Chapter 2
Hilbert Spaces

Throughout this book, $\mathcal{H}$ is a real Hilbert space with scalar (or inner) product $\langle \cdot \mid \cdot \rangle$. The associated norm is denoted by $\|\cdot\|$ and the associated distance by d, i.e.,

$$(\forall x \in \mathcal{H})(\forall y \in \mathcal{H}) \quad \|x\| = \sqrt{\langle x \mid x \rangle} \quad \text{and} \quad d(x,y) = \|x - y\|. \tag{2.1}$$

The identity operator on $\mathcal{H}$ is denoted by Id.

In this chapter, we derive useful identities and inequalities, and we review examples and basic results from linear and nonlinear analysis in a Hilbert space setting.

2.1 Notation and Examples

The *orthogonal complement* of a subset C of $\mathcal{H}$ is denoted by $C^\perp$, i.e.,

$$C^\perp = \big\{ u \in \mathcal{H} \;\big|\; (\forall x \in C) \;\; \langle x \mid u \rangle = 0 \big\}. \tag{2.2}$$

An orthonormal subset C of $\mathcal{H}$ is an *orthonormal basis* of $\mathcal{H}$ if $\overline{\operatorname{span}}\, C = \mathcal{H}$. The space $\mathcal{H}$ is *separable* if it possesses a countable orthonormal basis. Now let $(x_i)_{i \in I}$ be a family of vectors in $\mathcal{H}$ and let $\mathcal{I}$ be the class of nonempty finite subsets of I, directed by $\subset$. Then $(x_i)_{i \in I}$ is *summable* if there exists $x \in \mathcal{H}$ such that the net $(\sum_{i \in J} x_i)_{J \in \mathcal{I}}$ converges to x, i.e., by (1.26),

$$(\forall \varepsilon \in \mathbb{R}_{++})(\exists\, K \in \mathcal{I})(\forall J \in \mathcal{I}) \quad J \supset K \;\;\Rightarrow\;\; \Big\| x - \sum_{i \in J} x_i \Big\| \leqslant \varepsilon. \tag{2.3}$$

H.H. Bauschke, P.L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, CMS Books in Mathematics,
DOI 10.1007/978-3-319-48311-5_2

In this case we write $x = \sum_{i\in I} x_i$. For every family $(\alpha_i)_{i\in I}$ in $[0,+\infty]$, we have
$$\sum_{i\in I} \alpha_i = \sup_{J\in\mathcal{I}} \sum_{i\in J} \alpha_i. \tag{2.4}$$

Here are specific real Hilbert spaces that will be used in this book.

Example 2.1 Let I be a nonempty set. The *Hilbert direct sum* of a family of real Hilbert spaces $(\mathcal{H}_i, \|\cdot\|_i)_{i\in I}$ is the real Hilbert space
$$\bigoplus_{i\in I} \mathcal{H}_i = \Big\{ \boldsymbol{x} = (x_i)_{i\in I} \in \bigtimes_{i\in I} \mathcal{H}_i \;\Big|\; \sum_{i\in I} \|x_i\|_i^2 < +\infty \Big\} \tag{2.5}$$
equipped with the addition $(\boldsymbol{x},\boldsymbol{y}) \mapsto (x_i + y_i)_{i\in I}$, the scalar multiplication $(\alpha, \boldsymbol{x}) \mapsto (\alpha x_i)_{i\in I}$, and the scalar product
$$(\boldsymbol{x},\boldsymbol{y}) \mapsto \sum_{i\in I} \langle x_i \mid y_i \rangle_i\,, \tag{2.6}$$
where $\langle \cdot \mid \cdot \rangle_i$ denotes the scalar product of $\mathcal{H}_i$ (when I is finite, we shall sometimes adopt a common practice and write $\bigtimes_{i\in I}\mathcal{H}_i$ instead of $\bigoplus_{i\in I}\mathcal{H}_i$). Now suppose that, for every $i \in I$, $f_i\colon \mathcal{H}_i \to \left]-\infty,+\infty\right]$ and that, if I is infinite, $\inf_{i\in I} f_i \geqslant 0$. Then
$$\bigoplus_{i\in I} f_i\colon \bigoplus_{i\in I} \mathcal{H}_i \to \left]-\infty,+\infty\right] : (x_i)_{i\in I} \mapsto \sum_{i\in I} f_i(x_i). \tag{2.7}$$

Example 2.2 If each $\mathcal{H}_i$ is the Euclidean line $\mathbb{R}$ in Example 2.1, then we obtain $\ell^2(I) = \bigoplus_{i\in I}\mathbb{R}$, which is equipped with the scalar product $(x,y) = ((\xi_i)_{i\in I}, (\eta_i)_{i\in I}) \mapsto \sum_{i\in I} \xi_i\eta_i$. The standard unit vectors $(e_i)_{i\in I}$ of $\ell^2(I)$ are defined by
$$(\forall i \in I) \quad e_i\colon I \to \mathbb{R}\colon j \mapsto \begin{cases} 1, & \text{if } j = i; \\ 0, & \text{otherwise.} \end{cases} \tag{2.8}$$

Example 2.3 If $I = \{1,\ldots,N\}$ in Example 2.2, then we obtain the standard Euclidean space $\mathbb{R}^N$.

Example 2.4 Let M and N be strictly positive integers. Then $\mathbb{R}^{M\times N}$ denotes the Hilbert space of real $M \times N$ matrices equipped with the scalar product $(A,B) \mapsto \operatorname{tra}(A^{\mathsf{T}}B)$, where tra is the trace function. The associated norm is the *Frobenius norm* $\|\cdot\|_{\mathsf{F}}$.

Example 2.5 Let N be a strictly positive integer. The Hilbert space $\mathbb{S}^N$ is the subspace of $\mathbb{R}^{N\times N}$ that consists of all the symmetric matrices.

Example 2.6 Let $(\Omega, \mathcal{F}, \mu)$ be a (positive) measure space. A property is said to hold μ-*almost everywhere* (μ-a.e.) on Ω if there exists a set $C \in \mathcal{F}$ such that $\mu(C) = 0$ and the property holds on $\Omega \smallsetminus C$. Let $(\mathsf{H}, \langle\cdot\mid\cdot\rangle_{\mathsf{H}})$ be a separable real Hilbert space, and let $p \in \left[1,+\infty\right[$. Denote by $L^p((\Omega,\mathcal{F},\mu);\mathsf{H})$ the space of (equivalence classes of) Borel measurable functions $x\colon \Omega \to \mathsf{H}$ such that

$\int_\Omega \|x(\omega)\|_{\mathsf{H}}^p \mu(d\omega) < +\infty$. Then $L^2((\Omega,\mathcal{F},\mu);\mathsf{H})$ is a real Hilbert space with scalar product $(x,y) \mapsto \int_\Omega \langle x(\omega) \mid y(\omega)\rangle_{\mathsf{H}}\, \mu(d\omega)$.

Example 2.7 In Example 2.6, let $\mathsf{H} = \mathbb{R}$. Then we obtain the real Banach space $L^p(\Omega,\mathcal{F},\mu) = L^p((\Omega,\mathcal{F},\mu);\mathbb{R})$ and, for $p = 2$, the real Hilbert space $L^2(\Omega,\mathcal{F},\mu)$, which is equipped with the scalar product $(x,y) \mapsto \int_\Omega x(\omega)y(\omega)\mu(d\omega)$.

Example 2.8 In Example 2.6, let $T \in \mathbb{R}_{++}$, set $\Omega = [0,T]$, and let μ be the Lebesgue measure. Then we obtain the Hilbert space $L^2([0,T];\mathsf{H})$, which is equipped with the scalar product $(x,y) \mapsto \int_0^T \langle x(t) \mid y(t)\rangle_{\mathsf{H}}\, dt$. In particular, when $\mathsf{H} = \mathbb{R}$, we obtain the classical Lebesgue space $L^2([0,T]) = L^2([0,T];\mathbb{R})$.

Example 2.9 Let $(\Omega,\mathcal{F},\mathsf{P})$ be a *probability space*, i.e., a measure space such that $\mathsf{P}(\Omega) = 1$. A property that holds P-almost everywhere on Ω is said to hold *almost surely* (a.s.). A *random variable* (r.v.) is a measurable function $X\colon \Omega \to \mathbb{R}$, and its expected value is $\mathsf{E}X = \int_\Omega X(\omega)\mathsf{P}(d\omega)$, provided that the integral exists. In this context, Example 2.7 yields the Hilbert space

$$L^2(\Omega,\mathcal{F},\mathsf{P}) = \big\{X \text{ r.v. on } (\Omega,\mathcal{F},\mathsf{P}) \;\big|\; \mathsf{E}|X|^2 < +\infty\big\} \tag{2.9}$$

of random variables with finite second absolute moment, which is equipped with the scalar product $(X,Y) \mapsto \mathsf{E}(XY)$.

Example 2.10 Let $T \in \mathbb{R}_{++}$ and let $(\mathsf{H}, \langle\cdot \mid \cdot\rangle_{\mathsf{H}})$ be a separable real Hilbert space. For every $y \in L^2([0,T];\mathsf{H})$, the function $x\colon [0,T] \to \mathsf{H}\colon t \mapsto \int_0^t y(s)ds$ is differentiable almost everywhere (a.e.) on $]0,T[$ with $x'(t) = y(t)$ a.e. on $]0,T[$. We say that $x\colon [0,T] \to \mathsf{H}$ belongs to $W^{1,2}([0,T];\mathsf{H})$ if there exists $y \in L^2([0,T];\mathsf{H})$ such that

$$(\forall t \in [0,T]) \quad x(t) = x(0) + \int_0^t y(s)ds. \tag{2.10}$$

Alternatively,

$$W^{1,2}([0,T];\mathsf{H}) = \big\{x \in L^2([0,T];\mathsf{H}) \;\big|\; x' \in L^2([0,T];\mathsf{H})\big\}. \tag{2.11}$$

The scalar product of this real Hilbert space is $(x,y) \mapsto \int_0^T \langle x(t) \mid y(t)\rangle_{\mathsf{H}}\, dt + \int_0^T \langle x'(t) \mid y'(t)\rangle_{\mathsf{H}}\, dt$.

