# Lectures Notes on Convex Analysis

Leandro Cioletti

Departamento de Matemática - Universidade de Brasília
*Email address*: cioletti@mat.unb.br

2010 *Mathematics Subject Classification.* Primary: 46-XX, 46N10,47N10.

*Key words and phrases.* Convex Analysis, Functional Analysis, Locally Convex Spaces, Variational Calculus

# Contents

CHAPTER 1

# General Topology

*"Point set topology is a disease which the human race will soon recover"*
—attributed to H. Poincaré.

The aim of this chapter is only to fix some notations and recall important notions of General Topology. The material covered is completely standard and should be presented in most of (under)graduate text books on this subject. The presentation here follows closely the references [**2, 4, 5, 6**].

## 1.1. Topological Spaces

This chapter introduces what sometimes is called point set topology. It is not our intention to provide motivations or examples for the concepts introduced here and neither background on metric spaces, although they should appear sooner or latter. The goal is to conduct the discussion towards to understand and prove compactness theorems in infinite dimensional vector spaces in some suitable weak topologies. A lot of definitions will be required to achieve this goal and due to lack of time, very occasionally examples are given. So we encourage the reader to take a look at the above cited references for a "conventional" (motivation, examples and so on...) exposition.

Definition 1.1 (Topological Space). A topological space is an ordered pair $(X, \tau)$, where $X$ is a set and $\tau$ is a collection of subsets of $X$ (the topology) that obeys three axioms:

(1) $\emptyset, X \in \tau$;
(2) for all $U_1, \ldots, U_n \in \tau$ we have $U_1 \cap \ldots \cap U_n \in \tau$;
(3) if $\{U_\alpha\}_{\alpha\in I}$ is an arbitrary subcollection of $\tau$, then $\bigcup_{\alpha\in I} U_\alpha \in \tau$.

The sets in $\tau$ are called *open sets*.

When there is no danger of confusion we denote a topological space $(X, \tau)$ simply by its underlying set $X$. So when we say things like, let $X$ be a topological space the topology will be completely clear from the context or the statement will be valid for a general topology on $X$.

Proposition 1.2. *Let $I$ be an arbitrary index set. If $\tau_i$, for each $i \in I$, is a topology on $X$. Then the collection*

$$\tau \equiv \{U \subset X : U \in \tau_i,\ \forall i \in I\}$$

*is a topology on $X$. In other words, the intersection of any family of topologies on $X$ is a topology on $X$.*

Definition 1.3 (Closed Sets). Let $(X, \tau)$ be a topological space. We say that a set $A \subset X$ is a *closed set* if $A^c \equiv X \setminus A$ is an open set.

The proof of the following theorem is an obvious application of de Morgan's laws together with the definition of a topology on $X$, and will be omitted.

THEOREM 1.4. *If $\mathscr{C}$ is the collection of all closed sets in a topological space $(X,\tau)$, then*

*i)* $\emptyset, X \in \mathscr{C}$;
*ii)* *if* $A_1, \dots, A_n \in \mathscr{C}$, *then* $A_1 \cup \ldots \cup A_n \in \mathscr{C}$;
*iii)* *if $I$ is an arbitrary index set and* $A_\alpha \in \mathscr{C}$, $\forall \alpha \in I$. *Then* $\bigcap_{\alpha \in I} A_\alpha \in \mathscr{C}$.

*Conversely, given a set $X$ and any family $\mathscr{C}$ of subsets of $X$ satisfying* i)–iii)*, the collection of complements of members of $\mathscr{C}$ is a topology on $X$ in which the family of closed sets is exactly $\mathscr{C}$.*

This theorem is a result of, and illustrates, the obvious duality between the notions of open and closed set. More formally, any result about the open sets in a topological space becomes a result about closed sets upon replacing "open" by "closed" and interchanging $\bigcup$ by $\bigcap$.

DEFINITION 1.5 (Closure). Let $(X,\tau)$ be a topological space and $E \subset X$. The *closure* of $E$ in $(X,\tau)$ is the set

$$\overline{E} \equiv \bigcap \{A \subset X : A \text{ is closed and } E \subset A\}.$$

Note that $\overline{E}$ is the "smallest" closed set containing $E$, in the sense that it is contained in every closed set containing $E$.

LEMMA 1.6. *Let $(X,\tau)$ be a topological space and $A \subset B \subset X$. Then $\overline{A} \subset \overline{B}$.*

THEOREM 1.7. *Let $(X,\tau)$ be a topological space and $\mathscr{P}(X)$ the set of parts of $X$. The operation $\mathscr{P}(X) \ni A \longmapsto \overline{A} \in \mathscr{P}(X)$ has the following properties:*

*i)* $A \subset \overline{A}$;
*ii)* $\overline{\overline{A}} = \overline{A}$;
*iii)* $\overline{A \cup B} = \overline{A} \cup \overline{B}$;
*iv)* $\overline{\emptyset} = \emptyset$;
*v)* *$A$ is closed if and only if* $A = \overline{A}$.

*Moreover, if $\Phi : \mathscr{P}(X) \to \mathscr{P}(X)$ is a set-function satisfying* i)– v) *and we define the collection $\mathscr{C}$ of the sets satisfying* v)*. We have that the complements of members of $\mathscr{C}$ determines a topology on $X$, whose the closure operation is just the operation $\Phi$ we began with.*

Any operation $\mathscr{P}(X) \ni A \longmapsto \overline{A} \in \mathscr{P}(X)$ in a set $X$ satisfying *i*)– *v*) is called a *Kuratowski* closure operation. Thus every Kuratowski closure operation determines and is determined by some topology.

DEFINITION 1.8 (Interior). Let $(X,\tau)$ be a topological space and $E \subset X$. The *interior* of $E$ in $(X,\tau)$ is the set

$$\operatorname{Int}(E) \equiv \bigcup \{U \subset X : \ U \text{ is open and } U \subset E\}.$$

Evidently, $\operatorname{Int}(E)$ is an open set. It is the largest open set contained in $E$, in the sense that it contains any other open set contained in $E$. We also remark that it might be possible, for some sets to have $\operatorname{Int}(E) = \emptyset$.

The notions of interior and closure are dual to each other, in much the same way that are "open" and "closed". The strictly formal nature of this duality can be brought out by observing that

$$X \setminus \mathrm{Int}(E) = \overline{X \setminus E}$$
$$X \setminus \overline{E} = \mathrm{Int}(X \setminus E).$$

Thus any theorem about closures in a topological space can be translated to a theorem about interiors. The next results are, for example, of this duality.

LEMMA 1.9. *Let* $(X, \tau)$ *be a topological space and* $A \subset B \subset X$*. Then we have* $\mathrm{Int}(A) \subset \mathrm{Int}(B)$.

THEOREM 1.10. *Let* $(X, \tau)$ *be a topological space and* $\mathscr{P}(X)$ *the set of parts of* $X$*. The operation* $\mathscr{P}(X) \ni A \longmapsto \mathrm{Int}(A) \in \mathscr{P}(X)$ *has the following properties:*

*i)* $\mathrm{Int}(A) \subset A$;
*ii)* $\mathrm{Int}(\mathrm{Int}(A)) = \mathrm{Int}(A)$;
*iii)* $\mathrm{Int}(A \cap B) = \mathrm{Int}(A) \cap \mathrm{Int}(B)$;
*iv)* $\mathrm{Int}(X) = X$;
*v)* $A$ *is open if and only if* $A = \mathrm{Int}(A)$.

*Moreover, if* $\Psi : \mathscr{P}(X) \to \mathscr{P}(X)$ *is a set-function satisfying* i)– iv) *and we define* $\tau$ *as the collection of the sets satisfying* v)*, then* $\tau$ *is a topology on* $X$*. In this topology the interior of a set* $A$ *is just* $\Psi(A)$.

## 1.2. Neighborhoods

The characterizations we have so far to describe a topology (open sets, the closure operation and so on) are not the most convenient, and for this reason are rarely used. In this section we present the two most popular ways to describe topologies.

Very often the topology we wish to present is quite "regular", in the sense that the open sets containing one point look no different from the open sets containing any other(this is true for example, in the Euclidean spaces). In such cases one can describe the topology by describing what it is look like "around" one point, or a few points. Considerable saving of effort can result, and topologies will often be presented this way here, so we will present in what follows a detailed discussion of the "local" description of topologies and topological concepts.

DEFINITION 1.11 (Neighborhoods). Let $(X, \tau)$ be a topological space and $x \in X$ an arbitrary point. A *neighborhood* of $x$ is a subset $U \subset X$ satisfying:

- $x \in U$;
- there is $V \in \tau$ such that $x \in V$ and $V \subset U$.

Clearly, $U$ is a neighborhood of $x$ if and only if $x \in \mathrm{Int}(U)$.

DEFINITION 1.12 (Neighborhood System at $x$). Let $(X, \tau)$ be a topological space and $x \in X$ an arbitrary point. The collection

$$\mathscr{U}_x \equiv \{U \subset X : U \text{ is a neighborhood of } x\}$$

is called *neighborhood system at* $x$.

The next result lists some fundamental properties of a neighborhood system at $x$ in a topological space and provides a converse which says whenever neighborhoods has been assigned to each point of a set, satisfying these properties, one has a topology.

THEOREM 1.13. *Let $(X, \tau)$ be a topological space, $x \in X$ an arbitrary point, and $\mathscr{U}_x$ a* neighborhood *system at $x$. Then*

*i) if $U \in \mathscr{U}_x$, then $x \in U$;*
*ii) if $U, V \in \mathscr{U}_x$, then $U \cap V \in \mathscr{U}_x$;*
*iii) if $U \in \mathscr{U}_x$, then $\exists V \in \mathscr{U}_x$ such that for each $y \in V$, we have $U \in \mathscr{U}_y$;*
*iv) if $U \in \mathscr{U}_x$ and $U \subset V$, then $V \in \mathscr{U}_x$;*
*v) $A$ is open, if and only if $A$ contains a neighborhood of each of its points.*

*Conversely, if in a set $X$ a nonempty collection $\mathscr{U}_x$ of subsets of $X$ is assigned to each $x \in X$ satisfying i) –iv) and v) is used to define a collection $\tau$ of subsets of $X$, then $\tau$ is a topology on $X$, in which the neighborhood system at $x \in X$ is precisely the collection $\mathscr{U}_x$.*

DEFINITION 1.14 (Neighborhood base at $x$). Let $(X, \tau)$ be a topological space, $x \in X$ an arbitrary point, and $\mathscr{U}_x$ a *neighborhood* system at $x$. A subcollection $\mathscr{B}_x$ taken from $\mathscr{U}_x$, having the property that each $U \in \mathscr{U}_x$ contains some $V \in \mathscr{B}_x$ is called a *neighborhood base at $x$*. That is, $\mathscr{U}_x$ must be determined by $\mathscr{B}_x$ in the following sense

$$\mathscr{U}_x = \{U \subset X : \text{ there exists some } V \in \mathscr{B}_x \text{ such that } V \subset U\}.$$

Once a neighborhood base at $x$ has been chosen (there are many to choose from, all producing the same neighborhood system at $x$) its elements are called *basic neighborhoods*. Obviously, a neighborhood system at $x$ is itself always a neighborhood base at $x$. In general, we are interested in smaller basis.

We turn now to the problem of specifying a topology by giving a collection of basic neighborhoods at each point of the space.

The following theorem is used much more often than the corresponding Theorem 1.13 about neighborhood systems.

THEOREM 1.15. *Let $(X, \tau)$ be a topological space and for each $x \in X$, let $\mathscr{B}_x$ be a neighborhood base at $x$. Then*

*i) if $V \in \mathscr{B}_x$, then $x \in V$;*
*ii) if $V_1, V_2 \in \mathscr{B}_x$, then $\exists V_3 \in \mathscr{B}_x$ such that $V_3 \subset V_1 \cap V_2$;*
*iii) if $V \in \mathscr{B}_x$, there is some $V_0 \in \mathscr{B}_x$ such that for each $y \in V_0$ there is some $W \in \mathscr{B}_y$ with $W \subset V$;*
*iv) $A \subset X$ is open if and only if $A$ contains a basic neighborhood of each of its points.*

*Conversely, in a set $X$ if a collection $\mathscr{B}_x$ of subsets of $X$ is assigned to each $x \in X$ satisfying i)– iii) and if we define a collection $\tau$ by using v), then $\tau$ is a topology on $X$ in which $\mathscr{B}_x$ is a neighborhood base at $x$, for each $x \in X$.*

Since neighborhood bases are important descriptive devices in dealing with topologies, it will be useful to have neighborhood characterizations of all concepts so far introduced for topological spaces.

THEOREM 1.16. *Let $(X, \tau)$ be a topological space and suppose that $\mathscr{B}_x$, for each $x \in X$, is a fixed neighborhood base at $x$. Then*

*i)* $A \subset X$ *is open if and only if for each* $x \in A$ *there is some* $U \in \mathscr{B}_x$ *such that* $U \subset A$*;*
*ii)* $A \subset X$ *is closed if and only if for each point* $x \in X \setminus A$ *there is* $U \in \mathscr{B}_x$ *such that* $A \cap U = \emptyset$*;*
*iii)* $\overline{A} = \{x \in X : \forall U \in \mathscr{B}_x \text{ we have } U \cap A \neq \emptyset\}$*;*
*iv)* $\mathrm{Int}(A) = \{x \in X : \exists U \in \mathscr{B}_x \text{ satisfying } U \subset A\}$*;*
*v)* $\partial A = \{x \in X : \forall U \in \mathscr{B}_x \text{ we have } U \cap A \neq \emptyset \text{ and } U \cap (X \setminus A) \neq \emptyset\}$*.*

Roughly speaking the next theorem states that “small neighborhoods make large topologies”. This is intuitive since, smaller the neighborhoods in a space are, the easier it is for a set to contain neighborhoods of all its points and then the more open sets there will be.

THEOREM 1.17 (Hausdorff Criterion). *For each* $x \in X$ *let* $\mathscr{B}_x^1$ *be a neighborhood base at* $x$ *for a topology* $\tau_1$ *on* $X$*, and* $\mathscr{B}_x^2$ *be a neighborhood base at* $x$ *for a topology* $\tau_2$ *on* $X$*. Then* $\tau_1 \subset \tau_2$ *if and only if for each* $x \in X$ *given* $B_1 \in \mathscr{B}_x^1$ *there is some* $B_2 \in \mathscr{B}_x^2$ *such that* $B_2 \subset B_1$*.*

We close this section by introducing a concept that will play a major role in the subsequent sections.

DEFINITION 1.18 (Accumulation point). An *accumulation point* of a set $A$ in a topological space $(X, \tau)$ is a point $x \in X$ such that each neighborhood (basic neighborhood, if you prefer) of $x$ contains some point of $A$, other than $x$. The set of all accumulation points of $A$ is denoted by $A'$ and sometimes called *derived* set of $A$.

THEOREM 1.19. *Let* $(X, \tau)$ *be a topological space. Then for any* $A \subset X$ *we have* $\overline{A} = A \cup A'$*.*

DEFINITION 1.20 (Cluster point). Let $(X, \tau)$ be a topological space. A point $x \in X$ is a *cluster point of a sequence* $(x_n)_{n\in\mathbb{N}}$ if for any open set $A \ni x$ and any $N \in \mathbb{N}$ there exists $n \geq N$ so that $x_n \in A$.

Note that the notions of cluster and accumulation points are, in general, different. For example, consider the topological space $(X, \tau)$, where $X = \{0\} \cup \{n^{-1} : n \geq 1\}$ and $\tau$ is the discrete topology on $X$. Let $(x_n)_{n\in\mathbb{N}}$ be a constant sequence on $X$ given by $x_n = x \neq 0$, for all $n \in \mathbb{N}$. Then $x$ is a cluster point of $(x_n)_{n\in\mathbb{N}}$, but $x$ is not an accumulation point of $A = (x_n)_{n\in\mathbb{N}}$.

## 1.3. Bases and subbases

As we observed in last section, we can specify the neighborhood system at a point $x$ of a topological space $(X, \tau)$ by giving a somewhat smaller collection of sets, that is, a neighborhood base at $x$. In much the same way, the idea of *a base for a topology* $\tau$ will be a way to specify it without the needing of describe each and every of its open sets.

