Chapter 5 Convex Optimization in Function Space

5.1 Foundations of Convex Analysis

Let V be a vector space over $\mathbb{R}$ and $\|\cdot\| : V \to \mathbb{R}$ be a norm on V. We recall that $(V, \|\cdot\|)$ is called a Banach space, if it is complete, i.e., if any Cauchy sequence $\{v_k\}_{\mathbb{N}}$ of elements $v_k \in V, k \in \mathbb{N}$, converges to an element $v \in V$ ($\|v_k - v\| \to 0$ as $k \to \infty$).

Examples: Let Ω be a domain in $\mathbb{R}^d, d \in \mathbb{N}$. Then, the space $C(\Omega)$ of continuous functions on Ω is a Banach space with the norm

$$\|u\|_{C(\Omega)} := \sup_{x\in\Omega} |u(x)| .$$

The spaces $L^p(\Omega), 1 \le p < \infty$, of (in the Lebesgue sense) p-integrable functions are Banach spaces with the norms

$$\|u\|_{L^p(\Omega)} := \Big(\int_\Omega |u(x)|^p \, dx \Big)^{1/p} .$$

The space $L^\infty(\Omega)$ of essentially bounded functions on Ω is a Banach space with the norm

$$\|u\|_{L^\infty(\Omega)} := \operatorname*{ess\,sup}_{x\in\Omega} |u(x)| .$$

The (topologically and algebraically) dual space V^* is the space of all bounded linear functionals $\mu : V \to \mathbb{R}$. Given $\mu \in V^*$, for $\mu(v)$ we often write $\langle \mu, v\rangle$ with $\langle\cdot,\cdot\rangle$ denoting the dual product between V^* and V. We note that V^* is a Banach space equipped with the norm

$$\|\mu\| := \sup_{v\in V\setminus\{0\}} \frac{|\langle \mu, v\rangle|}{\|v\|} .$$

Examples: The dual of $C(\Omega)$ is the space $\mathcal{M}(\Omega)$ of Radon measures μ with

$$\langle \mu, v\rangle := \int_\Omega v \, d\mu \quad , \quad v \in C(\Omega) .$$

The dual of $L^1(\Omega)$ is the space $L^\infty(\Omega)$. The dual of $L^p(\Omega), 1 < p < \infty$, is the space $L^q(\Omega)$ with q being conjugate to p, i.e., $1/p + 1/q = 1$.
The dual of $L^\infty(\Omega)$ is the space of Borel measures.

A Banach space V is said to be reflexive, if $V^{**} = V$.
In view of the examples before, the spaces $L^p(\Omega), 1 < p < \infty$, are reflexive, but $C(\Omega)$ and $L^1(\Omega), L^\infty(\Omega)$ are not.

We denote by 2^{V^*} the power set of V^*, i.e., the set of all subsets of V^*.

Definition 5.1 (Weighted duality mapping)
Let $h : \mathbb{R}_+ \to \mathbb{R}_+$ be a continuous and non-decreasing function such that $h(0) = 0$ and $h(t) \to +\infty$ as $t \to +\infty$. Then, the mapping

$$J_h(u) \ := \ \{u \in V^* \mid \langle u, u^* \rangle = \|u\| \|u^*\| \ , \ \|u^*\| = h(\|u\|)\}$$

is called the weighted (or gauged) duality mapping, and h is referred to as the weight (or gauge).
The weighted duality mapping is surjective, if and only if V is reflexive.

Example: For $V = L^p(\Omega), V^* = L^q(\Omega), 1 < p, q < +\infty, 1/p + 1/q = 1$, and $h(t) = t^{p-1}$, we have

$$J_h(u)(x) \ := \ \begin{cases} |u(x)|^{p-1} \mathrm{sgn}(u(x)) \ , \ u(x) \neq 0 \\ 0 \ , \ u(x) = 0 \end{cases} .$$

Let V be a Banach space and $u_k \in V, k \in \mathbb{N}$, and $u \in V$.
The sequence $\{u_k\}_{\mathbb{N}}$ is said to converge strongly to u ($u_k \to u$ $(k \to \infty)$ or s-lim $u_k = u$), if $\|u_k - u\| \to 0$ $(k \to \infty)$.
The sequence $\{u_k\}_{\mathbb{N}}$ is said to converge weakly to u ($u_k \rightharpoonup u$ $(k \to \infty)$ or w-lim $u_k = u$), if $\langle \mu, u_k - u \rangle \to 0$ $(k \to \infty)$ for all $\mu \in V^*$.

Theorem 5.2 (Theorem of Eberlein/Shmulyan)
In a reflexive Banach space V a bounded sequence $\{u_k\}_{\mathbb{N}}, u_k \in V, k \in \mathbb{N}$, contains a weakly convergent subsequence, i.e., there exist a subsequence $\mathbb{N}' \subset \mathbb{N}$ and an element $u \in V$ such that $u_k \rightharpoonup u$ $(k \in \mathbb{N}' \to \infty)$.

In the sequel, we assume V to be a reflexive Banach space.

