Commentationes Mathematicae Universitatis Carolinae

Luděk Jokl

An application of the Hahn-Banach theorem in convex analysis

*Commentationes Mathematicae Universitatis Carolinae*, Vol. 22 (1981), No. 4, 799--807

Persistent URL: http://dml.cz/dmlcz/106121

COMMENTATIONES MATHEMATICAE UNIVERSITATIS CAROLINAE

22,4 (1981)

# AN APPLICATION OF THE HAHN-BANACH THEOREM IN CONVEX ANALYSIS

Luděk JOKL

Abstract: A new principle (Theorem 7) which is based on the Hahn-Banach theorem, is presented. It is shown that some well-known basic theorems of convex analysis follow at once from this principle.

Key words: Convex function, conjugate function, subdifferential, infimal convolution, normal cone, Hahn-Banach theorem.

Classification: Primary 47H99
Secondary 46A15, 46A55

---

Let X, Y be linear topological spaces over reals R, $X^*$, $Y^*$ the dual spaces of X, Y, respectively, $\langle x,x^*\rangle$ the pairing between X and $X^*$. Let $A:X \longrightarrow Y$ be a linear continuous operator. The operator $A^*:Y^* \rightarrow X^*$ is defined by

$$x \in X,\ y^* \in Y^* \Longrightarrow \langle x, A^*y\rangle = \langle Ax, y^*\rangle .$$

Let $f:X \rightarrow [-\infty,+\infty]$ be a convex function. The effective domain $\{x \in X: f(x) < +\infty\}$ of f we denote by dom f. By $\partial f(x)$ we denote the subdifferential of f at the point $x \in X$,

$$\partial f(x) = \{x^* \in X^*: h \in X \Longrightarrow \langle h,x^*\rangle \leqq f(x+h) - f(x)\}.$$

The symbol $f^*$ stands for the conjugate function of f, defined by

$$x^* \in X^* \Longrightarrow f^*(x^*) = \sup\{\langle x, x^*\rangle - f(x) : x \in X\}.$$

Moreover, we use the symbol $\square$ to denote the operation of infimal convolution. By the normal cone to a convex set C at $x_0 \in C$ we mean the set $N(x_0|C)$ defined by

$$N(x_0|C) = \{x^* \in X^* : x \in C \Longrightarrow \langle x - x_0, x^*\rangle \leqq 0\}.$$

Let $K_{\partial f(x_0)}$ be a convex cone generated by $\{0\} \cup \partial f(x_0)$.

The following theorems play an important role in convex analysis:

Theorem 1 (Moreau, Rockafellar, [2, Chapt. 4, § 2, Th.1]). Let $f:X \to ]-\infty; +\infty]$ and $g:X \to ]-\infty, +\infty]$ be convex functions. Suppose there exists a point of dom $\cap$ dom $g$ at which $f$ is continuous. Then, for every $x \in X$,

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

Theorem 2 (Moreau, Rockafellar, [2, Chapt. 3, § 4, Th.1]). Under the assumptions of Theorem 1 it holds

$$(f + g)^* = f^* \square g^*.$$

Moreover, for every $x^* \in \text{dom } (f + g)^*$ there exist $y^* \in \text{dom } f$ and $z^* \in \text{dom } g$ such that

$$y^* + z^* = x^*, \quad f^*(y^*) + g^*(z^*) = (f + g)^*(x^*).$$

Theorem 3 ([2, Chapt. 4, § 2, Th. 2]). Let $A:X \to Y$ be a linear continuous operator and $f:Y \to ]-\infty, +\infty]$ be a convex function. Suppose there exists $x_0 \in X$ such that $f$ is finite and continuous at the point $y_0 = Ax_0$. Then, for every $x \in X$

$$\partial(f \circ A)(x) = A^* \partial f(Ax).$$

Theorem 4 ([2, Chapt. 3, § 4, Th. 3]). Under the assumptions of Theorem 3 it holds

$$f(A)^* = A^* f^*$$

and for every $x^* \in \operatorname{dom}(fA)^*$ there exists $y^* \in \operatorname{dom} f^*$ such that

$$A^* y^* = x^* , \quad (fA)^* (x^*) = f^*(y^*).$$

Theorem 5 ([2, Chapt. 4, § 3, Prop. 2]). Let $f: X \longrightarrow ]-\infty, +\infty]$ be a convex function which is finite and continuous at the point $x_0 \in X$. Put

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

Suppose there exists $x \in X$ such that $f(x) < f(x_0)$. Then

$$N(x_0 | C) = K_{\partial f(x_0)}.$$

The purpose of this paper is to demonstrate that Theorems 1 - 5 follow immediately from the principle expressed below in Theorem 7. To prove this theorem we use the following version of the Hahn-Banach theorem:

Theorem 6 ([2, Chapt. 3, § 2, Th. 1, Chapt 4, § 2, Prop. 3]). Let $\varphi : X \longrightarrow [-\infty, +\infty]$ be a convex function such that $\varphi$ is bounded from above on a neighbourhood of the origin and $\varphi(0) = 0$. Then there exists a linear continuous functional $x^* \in X^*$ such that for every $x \in X$

$$\langle x, x^* \rangle \leqq \varphi(x).$$

It should be noted that Theorem 6 is equivalent to the Eidelheit theorem on separation of convex sets.

