Chapter 2

Convexity, Smoothness, and Duality Mappings

Geometric structures such as convexity and smoothness of Banach spaces play an important role in the existence and approximation of fixed points of nonlinear mappings. This chapter presents a substantial number of useful properties of duality mappings and Banach spaces having these geometric structures.

2.1 Strict convexity

Let X be a linear space. The *line segment* or *interval joining* the two points $x, y \in X$ is the set $[x, y] := \{\lambda x + (1-\lambda)y : 0 \leq \lambda \leq 1\}$, i.e., $[x, y] = co(\{x, y\})$ is convex hull of x and y.

The basic property of a norm of a Banach space X is that it is always convex, i.e.,

$$\|(1-\lambda)x + \lambda y\| \leq (1-\lambda)\|x\| + \lambda\|y\| \text{ for all } x, y \in X \text{ and } \lambda \in [0, 1].$$

A number of Banach spaces do not have equality when $x \neq y$, i.e.,

$$\|(1-\lambda)x + \lambda y\| < (1-\lambda)\|x\| + \lambda\|y\| \quad \text{for all } x, y \in X \text{ with } x \neq y \text{ and } \lambda \in (0, 1). \tag{2.1}$$

We use S_X to denote the unit sphere $S_X = \{x \in X : \|x\| = 1\}$ on Banach space X. If $x, y \in S_X$ with $x \neq y$, then (2.1) reduces to

$$\|(1-\lambda)x + \lambda y\| < 1 \text{ for all } \lambda \in (0, 1),$$

which says that the unit sphere S_X contains no line segments. This suggests strict convexity of norm.

R.P. Agarwal et al., *Fixed Point Theory for Lipschitzian-type Mappings with Applications*, Topological Fixed Point Theory and Its Applications 6, DOI 10.1007/978-0-387-75818-3_2,

Definition 2.1.1 *A Banach space X is said to be strictly convex if*

$$x, y \in S_X \text{ with } x \neq y \Rightarrow \|(1-\lambda)x + \lambda y\| < 1 \text{ for all } \lambda \in (0,1).$$

This says that the midpoint $(x+y)/2$ of two distinct points x and y in the unit sphere S_X of X does not lie on S_X. In other words, if $x, y \in S_X$ with $\|x\| = \|y\| = \|(x+y)/2\|$, then $x = y$.

Example 2.1.2 *Consider $X = \mathbb{R}^n, n \geq 2$ with norm $\|x\|_2$ defined by*

$$\|x\|_2 = \left(\sum_{i=1}^{n} x_i^2\right)^{1/2}, \quad x = (x_1, x_2, \cdots, x_n) \in \mathbb{R}^n.$$

Then X is strictly convex.

Example 2.1.3 *Consider $X = \mathbb{R}^n, n \geq 2$ with norm $\|\cdot\|_1$ defined by*

$$\|x\|_1 = |x_1| + |x_2| + \cdots + |x_n|, \quad x = (x_1, x_2, \cdots, x_n) \in \mathbb{R}^n.$$

Then X is not strictly convex. To see it, let

$$x = (1, 0, 0, \cdots, 0) \text{ and } y = (0, 1, 0, \cdots, 0).$$

It is easy to see that $x \neq y, \|x\|_1 = 1 = \|y\|_1$, but $\|x+y\|_1 = 2$.

Example 2.1.4 *Consider $X = \mathbb{R}^n, n \geq 2$ with norm $\|\cdot\|_\infty$ defined by*

$$\|x\|_\infty = \max_{1 \leq i \leq n} |x_i|, \quad x = (x_1, x_2, \cdots, x_n) \in \mathbb{R}^n.$$

Then X is not strictly convex. Indeed, for $x = (1, 0, 0, \cdots, 0)$ and $y = (1, 1, 0, \cdots, 0)$, we have, $x \neq y$, $\|x\|_\infty = 1 = \|y\|_\infty$, but $\|x+y\|_\infty = 2$.

The other equivalent conditions of strict convexity are given in the following:

Proposition 2.1.5 *Let X be a Banach space. Then the following are equivalent:*

(a) X is strictly convex.

(b) For each nonzero $f \in X^*$, there exists at most one point x in X with $\|x\| = 1$ such that $\langle x, f \rangle = f(x) = \|f\|_*$.

Proof. $(a) \Rightarrow (b)$. Let X be a strictly convex Banach space and f an element in X^*. Suppose there exist two distinct points x, y in X with $\|x\| = \|y\| = 1$ such that $f(x) = f(y) = \|f\|_*$. If $t \in (0,1)$, then

$$\begin{aligned}
\|f\|_* &= tf(x) + (1-t)f(y) \qquad (as\ f(x) = f(y) = \|f\|_*) \\
&= f(tx + (1-f)y) \\
&\leq \|f\|_* \|tx + (1-t)y\| \\
&< \|f\|_*, \qquad (as\ \|tx + (1-t)y\| < 1)
\end{aligned}$$

which is a contradiction. Therefore, there exists at most one point x in X with $\|x\| = 1$ such that $f(x) = \|f\|_*$.

(*b*) $\Rightarrow$ (*a*). Suppose $x, y \in S_X$ with $x \neq y$ such that $\|(x+y)/2\| = 1$. By Corollary 1.6.6, there exists a functional $j \in S_{X^*}$ such that

$$\|j\|_* = 1 \text{ and } \langle (x+y)/2, j\rangle = \|(x+y)/2\|.$$

Because $\langle x, j\rangle \leq 1$ and $\langle y, j\rangle \leq 1$, we have $\langle x, j\rangle = \langle y, j\rangle$. This implies, by hypothesis, that $x = y$. Therefore, (*b*) $\Rightarrow$ (*a*) is proved. ∎

Proposition 2.1.6 *Let X be a Banach space. Then the following statements are equivalent:*

(a) X is strictly convex.

(b) For every $1 < p < \infty$,

$$\|tx+(1-t)y\|^p < t\|x\|^p+(1-t)\|y\|^p \textit{ for all } x, y \in X,\ x \neq y \textit{ and } t \in (0,1).$$

Proof. (*a*) $\Rightarrow$ (*b*). Let X be strictly convex. Suppose $x, y \in X$ with $x \neq y$. By strict convexity of X,

$$\|tx + (1-t)y\|^p < (t\|x\| + (1-t)\|y\|)^p \text{ for all } t \in (0,1). \tag{2.2}$$

If $\|x\| = \|y\|$, then

$$\|tx + (1-t)y\|^p < \|x\|^p = t\|x\|^p + (1-t)\|y\|^p.$$

We now assume that $\|x\| \neq \|y\|$. Consider the function $t \mapsto t^p$ for $1 < p < \infty$. Then it is a convex function and

$$\left(\frac{a+b}{2}\right)^p < \frac{a^p + b^p}{2} \text{ for all } a, b \geq 0 \text{ and } a \neq b.$$

Hence from (2.2) with $t = 1/2$, we have

$$\left\|\frac{x+y}{2}\right\|^p \leq \left(\frac{\|x\| + \|y\|}{2}\right)^p < \frac{1}{2}(\|x\|^p + \|y\|^p). \tag{2.3}$$

If $t \in (0, 1/2]$, then from (2.2), we have

$$\begin{aligned} \|tx + (1-t)y\|^p &= \left\|2t\frac{x+y}{2} + (1-2t)y\right\|^p \\ &\leq \left(2t\left\|\frac{x+y}{2}\right\| + (1-2t)\|y\|\right)^p \\ &< 2t\left\|\frac{x+y}{2}\right\|^p + (1-2t)\|y\|^p \\ &\leq t\|x\|^p + (1-t)\|y\|^p. \qquad \text{(by (2.3))} \end{aligned}$$

The proof is similar if $t \in (1/2, 1)$.

(*b*) $\Rightarrow$ (*a*). It is obvious. ∎

Proposition 2.1.7 *Let X be a strictly convex Banach space. If $\|x+y\| = \|x\| + \|y\|$ for $0 \neq x \in X$ and $y \in X$, then there exists $t \geq 0$ such that $y = tx$.*

Proof. Let $x, y \in X \backslash \{0\}$ be such that $\|x+y\| = \|x\|+\|y\|$. From Corollary 1.6.6, there exists $j \in X^*$ such that

$$\langle x+y, j\rangle = \|x+y\| \text{ and } \|j\|_* = 1.$$

Because $\langle x, j\rangle \leq \|x\|$ and $\langle y, j\rangle \leq \|y\|$, we must have $\langle x, j\rangle = \|x\|$ and $\langle y, j\rangle = \|y\|$. This means that $\langle x/\|x\|, j\rangle = \langle y/\|y\|, j\rangle = 1$. By strict convexity of X, it follows from Proposition 2.1.5 that $x/\|x\| = y/\|y\|$. Therefore, result holds. ∎

We now present the existence and uniqueness of elements of minimal norm in convex subsets of strictly convex Banach spaces.

Proposition 2.1.8 *Let X be a strictly convex Banach space and C a nonempty convex subset of X. Then there is at most one point x in C such that $\|x\| = inf\{\|z\| : z \in C\}$.*

Proof. Suppose, there exist two points $x, y \in C, x \neq y$ such that

$$\|x\| = \|y\| = \inf\{\|z\| : z \in C\} = d\ (say).$$

If $t \in (0, 1)$, then by strict convexity of X we have that

$$\|tx + (1-t)y\| < d,$$

which is a contradiction, as $tx + (1-y) \in C$ by the convexity of C. ∎

Proposition 2.1.9 *Let C be a nonempty closed convex subset of a reflexive strictly convex Banach space X. Then there exists a unique point $x \in C$ such that $\|x\| = \inf\{\|z\| : z \in C\}$.*

Proof. Existence: Let $d := \inf\{\|z\| : z \in C\}$. Then there exists a sequence $\{x_n\}$ in C such that $\lim_{n\to\infty} \|x_n\| = d$. By the reflexivity of X, there exists a subsequence $\{x_{n_i}\}$ of $\{x_n\}$ that converges weakly to an element x in C. The weak lower semicontinuity (w-lsc) of the norm (see Theorem 1.9.10) gives

$$\|x\| \leq \lim_{n\to\infty} \|x_n\| = d.$$

Therefore, $d = \|x\|$.

Uniqueness: It follows from Proposition 2.1.8 . ∎

The following result has important applications in the existence and uniqueness of best approximations.

Proposition 2.1.10 *Let C be a nonempty closed convex subset of a reflexive strictly convex Banach space X. Then for $x \in X$, there exists a unique point $z_x \in C$ such that $\|x - z_x\| = d(x, C)$.*

Proof. Let $x \in C$. Because C is a nonempty closed convex subset Banach space X, then $D = \{y - x : y \in C\}$ is a nonempty closed convex subset of X. By Proposition 2.1.9, there exists a unique point $u_x \in D$ such that $\|u_x\| = \inf\{\|y - x\| : y \in C\}$. For $u_x \in D$, there exists a point $z_x \in C$ such that $u_x = z_x - x$. Thus, there exists a unique point $z_x \in C$ such that $\|z_x - x\| = d(x, C)$. ∎

2.2 Uniform convexity

The strict convexity of a normed space X says that the midpoint $(x + y)/2$ of the segment joining two distinct points $x, y \in S_X$ with $\|x - y\| \geq \varepsilon > 0$ does not lie on S_X, i.e.,

$$\left\| \frac{x + y}{2} \right\| < 1.$$

In such spaces, we have no information about $1 - \|(x + y)/2\|$, the distance of midpoints from the unit sphere S_X. A stronger property than strict convexity that provides information about the distance $1 - \|(x+y)/2\|$ is uniform convexity.

Definition 2.2.1 *A Banach space X is said to be uniformly convex* [1] *if for any ε, $0 < \varepsilon \leq 2$, the inequalities $\|x\| \leq 1, \|y\| \leq 1$ and $\|x - y\| \geq \varepsilon$ imply there exists a $\delta = \delta(\varepsilon) > 0$ such that $\|(x + y)/2\| \leq 1 - \delta$.*

This says that if x and y are in the closed unit ball $B_X := \{x \in X : \|x\| \leq 1\}$ with $\|x - y\| \geq \varepsilon > 0$, the midpoint of x and y lies inside the unit ball B_X at a distance of at least δ from the unit sphere S_X.

Example 2.2.2 *Every Hilbert space H is a uniformly convex space. In fact, the parallelogram law gives us*

$$\|x + y\|^2 = 2(\|x\|^2 + \|y\|^2) - \|x - y\|^2 \text{ for all } x, y \in H.$$

Suppose $x, y \in B_H$ with $x \neq y$ and $\|x - y\| \geq \varepsilon$. Then

$$\|x + y\|^2 \leq 4 - \varepsilon^2,$$

so it follows that

$$\|(x + y)/2\| \leq 1 - \delta(\varepsilon),$$

where $\delta(\varepsilon) = 1 - \sqrt{1 - \varepsilon^2/4}$. Therefore, H is uniformly convex.

Example 2.2.3 *The spaces ℓ_1 and ℓ_∞ are not uniformly convex. To see it, take $x = (1, 0, 0, 0 \cdots), y = (0, -1, 0, 0, \cdots) \in \ell_1$ and $\varepsilon = 1$. Then*

$$\|x\|_1 = 1, \|y\|_1 = 1, \|x - y\|_1 = 2 > 1 = \varepsilon.$$

[1] The concept of uniform convexity was introduced by Clarkson in 1936.

However, $\|(x+y)/2\|_1 = 1$ *and there is no* $\delta > 0$ *such that* $\|(x+y)/2\|_1 \leq 1-\delta$. *Thus,* ℓ_1 *is not uniformly convex.*

Similarly, if we take $x = (1,1,1,0,0,\cdots), y = (1,1,-1,0,0\cdots) \in \ell_\infty$ *and* $\varepsilon = 1$, *then*

$$\|x\|_\infty = 1, \|y\|_\infty = 1, \|x-y\|_\infty = 2 > 1 = \varepsilon.$$

Because $\|(x+y)/2\|_\infty = 1$, ℓ_∞ *is not uniformly convex.*

Observation

- The Banach spaces ℓ_p, ℓ_p^n (whenever n is a nonnegative integer), and $L_p[a,b]$ with $1 < p < \infty$ are uniformly convex.
- The Banach spaces ℓ_1, c, c_0, ℓ_∞, $L_1[a,b]$, $C[a,b]$ and $L_\infty[a,b]$ are not strictly convex.

We derive some consequences from the definition of uniform convexity.

Theorem 2.2.4 *Every uniformly convex Banach space is strictly convex.*

Proof. Let X be a uniformly convex Banach space. It easily follows from Definition 2.2.1 that X is strictly convex. $\blacksquare$

Remark 2.2.5 *The converse of Theorem 2.2.4 is not true in general. Let* $\beta > 0$ *and let* $X = c_o$ *with the norm* $\|\cdot\|_\beta$ *defined by*

$$\|x\|_\beta = \|x\|_{c_o} + \beta\left(\sum_{i=1}^{\infty}\left(\frac{x_i}{i}\right)^2\right)^{1/2}, \quad x = \{x_i\} \in c_o.$$

The spaces $(c_o, \|\cdot\|_\beta)$ *for* $\beta > 0$ *are strictly convex, but not uniformly convex, while* c_0 *with its usual norm is not strictly convex.*

Theorem 2.2.6 *Let* X *be a uniformly convex Banach space. Then we have the following:*

(a) For any r *and* ε *with* $r \geq \varepsilon > 0$ *and elements* $x, y \in X$ *with* $\|x\| \leq r$, $\|y\| \leq r$, $\|x-y\| \geq \varepsilon$, *there exists a* $\delta = \delta(\varepsilon/r) > 0$ *such that*

$$\|(x+y)/2\| \quad \leq \quad r[1-\delta(\varepsilon/r)].$$

(b) For any r *and* ε *with* $r \geq \varepsilon > 0$ *and elements* $x, y \in X$ *with* $\|x\| \leq r$, $\|y\| \leq r$, $\|x-y\| \geq \varepsilon$, *there exists a* $\delta = \delta(\varepsilon/r) > 0$ *such that*

$$\|tx + (1-t)y\| \leq r[1 - 2\min\{t, 1-t\}\delta(\varepsilon/r)] \text{ for all } t \in (0,1).$$

Proof. (a) Suppose that $\|x\| \leq r$, $\|y\| \leq r$ and $\|x-y\| \geq \varepsilon > 0$. Then we have that

$$\left\|\frac{x}{r}\right\| \leq 1, \left\|\frac{y}{r}\right\| \leq 1 \text{ and } \left\|\frac{x}{r} - \frac{y}{r}\right\| \geq \frac{\varepsilon}{r} > 0.$$

By the definition of uniform convexity, there exists $\delta = \delta(\varepsilon/r) > 0$ such that

$$\left\|\frac{x+y}{2r}\right\| \le 1-\delta,$$

which yields

$$\left\|\frac{x+y}{2}\right\| \le r(1-\delta).$$

(b) When $t = 1/2$, we are done by Part (a). If $t \in (0, 1/2]$, we have

$$\|tx + (1-t)y\| = \|t(x+y) + (1-2t)y\| \le 2t\|\frac{x+y}{2}\| + (1-2t)\|y\|. \quad (2.4)$$

From part (a), there exists a $\delta = \delta(\varepsilon/r) > 0$ such that

$$\left\|\frac{x+y}{2}\right\| \le r\left[1-\delta\left(\frac{\varepsilon}{r}\right)\right].$$

From (2.4), we have

$$\begin{aligned}\|tx+(1-t)y\| &\le 2t\left[1-\delta\left(\frac{\varepsilon}{r}\right)\right]r + (1-2t)r \quad (\text{as } \|y\| \le r)\\ &\le r\left[1-2t\delta\left(\frac{\varepsilon}{r}\right)\right].\end{aligned}$$

Now by the choice of $t \in [1/2, 1)$, we have

$$\begin{aligned}\|tx+(1-t)y\| &= \|(2t-1)x + (1-t)(x+y)\|\\ &\le (2t-1)\|x\| + 2(1-t)\left\|\frac{x+y}{2}\right\|\\ &\le (2t-1)r + 2(1-t)r\left[1-\delta\left(\frac{\varepsilon}{r}\right)\right]\\ &= r\left[1-2(1-t)\delta\left(\frac{\varepsilon}{r}\right)\right].\end{aligned}$$

Therefore,

$$\|tx+(1-t)y\| \le r\left[1-2\min\{t,1-t\}\delta\left(\frac{\varepsilon}{r}\right)\right]. \quad \blacksquare$$

Theorem 2.2.7 *Let X be a Banach space. Then the following are equivalent:*

(a) X is uniformly convex.

(b) For two sequences $\{x_n\}$ and $\{y_n\}$ in X,

$$\|x_n\| \le 1, \|y_n\| \le 1 \text{ and } \lim_{n\to\infty}\|x_n+y_n\| = 2 \Rightarrow \lim_{n\to\infty}\|x_n-y_n\| = 0. \quad (2.5)$$

Proof. $(a) \Rightarrow (b)$. Suppose X is uniformly convex. Let $\{x_n\}$ and $\{y_n\}$ be two sequences in X such that $\|x_n\| \leq 1$, $\|y_n\| \leq 1$ for all $n \in \mathbb{N}$ and $\lim_{n\to\infty} \|x_n + y_n\| = 2$. Suppose, for contradiction, that $\lim_{n\to\infty} \|x_n - y_n\| \neq 0$. Then for some $\varepsilon > 0$, there exists a subsequence $\{n_i\}$ of $\{n\}$ such that

$$\|x_{n_i} - y_{n_i}\| \geq \varepsilon.$$

Because X is uniformly convex, there exists $\delta(\varepsilon) > 0$ such that

$$\|x_{n_i} + y_{n_i}\| \leq 2(1 - \delta(\varepsilon)). \tag{2.6}$$

Because $\lim_{n\to\infty} \|x_n + y_n\| = 2$, it follows from (2.6) that

$$2 \leq 2(1 - \delta(\varepsilon)),$$

a contradiction.

$(b) \Rightarrow (a)$. Suppose condition (2.5) is satisfied. If X is not uniformly convex, for $\varepsilon > 0$, there is no $\delta(\varepsilon)$ such that

$$\|x\| \leq 1, \|y\| \leq 1, \|x - y\| \geq \varepsilon \Rightarrow \|x + y\| \leq 2(1 - \delta(\varepsilon)),$$

and then we can find sequences $\{x_n\}$ and $\{y_n\}$ in X such that

(i) $\|x_n\| \leq 1$, $\|y_n\| \leq 1$,

(ii) $\|x_n + y_n\| \geq 2(1 - 1/n)$,

(iii) $\|x_n - y_n\| \geq \varepsilon$.

Clearly $\|x_n - y_n\| \geq \varepsilon$, which contradicts the hypothesis, as (ii) gives $\lim_{n\to\infty} \|x_n + y_n\| = 2$. Thus, X must be uniformly convex. ∎

For the class of uniform convex Banach spaces, we have the following important results:

Theorem 2.2.8 *Every uniformly convex Banach space is reflexive.*

Proof. Let X be a uniformly convex Banach space. Let $S_{X^*} := \{j \in X^* : \|j\|_* = 1\}$ be the unit sphere in X^* and $f \in S_{X^*}$. Suppose $\{x_n\}$ is a sequence in S_X such that $f(x_n) \to 1$. We show that $\{x_n\}$ is a Cauchy sequence. Suppose, for contradiction, that there exist $\varepsilon > 0$ and two subsequences $\{x_{n_i}\}$ and $\{x_{n_j}\}$ of $\{x_n\}$ such that $\|x_{n_i} - x_{n_j}\| \geq \varepsilon$. The uniform convexity of X guarantees that there exists $\delta(\varepsilon) > 0$ such that $\|(x_{n_i} + x_{n_j})/2\| < 1 - \delta$. Observe that

$$|f((x_{n_i} + x_{n_j})/2)| \leq \|f\|_* \|(x_{n_i} + x_{n_j})/2\| < \|f\|_*(1 - \delta) = 1 - \delta$$

and $f(x_n) \to 1$, yield a contradiction. Hence $\{x_n\}$ is a Cauchy sequence and there exists a point x in X such that $x_n \to x$. Clearly, $x \in S_X$. In fact,

$$\|x\| = \|\lim_{n\to\infty} x_n\| = \lim_{n\to\infty} \|x_n\| = 1.$$

Using the James theorem (which states that a Banach space is reflexive if and only if for each $f \in S_{X^*}$, there exists $x \in S_X$ such that $f(x) = 1$), we conclude that X is reflexive. ∎

Remark 2.2.9 *Every finite-dimensional Banach space is reflexive, but it need not be uniformly convex, for example,* $X = \mathbb{R}^n, n \geq 2$ *with the norm* $\|x\|_1 = \sum_{i=1}^{n} |x_i|$.

Combining Proposition 2.1.9 and Theorems 2.2.4 and 2.2.8, we obtain the following interesting result:

Theorem 2.2.10 *Let C be a nonempty closed convex subset of a uniformly convex Banach space X. Then C has a unique element of minimum norm, i.e., there exists a unique element $x \in C$ such that $\|x\| = \inf\{\|z\| : z \in C\}$.*

We now introduce a useful property.

Definition 2.2.11 *A Banach space X is said to have the Kadec-Klee property if for every sequence $\{x_n\}$ in X that converges weakly to x where also $\|x_n\| \to \|x\|$, then $\{x_n\}$ converges strongly to x.*

Remark 2.2.12 *In Definition 2.2.11, the sequence $\{x_n\}$ can be replaced by the net $\{x_\alpha\}$ for the definition of the Kadec property.*

The following result has a very useful property:

Theorem 2.2.13 *Every uniformly convex Banach space has the Kadec-Klee property.*

Proof. Let X be a uniformly convex Banach space. Let $\{x_n\}$ be a sequence in X such that $x_n \rightharpoonup x \in X$ and $\|x_n\| \to \|x\|$. If $x = 0$, then $\lim_{n\to\infty} \|x_n\| = 0$, which yields that $\lim_{n\to\infty} x_n = 0$.

Suppose $x \neq 0$. Then we show that $x_n \to x$. Suppose, for contradiction, that $\lim_{n\to\infty} x_n \neq x$, i.e., $x_n/\|x_n\| \not\to x/\|x\|$. Then for $\varepsilon > 0$, there exists a subsequence $\{x_{n_i}/\|x_{n_i}\|\}$ of $\{x_n/\|x_n\|\}$ such that

$$\left\| \frac{x_{n_i}}{\|x_{n_i}\|} - \frac{x}{\|x\|} \right\| \geq \varepsilon > 0.$$

Because X is uniformly convex, there exists $\delta(\varepsilon) > 0$ such that

$$\frac{1}{2}\left\| \frac{x_{n_i}}{\|x_{n_i}\|} + \frac{x}{\|x\|} \right\| \leq 1 - \delta.$$

Because $x_n \rightharpoonup x$ and $\|x_n\| \to \|x\|$ imply $x_n/\|x_n\| \rightharpoonup x/\|x\|$, it follows that

$$\left\| \frac{x}{\|x\|} \right\| \quad \leq \quad \liminf_{n\to\infty} \frac{1}{2}\left\| \frac{x_{n_i}}{\|x_{n_i}\|} + \frac{x}{\|x\|} \right\| \leq 1 - \delta,$$

a contradiction. Therefore, $\{x_n\}$ converges strongly to $x \in X$. ∎

2.3 Modulus of convexity

Definition 2.3.1 *Let X be a Banach space. Then a function $\delta_X : [0,2] \to [0,1]$ is said to be the modulus of convexity of X if*

$$\delta_X(\varepsilon) = inf\left\{1 - \left\|\frac{x+y}{2}\right\| : \|x\| \le 1, \|y\| \le 1, \|x-y\| \ge \varepsilon\right\}.$$

It is easy to see that $\delta_X(0) = 0$ and $\delta_X(t) \ge 0$ for all $t \ge 0$.

Example 2.3.2 *For the case of a Hilbert space H (see Example 2.2.2),*

$$\delta_H(\varepsilon) = 1 - \sqrt{1 - \frac{\varepsilon^2}{4}},\ \varepsilon \in (0,2].$$

We now give the modulus of convexity for ℓ_p $(2 \le p < \infty)$ spaces. The following result gives an analogue of the parallelogram law in ℓ_p $(2 \le p < \infty)$ spaces.

