Basic definitions and general facts. Let us recall a few basic facts about differentiability of real-valued functions. Let $X$ be a normed space, $A \subset X$ an open set, $f: A \to \mathbb{R}$ a function, and $a \in A$ a point.

For a “direction” $v \in X$ (not necessarily of norm one), we shall consider the right directional derivative $f'_+(a,v)$, the left directional derivative $f'_-(a,v)$, and the (bilateral) directional derivative $f'(a,v)$, which are defined by:

$$f'_+(a,v) = \lim_{t \to 0^+} \frac{f(a + tv) - f(a)}{t},$$

$$f'_-(a,v) = \lim_{t \to 0^-} \frac{f(a + tv) - f(a)}{t},$$

$$f'(a,v) = \lim_{t \to 0} \frac{f(a + tv) - f(a)}{t}.$$

It is easy to see that $f'_-(a,v) = -f'_+(a,-v)$. Thus $f'(a,v)$ exists if and only if both $f'_+(a,\pm v)$ exist and $-f'_+(a,-v) = f'_+(a,v)$. Another easy fact is that $f'_+(a,0) = 0$ and $f'_+(a,\lambda v) = \lambda f'_+(a,v)$ for each $\lambda > 0$.

Definition 0.1. Let $X, A, f, a$ be as above. We shall say that $f$ is:

- **Gâteaux differentiable** at $a$ if there exists $x^* \in X^*$ such that $f'(a,v) = x^*(v)$ for each $v \in X$ (that is, $f'(a,\cdot)$ is everywhere defined, real-valued, linear and continuous);
- **Fréchet differentiable** at $a$ if there exists $x^* \in X^*$ such that for each $v \in X$ one has

$$f(a + tv) = f(a) + tx^*(v) + o(t)$$

as $t \to 0$.

The functional $x^*$ is called the Gâteaux/Fréchet differential (or derivative) of $f$ at $a$, and it is denoted by $f'(a)$.

It is easy to see that the notions of Gâteaux and Fréchet differentiability are local notions (i.e., they depend only on the values of $f$ at a neighborhood of the point) and they do not change if we pass to an equivalent norm on $X$.

Observation 0.2. The following assertions are equivalent:

(i) $f$ is Gâteaux differentiable at $a$;

(ii) $f'_+(a,\cdot) \in X^*$;

(iii) there exists $x^* \in X^*$ such that for each $v \in X$ one has

$$f(a + tv) = f(a) + tx^*(v) + o(t)$$

as $t \to 0$.

Observation 0.3. The following assertions are equivalent:

(i) $f$ is Fréchet differentiable at $a$;

(ii) there exists $x^* \in X^*$ such that

$$f(a + h) = f(a) + x^*(h) + o(||h||)$$

as $h \to 0$;
(iii) \( f \) is Gâteaux differentiable at \( a \) and the limit
\[
\lim_{t \to 0} \frac{f(a + tv) - f(a)}{t} = f'(a)(v)
\]
is uniform for \( \|v\| = 1 \).

It is well known that we have always the implications
\[
\downarrow \quad \downarrow
\]
\( f \) is Fréchet diff. at \( a \) \( \implies \) \( f \) is Gâteaux diff. at \( a \)
and no other implication holds. (Of course, the Gâteaux and Fréchet differential of \( f \) at \( a \) coincide whenever both of them exist.)

In finite-dimensional spaces we have the following important consequence of compactness of the unit sphere.

**Lemma 0.4.** Let \( X, A, f, a \) be as above. Assume that \( X = \mathbb{R}^d \) and \( f \) is Lipschitz on some neighborhood of \( a \). Then \( f \) is Fréchet differentiable at \( a \) if and only if \( f \) is Gâteaux differentiable at \( a \).

