Notes for the course Analyse 2.

Horia Cornean[1], 7/03/2012.

1 Banach's fixed point theorem

Definition 1.1. *Let (X, d) be a metric space. A map $F : X \to X$ is called a contraction if there exists $\alpha \in [0, 1)$ such that:*

$$d(F(x), F(y)) \leq \alpha d(x, y), \quad \forall x, y \in X. \tag{1.1}$$

A point $x \in X$ is a fixed point for F if $F(x) = x$.

Theorem 1.2. *Let (X, d) be a complete metric space and $F : X \to X$ a contraction. Then F has a unique fixed point.*

Proof. Vi start by showing uniqueness. Assume that there exist $a, b \in X$ such that $F(a) = a$ and $F(b) = b$. Then (1.1) implies that

$$0 \leq d(a, b) = d(F(a), F(b)) \leq \alpha d(a, b), \quad (1 - \alpha)d(a, b) \leq 0,$$

i.e. $d(a, b) = 0$ and $a = b$.

Now let us construct such a fixed point. Consider the sequence $\{y_n\}_{n\geq 1} \subset X$, where y_1 is arbitrary and $y_n := F(y_{n-1})$ for every $n \geq 2$. We will show two things:

(i). The sequence is Cauchy in X, thus convergent to a limit y_∞ because we assumed X to be complete;

(ii). y_∞ is a fixed point for F.

Let us start with (i). For every $\epsilon > 0$ we will construct $N(\epsilon) > 0$ such that for all $p \geq q \geq N(\epsilon)$ we have $d(y_q, y_p) < \epsilon$. In other words:

$$d(y_q, y_{q+k}) < \epsilon, \quad \forall k \geq 0, \quad \forall q \geq N(\epsilon). \tag{1.2}$$

If $k \geq 1$, the triangle inequality implies:

$$\begin{aligned} d(y_q, y_{q+k}) &\leq d(y_q, y_{q+1}) + d(y_{q+1}, y_{q+k}) \\ &\leq d(y_q, y_{q+1}) + d(y_{q+1}, y_{q+2}) + d(y_{q+2}, y_{q+k}) \\ &\leq \sum_{i=0}^{k-1} d(y_{q+i}, y_{q+i+1}). \end{aligned} \tag{1.3}$$

For every $n \geq 1$ we have:

$$d(y_n, y_{n+1}) = d(F(y_{n-1}), F(y_n)) \leq \alpha d(y_{n-1}, y_n) \leq \cdots \leq \alpha^{n-1} d(y_1, y_2), \quad \forall n \geq 1.$$

Thus $d(y_{q+i}, y_{q+i+1}) \leq \alpha^{q+i-1} d(y_1, y_2)$ for all $q \geq 1$ and $i \geq 0$. Together with (1.3), this implies:

$$d(y_q, y_{q+k}) \leq \alpha^{q-1} d(y_1, y_2)(1 + \cdots + \alpha^{k-1}) \leq \frac{\alpha^{q-1}}{1 - \alpha} d(y_1, y_2), \quad \forall k \geq 1.$$

Because $\alpha < 1$, then $\lim_{q\to\infty} \alpha^q = 0$ and (1.2) follows. We conclude that there exists $y_\infty \in X$ such that

$$\lim_{n\to\infty} d(y_n, y_\infty) = 0. \tag{1.4}$$

Now we prove (ii). For every $n \geq 1$ we have:

$$d(F(y_\infty), y_\infty) \leq d(F(y_\infty), F(y_n)) + d(F(y_n), y_\infty).$$

But $d(F(y_\infty), F(y_n)) \leq \alpha d(y_\infty, y_n) \to 0$ and $d(F(y_n), y_\infty) = d(y_{n+1}, y_\infty) \to 0$ when $n \to \infty$, thus $d(F(y_\infty), y_\infty) = 0$ and $F(y_\infty) = y_\infty$. □

[1]These notes are strongly inspired by the book *Principles of Mathematical Analysis* by Walter Rudin.

2 Local existence and uniqueness for first order ODE's

We start with some general facts about functional spaces.

2.1 Spaces of bounded/continuous functions

Proposition 2.1. *Let (A,d) be a metric space, $(Y, ||\cdot||)$ a normed space, and H an arbitrary non-empty subset of A. We define*

$$B(H;Y) := \{f : H \to Y : \sup_{x\in H} ||f(x)|| < \infty\}.$$

Define the map $||\cdot||_\infty : B(H;Y) \to \mathbb{R}_+$, $\quad ||f||_\infty := \sup_{x\in H} ||f(x)||$. Then the space $(B(H;Y), ||\cdot||_\infty)$ is a normed space, and the map $d_\infty(f,g) := ||f-g||_\infty$ defines a metric.

Proof. Clearly, $||f||_\infty = 0$ if and only if $f(x) = 0$ for all $x \in H$. Moreover,

$$||\lambda f||_\infty = \sup_{x\in H} ||\lambda f(x)|| = |\lambda| \sup_{x\in H} ||f(x)|| = |\lambda|\; ||f||_\infty.$$

Finally, let us prove the triangle inequality. Take $f, g \in B(H;Y)$; then for every $x \in H$ we apply the triangle inequality in $(Y, ||\cdot||)$:

$$||f(x) + g(x)|| \leq ||f(x)|| + ||g(x)|| \leq ||f||_\infty + ||g||_\infty.$$

Thus $||f||_\infty + ||g||_\infty$ is an upper bound for the set $\{||f(x) + g(x)|| : x \in H\}$, hence

$$\sup_{x\in H} ||f(x) + g(x)|| = ||f+g||_\infty \leq ||f||_\infty + ||g||_\infty.$$

Note that $d_\infty(f,g) := ||f-g||_\infty$ is the metric induced by the norm. □

Proposition 2.2. *Denote by $C(H;Y)$ the subset of $B(H;Y)$ where the functions are also continuous. Assume that $(Y, ||\cdot||)$ is a Banach space (a complete normed space). Then $(C(H;Y), ||\cdot||_\infty)$ is a Banach space, too.*

Proof. We need to prove that every Cauchy sequence is convergent. Assume that $\{f_n\}_{n\geq 1} \subset C(H;Y)$ is Cauchy, i.e. for every $\epsilon > 0$ one can find $N_C(\epsilon) > 0$ such that $||f_p - f_q||_\infty < \epsilon$ if $p, q > N_C(\epsilon)$. We have to show that the sequence has a limit f which belongs to $C(H;Y)$.