Proposition 2.3.3 *In ℓ_p $(2 \le p < \infty)$ spaces,*

$$\|x+y\|^p + \|x-y\|^p \le 2^{p-1}(\|x\|^p + \|y\|^p) \text{ for all } x, y \in \ell_p. \tag{2.7}$$

Proof. We observe from Lemma A.1.1 of Appendix A that for $a, b \in \mathbb{R}$ and $p \in [2, \infty)$

$$\begin{aligned}
|a+b|^p + |a-b|^p &\le [|a+b|^2 + |a-b|^2]^{p/2} \\
&\le [2|a|^2 + 2|b|^2]^{p/2} \\
&= 2^{p/2}(|a|^2 + |b|^2)^{p/2} \\
&\le 2^{p/2}2^{(p-2)/2}(|a|^p + |b|^p) \\
&= 2^{p-1}(|a|^p + |b|^p).
\end{aligned}$$

Hence for $x = \{x_i\}_{i=1}^\infty$, $y = \{y_i\}_{i=1}^\infty \in \ell_p$, we have

$$\sum_{i=1}^\infty |x_i + y_i|^p + \sum_{i=1}^\infty |x_i - y_i|^p \le 2^{p-1}\left(\sum_{i=1}^\infty |x_i|^p + \sum_{i=1}^\infty |y_i|^p\right),$$

which implies that points $x, y \in \ell_p$ $(2 \le p < \infty)$ satisfy the following analogue of the parallelogram law:

$$\|x+y\|^p + \|x-y\|^p \le 2^{p-1}(\|x\|^p + \|y\|^p). \qquad \blacksquare$$

Example 2.3.4 *For the ℓ_p $(2 \le p < \infty)$ space,*

$$\delta_{\ell_p}(\varepsilon) = 1 - \left(1 - \left(\frac{\varepsilon}{2}\right)^p\right)^{1/p},\quad \varepsilon \in (0,2).$$

To see this, let $\varepsilon \in (0,2)$ *and* $x, y \in \ell_p$ *such that* $\|x\| \leq 1$, $\|y\| \leq 1$ *and* $\|x-y\| \geq \varepsilon$. *Then from (2.7), we have*

$$\|x+y\|^p \leq 2^p - \|x-y\|^p,$$

which implies that

$$\begin{aligned}\left\|\frac{x+y}{2}\right\| &\leq \left(1-\left(\frac{\varepsilon}{2}\right)^p\right)^{1/p} = 1-\left[1-\left(1-\left(\frac{\varepsilon}{2}\right)^p\right)^{1/p}\right] \\ &\leq 1-\delta_{\ell_p}(\varepsilon),\end{aligned}$$

where $\delta_{\ell_p}(\varepsilon) \geq 1-\left(1-\left(\frac{\varepsilon}{2}\right)^p\right)^{1/p}$.

Observation

- $\delta_H(\varepsilon) = 1-\sqrt{1-(\varepsilon/2)^2}$.
- $\delta_{\ell_p}(\varepsilon) = 1-(1-(\varepsilon/2)^p)^{p/2}$.
- $\delta_{\ell_p}(\varepsilon)$, the modulus of convexity for ℓ_p $(1 < p \leq 2)$ satisfies the following implicit formula:
$$\left|1-\delta_{\ell_p}(\varepsilon)+\frac{\varepsilon}{2}\right|^p + \left|1-\delta_{\ell_p}(\varepsilon)-\frac{\varepsilon}{2}\right|^p = 2.$$
- $\delta_{\ell_p}(\varepsilon) > 0$ for all $\varepsilon > 0$ $(1 < p < \infty)$.
- $\delta_X(\varepsilon) \leq \delta_H(\varepsilon)$ for any Banach spaces X and any Hilbert space H, i.e., a Hilbert space is the most convex Banach space.

We now give some important properties of the modulus of convexity of Banach spaces.

Theorem 2.3.5 *A Banach space* X *is strictly convex if and only if* $\delta_X(2) = 1$.

Proof. Let X be a strictly convex Banach space with modulus of convexity δ_X. Suppose $\|x\| = \|y\| = 1$ and $\|x-y\| = 2$ with $x \neq -y$. By strict convexity of X, we have

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

a contradiction. Hence $x = -y$. Therefore, $\delta_X(2) = 1$.

Conversely, suppose $\delta_X(2) = 1$. Let $x, y \in X$ such that $\|x\| = \|y\| = \|(x+y)/2\| = 1$. Then

$$\left\|\frac{x-y}{2}\right\| = \left\|\frac{x+(-y)}{2}\right\| \leq 1-\delta_X(\|x-(-y)\|) = 1-\delta_X(2) = 0,$$

which implies that $x = y$. Thus, $\|x\| = \|y\|$ and $\|x+y\| = 2 = \|x\|+\|y\|$ imply that $x = y$. Therefore, X is strictly convex. ∎

Theorem 2.3.6 *A Banach space X is uniformly convex if and only if $\delta_X(\varepsilon) > 0$ for all $\varepsilon \in (0, 2]$.*

Proof. Let X be a uniformly convex Banach space. Then for $\varepsilon > 0$, there exists $\delta(\varepsilon) > 0$ such that

$$0 < \delta(\varepsilon) \leq 1 - \left\| \frac{x+y}{2} \right\|$$

for all $x, y \in X$ with $\|x\| \leq 1$, $\|y\| \leq 1$ and $\|x - y\| \geq \varepsilon$. Therefore, from the definition of modulus of convexity, we have $\delta_X(\varepsilon) > 0$.

Conversely, suppose X is a Banach space with modulus of convexity δ_X such that $\delta_X(\varepsilon) > 0$ for all $\varepsilon \in (0, 2]$. Let $x, y \in X$ such that $\|x\| = 1$, $\|y\| = 1$ with $\|x - y\| \geq \varepsilon$ for fixed $\varepsilon \in (0, 2]$. By the modulus of convexity $\delta_X(\varepsilon)$, we have

$$0 < \delta_X(\varepsilon) \leq 1 - \left\| \frac{x+y}{2} \right\|,$$

which implies that

$$\left\| \frac{x+y}{2} \right\| \leq 1 - \delta(\varepsilon),$$

where $\delta(\varepsilon) = \delta_X(\varepsilon)$, which is independent of x and y. Therefore, X is uniformly convex. ∎

Theorem 2.3.7 *Let X be a Banach space with modulus of convexity δ_X. Then we have the following:*

(a) For all ε_1 and ε_2 with $0 \leq \varepsilon_1 < \varepsilon_2 \leq 2$,

$$\delta_X(\varepsilon_2) - \delta_X(\varepsilon_1) \leq \frac{\varepsilon_2 - \varepsilon_1}{2 - \varepsilon_1}(1 - \delta_X(\varepsilon_1)) \leq \frac{\varepsilon_2 - \varepsilon_1}{2 - \varepsilon_1}.$$

In particular, δ_X is a continuous function on $[0, 2)$.

(b) $\delta_X(s)/s$ is a nondecreasing function on $(0, 2]$.

(c) δ_X is a strictly increasing function if X is uniformly convex.

Proof. (a) We define the set

$$S_{u,v} = \{(x, y) : x, y \in B_X; x - y = au, x + y = bv \text{ for some } u, v \in X \text{ and } a, b \geq 0\}$$

and the function

$$\delta_{u,v}(\varepsilon) = \inf \left\{ 1 - \left\| \frac{x+y}{2} \right\| : x, y \in S_{u,v}, \|x - y\| \geq \varepsilon \right\}.$$

Note that $\delta_{u,v}(0) = 0$. For given ε_1 and ε_2 in $(0, 2]$ and $\eta > 0$, we can choose (x_i, y_i) in $S_{u,v}$ such that

$$\|x_i - y_i\| \geq \varepsilon_i \text{ and } \delta_{u,v}(\varepsilon_i) + \eta \geq 1 - \left\| \frac{x_i + y_i}{2} \right\| \text{ for } i = 1, 2.$$

Now for $t \in [0,1]$, let $x_3 = tx_1 + (1-t)x_2$ and $y_3 = ty_1 + (1-t)y_2$. Because $x_i, y_i \in B_X$ for $i = 1, 2$, it follows that

$$\|x_3\| \le t\|x_1\| + (1-t)\|x_2\| \le 1$$

and

$$\|y_3\| \le t\|y_1\| + (1-t)\|y_2\| \le 1.$$

Because $(x_i, y_i) \in S_{u,v}$, there exist positive constants $a_i, b_i \ge 0$ with $i = 1, 2$ such that $x_i - y_i = a_i u$ and $x_i + y_i = b_i v$. Set $\alpha := ta_1 + (1-t)a_2$ and $\beta := tb_1 + (1-t)b_2$. Then

$$\begin{aligned} x_3 - y_3 &= t(x_1 - y_1) + (1-t)(x_2 - y_2) \\ &= ta_1 u + (1-t)a_2 u \\ &= (ta_1 + (1-t)a_2)u \\ &= \alpha u. \end{aligned}$$

Similarly, $x_3 + y_3 = \beta v$. Thus, (x_3, y_3) is in $S_{u,v}$.

Observe that

$$\begin{aligned} \|x_3 - y_3\| &= (ta_1 + (1-t)a_2)\|u\| \\ &= t\|x_1 - y_1\| + (1-t)\|x_2 - y_2\| \\ &\ge t\varepsilon_1 + (1-t)\varepsilon_2 \text{ by the choice of } x_i, y_i, \end{aligned}$$

and $\|x_3 + y_3\| = t\|x_1 + y_1\| + (1-t)\|x_2 + y_2\|$.

By the definition of the function $\delta_{u,v}(\cdot)$, we have

$$\begin{aligned} \delta_{u,v}(t\varepsilon_1 + (1-t)\varepsilon_2) &\le 1 - \left\|\frac{x_3 + y_3}{2}\right\| \\ &\le 1 - t\left\|\frac{x_1 + y_1}{2}\right\| - (1-t)\left\|\frac{x_2 + y_2}{2}\right\| \\ &= t\left(1 - \left\|\frac{x_1 + y_1}{2}\right\|\right) + (1-t)\left(1 - \left\|\frac{x_2 + y_2}{2}\right\|\right) \\ &\le t\left(\delta_{u,v}(\varepsilon_1) + \frac{\eta}{2}\right) + (1-t)\left(\delta_{u,v}(\varepsilon_2) + \frac{\eta}{2}\right) \\ &= t\delta_{u,v}(\varepsilon_1) + (1-t)\delta_{u,v}(\varepsilon_2) + \frac{\eta}{2}. \end{aligned}$$

Because η is arbitrary, it follows that $\delta_{u,v}(\varepsilon)$ is a convex function of ε.

Note that

$$\delta_X(\varepsilon) \le \delta_{u,v}(\varepsilon) \text{ for all } u, v$$

and

$$(x, y) \in S_{u,v} \text{ with } \|x\| \le 1 \text{ and } \|y\| \le 1 \text{ for some } u, v \in X;$$

and hence

$$\delta_X(\varepsilon) = \inf\{\delta_{u,v}(\varepsilon) : u, v \in X \setminus \{0\}\}.$$

Now for any real number $\varepsilon > 0$, there exist $u, v \in X$ such that

$$\delta_{u,v}(\varepsilon_1) \le \delta_X(\varepsilon_1) + \varepsilon.$$

Hence

$$\begin{aligned}\delta_{u,v}(\varepsilon_2) &= \delta_{u,v}\left(2\,\frac{\varepsilon_2-\varepsilon_1}{2-\varepsilon_1} + \left(1-\frac{\varepsilon_2-\varepsilon_1}{2-\varepsilon_1}\right)\varepsilon_1\right)\\ &\le \frac{\varepsilon_2-\varepsilon_1}{2-\varepsilon_1}\delta_{u,v}(2) + \left(1-\frac{\varepsilon_2-\varepsilon_1}{2-\varepsilon_1}\right)\delta_{u,v}(\varepsilon_1),\end{aligned}$$

which implies that

$$\begin{aligned}\delta_{u,v}(\varepsilon_2) - \delta_{u,v}(\varepsilon_1) &\le \frac{\varepsilon_2-\varepsilon_1}{2-\varepsilon_1}(\delta_{u,v}(2) - \delta_{u,v}(\varepsilon_1))\\ &\le \frac{\varepsilon_2-\varepsilon_1}{2-\varepsilon_1}(1-\delta_X(\varepsilon_1)).\end{aligned}$$

Then we have

$$\begin{aligned}\delta_X(\varepsilon_2) - \delta_X(\varepsilon_1) &\le \delta_{u,v}(\varepsilon_2) - \delta_{u,v}(\varepsilon_1) + \varepsilon\\ &\le \frac{\varepsilon_2-\varepsilon_1}{2-\varepsilon_1}(1-\delta_X(\varepsilon_1)) + \varepsilon.\end{aligned}$$

Because $\varepsilon > 0$ is arbitrary, we have

$$\delta_X(\varepsilon_2) - \delta_X(\varepsilon_1) \le \frac{\varepsilon_2-\varepsilon_1}{2-\varepsilon_1}\Big(1-\delta_X(\varepsilon_1)\Big).$$

Because $\delta_X(\varepsilon_1) \ge 0$, we have

$$\delta_X(\varepsilon_2) - \delta_X(\varepsilon_1) \le \frac{\varepsilon_2-\varepsilon_1}{2-\varepsilon_1},$$

which implies that $\delta_X(\cdot)$ is continuous on $[0,2)$.

(b) Fix $s \in (0,2]$ with $s \le \varepsilon$ and $x, y \in S_X$ and $\|x-y\| = \varepsilon$.

Set

$$t := \frac{s}{\varepsilon}, u := tx + (1-t)\frac{x+y}{\|x+y\|} \text{ and } v := ty + (1-t)\frac{x+y}{\|x+y\|}.$$

Then

$$u - v = t(x-y), \|u-v\| = s \text{ and } \frac{u+v}{2} = \frac{x+y}{\|x+y\|}\left(\frac{t}{2}\,\|x+y\| + 1 - t\right).$$

Thus,

$$\begin{aligned}\left\|\frac{x+y}{\|x+y\|} - \frac{u+v}{2}\right\| &= t - t\left\|\frac{x+y}{2}\right\|\\ &= 1 - \left(1 - t + t\left\|\frac{x+y}{2}\right\|\right)\\ &= 1 - \left\|\frac{u+v}{2}\right\|.\end{aligned}$$

Observe that

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

and

$$\begin{aligned}\left\|\frac{x+y}{\|x+y\|}-\frac{u+v}{2}\right\|\Big/\|u-v\| &= \left(1-\left\|\frac{u+v}{2}\right\|\right)\Big/s\\ &= \left(1-(1-t)-t\left\|\frac{x+y}{2}\right\|\right)\Big/s\\ &= \left(1-\left\|\frac{x+y}{2}\right\|\right)\Big/\|x-y\|.\end{aligned}$$

Hence

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

By taking the infimum over all possible x and y with $\varepsilon=\|x-y\|$ and $x,y\in S_X$, we obtain

$$\frac{\delta_X(s)}{s}\leq\frac{\delta_X(\varepsilon)}{\varepsilon}.$$

(c) Observe that

$$\frac{\delta_X(s)}{s}\leq\frac{\delta_X(t)}{t}\text{ for } s<t\leq 2 \text{ and } \delta_X(t)>0.$$

Hence

$$t\delta_X(s)\leq s\delta_X(t)<t\delta_X(t),$$

which implies that

$$\delta_X(s)<\delta_X(t).$$

Therefore, δ_X is a strictly increasing function. ∎

Remark 2.3.8 *The modulus of convexity δ_X need not be convex on [0,2] and need not be continuous at $t=2$.*

Theorem 2.3.9 *Let X be a Banach space with modulus of convexity δ_X. Then*

$$\|tx+(1-t)y\|\leq 1-2\min\{t,1-t\}\delta_X(\|x-y\|)$$

for all $x,y\in X$ with $\|x\|\leq 1, \|y\|\leq 1$ and all $t\in[0,1]$.

Proof. The result follows from Theorem 2.2.6(b). ∎

Corollary 2.3.10 *Let X be a Banach space with modulus convexity δ_X. Then*

$$\|(1-t)x+ty\|\leq 1-2t(1-t)\delta_X(\|x-y\|)$$

for all $x,y\in X$ with $\|x\|\leq 1, \|y\|\leq 1$ and all $t\in[0,1]$.

Proof. Because $t(1-t) \leq \min\{t,\ 1-t\}$ for all $t \in [0,1]$, the result follows Theorem 2.3.9. $\blacksquare$

Corollary 2.3.11 *Let X be a uniformly convex Banach space with modulus of convexity δ_X. If $r > 0$ and $x, y \in X$ with $\|x\| \leq r$, $\|y\| \leq r$, then*

$$\|tx + (1-t)y\| \leq r\left[1 - 2\min\{t, 1-t\}\delta_X\left(\frac{\|x-y\|}{r}\right)\right] \textit{ for all } t \in (0,1).$$

Theorem 2.3.12 *Let X be a uniformly convex Banach space X. Then there exists a strictly increasing continuous convex function $g : \mathbb{R}^+ \to \mathbb{R}^+$ with $g(0) = 0$ such that*

$$2t(1-t)g(\|x-y\|) \leq 1 - \|(1-t)x + ty\|$$

for all $x, y \in X$ with $\|x\| \leq 1, \|y\| \leq 1$ and all $t \in [0,1]$.

Proof. Let δ_X be the modulus of convexity of X. Define a function $g : \mathbb{R}^+ \to \mathbb{R}^+$ by

$$g(\lambda) = \begin{cases} \frac{1}{2}\int_0^\lambda \delta_X(s)ds & if \quad 0 \leq \lambda \leq 2, \\ g(2) + \frac{1}{2}\delta_X(2)(\lambda - 2) & if \quad \lambda > 2. \end{cases}$$

For $t \in (0,2]$, we have

$$0 < g(t) = \frac{1}{2}\int_0^t \delta_X(s)ds \leq \frac{t}{2}\,\delta_X(t) \leq \delta_X(t). \qquad (\text{as } \delta_X(s) \leq \delta_X(t))$$

From the definition of g, we have

$$g'(t) = \frac{1}{2}\delta_X(t) \text{ for all } t \in [0,2].$$

Hence g' is increasing with $g'(2) = \delta_X(2)/2 = 1/2$, and it follows that g is convex.

Now, let $\|x\| \leq 1$, $\|y\| \leq 1$ and $t \in [0,1]$. Then, we have (see Corollary 2.3.10)

$$\|(1-t)x + ty\| \leq 1 - 2t(1-t)\delta_X(\|x-y\|). \tag{2.8}$$

Hence from (2.8) we have

$$\begin{aligned} 2t(1-t)g(\|x-y\|) &= t(1-t)\int_0^{\|x-y\|} \delta_X(s)ds \\ &\leq t(1-t)\delta_X(\|x-y\|)\|x-y\| \\ &\leq 2t(1-t)\delta_X(\|x-y\|) \\ &\leq 1 - \|(1-t)x + ty\|. \end{aligned}$$

Moreover, for $rs < 2$, the function $s \mapsto g(rs)/s$ is increasing (as $(g(rs)/s)' = [rs\delta_X(rs)/2 - g(rs)]/s^2 \geq 0$). Therefore, g is a strictly increasing continuous convex function. ∎

Using Corollary 2.3.11, we obtain the following, which has important applications in approximation of fixed points of nonlinear mappings in Banach spaces.

Theorem 2.3.13 *Let X be a uniformly convex Banach space and let $\{t_n\}$ be a sequence of real numbers in (0,1) bounded away from 0 and 1. Let $\{x_n\}$ and $\{y_n\}$ be two sequences in X such that*

$$\limsup_{n\to\infty} \|x_n\| \leq a, \ \ \limsup_{n\to\infty} \|y_n\| \leq a \ \text{and} \ \limsup_{n\to\infty} \|t_n x_n + (1-t_n)y_n\| = a$$

for some $a \geq 0$. Then $\lim_{n\to\infty} \|x_n - y_n\| = 0$.

Proof. The case $a = 0$ is trivial. So, let $a > 0$. Suppose, for contradiction, that $\{x_n - y_n\}$ does not converge to 0. Then there exists a subsequence $\{x_{n_i} - y_{n_i}\}$ of $\{x_n - y_n\}$ such that $\inf_i \|x_{n_i} - y_{n_i}\| > 0$. Note $\{t_n\}$ is bounded away from 0 and 1, and there exist two positive numbers α and β such that $0 < \alpha \leq t_n \leq \beta < 1$ for all $n \in \mathbb{N}$. Because $\limsup_{n\to\infty} \|x_n\| \leq a$ and $\limsup_{n\to\infty} \|y_n\| \leq a$, we may assume an $r \in (a, a+1)$ for a subsequence $\{n_i\}$ such that $\|x_{n_i}\| \leq r$, $\|y_{n_i}\| \leq r$, $a < r$. Choose $r \geq \varepsilon > 0$ such that

$$2\alpha(1-\beta)\delta_X(\varepsilon/r) < 1 \text{ and } \|x_{n_i} - y_{n_i}\| \geq \varepsilon > 0 \text{ for all } i \in \mathbb{N}.$$

From Corollary 2.3.11, we have

$$\begin{aligned} \|t_{n_i}x_{n_i} + (1-t_{n_i})y_{n_i}\| &\leq r[1 - 2t_{n_i}(1-t_{n_i})\delta_X(\varepsilon/r)] \\ &\leq r[1 - 2\alpha(1-\beta)\delta_X(\varepsilon/r)] < a \text{ for all } i \in \mathbb{N}, \end{aligned}$$

which contradicts the hypothesis. ∎

We now present the following intersection theorem:

Theorem 2.3.14 (Intersection theorem) – *Let $\{C_n\}$ be a decreasing sequence of nonempty closed convex bounded subsets of a uniformly convex Banach space X. Then $\cap_{n\in\mathbb{N}} C_n$ is a nonempty closed convex subset of X.*

Proof. Let $z \in X$ be a point such that $z \notin C_1, r_n = d(z, C_n)$ and $r = \lim_{n\to\infty} r_n$. Let $\{\varepsilon_n\}$ be a sequence of positive numbers that decreases to zero. Set

$$\begin{aligned} D_n := B_{r+\varepsilon_n}[z] &= \{x \in C_n : \|z - x\| \leq r + \varepsilon_n\}, \\ d_n : &= diam(D_n), \\ d : &= \lim_{n\to\infty} d_n. \end{aligned}$$

Suppose x and y are two elements in D_n such that $\|x-y\| \geq d_n - \varepsilon_n$. Then Corollary 2.3.11 gives

$$\left\| z - \frac{x+y}{2} \right\| \leq \left(1 - \delta_X\left(\frac{\|x-y\|}{r+\varepsilon_n}\right)\right)(r+\varepsilon_n)$$

and hence

$$r_n \leq \left(1 - \delta_X\left(\frac{d_n-\varepsilon_n}{r+\varepsilon_n}\right)\right)(r+\varepsilon_n).$$

This yields a contradiction unless $d=0$. This in turn implies that $\cap_{n\in\mathbb{N}} D_n$ is nonempty, and so is $\cap_{n\in\mathbb{N}} C_n$. $\blacksquare$

Remark 2.3.15 *Theorem 2.3.14 remains valid if the sequence $\{C_n\}$ is replaced by an arbitrary decreasing net of nonempty closed convex bounded subsets of X.*

We now study a weaker type convexity of Banach spaces that is called locally uniform convexity.

Definition 2.3.16 *A Banach space X is said to be locally uniformly convex if for any $\varepsilon > 0$ and $x \in S_X$, there exists $\delta = \delta(x,\varepsilon) > 0$ such that*

$$\|x-y\| \geq \varepsilon \text{ implies that } \left\|\frac{x+y}{2}\right\| \leq 1-\delta \text{ for all } y \in S_X.$$

The *modulus of local convexity* of the Banach space X is

$$\delta_X(x,\varepsilon) = \inf\left\{1 - \frac{\|x+y\|}{2} : y \in S_X, \|x-y\| \geq \varepsilon\right\} \text{ for each } x \in S_X \text{ and } 0 < \varepsilon \leq 2.$$

One may easily see that the Banach space X is locally uniformly convex if $\delta_X(x,\varepsilon) > 0$ for all $x \in S_X$ and $\varepsilon > 0$.

Observation

- Every uniformly convex Banach space is locally uniformly convex.
- By Definition 2.3.16, every locally uniformly convex Banach space is strictly convex.

We now give interesting properties of locally uniformly convex Banach spaces:

Proposition 2.3.17 *Let X be a Banach space. Then the following are equivalent:*

(a) X is locally uniformly convex.

(b) Every sequence $\{x_n\}$ in S_X and $x \in S_X$ with $\|x_n + x\| \to 2$ implies that $x_n \to x$.

Proof. $(a) \Rightarrow (b)$. By locally uniformly convexity of X, $\delta_X(x,\varepsilon) > 0$ for all $\varepsilon > 0$. Therefore,

$$1 - \frac{\|x_n + x\|}{2} \to 0 \text{ implies that } \|x_n - x\| \to 0.$$

$(b) \Rightarrow (a)$. Let $\{x_n\}$ be a sequence in S_X such that $\|x_n + x\| \to 2$ implies that $x_n \to x$. Then

$$\|x_n - x\| \geq \varepsilon > 0 \text{ implies that } \left\|\frac{x_n + x}{2}\right\| < 1.$$

Hence, by the definition of modulus of locally uniform convexity, $\delta_X(x, \varepsilon) > 0$. Therefore, X is locally uniformly convex. $\blacksquare$

The following theorem is a generalization of Theorem 2.2.13.

Theorem 2.3.18 *Every locally uniformly convex Banach space has the Kadec-Klee property.*

Proof. Let X be a locally uniformly convex Banach space. Let $\{x_n\}$ be a sequence in X such that $x_n \rightharpoonup x \in X$ and $\|x_n\| \to \|x\|$. For $x = 0$, $\|x_n\| \to 0$ implies that $x_n \to 0$. Suppose $x \neq 0$. Then

$$\frac{x_n}{\|x_n\|} \rightharpoonup \frac{x}{\|x\|} \Rightarrow \frac{x_n}{\|x_n\|} + \frac{x}{\|x\|} \rightharpoonup 2\frac{x}{\|x\|}.$$

By w-lsc of the norm, we have

$$\begin{aligned} 2 = 2\left\|\frac{x}{\|x\|}\right\| &\leq \liminf_{n\to\infty} \left\|\frac{x_n}{\|x_n\|} + \frac{x}{\|x\|}\right\| \\ &\leq \limsup_{n\to\infty} \left(\frac{\|x_n\|}{\|x_n\|} + \frac{\|x\|}{\|x\|}\right) = 2, \end{aligned}$$

which implies that $\|x_n/(\|x_n\|) + x/(\|x\|)\| \to 2$. By Proposition 2.3.17, we conclude that $x_n/\|x_n\| \to x/\|x\|$. Therefore, $x_n \to x$. $\blacksquare$

2.4 Duality mappings

Definition 2.4.1 *Let X^* be the dual of a Banach space X. Then a multivalued mapping $J : X \to 2^{X^*}$ is said to be a (normalized) duality mapping if*

$$Jx = \{j \in X^* : \langle x, j\rangle = \|x\|^2 = \|j\|_*^2\}.$$

Example 2.4.2 *In a Hilbert space H, the normalized duality mapping is the identity. To see this, let $x \in H$ with $x \neq 0$. Note that $H = H^*$ and*

$$\langle x, x\rangle = \|x\| \cdot \|x\| \text{ implies } x \in Jx.$$

Suppose $y \in Jx$. Then by the definition of J, we have $\langle x, y\rangle = \|x\|\|y\|$ and $\|x\| = \|y\|$. Because

$$\|x - y\|^2 = \|x\|^2 + \|y\|^2 - 2\langle x, y\rangle,$$

it follows that $x = y$. Therefore, $Jx = \{x\}$.

