3

Classical Derivatives

3.1 Directional Derivatives

Convention. Throughout this chapter, unless otherwise specified, we assume that E and F are normed vector spaces, $D \subseteq E$ is nonempty and open, $\bar{x} \in D$, and $f : D \to F$.

We will recall some classical concepts and facts. To start with, we consider directional derivatives. We write

$$\Delta f(\bar{x}, y) := f(\bar{x} + y) - f(\bar{x}) \quad \forall\, y \in D - \bar{x}.$$

We use the following abbreviations: G-derivative for *Gâteaux derivative*, H-derivative for *Hadamard derivative*, F-derivative for *Fréchet derivative.*

Definition 3.1.1 Let $y \in E$. We call

$$\begin{aligned}
f_G(\bar{x}, y) &:= \lim_{\tau \downarrow 0} \tfrac{1}{\tau} \Delta f(\bar{x}, \tau y) && \textit{directional G-derivative,}\\
f_G^s(\bar{x}, y) &:= \lim_{\substack{\tau \downarrow 0\\ x \to \bar{x}}} \tfrac{1}{\tau} \Delta f(x, \tau y) && \textit{strict directional G-derivative,}\\
f_H(\bar{x}, y) &:= \lim_{\substack{\tau \downarrow 0\\ z \to y}} \tfrac{1}{\tau} \Delta f(\bar{x}, \tau z) && \textit{directional H-derivative,}\\
f_H^s(\bar{x}, y) &:= \lim_{\substack{\tau \downarrow 0\\ x \to \bar{x}\\ z \to y}} \tfrac{1}{\tau} \Delta f(x, \tau z) && \textit{strict directional H-derivative}
\end{aligned}$$

of f at $\bar{x}$ in the direction y, provided the respective limit exists.

Lemma 3.1.2

(a) *If $f_H(\bar{x}, y)$ exists, then $f_G(\bar{x}, y)$ also exists and both directional derivatives coincide.*
(b) *If f is locally L-continuous around $\bar{x}$, then $f_H(\bar{x}, y)$ exists if and only if $f_G(\bar{x}, y)$ exists.*

Proof. (a) is obvious. We verify (b). Assume that $f_G(\bar{x}, y)$ exists. Let $\epsilon > 0$ be given. Then there exists $\delta_1 > 0$ such that

$$\left\| \tfrac{1}{\tau} \Delta f(\bar{x}, \tau y) - f_G(\bar{x}, y) \right\| < \tfrac{\epsilon}{2} \quad \text{whenever } 0 < \tau < \delta_1.$$

Since f is locally L-continuous around $\bar{x}$, there further exist $\lambda > 0$ and $\delta_2 > 0$ such that

$$\|f(x_1) - f(x_2)\| \le \lambda \|x_1 - x_2\| \quad \forall x_1, x_2 \in \mathrm{B}(\bar{x}, \delta_2).$$

Now set

$$\delta_3 := \frac{\epsilon}{2\lambda} \quad \text{and} \quad \delta_4 := \min\left\{\delta_1, \frac{\delta_2}{\delta_3 + \|y\|}\right\}.$$

If $z \in \mathrm{B}(y, \delta_3)$ and $0 < \tau < \delta_4$, then we obtain $\|\tau y\| < \delta_2$ and

$$\|\tau z\| \le \tau(\|z - y\| + \|y\|) \le \tau(\delta_3 + \|y\|) \le \delta_2$$

and so

$$\begin{aligned} &\left\| \tfrac{1}{\tau} \Delta f(\bar{x}, \tau z) - f_G(\bar{x}, y) \right\| \\ &\le \tfrac{1}{\tau} \left\| f(\bar{x} + \tau z) - f(\bar{x} + \tau y) \right\| + \left\| \tfrac{1}{\tau} \Delta f(\bar{x}, \tau y) - f_G(\bar{x}, y) \right\| \\ &\le \lambda \|z - y\| + \tfrac{\epsilon}{2} < \epsilon. \end{aligned}$$

We conclude that $f_H(\bar{x}, y)$ exists and equals $f_G(\bar{x}, y)$. □

Lemma 3.1.3 *If the directional H-derivative $f_H(\bar{x}, \cdot)$ exists in a neighborhood of $y_0 \in E$, then it is continuous at y_0.*

Proof. Let $\rho_0 > 0$ be such that $f_H(\bar{x}, y)$ exists for each $y \in \mathrm{B}(y_0, \rho_0)$. Let $\epsilon > 0$ be given. Then there exists $\rho \in (0, \rho_0)$ such that

$$\left\| \tfrac{1}{\tau} \Delta f(\bar{x}, \tau y) - f_H(\bar{x}, y_0) \right\| \le \epsilon \quad \text{whenever } 0 < \tau < \rho \text{ and } y \in \mathrm{B}(y_0, \rho).$$

Letting $\tau \downarrow 0$, we obtain

$$\|f_G(\bar{x}, y) - f_H(\bar{x}, y_0)\| \le \epsilon \quad \forall\, y \in \mathrm{B}(y_0, \rho).$$

By Lemma 3.1.2 we have $f_G(\bar{x}, y) = f_H(\bar{x}, y)$ for each $y \in \mathrm{B}(y_0, \rho_0)$, and the assertion follows. □

Now we consider a proper function $f : E \to \overline{\mathbb{R}}$. If $D := \operatorname{int} \operatorname{dom} f$ is nonempty, then of course the above applies to $f|D$. In addition, we define the following directional derivatives:

$$\begin{aligned} \overline{f}_G(\bar{x}, y) &:= \limsup_{\tau \downarrow 0} \tfrac{1}{\tau} \Delta f(\bar{x}, \tau y) && \textit{upper directional G-derivative,} \\ \underline{f}_G(\bar{x}, y) &:= \liminf_{\tau \downarrow 0} \tfrac{1}{\tau} \Delta f(\bar{x}, \tau y) && \textit{lower directional G-derivative,} \\ \overline{f}_H(\bar{x}, y) &:= \limsup_{\substack{\tau \downarrow 0 \\ z \to y}} \tfrac{1}{\tau} \Delta f(\bar{x}, \tau z) && \textit{upper directional H-derivative,} \\ \underline{f}_H(\bar{x}, y) &:= \liminf_{\substack{\tau \downarrow 0 \\ z \to y}} \tfrac{1}{\tau} \Delta f(\bar{x}, \tau z) && \textit{lower directional H-derivative.} \end{aligned}$$

Notice that these directional derivatives generalize the *Dini derivates* of a function $f : I \to \mathbb{R}$ ($I \subseteq \mathbb{R}$ an interval) which are defined at $\bar{x} \in \operatorname{int} I$ by

$$D^+ f(\bar{x}) := \limsup_{h \downarrow 0} \frac{f(\bar{x}+h) - f(\bar{x})}{h}, \quad D_+ f(\bar{x}) := \liminf_{h \downarrow 0} \frac{f(\bar{x}+h) - f(\bar{x})}{h},$$
$$D^- f(\bar{x}) := \limsup_{h \uparrow 0} \frac{f(\bar{x}+h) - f(\bar{x})}{h}, \quad D_- f(\bar{x}) := \liminf_{h \uparrow 0} \frac{f(\bar{x}+h) - f(\bar{x})}{h}.$$

If $\bar{x} \in I$ is the left boundary point of I, then $D^+ f(\bar{x})$ and $D_+ f(\bar{x})$ still make sense; an analogous remark applies to the right boundary point of I. Notice that, among others, $D^+ f(\bar{x}) = \overline{f}_G(\bar{x}, 1)$. If $D^+ f(\bar{x}) = D_+ f(\bar{x})$, then this common value is called the *right derivative* of f at $\bar{x}$ and is denoted $f'_+(\bar{x})$. The *left derivative* $f'_-(\bar{x})$ is defined analogously.

3.2 First-Order Derivatives

Our aim in this section is to recall various kinds of derivatives. For this, the following notion will be helpful.