We first construct f. For every $x_0 \in H$ we consider the sequence $\{f_n(x_0)\}_{n\geq 1} \subset Y$. Note the difference between $\{f_n(x_0)\}_{n\geq 1}$ (a sequence of vectors from Y) and $\{f_n\}_{n\geq 1}$ (a sequence of functions from $C(H;Y)$). It is easy to see that $\{f_n(x_0)\}_{n\geq 1}$ is Cauchy in Y (exercise), and because Y is complete, then $\{f_n(x_0)\}_{n\geq 1}$ has a limit in Y. We denote it with $f(x_0)$. Moreover, since $\{f_n\}_{n\geq 1}$ is Cauchy it must be bounded, i.e. $||f_n||_\infty \leq M < \infty$ for all $n \geq 1$. Thus we have:

$$||f(x)|| = \lim_{n\to\infty} ||f_n(x)|| \leq M, \quad \forall x \in H,$$

therefore $||f||_\infty < \infty$.

Second, we prove the "uniform convergence" part, or the convergence in the norm $||\cdot||_\infty$. More precisely, it means that for every $\epsilon > 0$ we must construct $N_1(\epsilon) > 0$ so that:

$$\sup_{x\in H} ||f(x) - f_n(x)|| < \epsilon \quad \text{whenever} \quad n > N_1(\epsilon). \tag{2.1}$$

In order to do that, take an arbitrary point $x \in H$. For every $p, n \geq 1$ we have

$$\begin{aligned} ||f(x) - f_n(x)|| &\leq ||f(x) - f_p(x)|| + ||f_p(x) - f_n(x)|| \\ &\leq ||f(x) - f_p(x)|| + ||f_p - f_n||_\infty. \end{aligned} \tag{2.2}$$

If we choose $n, p > N_C(\epsilon/2)$, then we have $||f_p - f_n||_\infty < \epsilon/2$ and

$$||f(x) - f_n(x)|| \leq ||f(x) - f_p(x)|| + \epsilon/2, \quad n, p > N_C(\epsilon/2).$$

But the above left hand side does not depend on p, thus if we take $p \to \infty$ on the right hand side, we get:

$$||f(x) - f_n(x)|| \leq \epsilon/2 < \epsilon, \quad n > N_C(\epsilon/2). \tag{2.3}$$

Note that this inequality holds true *for every* x. This means that $\epsilon/2$ is an upper bound for the set $\{||f(x) - f_n(x)|| : x \in H\}$, hence (2.1) holds true with $N_1(\epsilon) = N_C(\epsilon/2)$.

Third, we must prove that f is a continuous function on H. Fix some point $a \in H$. Choose $\epsilon > 0$. Since $\lim_{n\to\infty} f_n(a) = f(a)$, we can find $N_2(\epsilon, a) > 0$ such that $||f_n(a) - f(a)|| < \epsilon$ whenever $n > N_2$. We define $n_1 := \max\{N_1(\epsilon/3), N_2(\epsilon/3, a)\}$. Because f_{n_1} is continuous at a, we can find $\delta(\epsilon, a) > 0$ so that for every $x \in H$ with $d(x, a) < \delta$ we have $||f_{n_1}(x) - f_{n_1}(a)|| < \epsilon/3$. Thus if $x \in H$ with $d(x, a) < \delta$ we have:

$$\begin{aligned} ||f(x) - f(a)|| &\leq ||f(x) - f_{n_1}(x)|| + ||f_{n_1}(x) - f_{n_1}(a)|| + ||f_{n_1}(a) - f(a)|| \\ &< 2||f - f_{n_1}||_\infty + ||f_{n_1}(x) - f_{n_1}(a)|| < \epsilon. \end{aligned} \tag{2.4}$$

Since a is arbitrary, we can conclude that f is continuous on H, thus belongs to $C(H; Y)$. Therefore we can rewrite (2.1) as:

$$||f - f_n||_\infty < \epsilon \quad \text{whenever} \quad n > N_1(\epsilon), \tag{2.5}$$

and the proof is over. □

Remark 2.3. *The "ordinary" convergence in the functional space $(C(H; Y), ||\cdot||_\infty)$ (given in (2.5)) is nothing but the uniform convergence of a sequence of functions defined on the set H (see (2.1)). One can find more details in Wade, exercise 10.6.6 in Chapter 10.6 (page 376).*

2.2 The main theorem

Let U be an open set in $\mathbb{R}^d$, $d \geq 1$, and $I \subset \mathbb{R}$ an open interval. Assume that there exist $\mathbf{y}_0 \in U$ and $r_0, \delta_0 > 0$ such that $\overline{B_{r_0}(\mathbf{y}_0)} \subset U$ and $[t_0 - \delta_0, t_0 + \delta_0] \subset I$.

We consider a continuous function $\mathbf{f} : I \times U \to \mathbb{R}^d$ for which there exists $L > 0$ such that

$$\|\mathbf{f}(t, \mathbf{x}) - \mathbf{f}(t, \mathbf{y})\| \leq L\|\mathbf{x} - \mathbf{y}\|, \quad \forall t \in [t_0 - \delta_0, t_0 + \delta_0], \quad \forall \mathbf{x}, \mathbf{y} \in \overline{B_{r_0}(\mathbf{y}_0)}. \tag{2.6}$$

We define the compact set $H_0 := [t_0 - \delta_0, t_0 + \delta_0] \times \overline{B_{r_0}(\mathbf{y}_0)} \subset \mathbb{R}^{n+1}$. Because $\mathbf{f}$ is continuous, the set $\mathbf{f}(H_0)$ is also compact (see Theorem 10.61 in Wade), hence bounded. Thus we can find $M < \infty$ such that

$$\sup_{(t,\mathbf{x}) \in H_0} \|\mathbf{f}(t, \mathbf{x})\| =: M < \infty. \tag{2.7}$$

Consider the initial value problem:

$$\mathbf{y}'(t) = \mathbf{f}(t, \mathbf{y}), \quad y(t_0) = \mathbf{y}_0. \tag{2.8}$$

Theorem 2.4. *Define*

$$\delta_1 := \min\{\delta_0, r_0/M, 1/L\}.$$

Then the equation (2.8) has a unique solution for $t \in]t_0 - \delta_1, t_0 + \delta_1[$.