DEFINITION 1.21 (Base for a topology). Let $(X, \tau)$ be a topological space. A collection $\mathscr{B} \subset \tau$ is a base for $\tau$ if

$$\tau = \left\{ \bigcup_{U \in \mathscr{C}} U : \mathscr{C} \subset \mathscr{B} \right\}.$$

As we can see if $\mathscr{B}$ is a base for a topology $\tau$ on $X$, then the topology can be recovered from $\mathscr{B}$ by taking all possible unions of subcollections from $\mathscr{B}$.

Proposition 1.22. *A collection $\mathscr{B}$ is a base for $(X,\tau)$ if and only if whenever $A$ is a open set and $x \in A$ is an arbitrary point, there is some $U \in \mathscr{B}$ such that $x \in U \subset A$.*

The following theorem is similar to theorems 1.7, 1.10, 1.13, and 1.15. That is, it list a few properties that a bases enjoy and provides the converse assertion: any collection of subsets of $X$ enjoying these properties provides a topology on $X$.

Theorem 1.23. *A collection $\mathscr{B}$ of subsets of a set $X$ is a base for a topology $\tau(\mathscr{B}) = \left\{\bigcup_{U\in\mathscr{C}} U : \mathscr{C} \subset \mathscr{B}\right\}$ on $X$ if and only if*

*i) $X = \bigcup_{U\in\mathscr{B}} U$;*
*ii) for any pair $U_1, U_2 \in \mathscr{B}$ with $x \in U_1 \cap U_2$, there exists $U_3 \in \mathscr{B}$ such that $x \in U_3 \subset U_1 \cap U_2$.*

The reader might well suspected that more than a casual similarity exists between the idea of a neighborhood base at each point of $X$ on the one hand and the notion of a base for the topology of $X$ on the other hand. Indeed, as the next theorem make clear, the only real difference between the two notions is that neighborhoods bases need not consist of open sets.

Theorem 1.24. *If $\mathscr{B}$ is a collection of open sets in $(X,\tau)$, $\mathscr{B}$ is a base for $\tau$ if and only if for each $x \in X$, the collection $\mathscr{B}_x \equiv \{U \in \mathscr{B} : x \in U\}$ is a neighborhood base at $x$.*

We can go one step further in reducing the size of the collection we must specify to describe a topology. Recall that the reduction from topology to a base was accomplished essentially by dropping the requirement that any union of elements of $\tau$ belongs to $\tau$. The further reduction to subbase is accomplished essentially by dropping the requirement that any finite intersection of elements of $\tau$ belongs to $\tau$.

Definition 1.25 (Subbase). If $(X,\tau)$ is a topological space, a *subbase* for $\tau$ is a collection $\mathscr{C} \subset \tau$ such that the collection of all finite intersections of elements from $\mathscr{C}$ forms a base for $\tau$.

Theorem 1.26. *Any collection $\mathscr{C}$ of subsets of $X$ is a subbase for some topology $\tau(\mathscr{C})$ on $X$. Moreover,*

$$\tau(\mathscr{C}) \equiv \bigcap_{\substack{\tau\supset\mathscr{C} \\ \tau \text{ is a topology on } X}} \tau,$$

*that is, the smallest topology on $X$ containing $\mathscr{C}$.*

## 1.4. Subspaces

A subset of a topological space inherits a topology of its own, in a obvious way. This topology and some of its easily developed properties will be presented here.

Definition 1.27 (Subspace). Let $(X,\tau)$ be a topological space and $Y \subset X$. The collection $\tau_Y \equiv \{U \cap Y : U \in \tau\}$ is a topology on $Y$, called *relative topology* for $Y$. The fact that a subset of $X$ is being given this topology is signified by referring to it as a subspace of $X$.

Note that any subspace of a discrete topological space is discrete and any subspace of a trivial space is trivial. A subspace of a subspace is a subspace. That is, if $Z \subset Y \subset X$, then the relative topology induced on $Z$ by the relative topology of $Y$ is just the relative topology of $Z$ in $X$.

The open sets in a subspace $Y$ of $X$ are the intersections with $Y$ of the open sets in $X$. Most, but no all, of the related topological notions are introduced into $Y$ in the same way, by intersection, as the following theorem shows.

THEOREM 1.28. *If $(Y, \tau_Y)$ is a subspace of $(X, \tau)$ then*

*i) $V \subset Y$ is open in $(Y, \tau_Y)$ if and only if $V = U \cap Y$, where $U$ is open in $(X, \tau)$;*
*ii) $B \subset Y$ is closed in $(Y, \tau_Y)$ if and only if $B = A \cap Y$, where $A$ is closed in $(X, \tau)$;*
*iii) If $A \subset Y$, then $\mathrm{Cl}_Y(A) = Y \cap \mathrm{Cl}_X(A)$, where $\mathrm{Cl}_Y(A)$ is the closure of $A$ in $(Y, \tau_Y)$.*
*iv) if $x \in Y$, then $V$ is a neighborhood of $x$ in $(Y, \tau_Y)$ if and only if $V = U \cap Y$, where $U$ is a neighborhood of $x$ in $(X, \tau)$.*
*v) if $x \in Y$, and $\mathscr{B}_x$ is neighborhood base at $x$ in $(X, \tau)$, then $\{U \cap Y : U \in \mathscr{B}_x\}$, is a neighborhood base at $x$ in $(Y, \tau_Y)$.*
*vi) if $\mathscr{B}$ is a base for $\tau$, then $\{U \cap Y : U \in \mathscr{B}\}$ is a base for $\tau_Y$.*

## 1.5. Continuous Functions

It is the aim this section to define continuous functions on a topological space and establish their elementary properties.

DEFINITION 1.29. Let $(X, \tau_X)$ and $(Y, \tau_Y)$ be topological spaces. A function $f : X \to Y$ is continuous at $x_0 \in X$ if and only if for each neighborhood $V$ of $f(x_0)$ in $Y$, there is a neighborhood $U$ of $x_0$ in $X$ such that $f(U) \subset V$. We say that $f$ is continuous on $X$ if $f$ is continuous at each $x_0 \in X$.

Is left to the reader as exercise to verify that the concept of continuous is not altered if we replace neighborhoods by basic neighborhoods throughout.

In the sequel we see an alternative, and somewhat surprising set of characterizations of continuity. This theorem, in one or another of its form, is more often used to check "global" continuity than the alternative, that is, check continuity at each point of $X$ individually. The fourth characterization, although not often used as a test for continuity, is interesting. It provide us with description of the continuous functions $f : X \to Y$ as precisely those functions which take the points close to a set $E$ in $X$ close it its image in $Y$.

THEOREM 1.30. *If $(X, \tau_X)$ and $(Y, \tau_Y)$ are topological spaces and $f : X \to Y$. Then the following are all equivalent:*

*i) $f$ is continuous;*
*ii) for each open set $V \subset Y$, we have that $f^{-1}(V)$ is open in $X$;*
*iii) for each closed set $B \subset Y$, we have that $f^{-1}(B)$ is closed in $X$;*
*iv) for each set $A \subset X$, we have that $f(\overline{A}) \subset \overline{f(A)}$.*

THEOREM 1.31. *Let $(X, \tau_X)$ and $(Y, \tau_Y)$ be a topological spaces and $f : X \to Y$ a function. If $\mathscr{C}$ is a subbase for the topology $\tau_Y$ and $f^{-1}(S)$ is an open set in $X$, for any $S \in \mathscr{C}$, then $f$ is a continuous function.*

PROOF. Let us first note that if the topology of the range space $Y$ is given by a basis $\mathscr{B}$, then to prove continuity of $f$ it suffices to show that the inverse image of every basis element is open. Indeed, an arbitrary open set $V \subset Y$ can be written as a union of basis elements

$$V = \bigcup_{i \in I} W_i.$$

Therefore

$$f^{-1}(V) = f^{-1}\Big(\bigcup_{i \in I} W_i\Big) = \bigcup_{i \in I} f^{-1}(W_i),$$

and so $f$ is continuous.

Since $\mathscr{C}$ is a subbase of $\tau_Y$ we have that and arbitrary base element $W \in \tau_Y$ can be written as a finite intersection $W = S_1 \cap \ldots \cap S_n$ of subbase elements. From the elementary properties of functions we have

$$f^{-1}(W) = f^{-1}(S_1) \cap \ldots \cap f^{-1}(S_n).$$

Thus proving that the inverse image of any base element of $\tau_Y$ is open set in $X$. From the observation at the beginning of this proof, we get that $f$ is continuous. ■

The following result is intuitive, easily proved and surprisingly important.

THEOREM 1.32. *If $X, Y$ and $Z$ are topological spaces and $f : X \to Y$ and $g : Y \to Z$ are continuous functions, then $g \circ f : X \to Z$ is continuous.*

DEFINITION 1.33. Let $f : X \to Y$ be a function and $A \subset X$. The restriction of $f$ to $A$ is a function from $A$ to $Y$, denoted by $f|_A$, given by $f|_A(a) = f(a)$ for all $a \in A$.

THEOREM 1.34. *Let $(X, \tau)$ be a topological space, $A \subset X$ be a generic subset endowed with the relative topology $\tau_A$ and $f : X \to Y$ is a continuous function. Then $f|_A : A \to Y$ is a continuous function.*

PROOF. If $V$ is a open set in $Y$, then $f|_A^{-1}(V) = A \cap f^{-1}(V)$, and the latter is open in the relative topology on $A$. ■

PROPOSITION 1.35. *If $(A, \tau_A)$ is a subspace of a topological space $(X, \tau)$, then the inclusion function $j : A \to X$ is a continuous function.*

LEMMA 1.36 (Local Formulation of Continuity). *Let $(X, \tau_X)$ and $(Y, \tau_Y)$ be general topological spaces. Suppose that $X = \cup_{\alpha \in \Gamma} A_\alpha$, where $A_\alpha \in \tau_X$, for all $\alpha \in \Gamma$. If $f : X \to Y$ is a function such that for all $\alpha \in \Gamma$ we have $f|_{A_\alpha}$ is continuous, then $f$ is continuous.*

PROOF. Suppose that $A_\alpha$ is an open set for each $\alpha \in \Gamma$. If $V$ is an open set in $(Y, \tau_Y)$, then follows from basic properties of inverse image that

$$f^{-1}(V) = \bigcup_{\alpha \in \Gamma} (f|_{A_\alpha})^{-1}(V).$$

From hypothesis, for each $\alpha \in \Gamma$, we have that $(f|_{A_\alpha})^{-1}(V) \in \tau_{A_\alpha}$, the relative topology. Since $A_\alpha$ is open in $X$ it follows that $\tau_{A_\alpha} \subset \tau$ and so $(f|_{A_\alpha})^{-1}(V) \in \tau$, thus finishing the proof. ■

When we write $f : X \to Y$, we specified the domain of $f$ as being $X$, but the image of $f$ is not determined, except that it must be contained in $Y$. The next theorem says essentially, that is not necessary to modify this procedure when dealing with continuous functions.

THEOREM 1.37. *Suppose that* $(Y, \tau_Y) \subset (Z, \tau_Z)$ *and* $f : X \to Y$ *is a function. Then* $f$ *is a continuous map from* $X$ *to* $Y$ *if and only if* $\tilde{f} : X \to Z$ *given by* $\tilde{f}(x) = f(x)$, *for all* $x \in X$, *is a continuous function.*

LEMMA 1.38 (The pasting Lemma). *Let* $(X, \tau_X)$ *and* $(Y, \tau_Y)$ *be topological spaces and* $A, B \subset X$ *closed sets in* $(X, \tau)$ *such that* $A \cup B = X$. *Let* $f : A \to Y$ *and* $g : B \to Y$ *be continuous functions. If* $f(x) = g(x)$ *for all* $x \in A \cap B$, *then* $f$ *and* $g$ *can be used to construct a new continuous function* $h : X \to Y$ *given by*

$$h(x) = \begin{cases} f(x), & \text{if } x \in A; \\ g(x), & \text{if } x \in B \setminus A. \end{cases}$$

PROOF. Let $C$ be a closed subset of $(Y, \tau_Y)$. From the elementary properties of inverse image we get that

$$h^{-1}(C) = f^{-1}(C) \cup g^{-1}(C)$$

Since $f$ is continuous, $f^{-1}(C)$ is a closed set in $A$ and, therefore closed in $X$, see Theorem 1.28 item *iii)*. Similarly, $g^{-1}(C)$ is closed in $B$ and therefore closed in $X$. Their union $h^{-1}(C)$ is thus closed in $X$. ■

Summarizing we have discussed the following techniques to construct continuous functions:

- Inclusion (Proposition 1.35).
- Compositions (Theorem 1.32).
- Restricting the domain (Theorem 1.34).
- Restricting or expanding the range (Theorem 1.37).
- Local formulation of Continuity (Lemma 1.36).
- The pasting lemma (Lemma 1.38).

## 1.6. Homeomorphisms

A function from a topological space $(X, \tau_X)$ to a topological space $(Y, \tau_Y)$ can make "information" disappear in two ways. The first in set-theoretical sense, which means $Y$ will have fewer (or at least, no more) points than $X$. The second is topological, meaning that the topological space $(Y, \tau_Y)$ will have fewer (or at least, no more) open sets than $(X, \tau_X)$, in the sense that each open $V \in Y$ is the image of an open set (for example, $U = f^{-1}(V)$ ) in $X$, but there may well be open sets $U \in \tau_X$ such that $f(U)$ is not open in $(Y, \tau_Y)$. The maps which preserve $X$ set-theoretically and topologically are called homeomorphisms.

DEFINITION 1.39 (Homeomorphism). Let $(X, \tau_X)$ and $(Y, \tau_Y)$ be topological spaces. A function $f : X \to Y$ is called a *homeomorphism* between $X$ and $Y$, if $f$ is injective, surjective (or onto) and continuous, and $f^{-1} : Y \to X$ is also continuous.

DEFINITION 1.40 (Embedding). Let $(X, \tau_X)$ and $(Y, \tau_Y)$ be topological spaces. A function $f : X \to Y$ is called a *embedding* of $X$ into $Y$, if $f$ is injective and continuous, and $f^{-1} : f(X) \to X$ is also continuous.

Note that $X$ is embedded into $Y$ by $f$ if and only if the function $f$ is a homeomorphism between $X$ and some subspace of $Y$.

Evidently, a continuous map $f : X \to Y$ is a homeomorphism if and only if there is a continuous map $g : Y \to X$ such that the compositions $g \circ f$ and $f \circ g$ are the identity maps on $X$ and $Y$, respectively.

The reader can easily verify the next theorem as a direct consequence of Theorem 1.30

THEOREM 1.41. *Let $(X, \tau_X)$ and $(Y, \tau_Y)$ be topological spaces and $f : X \to Y$ a bijection. The following are equivalent:*

*i) $f$ is a homeomorphism;*
*ii) if $U \subset X$, then $f(U)$ is open in $Y$ if and only if $U$ is open in $X$;*
*iii) if $A \subset X$, then $f(A)$ is closed in $Y$ if and only if $A$ is closed in $X$;*
*iv) if $A \subset X$, then $f(\overline{A}) = \overline{f(A)}$.*

Homeomorphic topological spaces are, for the purposes of a topologist, the same. That is, there is nothing about homeomorphic spaces $X$ and $Y$ having to do only with their respective topologies which we can use to distinguish them. If we denote "$X$ is homeomorphic to $Y$" by $X \sim Y$, then "$\sim$" defines an equivalence relation. Indeed,

i) $X \sim X$;
ii) if $X \sim Y$, then $Y \sim X$;
iii) if $X \sim Y$ and $Y \sim Z$, then $X \sim Z$.

In general, to prove that two spaces are homeomorphic, one constructs a homeomorphism. To establish that two spaces are not homeomorphic, one must find a topological property possessed by one and not the other. The definition of "topological property" makes it clear why this work. A *topological property* is a property of topological spaces which, if possessed by $X$, is possessed by all spaces homeomorphic to $X$.