Definition 3.2.1 A nonempty collection β of subsets of E is called *bornology* if the following holds:

$$\begin{aligned}
&\text{each } S \in \beta \text{ is bounded and } \bigcup_{S \in \beta} S = E,\\
&S \in \beta \quad \Longrightarrow \quad -S \in \beta,\\
&S \in \beta \text{ and } \lambda > 0 \quad \Longrightarrow \quad \lambda S \in \beta,\\
&S_1, S_2 \in \beta \quad \Longrightarrow \quad \exists S \in \beta : S_1 \cup S_2 \subset S.
\end{aligned}$$

In particular:

- The *G-bornology* β_G is the collection of all finite sets.
- The *H-bornology* β_H is the collection of all compact sets.
- The *F-bornology* β_F is the collection of all bounded sets.

We set

$$\mathrm{L}(E, F) := \text{ vector space of all continuous linear mappings } T : E \to F.$$

Definition 3.2.2 Let β be a bornology on E.

(a) The mapping $f : D \to F$ is said to be *β-differentiable* at $\bar{x}$ if there exists $T \in \mathrm{L}(E, F)$, the β-derivative of f at $\bar{x}$, such that

$$\lim_{\tau \to 0} \sup_{y \in S} \left\| \tfrac{1}{\tau}\big(f(\bar{x} + \tau y) - f(\bar{x})\big) - T(y) \right\| = 0 \quad \forall\, S \in \beta. \tag{3.1}$$

(b) The mapping $f : D \to F$ is said to be *strictly β-differentiable* at $\bar{x}$ if there exists $T \in \mathrm{L}(E, F)$, the *strict β-derivative* of f at $\bar{x}$, such that

$$\lim_{\substack{\tau \to 0 \\ x \to \bar{x}}} \sup_{y \in S} \left\| \tfrac{1}{\tau}\big(f(x + \tau y) - f(x)\big) - T(y) \right\| = 0 \quad \forall\, S \in \beta. \tag{3.2}$$

(c) In particular, f is said to be *G-differentiable* or *strictly G-differentiable* if (3.1) or (3.2), respectively, holds with $\beta = \beta_G$. In this case, T is called *(strict) G-derivative* of f at $\bar{x}$. Analogously we use *(strictly) H-differentiable* if $\beta = \beta_H$ and *(strictly) F-differentiable* if $\beta = \beta_F$. In the respective case, T is called *(strict) H-derivative* or *(strict) F-derivative* of f at $\bar{x}$.

Remark 3.2.3

(a) If the β-derivative T of f at $\bar{x}$ exists for some bornology β, then

$$T(y) = \lim_{\tau \to 0} \tfrac{1}{\tau}\big(f(\bar{x} + \tau y) - f(\bar{x})\big) = f_G(\bar{x}, y) \quad \forall\, y \in E.$$

Hence if two of the above derivatives exist, then they coincide. This justifies denoting them by the same symbol; we choose

$$f'(\bar{x}) := T.$$

Condition (3.1) means that we have

$$\lim_{\tau \to 0} \left\| \tfrac{1}{\tau}\big(f(\bar{x}+\tau y) - f(\bar{x})\big) - f'(\bar{x})y \right\| = 0 \ \text{ uniformly in } y \in S \text{ for each } S \in \beta.$$

An analogous remark applies to (3.2). Here and in the following we write $f'(\bar{x})y$ instead of $f'(\bar{x})(y)$. If $f : D \to \mathbb{R}$, then $f'(\bar{x}) \in E^*$ and as usual we also write $\langle f'(\bar{x}), y\rangle$ instead of $f'(\bar{x})(y)$.

(b) Now let E be a (real) Hilbert space with inner product $(x \,|\, y)$ and $f : E \to \mathbb{R}$ a functional. If f is G-differentiable at $\bar{x} \in E$, then the G-derivative $f'(\bar{x})$ is an element of the dual space E^*. By the Riesz representation theorem, there is exactly one $z \in E$ such that $\langle f'(\bar{x}), y\rangle = (z \,|\, y)$ for all $y \in E$. This element z is called gradient of f at $\bar{x}$ and is denoted $\nabla f(\bar{x})$. In other words, we have

$$\langle f'(\bar{x}), y\rangle = (\nabla f(\bar{x}) \,|\, y) \quad \forall\, y \in E.$$

Proposition 3.2.4 says that f is G-differentiable at $\bar{x}$ if and only if the directional G-derivative $y \mapsto f_G(\bar{x}, y)$ exists and is linear and continuous on E. An analogous remark applies to strict G-differentiability as well as (strict) H-differentiability. Recall that if $g : E \to F$, then

$$g(x) = \mathbf{o}(\|x\|),\ x \to o \quad \text{means} \quad \lim_{x \to o} \frac{g(x)}{\|x\|} = o.$$

Proposition 3.2.4

(i) *f is G-differentiable at $\bar{x}$ if and only if there exists $f'(\bar{x}) \in \mathrm{L}(E,F)$ such that $f'(\bar{x})y = f_G(\bar{x}, y)$ for all $y \in E$.*
(ii) *f is H-differentiable at $\bar{x}$ if and only if there exists $f'(\bar{x}) \in \mathrm{L}(E,F)$ such that $f'(\bar{x})y = f_H(\bar{x}, y)$ for all $y \in E$.*
(iii) *The following assertions are equivalent:*
  (a) *f is strictly H-differentiable at $\bar{x}$.*
  (b) *There exists $f'(\bar{x}) \in \mathrm{L}(E,F)$ such that $f'(\bar{x})y = f^s_H(\bar{x}, y)$ for $y \in E$.*
  (c) *f is locally L-continuous around $\bar{x}$ and strictly G-differentiable at $\bar{x}$.*
(iv) *f is F-differentiable at $\bar{x}$ if and only if there exists $f'(\bar{x}) \in \mathrm{L}(E,F)$ such that*

$$\big(f(\bar{x}+z) - f(\bar{x})\big) - f'(\bar{x})z = \mathbf{o}(\|z\|),\ z \to o.$$

(v) *f is strictly F-differentiable at $\bar{x}$ if and only if there exists $f'(\bar{x}) \in \mathrm{L}(E,F)$ such that*

$$\big(f(x+z) - f(x)\big) - f'(\bar{x})z = \mathbf{o}(\|z\|),\ z \to o,\ x \to \bar{x}.$$

Proof. We only verify (iii), leaving the proof of the remaining assertions as Exercise 3.8.4.
(a) $\Longrightarrow$ (c): Let (a) hold. We only have to show that f is locally L-continuous around $\bar{x}$. Assume this is not the case. Then there exist sequences (x_n) and (x'_n) in $\mathrm{B}(\bar{x}, \frac{1}{n})$ such that

$$\|f(x_n) - f(x'_n)\| > n\|x_n - x'_n\| \quad \forall\, n \in \mathbb{N}. \tag{3.3}$$

Setting $\tau_n := \sqrt{n}\|x_n - x'_n\|$ and $y_n := \frac{1}{\tau_n}(x'_n - x_n)$, we obtain as $n \to \infty$,

$$0 \le \tau_n \le \sqrt{n}(\|x_n - \bar{x}\| + \|\bar{x} - x'_n\|) < \frac{2}{\sqrt{n}} \to 0 \quad \text{and} \quad \|y_n\| = \frac{1}{\sqrt{n}} \to 0.$$

By (3.3) we have

$$\big\|\tfrac{1}{\tau_n}\Delta f(x_n, \tau_n y_n)\big\| > \tfrac{1}{\tau_n}\cdot n\|\tau_n y_n\| = \sqrt{n} \quad \forall\, n \in \mathbb{N},$$

and the continuity of $f'(\bar{x})$ implies that, with some $n_0 \in \mathbb{N}$, we obtain $\|f'(\bar{x})\|\cdot\|y_n\| < \frac{1}{2}$ for each $n > n_0$. It follows that

$$\begin{aligned}&\big\|\tfrac{1}{\tau_n}\Delta f(x_n, \tau_n y_n) - f'(\bar{x})y_n\big\|\\ &\ge \big\|\tfrac{1}{\tau_n}\Delta f(x_n, \tau_n y_n)\big\| - \|f'(\bar{x})\|\cdot\|y_n\| > \sqrt{n} - \tfrac{1}{2} \quad \forall\, n > n_0,\end{aligned}$$

which contradicts (3.2) for the compact set $S := \{o\} \cup \{y_n \,|\, n > n_0\}$.
(c) $\Longrightarrow$ (b): Let $y \in E$ and $\epsilon > 0$ be given. Since f is strictly G-differentiable at $\bar{x}$, there exists $\delta_1 > 0$ such that

$$\big\|\tfrac{1}{\tau}\Delta f(x, \tau y) - f'(\bar{x})y\big\| < \epsilon \quad \text{whenever } 0 < |\tau| < \delta_1,\ \|x - \bar{x}\| < \delta_1. \tag{3.4}$$

