} [e^{2r+1} \tilde{h}(ux(\tau),\tau,\varepsilon)] \| \| x(\tau) \| \\
\leq \sup_{\tau \geq 0} e^{\eta \tau} e^{2r+1-2j} b_{ij}(\tau) e^{-\gamma \tau} \| x \|_\gamma,
\]

which approaches 0 as \( \varepsilon \) approaches 0 uniformly on bounded subsets of \( X_\gamma \).

Finally we must compute the partial derivatives of \( \tilde{H}_0^{(i,j)}(x,0) \) at \( (x,0) \), \( 0 \leq i + j \leq r - 1 \). It is clear that \( D_1 \tilde{H}_0^{(i,j)}(x,0) = \tilde{H}_0^{(i+1,j)}(x,0) \) since \( \tilde{H}_0^{(i,j)}(x,0) = 0 \) for all \( x \) and \( \tilde{H}_0^{(i+1,j)}(x,0) = 0 \). To show that \( D_2 \tilde{H}_0^{(i,j)}(x,0) = \tilde{H}_0^{(i,j+1)}(x,0) = 0 \), we must show that for \( 0 \leq i + j \leq r - 1 \), \( \tilde{H}_0^{(i,j)}(x,\varepsilon) = o(|\varepsilon|) \) for fixed \( x \). This follows from (18) and (19) by noting that for \( j \leq r - 1, 2r + 1 - 2j \geq 3 \).

7. Proofs of Propositions 2.3, 2.4 and 2.5. We consider the mapping \( F(x,x_0,\varepsilon) := Lx_0 + KN(x,\varepsilon) \), where \( L, K, \) and \( N \) are as defined in §2. The numbers \( M \) and \( \alpha \) and the spaces \( X_\gamma \) are also as defined in §2. We assume for simplicity that the norms of the projections \( \pi_+ \) and \( \pi_- \) are one; of course, this can always be arranged by a linear change of coordinates.

**Lemma 7.1.** If \( 0 < j \leq \alpha \), then \( \|Lx_0\|_\gamma \leq M \|x_0\| \).

**Proof.**

\[
\|Lx_0\|_\gamma = \sup_{\tau \geq 0} e^{\eta \tau} \| e^{\gamma \tau} x_0 \| \leq \sup_{\tau \geq 0} e^{\eta \tau} M e^{-\alpha \tau} \|x_0\| \leq M \|x_0\|. \tag{20}
\]

**Lemma 7.2.** If \( \alpha > \gamma \geq \eta > 0 \), then \( K \) maps \( X_\gamma \) into \( X_\eta \). The norm of \( K \) as a linear map from \( X_\gamma \) to \( X_\eta \) is at most \( M (1/(\alpha - \gamma) + 1/(\alpha + \gamma)) \).

**Proof.**

\[
\|Kx\|_\eta = \sup_{\tau \geq 0} e^{\eta \tau} \| Kx(\tau) \|, \quad \mathrm{and}
\]

\[
e^{\eta \tau} \| Kx(\tau) \| \leq e^{\eta \tau} \left\{ \int_0^\tau M e^{-\alpha (\tau - \sigma)} e^{-\gamma \sigma} \| x \|_\gamma d\sigma + \int_\tau^\infty M e^{\alpha (\tau - \sigma)} e^{-\gamma \sigma} \| x \|_\gamma d\sigma \right\} \quad \text{since } \alpha > \gamma \geq \eta.
\]

Now Theorem 5.1(1) and (2) imply in particular that \( N \) maps \( X_\gamma \) into \( X_\eta \) if \( \gamma \geq \eta \). Proposition 2.3 follows from this fact together with Lemmas 7.1 and 7.2.

Proposition 2.4 follows from Lemma 7.2, Proposition 2.2, Theorem 5.1(1) and (2), and Theorem 6.1.

To prove Proposition 2.5, let \( \alpha_1 \in (0,\alpha) \) and let \( M_1 = M (1/(\alpha - \alpha_1) + 1/\alpha) \).

