NUMER. FUNCT. ANAL. AND OPTIMIZ., 14(5&6), 503–514 (1993)

ON SMOOTHNESS AND INVARIANCE PROPERTIES OF THE GAUSS-NEWTON METHOD

W.-J. Beyn

Fakultät für Mathematik
Universität Bielefeld
Postfach 100131
D-33501 Bielefeld, Germany

Key Words: Gauss-Newton method, parametrized equations, invariant manifolds, foliations.

AMS(MOS) subject classification: 65H10,65J05,58F18.

ABSTRACT

We consider systems of m nonlinear equations in $m + p$ unknowns which have p-dimensional solution manifolds. It is well-known that the Gauss-Newton method converges locally and quadratically to regular points on this manifold. We investigate in detail the mapping which transfers the starting point to its limit on the manifold. This mapping is shown to be smooth of one order less than the given system. Moreover, we find that the Gauss-Newton method induces a foliation of the neighborhood of the manifold into smooth submanifolds. These submanifolds are of dimension m, they are invariant under the Gauss-Newton iteration, and they have orthogonal intersections with the solution manifold.

1. INTRODUCTION

We consider a nonlinear system of m equations in $m + p$ unknowns

$$F(v) = 0 \tag{1.1}$$

where $F \in C^{k+1}(\mathfrak{R}^{m+p}, \mathfrak{R}^m)$, $k \geq 1$. Let $\bar{v} \in \mathfrak{R}^{m+p}$ be a regular solution of (1.1), i.e.

$$F(\bar{v}) = 0 \qquad \text{Rank}(F'(\bar{v})) = m$$

In a neighborhood of $\bar{v}$ we want to analyze the behavior of the *Gauss-Newton method*

$$v_{n+1} = T(v_n) \qquad T(v) = v - F'(v)^+F(v) \tag{1.2}$$

Here $F'(v)^+$ denotes the pseudo-inverse of $F'(v)$.

It is well-known that the intersection of the solution set $M = F^{-1}(0)$ with a neighborhood of $\bar{v}$ forms a p-dimensional C^{k+1}-submanifold of $\mathcal{R}^{m+p}$ and that it may be written as the graph of a C^{k+1}-function

$$w_\infty: N(F'(\bar{v})) \mapsto N(F'(\bar{v}))^\perp \qquad s \mapsto w_\infty(s)$$

see, e.g., Rheinboldt [10]. It is also well-known that the Gauss-Newton method (1.2) converges quadratically to a certain limit in this manifold

$$\lim_{n\to\infty} v_n =: T_\infty(v_0)$$

provided v_0 is sufficiently close to $\bar{v}$, cf. Deuflhard and Heindl [3], Allgower and Georg [1].

In continuation methods the Gauss-Newton method or one of its modifications is often used as a corrector iteration (see [11], [1], [4], [7], [2]). Therefore, it is important to know whether small perturbations of the predictor v_0 lead to small perturbations of the corrector $T_\infty(v_0)$. We will show that T_∞ is in fact a C^k-function in case of the Gauss-Newton method. This is not obvious, since the Gauss-Newton method involves an implicit parametrization of the solution manifold. It generates a sequence which ultimately approaches the solution in a normal direction.

Other corrector methods use for parametrization an explicit equation of the form

$$G(v, v_0) = 0 \qquad G: \mathcal{R}^{m+p} \times \mathcal{R}^{m+p} \mapsto \mathcal{R}^p \tag{1.3}$$

and then try to solve the quadratic system (1.1), (1.3) for v. E.g., for the pseudo-arclength method (see [9]) the solutions of (1.3) form a hyperplane orthogonal to a previous secant or tangent and for the simplified Gauss-Newton method (i.e., $T(v) = v - F'(v_0)^+F(v)$, see [11]) they form the m-dimensional subspace orthogonal to $N(F'(v_0))$. The smooth dependence of the solutions of (1.1), (1.3) on the predictor v_0 is then an easy consequence of the implicit function theorem.

One way of studying the smoothness of the map T_∞ is via an analysis of the set of all points v_0 which under the Gauss-Newton method converge to a fixed solution of (1.1). In Section 3 we will show that these sets are in fact C^k-manifolds of dimension m which intersect the solution manifold orthogonally. In this way we obtain a foliation of the space into T-invariant smooth submanifolds. In a certain sense this provides us with a complete picture of the relations between starting and limit values.

Such invariant foliations also appear in the theory of stable and unstable manifolds of diffeomorphisms (see Fenichel [5, 6]), but we notice that the Gauss-Newton operator is, in general, not a diffeomorphism. E.g., in the linear case it is in fact a projection.