Since f is locally L-continuous around $\bar{x}$, there further exist $\lambda > 0$ and $\delta_2 > 0$ such that

$$\|f(x_1) - f(x_2)\| < \lambda \|x_1 - x_2\| \quad \forall\, x_1, x_2 \in \mathrm{B}(\bar{x}, \delta_2). \tag{3.5}$$

Setting $x_1 := x + \tau z$ and $x_2 := x + \tau y$, we have the estimates

$$\|x_1 - \bar{x}\| \le \|x - \bar{x}\| + |\tau|(\|z - y\| + \|y\|) \quad \text{and} \quad \|x_2 - \bar{x}\| \le \|x - \bar{x}\| + |\tau|\|y\|$$

which show that $x_1, x_2 \in \mathrm{B}(\bar{x}, \delta_2)$ provided $|\tau|$, $\|z - y\|$, and $\|x - \bar{x}\|$ are sufficiently small. Under this condition, (3.4) and (3.5) imply that

$$\begin{aligned} &\left\| \tfrac{1}{\tau} \Delta f(x, \tau z) - f'(\bar{x}) y \right\| \\ &\le \tfrac{1}{|\tau|} \left\| f(x + \tau z) - f(x + \tau y) \right\| + \left\| \tfrac{1}{\tau} \Delta f(x, \tau y) - f'(\bar{x}) y \right\| \le \lambda \|z - y\| + \epsilon. \end{aligned}$$

This verifies (b).
(b) $\Longrightarrow$ (a): Let (b) hold and assume that (a) does not hold. Let $T \in \mathrm{L}(E, F)$ be given. Then for some compact subset S of E, the relation (3.2) does not hold. Hence there exist $\epsilon_0 > 0$ as well as sequences $\tau_n \downarrow 0$, $y_n \in S$, and $x_n \to \bar{x}$ such that

$$\left\| \tfrac{1}{\tau} \Delta f(x_n, \tau_n y_n) - T(y_n) \right\| > \epsilon_0 \quad \forall\, n \in \mathbb{N}.$$

Since S is compact, a subsequence of (y_n), again denoted (y_n), converges to some $y \in S$. It follows that for any $n > n_0$ we have

$$\begin{aligned} &\left\| \tfrac{1}{\tau} \Delta f(x_n, \tau_n y_n) - T(y) \right\| \\ &\ge \underbrace{\left\| \tfrac{1}{\tau} \Delta f(x_n, \tau_n y_n) - T(y_n) \right\|}_{>\ \epsilon_0} - \underbrace{\|T\| \cdot \|y_n - y\|}_{<\ \epsilon_0/2} > \frac{\epsilon_0}{2}; \end{aligned} \tag{3.6}$$

in this connection, we exploited that T is linear and continuous. However, the relation (3.6) contradicts (b). □

Proposition 3.2.5 *If $f : D \to F$ is H-differentiable at $\bar{x}$, then f is continuous at $\bar{x}$.*

Proof. See Exercise 3.8.5. □

3.3 Mean Value Theorems

We recall a variant of the classical mean value theorem (see, for instance, Walter [212]).

Proposition 3.3.1 (Mean Value Theorem in Terms of Dini Derivates) *Let I and J be intervals in $\mathbb{R}$ and let $A \subseteq I$ be a countable set. Further let $f : I \to \mathbb{R}$ be continuous, let $D \in \{D^+, D_+, D^-, D_-\}$, and assume that*

$$Df(x) \in J \quad \forall\, x \in I \setminus A.$$

Then

$$\frac{f(b)-f(a)}{b-a} \in J \quad \forall\, a,b \in I,\, a \neq b.$$

If $f : [a,b] \to \mathbb{R}$ is continuous on $[a,b]$ and differentiable on (a,b), then by the intermediate value theorem for derivatives the set $J := \{f'(x) \,|\, x \in (a,b)\}$ is an interval and so the usual mean value theorem follows from Proposition 3.3.1.

Now we return to the setting described by the convention at the beginning of the chapter. If $x, z \in E$, we write

$$[x,\, z] := \{\lambda x + (1-\lambda)z \,|\, 0 \leq \lambda \leq 1\}.$$

If $f : D \to F$ is G-differentiable on D (i.e., G-differentiable at any $x \in D$), then we may consider the mapping $f' : x \mapsto f'(x)$ of D to $\mathrm{L}(E,F)$.

Definition 3.3.2 Let f be G-differentiable on D. The mapping f' is said to be *radially continuous* if for all $x, y \in E$ such that $[x, x+y] \subseteq D$, the function $\tau \mapsto f'(x+\tau y)y$ is continuous on $[0,1]$.

Proposition 3.3.3 (Mean Value Theorem in Integral Form) *Let $f : D \to \mathbb{R}$ be G-differentiable and let f' be radially continuous. Then for all $x, y \in D$ such that $[x, x+y] \subseteq D$ one has*

$$f(x+y) - f(x) = \int_0^1 \langle f'(x+\tau y), y\rangle \,\mathrm{d}\tau. \tag{3.7}$$

Proof. For $\tau \in [0,1]$ let $\varphi(\tau) := f(x+\tau y)$. By assumption φ is continuously differentiable and $\varphi'(\tau) = \langle f'(x+\tau y), y\rangle$. The main theorem of calculus gives

$$\varphi(1) - \varphi(0) = \int_0^1 \varphi'(\tau)\,\mathrm{d}\tau,$$

which is (3.7). □

The above result is formulated for functionals only, in which case it will be used later. In Proposition 4.3.8 below we shall describe an important class of functionals to which the mean value formula (3.7) applies. We mention that, by an appropriate definition of the Riemann integral, the formula extends to a mapping $f : D \to F$ provided F is a Banach space.

If β is a bornology of E, we denote by $\mathrm{L}_\beta(E,F)$ the vector space $\mathrm{L}(E,F)$ equipped with the topology of uniform convergence on the sets $S \in \beta$. In particular, $\mathrm{L}_{\beta_F}(E,F)$ denotes $\mathrm{L}(E,F)$ equipped with the topology generated by the norm $\|T\| := \sup\{\|Tx\| \mid x \in \mathrm{B}_E\}$. In particular we write $E^*_\beta := \mathrm{L}_\beta(E,\mathbb{R})$.

Proposition 3.3.4 (Mean Value Theorem in Inequality Form) *Let $y \in E$ be such that $[\bar{x}, \bar{x}+y] \subseteq D$ and f is G-differentiable on $[\bar{x}, \bar{x}+y]$. Further let $T \in \mathrm{L}_{\beta_F}(E,F)$. Then one has*

$$\|\big(f(\bar{x}+y) - f(\bar{x})\big) - Ty\| \le \|y\| \sup_{0\le\tau\le 1} \|f'(\bar{x}+\tau y) - T\|.$$

Proof. Set $g(x) := f(x) - T(x-\bar{x}),\ x \in E$. By the Hahn–Banach theorem, there exists $v \in F^*$ satisfying $\|v\| = 1$ and $\big\langle v, \Delta g(\bar{x}, y)\big\rangle = \|\Delta g(\bar{x}, y)\|$. Now define $\varphi(\tau) := \big\langle v, g(\bar{x}+\tau y)\big\rangle,\ \tau \in [0,1]$. It is easy to see that φ is differentiable, and one has

$$\varphi'(\tau) = \big\langle v, g'(\bar{x}+\tau y)y\big\rangle = \big\langle v, f'(\bar{x}+\tau y)y - Ty\big\rangle. \tag{3.8}$$

By the classical mean value theorem, there exists $\tau \in (0,1)$ such that $\varphi'(\tau) = \varphi(1) - \varphi(0)$. This together with (3.8) and

$$\big|\big\langle v, f'(\bar{x}+\tau y)y - Ty\big\rangle\big| \le \|v\|\|f'(\bar{x}+\tau y)y - Ty\| \le \|f'(\bar{x}+\tau y) - T\|\|y\|$$

completes the proof. □

3.4 Relationship between Differentiability Properties

In this section we will study the interrelations between the various differentiability properties. First we introduce some terminology.