For a complex number, we define the “sign” function by

$$sgn\ \alpha = \begin{cases} 0 & \text{if } \alpha = 0, \\ \alpha/|\alpha| & \text{if } \alpha \neq 0. \end{cases}$$

Observation

- $|sgn\ \alpha| = \begin{cases} 0 & \text{if } \alpha = 0, \\ 1 & \text{if } \alpha \neq 0. \end{cases}$
- $\alpha\ sgn\ \overline{\alpha} = \begin{cases} 0 & \text{if } \alpha = 0, \\ \alpha\overline{\alpha}/|\alpha| = |\alpha| & \text{if } \alpha \neq 0. \end{cases}$

Example 2.4.3 *In the* ℓ_2 *space,*

$$Jx = (|x_1|sgn(x_1), |x_2|sgn(x_2), \cdots, |x_i|sgn(x_i), \cdots), \quad x = \{x_i\} \in \ell_2.$$

Example 2.4.4 *In the* $L_2[0,1]$ $(1 < p < \infty)$ *space, the duality mapping is given by*

$$Jx = \begin{cases} |x|\ sgn(x)/\|x\|, & \textit{if } x \neq 0, \\ 0 & \textit{if } x = 0. \end{cases}$$

Before giving fundamental properties of duality mappings, we need the following notations and definitions:

Let $T : X \to 2^{X^*}$ a multivalued mapping. The domain $Dom(T)$, range $R(T)$, inverse T^{-1}, and graph $G(T)$ of T are defined as

$$\begin{aligned} Dom(T) &= \{x \in X : Tx \neq \emptyset\}, \\ R(T) &= \cup_{x \in Dom(T)} Tx, \\ T^{-1}(y) &= \{x \in X : y \in Tx\}, \\ G(T) &= \{(x, y) \in X \times X^* : y \in Tx, x \in Dom(T)\}. \end{aligned}$$

The graph $G(T)$ of T is a subset of $X \times X^*$.

The mapping T is said to be

(i) *monotone* if $\langle x-y, j_x-j_y\rangle \geq 0$ for all $x, y \in Dom(T)$ and $j_x \in Tx, j_y \in Ty$.

(ii) *strictly monotone* if $\langle x - y, j_x - j_y\rangle > 0$ for all $x, y \in Dom(T)$ with $x \neq y$ and $j_x \in Tx, j_y \in Ty$.

(iii) *α-monotone* if there exists a continuous strictly increasing function $\alpha : [0, \infty) \to [0, \infty)$ with $\alpha(0) = 0$ and $\alpha(t) \to \infty$ as $t \to \infty$ such that

$$\langle x - y, j_x - j_y\rangle \geq \alpha(\|x - y\|)\|x - y\|$$

for all $x, y \in Dom(T),\ j_x \in Tx, j_y \in Ty$.

(iv) *strongly monotone* if T is α-monotone with $\alpha(t) = kt$ for some constant $k > 0$.

(v) *injective* if $Tx \cap Ty = \emptyset$ for $x \neq y$.

The monotone operator $T : Dom(T) \subset X \to 2^{X^*}$ is said to be *maximal monotone* if it has no proper monotone extensions, i.e., if for $(x, y) \in X \times X^*$

$$\langle x - z, y - j_z \rangle \geq 0 \text{ for all } z \in Dom(T) \text{ and } j_z \in Tz \text{ implies } y \in Tx.$$

The mapping $T : Dom(T) \subset X \to X^*$ is said to be *coercive* on a subset C of $Dom(T)$ if there exists a function $c : (0, \infty) \to [-\infty, \infty]$ with $c(t) \to \infty$ as $t \to \infty$ such that $\langle x, Tx \rangle \geq c(\|x\|)\|x\|$ for all $x \in C$.

In other words, T is coercive on C if $\frac{\langle x, Tx \rangle}{\|x\|} \to \infty$ as $\|x\| \to \infty, \ x \in C$.

Observation

- Every monotonically increasing mapping is monotone.
- If H is a Hilbert space and $T : H \to H$ is nonexpansive, then $I - T$ is monotone.

We are now in a position to establish fundamental properties of duality mappings in Banach spaces.

Proposition 2.4.5 *Let X be a Banach space and let $J : X \to 2^{X^*}$ be the normalized duality mapping. Then we have the following:*

(a) $J(0) = \{0\}$.

(b) For each $x \in X, Jx$ is nonempty closed convex and bounded subset of X^*.

(c) $J(\lambda x) = \lambda Jx$ for all $x \in X$ and real λ, i.e., J is homogeneous.

(d) J is multivalued monotone, i.e., $\langle x - y, j_x - j_y \rangle \geq 0$ for all $x, y \in X$, $j_x \in Jx$ and $j_y \in J(y)$.

(e) $\|x\|^2 - \|y\|^2 \geq 2\langle x - y, j \rangle$ for all $x, y \in X$ and $j \in Jy$.

(f) If X^* is strictly convex, J is single-valued.

(g) If X is strictly convex, then J is one-one, i.e., $x \neq y \Rightarrow Jx \cap Jy = \emptyset$.

(h) If X is reflexive with strictly convex dual X^*, then J is demicontinuous.

(i) If X is uniformly convex, then for $x, y \in B_r[0]$, $j_x \in Jx$, $j_y \in Jy$

$$\langle x - y, j_x - j_y \rangle \geq w_r(\|x - y\|)\|x - y\|,$$

where $w_r : \mathbb{R}^+ \to \mathbb{R}^+$ is a function satisfies the conditions:

$$w_r(0) = 0, w_r(t) > 0 \text{ for all } t > 0 \text{ and } t \leq s \Rightarrow w_r(t) \leq w_r(s).$$

Proof. (a) It is obvious.

(b) If $x = 0$, we are done by Part(a). If x is a nonzero element in X, then by the Hahn-Banach theorem (see Corollary 1.6.6), there exists $f \in X^*$ such that $\langle x, f \rangle = \|x\|$ and $\|f\|_* = 1$. Set $j := \|x\|f$. Then $\langle x, j \rangle = \|x\|\langle x, f \rangle = \|x\|^2$ and $\|j\|_* = \|x\|$, and it follows that Jx is nonempty for each $x \neq 0$.

Now suppose $f_1, f_2 \in Jx$ and $t \in (0, 1)$. Because

$$\langle x, f_1 \rangle = \|x\|\|f_1\|_*, \|x\| = \|f_1\|_*$$

and

$$\langle x, f_2\rangle = \|x\|\|f_2\|_*, \|x\| = \|f_2\|_*,$$

we obtain

$$\langle x, tf_1 + (1-t)f_2\rangle = \|x\|(t\|f_1\|_* + (1-t)\|f_2\|_*) = \|x\|^2.$$

Observe that

$$\begin{aligned}\langle x, tf_1 + (1-t)f_2\rangle &\leq \|tf_1 + (1-t)f_2\|_*\|x\| \\ &\leq (t\|f_1\|_* + (1-t)\|f_2\|_*)\|x\| \\ &= \|x\|^2.\end{aligned}$$

Then

$$\|x\|^2 \leq \|x\|\|tf_1 + (1-t)f_2\|_* \leq \|x\|^2,$$

which gives us

$$\|x\|^2 = \|x\|\|tf_1 + (1-t)f_2\|_*,$$

i.e.,

$$\|tf_1 + (1-t)f_2\|_* = \|x\|.$$

Thus,

$$\langle x, tf_1 + (1-t)f_2\rangle = \|x\|\|tf_1 + (1-t)f_2\|_* \text{ and } \|x\| = \|tf_1 + (1-t)f_2\|_*,$$

and this means that $tf_1 + (1-t)f_2 \in Jx$, i.e., Jx is a convex set.

Similarly, one can show that Jx is a closed and bounded set in X^*.

(c) For $\lambda = 0$, it is obvious that $J(0x) = 0Jx$. Assume that $j \in J(\lambda x)$ for $\lambda \neq 0$. First, we show that $J(\lambda x) \subseteq \lambda Jx$. Because $j \in J(\lambda x)$, we have

$$\langle \lambda x, j\rangle = \|\lambda x\|\|j\|_* \text{ and } \|\lambda x\| = \|j\|_*,$$

and it follows that $\langle \lambda x, j\rangle = \|j\|_*^2$. Hence

$$\langle x, \lambda^{-1}j\rangle = \lambda^{-1}\langle \lambda x, \lambda^{-1}j\rangle = \lambda^{-2}\langle \lambda x, j\rangle = \lambda^{-2}\|\lambda x\|\|j\|_* = \|\lambda^{-1}j\|_*^2 = \|x\|^2.$$

This shows that $\lambda^{-1}j \in Jx$, i.e., $j \in \lambda Jx$. Thus, we have $J(\lambda x) \subseteq \lambda Jx$. Similarly, one can show that $\lambda Jx \subseteq J(\lambda x)$. Therefore, $J(\lambda x) = \lambda Jx$.

(d) Let $j_x \in Jx$ and $j_y \in Jy$ for $x, y \in X$. Hence

$$\begin{aligned}\langle x - y, j_x - j_y\rangle &= \langle x, j_x\rangle - \langle x, j_y\rangle - \langle y, j_x\rangle + \langle y, j_y\rangle \\ &\geq \|x\|^2 + \|y\|^2 - \|x\|\|j_y\|_* - \|y\|\|j_x\|_* \\ &\geq \|x\|^2 + \|y\|^2 - 2\|x\|\|y\| \\ &= (\|x\| - \|y\|)^2 \geq 0. \end{aligned} \tag{2.9}$$

(e) Let $j \in Jx$, $x \in X$. Then

$$\begin{aligned}\|y\|^2 &- \|x\|^2 - 2\langle y - x, j\rangle \\ &= \|x\|^2 + \|y\|^2 - 2\langle y, j\rangle \\ &\geq \|x\|^2 + \|y\|^2 - 2\|x\|\|y\| \\ &= (\|x\| - \|y\|)^2 \geq 0. \end{aligned} \tag{2.10}$$

(f) Let $j_1, j_2 \in Jx$ for $x \in X$. Then

$$\langle x, j_1 \rangle = \|j_1\|_*^2 = \|x\|^2$$

and

$$\langle x, j_2 \rangle = \|j_2\|_*^2 = \|x\|^2.$$

Adding the above identities, we have

$$\langle x, j_1 + j_2 \rangle = 2\|x\|^2.$$

Because

$$2\|x\|^2 = \langle x, j_1 + j_2 \rangle \leq \|x\|\|j_1 + j_2\|_*,$$

this implies that

$$\|j_1\|_* + \|j_2\|_* = 2\|x\| \leq \|j_1 + j_2\|_*.$$

It now follows from the fact $\|j_1 + j_2\|_* \leq \|j_1\|_* + \|j_2\|_*$ that

$$\|j_1 + j_2\|_* = \|j_1\|_* + \|j_2\|_*.$$

Because X^* is strictly convex and $\|j_1 + j_2\|_* = \|j_1\|_* + \|j_2\|_*$, then there exists $\lambda \in \mathbb{R}$ such that $j_1 = \lambda j_2$. Because

$$\langle x, j_2 \rangle = \langle x, j_1 \rangle = \langle x, \lambda j_2 \rangle = \lambda \langle x, j_2 \rangle,$$

this implies that $\lambda = 1$ and hence $j_1 = j_2$. Therefore, J is single-valued.

(g) Suppose that $j \in Jx \cap Jy$ for $x, y \in X$. Because $j \in Jx$ and $j \in Jy$, it follows from $\|j\|_*^2 = \|x\|^2 = \|y\|^2 = \langle x, j \rangle = \langle y, j \rangle$ that

$$\|x\|^2 = \langle (x+y)/2, j \rangle \leq \|(x+y)/2\|\|x\|,$$

which gives that

$$\|x\| = \|y\| \leq \|(x+y)/2\| \leq \|x\|.$$

Hence $\|x\| = \|y\| = \|(x+y)/2\|$. Because X is strictly convex and $\|x\| = \|y\| = \|(x+y)/2\|$, we have $x = y$. Therefore, J is one-one.

(h) It suffices to prove demicontinuity of J on the unit sphere S_X. For this, let $\{x_n\}$ be a sequence in S_X such that $x_n \to z$ in X. Then $\|Jx_n\|_* = \|x_n\| = 1$ for all $n \in \mathbb{N}$, i.e., $\{Jx_n\}$ is bounded. Because X is reflexive and hence X^* is also reflexive. Then there exists a subsequence $\{Jx_{n_k}\}$ of $\{Jx_n\}$ in X^* such that $\{Jx_{n_k}\}$ converges weakly to some j in X^*. Because $x_{n_k} \to z$ and $Jx_{n_k} \rightharpoonup j$, then we have

$$\langle z, j \rangle = \lim_{k\to\infty} \langle x_{n_k}, Jx_{n_k} \rangle = \lim_{k\to\infty} \|x_{n_k}\|^2 = 1.$$

Moreover,

$$\begin{aligned} \|j\|_* &\leq \lim_{k\to\infty} \|Jx_{n_k}\|_* = \lim_{k\to\infty} (\|Jx_{n_k}\|_* \, \|x_{n_k}\|) \\ &= \lim_{k\to\infty} \langle x_{n_k}, Jx_{n_k} \rangle = \langle z, j \rangle = \|j\|_*. \end{aligned}$$

This shows that

$$\langle z, j\rangle = \|j\|_*\|z\| \text{ and } \|j\|_* = \|z\|.$$

This implies that $j = Jz$. Thus, every subsequence $\{Jx_{n_i}\}$ converging weakly to $j \in X^*$. This gives $Jx_n \rightharpoonup Jz$. Therefore, J is demicontinuous.

(i) Let $r > 0$ and $w_r : \mathbb{R}^+ \to \mathbb{R}^+$ a function defined by

$$\begin{cases} w_r(0) & = 0; \\ w_r(t) & = \inf\{\frac{\langle x-y, j_x - j_y\rangle}{\|x-y\|} : x, y \in B_r[0], \|x-y\| \geq t, j_x \in Jx, j_y \in Jy\} \\ & \quad \text{if } t \in (0, 2r]; \\ w_r(t) & = w_r(2r); \text{if } t \in (2r, \infty). \end{cases}$$

By (d), we have

$$\langle x - y, j_x - j_y\rangle \geq 0,$$

and it follows that $w_r(t) \geq 0$ for all $t \in \mathbb{R}^+$. It can be readily seen that w_r is nondecreasing. So it remains to prove that $w_r(t) > 0$ for all $t > 0$.

Suppose, for contradiction, that there exists $\lambda \in (0, 2r]$ such that $w_r(\lambda) = 0$. Then there exist sequences $\{x_n\}, \{y_n\}$ in $B_r[0]$ such that

$$\|x_n - y_n\| \geq \lambda > 0 \text{ and } \langle x_n - y_n, j_{x_n} - j_{y_n}\rangle \to 0,$$

where $j_{x_n} \in Jx_n, j_{y_n} \in Jy_n$. We know from (2.9) that

$$(\|x_n\| - \|y_n\|)^2 \leq \langle x_n - y_n, j_{x_n} - j_{y_n}\rangle.$$

We may assume that

$$\lim_{n\to\infty} \|x_n\| = \lim_{n\to\infty} \|y_n\| = a > 0 \text{ (say)}.$$

Notice

$$\begin{aligned} \langle x_n + y_n, j_{x_n} + j_{y_n}\rangle &= 2\|x_n\|^2 + 2\|y_n\|^2 - \langle x_n - y_n, j_{x_n} - j_{y_n}\rangle \\ &\to 4a^2 \end{aligned} \tag{2.11}$$

and

$$\limsup_{n\to\infty} \|x_n + y_n\| \leq \limsup_{n\to\infty}(\|x_n\| + \|y_n\|) = 2a.$$

Moreover, from (2.11), we have

$$\begin{aligned} 4a^2 &= \lim_{n\to\infty} \langle x_n + y_n, j_{x_n} + j_{y_n}\rangle \\ &\leq \liminf_{n\to\infty} \|x_n + y_n\|(\|x_n\| + \|y_n\|) = 2a \liminf_{n\to\infty} \|x_n + y_n\|, \end{aligned}$$

which implies that

$$2a \leq \liminf_{n\to\infty} \|x_n + y_n\|.$$

Thus, we have that $\lim_{n\to\infty} \|x_n + y_n\| = 2a$. By the uniform convexity of X (see Theorem 2.3.13), we obtain that $\lim_{n\to\infty} \|x_n - y_n\| = 0$, which contradicts our assumption that $\|x_n - y_n\| \geq \lambda > 0$. ∎

The inequalities given in the following results are very useful in many applications.

Proposition 2.4.6 *Let X be a Banach space and $J : X \to 2^{X^*}$ the duality mapping. Then we have the following:*

(a) $\|x+y\|^2 \geq \|x\|^2 + 2\langle y, j_x\rangle$ *for all* $x, y \in X$, *where* $j_x \in Jx$.

(b) $\|x+y\|^2 \leq \|y\|^2 + 2\langle x, j_{x+y}\rangle$ *for all* $x, y \in X$, *where* $j_{x+y} \in J(x+y)$.

Proof. (a) Replacing y by $x+y$ in (2.10), we get the inequality.

(b) Replacing x by $x+y$ in (2.10), we get the result. ∎

Proposition 2.4.7 *Let X be a Banach and $J : X \to 2^{X^*}$ a normalized duality mapping. Then for $x, y \in X$, the following are equivalent:*

(a) $\|x\| \leq \|x+ty\|$ *for all* $t > 0$.

(b) There exists $j \in Jx$ *such that* $\langle y, j\rangle \geq 0$.

Proof. (a) ⇒ (b). For $t > 0$, let $f_t \in J(x+ty)$ and define $g_t = f_t/\|f_t\|_*$. Hence $\|g_t\|_* = 1$. Because $g_t \in \|f_t\|_*^{-1} J(x+ty)$, it follows that

$$\begin{aligned} \|x\| &\leq \|x+ty\| = \|f_t\|_*^{-1}\langle x+ty, f_t\rangle \\ &= \langle x+ty, g_t\rangle = \langle x, g_t\rangle + t\langle y, g_t\rangle \\ &\leq \|x\| + t\langle y, g_t\rangle. \quad (\text{as } \|g_t\|_* = 1) \end{aligned}$$

By the Banach-Alaoglu theorem (which states that the unit ball in X^* is weak*ly-compact), the net $\{g_t\}$ has a limit point $g \in X^*$ such that

$$\|g\|_* \leq 1, \ \langle x, g\rangle \geq \|x\| \text{ and } \langle y, g\rangle \geq 0.$$

Observe that

$$\|x\| \leq \langle x, g\rangle \leq \|x\|\|g\|_* = \|x\|,$$

which gives that

$$\langle x, g\rangle = \|x\| \text{ and } \|g\|_* = 1.$$

Set $j = g\|x\|$, then $j \in Jx$ and $\langle y, j\rangle \geq 0$.

(b) ⇒ (a). Suppose for $x, y \in X$ with $x \neq 0$ there exists $j \in Jx$ such that $\langle y, j\rangle \geq 0$. Hence for $t > 0$,

$$\|x\|^2 = \langle x, j\rangle \leq \langle x, j\rangle + \langle ty, j\rangle = \langle x+ty, j\rangle \leq \|x+ty\|\|x\|,$$

which implies that

$$\|x\| \leq \|x+ty\|. \quad ∎$$

Observation

- $Dom(J) = X$.
- J is odd, i.e., $J(-x) = -Jx$.
- J is homogeneous (hence J is positive homogeneous, i.e., $J(\lambda x) = \lambda Jx$ for all $\lambda > 0$).
- J is bounded.

We now consider the duality mappings that are more general than the normalized duality mappings.

Definition 2.4.8 *A continuous strictly increasing function $\mu : \mathbb{R}^+ \to \mathbb{R}^+$ is said to be gauge function if $\mu(0) = 0$ and $\lim_{t\to\infty} \mu(t) = \infty$.*

Definition 2.4.9 *Let X be a normed space and μ a gauge function. Then the mapping $J_\mu : X \to 2^{X^*}$ defined by*

$$J_\mu(x) = \{j \in X^* : \langle x, j\rangle = \|x\|\|j\|_*, \|j\|_* = \mu(\|x\|)\}, \quad x \in X$$

is called the duality mapping with gauge function μ.

In the particular case $\mu(t) = t$, the duality mapping $J_\mu = J$ is called the normalized duality mapping .

In the case $\mu(t) = t^{p-1}$, $p > 1$, the duality mapping $J_\mu = J_p$ is called *the generalized duality mapping* and it is given by

$$J_p(x) := \{j \in X^* : \langle x, j\rangle = \|x\|\|j\|_*, \|j\|_* = \|x\|^{p-1}\}, \quad x \in X.$$

Note that if $p = 2$, then $J_p = J_2 = J$ is the normalized duality mapping.

Remark 2.4.10 *For the gauge function μ, the function $\Phi : \mathbb{R}^+ \to \mathbb{R}^+$ defined by*

$$\Phi(t) = \int_0^t \mu(s)ds$$

is a continuous convex strictly increasing function on $\mathbb{R}^+$. Therefore, Φ has a continuous inverse function Φ^{-1}.

Example 2.4.11 *Let $x = (x_1, x_2, \cdots) \in \ell_p \ \ (1 < p < \infty)$, set*

$$J_\mu(x) = (|x_1|^{p-1}sgn(x_1), |x_2|^{p-1}sgn(x_2), \cdots)$$

and let $\mu(t) = t^{p-1} = t^{p/q}$, where $1/p + 1/q = 1$. Observe that

$$\left(\sum_{i=1}^{\infty} |x_i|^{(p-1)q}\right)^{1/q} = \left(\sum_{i=1}^{\infty} |x_i|^p\right)^{1/q} \text{ and } \ J_\mu(x) \in \ell_q.$$

Moreover,

$$\mu(\|x\|) = \|x\|^{p/q} = \|J_\mu(x)\|_*$$

and

$$\begin{aligned}\langle x, J_\mu(x)\rangle &= \sum_{i=1}^{\infty} x_i|x_i|^{p-1}sgn(x_i) = \sum_{i=1}^{\infty} |x_i|^p = \|x\|^p \\ &= \|x\|\|x\|^{p-1} = \|x\|\mu(\|x\|) = \|x\|\|J_\mu(x)\|_*.\end{aligned}$$

Thus, J_μ is a duality mapping with gauge function μ. Therefore, the generalized duality mapping J_p in ℓ_p space is given by

$$J_p(x) = (|x_1|^{p-1}sgn(x_1), |x_2|^{p-1}sgn(x_2), \cdots), \quad x \in \ell_p.$$

One can easily see the following facts:

(i) $J_\mu(x)$ is a nonempty closed convex set in X^* for each $x \in X$,

(ii) J_μ is a function when X^* is strictly convex.

(iii) If $J_\mu(x)$ is single-valued, then

$$J_\mu(\lambda x) = \frac{sign(\lambda)\mu(\|\lambda x\|)}{\mu(\|x\|)} J_\mu(x) \text{ for all } x \in X \text{ and } \lambda \in \mathbb{R}$$

and

$$\langle x - y, J_\mu(x) - J_\mu(y)\rangle \geq (\mu(\|x\|) - \mu(\|y\|))(\|x\| - \|y\|) \text{ for all } x, y \in X.$$

We now give other interesting properties of the duality mappings J_μ in reflexive Banach spaces.

Theorem 2.4.12 *Let X be a Banach space and J_μ a duality mapping with gauge function μ. Then X is reflexive if and only if $\bigcup_{x\in X} J_\mu(x) = X^*$, i.e., J_μ is onto.*

Proof. Let X be reflexive and let $j \in X^*$. By the Hahn-Banach theorem, there is an $x \in S_X$ such that $\langle x, j\rangle = \|x\|$.

Because μ has the property of Darboux, there exists a constant $t \geq 0$ such that

$$\mu(||tx||) = \mu(t) = \|j\|_*.$$

Because $\langle tx, j\rangle = \|tx\|\|j\|_*$, it follows that $j \in J_\mu(tx)$.

Conversely, suppose that for each $j \in X^*$, there is $x \in X$ such that $j \in J_\mu(x)$. Set $y := x/\|x\|$. Then $\|y\| = 1$ and $\langle y, j\rangle = \|j\|_*$. Hence each continuous functional attains its supremum on the unit ball. By the James theorem, X is reflexive. ∎

Theorem 2.4.13 *Let X be a reflexive Banach space and J a duality mapping with gauge function μ. Then J^{-1} is the duality mapping with gauge μ^{-1}.*

Proof. From Theorem 2.4.12, we obtain

$$J^{-1}(j) = \{x \in X : j \in J_\mu(x)\} \neq \emptyset \text{ for all } j \in X^*.$$

Let J^* be the duality mapping on X^* with gauge μ^{-1}. Observe that $x \in J^{-1}(j)$ if and only if $\langle x, j\rangle = \|x\|\|j\|_*$ and $\|x\| = \mu^{-1}(\|j\|_*)$ or equivalently if and only if $x \in J^*(j)$. Thus,

$$J^*(j) = J^{-1}(j) = \{x \in X : \langle x, j\rangle = \|x\|\|j\|_*, \|x\| = \mu^{-1}(\|j\|_*)\}. \quad ∎$$

Corollary 2.4.14 *Let X be a reflexive Banach space and $J^* : X^* \to X$ the inverse of the normalized duality mapping $J : X \to X^*$. Then*

$$J^*J = I \text{ and } JJ^* = I^* \quad (\text{identity mappings on } X \text{ and } X^*, \text{ respectively}).$$

Theorem 2.4.15 *Let X be a Banach space and let J_μ be the duality mapping with gauge function μ. If X^* is uniformly convex, then J_μ is uniformly continuous on each bounded set in X, i.e., for $\varepsilon > 0$ and $K > 0$, there is a $\delta > 0$ such that*

$$\|x\| \leq K, \|y\| \leq K \text{ and } \|x - y\| < \delta \Rightarrow \|J_\mu(x) - J_\mu(y)\|_* < \varepsilon.$$

Proof. Because X^* is strictly convex, J_μ is single-valued. Suppose $\{x_n\}$ and $\{y_n\}$ are sequences in X such that $\|x_n\| \leq K, \|y_n\| \leq K$ and $\|x_n - y_n\| \to 0$.

Assume that $x_n \to 0$, then $y_n \to 0$. Moreover,

$$\|J_\mu(x_n)\|_* = \mu(\|x_n\|) \to 0 \quad \text{and} \quad \|J_\mu(y_n)\|_* = \mu(\|y_n\|) \to 0.$$

Hence $\|J_\mu(x_n) - J_\mu(y_n)\|_* \to 0$ and we are done.