Proof. Take some $0 < \delta < \delta_1$ and define the compact interval $K := [t_0 - \delta, t_0 + \delta] \subset \mathbb{R}$. Then any continuous function $\phi : K \to \mathbb{R}^d$ is automatically bounded, and since the Euclidian space $Y = \mathbb{R}^d$ is a Banach space, we can conclude from Proposition 2.2 that the space $(C(K; \mathbb{R}^d), d_\infty)$ of continuous functions defined on the compact K with values in $\mathbb{R}^d$ is a complete metric space.

Define

$$X := \{g \in C(K; \mathbb{R}^d) : \quad g(t) \in \overline{B_{r_0}(\mathbf{y}_0)}, \ \forall t \in K\}. \tag{2.9}$$

Lemma 2.5. *The metric space (X, d_∞) is complete.*

Proof. Consider a Cauchy sequence $\{f_n\}_{n\geq 1} \subset X$. Because $(C(K;\mathbb{R}^d), d_\infty)$ is complete, we can find $f_\infty \in C(H;\mathbb{R}^d)$ such that $\lim_{n\to\infty} d_\infty(f_n, f_\infty) = 0$. Thus for every $t \in H$ we have

$$f_\infty(t) = \lim_{n\to\infty} f_n(t), \quad \lim_{n\to\infty} \|f_n(t) - f_\infty(t)\| = 0.$$

Since by assumption $\|f_n(t) - \mathbf{y}_0\| \leq r_0$ for all t and n, we have

$$\|f_\infty(t) - \mathbf{y}_0\| = \lim_{n\to\infty} \|f_n(t) - \mathbf{y}_0\| \leq r_0, \quad \forall t \in K,$$

which implies that $f_\infty \in X$. □

Lemma 2.6. *Define the map $F : X \to C(K;\mathbb{R}^d)$*

$$[F(g)](t) := \mathbf{y}_0 + \int_{t_0}^t \mathbf{f}(s, g(s))ds, \quad \forall t \in K,$$

where $\mathbf{f}$ is given in (2.6). Then:
(i). The range of F belongs to X;
(ii). $F : X \to X$ is a contraction.

Proof. (i). Because $g(s) \in \overline{B_{r_0}(\mathbf{y}_0)}$ for all $s \in K$, we have that $(s, g(s)) \in H_0$ for all $s \in K$. Thus (see (2.7)) $\sup_{s\in K} \|\mathbf{f}(s, g(s))\| \leq M$ and

$$\|[F(g)](t) - \mathbf{y}_0\| \leq \left\| \int_{t_0}^t \mathbf{f}(s, g(s))ds \right\| \leq M\delta < r_0, \quad \forall t \in K,$$

which means that $[F(g)](t) \in \overline{B_{r_0}(\mathbf{y}_0)}$ for all $t \in K$.

(ii). Consider two functions $g, h \in X$. We have

$$d_\infty(F(g), F(h)) = \sup_{t\in K} \|[F(g)](t) - [F(h)](t)\|.$$

But the Lipschitz condition from (2.6) implies:

$$\begin{aligned} |F(g)](t) - [F(h)](t)| &= \left| \int_{t_0}^t [\mathbf{f}(s, g(s)) - \mathbf{f}(s, h(s))]ds \right| \leq (\delta\ L) \sup_{s\in K} ||g(s) - h(s)|| \\ &\leq (\delta\ L) d_\infty(g, h), \quad \forall t \in K. \end{aligned} \tag{2.10}$$

It means that $d_\infty(F(g), F(h)) \leq (\delta\ L)d_\infty(g, h)$ for all $g, h \in X$, and remember that $\delta L < 1$. Thus F is a contraction. □

Finishing the proof of Theorem 2.4. Vi have seen that F is a contraction on X. Then Theorem 1.2 implies that there exists a continuous function $\mathbf{y} : K \to \overline{B_{r_0}(\mathbf{y}_0)}$ such that

$$\mathbf{y}(t) = [F(\mathbf{y})](t) = \mathbf{y}_0 + \int_{t_0}^t \mathbf{f}(s, \mathbf{y}(s))ds, \quad t \in [t_0 - \delta, t_0 + \delta].$$

It means that $\mathbf{y}$ is differentiable for $t \in]t_0 - \delta, t_0 + \delta[$ and (2.8) is satisfied. □

Remark 2.7. *Choose $0 < \delta < \delta_1$. Define the sequence of functions $\mathbf{y}_k : [t_0 - \delta, t_0 + \delta] \to \mathbb{R}^d$, $k \geq 1$, where $\mathbf{y}_1(t) = \mathbf{y}_0$ and*

$$\mathbf{y}_{k+1}(t) = \mathbf{y}_0 + \int_{t_0}^t \mathbf{f}(s, \mathbf{y}_k(s))ds, \quad k \geq 1.$$

We see that $\mathbf{y}_{k+1} = F(\mathbf{y}_k)$, where F is given by Lemma 2.6. A direct use of Lemma 2.6 (ii) implies that $\{\mathbf{y}_k\}_{k\geq 1}$ converges uniformly on the interval $[t_0 - \delta, t_0 + \delta]$ towards a continuous function $\mathbf{y}_\infty$ which obeys the fixed point equation

$$\mathbf{y}_\infty(t) = \mathbf{y}_0 + \int_{t_0}^t \mathbf{f}(s, \mathbf{y}_\infty(s))ds,$$

thus solving (2.8). This is Picard's iteration method.

3 The inverse function theorem

In this section we only work with the Euclidian space $\mathbb{R}^d$, whose norm is defined by $||\mathbf{x}|| = \sqrt{\sum_{j=1}^d |x_j|^2}$.