2.2 Basic Identities and Inequalities

Fact 2.11 (Cauchy–Schwarz) *Let x and y be in $\mathcal{H}$. Then*

$$|\langle x \mid y\rangle| \leqslant \|x\|\, \|y\|. \tag{2.12}$$

Moreover, $\langle x \mid y\rangle = \|x\|\,\|y\| \Leftrightarrow (\exists\, \alpha \in \mathbb{R}_+)\ x = \alpha y$ or $y = \alpha x$.

Lemma 2.12 *Let x, y, and z be in $\mathcal{H}$. Then the following hold:*

(i) $\|x+y\|^2 = \|x\|^2 + 2\langle x \mid y\rangle + \|y\|^2$.
(ii) *Parallelogram identity:* $\|x+y\|^2 + \|x-y\|^2 = 2\|x\|^2 + 2\|y\|^2$.
(iii) *Polarization identity:* $4\langle x \mid y\rangle = \|x+y\|^2 - \|x-y\|^2$.
(iv) *Apollonius's identity:* $\|x-y\|^2 = 2\|z-x\|^2 + 2\|z-y\|^2 - 4\|z-(x+y)/2\|^2$.

Proof. (i): A simple expansion.

(ii)&(iii): It follows from (i) that

$$\|x-y\|^2 = \|x\|^2 - 2\langle x \mid y\rangle + \|y\|^2. \tag{2.13}$$

Adding this identity to (i) yields (ii), and subtracting it from (i) yields (iii).

(iv): Apply (ii) to the points $(z-x)/2$ and $(z-y)/2$. ☐

Lemma 2.13 *Let x and y be in $\mathcal{H}$. Then the following hold:*

(i) $\langle x \mid y\rangle \leqslant 0 \Leftrightarrow (\forall \alpha \in \mathbb{R}_+)\,\|x\| \leqslant \|x-\alpha y\| \Leftrightarrow (\forall \alpha \in [0,1])\,\|x\| \leqslant \|x-\alpha y\|$.
(ii) $x \perp y \Leftrightarrow (\forall \alpha \in \mathbb{R})\ \|x\| \leqslant \|x-\alpha y\| \Leftrightarrow (\forall \alpha \in [-1,1])\ \|x\| \leqslant \|x-\alpha y\|$.

Proof. (i): Observe that

$$(\forall \alpha \in \mathbb{R}) \quad \|x-\alpha y\|^2 - \|x\|^2 = \alpha\big(\alpha\|y\|^2 - 2\langle x \mid y\rangle\big). \tag{2.14}$$

Hence, the forward implications follow immediately. Conversely, if for every $\alpha \in \,]0,1]$, $\|x\| \leqslant \|x-\alpha y\|$, then (2.14) implies that $\langle x \mid y\rangle \leqslant \alpha\|y\|^2/2$. As $\alpha \downarrow 0$, we obtain $\langle x \mid y\rangle \leqslant 0$.

(ii): A consequence of (i), since $x \perp y \Leftrightarrow [\langle x \mid y\rangle \leqslant 0 \text{ and } \langle x \mid -y\rangle \leqslant 0]$. ☐

Lemma 2.14 *Let $(x_i)_{i\in I}$ and $(u_i)_{i\in I}$ be finite families in $\mathcal{H}$ and let $(\alpha_i)_{i\in I}$ be a family in $\mathbb{R}$ such that $\sum_{i\in I}\alpha_i = 1$. Then the following hold:*

(i) $\big\langle \sum_{i\in I}\alpha_i x_i \mid \sum_{j\in I}\alpha_j u_j\big\rangle + \sum_{i\in I}\sum_{j\in I}\alpha_i\alpha_j\langle x_i - x_j \mid u_i - u_j\rangle/2$
$= \sum_{i\in I}\alpha_i\langle x_i \mid u_i\rangle$.

(ii) $\big\|\sum_{i\in I}\alpha_i x_i\big\|^2 + \sum_{i\in I}\sum_{j\in I}\alpha_i\alpha_j\|x_i - x_j\|^2/2 = \sum_{i\in I}\alpha_i\|x_i\|^2$.

Proof. (i): We have

$$\begin{aligned}
2\Big\langle \sum_{i\in I}\alpha_i x_i \,\Big|\, \sum_{j\in I}\alpha_j u_j\Big\rangle
&= \sum_{i\in I}\sum_{j\in I}\alpha_i\alpha_j\big(\langle x_i \mid u_j\rangle + \langle x_j \mid u_i\rangle\big)\\
&= \sum_{i\in I}\sum_{j\in I}\alpha_i\alpha_j\big(\langle x_i \mid u_i\rangle + \langle x_j \mid u_j\rangle - \langle x_i - x_j \mid u_i - u_j\rangle\big)\\
&= 2\sum_{i\in I}\alpha_i\langle x_i \mid u_i\rangle - \sum_{i\in I}\sum_{j\in I}\alpha_i\alpha_j\langle x_i - x_j \mid u_i - u_j\rangle.
\end{aligned} \tag{2.15}$$

(ii): This follows from (i) when $(u_i)_{i\in I} = (x_i)_{i\in I}$. ☐

The following two results imply that Hilbert spaces are uniformly convex and strictly convex Banach spaces, respectively.

Corollary 2.15 *Let $x \in \mathcal{H}$, let $y \in \mathcal{H}$, and let $\alpha \in \mathbb{R}$. Then*

$$\|\alpha x + (1-\alpha)y\|^2 + \alpha(1-\alpha)\|x-y\|^2 = \alpha\|x\|^2 + (1-\alpha)\|y\|^2. \tag{2.16}$$

Corollary 2.16 *Suppose that x and y are distinct points in $\mathcal{H}$ such that $\|x\| = \|y\|$, and let $\alpha \in \,]0,1[$. Then $\|\alpha x + (1-\alpha)y\| < \|x\|$.*

Proof. An immediate consequence of Corollary 2.15. ☐

Lemma 2.17 *Let $(x,y) \in \mathcal{H} \times \mathcal{H}$. Then the following hold:*

(i) *Let $\alpha \in \,]0,1[$. Then*

$$\begin{aligned}\alpha^2\big(\|x\|^2 - \|(1-\alpha^{-1})x + \alpha^{-1}y\|^2\big)& \\ &= (2\alpha-1)\|x\|^2 + 2(1-\alpha)\langle x \mid y\rangle - \|y\|^2 \\ &= 2(1-\alpha)\langle x \mid y\rangle - \big(\|y\|^2 + (1-2\alpha)\|x\|^2\big) \\ &= \alpha\big(\|x\|^2 - \alpha^{-1}(1-\alpha)\|x-y\|^2 - \|y\|^2\big).\end{aligned}$$

(ii) *We have*

$$\begin{aligned}\|x\|^2 - \|2y-x\|^2 &= 4\big(\langle x \mid y\rangle - \|y\|^2\big) \\ &= 4\langle x-y \mid y\rangle \\ &= 2\big(\|x\|^2 - \|x-y\|^2 - \|y\|^2\big).\end{aligned}$$

Proof. (i): These identities follow from Lemma 2.12(i).
(ii): Divide by α^2 in (i) and set $\alpha = 1/2$. ☐

The following inequality is classical.

Fact 2.18 (Hardy–Littlewood–Pólya) (See [196, Theorems 368 and 369]) Let x and y be in $\mathbb{R}^N$, and let $x_\downarrow$ and $y_\downarrow$ be, respectively, their rearrangement vectors with entries ordered decreasingly. Then

$$\langle x \mid y\rangle \leqslant \langle x_\downarrow \mid y_\downarrow\rangle, \tag{2.17}$$

and equality holds if and only if there exists a permutation matrix P of size $n \times n$ such that $Px = x_\downarrow$ and $Py = y_\downarrow$.

2.3 Linear Operators and Functionals

Let $\mathcal{X}$ and $\mathcal{Y}$ be real normed vector spaces. We set

$$\mathcal{B}(\mathcal{X},\mathcal{Y}) = \big\{T\colon \mathcal{X} \to \mathcal{Y} \;\big|\; T \text{ is linear and continuous}\big\} \tag{2.18}$$

and $\mathcal{B}(\mathcal{X}) = \mathcal{B}(\mathcal{X}, \mathcal{X})$. Equipped with the norm

$$(\forall T \in \mathcal{B}(\mathcal{X}, \mathcal{Y})) \quad \|T\| = \sup \big\|T(B(0;1))\big\| = \sup_{\substack{x \in \mathcal{X}, \\ \|x\| \leqslant 1}} \|Tx\|, \tag{2.19}$$

$\mathcal{B}(\mathcal{X}, \mathcal{Y})$ is a normed vector space, and it is a Banach space if $\mathcal{Y}$ is a Banach space.

Example 2.19 Let $A \in \mathbb{R}^{M \times N}$. Then $A \in \mathcal{B}(\mathbb{R}^N, \mathbb{R}^M)$ and the operator norm of A given by (2.19) is the spectral norm of A, i.e., the largest singular value of A, and it is denoted by $\|A\|_2$. We have $\|A\|_2 \leqslant \|A\|_{\mathsf{F}}$.

Fact 2.20 *Let $\mathcal{X}$ and $\mathcal{Y}$ be real normed vector spaces and let $T \colon \mathcal{X} \to \mathcal{Y}$ be linear. Then T is continuous at a point in $\mathcal{X}$ if and only if it is Lipschitz continuous.*

Fact 2.21 (See [116, Proposition III.6.1]) Let $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$ be real normed vector spaces and let $T \colon \mathcal{X} \oplus \mathcal{Y} \to \mathcal{Z}$ be a bilinear operator. Then T is continuous if and only if

$$(\exists\, \beta \in \mathbb{R}_+)(\forall x \in \mathcal{X})(\forall y \in \mathcal{Y}) \quad \|T(x,y)\| \leqslant \beta \|x\|\, \|y\|. \tag{2.20}$$

The following result is also known as the *Banach–Steinhaus theorem*.