Definition 5.3 (Convex set, convex hull)
Let $u, v \in V$. By $[u, v] \subset V$ we denote the line-segment with endpoints u and v according to

$$[u, v] \ := \ \{\lambda u + (1 - \lambda) v \mid \lambda \in [0, 1]\} \ .$$

A set $A \subset V$ is called convex, if and only if for any $u, v \in A$ the segment $[u, v]$ is contained in A as well.
The convex hull co A of a subset $A \subset V$ is the convex combination of all elements of A, i.e.,

$$\mathrm{co}\ A \ := \ \{\sum_{i=1}^n \lambda_i u_i \mid n \in \mathbb{N}, \sum_{i=1}^n \lambda_i = 1, \lambda_i \geq 0, u_i \in A, 1 \leq i \leq n\} \ .$$

The closure of the convex hull $\overline{\mathrm{co}}\ A$ is said to be the closed convex hull.

Definition 5.4 (Affine hyperplane, supporting hyperplane, separation of sets)
Let $\mu \in V^*, \mu \neq 0$, and $\alpha \in \mathbb{R}$. The set of elements

$$\mathcal{H} \ := \ \{v \in V \mid \mu(v) = \alpha\}$$

is called an affine hyperplane. The convex sets

$$\begin{array}{ccc} \{v \in V \mid \mu(v) < \alpha\} & , & \{v \in V \mid \mu(v) \leq \alpha\} \ , \\ \{v \in V \mid \mu(v) > \alpha\} & , & \{v \in V \mid \mu(v) \geq \alpha\} \end{array}$$

are called open resp. closed half-spaces bounded by $\mathcal{H}$.
If $A \subset V$ and $\mathcal{H}$ is a closed, affine hyperplane containing at least one point $u \in A$ such that A is completely contained in one of the closed half-spaces determined by $\mathcal{H}$, then $\mathcal{H}$ is called a supporting hyperplane of A and u is said to be a supporting point of A.
An affine hyperplane $\mathcal{H}$ is said to separate (strictly separate) two sets $A, B \subset V$, if each of the closed (open) half-spaces bounded by $\mathcal{H}$ contains one of them, i.e.,

$$\begin{array}{lllll} \mu(u) \ \leq \ \alpha \ , \ u \in A & , & \mu(v) \ \geq \ \alpha \ , \ v \in B & \text{resp.} \\ \mu(u) \ < \ \alpha \ , \ u \in A & , & \mu(v) \ > \ \alpha \ , \ v \in B \ . \end{array}$$

Theorem 5.5 (Geometrical form of the Hahn-Banach theorem)
Let $A \subset V$ be an open, non-empty, convex set and M a non-empty affine subspace with $A \cap M = \emptyset$. Then, there exists a closed affine hyperplane $\mathcal{H}$ with $M \subset \mathcal{H}$ and $A \cap \mathcal{H} = \emptyset$.

Corollary 5.6 (Separation of convex sets)
(i) Let $A \subset V$ be an open, non-empty, convex set and $B \subset V$ a non-empty, convex set with $A \cap B = \emptyset$. Then, there exists a closed affine hyperplane $\mathcal{H}$ which separates A and B.
(ii) Let $A \subset V$ be a compact, non-empty convex set and $B \subset V$ a closed, non-empty, convex set with $A \cap B = \emptyset$. Then, there exists a closed affine hyperplane $\mathcal{H}$ which strictly separates A and B.

A consequence of Corollary 5.6 (i) is:

Corollary 5.7 (Boundary of convex sets)
Let $A \subset V$ be a convex set with non-empty interior. Then, any boundary point of A is a supporting point of A.

As a consequence of Corollary 5.6 (ii) we obtain:

Corollary 5.8 (Characterization of closed convex sets)
Any closed convex set $A \subset V$ is the intersection of the closed half-spaces which contain A.
In particular, every closed convex set is weakly closed.

The converse of Corollary 5.8 is known as Mazur's lemma:

Lemma 5.9 (Mazur's Lemma)
Let $\{u_k\}_{\mathbb{N}}, u_k \in V, k \in \mathbb{N}$, and $u \in V$ such that w-lim $u_k = u$. Then, there is a sequence $\{v_k\}_{\mathbb{N}}$ of convex combinations

$$v_k \; = \; \sum_{i=k}^{K} \lambda_i u_i \quad , \quad \sum_{i=k}^{K} \lambda_i = 1 \; , \; \lambda_i \geq 0 \; , \; k \leq i \leq K \; ,$$

such that s-lim $v_k = u$.

The combination of Corollary 5.8 and Lemma 5.9 gives:

Corollary 5.10 (Properties of convex sets)
A convex set $A \subset V$ is closed if and only if it is weakly closed.