Lemma 3.1. *Let A be a $d \times d$ matrix with real components $\{a_{jk}\}$. Define the quantity $||A||_{\mathrm{HS}} := \sqrt{\sum_{j=1}^d \sum_{k=1}^d |a_{jk}|^2}$. Then*

$$||A\mathbf{x}|| \leq ||A||_{\mathrm{HS}}\, ||x||, \qquad \forall \mathbf{x} \in \mathbb{R}^d. \tag{3.1}$$

Proof. From the Cauchy-Schwarz inequality we have:

$$|(A\mathbf{x})_j|^2 = \left(\sum_{k=1}^d a_{jk}x_k\right)^2 \leq \sum_{m=1}^d |a_{jm}|^2 \sum_{n=1}^d |x_n|^2 = \sum_{m=1}^d |a_{jm}|^2 ||\mathbf{x}||^2,$$

and the lemma follows after summation with respect to j. □

Lemma 3.2. *Let $K := \overline{B_\delta(\mathbf{x}_0)} = \{\mathbf{y} \in \mathbb{R}^d : \ ||\mathbf{y} - \mathbf{x}_0|| \leq \delta\}$ be a closed ball in $\mathbb{R}^d$. Let $\phi : \mathbb{R}^d \mapsto \mathbb{R}$ be a $C^1(K)$ map (which means that $\partial_j \phi$ exist for all j and are continuous functions on K). Denote by $||\partial_j \phi||_\infty = \sup_{\mathbf{x} \in K} |\partial_j \phi(\mathbf{x})| < \infty$. Then for every $\mathbf{u}, \mathbf{w} \in K$ we have:*

$$|\phi(\mathbf{u}) - \phi(\mathbf{w})| \leq \sqrt{\sum_{j=1}^d ||\partial_j \phi||_\infty^2}\ ||\mathbf{u} - \mathbf{w}||. \tag{3.2}$$

Proof. Define the real valued map $f(t) = \phi((1-t)\mathbf{w} + t\mathbf{u})$, $0 \leq t \leq 1$. Applying the chain rule we obtain:

$$f'(t) = \sum_{j=1}^d (u_j - w_j)(\partial_j \phi)((1-t)\mathbf{w} + t\mathbf{u}),$$

thus the Cauchy-Schwarz inequality implies:

$$|f'(t)| \leq \sqrt{\sum_{j=1}^d |\partial_j \phi)((1-t)\mathbf{w} + t\mathbf{u})|^2}\ ||\mathbf{u} - \mathbf{w}|| \leq \sqrt{\sum_{j=1}^d ||\partial_j \phi||_\infty^2}\ ||\mathbf{u} - \mathbf{w}||, \quad \forall 0 < t < 1.$$

Since $\phi(\mathbf{u}) - \phi(\mathbf{w}) = f(1) - f(0) = \int_0^1 f'(t)dt$, we obtain:

$$|\phi(\mathbf{u}) - \phi(\mathbf{w})| \leq \int_0^1 |f'(t)|dt \leq \sqrt{\sum_{j=1}^d ||\partial_j \phi||_\infty^2}\ ||\mathbf{u} - \mathbf{w}||$$

which proves (3.2). □

Lemma 3.3. *Let K be as above. Let $\mathbf{f} : \mathbb{R}^d \mapsto \mathbb{R}^d$ a vector valued map which is $C^1(K)$ (which means that $\partial_j f_k$ exist and are continuous functions on K). Define*

$$||\Delta \mathbf{f}||_{\infty,K} := \sqrt{\sum_{k=1}^d \sum_{j=1}^d ||\partial_j f_k||_\infty^2}.$$

Then we have:

$$||\mathbf{f}(\mathbf{u}) - \mathbf{f}(\mathbf{w})|| \leq ||\Delta \mathbf{f}||_{\infty,K} ||\mathbf{u} - \mathbf{w}||, \quad \forall \mathbf{u}, \mathbf{w} \in K. \tag{3.3}$$

Proof. Use (3.2) with ϕ replaced by f_k. We have:

$$|f_k(\mathbf{u}) - f_k(\mathbf{w})|^2 \leq \sum_{j=1}^{d} ||\partial_j f_k||_\infty^2 ||\mathbf{u} - \mathbf{w}||^2$$

and then sum over k. $\square$

Lemma 3.4. *Using the above notation, define* $\mathbf{g}(\mathbf{x}) = \mathbf{f}(\mathbf{x}) - [D\mathbf{f}(\mathbf{x}_0)]\mathbf{x}$*, where* $[D\mathbf{f}(\mathbf{x}_0)]$ *is the Jacobi matrix with elements* $[D\mathbf{f}(\mathbf{x}_0)]_{kj} = (\partial_j f_k)(\mathbf{x}_0)$*. Then for every* $\beta > 0$ *there exists a* $\delta_\beta > 0$ *such that for every* $0 < \delta < \delta_\beta$ *we have:*

$$||\mathbf{g}(\mathbf{u}) - \mathbf{g}(\mathbf{w})|| \leq \beta ||\mathbf{u} - \mathbf{w}||, \quad \forall \mathbf{u}, \mathbf{w} \in K. \tag{3.4}$$

Proof. A straightforward computation gives $\partial_j g_k(\mathbf{x}) = \partial_j f_k(\mathbf{x}) - \partial_j f_k(\mathbf{x}_0)$. Thus $||\partial_j g_k||_\infty$ can be made arbitrarily small when δ gets smaller, because $\mathbf{f}$ has continuous partial derivatives. It follows that $||\Delta \mathbf{g}||_{\infty,K} \leq \beta$ whenever δ gets smaller than some small enough δ_β, and then we can use (3.3) with $\mathbf{g}$ instead of $\mathbf{f}$. $\square$

Lemma 3.5. *Let* $\mathbf{a} \in \mathbb{R}^d$ *and let* $U \subset \mathbb{R}^d$ *be an open set with* $\mathbf{a} \in U$*. Let* $\mathbf{f}$ *be a* $C^1(U)$ *vector valued function, such that* $[D\mathbf{f}(\mathbf{a})]$ *is an invertible matrix. Then there exists* $r > 0$ *small enough such that the restriction of* $\mathbf{f}$ *to* $B_r(\mathbf{a})$ *is injective.*

