STUDIA MATHEMATICA 189 (1) (2008)

Mazur–Orlicz equality

by

FON-CHE LIU (Tamshui)

Dedicated to Andrzej Granas

Abstract. A remarkable theorem of Mazur and Orlicz which generalizes the Hahn–Banach theorem is here put in a convenient form through an equality which will be referred to as the Mazur–Orlicz equality. Applications of the Mazur–Orlicz equality to lower barycenters for means, separation principles, Lax–Milgram lemma in reflexive Banach spaces, and monotone variational inequalities are provided.

1. Introduction. All vector spaces considered here are real vector spaces.

A real-valued function q defined on a vector space $\mathbb{E}$ is called a *sublinear functional* if (i) $q(\lambda x) = \lambda q(x)$ for $\lambda > 0$ and $x \in \mathbb{E}$; (ii) $q(x+y) \le q(x) + q(y)$ for x and y in $\mathbb{E}$. Since $q(0) = q(\lambda 0) = \lambda q(0)$ for all $\lambda > 0$, and $q(0) = q(x + [-x]) \le q(x) + q(-x)$, it follows that $q(0) = 0$ and $-q(-x) \le q(x)$ for $x \in \mathbb{E}$. As usual, the algebraic dual of $\mathbb{E}$ will be denoted by $\mathbb{E}'$.

Mazur and Orlicz proved in [10] the following remarkable theorem:

THEOREM 1. *Let $\mathbb{E}$ be a vector space with a sublinear functional q defined on it. Suppose S is an arbitrary nonempty set, τ a map from S into $\mathbb{E}$, and θ a real-valued function on S. Then the following two statements are equivalent*:

(A) *There is $l \in \mathbb{E}'$ with $l \le q$ on $\mathbb{E}$ such that*

$$\theta(s) \le l(\tau(s)) \qquad \forall s \in S.$$

2000 *Mathematics Subject Classification*: 46A20, 46A22.

Key words and phrases: sublinear functional, means, barycenter, duality map, variational inequality.

The paper is an expanded version of a paper read at Conference of Fixed Point Theory and its Applications in honor of Andrzej Granas, held in Montréal from August 16 to 20, 2004.

(B) *For any finite subset* $\{s_1, \dots, s_l\}$ *of* S,

$$\sum_{j=1}^{l} \lambda_j \theta(s_j) \le q\Big(\sum_{j=1}^{l} \lambda_j \tau(s_j)\Big)$$

for all $\lambda_1 \ge 0, \dots, \lambda_l \ge 0$.

Remark 1. To give a flavor of Theorem 1, we infer from Theorem 1 that for each $x_0 \in \mathbb{E}$ there is $l \in \mathbb{E}'$ with $l \le q$ such that $l(x_0) = q(x_0)$. Indeed, let $S = \{x_0\}$, $\tau(x_0) = x_0$, and $\theta(x_0) = q(x_0)$ in Theorem 1. Then (B) holds trivially and hence (A) holds, i.e. there is $l \in \mathbb{E}'$ with $l \le q$ such that $l(\tau(x_0)) = l(x_0) \ge \theta(x_0) = q(x_0)$. But $l(x_0) \le q(x_0)$, because $l \le q$. Thus $l(x_0) = q(x_0)$.

Mazur and Orlicz [10] also gave extensive applications of Theorem 1, showing that it is an ingenious and useful form of the Hahn–Banach theorem. Different proofs and generalizations of this theorem abound in the literature (see, for example, [1], [4], [5], [8], [9], [11], [12]) and it is now usually referred to as the Mazur–Orlicz theorem.

In this note we consider a convenient form of the Mazur–Orlicz theorem, which we will later refer to as the Mazur–Orlicz equality (see Theorem 2). This form will be given in Section 2 together with some applications to lower barycenters of means (see below for definitions) and to some general separation principles. In Section 3, we shall apply the Mazur–Orlicz equality to obtain a generalization of the Lax–Milgram theorem in reflexive Banach spaces and to give a simple proof of a variational inequality of Hartman and Stampacchia [6].

The remaining part of this section is devoted to some necessary definitions and preliminaries. For a given nonempty set S, we denote by $P_{\mathrm{f}}(S)$ the family of all probability measures supported on finite sets in S. Hence if $p \in P_{\mathrm{f}}(S)$, then there is a finite set $\{s_1, \dots, s_l\}$ in S and $\lambda_1 \ge 0, \dots, \lambda_l \ge 0$ with $\sum_{j=1}^{l} \lambda_j = 1$ such that $p(A) = \sum_{s_j \in A} \lambda_j$; p will then be written as $\big\langle {s_1,\dots,s_l \atop \lambda_1,\dots,\lambda_l} \big\rangle$. If τ is a map from S into a vector space, we shall denote $\int_S \tau\, dp = \sum_{j=1}^{l} \lambda_j \tau(s_j)$ by $\tau(p)$ if $p = \big\langle {s_1,\dots,s_l \atop \lambda_1,\dots,\lambda_l} \big\rangle$.