Definition 3.4.1

(a) The mapping $f : D \to F$ is said to be *β-smooth* at $\bar{x}$ if f is β-differentiable for any x in an open neighborhood U of $\bar{x}$ and the mapping $f' : x \mapsto f'(x)$ of U to $\mathrm{L}_\beta(E,F)$ is continuous on U.
(b) The mapping $f : D \to F$ is said to be *continuously differentiable* at $\bar{x}$ if f is G-differentiable for any x in an open neighborhood U of $\bar{x}$ and the mapping $f' : x \mapsto f'(x)$ of U to $\mathrm{L}_{\beta_F}(E,F)$ is continuous at $\bar{x}$.
(c) If $f : D \to F$ is continuously differentiable at every point of D, then f is said to be a *C^1-mapping* on D, written $f \in \mathrm{C}^1(D,F)$.

We shall make use of the following abbreviations:

(G): f is G-differentiable at $\bar{x}$,
(SG): f is strictly G-differentiable at $\bar{x}$,
(CD): f is continuously differentiable at $\bar{x}$.

In analogy to (G), (SG) we use (H), (SH), (F), and (SF).

Proposition 3.4.2 *The following implications hold true:*

$$\begin{array}{ccccccc}
(CD) \Longrightarrow & (SF) & \overset{\Longrightarrow}{\longleftarrow} & (SH) & \overset{\Longrightarrow}{\leftarrow -} & (SG) \\
& \Downarrow & & \Downarrow & & \Downarrow \\
& (F) & \overset{\Longrightarrow}{\longleftarrow} & (H) & \overset{\Longrightarrow}{\leftarrow -} & (G)
\end{array}$$

In this connection, $\longleftarrow$ *means* implication provided E is finite dimensional*, and* $\leftarrow -$ *means* implication provided f is locally L-continuous around $\bar{x}$.

Proof. In view of the foregoing, it only remains to verify the implication (CD) $\Longrightarrow$ (SF). Thus let (CD) hold. Then there exists $\rho > 0$ such that f is G-differentiable on $\mathrm{B}(\bar{x}, 2\rho)$. If $x \in \mathrm{B}(\bar{x}, \rho)$ and $y \in \mathrm{B}(o, \rho)$, then $[x, x + y] \subset \mathrm{B}(\bar{x}, 2\rho)$. By Proposition 3.3.4 with $T := f'(\bar{x})$, we obtain

$$\|f(x + y) - f(x) - f'(\bar{x})y\| \le \|y\| \sup_{0 \le \tau \le 1} \|f'(x + \tau y) - f'(\bar{x})\|. \tag{3.9}$$

Now let $\epsilon > 0$ be given. Since f' is continuous at $\bar{x}$, there exists $\delta > 0$ such that

$$\sup_{0 \le \tau \le 1} \|f'(x + \tau y) - f'(\bar{x})\| < \epsilon \quad \forall\, x \in \mathrm{B}(\bar{x}, \delta) \quad \forall\, y \in \mathrm{B}(o, \delta).$$

This together with (3.9) implies (SF). □

Remark 3.4.3 By Proposition 3.4.2 it is clear that if f is continuously differentiable on an open neighborhood U of $\bar{x}$, then f is β-smooth at any $\bar{x} \in U$ for any bornology $\beta \subseteq \beta_F$. In particular, f is F-differentiable at any $\bar{x} \in U$ and the F-derivative f' is continuous from U to $\mathrm{L}_{\beta_F}(E, F)$.

Beside E and F let G be another normed vector space. Beside $f : D \to F$ let $g : V \to G$ be another mapping, where V is an open neighborhood of $\bar{z} := f(\bar{x})$ in F. Assume that $f(D) \subset V$. Then the composition $g \circ f : D \to G$ is defined.

Proposition 3.4.4 (Chain Rule) *Assume that f and g are H-differentiable at $\bar{x}$ and $\bar{z}$, respectively. Then $g \circ f$ is H-differentiable at $\bar{x}$, and there holds*

$$(g \circ f)'(\bar{x}) = g'(\bar{z}) \circ f'(\bar{x}).$$

An analogous statement holds true if H-differentiable is replaced by F-differentiable.

The *proof* is the same as in multivariate calculus. An analogous chain rule for G-differentiable mappings does not hold (see Exercise 3.8.3).

3.5 Higher-Order Derivatives

We again use the notation introduced at the beginning of the chapter. Assume that $f \in \mathrm{C}^1(D, F)$. If the (continuous) mapping $f' : D \to \mathrm{L}_{\beta_F}(E, F)$ is continuously differentiable on D, then f is said to be a *twice continuously differentiable mapping* on D, or a C^2*-mapping* on D, with *second-order derivative* $f'' := (f')'$. The set of all twice continuously differentiable mappings $f : D \to F$ is denoted $\mathrm{C}^2(D, F)$.

Notice that f'' maps D into $H := \mathrm{L}_{\beta_F}\big(E, \mathrm{L}_{\beta_F}(E, F)\big)$. Parallel to H we consider the vector space $\mathrm{B}(E, F)$ of all continuous bilinear mappings $b : E \times E \to F$, which is normed by

$$\|b\| := \sup\{\|b(y, z)\| \mid \|y\| \le 1,\ \|z\| \le 1\}. \tag{3.10}$$

If $h \in H$, then

$$b_h(y, z) := h(y)z \quad \forall\, (y, z) \in E \times E$$

defines an element $b_h \in \mathrm{B}(E, F)$. Conversely, given $b \in \mathrm{B}(E, F)$, set

$$h(y) := b(y, \cdot) \quad \forall\, y \in E.$$

Then $h \in H$ and $b_h = b$. Evidently the mapping $h \mapsto b_h$ is an isomorphism of H onto $\mathrm{B}(E, F)$. Therefore H can be identified with $\mathrm{B}(E, F)$. In this sense, we interpret $f''(\bar{x})$ as an element of $\mathrm{B}(E, F)$ and write $f''(\bar{x})(y, z)$ instead of $\big(f''(\bar{x})y\big)z$. If, in particular, $f \in \mathrm{C}^2(D, \mathbb{R})$, then $f''(\bar{x})$ is a continuous bilinear form on $E \times E$.

Proposition 3.5.1 (Taylor Expansion) *Assume that D is open and $f \in \mathrm{C}^2(D, \mathbb{R})$. Then for all $\bar{x} \in D$, $y \in D - \bar{x}$ one has*

$$f(\bar{x} + y) = f(\bar{x}) + \langle f'(\bar{x}), y\rangle + \tfrac{1}{2} f''(\bar{x})(y, y) + r(y), \quad \textit{where } \lim_{y \to o} \frac{r(y)}{\|y\|^2} = o.$$

In particular, there exist $\sigma > 0$ and $\epsilon > 0$ such that

$$f(\bar{x} + y) \ge f(\bar{x}) + \langle f'(\bar{x}), y\rangle - \sigma\|y\|^2 \quad \forall\, y \in \mathrm{B}(o, \epsilon). \tag{3.11}$$

Proof. The first assertion follows readily from the classical Taylor expansion of the function $\varphi(\tau) := f(\bar{x} + \tau y)$, $\tau \in [0, 1]$. From the first result we obtain (3.11) since in view of (3.10) we have

$$|\tfrac{1}{2} f''(\bar{x})(y, y)| \le \tfrac{1}{2}\|f''(\bar{x})\|\,\|y\|^2 \quad \forall\, y \in E,$$

and the limit property of r entails the existence of $\kappa > 0$ such that $|r(y)| \le \kappa\|y\|^2$ if $\|y\|$ is sufficiently small. □

We only mention that in an analogous manner, derivatives of arbitrary order n, where $n \in \mathbb{N}$, can be defined using n-linear mappings, which leads to higher-order Taylor expansions.

