Chapter 4

Properties of the metric projection

In the focus of this chapter are continuity and differentiability properties of the metric projections. We mainly restrict us to the case of a projection onto Chebyshev sets. First of all, we study geometric properties of Banach spaces which are helpful to analyze metric projections. In the focus of the geometric view on Banach spaces is the property of smoothness.

For the material on geometric Banach space theory we refer to Borwein and Lewis [7], Chidume [10], Cioranescu [11], Deimling [14], Kabalo [24], Hiriart-Urruty and Lemarechal [22, 21], Megginson [30] and Werner [41].

4.1 Uniform convexity

Next we consider a class of Banach spaces which share with Hilbert spaces the following property: their unit ball $\overline{B}_1$ is „rotund“ and „touches“ convex sets just in one point. We may say that the following considerations are devoted to the geometric properties of Banach spaces.

Definition 4.1. *A Banach space $\mathcal{X}$ is* **uniformly convex** *(or* **strongly rotund***) if and only if its* **modulus of convexity** $\delta_{\mathcal{X}}$ *defined by*

$$\delta_{\mathcal{X}}(\varepsilon) := \inf\{1 - \frac{1}{2}\|x+y\| : \|x\| = \|y\| = 1, \|x-y\| = \varepsilon\}\,,\ \varepsilon \in [0,2],$$

is positive for all $\varepsilon \in (0,2]$. □

Clearly, a uniformly convex space is strictly convex. Let us collect some properties of the modulus $\delta_{\mathcal{X}}$.

Lemma 4.2. *Let $\mathcal{X}$ be a Banach space with modulus of convexity $\delta_{\mathcal{X}}$. Then we have:*

(1) $\delta_{\mathcal{X}}(0) = 0, \delta_{\mathcal{X}}(\varepsilon) \leq \frac{1}{2}\varepsilon, \varepsilon \geq 0$.

(2) $(0,2] \ni \varepsilon \longmapsto \delta_{\mathcal{X}}(\varepsilon)\varepsilon^{-1} \in \mathbb{R}$ *is non-decreasing.*

(3) If $\mathcal{X}$ is a Hilbert space $\mathcal{H}$ then $\delta_{\mathcal{H}}(\varepsilon) = 1 - \sqrt{1 - \frac{1}{4}\varepsilon^2}, \varepsilon \in (0,2]$.

Proof:
Ad (1) Evidentely, $\delta_X(0) = 0$. Let $\varepsilon \in [0,2]$ and let $x, y \in \overline{S}_1$ with $\|x-y\| = \varepsilon$. Then

$$\frac{1}{2}\|x+y\| = \|x + \frac{1}{2}(y-x)\| \geq \|x\| - \frac{1}{2}\|y-x\| \geq 1 - \frac{1}{2}\varepsilon\,.$$

Ad (2) Let $0 < \eta \leq \varepsilon \leq 2$. Choose $x, y \in \overline{S}_1$ with $\|x-y\| = \varepsilon$. Set

$$u := \frac{\eta}{\varepsilon}x + (1-\frac{\eta}{\varepsilon})\frac{x+y}{\|x+y\|}, w := \frac{\eta}{\varepsilon}y + (1-\frac{\eta}{\varepsilon})\frac{x+y}{\|x+y\|}\,.$$

Then

$$\|u-w\| = \frac{\eta}{\varepsilon}\|x-y\| = \eta\,,\ \frac{1}{2}(u+w) = \frac{1}{2}\frac{\eta}{\varepsilon}(x+y) + \frac{1}{2}(1-\frac{\eta}{\varepsilon})\frac{x+y}{\|x+y\|}$$

and

$$\|\frac{x+y}{\|x+y\|} - \frac{1}{2}(u+w)\| = \frac{\eta}{\varepsilon} - \frac{1}{2}\frac{\eta}{\varepsilon}\|x+y\| = 1 - (1 - \frac{\eta}{\varepsilon} + \frac{1}{2}\frac{\eta}{\varepsilon}\|x+y\|) = 1 - \frac{1}{2}\|u+v\|\,.$$

Notice that

$$\|\frac{x+y}{\|x+y\|} - \frac{1}{2}(x+y)\| = 1 - \frac{1}{2}\|x+y\|\,.$$

Therefore

$$\begin{aligned}\frac{\|\frac{x+y}{\|x+y\|} - \frac{1}{2}(u+v)\|}{\|u-v\|} &= (\frac{\eta}{\varepsilon} - \frac{1}{2}\frac{\eta}{\varepsilon}\|x+y\|)/\eta \\ &= (1 - \frac{1}{2}\|x+y\|)/\varepsilon \;=\; \frac{\|\frac{x+y}{\|x+y\|} - \frac{1}{2}(x+y)\|}{\|x-y\|}\,.\end{aligned}$$

Now, we obtain

$$\begin{aligned}\frac{\delta_{\mathcal{X}}(\eta)}{\eta} &\leq \frac{1 - \frac{1}{2}\|u+v\|}{\|u-v\|} = \frac{\|\frac{x+y}{\|x+y\|} - \frac{1}{2}(u+v)\|}{\|u-v\|} \\ &= \frac{\|\frac{x+y}{\|x+y\|} - \frac{1}{2}(x+y)\|}{\|x-y\|} = \frac{1 - \frac{1}{2}\|x+y\|}{\|x-y\|} = \frac{1 - \frac{1}{2}\|x+y\|}{\varepsilon}\end{aligned}$$

Taking the infimum with respect to the choice of x, y we obtain the assertion.
Ad (3) Follows from the parallelogram identity. ■

Remark 4.3. *Let $\delta_{\mathcal{X}}$ be the modulus of convexity in the Banach space $\mathcal{X}$. There exists a constant $0 < L < 3.18$ such that if $0 < \eta \leq \varepsilon \leq 2$ then we have (see [16])*

$$\frac{\delta_{\mathcal{X}}(\eta)}{\eta^2} \leq 4L\frac{\delta_{\mathcal{X}}(\varepsilon)}{\varepsilon^2}$$

□

Theorem 4.4. *A uniformly convex Banach $\mathcal{X}$ space is an E-space.*

Proof:
We know already that $\mathcal{X}$ is strictly convex. To show the reflexivity of $\mathcal{X}$ we may use the criterion of James (see the preliminaries in the preface). To apply this criterion we have to show that for each functional $\lambda \in \mathcal{X}^*$ with $\|\lambda\| = 1$ there exists $x \in \overline{S}_1$ with $\langle \lambda, x \rangle = 1$. Let $\lambda \in \mathcal{X}^*$ with $\|\lambda\| = 1$. Then there exists a sequence $(x^n)_{n\in\mathbb{N}}$ in $\overline{S}_1$ with $\lim_n \langle \lambda, x^n \rangle = 1$. This implies $\lim_n \|x^n\| = 1$ and from $\lim_{m,n} \langle \lambda, x^n + x^m \rangle = 1$ we obtain

$$\lim_{n,m} \|\frac{1}{2}(x^n + x^m)\| = 1 .$$

Due to the uniform convexity, the sequence $(x^n)_{n\in\mathbb{N}}$ converges to some $x \in \overline{S}_1$. Clearly, $\langle \lambda, x \rangle = 1 = \|\lambda\|$.
To prove the third property of an E-space, let $x \in \overline{S}_1$ and let x be the weak limit of the sequence $(x^n)_{n\in\mathbb{N}}$ in $\overline{S}_1$. Assume that x is not the strong limit of this sequence. Then there exists a subsequence $(x^{n_k})_{k\in\mathbb{N}}$ and $\varepsilon > 0$ such that $\|x^{n_k} - x\| \geq \varepsilon$, $k \in \mathbb{N}$. By the uniform convexity there exists a $d > 0$ such that

$$\|\frac{1}{2}(x^{n_k} + x)\| \leq 1 - d, k \in \mathbb{N} .$$

Let $\lambda \in \Sigma(x)$. Then

$$1 = \langle \lambda, x \rangle = \lim_k \langle \lambda, \frac{1}{2}(x^{n_k} + x) \rangle \leq \|\lambda\|(1 - d) < 1$$

This is a contradiction. ■

Theorem 4.5. *Let $\mathcal{X}$ be a uniformly convex Banach space. Then every nonempty convex closed subset C of $\mathcal{X}$ is a Chebyshev set. Moreover, the best approximation problem* (2.1) *is strongly solvable.*

Proof:
This follows from Theorem 4.4 and Theorem 2.24. ■

Remark 4.6. *Notice that the property of uniform convexity is not stable under equivalent renorming the space.* □

Theorem 4.7 (Milman)**.** *A uniformly convex Banach space is reflexive.*

Proof:
Follows from Theorem 4.4. ■

For concrete examples of uniformly convex spaces one can show the reflexivity ad hoc. For instance, the spaces $l_p, 1 < p < \infty$, are uniformly convex and reflexive. The modulus of convexity is given by

$$\delta(\varepsilon) := 1 - \sqrt[q]{1 - (\frac{1}{2}\varepsilon)^q}, \; 0 < \varepsilon \leq 2,$$

where $\frac{1}{p} + \frac{1}{q} = 1$. To compute this result one uses a Hölder-type inequality. The reflexivity follows from the fact that the dual space of l_p is given by l_q where q is as above. Then we conclude that $l_p^{**} = l_q^* = l_p$. Obviously, the spaces $l_p, p = 1, \infty$ are whether uniformly convex nor reflexive.

Remark 4.8. *The class of reflexive Banach spaces is larger than the class of uniformly convex Banach spaces. This shows the famous result of M. Day (see [13]): There exists a separable reflexive strictly convex Banach space which is not isomorphic to a uniformly convex Banach space.*

On the other hand, a separable Banach space $\mathcal{X}$ can be renormed in such a way that it is strictly convex. This follows from the fact that there exists an injective linear mapping $T : \mathcal{X} \longrightarrow l_2$ which leads to the new norm $\|\cdot\|' : \mathcal{X} \ni x \longmapsto \|x\| + \|Tx\| \in \mathbb{R}$. Now, $(\mathcal{X}, \|\cdot\|')$ is strictly convex; see Koethe [26]. □

Remark 4.9. *As we know from Theorem 4.7, each uniformly convex Banach space is reflexive. Obviously, the converse does not hold; consider for example the two-dimensional space $\mathbb{R}^2$ endowed with the l_1 norm. A more sharp result concerning this question is the classical result of M. Day ([13]): There exist Banach spaces which are separable, reflexive, and strictly convex, but are not isomorphic to any uniformly convex space. This result has lead to the class of Banach spaces which are called* **superreflexive**. *Such spaces are characterized by the property that they admit an equivalent uniformly convex norm. We omit the question how a reflexive space which is not superreflexive looks like.* □

Here is a result which connects uniform convexity to the quantitative behavior of the duality map.

Theorem 4.10 (Prüss, 1981)**.** *Let $\mathcal{X}$ be a Banach space. Then $\mathcal{X}$ is uniformly convex if and only if for each $R > 0$ there exists a gauge-function γ_R such that*

$$\langle J_{\mathcal{X}}(x) - J_{\mathcal{X}}(y), x - y\rangle \geq \gamma_R(\|x - y\|)\|x - y\| \text{ for all } x, y \in \overline{B}_R\,. \tag{4.1}$$

Proof:
We follow [35].
Let $\mathcal{X}$ be uniformly convex and choose $R > 0$. Let γ_R be defined by $\gamma_R(0) = 0$, and

$$\gamma_R(r) := \inf\{\langle\lambda-\mu, x-y\rangle\|x-y\|^{-1} : x, y \in \overline{B}_R, \|x-y\| \geq r, \lambda \in J_{\mathcal{X}}(x), \mu \in J_{\mathcal{X}}(y)\}, r \in (0, 2R]\,,$$

We extend γ_R by

$$\gamma_R(r) := \gamma_R(2R)\,,\ r \geq 2R\,.$$

Obviously, γ_R is nondecreasing. Let us prove that $\gamma_R(r) > 0$ for $r > 0$.
Assume on the contrary that $\gamma_R(r) = 0$ for some $r \in (0, 2R]$. Then there exist sequences $(x^n)_{n\in\mathbb{N}}, (y^n)_{n\in\mathbb{N}}$ belonging to $\overline{B}_1$ and $\lambda_n \in J_{\mathcal{X}}(x^n), \mu_n \in J_{\mathcal{X}}(y^n)$ such that $\|x^n - y^n\| \geq r$ and $\lim_n\langle\lambda_n - \mu_n, x^n - y^n\rangle\|x^n - y^n\|^{-1} = 0$. Since we know (see (3.4))

$$(\|x^n\| - y^n\|)^2 \leq \langle\lambda_n - \mu_n, x^n - y^n\rangle$$

we may assume $\lim_n \|x^n\| = \lim_n \|y\| = a > 0$. Then

$$\liminf_n |\|x^n\|^{-1}x^n - \|y^n\|^{-1}y^n| = a^{-1}\liminf_n \|x^n - y^n\| \geq ra^{-1}\,.$$

Therefore, by the uniform convexity of $\mathcal{X}$,

$$\limsup_n \|x^n + y^n\| = a\limsup_n |\|x^n\|^{-1}x^n + \|y^n\|y^n| \leq 2a(1 - \delta)$$

for some $0 < \delta = \delta(r)$. On the other hand,

$$\|x^n\|^2 + \|y^n\|^2 - \langle \mu_n, x^n \rangle - \langle \lambda_n, y^n \rangle = \langle \lambda_n - \mu_n, x^n - y^n \rangle .$$

This implies $\lim_n \langle \lambda_n, x^n + y^n \rangle = 2a^2$ and leads to the contradiction

$$2a^2 = \lim_n \langle \lambda_n, x^n + y^n \rangle \leq a \limsup_n \|x^n + y^n\| \leq 2a^2(1-\delta) .$$

Let us prove the sufficiency. We first show that the range of J is closed. Let $(\lambda_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{X}^*$ with $\lambda_n \in J_{\mathcal{X}}(x^n), n \in \mathbb{N}$, and $\lim_n \lambda_n = \lambda \in \mathcal{X}^*$. Then $(x^n)_{n\in\mathbb{N}}$ is bounded, say $\|x^n\| \leq R, n \in \mathbb{N}$, due to the fact that $(\lambda_n)_{n\in\mathbb{N}}$ is bounded. Then we find γ_R with

$$\gamma_R(\|x^n - x^m\|)\|x^n - y^m\| \leq \langle \lambda_n - \lambda_m, x^n - x^m \rangle \leq \|x^n - x^m\|\|\lambda_n - \lambda_m\| , \; n, m \in \mathbb{N} .$$

Hence, $(x^n)_{n\in\mathbb{N}}$ is a Cauchy sequence and therefore convergent. Let $\lim_n x^n = z$. Then $\lambda \in J_{\mathcal{X}}(z)$ and J has closed range.
Therefore J is surjective since $\operatorname{ran}(J)$ is dense in $\mathcal{X}^*$ and $\mathcal{X}$ is reflexive. This follows from the Bishop-Phelps-theorem; see [6]. Now, we have $J^* = J^{-1}$. Due to

$$\gamma_R(\|x - y\|) \leq \|\lambda - \mu\| \text{ for } \lambda, \mu \in \mathcal{X}^*, \|\lambda\| \leq R, \|\mu\| \leq R, x \in J^*(\lambda), y \in J^*(\mu) ,$$

hence J^* is single-valued and uniformly continuous on bounded subsets of $\mathcal{X}^*$. Therefore $\mathcal{X} = (\mathcal{X}^*)^*$ is uniformly convex and the proof is complete. ■

4.2 Uniform smoothness

Definition 4.11. *A Banach space $\mathcal{X}$ is called* **uniformly smooth** *if the the following property holds:*

$$\forall \varepsilon \, \exists \delta > 0 \, \forall x, y \in \mathcal{X} (\|x\| = 1, \|y\| \leq \delta \implies \|x + y\| + \|x - y\| - 2 < \varepsilon \|y\|)$$

□

To make the definition more applicable let us introduce a modulus which can be used to characterize uniformly smooth spaces.

Definition 4.12. *Let $\mathcal{X}$ be a Banach space. The* **modulus of smoothness** *of $\mathcal{X}$ is the function $\rho_{\mathcal{X}} : [0, \infty) \longrightarrow [0, \infty)$ defined by*

$$\begin{aligned} \rho_{\mathcal{X}}(\tau) &:= \sup\{\frac{1}{2}(\|x + y\| + \frac{1}{2}\|x - y\|) - 1 : \|x\| = 1, \|y\| = \tau\} \\ &= \sup\{\frac{1}{2}(\|x + \tau y\| + \frac{1}{2}\|x - \tau y\|) - 1 : \|x\| = 1 = \|y\|\} \end{aligned}$$

□

Theorem 4.13. *Let $\mathcal{X}$ be a Banach space. Then the following conditions are equivalent:*

(a) $\mathcal{X}$ is uniformly convex.

(b) $\lim_{\tau\downarrow 0}\frac{\rho_{\mathcal{X}}(\tau)}{\tau}=0$.

