Solutions to Selected Exercises in Chapter 4

Exercises from Section 4.1

4.1.1. (a) Suppose $f : X \to [-\infty, +\infty]$ is lower semicontinuous. Let $\alpha \in \mathbb{R}$, and and $S := \{x : f(x) \le \alpha\}$. Suppose $(x_n) \subset S$ and $x_n \to \bar{x}$. Then $\liminf f(x_n) \ge f(\bar{x})$ because f is lower semicontinuous. Therefore, $\bar{x} \in S$ and S is closed.
Now suppose each lower level set of f is closed. Let $(x_n, t_n) \in \operatorname{epi} f$ and suppose $(x_n, t_n) \to (\bar{x}, \bar{t})$. For any $\epsilon > 0$, the set $S := \{x : f(x) \le \bar{t} + \epsilon\}$ is closed. Now $x_n \in S$ for all sufficiently large n, therefore $\bar{x} \in S$. This implies $f(\bar{x}) \le \bar{t} + \epsilon$. Because $\epsilon > 0$ was arbitrary, this implies $f(\bar{x}) \le \bar{t}$, and so $(\bar{x}, \bar{t}) \in \operatorname{epi} f$, and we conclude that $\operatorname{epi} f$ is closed.
Finally, suppose $\operatorname{epi} f$ is closed. Suppose $x_n \to \bar{x}$. If $\liminf f(x_n) = \infty$, then clearly $f(\bar{x}) \le \liminf f(x_n)$. Thus we suppose $\liminf f(x_n) < \infty$. Let α be any real number such that $\liminf f(x_n) < \alpha$. Then $(x_n, \alpha) \in \operatorname{epi} f$ for all sufficiently large n. Because $\operatorname{epi} f$ is closed, this implies $(\bar{x}, \alpha) \in \operatorname{epi} f$, and $f(\bar{x}) \le \alpha$. Hence $\liminf f(x_n) \le f(\bar{x})$. Thus f is lower semicontinuous.
(b) Suppose $\operatorname{epi} f$ is weakly closed, and suppose $x_\alpha \to x$ weakly. In the case $\liminf f(x_\lambda) = \infty$, it is clear $f(x) \le \liminf f(x_\lambda)$. So suppose $\liminf f(x_\lambda) < \infty$. Let $\alpha \in \mathbb{R}$ be such that $\liminf f(x_\lambda) < \alpha$. Then (x_λ, α) is eventually in $\operatorname{epi} f$. It then follows that $(x, \alpha) \in \operatorname{epi} f$ and so $f(x) \le \alpha$. It follows that f is weakly-lower semicontinuous.
Conversely, suppose f is weakly-lower semicontinuous, and suppose $(x_\lambda, t_\lambda) \in \operatorname{epi} f$ satisfies $(x_\lambda, t_\lambda) \to (\bar{x}, \bar{t})$ weakly. Then $\liminf f(x_\lambda) \le \liminf t_\lambda = \bar{t}$. Thus $(\bar{x}, \bar{t}) \in \operatorname{epi} f$, and we deduce that $\operatorname{epi} f$ is weakly closed.
The epigraph of a convex function is convex. Therefore $\operatorname{epi} f$ is weakly closed if and only if it is (norm) closed. Consequently, by the previous parts of this exercise, f is weakly-lower semicontinuous if and only if it is lower semicontinuous.
(c) This is similar to the corresponding parts of (a) and (b). □

4.1.2. (a) Clearly $\operatorname{cl} f \le f$ and $\operatorname{cl} f$ is lower semicontinuous since it is closed. Now suppose $g \le f$, and g is lower semicontinuous. Then $\operatorname{epi} f \subset \operatorname{cl} \operatorname{epi} f \subset \operatorname{epi} g$. Thus $g \le \operatorname{cl} f$. Consequently, $\operatorname{cl} f = \sup\{g : g \text{ is lower semicontinuous and } g \le f\}$.
For any $x_0 \in X$, it follows that

$$\operatorname{cl} f(x_0) \le \lim_{\delta \downarrow 0} \inf_{\|x - x_0\| < \delta} f(x) \tag{1}$$

(because $\operatorname{cl} f$ is lower semicontinuous and so $f(x_0) \le \liminf_{x \to x_0} f(x)$. On the other hand, suppose $\alpha < \lim_{\delta \downarrow 0} \inf_{\|x - x_0\| < \delta} f(x)$. Then we choose $\delta > 0$ so that $\inf_{\|x - x_0\| < \delta} f(x) \ge \beta > \alpha$. Consequently, $(x_0, \alpha) \notin \operatorname{cl} \operatorname{epi} f$. Thus $\alpha \le \operatorname{cl} f(x_0)$. It follows that

$$\operatorname{cl} f(x_0) \ge \lim_{\delta \downarrow 0} \inf_{\|x - x_0\| < \delta} f(x). \tag{2}$$

It follows from (1) and (2) that $f(x_0) = \lim_{\delta \downarrow 0} \inf_{\|x - x_0\| < \delta} f(x)$.
(b) For (i), consider $f(0) := -\infty$ and $f(t) := +\infty$ when $t \ne 0$. This function has no continuous affine minorant. For (ii), let c_{00} be the nonclosed linear subspace of all finitely supported sequences in c_0. Consider, for example, $f : c_0 \to (-\infty, +\infty]$ for which $f(x) := \sum x_i$ if $x \in B_{c_0} \cap c_{00}$, and $f(x) := +\infty$ otherwise. Then f has no affine minorant because $\inf_{B_{c_0}} f = -\infty$ (simply consider $f(-1, -1, \ldots, -1, 0, 0, \ldots)$ so the suprema of the affine minorants is identically equal $-\infty$ by convention. On the other hand, one can check that $(\operatorname{cl} f)(x) = -\infty$ if $x \in B_{c_0}$ and

$(\operatorname{cl} f)(x) = +\infty$ otherwise. Thus (ii) is not generally true. Part (iii) is true, indeed, $\operatorname{cl} f$ is lower semicontinuous and proper, so let $x_0 \in X$ and $a < \operatorname{cl} f(x_0)$. According to the epi-point separation theorem (4.1.21) there exist $\phi \in X^*$ so that $\phi(x) + a - \phi(x_0) \leq f(x)$ for all $x \in X$. Then $\alpha(\cdot) := \phi(\cdot) + a - \phi(x_0)$ is an affine function, and $\alpha(x_0) \geq a$. It follows that $\operatorname{cl} f$ is the suprema of affine functionals that minorize it.

(c) One can use the example from b(ii). A similar example is the function $f : c_0 \to (-\infty, +\infty]$ for which $f(x) := \sum x_i$ if $x \in c_{00}$, and $f(x) := +\infty$ otherwise. It is not hard to show that $\operatorname{cl} f \equiv -\infty$. □

4.1.3. See the solution to Exercise 2.1.5 which works in any normed linear space, and moreover shows the domain of a convex function is convex. □

4.1.4. Suppose $f : X \to (-\infty, +\infty]$ is convex. If $f \equiv +\infty$, then $\operatorname{epi} f = \emptyset$ is convex. Otherwise, let $(x, t), (y, s) \in \operatorname{epi} f$. Then for $0 \leq \lambda \leq 1$ we have

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y) \leq \lambda t + (1 - \lambda)s.$$

Therefore, $\lambda(x, t) + (1 - \lambda)(y, s) \in \operatorname{epi} f$ as desired.

Conversely, suppose $\operatorname{epi} f$ is convex, if $\operatorname{epi} f = \emptyset$ then $f \equiv +\infty$ is convex. Otherwise, suppose $x, y \in \operatorname{dom} f$. Then $(x, f(x)), (y, f(y)) \in \operatorname{epi} f$. Thus for $0 \leq \lambda \leq 1$, we have $\lambda(x, f(x)) + (1 - \lambda)(y, f(y)) \in \operatorname{epi} f$. This implies

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y),$$

as desired. □

4.1.5. (a) Let $S := \bigcap_\alpha C_\alpha$, where each C_α is convex. Suppose $x, y \in S$ and $0 \leq \lambda \leq 1$. Then $x, y \in C_\alpha$ for each α. Therefore, $\lambda x + (1 - \lambda)y \in C_\alpha$ for each α because C_α is convex. Then $\lambda x + (1 - \lambda)y \in S$, and thus S is convex.

(b) Given $f := \sup\{f_\alpha : \alpha \in A\}$, we have $\operatorname{epi} f = \bigcap_{\alpha \in A} \operatorname{epi} f_\alpha$. Therefore $\operatorname{epi} f$ is convex (resp. closed) if each f_α is convex (resp. closed). Thus f is convex (resp. closed) when each f_α is convex (resp. closed).