Let us finally remark that the behavior of the Gauss-Newton method and its invariant manifolds becomes much more complicated in the neighborhood of singular points, e.g., near simple bifurcation points. Here only partial results, such as the existence of small convergence cones around the branches (see [1, 9, 8]), are known. This case is currently under investigation.

2. DIFFERENTIABILITY OF THE OPERATOR T_∞

Let us first review the classical results on the convergence of the Gauss-Newton method. Without loss of generality we assume that the regular solution of (1.1) is $\bar{v} = 0$.

Theorem 2.1: Let $0 \in \mathcal{R}^{m+p}$ be a regular zero of $F \in C^{k+1}(\mathcal{R}^{m+p}, \mathcal{R}^m)$ with $k \geq 1$. Then there exist open neighborhoods $V_0 \subset V_1$ of $0 \in \mathcal{R}^{m+p}$ and a C^{k+1}-function

$$w_\infty : V_1 \cap N(F'^0) \mapsto N(F'^0)^\perp \qquad F'^0 := F'(0)$$

with the following properties

(i) $w_\infty(0) = 0$, $w'_\infty(0) = 0$ and

$$V_1 \cap F^{-1}(0) = \{s + w_\infty(s) : s \in V_1 \cap N(F'^0)\}$$

(ii) For each $v_0 \in V_0$ the Gauss-Newton sequence $v_n = T^n(v_0)$ exists, lies in V_1, and converges to some limit $T_\infty(v_0) \in V_1 \cap F^{-1}(0)$. For some $\alpha < 1$ and for all $n > 0$, $v_0 \in V_0$ the following estimates hold

$$\|v_n - T_\infty(v_0)\|_2 \leq C\alpha^{2^n} \qquad \|F(v_n)\|_2 \leq C\alpha^{2^n} \tag{2.1}$$

The geometrical situation of Theorem 2.1 is illustrated in Figure 1 for the case $m = 2, p = 1$.

Our basic result is

Theorem 2.2: Under the assumptions of Theorem 2.1 and for sufficiently small V_0 we have $T_\infty \in C^k(V_0, \mathcal{R}^{m+p})$.

In this section we will give an elementary proof of this Theorem for the case $k = 1$.

Let us first notice that Rank $F'^0 = m$ implies that $F'(v)$ has constant rank m throughout some neighborhood of 0. Calling this neighborhood V_1 again we obtain

$$F'(v)^+ = F'(v)^T(F'(v)F'(v)^T)^{-1} \qquad v \in V_1$$

and thus $F'(\cdot)^+ \in C^k(V_1, \mathcal{R}^{m+p,m+p})$.

Our next observation is that (2.1) implies the uniform convergence

$$\sup_{v_0 \in V_0} \|T^n(v_0) - T_\infty(v_0)\|_2 \to 0 \quad \text{as} \quad n \to \infty$$

and hence the continuity of T_∞.

We show that also the derivatives $T^{n\prime}$ converge uniformly on V_0 to some bounded continuous function

$$A_\infty : V_0 \mapsto \mathcal{R}^{m+p,m+p}$$

By a familiar result from analysis this implies that T_∞ is continuously differentiable in V_0 with $T'_\infty = A_\infty$.