Proof:
(a) $\Longrightarrow$ (b) Let $\varepsilon>0$. Then there exists $\delta>0$ such that

$$\frac{1}{2}(\|x+y\|+\|x-y\|)-1<\frac{1}{2}\varepsilon\|y\|$$

for all $x,y\in\mathcal{X}$ with $\|x\|=1,\|y\|=\delta$. Then $\rho_{\mathcal{X}}(\tau)<\frac{1}{2}\varepsilon\tau$ for all $\tau\in(0,\delta)$.
(b) $\Longrightarrow$ (a) Let $\varepsilon>0$. Then there exists $\delta>0$ such that $\rho_{\mathcal{X}}(\tau)<\frac{1}{2}\varepsilon\tau$ for all $\tau\in(0,\delta]$. Let $x,y\in\mathcal{X}$ with $\|x\|=1,\|y\|\leq\delta$. Set $\tau:=\|y\|$. Then

$$\|x+y\|+\|x-y\|-2<\varepsilon\tau=\varepsilon\|y\|.$$

■

Lemma 4.14. *Let $\mathcal{X}$ be a Banach space with modulus of smoothness $\rho_{\mathcal{X}}$. Then we have:*

(1) $\rho_{\mathcal{X}}(0)=0$.

(2) $\max\{0,\tau-1\}\leq\rho_{\mathcal{X}}(\tau)\leq\tau$ *for all* $\tau\in[0,\infty)$.

(3) $(0,\infty)\ni\tau\longmapsto\rho_{\mathcal{X}}(\tau)\tau^{-1}\in(0,\infty)$ *is non-decreasing.*

(4) $\rho_{\mathcal{X}}$ *is a continuous convex function from* $(0,\infty)$ *to* $(0,\infty)$.

Proof:
Ad (1) Obviuosly true.
Ad (2) Let $x,y\in\overline{S}_1$. For $\tau>0$ let $x\in\overline{S}_1$ and set $y:=\tau x$. Then $\frac{1}{2}(\|x+y\|+\|x-y\|)=\tau$. This shows $\rho_{\mathcal{X}}(\tau)\leq\tau$.
Let $\tau>0$. Since $\rho_{\mathcal{X}}(\tau)\leq\tau$ we have to show $\rho_{\mathcal{X}}(\tau)\geq\tau-1$. Let $x,y\in\overline{S}_1$. Then

$$\frac{1}{2}(\|x+\tau y\|+\|x-\tau y\|)-1\geq\frac{1}{2}\|2\tau y\|-1=\tau-1.$$

Ad (3) Let $x,y\in\overline{S}_1$. Let $0<s\leq t$. Applying Lemma 3.16 we obtain

$$\begin{aligned}\frac{\frac{1}{2}(\|x+sy\|+\|x-sy\|)-1}{s} &= \frac{1}{2}\Big(\frac{\|x+sy\|-\|x\|}{s}-\frac{\|x-sy\|-\|x\|}{-s}\Big)\\ &\leq \frac{1}{2}\Big(\frac{\|x+ty\|-\|x\|}{t}-\frac{\|x-ty\|-\|x\|}{-t}\Big)\\ &= \frac{\frac{1}{2}(\|x+ty\|+\|x-ty\|)-1}{t}\end{aligned}$$

Therefore

$$\frac{\rho_{\mathcal{X}}(s)}{s}\leq\frac{\rho_{\mathcal{X}}(t)}{t}.$$

Ad (4) It is sufficient to show that $\rho_{\mathcal{X}}$ is convex since a convex function is continuous on the interior of its domain of definition.
Let $x,y\in X$ with $\|x\|=\|y\|=1$. Consider

$$f_{x,y}(t):=\frac{\|x+ty\|+\|x-ty\|}{2}-1.$$

Then we obtain for $a \in [0,1]$

$$\begin{aligned} f_{x,y}(at+(1-a)s) &= \frac{\|x+(at+(1-a)s)y\| + \|x-(at+(1-a)s)y\|}{2} - 1 \\ &\leq \frac{a\|x+ty\| + (1-a)\|x+sy\| + a\|x-ty\| + (1-a)\|x-sy\|}{2} - 1 \\ &= af_{x,y}(t) + (1-a)f_{x,y}(s) \end{aligned}$$

Therefore $f_{x,y}$ is convex for all x, y.
Let $\varepsilon > 0$. Then there exist $x, y \in X$ with

$$\rho_{\mathcal{X}}(at+(1-a)s)-\varepsilon \leq f_{x,y}(at+(1-a)s) \leq af_{x,y}(t)+(1-a)f_{x,y}(s) \leq a\rho_{\mathcal{X}}(t)+(1-a)\rho_{\mathcal{X}}(s).$$

Since $\varepsilon > 0$ was arbitrary chosen, $\rho_{\mathcal{X}}$ is convex.

■

Theorem 4.15. *A uniformly smooth Banach space is smooth.*

Proof:
Assume that $\mathcal{X}$ is not smooth. Then there exist $x^0 \in \mathcal{X}$ and $\lambda, \mu \in X^*$, such that $\lambda \neq \mu, \|\lambda\| = \|\mu\| = 1$ and $\langle \lambda, x^0 \rangle = \|x^0\| = \langle \mu, x^0 \rangle$. We may assume $\|x^0\| = 1$. Let $y^0 \in X$ with $\|y^0\| = 1$ and $\langle \lambda - \mu, y^0 \rangle > 0$. We obtain for $t > 0$

$$0 < t\langle \lambda-\mu, y^0 \rangle = \frac{1}{2}(\langle \lambda, x^0+ty^0 \rangle + \langle \mu, x^0-ty^0 \rangle) - 1 \leq \frac{1}{2}(\|x^0+ty^0\| + \|x^0-ty^0\|) - 1. \quad (4.2)$$

Hence $0 < \langle \lambda - \mu, y^0 \rangle \leq \rho_{\mathcal{X}}(t)/t$. Since $t > 0$ was arbitrary chosen we conclude that $\mathcal{X}$ is not uniformly smooth. ■

Theorem 4.16 (Lindenstrauss-Tzafiri, 1978)**.** *Let $\mathcal{X}$ be a Banach space. For every $\tau > 0$ we have*

$$\rho_{\mathcal{X}^*}(\tau) = \sup\{\frac{1}{2}\tau\varepsilon - \delta_{\mathcal{X}}(\varepsilon) : 0 < \varepsilon \leq 2\},\ \rho_{\mathcal{X}}(\tau) = \sup\{\frac{1}{2}\tau\varepsilon - \delta_{\mathcal{X}^*}(\varepsilon) : 0 < \varepsilon \leq 2\} \quad (4.3)$$

Proof:
Let us proof the first formula.
Let $\tau > 0 < \varepsilon \leq 2, x, y \in \mathcal{X}$ with $\|x\| = \|y\| = 1$. From the Hahn-Banach theorem we obtain $\lambda, \mu \in \mathcal{X}^*$ mit $\|\lambda\| = \|\mu\| = 1$ such that

$$\langle \lambda, x+y \rangle = \|x+y\| \text{ and } \langle \mu, x-y \rangle = \|x-y\|.$$

Then

$$\begin{aligned} \|x+y\| + \tau\|x-y\| - 2 &= \langle \lambda, x+y \rangle + \tau\langle \mu, x-y \rangle = \|x-y\| - 2 \\ &= \langle \lambda+\tau\mu, x \rangle + \langle \lambda-\tau\mu, y \rangle - 2 \\ &\leq \|\lambda+\tau\mu\| + \|\lambda-\tau\mu\| - 2 \\ &\leq \sup\{\|\lambda+\tau\mu\| + \|\lambda-\tau\mu\| - 2 : \|x\| = \|y\| = 1\} \\ &= 2\rho_{\mathcal{X}^*}(\tau) \end{aligned}$$

If additional $\|x-y\| \geq \varepsilon$ then we obtain

$$\frac{1}{2}\tau\varepsilon - \rho_{\mathcal{X}^*}(\tau) \leq 1 - \frac{1}{2}\|x+y\|\,.$$

Therefore $\frac{1}{2}\tau\varepsilon - \rho_{\mathcal{X}^*}(\tau) \leq \delta_{\mathcal{X}}(\varepsilon)$ and since $0 < \varepsilon \leq 2$ is arbitrary chosen

$$\sup\{\frac{1}{2}\tau\varepsilon - \delta_{\mathcal{X}}(\varepsilon) : 0 < \varepsilon \leq 2\} \leq \rho_{\mathcal{X}^*}(\tau)\,.$$

Choose $\lambda, \mu \in X^*$ with $\|\lambda\| = \|\mu\| = 1$ and $\delta > 0$. For $\tau > 0$ there exist $x^0, y^0 \in \mathcal{X}$ with $x^0 \neq y^0$ and $\|x^0\| = \|y^0\| = 1$ such that

$$\|\lambda + \tau\mu\| \leq \langle \lambda + \tau\mu, x^0\rangle + \delta\,,\ \|\lambda - \tau\mu\| \leq \langle \lambda - \tau\mu, y^0\rangle + \delta\,.$$

Using these inequalities we obtain

$$\begin{aligned}
\|\lambda + \tau\mu\| + \|\lambda - \tau\mu\| - 2 &\leq \langle \lambda + \tau\mu, x^0\rangle + \langle \lambda - \tau\mu, y^0\rangle - 2 + 2\delta \\
&= \langle \lambda, x^0 + y^0\rangle + \langle \mu, x^0 - \tau y^0\rangle - 2 + 2\delta \\
&\leq \|x^0 + y^0\| - 2 + \tau|\langle \mu, x^0 - y^0\rangle| + 2\delta\,.
\end{aligned}$$

We set $\varepsilon_0 := |\langle \mu, x^0 - y^0\rangle|$. Then we have $0 < \varepsilon_0 \leq \|x^0 + y^0\| \leq 2$ and

$$\begin{aligned}
\frac{1}{2}(\|x^0 + \tau y^0\| + \|x^0 - \tau y^0\|) - 1 &\leq \frac{1}{2}\tau\varepsilon_0 + \delta - \delta_X(\varepsilon_0) \\
&\leq \delta + \sup\{\frac{1}{2}\tau\varepsilon - \delta_{\mathcal{X}}(\varepsilon) : 0 < \varepsilon \leq 2\}\,.
\end{aligned}$$

Since $\delta > 0$ is arbitrary chosen

$$\rho_{\mathcal{X}^*}(\tau) \leq \sup\{\frac{1}{2}\tau\varepsilon - \delta_{\mathcal{X}}(\varepsilon) : 0 < \varepsilon \leq 2\}$$

and the proof is complete.
Let us prove the second formula.
Let $\tau > 0$ and let $\lambda, \mu \in \mathcal{X}^*$ with $\|\lambda\| = \|\mu\| = 1$. For $\eta > 0$ there exist x^0, y^0 in $\mathcal{X}$ with $\|x^0\| = \|y^0\| = 1$ such that

$$\|\lambda + \mu\| - \eta \leq \langle \lambda + \mu, x^0\rangle\,,\ \|\lambda - \mu\| - \eta \leq \langle \lambda - \mu, y^0\rangle\,.$$

Then

$$\begin{aligned}
\|\lambda + \mu\| + \tau\|\lambda - \mu\| - 2 &\leq \langle \lambda + \mu, x^0\rangle + \tau\langle \lambda - \mu, y^0\rangle - 2\eta(1+\tau) \\
&\leq \sup\{\|x + \tau y\| + \|x - \tau y\| - 2 : \|x\| = \|y\| = 1\} + \eta(1+\tau) \\
&= 2\rho_{\mathcal{X}}(\tau) + \eta(1+\tau)\,.
\end{aligned}$$

If $0 < \varepsilon \leq \|\lambda - \mu\| \leq 2$ then we have

$$\frac{1}{2}\tau\varepsilon - \rho_{\mathcal{X}}(\tau) - \eta(1+\tau) \leq 1 - \|\frac{1}{2}(\lambda + \mu)\|,$$

was

$$\frac{1}{2}\tau\varepsilon - \rho_{\mathcal{X}}(\tau) - \eta(1+\tau) \leq \delta_{\mathcal{X}^*}(\varepsilon)\,.$$

Since $\eta > 0$ is arbitrary we obtain

$$\frac{1}{2}\tau\varepsilon - \rho_{\mathcal{X}}(\tau) \leq \delta_{\mathcal{X}^*}(\varepsilon)$$

for all $\varepsilon \in (0, 2]$ and finally

$$\sup\{\frac{1}{2}\tau\varepsilon - \rho_{\mathcal{X}}(\tau) : 0 < \varepsilon \leq 2\} \leq \rho_{\mathcal{X}}(\tau)\,.$$

Let $x, y \in \mathcal{X}$ with $\|x\| = \|y\| = 1$ and let $\tau > 0$. From the Hahn-Banach theorem we obtain $\lambda, \mu \in \mathcal{X}^*$ with $\|\lambda\| = \|\mu\| = 1$, such that

$$\langle \lambda, x + \tau y\rangle = \|x + \tau y\| \text{ and } \langle \mu, x - \tau y\rangle = \|x - \tau y\|\,.$$

Then

$$\begin{aligned}
\|x + \tau y\| + \|x - \tau y\| - 2 &= \langle \lambda, x + \tau y\rangle + \langle \mu, x - \tau y\rangle = \|x - \tau y\| - 2 \\
&= \langle \lambda + \mu, x\rangle + \tau\langle \lambda - \mu, y\rangle - 2 \\
&\leq \|\lambda + \tau\mu\| + \|\lambda - \tau\mu\| - 2
\end{aligned}$$

We define $\varepsilon_0 := |\langle \lambda - \mu, y\rangle|$. Then we have $0 < \varepsilon_0 \leq \|x - y\| \leq 2$ and

$$\begin{aligned}
\frac{1}{2}(\|x + \tau y\| + \|x - \tau y\|) - 1 &\leq \frac{1}{2}(\|\lambda + \mu\| + \tau|\langle \lambda - \mu, y\rangle|) - 1 \\
&= \frac{1}{2}\tau\varepsilon_0 - (1 - \frac{1}{2}\|\lambda + \mu\|) \\
&\leq \frac{1}{2}\tau\varepsilon_0 - \delta_{\mathcal{X}^*}(\varepsilon_0) \\
&\leq \sup\{\frac{1}{2}\tau\varepsilon - \delta_{\mathcal{X}^*}(\varepsilon) : 0 < \varepsilon \leq 2\}\,.
\end{aligned}$$

This implies

$$\sup\{\frac{1}{2}\tau\varepsilon - \rho_X(\tau) : 0 < \varepsilon \leq 2\} \geq \rho_{\mathcal{X}}(\tau)\,.$$

■

Let us add a few properties of the convexity and smoothness modulus.

Corollary 4.17. *Let $\mathcal{X}$ be a Banach space and let $\mathcal{H}$ be a Hilbert space.*

(1) $\delta_{\mathcal{X}}$ *is convex and continuous.*

(2) $\delta_{\mathcal{X}}$ *is strictly in increasing iff* $\mathcal{X}$ *is uniformly convex.*

(3) $\delta_{\mathcal{X}}(\varepsilon) \leq \delta_{\mathcal{H}}(\varepsilon), \varepsilon \in [0, 2]$.

(4) $\rho_{\mathcal{X}}(\tau) \geq \rho_{\mathcal{H}}(\tau), \tau \in (0, \infty)$.

(5) $\rho_{\mathcal{H}}(\tau) = \sqrt{1 + \tau^2} - 1$.

Proof:
For the proofs of $(1)-(4)$ we refer to the literature ([11, 30, 41]).
Ad (5) this follows from the second formula in Theorem 4.16.

■

Theorem 4.18 (Shmulian, 1940). *Let $\mathcal{X}$ be a Banach space. We have:*

(a) $\mathcal{X}$ is uniformly smooth if and only if $\mathcal{X}^*$ is uniformly convex.

(b) $\mathcal{X}$ is uniformly convex if and only if $\mathcal{X}^*$ is uniformly smooth.