**Proof.** We can suppose that \( f \) is Lipschitz on \( A \) with a Lipschitz constant \( L \). Suppose that \( f \) is Gâteaux differentiable at \( a \) with \( f'(a) = x^* \) but not Fréchet differentiable at \( a \). There exists a sequence \( \{h_n\} \subset X \setminus \{0\} \) such that \( \|h_n\| \to 0 \)
\[
D_n := \frac{f(a + h_n) - f(a) - x^*(h_n)}{\|h_n\|} \not\to 0.
\]
We can write \( h_n = t_nv_n \) where \( t_n = \|h_n\| \) and \( v_n = \frac{h_n}{\|h_n\|} \). By compactness of the unit sphere \( S_{\mathbb{R}^d} \), we can assume that \( v_n \to v_0 \in S_{\mathbb{R}^d} \). Then we obtain
\[
D_n = \frac{f(a + t_nv_n) - f(a)}{t_n} - x^*(v_n)
\]
\[
\leq \frac{f(a + t_nv_0) - f(a)}{t_n} - x^*(v_0) + \frac{|f(a + t_nv_n) - f(a + t_nv_0)|}{t_n} + |x^*(v_0) - x^*(v_n)|
\]
\[
\leq \frac{f(a + t_nv_0) - f(a)}{t_n} - x^*(v_0) + (L + \|x^*\|) \|v_n - v_0\| \to 0,
\]
a contradiction with (1). \( \square \)

**Derivatives and differentials of convex functions.** In what follows, \( A \) will be an open convex subset of a normed space \( X \), \( f: A \to \mathbb{R} \) a convex function, \( a \in A \) a point. The following proposition collects main important properties of right directional derivatives of a convex function at a point.

**Proposition 0.5.** Let \( X, A, f, a \) be as above.

\( a \) \( f'_+(a, v) \) exists finite for each \( v \in X \), and the functional \( p = f'_+(a, \cdot) \) is sublinear on \( X \).

\( b \) \( -p(-v) \leq p(v) \) for each \( v \in X \).
(c) The set
\[ V = \{ v \in X : f'(a, v) \text{ exists} \} = \{ v \in X : -p(-v) = p(v) \} \]
is linear (that is, \( V \) is a subspace of \( X \)) and the restriction \( p|_V \) is linear.

(d) If \( f \) is continuous, then \( p \) is Lipschitz (in particular, \( p|_V \in V^* \)).

Moreover, the properties (b),(c) hold for any sublinear functional \( p : X \to \mathbb{R} \).

Proof. (a) Existence of \( f'_+(a, v) \) follows immediately from the well known facts about existence of one-sided derivatives of a convex function of one real variable (indeed, \( f'_+(a, v) = \varphi'_+(0) \) where \( \varphi(t) = f(a + tv) \)). We already know that \( p \) is positively homogeneous. To show subadditivity, consider \( u, v \in X \) and compute
\[
p(u + v) = 2 p \left( \frac{u+v}{2} \right) = 2 \lim_{t \to 0^+} \frac{f \left( \frac{1}{2}(a+tu)+\frac{1}{2}(a+tv) ight) - f(a)}{t} = p(u) + p(v).
\]
(b) is easy: \( 0 = p(0) = p(v + (-v)) \leq p(v) + p(-v) \).

(c) Now, let \( u, v \in V \) and \( \lambda \in \mathbb{R} \). By positive homogeneity of \( p \), if \( \lambda \geq 0 \) then \( \lambda v \in V \) and \( p(\lambda v) = \lambda p(v) \). By definition of \( V \), we have \(-v \in V \) and \( p(-v) = -p(v) \). For \( \lambda < 0 \), we have \(-\lambda > 0 \) and hence \(-p(-\lambda v) = \lambda p(v) = -\lambda p(-v) = p(\lambda v) \), which shows that \( \lambda v \in V \) and \( p(\lambda v) = \lambda p(v) \). Finally, the inequalities
\[
p(u + v) \leq p(u) + p(v) = - \left[ p(-u) + p(-v) \right] \leq -p(-u - v) \leq p(u + v)
\]
imply that \( u + v \in V \) and \( p(u + v) = p(u) + p(v) \).