Proof. Assume the contrary: for every $r > 0$ we can find two different points $\mathbf{x}_r \neq \mathbf{y}_r$ in $B_r(\mathbf{a})$ such that $\mathbf{f}(\mathbf{x}_r) = \mathbf{f}(\mathbf{y}_r)$. Define $\mathbf{g}(\mathbf{x}) = \mathbf{f}(\mathbf{x}) - [D\mathbf{f}(\mathbf{a})]\mathbf{x}$ on $B_r(\mathbf{a})$. Then we have $\mathbf{g}(\mathbf{x}_r) - \mathbf{g}(\mathbf{y}_r) = [D\mathbf{f}(\mathbf{a})](\mathbf{y}_r - \mathbf{x}_r)$ or:

$$\mathbf{y}_r - \mathbf{x}_r = [D\mathbf{f}(\mathbf{a})]^{-1}(\mathbf{g}(\mathbf{x}_r) - \mathbf{g}(\mathbf{y}_r)), \quad \forall r > 0.$$

Now using (3.1) we have:

$$||\mathbf{y}_r - \mathbf{x}_r|| = ||[D\mathbf{f}(\mathbf{a})]^{-1}||_{\mathrm{HS}} \, ||\mathbf{g}(\mathbf{x}_r) - \mathbf{g}(\mathbf{y}_r)||, \quad \forall r > 0.$$

Choosing $\beta = \frac{1}{1 + ||[D\mathbf{f}(\mathbf{a})]^{-1}||_{\mathrm{HS}}}$, then (3.4) claims that there exists some $r_\beta > 0$ sufficiently small such that for every $r \leq r_\beta$ we have $||\mathbf{g}(\mathbf{x}_r) - \mathbf{g}(\mathbf{y}_r)|| \leq \beta ||\mathbf{y}_r - \mathbf{x}_r||$. It follows that:

$$||\mathbf{y}_r - \mathbf{x}_r|| \leq \frac{||[D\mathbf{f}(\mathbf{a})]^{-1}||_{\mathrm{HS}}}{1 + ||[D\mathbf{f}(\mathbf{a})]^{-1}||_{\mathrm{HS}}} ||\mathbf{y}_r - \mathbf{x}_r|| < ||\mathbf{y}_r - \mathbf{x}_r||, \quad \forall 0 < r < r_\beta,$$

which contradicts the assumption $||\mathbf{y}_r - \mathbf{x}_r|| \neq 0$. $\square$

Lemma 3.6. *Let* $\mathbf{f}$ *be as in Lemma 3.5, and consider the injective restriction of* f *to* $B_r(\mathbf{a})$*. Then by eventually making* r *even smaller we have that the Jacobi matrix* $[D\mathbf{f}(\mathbf{x})]$ *is invertible for every* $\mathbf{x} \in B_r(\mathbf{a})$*.*

Proof. The matrix $[D\mathbf{f}(\mathbf{x})]$ is invertible if and only if its determinant $\det[D\mathbf{f}(\mathbf{x})] \neq 0$. But the determinant is a continuous function of $\mathbf{x}$ since $\mathbf{f}$ is C^1. Because $[D\mathbf{f}(\mathbf{a})]$ is invertible, it follows that $|\det[D\mathbf{f}(\mathbf{a})]| > 0$. Being continuous at $\mathbf{a}$, the determinant has the property that $|\det[D\mathbf{f}(\mathbf{x})]| \geq |\det[D\mathbf{f}(\mathbf{a})]|/2 > 0$ on a small ball around $\mathbf{a}$. Thus $[D\mathbf{f}(\mathbf{x})]$ is invertible there. $\square$

Theorem 3.7. *Let* $\mathbf{f}$ *be* C^1 *on an open set containing* $\mathbf{a} \in \mathbb{R}^d$*, such that* $[D\mathbf{f}(\mathbf{a})]$ *is invertible. Then there exists* $r > 0$ *small enough such that the restriction of* f *to* $B_r(\mathbf{a})$ *is injective, and* $[D\mathbf{f}(\mathbf{x})]$ *is invertible for every* $\mathbf{x} \in B_r(\mathbf{a})$*. Moreover, if* $V := f(B_r(\mathbf{a}))$*, the following facts hold true:*

(i). The set V *is open in* $\mathbb{R}^d$*;*

(ii). The inverse $\mathbf{f}^{-1} : V \mapsto B_r(\mathbf{a})$ *is a* $C^1(V)$ *function, and we have:*

$$[D\mathbf{f}^{-1}(\mathbf{y})] = [D\mathbf{f}(\mathbf{f}^{-1}(\mathbf{y}))]^{-1}.$$

Proof. The ball $B_r(\mathbf{a})$ has already been constructed in Lemma 3.6, hence we only need to prove (i) and (ii).

We start with (i). Assume that $\mathbf{y}_0 \in V$, thus equal to $\mathbf{f}(\mathbf{x}_0)$ for some $\mathbf{x}_0 \in B_r(\mathbf{a})$. We will show that $\mathbf{y}_0$ is an interior point of V. This means that we must show the existence of a small ball $B_\epsilon(\mathbf{y}_0)$ which is completely contained in V. In other words, we have to prove that there exists a sufficiently small $\epsilon > 0$ such that for every $\mathbf{y} \in \mathbb{R}^d$ with $||\mathbf{y} - \mathbf{y}_0|| < \epsilon$ we can find a point $\mathbf{x}_\mathbf{y} \in B_r(\mathbf{a})$ such that $\mathbf{f}(\mathbf{x}_y) = \mathbf{y} \in V$, hence $B_\epsilon(\mathbf{y}_0) \subset V$.