3.6 Some Examples

For illustration and later purposes we collect some examples. Further examples are contained in the exercises.

Example 3.6.1 Let E be a normed vector space and $a : E\times E \to \mathbb{R}$ a bilinear functional. Recall that a is said to be symmetric if $a(x,y) = a(y,x)$ for all $x,y \in E$, and a is said to be bounded if there exists $\kappa > 0$ such that

$$|a(x,y)| \le \kappa\|x\|\,\|y\| \quad \forall\, x,y \in E.$$

Consider the quadratic functional $f : E \to \mathbb{R}$ defined by

$$f(x) := \tfrac{1}{2}a(x,x), \quad x \in E,$$

where a is bilinear, symmetric, and bounded. It is left as Exercise 3.8.6 to show that f is continuously differentiable on E and to calculate the derivative. In particular, if E is a Hilbert space with inner product $(x\,|\,y)$, then the functional

$$g(x) := \frac{1}{2}\|x\|^2 = \frac{1}{2}(x\,|\,x), \quad x \in E,$$

is continuously differentiable on E with $\langle g'(x), y\rangle = (x\,|\,y)$ for all $x,y \in E$. Hence $\nabla g(x) = x$ for any $x \in E$. Finally, concerning the norm functional $\omega(x) := \|x\| = \sqrt{2g(x)}$, the chain rule gives $\nabla\omega(x) = \frac{x}{\|x\|}$ for any $x \ne o$.

Example 3.6.2 Let again E denote a Hilbert space with inner product $(x\,|\,y)$ and define $g : E \to \mathbb{R}$ by

$$g(x) := \big(\delta^2 + 2\delta(u\,|\,x-\bar{x}) - \|x-\bar{x}\|^2\big)^{1/2},$$

where the positive constant δ and the element $u \in E$ are fixed. Choose $\epsilon > 0$ such that the term $(\cdots)$ is positive for each $x \in \mathring{\mathrm{B}}(\bar{x},\epsilon)$. Define $\psi : (0,+\infty) \to \mathbb{R}$ by $\psi(z) := z^{1/2}$ and $\varphi : \mathring{\mathrm{B}}(\bar{x},\epsilon) \to \mathbb{R}$ by

$$\varphi(x) := \delta^2 + 2\delta(u\,|\,x-\bar{x}) - \|x-\bar{x}\|^2.$$

Then we have $g = \psi\circ\varphi$, and the chain rule implies

$$(g'(x)\,|\,y) = \frac{\delta(u\,|\,y)}{\big(\delta^2 + 2\delta(u\,|\,x-\bar{x}) - \|x-\bar{x}\|^2\big)^{1/2}} \quad \forall\, x \in \mathring{\mathrm{B}}(\bar{x},\epsilon) \quad \forall\, y \in E.$$

In particular, $(g'(\bar{x})\,|\,y) = (u\,|\,y)$ for all $y \in E$, which means $\nabla g(\bar{x}) = u$. Moreover, it is easy to see that g is a C^2-mapping on $\mathring{\mathrm{B}}(\bar{x},\epsilon)$. This example will be used later in connection with proximal subdifferentials.

In view of Example 3.6.3, recall that an *absolutely continuous* function $x : [a,b] \to \mathbb{R}$ is differentiable almost everywhere, i.e., outside a Lebesgue null set $N \subseteq [a,b]$. Setting $\dot{x}(t) := 0$ for each $t \in N$, which we tacitly assume from now on, the function $\dot{x} : [a,b] \to \mathbb{R}$ belongs to $\mathcal{L}^1[a,b]$ and one has $\int_{[a,b]} \dot{x}(t)\,\mathrm{d}t = x(b) - x(a)$. In this connection, also recall that $\mathcal{L}^p[a,b]$, where $p \in [1,+\infty)$, denotes the vector space of all Lebesgue measurable functions $g : [a,b] \to \mathbb{R}$ such that $|g|^p$ is Lebesgue integrable over $[a,b]$. In addition, $\mathcal{L}^\infty[a,b]$ denotes the vector space of all Lebesgue measurable functions $g : [a,b] \to \mathbb{R}$ such that $\operatorname{ess\,sup}_{x\in[a,b]}|g(x)| < +\infty$. We denote by $\mathrm{AC}^\infty[a,b]$ the vector space of all absolutely continuous functions $x : [a,b] \to \mathbb{R}$ such that $\dot{x} \in \mathcal{L}^\infty[a,b]$. Notice that $\mathrm{AC}^\infty[a,b]$ is a Banach space with respect to the norm

$$\|x\|_{1,\infty} := \max\{\|x\|_\infty, \|\dot{x}\|_\infty\}.$$

Example 3.6.3 Let $E := \mathrm{AC}^\infty[a,b]$, where $a < b$, and consider the *variational functional*

$$f(x) := \int_a^b \varphi\big(t, x(t), \dot{x}(t)\big)\,\mathrm{d}t \quad \forall\, x \in \mathrm{AC}^\infty[a,b].$$

If $\bar{x} \in \mathrm{AC}^\infty[a,b]$ is fixed, we write $\overline{\varphi}(t) := \varphi\big(t, \bar{x}(t), \dot{\bar{x}}(t)\big)$ for any $t \in [a,b]$. Assume that the function $(t,x,v) \mapsto \varphi(t,x,v)$ is continuous on $[a,b]\times\mathbb{R}\times\mathbb{R}$ and has continuous first-order partial derivatives with respect to x and v there. We shall show that the functional f is continuously differentiable at any $\bar{x} \in \mathrm{AC}^\infty[a,b]$ and that

$$\langle f'(\bar{x}), y\rangle = \int_a^b \big(\overline{\varphi}_x(t)\cdot y(t) + \overline{\varphi}_v(t)\cdot \dot{y}(t)\big)\,\mathrm{d}t \quad \forall\, y \in \mathrm{AC}^\infty[a,b]. \tag{3.12}$$

Proof.

(I) *The directional G-derivative $f_G(\bar{x},y)$ exists for all $\bar{x}, y \in \mathrm{AC}^\infty[a,b]$.* In fact, we have

$$\begin{aligned} f_G(\bar{x},y) &= \frac{\partial}{\partial\tau} f(\bar{x}+\tau y)\Big|_{\tau=0} = \int_a^b \frac{\partial}{\partial\tau}\Big[\varphi(t,\ \bar{x}(t)+\tau y(t),\ \dot{\bar{x}}(t)+\tau\dot{y}(t))\Big]\Big|_{\tau=0}\mathrm{d}t \\ &= \int_a^b [\bar{\varphi}_x(t)y(t) + \bar{\varphi}_v(t)\dot{y}(t)]\,\mathrm{d}t. \end{aligned}$$

Notice that the assumptions on φ and $\bar{x}$ imply that the integrand in the last term is bounded, which allows differentiating under the integral sign.

(II) It is easy to verify that the functional $y \mapsto f_G(\bar{x},y)$ is linear and continuous. Hence the G-derivative is given by (3.12).