(d) If \( f \) is continuous, it is Lipschitz on a neighborhood of 0 with some constant \( L \). This easily implies that \( |p(v)| \leq L \|v\| \) for each \( v \in X \). For \( u, v \in X \), we have by subadditivity
\[
p(u) - p(v) = p(v + (u - v)) - p(v) \leq p(u - v) \leq L \|u - v\|.
\]
By interchanging the role of \( u \) and \( v \), we obtain that \( p \) is \( L \)-Lipschitz on \( X \).

Corollary 0.6. Let \( X, A, f, a \) be as above. If \( f'(a, v) \) exists for each \( v \in X \), then \( f \) is Gâteaux differentiable at \( a \).

Proof. In the notation of Proposition 0.5, \( V = X \). Thus \( f'(a, \cdot) \in X^* \).

Corollary 0.7. Let \( A \subset \mathbb{R}^d \) be an open convex set, \( f : A \to \mathbb{R} \) a convex function, \( a \in A \). If all partial derivatives
\[
\frac{\partial f}{\partial x_i}(a) \quad (i = 1, \ldots, d)
\]
exist, then \( f \) is Fréchet differentiable at \( a \).

Proof. Notice that, in the notation of Proposition 0.5, \( V = \mathbb{R}^d \) since \( V \) contains the canonical orthonormal basis of \( \mathbb{R}^d \). Moreover, \( f \) is continuous and hence locally Lipschitz on \( A \). By Corollary 0.6, \( f \) is Gâteaux differentiable at \( a \). Apply Lemma 0.4.
Example 0.8. Consider the continuous convex function

$$f : \ell_1 \to \mathbb{R}, \quad f(x) = \|x\|_1 = \sum_{i=1}^{+\infty} |x(i)|.$$ 

This function \(f\) is Gâteaux differentiable exactly at the points having all coordinates nonzero, and \(f\) is nowhere Fréchet differentiable.

Proof. Let \(e_k\) denote the \(k\)-th canonical unit vector in \(\ell_1\). If \(f\) is Gâteaux differentiable at a point \(a \in \ell_1\) then the functions \(\varphi_k(t) := f(a+te_k) = |a(k)+t|+\sum_{i \neq k} |a(i)|\) are differentiable at 0, and hence \(a(k) \neq 0\) for each \(k\).

Now, let us prove that \(f\) is Gâteaux differentiable at \(a\) whenever \(a(k) \neq 0\) for each \(k\). Since \(f(x)\) remains the same if we change signs of some of the coordinates of \(x\), we can suppose that \(a(k) > 0\) for each \(k\). Fix an arbitrary \(v \in \ell_1\) with \(\|v\|_1 = 1\). For such \(v\) and any \(t > 0\), we have

$$\frac{f(a+tv) - f(a)}{t} = \sum_{i=1}^{+\infty} \frac{|a(i)+tv(i)| - a(i)}{t}.$$ 

For \(n \in \mathbb{N}\) and \(0 < t < \min\{a(1), \ldots, a(n)\}\), we have \(t|v(i)| \leq t < a(i)\) whenever \(1 \leq i \leq n\), and hence \(a(i)+tv(i) > 0\) for such \(i\). Thus we can write

$$\frac{f(a+tv) - f(a)}{t} = \sum_{i=1}^{n} v(i) + \varepsilon_n(t)$$

where \(\varepsilon_n(t) = \sum_{i>n} \frac{|a(i)+tv(i)| - a(i)}{t}\). Notice that \(|\varepsilon_n(t)| \leq \sum_{i>n} \frac{a(i)+tv(i)| - a(i)}{t} = \sum_{i>n} |v(i)|\). Passing to limit as \(t \to 0^+\), we obtain

$$f'_+(a,v) = \sum_{i=1}^{n} v(i) + \lambda_n$$

where \(\lambda_n = \lim_{t \to 0^+} \varepsilon_n(t)\). Of course, \(\lambda_n \to 0\) as \(n \to +\infty\), since \(|\lambda_n| \leq \sum_{i>n} |v(i)|\).