Proof:
Ad (a) Let $\mathcal{X}$ be uniformly smooth. Assume that $\mathcal{X}^*$ is not uniformly convex.
Then there exists $\varepsilon_0 \in (0,2]$ with $\delta_{\mathcal{X}^*}(\varepsilon_0) = 0$ and we obtain from Theorem 4.16 for all $\tau > 0$ $0 < \frac{1}{2}\varepsilon_0 \leq \rho_{\mathcal{X}}(\tau)/\tau$. This implies that $\mathcal{X}$ is not uniformly smooth.
Let $\mathcal{X}^*$ be uniformly convex. Assume that $\mathcal{X}$ is not uniformly smooth.
Then $\lim_{\tau\downarrow 0} \rho_{\mathcal{X}}(\tau)/\tau \neq 0$. Hence, there exists $\varepsilon > 0$ such that for all $\delta > 0$ there exists t with $0 < t < \delta$ and $t\varepsilon \leq \rho_{\mathcal{X}}(t)$. Then there exists a sequence $(\tau_n)_{n\in\mathbb{N}}$ with $0 < \tau_n < 1, \rho_{\mathcal{X}}(\tau_n) > \frac{1}{2}\varepsilon\tau_n, n \in \mathbb{N}, \lim_n \tau_n = 0$. Due to Theorem 4.16 we obtain for each $n \in \mathbb{N}$ $\varepsilon_n \in (0,2]$ such that $\frac{1}{2}\varepsilon\tau_n \leq \frac{1}{2}\tau_n\varepsilon_n - \delta_{\mathcal{X}^*}(\varepsilon_n)$. This implies $0 < \delta_{\mathcal{X}^*}(\varepsilon_n) \leq \frac{1}{2}\tau_n(\varepsilon_n - \varepsilon)$. Morreover, $\varepsilon < \varepsilon_n$ und $\lim_n \delta_{\mathcal{X}^*}(\varepsilon_n) = 0$. Since $\delta_{\mathcal{X}^*}$ is nondecreasing we obtain $\delta_{\mathcal{X}^*}(\varepsilon) \leq \delta_{\mathcal{X}^*}(\varepsilon_n), n \in \mathbb{N}$. This shows $\delta_{\mathcal{X}^*}(\varepsilon) \leq 0$. Hence $\mathcal{X}^*$ is not uniformly convex.
Ad (b) The proof folows along the lines above by exchanging the roles of $\mathcal{X}$ and $\mathcal{X}^*$. ■

Lemma 4.19. *Let $\mathcal{X}$ be a uniformly smooth Banach space. Then its duality map J is single-valued and uniformly continuous on bounded sets.*

Proof:
Since $\mathcal{X}$ is uniformly smooth $\mathcal{X}^*$ is uniformly convex; see Theorem 4.18. Then we know that J is single-valued; see Lemma 3.14. It is enough to prove the uniformly boundeness property for the bounded set $\overline{S}_1$. Let $\varepsilon > 0$ and let $x, y \in \overline{S}_1$ with $\|x - y\| < 2\delta_{\mathcal{X}^*}(\varepsilon)$. Then

$$\begin{aligned}
\|J_{\mathcal{X}}(x) - J_{\mathcal{X}}(y)\| &\geq \langle J_{\mathcal{X}}(x) + J_{\mathcal{X}}(y), x\rangle \\
&= \langle J_{\mathcal{X}}(x), x\rangle + \langle J_{\mathcal{X}}(y), y\rangle + \langle J_{\mathcal{X}}(y), x - y\rangle \\
&\geq \|x\|^2 + \|y\|^2 - \|x - y\|\|J_{\mathcal{X}}(y)\| \;>\; 2(1 - \delta_{\mathcal{X}^*}).(\varepsilon)
\end{aligned}$$

This shows $\|J_{\mathcal{X}}(x) - J_{\mathcal{X}}(y)\| < \varepsilon$. ■

Let $\mathcal{H}$ be a Hilbert space. Then with the results in Theorem 4.16 we obtain

$$\delta_{\mathcal{H}}(\varepsilon) = 1 - \sqrt{1 - \frac{1}{4}\varepsilon^2} = \frac{1}{8}\varepsilon^2 + O(\varepsilon^4)\,,\ \rho_{\mathcal{H}}(\tau) = \sqrt{1 + \tau^2} - 1 = \frac{1}{2}\tau^2 + O(\tau^4)\,.$$

Hilbert spaces are the „most uniformly convex“ and „most uniformly smooth “ Banach spaces in the sense that if any space whose dimension is at least two, then $\delta_{\mathcal{H}}(\varepsilon) \geq \delta_{\mathcal{X}}(\varepsilon)$ and $\rho_{\mathcal{H}}(\tau) \leq \rho_{\mathcal{X}}(\tau)$; see for instance [34], Day [28], [30]. Quantitative versions of strict convexity and smoothness are of particular importance in Banach space theory. Here they can be used to quantify the continuity properties of the metric projection; see below.

Remark 4.20. *For uniformly convex smooth Banach spaces there exist criteria for convexity of their Chebyshev sets: if C be a Chebyschev subset of the uniformly convex smooth Banach space $\mathcal{X}$ then C is convex iff C is approximately compact.* □

4.3 Strongly smoothness

Let $\mathcal{X}$ be a Banach space with norm $\|\cdot\|$ and norm mapping $\nu : \mathcal{X} \ni x \longmapsto \|x\| \in \mathbb{R}$. If the norm is Gateaux differentiable in $x \in \mathcal{X}\backslash\{\theta\}$ then the limit

$$\lim_{t\to 0} \frac{\|x+ty\| - \|x\|}{t} \tag{4.4}$$

exists and is equal to the value $\nu'(x,y)$ of the uniquely determined functional $\nu'(x,\cdot) \in \mathcal{X}^*$; see Section 3.3.

Definition 4.21. *Let $\mathcal{X}$ be a Banach space.*

(a) The norm in $\mathcal{X}$ is called **Fréchet-differentiable** *at $x \in \mathcal{X}\backslash\{\theta\}$ if*

$$\lim_{t\to 0} \frac{\|x+ty\| - \|x\|}{t}$$

exists uniformly in $y \in \overline{S}_1$. The limit is called the **Fréchet derivative** *at x in direction y.*

(b) The norm in $\mathcal{X}$ is called **Fréchet-differentiable** *if the norm is Fréchet-differentiable at each $x \in \mathcal{X}\backslash\{\theta\}$.*

(c) $\mathcal{X}$ is called **strongly smooth** *if the norm is Fréchet-differentiable at each $x \in \mathcal{X}\backslash\{\theta\}$.*

(d) The norm in $\mathcal{X}$ is called **uniformly Fréchet-differentiable** *if the limit in*

$$\lim_{t\to 0} \frac{\|x+ty\| - \|x\|}{t}$$

exists uniformly for $(x,y) \in \overline{S}_1 \times \overline{S}_1$.

□

Let $\mathcal{X}$ be a Banach space with norm $\|\cdot\|$ and norm-mapping ν. It is clear that at a point x where the norm is Fréchet differentiable the norm is Gateaux differentiable and the Fréchet derivative in x in direction y is equal to $\nu'(x,y)$. Since the norm is homogeneous one has to check the Fréchet differentiability at $x \in \overline{S}_1$ only. In the general, Gateaux differentiability of the norm does not imply Fréchet differentiability. An example is the norm in l_1; see [30].

Lemma 4.22. *Let $\mathcal{X}$ be a Banach space with norm $\|\cdot\|$ and norm-mapping ν. Let $x \in \mathcal{X}\backslash\{\theta\}$. Then the following statements are equivalent:*

(a) The norm is Fréchet differentiable in x.

(b)

$$\lim_{v\to\theta} \frac{\|x+v\| - \|x\| - \nu'(x,v)}{\|v\|} = 0\,.$$

Proof:
Observe that $v = ty$ with $t = \|v\|, y = v\|v\|^{-1} \in \overline{S}_1$. ■

Lemma 4.23 (Shmulian, 1939). *Let $\mathcal{X}$ be a Banach space with norm-mapping ν. Then for $x \in \overline{S}_1$ the following conditions are equivalent:*

(a) The norm is Fréchet-differentiable in x.

(b) If $(\lambda_n)_{n\in\mathbb{N}}$ is a sequence in $\mathcal{X}^*$ with $\|\lambda_n\| \leq 1, n \in \mathbb{N}$, and $\lim_n \langle \lambda_n, x\rangle = 1$ then this sequence is convergent.

Proof:
(a) $\Longrightarrow$ (b) Let $\lambda := \nu'(x, \cdot)$ and let $(\lambda_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{X}^*$ with $\|\lambda_n\| \leq 1, n \in \mathbb{N}$, and $\lim_n \langle \lambda_n, x\rangle = 1$. Since

$$1 \geq \|\lambda_n\| \geq \langle \lambda_n, x\rangle, n \in \mathbb{N}, \text{ and } \lim_n \langle \lambda_n, x\rangle = 1$$

we may suppose that $\|\lambda_n\| = 1$ for all $n \in \mathbb{N}$.
Assume that $(\lambda_n)_{n\in\mathbb{N}}$ does not converge to λ. Then there exists $\varepsilon > 0$ and $(z^n)_{n\in\mathbb{N}}$ in $\mathcal{X}$ with $\|z^n\| = 1, n \in \mathbb{N}$, with $\langle \lambda_n - \lambda, z^n\rangle \geq 2\varepsilon$, $n \in \mathbb{N}$. We set $x^n := z^n(1 - \langle \lambda_n, x\rangle)\varepsilon^{-1}$. Then $\lim_n x^n = \theta$ and (see 4.22)

$$\begin{aligned}
\frac{\nu(x + x^n) - \nu(x) - \langle \lambda, x^n\rangle}{\|x^n\|} &\geq \frac{\langle \lambda_n, x + x^n\rangle - 1 - \langle \lambda, x^n\rangle}{\|x^n\|} \\
&= \frac{\langle \lambda_n, x\rangle - 1 + \langle \lambda_n - \lambda, z^n(1 - \langle \lambda_n, x\rangle)\varepsilon^{-1}\rangle}{(1 - \langle \lambda_n, x\rangle)\varepsilon^{-1}} \\
&= \langle \lambda_n - \lambda, z^n\rangle - \varepsilon \ \geq \ \varepsilon
\end{aligned}$$

This contradicts the Fréchet-differentiability of ν in x.
(b) $\Longrightarrow$ (a) Set $\lambda := \nu'_+(x, \cdot)$. Assume that there exists $\varepsilon > 0$ and $(x^n)_{n\in\mathbb{N}}$ with $\lim_n x^n = \theta$ such that

$$\left| \frac{\nu(x + x^n) - \nu(x) - \langle \lambda, x^n\rangle}{\|x^n\|} \right| \geq \varepsilon, \ n \in \mathbb{N}.$$

Then

$$\|x + x^n\| - \langle \lambda, x + x^n\rangle \geq \varepsilon\|x^n\|, \ n \in \mathbb{N}.$$

Choose $\lambda_n \in \overline{S}_1$ such that $\langle \lambda_n, x + x^n\rangle = \|x + x^n\|, n \in \mathbb{N}$. Since $\lim_n x^n = \theta$ we have $\lim_n \langle \lambda_n, x\rangle = \|x\| = 1$. But

$$\|\lambda_n - \lambda\| \geq \langle \lambda_n - \lambda, x^n\rangle\|x^n\|^{-1} \geq \frac{\langle \lambda, x\rangle - \langle \lambda_n, x\rangle}{\|x^n\|} + \varepsilon \geq \varepsilon$$

since $\langle \lambda, x\rangle = \|x\| \geq \langle \lambda_n, x\rangle$. This means that $(\lambda_n)_{n\in\mathbb{N}}$ does not converge to λ. ■

Corollary 4.24. *Let $\mathcal{X}$ be a Banach space and let the norm in the dual space $\mathcal{X}^*$ be Fréchet differentiable. Then $\mathcal{X}$ is reflexive.*

Proof:
We use the result of James which says that $\mathcal{X}$ is reflexive if and only if each nonzero functional $\lambda \in \mathcal{X}^*$ attains its norm at some $x \in \overline{S}_1$; see the preliminaries in the preface.
Let $\lambda \in \mathcal{X}^*, \|\lambda\| = 1$, and let $(x^n)_{n\in\mathbb{N}}$ be a sequence in $\overline{S}_1$ such that $\lim_n \langle \lambda, x^n\rangle = 1$. By applying Smulian's lemma, $\lim_n x^n = x \in \overline{S}_1$. Therefore

$$\langle \lambda, x\rangle = \lim_n \langle \lambda, x^n\rangle = 1 = \|\lambda\|.$$

If now $\mu \in \mathcal{X}^*\backslash\{\theta\}$ apply the consideration above to $\lambda := \mu\|\mu\|^{-1}$. ■

Theorem 4.25. *Let $\mathcal{X}$ be a Banach space with duality map J and norm mapping ν. Then we have the equivalence of the following conditions:*

(a) J *is single-valued and uniformly norm-to-norm continuous an bounded subsets.*

(b) The norm in $\mathcal{X}$ is uniformly Fréchet differentiable.

(c) $\mathcal{X}$ is uniformly smooth.

(d) $\mathcal{X}^*$ is uniformly convex.

Proof:
$(a) \implies (b)$ Let $x \in \overline{S}_1, y \in \mathcal{X}$. Let $t > 0$ be such that $\|x + ty\| > 0$. Then

$$\begin{aligned}
\langle J_{\mathcal{X}}(x), y\rangle &= \frac{\langle J_{\mathcal{X}}(x), x\rangle - \|x\| + \langle J_{\mathcal{X}}(x), ty\rangle}{t} \\
&= \frac{\langle J_{\mathcal{X}}(x), x + ty\rangle - \|x\|}{t} \\
&\leq \frac{\|x + ty\| - \|x\|}{t} \\
&= \frac{\|x + ty\|^2 - \|x\|\|x + ty\|}{t\|x + ty\|} \\
&\leq \frac{\langle J_{\mathcal{X}}(x + ty), x + ty\rangle - |\langle J_{\mathcal{X}}(x + ty), x\rangle|}{t\|x + ty\|} \\
&= \frac{t\langle J_{\mathcal{X}}(x + ty), y\rangle + \langle J_{\mathcal{X}}(x + ty), x\rangle - |\langle J_{\mathcal{X}}(x + ty), x\rangle|}{t\|x + ty\|} \\
&\leq \frac{\langle J_{\mathcal{X}}(x + ty), y\rangle}{\|x + ty\|}
\end{aligned}$$

Hence

$$\frac{\langle J_{\mathcal{X}}(x), y\rangle}{\|x\|} \leq \frac{\|x + ty\| - \|x\|}{t} \leq \frac{\langle J_{\mathcal{X}}(x + ty), y\rangle}{\|x + ty\|}.$$

Using the property (a) we read off that the norm is uniformly Fréchet differentiable.
$(b) \implies (c)$ Assuming (b) we can choose for $\varepsilon > 0$ a $\delta > 0$ so that for all $t \in (0, \delta)$ and all $x, y \in \overline{S}_1$

$$\left|\frac{\|x + ty\| - \|x\|}{t} - \nu'(x, y)\right| < \frac{1}{2}\varepsilon.$$

But this implies that for all $t \in (0, \delta)$ and all $x, y \in \overline{S}_1$

$$\begin{aligned}
\|x + ty\| - \|x - ty\| &= 2 + t\left(\frac{\|x + ty\| - \|x\|}{t} - \nu'(x, y)\right) \\
&\quad + t\left(\frac{\|x + t(-y)\| - \|x\|}{t} - \nu'(x, -y)\right) \\
&< 2 + t\varepsilon.
\end{aligned}$$

This implies the assertion in (c)
$(c) \implies (d)$ Let $\varepsilon > 0$. By (c) we can find $\delta > 0$ such that for all $x \in \overline{S}_1$ and $z \in \mathcal{X}$ with $\|z\| \leq \delta$ we have $\|x + z\| + \|x - z\| < 2 + \frac{1}{4}\varepsilon\|z\|$.

Let $\lambda, \mu \in \mathcal{X}^*$ with $\lambda\| = \|\mu\| = 1$ and $\|\lambda - \mu\| \geq \varepsilon$. Then there is a $z \in \mathcal{X}$ with $\|z\| \leq \frac{1}{2}\delta$ and $\langle \lambda - \mu, z\rangle \geq \frac{1}{2}\varepsilon\delta$. This implies

$$\begin{aligned}
\|\lambda + \mu\| &= \sup_{x \in \overline{S}_1} |\langle \lambda + \mu, z\rangle \\
&= \sup_{x \in \overline{S}_1} |\langle \lambda, x + z\rangle + \langle \mu, x - z\rangle - \langle \lambda - \mu, z\rangle| \\
&\leq \sup_{x \in \overline{S}_1} (\|x + z\| + \|x - z\|) - \frac{1}{2}\varepsilon\delta \\
&\leq 2 + \frac{1}{4}\varepsilon\|z\| - \frac{1}{2}\varepsilon\delta \ < \ 2 - \frac{1}{4}\varepsilon\delta .
\end{aligned}$$

(c) $\Longrightarrow$ (a) J is single-valued due to Lemma 3.14. Let $\varepsilon > 0$. Then there exists $\delta > 0$ such that for all $\lambda, \mu \in \overline{S}_1 \subset \mathcal{X}^*$ we have

$$\|\lambda + \mu\| > 2 - \delta \text{ implies } \|\lambda - \mu\| < \varepsilon .$$

Choose $x, y \in \overline{S}_1 \subset \mathcal{X}$ with $\|x\| + \mu\| = 1$ and $\|x - y\| = \delta$. Then with $\lambda \in J_{\mathcal{X}}(x), \mu \in J_{\mathcal{X}}(y)$ we obtain

$$\|\lambda + \mu\| \geq \langle \lambda + \mu, y\rangle = \langle \lambda, x\rangle + \langle \mu, y\rangle - \langle \lambda, x - y\rangle > 2 - \|x - y\| > 2 - \delta .$$

Therefore J is norm-to-norm continuous on unit sphere. Since J is homogeneous, i.e. $J_{\mathcal{X}}(ax) = aJ_{\mathcal{X}}(x)$ for all $x \in \mathcal{X}, a \in \mathbb{R}$, J is norm-to-norm continuous on the unit ball. This implies that J is uniformly continuous on each bounded set.