Lemma 2.22 (Uniform boundedness principle) *Let $\mathcal{X}$ be a real Banach space, let $\mathcal{Y}$ be a real normed vector space, and let $(T_i)_{i \in I}$ be a family of operators in $\mathcal{B}(\mathcal{X}, \mathcal{Y})$ that is* pointwise bounded, *i.e., $(\forall x \in \mathcal{X})$ $\sup_{i \in I} \|T_i x\| < +\infty$. Then $(T_i)_{i \in I}$ is* uniformly bounded, *i.e., $\sup_{i \in I} \|T_i\| < +\infty$.*

Proof. Apply Lemma 1.44(i) to $\big(\{x \in \mathcal{X} \mid \sup_{i \in I} \|T_i x\| \leqslant n\}\big)_{n \in \mathbb{N}}$. ☐

Definition 2.23 Let $T \colon \mathcal{H} \to \mathcal{H}$ be linear and let $\alpha \in \mathbb{R}_{++}$. Then T is:

(i) *monotone* if $(\forall x \in \mathcal{H})$ $\langle Tx \mid x \rangle \geqslant 0$;
(ii) *strictly monotone* if $(\forall x \in \mathcal{H} \smallsetminus \{0\})$ $\langle Tx \mid x \rangle > 0$;
(iii) *α-strongly monotone* if $(\forall x \in \mathcal{H})$ $\langle Tx \mid x \rangle \geqslant \alpha \|x\|^2$.

The Riesz–Fréchet representation theorem states that every continuous linear functional on the real Hilbert space $\mathcal{H}$ can be identified with a vector in $\mathcal{H}$.

Fact 2.24 (Riesz–Fréchet representation) *Let $f \in \mathcal{B}(\mathcal{H}, \mathbb{R})$. Then there exists a unique vector $u \in \mathcal{H}$ such that $(\forall x \in \mathcal{H})$ $f(x) = \langle x \mid u \rangle$. Moreover, $\|f\| = \|u\|$.*

If $\mathcal{K}$ is a real Hilbert space and $T \in \mathcal{B}(\mathcal{H}, \mathcal{K})$, then the *adjoint* of T is the unique operator $T^* \in \mathcal{B}(\mathcal{K}, \mathcal{H})$ that satisfies

$$(\forall x \in \mathcal{H})(\forall y \in \mathcal{K}) \quad \langle Tx \mid y \rangle = \langle x \mid T^* y \rangle\, . \tag{2.21}$$

Fact 2.25 *Let $\mathcal{K}$ be a real Hilbert space, let $T \in \mathcal{B}(\mathcal{H},\mathcal{K})$, and let* $\ker T = \{x \in \mathcal{H} \mid Tx = 0\}$ *be the kernel of T. Then the following hold:*

(i) $T^{**} = T$.
(ii) $\|T^*\| = \|T\| = \sqrt{\|T^*T\|}$.
(iii) *Suppose that $\mathcal{K} = \mathcal{H}$ and that $T = T^*$. Then*

$$\|T\| = \sup\{|\langle Tx \mid x\rangle| \mid x \in B(0;1)\}. \tag{2.22}$$

(iv) $(\ker T)^\perp = \overline{\operatorname{ran}}\, T^*$.
(v) $(\operatorname{ran} T)^\perp = \ker T^*$.
(vi) $\ker T^*T = \ker T$ *and* $\overline{\operatorname{ran}}\, TT^* = \overline{\operatorname{ran}}\, T$.
(vii) $(\operatorname{gra} T)^\perp = \{(u,v) \in \mathcal{H} \oplus \mathcal{K} \mid u = -T^*v\}$.

Fact 2.26 *Let $\mathcal{K}$ be a real Hilbert space and let $T \in \mathcal{B}(\mathcal{H},\mathcal{K})$. Then* $\operatorname{ran} T$ *is closed* $\Leftrightarrow$ $\operatorname{ran} T^*$ *is closed* $\Leftrightarrow$ $\operatorname{ran} TT^*$ *is closed* $\Leftrightarrow$ $\operatorname{ran} T^*T$ *is closed* $\Leftrightarrow$ $(\exists\, \alpha \in \mathbb{R}_{++})(\forall x \in (\ker T)^\perp)\ \|Tx\| \geqslant \alpha\|x\|$.

Suppose that $\mathcal{H} \neq \{0\}$. Let $f \colon \mathcal{H} \to \mathbb{R}$ be nonzero and linear, and let $\eta \in \mathbb{R}$. A *hyperplane* in $\mathcal{H}$ is a set of the form

$$\{x \in \mathcal{H} \mid f(x) = \eta\}, \tag{2.23}$$

and it is closed if and only if f is continuous; if it is not closed, it is dense in $\mathcal{H}$. Alternatively, let $u \in \mathcal{H} \smallsetminus \{0\}$. Then it follows from Fact 2.24 that a *closed hyperplane* in $\mathcal{H}$ is a set of the form

$$\{x \in \mathcal{H} \mid \langle x \mid u\rangle = \eta\}. \tag{2.24}$$

Moreover, a *closed half-space* with *outer normal* u is a set of the form

$$\{x \in \mathcal{H} \mid \langle x \mid u\rangle \leqslant \eta\}, \tag{2.25}$$

and an *open half-space* with *outer normal* u is a set of the form

$$\{x \in \mathcal{H} \mid \langle x \mid u\rangle < \eta\}. \tag{2.26}$$

The distance function to $C = \{x \in \mathcal{H} \mid \langle x \mid u\rangle = \eta\}$ is (see (1.47))

$$d_C \colon \mathcal{H} \to \mathbb{R}_+ : x \mapsto \frac{|\langle x \mid u\rangle - \eta|}{\|u\|}. \tag{2.27}$$

We conclude this section with an example of a discontinuous linear functional.

Example 2.27 Assume that $\mathcal{H}$ is infinite-dimensional and let H be a Hamel basis of $\mathcal{H}$, i.e., a maximally linearly independent subset. Then H is uncountable. Indeed, if $\mathcal{H} = \bigcup_{n\in\mathbb{N}} \operatorname{span}\{h_k\}_{0\leqslant k\leqslant n}$ for some Hamel basis

$H = \{h_n\}_{n\in\mathbb{N}}$, then Lemma 1.44(i) implies that some finite-dimensional linear subspace $\operatorname{span}\{h_k\}_{0\leqslant k\leqslant n}$ has nonempty interior, which is absurd. The Gram–Schmidt orthonormalization procedure thus guarantees the existence of an orthonormal set $B = \{e_n\}_{n\in\mathbb{N}}$ and an uncountable set $C = \{c_a\}_{a\in A}$ such that $B \cup C$ is a Hamel basis of $\mathcal{H}$. Thus, every point in $\mathcal{H}$ is a (finite) linear combination of elements in $B \cup C$ and, therefore, the function

$$f\colon \mathcal{H} \to \mathbb{R}\colon x = \sum_{n\in\mathbb{N}} \xi_n e_n + \sum_{a\in A} \eta_a c_a \mapsto \sum_{n\in\mathbb{N}} \xi_n \tag{2.28}$$

is well defined and linear. Now take $(\alpha_n)_{n\in\mathbb{N}} \in \ell^2(\mathbb{N}) \smallsetminus \ell^1(\mathbb{N})$ (e.g., $(\forall n \in \mathbb{N})$ $\alpha_n = 1/(n+1)$) and set

$$(\forall n \in \mathbb{N}) \quad x_n = \sum_{k=0}^{n} \alpha_k e_k. \tag{2.29}$$

Then $(x_n)_{n\in\mathbb{N}}$ converges strongly to some point $z \in \mathcal{H}$ and $f(x_n) \to +\infty$. This proves that f is discontinuous at z and hence discontinuous everywhere by Fact 2.20. Now set $(\forall n \in \mathbb{N})$ $y_n = (x_n - f(x_n)e_0)/\max\{f(x_n), 1\}$. Then $(y_n)_{n\in\mathbb{N}}$ lies in $C = \big\{x \in \mathcal{H} \;\big|\; f(x) = 0\big\}$ and $y_n \to -e_0$. On the other hand, $-e_0 \notin C$, since $f(-e_0) = -1$. As a result, the hyperplane C is not closed. In fact, as will be proved in Example 8.42, C is dense in $\mathcal{H}$.

2.4 Strong and Weak Topologies

The metric topology of $(\mathcal{H}, d)$ is called the *strong topology* (or *norm topology*) of $\mathcal{H}$. Thus, a net $(x_a)_{a\in A}$ in $\mathcal{H}$ converges strongly to a point x if $\|x_a - x\| \to 0$; in symbols, $x_a \to x$. When used without modifiers, topological notions in $\mathcal{H}$ (closedness, openness, neighborhood, continuity, compactness, convergence, etc.) will always be understood with respect to the strong topology.

Fact 2.28 *Let U and V be closed linear subspaces of $\mathcal{H}$ such that V has finite dimension or finite codimension. Then $U + V$ is a closed linear subspace.*

In addition to the strong topology, a very important alternative topology can be introduced.

Definition 2.29 The family of all finite intersections of open half-spaces of $\mathcal{H}$ forms the base of the *weak topology* of $\mathcal{H}$; $\mathcal{H}^{\text{weak}}$ denotes the resulting topological space.

Lemma 2.30 *$\mathcal{H}^{\text{weak}}$ is a Hausdorff space.*

Proof. Suppose that x and y are distinct points in $\mathcal{H}$. Set $u = x - y$ and $w = (x+y)/2$. Then $\big\{z \in \mathcal{H} \;\big|\; \langle z - w \mid u\rangle > 0\big\}$ and $\big\{z \in \mathcal{H} \;\big|\; \langle z - w \mid u\rangle < 0\big\}$ are disjoint weak neighborhoods of x and y, respectively. ☐

A subset of $\mathcal{H}$ is weakly open if it is a union of finite intersections of open half-spaces. If $\mathcal{H}$ is infinite-dimensional, nonempty intersections of finitely many open half-spaces are unbounded and, therefore, nonempty weakly open sets are unbounded. A net $(x_a)_{a\in A}$ in $\mathcal{H}$ *converges weakly* to a point $x \in \mathcal{H}$ if, for every $u \in \mathcal{H}$, $\langle x_a \mid u\rangle \to \langle x \mid u\rangle$; in symbols, $x_a \rightharpoonup x$. Moreover (see Section 1.7), a subset C of $\mathcal{H}$ is *weakly closed* if the weak limit of every weakly convergent net in C is also in C, and *weakly compact* if every net in C has a weak cluster point in C. Likewise (see Section 1.11), a subset C of $\mathcal{H}$ is *weakly sequentially closed* if the weak limit of every weakly convergent sequence in C is also in C, and *weakly sequentially compact* if every sequence in C has a weak sequential cluster point in C. Finally, let D be a nonempty subset of $\mathcal{H}$, let $\mathcal{K}$ be a real Hilbert space, let $T\colon D \to \mathcal{K}$, and let $f\colon \mathcal{H} \to [-\infty, +\infty]$. Then T is *weakly continuous* if it is continuous with respect to the weak topologies on $\mathcal{H}$ and $\mathcal{K}$, i.e., if, for every net $(x_a)_{a\in A}$ in D such that $x_a \rightharpoonup x \in D$, we have $Tx_a \rightharpoonup Tx$. Likewise, f is *weakly lower semicontinuous* at $x \in \mathcal{H}$ if, for every net $(x_a)_{a\in A}$ in $\mathcal{H}$ such that $x_a \rightharpoonup x$, we have $f(x) \leqslant \underline{\lim}\, f(x_a)$.