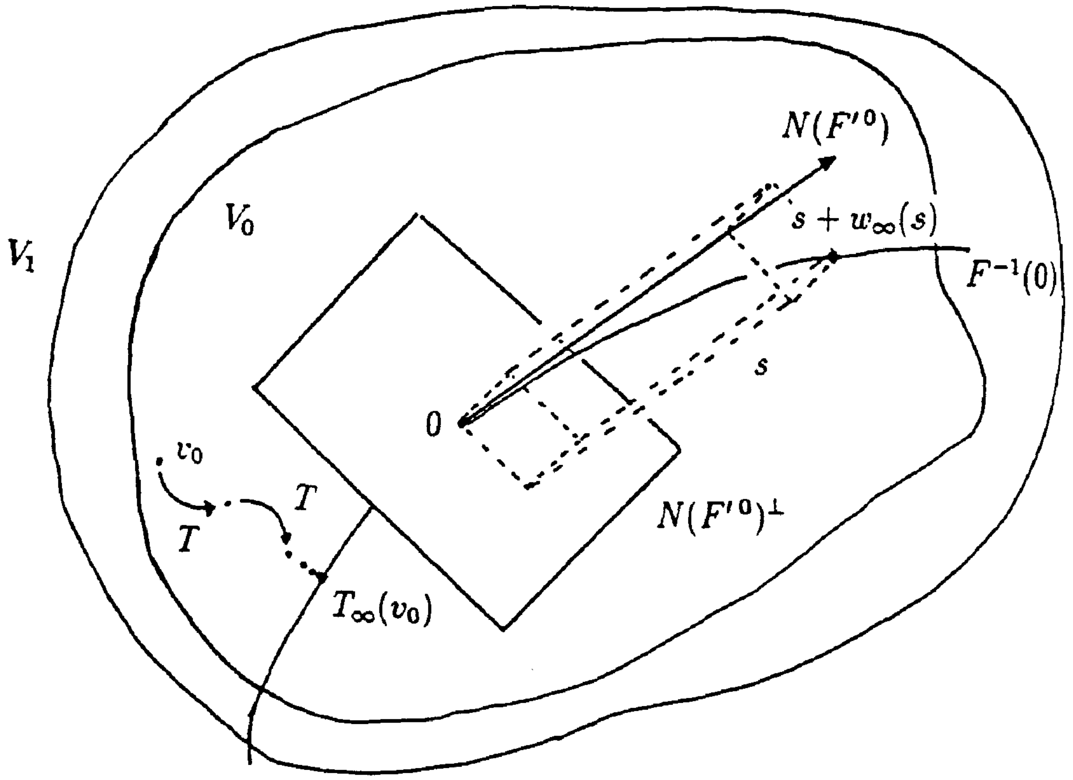

Figure 1 Geometry in the neighborhood of a regular point.

By the chain rule we have for $v \in V_1$

$$T^{n\prime}(v) = T'(T^{n-1}(v)) \cdot \cdots \cdot T'(v) = \prod_{i=0}^{n-1} T'(T^i(v)) \tag{2.2}$$

Here and in what follows we use the convention that in a matrix product

$$\prod_{i=0}^{n} A_i = A_n \cdot \cdots \cdot A_0$$

the factors are multiplied from the left with increasing index.

From (1.2) we obtain for $v \in V_1$

$$T'(v) = P(v) - S(v) \tag{2.3}$$

where

$$S(v) = \left(\frac{d}{dv} F'(v)^+\right) F(v)$$

and

$$P(v) = I - F'(v)^+ F'(v) \tag{2.4}$$

is the orthogonal projector onto $N(F'(v))$.

Using this expression and Theorem 2.1 we see that the factors in (2.2) are small perturbations of the projector $P(T_\infty(v))$. For such a product the following Lemma is useful.

Lemma 2.3: Let $Q \in \mathcal{R}^{q,q}$ be an orthogonal projector and let $E_i \in \mathcal{R}^{q,q}$, $i \in \mathcal{N}$ be matrices satisfying

$$\eta := \sum_{i=0}^{\infty} \|E_i\|_2 < \infty$$

Then the product $A_n = \prod_{i=0}^{n} (Q + E_i)$ converges to some $A_\infty \in \mathcal{R}^{q,q}$ and the following holds

$$\|A_m - A_n\|_2 \le e^{2\eta} \sum_{i=n}^{m} \|E_i\|_2 \qquad \text{for } m \ge n \tag{2.5}$$

$$(I - Q)A_\infty = 0 \tag{2.6}$$

Proof: We let $\eta_i = \|E_i\|_2$ and show that for $m \ge n$

$$\left\| \prod_{i=n}^{m} (Q + E_i) - Q \right\|_2 \le \sum_{j=n}^{m} \left\{ \eta_j \prod_{i=j+1}^{m} (1 + \eta_i) \right\} \tag{2.7}$$

The proof is by induction on m. For $m = n$ (2.7) is trivial. Now suppose that (2.7) is true for some m, then we can estimate as follows

$$\begin{aligned}
\left\| \prod_{i=n}^{m+1} (Q + E_i) - Q \right\|_2 &= \left\| (Q + E_{m+1}) \left(\prod_{i=n}^{m} (Q + E_i) - Q \right) + E_{m+1} Q \right\|_2 \\
&\le (1 + \eta_{m+1}) \sum_{j=n}^{m} \left\{ \eta_j \prod_{i=j+1}^{m} (1 + \eta_i) \right\} + \eta_{m+1} \\
&= \sum_{j=n}^{m+1} \left\{ \eta_j \prod_{i=j+1}^{m+1} (1 + \eta_i) \right\}
\end{aligned}$$

The limit

$$\sigma = \lim_{n \to \infty} \prod_{i=0}^{n} (1 + \eta_i)$$

exists since we have a convergent majorant

$$\ln \prod_{i=0}^{n} (1 + \eta_i) = \sum_{i=0}^{n} \ln(1 + \eta_i) \le \sum_{i=0}^{n} \eta_i \le \eta$$

In particular, $\sigma \le e^{\eta}$. Now we use (2.7) to estimate for $m > n$

$$\begin{aligned}
\|A_m - A_n\|_2 &= \left\| \left\{ \prod_{i=n+1}^{m} (Q + E_i) - Q + Q - I \right\} \prod_{i=0}^{n} (Q + E_i) \right\|_2 \\
&\le \left\| \prod_{i=n+1}^{m} (Q + E_i) - Q \right\|_2 \left\| \prod_{i=0}^{n} (Q + E_i) \right\|_2 \\
&\quad + \left\| (Q - I) E_n \prod_{i=0}^{n-1} (Q + E_i) \right\|_2 \\
&\le \sigma^2 \sum_{j=n+1}^{m} \eta_j + \eta_n \sigma \le \sigma^2 \sum_{j=n}^{m} \eta_j
\end{aligned}$$

This proves (2.5) and hence the convergence of A_n.

