JOURNAL OF FUNCTIONAL ANALYSIS 33, 165–174 (1979)

An Inverse Function Theorem in Fréchet-Spaces

S. ŁOJASIEWICZ, JR.

Institute of Mathematics, Polish Academy of Sciences, Solskiego 30, 31-027 Crakow, Poland

AND

E. ZEHNDER

Mathematisches Institut de Ruhr-Universität Bochum, 4630 Bochum 1, West Germany

Communicated by Peter D. Lax

Received March 15, 1978

A technical inverse function theorem of Nash-Moser type is proved for maps between Fréchet spaces allowing smoothing operators. A counterexample shows that the growth requirements on the rightinverse of the linearized map needed are minimal.

1. SETUP AND RESULT

It is our aim to prove a technical inverse function theorem for maps between Fréchet-spaces under minimal growth requirements on the rightinverse of the linearized map.

We consider a continuous map $\phi: E \to F$ between two Fréchet-spaces E and F, satisfying

$$\phi(0) = 0.$$

We are looking for conditions on ϕ, which guarantee a local inverse map ψ, satisfying $\phi \circ \psi = id$ in an open neighborhood V of $0 \in F$. If E and F are Banach spaces and if ϕ is of class C^1, then ϕ is a local C^1-diffeomorphism in some neighborhood of 0, provided $\phi'(0)$ is an isomorphism of E onto F. In contrast—the situation in Fréchet-spaces is quite different. Take for example $\phi: f \to \exp(f)$ from $C^\infty(\mathbb{R})$ into itself, or from the Fréchet-space of entire functions into itself. This map is smooth and injective, its derivative at every point is an isomorphism of the Fréchet-space, but the range of ϕ is clearly nowhere dense, hence there is no inverse map on an open set. Therefore in order to find an inverse one has to require additional conditions on E and ϕ. We formulate

0022-1236/79/080165-10$02.00/0

580/33/2-4

next some smoothness and growth conditions on ϕ and the inverse of the linearized map, which allow a local inverse of ϕ and which are, as it turns out, in a certain sense minimal.

Let E be a Fréchet-space with an increasing family of norms defining its topology

$$|x|_n \leqslant |x|_m, \qquad x \in E, \tag{1}$$

if $n \leqslant m$. We assume that E admits smoothing operators S_θ, $\theta \geqslant 1$, that is a one-parameter family of linear maps $S_\theta: E \to E$, such that the following estimates hold:

$$\begin{aligned} |(1 - S_\theta)(x)|_k &\leqslant C\theta^{-(n-k)} |x|_n \\ |S_\theta x|_n &\leqslant C\theta^{n-k} |x|_k \end{aligned} \tag{2}$$

for $x \in E$, $\theta \geqslant 1$ and $0 \leqslant k \leqslant n$. The constant $C > 0$ may depend on k and n. In all the following estimates we let C denote various constants which may always depend on the various norms $|\ |_n$ involved. These quantitative estimates (2), which are crucial for our purpose, single out a restricted class of norms among the increasing families of norms (1) defining the same topology of E. In view of (2) we also have the convexity estimates at our disposal:

$$|x|_l \leqslant C |x|_k^{1-\alpha} |x|_n^\alpha, \qquad l = (1-\alpha)k + \alpha n, \tag{3}$$

$x \in E$, $0 \leqslant k \leqslant n$ and $0 \leqslant \alpha \leqslant 1$. For example, the Fréchet-space $C^\infty(M)$, M a compact manifold, is such a graded Fréchet-space allowing smoothing operators, the norms being the C^k-norms, or the Hölder-norms or the Sobolev-norms. Another example is provided by the Fréchet-space of entire functions dealed with later on.