(III) *f is continuously differentiable at any $\bar{x} \in \mathrm{AC}^\infty[a,b]$.* For arbitrary $x, \bar{x}, y \in \mathrm{AC}^\infty[a,b]$ we have

$$\begin{aligned} &[f'(x) - f'(\bar{x})]y \\ &= \int_a^b [\varphi_x(t,x,\dot{x}) - \varphi_x(t,\bar{x},\dot{\bar{x}})]y\,\mathrm{d}t + \int_a^b [\varphi_v(t,x,\dot{x}) - \varphi_v(t,\bar{x},\dot{\bar{x}})]\dot{y}\,\mathrm{d}t \end{aligned}$$

and so

$$\begin{aligned}\|f'(x) - f'(\bar{x})\| &= \sup_{\|y\|_{1,\infty}\le 1} |[f'(x) - f'(\bar{x})]y| \\ &\le \int_a^b |\varphi_x(t,x,\dot{x}) - \varphi_x(t,\bar{x},\dot{\bar{x}})|\,\mathrm{d}t + \int_a^b |\varphi_v(t,x,\dot{x}) - \varphi_v(t,\bar{x},\dot{\bar{x}})|\,\mathrm{d}t, \\ &< \epsilon \quad \text{if } \|x-\bar{x}\|_{1,\infty} \text{ is sufficiently small.}\end{aligned}$$

Justification of the last line: According to hypothesis, φ_x and φ_v are continuous on $[a,b]\times\mathbb{R}\times\mathbb{R}$, hence uniformly continuous on the compact set

$$\{(t,\xi,\zeta)\in\mathbb{R}^3 \mid t\in[a,b],\ |\xi-\bar{x}(t)|\le 1,\ |\zeta-\dot{\bar{x}}(t)|\le 1\}.$$

Thus, for each $\epsilon>0$ there exists $\delta\in(0,1)$ such that

$$|\varphi_x(t,x(t),\dot{x}(t)) \;-\; \varphi_x(t,\bar{x}(t),\dot{\bar{x}}(t))| \;<\; \frac{\epsilon}{2(b-a)}$$

whenever $t\in[a,b]$, $|x(t)-\bar{x}(t)|\le\delta$, and $|\dot{x}(t)-\dot{\bar{x}}(t)|\le\delta$. An analogous estimate holds for φ_v. □

3.7 Implicit Function Theorems and Related Results

Now we make the following assumptions:

(A) E, F, and G are normed vector spaces.
U and V are open neighborhoods of $\bar{x}\in E$ and $\bar{y}\in F$, respectively.
$f: U\times V\to G$.

Define $g_1: U\to G$ by $g_1(x) := f(x,\bar{y})$, $x\in U$. We denote the derivative (in the sense of Gâteaux, Hadamard, or Fréchet) of g_1 at $\bar{x}$, whenever it exists, by $f_{|1}(\bar{x},\bar{y})$ or by $D_1f(\bar{x},\bar{y})$ and call it *partial derivative* of f, with respect to the first variable, at $(\bar{x},\bar{y})$. Notice that $f_{|1}(\bar{x},\bar{y})$ is an element of $\mathrm{L}(E,G)$. If $f_{|1}(x,y)$ exists, say, for all $(x,y)\in U\times V$, then

$$f_{|1}: (x,y)\mapsto f_{|1}(x,y), \quad (x,y)\in U\times V,$$

defines the mapping $f_{|1}: U\times V\to \mathrm{L}(E,G)$. An analogous remark applies to $f_{|2}(x,y)$ and $D_2f(\bar{x},\bar{y})$.

As in classical multivariate calculus, we have the following relationship.

Proposition 3.7.1 *Let the assumptions* (A) *be satisfied.*

(a) *If f is G-differentiable at $(\bar{x},\bar{y})$, then the partial G-derivatives $f_{|1}(\bar{x},\bar{y})$ and $f_{|2}(\bar{x},\bar{y})$ exist and one has*

$$f'(\bar{x},\bar{y})(u,v) = f_{|1}(\bar{x},\bar{y})u + f_{|2}(\bar{x},\bar{y})v \quad \forall\,(u,v)\in E\times F. \tag{3.13}$$

An analogous statement holds for H- and F-derivatives.

(b) *Assume that the partial G-derivatives* $f_{|1}$ *and* $f_{|2}$ *exist on* $U \times V$ *and are continuous at* $(\bar{x}, \bar{y})$. *Then* f *is F-differentiable at* $(\bar{x}, \bar{y})$ *and* (3.13) *holds true.*

Now we establish two *implicit function theorems*: one under standard hypotheses and one under relaxed differentiability hypotheses but with G finite dimensional.

Theorem 3.7.2 (Classical Implicit Function Theorem) *In addition to* (A), *let the following hold:*

(a) E, F, *and* G *are Banach spaces.*
(b) f *is continuous at* $(\bar{x}, \bar{y})$ *and* $f(\bar{x}, \bar{y}) = 0$.
(c) *The partial F-derivative* $f_{|2}$ *exists on* $U \times V$ *and is continuous at* $(\bar{x}, \bar{y})$.
(d) *The (continuous linear) mapping* $f_{|2}(\bar{x}, \bar{y}) : F \to G$ *is bijective.*

Then:

(i) *There exist neighborhoods* $U' \subseteq U$ *and* $V' \subseteq V$ *of* $\bar{x}$ *and* $\bar{y}$, *respectively, such that for each* $x \in U'$ *there is precisely one* $\varphi(x) \in V'$ *satisfying*

$$f\big(x, \varphi(x)\big) = o \quad \forall\, x \in U'.$$

(ii) *If* f *is continuous in a neighborhood of* $(\bar{x}, \bar{y})$, *then the function* $\varphi : x \mapsto \varphi(x)$ *is continuous in a neighborhood of* $\bar{x}$.
(iii) *If* f *is continuously differentiable in a neighborhood of* $(\bar{x}, \bar{y})$, *then* φ *is continuously differentiable in a neighborhood of* $\bar{x}$ *and there holds*

$$\varphi'(x) = -f_{|2}\big(x, \varphi(x)\big)^{-1} \circ f_{|1}\big(x, \varphi(x)\big). \tag{3.14}$$

Concerning the *proof* of the theorem, which is based on the Banach fixed point theorem, see for instance Dieudonné [53], Schirotzek [196], or Zeidler [222]. Observe that the assumptions on $f_{|2}$ guarantee that $f_{|2}\big(x, \varphi(x)\big)^{-1}$ exists as an element of $\mathrm{L}(G, F)$ provided $\|x - \bar{x}\|$ is sufficiently small.

Now we relax the differentiability assumptions on $f_{|2}$.

Proposition 3.7.3 *In addition to* (A), *let the following hold:*

(a) G *is finite dimensional.*
(b) f *is continuous in a neighborhood of* $(\bar{x}, \bar{y})$ *and* $f(\bar{x}, \bar{y}) = 0$.
(c) *The partial F-derivative* $f_{|2}(\bar{x}, \bar{y})$ *exists and is surjective.*

Then, for each neighborhood $V' \subseteq V$ *of* $\bar{y}$ *there exist a neighborhood* $U' \subseteq U$ *of* $\bar{x}$ *and a function* $\varphi : U' \to V'$ *such that the following holds:*

(i) $f\big(x, \varphi(x)\big) = o \quad \forall\, x \in U', \quad \varphi(\bar{x}) = \bar{y}$.
(ii) φ *is continuous at* $\bar{x}$.

Proof.

(I) Without loss of generality we may assume that $\bar{y} = o$. Further we set $T := f_{|2}(\bar{x}, o)$. By assumption, T is a continuous linear mapping of F onto the finite-dimensional space G. Hence there exists a finite-dimensional linear subspace $\tilde{F}$ of F such that the linear mapping $T^{-1} : G \to \tilde{F}$ satisfying $TT^{-1}(z) = z$ for any $z \in G$ is a linear isomorphism. In order to verify the assertions (i) and (ii), we may replace $f : E \times F \to G$ by its restriction to $E \times \tilde{F}$. But $\tilde{F}$ can be identified with G and so we may assume that $F = G$. Then T is a bijective linear mapping of G onto G.

(IIa) Let $\epsilon > 0$ be such that $\mathrm{B}_F(o, \epsilon) \subseteq V$ and f is continuous on the neighborhood $\mathrm{B}_E(\bar{x}, \epsilon) \times \mathrm{B}_F(o, \epsilon)$ of $(\bar{x}, o)$. Let $\alpha \in (0, \epsilon)$ be such that

$$|f(\bar{x}, y) - T(y)| \le \frac{\alpha}{2\|T^{-1}\|} \quad \forall\, y \in \mathrm{B}_F(o, \alpha). \tag{3.15}$$

Since f is continuous and $\mathrm{B}_F(o, \alpha)$ is compact, there further exists $\beta \in (0, \epsilon)$ such that

$$|f(\bar{x}, y) - f(x, y)| \le \frac{\alpha}{2\|T^{-1}\|} \quad \forall\, x \in \mathrm{B}_E(\bar{x}, \beta) \quad \forall\, y \in \mathrm{B}_F(o, \alpha). \tag{3.16}$$

(IIb) For any $x \in \mathrm{B}_E(\bar{x}, \beta)$ define $h_x : \mathrm{B}_F(o, \alpha) \to F$ by $h_x(y) := y - T^{-1} f(x, y)$. Notice that h_x is continuous.