(c) Let $g := \limsup f_\alpha$. Fix x, y and $0 \leq \lambda \leq 1$. Then

$$\begin{aligned} g(\lambda x + (1 - \lambda)y) &= \limsup_\alpha f_\alpha(\lambda x + (1 - \lambda)y) \\ &\leq \limsup[\lambda f_\alpha(x) + (1 - \lambda)f_\alpha(y)] \\ &\leq \lambda \limsup f_\alpha(x) + (1 - \lambda) \limsup_\alpha f_\alpha(y) \\ &= \lambda g(x) + (1 - \lambda)g(y). \end{aligned}$$

The proof for limits is similar, but easier.

(d) Consider $f(t) := t$ and $g(t) := -t$. Then $\min\{f, g\}(t) = -|t|$ is not convex. □

4.1.8. A proof for the convexity of d_C (that works in any normed space) can be found in Fact 2.1.6. Clearly the epigraph of δ_C is convex, so δ_C is convex. □

4.1.9. For c_0 let $A = c_0^+$, that is $A := \{(x_i) \in c_0 : x_i \geq 0,\ i \in \mathbb{N}\}$. Let B be the ray $\{tx - y : t \geq 0\}$ where $x = (4^{-i})$ and $y = (2^{-i})$. Fix $t \geq 0$. For $i > 2^i$, we have $t4^{-i} - 2^{-i} < 0$. Therefore, $A \cap B = \emptyset$.

Now suppose $\phi \in \ell_1 \setminus \{0\}$ is bounded below on A, say $\phi = (s_i)$. Then $s_i \geq 0$ for all i (otherwise if $s_{i_0} < 0$, then we choose $ne_{i_0} \in c_0^+$ and $\phi(ne_{i_0}) \to -\infty$ as $n \to \infty$. Since $\phi \neq 0$, we know $s_i > 0$ for some i. Therefore, $\phi(x) > 0$. Consequently,

$$\lim_{t\to\infty} \phi(tx - y) = t\phi(x) - \phi(y) = \infty.$$

Thus ϕ cannot separate A and B. The proof in the case of ℓ_p is similar. □

4.1.11. (a) Clearly a separating family is total since it can separate x from 0 when $x \neq 0$. Conversely, if S is total, and $x \neq y$, then we choose $\phi \in S$ so that $\phi(x - y) \neq 0$, and then $\phi(x) \neq \phi(y)$ so S is separating.
(b) Let $\{h_n\}_{n=1}^{\infty}$ be norm dense in S_X. Choose $\phi_n \in S_{X^*}$ so that $\phi_n(x_n) = 1$. Now let $x \in X\setminus\{0\}$. Let $h = x/\|x\|$. Now choose h_n such that $\|h_n - h\| < 1/2$. Then $\phi_n(h) > 1/2$, and so $\phi_n(x) > 0$.
(c) The standard basis $\{e_n\}_{n=1}^{\infty}$ of ℓ_1 will work. In general, any countable norm-dense set in B_X will work as a separating family for X^*. Indeed, if $\phi \in X^*$ vanishes on a norm-dense set in B_X, it vanishes on B_X and must be the zero-functional. □

4.1.13. Alaoglu's theorem (4.1.6) ensures $B_{X^{**}}$ is weak*-compact. Thus, the weak*-closure of B_X is contained in $B_{X^{**}}$. Suppose the containment is proper, say $x^{**} \in B_{X^{**}} \setminus C$ where C is the weak*-closure of B_X. Notice that C is convex. Thus, according to the weak*-separation theorem (4.1.22), there exists $\phi \in B_{X^*}$ such that $\phi(x^{**}) > \sup_C \phi = 1$ which is contraction. Thus $C = B_{X^{**}}$. □

4.1.18. Observe that f is convex as a maximum of two convex functions. Observe that $\operatorname{dom} \partial f = \{(x_1, x_1) : x_1 > 0\} \cup \{(0, x_2) : |x_2| > 1\}$ which is not convex. Indeed, f is continuous at each point in the first set of the union (which is the interior of the domain of f), while when $|x_2| \leq 1$, we have $\lim_{t\to 0^+}[f(0, x_2 + t) - f(0, x_2)]/t = -\infty$ and so ∂f is empty at those points, while if $|x_2| > 1$, $\partial f(0, x_0) = \{(s, x_2/|x_2| : s \leq 0\}$. □

4.1.19.(a) Let $x, y \in U$, $0 \leq \lambda \leq 1$ and let $\phi \in \partial f(\lambda x + (1 - \lambda)y)$. Now let $a := f(\lambda x + (1 - \lambda)y) - \phi(\lambda x + (1-\lambda)y)$. Then the subdifferential inequality implies $a + \phi(u) \leq f(u)$ for all $u \in U$, and so

$$\begin{aligned} f(\lambda x + (1 - \lambda)y) &= \phi(\lambda x + (1 - \lambda)y) + a \\ &= \lambda(\phi(x) + a) + (1 - \lambda)(\phi(y) + a) \\ &\leq \lambda f(x) + (1 - \lambda)f(y) \end{aligned}$$

as desired.

(b) In the case X is finite-dimensional, the converse to (a) is true. Indeed, f is continuous on U by Theorem 2.1.12 and hence the max formula (2.1.19) ensure $\partial f(x) \neq \emptyset$ for all $x \in U$. However, when X is infinite dimensional, it fails. Indeed, we let f be a discontinuous linear functional on X. Then $\partial f(x) = \emptyset$ because f is not bounded below on any open subset of X. □

4.1.20. (b) Suppose $0 \in \operatorname{int} C$, it is clear that $0 \in \operatorname{core} C$. Now suppose $0 \in \operatorname{core} C$. Given $h \in X$, there exists $\delta > 0$ such that $0 + th \in C$ for $0 \leq t \leq \delta$. Then for $t = 1/n$ where $1/n < \delta$ we have $\frac{1}{n}h \in C$, and so $h \in nC$. Thus $\displaystyle\bigcup_{n=1}^{\infty} nC = X$.

Now suppose C is absorbing. By the Baire category theorem, nC has nonempty interior for some n, thus $u + rB_X \subset nC$ for some $u \in X$ and $r > 0$. Because $0 \in C$, this implies $t(u + rB_X) \subset C$ for $0 \le t \le 1/n$. Because C is absorbing (and convex with $0 \in C$) $s(-u) \in C$ for $0 \le s \le 1/m$ for some $m \in N$. Now choose $k = \max\{n, m\}$. Then $\frac{1}{k}(-u) \in C$ and $\frac{1}{k}(u + rB_X) \subset C$. By the convexity of C,

$$\frac{r}{2k}B_X = \frac{1}{2}1k(-u) + \frac{1}{2}1k(u + rB_X) \subset C$$

and so $0 \in \operatorname{int} C$.

(c) This is straighforward. Clearly $(1/n, 1/n^3) \notin C$ for $n = 1, 2, \ldots$ and so $(0,0) \notin \operatorname{int} C$. We now show $(0,0) \in \operatorname{core} C$. Indeed, let $(x,y) \in \mathbb{R}^2$. If $y = 0$, then $t(x,y) \in C$ for all $t \ge 0$. If $x = 0$, again $t(x,y) \in C$ for all $t \ge 0$. Thus suppose $x \neq 0$, $y \neq 0$ and so $y = kx$ for some $k \neq 0$. For $|t| \le |k/x|$ we have

$$t(x,y) = t(x,kx) = (tx, tkx) \in C \ \text{ since } |tkx| \ge t^2x^2.$$

Thus, $(0,0) \in \operatorname{core} C$. □

4.1.21. (a) As suggested, define $\Lambda : c_{00} \to \mathbb{R}$ by $\Lambda(x) := \sum x_i$ where $x = (x_i)$. Then Λ is a linear functional, and we consider $u_n \in c_{00}$, such that the i-th coordinate of u_n is $1/n$ for $i = 1, 2, \ldots, n$, and 0 otherwise. Then $u_n \to 0$, but $\Lambda(u_n) = 1$ and $\Lambda(0) = 0$ and so Λ is not continuous.

(b) Let $\{e_n\}_{n=1}^{\infty}$ denote the standard coordinate (Schauder) basis of c_0. Extend this to an algebraic (Hamel) basis of c_0, say $\{b_\gamma\}_{\gamma\in\Gamma}$. Then an extension of of Λ to c_0 is $\Lambda(e_\gamma) = 1$ when $b_\gamma = e_n$ for some $n \in \mathbb{N}$ and $\Lambda(e_\gamma) = 0$ otherwise. □

4.1.22. (a) $\Rightarrow$ (c): Suppose $X = \mathbb{R}^n$, with standard basis $\{e_1, e_2, \ldots, e_n\}$, and $\bigcup_{k\ge1} kC = E$. Then there exist $\delta_i > 0$ such that $te_i \in C$ when $|t| \le \delta_i$. By the convexity of C, $(x_1, x_2, \ldots, x_n) \in C$ whenever $|x_i| \le \delta/n$ where $\delta := \min\{\delta_1, \delta_2, \ldots, \delta_n\}$. Thus $0 \in \operatorname{int} C$.