So the main question we need to answer is the solvability of the equation $\mathbf{f}(\mathbf{x}) = \mathbf{y}$. This equation is equivalent with:

$$0 = \mathbf{f}(\mathbf{x}) - \mathbf{y} = \mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{x}_0) - [D\mathbf{f}(\mathbf{x}_0)](\mathbf{x} - \mathbf{x}_0) + [D\mathbf{f}(\mathbf{x}_0)](\mathbf{x} - \mathbf{x}_0) + \mathbf{y}_0 - \mathbf{y}.$$

Denote by $\mathbf{g}(\mathbf{x}) = \mathbf{f}(\mathbf{x}) - [D\mathbf{f}(\mathbf{x}_0)]\mathbf{x}$. Then we have the equivalence between $0 = \mathbf{f}(\mathbf{x}) - \mathbf{y}$ and the equation

$$\mathbf{g}(\mathbf{x}) - \mathbf{g}(\mathbf{x}_0) + [D\mathbf{f}(\mathbf{x}_0)](\mathbf{x} - \mathbf{x}_0) + \mathbf{y}_0 - \mathbf{y} = 0.$$

Since $[D\mathbf{f}(\mathbf{x}_0)]$ is invertible, we can isolate $\mathbf{x}$ and write another equivalent equation:

$$\mathbf{x} = \mathbf{x}_0 + [D\mathbf{f}(\mathbf{x}_0)]^{-1}(\mathbf{y} - \mathbf{y}_0) - [D\mathbf{f}(\mathbf{x}_0)]^{-1}(\mathbf{g}(\mathbf{x}) - \mathbf{g}(\mathbf{x}_0)).$$

This looks like a fixed point equation. Indeed, let us denote by

$$F_\mathbf{y}(\mathbf{x}) := \mathbf{x}_0 + [D\mathbf{f}(\mathbf{x}_0)]^{-1}(\mathbf{y} - \mathbf{y}_0) - [D\mathbf{f}(\mathbf{x}_0)]^{-1}(\mathbf{g}(\mathbf{x}) - \mathbf{g}(\mathbf{x}_0)). \tag{3.5}$$

It follows that if we can find a fixed point for $F_\mathbf{y}$, it will also solve the equation $\mathbf{f}(\mathbf{x}) = \mathbf{y}$.

We note first that using (3.1) we have:

$$||F_\mathbf{y}(\mathbf{x}) - \mathbf{x}_0|| \leq ||[D\mathbf{f}(\mathbf{x}_0)]^{-1}||_{\mathrm{HS}}(||\mathbf{y} - \mathbf{y}_0|| + ||\mathbf{g}(\mathbf{x}) - \mathbf{g}(\mathbf{x}_0)||). \tag{3.6}$$

Choosing $\beta = \beta_1 = \frac{1}{3(1+||[D\mathbf{f}(\mathbf{x}_0)]^{-1}||_{\mathrm{HS}})}$ in (3.4), it follows that there exists a $\delta_1 > 0$ small enough, in any case smaller than $r - ||\mathbf{x}_0 - \mathbf{a}||$, such that for every $\delta < \delta_1$ and $||\mathbf{x} - \mathbf{x}_0|| < \delta$, we have $||\mathbf{g}(\mathbf{x}) - \mathbf{g}(\mathbf{x}_0)|| \leq \beta_1 ||\mathbf{x} - \mathbf{x}_0||$, thus:

$$||F_\mathbf{y}(\mathbf{x}) - \mathbf{x}_0|| \leq ||[D\mathbf{f}(\mathbf{x}_0)]^{-1}||_{\mathrm{HS}}||\mathbf{y} - \mathbf{y}_0|| + \frac{\delta}{3}, \quad \forall \mathbf{x} \in \overline{B_\delta(\mathbf{x}_0)} \subset B_r(\mathbf{a}). \tag{3.7}$$

In particular, if

$$||\mathbf{y} - \mathbf{y}_0|| < \frac{\delta}{3(1 + ||[D\mathbf{f}(\mathbf{x}_0)]^{-1}||_{\mathrm{HS}})} =: \epsilon_\delta \tag{3.8}$$

then (3.7) states that

$$||F_\mathbf{y}(\mathbf{x}) - \mathbf{x}_0|| \leq \frac{2\delta}{3} \leq \delta, \quad \forall \mathbf{x} \in \overline{B_\delta(\mathbf{x}_0)} \subset B_r(\mathbf{a}). \tag{3.9}$$

This proves that if δ is smaller than some critical value δ_1 and $||\mathbf{y} - \mathbf{y}_0|| < \epsilon_\delta$, then the map $F_\mathbf{y}$ invariates any closed ball $K := \overline{B_\delta(\mathbf{x}_0)}$, i.e. $F_\mathbf{y}(K) \subset K$.

Now we want to show that choosing δ even smaller, the map $F_{\mathbf{y}}$ becomes a contraction. Indeed, from its definition in (3.5) we have:

$$F_{\mathbf{y}}(\mathbf{u}) - F_{\mathbf{y}}(\mathbf{w}) = -[D\mathbf{f}(\mathbf{x}_0)]^{-1}(\mathbf{g}(\mathbf{u}) - \mathbf{g}(\mathbf{w})),$$

or

$$||F_{\mathbf{y}}(\mathbf{u}) - F_{\mathbf{y}}(\mathbf{w})|| \leq ||[D\mathbf{f}(\mathbf{x}_0)]^{-1}||_{\mathrm{HS}}\, ||\mathbf{g}(\mathbf{u}) - \mathbf{g}(\mathbf{w})||, \quad \forall \mathbf{u}, \mathbf{w} \in \overline{B_\delta(\mathbf{x}_0)}.$$

Use again (3.4) with $\beta = \frac{1}{2(1+||[D\mathbf{f}(\mathbf{x}_0)]^{-1}||_{\mathrm{HS}})}$: we obtain some $\delta_2 < \delta_1$ such that $F_{\mathbf{y}}(\overline{B_{\delta_2}(\mathbf{x}_0)}) \subset \overline{B_{\delta_2}(\mathbf{x}_0)}$ and

$$||F_{\mathbf{y}}(\mathbf{u}) - F_{\mathbf{y}}(\mathbf{w})|| \leq \frac{1}{2}||\mathbf{u} - \mathbf{w}||, \quad \forall \mathbf{u}, \mathbf{w} \in \overline{B_{\delta_2}(\mathbf{x}_0)}.$$

Banach's fixed point theorem states that there exists a unique solution $\mathbf{x}_y \in \overline{B_{\delta_2}(\mathbf{x}_0)}$ which solves the equation $F_{\mathbf{y}}(\mathbf{x}) = \mathbf{x}$. Going back to (3.9) we conclude that if $||\mathbf{y} - \mathbf{f}(x_0)|| < \epsilon_{\delta_2}$, then there exists a solution $\mathbf{x}_y \in \overline{B_{\delta_2}(\mathbf{x}_0)}$ which solves the equation $\mathbf{f}(\mathbf{x}_{\mathbf{y}}) = \mathbf{y}$. Since $\mathbf{f}$ is injective on its domain, this solution is also unique. Moreover, $B_{\epsilon_{\delta_2}}(\mathbf{y}_0) \subset V$. Since $\mathbf{y}_0$ was arbitrary, V is open.

