Subdifferentials of Lower Semicontinuous Functionals

9.1 Fréchet Subdifferentials: First Properties

In this section we study another kind of derivative-like concepts.

**Definition 9.1.1** Assume that $E$ is a Banach space, $f : E \to \mathbb{R}$ is proper and l.s.c., and $\bar{x} \in \text{dom } f$.

(a) The functional $f$ is said to be Fréchet subdifferentiable (F-subdifferentiable) at $\bar{x}$ if there exists $x^* \in E^*$, the F-subderivative of $f$ at $\bar{x}$, such that

$$\liminf_{y \to 0} \frac{f(\bar{x} + y) - f(\bar{x}) - \langle x^*, y \rangle}{\|y\|} \geq 0. \quad (9.1)$$

(b) The functional $f$ is said to be viscosity subdifferentiable at $\bar{x}$ if there exist $x^* \in E^*$, the viscosity subderivative of $f$ at $\bar{x}$, and a $C^1$-function $g : E \to \mathbb{R}$ such that $g'(\bar{x}) = x^*$ and $f - g$ attains a local minimum at $\bar{x}$.

If, in particular,

$$g(x) = \langle x^*, x - \bar{x} \rangle - \sigma \|x - \bar{x}\|^2$$

with some positive constant $\sigma$, then $x^*$ is called proximal subgradient of $f$ at $\bar{x}$. The sets

$$\partial_F f(\bar{x}) := \text{set of all F-subderivatives of } f \text{ at } \bar{x},$$

$$\partial_V f(\bar{x}) := \text{set of all viscosity subderivatives of } f \text{ at } \bar{x},$$

$$\partial_P f(\bar{x}) := \text{set of all proximal subgradients of } f \text{ at } \bar{x}$$

are called Fréchet subdifferential (F-subdifferential), viscosity subdifferential, and proximal subdifferential of $f$ at $\bar{x}$, respectively.

**Remark 9.1.2** Observe that the function $g$ in Definition 9.1.1(b) can always be chosen such that $(f - g)(\bar{x}) = 0$ (cf. Fig. 9.1).

We study the relationship between the different notions.
Proposition 9.1.3 Assume that $E$ is a Banach space, $f : E \to \mathbb{R}$ is proper and l.s.c., and $\bar{x} \in \text{dom } f$. Then $\partial_V f(\bar{x}) \subseteq \partial_F f(\bar{x})$.

Proof. See Exercise 9.8.1. \hfill \Box

Remark 9.1.4 Notice that $\partial_F f(\bar{x})$ and $\partial_V f(\bar{x})$ can be defined as above for any proper, not necessarily l.s.c. functional $f$. However, if $\partial_F f(\bar{x})$ (in particular, $\partial_V f(\bar{x})$) is nonempty, then in fact $f$ is l.s.c. at $\bar{x}$ (see Exercise 9.8.2).

The next result is an immediate consequence of the definition of the viscosity $F$-subdifferential and Proposition 9.1.3.

Proposition 9.1.5 (Generalized Fermat Rule) If the proper l.s.c. functional $f : E \to \mathbb{R}$ attains a local minimum at $\bar{x}$, then $o \in \partial_V f(\bar{x})$ and in particular $o \in \partial_F f(\bar{x})$.

We shall now show that we even have $\partial_V f(\bar{x}) = \partial_F f(\bar{x})$ provided $E$ is a Fréchet smooth Banach space. We start with an auxiliary result.

Lemma 9.1.6 Let $E$ be a Fréchet smooth Banach space and $\|\cdot\|$ be an equivalent norm on $E$ that is $F$-differentiable on $E \setminus \{o\}$. Then there exist a functional $d : E \to \mathbb{R}_+$ and a number $\alpha > 1$ such that:

(a) $d$ is bounded, $L$-continuous on $E$ and continuously differentiable on $E \setminus \{o\}$.

(b) $\|x\| \leq d(x) \leq \alpha \|x\|$ if $\|x\| \leq 1$ and $d(x) = 2$ if $\|x\| \geq 1$.