Remark 2.31 Strong and weak convergence of a net $(x_a)_{a\in A}$ in $\mathcal{H}$ to a point x in $\mathcal{H}$ can be interpreted in geometrical terms: $x_a \to x$ means that $d_{\{x\}}(x_a) \to 0$ whereas, by (2.27), $x_a \rightharpoonup x$ means that $d_C(x_a) \to 0$ for every closed hyperplane C containing x.

Example 2.32 Suppose that $\mathcal{H}$ is infinite-dimensional, let $(x_n)_{n\in\mathbb{N}}$ be an orthonormal sequence in $\mathcal{H}$, and let u be a point in $\mathcal{H}$. Bessel's inequality yields $\sum_{n\in\mathbb{N}} |\langle x_n \mid u\rangle|^2 \leqslant \|u\|^2$, hence $\langle x_n \mid u\rangle \to 0$. Thus $x_n \rightharpoonup 0$. However, $\|x_n\| \equiv 1$ and therefore $x_n \not\to 0$. Actually, $(x_n)_{n\in\mathbb{N}}$ has no Cauchy subsequence since, for any two distinct positive integers n and m, we have $\|x_n - x_m\|^2 = \|x_n\|^2 + \|x_m\|^2 = 2$. This also shows that the unit sphere $\big\{x \in \mathcal{H} \;\big|\; \|x\| = 1\big\}$ is closed but not weakly sequentially closed.

Suppose that $\mathcal{H}$ is infinite-dimensional. As seen in Example 2.32, an orthonormal sequence in $\mathcal{H}$ has no strongly convergent subsequence. Hence, it follows from Fact 1.39 that the closed unit ball of $\mathcal{H}$ is not compact. This property characterizes infinite-dimensional Hilbert spaces.

Fact 2.33 *The following are equivalent:*

(i) *$\mathcal{H}$ is finite-dimensional.*
(ii) *The closed unit ball $B(0;1)$ of $\mathcal{H}$ is compact.*
(iii) *The weak topology of $\mathcal{H}$ coincides with its strong topology.*
(iv) *The weak topology of $\mathcal{H}$ is metrizable.*

In striking contrast, compactness of closed balls always holds in the weak topology. This fundamental and deep result is known as the *Banach–Alaoglu–Bourbaki theorem.*

Fact 2.34 (Banach–Alaoglu–Bourbaki) *The closed unit ball $B(0;1)$ of $\mathcal{H}$ is weakly compact.*

Fact 2.35 (See [192, p. 181] and [2, Theorems 6.30&6.34]) The weak topology of the closed unit ball $B(0;1)$ of $\mathcal{H}$ is metrizable if and only if $\mathcal{H}$ is separable.

Lemma 2.36 *Let C be a subset of $\mathcal{H}$. Then C is weakly compact if and only if it is weakly closed and bounded.*

Proof. First, suppose that C is weakly compact. Then Lemma 1.12 and Lemma 2.30 assert that C is weakly closed. Now set $\mathcal{C} = \{\langle x \mid \cdot \rangle\}_{x \in C} \subset \mathcal{B}(\mathcal{H}, \mathbb{R})$ and take $u \in \mathcal{H}$. Then $\langle \cdot \mid u \rangle$ is weakly continuous. By Lemma 1.20, $\{\langle x \mid u \rangle\}_{x \in C}$ is a compact subset of $\mathbb{R}$, and it is therefore bounded by Lemma 1.41. Hence, $\mathcal{C}$ is pointwise bounded, and Lemma 2.22 implies that $\sup_{x \in C} \|x\| < +\infty$, i.e., that C is bounded. Conversely, suppose that C is weakly closed and bounded, say $C \subset B(0;\rho)$ for some $\rho \in \mathbb{R}_{++}$. By Fact 2.34, $B(0;\rho)$ is weakly compact. Using Lemma 1.12 in $\mathcal{H}^{\mathrm{weak}}$, we deduce that C is weakly compact. ☐

The following important fact states that weak compactness and weak sequential compactness coincide.

Fact 2.37 (Eberlein–Šmulian) *Let C be a subset of $\mathcal{H}$. Then C is weakly compact if and only if it is weakly sequentially compact.*

Corollary 2.38 *Let C be a subset of $\mathcal{H}$. Then the following are equivalent:*

(i) *C is weakly compact.*
(ii) *C is weakly sequentially compact.*
(iii) *C is weakly closed and bounded.*

Proof. Combine Lemma 2.36 and Fact 2.37. ☐

Lemma 2.39 *Let C be a bounded subset of $\mathcal{H}$. Then C is weakly closed if and only if it is weakly sequentially closed.*

Proof. If C is weakly closed, it is weakly sequentially closed. Conversely, suppose that C is weakly sequentially closed. By assumption, there exists $\rho \in \mathbb{R}_{++}$ such that $C \subset B(0;\rho)$. Since $B(0;\rho)$ is weakly sequentially compact by Fact 2.34 and Fact 2.37, it follows from Lemma 2.30 and Lemma 1.34 that C is weakly sequentially compact. In turn, appealing once more to Fact 2.37, we obtain the weak compactness of C and therefore its weak closedness by applying Lemma 1.12 in $\mathcal{H}^{\mathrm{weak}}$. ☐

Remark 2.40 As will be seen in Example 3.33, weakly sequentially closed sets need not be weakly closed.

Lemma 2.41 *Let $\mathcal{K}$ be a real Hilbert space, and let $T\colon \mathcal{H} \to \mathcal{K}$ be a continuous affine operator. Then T is weakly continuous.*

Proof. Set $L\colon x \mapsto Tx - T0$, let $x \in \mathcal{H}$, let $y \in \mathcal{K}$, and let $(x_a)_{a\in A}$ be a net in $\mathcal{H}$ such that $x_a \rightharpoonup x$. Then $L \in \mathcal{B}(\mathcal{H},\mathcal{K})$ and $\langle x_a \mid L^*y\rangle \to \langle x \mid L^*y\rangle$. Hence, $\langle Lx_a \mid y\rangle \to \langle Lx \mid y\rangle$, i.e., $Lx_a \rightharpoonup Lx$. We conclude that $Tx_a = T0 + Lx_a \rightharpoonup T0 + Lx = Tx$. ☐

Lemma 2.42 *The norm of $\mathcal{H}$ is weakly lower semicontinuous, i.e., for every net $(x_a)_{a\in A}$ in $\mathcal{H}$ and every x in $\mathcal{H}$, we have*

$$x_a \rightharpoonup x \quad \Rightarrow \quad \|x\| \leqslant \underline{\lim}\, \|x_a\|. \tag{2.30}$$

Proof. Take a net $(x_a)_{a\in A}$ in $\mathcal{H}$ and a point x in $\mathcal{H}$ such that $x_a \rightharpoonup x$. Then, by Cauchy–Schwarz, $\|x\|^2 = \lim |\langle x_a \mid x\rangle| \leqslant \underline{\lim}\, \|x_a\|\, \|x\|$. ☐

Lemma 2.43 *Let $\mathcal{G}$ and $\mathcal{K}$ be real Hilbert spaces and let $T\colon \mathcal{H}\oplus\mathcal{G} \to \mathcal{K}$ be a bilinear operator such that*

$$(\exists\, \beta \in \mathbb{R}_{++})(\forall x \in \mathcal{H})(\forall u \in \mathcal{G}) \quad \|T(x,u)\| \leqslant \beta\|x\|\, \|u\|. \tag{2.31}$$

Let $(x_a)_{a\in A}$ be a net in $\mathcal{H}$, let $(u_a)_{a\in A}$ be a net in $\mathcal{G}$, let $x \in \mathcal{H}$, and let $u \in \mathcal{G}$. Suppose that $(x_a)_{a\in A}$ is bounded, that $x_a \rightharpoonup x$, and that $u_a \to u$. Then $T(x_a,u_a) \rightharpoonup T(x,u)$.

Proof. Since $\sup_{a\in A}\|x_a\| < +\infty$ and $\|u_a - u\| \to 0$, we have $\|T(x_a, u_a - u)\| \leqslant \beta(\sup_{b\in A}\|x_b\|)\|u_a - u\| \to 0$. On the other hand, $T(\cdot,u) \in \mathcal{B}(\mathcal{H},\mathcal{K})$ by Fact 2.21. Thus, $T(\cdot,u)$ is weakly continuous by Lemma 2.41. In turn, $T(x_a - x, u) \rightharpoonup T(0,u) = 0$. Altogether $T(x_a,u_a) - T(x,u) = T(x_a, u_a - u) + T(x_a - x, u) \rightharpoonup 0$. ☐

Lemma 2.44 *Let $(x_a)_{a\in A}$ and $(u_a)_{a\in A}$ be nets in $\mathcal{H}$, and let x and u be points in $\mathcal{H}$. Suppose that $(x_a)_{a\in A}$ is bounded, that $x_a \rightharpoonup x$, and that $u_a \to u$. Then $\langle x_a \mid u_a\rangle \to \langle x \mid u\rangle$.*

Proof. Apply Lemma 2.43 to $\mathcal{G} = \mathcal{H}$, $\mathcal{K} = \mathbb{R}$, and $F = \langle\cdot \mid \cdot\rangle$. ☐

2.5 Weak Convergence of Sequences

Lemma 2.45 *Let $(x_n)_{n\in\mathbb{N}}$ be a bounded sequence in $\mathcal{H}$. Then $(x_n)_{n\in\mathbb{N}}$ possesses a weakly convergent subsequence.*

