Introduction

In the following X, Y and Z are (finite- or infinite-dimensional) Banach spaces.

Nonlinear equations are often written in the form (fixed point form)

$$x = G(x) \tag{1}$$

with $G : D \longrightarrow X$, $D \subset X$, or

$$F(x) = 0 \tag{2}$$

with $F : D \longrightarrow Y$, $D \subset X$.

A solution of (1) is called a **fixed point** of G.

Remark: An equation of the form (1) also fits into the form (2) with $F(x) = x - G(x)$. Any equation of the form (2) can be written in fixed point form (1) with $G(x) = x - H(x)F(x)$, $H : D \longrightarrow L(Y, X)$, $H(x)$ nonsingular.

If one wants to stress the role of the data, one writes

$$F(x) = y \tag{3}$$

with $F : D \longrightarrow Y$, $D \subset X$, or more generally

$$F(x, y) = 0 \tag{4}$$

with $F : D \longrightarrow Z$, $D \subset X \times Y$ and given data y.

The concept of a well-posed problem is of great importance. A very strong form of this concept applied to (3) is the concept of a **homeomorphism**: $F : D \to Y$, $D \subset X$, D open, is a homeomorphism, i.e. $F^{-1} : F(D) \longrightarrow X$ exists and F and F^{-1} are continuous on D and $F(D)$, respectively. Usually, the following local variant of this concept is considered:

Definition 0.1. *The mapping $F : D \longrightarrow Y$, $D \subset X$, D open, is a* ***local homeomorphism*** *at $x \in D$ if there exist open neighborhoods U and V of x and $F(x)$, respectively, such that $F : U \to V$ is a homeomorphism, i.e. $F^{-1} : V \longrightarrow U$ exists and F and F^{-1} are continuous on U and V, respectively.*

The following theorem provides sufficient conditions for the equation (3) to be well-posed in this sense:

We first recall the concept of F-derivative:

Definition 0.2. *A mapping $F : D \longrightarrow Y$, $D \subset X$, D open, is F-differentiable at $x \in D$ if there is a linear and bounded operator $F'(x) \in L(X, Y)$ such that*

$$\lim_{h \to 0} \frac{1}{\|h\|}(F(x+h) - F(x) - F'(x)h\| = 0.$$

Theorem 0.1 (Inverse function theorem)**.** *Suppose that $F : D \longrightarrow Y$, $D \subset X$, D open, has an F-derivative in D which is continuous at $x_0 \in D$ and that $F'(x_0)$ is nonsingular. Then F is a local homeomorphism at x_0. The inverse function F^{-1} has an F-derivative at $F(x_0)$ and*

$$(F^{-1})'(F(x_0)) = F'(x_0)^{-1}.$$

The extension to more general problems of the form (4) leads to:

Theorem 0.2 (Implicit function theorem)**.** *Let $F : D \longrightarrow Z$, $D \subset X \times Y$, D open, be continuous and let $(x_0, y_0) \in D$ with $F(x_0, y_0) = 0$. Assume that $\partial F/\partial x$ exists in D and is continuous at (x_0, y_0) and that $\partial F/\partial x(x_0, y_0)$ is nonsingular. Then there exist open neighborhoods U and V of x_0 and y_0, respectively, such that, for any $y \in V$ the equation*

$$F(x, y) = 0$$

has a unique solution $x = x(y) \in U$ and the mapping $x : V \longrightarrow X$ is continuous. Moreover, if $\partial F/\partial y$ exists at (x_0, y_0), then $x(y)$ is F-differentiable at y_0 and

$$x'(y_0) = -\left(\frac{\partial F}{\partial x}(x_0, y_0)\right)^{-1} \frac{\partial F}{\partial y}(x_0, y_0).$$

Remark: Although the notation $x = x(y)$ is formally incorrect (x is used to denote a variable as well as a function), we will nevertheless use it because of its suggestive character.

Notation:

$\|x\|$	norm of $x \in X$
$\|x\|_{\ell_2}$	Euclidean norm of a vector $x \in \mathbb{R}^n$
$\|A\|_{\ell_2}$	spectral norm of a matrix $A \in L(\mathbb{R}^n)$
$\|u\|_0$	L^2-norm of $u \in L^2(\Omega)$
$\|u\|_1$	H^1-norm of $u \in H^1(\Omega)$
(u, v)	inner product of $u, v \in X$, X Hilbert space
$x \cdot y$	Euclidean inner product of $x, y \in \mathbb{R}^n$
$(u, v)_1$	inner product of $u, v \in H^1(\Omega)$
$\langle f, v\rangle$	duality product of $f \in X^*$ and $v \in X$: $\langle f, v\rangle = f(v)$.