Choose \( \delta > 0 \) so small that if \( \|y\| < \delta \) and \( |\varepsilon| < \delta \), then

\[
\sum_{i=0}^{2r} |\varepsilon|^i \| Dg_i(y) \| + |\varepsilon|^{2r+1} \| D_1 h(y,t,\varepsilon) \| < \min \left( \frac{1}{2M_1}, \frac{\kappa}{M_1} \right).
\]
We can do this because \( Dg_0(0) = 0 \). If \( \|x\|_\gamma < \delta \) and \( |\varepsilon| < \delta \), then

\[
\|N(x, \varepsilon)\|_\gamma \leq \sup_{\tau \geq 0} \varepsilon^{\gamma} \left( \sum_{i=0}^{2r} |\varepsilon|^i \|g_i(x(\tau))\| + |\varepsilon|^{2r+1} \left\| h \left( x(\tau), \frac{\tau}{\varepsilon}, \varepsilon \right) \right\| \right)
\]

\[
\leq \sup_{\tau \geq 0} \varepsilon^{\gamma} \left( \sum_{i=0}^{2r} |\varepsilon|^i \int_0^1 \|Dg_i(ux(\tau))\| \, du \right.
\]

\[+ \left|\varepsilon\right|^{2r+1} \int_0^1 \left\| D_1 h \left( ux(\tau), \frac{\tau}{\varepsilon}, \varepsilon \right) \right\| \, du \right) \|x(\tau)\| \]

\[
\leq \frac{\delta}{2M_1}.
\]

If \( \gamma \in (0, \alpha_1) \), \( \|x_0\| < \delta/2M \), \( |\varepsilon| < \delta \), and \( \|x\|_\gamma \leq \delta \), then this computation and Lemmas 7.1 and 7.2 imply that

\[
\|F(x, x_0, \varepsilon)\|_\gamma \leq \|Lx_0\| + \|KN(x, \varepsilon)\| \leq M\delta/2M + M_1\delta/2M_1 = \delta.
\]

Thus we take \( \delta_1 := \min(\delta/2M, \delta) \). The estimate on \( \|D_1 F_\gamma^*(x, x_0, \varepsilon)\| \) uses (20) and the formula

\[
[D_1 N_\gamma^*(x, x_0, \varepsilon)x_1](\tau) = \left[ \sum_{i=0}^{2r} \varepsilon^i Dg_i(x(\tau)) + \varepsilon^{2r+1} D_1 h \left( x(\tau), \frac{\tau}{\varepsilon}, \varepsilon \right) \right] x_1(\tau),
\]

which follows from Theorems 5.1(1) and (2).

8. Contractions on a scale of Banach spaces. A scale of Banach spaces is a family of Banach spaces \( \{X_\gamma\} \), \( \gamma_0 < \gamma < \gamma_1 \), and one-to-one bounded linear maps \( J_\gamma : X_\gamma \to X_\eta \), \( \gamma > \eta \), such that \( J_\gamma \circ J_\xi = J_\xi \) whenever \( \xi > \gamma > \eta \). For example, the \( X_\gamma \) defined in §2, \( 0 < \gamma < \infty \), together with the natural injections \( J_\gamma^0 \) of \( X_\gamma \) into \( X_\eta \) for \( \gamma > \eta \), form a scale of Banach spaces.

Let \( \Lambda \) be an open subset of some Banach space. A family of mappings \( F_\gamma^0 : X_\gamma \times \Lambda \to X_\eta \), \( \gamma \geq \eta \), is called scale invariant if \( J_\gamma \circ F_\xi^0 = F_\gamma^0 \) whenever \( \xi \geq \gamma > \eta \), and \( F_\gamma^0(J_\gamma^0 x, \lambda) = F_\gamma^0(x, \lambda) \) whenever \( \xi > \gamma \geq \eta \).

**Theorem 8.1.** Let \( \{X_\gamma\} \) be a scale of Banach spaces and \( \{F_\gamma^0\} \) a scale invariant family of mappings. For each \( \gamma \) let \( Q_\gamma \) be a closed convex subset of \( X_\gamma \) such that \( J_\gamma Q_\gamma \subset Q_\gamma \) for \( \gamma > \eta \), and \( F_\gamma^0(Q_\gamma \times \Lambda) \subset Q_\gamma \) for all \( \gamma \). Assume:

1. \( D_1 F_\gamma^0(x, \lambda) \) exists for all \( \gamma \) and all \( (x, \lambda) \in Q_\gamma \times \Lambda \).
2. There exists a number \( \kappa \), \( 0 < \kappa < 1 \), such that \( \|D_1 F_\gamma^0(x, \lambda)\| \leq \kappa < 1 \) for all \( \gamma \) and all \( (x, \lambda) \in Q_\gamma \times \Lambda \).
3. \( F_\gamma^0 \) is \( C^* \), \( r \geq 1 \), if \( \gamma > \eta \).