Let $\phi: E \to F$ be a continuous map between two such graded Fréchet-spaces, locally defined in a neighborhood of 0, assume $\phi(0) = 0$. The growth and smoothness conditions on ϕ, formulated next, will be valid on the following open neighborhood $U \subset E$ of 0, $U = \{x \in E \mid |x|_l < 1\}$, for some fixed $l \geqslant 0$. Keeping in mind, that $\phi(0) = 0$, we require

$$|\phi(x)|_n \leqslant C |x|_{n+d_1} \tag{4}$$

for $x \in U$ and all $n \geqslant 0$ with some fixed $d_1 \geqslant 0$. Every smooth nonlinear partial differential operator on $C^\infty(M)$ for instance satisfies these growth conditions, $l = d_1$ is in this case the order of derivatives involved. We also assume $\phi: (U \subset E) \to F$ to be differentiable in the sense that in F the following limit

$$\lim_{t\to 0} \frac{1}{t} (\phi(x + tv) - \phi(x)) =: \phi'(x)v$$

exists for all $x \in U$ and $v \in E$. This derivative, ϕ': $(U \subset E) \times E \to F$ is required to satisfy:

$$| \phi'(x)v |_n \leqslant C(| x |_{n+d_2} | v |_l + | v |_{n+d_2})$$

for some fixed $d_2 \geqslant 0$ and for all $(x, v) \in U \times E$ and all $n \geqslant 0$. Again, smooth nonlinear partial differential operators on $C^\infty(M)$ for instance meet these conditions (5) as well as the conditions (8) below. By the way, if ϕ': $(U \subset E) \times E \to F$ is continuous, then (4) is a consequence of (5) by means of the Taylorformula. We shall assume, that the map $\phi'(x)$: $E \to F$, $x \in U$, possesses a rightinverse in the following sense. There is a map L: $(U \subset E) \times F \to E$ satisfying

$$\phi'(x) L(x) y = y, \qquad (x, y) \in U \times F \tag{6}$$

and the growth conditions

$$| L(x)y |_n \leqslant C(| x |_{\lambda n+d} | y |_d + | y |_{\lambda n+d}), \tag{7}$$

for all $(x, y) \in (U \subset E) \times F$, all $n \geqslant 0$ with some $d \geqslant 0$ and some $\lambda \geqslant 1$. Of course, if the norms $| \ |_n$ are labelled by integers, $n \in \mathbb{N}$, then λn is always understood to be the integer $[\lambda n]$, where for $a \in \mathbb{R}$, $[a]$ stands for the integer $[a] \leqslant a < [a] + 1$. We point out that, in contrast to the usual growth conditions on L, we allow $\lambda > 1$. In this case the norms on the right hand side of (7) blow up with increasing n by a factor λ. This means the loss of derivatives in solving the linearized problem (6) increases with n. Finally, for the remainder $R(x; v) := \phi(x + v) - \phi(x) - \phi'(x)v$, we require

$$| R(x; v)|_n \leqslant C(| x |_{n+d_2} | v |_l^2 + | v |_l | v |_{n+d_2}), \tag{8}$$

for all x, $x + v \in U$ and $n \geqslant 0$.

It is well known, that in the special case $\lambda = 1$, the above assumptions (4)–(8) on ϕ guarantee a local rightinverse ψ: $(V \subset F) \to U \subset E$ satisfying $\phi(\psi(y)) = y$. The map ψ even belongs to the same category as ϕ, namely it satisfies the growth estimates (4) with a different d_2 however, depending on d in (7), see [3]. Indeed many such inverse function theorems of Moser–Nash type, designed for quite different purposes are available nowadays if $\lambda = 1$, see for instance [2]–[7]. We shall prove the inverse function theorem for the more general case $\lambda > 1$. The proof illustrates the power of the Newton algorithm, introduced in this context by J. Moser [2]. We should mention that the proof requires merely some mild modifications of the crude standard techniques for which we refer in particular to the paper by R. Hamilton [3].