(IIc) We now show that h_x maps $\mathrm{B}_F(o, \alpha)$ into itself. Let any $y \in \mathrm{B}_F(o, \alpha)$ be given. We have

$$\|h_x(y)\| \le \|y - T^{-1} f(\bar{x}, y)\| + \|T^{-1}\big(f(\bar{x}, y) - f(x, y)\big)\|. \tag{3.17}$$

Furthermore, we obtain

$$\begin{aligned} &\|y - T^{-1} f(\bar{x}, y)\| = \|T^{-1}\big(T(y) - f(\bar{x}, y)\big)\| \\ &\le \|T^{-1}\| \cdot \|T(y) - f(\bar{x}, y)\| \underset{(3.15)}{\le} \frac{\alpha}{2} \end{aligned} \tag{3.18}$$

as well as

$$\|T^{-1}\big(f(\bar{x}, y) - f(x, y)\big)\| \underset{(3.16)}{\le} \|T^{-1}\| \cdot \frac{\alpha}{2\|T^{-1}\|} = \frac{\alpha}{2}.$$

Hence (3.17) shows that h_x maps $\mathrm{B}_F(o, \alpha)$ into itself.

(IId) In view of (IIb) and (IIc) the Brouwer fixed-point theorem applies, ensuring that h_x has a fixed point $\psi(x)$ in $\mathrm{B}_F(o, \alpha)$. This defines a mapping $\psi : x \mapsto \psi(x)$ of $\mathrm{B}_E(\bar{x}, \beta)$ into V satisfying

$$\psi(x) - T^{-1} f\big(x, \psi(x)\big) = h_x\big(\psi(x)\big) = \psi(x)$$

and so $f\big(x, \psi(x)\big) = o$ for any $x \in \mathrm{B}_E(\bar{x}, \beta)$.

(III) Let a neighborhood $V' \subseteq V$ of o be given. Choose $\nu \in \mathbb{N}$ such that $\mathrm{B}_F(o, \frac{1}{\nu}) \subseteq V'$ and set $V_i := \mathrm{B}_F(o, \frac{1}{\nu+i})$ for $i = 1, 2, \ldots$ By step (II) we know that for each i there exist a neighborhood U_i of $\bar{x}$ and a function $\psi_i : U_i \to V_i$ satisfying $f\big(x, \psi_i(x)\big) = o$ for any $x \in U_i$. Without loss of generality we may assume that U_{i+1} is a proper subset of U_i for $i = 1, 2, \ldots$ and that $\bigcap_{i=1}^{\infty} U_i = \{\bar{x}\}$. Now let $U' := U_1$ and define $\varphi : U' \to V'$ by

$$\varphi(\bar{x}) := o = \bar{y}, \quad \varphi(x) := \psi_i(x) \quad \text{whenever } x \in U_i \setminus U_{i+1}.$$

Then (i) holds by definition of φ. We verify (ii). Thus let $\eta > 0$ be given. Then we have $V_i \subseteq \mathrm{B}_F(o, \eta)$ for some i and $\psi_i : U_i \to V_i$. It follows that

$$\|\varphi(x) - o\| = \|\psi_i(x) - o\| \leq \eta \quad \text{whenever } x \in U_i \setminus U_{i+1}.$$

By the construction of U_i and V_i, we conclude that $\varphi(U_i) \subseteq \mathrm{B}_F\big(o, \eta\big)$. □

Theorem 3.7.4 (Halkin's Implicit Function Theorem) *In addition to* (A)*, let the following hold:*

(a) *G is finite dimensional.*
(b) *f is continuous in a neighborhood of $(\bar{x}, \bar{y})$ and $f(\bar{x}, \bar{y}) = 0$.*
(c) *f is F-differentiable at $(\bar{x}, \bar{y})$ and the partial F-derivative $f_{|2}(\bar{x}, \bar{y})$ is surjective.*

Then there exist a neighborhood U' of $\bar{x}$ and a function $\varphi : U' \to V$ satisfying:

(i) *$f\big(x, \varphi(x)\big) = o \quad \forall\, x \in U', \quad \varphi(\bar{x}) = \bar{y}$.*
(ii) *φ is F-differentiable at $\bar{x}$ and there holds*

$$f_{|1}(\bar{x}, \bar{y}) + f_{|2}(\bar{x}, \bar{y}) \circ \varphi'(\bar{x}) = o. \tag{3.19}$$

Proof.

(I) With the same argument as in step (I) of the proof of Proposition 3.7.3 we may assume without loss of generality that $F = G$. We may also assume that $\bar{x} = o$ and $\bar{y} = o$. We set $S := f_{|1}(o, o)$ and $T := f_{|2}(o, o)$. Notice that T is a bijective linear mapping of G onto G.
(II) By Proposition 3.7.3, there exist a neighborhood U' of $\bar{x} = o$ and a function $\varphi : U' \to V$ such that (i) holds and φ is continuous at o. We verify (ii). Since f is F-differentiable at o, there exists a function $r : U' \to F$ such that

$$f'(o, o)\big(x, \varphi(x)\big) + r\big(x, \varphi(x)\big) = o \quad \forall\, x \in U', \tag{3.20}$$

$$\lim_{\|x\|+\|y\| \to 0} \frac{r(x, y)}{\|x\| + \|y\|} = o. \tag{3.21}$$

By Proposition 3.7.1, (3.20) passes into

$$S(x) + T\bigl(\varphi(x)\bigr) + r\bigl(x, \varphi(x)\bigr) = o \quad \forall\, x \in U',$$

i.e.,

$$\varphi(x) = -T^{-1}S(x) - T^{-1}r\bigl(x, \varphi(x)\bigr) \quad \forall\, x \in U'. \tag{3.22}$$

(III) We estimate $\|\varphi(x)\|$. Let $\sigma > 0$ be such that $\mathrm{B}_E(o, \sigma) \subseteq U'$ and

$$\|r(x, y)\| \le \frac{(\|x\| + \|y\|)}{2\|T^{-1}\|} \quad \text{whenever } \|x\| \le \sigma,\ \|y\| \le \sigma. \tag{3.23}$$

Since φ is continuous at o, there further exists $\alpha \in (0, \sigma)$ such that $\|\varphi(x)\| \le \sigma$ for all $x \in \mathrm{B}_E(o, \alpha)$. It follows that

$$\begin{aligned} \|\varphi(x)\| &\underset{(3.22)}{\le} \|T^{-1}S\| \cdot \|x\| + \|T^{-1}\| \cdot \|r\bigl(x, ,\varphi(x)\bigr)\| \\ &\underset{(3.23)}{\le} \bigl(\|T^{-1}S\| + \frac{1}{2}\bigr) \cdot \|x\| + \frac{1}{2}\|\varphi(x)\| \quad \forall\, x \in \mathrm{B}_E(o, \alpha) \end{aligned}$$

and so

$$\|\varphi(x)\| \le (2\|T^{-1}S\| + 1) \cdot \|x\| \quad \forall\, x \in \mathrm{B}_E(o, \alpha). \tag{3.24}$$

We also have

$$\|T^{-1}r\bigl(x, \varphi(x)\bigr)\| \le \|T^{-1}\| \cdot \|r\bigl(x, \varphi(x)\bigr)\|.$$

The latter inequality, (3.21) and (3.24) show that $\|T^{-1}r\bigl(x, \varphi(x)\bigr)\|/\|x\|$ is arbitrarily small for all x in a sufficiently small neighborhood of $\bar{x} = o$. In view of (3.22), we conclude that φ is F-differentiable at o, with derivative $\varphi'(o) = -T^{-1}S$. □

To prepare the next result, recall (once more) that if the mapping $f : E \to G$ is F-differentiable at $\bar{x} \in E$, then with some neighborhood U of $\bar{x}$, one has

$$\begin{aligned} &f(x) = f(\bar{x}) + f'(\bar{x})(x - \bar{x}) + r(x) \quad \forall\, x \in U, \\ &\text{where } \lim_{x \to \bar{x}} \frac{r(x)}{\|x - \bar{x}\|} = o. \end{aligned}$$

Our aim now is to replace the correction term $r(x)$ for the function values on the right-hand side by a correction term $\rho(x)$ for the argument on the left-hand side:

$$\begin{aligned} &f\bigl(x + \rho(x)\bigr) = f(\bar{x}) + f'(\bar{x})(x - \bar{x}) \quad \forall\, x \in U, \\ &\text{where } \lim_{x \to \bar{x}} \frac{\rho(x)}{\|x - \bar{x}\|} = o. \end{aligned} \tag{3.25}$$

Theorem 3.7.5 says that this is possible under appropriate hypotheses.