Definition 5.11 (Convex function, strictly convex function, effective domain)
Let $A \subset V$ be a convex set and $f : A \to \overline{\mathbb{R}} := [-\infty, +\infty]$. Then, f is said to be convex if for any $u, v \in A$ there holds

$$f(\lambda u + (1-\lambda)v) \; \leq \; \lambda f(u) + (1-\lambda) f(v) \quad , \quad \lambda \in [0,1] \; .$$

A function $f : A \to \overline{\mathbb{R}}$ is called strictly convex if

$$f(\lambda u + (1-\lambda)v) \; < \; \lambda f(u) + (1-\lambda) f(v) \quad , \quad \lambda \in (0,1) \; .$$

A function $f : A \to \overline{\mathbb{R}}$ is called proper convex if $f(u) > -\infty, u \in A$, and $f \not\equiv +\infty$.

If $f : A \to \overline{\mathbb{R}}$ is convex, the convex set

$$\mathrm{dom}\, f \; := \; \{u \in A \mid f(u) < +\infty\}$$

is called the effective domain of f.

Definition 5.12 (Indicator function)
If $A \subset V$, the indicator function χ_A of A is defined by means of

$$\chi_A(u) \; := \; \begin{cases} 0 \; , \; u \in V \\ +\infty \; , \; u \notin V \end{cases} \; .$$

The indicator function of a convex set A is a convex function.

Definition 5.13 (Epigraph of a function)
Let $f : V \to \overline{\mathbb{R}}$ be a function. The set

$$\text{epi } f \ := \ \{(u,a) \in V \times \mathbb{R} \mid f(u) \le a\}$$

is called the epigraph of f. The projection of epi f onto V is the effective domain dom f.

Theorem 5.14 (Characterization of convex functions)
A function $f : V \to \overline{\mathbb{R}}$ is convex if and only if its epigraph is convex.

Proof: Let f be convex and assume $(u,a),(v,b) \in$ epi f. Then, for all $\lambda \in [0,1]$

$$f(\lambda u + (1-\lambda)v) \ \le \ \lambda f(u) + (1-\lambda) f(v) \le \lambda a + (1-\lambda) b \ ,$$

and hence, $\lambda(u,a) + (1-\lambda)(v,b) \in$ epi f.
Conversely, assume that epi f is convex. It suffices to verify the convexity of f on its effective domain. For that purpose, let $u,v \in$ dom f such that $a \ge f(u)$ and $b \ge f(v)$. Since $\lambda(u,a) + (1-\lambda)(v,b) \in$ epi f for every $\lambda \in [0,1]$ it follows that

$$f(\lambda u + (1-\lambda)v) \ \le \ \lambda a + (1-\lambda) b \ .$$

If both $f(u)$ and $f(v)$ are finite, we choose $a = f(u)$ and $b = f(v)$. If $f(u) = -\infty$ or $f(v) = -\infty$, it suffices to allow $a \to -\infty$ resp. $b \to -\infty$.

Definition 5.15 (Lower and upper semi-continuous functions)
A function $f : V \to \overline{\mathbb{R}}$ is called lower semi-continuous on V if there holds

$$\begin{aligned} &\{u \in V \mid f(u) \le a\} \text{ is closed for any } a \in \mathbb{R} \ , \\ &f(u) \ \le \ \liminf_{v \to u} f(v) \text{ for any } u \in V \ . \end{aligned}$$

A function $f : V \to \overline{\mathbb{R}}$ is called weakly lower semi-continuous on V if there holds

$$\begin{aligned} &\{u \in V \mid f(u) \le a\} \text{ is weakly closed for any } a \in \mathbb{R} \ , \\ &f(u) \ \le \ w - \liminf_{v \to u} f(v) \text{ for any } u \in V \ . \end{aligned}$$

A function $f : V \to \overline{\mathbb{R}}$ is called upper semi-continuous (weakly upper semi-continuous) on V if $-f$ is lower semi-continuous (weakly lower semi-continuous) on V.

Examples: (Lower/upper semi-continuous functions)
(i) Let $V := \mathbb{R}$ and

$$J(v) \ := \ \begin{cases} +1 \ , \ v < 0 \\ -1 \ , \ v \ge 0 \end{cases} \ .$$

Then J is lower semi-continuous on $\mathbb{R}$.

(ii) The weighted duality mapping $J_h : V \to 2^{V^*}$ is upper semi-continuous from V endowed with the strong topology onto V^* equipped with the weak-star topology (even for V^* equipped with the bounded weak-star topology).

(iii) The indicator function χ_A of a subset $A \subset V$ is lower semi-continuous (upper semi-continuous) if and only if A is closed (open).

Theorem 5.16 (Characterization of lower semi-continuous functions)
A function $f : V \to \overline{\mathbb{R}}$ is lower semi-continuous if and only if its epigraph epi f is closed.

Proof: Define $\Phi : V \times \mathbb{R} \to \overline{\mathbb{R}}$ by

$$\Phi(u,a) \;:=\; f(u) - a \quad , \quad (u,a) \in V \times \mathbb{R} \; .$$

Then, the lower semi-continuity of f and Φ are equivalent.
For every $r \in \mathbb{R}$, the section $\Phi(V \times [-\infty, r))$ is the set obtained from epi f by a simple translation. It is therefore closed, if and only if epi f is closed.