Consequently,

$$f'_+(a,v) = \sum_{i=1}^{+\infty} v(i).$$

By positive homogeneity, we conclude that the same formula holds for each \(v \in \ell_1\). Since \(f'_+(a,\cdot)\) is linear and continuous on \(\ell_1\), \(f\) is Gâteaux differentiable at \(a\).

It remains to show that \(f\) is not Fréchet differentiable at any such \(a\). Consider the vectors \(h_n = -2a(n)e_n\) \((n \in \mathbb{N})\) and observe that \(\|h_n\|_1 = 2a(n) \to 0\). If \(f\) were Fréchet differentiable at \(a\), we would have

$$0 = \lim_{n} \frac{f(a+h_n) - f(a) - f'_+(a,h_n)}{\|h_n\|} = \cdots = \lim_{n} \frac{|-a(n)| - a(n) + 2a(n)}{2a(n)} = 1$$

a contradiction. \(\square\)
Subdifferential. We shall need several times the following corollary of the Hahn-Banach theorem (see the last corollary in the chapter “Hahn-Banach theorems”).

**Corollary 0.9.** Let $X$ be a normed space, $f : X \to (-\infty, +\infty]$ a convex function, $H \subset X$ an affine set, and $h : H \to \mathbb{R}$ a continuous affine function such that $h \leq f$ on $H$. Assume that

$$\text{int}(\text{dom}(f)) \cap H \neq \emptyset,$$

and $f$ is continuous on $\text{int}(\text{dom}(f))$. Then there exists a continuous affine extension $\hat{h} : X \to \mathbb{R}$ of $h$, such that $\hat{h} \leq f$.

In what follows, unless otherwise specified,

$$\begin{align*}
A & \text{ is an open convex set in a normed space } X, \\
f & \text{ is a continuous convex function,} \\
a & \in A.
\end{align*}
$$

If necessary (e.g., for application of the above corollary), we consider $f$ extended by $+\infty$ outside of $A$.

**Definition 0.10.** The subdifferential of $f$ at $a$ is the set

$$\partial f(a) = \{ x^* \in X^* : f(x) \geq f(a) + x^*(x - a) \text{ for each } x \in A \}.$$ 

The elements of $\partial f(a)$ are called subgradients of $f$ at $a$.

The multivalued mapping $\partial f : A \to 2^{X^*}$ (where $2^{X^*}$ is the set of all subsets of $X^*$), $x \mapsto \partial f(x)$, is called the subdifferential map or simply the subdifferential of $f$.

The image $\partial f(E)$ of a set $E \subset A$ is the set

$$\partial f(E) = \bigcup_{x \in A} \partial f(x).$$

Geometrically, the subgradients of are those elements of $X^*$ that define affine functions $h : X \to \mathbb{R}$ such that $h(a) = f(a)$ and $h \leq f$ on $A$. We say that such $h$ supports $f$ at $a$.

**Proposition 0.11.** Assume (2). Then $\partial f(a) \neq \emptyset$.

**Proof.** By Corollary 0.9, applied to the affine set $H = \{a\}$ and the function $h(a) = f(a)$ on $H$, there exists a continuous affine extension $\hat{h} : X \to \mathbb{R}$ of $H$ such that $\hat{h} \leq f$ on $A$. Since $h(a) = f(a)$, $\hat{h}$ must be of the form $\hat{h}(x) = f(a) + x^*(x - a)$ for some $x^* \in X^*$. Then $x^* \in \partial f(a)$. $\square$

**Lemma 0.12.** Assume that (2) holds and $L \geq 0$ is given. Then $f$ is $L$-Lipschitz on $A$ if and only if $\partial f(A) \subset LB_{X^*}$.