Finally, we have

$$\|(I - Q)A_n\|_2 = \left\|(I - Q)E_n \prod_{i=0}^{n-1} (Q + E_i)\right\|_2 \leq \eta_n\sigma$$

from which (2.6) follows in the limit $n \to \infty$.

We can now complete the

Proof (Theorem 2.2, $k = 1$): Apply Lemma 2.3 to the projector

$$Q = P(T_\infty(v)) = I - F'(T_\infty(v))^+F'(T_\infty(v)) \qquad v \in V_0$$

and the matrices (see (2.3))

$$E_i = T'(T^i(v)) - Q = P(T^i(v)) - P(T_\infty(v)) - S(T^i(v))$$

From $F \in C^2$ and the regularity of $\bar{v} = 0$ we find that $P(v)$ is Lipschitz in V_0 and that $d/dvF'(v)^+$ is uniformly bounded in V_0. Hence, Theorem 2.1 yields

$$\|E_i\|_2 \leq C(\|T^i(v) - T_\infty(v)\|_2 + \|F(T^i(v)\|_2) \leq C\alpha^{2^i}$$

for all $i \in \mathcal{N}$ and $v \in V_0$.

From (2.5) we obtain that $A_n(v) := T^{n\prime}(v)$, $v \in V_0$ is a uniform Cauchy sequence. Moreover, $A_n(v)$ is uniformly bounded

$$\|A_n(v)\|_2 \leq \prod_{i=0}^{n} (1 + \|E_i\|_2) \leq \prod_{i=0}^{\infty} (1 + C\alpha^{2^i})$$

Therefore, A_n converges uniformly to some continuous bounded function A_∞: $V_0 \mapsto \mathcal{R}^{m+p,m+p}$, and Theorem 2.2 is proved.

Remark 2.4: Equation (2.6) implies

$$R(T'_\infty(v)) \subset N(F'(T_\infty(v)))$$

This is very natural, since T_∞ is in fact a map from $\mathcal{R}^{m+p}$ onto the manifold $M = V_1 \cap F^{-1}(0)$ which has the tangent space $N(F'(T_\infty(v)))$ at $T_\infty(v)$.

It is now tempting to establish $T_\infty \in C^k$ via the uniform convergence of the derivatives

$$T^{n(j)}(v) \qquad j = 1, \ldots, k \quad \text{as} \quad n \to \infty$$

However, this involves further differentiation of (2.2) and becomes rather awkward. We will therefore present a different approach in the next section which will allow us to obtain T_∞ from the implicit function theorem and which will give further insight into the behavior of the Gauss-Newton method.

3. INVARIANT MANIFOLDS FOR THE GAUSS-NEWTON OPERATOR

Let us reverse the question from Section 1, i.e., instead of asking for the limit $T_\infty(v)$ for a given v we would like to determine for a given solution (see Theorem 2.1)

$$v_\infty(s) := s + w_\infty(s) \qquad s \in U_1 := V_1 \cap N(F'^0)$$

the set of corresponding initial values

$$M_s := \{v \in V_0 : T_\infty(v) = v_\infty(s)\} \tag{3.1}$$

From this definition we see that M_s is positively invariant under the Gauss-Newton operator T. More precisely, we will show that M_s is an m-dimensional C^k-manifold which is a graph over $N(F'^0)^\perp$ and has tangent space $N(F'(v_\infty(s)))^\perp$ at $v_\infty(s)$ (see Figure 2).

In fact,

$$\bigcup_{s \in U_1} M_s$$

defines a smooth foliation of the neighborhood V_0 into T-invariant submanifolds.

Foliations of this type are generally known for stable and unstable manifolds of smooth invariant, hyperbolic manifolds of diffeomorphisms, see Fenichel [5, 6]. In our case the invariant manifold is the solution manifold $M = V_1 \cap F^{-1}(0)$ and its stable manifold is a full neighborhood of $\bar{v} = 0$. However, the GN-operator T is, in general, not a diffeomorphism. This prevents the application of the abstract results. On the other hand, we have a very special situation here, since the dynamics of T on the invariant manifold is trivial and since the manifold strongly attracts nearby points. This will admit for a rather simple proof.