Suppose $\{x_n\}$ does not converge strongly to zero. There exist $\alpha > 0$ and a subsequence $\{x_{n_k}\}$ of $\{x_n\}$ such that $\|x_{n_k}\| \geq \alpha$. Because $\|x_n - y_n\| \to 0$, one can assume that $\|y_{n_k}\| \geq \alpha/2$. Without loss of generality, we may assume that

$$\|x_n\| \geq \beta \text{ and } \|y_n\| \geq \beta \text{ for some } \beta > 0.$$

Set $u_n := x_n/\|x_n\|$ and $v_n := y_n/\|y_n\|$ so that $\|u_n\| = \|v_n\| = 1$ and

$$\begin{aligned}
\|u_n - v_n\| &= \left\| \frac{x_n\|y_n\| - \|x_n\|y_n}{\|x_n\|\|y_n\|} \right\| \\
&\leq \frac{1}{\beta^2} \Big\| x_n\|y_n\| - x_n\|x_n\| + x_n\|x_n\| - \|x_n\|y_n \Big\| \\
&\leq \frac{1}{\beta^2} \left(\Big| \|y_n\| - \|x_n\| \Big| \|x_n\| + \|x_n\|\|x_n - y_n\| \right) \\
&\leq \frac{1}{\beta^2} (\|y_n - x_n\|K + \|x_n - y_n\|K) \to 0 \text{ as } n \to \infty.
\end{aligned}$$

Because $\|J_\mu(u_n)\|_* = \mu(\|u_n\|) = \mu(1)$ and $\|J_\mu(v_n)\|_* = \mu(\|v_n\|) = \mu(1)$, we have

$$\begin{aligned}
\mu(1) + \mu(1) - \mu(1)\|u_n - v_n\| &\leq \langle u_n, J_\mu(u_n)\rangle + \langle v_n, J_\mu(v_n)\rangle + \langle u_n - v_n, J_\mu(v_n)\rangle \\
&= \langle u_n, J_\mu(u_n)\rangle + \langle u_n, J_\mu(v_n)\rangle \\
&= \langle u_n, J_\mu(u_n) + J_\mu(v_n)\rangle \\
&\leq \|J_\mu(u_n) + J_\mu(v_n)\|_* \leq 2\mu(1).
\end{aligned}$$

This shows that $\lim_{n\to\infty} \|J_\mu(u_n) + J_\mu(v_n)\|_* = 2\mu(1)$. Because X^* is uniformly convex, we have $\|J_\mu(u_n) - J_\mu(v_n)\|_* \to 0$ as $n \to \infty$. Hence

$$\begin{aligned}
&J_\mu(x_n) - J_\mu(y_n) \\
&\quad = [\mu(\|x_n\|)(J_\mu(u_n) - J_\mu(v_n)) + (\mu(\|x_n\|) - \mu(\|y_n\|))J_\mu(v_n)]/\mu(1),
\end{aligned}$$

and it follows that $\|J_\mu(x_n) - J_\mu(y_n)\|_* \to 0$ as $n \to \infty$. ∎

Observation

- If $J_\mu : X \to 2^{X^*}$ is a duality mapping with gauge function μ then
  (i) J_μ is norm to weak* upper semicontinuous.
  (ii) for each $x \in X$, the set $J_\mu(x)$ is convex and weakly closed in X^*;
  (iii) $J_\mu(-x) = -J_\mu(x)$ and $J_\mu(\lambda x) = \frac{\mu(\|\lambda x\|)}{\mu(\|x\|)} J_\mu(x)$ for all $x \in X, \lambda > 0$;
  (iv) each selection of J_μ is a homogeneous single-valued mapping $j : X \to X^*$ satisfying $j(x) \in J_\mu(x)$ for all $x \in X$,
  (v) J_μ is monotone, i.e., $\langle x - y, j_x - j_y\rangle \geq 0$ for all $x, y \in X$ and $j_x \in J_\mu(x)$, $j_y \in J_\mu(y)$;
  (vi) the strict convexity of X implies that J_μ is strictly monotone, i.e.,
  $$\langle x - y, j_x - j_y\rangle > 0 \text{ for all } x, y \in X \text{ and } j_x \in J_\mu(x), j_y \in J_\mu(y);$$
  (vii) the reflexivity of X and strict convexity of X^* imply that J_μ is single-valued monotone and demicontinuous.

One can easily see that the following are reflexive Kadec-Klee Banach spaces:

(a) a Banach space of finite-dimension,

(b) a reflexive Banach space that is locally uniformly convex,

(c) a uniformly convex Banach space.

We now conclude this section with an interesting result concerning a Banach space whose dual has the Kadec-Klee property.

Theorem 2.4.16 *Let X be a reflexive Banach space such that X^* has the Kadec-Klee property. Let $\{x_\alpha\}_{\alpha\in D}$ be a bounded net in X and $x, y \in w_w(\{x_\alpha\}_{\alpha\in D})$. Suppose $\lim_{\alpha\in D} \|tx_\alpha + (1-t)x - y\|$ exists for all $t \in [0, 1]$. Then $x = y$.*

Proof. Because $\lim_{\alpha\in D} \|tx_\alpha + (1-t)x - y\|$ exists (say, r), for each $\varepsilon > 0$, there exists $\alpha_0 \in D$ such that

$$\|tx_\alpha + (1-t)x - y\| \leq r + \varepsilon \text{ for all } \alpha \succeq \alpha_0.$$

It follows that for all $\alpha \succeq \alpha_0$ and $j(x - y) \in J(x - y)$,

$$\langle tx_\alpha + (1-t)x - y, j(x - y)\rangle \leq (r + \varepsilon)\|x - y\|.$$

Because $x \in \omega_w(\{x_\alpha\}_{\alpha\in D})$, we obtain

$$\begin{aligned}
\|x - y\|^2 &= \langle tx + (1-t)x - y, j(x - y)\rangle \\
&\leq \|x - y\|(\lim_{\alpha\in D} \|tx_\alpha + (1-t)x - y\| + \varepsilon), \\
&= (r + \varepsilon)\|x - y\|.
\end{aligned}$$

Taking the limit as $\varepsilon \to 0$, we obtain

$$\|x - y\| \leq r. \tag{2.12}$$

By Proposition 2.4.6 (b), we have

$$\|tx_\alpha + (1-t)x - y\|^2 \le \|x-y\|^2 + 2t\langle x_\alpha - x, j(tx_\alpha + (1-t)x - y)\rangle$$

for all $t \in (0,1]$ and $j(tx_\alpha + (1-t)x - y) \in J(tx_\alpha + (1-t)x - y)$. By (2.12), we have

$$\liminf_{\alpha \in D} \langle x_\alpha - x, j(tx_\alpha + (1-t)x - y)\rangle \ge 0.$$

Hence there exists a sequence $\{\alpha_n\}_{n\in\mathbb{N}}$ such that $\alpha_n \succeq \alpha_m$ for $n \ge m$ and

$$\left\langle x_\alpha - x, j\left(\frac{1}{n}x_\alpha + \left(1 - \frac{1}{n}\right)x - y\right)\right\rangle \ge -\frac{1}{n} \text{ for all } n \in \mathbb{N} \text{ and } \alpha \succeq \alpha_n. \quad (2.13)$$

Set $D_1 = \{\alpha : \alpha \succeq \alpha_1\}$. Without loss of generality, we may assume that $D = D_1$,

$$\omega_w(\{x_\alpha\}_{\alpha\in D}) = \omega_w\{x_\alpha\}_{\alpha \in D_1}$$

and

$$\lim_{\alpha\in D} \|tx_\alpha + (1-t)x - y\| = \lim_{\alpha \in D_1} \|tx_\alpha + (1-t)x - y\| \text{ for all } t \in [0,1].$$

Set $t_\alpha = \inf\{1/n : \alpha \succeq \alpha_n\}$ for all $\alpha \in D$.

We now consider two cases:

Case 1. $\alpha \in D$ *and* $t_\alpha > 0$.

Set $j_\alpha := j(t_\alpha x_\alpha + (1-t_\alpha)x - y)$. Then

$$\langle x - y, j_\alpha\rangle = \|t_\alpha x_\alpha + (1-t_\alpha)x - y\|^2 - t_\alpha\langle x_\alpha - x, j_\alpha\rangle \quad (2.14)$$

and

$$\|j_\alpha\| = \|t_\alpha x_\alpha + (1-t_\alpha)x - y\|. \quad (2.15)$$

By (2.13), we have

$$\langle x_\alpha - x, j_\alpha\rangle \ge -t_\alpha. \quad (2.16)$$

Case 2. $\alpha \in D$ *and* $t_\alpha = 0$.

In this case, we can choose a subsequence $\{j((1/n_k)x_\alpha + (1-1/n_k)x - y)\}_{k\in\mathbb{N}}$ which is weakly convergent to j, and set $j_\alpha := j$. It follows from (2.13) that

$$\langle x_\alpha - x, j_\alpha\rangle \ge 0. \quad (2.17)$$

Observe that

$$\begin{aligned} \|j_\alpha\| &\le \liminf_{k\to\infty} \left\| j\left(\frac{1}{n_k}x_\alpha + \left(1 - \frac{1}{n_k}\right)x - y\right)\right\| \\ &= \lim_{k\to\infty} \left\|\frac{1}{n_k}x_\alpha + \left(1 - \frac{1}{n_k}\right)x - y\right\| = \|x-y\|. \end{aligned}$$

On the other hand, we have

$$\begin{aligned}\langle x-y, j_\alpha\rangle &= \lim_{k\to\infty}\Big\langle x-y, j\Big(\frac{1}{n_k}x_\alpha+\Big(1-\frac{1}{n_k}\Big)x-y\Big)\Big\rangle \\ &= \lim_{k\to\infty}\Big(\|\frac{1}{n_k}x_\alpha+\Big(1-\frac{1}{n_k}\Big)x-y\|^2 \\ &\quad -\frac{1}{n_k}\Big\langle x_\alpha-x, j\Big(\frac{1}{n_k}x_\alpha+\Big(1-\frac{1}{n_k}\Big)x-y\Big)\Big\rangle\Big) \\ &= \|x-y\|^2. \end{aligned} \tag{2.18}$$

Therefore,

$$\|j_\alpha\| = \|x-y\| \tag{2.19}$$

and $j_\alpha \in J(x-y)$.

We note that by the Kadec-Klee property of X^*, the sequence $\{j((1/n_k)x_\alpha + (1-1/n_k)x-y)\}_{k\in\mathbb{N}}$ converges strongly to j_α.

Now from the net $\{x_\alpha\}_{\alpha\in D}$, we choose a subset $\{\alpha_\beta\}_{\beta\in\overline{D}}$ such that $\{x_{\alpha_\beta}\}_{\beta\in\overline{D}}$ converges weakly to $y \in w_w(\{x_\alpha\}_{\alpha\in D})$ and $\{j_{\alpha_\beta}\}_{\beta\in\overline{D}}$ converges weakly to $\overline{j}$. Then by (2.15) and (2.19) we get

$$\|\overline{j}\|_* \le \|x-y\|$$

and by (2.14) and (2.18), we get

$$\langle x-y, \overline{j}\rangle = \|x-y\|^2.$$

Hence $\overline{j} \in J(x-y)$. Because X is reflexive and X^* has the Kadec-Klee property, the space X^* has also the Kadec property and this implies that $\{j_{\alpha_\beta}\}_{\beta\in\overline{D}}$ converges strongly to $\overline{j}$. It follows from (2.16) and (2.17) that

$$\langle y-x, \overline{j}\rangle \ge 0,$$

i.e., $\|x-y\|^2 \le 0$. Therefore, $x = y$. ∎

Corollary 2.4.17 *Let X be a reflexive Banach space such that its dual X^* has the Kadec-Klee property. Let $\{x_n\}$ be a bounded sequence in X and $p, q \in \omega_w(\{x_n\})$. Suppose $\lim_{n\to\infty}\|tx_n+(1-t)p-q\|$ exists for all $t \in [0,1]$. Then $p = q$.*

2.5 Convex functions

Let X be a linear space and $f : X \to (-\infty, \infty]$ a function. Then

(i) f is said to be *convex* if $f(\lambda x + (1-\lambda)y) \le \lambda f(x) + (1-\lambda)f(y)$ for all $x, y \in X$ and $\lambda \in [0,1]$;

(ii) f is said to be *strictly convex* if $f(\lambda x+(1-\lambda)y) < \lambda f(x)+(1-\lambda)f(y)$ for all $\lambda \in (0,1)$ and $x, y \in X$ with $x \ne y$, $f(x) < \infty$, $f(y) < \infty$;

(iii) f is said to be *proper if* there exists $x \in X$ such that $f(x) < \infty$;

(iv) $Dom(f) = \{x \in X : f(x) < \infty\}$ is called *domain or effective domain*;

(v) f is said to be *bounded below* if there exists a real number α such that $\alpha \leq f(x)$ for all $x \in X$;

(vi) the set $epif = \{(x, \alpha) : x \in X, \alpha \in \mathbb{R}, f(x) \leq \alpha\}$ is called the *epigraph* of f.

Let C be a subset of X. Then the function i_C on X defined by

$$i_C(x) = \begin{cases} 0 & if \quad x \in C, \\ \infty & if \quad x \notin C \end{cases}$$

is called the *indicator function*.

Observation

- i_C is proper if and only if C is nonempty.
- $dom(i_C) = C$.
- The set C is convex if and only if its indicator function i_C is convex.
- The domain of each convex function is convex.

Let X be a topological space and $f : X \to (-\infty, \infty]$ a proper function. Then f is said to be *lower semicontinuous* (*l.s.c.*) at $x_0 \in X$ if

$$f(x_0) \leq \liminf_{x \to x_0} f(x_0) = \sup_{V \in U_{x_0}} \inf_{x \in V} f(x),$$

where U_{x_0} is a base of neighborhoods of the point $x_0 \in X$. f is said to be *lower semicontinuous on* X if it is lower semicontinuous on each point of X, i.e., for each $x \in X$

$$x_n \to x \Rightarrow f(x) \leq \liminf_{n \to \infty} f(x_n).$$

We now discuss some elementary properties of convex functions:

Proposition 2.5.1 *Let X be a linear space and $f : X \to (-\infty, \infty]$ a function. Then f is convex if and only if its epigraph is a convex subset of $X \times \mathbb{R}$.*

Proof. Suppose f is convex. Then for $(x, \alpha), (y, \beta)$ in $epif$, we have

$$f((1-t)x + ty) \quad \leq \quad (1-t)f(x) + tf(y) \leq (1-t)\alpha + t\beta \;\; \text{for all } t \in [0, 1].$$

This implies that $((1-t)x + ty, (1-t)\alpha + t\beta) \in epif$.

Conversely, suppose that $epif$ is convex. Then $Dom(f)$ is also convex. Because for $x, y \in Dom(f)$ and $(x, f(x))$, $(y, f(y)) \in epif$, we have

$$((1-t)x + ty, (1-t)f(x) + tf(y)) \in \; epif \text{ for all } t \in [0, 1].$$

Thus, by the definition of $epif$,

$$f((1-t)x + ty) \leq (1-t)f(x) + tf(y). \quad \blacksquare$$

Proposition 2.5.2 *Let X be a topological space and $f : X \to (-\infty, \infty]$ a function. Then the following statements are equivalent:*

(a) f is lower semicontinuous.

(b) For each $\alpha \in \mathbb{R}$, the level set $\{x \in X : f(x) \leq \alpha\}$ is closed.

(c) The epigraph of the function f, $\{(x, \alpha) \in X \times \mathbb{R} : f(x) \leq \alpha\}$ is closed.

Proof. We recall that

$$\liminf_{x \to x_0} f(x) = \sup_{V \in U_{x_0}} \inf_{x \in V} f(x).$$

$(a) \Rightarrow (b)$. Let $\alpha \in \mathbb{R}$ and let $x_0 \in X$ with $f(x_0) > \alpha$. Because f is lower semicontinuous, there exists $V_0 \in U_{x_0}$ such that $\inf_{x \in V_0} f(x) > \alpha$. Hence $V_0 \subset \{x \in X : f(x) > \alpha\}$. Consequently, $\{x \in X : f(x) > \alpha\}$ is open and hence $\{x \in X : f(x) \leq \alpha\}$ is closed.

$(b) \Rightarrow (a)$. Let $x_0 \in Dom(f)$, $\varepsilon > 0$ and $V_\varepsilon = \{x \in X : f(x) > f(x_0) - \varepsilon\}$. Because each level set of f is closed, it follows that $V_\varepsilon \in U(x_0)$. Because $\inf_{x \in V_\varepsilon} f(x) \geq f(x_0) - \varepsilon$, it follows that $\liminf_{x \to x_0} f(x) \geq f(x_0) - \varepsilon$. As ε is arbitrarily chosen, we conclude that (a) holds.

$(a) \Leftrightarrow (c)$. Define $\varphi : X \times \mathbb{R} \to (-\infty, \infty]$ by $\varphi(x, \alpha) = f(x) - \alpha$. Then, f is $l.s.c.$ on $X \Leftrightarrow \varphi$ is $l.s.c.$ on $X \times \mathbb{R}$. Because $epi f$ is a level set of φ, therefore, the conclusion holds. ∎

Proposition 2.5.3 *Let C be a nonempty closed convex subset of a Banach space X and $f : C \to (-\infty, \infty]$ a convex function. Then f is lower semicontinuous in the norm topology if and only if f is lower semicontinuous in the weak topology.*

Proof. Set $F_\alpha := \{x \in C : f(x) \leq \alpha\}$, $\alpha \in \mathbb{R}$. Then F_α is convex. Indeed, for $x, y \in F_\alpha$

$$\begin{aligned} f(\lambda x + (1-\lambda)y) &\leq \lambda f(x) + (1-\lambda) f(y) \\ &\leq \lambda\alpha + (1-\lambda)\alpha = a \text{ for all } \lambda \in [0,1]. \end{aligned}$$

It follows from Proposition 1.9.13 (which states that for a convex subset C in a normed space X, C is closed if and only if C is weakly closed) that F_α is closed if and only if F_α is weakly closed, i.e., F_α is closed in the weak topology. ∎

Before presenting an important result, we first establish a preliminary result:

Theorem 2.5.4 *Let X be a compact topological space and $f : X \to (-\infty, \infty]$ a lower semicontinuous function. Then there exists an element $x_0 \in X$ such that*

$$f(x_0) = \inf\{f(x) : x \in X\}.$$

Proof. Set $G_\alpha := \{x \in X : f(x) > \alpha\}$, $\alpha \in \mathbb{R}$. One may easily see that each G_α is open and $X = \bigcup_{\alpha\in\mathbb{R}} G_\alpha$. By compactness of X, there exists a finite family $\{G_{\alpha_i}\}_{i=1}^n$ of $\{G_\alpha\}_{\alpha\in\mathbb{R}}$ such that

$$X = \bigcup_{i=1}^{n} G_{\alpha_i}.$$

Suppose $\alpha_0 = \min\{\alpha_1, \alpha_2, \cdots, \alpha_n\}$. This gives $f(x) > \alpha_0$ for all $x \in X$. It follows that $\inf\{f(x) : x \in X\}$ exists. Let $m = \inf\{f(x) : x \in X\}$. Let β be a number such that $\beta > m$. Set $F_\beta := \{x \in X : f(x) \leq \beta\}$. Then F_β is a nonempty closed subset of X; and hence, by the intersection property, we have

$$\bigcap_{\beta > m} F_\beta \neq \emptyset.$$

Therefore, for any point x_0 of this intersection, we have $m = f(x_0)$. ∎

Theorem 2.5.5 *Let C be a weakly compact convex subset of a Banach space and $f : C \to (-\infty, \infty]$ a proper lower semicontinuous convex function. Then there exists $x_0 \in Dom(f)$ such that $f(x_0) = \inf\{f(x) : x \in C\}$.*

Proof. Because f is proper, there exists $u \in C$ such that $f(u) < \infty$. Then the set $C_0 = \{x \in C : f(x) \leq f(u)\}$ is nonempty. Because the set C_0 is closed and convex subset of C, it follows that C_0 is weakly compact. Applying Proposition 2.5.3, we have that f is lower semicontinuous in the weak topology. By Theorem 2.5.4, there exists $x_0 \in C_0 \subset C$ such that

$$f(x_0) = \inf\{f(x) : x \in C_0\} = \inf\{f(x) : x \in C\}. \quad ∎$$

Remark 2.5.6 *If f is strictly convex function in Theorem 2.5.5, then $x_0 \in C$ is the unique point such that $f(x_0) = \inf\limits_{x\in C} f(x)$.*

Recall that every closed convex bounded subset of a reflexive Banach space is weakly compact. Using this fact, we have

Theorem 2.5.7 *Let X be a reflexive Banach space and $f : X \to (-\infty, \infty]$ a proper lower semicontinuous convex function. Then for every nonempty closed convex bounded subset C of X, there exists a point $x_0 \in Dom(f)$ such that $f(x_0) = \inf\limits_{x\in C} f(x)$.*

In Theorem 2.5.7, the boundedness of C may be replaced by the weaker assumption

$$\lim_{x\in C, \|x\|\to\infty} f(x) = \infty.$$

Theorem 2.5.8 *Let C be a nonempty closed convex subset of a reflexive Banach space X and $f : C \to (-\infty, \infty]$ a proper lower semicontinuous convex function such that $f(x_n) \to \infty$ as $\|x_n\| \to \infty$. Then there exists $x_0 \in Dom(f)$ such that*

$$f(x_0) = \inf\{f(x) : x \in C\}.$$

Proof. Let $m = \inf\{f(x) : x \in C\}$. Choose a minimizing sequence $\{x_n\}$ in C, i.e., $f(x_n) \to m$. If $\{x_n\}$ is not bounded, there exists a subsequence $\{x_{n_i}\}$ of $\{x_n\}$ such that $\|x_{n_i}\| \to \infty$. From the hypothesis, we have $f(x_{n_i}) \to \infty$, which contradicts $m \neq \infty$. Hence $\{x_n\}$ is bounded. By the reflexivity X, there exists a subsequence $\{x_{n_j}\}$ of $\{x_n\}$ such that $x_{n_j} \rightharpoonup x_0 \in C$. Because f is lower semicontinuous in the weak topology, we have

$$m \leq f(x_0) \leq \liminf_{j\to\infty} f(x_{n_j}) = \lim_{n\to\infty} f(x_n) = m.$$

Therefore, $f(x_0) = m$. ∎

Differentiation of convex functions – Let X be a normed space and $\varphi : X \to (-\infty, \infty]$ a function. Then the limit

$$\lim_{t\to 0} \frac{\varphi(x+ty) - \varphi(x)}{t} = \inf_{t>0} \frac{\varphi(x+ty) - \varphi(x)}{t}$$

is said to be the *directional derivative* of φ at the point $x \in X$ in the direction $y \in X$. If it exists, it is denoted by $\varphi'(x, y)$.

The function φ is said to be *Gâteaux differentiable at a point* $x \in X$ if there exists a continuous linear functional j on X such that $\langle y, j\rangle = \varphi'(x, y)$ for all $y \in X$. The element j, denoted by $\varphi'(x)$ or $\nabla\varphi(x)$ (*i.e., grad*$\varphi(x)$) is called the *Gâteaux derivative* of φ at x.

One can easily see from the definition of Gâteaux derivative of φ that

(i) $\varphi'(x)(0) = 0$,

(ii) $\varphi'(x)(\lambda y) = \lambda \lim_{t\to 0} \frac{\varphi(x+t\lambda y) - \varphi(x)}{t} = \lambda\varphi'(x)(y)$ for all $\lambda \in \mathbb{R}$, i.e., $\varphi'(x)(\cdot)$ is homogeneous over $\mathbb{R}$.

Remark 2.5.9 *If the function φ is Gâteaux differentiable at $x \in X$, then there exists $j = \varphi'(x) \in X^*$ such that*

$$\frac{d}{dt}\varphi(x+ty)\bigg|_{t=0} = \langle y, \varphi'(x)\rangle = \langle y, j\rangle \text{ for all } y \in X.$$

Let X be a normed space and $\varphi : X \to (-\infty, \infty]$ a function. The function φ is said to be *Fréchet differentiable* at a point $x \in X$ if there exists a continuous linear functional j on X such that

$$\lim_{\|y\|\to 0} \frac{|\varphi(x+y) - \varphi(x) - \langle y, j\rangle|}{\|y\|} = 0.$$

In this case, the element j denoted by $d\varphi(x)$ is called the *Fréchet derivative* of φ at the point x.

Proposition 2.5.10 *Let X be a normed space and $\varphi : X \to (-\infty, \infty]$ a function. If φ is Fréchet differentiable at x, then φ is Gâteaux differentiable at x.*

Proof. Because φ is Fréchet differentiable at x,

$$\lim_{\|y\|\to 0} \frac{|\varphi(x+y)-\varphi(x)-d\varphi(x)y|}{\|y\|} = 0. \tag{2.20}$$

Set $y = ty_0$ for $t > 0$ and for any fixed $y_0 \neq 0$. From (2.20), we obtain

$$\lim_{t\to 0} \frac{|\varphi(x+ty_0)-\varphi(x)-td\varphi(x)y_0|}{t\|y_0\|} = 0,$$

which implies that

$$\lim_{t\to 0} \frac{\varphi(x+ty_0)-\varphi(x)}{t} = d\varphi(x)y_0.$$

Hence $d\varphi \in X^*$ and φ is Gâteaux differentiable at x. ∎

The following example shows that the converse of Proposition 2.5.10 is not true.

Example 2.5.11 *Let $X = \mathbb{R}^2$ be a normed space with norm $\|\cdot\|_2$ and $\varphi : X \to \mathbb{R}$ a function defined by*

$$\varphi(x,y) = \begin{cases} x^3y/(x^4+y^2) & \text{if } (x,y) \neq (0,0), \\ 0 & \text{if } (x,y) = (0,0). \end{cases}$$

One may easily see that φ is Gâteaux differentiable at 0 with Gâteaux derivative $\varphi'(0) = 0$. Because for $(h,k) \in X$, we have

$$\frac{|\varphi(h,k)|}{\|(h,k)\|_2} = \frac{|h^3k|}{(h^4+k^2)(h^2+k^2)^{1/2}} = \frac{1}{2(1+h^2)^{1/2}} \text{ for } k = h^2.$$

Therefore, φ is not Fréchet differentiable.

Observation

- Every Fréchet differentiable function is Gâteaux differentiable.
- If φ is Fréchet differentiable at x, then φ is continuous at x.
- If φ is Gâteaux differentiable at x, then φ is not necessarily continuous at x (e.g., the function $\varphi : \mathbb{R}^2 \to \mathbb{R}$ defined by

  $$\varphi(x,y) = \frac{2y\; exp(-x^{-2})}{y^2 + exp(-2x^{-2})},\ x \neq 0 \text{ and } \varphi(x,y) = 0,\ x = 0$$

  is Gâteaux differentiable at zero, but not continuous at zero).
- If φ is Gâteaux differentiable at x, then $\varphi(x+ty) \to \varphi(x)$ as $t \to 0$ (i.e., if $x_n \to x$ along a line, then $\varphi(x_n) \to \varphi(x)$).

Let X be a Banach space and $\varphi : X \to (-\infty, \infty]$ a proper convex function. Then an element $j \in X^*$ is said to be a *subgradient of* φ *at the point* $x \in X$ if

$$\varphi(x) - \varphi(y) \leq \langle x - y, j \rangle \text{ for all } y \in X.$$

The set (possibly nonempty)

$$\{j \in X^* : \varphi(x) - \varphi(y) \leq \langle x - y, j \rangle \text{ for all } y \in X\},$$

of subgradients of φ at $x \in X$ is called the *subdifferential of* φ *at* $x \in X$. Thus, the subdifferential of a proper convex function φ is a mapping $\partial\varphi : X \to 2^{X^*}$ (generally multivalued) defined by

$$\partial\varphi(x) = \{j \in X^* : \varphi(x) - \varphi(y) \leq \langle x - y, j \rangle \text{ for all } y \in X\}.$$

The domain of the subdifferential $\partial\varphi$ is denoted and defined by

$$\text{Dom}(\partial\varphi) = \{x \in X : \partial\varphi(x) \neq \emptyset\}.$$

Remark 2.5.12 *If* φ *is not the constant* ∞*, then* $Dom(\partial\varphi)$ *is a subset of* $Dom(\varphi)$*.*

Observation

- $\partial\varphi(x)$ is always for every $x \in X$ nonempty if φ is continuous.
- $\partial\varphi(x)$ is always a closed convex set in X^*.
- $\partial(\lambda\varphi(x)) = \lambda\partial\varphi(x)$, i.e., $\partial\varphi(x)$ is homogeneous.
- φ has a minimum value at $x_0 \in Dom(\partial\varphi)$ if and only if $0 \in \partial\varphi(x_0)$.
- $Dom(\partial\varphi) = Dom(\varphi)$ if φ is lower semicontinuous on X.
- For a lower semicontinuous proper convex function φ on a reflexive Banach space X, $\partial\varphi$ is maximal monotone.