(c) $\Rightarrow$ (b): Let $f : X \to \mathbb{R}$ be a linear functional. Then $\bigcup_{k\ge1} kf^{-1}(-1,1) = \bigcup_{k\ge1} f^{-1}(-k,k) = X$. By the hypothesis, $f^{-1}(-1,1)$ has 0 in its interior. Thus there exists $\delta > 0$ so that $f(\delta B_X) \subset (-1,1)$. Therefore

$$|f(x) - f(y)| = |f(x-y)| = \left| f\left(\delta\frac{x-y}{\|x-y\|}\right)\right| \cdot \frac{\|x-y\|}{\delta} \le \frac{1}{\delta}\|x-y\|.$$

That is, f has Lipschitz constant $\frac{1}{\delta}$.

(b) $\Rightarrow$ (a): Suppose X is not finite dimensional. Let $\{e_\gamma\}_{\gamma\in\Gamma}$ be an algebraic basis of X. Fix a countable set $\{e_{\gamma_1}, e_{\gamma_2}, \ldots\}$ in this basis and define a linear functional f on X by setting $f(e_{\gamma_k}) := k\|e_{\gamma_k}\|$. Then $f(v_k) \to 1$ while $v_k \to 0$ where $v_k = e_{\gamma_k}/(k\|e_{\gamma_k}\|)$. □

4.1.23. (a) Suppose C is a closed convex set, and suppose $\bar{x} = \sum_{i=1}^{\infty} \lambda_i x_i$ where $x_i \in C$ and $\lambda_i \ge 0$, $\sum \lambda_i = 1$. Fix $x_0 \in C$, and write

$$y_n := \sum_{i=1}^{n} \lambda_i x_i + \left(1 - \sum_{i=n+1}^{\infty}\right) x_0.$$

Then $y_n \to \bar{x}$, and so $\bar{x} \in C$. Thus closed convex sets in Banach spaces are convex series closed.

Suppose U is an open convex set and suppose $\bar{x} = \sum_{i=1}^{\infty} \lambda_i x_i$ where $x_i \in U$ and $\lambda_i \geq 0$, $\sum \lambda_i = 1$. Suppose $\bar{x} \notin U$. By the separation theorem (4.1.17) we choose $\phi \in X^*$ so that $\phi(\bar{x}) > \phi(x)$ for all $x \in U$. Then $\phi(\bar{x} - x_i) > 0$ for $i \in \mathbb{N}$, and therefore

$$0 > \phi\left(\bar{x} - \sum_{i=1}^{\infty} \lambda_i x_i\right) = 0.$$

This contradiction shows that open convex subsets of Banach spaces are convex series closed. It now follows easily that a G_δ-set, is convex series closed as it is an intersection of convex series closed sets.

(b) We will prove this by induction on the dimension of E. Indeed, this is true of convex sets in E where $\dim E = 1$ because those sets are intervals or singletons. Suppose we have shown the result for all Euclidean spaces of dimension less than n. Let $\dim E = n$ and suppose $C \subset E$ is a convex series closed set. Suppose $\bar{x} = \sum_{i=1}^{\infty} \lambda_i x_i$ where $x_i \in C$ and $\lambda_i \geq 0$, $\sum \lambda_i = 1$. Suppose $\bar{x} \notin \operatorname{ri} C$. Then then we choose $\phi \in E$ such that

$$r = \langle \phi, \bar{x} \rangle \geq \langle \phi, x \rangle \text{ for all } x \in C.$$

This implies $\phi(x_i) = \phi(\bar{x})$ for all i. Consequently, $x_i \subset C \cap \phi^{-1}(r)$ for $i \in \mathbb{N}$. Now $C \cap \phi^{-1}(r)$ is a translate of a convex set in a Euclidean space of dimension less than n. By the induction hypothesis, $\bar{x} \in C \cap \phi^{-1}(r)$, and so $\bar{x} \in C$. The result now follows using mathematical induction.

(c) Suppose S is a convex series closed set. If $\operatorname{int} \overline{S} = \emptyset$ there is nothing to do. By tranlating S, we assume $0 \in \operatorname{int} \overline{S}$ Thus, there exists $r > 0$ so that

$$0 \in rB_X \subset \overline{S} \subset S + \frac{r}{2} B_X.$$

Multiplying the previous inclusions by $\dfrac{1}{2^i}$ for $i = 1, 2, \ldots$ we obtain

$$\frac{r}{2^i} B_X \subset \frac{1}{2^i} S + \frac{r}{2^{i+1}} B_X.$$

It then follows that

$$\frac{r}{2} B_X \subset \frac{1}{2} S + \frac{1}{4} S + \ldots + \frac{1}{2^i} S + \frac{r}{2^{i+1}} B_X.$$

Thus for any $u \in \frac{r}{2} B_X$, there exist $s_1, s_2, \ldots, s_i \in S$ such that

$$u \in \frac{1}{2} s_1 + \frac{1}{4} s_2 + \ldots + \frac{1}{2^i} s_i + \frac{r}{2^{i+1}} B_X.$$

Let $u_n := \sum_{i=1}^{n} 2^{-i} s_i$ it follows that $u_n \to u$. Because S is convex series closed, it follows that $u \in S$. That is $\frac{r}{2} B_X \subset S$, or $0 \in \operatorname{int} S$ as desired.

(d) It is easy to verify an open map is onto, so suppose $T : X \to Y$ is onto. By translation and dilation it suffices to show $T(B_X)$ has nonempty interior. Suppose $\bar{y} \in Y \setminus \{0\}$, and choose $\bar{x} \in X$ such that $T\bar{x} = \bar{y}$. Then for $0 \leq t \leq 1/\|\bar{x}\|$, we have

$$t\bar{y} = tT\bar{x} = T(t\bar{x}) \in T(B_X) \text{ because } \|t\bar{x}\| \leq 1.$$

Thus $0 \in \operatorname{core} T(B_X) \subset \operatorname{core} \overline{T(B_X)}$, and so 0 is in the interior of $\overline{T(B_X)}$, and by part (c), $0 \in \operatorname{int} T(B_X)$ as desired.

Further notes Clearly the proof of part (b) used finite dimensionality in a crucial fashion. This is not artificial: for example c_{00} is not convex series closed in c_0. Moreover, this $c_{00} \subset c_0$ example confirms some additional property such as closed or open as used in (a) above is needed in infinite dimensional spaces. $\square$

4.1.28. Suppose $x, y \in X$ and $x \neq y$. Write $y = x + h$. By the three-slope inequality (2.1.1),

$$\begin{aligned} \frac{f(x+h) - f(x)}{\|h\|} &\leq \limsup_{n\to\infty} \frac{f(x+nh) - f(x)}{n\|h\|} \\ &\leq \limsup_{n\to\infty} \frac{K\|x+nh\| + |f(x)|}{n\|h\|} \\ &\leq \limsup_{n\to\infty} \frac{K\|x\| + Kn\|h\| + |f(x)|}{n\|h\|} = K. \end{aligned}$$

Thus $f(y) - f(x) \leq K\|y - x\|$. Similarly, $f(x) - f(y) \leq K\|x - y\|$ and we are done. $\square$

4.1.29. Consider $x := (1, 0)$. Then $r(1, 0) \notin C$ for all $r \geq 0$. Therefore, $\gamma_C(1, 0) = +\infty$. Now consider $x_n := (1, 1/n)$. Then $t(1, 1/n) \in C$ for all $t \geq \sqrt{n}$, and so $\gamma_C(1, 1/n) = 0$ for each n. Thus, $x_n \to x$, but $\liminf_{n\to\infty} \gamma_C(x_n) < \gamma_C(x)$. $\square$

4.1.30. For part (a), see the solution to Exercise 2.1.13. For part (b), suppose $x \in \operatorname{core} C$, then $0 \in \operatorname{core} C - x$. Thus for $u \in X$, we can write $u = th$ for some $t > 0$ and $h \in C - x$. Then $\gamma_{C-x}(u) \leq 1/t$ and the conclusion follows.

(c) Let $x \in X$ then $\Lambda(x) = \alpha$. If $\alpha \leq 1$, then $x \in C$, and if $\alpha > 1$, then $\alpha^{-1}x \in C$, and so $0 \in \operatorname{core} C$. However, γ_C is not continuous since there exists $x_n \to 0$ so that $\Lambda(x_n) > 1$, and hence $\gamma_C(x_n) \geq 1$. The closure of C contains 0 in its interior, hence it cannot be contained in $\{x : \gamma_C(x) \leq 1\}$.

