HAMILTON'S INVERSE FUNCTION THEOREM: PROOF OF THE CONVERGENCE OF THE ITERATION SCHEME

JEAN-CHRISTOPHE YOCCOZ

1. The setting

- E, F are two tame Frechet spaces, $S(t), t \geqslant 1$ are the corresponding smoothing operators;
- r_0 is a positive integer;
- f is a map defined on the ball $B = \{||x||_{r_0} < 1\}$ of E, with values in F satisfying $f(0) = 0$;
- f is C^2 tame and satisfies, for $x \in B$, $y \in E$, $r \geqslant 0$:

$$\begin{aligned} ||Df(x,y)||_r &\leqslant C_r(||x||_{r+r_0}||y||_{r_0} + ||y||_{r+r_0}), \\ ||D^2f(x,y,y)||_r &\leqslant C_r(||x||_{r+r_0}||y||^2_{r_0} + ||y||_{r_0}||y||_{r+r_0}). \end{aligned}$$

- There exists a continuous tame inverse of Df denoted by $L : B \times F \to E$ which satisfies, for $x \in B$, $z \in F$, $r \geqslant 0$:

$$||L(x,z)||_r \leqslant C_r(||x||_{r+r_0}||z||_{r_0} + ||z||_{r+r_0}).$$

2. The iteration scheme

Let $\varepsilon^* > 0$ and r_1 a positive integer to be chosen later (Actually, we will have $r_1 = 17r_0$). We start with $y \in F$ in the ball $B^* := \{||y||_{r_1} < \varepsilon^*\}$ and will construct $x \in B$ such that $f(x) = y$. Set $t_n := \exp(\frac{3}{2})^n$ for $n \geqslant 0$.

Let $x_0 = 0$; as long as $||x_n||_{2r_0} + ||f(x_n)||_{2r_0} < 1$, we define inductively

- $e_n = y - f(x_n)$;
- $\widetilde{\delta}_n = L(x_n, e_n)$;
- $\delta_n = S(t_n)\widetilde{\delta}_n$;
- $x_{n+1} = x_n + \delta_n$.

This is Newton's algorithm with the smoothing from $\widetilde{\delta}_n$ to δ_n added. We will prove that, if B^* is small enough, x_n is defined for all n and converge to a solution x of $f(x) = y$.

3. The basic estimates

The properties of L give, for $r \geqslant 0$

$$||\widetilde{\delta}_n||_r \leqslant C_r(||x_n||_{r+r_0}||e_n||_{r_0} + ||e_n||_{r+r_0}). \tag{1}$$

We assume that $r_1 \geqslant 2r_0$ and $\varepsilon^* < 1$. Then, the property $||f(x_n)||_{2r_0} < 1$ implies

$$||e_n||_{2r_0} < 2. \tag{2}$$

Then, as we have also $||x_n||_{2r_0} < 1$, we get from the previous inequality

$$||\widetilde{\delta}_n||_{r_0} \leqslant C. \tag{3}$$

The next estimate comes directly from the properties of the smoothing operators: for any $r' \geqslant r$, we have

$$||\delta_n||_{r'} \leqslant C_{r,r'} t_n^{r'-r} ||\widetilde{\delta}_n||_r. \tag{4}$$

Thus, we have

$$||\delta_n||_{r_0} \leqslant C. \tag{5}$$

To estimate e_{n+1}, we will use Taylor's formula at order 1 or 2. On one side

$$f(x_n + \delta_n) = f(x_n) + \int_0^1 Df(x_n + u\delta_n, \delta_n) du$$

which gives, in view of the properties of Df, for $r \geqslant 0$

$$||e_{n+1}||_r \leqslant ||e_n||_r + C_r(||\delta_n||_{r_0}(||x_n||_{r+r_0} + ||\delta_n||_{r+r_0}) + ||\delta_n||_{r+r_0}).$$

Using $||\delta_n||_{r_0} \leqslant C$, we get

$$||e_{n+1}||_r \leqslant ||e_n||_r + C'_r(||x_n||_{r+r_0} + ||\delta_n||_{r+r_0}). \tag{6}$$

This crude estimate will be useful in complement to the one coming from

$$f(x_n + \delta_n) = f(x_n) + Df(x_n, \delta_n) + \int_0^1 (1-u) D^2 f(x_n + u\delta_n, \delta_n, \delta_n) du.$$