CHAPTER 2

# Convex Analysis in Topological Vector Spaces

*"In these days the angel of topology and the devil of abstract algebra fight for the soul of every individual discipline of Mathematics."*

—attributed to H. Weil.

In this chapter we develop the basic theory of convex analysis on topological vector spaces. The exposition follow closely the classical reference [**3**].

## 2.1. Convex Sets and Their Separations

From now on $V$ denotes a vector space over $\mathbb{R}$. Given two distinct points $u, v \in V$ we define the line-segment joining then by the following set

$$[u, v] \equiv \{\lambda u + (1-\lambda)v : 0 \leq \lambda \leq 1\}$$

and the points $u$ and $v$ are called endpoints of $[u, v]$.

A set $\mathscr{A} \subset V$ is said to be convex if and only if for every pair of elements $u, v \in \mathscr{A}$ we have that $[u, v] \subset \mathscr{A}$. Note that the definition of convexity implies that for every finite subset $u_1, \ldots, u_n \in \mathscr{A}$ and for every family of positive real numbers $\lambda_1, \ldots, \lambda_n$ such that $\lambda_1 + \ldots + \lambda_n = 1$, we have

$$\sum_{i=1}^{n} \lambda_i u_i \in \mathscr{A}.$$

Of course, the whole space $V$ is a convex set, and we the empty set will be also considered as a convex convex set. It follows from the definition that an arbitrary intersection of convex sets is also a convex set. On the other hand, a union of convex sets may not be a convex set.

If $\mathscr{A}$ is an arbitrary subset of $V$, the intersection of all the convex sets of $V$ containing $\mathscr{A}$ is called the convex hull of $\mathscr{A}$ and is denoted by $\operatorname{co}\mathscr{A}$. Intuitively, this set is the smallest convex set containing $\mathscr{A}$. One can easily show that the convex hull of $\mathscr{A}$ can be characterized as follows

$$\operatorname{co}\mathscr{A} = \left\{ \sum_{i=1}^{n} \lambda_i u_i : n \in \mathbb{N}, \sum_{i=1}^{n} \lambda_i = 1, \lambda_i \geq 0, u_i \in \mathscr{A}, 1 \leq i \leq n \right\}. \tag{2.1}$$

Given a non-zero linear functional $\ell : V \to \mathbb{R}$ and a constant $\alpha \in \mathbb{R}$ the set

$$\mathscr{H} \equiv \{u \in V : \ell(u) = \alpha\}$$

is called an affine hyperplane. The sets

$$\{u \in V : \ell(u) < \alpha\} \qquad \{u \in V : \ell(u) > \alpha\}$$

are called open half-spaces bounded by $\mathscr{H}$. Similarly, we define closed half-spaces bounded by $\mathscr{H}$, replacing the strictly inequalities above by $\leq$ and $\geq$, respectively.

These half-spaces are convex sets and they are completely determined by $\mathscr{H}$ as a subset of $V$, which means that these sets do not depend on any particular choice of the pair $\ell$ and $\alpha$.

## 2.2. Separation of Convex Sets

We recall that a **topological vector space** (t.v.s.) is a vector space $V$ endowed with a topology $\tau$ for which the algebraic operations

$$\begin{aligned} V\times V \ni \quad & (u,v)\longmapsto u+v\in V,\\ \mathbb{R}\times V \ni \quad & (\lambda,u)\longmapsto \lambda u\in V \end{aligned}$$

are continuous. In a topological vector space $V$ the continuity of the sum operation implies that for any fixed $v\in V$ the neighborhood system at $v$ satisfies $\mathscr{U}_v = \mathscr{U}_0+\{v\}$. Indeed, we have for any fixed $v\in V$ that the mapping $T_v : V\to V$ given by $T_v(u)=u+v$ is a continuous function with a continuous inverse.

Example 2.1. Let $V\equiv C(\mathbb{R},\mathbb{R})$ the set of all continuous real-valued functions defined on $\mathbb{R}$. Let $\tau$ be the compact-open topology on $V$, which is generated by the sets of the form $\mathscr{B}(K,A)\equiv\{f\in V : f(K)\subset A\}$, where $K$ is a compact and $A$ is an open subset of the real line. Let $\Phi : V\times V\to V$ given by $\Phi(f,g)=f+g$. Endowing $V\times V$ with their natural product topology denoted by $\tau\times\tau$, the first part of our task consists in proving that $\Phi$ is a continuous function. As mentioned before it is enough to prove that the pre-image of any element of the subbase of $\tau$ is an open set in $\tau\times\tau$. That is, for any compact-open pair $K$ and $A$, we have to prove that $\Phi^{-1}(\mathscr{B}(K,A))\in\tau\times\tau$. In other words, if $(f,g)\in\Phi^{-1}(\mathscr{B}(K,A))$ is an arbitrary point, then we have to prove the existence of some neighborhood $\mathscr{N}$ of $(f,g)\in V\times V$ such that $\mathscr{N}\subset\Phi^{-1}(\mathscr{B}(K,A))$. To construct such neighborhood we first observe that for any $(f,g)\in\Phi^{-1}(\mathscr{B}(K,A))$ we have, from the definitions of $\Phi$ and $\mathscr{B}(K,A)$, that $(f+g)(K)\subset A$. Since $f$ and $g$ are continuous functions $(f+g)(K)$ is a non-empty compact subset of $\mathbb{R}$. The complement of $A$ is a closed subset and disjoint from $(f+g)(K)$. Therefore $\operatorname{dist}\big((f+g)(K),A^c\big)\equiv 8\varepsilon>0$. Consider the open sets

$$A_\varepsilon(x)\equiv f^{-1}\Big((-\varepsilon+f(x),f(x)+\varepsilon)\Big) \quad\text{and}\quad B_\varepsilon(x)\equiv g^{-1}\Big((-\varepsilon+g(x),g(x)+\varepsilon)\Big).$$

The unions over $x\in K$ of both collections $A_\varepsilon(x)$ and $B_\varepsilon(x)$ are open covers of $K$. Since $K$ is compact there are $x_1,\ldots,x_n\in K$ and $y_1,\ldots,y_m\in K$ so that

$$\bigcup_{i=1}^{n} A_\varepsilon(x_i)\supset K \quad\text{and}\quad \bigcup_{j=1}^{m} B_\varepsilon(y_j)\supset K.$$

Consider the following neighborhood $\mathscr{N}$ of $V\times V$ defined by

$$\mathscr{N}\equiv\bigcap_{i=1}^{n}\mathscr{B}\Big(\overline{A_\varepsilon(x_i)}\cap K, f(x_i)+(-2\varepsilon,2\varepsilon)\Big)\times\bigcap_{j=1}^{m}\mathscr{B}\Big(\overline{A_\varepsilon(y_j)}\cap K, g(y_j)+(-2\varepsilon,2\varepsilon)\Big).$$

From construction it is simple to check that the pair $(f,g)\in\mathscr{N}$. We claim that for any pair $(\varphi,\psi)\in\mathscr{N}$ we have $\Phi(\varphi,\psi)\in\mathscr{B}(K,A)$. Indeed, let $x\in K$ be an arbitrary point. Clearly, $x\in A_\varepsilon(x_i)\cap B_\varepsilon(y_j)$, for some $i\in\{1,\ldots,n\}$ and $j\in\{1,\ldots,m\}$.

Therefore $\varphi(x) = f(x_i)+r_1(x)$ and $\psi(x) = g(y_j)+r_2(x)$, where $|r_1(x)|, |r_2(x)| < 4\varepsilon$. And so we have

$$\begin{aligned}
\varphi(x) + \psi(x) &= f(x_i) + g(y_j) + r_1(x) + r_2(x) \\
&= f(x) + (f(x_i) - f(x)) + g(x) + (g(y_j) - g(x)) + r_1(x) + r_2(x) \\
&= f(x) + g(x) + R(x),
\end{aligned}$$

where $|R(x)| < 8\varepsilon$. Since $\operatorname{dist}\big((f+g)(K), A^c\big) = 8\varepsilon$ it follows that $\varphi(x) + \psi(x) = f(x) + g(x) + R(x) \in A$, thus proving the claim.

Now let $\Psi : \mathbb{R} \times V \to V$ be the multiplication by scalar function, that is, $\Psi(\lambda, f) = \lambda f$. Similarly to the sum operation, the continuity of $\Psi$ will be established if we prove that $\Psi^{-1}(\mathscr{B}(K, A))$ is an open set in $\mathbb{R}\times V$, for any choice of $K$ (compact) and $A$ (open).

Let $(\lambda, f) \in \Psi^{-1}(\mathscr{B}(K, A))$ and $\varepsilon > 0$ be such that

$$\varepsilon\Big(2\varepsilon + |\lambda| + \|f|_K\|_\infty\Big) < \operatorname{dist}(\lambda f(K), A^c).$$

Again, consider the sets $A_\varepsilon(x) \equiv f^{-1}\Big((-\varepsilon + f(x), f(x) + \varepsilon)\Big)$ and $x_1, \ldots, x_n$ be such that $\cup_{i=1}^n A_\varepsilon(x_i)$ is an open cover of $K$. Let $\mathscr{N}$ be the neighborhood on $\mathbb{R} \times V$ given by

$$\mathscr{N} \equiv (-\varepsilon + \lambda, \varepsilon + \lambda) \times \bigcap_{i=1}^{n} \mathscr{B}\Big(\overline{A_\varepsilon(x_i)} \cap K, f(x_i) + (-\varepsilon, \varepsilon)\Big).$$

From the definitions follow that $(\lambda, f) \in \mathscr{N}$. Let us prove that $\mathscr{N} \subset \Psi^{-1}(\mathscr{B}(K, A))$. Indeed, For any pair $(\beta, \varphi) \in \mathscr{N}$ we have

$$\begin{aligned}
|\beta\varphi(x) - \lambda f(x)| &\le |\beta||\varphi(x) - f(x)| + \|f|_K\|_\infty \cdot |\beta - \lambda| \\
&\le 2\varepsilon|\beta| + \varepsilon\|f|_K\|_\infty \\
&\le 2\varepsilon(|\beta - \lambda| + |\lambda|) + \varepsilon\|f|_K\|_\infty \\
&\le \varepsilon(2\varepsilon + |\lambda| + \|f|_K\|_\infty),
\end{aligned}$$

which implies that $\Psi(\beta, \varphi) \in \mathscr{B}(K, A)$ and finishes the proof that $(V, \tau)$ is a topological vector space.

Example 2.2. As in the previous example, we take $V = C(\mathbb{R}, \mathbb{R})$ and now the topology $\tau$ will be the topology of uniform convergence, that is, the one generated by the following neighborhoods system. For any fixed real function $f \in V$ and $\varepsilon > 0$ let $\mathscr{N}_f(\varepsilon) \equiv \{\varphi \in V : \|f - \varphi\|_\infty < \varepsilon\}$. Then $(V, \tau)$ is a not a topological vector space. To prove this it is enough to show that no finite intersections of elements of the subbase of $\mathbb{R} \times V$ is contained in $\Psi^{-1}(\mathscr{N}_{\lambda f}(\varepsilon))$ if $f$ is an unbounded function. Indeed, suppose that

$$(\lambda, f) \in \bigcap_{k=1}^{n} A_k \times \mathscr{N}_{f_k}(\varepsilon_k) \subset \Psi^{-1}(\mathscr{N}_{\lambda f}(\varepsilon)).$$

Then there is some $\delta > 0$ we have $((\lambda+\delta), f)$ is also a point in the above intersection. But this is an absurd, because $\Psi((\lambda + \delta), f) = \lambda f + \delta f$ does not belong to $\mathscr{N}_{\lambda f}(\varepsilon)$ since

$$\sup_{x\in\mathbb{R}} |\lambda f(x) + \delta f(x) - \lambda f(x)| = \sup_{x\in\mathbb{R}} |\delta f(x)| = +\infty.$$

A topological vector space $V$ is said to be a **locally convex space** (l.c.s.) if there exists a neighborhood base at the origin formed by convex neighborhoods. This is always the case of normed spaces: it is sufficient to observe that the set of all open balls centered at the origin is a family of convex sets and define a neighborhood base at the origin. Most of the t.v.s encountered in analysis are locally convex.

Let $V = (V, \tau)$ be a t.v.s. and $\mathscr{H}$ an affine hyperplane determined by the equation $\ell(u) = \alpha$, where $\ell : V \to \mathbb{R}$ is a non-zero linear functional and $\alpha$ is a real constant. It can be shown that the hyperplane $\mathscr{H}$ is topologically closed if and only if the function $\ell$ is continuous. Under this conditions the open(closed) half-spaces determined by $\mathscr{H}$ are topologically open(closed).

In a t.v.s., the closure of a convex set is convex and the interior of a convex set is also convex (possibly empty). More generally, if $\mathscr{A} \subset V$ is convex, $u \in \mathrm{Int}(\mathscr{A})$, and if $v \in \overline{\mathscr{A}}$, then $[u, v[\, \equiv \{\lambda u + (1-\lambda)v : 0 < \lambda \leq 1\} \subset \mathrm{Int}(\mathscr{A})$. From where we deduce that $\overline{\mathrm{Int}(\mathscr{A})} = \overline{\mathscr{A}}$, whenever $\mathrm{Int}(\mathscr{A}) \neq \emptyset$. This fact motivates the following definition: a point $u \in \mathscr{A}$ will be called **internal** if every line passing through $u$ intercept $\mathscr{A}$ in a line-segment $[v_1, v_2]$ such that $u \in ]v_1, v_2[\, \equiv \{\lambda v_1 + (1-\lambda)v_2 : 0 < \lambda < 1\}$. Therefore every interior point is internal, and by the above argument, if $\mathrm{Int}(\mathscr{A}) \neq \emptyset$ every internal point is interior.

If $\mathscr{A}$ is any subset of $V$, the intersection of all closed convex subsets containing $\mathscr{A}$ is the smallest closed convex subset containing $\mathscr{A}$. This set coincides with the closure of the convex hull of $\mathscr{A}$ and will be denoted by $\overline{\mathrm{co}}\,\mathscr{A}$. Note that, in general, $\overline{\mathrm{co}}\,\mathscr{A}$ is not the convex hull of the closure of $\mathscr{A}$.

An affine hyperplane is said to separate (strictly separate) two sets $\mathscr{A}$ and $\mathscr{B}$ if each of the closed (open) half-spaces bounded by $\mathscr{H}$ contains one of the them. This can be written analytically as follows. If $\ell(u) = \alpha$ is a linear equation determining $\mathscr{H}$, then we have

$$\ell(u) \leq \alpha, \quad \forall u \in \mathscr{A}, \qquad \ell(u) \geq \alpha, \quad \forall u \in \mathscr{B}.$$

To indicate strict separation, it is enough to replace "$\leq$" and "$\geq$" in the above inequalities by "$<$" and "$>$", respectively.

In the sequel we state the Hahn-Banach theorem in its geometric form, and some of its consequences for the separation of convex sets. We emphasize that these results are stated in the context of locally convex spaces which is the required generality to discuss some of the main results of this text. Their proofs can be found, for example, in the classical reference [**1**].

Theorem 2.3 (Hahn-Banach Theorem). *Let $V$ be a real topological vector space, $\mathscr{A}$ an open non-empty convex set, $\mathscr{M}$ a non-empty affine subspace which does not intersect $\mathscr{A}$. Then there exists a closed affine hyperplane $\mathscr{H}$ which contains $\mathscr{M}$ and does not intersect $\mathscr{A}$.*

Corollary 2.4. *Let $V$ be a real topological vector space, $\mathscr{A}$ an open non-empty convex set, $\mathscr{B}$ a non-empty convex set which does not intersect $\mathscr{A}$. Then there exists a closed affine hyperplane $\mathscr{H}$ which separates $\mathscr{A}$ and $\mathscr{B}$.*

An important application of Corollary 2.4 is the existence of supporting points under suitable assumptions, which is precisely stated in the sequel. First, let us introduce the concept of a supporting point. Let $\mathscr{A}$ be a subset of $V$. A point $u \in \mathscr{A}$ is called a **support point** of $\mathscr{A}$ if there exists a closed affine hyperplane $\mathscr{H}$ such that $u \in \mathscr{H}$ and $\mathscr{A}$ is contained in one of the closed half-spaces determined by $\mathscr{H}$. In this case $\mathscr{H}$ is called a support hyperplane. Note that $\mathscr{H}$ separates $\{u\}$ and $\mathscr{A}$. Before state the next corollary we should warn the reader that the existence of support point is a non-trivial question.