**Proof.** Let $f$ be $L$-Lipschitz on $A$. Given $x \in A$, let $r > 0$ be such that the closed ball $B(x, r)$ is contained in $A$. For each $x^* \in \partial f(x)$, we have $\|x^*\| = \frac{1}{r} \sup_{\|v\|=r} x^*(v) \leq \frac{1}{r} \sup_{\|v\|=r} [f(a + v) - f(a)] \leq \frac{1}{r} \sup_{\|v\|=r} L \|v\| = L.$

Now, let $\partial f(A) \subset LB_{X^*}$. For $x, y \in A$ and $x^* \in \partial f(x)$, we have $f(x) - f(y) \leq x^*(x - y) \leq \|x^*\| \|x - y\| \leq L \|x - y\|$. By interchanging the roles of $x$ and $y$, we conclude that $f$ is $L$-Lipschitz. $\square$
Proposition 0.14. Assume the values of $f$ which shows (2) is locally Lipschitz on $A$. (Indeed, $f$ is locally Lipschitz on $A$.)

Proof. We already know that $\partial f(a)$ is bounded (by Corollary 0.13). To see that it is also convex and $w^*$-closed, it suffices to notice that

$$\partial f(a) = \bigcap_{x \in A} \{ x^* \in X^* : x^*(x - a) \leq f(x) - f(a) \}$$

is intersection of a family of $w^*$-closed halfspaces in $X^*$.

Lemma 0.15. Assume (2). Then the subdifferential map $\partial f$ is monotone on $A$, that is,

$$(x^* - y^*)(x - y) \geq 0 \quad \text{whenever } x, y \in A, x^* \in \partial f(x) \text{ and } y^* \in \partial f(y).$$

Proof. Given $x, y \in A$, $x^* \in \partial f(x)$ and $y^* \in \partial f(y)$, we have

$$x^*(y - x) \leq f(y) - f(x)$$
$$y^*(x - y) \leq f(x) - f(y).$$

Summing the two inequalities, one easily gets $(x^* - y^*)(x - y) \geq 0$.

Proposition 0.16. Assume (2). For $x^* \in X^*$ the following assertions are equivalent:

(i) $x^* \in \partial f(a)$;

(ii) $x^*(v) \leq f'_+(a, v)$ for each $v \in X$;

(iii) $-f'_+(a, -v) \leq x^*(v) \leq f'_+(a, v)$ for each $v \in X$.

Proof. (ii) and (iii) are equivalent since $-x^*(v) = x^*(-v) \leq f'_+(a, -v)$.

(i) $\Rightarrow$ (ii). Let $x^* \in \partial f(a)$ and $v \in X$. There exists $\tau > 0$ such that $a + \tau v \in A$. Then, for each $t \in (0, \tau)$,

$$x^*(v) = \frac{1}{t} x^*(tv) \leq \frac{f(a + tv) - f(a)}{t}. $$

Letting $t \to 0^+$, we obtain the inequality in (ii).

(ii) $\Rightarrow$ (i). Let (ii) hold, and $x \in A$. Then we have

$$x^*(x - a) \leq f'_+(a, x - a) = \lim_{t \to 0^+} \frac{f(a + t(x - a)) - f(a)}{t} = \inf_{t \in [0,1]} \frac{f(a + t(x - a)) - f(a)}{t} \leq_1 f(x) - f(a),$$

which shows (i).

As a corollary of Proposition 0.16, we obtain that $\partial f(a)$ is fully determined by the values of $f$ in any neighborhood of $a$.

While Proposition 0.16 determines the subdifferential in terms of the directional derivatives, the following proposition shows that it is possible to do also the opposite: determine the directional derivatives in terms of the subdifferential.
Theorem 0.18. Assume $\hat{\text{convex functions of one real variable imply that}}$
\[ h(a) = f(a), \text{ we must have } \hat{h}(x) = f(a) + x^*(x - a) \text{ where } x^* \in X^*. \]
Then $x^*(v) = \hat{h}(a + v) - f(a) = h(a + v) - f(a) = f'_+(a, v).$
\[ \Box \]

Theorem 0.19. Assume (2). Then $f$ is Gâteaux differentiable at $a$ if and only if
\[ \partial f(a) = \{ f'(a) \}. \]
Proof. If $f$ is Gâteaux differentiable at $a$, then $-f'_+(a, -v) = f'_+(a, v) = f'(a)(v)$
for each $v \in X$. For any $x^* \in \partial f(a)$, Proposition 0.16 implies that $x^*(v) = f'(a)(v)$
for each $v$, that is $x^* = f'(a)$.