Theorem 7. Let U be a linear space, V a linear topological space, $f:V \to ]-\infty, +\infty]$ and $h:U \to ]-\infty, +\infty]$ convex functions, and $T:U \to V$ a linear mapping. If

(i) there is $u_0 \in \operatorname{dom} h$ such that $f$ is finite and continuous at the point $v_0 = Tu_0$ and

(ii) $\inf\{f(Tu) + h(u) : u \in U\} = 0$,

then there exists a linear continuous functional $v^* \in V^*$ such that

(1) $u \in U,\ v \in V \Rightarrow \langle v, v^*\rangle - f(v) \leq h(u) + \langle Tu, v^*\rangle$.

<u>Proof</u>. Put

$F = \{(v, \lambda) \in V \times R : f(v) \leq \lambda\}$,

$H = \{(Tu, \mu) \in V \times R : u \in U,\ \mu \leq -h(u)\}$,

$M = F - H$.

Because $F$ and $H$ are convex sets, $M$ is also convex. Therefore the function $\varphi : V \to [-\infty, +\infty]$ defined by

(2) $w \in V \Rightarrow \varphi(w) = \inf\{\lambda \in R : (w, \lambda) \in M\} =$

$= \inf\{f(v) + h(u) : u \in U,\ v \in V,\ v - Tu = w\}$

is convex. From (2) and (ii) it follows

$$\varphi(0) = \inf\{f(Tu) + h(u) : u \in U\} = 0.$$

From (2) we conclude that

(3) $u \in U,\ v \in V \Rightarrow \varphi(v - Tu) \leq f(v) + h(u)$.

According to (i) there exist a constant $\alpha \in R$ and a neighbourhood of the origin $N \subset V$, such that

(4) $w \in N \Rightarrow f(w + Tu_0) \leq \alpha$.

Let $w \in N$. Then according to (3) and (4)

$\varphi(w) = \varphi((w + Tu_0) - Tu_0) \leq f(w + Tu_0) + h(u_0) \leq \alpha + h(u_0)$.

Therefore the function $\varphi$ is bounded from above on $N$.

By Theorem 6 there exists a linear continuous functional $v^* \in V^*$ such that

(5) $w \in V \Rightarrow \langle w, v^* \rangle \leqq \varphi(w)$.

Let $u \in U$, $v \in V$. Putting $w = v - Tu$ in (5), then according to (3) we have that

$$\langle v - Tu, v^* \rangle \leqq f(v) + h(u),$$

which implies (1).

Theorem 1 contains a non-trivial part, which can be expressed as

Lemma. In addition to the assumptions of Theorem 1 suppose that $f(0) = g(0) = 0$. Then

$$\partial(f + g)(0) \subset \partial f(0) + \partial g(0).$$

Proof. Let $w^* \in \partial(f + g)(0)$. Therefore

$$\inf \{ f(x) + g(x) - \langle x, w^* \rangle : x \in X \} = 0.$$

Now we put $U = V = X$, $Tx = x$, $h(x) = g(x) - \langle x, w^* \rangle$ in Theorem 7. By Theorem 7 there exists $v^* \in X^*$ with the property

$$x \in X,\ y \in X \Rightarrow \langle y, v^* \rangle - f(y) \leqq g(x) - \langle x, w^* - v^* \rangle .$$

From this relation it follows

$$v^* \in \partial f(0),\ w^* - v^* \in \partial g(0),$$

hence, $w^* = v^* + (w^* - v^*) \in \partial f(0) + \partial g(0)$.

Proof of Theorem 2. First of all

(6) $$(f + g)^* \leqq f^* \square\, g^* .$$

By the assumptions of Theorem 2

(7) $$(f + g)^* > -\infty .$$

If $x^* \notin$ dom $(f + g)^*$ , then by (6)

$$(8) \qquad (f + g)^* (x^*) = (f^* \square g^*)(x^*).$$

Let $x^* \in$ dom $(f + g)^*$ . Put

$$(9) \qquad \alpha = (f + g)^* (x^*).$$

By (7) we see that $\alpha \in R$. Now from (9) it follows

$$\inf \{f(x) + g(x) + \alpha - \langle x,x^*\rangle : x \in X\} = 0.$$

We put $U = V = X$, $Tx = x$, $h(x) = g(x) + \alpha - \langle x,x^*\rangle$ in Theorem 7. By this theorem there exists $y^* \in X^*$ with the property

$$x \in X,\ y \in X \Rightarrow (\langle y,y^*\rangle - f(y)) + (\langle x,x^* - y^*\rangle - g(x)) \leqq \alpha .$$

Therefore

$$f^*(y^*) + g^*(x^* - y^*) \leqq \alpha .$$

Hence

$$y^* \in \text{dom } f^* ,\ x^* - y^* \in \text{dom } g^* .$$

The definition of the infimal convolution, (6) and (9) imply that

$$\alpha \leqq (f^* \square g^*)(x^*) \leqq f^*(y^*) + g^*(x^* - y^*) \leqq \alpha .$$

The theorem is proved.