Corollary 2.5. *Let $V$ be a real topological vector space and $\mathscr{A} \subset V$ a* ***closed convex*** *subset such that* $\mathrm{Int}(\mathscr{A}) \neq \emptyset$. *Then every boundary point of $\mathscr{A}$ is a support point of $\mathscr{A}$.*

We shall remark that there is an example due to V. Klee, showing the existence of a nonempty closed convex set in a suitable infinite dimensional Fréchet space (which is a t.v.s.) having no support points whatsoever.

Corollary 2.6. *Let $V$ be a real* ***locally convex space****, $\mathscr{C}$ and $\mathscr{B}$ two non-empty disjoint convex sets with one compact and the other closed. Then there exists a closed affine hyperplane $\mathscr{H}$ which strictly separates $\mathscr{C}$ and $\mathscr{B}$.*

We also mention an interesting application of Corollary 2.6 which is the content of the following corollary

Corollary 2.7. *Let $V$ be a real* ***locally convex space****. Then every closed convex set is the intersection of the closed half-spaces which contain it.*

All these results have fundamental importance in analysis because they allow a convenient theory of duality. Thus if $(V, \tau)$ is a **Hausdorff locally convex space** the Hahn-Banach theorem allows us to ensure the existence of at least one non-zero continuous linear functional defined on $V$. Indeed, it is sufficient to consider two distinct points $u_1, u_2 \in V$ and strictly separating them by a **closed** affine hyperplane $\mathscr{H}$ (Corollary 2.6). Let $\ell : V \to \mathbb{R}$ and $\alpha$ be a functional and a constant determining $\mathscr{H}$. As we mentioned before $\ell$ has to be continuous since $\mathscr{H}$ is closed and $\ell(u_1) \neq \ell(u_2)$ which ensures that $\ell$ is a non-zero functional. The vector space $V^*$ of all continuous (with respect to $\tau$) linear functionals defined on $V$ is called the topological dual of $V$. Sometimes we call $V^*$ simply the dual of $V$. The elements of $V^*$ will be denoted by $u^*$ and $v^*$. If $u \in V$ and $u^* \in V^*$ we will sometimes write $\langle u, u^* \rangle$ instead of $u^*(u)$.

The next example shows that some topological spaces that are not Hausdorff locally convex space are not suitable spaces for working with duality.

Example 2.8. There is a distance $d_p$ on $L^p(X, \mathscr{F}, \mu)$, with $0 < p < 1$ for which $(L^p(X, \mathscr{F}, \mu), \tau)$ is a topological vector space, where $\tau$ is the topology induced by the distance $d_p$. But its topological dual is composed only by the null vector.

Definition 2.9. A *dual pair* is a pair of vector spaces $X$ and $Y$, over the same field $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$, where $Y$ is a subspace of the algebraic dual of $X$ satisfying: for each $x \in X \setminus \{0\}$ there is some $\ell \in Y$ with $\ell(x) \neq 0$.

The notation $\langle X, Y \rangle$ will only be used when $X$ and $Y$ form a dual pair, in the sense of the above definition. To emphasize this duality, we will write $\langle x, \ell \rangle$ rather than $\ell(x)$ for $x \in X$ and $\ell \in Y$.

Although the definition of dual pair looks asymmetric in $X$ and $Y$, the fact that the bilinear map $X \times Y \ni (x, \ell) \longmapsto \langle x, \ell \rangle$ is non-degenerated in the sense that for each $x \in X \setminus \{0\}$ there is some $\ell \in Y$ such that $\langle x, \ell \rangle \neq 0$, and for each $l \in Y \setminus \{0\}$ there is some $x \in X$ such that $\langle x, \ell \rangle \neq 0$, shows that $X$ and $Y$ play a symmetrical roles.

Definition 2.10. Given a dual pair $\langle X, Y\rangle$, the $Y$-weak topology on $X$, denoted by $\sigma(X,Y)$ is the weakest topology on $X$ in which the maps $X \ni x \longmapsto \langle x, \ell\rangle \in \mathbb{R}$ are continuous for all $\ell \in Y$.

Let $\langle X, Y\rangle$ be a dual pair and $\sigma(X,Y)$ the $Y$-weak topology on $X$. Then $(X, \sigma(X,Y))$ is a Hausdorff topological space. Indeed, let $x \neq y$ and $\ell \in Y$ such that $\langle x, \ell\rangle < \langle y, \ell\rangle$. Let $m \in \mathbb{R}$ be a constant such that $\langle x, \ell\rangle < m < \langle y, \ell\rangle$. Then $x \in \ell^{-1}\big((-\infty, m)\big)$ and $y \in \ell^{-1}\big((m, +\infty)\big)$. Since the last two sets are disjoint and open with respect to $\sigma(X,Y)$ the claim is proved.

We shall remark that when $X$ is a Banach space then $X$ and $X^*$ defines a dual pair $\langle X, X^*\rangle$ and $\sigma(X, X^*)$ is the usual weak topology on $X$. On the other hand, we can think about $x \in X$ as a linear functional defined on $X^*$ which acts on $\ell \in X^*$ as follows $\ell \longmapsto \langle x, \ell\rangle$. Again, we have another dual pair $\langle X^*, X\rangle$ and the topology $\sigma(X^*, X)$ on $X^*$ is called the weak-$*$-topology.

Now let us get back to our previous notation. If $(V, \tau)$ is a Hausdorff locally convex space then Corollary 2.6 implies that $V$ and $V^*$ defines a dual pair in the sense of Definition 2.9. Since any element $u^* \in V^*$ is, by definition, continuous with respect to $\tau$ and $\sigma(V, V^*)$ is the weakest topology in which all the elements of $V^*$ are continuous, we have that $\sigma(V, V^*) \subset \tau$. Therefore if $\mathscr{A} \subset V$ is a closed subset with respect to $\sigma(V, V^*)$, then $\mathscr{A}^c \in \sigma(V, V^*) \subset \tau$, which means $\mathscr{A}$ is a closed subset, with respect to the topology $\tau$. In other words, in a Hausdorff locally convex space, any weakly closed subset of $V$ is a closed subset with respect to $\tau$. The converse is false, in general. However, it follows from Corollary 2.7 that every **closed convex** set in $\tau$ is closed in $\sigma(V, V^*)$. Therefore in a Hausdorff locally convex space, any weakly closed convex set is a closed convex set, with respect to $\tau$.

When $\tau$ is induced by a norm $\|\cdot\|$ on $V$, we shall sometimes use these above observations to obtain the following result.

Lemma 2.11 (Mazur's Lemma). *Let $(V, \|\cdot\|)$ be a normed space and $(u_n)_{n\in\mathbb{N}}$ a sequence converging weakly to $\bar{u} \in V$. Then there is a sequence of convex combinations $(v_n)_{n\in\mathbb{N}}$ such that*

$$v_n = \sum_{k=n}^{N} \lambda_k^n u_k, \quad \textit{where } \sum_{k=n}^{N} \lambda_k^n = 1,\ \lambda_k^n \geq 0,\ n \leq k \leq N(n)$$

*which converges to $\bar{u}$ in norm.*

Proof. By hypothesis, for every $n \in \mathbb{N}$, we have that $\bar{u}$ belongs to the weak closure of $\cup_{k=n}^{\infty}\{u_k\}$ and consequently to the weak closure of $\operatorname{co} \cup_{k=n}^{\infty}\{u_k\}$. By taking the closure of $\operatorname{co} \cup_{k=n}^{\infty}\{u_k\}$, with respect to the topology induced by the norm, we get $\overline{\operatorname{co}} \cup_{k=n}^{\infty}\{u_k\}$. As mentioned before this is a (norm) closed convex subset of $V$. By Corollary 2.7 the set $\overline{\operatorname{co}} \cup_{k=n}^{\infty}\{u_k\}$ is the intersection of closed half-spaces which contain it and therefore it is a $\sigma(V, V^*)$ closed subset.

Since $\overline{\operatorname{co}} \cup_{k=n}^{\infty}\{u_k\}$ is a $\sigma(V, V^*)$ closed subset containing $\operatorname{co} \cup_{k=n}^{\infty}\{u_k\}$ follows from the very definition of closure that

$$\overline{\operatorname{co} \cup_{k=n}^{\infty}\{u_k\}}^{\sigma(V,V^*)} \subset \overline{\operatorname{co}} \cup_{k=n}^{\infty}\{u_k\}$$

which implies $\bar{u} \in \overline{\operatorname{co}} \cup_{k=n}^{\infty}\{u_k\}$, for any $n \in \mathbb{N}$.

For each $n \in \mathbb{N}$ there is a point $\bar{u}_n \in \overline{\text{co}} \cup_{k=n}^{\infty} \{u_k\}$ such that $\|\bar{u} - \bar{u}_n\| < 1/n$. Now it suffices to chose $v_n \in \text{co} \cup_{k=n}^{\infty} \{u_k\}$ such that $\|v_n - \bar{u}_n\| < 1/n$, and use the characterization (2.1) of the convex hull. ■

## 2.3. Convex Functions

The set of the extended real numbers will be denoted by $\overline{\mathbb{R}} \equiv \mathbb{R} \cup \{\pm\infty\}$. From now on will be convenient to consider functions defined on a real vector space $V$ or a convex subset of it, and taking values in $\overline{\mathbb{R}}$.

Definition 2.12 (Convex Functions). Let $\mathscr{A} \subset V$ be a convex subset of $V$. A function $F : \mathscr{A} \to \overline{\mathbb{R}}$ is said to be convex if for every $u, v \in \mathscr{A}$ we have

$$F(\lambda u + (1-\lambda)v) \leq \lambda F(u) + (1-\lambda)F(v), \quad \forall \lambda \in [0,1], \tag{2.2}$$

whenever the right-hand side is well-defined.

The last sentence in the above definition means that the inequality 2.2 must be valid unless $F(u) = -F(v) = \pm\infty$.

By induction, it can be shown that if $F$ is convex, then for every $u_1, \ldots, u_n \in \mathscr{A}$ and $\lambda_1, \ldots, \lambda_n \geq 0$ with $\lambda_1 + \ldots + \lambda_n = 1$, we have

$$F(\lambda_1 u_1 + \ldots + \lambda_n u_n) \leq \lambda_1 F(u_1) + \ldots + \lambda_n F(u_n),$$

whenever the right-hand side is well-defined.

Note that if $F : V \to \overline{\mathbb{R}}$ is convex, then the so-called **sections**

$$\{u \in V : F(u) \leq a\} \quad \text{and} \quad \{u \in V : F(u) < a\}$$

are convex sets for each $a \in \overline{\mathbb{R}}$. The converse is false, in general. For instance, if $F$ is convex and $\phi : \overline{\mathbb{R}} \to \overline{\mathbb{R}}$ is an increasing function then $\phi \circ F : V \to \overline{\mathbb{R}}$ will also have convex sections but will not be convex, in general.

For every (not necessarily convex) function $F : V \to \overline{\mathbb{R}}$, we call the (very special) section

$$\operatorname{dom} F \equiv \{u \in V : F(u) < +\infty\}$$

the **effective domain** of $F$. From the previous observations follows that the effective domain of a convex function is always a convex subset of $V$.

As mentioned before is very convenient in convex analysis allow the functions take infinite values. The first strong reason for that is the following extension result. Let $\mathscr{A} \subset V$ be a convex set and $F : \mathscr{A} \to \mathbb{R}$ (finite valued function!) a convex function. Then $F$ can be extended to a convex function $\widetilde{F} : V \to \overline{\mathbb{R}}$ given by

$$\widetilde{F}(u) = \begin{cases} F(u), & \text{if } u \in \mathscr{A}; \\ +\infty, & \text{otherwise.} \end{cases}$$

Because of this result is common in convex analysis to consider that the convex functions are everywhere defined.

There is a further advantage. For any $\mathscr{A}$ subset of $V$, we define the **indicator function** of $\mathscr{A}$ as the function $\chi_{\mathscr{A}} : V \to \overline{\mathbb{R}}$ given by

$$\chi_{\mathscr{A}}(u) = \begin{cases} 0, & \text{if } u \in \mathscr{A}; \\ +\infty, & \text{otherwise.} \end{cases}$$

Notice that $\mathscr{A}$ is a convex subset of $V$ if and only if $\chi_{\mathscr{A}}$ is a convex function. Thus the study of convex sets is naturally reduced to the study of convex functions. On the other hand, convex functions assuming the value $-\infty$ are very special. In the following sense. Suppose that $F : V \to \overline{\mathbb{R}}$ is a convex function and there is some point $u \in V$ such that $F(u) = -\infty$. Let $L(u)$ be a generic half-line starting from $u$, which is given by $L(u) = \{u + tv : t \geq 0\}$, for some $v \in V$. Then either, for all $t > 0$ we have $F(u + tv) = -\infty$ or there is some $t_0 \in (0, +\infty)$ such that

$$F(u+tv) = \begin{cases} -\infty & \text{if } t \in [0, t_0); \\ \text{any finite value,} & \text{if } t = t_0; \\ +\infty, & \text{if } t \in (t_0, \infty). \end{cases}$$

To distinguish these very special cases, we shall say that a convex function $F : V \to \overline{\mathbb{R}}$ is **proper** if it nowhere takes the value $-\infty$ and it is not identically constant equal to $+\infty$.

Definition 2.13. The epigraph of a function $F : V \to \overline{\mathbb{R}}$ is the set

$$\operatorname{epi} F \equiv \{(u, a) \in V \times \mathbb{R} : F(u) \leq a\}.$$

The epigraph of $F$ is the set of points of $V \times \mathbb{R}$ (second coordinate has to be finite) which lies above the graph of $F$. Note that the projection of $\operatorname{epi} F$ on $V$ is none other than $\operatorname{dom} F$. The concept of epigraph is of paramount importance in the theory presented in these lectures notes. One of the reasons is given by the next proposition.

Proposition 2.14. *A function $F : V \to \overline{\mathbb{R}}$ is convex if and only if its epigraph is a convex subset of $V$.*

Proof. Suppose that $F$ is a convex function and take $(u, a)$ and $(v, b) \in \operatorname{epi} F$. Then, necessarily, $F(u) \leq a < +\infty$ and $F(v) \leq b < +\infty$, and for all $\lambda \in [0, 1]$ we have from (2.2) that

$$F(\lambda u + (1-\lambda)v) \leq \lambda F(u) + (1-\lambda)F(v) \leq \lambda a + (1-\lambda)b,$$

which means precisely that $\lambda(u, a) + (1-\lambda)(v, b) \in \operatorname{epi} F$.

Conversely, let $\operatorname{epi} F$ be a convex set. Theb its projection on $V$ which is $\operatorname{dom} F$ is a convex subset of $V$ and therefore is sufficient to verify (2.2) over $\operatorname{dom} F$. Let us take $u, v \in \operatorname{dom} F$ and $a \geq F(u)$ and $b \geq F(v)$. By hypothesis for all $\lambda \in [0, 1]$, we have $\lambda(u, a) + (1-\lambda)(v, b) \in \operatorname{epi} F$ and therefore

$$F(\lambda u + (1-\lambda)v) \leq \lambda a + (1-\lambda)b.$$

If both $F(u)$ and $F(v)$ are finite, it is sufficient to take $a = F(u)$ and $b = F(v)$. If either $F(u)$ or $F(v)$ is equal to $-\infty$ it is sufficient to take the limit when $a$ or $b$ tends to $-\infty$ and the inequality (2.2) is proved. ■

## 2.4. Pointwise Supremum of Continuous Affine Functions

As before $(V, \tau)$ denotes a locally convex space. Recall that a continuous affine function on $V$ is a function of the form $V \ni u \longmapsto \ell(u) + \alpha$, where $\ell \in V^*$ and $\alpha \in \mathbb{R}$ is a constant.