Without loss of generality we again assume $\bar{v} = 0$ and consequently $w_\infty(0) = 0$. For a fixed $\rho \in (0, 1)$ we consider the space of sequences in $\mathfrak{R}^{m+p}$ which decay like ρ^n

$$S_\rho = \{\{u_n\}_{n \in \mathcal{N}} : \rho^{-n}\|u_n\|_2 \quad \text{is bounded for } n \in \mathcal{N}\}$$

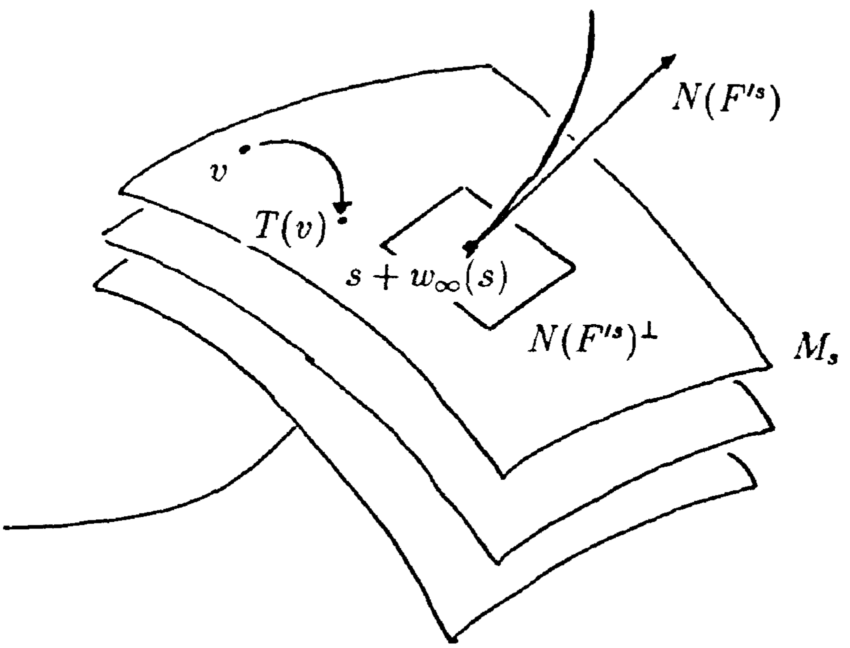

Figure 2 Invariant foliation induced by the Gauss-Newton method.

For brevity, we will write $u_{\mathcal{N}}$ instead of $\{u_n\}_{n\in\mathcal{N}}$. S_ρ becomes a Banach space under the norm

$$\|u_{\mathcal{N}}\|_\rho = \sup\{\rho^{-n}\|u_n\|_2 : n \in \mathcal{N}\}$$

Of course, any norm in $\mathcal{R}^{m+p}$ other than $\|\cdot\|_2$ leads to an equivalent norm in S_ρ.

Finally, we abbreviate the orthogonal projectors (see (2.4))

$$P_s = P(v_\infty(s)) = I - F'(v_\infty(s))^+ F'(v_\infty(s)) \qquad P_0 = P(0) = I - F'^{0+}F'^0$$

Our idea is to apply the implicit function theorem to the following equation

$$\Gamma(u_{\mathcal{N}}, w, s) = 0 \tag{3.2}$$

where the operator

$$\Gamma : S_\rho \times N(P_0) \times N(P_0)^\perp \mapsto S_\rho \times N(P_0)$$

is defined by

$$\Gamma(u_{\mathcal{N}}, w, s) = \begin{pmatrix} \{T(u_n + v_\infty(s)) - v_\infty(s) - u_{n+1}\}_{n\in\mathcal{N}} \\ (I - P_0)u_0 - w \end{pmatrix} \tag{3.3}$$

Suppose that, for given (w, s), we have a solution $u_{\mathcal{N}} = \tilde{u}_{\mathcal{N}}(w, s)$ of (3.2), then $v_n = u_n + v_\infty(s)$ is a Gauss-Newton sequence with $\lim_{n\to\infty} v_n = v_\infty(s)$. The difference of its starting value v_0 to the limit $v_\infty(s)$ has the prescribed component w in the subspace $N(P_0) = N(F'^0)^\perp$. Therefore, we expect the fiber M_s to have the representation

$$M_s = \{v_\infty(s) + \tilde{u}_0(w, s) : w \in W_0 \subset N(F'^0)^\perp\} \tag{3.4}$$

for some neighborhood W_0 of zero.