Let us now prove (ii). For any $\mathbf{y} \in V$ we constructed $\mathbf{x}_y = \mathbf{f}^{-1}(\mathbf{y})$ which solves $\mathbf{f}(\mathbf{x}_{\mathbf{y}}) = \mathbf{y}$ and $F_{\mathbf{y}}(\mathbf{x}_{\mathbf{y}}) = \mathbf{x}_{\mathbf{y}}$. The fixed point equation rewrites as:

$$\mathbf{x}_{\mathbf{y}} - \mathbf{x}_0 = [D\mathbf{f}(\mathbf{x}_0)]^{-1}(\mathbf{y} - \mathbf{y}_0) - [D\mathbf{f}(\mathbf{x}_0)]^{-1}(\mathbf{g}(\mathbf{x}_{\mathbf{y}}) - \mathbf{g}(\mathbf{x}_0)). \tag{3.10}$$

We know that as soon as $||\mathbf{y} - \mathbf{y}_0|| < \epsilon_{\delta_2}$, the point $\mathbf{x}_{\mathbf{y}}$ belongs to the ball around $\mathbf{x}_0$ where $||\mathbf{g}(\mathbf{x}_{\mathbf{y}}) - \mathbf{g}(\mathbf{x}_0)|| \leq \frac{1}{2(1+||[D\mathbf{f}(\mathbf{x}_0)]^{-1}||_{\mathrm{HS}})}||\mathbf{x}_{\mathbf{y}} - \mathbf{x}_0||$. Using this in (3.10) we get:

$$||\mathbf{x}_{\mathbf{y}} - \mathbf{x}_0|| \leq ||[D\mathbf{f}(\mathbf{x}_0)]^{-1}||_{\mathrm{HS}}\, ||\mathbf{y} - \mathbf{y}_0|| + \frac{1}{2}||\mathbf{x}_{\mathbf{y}} - \mathbf{x}_0||,$$

or $||\mathbf{x}_{\mathbf{y}} - \mathbf{x}_0|| \leq 2||[D\mathbf{f}(\mathbf{x}_0)]^{-1}||_{\mathrm{HS}}\, ||\mathbf{y} - \mathbf{y}_0||$ for $||\mathbf{y} - \mathbf{y}_0||$ smaller than some critical value ϵ_{δ_2}. In other words, it means that

$$\lim_{\mathbf{y} \to \mathbf{y}_0} \mathbf{f}^{-1}(\mathbf{y}) = \mathbf{f}^{-1}(\mathbf{y}_0), \quad ||\mathbf{f}^{-1}(\mathbf{y}) - \mathbf{f}^{-1}(\mathbf{y}_0)|| \leq C||\mathbf{y} - \mathbf{y}_0||,$$

which shows that $\mathbf{f}^{-1}$ is continuous on V. Moreover, (3.4) and the above estimate show that

$$\lim_{\mathbf{y} \to \mathbf{y}_0} \frac{||\mathbf{g}(\mathbf{x}_{\mathbf{y}}) - \mathbf{g}(\mathbf{x}_0)||}{||\mathbf{y} - \mathbf{y}_0||} = 0.$$

Finally, we conclude from (3.10) that

$$\lim_{\mathbf{y} \to \mathbf{y}_0} \frac{||\mathbf{f}^{-1}(\mathbf{y}) - \mathbf{f}^{-1}(\mathbf{y}_0) - [D\mathbf{f}(\mathbf{x}_0)]^{-1}(\mathbf{y} - \mathbf{y}_0)||}{||\mathbf{y} - \mathbf{y}_0||} = 0$$

which shows that $[D\mathbf{f}^{-1}(\mathbf{y}_0)] = [D\mathbf{f}(\mathbf{x}_0)]^{-1}$ for every pair $\mathbf{f}(\mathbf{x}_0) = \mathbf{y}_0$, and we are done. $\square$

4 The implicit function theorem

In this section $d = m + n$ with $1 \leq m, n < d$. A vector $\mathbf{x} \in \mathbb{R}^d$ can be uniquely decomposed as $\mathbf{x} = [\mathbf{u}, \mathbf{w}]$ with $\mathbf{u} \in \mathbb{R}^m$ and $\mathbf{w} \in \mathbb{R}^n$.

Lemma 4.1. *Assume that D is a $d \times d$ matrix which has the following triangular form:*

$$D = \begin{bmatrix} A & B \\ 0_{n\times m} & 1_{n\times n} \end{bmatrix},$$

where A is an $m \times m$ matrix and B is an arbitrary $n \times m$ matrix. Then D is invertible if and only if A is invertible and

$$D^{-1} = \begin{bmatrix} A^{-1} & -A^{-1}B \\ 0_{n\times m} & 1_{n\times n} \end{bmatrix}. \tag{4.1}$$

Proof. We check by direct computation that $DD^{-1} = D^{-1}D = 1_{d\times d}$. $\square$

Lemma 4.2. *Let $U \in \mathbb{R}^d$ be an open set and $\mathbf{h} : U \mapsto \mathbb{R}^m$ be a $C^1(U;\mathbb{R}^m)$ function. Denote by $[D_{\mathbf{u}}h](\mathbf{x})$ the partial $m\times m$ Jacobi matrix when $\mathbf{w}$ is kept fixed, and by $[D_{\mathbf{w}}h](\mathbf{x})$ the partial $m \times n$ Jacobi matrix when $\mathbf{u}$ is kept fixed. Define the function $\mathbf{f} : U \mapsto \mathbb{R}^d$ by the formula $\mathbf{f}([\mathbf{u},\mathbf{w}]) = [\mathbf{h}([\mathbf{u},\mathbf{w}]),\mathbf{w}]$. Then $\mathbf{f} \in C^1(U;\mathbb{R}^d)$ and*

$$[D\mathbf{f}](\mathbf{x}) = \begin{bmatrix} [D_{\mathbf{u}}h](\mathbf{x}) & [D_{\mathbf{w}}h](\mathbf{x}) \\ 0_{n\times m} & 1_{n\times n} \end{bmatrix}. \tag{4.2}$$

Proof. Direct computation. $\square$

We can now formulate the implicit function theorem.