Let \( x_\gamma(\lambda) \) denote the unique fixed point of \( F_\gamma^0(\cdot, \lambda) \) in \( Q_\gamma \). Then \( J_\gamma x_\gamma(\lambda) = x_\gamma(\lambda) \) whenever \( \gamma > \eta \), and \( x_\gamma(\lambda) \) is \( C^* \) for each \( \gamma \). Moreover, \( Dz_\gamma(\lambda) \) is the unique fixed point of the equation

\[
A = D_1 F_\gamma^0(x_\gamma(\lambda), \lambda)A + D_2 F_\xi^0(x_\xi(\lambda), \lambda)
\]

for any \( \xi > \gamma \).

In using this theorem to prove Theorem 1.1, one refers to Proposition 2.4: \( X_\gamma \) and \( F_\gamma^0 \) are as in §2, \( (\gamma_0, \gamma_1) := (0, \alpha_1) \), \( \Lambda \) is the ball of radius \( \delta_1 \) in \( E_- \times \mathbb{R} \), and \( Q_\gamma \).
is the ball of radius $\delta$ in $X_\gamma$. Propositions 2.3, 2.4 and 2.5 say that the hypotheses of Theorem 8.1 are satisfied.

In stating Theorem 8.1, we have made no attempt to give the weakest hypotheses that permit the conclusion that some $x_\gamma(\lambda)$ is $C^r$. In particular, Theorem 8.1 as stated does not apply to the proof of the center manifold theorem [8]: there $F^n_\eta$ is $C^r$ only for $\gamma$ sufficiently greater than $\eta$. All the ideas necessary to prove Theorem 8.1 are in [8]. We include the proof for the reader’s convenience.

In the proof, $L^k(X,Y)$ denotes the Banach space of $k$-multilinear maps between Banach spaces $X$ and $Y$, with the usual operator norm. The symbol $\| \|$ is used to denote the norm in any Banach space, including the spaces $L^k(X,Y)$. Which norm is meant should always be clear from the expression between the vertical lines.

PROOF. For each $\gamma \in (\gamma_0, \gamma_1)$ and each $\lambda \in \Lambda$, $F^n_\gamma|Q_\gamma \times \{\lambda\}$ is a contraction with contraction constant $\kappa$. By the contraction mapping theorem, $F^n_\gamma|Q_\gamma \times \{\lambda\}$ has a unique fixed point $x_\gamma(\lambda)$. The first part of the conclusion follows from

$$F^n_\eta(J^n_\gamma x_\gamma(\lambda), \lambda) = F^n_\eta(x_\gamma(\lambda), \lambda) = J^n_\gamma \circ F^n_\gamma(x_\gamma(\lambda), \lambda) = J^n_\gamma x_\gamma(\lambda).$$

We shall now prove the second part of the conclusion in the case $r = 1$ in a sequence of steps.

1. The map $x_\gamma(\lambda)$ is Lipschitz continuous for all $\eta$. Let $\gamma > \eta$. Then

$$\|x_\eta(\mu) - x_\eta(\lambda)\| = \|F^n_\eta(x_\eta(\mu), \mu) - F^n_\eta(x_\eta(\lambda), \lambda)\| \leq \|F^n_\eta(x_\eta(\mu), \mu) - F^n_\eta(x_\eta(\lambda), \mu)\|$$

$$+ \|F^n_\eta(x_\eta(\lambda), \mu) - F^n_\eta(x_\eta(\lambda), \lambda)\| \leq \kappa \|x_\eta(\mu) - x_\eta(\lambda)\| + \|F^n_\eta(x_\gamma(\lambda), \mu) - F^n_\eta(x_\gamma(\lambda), \lambda)\|.$$ 

(To justify the last step, note that $F^n_\eta(x_\gamma(\lambda), \mu) = F^n_\eta(J^n_\gamma x_\eta(\lambda), \mu) = F^n_\eta(x_\gamma(\lambda), \mu).$) Therefore

$$\|x_\eta(\mu) - x_\eta(\lambda)\| \leq (1 - \kappa)^{-1} \int_0^1 \|D_2 F^n_\eta(x_\gamma(\lambda), u\mu + (1 - u)\lambda)\| \, du |\mu - \lambda|.$$

Since $F^n_\eta$ is $C^1$, for fixed $\lambda$ there is a constant $b(\lambda)$ such that

$$\|D_2 F^n_\eta(x_\gamma(\lambda), \mu) + (1 - u)\lambda)\| \leq b(\lambda)(1 - \kappa)^{-1}|\mu - \lambda|$$

for all $\mu$ sufficiently near $\lambda$ and for all $u \in [0,1]$. Therefore $\|x_\eta(\mu) - x_\eta(\lambda)\| \leq b(\lambda)(1 - \kappa)^{-1}|\mu - \lambda|$ for all $\mu$ sufficiently near $\lambda$.