Now, assume that $f$ is not Gâteaux differentiable at $a$. By Corollary 0.6, there
exists $v \in X$ such that $f'_+(a, v) \neq f'_+(a, -v)$. By Proposition 0.17, there exist
$x_1^*, x_2^* \in \partial f(a)$ such that $x^*_1(v) = f'_+(a, v)$ and $x^*_2(v) = f'_+(a, -v)$. Then $x^*_1(v) - x^*_2(v) = f'_+(a, v) + f'_+(a, -v) \neq 0$, and hence $\partial f(a)$ is not a singleton.
\[ \Box \]

Let us show that the subdifferential of a Gâteaux differentiable function satisfies a kind
of a continuity in the $w^*$-topology.

Theorem 0.19. Assume (2). If $f$ is Gâteaux differentiable at $a$, then the following property
holds:
\[ \| x_n - a \| \to 0, \ x_n \in A, \ x_n^* \in \partial f(x_n) \implies \ x_n^* \overset{w^*}{\rightharpoonup} f'(a). \]
Proof. Assume that $x_n^*$ do not converge in the $w^*$-topology to $f'(a)$. This means that there
exists $v \in X$ such that $x_n^*(v) \not\to f'(a)(v)$. Passing to a subsequence if necessary, we can
(and do) suppose that, for some $\varepsilon > 0$,
\[ \| x_n^*(v) - f'(a)(v) \| \geq \varepsilon \quad \text{for each } n. \]
By Corollary 0.13, the sequence \{ $x_n^*$ \} is bounded, and hence (by Alaoglu’s theorem) it has a
$w^*$-accumulation point $x_n^*$. This means that each $w^*$-neighborhood of $x_n^*$ contains infinitely
many $x_n^*$’s.

Fix an arbitrary $y \in A$. Since the sets
\[ W_n = \{ z^* \in X^* : |x_n^*(y - a) - x_0^*(y - a)| < \frac{1}{n}, \ |x_n^*(v) - x_0^*(v)| < \frac{1}{n} \} \quad (n \in \mathbb{N}) \]
are $w^*$-neighborhoods of $x_0^*$, it is easy to see that we can suppose (by passing to a further
subsequence) that $x_n^*(y - a) \to x_0^*(y - a)$ and $x_n^*(v) \to x_0^*(v)$. By definition of subdifferential,
\[ f(y) - f(x_n) \geq x_n^*(y - x_n) = x_n^*(y - a) + x_n^*(a - x_n). \]
Notice that $|x_n^*(a - x_n)| \leq \| x_n^* \| \| a - x_n \| \to 0$ since \{ $x_n^*$ \} is bounded. Thus we can pass to
limits to obtain
\[ f(y) - f(a) \geq x_0^*(y - a). \]
Since $y \in A$ was arbitrary, we have $x_0^* \in \partial f(a)$, that is, $x_0^* = f'(a)$. Consequently, $x_n^*(v) \to f'(a)(v)$. But this contradicts (3).
\[ \Box \]
Now, we are going to characterize Fréchet differentiability in terms of the subdifferential. We shall need the following definition.

**Definition 0.20.** Let \( X, Y \) be a normed spaces, \( A \subset X \) a set, \( \Phi : A \rightarrow 2^Y \) a multivalued mapping, and \( a \in A \).

(a) The image by \( \Phi \) of a set \( E \subset A \) is defined as the set \( \Phi(E) = \bigcup_{x \in E} \Phi(x) \).