The following results are of fundamental importance in the study of convex functions. We begin with a basic result.

Proposition 2.5.13 *Let* C *be a nonempty closed convex subset of a Banach space* X *and* i_C *the indicator function of* C*, i.e.,*

$$i_C(x) = \begin{cases} 0 & if\ x \in C, \\ \infty & otherwise. \end{cases}$$

Then $\partial i_C(x) = \{j \in X^* : \langle x - y, j \rangle \geq 0 \text{ for all } y \in C\},\ x \in C.$

Proof. Because the indicator function is convex and lower semicontinuous function on X, by the subdifferentiability of i_C, we have

$$\partial i_C(x) = \{j \in X^* : i_C(x) - i_C(y) \leq \langle x - y, j \rangle \text{ for all } y \in C\}. \qquad \blacksquare$$

Remark 2.5.14 *$Dom(i_C) = Dom(\partial i_C) = C$ and $\partial i_C(x) = \{0\}$ for each $x \in int(C)$.*

We now give a relation between Gâteaux differentiability and subdifferentiability.

Theorem 2.5.15 *Let X be a Banach space and $\varphi : X \to (-\infty, \infty]$ a proper convex function. If φ is Gâteaux differentiable at a point $x_0 \in X$, then $\partial\varphi(x_0) = \{\varphi'(x_0)\}$, i.e., the subdifferential of φ at $x_0 \in X$ is a singleton set $\{\varphi'(x_0)\}$ in X^*.*

Conversely, if φ is continuous at x_0 and $\partial\varphi(x_0)$ contains a singleton element, then φ is Gâteaux differentiable at x_0 and $\varphi'(x_0) = \partial\varphi(x_0)$.

Proof. Let φ be Gâteaux differentiable at $x_0 \in X$. Then

$$\langle y, \varphi'(x_0)\rangle = \lim_{t\to 0} \frac{\varphi(x_0 + ty) - \varphi(x_0)}{t} \quad \text{for all } y \in X.$$

Notice

$$\varphi(x_0 + \lambda(z - x_0)) = \varphi((1-\lambda)x_0 + \lambda z) \le (1-\lambda)\varphi(x_0) + \lambda\varphi(z) \text{ for all } \lambda \in (0,1).$$

Set $y := z - x_0$. Then, we have

$$\varphi(x_0 + \lambda y) \le \varphi(x_0) + \lambda[\varphi(x_0 + y) - \varphi(x_0)].$$

Thus,

$$\frac{\varphi(x_0 + \lambda y) - \varphi(x_0)}{\lambda} \le \varphi(x_0 + y) - \varphi(x_0),$$

which implies that

$$\varphi(x_0) - \varphi(x_0 + y) \quad \le \quad -\langle y, \varphi'(x_0)\rangle = \langle x_0 - (x_0 + y), \varphi'(x_0)\rangle \text{ for all } y \in X,$$

i.e., $\varphi'(x_0) \in \partial\varphi(x_0)$.

Now, let $j_{x_0} \in \partial\varphi(x_0)$. Then, we have

$$\varphi(x_0) - \varphi(u) \le \langle x_0 - u, j_{x_0}\rangle \text{ for all } u \in X.$$

Therefore,

$$\frac{\varphi(x_0 + \lambda h) - \varphi(x_0)}{\lambda} \ge \langle h, j_{x_0}\rangle \text{ for all } \lambda > 0,$$

it follows that

$$\langle h, \varphi'(x_0) - j_{x_0}\rangle \ge 0 \text{ for all } h \in X,$$

i.e., $j_{x_0} = \varphi'(x_0)$. Thus, φ is Gâteaux differentiable at x_0 and $\varphi'(x_0) = \partial\varphi(x_0)$. ∎

Corollary 2.5.16 *Let X be a Banach space and $\varphi : X \to (-\infty, \infty]$ a proper convex function. Then φ is Gâteaux differentiable at $x \in int(dom(\varphi))$ if and only if it has a unique subgradient $\partial\varphi(x) = \{\varphi'(x)\}$, i.e., the subdifferential of φ at x is a singleton set in X^*. In this case*

$$\frac{d}{dt}\varphi(x+ty)\Big|_{t=0} = \langle y, \partial\varphi(x)\rangle = \langle y, \varphi'(x)\rangle \text{ for all } y \in X.$$

Theorem 2.5.17 *Let X be a Banach space, $J_\mu : X \to 2^{X^*}$ a duality mapping with gauge function μ, and $\Phi(\|x\|) = \int_0^{\|x\|} \mu(s)ds$, $0 \neq x \in X$. Then*

$$J_\mu(x) = \partial\Phi(\|x\|).$$

Proof. Because μ is a strictly increasing and continuous function, it follows that Φ is differentiable and hence $\Phi'(t) = \mu(t)$, $t \geq 0$. Then Φ is a convex function.

First, we show $J_\mu(x) \subseteq \partial\Phi(\|x\|)$. Let $x \neq 0$, and $j \in J_\mu(x)$. Then $\langle x, j\rangle = \|x\|\|j\|_*, \|j\|_* = \mu(\|x\|)$. In order to prove $j \in \partial\Phi(\|x\|)$, i.e., $\Phi(\|x\|) - \Phi(\|y\|) \leq \langle x-y, j\rangle$ for all $y \in X$, we assume that $\|y\| > \|x\|$. Then

$$\|j\|_* = \mu(\|x\|) = \Phi'(\|x\|) \leq \frac{\Phi(\|y\|) - \Phi(\|x\|)}{\|y\| - \|x\|},$$

which yields

$$\begin{aligned} \Phi(\|x\|) - \Phi(\|y\|) &\leq \|j\|_*(\|x\| - \|y\|) \\ &\leq \langle x, j\rangle - \langle y, j\rangle \\ &= \langle x-y, j\rangle. \end{aligned}$$

In a similar way, if $\|x\| > \|y\|$, we have

$$\Phi(\|x\|) - \Phi(\|y\|) \leq \langle x-y, j\rangle.$$

In the case when $\|x\| = \|y\|$, we have

$$\begin{aligned} \langle y-x, j\rangle &= \langle y, j\rangle - \langle x, j\rangle \\ &\leq \|y\|\|j\|_* - \|x\|\|j\|_* \quad (\text{as } \langle x, j\rangle = \|x\|\|j\|_*) \\ &\leq \|j\|_*(\|y\| - \|x\|), \end{aligned}$$

and it follows that

$$\Phi(\|x\|) - \Phi(\|y\|) = 0 = \|j\|_*(\|x\| - \|y\|) \leq \langle x-y, j\rangle.$$

Hence $j \in \partial\Phi(\|x\|)$. Thus, $J_\mu(x) \subseteq \partial\Phi(\|x\|)$ for all $x \neq 0$.

We now prove $\partial\Phi(\|x\|) \subseteq J_\mu(x)$ for all $x \neq 0$. Suppose $j \in \partial\Phi(\|x\|)$ for $0 \neq x \in X$. Then

$$\begin{aligned} \|x\|\|j\|_* &= \sup\{\langle y, j\rangle\|x\| : \|y\| = 1\} \\ &= \sup\{\langle y, j\rangle : \|x\| = \|y\| = 1\} \\ &\leq \sup\{\langle y, j\rangle : \|x\| = \|y\|\} \\ &\leq \sup\{\langle x, j\rangle + \Phi(\|y\|) - \Phi(\|x\|) : \|x\| = \|y\|\} \\ &\leq \|x\|\|j\|_*. \quad (\text{as } \langle y, j\rangle \leq \langle x, j\rangle + \Phi(\|y\|) - \Phi(\|x\|)). \end{aligned}$$

Thus, $\langle x, j \rangle = \|x\|\|j\|_*$. To see $j \in J_\mu(x)$, we show that $\|j\|_* = \mu(\|x\|) = \Phi'(\|x\|)$. Because

$$\Phi(\|x\|) - \Phi(t\|x\|) \leq \langle x - tx, j \rangle = (1-t)\|x\|\|j\|_* \ \text{ for all } t > 0,$$

this implies that

$$\|j\|_* \leq \frac{\Phi(t\|x\|) - \Phi(\|x\|)}{t\|x\| - \|x\|}. \tag{2.21}$$

It follows from (2.21) that

$$\|j\|_* \leq \frac{\Phi(t\|x\|) - \Phi(\|x\|)}{t\|x\| - \|x\|} \qquad \text{if } t > 1$$

and

$$\frac{\Phi(\|x\|) - \Phi(t\|x\|)}{\|x\| - t\|x\|} \leq \|j\|_* \qquad \text{if } t < 1.$$

Taking the limit as $t \to 1$, we get

$$\|j\|_* = \Phi'(\|x\|) = \mu(\|x\|).$$

Thus, $\partial\Phi(\|x\|) \subseteq J_\mu(x)$. Therefore, $J_\mu(x) = \partial\Phi(\|x\|)$ for all $x \neq 0$. ∎

Remark 2.5.18 *Both the sets $J_\mu(x)$ and $\partial\Phi(\|x\|)$ are equal to $\{0\}$ if $x = 0$.*

Corollary 2.5.19 *For $p \in (1, \infty)$, the generalized duality mapping J_p is the subdifferential of the functional $\|\cdot\|^p/p$.*

Proof. Define $\mu(t) = t^{p-1}$, $p > 1$. Hence

$$\Phi(t) = \int_0^t \mu(s)ds = \int_0^t s^{p-1}ds = \frac{t^p}{p}.$$

Therefore, $J_p(\cdot) = \partial(\|\cdot\|^p/p)$. ∎

Corollary 2.5.20 *Let X be a Banach space and $\varphi(x) = \|x\|^2/2$. Then the subdifferential $\partial\varphi$ coincides with the normalized duality mapping $J : X \to 2^{X^*}$ defined by*

$$Jx = \{j \in X^* : \langle x, j \rangle = \|x\|\|j\|_*, \|j\|_* = \|x\|\}, x \in X.$$

Theorem 2.5.21 *Let X be a Banach space. Then*

$$\partial\|x\| = \{j \in X^* : \langle x, j \rangle = \|x\|\|j\|_*, \|j\|_* = 1\} \text{ for all } x \in X \setminus \{0\}.$$

Proof. Let $j \in \partial\|x\|$. Then

$$\langle y - x, j\rangle \le \|y\| - \|x\| \le \|y - x\| \text{ for all } y \in X. \tag{2.22}$$

It follows that $j \in X^*$ and $\|j\|_* \le 1$. It is clear from (2.22) that $\|x\| \le \langle x, j\rangle$, which gives

$$\langle x, j\rangle = \|x\| \text{ and } \|j\|_* = 1.$$

Thus,

$$\partial\|x\| \subseteq \{j \in X^* : \langle x, j\rangle = \|x\| \text{ and } \|j\|_* = 1\}.$$

Now suppose $j \in X^*$ such that $j \in \{f \in X^* : \langle x, f\rangle = \|x\| \text{ and } \|f\|_* = 1\}$. Then $\langle x, j\rangle = \|x\|$ and $\|j\|_* = 1$. Thus,

$$\langle y - x, j\rangle = \langle y, j\rangle - \|x\| \le \|y\| - \|x\| \text{ for all } y \in X,$$

i.e., $j \in \partial\|x\|$. It follows that

$$\{j \in X^* : \langle x, j\rangle = \|x\| \text{ and } \|j\|_* = 1\} \subseteq \partial\|x\|.$$

Therefore, $\partial\|x\| = \{j \in X^* : \langle x, j\rangle \text{ and } \|j\|_* = 1\}$. ∎

Using Corollary 2.5.19, we establish an inequality in a general Banach space that is a generalization of the inequality given in Proposition 2.4.6(b).

Theorem 2.5.22 *Let X be a Banach space and let $J_p : X \to 2^{X^*}$, $1 < p < \infty$ be the generalized duality mapping. Then for any $x, y \in X$, there exists $j_p(x+y) \in J_p(x+y)$ such that $\|x+y\|^p \le \|x\|^p + p\langle y, j_p(x+y)\rangle$.*

Proof. By Corollary 2.5.19, J_p is the subdifferential of the functional $\|\cdot\|^p/p$. By the subdifferentiability of $\|\cdot\|^p/p$, for $x, y \in X$, there exists $j_p(x+y) \in J_p(x+y)$ such that $\|x+y\|^p \le \|x\|^p + p\langle y, j_p(x+y)\rangle$. ∎

The following result is very useful in the approximation of solution of nonlinear operator equations.

Theorem 2.5.23 *Let X be a Banach space and $J_\mu : X \to 2^{X^*}$ a duality mapping with gauge function μ. If J_μ is single-valued, then*

$$\Phi(\|x+y\|) = \Phi(\|x\|) + \int_0^1 \langle y, J_\mu(x+ty)\rangle dt \textit{ for all } x, y \in X.$$

Proof. Because J_μ is single-valued, it follows from Theorem 2.5.17 that $\partial\Phi(\|x\|) = \{J_\mu(x)\}$. Hence Corollary 2.5.16 implies that J_μ is the Gâteaux gradient of $\Phi(\|x\|)$, i.e.,

$$\frac{d}{dt}\Phi(\|x+ty\|)\Big|_{t=0} = \langle y, J_\mu(x)\rangle.$$

Hence

$$\frac{d}{dt}\Phi(\|x+ty\|)\Big|_{t=r} = \frac{d}{ds}\Phi(\|x+ry+sy\|)\Big|_{s=0} = \langle y,\ J_\mu(x+ry)\rangle, \quad r\in\mathbb{R}.$$

Because the function $t \mapsto \langle y,\ J_\mu(x+ty)\rangle$ is continuous, hence

$$\int_0^1 \langle y, J_\mu(x+ry)\rangle dr = \int_0^1 \frac{d}{dt}\Phi(\|x+ty\|)\Big|_{t=r} dr = \Phi(\|x+y\|)-\Phi(\|x\|).$$ ∎

Corollary 2.5.24 *Let X be a Banach space. If X^* is strictly convex, then we have the following:*

(*a*) $\Phi(\|x+y\|) = \Phi(\|x\|) + \int_0^1 \langle y, J_\mu(x+ty)\rangle dt$ *for all* $x, y \in X$;

(*b*) $\|x+y\|^p = \|x\|^p + p\int_0^1 \langle y, J_p(x+ty)\rangle dt$ *for all* $x, y \in X$ *and* $p > 1$;

(*c*) $\|x+y\|^2 = \|x\|^2 + 2\int_0^1 \langle y, J(x+ty)\rangle dt$ *for all* $x, y \in X$.

Proposition 2.5.25 *Let X be a Banach space with strictly convex dual and C a nonempty convex subset of X. Let x_0 be an element in C and $J_\mu : X \to X^*$ a duality mapping with gauge function μ. Then*

$$\|x_0\| = \inf_{x\in C}\|x\| \text{ if and only if } \langle x_0 - x, J_\mu(x_0)\rangle \le 0 \text{ for all } x \in C.$$

Proof. Let x_0 be a point in C such that $\langle x_0 - x, J_\mu(x_0)\rangle \le 0$ for all $x \in C$. Then

$$\|x_0\|\|J_\mu(x_0)\|_* = \langle x_0, J_\mu(x_0)\rangle \le \|x\|\|J_\mu(x_0)\|_* \text{ for all } x\in C.$$

Therefore, $\|x_0\| = \inf_{x\in C}\|x\|$.

Conversely, suppose that $x_0 \in C$ such that $\|x_0\| = \inf_{x\in C}\|x\|$. Then

$$\|x_0\| \le \|x_0 + t(x-x_0)\| \text{ for all } x\in C \text{ and } t\in[0,1],$$

which implies that

$$\Phi(\|x_0\|) - \Phi(\|x_0 + t(x-x_0)\|) \le 0.$$

Because $J_\mu(z) = \partial\Phi(\|z\|)$, it follows that

$$\Phi(\|x_0 + t(x-x_0)\|) - \Phi(\|x_0\|) \le \langle x_0 + t(x-x_0) - x_0, J_\mu(x_0 + t(x-x_0))\rangle,$$

which implies that

$$t\langle x_0 - x, J_\mu(x_0 + t(x-x_0))\rangle \le \Phi(\|x_0\|) - \Phi(\|x_0 + t(x-x_0)\|) \le 0.$$

Thus,

$$\langle x_0 - x, J_\mu(x_0 + t(x-x_0))\rangle \le 0.$$

Letting $t \to 0$, we obtain $\langle x_0 - x, J_\mu(x_0)\rangle \le 0$. ∎

2.6 Smoothness

Let C be a nonempty closed convex subset of a normed space X such that the origin belongs to the interior of C. A linear functional $j \in X^*$ is said to be *tangent* to C at a point $x_0 \in \partial C$ if $j(x_0) = \sup\{j(x) : x \in C\}$. If $H = \{x \in X : j(x) = 0\}$ is the hyperplane, then the set $H + x_0$ is called a *tangent hyperplane* to C at x_0.

Definition 2.6.1 *A Banach space X is said to be smooth if for each $x \in S_X$, there exists a unique functional $j_x \in X^*$ such that $\langle x, j_x \rangle = \|x\|$ and $\|j_x\| = 1$.*

Geometrically, the smoothness condition means that at each point x of the unit sphere, there is exactly one supporting hyperplane $\{j_x = 1\}$. This means that the hyperplane $\{j_x = 1\}$ is tangent at x to the unit ball, and this unit ball is contained in the half space $\{j_x \leq 1\}$.

Observation

- $\ell_p,\ L_p$ $(1 < p < \infty)$ are smooth Banach spaces.
- $c_0,\ \ell_1,\ L_1,\ \ell_\infty, L_\infty$ are not smooth.

Differentiability of norms of Banach spaces – Let X be a normed space and $S_X = \{x \in X : \|x\| = 1\}$, the unit sphere of X. Then the norm of X is *Gâteaux differentiable* at point $x_0 \in S_X$ if for $y \in S_X$

$$\frac{d}{dt}(\|x_0 + ty\|)\bigg|_{t=0} = \lim_{t \to 0} \frac{\|x_0 + ty\| - \|x_0\|}{t}$$

exists (say, $\langle y, \triangledown\|x_0\|\rangle$). $\triangledown\|x_0\|$ is called the *gradient of the norm* $\varphi(x) = \|x\|$ at $x = x_0$. The norm of X is said to *Gâteaux differentiable* if it is Gâteaux differentiable at each point of S_X. The norm of X is said to be *uniformly Gâteaux differentiable* if for each $y \in S_X$, the limit is approached uniformly for $x \in S_X$.

Example 2.6.2 *Let H be a Hilbert space. Then the norm of H is Gâteaux differentiable with $\triangledown\|x\| = x/\|x\|$, $x \neq 0$. Indeed, for each $x \in X$ with $x \neq 0$, we have*

$$\begin{aligned}\lim_{t \to 0} \frac{\|x + ty\| - \|x\|}{t} &= \lim_{t \to 0} \frac{\|x + ty\|^2 - \|x\|^2}{t(\|x + ty\| + \|x\|)} \\ &= \lim_{t \to 0} \frac{2t\langle y, x\rangle + t^2\|y\|^2}{t(\|x + ty\| + \|x\|)} = \langle y, x/\|x\|\rangle.\end{aligned}$$

Therefore, the norm of H is Gâteaux differentiable with $\triangledown\|x\| = x/\|x\|$.

Remark 2.6.3 *In view of Example 2.6.2, we have the following:*
(i) at $x \neq 0, \varphi(x) = \|x\|$ is Gâteaux differentiable with $\triangledown\|x\| = x/\|x\|$,
(ii) at $x = 0, \varphi(x) = \|x\|$ is not differentiable, but it is subdifferentiable. Indeed,

$$\begin{aligned}\partial\varphi(0) = \partial\|0\| &= \{j \in H : \langle x, j\rangle \leq \|x\| \ for\ all\ x \in H\} \\ &= \{j \in H : \|j\|_* \leq 1\}.\end{aligned}$$

Theorem 2.6.4 *Let X be a Banach space. Then we have the following:*

(a) If X^* is strictly convex, then X is smooth.

(b) If X^* is smooth, then X is strictly convex.

Proof. (a) Suppose X is not smooth. There exist $x_0 \in S_X$ and $j_1, j_2 \in S_{X^*}$ with $j_1 \neq j_2$ such that $\langle x_0, j_1\rangle = \langle x_0, j_2\rangle = 1$. This means that x_0 determines a continuous linear functional on X^* that takes its maximum value on B_{X^*} at two distinct points j_1 and j_1. Hence X^* is not strictly convex.

(b) Suppose X is not strictly convex. There exist $j \in S_{X^*}$ and $x, y \in S_X$ with $x \neq y$ such that $\langle x, j\rangle = \langle y, j\rangle = 1$. Thus, two supporting hyperplanes pass through $j \in S_{X^*}$ such that

$$\langle x, f\rangle = \langle y, f\rangle = 1, f \in X^*.$$

Therefore, X^* is not smooth. ∎

It is well-known that for a reflexive Banach space X, the dual spaces X and X^* can be equivalently renormed as strictly convex spaces such that the duality is preserved. Using the above fact, we have

Theorem 2.6.5 *Let X be a reflexive Banach space. Then we have the following:*
(a) X is smooth if and only if X^* is strictly convex.
(b) X is strictly convex if and only if X^* is smooth.

The following theorem establishes a relation between smoothness and Gâteaux differentiability of the norm.

Theorem 2.6.6 *A Banach space X is smooth if and only if the norm is Gâteaux differentiable on $X\backslash\{0\}$.*

Proof. Because the proper convex continuous function φ is Gâteaux differentiable if and only if it has a unique subgradient, we have

norm is Gâteaux differentiable at x

$\Leftrightarrow \partial\|x\| = \{j \in X^* : \langle x, j\rangle = \|x\|, \|j\|_* = 1\}$ is singleton

$\Leftrightarrow$ there exists a unique $j \in X^*$ such that $\langle x, j\rangle = \|x\|$ and $\|j\|_* = 1$

$\Leftrightarrow$ smooth. ∎

Next, we establish a relation between smoothness of a Banach space and a property of the duality mapping with gauge function μ.

Theorem 2.6.7 *Let X be a Banach space. Then X is smooth if and only if each duality mapping J_μ with gauge function μ is single-valued; in this case*

$$\frac{d}{dt}\Phi(\|x + ty\|)\bigg|_{t=0} = \langle y, J_\mu(x)\rangle \text{ for all } x, y \in X. \tag{2.23}$$

Proof. The Banach space X is smooth if and only if there exists a unique $j \in X^*$ satisfying

$$\langle x\mu(||x||), j\rangle = \|x\|\mu(\|x\|) \text{ and} \|j\|_* = 1;$$

in this case $\mu(\|x\|)j = J_\mu(x) = \partial\Phi(\|x\|)$, and hence by Corollary 2.5.16, we obtain the formula (2.23). ∎

Corollary 2.6.8 *Let X be a Banach space and $J_\mu : X \to 2^{X^*}$ a duality mapping with gauge function μ. Then $j \in J_\mu(x)$, $x \in X$ if and only if $H = \{y \in X : \langle y, j\rangle = \|x\|\mu(\|x\|)\}$ is a supporting hyperplane for the closed ball $B_{\|x\|}[0]$ at x.*

Corollary 2.6.9 *Let X be a Banach space and $J : X \to 2^{X^*}$ a duality mapping. Then the following are equivalent:*

(a) X is smooth.

(b) J is single-valued.

(c) The norm of X is Gâteaux differentiable with $\bigtriangledown\|x\| = \|x\|^{-1}Jx$.

We now study the continuity property of duality mappings.

Theorem 2.6.10 *Let X be a Banach space and $J : X \to X^*$ a single-valued duality mapping. Then J is norm to weak* continuous.*

Proof. We show that $x_n \to x \Rightarrow Jx_n \to Jx$ in the weak* topology. Let $x_n \to x$ and set $f_n := Jx_n$. Then

$$\langle x_n, f_n\rangle = \|x_n\|\|f_n\|_*, \|x_n\| = \|f_n\|_*.$$

Because $\{x_n\}$ is bounded, $\{f_n\}$ is bounded in X^*. Then there exists a subsequence $\{f_{n_k}\}$ of $\{f_n\}$ such that $f_{n_k} \to f \in X^*$ in the weak* topology. Because the norm of X^* is lower semicontinuous in weak* topology, we have

$$\|f\|_* \le \liminf_{k\to\infty} \|f_{n_k}\|_* = \liminf_{k\to\infty} \|x_{n_k}\| = \|x\|.$$

Because $\langle x, f - f_{n_k}\rangle \to 0$ and $\langle x - x_{n_k}, f_{n_k}\rangle \to 0$, it follows from the fact

$$\begin{aligned} |\langle x, f\rangle - \|x_{n_k}\|^2| &= |\langle x, f\rangle - \langle x_{n_k}, f_{n_k}\rangle| \\ &\le |\langle x, f - f_{n_k}\rangle| + |\langle x - x_{n_k}, f_{n_k}\rangle| \to 0 \end{aligned}$$

that

$$\langle x, f\rangle = \|x\|^2.$$

As a result

$$\|x\|^2 = \langle x, f\rangle \le \|f\|_*\|x\|.$$

Thus, we have $\langle x, f\rangle = \|x\|^2, \|x\| = \|f\|_*$. Therefore, $f = Jx$. ∎

Theorem 2.6.11 *Let X be a Banach space with a uniformly Gâteaux differentiable norm. Then the duality mapping $J : X \to X^*$ is uniformly demicontinuous on bounded sets, i.e., J is uniformly continuous from X with its norm topology to X^* with the weak* topology.*

Proof. Suppose the result is not true. Then there exist sequences $\{x_n\}$ and $\{z_n\}$, a point y_0 and a positive ε such that

$$\|x_n\| = \|z_n\| = \|y_0\| = 1, z_n - x_n \to 0 \text{ and } \langle y_0, Jz_n - Jx_n\rangle \geq \varepsilon \text{ for all } n \in \mathbb{N}.$$

Set

$$a_n := t^{-1}(\|x_n + ty_0\| - \|x_n\| - t\langle y_0, Jx_n\rangle)$$

and

$$b_n := t^{-1}(\|z_n - ty_0\| - \|z_n\| + t\langle y_0, Jz_n\rangle).$$

If $t > 0$ is sufficiently small, then both a_n and b_n are less than $\varepsilon/2$. On the other hand, we have

$$\begin{aligned} a_n &\geq t^{-1}(\langle x_n + ty_0, Jz_n\rangle - \langle x_n + ty_0, Jx_n\rangle) \\ &= \langle y_0, Jz_n - Jx_n\rangle + t^{-1}\langle x_n, Jz_n - Jx_n\rangle \end{aligned}$$

and

$$\begin{aligned} b_n &\geq t^{-1}(\langle z_n - ty_0, Jx_n\rangle - \langle z_n - ty_0, Jz_n\rangle) \\ &= \langle y_0, Jz_n - Jx_n\rangle - t^{-1}\langle z_n, Jz_n - Jx_n\rangle. \end{aligned}$$

Thus,

$$\begin{aligned} a_n + b_n &\geq 2\langle y_0, Jz_n - Jx_n\rangle + t^{-1}\langle x_n - z_n, Jz_n - Jx_n\rangle \\ &\geq 2\varepsilon - 2t^{-1}\|x_n - z_n\|, \end{aligned}$$

a contradiction by choosing $t = 2\|x_n - z_n\|/\varepsilon$ for sufficiently large n. ∎

2.7 Modulus of smoothness

Recall that the modulus of convexity of a Banach space X is a function $\delta_X : [0, 2] \to [0, 1]$ defined by

$$\delta_X(t) = \inf\{1 - \|(x + y)/2\| : x, y \in X, \|x\| \leq 1, \|y\| \leq 1, \|x - y\| \geq t\}.$$

We now introduce the modulus of smoothness of a Banach space.