THEOREM. *Assume* ϕ: $(U \subset E) \to F$, $\phi(0) = 0$, *satisfies the growth and smoothness assumptions* (4)–(8) *with* $1 \leqslant \lambda < 2$. *Then there are constants* s_0,

δ and $C > 0$, where $s_0 = O_2((2-\lambda)^{-1})$ and there is a map $\psi: (V \subset F) \to U$, which is defined in the open neighborhood $V := \{y \in F \mid |y|_{s_0} < \delta\}$ and which satisfies $\psi(0) = 0$,

$$\phi(\psi(y)) = y, \qquad y \in V \tag{9}$$

and the estimate $|\psi(y)|_l \leqslant C|y|_{s_0}$.

Moreover, if $\phi'(x): E \to F$, $x \in U$, is a bijection (that is if $L(x)\phi'(x)v = v$, $(x, v) \in U \times E$), then the inverse map $\psi: V \to U$ is unique and it is a continuous map. If in addition ϕ'' exists in U, and if $\phi'': U \times E \times E \to F$ is continuous, then ψ is differentiable on V and $\psi'(y)w = L(\psi(y))w$ for all $(y, w) \in V \times F$. In particular, $\psi': V \times F \to E$ is continuous, if $L: U \times F \to E$ is continuous.

The above existence statement, and this is our main point, is optimal in the sense, that it does not hold true anymore for $\lambda \geqslant 2$. A counterexample for $\lambda = 2$ is described below.

Remarks. (a) The above inverse map ψ does in general not allow estimates like (4) anymore, if $\lambda > 1$.

(b) The existence statement can be refined as follows. Replace the estimates (4), (5) and (8) by the weaker growth conditions admitting on the right hand side instead of the n-norms the $\mu \cdot n$-norms for some $\mu > 1$. For instance replace (4) by

$$|\phi(x)|_n \leqslant C|x|_{\mu n + d_1}. \tag{10}$$

Then the proof below still establishes a rightinverse as long as $\lambda \cdot \mu < 2$. If however one knows in addition, that $|L(x)\phi(x)|_n \leqslant C|x|_{\lambda n + d}$, then $\lambda < 2$ is the only restriction. Observe that in view of (3) the estimates $|\phi(x)|_n \leqslant C(|x|_{n+d_3})^\mu$ lead to (10).

(c) The restriction $\lambda < 2$ of the statement is related to the Newton iteration method, whose formally quadratic convergence enters crucially the construction of ψ.

2. Proof of the Theorem

To simplify the notation we may assume

$$d_1 = d_2 = 0 \qquad \text{and} \qquad l = d, \tag{11}$$

by relabeling the norms and increasing the value of $d > 0$ in (7). Recall, the l-norm describes the neighborhood $U := \{x \in E \mid |x|_l < 1\} \subset E$ in which the estimates are valid. We now write down the modified Newton algorithm for the sequence $x_p \in E$, $p \geqslant 0$, which as we shall demonstrate will converge to a solution x of $\phi(x) = y$ for small enough $y \in F$. (Recall $\phi(0) = 0$). Starting with $x_0 = 0$, we put for $p \geqslant 0$

$$\begin{aligned} x_{p+1} &= x_p + \Delta x_p \\ \Delta x_p &:= S_{\theta_p} L(x_p)\, z_p\,, \qquad z_p = y - \phi(x_p). \end{aligned} \tag{12}$$

The sequence θ_p in the smoothing operators are defined by $\theta_p := 2^{(\tau^p)}$, with $1 \leqslant \lambda < \tau < 2$. We have used the assumption $\lambda < 2$. We fix $\tau := 2^{-1}(\lambda + 2)$ and observe

$$\theta_{p+1} = \theta_p{}^{\tau}. \tag{13}$$

We next establish bounds for the norms $| x_p |_n$, $n \geqslant 0$, which will be valid as long as $| x_p |_d < 1$ and $| y |_d \leqslant 1$.