For a given set S, let $B(S)$ be the vector space of all bounded real-valued functions defined on S. The sublinear functional sup is defined by

$$\sup(f) = \sup_{s \in S} f(s)$$

for $f \in B(S)$; $\sup(|f|)$ is denoted by $\|f\|_\infty$ and called the sup-norm of f. It is clear that $B(S)$ is a Banach space with the sup-norm. Let $\mathbb{E}$ be a vector subspace of $B(S)$ and suppose $\mathbb{E}$ satisfies the following conditions:

(i) $\mathbb{E}$ contains all the constant functions;
(ii) if $f \in \mathbb{E}$, then $f^+ = f \vee 0 \in \mathbb{E}$.

From (ii) it follows that $f^- = -(f \wedge 0) = f^+ - f$ is in $\mathbb{E}$. Such a vector subspace $\mathbb{E}$ of $B(S)$ is called a $\vee$-*subspace* of $B(S)$. In particular, if S is a topological space, then the space of all bounded continuous functions is a $\vee$-subspace of $B(S)$. If $\mathbb{E}$ is a $\vee$-subspace of $B(S)$, then $l \in \mathbb{E}'$ is called a *mean* on $\mathbb{E}$ if

(i) $l(f) \geq 0$ when $f \geq 0$;
(ii) $l(1) = 1$;
(iii) $l(f) \leq \|f\|_\infty$.

It is easily verified that a linear form l is a mean on $\mathbb{E}$ if and only if $l \leq \sup$ on $\mathbb{E}$.

Let l be a mean on a $\vee$-subspace $\mathbb{E}$ of $B(S)$ and $\mathbb{F} \subset \mathbb{E}$. Then $x_0 \in S$ will be called a *lower barycenter* of l relative to $\mathbb{F}$ if

$$f(x_0) \leq l(f) \quad \forall f \in \mathbb{F}.$$

Note that if $-\mathbb{F} \subset \mathbb{F}$, then x_0 is a lower barycenter of l relative to $\mathbb{F}$ if and only if

$$f(x_0) = l(f) \quad \forall f \in \mathbb{F};$$

in this case we call x_0 a *barycenter* of l relative to $\mathbb{F}$. In particular, if $\mathbb{F}$ is a vector subspace of $\mathbb{E}$, then a lower barycenter of l relative to $\mathbb{F}$ is always a barycenter of l relative to $\mathbb{F}$.

A family $\mathbb{F}$ of functions defined on a set S is called *jointly convex-like* on S if for any finite family $\{f_1, \ldots, f_l\} \subset \mathbb{F}$ and for any $p \in P_{\mathrm{f}}(S)$ there is $x \in S$ such that

$$f_j(x) \leq f_j(p), \quad j = 1, \ldots, l.$$

The smallest topology on S with respect to which each function in $\mathbb{F}$ is lower semicontinuous will be denoted by $\underline{\mathcal{T}}(\mathbb{F})$. If S is compact for the topology $\underline{\mathcal{T}}(\mathbb{F})$, then the pair $(S, \mathbb{F})$ is called *lower compact*.

2. Mazur–Orlicz equality. We now formulate and prove the Mazur–Orlicz equality alluded to in Section 1. Some applications of it will also be considered in this section.

THEOREM 2 (Mazur–Orlicz equality). *Let* $\mathbb{E}$, q, *and* τ *be as in Theorem* 1. *Then*

$$\max_{\substack{l \in \mathbb{E}' \\ l \leq q}} \inf_{s \in S} l(\tau(s)) = \inf_{p \in P_{\mathrm{f}}(S)} q(\tau(p)). \tag{1}$$

Proof. For any $p \in P_{\mathrm{f}}(S)$ and $l \in \mathbb{E}'$ with $l \leq q$ we have (existence of such l is due to Banach, see also Remark 1)

$$\inf_{s \in S} l(\tau(s)) \leq l(\tau(p)) \leq q(\tau(p)),$$

hence

$$\sup_{\substack{l\in\mathbb{E}'\\ l\le q}}\inf_{s\in S} l(\tau(s)) \le \inf_{p\in P_{\mathrm{f}}(S)} q(\tau(p)).$$

If $\inf_{p\in P_{\mathrm{f}}(S)} q(\tau(p)) = -\infty$, then for $l\in\mathbb{E}'$ with $l\le q$ we have

$$\inf_{s\in S} l(\tau(s)) = -\infty,$$

and consequently

$$\max_{\substack{l\in\mathbb{E}'\\ l\le q}}\inf_{s\in S} l(\tau(s)) = \inf_{p\in P_{\mathrm{f}}(S)} q(\tau(p)) = -\infty.$$