Theorem 3.1: Under the assumptions of Theorem 2.1 the fibers M_s, $s \in U_1$ are m-dimensional C^k-submanifolds which are positively invariant under the Gauss-Newton iteration. More precisely, for suitable neighborhoods $V_0 \subset \mathcal{R}^{m+p}$, $U_1 \subset N(F'^0)$, $W_0 \subset N(F'^0)^\perp$ these fibers can be represented as in (3.4), where

$$\tilde{u}_0 : W_0 \times U_1 \mapsto \mathcal{R}^{m+p}$$

is a C^k-function satisfying

$$\tilde{u}_0(0, s) = 0 \qquad \text{for } s \in U_1 \tag{3.5}$$

$$R\left(\frac{\partial \tilde{u}_0}{\partial w}(0, s)\right) = N(P_s) \qquad \frac{\partial \tilde{u}_0}{\partial w}(0, 0) = I_{\hookrightarrow} \tag{3.6}$$

Here $I_{\hookrightarrow} : N(P_0) \mapsto \mathcal{R}^{m+p}$ is the canonical embedding.

Proof: From the Lipschitz boundedness of T we find for $u_{\mathcal{N}} \in S_\rho$ sufficiently small

$$\begin{aligned}\rho^{-n}\|T(u_n + v_\infty(s)) - v_\infty(s)\|_2 &= \rho^{-n}\|T(u_n + v_\infty(s)) - T(v_\infty(s))\|_2 \\ &\le C\rho^{-n}\|u_n\|_2 \le C\|u_{\mathcal{N}}\|_\rho\end{aligned}$$

So $\Gamma(u_{\mathcal{N}}, w, s)$ is in fact an element of $S_\rho \times N(P_0)$. Next, we show that the mapping

$$H(u_{\mathcal{N}}, s) = \{T(u_n + v_\infty(s)) - v_\infty(s)\}_{n\in\mathcal{N}}$$

is in $C^k(\Omega, S_\rho)$ for some neighborhood $\Omega \subset S_\rho \times N(P_0)^\perp$, which then implies $\Gamma \in C^k$.

From (1.2) we obtain

$$\begin{aligned} &T(u_n + v_\infty(s)) - v_\infty(s) \\ &= u_n - F'(u_n + v_\infty(s))^+ F(u_n + v_\infty(s)) \\ &= u_n - F'(u_n + v_\infty(s))^+ \int_0^1 F'(v_\infty(s) + tu_n)\,dt\,u_n \end{aligned} \tag{3.7}$$

Since F is in C^{k+1} we see from this formula that H can be written as a composition of C^k mappings and hence is in C^k itself.

For the partial derivative $H_u(u_{\mathcal{N}}, s)$ we find from (2.3)

$$H_u(u_{\mathcal{N}}, s)h_{\mathcal{N}} = \{(P(u_n + v_\infty(s)) - S(u_n + v_\infty(s)))h_n\}_{n\in\mathcal{N}} \tag{3.8}$$

For the operator Γ we obtain $\Gamma(0, 0, 0) = 0$ from (3.3) and from (3.8)

$$\Gamma_u(0, 0, 0)h_{\mathcal{N}} = (\{P_0h_n - h_{n+1}\}_{n\in\mathcal{N}}, (I - P_0)h_0)$$

First, suppose that $h_{\mathcal{N}} \in S_\rho$ is a solution of the homogeneous equation $\Gamma_u(0, 0, 0)h_{\mathcal{N}} = 0$. Then we have $(I - P_0)h_0 = 0$ and $h_n = P_0h_0$ for all $n \geq 0$. But h_n converges to zero, which implies $P_0h_0 = 0$ and hence $h_0 = 0$ and $h_{\mathcal{N}} = 0$.

Next, we solve the inhomogeneous equation

$$\Gamma_u(0, 0, 0)h_{\mathcal{N}} = (r_{\mathcal{N}}, w) \in S_\rho \times N(P_0) \tag{3.9}$$

Because of $\|r_n\|_2 \leq \rho^n$ the limit

$$r_\infty = \sum_{j=0}^{\infty} r_j$$

exists and we determine h_0 from the linear system

$$P_0h_0 = P_0r_\infty \qquad (I - P_0)h_0 = w$$

With this h_0 we set

$$h_n = P_0\left(h_0 - \sum_{j=0}^{n-2} r_j\right) - r_{n-1} \qquad n \geq 1$$

and easily verify $P_0h_n - h_{n+1} = r_n$. Moreover, for $n \geq 1$ we can estimate as follows

$$\begin{aligned} \rho^{-n}\|h_n\|_2 &\leq \rho^{-n}\left[\left\|P_0\left(h_0 - \sum_{j=0}^{n-2} r_j\right)\right\|_2 + \|r_{n-1}\|_2\right] \\ &\leq \rho^{-n}\left[\sum_{j=n-1}^{\infty} \|P_0r_j\|_2 + \|r_{n-1}\|_2\right] \\ &\leq \rho^{-n}\left[\sum_{j=n-1}^{\infty} \rho^j\|r_{\mathcal{N}}\|_\rho + \rho^{n-1}\|r_{\mathcal{N}}\|_\rho\right] \\ &= \rho^{-1}\left[\frac{1}{1-\rho} + 1\right]\|r_{\mathcal{N}}\|_\rho \end{aligned}$$