Chapter 1

Iterative methods

An iterative method generates a sequence

$$x^0, x^1, \ldots, x^k, \ldots$$

of approximative solutions by

$$x^{k+1} = G_k(x^k, x^{k-1}, \ldots, x^0), \quad k = 0, 1, \ldots,$$

with $G_k : D_k \longrightarrow X$, $D_k \subset X^{k+1}$.

An important class of iterative methods are m-step methods:

Definition 1.1 (m-step methods)**.** *Let $m \in \mathbb{N}$.*

1. *A sequence of operators (G_k) with $G_k : D_k \longrightarrow X$, $D_k \subset X^m$ defines an* **m*-step method*** *for initial values from a non-empty set $D_* \subset D_0$, if the sequence*

$$x^{k+1} = G_k(x^k, x^{k-1}, \ldots, x^{k-m+1}), \quad k \geq m-1$$

is well-defined for all $(x^0, x^{-1}, \ldots, x^{-m+1}) \in D_*$.

2. *An m-step method is **stationary**, if*

$$G_k = G, \quad D_k = D$$

for an operator $G\colon D \to X$, $D \subset X^m$.

Example: A stationary one-step method is of the form

$$x^{k+1} = G(x^k).$$

Definition 1.2 (Convergence)**.** *Let the sequence* (G_k) *define an iterative method for initial values from a set* $D_* \subset D_0$*.*

1. *The iterative method* ***converges*** *to* $x^* \in X$ *for initial values from* D_* *if*

$$x^k \to x^*$$

for all initial values $(x^0, x^{-1}, \ldots, x^{-m+1}) \in D_*$*.*

2. *The iterative method is called* ***locally convergent*** *if there is an element* $x^* \in X$ *and a neighborhood* U *of* x^* *such that the iterative method converges to* x^* *for all initial values from* $D_* = U^m$*.*

Definition 1.3 (quotient-convergence order, q-order)**.** *Let* (G_k) *define an iterative method and let* x^* *be the limit of a sequence* (x^k) *generated by the iterative method. The the sequence* (x^k)

1. ***converges*** **q-*linearly*** *(converges with* q*-order* 1*) if there are constants* $q \in [0, 1)$ *and* $K \in \mathbb{N}_0$ *such that*

$$\|x^{k+1} - x^*\| \le q\,\|x^k - x^*\| \quad \text{for all } k \ge K;$$

2. ***converges with*** **q-*order*** $\mathbf{p} > \mathbf{1}$ *if there are constants* $q > 0$ *and* $K \in \mathbb{N}_0$ *such that*

$$\|x^{k+1} - x^*\| \le q\,\|x^k - x^*\|^p \quad \text{for all } k \ge K.$$

Remark: In the case $p = 2$ and $p = 3$ we speak of **q-quadratic** and **q-cubic** convergence, respectively.

If (x^k) converges with q-order p, then the following **quotient convergence factor (in short: q-factor**) $q_p\{(x^k)\}$ is introduced:

$$q_p\{(x^k)\} = \limsup_{k\to\infty} \frac{\|x^{k+1} - x^*\|}{\|x^k - x^*\|^p}.$$

$q_p\{(x^k)\}$ is the infimum of all possible constants q in the definition from above. An interesting special case is $q_1\{(x^k)\} = 0$. That means q-linear convergence with arbitrarily small q.