LEMMA 1. *For every $n \geqslant d$, there exists a constant K, $K = K_n$, such that for all $y \in E$ with $| y |_d \leqslant 1$, we have*

$$| x_p |_n \leqslant K \theta_p^{L(n)} \, | y |_n \tag{14}$$

for all $p \geqslant 0$, as long as $| x_p |_d < 1$.

$$L(n) := n \frac{1}{\lambda} \left(\frac{\lambda - 1}{\tau - 1} \right) + l_0\,, \qquad \text{with} \quad l_0 = \frac{1}{\lambda} \frac{d + \lambda}{\tau - 1}\,.$$

Proof. Let $n \geqslant d$ and assume $| y |_d \leqslant 1$ and $| x_j |_d < 1$, $j = 1, 2, \ldots, p$. By (4) and (11) we can estimate $z_j = y - \phi(x_j)$ as follows: $| z_j |_n \leqslant | y |_n + | \phi(x_j) |_n \leqslant C(| y |_n + | x_j |_n)$. As long as $| y |_d \leqslant 1$ and $| x_j |_d < 1$, we have $| z_j |_d \leqslant C$. Define k (depending on n) by

$$n = \lambda(n - k) + d.$$

Respectively $k := n - [(n - d)/\lambda]$ in case the norms are labelled by integers. We estimate $x_j := S_{\theta_j} L(x_j)\, z_j$ by means of (2) and (7):

$$\begin{aligned} | \Delta x_j |_n &\leqslant C \theta_j{}^k \, | L(x_j)\, z_j |_{n-k} \\ &\leqslant C \theta_j{}^k (| x_j |_n \, | z_j |_d + | z_j |_n) \\ &\leqslant C \theta_j{}^k (| x_j |_n + | y |_n), \end{aligned}$$

where $C \geqslant 1$ is a constant independent of j. Repeated use of these estimates yields, with $\theta_j = 2^{(\tau^j)}$,

$$| x_{p+1} |_n \leqslant (p + 1)\, C^{p+1} 2^{k(\tau^{p+1} - 1)/(\tau - 1)} \, | y |_n \,.$$

But by definition $(\tau - 1)L(n) > k$, hence there is a constant $K > 0$ such that

$$(p+1)\, C^{p+1} 2^{k\tau^{p+1}/(\tau-1)} \leqslant K 2^{L(n)\tau^{p+1}} = K\theta_{p+1}^{L(n)}$$

for all $p \geqslant 0$. Therefore $| x_{p+1} |_n \leqslant K\theta_{p+1}^{L(n)} \, | y |_n$ as we wanted to prove. ∎

Now the low norms are estimated carefully, in order to prove that for $y \in F$ sufficiently small, the sequence $| z_p |_d$ converges to zero:

LEMMA 2. *There exist M, s_0, $\delta > 0$, such that if $y \in F$ satisfies $| y |_{s_0} \leqslant \delta$ we have*

$$| z_p |_d \leqslant M\theta_p^{-\mu} \, | y |_{s_0}, \qquad \mu := \frac{2+\tau}{2-\tau}\, d \tag{15}$$

for all $p \geqslant 0$ as long as $| x_p |_d < 1$. We find $s_0 = O_2((2-\lambda)^{-1})$.

Proof. Induction in p. By definition $x_{p+1} = x_p + \Delta x_p$, hence $\phi(x_{p+1}) = \phi(x_p) + \phi'(x_p)\,\Delta x_p + R(x_p, \Delta x_p)$ and with (6) and (12) we find $z_{p+1} = y - \phi(x_{p+1}) = \phi'(x_p)(1 - S_{\theta_p}) L(x_p)\, z_p - R(x_p, \Delta x_p)$. We estimate the first term, using (5) with $d_2 = 0$, $| x_p |_d \leqslant 1$, using (2) and (7) and abbreviating $s_0 := \lambda s + d$:

$$\begin{aligned} &| \phi(x_p)(1 - S_{\theta_p})\, L(x_p)\, z_p |_d \\ &\quad\leqslant C\, |(1 - S_{\theta_p})\, L(x_p)\, z_p |_d \\ &\quad\leqslant C\theta_p^{-(s-d)}\, | L(x_p)\, z_p |_s \\ &\quad\leqslant C\theta_p^{-(s-d)}(| x_p |_{s_0}\, | z_p |_d + | z_p |_{s_0}). \end{aligned}$$