Combining this with

$$\|h_0\|_2 \leq C(\|r_\infty\|_2 + \|w\|_2) \leq C\left(\frac{1}{1-\rho}\|r_{\mathcal{N}}\|_\rho + \|w\|_2\right)$$

finally yields

$$\|h_{\mathcal{N}}\|_p \leq C(\|r_{\mathcal{N}}\|_p + \|w\|_2)$$

Therefore, $\Gamma_u(0, 0, 0)$ is a linear homeomorphism and we can apply the implicit function theorem.

There exist zero-neighborhoods $\tilde{\Omega} \subset S_p$, $W_0 \subset N(P_0)$ and $U_2 \subset N(P_0)^\perp$ and a function $\tilde{u}_{\mathcal{N}} \in C^k(W_0 \times U_2, \tilde{\Omega})$ which satisfies $\tilde{u}_{\mathcal{N}}(0, 0) = 0$ and for all $w \in W_0$, $s \in U_2$

$$\Gamma(u_{\mathcal{N}}, w, s) = 0 \qquad u_{\mathcal{N}} \in \tilde{\Omega} \Longleftrightarrow u_{\mathcal{N}} = \tilde{u}_{\mathcal{N}}(w, s)$$

Let us first prove (3.5), (3.6). From the definition (3.3) we have

$$\Gamma(0, 0, s) = (\{T(v_\infty(s)) - v_\infty(s)\}_{n \in \mathcal{N}}, 0) = (0, 0)$$

hence by the uniqueness of the implicit function $\tilde{u}_{\mathcal{N}}(0, s) = 0$. Differentiating the implicit equation with respect to w at $w = 0$ we find

$$\Gamma_u(0, 0, s)Y_{\mathcal{N}} + \Gamma_w(0, 0, s) = 0 \qquad \text{for } Y_{\mathcal{N}} = \frac{\partial \tilde{u}_{\mathcal{N}}}{\partial w}(0, s)$$

More explicitly,

$$P_s Y_n - Y_{n+1} = 0 \qquad \text{for } n \geq 0$$
$$(I - P_0)Y_0 - I_{\hookrightarrow} = 0$$

Since $Y_n = P_s Y_0$ converges to zero, we have that Y_0 solves the system

$$P_s Y_0 = 0 \qquad (I - P_0)Y_0 = I_{\hookrightarrow} \tag{3.10}$$

For $\|P_s - P_0\|_2 < 1$ this system has the unique solution

$$Y_0 = (I - (P_0 - P_s))^{-1} I_{\hookrightarrow}$$

Notice that $\|P_0 - P_s\|_2 < 1$ implies that the subspaces $N(P_0)$ and $R(P_s)$ are complementary.

Since $I_{\hookrightarrow}$ has rank m the same holds for Y_0. Therefore, we obtain $R(Y_0) = N(P_s)$. In case $s = 0$ we even have $Y_0 = I_{\hookrightarrow}$.

Next, we consider the C^k mapping

$$\Phi: W_0 \times U_2 \mapsto \mathcal{R}^{m+p} \qquad \Phi(w, s) = v_\infty(s) + \tilde{u}_0(w, s)$$

Using (2.3) we obtain

$$\Phi(0, 0) = 0 \qquad \Phi'(0, 0)(w, s) = I_{\hookrightarrow} w + v'_\infty(0)s = w + s$$

Hence, Φ is a local C^k diffeomorphism and we can choose neighborhoods $\tilde{W}_0 \subset W_0$, $\tilde{U}_1 \subset U_2$ with the following properties. The mapping

$$\Phi: \tilde{W}_0 \times \tilde{U}_1 \mapsto \tilde{V}_0$$

is diffeomorphic, $\tilde{V}_0$ is contained in V_0 and $v_\infty: \tilde{U}_1 \mapsto \mathcal{R}^{m+p}$ is one-to-one (see Theorem 2.1).

We claim that the representation (3.4) holds for the fibers M_s, $s \in \tilde{U}_1$ from (3.1) with $\tilde{V}_0$ in place of V_0. For a given $v \in M_s$ with $v \in \tilde{V}_0$, $s \in \tilde{U}_1$ we find

$T_\infty(v) = v_\infty(s)$ and there exists a unique pair $(w, \tilde{s}) \in \tilde{W}_0 \times \tilde{U}_1$ satisfying $\Phi(w, \tilde{s}) = v$. By our construction of $\bar{u}_N$ we also have $T_\infty(v) = v_\infty(\tilde{s})$. Thus, $v_\infty(s) = v_\infty(\tilde{s})$ and $s = \tilde{s}$, from which our assertion follows.