Definition 1.4 (q-superlinear convergence)**.** *Let* (G_k) *define an iterative method and let* x^* *be the limit of a sequence* (x^k) *generated by the iterative method. Then* (x^k) *converges* ***q-superlinearly*** *if* $q_1\{(x^k)\} = 0$*.*

For simplicity assume that $K = 0$. Then the q-linear convergence implies

$$\|x^k - x^*\| \leq q\|x^{k-1} - x^*\| \leq \ldots \leq q^k \|x^0 - x^*\|.$$

So, the error can be estimated by a geometric sequence of the form $c\,q^k$, $q < 1$. For a sequence with q-order p we have

$$\|x^k - x^*\| \leq q^{1+p+\cdots+p^{k-1}} \, \|x^0 - x^*\|^{p^k} = \frac{1}{q^{1/(p-1)}} \left[q^{1/(p-1)} \|x^0 - x^*\| \right]^{p^k}.$$

So, the error can be bounded by a sequence of the form $c\,\bar{q}^{p^k}$, $\bar{q} < 1$ if x^0 is sufficiently close to x^*. This motivates the following definition:

Definition 1.5 (root-convergence order, r-order)**.** *Let* (G_k) *define an iterative method and let* x^* *be the limit of a sequence* (x^k) *generated by the iterative method. The the sequence* (x^k)

1. ***converges r-linearly*** *(converges with* r*-order 1) if there are constants* $c \geq 0$*,* $q \in [0, 1)$ *and* $K \in \mathbb{N}_0$ *such that*

$$\|x^k - x^*\| \leq c\,q^k \quad \textit{for all } k \geq K;$$

2. ***converges with r-order*** $\mathbf{p} > \mathbf{1}$ *if there are constants* $c > 0$*,* $q \in [0, 1)$ *and* $K \in \mathbb{N}_0$ *such that*

$$\|x^k - x^*\| \leq c\,q^{p^k} \quad \textit{for all } k \geq K.$$

The r-convergence order measures only the asymptotic behavior of the error for $k \to \infty$ while the q-convergence order measure the asymptotic behavior of two consecutive errors.

Remark: In the case $p = 2$ and $p = 3$ we speak of **r-quadratic** and **r-cubic** convergence, respectively.

If the sequence (x^k) is convergent with r-order p, then the following **root-convergence factor (in short: r-factor) $\mathbf{r_p}\{(\mathbf{x^k})\}$** is introduced

$$r_p\{(x^k)\} = \begin{cases} \limsup_{k\to\infty} \|x^k - x^*\|^{1/k} & \text{for } \; p = 1, \\ \limsup_{k\to\infty} \|x^k - x^*\|^{1/p^k} & \text{for } \; p > 1. \end{cases}$$

$r_p\{(x^k)\}$ is the infimum of all possible constants q in the definition from above. An interesting special case is $r_1\{(x^k)\} = 0$. That means r-linear convergence with arbitrarily small q.

Definition 1.6 (r-superlinear convergence)**.** *Let (G_k) define an iterative method and let x^* be the limit of a sequence (x^k) generated by the iterative method. Then (x^k) converges **r-superlinearly** if $r_1\{(x^k)\} = 0$.*

The Banach fixed point theorem can be seen as an important statement on stationary one-step methods:

Theorem 1.1 (The Banach fixed point theorem)**.** *Let the mapping $G : D \longrightarrow X$, $D \subset X$, D closed, be contractive, i.e. there is a constant $q \in [0, 1)$ with*

$$\|G(y) - G(x)\| \le q\,\|y - x\| \quad \textit{for all } x, y \in D,$$

and assume that $G(D) \subset D$. Then:

1. *The equation*

   $$x = G(x)$$

   has a unique solution x^* in D.

2. *The sequence (x^k), given by*

   $$x^{k+1} = G(x^k)$$

   converges to x for all initial values $x^0 \in D$ and

   $$\|x^{k+1} - x^*\| \le q\,\|x^k - x^*\|,$$

   i.e. (x^k) converges q-linearly and, therefore, also r-linearly:

   $$\|x^k - x^*\| \le q^k\,\|x^0 - x^*\|.$$

3. Additionally we have

$$\|x^k - x^*\| \leq \frac{q^k}{1-q}\,\|x^1 - x^0\|,$$

$$\|x^k - x^*\| \leq \frac{q}{1-q}\,\|x^k - x^{k-1}\|.$$

The following theorem is a local variant of the Banach fixed point theorem. The existence of a fixed point is assumed and only the convergence of the iterative method is considered:

Theorem 1.2 (Ostrowski)**.** *Suppose that the mapping $G : D \longrightarrow X$, $D \subset X$, D open, has a fixed point $x^* \in D$ and is F-differentiable at x^*. If the spectral radius of $G'(x^*)$ satisfies $\rho(G'(x^*)) < 1$, then the stationary one-step method generated by G is locally convergent and r-linearly convergent with $r_1\{(x^k)\} \leq \rho(G'(x^*))$. Under the stronger condition $\|G'(x^*)\| < 1$ the q-linear convergence with $q_1\{(x^k)\} \leq \|G'(x^*)\|$ follows.*