The set of all continuous affine functions will be denoted by

$$\mathscr{J}(V) \equiv \{\Phi : V \to \mathbb{R} : \Phi(u) = \ell(u) + \alpha \text{ for some } \ell \in V^* \text{ and } \alpha \in \mathbb{R}\}.$$

Proposition 2.15. *If $\mathscr{N}$ is a family of continuous affine functionals contained in $\mathscr{J}(V)$ then the function $F : V \to \overline{\mathbb{R}}$ given by*

$$F(u) \equiv \sup_{\Phi \in \mathscr{N}} \Phi(u), \qquad \forall u \in V$$

*is a convex and lower semicontinuous function.*

We remark that when $\mathscr{N}$, in the above proposition, is the empty family then we set $F \equiv -\infty$.

Definition 2.16. The set of all function $F : V \to \overline{\mathbb{R}}$ which are pointwise supremum of a family of continuous affine functions is denoted by $\Gamma(V)$. We say that $F \in \Gamma_0(V)$ if $F \in \Gamma(V)$, $F \not\equiv +\infty$ and $F \not\equiv -\infty$.

It follows immediately from the above proposition that if $F \in \Gamma(V)$ then $F$ is convex and lower semicontinuous function. Conversely, we have the following result which is one of the most important results of this section.

Theorem 2.17. *The following properties are equivalent to each other:*

*(1)* $F \in \Gamma(V)$*;*
*(2)* $F : V \to \overline{\mathbb{R}}$ *is a convex lower semicontinuous function and if for some* $u \in V$ *we have* $F(u) = -\infty$ *then* $F \equiv -\infty$*.*

Proof. We first prove *(1)⟹(2)*. Given $F \in \Gamma(V)$ let $\mathscr{N} \subset \mathscr{J}(V)$ such that

$$F(u) = \sup_{\Phi \in \mathscr{N}} \Phi(u), \qquad \forall u \in V.$$

Since for any $\Phi \in \mathscr{J}(V)$ follows that $\Phi(u) \in \mathbb{R}$, for all $u \in V$. Therefore $F(u) = -\infty$ if and only if $\mathscr{N} = \emptyset$. Indeed,

$$-\infty < \Phi(u) \leq \sup_{\Phi \in \mathscr{N}} \Phi(u) \equiv F(u).$$

On the other hand, if $\mathscr{N} = \emptyset$ then

$$F(u) \equiv \sup_{\Phi \in \mathscr{N}} \Phi(u) = \sup_{\Phi \in \emptyset} \Phi(u) = -\infty, \qquad \forall u \in V.$$

Now we prove *(2)⟹(1)*. If $F \equiv +\infty$ then it easy to see that

$$F(u) = \sup_{\Phi \in \mathscr{J}(V)} \Phi(u) = +\infty, \qquad \forall u \in V.$$

On the other hand, if there exists $u \in V$ such that $F(u) = -\infty$ we have from *(2)* that $F \equiv -\infty$ and so

$$F(u) = \sup_{\Phi \in \emptyset} \Phi(u) = -\infty, \qquad \forall u \in V.$$

We proceed by assuming that $F$ is a lower semicontinuous convex function satisfying $F(V) \subset (-\infty, +\infty]$ and $F \not\equiv +\infty$. In this case we can decompose $V$ as follows

$$V = F^{-1}(-\infty, +\infty) \cup F^{-1}(\{+\infty\}),$$

and $F^{-1}(-\infty, +\infty)$ is not empty. Let $u \in F^{-1}(-\infty, +\infty)$ and $a \in \mathbb{R}$ such that $a < F(u)$. Since $F$ is a semicontinuous convex function follows that its epigraph is a closed (with respect to $\tau$) convex subset of $V$. Since $(u, a) \notin \operatorname{epi} F$ we can apply Corollary 2.6 to conclude the existence of a closed hyperplane $\mathscr{H} \subset V \times \mathbb{R}$ strictly separating $(u, a)$ and $\operatorname{epi} F$, see Figure 1. Note that

$$\mathscr{H} \equiv \{(v, b) \in V \times \mathbb{R} : \ell(v) + \alpha b = \beta\},$$

where $\ell \in V^*$ and the constants $\alpha, \beta \in \mathbb{R}$.

The strict separation guarantees that

$$\ell(u) + \alpha a < \beta < \ell(v) + b\alpha, \qquad \forall (v, b) \in \operatorname{epi} F. \tag{2.3}$$

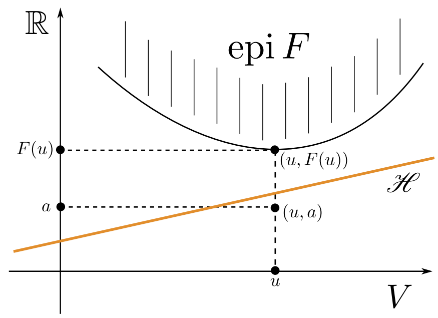


FIGURE 1. Graph of $F$ and its epigraph

Since $F(u) < +\infty$ follows that $(u, F(u)) \in \operatorname{epi} F$ and we get from the last inequality $\ell(u)+\alpha a < \ell(u)+\alpha F(u)$. This implies $\alpha(a-F(u)) < 0$, which in turn implies $\alpha > 0$, because $a < F(u)$. By using that $\alpha > 0$ we obtain from $\ell(u) + \alpha a < \ell(u) + \alpha F(u)$ the following inequality

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

Consider the continuous affine functional $\Phi : V \to \mathbb{R}$ given by

$$\Phi(v) = \frac{\beta}{\alpha} - \frac{1}{\alpha}\ell(v).$$

We claim that $\Phi(v) < F(v)$ for all $v \in V$. The claim is obvious if $v \in V$ is such that $F(v) = +\infty$. Let us now focus on the points $v$ such that $F(v) < +\infty$ (recall that $F$ does not take the value $-\infty$). Since $(v, \Phi(v)) \in \mathscr{H}$ and $(v, F(v)) \in \operatorname{epi} F$ follows from (2.3) that $\ell(v) + \alpha\Phi(v) = \beta < \ell(v) + \alpha F(v)$ by using again that $\alpha > 0$ we conclude from this last inequality that $\Phi(v) < F(v)$ thus proving the claim.

From the above argument we see that given $u \in F^{-1}(-\infty, +\infty)$ and $a < F(u)$ there is at least one continuous affine functional $\Phi \equiv \Phi_{u,a}$ such that $\Phi_{u,a}(v) < F(v)$ for all $v \in V$ and moreover $|F(u) - \Phi_{u,a}(u)| < F(u) - a$. Therefore if we consider the family $\mathscr{N}(u) \equiv \{\Phi_{u,a} : a \in \mathbb{R} \text{ and } a < F(u)\}$ then we have

$$F(u) = \sup_{\Phi \in \mathscr{N}(u)} \Phi(u) \qquad \text{and} \qquad \forall v \in V \quad \Phi(v) < F(v) \quad \text{for any } \Phi \in \mathscr{N}(u).$$

Now consider the family of continuous affine functionals

$$\mathscr{N} \equiv \bigcup_{u \in F^{-1}(-\infty, +\infty)} \mathscr{N}(u).$$

The extended real valued function on $V$, given by the pointwise supremum

$$v \longmapsto \sup_{\Phi \in \mathscr{N}} \Phi(v)$$

has the following properties. For any $\Phi \in \mathscr{N}$, we have $\Phi(v) \leq F(v)$, for all $v \in V$ and furthermore

$$\sup_{\Phi \in \mathscr{N}} \Phi(u) = F(u), \; \forall u \in F^{-1}(-\infty, +\infty). \tag{2.4}$$

To finish the proof we need to extend our analysis to the points $u \in F^{-1}(\{+\infty\})$. As the reader will see the remainder of the argument is similar to one presented above.

In this part of the argument we fix a pair $(u, a)$, where $a \in \mathbb{R}$ is arbitrarily chosen. The idea is to use again the Corollary 2.6 to construct a continuous affine functional $\Phi_{u,a}$ so that the following conditions hold

$$\Phi_{u,a}(v) \leq F(v),\ \forall v \in V, \qquad \text{and} \qquad a < \Phi_{u,a}(u). \tag{2.5}$$

Given a point $(u, a) \in V \times \mathbb{R}$ let $\mathscr{H}$ be a closed hyperplane in $V \times \mathbb{R}$ strictly separating $(u, a)$ from the closed set epi $F$. The equation defining $\mathscr{H}$ is of the form

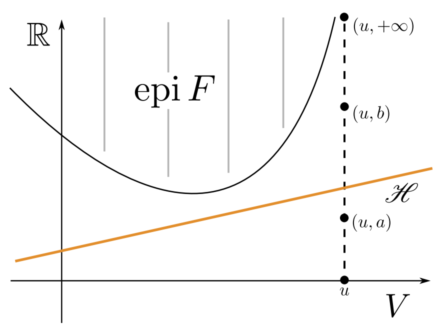


FIGURE 2. The case $F(u) = +\infty$.

$$\ell(v) + \alpha b = \beta, \tag{2.6}$$

where $\ell \in V^*$ and $\alpha, \beta \in \mathbb{R}$. Let us first consider the case $\alpha \neq 0$. Note that there is no loss of generality in assuming $\alpha > 0$. Indeed, if this is not the case just multiply both sides of the equation, defining $\mathscr{H}$, by minus one. Since $\mathscr{H}$ separates $(u, a)$ from epi $F$ it follows that

$$\ell(u) + \alpha a < \beta < \ell(v) + \alpha F(v), \qquad \forall (v, F(v)) \in \operatorname{epi} F. \tag{2.7}$$

The second inequality above implies

$$\Phi_{u,a}(v) \equiv \frac{\beta}{\alpha} - \frac{1}{\alpha}\ell(v) < F(v), \qquad \forall v \in V,$$

while the first one gives

$$a < \frac{\beta}{\alpha} - \frac{1}{\alpha}\ell(u) \equiv \Phi_{u,a}(u).$$

On the other hand, if $\alpha = 0$ then the equation of the hyperplane $\mathscr{H}$ is given simply by $\ell(v) = \beta$. Therefore

$$\ell(u) < \beta < \ell(v), \qquad \forall v \in V \text{ such that } (v, F(v)) \in \operatorname{epi} F. \tag{2.8}$$

Take any continuous affine functional $\Phi \in \mathscr{N}$, constructed as above. Recall that $\Phi(v) \leq F(v)$ for all $v \in V$. Let $\gamma \in \mathbb{R}$ and $m \in V^*$ be so that $\Phi(v) = \gamma - m(v)$. Given $k \geq 0$, we define a continuous affine functional

$$\Phi_{u,k}(v) \equiv \gamma - m(v) + k(\beta - \ell(v))$$

From (2.8) it follows that $k(\beta - \ell(v)) < 0$, for all $v \in \operatorname{dom} F$, and $\gamma - m(v) < F(v)$ for all $v \in V$. Therefore we get from these two observations that

$$\Phi_{u,k}(v) \equiv \gamma - m(v) + k(\beta - \ell(v)) < F(v), \qquad \forall v \in V.$$

On the other hand, the first inequality in (2.8) implies $\beta - \ell(u) > 0$ and so it is possible to find at least one $k \equiv k(a) > 0$ so that $a < \gamma - m(u) + k(\beta - \ell(u))$. Let $\Phi_{u,a} \equiv \Phi_{u,k(a)}$. Then $a < \gamma - m(u) + k(\beta - \ell(u)) = \Phi_{u,a}(u)$, and $\Phi_{u,a}(v) \leq F(v)$ for all $v \in V$.

From the previous arguments, for any choice of $u \in F^{-1}(\{+\infty\})$ and $a \in \mathbb{R}$, we have that

$$\mathscr{M}(u,a) \equiv \{\Phi_{u,a} \in \mathscr{J}(V) : \Phi_{u,a}(v) \leq F(v)\ \forall v \in V \text{ and } a \leq \Phi_{u,a}(u)\}$$

is not an empty family. Define

$$\mathscr{M} \equiv \bigcup_{\substack{u \in F^{-1}(\{+\infty\}) \\ a \in \mathbb{R}}} \mathscr{M}(u,a).$$

Then for any $u \in F^{-1}(\{+\infty\})$ we have

$$\sup_{\Phi \in \mathscr{M}} \Phi(u) \geq \sup_{a \in \mathbb{R}} \Phi_{u,a}(u) \geq \sup_{a \in \mathbb{R}} a = +\infty. \tag{2.9}$$

Finally, from the above constructions we have

$$F(u) = \sup_{\Phi \in \mathscr{N} \cup \mathscr{M}} \Phi(u), \qquad \forall u \in V. \tag{2.10}$$

In fact, for any $\Phi \in \mathscr{N} \cup \mathscr{M}$, we have $\Phi(u) \leq F(u)$ for all $u \in V$. From (2.4) we have for any $u \in F^{-1}(-\infty, +\infty)$

$$F(u) = \sup_{\Phi \in \mathscr{N}} \Phi(u) \leq \sup_{\Phi \in \mathscr{N} \cup \mathscr{M}} \Phi(u) \leq F(u).$$

On the other hand, if $u \in F^{-1}(\{+\infty\})$ then follows from (2.9)

$$F(u) = +\infty = \sup_{\Phi \in \mathscr{M}} \Phi(u) \leq \sup_{\Phi \in \mathscr{N} \cup \mathscr{M}} \Phi(u)$$

thus proving (2.10) and finishing the proof of the theorem. ■

## 2.5. Γ-regularization

PROPOSITION 2.18. *Let $F, G : V \to \overline{\mathbb{R}}$ functions. Then the following statements are equivalents:*

*(1) for all $u \in V$ we have*

$$G(u) = \sup_{\substack{\Phi \in \mathscr{J}(V) \\ \Phi \leq F}} \Phi(u);$$

*(2) $G$ is the greatest minorant of $F$ in $\Gamma(V)$.*

PROOF. We first prove that *(1)$\Longrightarrow$(2)*. Let $\widetilde{G} \in \Gamma(V)$ a minorant for $F$. We have to prove that $\widetilde{G} \leq G \leq F$. Indeed, since $\widetilde{G} \in \Gamma(V)$ we have from Theorem 2.17

$$\widetilde{G}(u) = \sup_{\Phi \in \mathscr{N}} \Phi(u),$$

where $\mathscr{N} \subset \mathscr{J}(V)$ is a family of continuous affine functions, satisfying $\Phi(v) \leq F(v)$, for all $\Phi \in \mathscr{N}$ and for all $v \in V$. Therefore $\mathscr{N} \subset \{\Phi \in \mathscr{J}(V) : \Phi \leq F\}$ and so the claim follows from the elementary properties of the sumpremum.

Conversely *(2)⟹(1)*. By Theorem 2.17 we have that any minorant $\widetilde{G}$ of $F$ in $\Gamma(V)$ satisfies

$$\widetilde{G}(u) \equiv \sup_{\Phi \in \mathscr{N}} \Phi(u), \qquad \forall u \in V,$$

and for all $\Phi \in \mathscr{N}$ we have $\Phi \leq F$. Therefore

$$\widetilde{G}(u) \equiv \sup_{\Phi \in \mathscr{N}} \Phi(u) \leq \sup_{\substack{\Phi \in \mathscr{J}(V) \\ \Phi \leq F}} \Phi(u) = G(u), \qquad \forall u \in V,$$

■

Definition 2.19. Let $F : V \to \overline{\mathbb{R}}$ be a function. The Γ-regularization of $F$ is the function $G : V \to \overline{\mathbb{R}}$ given by

$$G(u) = \sup_{\substack{\Phi \in \mathscr{J}(V) \\ \Phi \leq F}} \Phi(u), \qquad \forall u \in V.$$

The previous proposition allow us to interpret the Γ-regularization of a function $F$ on $V$ as being the greatest minorant of $F$ in $\Gamma(V)$.

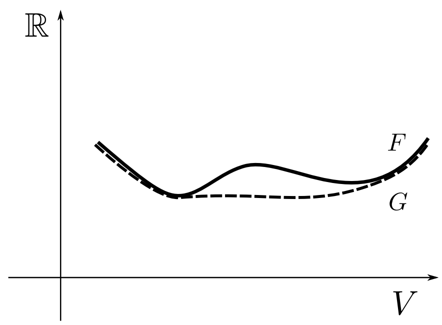


Figure 3. The function $G$ is the Γ-regularization of $F$

We remark that if $F \in \Gamma(V)$ then $F$ is, itself, its Γ-regularization. In general, the epigraph of the Γ-regularization of $F$ is the closure of the convex hull of the epigraph of $F$. A more precise statement is provided by the following proposition.