(d) Because $0 \in \operatorname{core} C$, and C is closed, there exists $r > 0$ so that $rB_X \subset C$. Then $\gamma_C(x) \leq \frac{1}{r}\|x\|$ for all $x \in X$. Consequently, γ_C is Lipschitz with Lipschitz constant $1/r$ by Exercise 4.1.28. Clearly, $C \subset \{x \in X : \gamma_C(x) \leq 1\}$. On the other hand if $x \notin C$, because C is closed it follows from the separation theorem that $\gamma_C(x) > 1$. Thus $C = \{x \in X : \gamma_C(x) \leq 1\}$. Then, clearly, $\{x \in X : \gamma_C(x) < 1\} \subset \operatorname{int} C$ and $\{x \in X : \gamma_C(x) = 1\}$ is the boundary of C. $\square$

4.1.31. It is clear that $\sigma_{C_1}(\cdot) \leq \sigma_{C_2}(\cdot)$ when $C_1 \subset C_2$. Conversely, suppose $C_1 \not\subset C_2$. Then we choose $\bar{x} \in C_1 \setminus C_2$. By the basic separation theorem (4.1.12) there exists $x^* \in X^*$ such that $x^*(\bar{x}) > \sup_{C_2} x^*$. Then $\sigma_{C_1}(x^*) > \sigma_{C_2}(x^*)$. By contraposition, $\sigma_{C_1}(\cdot) \leq \sigma_{C_2}(\cdot)$ implies $C_1 \subset C_2$.

Further notes. Let $C_1 := B_X$ and $C_2 := \operatorname{int} B_X$. Then $\sigma_{C_1} \leq \sigma_{C_2}$ but $C_1 \not\subset C_2$, so we cannot remove the closure assumption. More drastically, let $C_2 := \operatorname{span}\{e_1, e_2, e_3, \ldots\} \cap 2B_X$ and $C_1 := B_X$ where $X = c_0$. Then $\sigma_{C_1} < \sigma_{C_2}$ but $C_1 \not\subset C_2$. Convexity is crucial as well. Indeed, let $C_1 := B_X$ and $C_2 := 2S_X$. Then $\sigma_{C_1}(\cdot) < \sigma_{C_2}(\cdot)$, but $C_1 \not\subset C_2$. $\square$

4.1.32. Suppose f and g are proper convex functions on X with $f \geq -g$ that additionally satisfy either

$$\operatorname{dom} f \cap \operatorname{cont} g \neq \emptyset$$

or f and g are both lower semicontinuous and

$$0 \in \operatorname{core}(\operatorname{dom} g - \operatorname{dom} f).$$

Then there is an affine function α, say $\alpha = x^* + r$ for some $x^* \in X^*$ and $r \in R$ such that $-g \leq \alpha \leq f$ on X. See Figure 1.2 for a sketch. If $\bar{x} \in X$ is such that $f(\bar{x}) = -g(\bar{x})$, then $-x^* \in \partial g(\bar{x})$ and $x^* \in f(\bar{x})$. $\square$

4.1.33. Let C be a closed convex set. Then $x^* \in N_C(\bar{x})$ if and only if $x^* \in \partial\delta_C(\bar{x})$ if and only if $\langle x^*, y - x\rangle \leq \delta_C(y) - \delta_C(x)$ for all $y \in X$ if and only if $\langle x^*, y - x\rangle \leq \delta_C(y) - \delta_C(x)$ for all $y \in C$ as desired. $\square$

4.1.34. Let C be a nonempty closed convex set and let K be a nonempty compact convex set that is disjoint from C. Consider the distance function d_C. Then d_C is continuous and so it attains its minimum on K; say $d_C(x_0) = r > 0$ where $x_0 \in K$. Consider $S := K + \frac{r}{2}B_X$ and so $S := \{x + y : x \in K, y \in \frac{r}{2}B_X\}$. Then S convex, has nonempty interior and is disjoint from C. By the separation theorem (4.1.17) we find $\phi \in S_{X^*}$ such that $\sup_{\operatorname{int} S}\phi \leq \inf_C \phi$. Now $\sup_K \phi + r/2 \leq \sup_{\operatorname{int} S}\phi$ and so we are done. $\square$

4.1.38. Suppose $0 \in \partial f(\bar{x})$. Then $f(y) - f(\bar{x}) \geq \langle 0, y - x\rangle = 0$, and so $f(y) \geq f(\bar{x})$ for all $y \in X$. Conversely, suppose f has a local minimum at $\bar{x}$. Then for each $y \in X$, there exists $\delta > 0$, depending on y, so that $f(\bar{x} + th) \geq f(\bar{x})$ where $h := y - \bar{x}$ and $0 \leq t \leq \delta$. In the case $\delta > 1$, we have $f(y) = f(\bar{x} + h)$ and so $f(y) \geq f(\bar{x})$ as desired. In the case $\delta < 1$, we observe

$$\begin{aligned} f(\bar{x}) \leq f(\bar{x} + \delta h) &= f((1-\delta)\bar{x} + \delta(\bar{x} + h)) \\ &\leq (1-\delta)f(\bar{x}) + \delta f(y) \end{aligned}$$

to conclude $f(y) \geq f(\bar{x})$, as desired. $\square$

4.1.39. Suppose $f : X \to [-\infty, +\infty]$ is a convex function, and let $x \in X$. If $\partial f(x)$ is empty, there is nothing to do. Suppose $\partial f(x)$ is not empty. Let $\phi_\alpha \in \partial f(x)$ and suppose $\phi_\alpha \to_{w^*} \phi$. Then for any $y \in X$,

$$\phi(y) - \phi(x) = \lim_\alpha \phi_\alpha(y) - \phi(x) \leq f(y) - f(x).$$

Thus $\partial f(x)$ is weak*-closed. Let $\phi, \Lambda \in \partial f(x)$. Suppose $0 \leq \lambda \leq 1$, then for $y \in X$,

$$\begin{aligned} (\lambda\phi + (1-\lambda)\Lambda)(y) - (\lambda\phi + (1-\lambda)\Lambda)(x) &= \lambda\phi(y - x) + (1-\lambda)\Lambda(y - x) \\ &\leq \lambda[f(y) - f(x)] \\ &\qquad + (1-\lambda)[f(y) - f(x)] \\ &= f(y) - f(x). \end{aligned}$$

This completes the proof. $\square$

4.1.40. Proposition 4.1.4 shows f is continuous at x_0 if and only if it is Lipschitz in a neighborhood of x_0 and the proof of Proposition 4.1.25 shows this occurs if and only if ∂f is bounded on a neighborhood of x_0. Exercise 4.1.39 ensures that ∂f is weak*-closed. Because $\partial f(x)$ is bounded when f is continuous at x, it follows from Alaoglu's theorem that ∂f is weak*-compact. $\square$

4.1.45.(a) Let $f(t) := -\sqrt{t}$ and $g(t) = -\sqrt{-t}$. Then $f + g = \delta_{(0)}$, so $\partial(f + g)(0) = \mathbb{R}$. Howerver, $\partial f(0) = \emptyset = \partial g(0)$.

(b) Let $g(x,y) := -\sqrt{x}$ and $A : \mathbb{R} \to \mathbb{R}^2$ be defined by $A(x,y) := (0,y)$. Then $g \circ A = 0$ ans so $\partial(g \circ A)(0) = \{0\}$, but $\partial g(0) = \emptyset$ and so $A^*\partial g(0) = \emptyset$. □

4.1.46. (a) Follows because continuous convex functions are locally Lipschitz and directionally differentiable on the interior of their domains, and hence so is the difference of two such functions.

(b) Using the inner product,

$$\begin{aligned} f(x) &= \frac{1}{2}\sup\{\langle x,x\rangle - \langle x-y, x-y\rangle : y \in F\} \\ &= \frac{1}{2}\sup\{\langle x,x\rangle - \langle x,x\rangle + 2\langle x,y\rangle - \langle y,y\rangle : y \in F\} \\ &= \sup\left\{\langle x,y\rangle - \frac{1}{2}\|y\|^2 : y \in F\right\}. \end{aligned}$$

Thus f is convex as a supremum of convex functions.

(c) See Section 4.4 for properties of conjugate functions. Part (b) shows $f = \left(\frac{1}{2}\|\cdot\|^2 + \delta_F(\cdot)\right)^*$ and then using the fact that f^{**} is the convex closure of f we obtain

$$f^*(x) = \left(\frac{1}{2}\|\cdot\|^2 + \delta_F(\cdot)\right)^{**} = \overline{\text{conv}}\left(\frac{1}{2}\|\cdot\|^2 + \delta_F(\cdot)\right) = \frac{1}{2}\|\cdot\|^2 + \delta_{\overline{\text{conv}}(F)}(\cdot)$$

□

4.1.47. Let $g(t) := t\sin(1/t)$ for $t > 0$, and $g(t) := 0$ otherwise. Then g is continuous but not of bounded variation in any neigborhood of 0. Indeed, consider points $t_n := 1/(2n\pi - 3\pi/2)$ and $s_n := 1/(2n\pi - \pi/2)$. Then $|g(t_n) - g(s_n)| \geq 1/n\pi$. Given any neighborhood U of 0, there exists N such that $s_n, t_n \in U$ for all $n > N$ and the claim follows. Now let $f := \int_0^x g(t)dt$. Then $f'(x) = g(x)$ for all x, and so f is locally Lipschitz because f' is continuous, but f is not locally DC otherwise we would write $f = h - k$ where h and k are convex functions on a neighborhood of 0. Then $f' = h'_+ - k'_+$ is of bounded variation because h_+ and k_+ are monotone.
For an example where f is twice differentiable, proceed similarly with $g(t) := t^2\sin(1/t^2)$ for $t > 0$ and $g(t) := 0$ otherwise. Then $g'(0) = 0$ and so g is everywhere differentiable and g is not of bounded variation as one can check with points $t_n := 1/\sqrt{2n\pi - 3\pi/2}$ and $s_n := 1/\sqrt{2n\pi - \pi/2}$. The remaining details are similar. □