■

Remark 4.26. *If the the norm in a Banach space is Fréchet differentiable then the duality map is uniformly continuous on bounded subsets from the norm-topology to the weak*-topology; see [10].* □

Theorem 4.27. *Let $\mathcal{X}$ be a Banach space. Then the following statements are equivalent:*

(a) $\mathcal{X}$ is an E-space.

(b) $\mathcal{X}^*$ is strongly smooth.

Proof:
(a) $\Longrightarrow$ (b) We know by definition that $\mathcal{X}$ is reflexive and strictly convex and therefore $\mathcal{X}^*$ smooth by Theorem 3.27. We want to use Lemma 4.23 to verify that the norm in $\mathcal{X}^*$ is Fréchet differentiable at each $\lambda \in \mathcal{X}^*$.
Let $\lambda \in \mathcal{X}^*$ with $\|\lambda\| = 1$ and let $(x^n)_{n\in\mathbb{N}}$ be a sequence in $\overline{B}_1$ with $\lim_n \langle \lambda, x^n\rangle = 1$. Now, $1 \geq \|x^n\| \geq \langle \lambda, x^n\rangle, n \in \mathbb{N}$, which implies $\lim_n \|x^n\| = 1$. Let $z^n := x^n\|x^n\|^{-1}, n \in \mathbb{N}$, and let z be a weak cluster point of the sequence $(z^n)_{n\in\mathbb{N}}$ in $\overline{S}_1$; it exists due to the reflexivity of $\mathcal{X}$. Then $\|z\| \leq 1$, but $\langle \lambda, z\rangle = \lim_n \langle \lambda, x^n\rangle\|x^n\|^{-1} = 1$. So $\|z\| = 1$. Since $\mathcal{X}$ is an E-space we obtain that z is cluster point of $(x^n)_{n\in\mathbb{N}}$ with respect to the norm convergence. Because $\mathcal{X}^*$ is smooth, z is uniquely determined by the conditions $z \in \overline{S}_1, \langle \lambda, z\rangle = 1$. Therefore, $\lim_n z^n = z$, and so $\lim_n x^n = x$ also.
(b) $\Longrightarrow$ (a) This is proved similarly. We omit the proof. ■

4.4 Radial projection and norm inequalities

Here we want to study the projection onto the unit ball in a Banach space $\mathcal{X}$. Such a projection is the mapping

$$P_{rad} : \mathcal{X} \ni x \longmapsto \begin{cases} \frac{x}{\|x\|} & \text{if } \|x\| \geq 1 \\ x & \text{if } \|x\| \leq 1 \end{cases} \in \overline{B}_1 \subset \mathcal{X}. \tag{4.5}$$

Lemma 4.28. *Let $\mathcal{X}$ be a Banach space and let $x \in \mathcal{X}\backslash\{\theta\}$. Then $P_{rad} \in P_{\overline{B}_1}$, especially $\overline{B}_1$ is a proximal subset.*

Proof:
We want to use the characterization result 3.32. Let $\lambda \in J_{\mathcal{X}}(x - \frac{x}{\|x\|})$. Then $\lambda = (1 - \frac{1}{\|x\|})\mu$ with $\mu \in J_{\mathcal{X}}(x)$. We want to show that $\lambda \in N(\frac{x}{\|x\|}, \overline{B}_1)$. This follows from

$$\langle \lambda, u - \frac{x}{\|x\|} \rangle = (1 - \frac{1}{\|x\|})(\langle \mu, u \rangle - \langle \mu, \frac{x}{\|x\|} \rangle) \leq (1 - \frac{1}{\|x\|})(\|\mu\|\|u\| - \|x\|) \leq 0 \text{ for all } u \in \overline{B}_1.$$

■

As a consequence of Lemma 4.28, $P_{\overline{B}_1}(x) = P_{rad}(x)$ for all $x \in \mathcal{X}\backslash\{\theta\}$ if $\mathcal{X}$ is strictly convex; see Theorem 2.15. Thus in a Hilbert space the radial projection $P_{rad}(x)$ is the best approximation of x in the unit ball. Thus, in this Hilbert space case we know that P_{rad} is nonexpansive.

As we see in the following, the nonexpansivity property of P_{rad} does not hold in the Banach space case in general. To study this question, let us begin with a few famous inequalities concerning the triangle inequality in Hilbert and Banach spaces which may be used to compute the Lipschitz constant of the projection P_{rad}.

Theorem 4.29 (Dunkl and Williams, 1964)**.** *Let $\mathcal{H}$ be a Hilbert space and let $x, y \in \mathcal{H}, x \neq \theta, y \neq \theta$. Then*

$$\|\frac{x}{\|x\|} - \frac{y}{\|y\|}\| \leq \frac{2}{\|x\| + \|y\|}\|x - y\| \tag{4.6}$$

Proof:

$$\begin{aligned} \|\frac{x}{\|x\|} - \frac{y}{\|y\|}\|^2 &= \langle \frac{x}{\|x\|} - \frac{y}{\|y\|} | \frac{x}{\|x\|} - \frac{y}{\|y\|} \rangle \\ &= 2 - 2\langle \frac{x}{\|x\|} | \frac{y}{\|y\|} \rangle \\ &= \frac{1}{\|x\|\|y\|}(2\|x\|\|y\| - \langle x|y \rangle) \end{aligned}$$

This implies

$$\|x - y\|^2 - \Big(\frac{\|x\| + \|y\|}{2}\Big)^2 \|\frac{x}{\|x\|} - \frac{y}{\|y\|}\|^2 = \frac{(\|x\| - \|y\|)^2}{4\|x\|\|y\|}\Big((\|x\| + \|y\|)^2 - \|x - y\|^2\Big)$$

The righthand side in the last equality is nonnegative as one easily can check.

■

Theorem 4.30 (Maligranda, 2006). *Let $\mathcal{X}$ be a Banach space and let $x, y \in \mathcal{X}, x \neq \theta, y \neq \theta$.*

(a) $\|x+y\| \leq \|x\| + \|y\| - (2 - \|\frac{x}{\|x\|} + \frac{y}{\|y\|}\|) \min(\|x\|, \|y\|)$

(b) $\|x+y\| \geq \|x\| + \|y\| - (2 - \|\frac{x}{\|x\|} + \frac{y}{\|y\|}\|) \max(\|x\|, \|y\|)$

If either $\|x\| = \|y\| = 1$ or $y = ax$ with $a > 0$, then equality holds in both (a) or (b).

Proof:
We follow Maligranda [29].
Without loss of generality assume $\|y\| \geq \|x\|$, i.e. $\min(\|x\|, \|y\|) = \|x\|$. Ad (a)

$$\begin{aligned}
\|x+y\| &= \Big\|\frac{\|x\|}{\|x\|}x + \frac{\|x\|}{\|y\|}y + (1 - \frac{\|x\|}{\|y\|})y\Big\| \\
&\leq \|x\|\|\frac{x}{\|x\|} + \frac{y}{\|y\|}\| + \|y\| - \|x\| \\
&= \|y\| + \Big(\|\frac{x}{\|x\|} + \frac{y}{\|y\|}\| - 1\Big)\|x\| \\
&= \|x\| + \|y\| + \Big(\|\frac{x}{\|x\|} + \frac{y}{\|y\|}\| - 2\Big)\|x\|.
\end{aligned}$$

Ad (b)

$$\begin{aligned}
\|x+y\| &= \Big\|\frac{\|y\|}{\|y\|}x + \frac{\|y\|}{\|x\|}x + \Big(1 - \frac{\|y\|}{\|x\|}\Big)x\Big\| \\
&\geq \|y\|\|\frac{y}{\|y\|} + \frac{x}{\|x\|}\| - |\|x\| - \|y\||\\
&= \|y\|\|\frac{y}{\|y\|} + \frac{x}{\|x\|}\| - \|y\| + \|x\| \\
&= \|x\| + \|y\| - \Big(2 - \|\frac{x}{\|x\|} + \frac{y}{\|y\|}\|\Big)\|y\|
\end{aligned}$$

The additional assertion concerning the equality can easily be verified. ■

Corollary 4.31. *Dunkl and Williamson, 1964] Let $\mathcal{X}$ be a Banach space and let $x, y \in \mathcal{X}, x \neq \theta, y \neq \theta$. Then*

$$\|\frac{x}{\|x\|} - \frac{y}{\|y\|}\| \leq \frac{4}{\|x\| + \|y\|}\|x - y\| \tag{4.7}$$

Proof:
From (b) in Theorem 4.30 we obtain

$$\|\frac{x}{\|x\|} - \frac{y}{\|y\|}\| \leq \frac{\|x-y\| + |\|x\| - \|y\||}{\max(\|x\|, \|y\|)}$$

which implies (4.7). ■

Example 4.32. *Consider* $\mathbb{R}^2$ *endowed with the* l_1*-norm. Take* $x = (1, \varepsilon)$ *and* $y := (1, 0)$ *where* ε *is small. Then*

$$\|\frac{x}{\|x\|} - \frac{y}{\|y\|}\| \frac{\|x\| + \|y\|}{\|x - y\|} = \frac{4 + 2\varepsilon}{1 + \varepsilon}$$

where

$$\lim_{\varepsilon \downarrow 0} \frac{4 + 2\varepsilon}{1 + \varepsilon} = 4\,.$$

If the norm in $\mathbb{R}^2$ *is given by the* l_∞*-norm then we obtain for* $x := (1, 1)$ *and* $y := (1 - \varepsilon, 1 + \varepsilon)$*, where* $\varepsilon > 0$ *is small,*

$$\lim_{\varepsilon \downarrow 0} \|\frac{x}{\|x\|} - \frac{y}{\|y\|}\| \frac{\|x\| + \|y\|}{\|x - y\|} = \lim_{\varepsilon \downarrow 0} \frac{4 + 2\varepsilon}{1 + \varepsilon} = 4\,.$$

□

Let $\mathcal{X}$ be a Banach space and let $x, y \in \mathcal{X} \backslash \{\theta\}$. The quantity

$$\alpha(x, y) := \|\frac{x}{\|x\|} - \frac{y}{\|y\|}\|$$

is called the **angular distance** or **Clarkson distance** of x, y. As we see from the results in Theorem 4.30 this quantity controls the triangle inequality. In [12] a detailed analysis of the triangle inequality bas been made for uniformly convex Banach spaces. From (b) in Theorem 4.30 we obtain (see the proof of 4.31

$$\alpha(x, y) \leq \frac{2\|x - y\|}{\max(\|x\|, \|y\|}\,. \tag{4.8}$$

This a sligthly strengthening of the Dunkl-Williams inequality (4.6) in Theorem 4.29. Kirk and Smiley showed in [25] that this inequality characterizes inner product spaces: If in a Banach space $\mathcal{X}$ the inequality (4.8) holds for all $x, y \in \mathcal{X} \backslash \{\theta\}$ then $\mathcal{X}$ is actually a Hilbert space; see also [32, 37].

This observation may be seen a motivation to introduce the so called **Dunkl-Williams-constant**:

$$\mathrm{DW}(\mathcal{X}) := \sup\{\alpha(x, y) \frac{\|x\| + \|y\|}{\|x - y\|} : x, y \in \mathcal{X} \backslash \{\theta\}, x \neq y\}\,. \tag{4.9}$$

Observe that $2 \leq \mathrm{DW}(\mathcal{X}) \leq 4$. We know from example 4.32 that $\mathrm{DW}(\mathbb{R}^2, \|\cdot\|_1) = \mathrm{DW}(\mathbb{R}^2, \|\cdot\|_\infty) = 4$. The Dunkl-Williams-constant may be considered as a measure how much a Banach space is close to a Hilbert space.

Let us come back to the radial projection P_{rad}. It is obvious that the inequalities above play an important role for studying the Lipschitz continuity of P_{rad}.

Theorem 4.33. *Let* $\mathcal{X}$ *be a Banach space. Then*

$$\|P_{rad}(x) - P_{rad}(y)\| \leq \frac{4}{\|x\| + \|y\|} \|x - y\| \text{ for all } x, y \in \mathcal{X}\,. \tag{4.10}$$

Proof:
Let $x, y \in \mathcal{X}$.
First case: $x, y \notin \overline{B}_1$. Then (4.10) follows from Corollary 4.31.
Second case: $x \in \overline{B}_1, y \notin \overline{B}_1$. Then

$$\begin{aligned} \|P_{rad}(x) - P_{rad}(y)\| &\leq \|x - \frac{y}{\|y\|} \leq \|x - y\| + \|y - \frac{y}{\|y\|}\| \\ &= \|x - y\| + \|y\| - 1 \leq \|x - y\| + \|y\| - \|x\| \leq 2\|x - y\| . \end{aligned}$$

■

From Theorem 4.33 we obtain a Lipschitz constant 2 for the radial projection. This constant cannot improved in general as the following example shows.

Example 4.34. *Consider* $\mathbb{R}^2$ *endowed with the* l_1*-norm. Take* $x = (0, 2)$ *and* $y := (\varepsilon, 2)$ *where* ε *is small. Then* $P_{rad}(x) = (0, 1), P_{rad}(y) = (\varepsilon, 1 - \varepsilon)$ *and*

$$2\varepsilon = \|P_{rad}(x) - P_{rad}(y)\| = 2\varepsilon = 2\|x - y\| .$$

□

4.5 Continuity of a single-valued metric projection

The considerations in the last chapters show that the metric projection operators have extremely good properties in Hilbert spaces, especially metric projections onto subspaces: they are linear orthogonal projectors, non-expansive, uniformly continuous, self-adjoint and idempotent. In the last section we have seen a big difference concerning metric projections in Hilbert spaces and Banach spaces: metric projections in a Banach space are not nonexpansive in general. Additionally, we have to realize that the metric projection is set-valued in general. What should be a definition of continuity of such a map?

Before we investigate the general case in the next section let us study continuity properties of single-valued metric projections in this section.

Theorem 4.35. *Let* $\mathcal{X}$ *be a a Banach space and let* C *be an approximately compact Chebyshev subset of* $\mathcal{X}$. *Then* P_C *is continuous.*

Proof:
Let $x \in \mathcal{X}$ and let $(x^n)_{n\in\mathbb{N}}$ with $x = \lim_n x^n$. Since C is a Chebyshev set there exist for every x^n a $u^n \in C$ with $\|u^n - x^n\| = \text{dist}(x^n, C)$; $u^n = P_C(x^n)$. Then $\lim_n \|u^n - x\| = \text{dist}(x, C)$ since $\text{dist}(\cdot, C)$ is continuous; see Lemma lem:nonexpdist. From the fact that C is approximately compact we obtain a subsequence $(u^{n_k})_{k\in\mathbb{N}}$ and $u \in \mathcal{X}$ with $w = \lim_k x^{n_k}$. Since C as a Chebyshev set is closed we have $u \in C$. Due to $\|u - x\| = \lim_k \|u^{n_k} - x\| = \text{dist}(x, C)$ we have $u \in P_C(x)$. Since C is a Chebyshev set we have $\{u\} = P_C(x)$. From this uniqueness we conclude that $(u^n)_{n\in\mathbb{N}}$ converges to $P_C(x)$. ■

Corollary 4.36. *Let* $\mathcal{X}$ *be a Banach space and let* C *be a compact Chebyshev subset of* $\mathcal{X}$. *Then* P_C *is continuous.*

Proof:
Follows from Theorem 4.35 since a compact set is approximately compact. ■

Corollary 4.37. *Let $\mathcal{X}$ be a finite-dimensional Banach space and let C be a Chebyshev subset of $\mathcal{X}$. Then P_C is continuous.*

Proof:
Let $x \in \mathcal{X}$. Then $P_C(x) \in C \cap \overline{B}_r(x)$ with $r > \operatorname{dist}(x, C)$. Now, $C \cap \overline{B}_r(x)$ is closed and bounded, therefore compact. Apply now Corollary 4.36. ■

Now, we come to continuity results which use geometric properties of the space under consideration. In the light of Corollary 4.36 it is natural to ask whether every Chebyshev set has a continuous metric projection mapping. But it is surprising that this question may fail even in a reflexive space. Unfortunately, finding an example of a Chebyshev set with a discontinuous metric projection mapping is not as straightforward as one might first imagine. Brown gave an example of cheyshevian subspace in a smooth reflexive Banach space with a discontinuous metric projection; see [8].[1]

Theorem 4.38. *Let $\mathcal{X}$ be a strictly convex Banach space and let C be a nonempty compact convex subset of $\mathcal{X}$. Then C is a Chebyshev set and $P_C : X \longrightarrow C$ is continuous.*

Proof:
Clearly, C is closed and a Chebyshev set.
Let $x \in \mathcal{X}$ and let $(x^n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{X}$ with $\lim_n x^n = x$. Then the sequence $(P_C(x^n))_{n\in\mathbb{N}}$ belongs to the compact set C. Therefore, there exists a subsequence $(P_C(x^{n_k}))_{k\in\mathbb{N}}$ and $z \in C$ with $\lim_k P(x^{n_k}) = z$.
Let $u \in C$. Then $\|x^{n_k} - P_C(x^{n_k})\| \leq \|x^{n_k} - u\|$ and therefore $\|x - z\| \leq \|x - u\|$ since $\lim_k (x^{n_k}, P_C(x^{n_k})) = (x, z)$. This implies $z \in P_C(x)$ and C is a Chebyshev set $z = P_C(x)$. ■

Theorem 4.39. *Let $\mathcal{X}$ be an E-space and let C be a nonempty closed convex subset. Then the metric pojection $P_C : X \longrightarrow C$ is continuous.*

Proof: Let $x \in \mathcal{X}$ and let $x = \lim_n x^n$. Then

$$\|P_C(x^n) - x\| \leq \|P_C(x^n) - x^n\| + \|x^n - x\| \leq \|P_C(x) - x^n\| + \|x^n - x\| .$$

Therefore, $\lim_n \|P_C(x^n) - x\| = \operatorname{dist}(x, C)$. Then $\lim_n P_C(x^n) = P_C(x)$; see Corollary 2.20. ■

There is no hope that strengthening the E-property to uniformly convexity would result in the uniform continuity of the metric projection since this fails even in the finite-dimensional case.