Proof. Let $b : E \to \mathbb{R}$ be the bump functional of Lemma 8.4.1. Define $d : E \to \mathbb{R}_+$ by $d(o) := 0$ and

$$d(x) := \frac{2}{s(x)}, \quad \text{where} \quad s(x) := \sum_{n=0}^{\infty} b(nx) \quad \text{for } x \neq o.$$
We show that $d$ has the stated properties:

Ad (b). First notice that the series defining $s$ is locally a finite sum. In fact, if $x \neq o$, then we have

$$b(nx) = 0 \quad \forall x \in B(x, \|x\|/2) \quad \forall n \geq 2\|x\|. \quad (9.2)$$

Moreover, $s(x) \geq b(o) = 1$ for any $x \neq o$. Hence $d$ is well defined. We have

$$d(E) \subseteq [0, 2] \quad \text{and} \quad d(x) = 2 \quad \text{whenever} \quad \|x\| \geq 1.$$ 

Further it is clear that

$$[x \neq o \text{ and } b(nx) \neq 0] \quad \implies \quad n < 1/\|x\| \quad (9.3)$$

and so, since $0 \leq b \leq 1$, we conclude that $s(x) \leq 1 + 1/\|x\|$. Hence $d(x) \geq 2\|x\|/(1 + \|x\|)$, which shows that $d(x) \geq \|x\|$ whenever $\|x\| \leq 1$. Since $b(o) = 1$ and $b$ is continuous at $o$, there exists $\eta > 0$ such that $b(x) \geq 1/2$ whenever $\|x\| \leq \eta$. Let $x \in E$ and $m \geq 1$ be such that $\eta/(m + 1) < \|x\| \leq \eta/m$.

It follows that

$$s(x) \geq \sum_{n=1}^{m} b(nx) \geq \frac{m + 1}{2} > \frac{\eta}{2\|x\|}$$

and so $d(x) < (4/\eta)\|x\|$ whenever $\|x\| \leq \eta$. This and the boundedness of $d$ imply that $d(x)/\|x\|$ is bounded on $E \setminus \{o\}$. This verifies (b).

Ad (a). Since by (9.2) the sum defining $s$ is locally finite, the functional $d$ is continuously differentiable on $E \setminus \{o\}$. For any $x \neq o$ we have

$$d'(x) = -2 \left( \sum_{n=0}^{\infty} nb'(nx) \right) \left( \sum_{n=0}^{\infty} b(nx) \right)^{-2} = -\frac{(d(x))^2}{2} \sum_{n=0}^{\infty} nb'(nx).$$

Since $b$ is L-continuous, $\lambda := \sup\{\|b'(x)\| \mid x \in E\}$ is finite and we obtain for any $x \neq o$,

$$\left\| \sum_{n=0}^{\infty} nb'(nx) \right\| \leq \lambda \sum_{n=0}^{\|x\|^{-1}} n \leq \lambda \left( 1 + \frac{1}{\|x\|} \right)^2,$$

here the first inequality holds by (9.3). This estimate together with (b) yields

$$\|d'(x)\| \leq \lambda \max\{\alpha, 2\}^2(\|x\| + 1)^2,$$

showing that $d'$ is bounded on $B(o, 1) \setminus \{o\}$. Since $d'$ is zero outside $B(o, 1)$, it follows that $d'$ is bounded on $E \setminus \{o\}$. Hence $d$ is L-continuous on $E$. This verifies (a).

Now we can supplement Proposition 9.1.3.

**Theorem 9.1.7** Let $E$ be a Fréchet smooth Banach space, $f : E \to \overline{\mathbb{R}}$ be a proper l.s.c. functional, and $\bar{x} \in \text{dom} \ f$. Then $\partial_{V} f(\bar{x}) = \partial_{F} f(\bar{x})$. 