Corollary 5.17 (Lower semi-continuity of convex functions)
Every lower semi-continuous function $f : V \to \overline{\mathbb{R}}$ is weakly lower semi-continuous.

Proof: By Theorem 5.16, the epigraph epi f is a closed convex set and hence, it is weakly closed by Corollary 5.10.

Definition 5.18 (Lower semi-continuous regularization)
Let $f : V \to \overline{\mathbb{R}}$. The largest lower semi-continuous minorant $\bar{f}$ of f is said to be the lower semi-continuous regularization of f.

Corollary 5.19 (Properties of the lower semi-continuous regularization)
If $f : V \to \overline{\mathbb{R}}$ and $\bar{f}$ is its lower semi-continuous regularization, there holds

$$\begin{aligned} \text{epi } \bar{f} &= \overline{\text{epi } f} \; , \\ \bar{f}(u) &= \liminf_{v \to u} f(v) \; . \end{aligned}$$

Definition 5.20 (Pointwise supremum of continuous affine functions)
Let $\ell \in V^*$ and $\alpha \in R$. A function $g : V \to \mathbb{R}$ of the form $g(v) = \ell(v) + \alpha$ is called an affine continuous function. We denote by $\Gamma(V)$

the set of functions $f : V \to \overline{\mathbb{R}}$ which are the pointwise supremum of a family of continuous affine functions and by $\Gamma_0(V)$ the subset $\Gamma_0(V) := \{f \in \Gamma(V) | f \not\equiv -\infty, f \not\equiv +\infty\}$.

Theorem 5.21 (Characterization of function in $\Gamma(V)$)
For a function $f : V \to \overline{\mathbb{R}}$ there holds $f \in \Gamma(V)$, if and only if f is a lower semi-continuous convex function, and if f attains the value $-\infty$, then $f \equiv -\infty$.

Proof: The necessity follows from the fact that the pointwise supremum of an empty family is $-\infty$. Therefore, if the family under consideration is non-empty, f can not take the value $-\infty$.
Conversely, assume that f is a lower semi-continuous convex function with $f \not\equiv -\infty$. If $f \equiv +\infty$, it obviously is the pointwise supremum of all continuous affine functions. Hence, we consider the case when $f \not\equiv +\infty$.
We show that for every $\bar{u} \in V$ and every $\bar{a} \in \mathbb{R}$ such that $\bar{a} < f(\bar{u})$ there exists a continuous affine function g with $\bar{a} \le g(\bar{u}) \le f(\bar{u})$.
Since epi f is a closed convex set with $(\bar{u}, \bar{a}) \notin$ epi f, there exist $\ell \in V^*$ and $\alpha, \beta \in \mathbb{R}$ such that the closed affine hyperplane

$$\mathcal{H} \ := \ \{(u,a) \in V \times \mathbb{R} \mid \ell(u) + \alpha a = \beta\}$$

separates $(\bar{u}, \bar{a})$ and epi F, i.e.,

$$\begin{array}{lll} (*) & \ell(\bar{u}) \ + \ \alpha\bar{a} \ < \ \beta \ , & \\ (**) & \ell u \ + \ \alpha a \ > \ \beta \quad , & (u,a) \in \text{epi } f \ . \end{array}$$

Case I: $f(\bar{u}) < +\infty$
In this case, we may choose $u = \bar{u}$ and $a = f(\bar{u})$. Then $(*)$ and $(**)$ imply

$$\alpha\Big(f(\bar{u}) - \bar{a}\Big) \ > \ 0 \ ,$$

whence $\alpha > 0$. Dividing $(*)$ and $(**)$ by α yields

$$\bar{a} \ < \ \frac{\beta}{\alpha} \ - \ \frac{1}{\alpha}\ell(\bar{u}) \ < \ f(\bar{u}) \ .$$

Hence, the continuous affine function

$$g(\cdot) := \frac{\beta}{\alpha} \ - \ \frac{1}{\alpha}\ell(\cdot)$$

does the job.
Case II: $f(\bar{u}) = +\infty$
If $\alpha \neq 0$, we may argue as in Case I. If $\alpha = 0$, we set $g(\cdot) := \beta - \ell(\cdot)$. In view of $(*)$ and $(**)$ we have

$$(\diamond) \quad g(\bar{u}) \ > \ 0 \quad , \quad g(u) \ < \ 0 \ , \ u \in \text{dom } f \ .$$

Therefore, there exist $m \in V^*$ and $\gamma \in \mathbb{R}$ such that for $\tilde{g}(\cdot) := \gamma - m(\cdot)$ there holds

$$\tilde{g}(u) \ < \ f(u) \ , \ u \in V \ .$$

Due to ($\diamond$), for every $\kappa > 0$

$$\bar{g}_\kappa(u) \ := \ \tilde{g}(u) + \kappa\Big(\beta - \ell(u)\Big) \ < \ f(u) \ , \ u \in V \ .$$

Finally, we choose $\kappa > 0$ so large that

$$\bar{g}_\kappa(\bar{u}) \ > \ \bar{a} \ ,$$

which shows that the corresponding $\bar{g}_\kappa$ does the job. $\square$

Definition 5.22 (Γ regularization)
The largest minorant $G \in \Gamma(V)$ of $f : V \to \overline{\mathbb{R}}$ in $\Gamma(V)$ is called the Γ regularization of f.