Proof. Let $\|G'(x^*)\| < 1$. The existence of the F-derivative of G at x^* implies that, for each $\varepsilon > 0$ there is a $\delta > 0$ with

$$\|G(x) - G(x^*) - G'(x^*)(x - x^*)\| \leq \varepsilon\|x - x^*\| \quad \text{for all } x \in \overline{K}(x^*, \delta)$$

with $\overline{K}(x^*, \delta) = \{y \in X : \|y - x^*\| \leq \delta\}$. Therefore,

$$\begin{aligned} \|G(x) - x^*\| &\leq \|G(x) - G(x^*) - G'(x^*)(x - x^*)\| + \|G'(x^*)\|\|x - x^*\| \\ &\leq (\varepsilon + \|G'(x^*)\|)\|x - x^*\|. \end{aligned}$$

For sufficiently small ε we have

$$\|G(x) - x^*\| \leq q\,\|x - x^*\|$$

with $q = \varepsilon + \|G'(x^*)\| < 1$. This shows the q-linear convergence and $q_1\{(x^k)\} \leq \|G'(x^*)\|$.

If $\rho(G'(x^*)) = \lim_{k\to\infty} \|G'(x^*)^k\|^{1/k} < 1$, then there is an $M \in \mathbb{N}$ with

$$\|G'(x^*)^m\| < 1 \quad \text{for all } m \geq M.$$

For the fixed point equation

$$x = G^m(x)$$

($G^m(x) = G(G(\ldots G(x)\ldots)))$ the q-linear convergence of the sub-sequences $(x^{km+l})_{k\in\mathbb{N}}$ towards the fixed point x^* follows from the first part of the proof for all $l = 0, 1, \ldots, m-1$ with $q_1\{(x^{km+l})\} \leq \|G'(x^*)^m\|$.

Therefore, for arbitrary but fixed $m \in \mathbb{N}$ and $\varepsilon > 0$, there is a number $K \in \mathbb{N}$, such that

$$\|x^{km+l} - x^*\| \leq [\|G'(x^*)^m\| + \varepsilon]^{k-K} \|x^{Km+l} - x^*\|.$$

for all $k \geq K$ and $l < m$. Hence

$$\|x^{km+l} - x^*\|^{1/(km+l)} \leq [\|G'(x^*)^m\| + \varepsilon]^{(k-K)/(km+l)} \|x^{Km+l} - x^*\|^{1/(km+l)}.$$

This implies

$$\limsup_{n\to\infty} \|x^n - x^*\|^{1/n} \leq [\|G'(x^*)^m\| + \varepsilon]^{1/m}.$$

With $m \to \infty$ and $\varepsilon \to 0$ we finally have

$$r_1\{(x^k)\} \leq \rho(G'(x^*)).$$

□

In particular, the Ostrowski theorem implies the r-superlinear convergence if $\rho(G'(x^*)) = 0$, and the q-superlinear convergence if $G'(x^*) = 0$.

Theorem 1.3. *Suppose that the mapping $G : D \longrightarrow X$, $D \subset X$, D open, has a fixed point $x^* \in D$, and is m times differentiable in D and the derivative $G^{(m)}$ satisfies the Lipschitz-condition:*

$$\|G^{(m)}(x) - G^{(m)}(x^*)\| \leq \gamma\|x - x^*\| \quad \textit{for all } x \in D.$$

Then we have: If

$$G^{(k)}(x^*) = 0, \quad \textit{for all } k = 1, 2, \ldots, m,$$

then the one-step method generated by G is locally convergent and has q-order $m + 1$.

Proof. By estimating the remainder of the Taylor expansion it follows

$$\begin{aligned}\|G(x)-x^*\| &= \|G(x)-\sum_{k=0}^{m}\frac{1}{k!}G^{(k)}(x^*)(x-x^*)^k\| \\ &\le \frac{1}{m!}\|[G^{(m)}(x^*+t(x-x^*))-G^{(m)}(x^*)](x-x^*)^m\| \\ &\le \frac{\gamma}{m!}\|x-x^*\|^{m+1}\end{aligned}$$

for some $t\in(0,1)$. This immediately implies the statements. □

Remark: For $m=1$ the last theorem implies the q-quadratic convergence, if $G'(x^*)=0$ and if G' satisfies the Lipschitz-condition.