Proof. In view of Proposition 9.1.3 it remains to show that \( \partial_F f(\bar{x}) \subseteq \partial_V f(\bar{x}) \). Thus let \( x^* \in \partial_F f(\bar{x}) \). Replacing \( f \) with the functional \( \bar{f} : E \to \mathbb{R} \) defined by
\[
\bar{f}(y) := \sup \{ f(\bar{x} + y) - f(\bar{x}) - \langle x^*, y \rangle, -1 \}, \quad y \in E,
\]
we have \( \rho \in \partial_V \bar{f}(o) \). We show that \( o \in \partial_V \bar{f}(o) \). Notice that \( \bar{f}(\bar{x}) = 0 \) and \( \bar{f} \) is bounded below. By (9.1) we obtain
\[
\liminf_{y \to o} \frac{\bar{f}(y)}{\|y\|} \geq 0. \tag{9.4}
\]
Define \( \rho : \mathbb{R}_+ \to \mathbb{R} \) by \( \rho(t) := \inf \{ \bar{f}(y) \| y \| \leq t \} \). Then \( \rho \) is nonincreasing, \( \rho(0) = 0 \) and \( \rho \leq 0 \). This and (9.4) give
\[
\lim_{t \to 0} \frac{\rho(t)}{t} = 0. \tag{9.5}
\]
Define \( \rho_1 \) and \( \rho_2 \) on \( (0, +\infty) \) by
\[
\rho_1(t) := \int_t^{e^t} \frac{\rho(s)}{s} \, ds, \quad \rho_2(t) := \int_t^{e^t} \frac{\rho_1(s)}{s} \, ds.
\]
Since \( \rho \) is nonincreasing, we have
\[
\rho_1(et) = \int_{et}^{e^{et}} \frac{\rho(s)}{s} \, ds \geq \rho(e^2t) \int_{et}^{e^t} \frac{1}{s} \, ds = \rho(e^2t). \tag{9.6}
\]
Since \( \rho_1 \) is also nonincreasing, we obtain analogously \( \rho_1(et) \leq \rho_2(t) \leq 0 \). This and (9.5) yield
\[
\lim_{t \downarrow 0} \frac{\rho_2(t)}{t} = \lim_{t \downarrow 0} \frac{\rho_1(t)}{t} = \lim_{t \downarrow 0} \frac{\rho(t)}{t} = 0. \tag{9.7}
\]
Now define \( \tilde{g} : E \to \mathbb{R} \) by \( \tilde{g}(x) := \rho_2(d(x)) \) for \( x \neq o \) and \( \tilde{g}(o) := 0 \), where \( d \) denotes the functional in Lemma 9.1.6. Recall that \( d(x) \neq 0 \) whenever \( x \neq o \). Since \( \rho_1 \) is continuous on \( (0, +\infty) \) and \( \rho_2 \) is continuously differentiable on \( (0, +\infty) \), the chain rule implies that \( \tilde{g} \) is continuously differentiable on \( E \setminus \{o\} \) with derivative
\[
\tilde{g}'(x) = \frac{\rho_1(\text{ed}(x)) - \rho_1(d(x))}{d(x)} \cdot d'(x), \quad x \neq o.
\]
The properties of \( d \) and (9.7) further imply that \( \lim_{x \to o} \|\tilde{g}'(x)\| = 0 \). Therefore it follows as a consequence of the mean value theorem that \( \tilde{g} \) is also \( F \)-differentiable at \( o \) with \( \tilde{g}'(o) = o \), and \( \tilde{g}' \) is continuous at \( o \). Since \( \rho \) is nonincreasing, we have \( \rho_2(t) \leq \rho_1(t) \leq \rho(t) \); here, the second inequality follows analogously as (9.6) and the first is a consequence of the second. Let \( \|x\| \leq 1 \). Then \( \|x\| \leq d(x) \), and since \( \rho_2 \) is nonincreasing (as \( \rho_1 \) is nonincreasing), we obtain
\[
(f - g)(x) = f(x) - \rho_2(d(x)) \geq \tilde{f}(x) - \rho_2(\|x\|) \geq \tilde{f}(x) - \rho(\|x\|) \geq 0.
\]
Since \(0 = (f - g)(o)\), we see that \(f - g\) attains a local minimum at \(o\). Hence \(o \in \partial_V f(o)\) and so \(x^* \in \partial_V f(\bar{x})\).

\textbf{Remark 9.1.8} Let \(E, f\), and \(\bar{x}\) be as in Theorem 9.1.7. Further let \(x^* \in \partial_V f(\bar{x})\), which by Theorem 9.1.7 is equivalent to \(x^* \in \partial_F f(\bar{x})\). Then there exists a concave \(C^1\) function \(g : E \to \mathbb{R}\) such that \(g'(\bar{x}) = x^*\) and \(f - g\) attains a local minimum at \(\bar{x}\) (cf. Fig. 9.1); see Exercise 9.8.4.

In order to have both the limit definition and the viscosity definition of \(F\)-subderivatives at our disposal, we shall in view of Theorem 9.1.7 assume that \(E\) is a Fréchet smooth Banach space and we denote the common \(F\)-subdifferential of \(f\) at \(\bar{x}\) by \(\partial_F f(\bar{x})\).

The relationship to classical concepts is established in Proposition 9.1.9. In this connection recall that

\[
\partial_F f(\bar{x}) \subseteq \partial_F f(\bar{x}).
\]

\textbf{Proposition 9.1.9} Assume that \(E\) is a Fréchet smooth Banach space and \(f : E \to \mathbb{R}\) is proper and l.s.c.

(a) If the directional \(G\)-derivative \(f_G(\bar{x}, \cdot)\) of \(f\) at \(\bar{x} \in \text{dom } f\) exists on \(E\), then for any \(x^* \in \partial_F f(\bar{x})\) (provided there exists one),

\[
\langle x^*, y \rangle \leq f_G(\bar{x}, y) \quad \forall y \in E.
\]

If, in particular, \(f\) is \(G\)-differentiable at \(\bar{x} \in \text{dom } f\), then \(\partial_F f(\bar{x}) \subseteq \{f'(\bar{x})\}\).