Corollary 4.40. *Let $\mathcal{X}$ be a uniformly convex Banach space and let C be a nonempty closed convex subset. Then the metric pojection $P_C : X \longrightarrow C$ is continuous.*

Proof:
This is a consequence of Theorem 4.39 and Theorem 4.4. ■

[1] Another example of a discontinuous metric projection is due to Kripke (unpublished).

Theorem 4.41. *Let $\mathcal{X}$ be a strictly convex and reflexive Banach space and let C be a nonempty convex closed subset of $\mathcal{X}$. Then C is a Chebyshev set and the metric projection $\mathsf{P}_{\mathsf{C}} : \mathcal{X} \longrightarrow \mathcal{X}$ is continuous.*

Proof:
See [27]. ■

Theorem 4.42. *Let $\mathcal{X}$ be a reflexive Banach space with a Kadek-Klee norm and let C be a weakly closed Chebyshev subset of $\mathcal{X}$. Then the metric projection P_{C} is continuous.*

Proof:
Let $x \in \mathcal{X}$. Let $(x^n)_{n\in\mathbb{N}}$ be a sequence with $\lim_n x^n = x$. Since P_{C} is locally bounded (see Lemma 4.63) and $\mathcal{X}$ is reflexive, there exists a subsequence $(x^{n_k})_{k\in\mathbb{N}}$ and a point $y \in \mathsf{C}$ such that $(\mathsf{P}_{\mathsf{C}}(x^{n_k}))_{k\in\mathbb{N}}$ converges to y with respect to the weak topology. Since C is weakly closed, $y \in \mathsf{C}$. Therefeore,

$$\begin{aligned} \operatorname{dist}(x, \mathsf{C}) \;\le\; \|x - y\| \;&\le\; \liminf_k \|x^{n_k} - \mathsf{P}_{\mathsf{C}}(x^{n_k})\| \\ &\le\; \limsup_k \|x^{n_k} - \mathsf{P}_{\mathsf{C}}(x^{n_k})\| \;=\; \lim_k \operatorname{dist}(x^{n_k}, \mathsf{C}) \;=\; \operatorname{dist}(x, \mathsf{C}) \end{aligned}$$

since $(x^{n_k} - \mathsf{P}_{\mathsf{C}}(x^{n_k}))_{n\in\mathbb{N}}$ converges to $x - y$ with respect to the weak topology on $\mathcal{X}$ and the norm in $\mathcal{X}$ is weakly lower semicontinuous. Now, because C is a Chebyshev set it follows that $y = \mathsf{P}_{\mathsf{C}}(x)$. Furthermore, due to Kadek-Klee property $(x^{n_k} - \mathsf{P}_{\mathsf{C}}(x^{n_k}))_{n\in\mathbb{N}}$ converges in norm to $x - y = x - \mathsf{P}_{\mathsf{C}}(x)$. Therefore, $\mathsf{P}_{\mathsf{C}}(x^{n_k}))_{n\in\mathbb{N}}$ converges in norm to $\mathsf{P}_{\mathsf{C}}(x)$. Due to this uniqueness of the cluster point of $\mathsf{P}_{\mathsf{C}}(x^n))_{n\in\mathbb{N}}$ we conclude that $(\mathsf{P}_{\mathsf{C}}(x^n))_{n\in\mathbb{N}}$ converges in norm to $\mathsf{P}_{\mathsf{C}}(x)$. ■

Corollary 4.43. *Let $\mathcal{X}$ be a uniformly convex Banach space and let C be a weakly closed Chebyshev subset of $\mathcal{X}$. Then the metric projection P_{C} is continuous.*

Proof:
A uniformly convex Banach space has a Kadek-Klee norm. ■

A famous problem in approximation theory is the convexity of Chebyshev sets. We have seen that a Banach space $\mathcal{X}$ has the property that every closed convex subset C is a Chebyshev set if and only if $\mathcal{X}$ is reflexive and strictly convex. It is natural to ask what are the Banach spaces $\mathcal{X}$ in which the converse property holds, i.e. in which every Chebyshev set C is convex. This problem has been solved for 2- and 3-dimensional spaces $\mathcal{X}$. For Banach spaces of finite dimension ≥ 4 only sufficient conditions are known. For infinite-dimensional Banach spaces $\mathcal{X}$ the problem is considerable more difficult, even the answer to the following problem beeing unknown:

In a Hilbert space, is every Chebyshev set necessarily convex?

At the beginning of the analysis of this problem the approach was to add a condition on the set itself in order to derive convexity. An important step was the idea of Klee of imposing continuity conditions on the metric projection. It was quickly realized that continuity properties of the metric projection play a key role in determining the convexity of Chebyshev sets. Here is the most important result in this concept.

Theorem 4.44 (Vlasov)**.** *Let $\mathcal{X}$ be a Banach space and let $C \subset \mathcal{X}$ be a Chebyshev set. If X^* is strictly convex and P_C is continuous then C is convex.*

Proof:
[40].

■

Corollary 4.45. *If C is a Chebyshev set in a Hilbert space with a continuous metric projection then C is convex.*

Proof:
Follows from Theorem 4.44. ■

Corollary 4.46. *If C is a weakly closed Chebyshev set in a smooth uniformly convex Banach space $\mathcal{X}$ then C is convex.*

Proof:
Follows from ■

Lemma 4.47. *Let C be a nonempty weakly closed subset of a Hilbert space $\mathcal{H}$. Then the metric projection P_C is single-valued and continuous.*

Proof:

■

Theorem 4.48. *Let C be a nonempty weakly closed subset of a Hilbert space $\mathcal{H}$. Then the following conditions are equivalent:*

(a) C *is a Chebyshev set.*

(b) C *is a convex set.*

Proof:
Ad $(a) \implies (b)$ Consider the function

$$f : \mathcal{H} \ni x \longmapsto \frac{1}{2}\|x\|^2 + \delta_C(x) \in \mathbb{R} \cup \{\infty\}.$$

Notice that f^* is convex. Then for $y \in \mathcal{H}$

$$\begin{aligned}
f^*(y) &= \sup_{u \in C}(\langle y|u\rangle - \frac{1}{2}\|y\|^2) \\
&= \frac{1}{2}\langle y|y\rangle + \frac{1}{2}{}_{u\in C}(-\frac{1}{2}\langle y|y\rangle + 2\langle y|u\rangle - \langle u|u\rangle) \\
&= \frac{1}{2}\|y\|^2 - \frac{1}{2}\inf_{u\in C}\|y-u\|^2 \;=\; \frac{1}{2}\|y\|^2 - \frac{1}{2}\mathrm{dist}(y, C)^2 \\
&= \frac{1}{2}\|y\|^2 - \frac{1}{2}\|y - P_C(y)\|^2 \;=\; \langle y|P_C(y)\rangle - \frac{1}{2}\|P_C(y)\|^2 \\
&= \langle y|P_C(y)\rangle - f(P_C(y))
\end{aligned}$$

This shows $\mathrm{dom}(f^*) = \mathcal{H}$ and $f^*(y) = \langle y|P_C(y)\rangle - f(P_C(y)), y \in \mathcal{H}$. We conclude $P_C(y) \in \partial f^*(y)$ for all $y \in \mathcal{H}$. Now suppose $z \in \partial f^*(y)$. Set $y_n := y + \frac{1}{n}(z - P_C(y))$. Then $\lim_n y_n = y$ and hence $\lim_n P_C(y_n) = P_C(y)$ by Lemma 4.47. From

$$0 \leq \langle y_n - y|P_C(y_n) - z\rangle = \frac{1}{n}\langle z - P_C(y)|P_C(y_n) - z\langle\,,\, n \in \mathbb{N}\,,$$

which is a consequence of the defining property of a subdifferential we conclude

$$0 \leq \liminf_n \langle z - P_C(y)|P_C(y_n) - z\rangle = -\|z - P_C(y)\|^2\,.$$

Consequentely, $z = P_C(y)$. This shows $\partial f^*(y) = \{P_C(y)\}$ for all $y \in \mathcal{H}$. Since f^* is continuous and $\partial f^*(y)$ is a singleton f^* is Gateaux differentiable. Now, $-\infty < f^{**}(x) \leq f(x), x \in Hs$. Moreover, f is weakly lower semicontinuous since C is weakly closed. This implies that f is convex and hence $\mathrm{dom}(f) = C$ is convex.
Ad $(b) \implies (a)$ We may assume without loss of generality, $\theta \in C$. Choose any $x \in \mathcal{H}$ and consider the functions

$$f : \mathcal{H} \ni z \longmapsto -\langle x|z\rangle + \delta_{\overline{B}_1}(z) \in \mathbb{R} \cup \{\infty\},\ g : \mathcal{H} \ni z \longmapsto \sup_{u \in C}(z) \in \mathbb{R} \cup \{\infty\}.$$

Since f is continuous at $\theta \in \mathrm{dom}(f) \cap \mathrm{dom}(g)$ we have no duality gap and we know

$$p := \inf_{z \in \mathcal{H}}(f(z) + g(z)) = d := \max_{y \in \mathcal{H}}(-f^*(y) - g^*(-y)) < \infty\,.$$

Since $f^*(y) = \|x + y\|, g^*(y) = \delta_C(y), y \in \mathcal{H}$, and since C is closed, we obtain

$$d = \max_{y \in \mathcal{H}} -\|x + y\| - \delta_C(-y)) = -\mathrm{dist}(x, y)\,.$$

Choose any $u \in C$. Observe that $0 \leq \mathrm{dist}(x, C) \leq \|x + y\|$. Therefore, $P_C(x)$ must be nonempty. Uniqueness follows from the convexity of C. ■

Remark 4.49. *If C is a weakly closed Chebyshev set in a smooth uniformly convex Banach space $\mathcal{X}$ then the metric projection is continuous. As a consequence, a nonempty weakly closed subset is then convex if and only if C is a Chebyshev set.* □

Corollary 4.50. *If C is a weakly closed Chebyshev set in a Hilbert space then C is convex.*

Proof:

■

Theorem 4.51 (Asplund,). *Let $\mathcal{X}$ be a Hilbert space and let $C \subset \mathcal{X}$ be a Chebyshev set. If $P_C : \mathcal{H} \longrightarrow \mathcal{H}$ is continuous when $\mathcal{H}$ is endowed with the norm-topology in the domain and $\mathcal{H}$ is endowed with the weak topology in the range of the mapping then C is convex.*

Proof:
[?].

■

Remark 4.52. *The metric projection* $P := P_U$ *onto a closed linear subspace* U *of a Hilbert space* $\mathcal{H}$ *is continuous and it has the the property* $P \circ P = P$ *(Idempotencity). Moreover, it is continuous and nonexpansive. In a Banach space* $\mathcal{X}$ *it is not clear that a projection* P *onto a closed subspace is continuous. A necessary und sufficient condition is that the subspace* U *is complemented. This means that there exists a closed subspace* V *such that* $X = U \oplus V$. *But it is known that not each closed subspace is complemented.* □

Let us add a surprising result concerning the linearity of a metric projection. We need two definitions.

Definition 4.53. *Let* $\mathcal{X}$ *be a Banach space and let* U *be a proximal subspace of* $\mathcal{X}$.

(1) $U^0 := \{x \in \mathcal{X} : \theta \in P_U(x)\}$ *is called the kernel of* P_U.

(2) $\|P_U\| := \sup\{\|y\| : y \in P_U(x), x \in \overline{S}_1\}$.

□

Let $\mathcal{X}$, as in Definition 4.53. Obviously, $U^0 = \{x \in \mathcal{X} : \|x\| = \operatorname{dist}(x, U)$. Moreover, $\|P_U\| \in [1, 2]$ since $P_U|_U = I$, and $\|P_U(x)\| \leq \|P_U(x) - x\| + \|x\| \leq 2\|x\|, x \in \mathcal{X}$.

Lemma 4.54 (Holmes, Kripke, 1968)**.** *Let* $\mathcal{X}$ *be a Banach space and let* U *be a Chebyshev subspace of* $\mathcal{X}$.

(1) If P_U *is continuous then* $I + P_U$ *is a homoemorphism of* $\mathcal{X}$ *into itself.*

(2) If P_U *is linear, then* $\|P_U\| = 1$ *if and only if* U^0 *is a proximal subspace and* $x - P_U(x) \in P_{U^0}(x)$ *for each* $x \in \mathcal{X}$.

(3) If P_U *is linear and* U^0 *is a Chebyshev subspace, then the following statements are equivalent:*

(i) $\|P_U\| = 1$.

(ii) $P_{U^0} = I - P_U$.

(iii) $(U^0)^0 = U$.

Proof:
See [23] for a proof. ■

As we know, in a Hilbert space the metric projection onto a closed linear subspace is always linear and has norm one. In general normed spaces however, with the exception of Chebyshev subspaces of codimension one, Chebyshev subspaces having linear metric projections are relatively scarce. For example, the space $C[0, 1]$ of continuous functions endowed with the supremum norm has none with finite dimension; see [9]. Deutsch and Lambert (see [15]) gave for each $r \in [0, 2]$ an example of a one-dimensional subspace U in $\mathbb{R}^2$, endowed with the l_1-norm which has the properties: U, U^0 are Chebyshev subspaces, P_U, P_{U^0} are linear, $\|P_U\| = r$.

4.6 Continuity properties of a multi-valued metric projections

Let $\mathcal{X}$ be a Banach space and let C be a proximal subset of $\mathcal{X}$. Then the metric projection is a mapping $P_C : \mathcal{X} \rightrightarrows \mathcal{X}$ with nonempty closed convex subsets $P_C(x), x \in \mathcal{X}$; clearly, $\#P_C(x) = 1$ for $x \in C$. To study continuity propertis of this mapping one has to introduce concepts of neighborhoods and/or convergence of families of sets.

Let X, Y be a topological spaces with the system $\mathcal{T}_{\mathcal{X}}, \mathcal{T}_{\mathcal{Y}}$ of open sets. Let $G : X \rightrightarrows Y$. The most important continuity concept for set-valued mapping is that of lower semi-continuity since in this concept under some additional circumstances Michael's selection theorem may be used to obtain a continuous mapping $S : X \longrightarrow Y$ such that $S(x) \in G(x)$ for all $x \in X$. For example, the radial projection is a continuous selection for the metric projection onto the closed unit ball of Banach space; see Section 4.4.

Definition 4.55. *Let X, Y be a topological spaces with the system $\mathcal{T}_{\mathcal{X}}, \mathcal{T}_{\mathcal{Y}}$ of open sets. Let $G : X \rightrightarrows Y$.*

(a) G *is called* **upper semicontinuous** *if for every open set W in Y the set $\{x \in X : G(x) \subset W\}$ is open in X.*

(b) G *is called* **lower semicontinuous** *if for every open set W in Y the set $\{x \in X : G(x) \cap W\}$ is open in X.*

(c) G *is* **continuous** *if it is both upper semicontinuous and lower semicontinuous.*

(d) $dom(G) := \{x \in X : G(x) \neq \emptyset\}$.

(e) $graph(G) := \{(x, y) \in X \times Y : y \in G(x)\}$ *is called the* **graph** *of* G.

□

A **paracompact topological space** X is defined as follows: Every open covering of X has an open refinement $\mathcal{U} = \cup_{i=1}^{\infty} U_i$ where each U_i is a locally finite collection of open subsets of X. Every compact topological space, every metric space is paracompact.

Theorem 4.56 (Michael, 1953)**.** *Let X be a paracompact topological Hausdorff space and let $F : X \rightrightarrows X$ be a lower semicontinuous mapping with nonempty closed convex subsets $F(x), x \in X$. Then there exists a continuous mapping $f : X \longrightarrow X$ with $f(x) \in F(x)$ for all $x \in X$.*

Let us collect some topological concepts for the convergence of families of sets.

Definition 4.57. *Let $\mathcal{X}$ be a Banach space and let $(C_n)_{n\in\mathbb{N}}$ be sequence of sets.*

(a) $(C_n)_{n\in\mathbb{N}}$ *is said to be* **Hausdorff-convergent** *to a set C if*

(i) for each $\varepsilon > 0$ there exists $N \in \mathbb{N}$ with $C_n \subset \{x \in \mathcal{X} : dist(x, A) \leq \varepsilon\}$ for all $n \geq N$;

(ii) for each $\varepsilon > 0$ *there exists* $N \in \mathbb{N}$ *with* $C \subset \{x \in \mathcal{X} : dist(x, A_n) \leq \varepsilon\}$ *for all* $n \geq N$.