Definition 2.7.1 *Let X be a Banach space. Then a function $\rho_X : \mathbb{R}^+ \to \mathbb{R}^+$ is said to be the modulus of smoothness of X if*

$$\begin{aligned} \rho_X(t) &= \sup\left\{\frac{\|x + y\| + \|x - y\|}{2} - 1 : \|x\| = 1, \|y\| = t\right\} \\ &= \sup\left\{\frac{\|x + ty\| + \|x - ty\|}{2} - 1 : \|x\| = \|y\| = 1\right\}, \quad t \geq 0. \end{aligned}$$

It is easy to check that $\rho_X(0) = 0$ and $\rho_X(t) \geq 0$ for all $t \geq 0$.

The following result contains important properties of the modulus of smoothness.

Proposition 2.7.2 *Let ρ_X be the modulus of smoothness of a Banach space X. Then ρ_X is an increasing continuous convex function.*

Proof. Because for fixed $x, y \in X$ with $\|x\| = 1,\ \|y\| = 1$, the function

$$f(t) = \frac{\|x + ty\| + \|x - ty\|}{2} - 1,\ t \in \mathbb{R}$$

is convex and continuous on $\mathbb{R}$, it follows that the modulus of smoothness ρ_X is also continuous and a convex function.

Moreover, $f(-t) = f(t)$ for each $t \in \mathbb{R}$, f is nondecreasing on $\mathbb{R}^+$. Hence ρ_X is nondecreasing. ∎

The following theorem gives us an important relation between the modulus of convexity of X (respectively, X^*) and that of smoothness of X^* (respectively, X).

Theorem 2.7.3 *Let X be a Banach space. Then we have the following:*

(a) $\rho_{X^*}(t) = \sup\left\{\frac{t\varepsilon}{2} - \delta_X(\varepsilon) : 0 \le \varepsilon \le 2\right\}$ *for all* $t > 0$.

(b) $\rho_X(t) = \sup\left\{\frac{t\varepsilon}{2} - \delta_{X^*}(\varepsilon) : 0 \le \varepsilon \le 2\right\}$ *for all* $t > 0$.

Proof. (a) By the definition of modulus of smoothness of X^*, we have

$$\begin{aligned} 2\rho_{X^*}(t) &= \sup\{\|x^* + ty^*\|_* + \|x^* - ty^*\|_* - 2 : x^*, y^* \in S_{X^*}\} \\ &= \sup\{\langle x, x^*\rangle + t\langle x, y^*\rangle + \langle y, x^*\rangle - t\langle y, y^*\rangle - 2 : x, y \in S_X, x^*, y^* \in S_{X^*}\} \\ &= \sup\{\|x + y\| + t\|x - y\| - 2 : x, y \in S_X\} \\ &= \sup\{\|x + y\| + t\varepsilon - 2 : x, y \in S_X, \|x - y\| = \varepsilon, 0 \le \varepsilon \le 2\} \\ &= \sup\{t\varepsilon - 2\delta_X(\varepsilon) : 0 \le \varepsilon \le 2\}. \end{aligned}$$

Part (b) can be obtained in the same manner. ∎

As an immediate consequence of Theorem 2.7.3 (b), we have

Corollary 2.7.4 *Let X be a Banach space. Then $\rho_X(t)/t$ is increasing function and $\rho_X(t) \le t$ for all $t > 0$.*

Theorem 2.7.3 allows us to estimate ρ_X for Hilbert spaces. Indeed, we have

Proposition 2.7.5 *Let H be a Hilbert space. Then for $t > 0$*

$$\rho_H(t) = \sup\{t\varepsilon/2 - 1 + (1 - \varepsilon^2/4)^{1/2} : 0 < \varepsilon \le 2\} = (1 + t^2)^{1/2} - 1.$$

Observation

- If X is a Banach space and H is a Hilbert space, then $\rho_X(t) \geq \rho_H(t) = \sqrt{1+t^2}-1$ for all $t \geq 0$.

Let X be a Banach space. Then the *characteristic of convexity* or the *coefficient of convexity* of the Banach space X is the number

$$\epsilon_0(X) = \sup\{\varepsilon \in [0,2] : \delta_X(\varepsilon) = 0\}.$$

The Banach space X is said to be *uniformly convex* if $\epsilon_0(X) = 0$ and *uniformly nonsquare* if $\epsilon_0(X) < 2$. One may easily see that the modulus of convexity δ_X is strictly increasing on $[\epsilon_0, 2]$.

Example 2.7.6 *Let $X = \mathbb{R}^2$ with norm $\|\cdot\|_\infty$ defined by*

$$\|x\|_\infty = \|(x_1, x_2)\|_\infty = \max\{|x_1|, |x_2|\}.$$

Then X has a square-shaped unit ball for which $\delta_X(\varepsilon) = 0$ for $\varepsilon \in [0,2]$. Hence $\epsilon_0(X) = 2$.

The following theorem gives an important relation between the modulus of smoothness of a Banach space and the characteristic of convexity of its dual space.

Theorem 2.7.7 *Let X be a Banach space. Then the following statements are equivalent:*

(a) $\lim_{t\to 0} \frac{\rho_X(t)}{t} < \varepsilon/2$ *for all* $\varepsilon \leq 2$.

(b) $\epsilon_0(X^*) < \varepsilon$ *for all* $\varepsilon \leq 2$.

Proof. (a) $\Rightarrow$ (b). Let $\varepsilon \in [0,2]$. Suppose, for contradiction, that $\epsilon_0(X^*) \geq \varepsilon$. Then there exist $\{f_n\}$ and $\{g_n\}$ in S_{X^*} such that

$$\|f_n - g_n\|_* \geq \varepsilon \text{ and } \lim_{n\to\infty} \|f_n + g_n\|_* = 2. \tag{2.24}$$

From the definition of ρ_X, we get

$$\rho_X(t) \geq \left\|\frac{x+ty}{2}\right\| + \left\|\frac{x-ty}{2}\right\| - 1 \text{ for all } t > 0 \text{ and } x, y \in S_X.$$

Therefore,

$$\rho_X(t) \geq \left|\frac{f(x)+g(x)}{2}\right| + t\left|\frac{f(y)-g(y)}{2}\right| - 1 \text{ for all } f, g \in S_{X^*}.$$

Because x and y were arbitrary, we get

$$\rho_X(t) \geq \left\|\frac{f+g}{2}\right\|_* + t\left\|\frac{f-g}{2}\right\|_* - 1.$$

In particular, we have

$$\rho_X(t) \geq \left\|\frac{f_n + g_n}{2}\right\|_* + t\left\|\frac{f_n - g_n}{2}\right\|_* - 1 \text{ for all } n \in \mathbb{N}.$$

It follows from (2.24) that

$$\rho_X(t) \geq \frac{t\varepsilon}{2}.$$

(b) $\Rightarrow$ (a). Assume that $\epsilon_0(X^*) < \varepsilon$ and let $\varepsilon' \in (\epsilon_0(X^*), \varepsilon)$. Set $t' = \delta_{X^*}(\varepsilon')$ and consider $t \in [0,2]$. There are two possibilities :

(i) Assume that $t < \varepsilon'$. Then $t\lambda/2 < \lambda\varepsilon'/2$ and so $t\lambda/2 - \delta_{X^*}(t) < \lambda\varepsilon'/2$.

(ii) Assume that $\varepsilon' \leq t$. Then $\delta_{X^*}(t) \geq \delta_{X^*}(\varepsilon') = t'$, because the modulus of convexity is an increasing function. Therefore,

$$\frac{\lambda t}{2} \leq \lambda < t' < \delta_{X^*}(t) \text{ for any } \lambda < t'.$$

This implies that

$$\frac{t\lambda}{2} - \delta_{X^*}(t) < 0.$$

Therefore, in any case we have for $\lambda < t'$

$$\sup\left\{\frac{t\lambda}{2} - \delta_{X^*}(t) : t \in [0,2]\right\} \leq \frac{\lambda\varepsilon'}{2}.$$

Using Theorem 2.7.3, we get $\rho_X(\lambda) \leq \lambda\epsilon'/2$, which gives that $\lim\limits_{\lambda\to 0} \rho_X(\lambda)/\lambda \leq \varepsilon'/2$. Our choice of ε' implies that (b) is true. ∎

Let X be a Banach space. Then the *characteristic of smoothness* of X is the number

$$\rho_0(X) = \lim_{t\to 0} \frac{\rho_X(t)}{t}.$$

The following theorem allows us to estimate $\rho_0(X)$ for Banach spaces X.

Theorem 2.7.8 *Let X be a Banach space. Then*

$$\rho_0(X) = \rho'_X(0) = \lim_{t\to 0} \frac{\rho_X(t)}{t} = \frac{\epsilon_0(X^*)}{2}.$$

Proof. Assume first that $\epsilon_0(X^*) = 2$. Then $\delta_{X^*}(\varepsilon) = 0$ for every $\varepsilon \in [0,2]$. Therefore, using Theorem 2.7.3, we get $\rho_X(t) = t$ for every $t > 0$. Hence

$$\lim_{t\to 0} \frac{\rho_X(t)}{t} = 1 = \frac{\epsilon_0(X^*)}{2}.$$

Now if we assume that $\epsilon_0(X^*) < 2$, then from Theorem 2.7.7 we get the desired conclusion. ∎

Using Theorem 2.7.3 and 2.7.8, we have

Theorem 2.7.9 *Let X be a Banach space. Then we have the following:*

(a) $\rho_0(X) = \epsilon_0(X^*)/2$.

(b) $\rho_0(X^*) = \epsilon_0(X)/2$.

2.8 Uniform smoothness

Recall that the Banach space X is uniformly convex if $\delta_X(\varepsilon) > 0$ for all $\varepsilon \in (0, 2]$.

We now define uniform smoothness of a Banach space.

Definition 2.8.1 *A Banach space X is said to be uniformly smooth if*

$$\rho'_X(0) = \lim_{t\to 0} \frac{\rho_X(t)}{t} = 0.$$

Example 2.8.2 *The ℓ_p spaces $(1 < p \leq 2)$ are uniformly smooth. In fact,*

$$\lim_{t\to 0} \frac{\rho_{\ell_p}(t)}{t} = \lim_{t\to 0} \frac{(1+t^p)^{1/p} - 1}{t} = 0.$$

Uniform smoothness has a close relation with differentiability of norm.

Theorem 2.8.3 *Every uniformly smooth Banach space X is smooth.*

Proof. Suppose, for contradiction, that X is not smooth. Then there exist $x \in X\backslash\{0\}$, and $i, j \in X^*$ such that $i \neq j$, $\|i\| = \|j\| = 1$ and $\langle x, i\rangle = \langle x, j\rangle = \|x\|$. Let $y \in X$ such that $\|y\| = 1$ and $\langle y, i - j\rangle > 0$. For each $t > 0$, we have

$$\begin{aligned} 0 &< t\langle y, i - j\rangle \\ &= t\langle y, i\rangle - t\langle y, j\rangle \\ &= \frac{\langle x + ty, i\rangle + \langle x - ty, j\rangle}{2} - 1 \\ &\leq \frac{\|x + ty\| + \|x - ty\|}{2} - 1, \end{aligned}$$

and it follows that

$$0 < \langle y, i - j\rangle \leq \frac{\rho_X(t)}{t} \text{ for each } t > 0.$$

Hence X is not uniformly smooth. ∎

Next, we establish the duality between uniform convexity and uniform smoothness.

Theorem 2.8.4 *Let X be a Banach space. Then X is uniformly smooth if and only if X^* is uniformly convex.*

Proof. Recall that

$$\rho_X(t) = \sup\left\{\frac{t\varepsilon}{2} - \delta_{X^*}(\varepsilon) : 0 < \varepsilon \leq 2\right\} \text{ for all } t > 0. \tag{2.25}$$

Suppose, for contradiction, that X^* is not uniformly convex. Then there exists $\varepsilon_0 \in (0,2]$ with $\delta_{X^*}(\varepsilon_0) = 0$. From (2.25), we have

$$\frac{t\varepsilon_0}{2} - \delta_{X^*}(\varepsilon_0) \leq \rho_X(t)$$

which gives us that

$$0 < \frac{\varepsilon_0}{2} \leq \frac{\rho_X(t)}{t} \text{ for all } t > 0,$$

and this means that X is not uniformly smooth.

Conversely, assume that X is not uniformly smooth. Then $\rho_X'(0) = \lim_{t\to 0} \frac{\rho_X(t)}{t} \neq 0$. Hence for $\varepsilon > 0$ with $\lim_{t\to 0} \frac{\rho_X(t)}{t} = \varepsilon$, there exists a sequence $\{t_n\}$ in $(0,1)$ such that

$$t_n \to 0 \text{ and } \lim_{n\to\infty} \frac{\rho_X(t_n)}{t_n} = \varepsilon.$$

From (2.25), there exists a sequence $\{\varepsilon_n\}$ in $(0,2]$ such that

$$\frac{\varepsilon}{2} t_n \leq \frac{t_n \varepsilon_n}{2} - \delta_{X^*}(\varepsilon_n),$$

which implies that

$$0 < \delta_{X^*}(\varepsilon_n) \leq \frac{t_n}{2}(\varepsilon_n - \varepsilon).$$

It follows from the condition $t_n < 1$ that $\varepsilon < \varepsilon_n$. Because δ_{X^*} is a nondecreasing function, we have $\delta_{X^*}(\varepsilon) \leq \delta_{X^*}(\varepsilon_n) \to 0$, i.e., X^* is not uniformly convex. ∎

Theorem 2.8.5 *Let X be a Banach space. Then X is uniformly convex if and only if X^* is uniformly smooth.*

Proof. Notice

$$\rho_{X^*}(t) = \sup\left\{\frac{t\varepsilon}{2} - \delta_X(\varepsilon) : 0 < \varepsilon \leq 2\right\} \text{ for all } t > 0.$$

By interchanging the roles of X and X^*, we obtain the result by Theorem 2.8.4. ∎

Theorem 2.8.6 *Every uniformly smooth Banach space is reflexive.*

Proof. Let X be a uniformly smooth Banach space. Then X^* is uniformly convex and hence X^* is reflexive. It follows from Theorem 1.9.26 (which states that the reflexivity of X^* implies the reflexivity of X) that X is reflexive. ∎

Fréchet differentiability of norm and uniform smoothness

Uniform smoothness can be characterized by uniform Fréchet differentiability of the norm.

The norm of a Banach space X is said to *Fréchet differentiable* if for each $x \in S_X$, $\lim_{t\to 0} \frac{\|x+ty\| - \|x\|}{t}$ exists uniformly for $y \in S_X$.

In the other words, there exists a function $\varepsilon_x(s)$ with $\varepsilon_x(s) \to 0$ as $s \to 0$ such that

$$\Big| \|x+ty\| - \|x\| - t\langle y, Jx\rangle \Big| \leq |t|\varepsilon_x(|t|) \text{ for all } y \in S_X.$$

In this case, the norm is Gâteaux differentiable and

$$\lim_{t\to 0} \sup_{y\in S_X} \left| \frac{\frac{1}{2}\|x+ty\|^2 - \frac{1}{2}\|x\|^2}{t} - \langle y, Jx\rangle \right| = 0 \text{ for all } x \in X.$$

On the other hand,

$$\frac{1}{2}\|x\|^2 + \langle h, Jx\rangle \leq \frac{1}{2}\|x+h\|^2 \leq \frac{1}{2}\|x\|^2 + \langle h, Jx\rangle + b(\|h\|)$$

for all bounded $x, h \in X$, where b is a function defined on $\mathbb{R}^+$ such that $\lim_{t\to 0} \frac{b(t)}{t} = 0$.

We say that the norm of a Banach space X is *uniformly Fréchet differentiable* if

$$\lim_{t\to 0} \frac{\|x+ty\| - \|x\|}{t} \text{ exists uniformly for all } x, y \in S_X.$$

We now establish some results concerning Fréchet differentiability of the norm of Banach spaces.

Theorem 2.8.7 *Let X be a Banach space with a Fréchet differentiable norm. Then the duality mapping $J : X \to X^*$ is norm to norm continuous.*

Proof. It is sufficient to prove that $x_n \to x \in S_X \Rightarrow Jx_n \to Jx \in S_{X^*}$. Let $\{x_n\}$ be a sequence in S_X such that $x_n \to x$. Then $x \in S_X$. Because X has a Fréchet differentiable norm,

$$\lim_{t\to 0} \frac{\|x+ty\| - \|x\|}{t} = \langle y, Jx\rangle \text{ uniformly in } y \in S_X,$$

i.e., for any $\varepsilon > 0$, there exists $\delta > 0$ such that

$$\left| \frac{\|x+ty\| - \|x\|}{t} - \langle y, Jx\rangle \right| < \varepsilon \text{ for all } y \in S_X \text{ and all } t \text{ with } 0 < |t| \leq \delta.$$

Hence

$$\|x+ty\| - \|x\| < t(\langle y, Jx\rangle + \varepsilon) \text{ and } \|x-ty\| - \|x\| < -t(\langle y, Jx\rangle - \varepsilon),$$

so that

$$\|x+ty\| - 1 < t(\langle y, Jx\rangle + \varepsilon) \text{ and } \|x-ty\| - 1 < t(\varepsilon - \langle y, Jx\rangle).$$

Note

$$\begin{aligned} 0 \leq 1 - \langle x, Jx_n\rangle &= \langle x_n, Jx_n\rangle - \langle x, Jx_n\rangle \\ &\leq \langle x_n - x, Jx_n\rangle \\ &\leq \|x_n - x\|\|Jx_n\|_* = \|x_n - x\| \to 0, \end{aligned}$$

i.e., $\langle x, Jx_n\rangle \to 1$ as $n \to \infty$. Then there exists $n_0 \in \mathbb{N}$ such that

$$1 \leq \langle x, Jx_n\rangle + t\varepsilon \text{ for all } n \geq n_0.$$

Because

$$\begin{aligned} 1 - t\varepsilon \leq \langle x, Jx_n\rangle &= \langle x, Jx + Jx_n\rangle - 1 \\ &= \langle x+ty, Jx\rangle + \langle x-ty, Jx_n\rangle - t\langle y, Jx - Jx_n\rangle - 1 \\ &\leq \|x+ty\|\|Jx\|_* + \|x-ty\|\|Jx_n\|_* - t\langle y, Jx - Jx_n\rangle - 1 \\ &\leq t\langle y, Jx\rangle + t\varepsilon + 1 + 1 + t\varepsilon - t\langle y, Jx\rangle - t\langle y, Jx - Jx_n\rangle - 1 \\ &= 2t\varepsilon - t\langle y, Jx - Jx_n\rangle + 1, \end{aligned}$$

this implies that

$$\langle y, Jx - Jx_n\rangle \leq 3\varepsilon \text{ for all } y \in S_X.$$

Similarly, we can show that

$$\langle y, Jx_n - Jx\rangle \leq 3\varepsilon \text{ for all } y \in S_X.$$

Thus,

$$\left|\langle y, Jx_n - Jx\rangle\right| \leq 3\varepsilon \text{ for all } n \geq n_0 \text{ and } y \in S_X$$

which gives us

$$\|Jx_n - Jx\|_* < 3\varepsilon \text{ for all } n \geq n_0.$$

Therefore, $x_n \to x$ in X implies $Jx_n \to Jx$ in X^*. ∎

Theorem 2.8.8 *Let X be a Banach space. Then the following are equivalent:*

(a) X has a uniformly Fréchet differentiable norm.

(b) X^* is uniformly convex.

Proof. $(a) \Rightarrow (b)$. Suppose the norm of X is uniformly Fréchet differentiable. Then for any $\varepsilon > 0$, there exists $\delta > 0$ such that

$$\left| \frac{\|x+ty\| - \|x\|}{t} - \langle y, Jx\rangle \right| < \frac{\varepsilon}{8} \text{ for all } x, y \in S_X \text{ and all } t \text{ with } 0 < |t| \le \delta.$$

Then for fixed t with $0 < t < \delta$, we have

$$\|x+ty\| < \frac{t\varepsilon}{8} + t\langle y, Jx\rangle + 1$$

and

$$\|x-ty\| < \frac{t\varepsilon}{8} - t\langle y, Jx\rangle + 1.$$

As a result

$$\|x+ty\| + \|x-ty\| < \frac{t\varepsilon}{4} + 2 \text{ for all } x, y \in S_X.$$

Now, let $i, j \in S_{X^*}$ with $\|i-j\|_* \ge \varepsilon > 0$, then there exists $y_0 \in S_X$ such that

$$\langle y_0, i-j\rangle > \frac{\varepsilon}{2}.$$

Note

$$\begin{aligned}
\|i+j\|_* &= \sup_{x\in S_X} \langle x, i+j\rangle \\
&= \sup_{x\in S_X} \left(\langle x+ty_0, i\rangle + \langle x-ty_0, j\rangle - \langle ty_0, i-j\rangle\right) \\
&< \sup_{x\in S_X} \left(\|x+ty_0\| + \|x-ty_0\| - \frac{t\varepsilon}{2}\right) \\
&\le \frac{t\varepsilon}{4} + 2 - \frac{t\varepsilon}{2} \\
&\le 2 - \frac{t\varepsilon}{2}.
\end{aligned}$$

This implies $\|(i+j)/2\|_* < 1 - \delta(\varepsilon)$. Hence X^* is uniformly convex.

$(b) \Rightarrow (a)$. Let $x, y \in S_X$. Then for $t > 0$,

$$\begin{aligned}
\frac{\langle y, Jx\rangle}{\|x\|} &= \frac{\langle x+ty, Jx\rangle - \|x\|^2}{t\|x\|} \\
&\le \frac{\|x+ty\|\|x\| - \|x\|^2}{t\|x\|} \\
&= \frac{\|x+ty\| - \|x\|}{t} \\
&= \frac{\|x+ty\|^2 - \|x+ty\|\|x\|}{t\|x+ty\|} \\
&\le \frac{\langle x+ty, J(x+ty)\rangle - \langle x, J(x+ty)\rangle}{t\|x+ty\|} \\
&= \frac{\langle y, J(x+ty)\rangle}{\|x+ty\|}
\end{aligned}$$

and for $t < 0$,

$$\frac{\langle y, J(x+ty)\rangle}{\|x+ty\|} \leq \frac{\|x+ty\| - \|x\|}{t} \leq \frac{\langle y, Jx\rangle}{\|x\|}.$$

By Theorem 2.4.15, X has a uniformly Fréchet differentiable norm. ∎

Theorem 2.8.9 *Let X be a Banach space with uniformly Fréchet differentiable norm. Then the duality mapping $J : X \to X^*$ is uniformly continuous on each bounded set in X.*

Proof. Because X^* is uniformly convex, the result follows from Theorem 2.4.15 ∎

We now study the duality mapping from X^* to X. To do so, we define the *conjugate function* $f^* : X^* \to (-\infty, \infty]$ of any function $f : X \to (-\infty, \infty]$ by

$$f^*(j) = \sup\{\langle x, j\rangle - f(x) : x \in X\}, \quad j \in X^*. \tag{2.26}$$

The conjugate of f^*, i.e., the function on X defined by

$$f^{**}(x) = \sup\{\langle x, j\rangle - f^*(j) : j \in X^*\}, \quad x \in X$$

is called the *biconjugate* of f.

Observation

- f is lower semicontinuous proper convex on X if and only if $f^{**} = f$.

Example 2.8.10 *Let C be a nonempty subset of normed space X. Then the conjugate of the indicator function i_C of C is given by*

$$i_C^*(j) = \sup\{\langle x, j\rangle : x \in C\}, \quad j \in X^*.$$

The function i_C^* is called the *support function* of C.

We now give some basic properties of conjugate functions.

Proposition 2.8.11 *Let f^* be the conjugate function f. Then*

$$f(x) + f^*(j) \geq \langle x, j\rangle \text{ for all } x \in X, j \in X^*. \tag{2.27}$$

Proof. It easily follows from (2.26). ∎

The inequality (2.27) is known as the *Young inequality.* Observe also that if f is a proper function, then the relation (2.26) can be written as

$$f^*(j) = \sup\{\langle x, j\rangle - f(x) : x \in Dom(f)\}, \; j \in X^*.$$

Proposition 2.8.12 *Let f^* be the conjugate function of f. Then*

$$(cf)^*(j) = cf^*(c^{-1}j) \text{ for all } c > 0 \text{ and } j \in X^*.$$

Proof. For $j \in X^*$, we have

$$\begin{aligned} (cf)^*(j) &= \sup\{\langle x, j\rangle - (cf)(x) : x \in X\} \\ &= c\sup\{c^{-1}\langle x, j\rangle - f(x) : x \in X\} \\ &= c\sup\{\langle x, c^{-1}j\rangle - f(x) : x \in X\} \\ &= cf^*(c^{-1}j). \quad \blacksquare \end{aligned}$$

Proposition 2.8.13 *Let X be a normed space and $f : X \to (-\infty, \infty]$ a proper convex function. Then the following statements are equivalent:*

(a) $j \in \partial f(x)$ for $x \in X$.
(b) $f(x) + f^*(j) \leq \langle x, j\rangle$.
(c) $f(x) + f^*(j) = \langle x, j\rangle$.

Proof. $(b) \Leftrightarrow (c)$. The Young inequality (2.27) shows that (b) and (c) are equivalent.

$(c) \Leftrightarrow (a)$. Suppose condition (c) holds. Then from the Young inequality (2.27), we find that

$$f(y) - f(x) \geq \langle y - x, j\rangle \text{ for all } y \in X,$$

i.e., $j \in \partial f(x)$.