(b) If \(f \in C^1(U)\), where \(U \subseteq E\) is nonempty and open, then \(\partial_F f(x) = \{f'(x)\}\) for any \(x \in U\).

(c) If \(f \in C^2(U)\), where \(U \subseteq E\) is nonempty and open, then \(\partial_F f(x) = \partial_F f(x) = \partial_v f(x) = \{f'(x)\}\) for any \(x \in U\).

(d) If \(f\) is convex, then \(\partial_F f(x) = \partial_F f(x) = \partial f(x)\) for any \(x \in \text{dom } f\).

(e) If \(f\) is locally \(L\)-continuous on \(E\), then \(\partial_F f(x) \subseteq \partial f(x)\) for any \(x \in E\).

\textbf{Proof.}

(a) Let \(x^* \in \partial_F f(\bar{x})\) be given. Then there exist a \(C^1\) function \(g\) and a number \(\epsilon > 0\) such that \(g'(\bar{x}) = x^*\) and for each \(x \in B(\bar{x}, \epsilon)\) we have

\[
(f - g)(x) \geq (f - g)(\bar{x}) \quad \forall x \in B(\bar{x}, \epsilon).
\]

Now let \(y \in E\). Then for each \(\tau > 0\) sufficiently small we have \(\bar{x} + \tau y \in B(\bar{x}, \epsilon)\) and so

\[
\frac{1}{\tau}(f(\bar{x} + \tau y) - f(\bar{x})) \geq \frac{1}{\tau}(g(\bar{x} + \tau y) - g(\bar{x})).
\]

Letting \(\tau \downarrow 0\) it follows that \(f_G(\bar{x}, y) \geq (g'(\bar{x}), y) = \langle x^*, y \rangle\). If \(f\) is \(G\)-differentiable at \(\bar{x}\), then by linearity the latter inequality passes into \(f'(\bar{x}) = x^*\).
(b) It is obvious that \( f'(x) \in \partial F f(x) \) for each \( x \in U \). This and (a) imply \( \partial F f(x) = \{ f'(x) \} \) for each \( x \in U \).

(c) By Proposition 3.5.1 we have \( f'(x) \in \partial F f(x) \), which together with (a) and (9.8) verifies the assertion.

(d) It is evident that \( \partial f(\bar{x}) \subseteq \partial F f(\bar{x}) \subseteq \partial F f(\bar{x}) \) for each \( \bar{x} \in \text{dom } f \). Now let \( x^* \in \partial F f(\bar{x}) \) be given. As in the proof of (a) let \( g \) and \( \epsilon \) be such that (9.9) holds. Further let \( x \in E \). If \( \tau \in (0,1) \) is sufficiently small, then
\[
(1 - \tau) f(\bar{x}) + \tau f(x) \geq f((1 - \tau)\bar{x} + \tau x) \geq f(\bar{x}) + g((1 - \tau)\bar{x} + \tau x) - g(\bar{x}).
\]

It follows that
\[
f(x) - f(\bar{x}) \geq \frac{g(\bar{x} + \tau(x - \bar{x})) - g(\bar{x})}{\tau}.
\]
Letting \( \tau \downarrow 0 \), we see that
\[
f(x) - f(\bar{x}) \geq \langle g'(\bar{x}), x - \bar{x} \rangle = \langle x^*, x - \bar{x} \rangle.
\]
Since \( x \in E \) was arbitrary, we conclude that \( x^* \in \partial f(\bar{x}) \).

(e) See Exercise 9.8.5. \( \square \)

In Sect. 9.5 we shall establish the relationship between the Fréchet subdifferential and the Clarke subdifferential.

### 9.2 Approximate Sum and Chain Rules

**Convention.** Throughout this section, we assume that \( E \) is a Fréchet smooth Banach space, and \( \| \cdot \| \) is a norm on \( E \) that is F-differentiable on \( E \setminus \{ 0 \} \).

Recall that we write \( \omega(x) := \| x - \bar{x} \| \), and in particular \( \omega(x) := \| x \| \), \( x \in E \).

One way to develop subdifferential analysis for l.s.c. functionals is to start with sum rules. It is an easy consequence of the definition of the F-subdifferential that we have
\[
\partial F f_1(\bar{x}) + \partial F f_2(\bar{x}) \subseteq \partial F (f_1 + f_2)(\bar{x}).
\]
But the reverse inclusion
\[
\partial F (f_1 + f_2)(\bar{x}) \subseteq \partial F f_1(\bar{x}) + \partial F f_2(\bar{x}) \tag{9.10}
\]
do not hold in general.
Nonsmooth Analysis
Schirrotzek, W.
2007, XII, 378 p. 31 illus., Softcover
ISBN: 978-3-540-71332-6