(b) We say that \( \Phi \) is single-valued and continuous at \( a \) if \( \Phi(a) = \{ y_0 \} \) and

\[
\forall \varepsilon > 0 \exists \delta > 0 : [ x \in A, \| x - a \| < \delta \Rightarrow \Phi(x) \subset B^0(y_0, \varepsilon) ].
\]

**Exercise 0.21.** In the notation of the above definition, the following assertions are equivalent.

(i) \( \Phi \) is single-valued and continuous at \( a \).

(ii) \( \Phi(a) = \{ y_0 \} \) and \( [ A \ni x_n \rightarrow a, y_n \in \Phi(x_n) \Rightarrow y_n \rightarrow y_0 ] \).

(iii) \( \text{osc}(\Phi, a) := \lim_{\delta \to 0+} \text{diam} \{ \Phi(B(a, \delta)) \} = 0 \).

(\( \text{osc}(\Phi, a) \) is called the oscillation of \( \Phi \) at \( a \).)

**Theorem 0.22.** Assume (2). Then \( f \) is Fréchet differentiable at \( a \) if and only if \( \partial f \) is single-valued and continuous at \( a \).

**Proof.** If part. Let \( \partial f \) be single-valued and continuous at \( a \). For \( x = a + h \in A \), let \( x_n^* \in \partial f(a + h) \). (Thus \( \partial f(a) = \{ x_0^* \} \).) Then we have

\[
0 \leq \frac{f(a + h) - f(a) - x_0^*(h)}{\| h \|} \leq \frac{x_0^*(h) - x_n^*(h)}{\| h \|} \leq \| x_n^* - x_0^* \| \to 0
\]

as \( h \to 0 \), which means that \( f \) is Fréchet differentiable at \( a \).

Only if part. Let \( f \) be Fréchet differentiable at \( a \) with \( f'(a) = x_0^* \), that is,

\[
\varepsilon(v) := \frac{f(a + v) - f(a) - x_0^*(v)}{\| v \|} \to 0 \quad \text{as} \quad \| v \| \to 0.
\]

Let \( \rho > 0 \) be such that \( B(a, \rho) \subset A \) and \( f \) is Lipschitz (with a constant \( L \geq 0 \)) on \( B(a, \rho) \). Consider sequences \( \{ h_n \} \subset B(0, \rho) \setminus \{ 0 \} \) and \( \{ x_n^* \} \subset X^* \) such that \( h_n \to 0 \) and \( x_n^* \in \partial f(a + h_n) \) for each \( n \). We want to show that \( x_n^* \to f'(a) \).

Observe that \( (x_n^* - x_0^*)(v) = -x_0^*(v) + x_n^*((a + v) - (a + h_n) + x_n^*(h_n) \leq [\| v \| \varepsilon(v) - f(a + v) + f(a)] + [f(a + v) - f(a + h_n)] + L\| h_n \| = [\| v \| \varepsilon(v) + f(a) - f(a + h_n) + L\| h_n \|].

Consequently, for each \( r \in (0, \rho) \) we have

\[
\| x_n^* - x_0^* \| = \sup_{\| v \| = r} (x_n^* - x_0^*)(v/r) \leq \sup_{\| v \| = r} \varepsilon(v) + \frac{f(a) - f(a + h_n) + L\| h_n \|}{r}.
\]

Hence

\[
\limsup_n \| x_n^* - x_0^* \| \leq \sup_{\| v \| = r} \varepsilon(v) \quad (r \in (0, \rho)).
\]

Since \( \sup_{\| v \| = r} \varepsilon(v) \to 0 \) as \( r \to 0^+ \), we conclude that \( \| x_n^* - x_0^* \| \to 0 \). \( \square \)

**Corollary 0.23.** Assume (2). If \( f \) is Fréchet differentiable at each point of \( A \), then \( f \in C^1(A) \).
**Corollary 0.24.** Let \( f \) be a convex function on an open convex set \( A \subset \mathbb{R}^d \). If \( f \) has all partial derivatives at each point of \( A \), then \( f \in C^1(A) \).

(Indeed, by Corollary 0.7, \( f \) is Fréchet differentiable on \( A \).)