Remark: The existence of $G^{(m+1)}(x^*)$ guarantees the Lipschitz-condition for $G^{(m)}$ in a neighborhood of x^*.

For an equation of the form $F(x)=0$ we obtain a fixed point form $x=G(x)$ with

$$G(x)=x-A(x)^{-1}F(x). \qquad (1.1)$$

for nonsingular $A(x)\in L(X,Y)$. We study the differentiability of such a mapping G in the next lemma:

Lemma 1.1. *Assume that $F:D\longrightarrow Y$, $D\subset X$ open, is F-differentiable at$x^*\in D$ with $F(x^*)=0$. Let $A:D\longrightarrow L(X,Y)$ be continuous at x^* with nonsingular $A(x^*)$. Then there is a neighborhood $U\subset D$ of x^*, for which*

$$G:U\longrightarrow X,\quad G(x)=x-A(x)^{-1}F(x)$$

is well-defined. Moreover, G is F-differentiable at x^* differenzierbar and

$$G'(x^*)=I-A(x^*)^{-1}F'(x^*).$$

If A is F-differentiable, the proof is trivial. If A is continuous the proof goes as follows:

Proof. $A(x^*)$ is nonsingular, A is continuous at x^*. Then $A(x)$ is nonsingular in a neighborhood U of x^*. This shows that G is well-defined.

For $x \in U$ we have:

$$\begin{aligned}
&\|G(x) - G(x^*) - [I - A(x^*)^{-1}F'(x^*)](x - x^*)\| \\
&= \|A(x^*)^{-1}F'(x^*)(x - x^*) - A(x)^{-1}F(x)\| \\
&\leq \|[A(x^*)^{-1} - A(x)^{-1}]F'(x^*)(x - x^*)\| + \|A(x)^{-1}[F(x) - F(x^*) - F'(x^*)(x - x^*)]\|.
\end{aligned}$$

For arbitrary $\varepsilon > 0$ there is a $\delta > 0$ such that

$$\|A(x) - A(x^*)\| \leq \varepsilon \quad \text{for all } x \in \overline{K}(x^*, \delta)$$

and

$$\|F(x) - F(x^*) - F'(x^*)(x - x^*)\| \leq \varepsilon\|x - x^*\| \quad \text{f'ur alle } x \in \overline{K}(x^*, \delta).$$

Then we have

$$\begin{aligned}
\|A(x^*)^{-1} - A(x)^{-1}\| &= \|A(x)^{-1}[A(x) - A(x^*)]A(x^*)^{-1}\| \\
&\leq \beta\,\varepsilon\,\|A(x)^{-1}\|
\end{aligned}$$

with $\beta = \|A(x^*)^{-1}\|$ and

$$\begin{aligned}
\|A(x)^{-1}\| &= \|\,[A(x^*) + \Delta A]^{-1}\,\| = \|\,\left[I + A(x^*)^{-1}\Delta A\right]^{-1} A(x^*)^{-1}\| \\
&\leq \frac{\|A(x^*)^{-1}\|}{1 - \|A(x^*)^{-1}\Delta A\|} \leq \frac{\beta}{1 - \beta\,\varepsilon}
\end{aligned}$$

Therefore,

$$\begin{aligned}
&\|G(x) - G(x^*) - [I - A(x^*)^{-1}F'(x^*)](x - x^*)\| \\
&\leq \left(\frac{\beta^2\varepsilon}{1 - \beta\varepsilon}\|F'(x^*)\| + \frac{\beta\varepsilon}{1 - \beta\varepsilon}\right)\|x - x^*\|,
\end{aligned}$$

which implies that $I - A(x^*)^{-1}F'(x^*)$ is the F-derivative of G at x^*. $\square$

This lemma shows that

$$\rho(I - A(x^*)^{-1}F'(x^*)) < 1.$$

implies the local convergence of the corresponding one-step method. If

$$A(x^*) = F'(x^*),$$

we have $G'(x^*) = 0$, which implies the q-superlinear convergence. One way to guarantee this condition is the choice

$$A(x) = F'(x).$$

This leads to Newton's method

$$x^{k+1} = x^k - F'(x^k)^{-1} F(x^k).$$