We will write $\lim_H C_n = C$.

(b) $(C_n)_{n\in\mathbb{N}}$ *is said to* **converge Wijsman** *to a set* C *if*

$$\lim_n dist(x, C_n) = dist(x, C) \text{ for all } x \in \mathcal{X}.$$

We will write $\lim_W C_n = C$.[2]

(c) $(C_n)_{n\in\mathbb{N}}$ *is said to* **converge Mosco** *to a set* C *if*

(i) for each $x \in C$ *there exist* $N \in \mathbb{N}$ *such that for each* $n \geq N$ *a* $x \in C_n$ *exists such that* $(x^n)_{n\in\mathbb{N}}$ *converges to* x;

(ii) if there exists a subsequence $(n_k)_{k\in\mathbb{N}}$ *and an associated sequence* $(x^{n_k})_{k\in\mathbb{N}}$ *with* $x^{n_k} \in C_{n_k}, k \in \mathbb{N}$, *such that* $(x^{n_k})_{k\in\mathbb{N}}$ *converges weakly to a point* $x \in \mathcal{X}$ *then* $x \in C$.

We will write $\lim_M C_n = C$.

(d) $(C_n)_{n\in\mathbb{N}}$ *is said to be* **slice-convergent** *to a set* C *if*

$$dist(B, C_n) \text{ converges to } dist(B, C) \text{ for every bounded convex set } B.$$

We will write $\lim_S C_n = C$.

□

It is appearent that Wijsman convergence depends on the precise norm used, while Mosco convergence is preserved by equivalent renorming.

Lemma 4.58. *Let* $\mathcal{X}$ *be a Banach space and let* $(C_n)_{n\in\mathbb{N}}$ *be sequence of sets.*

(1) If $\lim_W C_n = \emptyset$ *then* $\lim_M C_n = \emptyset$.

(2) If $\mathcal{X}$ *is reflexive and if* $\lim_M C_n = \emptyset$ *then* $\lim_W C_n = \emptyset$.

(3) If $\mathcal{X}$ *is reflexive and if* $\lim_M C_n = C$ *then* $\lim_W C_n = C$.

Proof:

■

Remark 4.59. *If* $\mathcal{X}$ *is nt reflexive then there exists a sequence* $(C_n)_{n\in\mathbb{N}}$ *of compact convex sets such that* $\lim_M C_n$ *exists and is a nonempty compact convex set, but* $\lim_W C_n$ *fails to exist. Thus, a Banach space is reflexive if and only if Mosco convergence implies Wijsman convergence of closed convex subsets of* $\mathcal{X}$; *see [?].* □

Lemma 4.60. *Let* $\mathcal{X}$ *be a Banach space and let* $(C_n)_{n\in\mathbb{N}}$ *be sequence of closed sets and let* $(r_k)_{k\in\mathbb{N}}$ *an increasing sequence of positive real numbers. Assume that* $D_k := \lim_M(C_n \cap B_{r_k})$ *exists for all* $k \in \mathbb{N}$. *Then* $C := \lim_M C_n$ *exists and* $C = \cup_{k\in\mathbb{N}} D_k$. *Moreover,* C *is closed.*

[2] We use the convention $\text{dist}(x, \emptyset) = \infty$.

Proof:

■

There is a connection to the Vietoris-topology and the topology induced by the Hausdorff metric in the case of closed-valued/compact-valued mappings.

Lemma 4.61. *Let $\mathcal{X}$ be a Banach space and let C be a nonempty closed subset of $\mathcal{X}$. If $(x^n, y^n)_{n\in\mathbb{N}}$ is a sequence in $\mathcal{X} \times \mathcal{X}$ with $y^n \in P_C(x^n), n \in \mathbb{N}$, and $\lim_n(x^n, y^n) = (x, y)$ then $y \in P_C(x)$.*

Proof:
Consider the sequence as above. Firstly, $y \in C$ since C is closed. Furthermore,

$$\|x - y\| = \lim_n \|x^n - y^n\| = \lim_n \operatorname{dist}(x^n, C) = d(x, C)$$

since $\operatorname{dist}(\cdot, C)$ is continuous due to Lemma 2.1. Thus, $y \in P_C(x)$. ■

Definition 4.62. *Let $\mathcal{X}, \mathcal{Y}$ be Banach spaces and let $\Phi : \mathcal{X} \rightrightarrows \mathcal{Y}$. Φ is called* **locally bounded** *if for every $x \in X$ there exist $r > 0, R > 0$ such that $\Phi(B_r(x)) \subset B_R$.* □

Lemma 4.63. *Let $\mathcal{X}$ be a Banach space and let C be a nonempty subset of $\mathcal{X}$. Then the metrix projection mapping $P_C : \mathcal{X} \rightrightarrows \mathcal{X}$ is locally bounded.*

Proof:
Let $x \in \mathcal{X}$ and let $r := 1, R := \operatorname{dist}(x, C) + \|x\| + 2$. Suppose $y \in P_C(B_r(x))$, i.e. $y \in P_C(u)$ for some $u \in B_r(x)$. Using the nonexpansivity of the distance function we obtain

$$\begin{aligned}
\|y\| &\leq \|y - u\| + \|u - x\| + \|x\| \\
&= \operatorname{dist}(u, C) + \|u - x\| + \|x\| \\
&\leq \operatorname{dist}(x, C) + \|u - x\| + \|u - x\| + \|x\| \\
&\leq \operatorname{dist}(x, C) + 2 + \|x\| \;=\; R.
\end{aligned}$$

Hence, $P_C(B_r(x)) \subset B_R$. ■

Definition 4.64. *Let $\mathcal{X}, \mathcal{Y}$ be Banach spaces and let $\Phi : \mathcal{X} \rightrightarrows \mathcal{Y}$. Φ is called* **continuous** *at $x \in \mathcal{X}$ if $\Phi(x)$ is a singleton and for any sequence $(x^n, y^n)_{n\in\mathbb{N}}$ in the graph of Φ with $\lim_n x^n = x$ we have $\lim_n y^n = z$ with $z \in \Phi(x)$.* □

Theorem 4.65. *Let $\mathcal{X}$ be a Banach space and let C be a boundedly compact subset of $\mathcal{X}$. If $P_C(x)$ is a singleton for $x \in \mathcal{X}$, then the metric projection mapping P_C is continuous at x.*

Proof:
Let $x \in \mathcal{X}$ and let $P_C(x) = \{z\}$. Let $(x^n, y^n)_{n\in\mathbb{N}}$ be a sequence in the graph of P_C with $\lim_n x^n = x$. Since P_C is locally bounded and since C is boundedly compact there exists a subsequence $(x^{n_k}, y^{n_k})_{k\in\mathbb{N}}$ such that $\lim_k y^{n_k} = y$ for some $y \in C$. Since $\lim_k(x^{n_k}, y^{n_k}) = (x, y)$ we conclude by Lemma 4.61 $y \in P_C(x)$. Thus, $y = z$ as required. ■

4.7 Differentiability properties of the distance mapping

Let $\mathcal{X}$ be a Banach space and let C be a nonempty closed subset of $\mathcal{X}$. We know that the distance mapping

$$\mathcal{X} \ni x \longmapsto \operatorname{dist}(x, C) := \inf_{u \in C} \|x - u\| \in \mathbb{R}$$

is nonexpansive. As a consequence, the differential quotients

$$\frac{\operatorname{dist}(x + th, C) - \operatorname{dist}(x, C)}{t}, \; t \neq 0, x \in \mathcal{X} \backslash \{\theta\}, h \in \mathcal{X},$$

are bounded by 1. Therfore, one should ask for differentiability properties of the distance mapping.

Theorem 4.66. *Let $\mathcal{H}$ be a Hilbert space and let C is Chebyshev subset of $\mathcal{H}$. Then the distance function dist$(\cdot, C)$ is Fréchet differentiable and $\|D\,dist(\cdot, C)(x)\| = 1$.*

Proof:

■

Let $f : \mathcal{X} \longrightarrow \mathbb{R}$ be a Lipschitz continuous function. **Clarke's generalized directional derivative** of f at a point x in the direction $h \in \mathcal{X}$, denoted by $f^\circ(x; h)$, is given by :

$$f^\circ(x; h) := \limsup_{z \to x, t \downarrow 0} \frac{f(z + th) - f(z)}{t}. \tag{4.11}$$

Clarke's generalized subdifferential of f at x is given by

$$\partial f(x) := \{\lambda \in X^* : f^\circ(x; h) \geq f(y) \text{ for all } y \in \mathcal{X}\}. \tag{4.12}$$

Theorem 4.67. *Let $\mathcal{X}$ be a Banach space such that X^* is smooth. Then a nonempty closed subset C of $\mathcal{X}$ is convex if its distance function dist$(\cdot, C)$ satisfies*

$$\limsup_{y \to \theta} \frac{dist(x + y, C) - dist(x, C)}{\|y\|} = 1 \text{ for all } x \in \mathcal{X} \backslash C. \tag{4.13}$$

In particular, this differentiability condition is satisfied if dist$(\cdot, C)$ is Gateaux-differentiable at x and dist$(x, C) \in \overline{S}_1$ of the dual space X^*.

Proof:
See [5]. ■

Theorem 4.68. *Let $\mathcal{X}$ be a Banach space such that X^* is strictly convex. Let $C \subset \mathcal{X}$ be a Chebyshev set, let $x \in C$ such that $\partial\, dist(x, C)$ is a singleton. Then the following assertions are equivalent:*

(a) C *is convex.*

(b) *dist$(\cdot, C)$ is convex.*

(c) *dist$(\cdot, C)$ is Gateaux differentiable at x.*

(d) There exists $h \in \overline{S}_1$ *such that* $\lim_{t\downarrow 0} \dfrac{dist(x+th)-dist(x)}{t} = 1$.

(e) $\limsup_{y\to\theta} \dfrac{dist(x+y)-dist(x)}{\|y\|} = 1$.

Proof:
$(a) \implies (b)$ Since C is convex it can easily be seen that $\mathrm{dist}(\cdot, C)$ is convex.
$(b) \implies (c)$ Since $\mathrm{dist}(\cdot, C)$ is convex and continuous at x and $\partial\mathrm{dist}(\cdot, C)$ is a singleton, $\mathrm{dist}(\cdot, C)$ is Gateaux differentiable at x and $\partial\mathrm{dist}(\cdot, C)(x) = \{d'_C(x)\}$.
$(c) \implies (d)$ Since C is a chebyshev set we have $\|x - P_C(x)\| = \mathrm{dist}(x, C)$. It follows from Gateaux differentiability of $\mathrm{dist}(\cdot, C)$ that

$$\liminf_{t\downarrow 0} \frac{\mathrm{dist}(x+th, C) - \mathrm{dist}(x, C)}{t}$$

exists for every $y \in \mathcal{X}$. For each $t > 0$ we have

$$\mathrm{dist}(x + t(x - P_C(x), C) - \mathrm{dist}(x, C) \leq t\mathrm{dist}(x, C).$$

Hence for $y := x - P_C(x)$ we set

$$\liminf_{t\downarrow 0} \frac{\mathrm{dist}(x + t(x - P_C(x), C) - \mathrm{dist}(x, C)}{t} = \mathrm{dist}(x, C).$$

Since $x \in \mathcal{X}\backslash C$, $\mathrm{dist}(x, C) > 0$ and if $t' = t/\mathrm{dist}(x, C)$ as $t \downarrow 0$. Then by the above

$$\liminf_{t'\downarrow 0} \frac{\mathrm{dist}(x + t'(x - P_C(x), C) - \mathrm{dist}(x, C)}{t} = \mathrm{dist}(x, C).$$

If now $z = (x - P_C(x))\|x - P_C(x)\|^{-1}$ then we have

$$\liminf_{t\downarrow 0} \frac{\mathrm{dist}(x + tz, C) - \mathrm{dist}(x, C)}{t} = 1.$$

On the other hand, $\mathrm{dist}(\cdot, C))$ is a Lipschitz function and so

$$\limsup_{t\downarrow 0} \frac{\mathrm{dist}(x + y, C) - \mathrm{dist}(x, C)}{\|y\|} = 1.$$

$(d) \implies (e)$ Since $\mathrm{dist}(\cdot, C)$ is a Lipschitz function

$$\limsup_{t\downarrow 0} \frac{\mathrm{dist}(x + y, C) - \mathrm{dist}(x, C)}{\|y\|} \leq 1.$$

On the other hand, for each $v \in \overline{S}_1$

$$\lim_{t\downarrow 0} \frac{\mathrm{dist}(x + tv, C) - \mathrm{dist}(x, C)}{t} \leq \limsup_{y\to\theta} \frac{\mathrm{dist}(x + y, C) - \mathrm{dist}(x, C)}{\|y\|},$$

in particular for $v := z$ in (d), we have

$$1 \leq \limsup_{y\to\theta} \frac{\mathrm{dist}(x + y, C) - \mathrm{dist}(x, C)}{\|y\|}.$$

$(e) \implies (a)$ Follows from ■

Theorem 4.69. *Let $\mathcal{X}$ be a Banach space such that X^* is smooth. Then a nonempty closed subset C of $\mathcal{X}$ is convex if for some $r > 0$ the set*

$$P_{C,r} := \{x^0 - rz : x^0 \in P_C(C), p(x^0) \in C, \|x^0 - p(x^0)\| = dist(x^0, C) > r, z = (x^0 - p(x^0))\|x^0 - p(x^0)\|^{-1}\}$$

is dense in $\mathcal{X}\backslash C$.

Proof:
See [18, 19]. ■

Lemma 4.70. *Let $\mathcal{X}$ be a Banach space and let C be a nonempty subset of $\mathcal{X}$. If the distance mapping $dist(\cdot, C)$ is Gateaux-differentiable in $x \in \mathcal{X}\backslash\overline{C}$ and $w \in P_C(x)$ then*

$$d\,dist(x)(x - w) = \|x - w\| \text{ and } \|d\,dist(x)\| = 1\,. \tag{4.14}$$

Proof:
See [17]

■

4.8 Differentiability properties of the metric projections

Let $\mathcal{H}$ be a Hilbert space and let C be a nonempty closed convex subset o $\mathcal{H}$. The the metric projection $_C$ is welldefined and one may ask for differentiability properties of P_C.

We say that P_C is **directionally differentiable** at a point $x \in \mathcal{H}$ if

$$d_+P_C(x)(h) := \lim_{t\downarrow 0} \frac{P_C(x + th) - P_C(x)}{t} \tag{4.15}$$

exists for each $h \in \mathcal{H}$. If the mapping $h \longmapsto d_+P_C(x)$ is linear and continuous then P_C is called **Gateaux-differentiable** at $x \in \mathcal{H}$. This mapping, when it exists, is denoted by $dP_C(x)$. If the above limit exists uniformly for $h \in \overline{B}_1$ and if $h \longmapsto d_+P_C(x)$ is linear and continuous we say that P_C is **Fréchet-differentiable** at x and denote the corresponding linear operator by $DP_C(x)$.

Lemma 4.71. *Let $\mathcal{H}$ be a Hilbert space and let C is Chebyshev subset of $\mathcal{H}$. Then we define*

$$W : \mathcal{H}\backslash C \ni x \longmapsto \frac{1}{2}(\|x\|^2 - dist(x, C)^2 \in \mathbb{R}\,. \tag{4.16}$$

Then W is Fréchet differentiable and $DW(x) = P_C(x), x \in \mathcal{H}$. Moreover, W is convex and

$$W(x) = \sup_{u\in C} \frac{1}{2}(2\langle x|u\rangle - \|u\|^2)\,,\ x \in \mathcal{H}\backslash C\,. \tag{4.17}$$

Proof:

■

By the comments above, the metric projection $P_{T(x,C)}$ exists. Zarantanello [42] and Haraux [20] (see also [39]) have proved the following result.