Proof. First, recall from Lemma 2.30 that $\mathcal{H}^{\mathrm{weak}}$ is a Hausdorff space. Now set $\rho = \sup_{n\in\mathbb{N}}\|x_n\|$ and $C = B(0;\rho)$. Fact 2.34 and Fact 2.37 imply that C is weakly sequentially compact. Since $(x_n)_{n\in\mathbb{N}}$ lies in C, the claim follows from Definition 1.33. ☐

Lemma 2.46 *Let $(x_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{H}$. Then $(x_n)_{n\in\mathbb{N}}$ converges weakly if and only if it is bounded and possesses at most one weak sequential cluster point.*

Proof. Suppose that $x_n \rightharpoonup x \in \mathcal{H}$. Then it follows from Lemma 2.30 and Fact 1.9 that x is the unique weak sequential cluster point of $(x_n)_{n\in\mathbb{N}}$. Moreover, for every $u \in \mathcal{H}$, $\langle x_n \mid u\rangle \to \langle x \mid u\rangle$ and therefore $\sup_{n\in\mathbb{N}} |\langle x_n \mid u\rangle| < +\infty$. Upon applying Lemma 2.22 to the sequence of continuous linear functionals $(\langle x_n \mid \cdot\rangle)_{n\in\mathbb{N}}$, we obtain the boundedness of $(\|x_n\|)_{n\in\mathbb{N}}$. Conversely, suppose that $(x_n)_{n\in\mathbb{N}}$ is bounded and possesses at most one weak sequential cluster point. Then Lemma 2.45 asserts that it possesses exactly one weak sequential cluster point. Moreover, it follows from Fact 2.34 and Fact 2.37 that $(x_n)_{n\in\mathbb{N}}$ lies in a weakly sequentially compact set. Therefore, appealing to Lemma 2.30, we apply Lemma 1.35 in $\mathcal{H}^{\text{weak}}$ to obtain the conclusion. ☐

Lemma 2.47 *Let $(x_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{H}$ and let C be a nonempty subset of $\mathcal{H}$. Suppose that, for every $x \in C$, $(\|x_n - x\|)_{n\in\mathbb{N}}$ converges and that every weak sequential cluster point of $(x_n)_{n\in\mathbb{N}}$ belongs to C. Then $(x_n)_{n\in\mathbb{N}}$ converges weakly to a point in C.*

Proof. By assumption, $(x_n)_{n\in\mathbb{N}}$ is bounded. Therefore, in view of Lemma 2.46, it is enough to show that $(x_n)_{n\in\mathbb{N}}$ cannot have two distinct weak sequential cluster points in C. To this end, let x and y be weak sequential cluster points of $(x_n)_{n\in\mathbb{N}}$ in C, say $x_{k_n} \rightharpoonup x$ and $x_{l_n} \rightharpoonup y$. Since x and y belong to C, the sequences $(\|x_n - x\|)_{n\in\mathbb{N}}$ and $(\|x_n - y\|)_{n\in\mathbb{N}}$ converge. In turn, since

$$(\forall n \in \mathbb{N}) \quad 2\,\langle x_n \mid x - y\rangle = \|x_n - y\|^2 - \|x_n - x\|^2 + \|x\|^2 - \|y\|^2, \tag{2.32}$$

$(\langle x_n \mid x - y\rangle)_{n\in\mathbb{N}}$ converges as well, say $\langle x_n \mid x - y\rangle \to \ell$. Passing to the limit along $(x_{k_n})_{n\in\mathbb{N}}$ and along $(x_{l_n})_{n\in\mathbb{N}}$ yields, respectively, $\ell = \langle x \mid x - y\rangle = \langle y \mid x - y\rangle$. Therefore, $\|x - y\|^2 = 0$ and hence $x = y$. ☐

Proposition 2.48 *Suppose that $(y_n)_{n\in\mathbb{N}}$ is an orthonormal sequence in $\mathcal{H}$ and let $(x_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{H}$ such that $x_n - y_n \to 0$. Then $x_n \rightharpoonup 0$.*

Proof. This follows from Example 2.32. ☐

The next result provides a partial converse to Proposition 2.48.

Proposition 2.49 *Suppose that $(x_n)_{n\in\mathbb{N}}$ is a sequence in $\mathcal{H}$ such that $x_n \rightharpoonup 0$ and $(\forall n \in \mathbb{N})$ $\|x_n\| = 1$. Then there exist an orthonormal sequence $(y_n)_{n\in\mathbb{N}}$ in $\mathcal{H}$ and a subsequence $(x_{k_n})_{n\in\mathbb{N}}$ of $(x_n)_{n\in\mathbb{N}}$ such that $x_{k_n} - y_n \to 0$.*

Proof. Let $(\varepsilon_n)_{n\in\mathbb{N}}$ be a sequence in $]0, 1/2[$ such that $\varepsilon_n \to 0$. Set $V = \overline{\operatorname{span}}\,\{x_n\}_{n\in\mathbb{N}}$ and let $(e_n)_{n\in\mathbb{N}}$ be an orthonormal basis of V. Let $l_0 \in \mathbb{N}$. Since $x_n \rightharpoonup 0$, there exists $k_0 \in \mathbb{N}$ such that $u_0 = \sum_{i=0}^{l_0} \langle x_{k_0} \mid e_i\rangle\, e_i \in B(0; \varepsilon_0)$. Because $(e_n)_{n\in\mathbb{N}}$ is an orthonormal basis of V, there exists $l_1 \in \mathbb{N}$ such that $l_1 > l_0$ and $w_0 = \sum_{i\geqslant l_1+1} \langle x_{k_0} \mid e_i\rangle\, e_i \in B(0; \varepsilon_0)$. We continue in this fashion and thus obtain inductively two strictly increasing sequences $(l_n)_{n\in\mathbb{N}}$ and $(k_n)_{n\in\mathbb{N}}$ in $\mathbb{N}$ such that

$$(\forall n \in \mathbb{N}) \quad u_n = \sum_{i=0}^{l_n} \langle x_{k_n} \mid e_i\rangle\, e_i \in B(0; \varepsilon_n) \tag{2.33}$$

and

$$(\forall n \in \mathbb{N}) \quad w_n = \sum_{i \geqslant l_{n+1}+1} \langle x_{k_n} \mid e_i \rangle \, e_i \in B(0; \varepsilon_n). \tag{2.34}$$

Now set

$$(\forall n \in \mathbb{N}) \quad v_n = \sum_{i=l_n+1}^{l_{n+1}} \langle x_{k_n} \mid e_i \rangle \, e_i = x_{k_n} - u_n - w_n. \tag{2.35}$$

Then $(\forall n \in \mathbb{N})$ $1 = \|x_{k_n}\| \geqslant \|v_n\| \geqslant \|x_{k_n}\| - \|u_n\| - \|w_n\| \geqslant 1 - 2\varepsilon_n > 0$ and $\|x_{k_n} - v_n\| \leqslant \|u_n\| + \|w_n\| \leqslant 2\varepsilon_n$. Moreover, $(v_n)_{n\in\mathbb{N}}$ is an orthogonal sequence since $(e_n)_{n\in\mathbb{N}}$ is an orthonormal basis of V. Finally, set $(\forall n \in \mathbb{N})$ $y_n = v_n/\|v_n\|$. Then $(y_n)_{n\in\mathbb{N}}$ is an orthonormal sequence in $\mathcal{H}$ and $(\forall n \in \mathbb{N})$ $\|x_{k_n} - y_n\| \leqslant \|x_{k_n} - v_n\| + \|v_n - y_n\| = \|x_{k_n} - v_n\| + (1 - \|v_n\|) \leqslant 4\varepsilon_n \to 0$. ☐

Proposition 2.50 *Let $(e_i)_{i\in I}$ be a family in $\mathcal{H}$ such that $\overline{\operatorname{span}}\,\{e_i\}_{i\in I} = \mathcal{H}$, let $(x_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{H}$, and let x be a point in $\mathcal{H}$. Then the following are equivalent:*

(i) $x_n \rightharpoonup x$.
(ii) $(x_n)_{n\in\mathbb{N}}$ *is bounded and* $(\forall i \in I)$ $\langle x_n \mid e_i \rangle \to \langle x \mid e_i \rangle$ *as* $n \to +\infty$.

Proof. (i)⇒(ii): Lemma 2.46.

(ii)⇒(i): Set $(y_n)_{n\in\mathbb{N}} = (x_n - x)_{n\in\mathbb{N}}$. Lemma 2.45 asserts that $(y_n)_{n\in\mathbb{N}}$ possesses a weak sequential cluster point y, say $y_{k_n} \rightharpoonup y$. In view of Lemma 2.46, it suffices to show that $y = 0$. For this purpose, fix $\varepsilon \in \mathbb{R}_{++}$. Then there exists a finite subset J of I such that $\|y - z\| \sup_{n\in\mathbb{N}} \|y_{k_n}\| \leqslant \varepsilon$, where $z = \sum_{j\in J} \langle y \mid e_j \rangle \, e_j$. Thus, by Cauchy–Schwarz,

$$\begin{aligned}(\forall n \in \mathbb{N}) \quad |\langle y_{k_n} \mid y \rangle| &\leqslant |\langle y_{k_n} \mid y - z \rangle| + |\langle y_{k_n} \mid z \rangle| \\ &\leqslant \varepsilon + \sum_{j\in J} |\langle y \mid e_j \rangle| \, |\langle y_{k_n} \mid e_j \rangle| \,. \end{aligned} \tag{2.36}$$

Hence $\overline{\lim}\, |\langle y_{k_n} \mid y \rangle| \leqslant \varepsilon$. Letting $\varepsilon \downarrow 0$ yields $\|y\|^2 = \lim \langle y_{k_n} \mid y \rangle = 0$. ☐

Lemma 2.51 *Let $(x_n)_{n\in\mathbb{N}}$ and $(u_n)_{n\in\mathbb{N}}$ be sequences in $\mathcal{H}$, and let x and u be points in $\mathcal{H}$. Then the following hold:*

(i) $\big[x_n \rightharpoonup x$ *and* $\overline{\lim}\, \|x_n\| \leqslant \|x\|\big] \;\Leftrightarrow\; x_n \to x$.
(ii) *Suppose that $\mathcal{H}$ is finite-dimensional. Then* $x_n \rightharpoonup x \;\Leftrightarrow\; x_n \to x$.
(iii) *Suppose that $x_n \rightharpoonup x$ and $u_n \to u$. Then* $\langle x_n \mid u_n \rangle \to \langle x \mid u \rangle$.