Theorem 5.23 (Properties of the Γ regularization)
Let $G \in \Gamma(V)$ be the Γ regularization of $f : V \to \overline{\mathbb{R}}$. If there exists a continuous affine function $\Phi : V \to \mathbb{R}$ such that $\Phi(u) < f(u), u \in V$, there holds

$$\text{epi } G \ = \ \overline{\text{co}} \text{ epi } f \ .$$

Example: Let $A \subset V$ and χ_A be its indicator function. Then, the Γ regularization of χ_A is the indicator function of its closed convex envelope.

Corollary 5.24 (Lower semi-continuous and Γ regularization)
For $f : V \to \overline{\mathbb{R}}$ let $\bar{f}$ and G be its lower semi-continuous and Γ regularization, respectively. Then, there holds

$$G(u) \ \le \ \bar{f}(u) \ \le \ f(u) \ , \ u \in V \ .$$

If f is convex and admits a continuous affine minorant, then

$$G \ = \ \bar{f} \ .$$

Definition 5.25 (Polar functions)
If $f : V \to \overline{\mathbb{R}}$, then the function $f^* : V^* \to \overline{\mathbb{R}}$ defined by

$$f^*(u^*) \ := \ \sup_{u \in V} \ \Big(\langle u^*, u \rangle - f(u) \Big)$$

is called the polar or conjugate function of f.

Example: Let $A \subset V$ and let χ_A be the indicator function of A. Then, its polar χ_A^* is given by

$$\chi_A^*(u^*) \;=\; \sup_{u \in V} \Big(\langle u^*, u \rangle - \chi_A(u) \Big) \;=\; \sup_{u \in A} \langle u^*, u \rangle \;.$$

It is a lower semi-continuous convex function which is called the support function of A.

Definition 5.26 (Gateaux-differentiability, Gateaux derivative)
A function $f : V \to \overline{\mathbb{R}}$ is called Gateaux-differentiable in $u \in V$, if

$$f'(u; v) \;=\; \lim_{\lambda \to 0_+} \frac{f(u + \lambda v) - f(u)}{\lambda}$$

exists for all $v \in V$. $f'(u; v)$ is said to be the Gateaux-variation of f in $u \in V$ with respect to $v \in V$.
Moreover, if there exists $f'(u) \in V^*$ such that

$$f'(u; v) \;=\; f'(u)(v) \;=\; \langle f'(u), v \rangle \;,\; v \in V \;,$$

then $f'(u)$ is called the Gateaux-derivative of f in $u \in V$.

There are of course functions which are not Gateaux-differentiable. An easy example is given by

$$f(x) \;:=\; |x| \;,\; x \in \mathbb{R} \;,$$

which obviously is not differentiable in $x = 0$.
However, the concept of differentiability can be relaxed by admitting all tangents at the point of non-differentiability which support the epigraph of the function:

Definition 5.27 (Subdifferentiability, subgradient, subdifferential)
A function $f : V \to \overline{\mathbb{R}}$ is said to be subdifferentiable at $u \in V$, if f has a continuous affine minorant ℓ which is exact at u. Obviously, $f(u)$ must be finite, and ℓ has to be of the form

$$\ell(v) \;=\; \langle u^*, v - u \rangle \;+\; f(u) \;=\; \langle u^*, v \rangle \;+\; f(u) \;-\; \langle u^*, u \rangle \;. \tag{5.1}$$

The constant term is the greatest possible, whence

$$f(u) \;-\; \langle u, u^* \rangle \;=\; -f^*(u^*) \;. \tag{5.2}$$

The slope $u^* \in V^*$ of ℓ is said to be the subgradient of f at u, and the set of all subgradients at u will be dnoted by $\partial f(u)$. We have the following characterization

$$\begin{aligned} &u^* \in \partial f(u) \text{ if and only if } f(u) \text{ is finite and} \\ &\langle u^*, v - u \rangle \;+\; f(u) \;\le\; f(v) \;,\; v \in V \;. \end{aligned} \tag{5.3}$$

Example: For the function $f(x) = |x|, x \in \mathbb{R}$, we have

$$\partial f(x) \;=\; \begin{cases} \{-1\} \;,\; x < 0 \\ [-1, +1] \;,\; x = 0 \\ \{+1\} \;,\; x > 1 \end{cases} \;.$$

We see in this example that at points where f only has one subgradient, it coincides with the Gateaux derivative. This property holds true in general:

Definition 5.28 (Subdifferential and Gateaux derivative)
Let $f : V \to \overline{\mathbb{R}}$ be a convex function. If f is Gateaux differentiable at $u \in V$ with Gateaux derivative $f'(u)$, then it is subdifferentiable at $u \in V$ with $\partial f(u) = \{f'(u)\}$.
On the other hand, if f is continuous and finite at $u \in V$ and only has one subgradient, then f is Gateaux differentiable at u with $\{f'(u)\} = \partial f(u)$.