Here, from the definition of δ_n we get

$$e_{n+1} = Df(x_n, (1-S(t_n))L(x_n, e_n)) - \int_0^1 (1-u) D^2 f(x_n + u\delta_n, \delta_n, \delta_n) du = e'_{n+1} - e''_{n+1}$$

The properties of $D^2 f$ give, for any $r \geqslant 0$

$$||e''_{n+1}||_r \leqslant C_r(||\delta_n||_{r_0}^2(||x_n||_{r+r_0} + ||\delta_n||_{r+r_0}) + ||\delta_n||_{r+r_0}||\delta_n||_{r_0}).$$

Using again $||\delta_n||_{r_0} \leqslant C$, we get

$$||e''_{n+1}||_r \leqslant C''_r(||\delta_n||_{r_0}^2||x_n||_{r+r_0} + ||\delta_n||_{r+r_0}||\delta_n||_{r_0}). \tag{7}$$

To estimate e'_{n+1}, we will use the approximation property of $S(t_n)$.

4. Convergence of the iteration scheme

Lemma 4.1. *Let $r \geqslant r_0$. There exists a constant $A = A(r)$ such that, for every $n \geqslant 0$ such that x_{n+1} is defined, one has*

$$||\widetilde{\delta}_n||_{r-r_0} \leqslant A\, t_n^{5r_0} ||y||_r, \tag{8}$$

$$||\delta_n||_{r+r_0} \leqslant A\, t_n^{7r_0} ||y||_r, \tag{9}$$

$$||x_{n+1}||_{r+r_0} \leqslant A\, t_n^{7r_0} ||y||_r, \tag{10}$$

$$||e_{n+1}||_r \leqslant A\, t_n^{7r_0} ||y||_r \tag{11}$$

Proof. We have $x_0 = 0$ and $e_0 = y$. For $n \geqslant 0$, we write

$$\begin{aligned} ||\widetilde{\delta}_n||_{r-r_0} &= A_1(n,r)\, t_n^{5r_0}||y||_r, \\ ||\delta_n||_{r+r_0} &= A_2(n,r)\, t_n^{7r_0}||y||_r, \\ ||x_{n+1}||_{r+r_0} &= A_3(n,r)\, t_n^{7r_0}||y||_r, \\ ||e_{n+1}||_r &= A_4(n,r)\, t_n^{7r_0}||y||_r. \end{aligned}$$

We write C_r for various constants depending only on r.

From (1), we have $A_1(0,r) \leqslant C_r$. From (4), we have

$$A_2(n,r) \leqslant C_r\ A_1(n,r) \tag{12}$$

Next we have $A_3(0,r) = A_2(0,r)$ and, for $n > 0$

$$A_3(n,r) \leqslant A_2(n,r) + A_3(n-1,r)(t_{n-1}t_n^{-1})^{7r_0}, \tag{13}$$

with $t_{n-1}t_n^{-1} = t_n^{-1/3}$. From (6), we have also

$$A_4(n,r) \leqslant C_r(A_2(n,r) + t_n^{-7r_0/3}(A_3(n-1,r) + A_4(n-1,r))). \tag{14}$$

Finally, from (1), we get, for $n > 0$

$$A_1(n,r) \leqslant C_r t_n^{-7r_0/3}(A_3(n-1,r) + A_4(n-1,r)). \tag{15}$$

Whatever the values of the constants C_r, the inequalities (12)-(15) imply that the sequences $A_i(n,r)$, $1 \leqslant i \leqslant 4$, are bounded from above by a constant depending only on r. □

Lemma 4.2. *There exists a constant A^*, and, for any $r \geqslant 8\,r_0$, a constant $A^*(r)$, such that, for all $n \geqslant 0$ such that x_{n+1} is defined, one has*

$$||e_{n+1}||_{r_0} \leqslant A^* t_n^{3r_0}||e_n||_{r_0}^2 + A^*(r) t_n^{8r_0-r}||y||_r. \tag{16}$$

Proof. Write $e_{n+1} = e'_{n+1} - e''_{n+1}$ as above. We have, from (7), (4), (1)

$$||e''_{n+1}||_{r_0} \leqslant C||\delta_n||_{r_0}||\delta_n||_{2r_0} \leqslant C' t_n^{3r_0}||\widetilde{\delta}_n||_0^2 \leqslant A^* t_n^{3r_0}||e_n||_{r_0}^2$$

On the other hand, from the properties of Df , $S(t_n)$ and L, we have

$$\begin{aligned} ||e'_{n+1}||_{r_0} &\leqslant C||(1-S(t_n))L(x_n,e_n)||_{2r_0} \\ &\leqslant C'(r) t_n^{3r_0-r}||L(x_n,e_n)||_{r-r_0} \\ &\leqslant C''(r) t_n^{3r_0-r}(||x_n||_r + ||e_n||_r) \\ &\leqslant A^*(r) t_n^{8r_0-r}||y||_r, \end{aligned}$$

where the last inequality follows from Lemma 4.1. The proof of the lemma is complete. □

We now take $r_1 := 17r_0$.