Proposition 2.20. *Let $F : V \to \overline{\mathbb{R}}$ be a function and $G : V \to \overline{\mathbb{R}}$ its Γ-regularization. If there exists some $\Psi \in \mathscr{J}(V)$ such that $\Psi \leq F$ everywhere, then* $\operatorname{epi} G = \overline{\operatorname{co}} \operatorname{epi} F$.

Proof. Since we are assuming the existence of a function $\Psi \in \mathscr{J}(V)$ such that $\Psi \leq F$ everywhere, we have that $\operatorname{epi} F \subset \operatorname{epi} \Psi$. Since $\Psi$ is a convex lower semicontinuous functions its epigraph is closed convex subset of $V$. Therefore $\overline{\operatorname{co}} \operatorname{epi} F \subset \operatorname{epi} \Psi$. This observation ensures that the expression below provide us a function in $\Gamma(V)$, which does not take the value $-\infty$,

$$\widetilde{G}(u) = \sup_{\substack{\Phi \in \mathscr{J}(V) \\ \overline{\operatorname{co}} \operatorname{epi} F \subset \operatorname{epi} \Phi}} \Phi(u), \qquad \forall u \in V.$$

We claim that $\operatorname{epi} \widetilde{G} = \overline{\operatorname{co}} \operatorname{epi} F$. In fact, we first observe that

$$\overline{\operatorname{co}} \operatorname{epi} F \subset \bigcap_{\substack{\Phi \in \mathscr{J}(V) \\ \overline{\operatorname{co}} \operatorname{epi} F \subset \operatorname{epi} \Phi}} \operatorname{epi} \Phi = \operatorname{epi} \widetilde{G}.$$

Suppose by absurd that there is some point $(u, \widetilde{G}(u)) \in \operatorname{epi} \widetilde{G}$ which is not in $\overline{\mathrm{co}} \operatorname{epi} F$. Then By Corollary 2.6 there is some closed hyperplane $\mathscr{H} \subset V \times \mathbb{R}$ strictly separating $(u, \widetilde{G}(u))$ from $\overline{\mathrm{co}} \operatorname{epi} F$. By using similar reasoning as in the first part of the proof of Theorem 2.17 one can see that this hyperplane is the graph of some continuous affine functional $\Phi_0 : V \to \mathbb{R}$, and the separation implies that $\overline{\mathrm{co}} \operatorname{epi} F \subset \operatorname{epi} \Phi_0$ and $\tilde{G}(u) < \Phi_0(u)$, which is an absurd by the definition of $\widetilde{G}$. Therefore

$$\overline{\mathrm{co}} \operatorname{epi} F = \operatorname{epi} \widetilde{G}.$$

By the definition of $\Gamma$-regularization we have that $\widetilde{G} \leq G \leq F$. And so

$$\operatorname{epi} F \subset \operatorname{epi} G \subset \operatorname{epi} \widetilde{G} = \overline{\mathrm{co}} \operatorname{epi} F.$$

Since $\operatorname{epi} G$ is a closed convex subset of $V$ follow from the above relations that $\overline{\mathrm{co}} \operatorname{epi} F \subset \operatorname{epi} G$ and therefore $\operatorname{epi} G = \overline{\mathrm{co}} \operatorname{epi} F$ thus proving the proposition. ■

By using the above proposition we can see immediately that for any $\mathscr{A} \subset V$ the $\Gamma$-regularization of $\chi_A$ is given by $\chi_{\overline{\mathrm{co}}\, \mathscr{A}}$. The next proposition explain what is the order relation between the $\Gamma$-regularization and the lower semicontinuous regularization, introduced before.

Proposition 2.21. *Let $F : V \to \overline{\mathbb{R}}$ be an arbitrary function, $\bar{F} : V \to \overline{\mathbb{R}}$ its lower semincontinuous regularization and $G : V \to \overline{\mathbb{R}}$ its $\Gamma$-regularization. Then*

*(1) $G \leq \bar{F} \leq F$;*
*(2) if $F$ is convex and admits a continuous affine minorant then $\bar{F} = G$.*

Proof. We first prove *(1)*. By definition $\bar{F}$ is the largest lower semicontinuous minorant of $F$ and since $G$ is l.s.c and bounded by $F$ the statement is proved.

Now prove *(2)*. From the hypothesis it follows that $F$ does not take the value $-\infty$ and therefore follows from Theorem 2.17 that

$$F(u) = \sup_{\substack{\Phi \in \mathscr{I}(V) \\ \Phi \leq F}} \Phi(u), \qquad \forall u \in V.$$

The expression on the right hand side above is exactly $G(u)$, and follows from item *(1)* that $G(u) = \bar{F}(u) = F(u)$ for all $u \in V$, finishing the proof of the proposition. ■

## 2.6. The Legendre-Fenchel Transform

As in the previous section we will work with a topological vector space $(V, \tau)$ which is locally convex and Hausdorff. The space $V$ is placed in duality with its topological dual $V^*$ by the cannonical bilinear form denoted by $\langle \cdot, \cdot \rangle$. We will also consider sometimes that $V$ is endowed with its weak topology $\sigma(V, V^*)$ and analogously $V^*$ will be endowed with the weak-$*$-topology $\sigma(V^*, V)$, which render these spaces locally convex and Hausdorff.

Definition 2.22. The Legendre-Fenchel transform of a function $F : V \to \overline{\mathbb{R}}$ is defined as the function $F^* : V^* \to \overline{\mathbb{R}}$ given by the following expression

$$F^*(u^*) \equiv \sup_{u \in V} \langle u, u^* \rangle - F(u).$$

Note that for a fixed $u \in V$ the mapping $u^* \longmapsto \langle u, u^*\rangle - F(u)$ is a continuous affine functional on $(V^*, \sigma(V^*, V))$ and therefore $F^*$ is nothing more than a pointwise supremum of continuous affine functionals on $V^*$ and so $F^* \in \Gamma(V^*)$. In particular, $F^*$ is always a convex and lower semicontinuous function, with respect to the weak-$*$-topology.

We observe that if $F$ is the constant function $+\infty$ then $F^*$ is the constant function $-\infty$. On the other hand, if $F$ has an affine minorant $\Phi : V \to \mathbb{R}$ of the form $\Phi(u) = \langle u, u^*\rangle - \alpha$, where $u^* \in V^*$ and $\alpha \in \mathbb{R}$, then we have the following inequality $\langle u, u^*\rangle - \alpha \leq F(u)$ for all $u \in V$ and consequently

$$-\infty < F^*(u^*) \equiv \sup_{u \in V} \langle u, u^*\rangle - F(u) \leq \alpha.$$

More generally, the Legendre-Transform sastify the following properties:

- $F^*(0) = -\inf\limits_{u \in V} F(u)$;
- if $F \leq G$, then $G^* \leq F^*$;
- $(\inf\limits_{i \in I} F_i)^* = \sup\limits_{i \in I} F_i^*$;
- $(\sup\limits_{i \in I} F_i)^* \leq \inf\limits_{i \in I} F_i^*$,

where $(F_i)_{i \in I}$ is an arbitrary family of extended real valued functions on $V$. Moreover,

- $(\lambda F)^*(u^*) = \lambda F^*(u^*/\lambda)$, for all $\lambda > 0$;
- $(F + \alpha)^* = F^* - \alpha$, for all $\alpha \in \mathbb{R}$.

If for a fixed $a \in V$ we define $F_a : V \to \overline{\mathbb{R}}$ by $F_a(u) = F(u - a)$, then we have

- $(F_a)^*(u^*) = F^*(u^*) + \langle a, u^*\rangle$.

We can use the symmetric role played by $V$ and $V^*$ on the dual pair $\langle V, V^*\rangle$ and "repeat" the process in order to define $F^{**}$, a kind of a double Legendre-Transform of $F$. More precise, for each $u \in V$ we define

$$F^{**}(u) \equiv \sup_{u^* \in V^*} \langle u, u^*\rangle - F^*(u^*).$$

By arguing as above we can see immediately that $F^{**} \in \Gamma(V)$. Of course, It is natural to ask whether some relation between $F$ and $F^{**}$ do exists. The answer is the following.

Proposition 2.23. *Let $F : V \to \overline{\mathbb{R}}$ be an arbitrary function and $F^{**}$ as defined above. The function $F^{**}$ is the $\Gamma$-regularization of $F$. In particular, if $F \in \Gamma(V)$, then $F^{**} = F$.*

Proof. Let $G$ be the $\Gamma$-regularization of $F$, that is

$$G(u) = \sup_{\substack{\Phi \in \mathscr{I}(V) \\ \Phi \leq F}} \Phi(u), \qquad \forall u \in V.$$

For each fixed $u^* \in V^*$ defines the continuous affine functional $\Phi_{u^*} : V \to \mathbb{R}$ given by $\Phi_{u^*}(u) \equiv \langle u, u^* \rangle - F^*(u^*)$. Note that $\Phi_{u^*}$ is everywhere less than $F$. Indeed, for any $u \in V$ we have

$$\Phi_{u^*}(u) = \langle u, u^* \rangle - \sup_{v \in V} \big\{ \langle v, u^* \rangle - F(v) \big\} \leq \langle u, u^* \rangle - \langle u, u^* \rangle + F(u) = F(u).$$

Therefore the collection $\mathscr{N} \equiv \{\Phi_{u^*} : u^* \in V^*\}$ is a collection of continuous affine functionals such that each one of its elements is dominated by $F$. We observe that the elements in $\mathscr{N}$ are maximal in the following sense. If $\Phi(u) = \langle u, u^* \rangle - \alpha \leq F(u)$ for all $u \in V$, then $\Phi(u) \leq \Phi_{u^*}(u)$. Indeed, if for all $u \in V$ we have $\langle u, u^* \rangle - \alpha \leq F(u)$, then $\langle u, u^* \rangle - F(u) \leq \alpha$ and so $F^*(u^*) \leq \alpha$ which implies $\Phi(u) = \langle u, u^* \rangle - \alpha \leq \langle u, u^* \rangle - F^*(u^*) = \Phi_{u^*}(u)$ thus proving the claim. Therefore we have

$$F^{**}(u) = \sup_{u^* \in V^*} \langle u, u^* \rangle - F^*(u^*) = \sup_{u^* \in V^*} \Phi_{u^*}(u) = \sup_{\substack{\Phi \in \mathscr{J}(V) \\ \Phi \leq F}} \Phi(u) = G(u), \quad \forall u \in V.$$

In particular, if $F \in \Gamma(V)$, then $F = G = F^{**}$. ■

Corollary 2.24. *For any function $F : V \to \overline{\mathbb{R}}$, we have $F^* = (F^{**})^*$.*

Proof. Since $F^{**}$ is the $\Gamma$-regularization of $F$, we have $F^{**} \leq F$, and so $F^* \leq (F^{**})^*$. On the other hand, for any fixed $u^* \in V^*$, follows from the definition of the Legendre-Fenchel transform that

$$\langle u, u^* \rangle - F^{**}(u) \leq F^*(u^*), \qquad \forall u \in V$$

whence

$$(F^{**})^*(u^*) = \sup_{u \in V} \big\{ \langle u, u^* \rangle - F^{**}(u) \big\} \leq F^*(u^*)$$

which proves that $(F^{**})^* \leq F^*$ and finish the proof of the corollary. ■

We have seen that $F \in \Gamma(V)$ if and only if $F^{**} = F$. Therefore the map taking $F \in \Gamma(V)$ to its Legendre-Fenchel transform defines a bijective mapping. Indeed, if $F_1, F_2 \in \Gamma(V)$ are such that $F_1^* = F_2^*$, then we have from Proposition 2.23 that $F_1 = F_1^{**} = F_2^{**} = F_2$, proving that the mapping is injective. To see that this map is also onto, take $G \in \Gamma(V^*)$. Then $G^* \in \Gamma(V)$, and another application of Proposition 2.23 shows that $(G^*)^* = G^{**} = G$ and so the mapping is onto.

This observation motivates the following definition.

Definition 2.25. The Legendre-Fenchel transform establishes a bijection between $\Gamma(V)$ and $\Gamma(V^*)$. The functions $F \in \Gamma(V)$ and $G \in \Gamma(V^*)$ are said to be **in duality** if they are corresponding, with respect to the bijection mentioned above, that is, $F = G^*$ and $G = F^*$.

The constant functions $\pm\infty$ on $V$ and $\pm\infty$ on $V^*$ are in duality. Thus $F \in \Gamma_0(V)$ if and only if $F^* \in \Gamma_0(V^*)$. In other words, the Legendre-Fenchel transform establishes a one-to-one correspondence between $\Gamma_0(V)$ and $\Gamma_0(V^*)$.

## 2.7. Subdifferentiability

As in the previous section we still work with a topological vector space $(V, \tau)$ which is locally convex and Hausdorff.

Before proceed let us introduce an important terminology for this section. If $\Phi$ is a continuous affine functional on $V$ given by $\Phi(u) = \langle u, u^* \rangle + \beta$, for all $u \in V$, then the continuous linear functional $u^*$ is called **slope** of $\Phi$.

DEFINITION 2.26. Given a function $F : V \to \overline{\mathbb{R}}$ we say that a continuous affine functional $\Phi \in \mathscr{J}(V)$ satisfying $\Phi \leq F$ is exact at a point $u \in V$ if $\Phi(u) = F(u)$.

PROPOSITION 2.27. *Let $F : V \to \overline{\mathbb{R}}$ be a function and suppose that $\Phi \in \mathscr{J}(V)$ satisfies $\Phi \leq F$ and is exact at a point $u \in V$. Then there is some $u^* \in V^*$ such that $\Phi(v) = \langle v - u, u^* \rangle + F(u)$, for all $v \in V$.*

PROOF. Since $\Phi \in \mathscr{J}(V)$ there are some $u^* \in V^*$ and a constant $\alpha \in \mathbb{R}$ such that $\Phi(v) = \langle v, u^* \rangle + \alpha$, for all $v \in V$. Since $\Phi$ is exact at $u \in V$, we get that $F(u) = \Phi(u) = \langle u, u^* \rangle + \alpha$. And so $\alpha = F(u) - \langle u, u^* \rangle$ which immediately implies $\Phi(v) = \langle v - u, u^* \rangle + F(u)$, for all $v \in V$. ■

Let $F : V \to \overline{\mathbb{R}}$ be an arbitrary function. Suppose that $F$ is bounded everywhere from below by a continuous affine functional $\Phi \in \mathscr{J}(V)$, that is, $\Phi \leq F$, and $\Phi(v) = \langle v, u^* \rangle + \alpha$, for some $u^* \in V^*$ and $\alpha \in \mathbb{R}$. We claim that if $\Psi \in \mathscr{J}(V)$ is any other continuous affine functional satisfying both $\Psi \leq F$ and $\Psi(v) = \langle v, u^* \rangle + \beta$ ($\Psi$ has the same slope as $\Phi$) for some $\beta \in \mathbb{R}$, then its constant term $\beta \leq -F^*(u^*)$.

Clearly, to prove the claim it is enough to show that

$$\beta \leq \inf_{v \in V} F(v) - \langle v, u^* \rangle.$$

This inequality follows from the following argument. Suppose that it is possible to have $\beta > \inf_{v \in V} F(v) - \langle v, u^* \rangle$. By taking $\varepsilon \equiv (1/2)(\beta - \inf_{v \in V} F(v) - \langle v, u^* \rangle) > 0$ and using the definition of infimum, we can find an element $u \in V$ such that

$$\begin{aligned} F(u) - \langle u, u^* \rangle &< \inf_{v \in V} F(v) - \langle v, u^* \rangle) + \varepsilon \\ &< \inf_{v \in V} F(v) - \langle v, u^* \rangle) + \frac{\beta}{2} - \frac{1}{2}(\inf_{v \in V} F(v) - \langle v, u^* \rangle) \\ &= \frac{\beta}{2} + \frac{1}{2}(\inf_{v \in V} F(v) - \langle v, u^* \rangle)) \\ &< \frac{\beta}{2} + \frac{\beta}{2} = \beta \end{aligned}$$

which contradicts the assumption $\Psi \leq F$. Therefore we have

$$\beta \leq \inf_{v \in V} F(v) - \langle v, u^* \rangle = -\sup_{v \in V} \langle v, u^* \rangle - F(v) = -F^*(u^*).$$

The above observation allow us to conclude the following. Among all continuous affine functionals $\Phi \leq F$ having the same slope $u^* \in V^*$, the maximal is the continuous affine functional defined by $v \longmapsto \langle v, u^* \rangle - F^*(u^*)$.