Theorem 3.7.5 (Halkin's Correction Theorem) *Let E and G be normed vector spaces with G finite dimensional. Further let $f : E \to G$ and $\bar{x} \in E$. Assume the following:*

(a) *f is continuous in a neighborhood of $\bar{x}$.*
(b) *The F-derivative $f'(\bar{x})$ exists and is surjective.*

Then there exist a neighborhood U of $\bar{x}$ and a function $\rho : U \to E$ such that (3.25) holds. The function ρ satisfies $\rho(\bar{x}) = o$ and is F-differentiable at $\bar{x}$ with $\rho'(\bar{x}) = o$.

Proof. Let F be the finite-dimensional linear subspace of E which $f'(\bar{x})$ maps onto G. Define $\tilde{f} : E \times F \to G$ by

$$\tilde{f}(x, y) := f(x + y) - f'(\bar{x})(x - \bar{x}) - f(\bar{x}).$$

Notice that $\tilde{f}$ is F-differentiable at $(\bar{x}, o)$ and that

$$\tilde{f}_{|1}(\bar{x}, o) = o, \quad \tilde{f}_{|2}(\bar{x}, o) = f'(\bar{x}). \tag{3.26}$$

Hence Theorem 3.7.4 applies to $\tilde{f}$ at (x, o). Thus there exist a neighborhood U of $\bar{x}$ and a function $\varphi : U \to F$ that is F-differentiable at $\bar{x}$ and is such that

$$\tilde{f}\big(x, \varphi(x)\big) = o \quad \forall\, x \in U, \quad \varphi(\bar{x}) = o,$$
$$\tilde{f}_{|1}(\bar{x}, o) + \tilde{f}_{|2}(\bar{x}, o) \circ \varphi'(\bar{x}) = o.$$

Setting $\rho := \varphi$, the definition of $\tilde{f}$ gives

$$f\big(x + \rho(x)\big) = f(\bar{x}) + f'(\bar{x})(x - \bar{x}) \quad \forall\, x \in U.$$

Moreover, by (3.26) we have $f'(\bar{x}) \circ \rho'(\bar{x}) = o$. Since $f'(\bar{x}) : F \to G$ is bijective, it follows that $\rho'(\bar{x}) = o$. From this and $\rho(\bar{x}) = o$ we finally deduce that $\rho(x)/\|x - \bar{x}\| \to o$ as $x \to \bar{x}$. □

Theorem 3.7.5 will be a key tool for deriving a multiplier rule for a nonsmooth optimization problem in Sect. 12.3.

Theorem 3.7.6 (Halkin's Inverse Function Theorem) *Let E be a finite-dimensional normed vector space. Further let $f : E \to E$ and $\bar{x} \in E$. Assume the following:*

(a) *f is continuous in a neighborhood of $\bar{x}$.*
(b) *The F-derivative $f'(\bar{x})$ exists and is surjective.*

Then there exist a neighborhood U of $\bar{x}$ and a function $\varphi : U \to E$ such that the following holds:

(i) $f\big(\varphi(x)\big) = x \quad \forall\, x \in U, \quad \varphi\big(f(\bar{x})\big) = \bar{x}$.
(ii) *φ is F-differentiable at $f(\bar{x})$, with $\varphi'\big(f(\bar{x})\big) = f'(\bar{x})^{-1}$.*

Proof. Define $\tilde{f} : E \times E \to E$ by $\tilde{f}(u,v); = u - f(v)$ and set $\bar{u} := f(\bar{x})$, $\bar{v} := \bar{x}$. Then $\tilde{f}$ is F-differentiable at $(\bar{u},\bar{v})$, with $\tilde{f}_{|1}(\bar{u},\bar{v}) = \mathrm{id}_E$ and $\tilde{f}_{|2}(\bar{u},\bar{v}) = -f'(\bar{x})$. By Theorem 3.7.4 applied to $\tilde{f}$ at $(\bar{u},\bar{v})$, there exist a neighborhood U of $\bar{u}$ and a function $\varphi : U \to E$ such that $\tilde{f}\big(u,\varphi(u)\big) = o$ for any $u \in U$ and $\varphi(\bar{u}) = \bar{v}$. Moreover, φ is F-differentiable at $\bar{u}$ and satisfies

$$\tilde{f}_{|1}(\bar{u},\bar{v}) + \tilde{f}_{|2}(\bar{u},\bar{v}) \circ \varphi'(\bar{u}) = o.$$

It is obvious that φ meets the assertions of the theorem. □

3.8 Bibliographical Notes and Exercises

The subject of this chapter is standard. We refer to Dieudonné [53], Schirotzek [196], Schwartz [197], and Zeidler [222] for differential calculus in Banach spaces and to Zeidler [221, 224] for differentiability properties of integral functionals on Sobolev spaces. The results from Proposition 3.7.3 to the end of Sect. 3.7 are due to Halkin [82]. See also the Bibliographical Notes to Chap. 4.

Exercise 3.8.1 Define $g : \mathbb{R}^2 \to \mathbb{R}$ by

$$g(x_1,x_2) := \begin{cases} \frac{x_1^3}{x_2} & \text{if } x_2 \neq 0, \\ 0 & \text{if } x_2 = 0. \end{cases}$$

Show that g is G-differentiable but not H-differentiable at $\bar{x} = (0,0)$.

Exercise 3.8.2 Show that the function

$$f(x) := x^2 \sin(1/x) \text{ if } x \in \mathbb{R} \setminus \{0\}, \quad f(x) := 0 \text{ if } x = 0$$

is F-differentiable but not continuously differentiable at $\bar{x} = 0$.

In Sect. 4.6 we shall show that the maximum norm on $\mathrm{C}[a,b]$, where $a < b$, is H-differentiable at certain points but nowhere F-differentiable; compare this and the preceding two examples with Proposition 3.4.2.

Exercise 3.8.3 Define $f : \mathbb{R}^2 \to \mathbb{R}^2$ by $f(x_1,x_2) := (x_1,x_2^3)$ and let $g : \mathbb{R}^2 \to \mathbb{R}$ be the function of Exercise 3.8.1. Then f is F-differentiable (and so G-differentiable) on $\mathbb{R}^2$ and g is G-differentiable at $\bar{x} = (0,0)$. Is the composite function $g \circ f$ G-differentiable at $\bar{x}$?

Exercise 3.8.4 Carry out the omitted proofs for Proposition 3.2.4.

Exercise 3.8.5 Prove Proposition 3.2.5.

Exercise 3.8.6 Show that the functional $f(x) := \frac{1}{2}a(x,x)$, $x \in E$, where $a : E \times E \to \mathbb{R}$ is bilinear, symmetric, and bounded, is continuously differentiable on E and calculate its derivative (cf. Example 3.6.1).

Exercise 3.8.7 Assume that $\varphi : [a,b] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is continuous and possesses continuous partial derivatives with respect to the second and the third variable. Modeling the proof in Example 3.6.3, show that the functional $f : \mathrm{C}^1[a,b] \times [a,b] \times [a,b]$ defined by

$$f(x,\sigma,\tau) := \int_\sigma^\tau \varphi\big(t, x(t), \dot{x}(t)\big)\,\mathrm{d}t, \quad x \in \mathrm{C}^1[a,b], \quad \sigma,\tau \in (a,b),$$

is continuously differentiable and calculate its derivative. (Functionals of this kind appear in variable-endpoint problems in the classical calculus of variations.)