4.1.49. See solution to Exercise 2.4.20. □

Exercises from Section 4.2

4.2.1. Let $C := \{x \in c_0 : |x_n| \leq n^{-2}, n \in \mathbb{N}\}$ where we denote $x := (x_n)$. Let $(e_n)_{n=1}^\infty$ denote the standard coordinate basis of c_0. Let $f := \delta_C$. Then $0 \in \partial f(0)$, and so we suppose $\phi \in \partial f(0)$. Then

$$\langle \phi, te_n - 0\rangle \leq f(te_n) - f(0) = 0 \quad \text{whenever } |t| \leq 1/n.$$

Therefore, $\langle \phi, e_n\rangle = 0$ for all $n \in \mathbb{N}$. Thus $\phi = 0$. This shows $\partial f(0) = \{0\}$.
However, for $h := (n^{-1})$, the $(k+1)$-th coordinate of $k^{-1}h = \frac{1}{k(k+1)} > 1/(k+1)^2$, and so $f(k^{-1}h) = \infty$. Therefore

$$\lim_{k\to\infty} f\left(0 + \frac{1}{k}h\right) - f(0) = \infty$$

and f is not Gâteaux differentiable at 0.

(b) Observe that $0 \in \partial d_C(0) \subset \partial f(0)$. Therefore, $\partial d_C(0) = \{0\}$ and so d_C is Gâteaux differentiable at 0 because its subdifferential is a singleton (Corollary 4.2.5).

Further notes. Observe that the function d_C in (b) is not Fréchet differentiable at 0. If it were, its derivative must equal 0 (the Gâteaux derivative). However,

$$\liminf_{n\to\infty} \frac{d_C\left(\frac{2}{n^2}e_n\right) - d_C(0)}{\frac{2}{n^2}} = \liminf_{n\to\infty} \frac{\frac{1}{n^2}}{\frac{2}{n^2}} = 1/2$$

and so if the Fréchet derivative of d_C exists at 0, it cannot be 0, and hence does not exist. □

4.2.2. Let ϕ be a discontinuous linear functional on X, and define $f := \phi^2$. Fix $h \in X$. Then $\phi(h) = \alpha$ for some $\alpha \in \mathbb{R}$, and so $f(th) = t^2\alpha^2$ for any $t \in \mathbb{R}$. Consequently,

$$\lim_{t\to 0}[f(0+th) - f(0)]/t = \lim_{t\to 0}(t^2\alpha^2)/t = 0.$$

Thus $\nabla f(0) = 0$ as a Gâteaux derivative.
No, f is not lower semicontinuous on X, otherwise f would be continuous at 0 because it is Gâteaux differentiable at 0 (Proposition 4.2.2). □

4.2.3. Let $\Lambda : X \to \mathbb{R}$ be a discontinuous linear functional. Let $f : X \to [0, +\infty]$ be defined by $f(x) := \Lambda^2(x)$ when $|\Lambda(x)| \leq 1$ and $f(x) := +\infty$ otherwise. We claim the Gâteaux derivative of f at 0 is 0. Indeed, fix $h \in S_X$; then $\Lambda(x) = k_x$ where $k_x \in \mathbb{R}$, then

$$\lim_{t\to 0} \frac{f(0+th) - f(0)}{t} = \lim_{t\to 0} \frac{t^2 k_x^2}{t} = \lim_{t\to 0} t k_x^2 = 0.$$

Because Λ is not continuous, $\{x : |\Lambda(x)| \leq 1\}$ has empty interior. □

4.2.4. (a) $\Rightarrow$ (b): Let $D \subset X$ be dense in X such that $D \subset \bigcup_{n=1}^{\infty} nC$. Given $h \in D$, then $\frac{1}{n}h \in C$ for some $n \in \mathbb{N}$. Suppose $\phi \in \partial\delta_C(0)$. Then

$$\phi\left(\frac{1}{n}h\right) - \phi(0) \leq \delta_C\left(\frac{1}{n}h\right) - \delta_C(0) = 0.$$

Therefore, $\phi(x) \leq 0$ for all x in the dense subset D of X. Therefore $\phi = 0$.
(b) $\Rightarrow$ (c): Let $f := d_C$. Observe that $0 \in \partial f(0) \subset \partial\delta_C(0) = \{0\}$, and so $\partial f(0) = \{0\}$.
(c) $\Rightarrow$ (a): We suppose (a) is not true and proceed by contraposition. Because $\bigcup_{n=1}^{\infty} nC$ is not dense in X, the complement of the closure is a nonempty open set, and so we choose $h \in X$ and $r > 0$ such that $(h + rB_X) \cap (\bigcup_{n=1}^{\infty} nC) = \emptyset$. Then $\frac{1}{k}(h + rB_X) \cap (\bigcup_{n=1}^{\infty} nC) = \emptyset$ for all $k \in \mathbb{N}$, for otherwise we would have

$$\frac{1}{k}h + \frac{1}{k}ru = m\bar{x} \;\text{ for some } k, m \in \mathbb{N},\; u \in B_X,\; \bar{x} \in C.$$

Then $h + ru = km\bar{x}$ and so $h + ru \in kmC$ which is a contradiction. Consequently, $d_C\left(\frac{1}{k}h\right) \geq \frac{1}{k}r$ for all $k \in \mathbb{N}$ (since $\frac{1}{k}(h + rB_X) \cap (\bigcup_{n=1}^{\infty} nC) = \emptyset$ for all $k \in \mathbb{N}$). Therefore, $f'(0; h) \geq r$, and the max formula (4.1.10) ensures $\partial f(0) \neq \{0\}$, thus (c) does not hold. □

4.2.5.(a) $\Rightarrow$ (b): Let K be as given, and let $f := \delta_K$, the indicator function of K. Then f is not continuous at 0 since 0 is not in the interior of the domain of f. However, $\partial f(0) = \{0\}$ by Exercise 4.2.4.

(b) $\Rightarrow$ (a): Here is an argument slightly different from the suggestion in the hint. Let f be a function as given in (b), and let $\phi \in \partial f(0)$. By replacing f with $f - \phi - f(0)$, we may assume that $\partial f(0) = \{0\}$ and $f(x) \geq f(0) = 0$ for all $x \in X$. Now f is not continuous at 0, because f is not Gâteaux differentiable at 0 (Proposition 4.2.2). Let $K = \{x \in B_X : f(x) \leq 1\}$. Then K is a closec covnex set with $0 \in K$, but $0 \notin \operatorname{int} K$.

If $\bigcup_{n=1}^{\infty} nK$ is norm dense in X, then we are done. So we suppose not, and choose $h \in X$ and $r > 0$ such that $h + rB_X \cap (\bigcup_{n=1}^{\infty} nC) = \emptyset$. Choose $\phi_n \in S_{X^*}$ separating nC and $h + rB_X$ so that $\sup_{nC} \Lambda_n + r \leq \Lambda_n(h)$. By the weak*-compactness of B_{X^*} we let Λ be the weak*-limit point of a convergent subnet of Λ_n. Then $\Lambda(h) \geq r$ (since $0 \in nC$, $\sup_{nC} \Lambda_n \geq 0$ and so $\Lambda_n h \geq r$ for each n). Moreover, $\sup_C \Lambda = 0$, otherwise if $\Lambda(\bar{x}) = \alpha > 0$ for some $\bar{x} \in C$, then $\Lambda_n(\bar{x}) > \alpha/2$ for infinitely many n, and so $r \geq \limsup_{nC} \Lambda_n \to \infty$ which is impossible. Therefore, $\Lambda(x) \leq 0$ for all $x \in C$.

We will obtain a contradiction by showing $\Lambda \in \partial f(0)$ (since $\Lambda \neq 0$). Indeed, suppose $x \in \operatorname{dom} f$, then choose $0 < t < 1$ so small that $\|tx\| \leq 1$ and $tf(x) \leq 1$. Then, $f(tx) \leq (1-t)f(0) + tf(x) \leq 1$ and so $tx \in C$. Thus $\lambda(tx) \leq 0$. Consequently, $\Lambda(x) - \Lambda(0) \leq 0 \leq f(x) - f(0)$ as required. □

4.2.6. (a) Let C be a closed convex set with empty interior and let $\bar{x} \in C$. Replacing C with $C - \bar{x}$ it suffices to show f is not Fréchet differentiable at 0 where $f := d_C$. Because $0 \in \partial f(0)$, it is the candidate for the derivative. Suppose f is Fréchet differentiable at 0. Then there exists $n \in \mathbb{N}$ so that

$$|f(0+h) - f(0) - \langle 0, h\rangle| \leq \frac{1}{4}\|h\| \quad \text{whenever } \|h\| \leq \frac{1}{n}.$$

Thus $d_C(h) \leq \|h\|/4$ whenever $\|h\| \leq 1/n$. Now consider the closed convex set nC which also has empty interior. Fix $h \in S_X$. Then we find $x_0 \in C$ such that $\|x_0 - n^{-1}h\| < 1/(3n)$. Thus $nx_0 \in nC$ and $\|nx_0 - h\| < 1/3$. Thus $d_{nC}(h) < 1/3$ for each $h \in S_X$.