Conversely, suppose that there exist $u^* \in V^*$ such that $F^*(u^*)$ is finite. Then any functional of the form $\Phi(v) = \langle v, u^* \rangle + \beta$ with $\beta \leq -F^*(u^*)$ satisfies $\Phi(v) \leq F(v)$ for all $v \in V$. Indeed, for any $v \in V$, we have

$$\begin{aligned} \langle v, u^* \rangle + \beta \leq \langle v, u^* \rangle - F^*(u^*) &= \langle v, u^* \rangle - \left( \sup_{u \in V} \langle u, u^* \rangle - F(u) \right) \\ &\leq \langle v, u^* \rangle - \langle v, u^* \rangle + F(v) = F(v). \end{aligned}$$

Now let us show another interesting fact. Let $\Phi \in \mathscr{J}(V)$ and suppose that $\Phi \leq F$, exact at a point $u \in V$, and defined for each $v \in V$ by the following expression $\Phi(v) = \langle v, u^* \rangle + \beta$, where $u^* \in V^*$ and $\beta \in \mathbb{R}$. Then we must have $\beta = -F^*(u^*)$, and consequently $\Phi$ is maximal among all the continuous affine functionals dominated by $F$ and having slope $u^*$. Moreover the following identity holds

$$F(u) + F^*(u^*) = \langle u, u^* \rangle.$$

In order to prove these claims, we first observe that the condition $\Phi \leq F$ implies $F^* \leq \Phi^*$ and consequently $F^*(u^*) \leq \Phi^*(u^*) = \sup_{v \in V} \langle v, u^* \rangle - \Phi(v) = \sup_{v \in V} \langle v, u^* \rangle - \langle v, u^* \rangle + \beta = \beta$. Since $F$ is exact at $u$, we have that $F(u)$ is finite. And so follows from the definition of Legendre-Fenchel transform and previous inequality that

$$-\infty < \langle u, u^* \rangle - F(u) \leq F^*(u^*) < \beta$$

thus showing that $F^*(u^*)$ must be finite.

From the previous arguments we conclude that

$$\Phi(v) \leq \langle v, u^* \rangle - F^*(u^*) \leq F(v), \qquad \forall v \in V.$$

Since we are assuming $F$ is exact at $u$, follows from the previous inequality that

$$F(u) = \Phi(u) \leq \langle u, u^* \rangle - F^*(u^*) \leq F(u). \tag{2.11}$$

Therefore $\Phi(u) = \langle u, u^* \rangle - F^*(u^*)$. But on the other hand, $\Phi(u) = \langle u, u^* \rangle + \beta$ and so $\beta = -F^*(u^*)$. Another consequence of (2.11) is the identity

$$F(u) + F^*(u^*) = \langle u, u^* \rangle.$$

Now we introduce the most important concept of this section.

Definition 2.28. A function $F : V \to \overline{\mathbb{R}}$ is said to be **subdifferentiable** at a point $u \in V$ if there exists some continuous affine functional $\Phi \leq F$ which is exact at $u$. The slope $u^*$ of $\Phi$ is called a **subgradient** of $F$ at $u$, and the set of all subgradients of $F$ at $u$ is called subdifferential of $F$ at $u$ and denoted by $\partial F(u)$.

We use the notation $\partial F(u) = \emptyset$ to say that $F$ is not subdifferentiable at $u$. Therefore for any $u \in V$ we always have $\partial F(u) \subset V^*$. We have the following characterization

$$u^* \in \partial F(u) \quad \Longleftrightarrow \quad F(u) \in \mathbb{R} \text{ and } \langle v - u, u^* \rangle + F(u) \leq F(v),\ \forall v \in V. \tag{2.12}$$

If $\Phi \in \mathscr{J}(V)$ is an arbitrary continuous affine functional dominated by $F$, then we have $\Phi \leq F^{**}$, which is the $\Gamma$-regularization of $F$. If in addition $\Phi$ is exact at $u$, then $F(u) = \Phi(u) \leq F^{**}(u) \leq F(u)$ and so $F(u) = F^{**}(u)$. From this observation and (2.12) we have

$$\partial F(u) \neq \emptyset \quad \Longrightarrow \quad F(u) = F^{**}(u).$$

Furthermore, we can show that

$$F(u) = F^{**}(u) \quad \Longrightarrow \quad \partial F(u) = \partial F^{**}(u).$$

In fact, for the case $F(u) = F^{**}(u) = \pm\infty$, we have both $\partial F(u)$ and $\partial F^{**}(u)$ are the empty set. Let us now consider the case $F(u) = F^{**}(u)$ finite. If $u^* \in \partial F(u)$ then follow from (2.12) that $\langle v - u, u^* \rangle + F(u) \leq F(v)$, for all $v \in V$. Since $F^{**}$ is the $\Gamma$-regularization of $F$ the previous inequality implies $\langle v - u, u^* \rangle + F(u) \leq F^{**}(v)$, for all $v \in V$. By replacing $F(u)$ in the last inequality by $F^{**}(u)$ we get $\langle v - u, u^* \rangle + F^{**}(u) \leq F^{**}(v)$, for all $v \in V$. This last inequality together with

(2.12) provide $u^* \in \partial F^{**}(u)$, thus showing that $\partial F(u) \subset \partial F^{**}(u)$. On the other hand, if $u^* \in \partial F^{**}(u)$ then we have from (2.12) that $\langle v-u, u^*\rangle + F^{**}(u) \leq F^{**}(v)$, for all $v \in V$. Since $F(u) = F^{**}(u)$ and $F^{**}(v) \leq F(v)$ for all $v \in V$, we conclude that $\langle v-u, u^*\rangle + F(u) \leq F(v)$, for all $v \in V$. One more application of (2.12) yields $u^* \in \partial F(u)$ and therefore $\partial F^{**}(u) \subset \partial F(u)$ which finishes the proof.

We now have a very interesting consequence of the definition of subdifferentiability, which make us anticipate the role of subdifferentials in optimization problems.

Proposition 2.29. *Let $F : V \to \overline{\mathbb{R}}$ be an arbitrary function. Then the following statements are equivalent*

$$F(u) = \inf_{v\in V} F(v) \qquad \Longleftrightarrow \qquad 0 \in \partial F(u). \tag{2.13}$$

Proof. Suppose initially that $0 \in \partial F(u)$. From the equivalence (2.12) we have that $\langle v-u, 0\rangle + F(u) \leq F(v)$ for all $v \in V$ and so $F(u) = \inf_{v\in V} F(v)$. Conversely, if $u \in V$ is such that $F(u) = \inf_{v\in V} F(v)$, then the following inequality holds $\langle v-u, 0\rangle + F(u) \leq F(v)$ for all $v \in V$ and from (2.12) follows that $0 \in \partial F(u)$. ∎

Proposition 2.30. *Let $F : V \to \overline{\mathbb{R}}$ be an arbitrary function and $F^*$ its Legendre-Fenchel transform. Then*

$$u^* \in \partial F(u) \qquad \Longleftrightarrow \qquad F(u) + F^*(u^*) = \langle u, u^*\rangle.$$

Proof. Suppose that $u^* \in \partial F(u)$. Then by (2.12) we have $\langle v-u, u^*\rangle + F(u) \leq F(v)$ for all $v \in V$. This implies $\langle v, u^*\rangle - F(v) \leq \langle u, u^*\rangle - F(u)$ for all $v \in V$. By taking the supremum over $V$ we get $F^*(u^*) = \langle u, u^*\rangle - F(u)$, which is equivalent $F(u) + F^*(u^*) = \langle u, u^*\rangle$.

Conversely, assume that $F(u) + F^*(u^*) = \langle u, u^*\rangle$. From the definition of the Legendre-Fenchel transform we have that

$$\langle u, u^*\rangle - F(u) = F^*(u^*) = \sup_{v\in V}\langle v, u^*\rangle - F(v) \geq \langle v, u^*\rangle - F(v), \qquad \forall v \in V.$$

Consider the continuous affine functional $\Phi : V \to \mathbb{R}$ given by $\Phi(v) = \langle v-u, u^*\rangle + F(u)$, for all $v \in V$. Then follows from the last inequality that $\Phi(v) \leq F(v)$ and $\Phi(u) = F(u)$. Then $\Phi$ is exact at $u$ and so its slope $u^*$ is a subgradient of $F$ at $u$, that is, $u^* \in \partial F(u)$. ∎

Corollary 2.31. *The set $\partial F(u)$ (which is possibly empty) is convex and closed in $V^*$, with respect to $\sigma(V^*, V)$.*

Proof. From the definition of the Legendre-Fenchel transform we have the following inequality $-F(u) \leq F^*(u^*) - \langle u, u^*\rangle$, for any $u \in V$ and $u^* \in V^*$. This observation together with Proposition 2.30 implies

$$\partial F(u) = \{u^* \in V^* : F^*(u^*) - \langle u, u^*\rangle \leq -F(u)\}.$$

Since $F^* \in \Gamma(V^*)$ and for any fixed $u \in V$ we have $u^* \longmapsto \langle u, u^*\rangle$ is a continuous mapping follows that $u^* \longmapsto F^*(u^*) - \langle u, u^*\rangle$ is a lower semicontinuous map. Therefore $\partial F(u)$ is closed in $V^*$, with respect to $\sigma(V^*, V)$.

Now we prove that $\partial F(u)$ is a convex set. Indeed, recall that $F^* \in \Gamma(V^*)$, in particular, $F^*$ is a convex function. Given $u^*, v^* \in \partial F(u)$ and $\lambda \in [0,1]$ we have

$$\begin{aligned} F^*(\lambda u^* + (1-\lambda)v^*) - \langle u, \lambda u^* + (1-\lambda)v^*\rangle& \\ &\leq \lambda(F^*(u^*) - \langle u, u^*\rangle) + (1-\lambda)(F^*(v^*) - \langle v, u^*\rangle) \\ &\leq \lambda(-F(u)) + (1-\lambda)(-F(u)) = -F(u) \end{aligned}$$

which shows that $\lambda u^* + (1-\lambda)v^* \in \partial F(u)$ and finish the proof. ■

Corollary 2.32. *For any function $F : V \to \overline{\mathbb{R}}$ we have*

*(1) $u^* \in \partial F(u) \Longrightarrow u \in \partial F^*(u^*)$.*
*(2) if $F \in \Gamma(V)$, then $u^* \in \partial F(u) \Longleftrightarrow u \in \partial F^*(u^*)$.*

Proof. Proof of *(1)*. Let $u^* \in \partial F(u)$. From Proposition 2.30 we have the following equality $F(u) + F^*(u^*) = \langle u, u^*\rangle$. Recalling $F^{**} \leq F$, we have

$$F^{**}(u) + F^*(u^*) \leq \langle u, u^*\rangle.$$

On the other hand, the definition of $F^{**}$ ensures the validity of reverse inequality and so $F^{**}(u) + F^*(u^*) = \langle u, u^*\rangle$. To get the conclusion, $u \in \partial F^*(u^*)$, we apply the Proposition 2.30 with the roles of $V$ and $V^*$ interchanged.

Proof of *(2)*. If $F \in \Gamma(V)$ then $F = F^{**}$. And so for any pair $u, u^*$ with $u^* \in \partial F(u)$ we have $F^{**}(u) + F^*(u^*) = \langle u, u^*\rangle$ and another application of Proposition 2.30 shows that $u \in \partial F^*(u^*)$. Similarly, if $u \in \partial F^*(u^*)$, then follows again from the identities $F^{**}(u) + F^*(u^*) = \langle u, u^*\rangle$ and $F = F^{**}$ that $u^* \in \partial F(u)$. ■

Proposition 2.33. *Let $F : V \to \overline{\mathbb{R}}$ a convex function and suppose that $F$ is continuous at $u \in V$. Then $\partial F(v) \neq \emptyset$ for all $v \in \operatorname{int}(\operatorname{dom} F)$. In particular, $\partial F(u) \neq \emptyset$.*

Proof. Since $F(u) \in \mathbb{R}$ and $F$ is continuous at $u$ follows that $F$ is uniformly bounded in some neighborhood $\mathcal{V}$ containing $u$. Furthermore $F$ is finite and continuous at any point of $\operatorname{int}(\operatorname{dom} F)$. Therefore it is enough to show that $\partial F(u) \neq \emptyset$.

The convexity of $F$ implies $\operatorname{epi} F$ is convex a subset of $V \times \mathbb{R}$. Moreover, the continuity of $F$ implies $\operatorname{int}(\operatorname{epi} F) \neq \emptyset$. Indeed, for each $v \in \operatorname{dom} F$ there exists neighborhood $\mathcal{O} \subset V$, with $v \in \mathcal{O}$ and a constant $c$ such that $|F(w)| \leq c$, for all $w \in \mathcal{O}$. Therefore the set $\mathcal{O} \times (c, +\infty) \subset V \times \mathbb{R}$ is an open set contained in $\operatorname{epi} F$.

As consequence of Hanh-Banach theorem we can separate $(u, F(u)) \in \partial(\operatorname{epi} F)$ from $\operatorname{int}(\operatorname{epi} F)$ by a closed affine hyperplane $\mathscr{H} \subset V \times \mathbb{R}$. To be more precise, $\mathscr{H}$ is a supporting hyperplane of $\operatorname{epi} F$, containing $(u, F(u))$.

The hyperplane $\mathscr{H} = \{(v, a) \in V \times \mathrm{R} : \langle v, u^*\rangle + a\alpha = \beta\}$, for some $u^* \in V^*$ and $\alpha, \beta \in \mathbb{R}$. Note that the coefficients are not all zero and the following hold

$$\langle v, u^*\rangle + a\alpha \geq \beta, \qquad \forall (v, a) \in \operatorname{epi} F; \tag{2.14}$$

$$\langle u, u^*\rangle + F(u)\alpha = \beta \qquad . \tag{2.15}$$

More precisely, the coefficient $\alpha$ can not be zero. Indeed, if $\alpha = 0$, then we have from the above expressions $\langle u, u^*\rangle = \beta \leq \langle v, u^*\rangle \Longrightarrow 0 \leq \langle v - u, u^*\rangle$, for all $v \in \operatorname{dom} F$. In particular, the last inequality holds for $v - u \in \mathcal{O}$, the open neighborhood of $u$ determined above. Actually, the neighborhood $\mathcal{O}$ can be taken as a symmetric neighborhood of $u$ of the form $\{u\} + (\mathscr{V} \cap -\mathscr{V})$, where $\mathscr{V}$ is a neighborhood of the origin. But this implies immediately that $u^* \equiv 0$, which is an absurd.

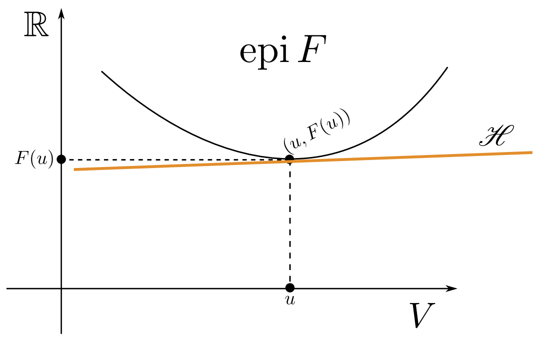


FIGURE 4. The supporting hyperplane $\mathscr{H}$ of $\operatorname{epi} F$, containing $(u, F(u))$.