Proof. (i): Suppose that $x_n \rightharpoonup x$ and that $\overline{\lim}\, \|x_n\| \leqslant \|x\|$. Then it follows from Lemma 2.42 that $\|x\| \leqslant \underline{\lim}\, \|x_n\| \leqslant \overline{\lim}\, \|x_n\| \leqslant \|x\|$, hence $\|x_n\| \to \|x\|$. In turn, $\|x_n - x\|^2 = \|x_n\|^2 - 2\langle x_n \mid x \rangle + \|x\|^2 \to 0$. Conversely, suppose that $x_n \to x$. Then $\|x_n\| \to \|x\|$ by continuity of the norm. On the other hand, $x_n \rightharpoonup x$ since for every $n \in \mathbb{N}$ and every $u \in \mathcal{H}$, the Cauchy–Schwarz inequality yields $0 \leqslant |\langle x_n - x \mid u \rangle| \leqslant \|x_n - x\| \, \|u\|$.

(ii): Set $\dim \mathcal{H} = m$ and let $(e_k)_{1\leqslant k\leqslant m}$ be an orthonormal basis of $\mathcal{H}$. Now assume that $x_n \rightharpoonup x$. Then $\|x_n - x\|^2 = \sum_{k=1}^m |\langle x_n - x \mid e_k\rangle|^2 \to 0$.

(iii): Combine Lemma 2.44 and Lemma 2.46. $\square$

The combination of Lemma 2.42 and Lemma 2.51(i) yields the following characterization of strong convergence.

Corollary 2.52 *Let $(x_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{H}$ and let x be in $\mathcal{H}$. Then $x_n \to x \Leftrightarrow \big[x_n \rightharpoonup x$ and $\|x_n\| \to \|x\|\big]$.*

We conclude this section with a consequence of Ostrowski's theorem (Theorem 1.49).

Lemma 2.53 *Suppose that $\mathcal{H}$ is finite-dimensional and let $(x_n)_{n\in\mathbb{N}}$ be a bounded sequence in $\mathcal{H}$ such that $x_n - x_{n+1} \to 0$. Then the set of cluster points of $(x_n)_{n\in\mathbb{N}}$ is compact and connected.*

2.6 Differentiability

In this section, $\mathcal{K}$ is a real Banach space.

Definition 2.54 Let C be a nonempty subset of $\mathcal{H}$, let $T\colon C \to \mathcal{K}$, and suppose that $x \in C$ is such that $(\forall y \in \mathcal{H})(\exists\, \alpha \in \mathbb{R}_{++})\ [x, x+\alpha y] \subset C$. Then T is *Gâteaux differentiable* at x if there exists an operator $\mathsf{D}T(x) \in \mathcal{B}(\mathcal{H},\mathcal{K})$, called the *Gâteaux derivative* of T at x, such that

$$(\forall y \in \mathcal{H}) \quad \mathsf{D}T(x)y = \lim_{\alpha\downarrow 0} \frac{T(x+\alpha y) - T(x)}{\alpha}. \tag{2.37}$$

Moreover, T is Gâteaux differentiable on C if it is Gâteaux differentiable at every point in C. Higher-order Gâteaux derivatives are defined inductively. Thus, the *second Gâteaux derivative* of T at x is the operator $\mathsf{D}^2T(x) \in \mathcal{B}(\mathcal{H}, \mathcal{B}(\mathcal{H},\mathcal{K}))$ that satisfies

$$(\forall y \in \mathcal{H}) \quad \mathsf{D}^2T(x)y = \lim_{\alpha\downarrow 0} \frac{\mathsf{D}T(x+\alpha y) - \mathsf{D}T(x)}{\alpha}. \tag{2.38}$$

The Gâteaux derivative $\mathsf{D}T(x)$ in Definition 2.54 is unique whenever it exists (Exercise 2.23). Moreover, since $\mathsf{D}T(x)$ is linear, for every $y \in \mathcal{H}$, we have $\mathsf{D}T(x)y = -\mathsf{D}T(x)(-y)$, and we can therefore replace (2.37) by

$$(\forall y \in \mathcal{H}) \quad \mathsf{D}T(x)y = \lim_{0\neq\alpha\to 0} \frac{T(x+\alpha y) - T(x)}{\alpha}. \tag{2.39}$$

Remark 2.55 Let C be a subset of $\mathcal{H}$, let $f\colon C \to \mathbb{R}$, and suppose that f is Gâteaux differentiable at $x \in C$. Then, by Fact 2.24, there exists a unique vector $\nabla f(x) \in \mathcal{H}$ such that

$$(\forall y \in \mathcal{H}) \quad \mathsf{D}f(x)y = \langle y \mid \nabla f(x)\rangle . \tag{2.40}$$

We call $\nabla f(x)$ the *Gâteaux gradient* of f at x. If f is Gâteaux differentiable on C, the *gradient operator* is $\nabla f \colon C \to \mathcal{H} \colon x \mapsto \nabla f(x)$. Likewise, if f is twice Gâteaux differentiable at x, we can identify $\mathsf{D}^2 f(x)$ with an operator $\nabla^2 f(x) \in \mathcal{B}(\mathcal{H})$ in the sense that

$$(\forall y \in \mathcal{H})(\forall z \in \mathcal{H}) \quad (\mathsf{D}^2 f(x)y)z = \big\langle z \mid \nabla^2 f(x)y\big\rangle . \tag{2.41}$$

If the convergence in (2.39) is uniform with respect to y on bounded sets, then $x \in \operatorname{int} C$ and we obtain the following notion.

Definition 2.56 Let $x \in \mathcal{H}$, let $C \in \mathcal{V}(x)$, and let $T \colon C \to \mathcal{K}$. Then T is *Fréchet differentiable* at x if there exists an operator $\mathsf{D}T(x) \in \mathcal{B}(\mathcal{H}, \mathcal{K})$, called the *Fréchet derivative* of T at x, such that

$$\lim_{0 \neq \|y\| \to 0} \frac{\|T(x+y) - Tx - \mathsf{D}T(x)y\|}{\|y\|} = 0. \tag{2.42}$$

Moreover, T is Fréchet differentiable on C if it is Fréchet differentiable at every point in C. Higher-order Fréchet derivatives are defined inductively. Thus, the *second Fréchet derivative* of T at x is the operator $\mathsf{D}^2 T(x) \in \mathcal{B}\big(\mathcal{H}, \mathcal{B}(\mathcal{H}, \mathcal{K})\big)$ that satisfies

$$\lim_{0 \neq \|y\| \to 0} \frac{\|\mathsf{D}T(x+y) - \mathsf{D}Tx - \mathsf{D}^2 T(x)y\|}{\|y\|} = 0. \tag{2.43}$$

The *Fréchet gradient* of a function $f \colon C \to \mathbb{R}$ at $x \in C$ is defined as in Remark 2.55. Here are some examples.

Example 2.57 Let $L \in \mathcal{B}(\mathcal{H})$, let $u \in \mathcal{H}$, let $x \in \mathcal{H}$, and set $f \colon \mathcal{H} \to \mathbb{R} \colon y \mapsto \langle Ly \mid y\rangle - \langle y \mid u\rangle$. Then f is twice Fréchet differentiable on $\mathcal{H}$ with $\nabla f(x) = (L + L^*)x - u$ and $\nabla^2 f(x) = L + L^*$.

Proof. Take $y \in \mathcal{H}$. Since

$$\begin{aligned} f(x+y) - f(x) &= \langle Lx \mid y\rangle + \langle Ly \mid x\rangle + \langle Ly \mid y\rangle - \langle y \mid u\rangle \\ &= \langle y \mid (L + L^*)x\rangle - \langle y \mid u\rangle + \langle Ly \mid y\rangle , \end{aligned} \tag{2.44}$$

we have

$$\big| f(x+y) - f(x) - \langle y \mid (L+L^*)x - u\rangle \big| = |\langle Ly \mid y\rangle| \leqslant \|L\| \, \|y\|^2. \tag{2.45}$$

In view of (2.42), f is Fréchet differentiable at x with $\nabla f(x) = (L + L^*)x - u$. In turn, (2.43) yields $\nabla^2 f(x) = L + L^*$. ☐

Proposition 2.58 *Let $f \colon \mathcal{H} \to \mathbb{R}$ be Gâteaux differentiable, let $L \in \mathcal{B}(\mathcal{H})$, and suppose that $\nabla f = L$. Then $L = L^*$, $f \colon x \mapsto f(0) + (1/2)\langle Lx \mid x\rangle$, and f is twice Fréchet differentiable.*

Proof. Fix $x \in \mathcal{H}$ and set $\phi\colon \mathbb{R} \to \mathbb{R}\colon t \mapsto f(tx)$. Then $(\forall t \in \mathbb{R})$ $\phi'(t) = \langle x \mid \nabla f(tx)\rangle = \langle x \mid L(tx)\rangle = t\,\langle x \mid Lx\rangle$. It follows that $f(x) - f(0) = \phi(1) - \phi(0) = \int_0^1 \phi'(t)\,dt = \int_0^1 t\,\langle Lx \mid x\rangle\,dt = (1/2)\,\langle Lx \mid x\rangle$. We deduce from Example 2.57 that f is twice Fréchet differentiable and that $L = \nabla f = (L+L^*)/2$. Hence, $L^* = L$. ☐

Example 2.59 Let $F\colon \mathcal{H}\times\mathcal{H} \to \mathbb{R}$ be a symmetric bilinear form such that, for some $\beta \in \mathbb{R}_+$,

$$(\forall x \in \mathcal{H})(\forall y \in \mathcal{H}) \quad |F(x,y)| \leqslant \beta\|x\|\,\|y\|, \tag{2.46}$$

let $\ell \in \mathcal{B}(\mathcal{H},\mathbb{R})$, let $x \in \mathcal{H}$, and set $f\colon \mathcal{H} \to \mathbb{R}\colon y \mapsto (1/2)F(y,y) - \ell(y)$. Then f is Fréchet differentiable on $\mathcal{H}$ with $\mathsf{D}f(x) = F(x,\cdot) - \ell$.