2. For fixed $\lambda$ and $\xi > \gamma$, consider the equation

$$(21) \quad A = D_1 F^n_\gamma(x_\gamma(\lambda), \lambda)A + D_2 F^n_\gamma(x_\gamma(\lambda), \lambda),$$

where $A \in L(\Lambda, X_\gamma)$.

We note three facts:

(a) $D_2 F^n_\gamma(x_\gamma(\lambda), \lambda)$ is independent of $\xi$.

(b) $J^n_\gamma D_1 F^n_\gamma(x_\gamma(\lambda), \lambda) = D_1 F^n_\gamma(x_\gamma(\lambda), \lambda)J^n_\gamma$.

(c) $J^n_\gamma D_2 F^n_\gamma(x_\gamma(\lambda), \lambda) = D_2 F^n_\gamma(x_\gamma(\lambda), \lambda)$.

The second follows from the chain rule:

$$J^n_\gamma D_1 F^n_\gamma(x, \lambda) = D_1 J^n_\gamma F^n_\gamma(x, \lambda) = D_1 F^n_\gamma(x, \lambda) = D_1 F^n_\gamma(J^n_\gamma x, \lambda) J^n_\gamma.$$
The third fact also follows from the chain rule. The first follows from differentiation with respect to $\lambda$ of the identity

$$F^{\gamma}(x_\xi(\lambda), \lambda) = F^{\omega}(x_\omega(\lambda), \lambda), \quad \omega > \xi > \gamma.$$  

Note that

$$D_1 F^{\gamma}_\xi(x_\xi(\lambda), \lambda) \frac{dx_\xi}{d\lambda} = D_1 F^{\gamma}_\omega(x_\xi(\lambda), \lambda) J^\xi \frac{dx_\omega}{d\lambda} = D_1 F^{\gamma}_\omega(x_\omega(\lambda), \lambda) \frac{dx_\omega}{d\lambda}.$$  

Now the equation (21), which is linear in $A$, has the unique fixed point $A_\eta(\lambda)$. By (a), $A_\gamma(\lambda)$ does not depend on $\xi$. Moreover, applying $J^\eta_\gamma$ to both sides of (21), we find, using (b) and (c):

$$J^\eta_\gamma A_\gamma(\lambda) = D_1 F^{\eta}_\eta(x_\eta(\lambda), \lambda) J^\eta_\gamma A_\gamma(\lambda) + D_2 F^{\eta}_\xi(x_\xi(\lambda), \lambda).$$  

Therefore $A_\eta(\lambda) = J^\eta_\gamma A_\gamma(\lambda)$.

3. We shall show that $x_\eta(\lambda)$ is differentiable with derivative $A_\eta(\lambda)$. Let $\xi > \gamma > \eta$. Then

$$x_\eta(\mu) - x_\eta(\lambda) - A_\eta(\lambda)(\mu - \lambda) = \int_{\xi}^{\mu} \{D_1 F^\eta_\eta(x_\eta(\lambda), \lambda) - F^{\eta}_\eta(x_\eta(\lambda), \lambda)\} du$$

$$= \int_{\xi}^{\mu} \{J^\eta_\gamma F^\eta_\eta(x_\eta(\lambda), \lambda) - D_1 F^\eta_\eta(x_\eta(\lambda), \lambda)A_\eta(\lambda)(\mu - \lambda)\}$$

$$= \{F^{\eta}_\gamma(x_\eta(\lambda), \mu) - F^{\eta}_\gamma(x_\eta(\lambda), \mu) - D_1 F^{\eta}_\eta(x_\eta(\lambda), \lambda)A_\gamma(\lambda)(\mu - \lambda)\}$$

$$+ \{D_1 F^{\eta}_\eta(x_\eta(\lambda), \lambda)(x_\eta(\mu) - x_\eta(\lambda)) - A_\gamma(\lambda)(\mu - \lambda)) + R(\lambda, \mu)$$

$$= D_1 F^{\eta}_\eta(x_\eta(\lambda), \lambda)(x_\eta(\mu) - x_\eta(\lambda)) - A_\eta(\lambda)(\mu - \lambda) + R(\lambda, \mu).$$