Since $| x_p |_d \leqslant 1$ and $| y |_d \leqslant 1$, we have $| z_p |_d \leqslant C$, from (4) we conclude $| z_p |_{s_0} \leqslant | y |_{s_0} + C\, | x_p |_{s_0}$. Hence applying Lemma 1 to $| x_p |_{s_0}$, we can estimate further

$$\leqslant C\theta_p^{-(s-d)}(\theta_p^{L(s_0)} + 1)\, | y |_{s_0}.$$

By definition of $L(s_0)$, $s_0 = \lambda s + d$, we have $s - d - L(s_0) = s(1 - (\lambda-1)/(\tau-1)) - d(\lambda - 1)/\lambda(\tau - 1) - l_0 - d$. Because of $\lambda < \tau$ we therefore can pick $s > 0$ such that

$$s - d - L(s_0) \geqslant \mu\tau.$$

Recalling $\tau = 2^{-1}(\lambda + 2)$ we find $s_0 = O_2((2-\lambda)^{-1})$. Clearly $(s - d) \geqslant \mu\tau$, hence

$$| \phi'(x_p)(1 - S_{\theta_p})\, L(x_p)\, z_p |_d \leqslant C\theta_{p+1}^{-\mu}\, | y |_{s_0},$$

the constant C being independent of p. In order to estimate the second term of z_{p+1}, we use $|\Delta x_p|_d = |S_{\theta_p} L(x_p) z_p|_d \leqslant C\theta_p{}^d |L(x_p) z_p|_0 \leqslant C\theta_p{}^d |z_p|_d$ which follows by (2) and (7) with $|x_p|_d \leqslant 1$. Therefore in view of (8) with $d_2 = 0$, we get

$$|R(x_p ; \Delta x_p)|_d \leqslant C |\Delta x_p|_d^2 \leqslant C\theta_p^{2d} |z_p|_d^2 .$$

We finally use the induction hypotheses, $|z_p|_d \leqslant M\theta_p^{-\mu} |y|_{s_0}$, and $\theta_p^{2d}\theta_p^{-2\mu} \leqslant \theta_{p+1}^{-\mu}$ (by definition of μ) to conclude $|R(x_p ; \Delta x_p)|_d \leqslant CM^2\theta_{p+1}^{-\mu} |y|_{s_0}^2$. Summarizing we have shown so far, that

$$|z_{p+1}|_d \leqslant C(1 + M^2 |y|_{s_0}) \theta_{p+1}^{-\mu} |y|_{s_0} ,$$

for some C independent of p. We may assume $M > C$, with C as in the previous estimate. Define $\delta := \min\{1, (M - C) C^{-1}M^{-2}\}$. Therefore, if we restrict $|y|_{s_0} \leqslant \delta$, we find $|z_{p+1}|_d \leqslant M\theta_{p+1}^{-\mu} |y|_{s_0}$ which completes the induction. ∎

From Lemma 2 we deduce inductively for the sequence $x_p \in E$, $p \geqslant 0$, that $|x_p|_d < 1$, if $|y|_{s_0} < \delta$ for δ sufficiently small. Indeed, if $|x_j|_d < 1$ for $0 \leqslant j \leqslant p$, we know $|\Delta x_j|_d \leqslant C\theta_j{}^d |z_j|_d$, so by Lemma 2, $|\Delta x_j|_d \leqslant C\theta_j^{-(\mu-d)} |y|_{s_0}$ for some C independent of j. Therefore $|x_{p+1}|_d \leqslant \sum_{j=0}^{p} |\Delta x_j|_d \leqslant C(\sum_{j=0}^{\infty} \theta_j^{-(\mu-d)}) |y|_{s_0}$. But $\mu > d$, and so we get

$$|x_{p+1}|_d \leqslant C |y|_{s_0} < C\delta < 1, \tag{16}$$

choosing δ smaller, if necessary. Hence (16) holds true for all $p \geqslant 0$. In the following we shall always assume $|y|_{s_0} < \delta$ with this particular choice of δ. Lemma 1 and Lemma 2 are then valid for all $p \geqslant 0$.