Theorem 4.72. *Let $\mathcal{H}$ be a Hilbert space and let C be a nonempty closed convex subset of $\mathcal{H}$. Then for any $x \in C$ we have*

$$P_C(x+th) = x + tP_{T(x,C)}(h) + o(t)\,,\ t > 0, h \in \mathcal{H}\,. \tag{4.18}$$

Proof:
Let $x \in C$ and $h \in \mathcal{H}$. We have to prove

$$d_+P_C(x)(h) = \lim_{t\downarrow 0} \frac{P_C(x+th) - P_C(x)}{t} = P_{T(x,C)}(h)\,. \tag{4.19}$$

In a first step we prove for $t > 0$

$$\frac{P_C(x+th) - P_C(x)}{t} = P_{(C-x)/t}(h)\,. \tag{4.20}$$

We have for $t > 0$ and $y \in C$

$$\begin{aligned} 0 &\geq \langle (x+th) - P_C(x+th) | y - P_C(x+th)\rangle \\ &= \langle (x+th) - P_C(x+th) | y - (P_C(x+th) + x - x)\rangle \\ &= t^2\langle \frac{x - P_C(x+th)}{t} + h | \frac{y-x}{t} - \frac{P_C(x+th) - x}{t}\rangle \\ &= t^2\langle h - \frac{P_C(x+th) - x}{t} | \hat{y} - \frac{P_C(x+th) - x}{t}\rangle \end{aligned}$$

where $\hat{y} = (C - x)/t$.
In a second step we verify

$$\lim_{t\downarrow 0} \inf\{\|h - u\| : u \in (C-x)/t\} = \inf\{\|h - v\| : v \in T(x, C)\} \tag{4.21}$$

which means $\lim_{t\downarrow 0} \|h - P_{(C-x)/t}(h)\| = \|h - P_{T(x,C)}(h)\|$. Notice that the metric projections $P_{(C-x)/t}(h), P_{T(x,C)}(h)$ exist since $(C-x)/t, T(x, C)$ are closed convex sets.
If $0 \neq t_1 \leq t_2$ then $(C-x)/t_2 \subset (C-x)/t_1$ due to the fact that C is convex. This shows that $q := \lim_{t\downarrow 0} \|h - P_{(C-x)/t}(h)\|$ exist. Moreover since $(C-x)/t \subset T(x, C)$ for all $t > 0$ we have $q \geq \|h - P_{T(x,C)}(h)\|$. Since $P_{T(x,C)}(h) \in T(x, C)$ we obtain $q = \|h - P_{T(x,C)}(h)\|$. Now, we obtain

$$\begin{aligned} \|P_{(C-x)/t} - P_{T(x,C)}(h)\|^2 &= 2\|h - P_{(C-x)/t}\|^2 + 2\|h - P_{T(x,C)}\|^2 - 4\|h \\ &\quad -\frac{1}{2}(P_{(C-x)/t} + P_{T(x,C)})\|^2 \\ &\leq 2\|h - P_{(C-x)/t}\|^2 + 2\|h - P_{T(x,C)}\|^2 - 4\|h - P_{T(x,C)}\|^2 \end{aligned}$$

and we conclude that $\lim_{t\downarrow 0} \|P_{(C-x)/t} - P_{T(x,C)}(h)\|^2 = 0$ which implies finally

$$\lim_{t\downarrow 0} \frac{P_C(x+th) - P_C(x)}{t} = P_{T(x,C)}(h)\,.$$

■

We conclude from the theorem above that at each point $x \in C$, the map P_C has a directional derivative in every direction $h \in \mathcal{H}$ given by $P_{T(x,c)}(h)$. Since the metric

projection $P_{T(x,C)}$ will be linear only in the very special case when $T(X, C)$ is a linear subspace, it is clear that $dP_C(x)$ does not generally exist at points x of C. In particular, if $C \neq \mathcal{H}$ and C has nonempty interior and $x \in \partial C$, then $T(x, C)$ will not be a subspace, so $dP_C(x)$ will not exist.

Rademacher's theorem (see [36] and [43]) states that a Lipschitz continuous function from $\mathbb{R}^n$ into $\mathbb{R}^m$ is dfferentiable almost everywhere. The theorem of Alexandrov (see [1] and [31]) says that each continuous convex function on $\mathbb{R}^n$ is almost everywhere twice differentiable. For $n = 1$ this reduces to Lebesgue's theorem about the differentiability almost everywhere of a monotone function. Alexandrov's theorem exploits Asplund (see [2]) as follows:

Theorem 4.73 (Asplund, 1973). *Let C be a nonempty closed convex subset the inner product space $\mathcal{H} := \mathbb{R}^n$ and let P_C be the associated metric projection from $\mathcal{H}$ into C. Then P_C is differentable almost everywhere.*

Proof:
Define

$$f : \mathcal{H} \longrightarrow \mathbb{R}, x \longmapsto \sup_{y \in C}(\langle x|y\rangle - \frac{1}{2}\|y\|^2).$$

Obviously, f is convex, actually f is the Fenchel-conjugate of $j : \mathcal{H} \ni u \longmapsto \frac{1}{2}\|u\|^2 \in \mathbb{R}$. One can show that

$$f(x) = \frac{1}{2}\|x\|^2 - \inf_{y \in C} \frac{1}{2}\|x - y\|^2 = \frac{1}{2}\|x\|^2 - \frac{1}{2}\|x - P_C(x)\|^2, x \in \mathcal{H}.$$

From this we conclude that $P_C(x)$ is in the subdifferential of f at x; see the chapter concerning the fundamentals facts on convex analysis. Since we know from Alexandrov's theorem that f is twice (Fréchet-)differentiable almost everywhere the metric projection is almost everywhere (Fréchet-)differentiable. ■

Let C be a nonempty closed convex subset the inner product space $\mathcal{H} := \mathbb{R}^n$ and let P_C be the associated metric projection from $\mathcal{H}$ into C. It is known that if $x \in C$, then P_C is directionally differentiable at x and the corresponding directional derivatives

$$d_+P_C(x)(h) = \lim_{t \uparrow 0} \frac{P_C(x + th) - P_C(x)}{t}$$

are given by the metric projection of h onto the support cone $S_C(x)$ to C at x. In the case $x \notin C$ directional differentiability of $_C$ at x is guaranteed only under certain conditions on smoothness of the boundary of C. Here is an example in $\mathbb{R}^3$ (constructed by Shapiro in [38]) whose metric projection fails to be directionally differentiable; see also [33].

Example 4.74. *Consider a sequence $(\alpha_n)_{n \in \mathbb{N}}$ of positive numbers defined by $\lambda_1 := \pi/2$ and $\alpha_{n+1} := c^n\alpha_1, n \in \mathbb{N}$, with $c \in (0, 1)$. Then this sequence is monotone decreasing and $\lim_n \alpha_n = 0$. Let C be defined as the convex hull of $\{(\cos(\alpha_n), \sin(\alpha_n)) : n \in \mathbb{N}\} \cup \{(0, 0), (1, 0)\}$. Let $x := (2, 0), d := (0, 1)$. We will show that $P_C(x)$ is not directionally differentiable in x in direction d.*
Consider for $n = 2, 3, \ldots$

$$t_n := \sin(\alpha_n) + (2 - \cos(\alpha_n)) \tan(\frac{1}{2}(\alpha_+ \alpha_{n-1})), \; s_n := \sin(\alpha_n) + (2 - \cos(\alpha_n)) \tan(\frac{1}{2}(\alpha_+ \alpha_{n+1})).$$

Then we have $P_C(x) = (1, 0)$ *and*

$$P_C(x + t_n d) = P_C(x + s_n d) = (\cos(\alpha_n), \sin(\alpha_n)), \; n = 2, 3, \dots .$$

Since

$$\lim_n \frac{\sin(\alpha_n)}{t_n} = \frac{2}{3 + c^{-1}}, \; \lim_n \frac{\sin(\alpha_n)}{s_n} = \frac{2}{3 + c}$$

we obtain

$$\lim_n \frac{\pi_2(P_C(x + t_n d) - \pi_2(P_C(x))}{t_n} \neq \lim_n \frac{\pi_2(P_C(x + s_n d) - \pi_2(P_C(x))}{s_n},$$

where $\pi_2(u)$ *denotes the second coordinate* u_2 *of* $u = (u_1, u_2)$. □

4.9 Appendix: Subdifferential

Let X be a Banach space and let $f : X \longrightarrow \hat{\mathbb{R}}$ be proper, convex and closed. We consider the **epsilon-subdifferential**:

$$\partial_\varepsilon f(x^0) := \{\lambda \in X^* | f(y) \geq f(x^0) + \langle \lambda, y - x^0 \rangle - \varepsilon \text{ for all } y \in X\}, \; \varepsilon \geq 0, x^0 \in \mathrm{dom}(f). \quad (4.22)$$

For $\varepsilon = 0$ we write $\partial f(x^0) := \partial_0 f(x^0)$. The epsilon-subdifferential defines a point-to-set operator $\partial_\varepsilon f : X \ni x \longmapsto \partial_\varepsilon f(x) \in \mathrm{POT}(X^*)$; we use the notation from the theory of set-valued mappings:

$$\partial_\varepsilon f : X \rightrightarrows X^*$$

It can easily proved that for all $x \in X, \varepsilon \geq 0$ the set $\partial_\varepsilon f(x)$ is convex and weak*-closed (possible empty). Moreover, the following properties are useful:

$$\forall x \in X \forall \varepsilon_1, \varepsilon_2 \geq 0 (\varepsilon_1 \leq \varepsilon_2 \implies \partial_{\varepsilon_1} f(x) \subset \partial_{\varepsilon_2} f(x)) \quad (4.23)$$

$$\forall x \in X \forall \varepsilon \geq 0 \, (\partial_\varepsilon f(x) = \bigcap_{\eta > \varepsilon} \partial_\eta f(x)) \quad (4.24)$$

Theorem 4.75. *Let* X *be a Banach space, let* $f : X \longrightarrow \hat{\mathbb{R}}$ *be convex and suppose that* f *is continous in* $x^0 \in dom(f)$. *Then*

(1) $\partial f(x^0) \neq \emptyset$.

(2) $\partial f(x^0)$ *is convex and* $\sigma(X^*, X)$*-closed.*

(3) $\partial f(x^0)$ *is bounded (as a subset of* X^**).*

Notice hat the assumption that f is continuous in $x^0 \in \mathrm{dom}(f)$ cannot be replaced by the assumption that f is lower semicontinuous in $x^0 \in \mathrm{dom}(f)$ as wes see in the example

$$f : \mathbb{R} \ni x \longmapsto \begin{cases} -\sqrt{1 - x^2} & \text{if } |x| \leq 1 \\ \infty & \text{else} \end{cases} \in \mathbb{R} \cup \{\infty\}.$$

Corollary 4.76. *Let* X *be a Banach space, let* $f : X \longrightarrow \hat{\mathbb{R}}$ *be convex,* $U \subset dom(f)$ *convex and open, and suppose that* f *is bounded from above in a neighborhood* V *of a point* $x \in U$. *Then* $\partial f(x^0) \neq \emptyset$ *for all* $x^0 \in U$.

Theorem 4.77. *Let X be a Banach space and let $f : X \longrightarrow \hat{\mathbb{R}}$ be proper, convex and closed. If $x^0 \in dom(f), \varepsilon > 0$ then $\partial_\varepsilon f(x^0)$ is nonempty.*

Proof:
The set $A := \mathrm{epi}(f)$ is nonempty, convex and closed and we have:

$$(x^0, f(x^0) - \varepsilon) \notin \mathrm{epi}(f)\,,\ B := \{(x^0, f(x^0) - \varepsilon)\} \text{ convex, compact}$$

Therefore we may apply a separation theorem (see for instance [3]) to the sets A, B and obtain $\mu \in X^*, \alpha \in \mathbb{R}$ such that

$$\langle \mu, x^0\rangle + \alpha(f(x^0) - \varepsilon) < \langle \mu, y\rangle + \alpha t \text{ for all } y \in \mathrm{dom}(f), t \geq f(y)\,.$$

If we choose $y = x^0, t = f(x^0)$ we obtain $-\alpha\varepsilon < 0$ which implies $\alpha > 0$. With $\lambda := -1/\alpha\,\mu$ and $t = f(y)$ we have

$$-\langle \lambda, x^0\rangle + f(x^0) - \varepsilon < -\langle \lambda, y\rangle + f(y) \text{ for all } y \in X\,.$$

This shows $\lambda \in \partial_\varepsilon f(x^0)$. ■

Let X be a Banach space. We set $N : X \ni x \longmapsto \|x\| \in \mathbb{R}$. Then the subfifferential of the norm function N has the form

$$\partial N(x) = \begin{cases} \frac{x}{\|x\|} & \text{if } x \neq \theta \\ \overline{B}_1 & \text{if } x = \theta \end{cases} \tag{4.25}$$

Theorem 4.78 (Bronsted and Rockafellar, 1965)**.** *Let X be a Banach space and let $f : X \longrightarrow \hat{\mathbb{R}}$ be proper, convex and closed. If $x^0 \in dom(f), \varepsilon > 0, \gamma > 0$ and $\xi_0 \in \partial_\varepsilon f(x^0)$ then there exists $x \in dom(f), \xi \in X^*$ with*

$$\xi \in \partial f(x^0)\,,\ \|x - x^0\| \leq \varepsilon/\gamma\,,\ \|\xi - \xi_0\| \leq \gamma\,. \tag{4.26}$$

In particular, the domain $dom(\partial f)$ is dense in $dom(f)$.

Proof:
We have

$$\langle \xi_0, y - x^0\rangle \leq f(y) - f(x^0) + \varepsilon \text{ for all } y \in X\,.$$

We consider the mapping $g : X \ni y \longmapsto f(y) - \langle \xi_0, y\rangle \in (-\infty, \infty]$. Then g is proper, convex and closed and we have $g(x^0) \leq \inf_{x\in X} g(x) + \varepsilon$. Now, we may apply Ekeland's variational principle. This gives $z \in \mathrm{dom}(g) = \mathrm{dom}(f)$ such that

$$\|z - x^0\| \leq \varepsilon/\gamma, g(z) \leq g(y) + \gamma\|y - z\| \text{ for all } y \in X\,.$$

With the mapping $h : X \ni y \longmapsto \gamma\|y - z\| \in \mathbb{R}$ we have $z \in \mathrm{dom}(g) \cap \mathrm{dom}(h)$ and

$$\theta \in \partial(g + h)(z) = \partial g(z) + \partial h(z)$$

since g is convex and closed and h is continuous. Therefore, there exist $\mu\partial g(z), \eta \in \partial h(z)$ with $\theta = \mu + \eta$. Clearly, $\partial g(z) = \partial f(z) - \xi_0$. Moreover,

$$\partial h(z) = \{\zeta \in X^* | \|\zeta\| \leq \gamma\}$$

which is proved as follows.
If $\zeta \in \partial h(z)$ then

$$\langle \zeta, y - z \rangle \leq h(y) - h(z) = \gamma \|y - z\| \text{ for all } y \in X$$

and therefore $\|\zeta\| \leq \gamma$.
If $\zeta \in X^*$ with $\|\zeta\| \leq \gamma$ then $\zeta \in \partial h(z)$ since

$$\langle \zeta, y - z \rangle \leq \|\zeta\| \|y - z\| \leq \gamma \|y - z\| = h(y) - h(z) \text{ for all } y \in X.$$

Thus $\mu = \xi - \xi_0$ with $\xi \in \partial f(z)$ and $\mu = -\eta$, hence $\|\xi - \xi_0\| \leq \gamma$. Now, we have the result by choosing $x := z$.

The assertion concerning the domain $\mathrm{dom}(\partial f)$ is an easy consequence. ■

The connection of subdifferentiability and conjugation is content of the following theorem.

Theorem 4.79. *Let X be a Banach space, let $f : X \longrightarrow \hat{\mathbb{R}}$ be convex and let $x^0 \in dom(f), \lambda \in X^*$. Then the following conditions are equivalent:*

(a) $\lambda \in \partial f(x^0)$.

(b) $f(x^0) + f^*(\lambda) = \langle \lambda, x^0 \rangle$.

(c) $f(x^0) - \langle \lambda, x^0 \rangle \leq f(x) - \langle \lambda, x \rangle$ *for all* $x \in X$.

Theorem 4.80 (Moreau-Rockafellar, 1965)**.** *Let X be a Banach space, let $f, g : X \longrightarrow \hat{\mathbb{R}}$ convex and suppose that f or g is continuous in a point $z \in dom(f) \cap dom(g)$. Then we have for all $x^0 \in X$*

$$\partial(f + g)(x^0) = \partial f(x^0) + \partial g(x^0) \tag{4.27}$$

Corollary 4.81. *Let X be a Banach space and let $F_i : X \longrightarrow \hat{\mathbb{R}}$ be convex, $i = 1, \ldots, m$. We assume that $F_1, \ldots, F_{m-1}$ are continuous in a point in $z \bigcap_{i=1}^m dom(F_i)$. Then we have for all $x^0 \in X$*

$$\partial(\sum_{i=1}^m F_i)(x^0) = \sum_{i=1}^m \partial F_i(x^0)$$

Lemma 4.82. *Let X be a Banach space and let $f : X \longrightarrow \hat{\mathbb{R}}$ be convex. Then*

$$\partial(tf)(x^0) = t\partial f(x^0) \text{ for all } x^0 \in X, t > 0.$$

Theorem 4.83. *Let X, Y be Banach spaces, let $A : X \longrightarrow Y$ be a linear and continuous mapping ans let $f : Y \longrightarrow \hat{\mathbb{R}}$ be convex. Suppose that there exists a point $z \in X$ such that f is continuous in Az. Then*

$$\partial(f \circ A)(x^0) = A^* \partial f(Ax^0) \text{ for all } x^0 \in X. \tag{4.28}$$

Definition 4.84. *Let X be a Banach space, let $f : X \longrightarrow \hat{\mathbb{R}}$ and let $x^0 \in dom(f)$, $v \in X$. The limit*

$$f'_+(x^0, v) := \lim_{h \downarrow 0} \frac{1}{h}(f(x^0 + hv) - f(x^0))$$

(in $\bar{\mathbb{R}}$) is called the (onesided) **directional derivative** *of f in x^0 in direction v.* □

Notice that $f'_+(x^0,\cdot) : X \longrightarrow \bar{\mathbb{R}}$ is convex and positively homogeneous.