Here

$$R(\lambda, \mu) = \int_0^1 [D_1 F^\eta_\eta(u x_\gamma(\mu) + (1 - u) x_\gamma(\lambda), \mu) - D_1 F^\eta_\eta(x_\gamma(\lambda), \lambda)](x_\gamma(\mu) - x_\gamma(\lambda)) du$$

$$+ \int_0^1 [D_2 F^\eta_\xi(x_\xi(\lambda), \lambda) - D_2 F^\eta_\xi(x_\xi(\lambda), \lambda)](\mu - \lambda) du.$$  

Since $x_\gamma$, $x_\xi$, $D_1 F^\eta_\gamma$, and $D_2 F^\eta_\xi$ are continuous, given $\epsilon > 0$ and $\lambda$ there is a $\delta > 0$ such that if $|\mu - \lambda| < \delta$ and $0 < u < 1$, then the norms of the expressions in square brackets are less than $\epsilon$. Moreover, for all sufficiently small $\delta$, $\|x_\gamma(\mu) - x_\gamma(\lambda)\| \leq c|\mu - \lambda|$ for some constant $c$ whenever $|\mu - \lambda| < \delta$. Therefore

$$\|R(\lambda, \mu)\| < (c + \|J^\eta_\gamma\|)\|\mu - \lambda\|$$

if $|\mu - \lambda| < \delta$. Therefore

$$\|x_\eta(\mu) - x_\eta(\lambda) - A_\eta(\lambda)(\mu - \lambda)\| < (1 - \kappa)^{-1}(c + \|J^\eta_\gamma\|)\|\mu - \lambda\|$$

if $|\mu - \lambda| < \delta$.

4. $A_\eta(\lambda)$ is a continuous function of $\lambda$. Let $\gamma > \eta$. $A_\eta(\lambda)$ is the fixed point of

$$G(A, \lambda) := D_1 F^\eta_\eta(x_\eta(\lambda), \lambda) A + D_2 F^\eta_\xi(x_\eta(\lambda), \lambda).$$  

License or copyright restrictions may apply to redistribution; see http://www.ams.org/journal-terms-of-use
Therefore $A_\eta(\lambda)$ is continuous if for each $A \in A_\eta(\Lambda)$, the map $\mu \to G(A, \mu)$ is continuous. (The argument is similar to step 1.) Since $A_\eta(\lambda) = J_\eta A_\eta(\lambda)$, we have

$$G(A_\eta(\lambda), \mu) = D_1 F^\eta_\eta(x_\eta(\mu), \mu) A_\eta(\lambda) + D_2 F^\eta_\eta(x_\eta(\mu), \mu)$$

$$= D_1 F^\eta_\eta(x_\eta(\mu), \mu) A_\eta(\lambda) + D_2 F^\eta_\eta(x_\eta(\mu), \mu),$$

which is a continuous function of $\mu$ because $\gamma > \eta$ and $x_\eta(\mu)$ is continuous.

This completes the proof of the theorem in the case $r = 1$.

Suppose the theorem has been proved for some integer $r \geq 1$. Suppose moreover that we have shown:

(A) $D^r x_\eta(\lambda) \in L^r(\Lambda, X_\eta)$ satisfies an equation of the form

$$A = D_1 F^\eta_\eta(x_\eta(\lambda), \lambda) A + H^\eta_\eta(\lambda),$$

where $\gamma > \eta$ and $H^\eta_\eta(\lambda) \in L^r(\Lambda, X_\eta)$ has the following form. Let $i = (i_1, \ldots, i_j)$ denote a $j$-tuple of integers. Then

$$H^\eta_\eta(\lambda) = \sum a_i D^i_1 D^{-|i|}_2 F^\eta_\eta(x_\eta(\lambda), \lambda) D^{i_1} x_1 \cdots D^{i_j} x_j.$$

The sum is over all $j$-tuples $i$ with $0 \leq j \leq r$ and $0 \leq |i| \leq r$, except the 1-triple $i = (r)$. The possibility $i = \emptyset$ is allowed; in that case the corresponding term is $D^r_2 F^\eta_\eta(x_\eta(\lambda), \lambda)$. The coefficients $a_i$ are not important for our purposes.

(B) $H^\eta_\eta(\lambda) = J^\eta_\eta H^\xi_\xi$ whenever $\xi > \gamma > \eta$.