The trick now is to improve the estimate (15) to any power of θ_p at the cost of course of arbitrary high norms of y.

LEMMA 3. *For every $a \geqslant 0$ there are constants $C = C(a) > 0$ and $n(a) > 0$ such that*

$$|z_p|_d \leqslant C |y|_{n(a)} \theta_p^{-a} \tag{17}$$

for all $p \geqslant 0$ and all $y \in E$ with $|y|_{s_0} < \delta$.

Proof. The statement is obviously true for $0 \leqslant a \leqslant \mu$ (Lemma 2). Let $a \geqslant \mu$ and assume the statement to hold true for this a, we shall prove it for $a + d$. We know

$$|z_{p+1}|_d \leqslant C |(1 - S_{\theta_p}) L(x_p) z_p|_d + |R(x_p , \Delta x_p)|_d .$$

Proceeding as in Lemma 2 we pick n_0, $n_0 = \lambda n + d$ with $n - d - L(n_0) \geqslant \tau(a + d)$, such that

$$|(1 - S_\theta) L(x_p) z_p|_d \leqslant C\theta_{p+1}^{-(a+d)} |y|_{n_0}.$$

On the other hand, $| R(x_p, \Delta x_p)|_d \leqslant C | \Delta x_p|_d^2 \leqslant C\theta_p^{2d} | z_p|_d^2$ can be estimated by the induction assumption:

$$| R(x_p, \Delta x_p)|_d \leqslant C | y |_{n(a)}^2 \theta_p^{-2(a-d)}.$$

By the convexity estimate (3) and using $|y|_0 \leqslant 1$, we estimate $|y|_{n(a)}^2 \leqslant C |y|_{2n(a)}$. Also, $2a - 2d \geqslant \tau(a + d)$, if $a \geqslant (2 - \tau)^{-1}(\tau + 2)d$, hence in particular if $a \geqslant \mu$ by our choice of μ. Therefore $| R(x_p, \Delta x_p)|_d \leqslant C |y|_{2n(a)} \theta_{p+1}^{-(a+d)}$. We proved, $| z_{p+1}|_d \leqslant C |y|_{n(a+d)} \theta_{p+1}^{-(a+d)}$ for all $p \geqslant 0$, with $n(a + d) := \max\{n_0, 2n(a)\}$. Trivially $| z_0 |_d \leqslant |y|_d \leqslant C |y|_{n(a+d)} \theta_0^{-(a+d)}$, by changing the constant if necessary. This proves the lemma. ∎

From Lemma 3 and Lemma 1 we get together with the convexity estimate (3) the improved estimates for the higher norms:

LEMMA 4. *For every $n \geqslant 0$ and every $b \geqslant 0$ there are constants $C = C(n, b) > 0$ and $\sigma(n, b) > 0$, such that for all $y \in F$ with $|y|_{s_0} < \delta$:*

$$| \Delta x_p|_n \leqslant C |y|_{\sigma(n,b)} \theta_p^{-b}$$

$$| z_p|_n \leqslant C |y|_{\sigma(n,b)} \theta_p^{-b}$$

for all $p \geqslant 0$.

Proof. By (3), for $m > n$

$$| \Delta x_p|_n \leqslant C | \Delta x_p|_0^{(1-n/m)} | \Delta x_p|_m^{n/m}.$$

From Lemma 3, we conclude for every $a \geqslant 0$, $| \Delta x_p|_0 \leqslant C |y|_{n(a)} \theta_p^{-a}$. Lemma 1 gives $| \Delta x_p|_m \leqslant | x_{p+1}|_m + | x_p|_m \leqslant C\theta_{p+1}^{L(m)} |y|_m = C\theta_p^{\tau L(m)} |y|_m$. If $b > 0$ is given, choose $m = 2n$, $a = 2b + \tau L(2n)$, and get $| \Delta x_p|_n \leqslant C |y|_{\sigma(n,b)} \theta_p^{-b}$, with $\sigma(n, b) = \max\{n(a), 2n\}$. Similarly for $| z_p|_n$, which proves Lemma 4. ∎