We have seen that if f has a subgradient $u^* \in \partial f(u), u \in V,$, then (5.2) holds true. Conversely, if we assume (5.2), the continuous affine function ℓ as given by (5.1) is everythere less than f and is exact at u. Hence, we have shown:

Theorem 5.29 (Characterization of subgradients)
Assume $f : V \to \overline{\mathbb{R}}$ and denote by $f^* : V^* \to \overline{\mathbb{R}}$ its polar. Then, there holds

$$u^* \in \partial f(u) \quad \Longleftrightarrow \quad f(u) \;+\; f^*(u^*) \;=\; \langle u^*, u\rangle \;. \tag{5.4}$$

The previous result immediately leads us to the following characterization of the subdifferential: Hence, we have shown:

Theorem 5.30 (Characterization of subdifferentials)
If $f : V \to \overline{\mathbb{R}}$ is subdiffernetiable at $u \in V$, then the subdifferential $\partial f(u)$ is convex and weakly* closed in V^*.

Proof: Due to the definition of the polar function there holds

$$f^*(u^*) \;-\; \langle u^*, u\rangle \;\geq\; -f(u) \;.$$

Consequently, in view of (5.4) we have

$$\partial f(u) \;=\; \{u^* \in V^* \mid f^*(u^*) - \langle u^*, u\rangle \leq -f(u)\} \;.$$

Now, let $\{u_n^*\}_{\mathbb{N}}$ be sequence of elements $u_n^* \in \partial f(u), n \in \mathbb{N}$, such that $u_n^* \rightharpoonup^* u^*$ as $n \to \infty$. Then, $\langle u_n^*, u\rangle \to \langle u^*, u\rangle$ and $f^*(u_n^*) \to f^*(u^*)$ as $n \to \infty$, since $f^* \in \Gamma(V^*)$. Consequently, $u^* \in \partial f(u)$. $\square$

Theorem 5.31 (Subdifferential calculus)
(i) Let $f : V \to \overline{\mathbb{R}}$ and $\lambda > 0$. Then, there holds

$$\partial(\lambda f)(u) \ = \ \lambda \partial f(u) \ , \ u \in V \ . \tag{5.5}$$

(ii) Let $f_i : V \to \overline{\mathbb{R}}, 1 \le i \le 2$. Then, there holds

$$\partial(f_1 + f_2)(u) \ \supset \ \partial f_1(u) \ + \ \partial f_2(u) \ , \ u \in V \ . \tag{5.6}$$

(iii) Let $f_i \in \Gamma(V) \to \overline{\mathbb{R}}, 1 \le i \le 2$. If there exists $\tilde{u} \in \mathrm{dom}\ f_1 \cap \mathrm{dom}\ f_2$ where f_1 is continuous, there holds

$$\partial(f_1 + f_2)(u) \ = \ \partial f_1(u) \ + \ \partial f_2(u) \ , \ u \in V \ . \tag{5.7}$$

(iv) Let Y be another Banach space with dual Y^* and $A : V \to Y$ be a continuous linear mapping with adjoint $A^* : Y^* \to V^*$ and $f \in \Gamma(Y)$. Assume that there exists $A\tilde{u} \in Y$ where f is continuous and finite. Then, there holds

$$\partial(f \circ A)(u) \ = \ A^* \partial f(u) \ , \ u \in V \ . \tag{5.8}$$

The notion of subdifferentiability allows us to consider optimization problems for subdifferentiable functions:

$$\inf_{v \in V} f(v) \ .$$

Obviously, a necessary optimality condition for $u \in V$ to be a minimizer of f is

$$0 \in \partial f(u) \ .$$

Another important example is that of a constrained optimization problem for a Gateaux differentiable function f:

$$\inf_{v \in K} f(v) \ ,$$

where $K \subset V$ is supposed to be a closed convex set. Then, we can restate the constrained as an unconstrained problem by means of the indicator function I_K of K:

$$\inf_{v \in V} \Big(f(v) \ + \ I_K(v) \Big)$$

and get the necessary optimality condition

$$0 \in f'(u) \ + \ \partial I_K(u) \ .$$

The subdifferential $\partial f(\cdot)$ is a particular example of a multivalued mapping from V into 2^{V^*}. Earlier, we have come across the weighted duality mapping J_h (with weight h) as a further example. Actuality, the duality mapping also represents a subdifferential:

Lemma 5.32 (Duality mapping as subdifferential)
Let $J_h : V \to 2^{V^*}$ be the duality mapping with weight h. Define $H(t) := \int_0^t h(s)ds$. and $j_h = H \circ \|\cdot\|$. Then, $J_h = \partial j_h$.