Theorem 4.3. *Let $U \in \mathbb{R}^d$ be an open set and $\mathbf{h} : U \mapsto \mathbb{R}^m$ be a $C^1(U;\mathbb{R}^m)$ function. Assume that there exists a point $\mathbf{a} = [\mathbf{u_a},\mathbf{w_a}] \in U$ such that $\mathbf{h}(\mathbf{a}) = 0$ and the $m \times m$ partial Jacobi matrix $[D_{\mathbf{u}}h](\mathbf{a})$ is invertible. Then there exists an open set $E \in \mathbb{R}^n$ containing $\mathbf{w_a}$, and a map $\mathbf{g} : E \mapsto \mathbb{R}^m$ in $C^1(E;\mathbb{R}^m)$, such that $\mathbf{g}(\mathbf{w_a}) = \mathbf{u_a}$ and $\mathbf{h}([\mathbf{g}(\mathbf{w}),\mathbf{w}]) = 0$ for all $\mathbf{w} \in E$. Moreover, $[D_{\mathbf{u}}h]([\mathbf{g}(\mathbf{w}),\mathbf{w}])$ is invertible for all $\mathbf{w} \in E$ and*

$$[D\mathbf{g}](\mathbf{w}) = -[D_{\mathbf{u}}h]^{-1}([\mathbf{g}(\mathbf{w}),\mathbf{w}])\ [D_{\mathbf{w}}h]([\mathbf{g}(\mathbf{w}),\mathbf{w}]). \tag{4.3}$$

Proof. As in Lemma 4.2 we define $\mathbf{f}([\mathbf{u},\mathbf{w}]) = [\mathbf{h}([\mathbf{u},\mathbf{w}]),\mathbf{w}])$ on U. From Lemma 4.1 and (4.2) we conclude that $[D\mathbf{f}](\mathbf{a})$ is invertible. The inverse function theorem 3.7 provides us with a ball $B_r(\mathbf{a})$ where $\mathbf{f}$ is injective, $[D\mathbf{f}](\mathbf{x})$ is invertible on $B_r(\mathbf{a})$, and the set $V = \mathbf{f}(B_r(\mathbf{a}))$ is open in $\mathbb{R}^d$. Of course, the vector $[0,\mathbf{w_a}] = \mathbf{f}(\mathbf{a}) \in V$. Since V is open, there exists $\epsilon > 0$ such that the d dimensional ball $B_\epsilon(\mathbf{f}(\mathbf{a})) = \{\mathbf{y} \in \mathbb{R}^d :\ ||\mathbf{y} - [0,\mathbf{w_a}]|| < \epsilon\} \subset V$.

In general, if $[\mathbf{u},\mathbf{w}] \in V$, assume that $\mathbf{f}^{-1}([\mathbf{u},\mathbf{w}]) = [\mathbf{u}',\mathbf{w}'] \in B_r(\mathbf{a})$. Then we must have:

$$[\mathbf{u},\mathbf{w}] = \mathbf{f}(\mathbf{f}^{-1}([\mathbf{u},\mathbf{w}])) = [\mathbf{h}(([\mathbf{u}',\mathbf{w}'])),\mathbf{w}']$$

which shows that $\mathbf{w} = \mathbf{w}'$ and $\mathbf{u} = \mathbf{h}([\mathbf{u}',\mathbf{w}])$. This proves that $\mathbf{f}^{-1}([\mathbf{u},\mathbf{w}])$ is also of the form $[\mathbf{u}',\mathbf{w}]$.

Let us define $E := \{\mathbf{w} \in \mathbb{R}^n :\ [0,\mathbf{w}] \in B_\epsilon(\mathbf{f}(\mathbf{a}))\}$, which is nothing but the n dimensional open ball $B_\epsilon(\mathbf{w_a}) \subset \mathbb{R}^n$. Then for every $\mathbf{w} \in E$ we have that $[0,\mathbf{w}] \in V$ and

$$[0,\mathbf{w}] = \mathbf{f}(\mathbf{f}^{-1}([0,\mathbf{w}])) = [h(\mathbf{f}^{-1}([0,\mathbf{w}])),\mathbf{w}]. \tag{4.4}$$

We already know that $\mathbf{f}^{-1}([0,\mathbf{w}])$ must be a vector of the form $[\mathbf{u}',\mathbf{w}]$, where $\mathbf{u}'$ is nothing but the vector obtained from the first m components of $\mathbf{f}^{-1}([0,\mathbf{w}])$. Denote it by $\mathbf{g}(\mathbf{w})$. Then (4.4) implies that $h([\mathbf{g}(\mathbf{w}),\mathbf{w}]) = 0$ on E. Moreover,

$$[\mathbf{g}(\mathbf{w_a}),\mathbf{w_a}] = \mathbf{f}^{-1}([0,\mathbf{w_a}]) = \mathbf{a} = [\mathbf{u_a},\mathbf{w_a}]$$

which implies $\mathbf{g}(\mathbf{w_a}) = \mathbf{u_a}$. Finally, since $(\mathbf{g}(\mathbf{w}))_j = (\mathbf{f}^{-1}([0,\mathbf{w}]))_j$ for $1 \le j \le m$ and $\mathbf{f}^{-1}([0,\mathbf{w}])$ is a $C^1(E;\mathbb{R}^d)$ map, then $\mathbf{g}$ is $C^1(E;\mathbb{R}^m)$. The formula (4.3) can be easily obtained by applying the chain rule to the equality $\mathbf{h}(\mathbf{g}(\mathbf{w}),\mathbf{w}) = 0$.

$\square$