We are in business. From Lemma 5 we conclude, that x_p is a Cauchy-Sequence in E, therefore $\lim_{n\to\infty} x_n =: x \in E$. On the other hand $z_n := y - \phi(x_n) \to 0$ in F. Since ϕ is continuous, $\lim_{p\to\infty} \phi(x_p) = \phi(x)$ and therefore $\phi(x) = y$. Denoting with $x := \psi(y)$ this solution, we have established the existence part of the theorem. The estimate $| \psi(y)|_d \leqslant C |y|_{s_0}$ follows from (16). The moreover part follows with our estimates by standard manipulations and will be omitted.

3. A Counterexample for $\lambda = 2$

Let E be the Fréchet-space of entire functions $x(z) = \sum_{n\geqslant 0} x_n z^n$, $z \in C$, or equivalently the sequence space $x = (x_n)_{n\geqslant 0}$ with $\lim_{n\to\infty} | x_n |^{1/n} = 0$. The norms $\| x \|_r$ defining the topology of E being given by

$$\| x \|_r^2 := \sum_{n\geqslant 0} | x_n |^2 r^{2n}, \qquad r \geqslant 1.$$

This graded Fréchet-space allows smoothing operators: define for $j \in Z$, $j \geqslant 0$ the truncation operators $T_j \colon E \to E$ by $(T_j(x))_n = x_n$ if $0 \leqslant n \leqslant j$ and $(T_j(x))_n = 0$ for $n \geqslant j + 1$. It follows immediately for all $j \geqslant 0$ and $1 \leqslant \rho \leqslant r$ and $x \in E$:

$$\| T_j(x)\|_r \leqslant \left(\frac{r}{\rho}\right)^j \| x \|_\rho$$

$$\|(1 - T_j)(x)\|_\rho \leqslant \left(\frac{\rho}{r}\right)^{j+1} \| x \|_r .$$

Therefore, defining the norms $| \ |_n$, $n \in \mathbb{R}$, $n \geqslant 0$ and the operators $S_\theta \colon E \to E$, $\theta \geqslant 1$ by

$$| x |_n := \| x \|_{e^n} , \qquad S_\theta := T_{[\log \theta]} ,$$

where $[\log \theta]$ stands for the integer $[\log \theta] \leqslant \log \theta \leqslant [\log \theta] + 1$, the required estimates (1) and (2) are immediate.

The map $\phi \colon E \to E$ is then defined as follows:

$$\phi(x) = A(x) + \tfrac{1}{4}B(x, x), \qquad x \in E.$$

A being the linear map defined as $y = A(x)$ with $y_0 = x_0$, $y_{2^n} = x_{2^{n+1}}$ for $n \geqslant 0$, $y_{2^{n+1}} = x_{2^{n-1}}$ for $n \geqslant 1$, and $y_k = x_{k-1}$ otherwise. $z = B(x, y)$, $x, y \in E$, is given by $z_{2^n} = x_{2^n} y_{2^n}$, $n \geqslant 0$ and $z_k = 0$ otherwise.

We next verify that this ϕ satisfies all our assumptions (4)–(8) with $d_1 = d_2 = l = d = 0$ and $\lambda = 2$. Clearly $\| A(x)\|_r \leqslant r^2 \| x \|_r$ and $\| B(x, y)\|_r \leqslant \| x \|_r \| y \|_1$ (resp. $\leqslant \| x \|_1 \| y \|_r$). Hence, if $\| x \|_1 < 1$ we conclude $\| \phi(x)\|_r \leqslant (r^2 + \frac{1}{4}) \| x \|_r$ for $r \geqslant 1$, and so $| \phi(x)|_n \leqslant (e^{2n} + \frac{1}{4}) | x |_n$ for all $x \in E$ with $| x |_0 < 1$ and all $n \geqslant 0$. Similarly we find for the derivative