Proof. Take $y \in \mathcal{H}$. Then,

$$\begin{aligned} f(x+y) - f(x) &= \tfrac{1}{2}F(x+y,x+y) - \ell(x+y) - \tfrac{1}{2}F(x,x) + \ell(x) \\ &= \tfrac{1}{2}F(y,y) + F(x,y) - \ell(y). \end{aligned} \tag{2.47}$$

Consequently, (2.46) yields

$$2\big|f(x+y) - f(x) - \big(F(x,y) - \ell(y)\big)\big| = |F(y,y)| \leqslant \beta\|y\|^2, \tag{2.48}$$

and we infer from (2.42) and (2.46) that $\mathsf{D}f(x)y = F(x,y) - \ell(y)$. ☐

Example 2.60 Let $\mathcal{K}$ be a real Hilbert space, let $L \in \mathcal{B}(\mathcal{H},\mathcal{K})$, let $r \in \mathcal{K}$, let $x \in \mathcal{H}$, and set $f\colon \mathcal{H} \to \mathbb{R}\colon y \mapsto \|Ly - r\|^2$. Then f is twice Fréchet differentiable on $\mathcal{H}$ with $\nabla f(x) = 2L^*(Lx - r)$ and $\nabla^2 f(x) = 2L^*L$.

Proof. Set $F\colon \mathcal{H}\times\mathcal{H} \to \mathbb{R}\colon (y,z) \mapsto (1/2)\,\langle L^*Ly \mid z\rangle$, $\ell\colon \mathcal{H} \to \mathbb{R}\colon y \mapsto \langle y \mid L^*r\rangle$, and $\alpha = (1/2)\|r\|^2$. Then $(\forall y \in \mathcal{H})$ $f(y) = 2(F(y,y) - \ell(y) + \alpha)$. Hence we derive from Example 2.59 that $\nabla f(x) = 2L^*(Lx - r)$, and from (2.43) that $\nabla^2 f(x) = 2L^*L$. ☐

Lemma 2.61 *Let $x \in \mathcal{H}$, let $C \in \mathcal{V}(x)$, and let $T\colon C \to \mathcal{K}$. Suppose that T is Fréchet differentiable at x. Then the following hold:*

(i) *T is Gâteaux differentiable at x and the two derivatives coincide.*
(ii) *T is continuous at x.*

Proof. Denote the Fréchet derivative of T at x by L_x.

(i): Let $\alpha \in \mathbb{R}_{++}$ and $y \in \mathcal{H} \smallsetminus \{0\}$. Then

$$\left\|\frac{T(x+\alpha y) - Tx}{\alpha} - L_x y\right\| = \|y\|\,\frac{\|T(x+\alpha y) - Tx - L_x(\alpha y)\|}{\|\alpha y\|} \tag{2.49}$$

converges to 0 as $\alpha \downarrow 0$.

(ii): Fix $\varepsilon \in \mathbb{R}_{++}$. By (2.42), we can find $\delta \in \left]0, \varepsilon/(\varepsilon + \|L_x\|)\right]$ such that $(\forall y \in B(0;\delta))\ \|T(x+y) - Tx - L_x y\| \leqslant \varepsilon\|y\|$. Thus $(\forall y \in B(0;\delta))\ \|T(x+y) - Tx\| \leqslant \|T(x+y) - Tx - L_x y\| + \|L_x y\| \leqslant \varepsilon\|y\| + \|L_x\|\,\|y\| \leqslant \delta(\varepsilon + \|L_x\|) \leqslant \varepsilon$. It follows that T is continuous at x. ☐

Fact 2.62 (See [146, Proposition 5.1.8]) Let $T\colon \mathcal{H} \to \mathcal{K}$ and let $x \in \mathcal{H}$. Suppose that the Gâteaux derivative of T exists on a neighborhood of x and that $\mathsf{D}T$ is continuous at x. Then T is Fréchet differentiable at x.

Fact 2.63 (See [146, Theorem 5.1.11]) Let $x \in \mathcal{H}$, let U be a neighborhood of x, let $\mathcal{G}$ be a real Banach space, let $T\colon U \to \mathcal{G}$, let V be a neighborhood of Tx, and let $R\colon V \to \mathcal{K}$. Suppose that R is Fréchet differentiable at x and that T is Gâteaux differentiable at Tx. Then $R \circ T$ is Gâteaux differentiable at x and $\mathsf{D}(R \circ T)(x) = (\mathsf{D}R(Tx)) \circ \mathsf{D}T(x)$. If T is Fréchet differentiable at x, then so is $R \circ T$.

Item (i) in the next result is known as the *descent lemma*.

Lemma 2.64 *Let U be a nonempty open convex subset of $\mathcal{H}$, let $\beta \in \mathbb{R}_{++}$, let $f\colon U \to \mathbb{R}$ be a Fréchet differentiable function such that ∇f is β-Lipschitz continuous on U, and let x and y be in U. Then the following hold:*

(i) $|f(y) - f(x) - \langle y - x \mid \nabla f(x)\rangle| \leqslant (\beta/2)\|y - x\|^2$.
(ii) $|\langle x - y \mid \nabla f(x) - \nabla f(y)\rangle| \leqslant \beta\|y - x\|^2$.

Proof. (i): Set $\phi\colon [0,1] \to \mathbb{R}\colon t \mapsto f(x + t(y - x))$. Then, by Cauchy–Schwarz,

$$\begin{aligned}
&\big|f(y) - f(x) - \langle y - x \mid \nabla f(x)\rangle\big| \\
&\qquad = \left|\int_0^1 \phi'(t)\,dt - \langle y - x \mid \nabla f(x)\rangle\right| \\
&\qquad = \left|\int_0^1 \langle y - x \mid \nabla f(x + t(y - x)) - \nabla f(x)\rangle\,dt\right| \\
&\qquad \leqslant \int_0^1 \|y - x\|\beta\|t(y - x)\|\,dt \\
&\qquad = \frac{\beta}{2}\|y - x\|^2, \qquad (2.50)
\end{aligned}$$

as claimed.

(ii): This follows from the Cauchy–Schwarz inequality. ☐

Example 2.65 Suppose that $\mathcal{H} \neq \{0\}$ and let $f\colon \mathcal{H} \to \mathbb{R}\colon x \mapsto \|x\|$. Then $f = \sqrt{\|\cdot\|^2}$ and, since Example 2.60 asserts that $\|\cdot\|^2$ is Fréchet differentiable with gradient operator $\nabla\|\cdot\|^2 = 2\mathrm{Id}$, it follows from Fact 2.63 that f is Fréchet differentiable on $\mathcal{H} \smallsetminus \{0\}$ with $(\forall x \in \mathcal{H} \smallsetminus \{0\})\ \nabla f(x) = x/\|x\|$. On the other hand, f is not Gâteaux differentiable at $x = 0$ since, although the limit in (2.37) exists, it is not linear with respect to y: $(\forall y \in \mathcal{H})\ \lim_{\alpha\downarrow 0} (\|0 + \alpha y\| - \|0\|)/\alpha = \|y\|$.

Fact 2.66 (See [146, Proposition 5.1.22]) Let $x \in \mathcal{H}$, let U be a neighborhood of x, let $\mathcal{K}$ be a real Banach space, and let $T\colon U \to \mathcal{K}$. Suppose that T is twice Fréchet differentiable at x. Then $(\forall (y,z) \in \mathcal{H}\times\mathcal{H})$ $(\mathsf{D}^2T(x)y)z = (\mathsf{D}^2T(x)z)y$.

Example 2.67 Let $x \in \mathcal{H}$, let U be a neighborhood of x, and let $f\colon U \to \mathbb{R}$. Suppose that f is twice Fréchet differentiable at x. Then, in view of Fact 2.66 and (2.41), $\nabla^2 f(x)$ is self-adjoint.

Exercises

Exercise 2.1 Let x and y be points in $\mathcal{H}$. Show that the following are equivalent:

(i) $\|y\|^2 + \|x-y\|^2 = \|x\|^2$.
(ii) $\|y\|^2 = \langle x \mid y\rangle$.
(iii) $\langle y \mid x-y\rangle = 0$.
(iv) $(\forall \alpha \in [-1,1])$ $\|y\| \leqslant \|\alpha x + (1-\alpha)y\|$.
(v) $(\forall \alpha \in \mathbb{R})$ $\|y\| \leqslant \|\alpha x + (1-\alpha)y\|$.
(vi) $\|2y - x\| = \|x\|$.

Exercise 2.2 Consider $\mathcal{X} = \mathbb{R}^2$ with the norms $\|\cdot\|_1\colon \mathcal{X} \to \mathbb{R}_+\colon (\xi_1,\xi_2) \mapsto |\xi_1| + |\xi_2|$ and $\|\cdot\|_\infty\colon \mathcal{X} \to \mathbb{R}_+\colon (\xi_1,\xi_2) \mapsto \max\{|\xi_1|,|\xi_2|\}$. Show that neither norm satisfies the parallelogram identity.

Exercise 2.3 Let x and y be points in $\mathcal{H}$, and let α and β be real numbers. Show that

$$\|\alpha x + \beta y\|^2 + \alpha\beta\|x-y\|^2 = \alpha(\alpha+\beta)\|x\|^2 + \beta(\alpha+\beta)\|y\|^2. \tag{2.51}$$

Exercise 2.4 Set

$$\varDelta\colon \mathcal{H}\times\mathcal{H} \to \mathbb{R}_+\colon (x,y) \mapsto \left\| \frac{x}{1+\|x\|} - \frac{y}{1+\|y\|} \right\|. \tag{2.52}$$

Show that $(\mathcal{H}, \varDelta)$ is a metric space.

Exercise 2.5 Define $\varDelta$ as in Exercise 2.4, let $(x_n)_{n\in\mathbb{N}}$ be a bounded sequence in $\mathcal{H}$, and let $x \in \mathcal{H}$. Show that $x_n \to x$ if and only if $\varDelta(x_n, x) \to 0$.

Exercise 2.6 Define $\varDelta$ as in Exercise 2.4, let $(x_n)_{n\in\mathbb{N}}$ be a bounded sequence in $\mathcal{H}$, and let $x \in \mathcal{H}$. Show that $x_n \to x$ if and only if $\varDelta(x_n, x) \to 0$.

Exercise 2.7 Construct a monotone operator $T \in \mathcal{B}(\mathcal{H})$ such that T is not self-adjoint.