On the other hand, because nC has empty interior, we know $\frac{1}{4}B_X \not\subset nC$. Thus we choose y_0 with $\|y_0\| \leq 1/4$ so that $y_0 \notin nC$. According to the basic separation theorem (4.1.12) there exists $\phi \in S_{X^*}$ so that $\sup_{nC} \phi < \phi(y_0) < 1/3$. This contradicts the property that $d_{nC}(h) < 1/3$ for each $h \in S_X$ (since we can take $h_0 \in S_X$ with $\phi(h_0) > 2/3$ and then if $x \in nC$ satisfies $\|x - h_0\| < 1/3$, we get the contradiction $\phi(x) > 1/3$). Hence we conclude d_C is not Fréchet differentiable at 0.

(b) Exercise 4.2.1(b) gives an explicit example on c_0 where d_C is Gâteaux differentiable at $0 \in C$, but C is a closed convex set with empty interior. Exercise 4.2.4 can be used to produce similar examples on any infinite dimensional separable Banach space using a set C as in Example 4.2.6.

Further notes. The fact that d_C can be Gâteaux differentiable at some $x \in C$ for C as in (a) makes the statement not quite as obvious as it might seem at first glance. However, when C is a closed convex set with nonempty interior, then d_C is not Gâteaux differentiable at any x in the boundary of C. Indeed, let $\bar{x} \in \operatorname{bnd} C$. According to the separation theorem (4.1.15), there exists $\phi \in S_{X^*}$ such that $\sup_C \phi = \phi(\bar{x})$ and $\langle \phi, x\rangle < \langle \phi, \bar{x}\rangle$ for every $x \in \operatorname{int} C$. Thus $\phi \neq 0$, and $\phi \in \partial d_C(\bar{x})$; indeed for fixed $x \in X$ and any $y \in C$,

$$\langle \phi, x - \bar{x}\rangle = \langle \phi, x - y\rangle + \langle \phi, y - \bar{x}\rangle \leq \langle \phi, x - y\rangle \leq \|x - y\|.$$

Taking the infimum over $y \in C$, shows $\phi \in d_C(\bar{x})$ as claimed. Because $0 \in \partial d_C(\bar{x})$, it follows that $\partial d_C(\bar{x})$ is not a singleton, and so d_C is not Gâteaux differentiable at $\bar{x}$. □

4.2.7. This modifies the proof given for the Fréchet differentiable case, essentially replacing norm convergence with weak*-convergence.

(a) ⇒ (e): Suppose that (e) does not hold, then there exist $\epsilon_n \to 0^+$, $\phi_n \in \partial_{\epsilon_n} f(x_0)$, $\phi \in \partial f(x_0)$, $h \in S_X$ and $\epsilon > 0$ such that

$$\langle \phi_n - \phi, h \rangle > \epsilon \quad \text{for all } n.$$

Let $t_n = 2\epsilon_n/\epsilon$. Then

$$\begin{aligned} \frac{t_n \epsilon}{2} &\leq t_n \epsilon - \epsilon_n \leq \phi_n(t_n h) - \phi(t_n h) - \epsilon_n \\ &\leq f(x_0 + t_n h) - f(x_0) - \phi(t h_n). \end{aligned}$$

and so (a) does not hold.

(e) ⇒ (d): Suppose $\phi_n \in \partial_{\epsilon_n} f(x_n)$, $\Lambda_n \in \partial_{\epsilon_n} f(y_n)$ where (x_n) and (y_n) converge in norm to x_0 and $\epsilon_n \to 0^+$. Because f is continuous at x_0, we know that f has Lipschitz constant, say M, on $B_{2r}(x_0)$ for some $r > 0$ (Proposition 4.1.4). In the case $\|x_n - x_0\| \leq r$, one can check that $\|\phi_n\| \leq M + \epsilon_n/r$. Consequently, $\phi_n(x_n) - \phi_n(x_0) \to 0$. Thus, for $y \in X$, one has

$$\begin{aligned} \phi_n(y) - \phi_n(x_0) &= \phi_n(y) - \phi_n(x_n) + \phi_n(x_n) - \phi_n(x_0) \\ &\leq f(y) - f(x_n) + \epsilon_n + \phi_n(x_n) - \phi_n(x_0) \\ &\leq f(y) - f(x_0) + \epsilon_n + |f(x_n) - f(x_0)| + |\phi_n(x_0) - \phi_n(x_0)|. \end{aligned}$$

Consequently, $\phi_n \in \partial_{\epsilon_n'} f(x_0)$ where $\epsilon_n' = \epsilon_n + |f(x_n) - f(x_0)| + |\phi_n(x_n) - \phi_n(x_0)|$ and $\epsilon_n' \to 0^+$. According to (d), $\phi_n \to_{w^*} \phi$. Similarly, $\Lambda_n \to_{w^*} \phi$ and so (c) holds.

(d) ⇒ (c): Letting $\Lambda_n = \phi$ and $y_n = x_0$, we see that (c) follows directly from (d).

(c) ⇒ (b): This follows because the continuity of f at x_0, implies f is continuous in a neighborhood of x_0 (Proposition 4.1.4), and so by the max formula (4.1.10), $\partial f(x_n)$ is eventually nonempty whenever $x_n \to x_0$.

(b) ⇒ (a) We prove the contrapositive. So we suppose f is not Gâteaux differentiable at x_0. Then there exist $t_n \downarrow 0$, $h \in B_X$ and $\epsilon > 0$ such that

$$f(x_0 + t_n h) - f(x_0) - \phi(t_n h) > \epsilon t_n \quad \text{where} \quad \phi \in \partial f(x_0).$$

Let $\phi_n \in \partial f(x_0 + t_n h)$ (for sufficiently large n by the max formula (4.1.10)). Now,

$$\phi_n(t_n h) \geq f(x_0 + t_n h) - f(x_0) \geq \phi(t_n h) + \epsilon t_n$$

and so $\phi_n \not\to_{w^*} \phi$. Thus (b) fails. □

4.2.9. (a) Let for $f := \|\cdot\|_1$ the usual norm on ℓ_1, consider $x^* := (2^{-i}) \in \ell_1$. Then $\phi \in \partial f(x^*)$ if and only if $\|\phi\| = 1$ and $\phi(x^*) = 1$. Thus $\phi = (1, 1, 1, \ldots) \in \ell_\infty$. Therefore, $\partial f(x^*) \cap c_0 = \emptyset$.

(b) Observe that the proof of the weak* epi-separation theorem (4.1.23) works the same $x_0^* \in \operatorname{dom} f$ as it does for $x_0^* \in \operatorname{cont} f$. Thus let $x_0^* \in \operatorname{dom} f$. Given $\epsilon > 0$, apply the weak* epi-separation theorem (4.1.23) with $\alpha = f(x_0^*) - \epsilon$ to find $x_0 \in X$ such that

$$\langle x_0, x^* - x_0 \rangle \leq f(x^*) - \alpha = f(x^*) - f(x_0^*) + \epsilon \quad \text{for all } x^* \in X^*.$$

Therefore $x_0 \in \partial_\epsilon f(x_0^*) \cap X$ as desired. □

Exercises from Section 4.3

4.3.1. The 'if' part was done, as suggested for the converse let (x_i) be any Cauchy sequence in X. Define the function $f(x) := \lim_{i\to\infty} d(x_i, x)$. Then f is lower semicontinuous with $\inf_X f = 0$. Let $\epsilon \in (0, 1)$. Choose $z \in X$ such that $f(z) < \epsilon$. Applying Ekeland's variational principle (4.3.1) with $\lambda = 1$, there exists $y \in X$ such that

$$f(y) + \epsilon d(z, y) \le f(z) \quad \text{and} \quad f(x) + \epsilon d(x, y) \ge f(y) \text{ for all } x \in X.$$

In particular, $f(y) < \epsilon$ and $f(x_n) + \epsilon d(x_n, y) \ge f(y)$. Taking the limit on the left-hand side of this equation, we get $\epsilon f(y) \ge f(y)$. Because $f(y) \ge 0$ this implies $f(y) = 0$, that is, $x_n \to y$. Thus (X, d) is complete. □