Lemma 4.3. *There exists a constant C^* such that, if $||y||_{r_1}$ is small enough, we have*

$$||e_n||_{r_0} \leqslant C^* t_n^{-6r_0}||y||_{r_1}. \tag{17}$$

Proof. This clearly holds for $n = 0$ if $C^* \geqslant 1$. We then proceed by induction, using the previous lemma with $r = r_1$:

$$\begin{aligned} ||e_{n+1}||_{r_0} &\leqslant A^* t_n^{3r_0} ||e_n||_{r_0}^2 + A^*(r_1) t_n^{-9r_0} ||y||_{r_1} \\ &\leqslant A^*(C^*)^2 t_n^{-9r_0} ||y||_{r_1}^2 + A^*(r_1) t_n^{-9r_0} ||y||_{r_1} \\ &= C^* \, t_{n+1}^{-6r_0} ||y||_{r_1} (A^* C^* ||y||_{r_1} + \frac{A^*(r_1)}{C^*}). \end{aligned}$$

It is therefore sufficient to take $C^* \geqslant 2A^*(r_1)$ and then $||y||_{r_1} < \frac{1}{2A^*C^*}$. □

Lemma 4.4. *If $||y||_{r_1}$ is small enough, the sequence (x_n) is defined for all $n \geqslant 0$ and we have, with appropriate constants C and any $0 \leqslant k \leqslant 5$*

$$||\widetilde{\delta}_n||_0 \leqslant C \, t_n^{-6r_0} ||y||_{r_1}, \tag{18}$$

$$||\delta_n||_{kr_0} \leqslant C \, t_n^{(k-6)r_0} ||y||_{r_1}, \tag{19}$$

$$||e_n||_{2r_0} \leqslant C \, t_n^{-5r_0} ||y||_{r_1}. \tag{20}$$

Proof. The first inequality follows from (1) and Lemma 4.3. Then the second inequality follows from (4). This proves in particular that $||x_n||_{2r_0}$ remains very small. From Hadamard interpolation inequalities and Lemmas 4.1 and 4.3, we have

$$||e_n||_{2r_0} \leqslant C \, ||e_n||_{r_0}^{15/16} ||e_n||_{r_1}^{1/16} \leqslant C \, t_n^{-5r_0} ||y||_{r_1}.$$

Therefore, $||f(x_n)||_{2r_0} = ||y - e_n||_{2r_0}$ remains also very small. This proves that the sequence (x_n) is defined for all $n \geqslant 0$. □

We now assume that $||y||_{r_1} < \varepsilon^*$, with ε^* small enough so that the conclusions of the last lemma are satisfied.

Lemma 4.5. *The sequence (x_n) converge in E to a limit x such that $f(x) = y$.*

Proof. Let $r \geqslant r_0$. We have

$$||\delta_n||_{3r} \leqslant A \, t_n^{7r_0} ||y||_{3r-r_0}, \quad A = A(3r - r_0),$$

from Lemma 4.1 and

$$||\delta_n||_0 \leqslant C t_n^{-6r_0} ||y||_{r_1}$$

from Lemma 4.4, hence

$$||\delta_n||_r \leqslant C_r t_n^{-5r_0/3} ||y||_{3r-r_0}^{1/3} ||y||_{r_1}^{2/3}$$

by interpolation. This proves the convergence of (x_n) to a limit x. A similar estimate is obtained from $||e_n||_r$, which proves that (e_n) converge to 0 in F. As f is continuous, this proves that $f(x) = y$. □

We have thus constructed a map $g : B^* \to B$ which satisfies $f \circ g(y) = y$ for $y \in B^*$.

It remains to prove that $g \circ f(x) = x$ for x close to 0, that g is continuous and tame, and that g is Gateaux differentiable with $Dg(y, v) = L(g(y), v) \ldots$ This is left to the reader.