From the last paragraph we learnt that the coefficient $\alpha$ appearing in the equation of $\mathscr{H}$ is not zero. Moreover, by taking any $(u, a) \in \operatorname{epi} F$, we get from (2.14) and (2.15) $\langle u, u^*\rangle + \alpha F(u) \leq \langle u, u^*\rangle + \alpha a \Longrightarrow 0 \leq \alpha(a - F(u)) \Longrightarrow \alpha > 0$. By using that *alpha* $> 0$ and dividing (2.14) and (2.15) by $\alpha$ we obtain the following:

$$\frac{\beta}{\alpha} - \langle v, \frac{u^*}{\alpha}\rangle \leq F(v), \ \forall v \in \operatorname{dom} F;$$

$$\frac{\beta}{\alpha} - \langle u, \frac{u^*}{\alpha}\rangle = F(u) \quad \Longleftrightarrow \quad \frac{\beta}{\alpha} = F(u) + \langle u, \frac{u^*}{\alpha}\rangle.$$

By replacing the last equality on the previous inequality we get

$$F(u) + \langle u, \frac{u^*}{\alpha}\rangle - \langle v, \frac{u^*}{\alpha}\rangle \leq F(v), \ \forall v \in \operatorname{dom} F;$$

$$\Downarrow$$

$$\langle u - v, \frac{u^*}{\alpha}\rangle + F(u) \leq F(v), \quad \forall v \in V.$$

From the last inequality and (2.12) follows that $u^*/\alpha \in \partial F(u)$, thus showing that this set is not empty. ■

## 2.8. Subdifferentials and Gâteaux Derivatives

Let $F : V \to \overline{\mathbb{R}}$ be a function. Given $u, v \in V$ if there exist the limit

$$\lim_{\lambda \downarrow 0} \frac{F(u + \lambda v) - F(u)}{\lambda} \equiv F'(u; v)$$

as an extended real number, we call it the directional derivative of $F$ at $u$ in the direction $v$. If there some $u^* \in V^*$ such that for all $v \in V$, we have $F'(u; v) = \langle v, u^*\rangle$ we say that $F$ is Gâteaux differentiable at $u$ and its Gâteaux derivative is denoted by $F'(u) = u^*$. Note that if $F$ is Gâteaux differentiable at a point $u$, then $F(u) \in \mathbb{R}$, that is, $F(u)$ is finite as well as $F'(u; v) = \langle v, u^*\rangle$.

Of course, if the Gâteaux derivative of a function $F$ exists at a point $u$, then it is unique. In fact,

$$\lim_{\lambda \downarrow 0} \frac{F(u + \lambda v) - F(u)}{\lambda} = \langle v, F'(u)\rangle.$$

LEMMA 2.34. *Let $u, v \in V$ be fixed. If $F : V \to \overline{\mathbb{R}}$ is a* ***convex*** *function, then the mapping*

$$(-\infty, 0) \cup (0, +\infty) \ni \lambda \longmapsto \frac{F(u + \lambda v) - F(u)}{\lambda}$$

*defines a non-decreasing function.*

PROOF. Let us first prove that the above map is non-decreasing on $(0, +\infty)$. Define an auxiliary function $G : (0, +\infty) \to \overline{\mathbb{R}}$, given by

$$G(\lambda) \equiv \frac{F(u + \lambda v) - F(u)}{\lambda}.$$

Therefore all we have to do is to show that $G(\lambda) \leq G(\lambda + \varepsilon)$ for any given $\varepsilon > 0$.

Note that the vector $u + \lambda v$ can be written as a convex combination of the vectors $u + (\lambda + \varepsilon)v$ and $u$ as follows

$$u + \lambda v = \frac{\lambda}{\lambda + \varepsilon}(u + (\lambda + \varepsilon)v) + \frac{\varepsilon}{\lambda + \varepsilon}u.$$

Since we are assuming that $F$ is a convex function the following estimate holds

$$F(u + \lambda v) \leq \frac{\lambda}{\lambda + \varepsilon}F(u + (\lambda + \varepsilon)v) + \frac{\varepsilon}{\lambda + \varepsilon}F(u).$$

Dividing both sides of the above inequality by $\lambda$ we get

$$\begin{aligned}
\frac{F(u + \lambda v)}{\lambda} &\leq \frac{1}{\lambda + \varepsilon}F(u + (\lambda + \varepsilon)v) + \frac{\varepsilon}{\lambda(\lambda + \varepsilon)}F(u) \\
&= \frac{1}{\lambda + \varepsilon}F(u + (\lambda + \varepsilon)v) + \frac{(\lambda + \varepsilon - \lambda)}{\lambda(\lambda + \varepsilon)}F(u) \\
&= \frac{1}{\lambda + \varepsilon}F(u + (\lambda + \varepsilon)v) + \frac{1}{\lambda}F(u) - \frac{1}{\lambda + \varepsilon}F(u)
\end{aligned}$$

From the above inequality it follows

$$G(\lambda) = \frac{F(u + \lambda v)}{\lambda} - \frac{1}{\lambda}F(u) \leq \frac{1}{\lambda + \varepsilon}F(u + (\lambda + \varepsilon)v) - \frac{1}{\lambda + \varepsilon}F(u) = G(\lambda + \varepsilon).$$

Now we prove the monotonicity for $\lambda \in (-\infty, 0)$. We first observe that

$$\frac{F(u + \lambda v) - F(u)}{\lambda} = (-1)\frac{F(u - \lambda(-v)) - F(u)}{-\lambda}, \tag{2.16}$$

By using the similar reasoning as in the previous case we can show that the map

$$(0, +\infty) \ni (-\lambda) \longmapsto \frac{F(u - \lambda(-v)) - F(u)}{-\lambda}$$

defines a non-decreasing function. So it follows from this observation and identity (2.16) that

$$(-\infty, 0) \ni \lambda \longmapsto \frac{F(u + \lambda v) - F(u)}{\lambda}$$

is actually a non-decreasing function. ■

As consequence of the above lemma for a convex function $F$ the following limit always exists (as an extended real number)

$$\lim_{h \downarrow 0} \frac{F(u + \lambda v) - F(u)}{\lambda} \equiv F'(u; v) \in [-\infty, +\infty].$$

In what follows we establish a relationship between Gâteaux-differentiability and subdifferentials of a **convex** function. We will see that if $F$ is Gâteaux-differentiable function at a point $u \in V$, then we have uniqueness of the subgradients of $F$ at $u$.

THEOREM 2.35. *Let $F : V \to \overline{\mathbb{R}}$ be a* ***convex*** *function. If $F$ is Gâteaux-differentiable at $u \in V$, then the subdifferential of $F$ at $u$ is not empty and moreover $\partial F(u) = \{F'(u)\}$.*

*Conversely, if $F(u)$ is finite, $F$ is continuous at $u \in V$ and $\#\partial F(u) = 1$, then $F$ is Gâteaux-differentiable at $u$ and $\partial F(u) = \{F'(u)\}$.*

PROOF. Suppose that $F$ is Gâteaux-differentiable at $u \in V$. For any $v \in V$ let $w = v - u$. Since $F$ is convex, by hypothesis, it follows from Lemma 2.34

$$\langle w, F'(u)\rangle = F'(u; w) = \lim_{\lambda \downarrow 0} \frac{F(u + \lambda w) - F(u)}{\lambda} \le F(u + w) - F(u).$$

From this inequality and the definition of $w$ we get

$$\langle v - u, F'(u)\rangle \le F(v) - F(u) \quad \Longrightarrow \quad \langle v - u, F'(u)\rangle + F(u) \le F(v), \qquad \forall v \in V.$$

By using the characterization (2.12) follows that $F'(u) \in \partial F(u)$. Remains to prove that if $u^* \in \partial F(u)$, then $u^* = F'(u)$. Indeed, for any $w \in V$ and $\lambda > 0$ we have by the characterization (2.12) that

$$\lambda\langle w, u^*\rangle \le F(u+\lambda w) - F(u) \quad \Longrightarrow \quad \langle w, u^*\rangle \le \lim_{\lambda \downarrow 0} \frac{F(u + \lambda w) - F(u)}{\lambda} = \langle w, F'(u)\rangle.$$

Therefore, for any $w \in V$, we have $\langle w, u^*\rangle \le \langle w, F'(u)\rangle$, which is equivalent to $\langle w, u^* - F'(u)\rangle \le 0$. Since the last inequality is valid for any $w \in V$ follows that $u^* - F'(u) = 0$, thus proving the claim.

Let us turn to the converse. Now, we assume that $F(u)$ is finite and continuous at $u \in V$ and the subdifferential of $F$ at $u$ has a unique subgradient.

Since $F$ is convex for a fixed $v \in V$, we obtain from Lemma 2.34 the following inequalities

$$F'(u; v) = \lim_{\lambda \downarrow 0} \frac{F(u + \lambda v) - F(u)}{\lambda} \le \frac{F(u + \lambda v) - F(u)}{\lambda}$$

and so $F(u) + \lambda F'(u; v) \le F(u + \lambda v)$ for any $\lambda > 0$. Similarly, if $\lambda < 0$ then

$$\frac{F(u + \lambda v) - F(u)}{\lambda} \le F'(u; v)$$

and so the inequality $F(u)+\lambda F'(u; v) \le F(u+\lambda v)$ is also valid for $\lambda < 0$. Therefore

$$F(u) + \lambda F'(u; v) \le F(u + \lambda v), \qquad \forall \lambda \in \mathbb{R}.$$

Geometrically this inequality says that the line $\gamma : \mathbb{R} \to V \times \mathbb{R}$ given by

$$\gamma(\lambda) = \Big(u + \lambda v, F(u) + \lambda F'(u; v)\Big)$$

does not intersect $\operatorname{int}(\operatorname{epi} F)$, which is a non-empty convex and open subset of $V \times \mathbb{R}$, because $F$ is continuous at $u$ and $F(u)$ is finite. Since $\gamma(\mathbb{R})$ is non-empty affine subspace, then a direct application of the Hanh-Banach Theorem (Theorem 2.3) ensures the existence of some closed hyperplane $\mathscr{H} \subset V \times \mathbb{R}$, separating $\gamma(\mathbb{R})$ from $\operatorname{int}(\operatorname{epi} F)$. Furthermore Theorem 2.3 ensures that we can chose the hyperplane $\mathscr{H}$ so that $\gamma(\mathbb{R}) \subset \mathscr{H}$.

The affine hyperplane $\mathscr{H} \equiv \{(w, a) \in V \times \mathbb{R} : \langle w, u^* \rangle + \alpha a = \beta\}$. By using the continuity of $F$ at $u$ we can find an open neighborhood of $u \in V$ of the form $\mathcal{O} \equiv \{u\} + (\mathscr{V} \cap -\mathscr{V})$, where $\mathscr{V}$ is a neighborhood of the origin, such that $|F(w)| \leq c$ for some positive constant $c \in \mathbb{R}$. From this observation we conclude that $\mathcal{O} \times (c, +\infty)$ is an open subset of $\operatorname{int}(\operatorname{epi} F)$, containing the point $(u, d)$, where $d = c + \varepsilon > F(u)$. Since the point $(u, F(u)) \in \gamma(\mathbb{R})$ follows from the equation of $\mathscr{H}$ that $\langle u, u^* \rangle + \alpha F(u) = \beta < \langle u, u^* \rangle + \alpha d$ and therefore $0 < \alpha(d - F(u))$ and so $\alpha > 0$. Consider $\Phi \in \mathscr{J}(V)$ given by

$$\Phi(w) = \frac{\beta}{\alpha} - \frac{1}{\alpha} \langle w, u^* \rangle.$$

Then we have $\langle w, u^* \rangle + \alpha \Phi(w) = \beta$ and therefore $\mathscr{H}$ is the graph of the continuous affine functional $\Phi$, that is, $\mathscr{H} = \{\big(v, \Phi(v)\big) \in V \times \mathbb{R} : v \in V\}$. We claim that $\Phi$ satisfies $\Phi \leq F$ and $\Phi$ is exact at $u$. Since $\operatorname{epi} F$ is a convex set with non-empty interior, we have that $\overline{\operatorname{int}(\operatorname{epi} F)} = \overline{\operatorname{epi} F}$, see page 14. Since

$$\operatorname{int}(\operatorname{epi} F) \subset \{(w, a) \in V \times \mathbb{R} : \langle w, u^* \rangle + \alpha a \geq \beta\}$$

and the last set is closed in $(V, \tau)$ because in a locally convex Hausdorff space every weakly closed **convex** set is a closed convex set. Therefore

$$\overline{\operatorname{epi} F} = \overline{\operatorname{int}(\operatorname{epi} F)} \subset \{(w, a) \in V \times \mathbb{R} : \langle w, u^* \rangle + \alpha a \geq \beta\}.$$

In particular, for all $w \in V$, we have $\beta \leq \langle w, u^* \rangle + \alpha F(w)$. On the other hand, we have observed above that $\langle u, u^* \rangle + \alpha F(u) = \beta$. Placing these two inequalities together we get

$$\langle u, u^* \rangle + \alpha F(u) = \beta \leq \langle w, u^* \rangle + \alpha F(w)$$

$$\Downarrow$$

$$\frac{1}{\alpha} \langle u - w, u^* \rangle + F(u) \leq F(w)$$

$$\Downarrow$$

$$\left( \frac{1}{\alpha} \langle u, u^* \rangle + F(u) \right) - \tfrac{1}{\alpha} \langle w, u^* \rangle \leq F(w)$$

As observed above the expression in parenthesis is equal to $\beta/\alpha$ and therefore we conclude that

$$\Phi(w) = \frac{\beta}{\alpha} - \frac{1}{\alpha} \langle w, u^* \rangle \leq F(w).$$

From last two expressions follows that $\Phi$ is exact at $u$ and so from definition $(-1/\alpha) u^* \in \partial F(u)$. Recalling that $\partial F(u)$ is a singleton we have that $\{(-1/\alpha) u^*\} = \partial F(u)$.

Since $\gamma(\mathbb{R}) \subset \mathscr{H}$ we have every point $\gamma(\lambda)$ is a point in $\mathscr{H}$. As observed above this point has to be the form $(w, \Phi(w))$ and so using the explicit expression for $\gamma(\lambda)$ we get

$$\Big(u + \lambda v, F(u) + \lambda F'(u; v)\Big) = \Big(u + \lambda v, \Phi(u + \lambda v)\Big).$$

The above equality implies

$$F(u) + \lambda F'(u; v) = \Phi(u + \lambda v) = \frac{\beta}{\alpha} - \frac{1}{\alpha} \langle u + \lambda v, u^* \rangle = F(u) - \frac{\lambda}{\alpha} \langle v, u^* \rangle.$$

Thus we have $\lambda F'(u;v) = -\frac{\lambda}{\alpha}\langle v, u^*\rangle$ which immediately implies $F'(u;v) = -\frac{1}{\alpha}\langle v, u^*\rangle$ and consequently $F$ is Gâteaux differentiable at $u$ and $F'(u) = (-1/\alpha)u^* = \partial F(u)$ thus finally finishing the proof. ■

# Bibliography


1. N. Bourbaki, *Topological vector spaces. Chapters 1–5*, Elements of Mathematics (Berlin), Springer-Verlag, Berlin, 1987, Translated from the French by H. G. Eggleston and S. Madan. MR 910295
2. ______, *General topology. Chapters 1–4*, Elements of Mathematics (Berlin), Springer-Verlag, Berlin, 1998, Translated from the French, Reprint of the 1989 English translation. MR 1726779
3. I. Ekeland and R. Temam, *Convex analysis and variational problems*, North-Holland Publishing Co., Amsterdam-Oxford; American Elsevier Publishing Co., Inc., New York, 1976, Translated from the French, Studies in Mathematics and its Applications, Vol. 1. MR 0463994
4. J. R. Munkres, *Topology*, Prentice Hall, Inc., Upper Saddle River, NJ, 2000, Second edition. MR 3728284
5. B. Simon, *Real analysis*, A Comprehensive Course in Analysis, Part 1, American Mathematical Society, Providence, RI, 2015, With a 68 page companion booklet. MR 3408971
6. S. Willard, *General topology*, Dover Publications, Inc., Mineola, NY, 2004, Reprint of the 1970 original Addison-Wesley, Reading, MA.