We can now easily complete the

Proof (Theorem 2.2): At the end of the proof of Theorem 3.1 we found that, for v in some sufficiently small neighborhood,

$$T_\infty(v) = v_\infty(s)$$

where (w, s) satisfies $\Phi(w, s) = v$. If we let Q be the projection from $N(P_0) \times N(P_0)^\perp$ onto its second factor, we may, therefore, write

$$T_\infty = v_\infty \circ Q \circ \Phi^{-1}$$

Thus, the smoothness of T_∞ follows from the smoothness of v_∞ and Φ^{-1}.

Remark 3.2: a. The results of [5, 6] suggest that the local foliation of the regular solution manifold can in fact be generalized to compact submanifolds, e.g., in the case of one parameter to a closed arc of the regular solution curve. However, we have not considered the technical details in the proof of such a result.

b. The implicit function theorem also guarantees the unique solvability of the variational equation

$$\Gamma_u(u_N, w, s)h_N = (r_N, \gamma) \qquad \gamma \in N(P_0),\ r_N \in S_\rho$$

for $\|u_N\|_\rho, \|w\|_2, \|s\|_2$ sufficiently small. More explicitly, this system reads

$$h_{n+1} = T'(u_n + v_\infty(s))h_n - r_n \qquad n \geq 0 \tag{3.11}$$

$$(I - P_0)h_0 = \gamma \tag{3.12}$$

Noticing that $T'(u_n + v_\infty(s)) = P_s + E_n$ where $E_n = O(\rho^n)$ we find from (3.11)

$$h_{n+1} = \prod_{i=0}^{n} (P_s + E_i)h_0 - \sum_{j=0}^{n} \left(\prod_{i=j+1}^{n} (P_s + E_i) \right) r_j$$

Now Lemma 2.3 may be used to conclude that $h_n \to A_\infty h_0 - r_\infty$ as $n \to \infty$, where

$$A_\infty = \prod_{i=0}^{\infty} (P_s + E_i) \qquad r_\infty = \sum_{j=0}^{\infty} \left(\prod_{i=j+1}^{\infty} (P_s + E_i) \right) r_j$$

Since $(I - P_s)A_\infty = 0$ and $(I - P_s)r_\infty = 0$ we can achieve $h_N \in S_\rho$ if we determine h_0 from the linear equations (3.12) and $P_s A_\infty h_0 - P_s r_\infty = 0$. These linear equations are uniquely solvable if s is sufficiently small (see (3.10)). We see that in the direct approach of Section 2 we had to solve variational equations of the general type (3.11), while with the implicit function theorem only a very special case had to be considered, see (3.9).

REFERENCES

1. E. L. Allgower and K. Georg, *Numerical continuation methods, an introduction*, Springer Series in Computational Mathematics 13, Springer, 1990.
2. G. Bader and P. Kunkel, Continuation and collocation for parameter dependent boundary value problems, *SIAM J. Sci. Stat. Comp.* *10* (1989), 72–88.
3. P. Deuflhard and G. Heindl, Affine invariant convergence theorems for Newton's method and extensions to related methods, *SIAM J. Numer. Anal.* *16* (1979), 1–10.

4. P. Deuflhard, B. Fiedler, and P. Kunkel, Efficient numerical pathfollowing beyond critical points, *SIAM J. Numer. Anal. 24* (1984), 912–927.
5. N. Fenichel, Asymptotic stability with rate constants, *Indiana Univ. Math. J. 23* (1974), 1109–1137.
6. N. Fenichel, Asymptotic stability with rate constants II, *Indiana Univ. Math. J. 26* (1977), 81–93.
7. H. Jarausch and W. Mackens, *Numerical treatment of bifurcation branches by adaptive condensation*, in Numerical methods for bifurcation problems, T. Küpper, H. D. Mittelmann, and H. Weber, eds. ISNM 70, Birkhäuser, 1984, 296–309.
8. A. D. Jepson and D. W. Decker, Convergence cones near bifurcation, *SIAM J. Numer. Anal. 23* (1986), 959–975.
9. H. B. Keller, *Numerical solution of bifurcation and nonlinear eigenvalue problems*, in Application of bifurcation theory, P. H. Rabinowitz, ed., Academic Press, 1977.
10. W. C. Rheinboldt, *Numerical analysis of parametrized nonlinear equations*, University of Arkansas, Lecture notes in the Math. Sci. 7, Wiley, 1986.
11. W. C. Rheinboldt, On the computation of multi-dimensional solution manifold of parametrized equations. *Numer. Math. 53* (1988), 165–181.

Received: February 3, 1993
Accepted: May 1993