Using a similar argument, it follows that (c) $\Rightarrow$ (a). $\blacksquare$

Proposition 2.8.14 *Let X be a normed space and $f : X \to (-\infty, \infty]$ a lower semicontinuous proper convex function. Then $j \in \partial f(x) \Leftrightarrow x \in \partial f^*(j)$.*

Proof. Because f is a lower semicontinuous convex function, $f^{**} = f$. Observe that

$$\begin{aligned} j \in \partial f(x) &\Leftrightarrow f(x) + f^*(j) = \langle x, j\rangle \\ &\Leftrightarrow f^{**}(x) + f^*(j) = \langle x, j\rangle \\ &\Leftrightarrow x \in \partial f^*(j). \quad \blacksquare \end{aligned}$$

Proposition 2.8.15 *Let X be a Banach space. If $f(x) = \|x\|^p/p$, $p > 1$, then*

$$f^*(j) = \|j\|_*^q/q, \ 1/p + 1/q = 1.$$

Proof. Because $J_p(x) = \partial(\|x\|^p/p) = \{j \in X^* : \langle x, j\rangle = \|x\|\|j\|_*, \|j\|_* = \|x\|^{p-1}\}$, we have

$$f^*(j) = \sup_{x\in X}\{\langle x, j\rangle - f(x)\rangle = \sup_{x\in X}\{\|x\|^p - \|x\|^p/p\} = \sup_{x\in X}\{\|x\|^p/q\}.$$

Note $\|j\|_* = \|x\|^{p-1}$ so $\|j\|_*^q = \|x\|^{q(p-1)} = \|x\|^p$. Therefore, $f^*(j) = \|j\|_*^q/q$. $\blacksquare$

Theorem 2.8.16 *Let $p > 1$. Let X be a uniformly smooth Banach space and let $J_p : X \to X^*$ and $J_q^* : X^* \to X$ be the duality mappings with gauge functions $\mu_p(t) = t^{p-1}$ and $\mu_q(t) = t^{q-1}$, respectively. Then $J_p^{-1} = J_q^*$.*

Proof. The uniform smoothness of X implies that X is reflexive (see Theorem 2.8.6) and that X^* is uniformly convex and reflexive. Note also J_μ is surjective if and only if X is reflexive. Because J_p is single-valued, it follows that the inverse $J_p^{-1} : X^* = Dom\ (J_p^{-1}) \to X = X^{**}$ exists and is given by

$$J_p^{-1}(j) = \{x \in X : j = J_p(x)\} \text{ for all } j \in X^*.$$

Now, let $\Phi(t) = t^p/p, t > 0$. It is easy to see that $\Phi(\|\cdot\|) = \|\cdot\|^p/p$ is a continuous convex function and that its conjugate is given by $\Phi^*(\|j\|_*) = \|j\|_*^q/q$ for all $j \in X^*$. Note $J_p(x) = \partial\Phi(\|x\|)$ and $J_q^*(j) = \partial\Phi^*(\|j\|_*)$ for all $x \in X, j \in X^*$. Using Proposition 2.8.14, we have

$$j \in \partial\Phi(\|x\|) \text{ if and only if } x \in \partial\Phi^*(\|j\|_*).$$

Therefore, $J_p^{-1}(j) = J_q^*(j)$ for all $j \in X^*$. ∎

The following inequality is very useful in the existence and approximation of solutions of nonlinear operator equations.

Theorem 2.8.17 *Let X be a Banach space. Then the following are equivalent:*

(a) X is uniformly convex.

(b) For any p, $1 < p < \infty$ and $r > 0$, there exists a strictly increasing convex function $g_r : \mathbb{R}^+ \to \mathbb{R}^+$ such that $g_r(0) = 0$ and

$$\|tx + (1-t)y\|^p \leq t\|x\|^p + (1-t)\|y\|^p - t(1-t)g_r(\|x-y\|) \tag{2.28}$$

for all $x, y \in B_r[0]$ and $t \in [0,1]$.

Proof. (a) ⇒ (b). Let X be a uniformly convex Banach space. Assume that $1 < p < \infty$. It suffices to prove that (2.28) is true for $r = 1$. Now we define a function γ by

$$\begin{aligned}\gamma(\varepsilon) &= \inf\{2^{p-1}(\|x\|^p + \|y\|^p) - \|x+y\|^p : x, y \in B_X \text{ and } \|x-y\| \geq \varepsilon\} \\ &\quad \text{for all } \varepsilon \in (0,2].\end{aligned}$$

Because

$$\left(\frac{a+b}{2}\right)^p < \frac{a^p + b^p}{2} \text{ for all } a, b \geq 0 \text{ and } a \neq b, \tag{2.29}$$

we have

$$\gamma(\varepsilon) \geq 0 \text{ for all } 0 < \varepsilon \leq 2.$$

Suppose that $\gamma(\varepsilon) = 0$ for some $\varepsilon > 0$. Then there exist sequences $\{x_n\}$ and $\{y_n\}$ in B_X such that $\|x_n - y_n\| \geq \varepsilon$ for all $n \in \mathbb{N}$ and

$$\lim_{n\to\infty} 2^{p-1}(\|x_n\|^p + \|y_n\|^p) - \|x_n + y_n\|^p = 0.$$

We may assume a subsequence of $\{x_n\}$ denoted by $\{x_n\}$ such that

$$a = \lim_{n\to\infty} \|x_n\|, \ b = \lim_{n\to\infty} \|y_n\| \text{ and } c = \lim_{n\to\infty} \|x_n + y_n\|$$

exist. Thus,

$$\left(\frac{a+b}{2}\right)^p = \frac{a^p + b^p}{2},$$

i.e., equality of inequality (2.29) holds with $c = a + b$. For $a = b > 0$, $c = 2a = \lim_{n\to\infty} \|x_n + y_n\|$, it follows from Theorem 2.2.7 that $\lim_{n\to\infty} \|x_n - y_n\| = 0$, a contradiction. Therefore,

$$\gamma(\varepsilon) > 0 \text{ for all } \varepsilon, \ 0 < \varepsilon \le 2.$$

Now set

$$\mu(\varepsilon) := \inf\left\{\frac{\lambda\|x\|^p + (1-\lambda)\|y\|^p - \|\lambda x + (1-\lambda)y\|^p}{\lambda(1-\lambda)}\right\},$$

where the infimum is taken over all $x, y \in B_X$ with $\|x - y\| \ge \varepsilon$ and $\lambda \in (0,1)$. Note $\mu(\varepsilon) \ge \gamma(\varepsilon)/2^{p-1} > 0$ for all ε, $0 < \varepsilon \le 2$. Thus, it suffices to take as g_1 the double dual Young's function μ^{**}.

$(b) \Rightarrow (a)$. Suppose (2.28) is satisfied. For $x, y \in B_X$ and $\|x - y\| = \varepsilon$, we have

$$\begin{aligned} \left\|\frac{x+y}{2}\right\| &\le 1 - \frac{1}{4}g_1(\varepsilon) \\ &\le 1 - \delta_X(\varepsilon), \end{aligned}$$

i.e., $\delta_X(\varepsilon) \ge g_1(\varepsilon)/4$, which shows that X is a uniformly convex Banach space. $\blacksquare$

2.9 Banach limit

In this section, we generalize the concept of limit by introducing Banach limits and we discuss its properties.

Let $\ell : c \to \mathbb{K}$ be the "limit functional" defined by

$$\ell(x) = \lim_{i\to\infty} x_i \text{ for } x = \{x_i\} \in c.$$

Then ℓ is a linear functional on c. In order to extend limit ℓ on ℓ_∞, use the following notations and results.

Let S be a nonempty set and let $B(S)$ be the Banach space of all bounded real-valued functions on S with supremum norm.

Example 2.9.1 *Let* $S = \mathbb{N} = \{1, 2, 3, \cdots\}$. *Then* $B(S) = l_\infty$.

Let X be a subspace of $B(S)$ and let j be an element of X^*. Let e be a constant function on X defined by $e(s) = 1$ for all $s \in S$. We will denote $j(e)$ by $j(1)$. When X contains constants, a linear functional j on X is called a *mean* on X if $\|j\|_* = j(1) = 1$.

The following example shows that there is a subspace of ℓ_∞ for which the mean exists.

Example 2.9.2 *Let* $\ell_\infty = \{x = \{x_i\} : \sup_{i\in\mathbb{N}} |x_i| < \infty\}$ *and* X *a subset of* ℓ_∞ *such that*

$$X = \left\{ x = \{x_i\} \in \ell_\infty : \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} x_i \text{ exists} \right\}.$$

Then X *is a linear subspace of* ℓ_∞*. In fact, for* $x = \{x_i\}$ *and* $y = \{y_i\}$ *in* X*, we have*

$$\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} x_i \text{ exists and } \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} y_i \text{ exists.} \tag{2.30}$$

Hence for scalars α, β*, we have*

$$\alpha x + \beta y = (\alpha x_1 + \beta y_1, \cdots, \alpha x_i + \beta y_i, \cdots).$$

Using (2.30), we obtain that

$$\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} (\alpha x_i + \beta y_i) = \alpha(\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} x_i) + \beta(\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} y_i)$$

exists. It follows that X *is a linear subspace of* ℓ_∞*. We now define* $j : X \to \mathbb{R}$ *by*

$$j(x) = \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} x_i \text{ for all } x \in X.$$

Note $j(1) = 1$ *and*

$$\begin{aligned} |j(x)| &= \left| \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} x_i \right| \\ &\le \limsup_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} |x_i| \\ &\le \|x\|_\infty, \end{aligned}$$

and it follows that $\|j\|_* = 1$*. Therefore,* j *is linear and* $\|j\|_* = j(1) = 1$*, i.e.,* j *is a mean on* X*.*

We now give an equivalent condition for mean.

Theorem 2.9.3 *Let X be a subspace of $B(S)$ containing constants and $j \in X^*$. Then the following are equivalent:*

(a) j is a mean on X, i.e., $\|j\|_* = j(1) = 1$.

(b) The inequalities

$$\inf_{s \in S} x(s) \leq j(x) \leq \sup_{s \in S} x(s)$$

hold for each $x \in X$.

Proof. $(a) \Rightarrow (b)$. First, we show that $j(x) \geq 0$ for all $x \geq 0$. Suppose, for contraction, that $j(x) < 0$. Choose a positive number K with $x \leq K$. Then

$$j(K - x) = Kj(1) - j(x) = K - j(x) > K.$$

Because

$$j(K - x) \leq \|j\|_* \|K - x\| = \|K - x\| = \sup_{s \in S} |K - x(s)| \leq K,$$

it follows that

$$K < j(K - x) \leq K,$$

a contradiction. Therefore, $j(x) \geq 0$.

Observe that

$$\inf_{s \in S} x(s) \leq x \leq \sup_{s \in S} x(s) \text{ for each } x \in X.$$

Because $j(x) \geq 0$ for $x \geq 0$, we have

$$\inf_{s \in S} x(s) = j(\inf_{s \in S} x(s)) \leq j(x) \leq j(\sup_{s \in S} x(s)) = \sup_{s \in S} x(s).$$

$(b) \Rightarrow (a)$. For $x = 1$, we have $1 \leq j(1) \leq 1$ and hence $j(1) = 1$. Note for each $x \in X$,

$$j(x) \leq \sup_{s \in S} x(s) \leq \sup_{s \in S} |x(s)| = \|x\|$$

and

$$-j(x) = j(-x) \leq \| - x\| = \|x\|,$$

so $|j(x)| \leq \|x\|$ for each $x \in X$. Thus, $\|j\|_* = 1$. Therefore, $\|j\|_* = j(1) = 1$, i.e., j is a mean on X. ∎

Let $f \in \ell_\infty$. We denote $f_n(x_{n+m})$ for $f(x_{m+1}, x_{m+2}, x_{m+3}, \cdots, x_{m+n}, \cdots)$, $m = 0, 1, 2, \cdots$. A continuous linear functional j on l_∞ is called a *Banach limit* if

(L_1) $\|j\|_* = j(1) = 1$,

(L_2) $j_n(x_n) = j_n(x_{n+1})$ for each $x = (x_1, x_2, \cdots) \in l_\infty$.

It is denoted by LIM.

Theorem 2.9.4 (The existence of Banach limits) – *There exists a linear continuous functional j on l_∞ such that $\|j\|_* = j(1) = 1$ and $j_n(x_n) = j_n(x_{n+1})$ for each $x = \{x_n\}_{n\in\mathbb{N}} \in \ell_\infty$.*

Proof. Let $p : \ell_\infty \to \mathbb{R}$ be the functional defined by

$$p(x) = \limsup_{n\to\infty} \frac{x_1 + x_2 + \cdots + x_n}{n}.$$

Then

$$-p(-x) = \liminf_{n\to\infty} \frac{x_1 + x_2 + \cdots + x_n}{n}.$$

For $x \in c$, we have

$$\ell(x) = \lim_{n\to\infty} x_n = \lim_{n\to\infty} \frac{x_1 + x_2 + \cdots + x_n}{n} = p(x).$$

Moreover,

$$p(x + y) \leq p(x) + p(y) \text{ for all } x, y \in c$$

and

$$p(\alpha x) = \alpha p(x) \text{ for all } x \in c \text{ and } \alpha \geq 0.$$

Thus, p is a sublinear functional with $\ell(x) = p(x)$. By the Hahn-Banach theorem, there is an extension $L : \ell_\infty \to \mathbb{R}$ of ℓ (from c to ℓ_∞) such that

$$L(x) \leq \ell(x) \text{ for all } x \in \ell_\infty$$

and

$$-p(-x) \leq L(x) \leq p(x) \text{ for all } x \in \ell_\infty.$$

Thus, we have

$$p(1,\ 1,\ 1, \cdots) = 1$$

and

$$p((x_1, x_2, \cdots, x_n, \cdots) - (x_2, x_3, \cdots,\ x_{n+1}, \cdots)) = \limsup_{n\to\infty} \frac{x_1 - x_{n+1}}{n} = 0.$$

Hence

$$L((x_1, x_2, \cdots, x_n, \cdots) - (x_2, x_3, \cdots, x_{n+1}, \cdots)) = 0,$$

which implies that

$$L(x_1, x_2, \cdots, x_n, \cdots) = L(x_2, x_3, \cdots, x_{n+1} \cdots)$$
$$\text{for all } x = (x_1, x_2, \cdots x_n, \cdots) \in \ell_\infty.$$

Therefore, L is a Banach limit. ∎

Observation

- Every Banach limit is a positive functional on ℓ_∞, i.e., $LIM_n(x) \geq 0$ for all $x \in \ell_\infty$.
- $LIM(1, 1, \cdots 1, \cdots) = 1$.

We now give elementary properties of Banach limits.

Proposition 2.9.5 *Let LIM be a Banach limit. Then*

$$\liminf_{n\to\infty} x_n \le LIM(x) \le \limsup_{n\to\infty} x_n \text{ for each } x = (x_1, x_2, \cdots) \in l_\infty.$$

Moreover, if $x_n \to a$, then $LIM(x) = a$.

Proof. For each $m \in \mathbb{N}$, we have

$$LIM_n(x_n) = LIM_n(x_{n+1}) = \cdots = LIM_n(x_{n+(m-1)}) \ge \inf_{n\ge m} x_n$$

and hence $LIM_n(x_n) \ge \sup_{m\in\mathbb{N}} \inf_{n\ge m} x_n = \liminf_{n\to\infty} x_n$.

Similarly, since $LIM_n(x_n) \le \sup_{n\ge m} x_n$, we have $LIM_n(x_n) \le \limsup_{n\to\infty} x_n$. Therefore,

$$\liminf_{n\to\infty} x_n \le LIM(x) \le \limsup_{n\to\infty} x_n \text{ for each } x = (x_1, x_2, \cdots) \in l_\infty.$$

Letting $x_n \to a$, we have $\liminf_{n\to\infty} x_n = \limsup_{n\to\infty} x_n = a$ and hence $LIM(x) = a$. ∎

Proposition 2.9.6 *Let a be a real number and let $(x_1, x_2, \cdots) \in \ell_\infty$. Then the following are equivalent:*

(a) $LIM_n(x_n) \le a$ for all Banach limits LIM.

(b) For each $\varepsilon > 0$, there exists $m_0 \in \mathbb{N}$ such that

$$\frac{x_n + x_{n+1} + \cdots + x_{n+m-1}}{m} < a + \varepsilon \quad \textit{for all} \quad m \ge m_0 \textit{ and } n \in \mathbb{N}. \tag{2.31}$$

Proof. $(a) \Rightarrow (b)$. Suppose that for $\{x_n\} \in \ell_\infty$, we have $LIM_n(x_n) \le a$ for all Banach limits LIM. Define a sublinear functional $q : \ell_\infty \to \mathbb{R}$ by

$$q(y_1, y_2, \cdots) = \limsup_{m\to\infty} \left(\sup_{n\in\mathbb{N}} \frac{1}{m} \sum_{i=n}^{n+m-1} y_i \right), \ \{y_n\} \in \ell_\infty.$$

By the Hahn-Banach theorem, there exists a linear functional $j : \ell_\infty \to \mathbb{R}$ such that

$$j \le q \text{ and } j_n(x_n) = q_n(x_n).$$

It is easy to see that j is a Banach limit. From the assumption, we have

$$q_n(x_n) = \limsup_{m\to\infty} \left(\sup_{n\in\mathbb{N}} \frac{1}{m} \sum_{i=n}^{n+m-1} x_i \right) \le a.$$

Thus, for $\varepsilon > 0$, there exists $m_0 \in \mathbb{N}$ such that

$$\frac{x_n + x_{n+1} + \cdots + x_{n+m-1}}{m} < a + \varepsilon \text{ for all } m \ge m_0 \text{ and } n \in \mathbb{N}.$$

$(b) \Rightarrow (a)$. Suppose for each $\varepsilon > 0$, there exists $m_0 \in \mathbb{N}$ such that (2.31) holds.
Let LIM be a Banach limit. Then

$$LIM_n(x_n) = LIM_n\left(\frac{x_n + x_{n+1} + \cdots + x_{n+m_0-1}}{m_0}\right) \le a + \varepsilon.$$

Because ε is an arbitrary positive real number, we have $LIM_n(x_n) \le a$. ∎

Proposition 2.9.7 *Let a be a real number and let $(x_1, x_2, \cdots) \in \ell_\infty$ such that $LIM_n(x_n) \le a$ for all Banach limits LIM and $\limsup_{n\to\infty}(x_{n+1} - x_n) \le 0$. Then $\limsup_{n\to\infty} x_n \le a$.*

Proof. Let $\varepsilon > 0$. By Proposition 2.9.6, there exists $m \ge 2$ such that

$$\frac{x_n + x_{n+1} + \cdots + x_{n+m-1}}{m} < a + \frac{\varepsilon}{2} \text{ for all } n \in \mathbb{N}.$$

Choose $n_0 \in \mathbb{N}$ such that

$$x_{n+1} - x_n < \frac{\varepsilon}{m-1} \text{ for all } n \ge n_0.$$

Let $n \ge n_0 + m$. Observe that

$$\begin{aligned} x_n &= x_{n-i} + (x_{n-i+1} - x_{n-i}) + \cdots + (x_n - x_{n-1}) \\ &\le x_{n-i} + \frac{i\varepsilon}{m-1} \text{ for each } i = 0, 1, \cdots, m-1. \end{aligned}$$

Thus,

$$\limsup_{n\to\infty} x_n \le a + \varepsilon.$$

Because ε is arbitrary positive number, we get the conclusion. ∎

We note that if a linear functional j on l_∞ satisfying:

$$\liminf_{n\to\infty} x_n \le j(x) \le \limsup_{n\to\infty} x_n \text{ for each } x = (x_1, x_2, \cdots) \in l_\infty,$$

then j is a mean on ℓ_∞. Thus, every Banach limit on ℓ_∞ is a mean on ℓ_∞.

Let X be a Banach space, $\{x_n\}$ a bounded sequence in X, and LIM a Banach limit. Then a point $x_0 \in X$ is said to be a *mean point* of $\{x_n\}$ concerning a Banach limit LIM if

$$LIM_n\langle x_n, j\rangle = \langle x_0, j\rangle \text{ for all } j \in X^*.$$

We establish two preliminary results related to mean points.

Proposition 2.9.8 ***(Existence of mean points)*** *– Let X be a reflexive Banach space and $\{x_n\}$ a bounded sequence in X. Then, for a Banach limit LIM, there exists a point x_0 in X such that*

$$LIM_n\langle x_n, j\rangle = \langle x_0, j\rangle \textit{ for all } j \in X^*.$$

Proof. Note the function $LIM_n\langle x_n, j\rangle$ is linear in j. Further, as

$$|LIM_n\langle x_n, j\rangle| \le (\sup_{n\in\mathbb{N}} \|x_n\|) \cdot \|j\|_*,$$

the function $LIM_n\langle x_n, j\rangle$ is also bounded in j. So, we have $j_0^* \in X^{**}$ such that

$$LIM_n\langle x_n, j\rangle = \langle j_0^*, j\rangle \text{ for every } j \in X^*.$$

Because X is reflexive, there exists $x_0 \in X$ such that $LIM_n\langle x_n, j\rangle = \langle x_0, j\rangle$ for all $j \in X^*$. ∎

Proposition 2.9.9 *Let $\{x_n\}$ be a bounded sequence in a Banach space X and $x_0 \in X$ a mean point of $\{x_n\}$ concerning a Banach limit LIM. Then $x_0 \in \bigcap_{n=1}^{\infty} \overline{co}(\{x_k\}_{k\ge n})$.*

Proof. If not, there exists $n_0 \in \mathbb{N}$ such that $x_0 \notin \overline{co}\{x_n : n \ge n_0\}$. By the separation theorem, we obtain a point $j \in X^*$ such that

$$\langle x_0, j\rangle < \inf\{\langle z, j\rangle : z \in \overline{co}\{x_n : n \ge n_0\}\}.$$

Thus, we have

$$\begin{aligned} LIM_n\langle x_n, j\rangle &= \langle x_0, j\rangle < \inf\{(x_n, j) : n \ge n_0\} \\ &\le LIM_n\{\langle x_n, j\rangle : n \ge n_0\} = LIM_n\langle x_n, j\rangle, \end{aligned}$$

a contradiction. ∎

We now characterize the sequences in ℓ_∞ for which all Banach limits coincide. It is obvious that for any element $x \in c$,

$$LIM(x) = \ell(x) = \lim_{n\to\infty} x_n \text{ for all Banach limit } LIM.$$

However, there exist nonconvergent sequences for which all Banach limits coincide.

Example 2.9.10 *Let $x = (1, 0, 1, 0, \cdots) \in \ell_\infty$. Then*

$$(x_1, x_2, \cdots x_n, \cdots) + (x_2, x_3, \cdots, x_{n+1}, \cdots) = (1, 1, 1, \cdots),$$

and it follows that

$$LIM_n(x_n) + LIM_n(x_{n+1}) = LIM_n(1) = 1 \textit{ for all } LIM.$$

Using (L_2), we have

$$LIM_n(x_n) = \frac{1}{2} \textit{ for all Banach limit } LIM.$$

A bounded sequence $x = \{x_i\}$ is said to be *almost convergent* if all its Banach limits have the same value at x. Equivalently, $x = \{x_i\} \in \ell_\infty$ is almost convergent if

$$\lim_{i\to\infty} \frac{x_n + x_{n+1} \cdots + x_{n+i-1}}{i} \text{ exists uniformly in } n.$$

We have seen in Example 2.9.10 that the sequence $(1, 0, 1, 0, \cdots)$ is not convergent, but it is almost convergent.

In optimization theory, the structure of M defined in our next result is of much interest.

Theorem 2.9.11 *Let C be a nonempty closed convex subset of a reflexive Banach space X, $\{x_n\}$ a bounded sequence in C, LIM a Banach limit, and φ a real-valued function on C defined by $\varphi(z) = LIM_n\|x_n - z\|^2$, $z \in C$. Then the set M defined by*

$$M = \{u \in C : LIM_n\|x_n - u\|^2 = \inf_{z\in C} LIM_n\|x_n - z\|^2\} \tag{2.32}$$

is a nonempty closed convex bounded set. Moreover, if X is uniformly convex, then M has exactly one point.

Proof. First, we show that φ is continuous and convex. Let $\{y_m\}$ be a sequence in C such that $y_m \to y \in C$. Set $L := \sup\{\|x_n - y_m\| + \|x_n - y\| : m, n \in \mathbb{N}\}$. Observe that

$$\begin{aligned} \|x_n - y_m\|^2 - \|x_n - y\|^2 &\leq (\|x_n - y_m\| + \|x_n - y\|)(\|x_n - y_m\| - \|x_n - y\|) \\ &\leq L|\ \|x_n - y_m\| - \|x_n - y\|\ | \\ &\leq L\|y_m - y\| \text{ for all } n, m \in \mathbb{N}. \end{aligned}$$

Then

$$LIM_n\|x_n - y_m\|^2 \leq LIM_n\|x_n - y\|^2 + L\|y_m - y\|.$$

Similarly we have

$$LIM_n\|x_n - y\|^2 \leq LIM_n\|x_n - y_m\|^2 + L\|y_m - y\|.$$

Thus, we have

$$|\varphi(y_m) - \varphi(x)| \leq L\|y_m - x\|.$$

Hence φ is continuous on C. Now, let $x, y \in C$ and $\lambda \in [0, 1]$. It is easy to see that

$$\varphi((1 - \lambda)x + \lambda y) \leq (1 - \lambda)\varphi(x) + \lambda\varphi(y).$$

Hence φ is convex.

Using the fact $((a + b)/2)^2 \leq (a^2 + b^2)/2$ for all $a, b \geq 0$, we have

$$\|y_m\|^2 \leq 2\|y_m - x_n\|^2 + 2\|x_n\|^2,$$

and hence

$$\|y_m\|^2 \leq 2\varphi(y_m) + 2\sup_{n\in\mathbb{N}}\|x_n\|^2,$$

i.e., $\varphi(y_m) \to \infty$ as $\|y_m\| \to \infty$. Thus, φ is a continuous convex functional and $\varphi(z) \to \infty$ as $\|z\| \to \infty$. Because X is reflexive, φ attains its infimum over C by Theorem 2.5.8. Then M is a nonempty closed convex set. Moreover, M is bounded. Indeed, let $u \in M$. Because

$$\|u\|^2 \leq 2\|u - x_n\|^2 + 2\|x_n\|^2 \text{ for all } n \in \mathbb{N},$$

this implies that

$$\|u\|^2 \quad \leq \quad 2\varphi(u) + 2K = 2\inf_{z\in C}\varphi(z) + 2K$$

for some $K \geq 0$.

Now, suppose X is uniformly convex. Let $z_1, z_2 \in M$. Then $(z_1+z_2)/2 \in M$. Choose $r > 0$ large enough so that $\{x_n\} \cup M \subset B_r[0]$. Then $x_n - z_1, x_n - z_2 \in B_{2r}[0]$ for all $n \in \mathbb{N}$. By Theorem 2.8.17, we have

$$\left\|x_n - \frac{z_1+z_2}{2}\right\|^2 \leq \frac{1}{2}\|x_n - z_1\|^2 + \frac{1}{2}\|x_n - z_2\|^2 - \frac{1}{4}g_{2r}(\|z_1 - z_2\|).$$

If $z_1 \neq z_2$, we have

$$\begin{aligned}\inf_{z\in C}\varphi(z) \leq \varphi\left(\frac{z_1+z_2}{2}\right) \quad &\leq \quad \frac{1}{2}\varphi(z_1) + \frac{1}{2}\varphi(z_2) - \frac{1}{4}g_{2r}(\|z_1 - z_2\|)\\ &= \quad \inf_{z\in C}\varphi(z) - \frac{1}{4}g_{2r}(\|z_1 - z_2\|)\\ &< \quad \inf_{z\in C}\varphi(z),\end{aligned}$$

a contradiction. Therefore, M has exactly one element. ▮

Let LIM be a Banach limit and let $\{x_n\}$ be a bounded sequence in a Banach space X. We observe that if $\psi : X \to \mathbb{R}$ is bounded, Gâteaux differentiable uniformly on bounded sets, then a function $f : X \to \mathbb{R}$ defined by $f(z) = LIM_n\psi(x_n + z)$ is Gâteaux differentiable with Gâteaux derivative given by $\langle y, f'(z)\rangle = LIM_n\langle y, \psi'(x_n + z)\rangle$ for each $y \in X$.