We formulate the non-trivial part of Theorem 3 as

<u>Lemma</u>. In addition to the assumptions of Theorem 3 suppose that $f(0) = 0$. Then

$$\partial(f \circ A)(0) \subset A^* \partial f(0).$$

<u>Proof</u>. Let $x^* \in \partial(f \circ A)(0)$. Then

$$\inf \{f(Ax) - \langle x,x^*\rangle : x \in X\} = 0.$$

Now we put $U = X$, $V = Y$, $T = A$, $h(x) = -\langle x,x^*\rangle$ in Theorem

7. By this theorem there exists $y^* \in Y^*$ with the property

$$x \in Y,\ y \in Y \Rightarrow \langle y, y^* \rangle - f(y) \leqq -\langle x, x^* \rangle + \langle Ax, y^* \rangle,$$

i.e.

$$x \in X,\ y \in Y \Rightarrow \langle y, y^* \rangle - f(y) \leqq \langle x, A^*y^* - x^* \rangle .$$

From this relation it follows immediately

$$x^* = A^*y^*,\ y^* \in \partial f(0).$$

Our lemma is proved.

<u>Proof of Theorem 4</u>. It holds $-\infty < (fA)^* \leqq A^* f^*$. Let $x^* \in \operatorname{dom} (fA)^*$. Put $\alpha = (fA)^* (x^*)$. Then $\alpha \in R$ and thus

$$\inf \{f(Ax) + \alpha - \langle x, x^* \rangle : x \in X\} = 0.$$

Define U, V, T in Theorem 7 in the same way as in the proof of Theorem 3. Next put $h(x) = \alpha - \langle x, x^* \rangle$ and proceed similarly as in the proof of Theorem 2.

Theorem 5 contains the non-trivial part, which we state as the following

<u>Lemma</u>. Let $f: X \rightarrow [-\infty, +\infty]$ is a convex function continuous at the origin, and $f(0) = 0$. Put

$$C = \{x \in X : f(x) \leqq 0\}.$$

If there exists $x_1 \in X$ such that $f(x_1) < 0$, then

$$N(0|C) \subset K_{\partial f(0)}.$$

<u>Proof</u>. Without the loss of generality one can assume that

$$x_1 \in \operatorname{int} C.$$

Let $0 \neq x^* \in N(0|C)$. Then

(10) $$\langle x_1, x^* \rangle < 0 ,$$

$$x \in \operatorname{Ker} x^* \Rightarrow f(x) \geqq 0,$$

i.e.

$$\inf \{ f(u) : u \in \operatorname{Ker} x^* \} = 0.$$

Set $U = \operatorname{Ker} x^*$ , $V = X$, $h(x) = 0$ in Theorem 7. Let T be the canonical injection of $\operatorname{Ker} x^*$ into X. By Theorem 7 there exists $y^* \in X^*$ with the property

$$x \in \operatorname{Ker} x^* ,\ y \in X \Rightarrow \langle y, y^* \rangle - f(y) \leqq \langle x, y^* \rangle .$$

Hence we conclude that

$$y^* \in \partial f(0),$$

(11) $\operatorname{Ker} x^* \subset \operatorname{Ker} y^*$ ,

(12) $\langle x_1, y^* \rangle \leqq f(x_1) < 0.$

According to (11) there exists $t \in R$ such that $y^* = tx^*$. By (12) and (10)

$$t = \langle x_1, y^* \rangle / \langle x_1, x^* \rangle > 0.$$

Hence $x^* = t^{-1} y^* \in t^{-1} \partial f(0) \subset K_{\partial f(0)}$,

which finishes the proof.

## R e f e r e n c e s


[1] R.T. ROCKAFELLAR: Convex analysis, Princeton University Press, 1970.

[2] A.D. IOFFE, V.M. TICHOMIROV: Teorija ekstremalnych zadač, Nauka, Moskva, 1974.

[3] V. BARBU, Th. PRECUPANU: Convexity and optimization in Banach spaces, Editura Academiei, Bucuresti-Sijthoof & Noordhoff Int. Publ. b.v., Alpen aan de Rijn, 1978.

Č V U T  Praha
Thákurova 7, 16629 Praha 6
Československo

(Oblatum 10.4. 1981)