Exercise 2.8 Suppose that $\mathcal{H} \neq \{0\}$ and define on $\mathcal{H} \smallsetminus \{0\}$ a relation by $x \equiv y \Leftrightarrow x \in \mathbb{R}_{++}y$. Show that $\equiv$ is an equivalence relation. For every $x \in \mathcal{H} \smallsetminus \{0\}$, let $[x] = \big\{y \in \mathcal{H} \smallsetminus \{0\} \bigm| x \equiv y\big\}$ be the corresponding equivalence class. The quotient set $\operatorname{hzn} \mathcal{H} = \big\{[x] \bigm| x \in \mathcal{H} \smallsetminus \{0\}\big\}$ is the *horizon* of $\mathcal{H}$ and $\operatorname{csm} \mathcal{H} = \mathcal{H} \cup \operatorname{hzn} \mathcal{H}$ is the *cosmic closure* of $\mathcal{H}$. Show that the function Δ of Exercise 2.4 extends to a distance on $\operatorname{csm} \mathcal{H}$ by defining

$$(\forall [x] \in \operatorname{hzn} \mathcal{H})(\forall y \in \mathcal{H}) \quad \Delta([x], y) = \Delta(y, [x]) = \left\| \frac{x}{\|x\|} - \frac{y}{1 + \|y\|} \right\| \tag{2.53}$$

and

$$\begin{aligned}(\forall [x] \in \operatorname{hzn} \mathcal{H})(\forall [y] \in \operatorname{hzn} \mathcal{H}) \quad \Delta([x], [y]) &= \Delta([y], [x]) \\ &= \left\| \frac{x}{\|x\|} - \frac{y}{\|y\|} \right\|. \end{aligned} \tag{2.54}$$

Exercise 2.9 Consider Exercise 2.8 and its notation. Let $(x_n)_{n \in \mathbb{N}}$ be a sequence in $\mathcal{H}$ and let $[x] \in \operatorname{hzn} \mathcal{H}$. Show that $\Delta(x_n, [x]) \to 0$ if and only if there exists a sequence $(\lambda_n)_{n \in \mathbb{N}}$ in $\mathbb{R}_{++}$ such that $\lambda_n \to 0$ and $\lambda_n x_n \to x$.

Exercise 2.10 Consider Exercise 2.8 and its notation. Let $([x_n])_{n \in \mathbb{N}}$ be a sequence in $\operatorname{hzn} \mathcal{H}$ and let $[x] \in \operatorname{hzn} \mathcal{H}$. Show that $\Delta([x_n], [x]) \to 0$ if and only if there exists a sequence $(\lambda_n)_{n \in \mathbb{N}}$ in $\mathbb{R}_{++}$ such that $\lambda_n x_n \to x$.

Exercise 2.11 Suppose that $\mathcal{H}$ is finite-dimensional and consider Exercise 2.8 and its notation. Show that $\operatorname{csm} \mathcal{H}$ is compact and sequentially compact with respect to the distance Δ.

Exercise 2.12 Let N be a strictly positive integer, set $I = \{1, \ldots, N\}$, and suppose that $(x_i)_{i \in I}$ are points in $\mathcal{H}$ such that $(\forall i \in I)\ \|x_i\| = 1$. Show the following:

(i) $\| \sum_{i \in I} x_i \|^2 = N + 2 \sum_{1 \leqslant i < j \leqslant N} \langle x_i \mid x_j \rangle$.
(ii) Suppose that, for every $(i, j) \in I \times I$ such that $i \neq j$ we have $\langle x_i \mid x_j \rangle = -1/(N-1)$. Then $\sum_{i \in I} x_i = 0$.
(iii) Suppose that $\sum_{i \in I} x_i = 0$. Then $2 \sum_{1 \leqslant i < j \leqslant N-1} \langle x_i \mid x_j \rangle = 2 - N$.
(iv) Suppose $N = 3$. Then $x_1 + x_2 + x_3 = 0$ if and only if $\langle x_1 \mid x_2 \rangle = \langle x_1 \mid x_3 \rangle = \langle x_2 \mid x_3 \rangle = -1/2$.

Exercise 2.13 Let x and y be points in $\mathcal{H}$, and let α and β be real numbers in $\mathbb{R}_+$. Show that $4 \langle \alpha x - \beta y \mid y - x \rangle \leqslant \alpha \|y\|^2 + \beta \|x\|^2$.

Exercise 2.14 Let x and y be in $\mathcal{H}$, and let α and β be in $\mathbb{R}$. Show that

$$\begin{aligned}\alpha(1-\alpha)\|\beta x + (1-\beta)y\|^2 + \beta(1-\beta)\|\alpha x - (1-\alpha)y\|^2 \\ = (\alpha + \beta - 2\alpha\beta)\big(\alpha\beta\|x\|^2 + (1-\alpha)(1-\beta)\|y\|^2\big). \end{aligned} \tag{2.55}$$

Exercise 2.15 Let x, y, and z be points in $\mathcal{H}$ such that $\|2x - y - z\| = \|2y - x - z\| = \|2z - x - y\|$. Show that $\|x - y\| = \|y - z\| = \|z - x\|$.

Exercise 2.16 Suppose that $\mathcal{H}$ is infinite-dimensional. Show that every weakly compact set has an empty weak interior.

Exercise 2.17 Provide an unbounded convergent net in $\mathbb{R}$ and compare with Lemma 2.46.

Exercise 2.18 Construct a sequence in $\mathcal{H}$ that converges weakly and possesses a strong sequential cluster point, but that does not converge strongly.

Exercise 2.19 Let C be a subset of $\mathcal{H}$ such that $(\forall n \in \mathbb{N})$ $C \cap B(0; n)$ is weakly sequentially closed. Show that C is weakly sequentially closed and compare with Lemma 1.40.

Exercise 2.20 Show that the conclusion of Lemma 2.51(iii) fails if the strong convergence of $(u_n)_{n\in\mathbb{N}}$ is replaced by weak convergence.

Exercise 2.21 (Opial's condition) Let $(x_n)_{n\in\mathbb{N}}$ be a weakly convergent sequence in $\mathcal{H}$ and let $x \in \mathcal{H}$. Show that $x_n \rightharpoonup x$ if and only if

$$(\forall y \in \mathcal{H}) \quad \underline{\lim}\, \|x_n - y\|^2 = \|x - y\|^2 + \underline{\lim}\, \|x_n - x\|^2. \tag{2.56}$$

In particular, if $x_n \rightharpoonup x$ and $y \in \mathcal{H} \smallsetminus \{x\}$, then $\underline{\lim}\, \|x_n - y\| > \underline{\lim}\, \|x_n - x\|$. This implication is known as *Opial's condition*.

Exercise 2.22 Suppose that $\mathcal{H}$ is infinite-dimensional and let $(e_n)_{n\in\mathbb{N}}$ be an orthonormal sequence in $\mathcal{H}$. Construct a bounded sequence $(x_n)_{n\in\mathbb{N}}$ in $\mathcal{H}$ such that $x_n - x_{n+1} \to 0$ and the set of strong cluster points of $(x_n)_{n\in\mathbb{N}}$ is $\{e_0, -e_0\}$. Compare to the Ostrowski results (Theorem 1.49 and Lemma 2.53).

Exercise 2.23 Show that if the derivative $\mathsf{D}T(x)$ exists in Definition 2.54, then it is unique.

Exercise 2.24 Let D be a nonempty open interval in $\mathbb{R}$, let $f \colon D \to \mathbb{R}$, and let $x \in D$. Show that the notions of Gâteaux and Fréchet differentiability of f at x coincide with classical differentiability, and that the Gâteaux and Fréchet derivatives coincide with the classical derivative

$$f'(x) = \lim_{0 \neq h \to 0} \frac{f(x+h) - f(x)}{h}. \tag{2.57}$$

Exercise 2.25 Consider the function

$$f \colon \mathbb{R}^2 \to \mathbb{R} \colon (\xi_1, \xi_2) \mapsto \begin{cases} \dfrac{\xi_1^2 \xi_2^4}{\xi_1^4 + \xi_2^8}, & \text{if } (\xi_1, \xi_2) \neq (0,0); \\ 0, & \text{if } (\xi_1, \xi_2) = (0,0). \end{cases} \tag{2.58}$$

Show that f is Gâteaux differentiable, but not continuous, at $(0, 0)$.

Exercise 2.26 Consider the function

$$f\colon \mathbb{R}^2 \to \mathbb{R}\colon x = (\xi_1, \xi_2) \mapsto \begin{cases} \dfrac{\xi_1 \xi_2^4}{\xi_1^2 + \xi_2^4}, & \text{if } (\xi_1, \xi_2) \neq (0,0); \\ 0, & \text{if } (\xi_1, \xi_2) = (0,0). \end{cases} \tag{2.59}$$

Show that f is Fréchet differentiable at $(0,0)$ and that ∇f is not continuous at $(0,0)$. Conclude that the converse of Fact 2.62 does not hold.

Exercise 2.27 Consider the function

$$f\colon \mathbb{R}^2 \to \mathbb{R}\colon (\xi_1, \xi_2) \mapsto \begin{cases} \dfrac{\xi_1 \xi_2^3}{\xi_1^2 + \xi_2^4}, & \text{if } (\xi_1, \xi_2) \neq (0,0); \\ 0, & \text{if } (\xi_1, \xi_2) = (0,0). \end{cases} \tag{2.60}$$

Show that, at $(0,0)$, f is continuous and Gâteaux differentiable, but not Fréchet differentiable.

Exercise 2.28 Consider the function

$$f\colon \mathbb{R}^2 \to \mathbb{R}\colon (\xi_1, \xi_2) \mapsto \begin{cases} \dfrac{\xi_1 \xi_2^2}{\xi_1^2 + \xi_2^2}, & \text{if } (\xi_1, \xi_2) \neq (0,0); \\ 0, & \text{if } (\xi_1, \xi_2) = (0,0). \end{cases} \tag{2.61}$$

Show that f is continuous and that, at $(0,0)$, the limit on the right-hand side of (2.37) exists but it is not linear as a function of (η_1, η_2). Conclude that f is not Gâteaux differentiable at $(0,0)$.

http://www.springer.com/978-3-319-48310-8

Convex Analysis and Monotone Operator Theory in
Hilbert Spaces
Bauschke, H.H.; Combettes, P.L.
2017, XIX, 619 p., Hardcover
ISBN: 978-3-319-48310-8