Proof: The result follows from Theorem 5.31 (iv). $\square$

Definition 5.33 (Generalized Moreau-Yosida approximation)
Let $M : V \to 2^{V^*}$ be a multivalued mapping. Then, its generalized Moreau-Yosida approximation $M_\lambda, \lambda > 0$, is given by

$$M_\lambda \ := \ \Big(M^{-1} \ + \ \lambda J_\lambda^{-1}\Big)^{-1} \ . \tag{5.9}$$

M is said to be regularizable, if for any $\lambda > 0$ the multivalued map $M^{-1} + \lambda J_h^{-1}$ is surjective, i.e.,

$$(M^{-1} + \lambda J_h^{-1})(V^*) \ = \ V \ .$$

In this case, dom $M_\lambda = V$.

The generalized Moreau-Yosida approximation can be computed by means of the Moreau-Yosida resolvent:

Definition 5.34 (Moreau-Yosida resolvent)
Let $M : V \to 2^{V^*}$ be a multivalued mapping and $\lambda > 0$. The Moreau-Yosida resolvent (Moreau-Yosida proximal map) $P_\lambda^M : V \to V$ is given by

$$P_\lambda^M(w) \ = \ \{v \in V \mid 0 \in J_h(\frac{v-w}{\lambda}) + M(v)\} \ , \ w \in V \ . \tag{5.10}$$

Example: If $K \subset V$ is a closed convex set and I_K its indicator function, then $P_\lambda^{\partial I_K}(w), w \in V$, is the metric projection of w onto K.

For a lower semi-continuous proper convex function f with subdifferential ∂f, we have the following characterization of the Moreau-Yosida resolvent:

Theorem 5.35 (Moreau-Yosida resolvent of a subdifferentiable function)
Let $f : V \to \overline{\mathbb{R}}$ be a lower semi-continuous proper convex function with subdifferential ∂f. Then, for $w \in V$, the Moreau-Yosida resolvent $P_\lambda^{\partial f}(w)$ is the set of minimizers of

$$\inf_{v \in V} \ f(v) \ + \ \lambda j_h(\frac{v-w}{\lambda}) \ .$$

Proof: The function $j_{w,\lambda} : V \to \overline{\mathbb{R}}$ as given by

$$j_{w,\lambda}(v) \ := \ \lambda j_h(\frac{v-w}{\lambda}) \ , \ v \in V \ ,$$

is finite, convex and continuous. Then, Theorem 5.31 implies

$$0 \in \partial(f + j_{w,\lambda})(v) \;=\; \partial f(v) \;+\; \partial j_{w,\lambda}(v) \;=\; \partial f(v) \;+\; J_h(\frac{v-w}{\lambda}) \; .$$

□

Theorem 5.36 (Moreau-Yosida approximation and Moreau-Yosida resolvent, Part I)
For any $\lambda > 0$ there holds

$$\text{dom } M_\lambda \;=\; \text{dom } P_\lambda^M \; , \tag{5.11}$$

and for any $w \in V$ we have

$$M_\lambda(w) \;=\; \bigcup_{v \in P_\lambda^M(w)} \Big(J_h(\frac{w-v}{\lambda}) \cap M(v) \Big) \; . \tag{5.12}$$

Note that $J_h(-v) = -J_h(v), v \in V$.
Proof: For $w \in \text{dom } P_\lambda^M$ and $v \in P_\lambda^M(w)$ there exists

$$v^* \in J_h(\frac{w-v}{\lambda}) \cap M(v) \; ,$$

and hence,

$$v \in M^{-1}(v^*) \quad , \quad \lambda^{-1}(w-v) \in J_h^{-1}(v^*) \; .$$

It follows that

$$w \in \Big(M^{-1} + \lambda J_h^{-1}\Big)(v^*) \quad \Longleftrightarrow \quad v^* \in \Big(M^{-1} + \lambda J_h^{-1}\Big)^{-1}(w) \; ,$$

which proves $v^* \in M_\lambda(w)$.
On the other hand, if $v^* \in M_\lambda(w)$, there exist $v \in M^{-1}(v^*)$ and $z \in J_h^{-1}(v^*)$ such that $w = v + \lambda z$. We deduce

$$v^* \in J_h(\lambda^{-1}(w-v)) \cap M(v) \; ,$$

whence $v \in P_\lambda^M(w)$. □

Corollary 5.37 (Moreau-Yosida approximation and Moreau-Yosida resolvent, Part II)
If J_h is single-valued, then for $\lambda > 0$ and $w \in V$ there holds

$$M_\lambda(w) \;=\; J_h(\lambda^{-1}w - \lambda^{-1}P_\lambda^M(w)) \; . \tag{5.13}$$

Proof: Since $M_\lambda(w) \subset J_h(\lambda^{-1}w - \lambda^{-1}P_\lambda^M(w))$ follows from the previous result, we only have to show $M_\lambda(w) \supset J_h(\lambda^{-1}w - \lambda^{-1}P_\lambda^M(w))$. For that purpose, let $w \in \text{dom } P_\lambda^M$ and $v \in P_\lambda^M(w)$ such that

$$v^* \in J_h(\lambda^{-1}w - \lambda^{-1}v) \; .$$

Let $z^* \in J_h(\lambda^{-1}(w-v)) \cap M(v)$. Since $J_h(\lambda^{-1}(w-v))$ consists of a single element, we must have $v^* = z^*$, whence

$$v^* \in J_h(\lambda^{-1}(w-v)) \cap M(v) \subset M_\lambda(w) \ .$$

$\square$

Example: We recall the example $f(x) = |x|, x \in \mathbb{R}$, where

$$\partial f(x) \ = \ \begin{cases} \{-1\} \ , \ x < 0 \\ [-1,+1] \ , \ x = 0 \\ \{+1\} \ , \ x > 1 \end{cases} \ .$$