$$\phi'(x)v = A(v) + \tfrac{1}{2}B(x, v)$$

the estimate $| \phi'(x)v |_n \leqslant (e^{2n} + \frac{1}{2}) | v |_n$ for all $n \geqslant 0$ and all $(x, v) \in E \times E$ with $| x |_0 < 1$. The remainder, $R(x; v) = \frac{1}{4}B(v, v)$, satisfies $| R(x; v)|_n \leqslant \frac{1}{4} | v |_n | v |_0$ for all $(x, v) \in E \times E$. A direct computation shows that $\phi'(x) \colon E \to E$

is a bijection, hence $\phi'(x)$ being continuous is an isomorphism by the closed graph theorem. We do not need this information, what we need are growth estimates of the inverse $L(x; y)$ in the open set $|x|_0 < 1$ only where we can apply the contraction principle. For $y \in E$ given, we have to solve $y = \phi'(x)v = A(v) + \frac{1}{2}B(x, v)$. From $A^{-1}(y) = v + \frac{1}{2}A^{-1}(B(x, v))$ we conclude with $\| A^{-1}(y)\|_r \leqslant \| y \|_{r^2}$, that $\| v \|_r \leqslant \| y \|_{r^2} + \frac{1}{2} \| x \|_{r^2} \| v \|_1$. Therefore, if $r = 1$ and $\| x \|_1 < 1$, we find $\| v \|_1 \leqslant \| y \|_1 + \frac{1}{2} \| v \|_1$ and so $\| v \|_1 \leqslant 2 \| y \|_1$. Hence $\| v \|_r \leqslant \| y \|_{r^2} + \| x \|_{r^2} \| y \|_1$ and consequently $| v |_n := | L(x, y)|_n \leqslant | y |_{2n} + | x |_{2n} | y |_0$, for all $(x, y) \in E \times E$ with $| x |_0 < 1$. We have checked that ϕ meets our assumptions (4)–(8) with $\lambda = 2$.

There is no local inverse ψ of the map ϕ in any open neighborhood of 0, since in every open neighborhood U of 0 we find a $y \in U$ with $y \notin \phi(E)$. Indeed, pick $y = (y_n)_{n \geqslant 0}$, $y_3 = 4\epsilon > 0$ and $y_n = 0$ otherwise. A simple direct computation yields a unique sequence $x = (x_n)_{n \geqslant 0}$ formally satisfying $\phi(x) = y$. It is given by $x_{2m} = (-4)(-\epsilon)^{(2^m)}$, for $m \geqslant 0$ and $x_n = 0$ otherwise. Clearly $x \notin E$ as $\lim_{m \to \infty} | x_{2m} |^{(1/2^m)} = \epsilon \neq 0$. A little more work shows, that the smooth injective map ϕ, whose differential $\phi'(x)$, $x \in E$ is an isomorphism, has a nowhere dense range, compare also [1].

References

1. S. Lojasiewicz, Jr., An example of a continuous injective polynomial map with nowhere dense range whose differential at each point is an isomorphism, *Bull. Acad. Polon. Sci.* **24**, No. 12 (1976), 1109–1111.
2. J. Moser, A new technique for the construction of solutions of nonlinear differential equations, *Proc. Nat. Acad. Sci. USA* **47** (1961), 1824–1831.
3. R. S. Hamilton, The inverse function theorem of Nash and Moser, Preprint, Cornell Univ., 1974.
4. F. Sergeraert, Un théorème de fonctions implicites sur certains espaces de Fréchet et quelques applications, *Ann. Sci. Ecole Norm. Sup. Sér. 4* **5** (1972), 599–660.
5. H. Jacobowitz, Implicit function theorems and isometric imbeddings, *Ann. of Math.* **95** (1972), 191–225.
6. L. Hörmander, The boundary problems of physical geodesy, *Arch. Rational Mech. Anal.* **62**, No. 1 (1976), 1–52.
7. E. Zehnder, Generalized implicit function theorems with applications to some small divisor problems, I, *Comm. Pure Appl. Math.* **28** (1975), 91–140.