Using the above facts, we give the following result, which will be used in convergence of sequences $\{x_n\}$ in Banach spaces with Gâteaux differentiable norm.

Theorem 2.9.12 *Let X be a Banach space with a uniformly Gâteaux differentiable norm and $\{x_n\}$ a bounded sequence in X. Let LIM be a Banach limit and $u \in X$. Then*

$$LIM_n\|x_n - u\|^2 = \inf_{z\in X} LIM_n\|x_n - z\|^2$$

if and only if

$$LIM_n\langle z, J(x_n - u)\rangle = 0 \text{ for all } z \in X.$$

Proof. Let $u \in X$ be such that $LIM_n\|x_n - u\|^2 = \inf_{z \in X} LIM_n\|x_n - z\|^2$. Then u minimizes the continuous convex function $\phi : X \to \mathbb{R}^+$ defined by $\phi(z) = LIM_n\|x_n - z\|^2$, so we have $\phi'(u) = 0$.

Note that the norm of X is Gâteaux differentiable, and Jx is the subdifferential of the convex function $\varphi(x) = \|x\|^2/2$ at x as the Gâteaux differential of φ. Hence

$$LIM_n\langle z, J(x_n - u)\rangle = \langle z, \phi'(u)\rangle = 0 \text{ for all } z \in X.$$

Conversely, suppose that $LIM_n\langle u - z, J(x_n - u)\rangle = 0$ for all $z \in X$. If $x \in X$,

$$\|x_n - x\|^2 - \|x_n - u\|^2 \geq 2\langle u - x, J(x_n - u)\rangle \text{ for all } n \in \mathbb{N}.$$

Because $LIM_n\langle u - x, J(x_n - u)\rangle = 0$ for all $x \in X$, we obtain

$$LIM_n\|x_n - u\|^2 = \inf_{x \in X} LIM_n\|x_n - x\|^2. \qquad \blacksquare$$

Corollary 2.9.13 *Let X be a Banach space with a uniformly Gâteaux differentiable norm and C a nonempty closed convex subset of X. Let $\{x_n\}$ be a bounded sequence in C. Let LIM be a Banach limit and $u \in C$. Then*

$$u \in M \textit{ if and only if } LIM_n\langle z, J(x_n - u)\rangle \leq 0 \textit{ for all } z \in C.$$

2.10 Metric projection and retraction mappings

Let C be a nonempty subset of a normed space X and let $x \in X$. An element $y_0 \in C$ is said to be a *best approximation* to x if

$$\|x - y_0\| = d(x, C),$$

where $d(x, C) = \inf_{y \in C} \|x - y\|$. The number $d(x, C)$ is called *the distance from x to C* or *the error in approximating x by C*.

The (possibly empty) set of all best approximations from x to C is denoted by

$$P_C(x) = \{y \in C : \|x - y\| = d(x, C)\}.$$

This defines a mapping P_C from X into 2^C and is called the *metric projection* onto C. The metric projection mapping is also known as the *nearest point projection mapping, proximity mapping*, and *best approximation operator.*

The set C is said to be a *proximinal* [2] (respectively, *Chebyshev*) *set* if each $x \in X$ has at least (respectively, exactly) one best approximation in C.

[2] The term "proximinal" is a combination of the words "proximity" and "minimal" and was coined by Killgrove.

Observation

- C is proximinal if $P_C(x) \neq \emptyset$ for all $x \in X$.
- C is Chebyshev if $P_C(x)$ is singleton for each $x \in X$.
- The set of best approximations is convex if C is convex.

Some fundamental results on proximinal sets are the following:

First, we observe that every proximinal set must be closed.

Proposition 2.10.1 *Let C be a proximinal subset of a Banach space X. Then C is closed.*

Proof. Suppose, for contradiction, that C is not closed. Then there exists a sequence $\{x_n\}$ in C such that $x_n \to x$ and $x \notin C$, but $x \in X$. It follows that

$$d(x, C) \leq \|x_n - x\| \to 0,$$

so that $d(x, C) = 0$. Because $x \notin C$, it means that

$$\|x - y\| > 0 \quad \text{for all} \quad y \in C.$$

This implies $P_C(x) = \emptyset$. This contradicts $P_C(x) \neq \emptyset$. ∎

Theorem 2.10.2 (The existence of best approximations) – *Let C be a nonempty weakly compact convex subset of a Banach space X and $x \in X$. Then x has a best approximation in C, i.e., $P_C(x) \neq \emptyset$.*

Proof. The function $f : C \to \mathbb{R}^+$ defined by

$$f(y) = \|x - y\|, \quad y \in C$$

is obviously lower semicontinuous. Because C is weakly compact, we can apply Theorem 2.5.5, and then there exists $y_0 \in C$ such that $\|x - y_0\| = \inf_{y \in C} \|x - y\|$. ∎

Corollary 2.10.3 *Let C be a nonempty closed convex subset of a reflexive Banach space X. Then each element $x \in X$ has a best approximation in C.*

Theorem 2.10.4 (The uniqueness of best approximations) – *Let C be a nonempty convex subset of a strictly convex Banach space X. Then for each element $x \in X$, C has at most one best approximation.*

Proof. Suppose, for contradiction, that $y_1, \ y_2 \in C$ are best approximations to $x \in X$. Because the set of best approximations is convex, it follows that $(y_1 + y_2)/2$ is also a best approximation to x. Set $r := d(x, C)$. Then

$$0 \leq r = \|x - y_1\| = \|x - y_2\| = \|x - (y_1 + y_2)/2\|,$$

and it follows that

$$\|(x - y_1) + (x - y_2)\| = 2r = \|x - y_1\| + \|x - y_2\|.$$

By the strict convexity of X we have

$$x - y_1 = t(x - y_2), \quad t \geq 0.$$

Taking the norm in this relation, we obtain $r = tr$, i.e., $t = 1$, which gives us $y_1 = y_2$. ∎

The following example shows that the strict convexity cannot be dropped in Theorem 2.10.4.

Example 2.10.5 *Let $X = \mathbb{R}^2$ with norm $\|\cdot\|_1$. It is easy to check that X is not strictly convex. Now, let*

$$C = \{(x, y) \in \mathbb{R}^2 : \|(x, y)\|_1 \leq 1\} = \{(x, y) \in \mathbb{R}^2 : |x| + |y| \leq 1\}.$$

Then C is a closed convex set. The distance from $z = (-1, -1)$ to the set C is one, and this distance is realized by more than one point of C.

In Theorem 2.10.4, uniqueness of best approximations need not be true for nonconvex sets.

Example 2.10.6 *Let $X = \mathbb{R}^2$ with the norm $\|\cdot\|_2$ and $C = S_X = \{(x, y)\} \in \mathbb{R}^2 : x^2 + y^2 = 1\}$. Then X is strictly convex and C is a nonconvex set. However, all points of C are best approximations to $(0, 0) \in X$.*

Theorem 2.10.7 *If in a Banach space X, every element possesses at most a best approximation with respect to every convex set, then X is strictly convex.*

Proof. Suppose, for contradiction, that X is not strictly convex. Then there exist $x, y \in X,\ x \neq y$ with

$$\|x\| = \|y\| = \|(x + y)/2\| = 1.$$

Furthermore,

$$\|tx + (1 - t)y\| = 1 \text{ for all } t \in [0, 1].$$

Set $C := co(\{x, y\})$. Then $\|0 - z\| = d(0, C)$ for all $z \in C$. This means that every element of C is the best approximation to zero and this clearly contradicts the uniqueness. ∎

From Corollary 2.10.3 and Theorem 2.10.4 (see also Proposition 2.1.10), we obtain some important results:

Theorem 2.10.8 *Let C be a nonempty weakly compact convex subset of a strictly convex Banach space X. Then for each $x \in X$, C has the unique best approximation, i.e., $P_C(\cdot)$ is a single-valued metric projection mapping from X onto C.*

Corollary 2.10.9 *Let C be a nonempty closed convex subset of a strictly convex reflexive (e.g., uniformly convex) Banach space X and let $x \in X$. Then there exists a unique element $x_0 \in C$ such that $\|x - x_0\| = d(x, C)$.*

Observation

- Every closed convex subset C of a reflexive Banach space is proximinal.
- Every closed convex subset C of a reflexive strictly convex Banach is a Chebyshev set.
- For every Chebyshev set C, we have
  (i) $P_C(x)$ is singleton set, i.e., P_C is a function from X onto C.
  (ii) $\|x - P_C(x)\| = d(x, C)$ for all $x \in X$.

We now study useful properties of metric projection mappings.

Theorem 2.10.10 *Let C be a subset of a normed space X and $\overline{x} \in X$. Then $P_C(\overline{x}) \subseteq \partial C$.*

Proof. Let $y \in P_C(\overline{x})$. Suppose $y \in int(C)$. Then there exists an $\varepsilon > 0$ such that $B_\varepsilon(y) \subset C$. For each $n \in \mathbb{N}$, let $z_n = (1/n)\overline{x} + (1 - 1/n)y$. Then

$$\|z_n - y\| = (1/n)\|\overline{x} - y\|.$$

For sufficiently large $N \in \mathbb{N}$, $\|z_N - y\| < \varepsilon$. Thus, $z_N \in B_\varepsilon(y) \subset C$. On the other hand,

$$\|\overline{x} - z_N\| = (1 - 1/N)\|\overline{x} - y\| < \|\overline{x} - y\| = d(\overline{x}, C),$$

which contradicts the fact that $y \in P_C(\overline{x})$. Therefore, $y \in \partial C$. ∎

Corollary 2.10.11 *Let C be a nonempty closed convex subset of a strictly convex reflexive Banach space X and let $x \in X$. Then we have the following:*

(a) If $x \in C$, then $P_C(x) = x$.

(b) If $x \notin C$, then $P_C(x) \in \partial C$.

Theorem 2.10.12 *Let C be a nonempty closed convex subset of a reflexive strictly convex Banach space X. If X has the Kadec-Klee property, then the projection mapping P_C of X onto C is continuous.*

Proof. Suppose, for contradiction, that P_C is not continuous. Then for the sequence $\{x_n\}$ in X with $\lim_{n\to\infty} x_n = x \in X$, there exists $\varepsilon > 0$ such that

$$\|P_C(x_n) - P_C(x)\| \geq \varepsilon \quad \text{for all} \quad n \in \mathbb{N}.$$

Because

$$|d(x_n, C) - d(x, C)| \leq \|x_n - x\|,$$

it follows that

$$\big|\|x_n - P_C(x_n)\| - \|x - P_C(x)\|\big| \leq \|x_n - x\|.$$

This implies that

$$\lim_{n\to\infty} \|x_n - P_C(x_n)\| = \|x - P_C(x)\|. \tag{2.33}$$

Because $\{P_C(x_n)\}$ is bounded in C by (2.33), there exists a subsequence $\{P_C(x_{n_i})\}$ of $\{P_C(x_n)\}$ such that $w - \lim_{i\to\infty} P_C(x_{n_i}) = z \in C$. Note

$$w - \lim_{i\to\infty} (x_{n_i} - P_C(x_{n_i})) = x - z. \tag{2.34}$$

By w-lsc of the functional $\|\cdot\|$, we have

$$\|x - z\| \leq \liminf_{i\to\infty} \|x_{n_i} - P_C(x_{n_i})\| = \|x - P_C(x)\|.$$

This implies $z = P_C(x)$ by definition of the function P_C. From (2.33) and (2.34)

$$w - \lim_{i\to\infty} (x_{n_i} - P_C(x_{n_i})) = x - P_C(x) \text{ and } \lim_{i\to\infty} \|x_{n_i} - P_C(x_{n_i})\| = \|x - P_C(x)\|.$$

Because X has the Kadec-Klee property, we obtain

$$\lim_{i\to\infty} (x_{n_i} - P_C(x_{n_i})) = x - P_C(x),$$

which implies that $\lim_{i\to\infty} P_C(x_{n_i}) = P_C(x)$, which is a contradiction to the assumption that $\|P_C(x_n) - P_C(x)\| \geq \varepsilon$. ∎

Then following Proposition 2.5.25, we have

Theorem 2.10.13 *Let C be a nonempty convex subset of a smooth Banach space X and let $x \in X$ and $y \in C$. Then the following are equivalent:*

(a) y is a best approximation to x: $\|x - y\| = d(x, C)$.

(b) y is a solution of the variational inequality:

$$\langle y - z, J_\mu(x - y)\rangle \geq 0 \text{ for all } z \in C,$$

where J_μ is a duality mapping with gauge function μ.

As an immediate consequence of Theorem 2.10.13, we have

Corollary 2.10.14 *Let C be a nonempty convex subset of a Hilbert space H and P_C be the metric projection mapping from H onto C. Let x be an element in H. Then the following are equivalent:*

(a) $\|x - P_C(x)\| = d(x, C)$.

(b) $\langle x - P_C(x), P_C(x) - z\rangle \geq 0$ for all $z \in C$.

Proposition 2.10.15 *Let C be a nonempty closed convex subset of a Hilbert space X and P_C the metric projection from X onto C. Then the following hold:*

(a) P_C is "idempotent": $P_C(P_C(x)) = P_C(x)$ for all $x \in X$.

(b) P_C is "firmly nonexpansive":

$$\langle x-y, P_C(x)-P_C(y)\rangle \geq \|P_C(x)-P_C(y)\|^2 \quad \textit{for all } x,y \in X.$$

(c) P_C is "nonexpansive": $\|P_C(x)-P_C(y)\| \leq \|x-y\|$ for all $x,y \in X$.

(d) P_C is "monotone": $\langle P_C(x)-P_C(y), x-y\rangle \geq 0$ for all $x,y \in X$.

(e) P_C is "demiclosed": $x_n \rightharpoonup x_0$ and $P_C(x_n) \to y_0 \Rightarrow P_C(x_0) = y_0$.

Proof. (a) Observe that $P_C(x) \in C$ for all $x \in X$ and $P_C(z) = z$ for all $z \in C$. Then $P_C(P_C(x)) = P_C(x)$ for all $x \in X$, i.e., $P_C^2 = P_C$.

(b) Set $j := P_C(x) - P_C(y)$ for $x, y \in X$. We have

$$\langle x-y, j\rangle = \langle x-P_C(x), j\rangle + \langle j, j\rangle + \langle P_C(y)-y, j\rangle.$$

Because from Corollary 2.10.14, we get

$$\langle x-P_C(x), j\rangle \geq 0 \text{ and } \langle y-P_C(y), j\rangle \geq 0,$$

it follows that

$$\langle x-y, j\rangle \geq \|j\|^2.$$

(c) This is an immediate consequence of (b).

(d) It follows from (b).

(e) From Corollary 2.10.14, we have

$$\langle x_n - P_C(x_n), P_C(x_n) - z\rangle \geq 0 \quad \text{for all} \quad z \in C.$$

Because $x_n \rightharpoonup x_0$ and $P_C(x_n) \to y_0$, we have

$$\langle x_0 - y_0, y_0 - z\rangle \geq 0 \text{ for all } z \in C.$$

Using Theorem 2.10.13, we obtain $\|x_0 - y_0\| = d(x_0, C)$. Therefore, $P_C(x_0) = y_0$. ∎

Remark 2.10.16 *Proposition 2.10.15(c) shows that in a Hilbert space, a metric projection operator is not only continuous, but also it is Lipschitz continuous and hence it is uniformly continuous.*

The following result is of fundamental importance. It shows that every point on line segment joining $x \in X$ to its best approximation $P_C(x) \in C$ has $P_C(x)$ as its best approximation.

Proposition 2.10.17 *Let C be a Chebyshev set in a Hilbert space H and $x \in H$. Then $P_C(x) = P_C(y)$ for all $y \in co(\{x, P_C(x)\})$.*

Proof. Suppose, for contradiction, that there exist $y \in co(\{x, P_C(x)\})$ and $z \in C$ such that

$$\|y - z\| < \|y - P_C(x)\|.$$

Set $y := \lambda x + (1-\lambda)P_C(x)$ for some $\lambda \in (0,1)$. Then

$$\begin{aligned} \|x - z\| &\le \|x - y\| + \|y - z\| \\ &< \|x - y\| + \|y - P_C(x)\| \\ &= (1-\lambda)\|x - P_C(x)\| + \lambda\|x - P_C(x)\| = d(x, C), \end{aligned}$$

a contradiction. ∎

If C is a Chebyshev set in a Hilbert space H, then

$$P_C[\lambda x + (1-\lambda)P_C(x)] = P_C(x),\ x \in H,\ 0 \le \lambda \le 1.$$

Motivated by this fact, we introduce the following:

A Chebyshev subset C of a normed space X is said to be *sun* if

$$P_C[\lambda x + (1-\lambda)P_C(x)] = P_C(x) \text{ for all } x \in X \text{ and } \lambda \ge 0.$$

In other words, C is a sun if and only if each point on the ray from $P_C(x)$ through x also has $P_C(x)$ as its best approximation in C.

Let C be a nonempty subset of a topological space X and D a nonempty subset of C. Then a continuous mapping $P : C \to D$ is said to be a *retraction* if $Px = x$ for all $x \in D$, i.e., $P^2 = P$. In such case, D is said to be a *retract* of C.

Example 2.10.18 *Every closed convex subset C of $\mathbb{R}^n$ is a retract of $\mathbb{R}^n$.*

We have seen in Theorem 2.10.8 that for every weakly compact convex subset C of a strictly convex Banach space, there exists a metric projection mapping $P_C : X \to C$ that may not be continuous. However, every single-valued metric projection mapping is a retraction if it is continuous.

Theorem 2.10.19 *Every closed convex subset C of a uniformly convex Banach space X is a retract of X.*

Proof. By Theorem 2.10.8, there exists a metric projection mapping $P_C : X \to C$ such that $P_C(x) = x$ for all $x \in C$. By Theorem 2.10.12, P_C is continuous. Therefore, P_C is retraction. ∎

We now show that every retraction P with condition (2.35) is sunny nonexpansive (and hence continuous).

Proposition 2.10.20 *Let C be a nonempty convex subset of a smooth Banach space X and D a nonempty subset of C. If P is a retraction of C onto D such that*

$$\langle x - Px, J(y - Px)\rangle \le 0 \text{ for all } x \in C \text{ and } y \in D, \tag{2.35}$$

then P is sunny nonexpansive.

Proof. <u>*P is sunny:*</u> For $x \in C$, set $x_t := Px + t(x - Px)$ for all $t > 0$. Because C is convex, it follows that $x_t \in C$ for all $t \in (0, 1]$. Hence

$$\langle x - Px, J(Px - Px_t)\rangle \geq 0 \text{ and } \langle x_t - Px_t, J(Px_t - Px)\rangle \geq 0. \tag{2.36}$$

Because $x_t - Px = t(x - Px)$ and $\langle t(x - Px), J(Px - Px_t)\rangle \geq 0$, we have

$$\langle x_t - Px, J(Px - Px_t)\rangle \geq 0. \tag{2.37}$$

Combining (2.36) and (2.37), we get

$$\begin{aligned} \|Px - Px_t\|^2 &= \langle Px - x_t + x_t - Px_t, J(Px - Px_t)\rangle \\ &\leq -\langle x_t - Px, J(Px - Px_t)\rangle + \langle x_t - Px_t, J(Px - Px_t)\rangle \\ &\leq 0. \end{aligned}$$

Thus, $Px = Px_t$. Therefore, P is sunny.

<u>*P is nonexpansive :*</u> For $x, z \in C$, we have from (2.35) that

$$\langle x - Px, J(Px - Pz)\rangle \geq 0 \text{ and } \langle z - Pz, J(Pz - Px)\rangle \geq 0.$$

Hence

$$\langle x - z - (Px - Pz), J(Px - Pz)\rangle \geq 0.$$

This implies that

$$\langle x - z, J(Px - Pz)\rangle \geq \|Px - Pz\|^2$$

and hence P is nonexpansive. ∎

We now give equivalent formulations of sunny nonexpansive retraction mappings.

Proposition 2.10.21 *Let C be a nonempty convex subset of a smooth Banach space X, D a nonempty subset of C, and $P : C \to D$ a retraction. Then the following are equivalent:*

(a) P is the sunny nonexpansive.

(b) $\langle x - Px, J(y - Px)\rangle \leq 0$ for all $x \in C$ and $y \in D$.

(c) $\langle x - y, J(Px - Py)\rangle \geq \|Px - Py\|^2$ for all $x, y \in C$.

Proof. (a)⇒ (b). Let P be the sunny nonexpansive retraction and $x \in C$. Then $Px \in D$ and there exists a point $z \in D$ such that $Px = z$. Set $M := \{z + t(x - z) : t \geq 0\}$. Then M is nonempty convex set. Hence for $v \in M$

$$\begin{aligned} \|y - z\| &= \|Py - Pv\| \quad (\text{as } P \text{ is sunny, i.e., } Pv = z) \\ &\leq \|y - v\| = \|y - z + t(z - x)\| \text{ for all } y \in D. \end{aligned}$$

Hence from Proposition 2.4.7, we have

$$\langle x - Px, J(y - Px)\rangle \leq 0.$$

(b) $\Rightarrow$ (a). It follows from Proposition 2.10.20.

(b) $\Rightarrow$ (c). Let $x, y \in C$. Then $Px, Py \in D$ and hence from (b), we have

$$\langle x - Px, J(Py - Px)\rangle \le 0 \text{ and } \langle y - Py, J(Px - Py)\rangle \le 0.$$

Combining the above inequalities, we get

$$\langle Px - Py - (x - y), J(Px - Py)\rangle \le 0.$$

Hence

$$\begin{aligned} \|Px - Py\|^2 &= \langle Px - Py, J(Px - Py)\rangle \\ &= \langle Px - Py - (x - y), J(Px - Py)\rangle + \langle x - y, J(Px - Py)\rangle \\ &\le \langle x - y, J(Px - Py)\rangle. \end{aligned}$$

(c) $\Rightarrow$ (b). Suppose (c) holds. Let $x \in C$ and $y \in D$. Replacing y by $y = Py$ in (c), we have

$$\langle x - Py, J(Px - P^2y) \ge \|Px - P^2y\|^2,$$

which implies that

$$\langle x - y, J(Px - y)\rangle \ge \|Px - y\|^2.$$

Therefore,

$$\begin{aligned} \langle x - Px, J(Px - y)\rangle &= \langle x - y, J(Px - y)\rangle + \langle y - Px, J(Px - y) \\ &\ge \|Px - y\|^2 - \|Px - y\|^2 = 0. \qquad \blacksquare \end{aligned}$$

Finally, we give uniqueness of sunny nonexpansive retraction mappings.

Proposition 2.10.22 *Let C be a nonempty convex subset of a smooth Banach space X and D a nonempty subset of C. If P is a sunny nonexpansive retraction from C onto D, then P is unique.*

Proof. Let Q be another sunny nonexpansive retraction from C onto D. Then, we have, for each $x \in C$

$$\langle x - Px, J(y - Px)\rangle \le 0 \text{ and } \langle x - Qx, J(y - Qx)\rangle \le 0 \quad \text{for all } y \in D.$$

In particular, because Px and Qx are in D, we have

$$\langle x - Px, J(Qx - Px)\rangle \le 0 \text{ and } \langle x - Qx, J(Px - Qx)\rangle \le 0,$$

which imply that $\|Px - Qx\|^2 \le 0$. Therefore, $Px = Qx$ for all $x \in C$. $\blacksquare$

Bibliographic Notes and Remarks

The results of Sections 2.1∼2.3 are well-known. These results are adapted from Cioranescu [41], Goebel and Kirk [59], Goebel and Reich [60], Istratescu [73], Martin [106], and Prus [121].

The results in Sections 2.4∼2.8 are based on Barbu [11], Goebel and Kirk [59], Prus [121], and Showalter [145]. Theorem 2.4.16 follows from Kaczor [82]. Theorem 2.6.11 is proved in Reich [127]. The interested reader should also consult the details on duality mappings in Cioranescu [41] and its review by Reich [131].

Some results in Section 2.9 are based on some results of Ha and Jung [65], Jung and Park [80], and Shioji and Takahashi [144]. Theorem 2.9.11 is a cornerstone of the "Optimization Method" expounded in Reich [126, 128]. The uniqueness of the minimizer is shown in the paper by Reich [130].

The results in Section 2.10 are adapted from several sources (Goebel and Kirk [59], Goebel and Reich [60], and Takahashi [155]). Theorem 2.10.12 is an important improvement of Proposition 3.2 of Martin [106]. Proposition 2.10.21 was first proved in Reich [123], where the term "sunny nonexpansive retraction" was coined. More updated information on (sunny) nonexpansive retractions can be found in Reich and Kopecká [132].

Exercises

2.1 Let X be a strictly convex Banach space and let $x, y \in X$ with $x \neq y$. If $\|x-z\| = \|x-w\|, \|z-y\| = \|w-y\|$ and $\|x-y\| = \|x-z\| + \|z-y\|$, show that $z = w$.

2.2 Let X be a uniformly convex Banach space and let δ_X be the modulus of convexity of X. Let $0 < \varepsilon < r \leq 2R$. Show that $\delta_X(\varepsilon/R) > 0$ and

$$\|\lambda x + (1-\lambda)y\| \leq r\left\{1 - 2\lambda(1-\lambda)\delta_X\left(\frac{\varepsilon}{R}\right)\right\}$$

for all $x, y \in X$ with $\|x\| \leq r, \|y\| \leq r$ and $\|x-y\| \geq \varepsilon$ and $\lambda \in [0,1]$.

2.3 Let X be a Banach space. Show that X is uniformly convex if and only if $\gamma(t) > 0$ for all $t \in (0,2]$, where

$$\gamma(t) = \inf\{\langle x-y, x^*-y^*\rangle : x, y \in S_X, \|x-y\| \geq t, x^* \in J(x), y^* \in J(y)\}.$$

2.4 If $1 < p < \infty$, and if the $X_n's$ are all strictly convex Banach spaces, show that

$$(\prod_{n\in\mathbb{N}} X_n)_p = \{x = \{x_n\} : x_n \in X_n \text{ for all } n \in \mathbb{N} \text{ and } \sum_{n\in\mathbb{N}} \|x_n\|_{x_n}^p < \infty\}$$

endowed with norm

$$\|x\| = (\sum_{n\in\mathbb{N}} \|x_n\|_{x_n}^p)^{1/p}$$

is strictly convex.

2.5 On $L^2([0,1],dt)$, we consider the norm

$$\|f\| = \left[\frac{1}{2}(\|f\|_2^2 + \|f\|_1^2)\right]^{1/2}.$$

Show that this norm is equivalent to $\|\cdot\|_2$, but is not smooth.

2.6 On ℓ_1, we consider the norm $\|x\| = (\|x\|_1^2 + \|x\|_2^2)^{1/2},\ x = \{x_n\}_{n\in\mathbb{N}}$ (where $\|x\|_1 = \sum_{n\in\mathbb{N}} |x_n|, \|x\|_2 = (\sum_{n\in\mathbb{N}} |x_n|^2)^{1/2}$).
Show that this norm is equivalent to the ℓ_1-norm and that it is strictly convex.

2.7 Let C be a nonempty closed convex subset of a strictly convex Banach space X and D a nonempty subset of C. Let $x \in C$ and P be a sunny nonexpansive retraction of C onto D such that $\|Px - y\| = \|x - y\|$ for some $y \in D$. Then $Px = x$.