Corollary 5.37 allows to compute the Moreau-Yosida approximation $(\partial f)_\lambda$. In case of the duality mapping J_h with weight $h(t) = t^{p-1}, 1 < p < +\infty$, we obtain

$$(\partial f)_\lambda(w) \ = \ \begin{cases} -1 \ , \ w < -\lambda \\ , \\ \{|\frac{w}{\lambda}|^{p-2} \ \frac{w}{\lambda}\} \ , \ w \in [-\lambda, +\lambda] \\ , \\ +1 \ , \ w > \lambda \end{cases} \ .$$

5.2 Convex Optimization Problems

We assume that $(V, \|\cdot\|)$ is a reflexive Banach space.

Definition 5.38 (Coercive functionals)

A functional $J : V \to \overline{\mathbb{R}}$ is said to be coercive, if

$$J(v) \to +\infty \quad \text{for } \|v\|_V \to +\infty \ .$$

Theorem 5.39 Solvability of unconstrained minimization problems

Suppose that $J : V \to (-\infty, +\infty], J \not\equiv +\infty$, is a weakly semi-continuous, coercive functional. Then, the unconstrained minimization problem

$$J(u) \ = \ \inf_{v \in V} J(v) \tag{5.14}$$

admits a solution $u \in V$.

Proof: Let $c := \inf_{v \in V} J(v)$ and assume that $\{v_n\}_{n \in \mathbb{N}}$ is a minimizing sequence, i.e., $J(v_n) \to c$ $(n \to \infty)$.

Since $c < +\infty$ and in view of the coercivity of J, the sequence $\{v_n\}_{n \in \mathbb{N}}$ is bounded. Consequently, in view of Theorem 5.1 there exist a subsequence $\mathbb{N}' \subset \mathbb{N}$ and $u \in V$ such that $v_n \rightharpoonup u$ $(n \in \mathbb{N}')$. The weak lower semi-continuity of J implies

$$J(u) \ \le \ \inf_{n \in \mathbb{N}'} J(v_n) \ = \ c \ ,$$

whence $J(u) = c$.

Theorem 5.40 (Existence and uniqueness)
Suppose that $J : V \to \overline{\mathbb{R}}$ is a proper convex, lower semi-continuous, coercive functional. Then, the unconstrained minimization problem (5.14) has a solution $u \in V$.
If J is strictly convex, then the solution is unique.
Proof: The existence follows from Theorem 5.39.
For the proof of the uniqueness let $u_1 \neq u_2$ be two different solutions. Then there holds

$$J(\frac{1}{2}(u_1 + u_2) \ < \ \frac{1}{2} \, J(u_1) \ + \ \frac{1}{2} \, J(u_2) \ = \ \inf_{v \in V} J(v) \ ,$$

which is a contradiction.

We recall that in the finite dimensional case $V = \mathbb{R}^n$, a necessary optimality condition for (5.14) is that $\nabla J(u) = 0$, provided J is continuously differentiable. This can be easily generalized to the infinite dimensional case.

Theorem 5.41 (Necessary optimality condition)
Assume that $J : V \to \overline{\mathbb{R}}$ is Gateaux-differentiable in $u \in V$ with Gateaux-derivative $J'(u) \in V^*$. Then, the variational equation

$$\langle J'(u), v \rangle \ = \ 0 \quad , \quad v \in V \tag{5.15}$$

is a necessary condition for $u \in V$ to be a minimizer of J.
If J is convex, then this condition is also sufficient.
Proof: Let $u \in V$ be a minimizer of J. Then, there holds

$$J(u \pm \lambda v) \ \ge \ J(u) \quad , \quad \lambda > 0 \ , \ v \in V \ ,$$

whence

$$\langle J'(u), \pm v \rangle \ \ge \ 0 \quad , \quad v \in V \ ,$$

and thus

$$\langle J'(u), v \rangle \ = \ 0 \quad , \quad v \in V \ .$$

If J is convex and (5.2) holds true, then

$$J(u + \lambda(v - u)) \ = \ J(\lambda v + (1 - \lambda) u) \ \le \ \lambda J(v) + (1 - \lambda) J(u) \ ,$$

and hence,

$$\begin{aligned} 0 \quad &= \quad < J'(u), v - u >_{V',V} \; = \; \lim_{\lambda \to 0_*} \frac{J(u + \lambda(v-u)) - J(u)}{\lambda} \quad \leq \\ &\leq \quad J(v) - J(u) \; . \end{aligned}$$