4.3.2. The contractivity of the mapping implies there can be at most one fixed point. As suggested, let $f(x) := d(x, \phi(x))$. Then f is lower semicontinuous and bounded below on X. Let $z \in X$ be such that $f(z) < \inf_X f + \epsilon$. Applying Ekeland's variational principle (4.3.1) to f with $\lambda = 1$ and $0 < \epsilon < 1 - k$, there exists $y \in X$ such that $f(x) + \epsilon d(x, y) \ge f(y)$ for all $x \in X$. In particular, with $x = \phi(y)$, we have

$$d(\phi(y), \phi(\phi(y)) + \epsilon d(\phi(y), y)) \ge d(y, \phi(y)),$$

or in other words, $d(\phi(y), \phi(\phi(y)) \ge (1 - \epsilon)d(y, \phi(y)) > kd(y, \phi(y))$ which is a contradiction. □

4.3.5. Let $0 < \epsilon < 1$. Suppose $x_0 \in S_X$ and $\phi_0 \in S_{X^*}$ satisfy $\phi_0(x_0) > 1 - \epsilon^2/2$. Then $\phi_0 \in \partial_{\epsilon^2/2}\|x_0\|$. According to the Brøndsted–Rockafellar theorem (4.3.2) there exists $x \in X$ with $\|x - x_0\| \le \epsilon/2$ and $\phi \in \partial\|x\|$ with $\|\phi - \phi_0\| \le \epsilon$. Now let $\bar{x} = x/\|x\|$ and $\bar{\phi} = \phi$. Then $\|\bar{x}\| = 1$, $\|\bar{\phi}\| = 1$ and $\bar{\phi}(\bar{x}) = 1$ as desired. □

4.3.6.(a) Let $0 < \epsilon < 1$. Fix $x_0 \in \operatorname{bnd}(C)$ and choose $x_1 \in X \setminus C$ such that $\|x_1 - x_0\| < \epsilon$. By the basic separation theorem (4.1.12) we choose $x_0^* \in S_{X^*}$ so that $\sigma_C(x_0^*) < \langle x_0^*, x_1\rangle$. Because $\|x_1 - x_0\| < \epsilon$, it follows that $\langle x_0^*, x_0\rangle > \sigma_C(x_0^*) - \epsilon$ and thus $x_0^* \in \partial_\epsilon f(x_0)$ where $f := \delta_C$. Applying the Brøndsted–Rockafellar theorem (4.3.2) with $\lambda = \sqrt{\epsilon}$ we obtain $x \in \operatorname{dom} f$ and $x^* \in \partial f(x)$ such that $\|x - x_0\| \le \sqrt{\epsilon}$ and $\|x^* - x_0^*\| \le \sqrt{\epsilon}$. In particular, $x^* \ne 0$, and $\langle x^*, x\rangle = \sigma_C(x^*)$.

(b) Suppose $\sigma_C(x_0^*) < \infty$ and let $0 < \epsilon < \|x_0\|^2$. Choose $x_0 \in C$ such that $\langle x_0^*, x_0\rangle > \sigma_C(x_0^*) - \epsilon$. Then $x_0^* \in \partial_\epsilon f(x_0)$ where $f := \delta_C$. As in (a), we apply the Brøndsted–Rockafellar theorem (4.3.2) with $\lambda = \sqrt{\epsilon}$ to find $x \in \operatorname{dom} f$ and $x^* \in \partial f(x)$ such that $\|x - x_0\| \le \sqrt{\epsilon}$ and $\|x^* - x_0^*\| \le \sqrt{\epsilon}$. In particular, $x^* \ne 0$, and $\langle x^*, x\rangle = \sigma_C(x^*)$. □

Exercises from Section 4.4

4.4.11. See solution to Exercise 2.3.12(a)(i). □

4.4.12. Using the definition and subdifferential sum rule (4.1.19), we observe

$$N_{C_1\cap C_2}(x) = \partial(\delta_{C_1}(x) + \delta_{C_2}(x)) = \partial\delta_{C_1}(x) + \partial\delta_{C_2}(x) = N_{C_1}(x) + N_{C_2}(x).$$

□

4.4.14. See solution to Exercise 2.3.12(a)(iii). $\square$

4.4.20. (a) Suppose X is reflexive, then the properness and coercivity of f imply that $K := \{x : f(x) \leq M\}$ is nonempty and bounded for some $M > 0$. Because f is lower semicontinuous and convex, K is weakly closed and convex. Therefore K weakly compact. Now f is bounded below on K, therefore, let $(x_n) \subset K$ be such that $f(x_n) \to \inf_K f$. According to the Eberlein-Šmulian theorem, (x_{n_k}) converges weakly to $\bar{x}$ for some subsequence and some $\bar{x}$. Because f is weakly lower semicontinuous, this implies $\liminf f(x_{n_k}) \geq f(\bar{x})$. Thus $f(\bar{x}) = \inf_K f$. Outside of K, $f(x) > M$, and so $f(\bar{x})$ is an absolute minimum for f.
Conversely, suppose X is not reflexive. According to James' theorem (4.1.27) there is a functional $\phi \in S_{X^*}$ that does not attain its norm on B_X. Then $f := \phi + \delta_{B_X}$ does not attain its minimum on X. Indeed, $\inf_X f = -1$, but there is no $x \in B_X$ such that $|\phi(x)| = 1$.
(b) For each $\phi \in X^*$, $f - \phi$ is supercoercive. Now $f - \phi$ attains its minimum at $\bar{x}$ for some $\bar{x} \in X$. Thus, $\phi \in \partial f(\bar{x})$.
(c) No. For example let f be defined by $f(x) := (\sup_{n\in\mathbb{N}} |x(n)-1|)^2 + \sum_{n=1}^{\infty} \frac{1}{2^n}|x(n)-1|^2$ where $x = (x(n)) \in c_0$. Then f is a continuous (in fact bounded on bounded sets) and supercoercive convex function that does not attain its minimum. Indeed, $f(x) > 1$ for all $x \in c_0$, and for $x_n := (1, 1, \ldots, 1, 0, \ldots)$, we have $\lim_{n\to\infty} f(x_n) = 1$. Consequently, $0 \notin \operatorname{range} \partial f$. More generally, in any nonreflexive space, consider $f := \|\cdot\|^2$. Using James' theorem (4.1.27), one can show that the subdifferential map is not onto. $\square$

4.4.21. (a) Because f is Lipschitz, there exists $K > 0$ so that $\operatorname{dom} f^* \subset KB_{X^*}$. According to the hypothesis, we choose $N \in \mathbb{N}$ so that $f_n^*(\phi) \leq f^*(\phi) + \epsilon$ for each $\phi \in \partial f(X) \subset KB_{X^*}$, and $n \geq N$. Now let $x \in X$, and let $\phi \in \partial f(x)$. Then for each $n \geq N$, one has

$$f^*(\phi) = \phi(x) - f(x) \geq f_n^*(\phi) - \epsilon \geq \phi(x) - f_n(x) - \epsilon.$$

Therefore, $f(x) - \epsilon \leq f_n(x) \leq f(x)$ for all $x \in X$, and all $n \geq N$.
(b) Let $x_0 \in X$. Given $\epsilon > 0$, and any number $\alpha < f(x_0)$ it suffices to show that there exists N such that $f_n(x_0) > \alpha - \epsilon$ for all $n \geq N$. Now choose $\phi \in X^*$ such that $\phi(x) - \phi(x_0) \leq f(x) - \alpha$ for all $x \in X$. Then $f^*(\phi) \leq \phi(x_0) - \alpha$ and so $\phi \in \operatorname{dom} f^*$. Now choose $N \in \mathbb{N}$ so that $f_n^*(\phi) \leq f^*(\phi) - \epsilon$ for $n \geq N$. Now, for all $n \geq N$,

$$\phi(x_0) - \alpha \geq f^*(\phi) \geq f_n^*(\phi) - \epsilon \geq \phi(x_0) - f_n(x_0) - \epsilon.$$

Therefore, $f_n(x_0) \geq \alpha - \epsilon$.
(c) We conclude that $f \,\square\, n\|\cdot\|^2$ converges uniformly (resp. pointwise) to f provided that f is Lipschitz (resp. lower semicontinuous proper) and convex because

$$(f \,\square\, n\|\cdot\|^2)^* = f^* + \frac{1}{2n}\|\cdot\|_*^2,$$

converges uniformly (resp. pointwise) to f^* (where $\|\cdot\|_*$ denotes the dual norm to $\|\cdot\|$). $\square$

4.4.22. We first prove Fact 4.4.4(b). Suppose f is continuous at x_0, then there exist $\delta > 0$ and $M > 0$ so that $f(x) \leq M$ for $x \in x_0 + \delta B_X$ (Proposition 4.1.4). Now suppose $x^{**} \in X^{**}$ and $\|x^{**} - x_0\| \leq \delta$. According to Goldstine's theorem (Exercise 4.1.13) there is a net $(x_\alpha) \subset \delta B_X$ with $x_\alpha \to_{w^*} (x^{**} - x_0)$. The weak*-lower -semicontinuity of f^{**} implies that

$$f^{**}(x^{**}) \leq \liminf_\alpha f^{**}(x_0 + x_\alpha) = \liminf_\alpha f(x_0 + x_\alpha) \leq M.$$

Thus f^{**} is bounded above on a neighborhood of x^{**}, and thus it is continuous at x^{**} (Proposition 4.1.4).