Theorem 4.85 (Moreau-Psenichnii, 1965)**.** *Let X be a Banach space, let $f : X \longrightarrow \hat{\mathbb{R}}$ be proper, convex and suppose that f is continuous in $x^0 \in dom(f)$. Then we have for all $v \in X$:*

(1) $f'_+(x^0, v) = \max_{\lambda\in\partial f(x^0)}\langle\lambda, v\rangle$.

(2) $-f'_+(x^0, -v) = \min_{\lambda\in\partial f(x^0)}\langle\lambda, v\rangle$.

Remark 4.86. *If we in Theorem 10.54 ommit the assumption that f is continuous one can prove the following result (see [4]):*

$$f'_+(x^0, v) = \inf_{\varepsilon\downarrow 0} \sup_{\lambda\in\partial_\varepsilon f(x^0)} \langle\lambda, v\rangle . \tag{4.29}$$

□

Definition 4.87. *Let X be a Banach space, let $f : X \longrightarrow \hat{\mathbb{R}}$ and let $x^0 \in dom(f)$. If the limit*

$$df(x^0)(h) := \lim_{t\to 0}(f(x^0 + th) - f(x^0))/t\,,\ h \in X$$

exists in $\mathbb{R}$ for all $h \in X$ and if $df(x^0) : X \longrightarrow \mathbb{R}$ is linear and bounded (i.e. $df(x^0) \in X^*$) then f is called **Gateaux-differentiable** *in x^0, and $df(x^0)$ is the* **Gateaux derivative** *in x^0.* □

Corollary 4.88. *Let X be a Banach space, let $f : X \longrightarrow \hat{\mathbb{R}}$ be convex and let $x^0 \in dom(f)$. Then the following conditions areequivalent:*

(a) *f is Gateaux-differentiable in x^0.*

(b) *$\partial f(x^0)$ is a singleton.*

Lemma 4.89. *Let $\mathcal{X}$ be a Banach space and let $j_p : X \ni x \longmapsto \frac{1}{p}\|x\|^p \in \mathbb{R}$ with $1 < p < \infty$. Then for each $x \in \|Xs$ holds:*

$$J_p(x) := \partial j_p(x) = \{\lambda \in X^* : \langle\lambda, x\rangle = \|x\|\|\lambda\|, \|\lambda\| = \|x\|^{p-1} . \tag{4.30}$$

Proof:
Siehe Kien ■

In the case $p = 2$ we use the short notation $J = J_2$.

Let us prove the **sandwich theorem**. Here we need

Definition 4.90. *Let $\mathcal{X}$ be a Banach space and let S be a subset of $\mathcal{X}$. We say $x \in core(S)$ if*

$$\mathcal{X} = \cup_{r>0} r(S - x) .$$

□

Theorem 4.91 (Sandwich Theorem). *Let $\mathcal{X}$ be a Banach space and let $f, -g \in \Gamma_0(\mathcal{X})$ and suppose*

$$f(x) \geq g(x) \text{ for all } x \in \mathcal{X}\,.$$

Assume that the following **constraint qualification (CQ)** *holds:*

$$\theta \in core(dom(f) - dom(-g))\,. \tag{4.31}$$

Then there is n affine function a such that

$$f(x) \geq a(x) \geq g(x) \text{ for all } x \in \mathcal{X}\,.$$

Proof:
Consider $h : \mathcal{X} \ni z \longmapsto \inf_{u \in \mathcal{X}}(f(u) - g(u - z)) \in \hat{\mathbb{R}}$. Then h is convex by Lemma **??** and continuous at $z = \theta$. Hence there is some $-\lambda \in \partial h(\theta)$, and this provides the linear part of the assured affine function. Indeed, we have

$$f(x) - g(z - u) \geq h(z) - h(\theta) \geq \langle \lambda, z \rangle\,.$$

■

Corollary 4.92. *Let $\mathcal{X}$ be a Banach space and let $f, g \in \Gamma_0(\mathcal{X})$ and let*

$$\theta \in core(dom(f) + dom(g))\,. \tag{4.32}$$

Then

$$\partial(f + g) = \partial f + \partial g\,.$$

Proof:

■

Corollary 4.93. *Let $\mathcal{X}$ be a Banach space with duality mapping $J_{\mathcal{X}}$, let $f \in \Gamma_0(\mathcal{X})$, let $j := \frac{1}{2}\|\cdot\|^2$, and let $h := f \,\square\, j$. Then*

(a) $dom(h) = \mathcal{X}$ *and* h *is continuous at each* $x \in \mathcal{X}$.

(b) $\lambda \in \partial f(v) + J_{\mathcal{X}}(v)$ *if and only if* $h^*(\lambda) + h(v) - \langle \lambda, v \rangle \leq 0$.

Proof:

■

4.10 Conclusions and comments

4.11 Exercises

1.) Let $\mathcal{X}$ be a normed space. Then the following conditions are equivalent:

(a) $\mathcal{X}$ is uniformly convex.

(b) If $(x^n)_{n\in\mathbb{N}}, (y^n)_{n\in\mathbb{N}}$ are sequences in $\mathcal{X}$ with $\|x^n\| = \|y^n\| = 1, n \in \mathbb{N}$, then we have:

$$\text{From } \lim_n \|\frac{1}{2}(x^n + y^n)\| = 1 \text{ we conclude } \lim_n (x^n - y^n) = \theta\,.$$

(c) If $(x^n)_{n\in\mathbb{N}}, (y^n)_{n\in\mathbb{N}}$ are sequences in $\mathcal{X}$ with $\limsup_n \|x^n\| \leq 1, \limsup_n \|y^n\| \leq 1, n \in \mathbb{N}$, we have:

$$\text{From } \lim_n (\|\frac{1}{2}(x^n + y^n)\|) = 1 \text{ we conclude } \lim_n (x^n - y^n) = \theta\,.$$

2.) Let $\mathcal{X}$ be a uniformly convex Banach space. If $(x^n)_{n\in\mathbb{N}}$ is a sequence in $\mathcal{X}$ with

$$\alpha := \lim_n \|x^n\| = \lim_{n,m} \|\frac{1}{2}(x^n + x^m)\|\,,$$

Then this sequence is convergent.

3.) Let $\mathcal{X}$ be a uniformly convex Banach space and let $(x^n)_{n\in\mathbb{N}}$ be a sequence in X. Then the following conditions are equivalent:

(a) $\lim_n x^n = x\,.$

(b) $w - \lim_n x^n = x\,,\ \lim_n \|x^n\| = \|x\|\,.$

4.) We say that the Banach space $\mathcal{X}$ has a quadratic modulus of smoothness if with a positive constant k we have

$$\|x + y\| + \|x - y\| \leq 2 + k\|y\|^2 \text{ for all } x, y \in Xs\,.$$

Show that the Banach spaces $l_p, 2 \leq p < \infty$, has a quadratic modulus of smoothness.

5.) Show: If a Banach space has a quadratic modulus of smoothness then the associated duality map is Lipschitz continuous.

6.) Let $\mathcal{X} := C[0,1]$ endowed with the norm $\|f\| := (\|f\|_\infty^2 + \|f\|_2^2)^{\frac{1}{2}}$. Then $\mathcal{X}$ is strictly convex but nor uniformly convex.

7.) Let $\mathcal{H}$ be a Hilbert space and let C be nonempty closed convex subset of $\mathcal{H}$. Let P_C be the metric projection. Show for $x, y \in \mathcal{H}$:

$$\|P_C(x) - P_C(y)\| < \|x - y\| \text{ or } P_C(x) - P_C(y) = x - y\,.$$

8.) Let $\mathcal{X}$ be a Banach space and let U, V closed subspaces of $\mathcal{X}$.

(a) Show that if U is a finite dimensional subspace then $U + V$ is closed.

(b) Consider now the space $\mathcal{X} := c_0$ of real sequences converging to zero endowed with the supremum norm. Let

$$\begin{aligned} U &:= \{(x^n)_{n\in\mathbb{N}} \in c_0 : x^n = nx^{n-1}, n \text{ even}\}, \\ V &:= \{(x^n)_{n\in\mathbb{N}} \in c_0 : x^n = 0, n \text{ odd}\}. \end{aligned}$$

Show that $U + V$ is dense in c_0 but not closed.

9.) Let $\mathbb{R}^2$ endowed with the euclidean norm. Consider the sequence $(C_n)_{n\in\mathbb{N}}$ with $C_n := \{x, x/n) : x \in \mathbb{R}\}, n \in \mathbb{N}$. Show taht $C := \lim_M C_n$ exists.

10.) Let Hs be the Hilbert space l_2 and consider the sets $C_n := \text{span}(\{e^1, e^n\}), n \in \mathbb{N}$. What can be said about $\lim_M C_n, \lim_W C_n, \lim_S C_n$?

11.) Show that the spaces c_0 and l_1 are isomorphic. Here c_0 is the space of real sequences converging to zero endowed with the supremum norm.

12.) Show:

(a) The closed unit ball in c_0 has no extreme point.

(b) The closed unit ball in c has an extreme point.

Here c is the space of converging real sequences and c_0 is the space of real sequences converging to zero; both spaces are endowed with the supremum norm. Is it possible that c_0, c are isometric isomorph?

13.) Let $\mathcal{X}$ be a Banach space and let $F : \mathcal{X} \rightrightarrows \mathcal{X}$ be lower semicontinuous. Moreover let $F(x)$ not empty for all $x \in \mathcal{X}$. Show that

$$\text{co}(F) : \mathcal{X} \rightrightarrows \mathcal{X}, \text{co}(F)(x) := \text{co}(F(x)),$$

is lower semicontinuous.

14.) Let $\mathcal{X}$ be a Banach space and let $F : \mathcal{X} \rightrightarrows \mathcal{X}$ be lower semicontinuous. Moreover let $F(x)$ not empty for all $x \in \mathcal{X}$. Show that

$$\overline{\text{co}}(F) : \mathcal{X} \rightrightarrows \mathcal{X}, \overline{\text{co}}(F)(x) := \overline{\text{co}(F(x))},$$

is lower semicontinuos.

15.) Let $\mathcal{X}$ be a Banach space and let $f : \mathcal{X} \longrightarrow \mathbb{R}$. Show that f is lower semicontinuous if and only if $x \longmapsto \{tin\mathbb{R} : f(x) \leq t\}$ is upper semicontinuous.

16.) consider the Hilbert space l_2 with the standard basis vectors $e^k, k \in \mathbb{N}$. Let sequences be $(x^n)_{n\in\mathbb{N}}, (u^n)_{n\in\mathbb{N}} \in l_2$ defined by

$$x^n := 1/n\, e^n, u^n := (1 + 1/n)e^n,\ n \in \mathbb{N}.$$

Moreover, $w := e^1$. Set $C := \{w\} \cup \{u^n : n \in \mathbb{N}\}$.

(a) Show $P_C(x^n) = u^n$, $n \in \mathbb{N}, n \geq 2$.

(b) Is P_C continuous in $\theta \in l_2$?

(c) Is C is closed ?

Bibliography

[1] A.D. Alexandrov. Almost everywhere existence of the second differential of a convex function and some properties of convex surfaces connected with it. *Leningrad Univ. Annnals Math.*, Ser. 6:3–35, 1939.

[2] E. Asplund. Differentiability of the metric projection in infinite-dimensional euclidean spaces. *Proc. Amer. Math. Soc.*, 38:218–219, 1973.

[3] J. Baumeister. Funktionalanalysis, 2013. Skriptum einer Vorlesung, Goethe–Universität Frankfurt/Main.

[4] J.M. Borwein. A note on ε-subgradients and maximal monotonicity. *Pacific Journal of Mathematics*, 103:307–314, 1982.

[5] J.M. Borwein, S.P. Fitzpatrick, and J.R. Giles. The differentiability of real functions on normed linear spaces using generalized subgradients. *J. Math. Anal. Appl.*, 128:512–534, 1987.

[6] J.M. Borwein and A.S. Lewis. Partially-finite convex programming, part i: Quasi-relative interiors and duality theory. *Mathematical Programming*, 57:15–48, 1992.

[7] J.M. Borwein and A.S. Lewis. *Convex Analysis and Nonlinear Optimization. Theory and Examples.* Springer, New York, 2006.

[8] A.L. Brown. A rotund and reflexive space having a subspace of codimension two with a discontinuous metric projection. *Michigan Math. J.*, 21:15–151, 1874.

[9] E.W. Cheney and K.H. Price. Minimal projections. In A. Talbot, editor, *Approximation Theory*, pages 261–289, New York, 1970. Academic Press.

[10] C. Chidume. *Geometric properties of Banach spaces and nonlinear iterations.* Springer, New York, 2009.

[11] I. Cioranescu. *Geometry of Banach Spaces, Duality Mappings and Nonlinear Problems.* Kluwer, Dordrecht, 1990.

[12] J.A. Clarkson. Uniformly convex spaces. *Trans. Amer. Math. Soc.*, pages 396–414, 1936.

[13] M. Day. Reflexive Banach spaces not isomorphic to uniformly convex spaces. *Bull. Amer. Math. Soc.*, pages 313–317, 1941.

[14] K. Deimling. *Nonlinear Functional analysis.* Springer, Berlin, 1985.

[15] F. Deutsch and J.M Lambert. On continuity of metric projections. *J. of Approximation Theory*, 29:116–131, 1980.

[16] T. Figiel. On the moduli of convexity and smoothness. *Studia Mathematica*, 56:121–155, 1976.

[17] S.P. Fitzpatrick. Metric projections and the differentiability of distance functions. *Bull. Austr. Math. Soc.*, 22:291–312, 1980.

[18] J.B. Giles. Differentiability of distance functions and a proximal property inducing convexity. *Proc. of the Amer. Math. Soc.*, 104:458–464, 1988.

[19] J.B. Giles. Distance function property implying differentiability. *Bull. Austral. Math. Soc.*, 39:59–70, 1989.

[20] A. Haraux. How to differentiate the projection on a convex set in a Hilbert space. some applications to variational inequalities. *J. Math. Soc. Japan*, pages 615–631, 1977.

[21] J.-B. Hiriart-Urruty and C. Lemarechal. *Fundamentals of Convex Analysis.* Springer, Berlin, 2001.

[22] J.-B. Hirriart-Urruty and C. Lemarechal. *Convex analysis and minimization algorithms. Vol I, II.* Springer, Berlin, 2001.

[23] R.B. Holmes and B.R Kripke. Smoothness of approximation. *Michigan Math. J.*, 15:225–234, 1968.

[24] W. Kaballo. *Grundkurs Funktionalanalysis.* Spektrum/Akademischer Verlag, Heidelberg, 2010.

[25] W.A. Kirk and M.F. Smiley. Another characterization of inner product spaces. *Amer. Math. Monthly*, 71:890–891, 1964.

[26] G. Koethe. *Topological linear spaces.* Springer, Berlin, 1960.

[27] J. Li. The metric projection and its applications to solving variational inequalities in Banach spaces. *Fixed point Theory*, 5:285–298, 2004.

[28] J. Lindenstrauss and L. Tzafriri. *Classical Banach spaces I and II.* Springer, Berlin, 1977.

[29] L. Maligranda. Simple norm inequalities. *Amer. Math. Monthly*, 113:256–260, 2006.

[30] R.E. Megginson. *An introduction to Banach space theory.* Springer, New York, 1998.

[31] F. Mignot. Controle dans inequations variationelles elliptiques. *J. Functional Analysis*, 2:130–185, 1976.

[32] M.S. Moslehian, F: Ddipour, R. Rajic, and A. Maric. A glimpse at the Dunkl-Williams inequality. *Banach J. Math. Anal.*, 5:138–151, 2011.

[33] D. Noll. Directionally differentiability of the metric projection in Hilbert space. *Pacific Journal of Mathematics*, 170:567–592, 1995.

[34] G. Nordlander. The modulus of convexity in normed linear spaces. *Arkiv for Matematik*, 100:15–17, 1960.

[35] J. Pruess. A characterization of uniform convexity and application to accretive operators. *Hiroshima Math. J.*, 11:229–234, 1981.

[36] H. Rademacher. über partielle und totale Differenzierbarkeit I. *Math. Annalen*, 79:340–359, 1919.

[37] R.M. Rakestraw. A characterization of inner product spaces. *J. Math. Anal. Appl.*, 68:267–272, 1979.

[38] A. Shapiro. Directionally nondifferentiable metric projection. *Journal of Optimization Theory and applications*, 81:203–204, 1994.

[39] A. Shapiro. Differentiability properties of metric projections onto convex sets. *Journal of Optimization Theory and applications*, 169:953–964, 2016.

[40] I. Vlasov. Chebyshev sets in Banach spaces. *Soviet Math. Dokl.*, pages 1373–1374, 2.

[41] D. Werner. *Funktionalanalysis*. Springer, 2002.

[42] E.H. Zarantonello. Projections on convex sets in Hilbert space and spectral theory. Technical report, Math. Res. Center, Wisconsin, 1971.

[43] W.P. Ziemer. *Weakly differentiable functions*. Springer, New York, 1989.