(C) $H^\eta_\eta(\lambda)$ is independent of $\gamma$.

We shall show that the theorem is true for the integer $r + 1$, and that (A), (B) and (C) hold with $r$ replaced by $r + 1$. Thus the theorem is true by induction. (Note that (A), (B), and (C) hold for $r = 1$ by step 2.)

Assume $F^\eta_\eta$ is $C^{r+1}$ for $\gamma > \eta$, the theorem is true for the integer $r \geq 1$, and (A), (B) and (C) hold. Define $Y_\gamma := L^r(\Lambda, X_\gamma)$; for $\gamma > \eta$ define $J^\eta_\gamma : Y_\gamma \to Y_\eta$ by $J^\eta_\gamma A := J^\eta_\eta \circ A$; for $\gamma > \eta$ define $F^\eta_\gamma : Y_\gamma \times \Lambda \to Y_\eta$ by

$$F^\eta_\gamma(\lambda) := D_1 F^\eta_\eta(x_\eta(\lambda), \lambda) A + H^\eta_\eta(\lambda).$$

Define $\tilde{F}^\eta_\gamma : Y_\gamma \times \Lambda \to Y_\eta$ by

$$\tilde{F}^\eta_\gamma(\lambda) := D_1 F^\eta_\eta(x_\eta(\lambda), \lambda) A + H^\eta_\eta(\lambda)$$

for any $\gamma > \eta$. $\tilde{F}^\eta_\gamma$ is well defined by (B).

Clearly $\{Y_\gamma\}$ with the maps $\{J^\eta_\gamma\}$ is a scale of Banach spaces. We claim that $\{\tilde{F}^\eta_\gamma\}$ is a scale invariant family of mappings of $\{Y_\gamma\}$ that satisfies the conditions of the theorem with $r = 1$. To show scale invariance, we note:

$$J^\eta_\gamma \tilde{F}^\eta_\xi(\lambda, \lambda) = J^\eta_\gamma D_1 F^\eta_\xi(x_\xi(\lambda), \lambda) A + J^\eta_\gamma H^\eta_\xi(\lambda)$$

$$= D_1 F^\eta_\xi(x_\xi(\lambda), \lambda) A + H^\eta_\xi(\lambda) = \tilde{F}^\eta_\xi(\lambda).$$

We have used (B) in the first computation and (C) in the second.

The "closed convex subset" of the theorem is $Y_\gamma$. Assumptions (1) and (2) of the theorem are clear: note that $\tilde{F}^\eta_\gamma$ is linear in $A$. Since $F^\eta_\gamma$ is $C^{r+1}$ for $\gamma > \eta$, each
partial derivative of $F^\eta_\gamma$ that appears in the definition of $H^\eta_\gamma$ is at least $C^1$. Also, $x_\gamma(\lambda)$ is $C^r$ by the induction hypothesis, and only derivatives of $x_\gamma(\lambda)$ through order $r-1$ appear in the definition of $H^\eta_\gamma$. Therefore $H^\eta_\gamma$ is $C^1$ for $\gamma > \eta$. It follows from the case $r = 1$ of the theorem that the unique fixed point $A_\eta(\lambda)$ of $\tilde{F}^\eta_\gamma$ is $C^1$; since $A_\eta(\lambda) = D^r x_\eta(\lambda)$, $x_\eta(\lambda)$ is $C^{r+1}$. The last line of the theorem implies that $D^{r+1} x_\eta(\lambda)$ satisfies the equation.

\[ B = D_1 \tilde{F}^\eta_\gamma (A_\eta(\lambda), \lambda) B + D_2 \tilde{F}^\eta_\gamma (A_\gamma(\lambda), \lambda) \]

\[ = D_1 F^\eta_\gamma (x_\eta(\lambda), \lambda) B + K^\eta_\gamma (\lambda), \]

where

\[ K^\eta_\gamma (\lambda) = D_1^2 F^\eta_\gamma (x_\gamma(\lambda), \lambda) D x_\gamma D^r x_\gamma \]

\[ + D_1 D_2 F^\eta_\gamma (x_\gamma(\lambda), \lambda) D^r x_\gamma + DH^\eta_\gamma (\lambda). \]

The arguments of step 2, applied to $\tilde{F}$, show that $K^\eta_\gamma$ satisfies (B) and (C). From the formula for $K^\eta_\gamma$ is also satisfies (A). □

REFERENCES