Conversely, if f^{**} is continuous at $x_0 \in X$, then so is f because $f^{**}|_X = f$ (Proposition 4.4.2(a)).

We now prove Fact 4.4.4(c). Suppose f is Fréchet differentiable at x_0. Let $\epsilon > 0$, the according to Proposition 4.2.7 there exists $\delta > 0$ so that

$$f(x_0 + h) + f(x_0 - h) - 2f(x_0) \le \epsilon\|h\| \text{ if } \|h\| < \delta.$$

Now suppose $h \in X^{**}$ and $\|h\| < \delta$. Then there exist $h_\alpha \in X$ with $\|h_\alpha\| = \|h\| < \delta$ and $h_\alpha \to_{w^*} h$ by Goldstine's theorem (Exercise 4.1.13). Using the weak*-lower semicontinuity of f^{**} and the fact $f^{**}|_X = f$ (Proposition 4.4.2(a)) we have

$$\begin{aligned} f^{**}(x_0 + h) &+ f^{**}(x_0 - h) - 2f^{**}(x_0) \\ &\le \lim_\alpha f^{**}(x_0 + h_\alpha) + f^{**}(x_0 - h_\alpha) - 2f^{**}(x_0) \\ &= \lim_\alpha f(x_0 + h_\alpha) + f(x_0 - h_\alpha) - 2f(x_0) \le \epsilon\|h\|. \end{aligned}$$

Applying Proposition 4.2.7 we conclude that f^{**} is Fréchet differentiable at x_0.

The converse, as in (b), follows from Proposition 4.4.2(a). □

4.4.23. (a) (i) ⇒ (ii): If f is supercoercive, it is clear $f - y^*$ is supercoercive for each $y^* \in X^*$. The Moreau-Rockafellar theorem (4.4.11) implies the equivalence of (ii) and (iii).

We now show (ii) ⇒ (i) when X is finite dimensional. Indeed, suppose by what of contradiction there exists $(x_n) \subset X$ such that $\|x_n\| \to \infty$ but $\limsup f(x_n)/\|x_n\| \le K$ for some $K > 0$. Let $u_n = x_n/\|x_n\|$, and by passing to a subsequence as necessary, we suppose $u_n \to u$. Now let $\phi \in S_X$ be such that $\phi(u) = 1$. Then let $y^* = 3K\phi$. Now $f - y^*$ is coercive, and for n sufficiently large $y^*(u_n) > 2K$ and $(f - y^*)(x_n) > 0$ and so $f(x_n) > y^*(x_n) \ge 2K\|x_n\|$ which is a contradiction.

For (b), define the conjugate function f^* by $f^*(x) := \|x\|^2 + \sum_{n=1}^{\infty}(x_i)^{2n}$, f^* is a continuous convex function and supercoercive. However, f^* is not bounded on $2B_{\ell_2}$, since $f(2e_n) = 2^{2n}$. Therefore, $f = f^{**}$ cannot be supercoercive. For (c), consider $f := \sqrt{\|\cdot\|}$. □

Exercises from Section 4.5

4.5.1. An elementary proof is as follows. Let $\epsilon_n \to 0^+$. By the definition of the conjugate choose $x_n \in X$ such that $x^*(x_n) - f(x_n) \ge f^*(x^*) - \epsilon_n$ (then $f(x_n)$ is necessarily real-valued) and $f(x_n) + f^*(x^*) \le x^*(x_n) + \epsilon_n$. Now $f^{**} \le f$ and f^{**} is proper, so $f^{**}(x_n) + f^*(x^*) \le x^*(x_n) + \epsilon_n$. According to Proposition 4.4.1(b), $x_n \in \partial_{\epsilon_n} f^*(x^*)$. By Šmulian's theorem, $x_n \to \nabla f^*(x^*)$. Therefore, $\nabla f^*(x^*) \in X$, and we let $x := \nabla f^*(x^*)$. According to the Fenchel–Young Proposition 4.4.1(a), we have $f^{**}(x) = \langle x^* x\rangle - f^*(x^*)$ and so

$$\begin{aligned} f(x) \ge f^{**}(x) &= \langle x^*, x\rangle - f^*(x^*) = \lim_{n\to\infty} \langle x^*, x_n\rangle - f^*(x^*) \\ &\ge \liminf_{n\to\infty} f(x_n) - \epsilon_n \ge f(x) \text{ using lower semicontinuous property of } f. \end{aligned}$$

Thus $f(x) = f^{**}(x)$ as desired. □

4.5.2. (a) Choose $(x_n) \subset C$ such that $\|x_n - x\| \to d_C(x)$. Because (x_n) is bounded the Eberlein-Šmulian theorem ensures there is a subsequence (x_{n_k}) that converges weakly to $\bar{x}$. Because C is closed and convex, it is weakly closed and so $\bar{x} \in C$. The weak lower semicontinuity of the norm implies $\liminf_{n\to\infty} \|x_n - x\| \geq \|\bar{x} - x\|$ and so $\bar{x} \in P_C(x)$ as desired.

(b) Fix $x \in X$, and suppose $P_C(x)$ is not empty. If $x \in C$, then $P_C(x) = \{x\}$ is a singleton as desired. So, suppose $x \notin C$, and suppose distinct $y, z \in P_C(x)$. Then $\frac{1}{2}(y + z) \in C$ by convexity, and the strict convexity of the norm implies

$$\left\| x - \frac{1}{2}(y+z) \right\| < \frac{1}{2}\|x - y\| + \frac{1}{2}\|x - z\| = P_C(x).$$

This contradiction implies $P_C(x)$ is a singleton.

Parts (a) and (b) together show that every closed convex subset of a strictly convex reflexive Banach space is a Čebyšev set. □

4.5.3. Let $\phi \in S_X$ be a functional that does not attain its norm on B_X, and let $C := \{x \in X : \phi(x) = 0\}$. Fix $\bar{x} \in X$ such that $\phi(\bar{x}) \neq 0$. Say, $\phi(\bar{x}) = \alpha$. It is elementary to check $d_C(\bar{x}) = |\alpha|$. Because ϕ does not attain its norm, we have $\alpha = |\phi(\bar{x} - y)| < \|x - y\|$ for all $y \in C$. Thus $P_C(\bar{x}) = \emptyset$. □

4.5.4. Let A be a closed proximal subset of X. Now observe that $\nabla d_A^2(x) = 0$ whenever $x \in A$, and $\nabla d_A^2(x_n) \to 0$ whenever $d_A(x_n) \to 0$. When $x \notin A$, then $\nabla d_A^2(x) = 2d_A(x)\nabla d_A(x)$, and so by the hypothesis of the exercise, ∇d_A is norm-to-weak* continuous at any $x \notin A$.

Suppose $x \notin A$. Then d_A is Gâteaux differentiable at x. Let $r := d_A(x)$, and let $\bar{x} \in P_A(x)$. Then $r = \|x - \bar{x}\| > 0$. Let $h := \frac{1}{r}(\bar{x} - x)$. Then

$$\nabla d_A(x)(h) = \lim_{t\to 0} \frac{d_A(x + th) - d_A(x)}{t}.$$

However,

$$\begin{aligned} \lim_{t\to 0} \frac{d_A(x + th) - d_A(x)}{t} &\leq \lim_{t\to 0^+} \frac{\|x + th - \bar{x}\| - r}{t} \\ &= \lim_{t\to 0^+} \frac{\left(1 - \frac{t}{r}\right)\|x - \bar{x}\| - r}{t} = -1. \end{aligned}$$

Thus $\|\nabla d_A(x)\| \geq 1$. However, because d_A has Lipschitz constant 1, we conclude $\|\nabla d_A(x)\| = 1$. Now suppose $x_n \to x$. Then eventually $x_n \notin A$, and $\|\nabla d_A(x_n)\| = 1$. Because the dual norm has the weak*-Kadec property, we deduce $d_A(x_n) \to d_A(x)$, and thus ∇d_A^2 is norm-to-norm continuous. □

4.5.5. Let X be a reflexive Banach space. Suppose the dual norm on X^* is strictly convex. Let $x \in X \setminus \{0\}$. Suppose $\phi, \Lambda \in S_{X^*}$ satisfy $\phi(x) = \|x\| = \Lambda(x)$. Then $\|\phi + \Lambda\| \geq (\phi + \Lambda)(x/\|x\|) = 2$. By the strict convexity of the dual norm, $\phi = \Lambda$. Thus $\partial\|x\|$ is a singleton, and so $\|\cdot\|$ is Gâteaux differentiable at x.

Conversely, suppose $x, y \in S_{X^*}$ are such that $\|x + y\| = 2$. Because of reflexivity, we can choose $\phi \in S_X$ so that $\phi(x + y) = 2$. Thus $x, y \in \partial\|\phi\|$. By Gâteaux differentiability, $x = y$. □