If $\beta = \inf_{p\in P_{\mathrm{f}}(S)} q(\tau(p)) > -\infty$, then β is a finite number and

$$q(\tau(p)) \ge \beta \quad \forall p\in P_{\mathrm{f}}(S).$$

Now if we let $\theta(s)=\beta$ for $s\in S$, then since q is sublinear, statement (B) in Theorem 1 holds. Then statement (A) holds by Theorem 1, which means that there is $l\in\mathbb{E}'$ with $l\le q$ such that

$$\inf_{s\in S} l(\tau(s)) \ge \beta.$$

Hence $\inf_{s\in S} l(\tau(s)) = \beta$, concluding the proof. ■

Remark 2. We have shown that the Mazur–Orlicz equality follows from the Mazur–Orlicz theorem (Theorem 1). Now we show the converse. So suppose Theorem 2 holds; we shall show that Theorem 1 follows. Since in Theorem 1, (B) follows from (A) trivially, we need only prove that (A) follows from (B). For this purpose, let $\widehat{\mathbb{E}} = \mathbb{E}\oplus\mathbb{R}$, $\widehat{q}(x,t) = q(x)+t$ if $x\in\mathbb{E}$ and $t\in\mathbb{R}$, and $\widehat{\tau}(s) = (\tau(s), -\theta(s))$ for $s\in S$. By Theorem 2 with $\mathbb{E}$, τ, q replaced by $\widehat{\mathbb{E}}$, $\widehat{\tau}$, and $\widehat{q}$ respectively, there is $\widehat{l}\in\widehat{\mathbb{E}}'$ with $\widehat{l}\le\widehat{q}$ such that

$$\inf_{s\in S}\widehat{l}(\widehat{\tau}(s)) = \inf_{p\in P_{\mathrm{f}}(S)}\widehat{q}(\widehat{\tau}(p)),$$

or

$$\widehat{l}(\widehat{\tau}(s)) \ge \inf_{p\in P_{\mathrm{f}}(S)}\widehat{q}(\widehat{\tau}(p)) \quad \forall s\in S. \tag{2}$$

Since $\widehat{l}(x,t) = l(x)+\alpha t$ where $l\in\mathbb{E}'$ and $\alpha\in\mathbb{R}$, it is easily verified from $\widehat{l}\le\widehat{q}$ that $l\le q$ and $\alpha=1$. Then

$$\widehat{l}(\widehat{\tau}(s)) = l(\tau(s)) - \theta(s) \quad \forall s\in S. \tag{3}$$

Now for $p\in P_{\mathrm{f}}(S)$, $\widehat{q}(\widehat{\tau}(p)) = q(p)-\theta(p)\ge 0$ by (B), hence $\inf_{p\in P_{\mathrm{f}}(S)}\widehat{q}(\widehat{\tau}(p)) \ge 0$. It then follows from (2) and (3) that $\theta(s)\le l(\tau(s))$ for $s\in S$. Hence (A) holds. Therefore the Mazur–Orlicz equality is equivalent to the Mazur–Orlicz theorem.

Corollary 1. *Let $f_1, \ldots, f_n$ be real-valued functions defined on a set S. Then*

$$\max_{\lambda\in\Delta^{n-1}} \inf_{s\in S} \sum_{j=1}^{n} \lambda_j f_j(s) = \inf_{p\in P_{\mathrm{f}}(S)} \max_{1\le j\le n} f_j(p), \tag{4}$$

$$\min_{\lambda\in\Delta^{n-1}} \sup_{s\in S} \sum_{j=1}^{n} \lambda_j f_j(s) = \sup_{p\in P_{\mathrm{f}}(S)} \min_{1\le j\le n} f_j(p), \tag{5}$$

where Δ^{n-1} is the standard simplex in $\mathbb{R}^n$.

Proof. (4) follows from Theorem 2 with $\mathbb{E}=\mathbb{R}^n$, $\tau(s) = (f_1(s), \ldots, f_n(s))$ and $q(x) = \max_{1\le j\le n} x_j$ for $x = (x_1, \ldots, x_n)$ in $\mathbb{R}^n$; while (5) follows from (4) by replacing each f_j by $-f_j$. ■

Corollary 1 has been obtained in [9] by using a generalized form of the Mazur–Orlicz theorem and has been applied there to prove a minimax theorem of Ky Fan [3]. We now use Corollary 1 to show the existence of lower barycenters for means.

Theorem 3. *Let S be a nonempty set and $\mathbb{E}$ a $\vee$-subspace of $B(S)$. Let $\mathbb{F} \subset \mathbb{E}$. Suppose $\mathbb{F}$ is jointly convex-like on S and $(S, \mathbb{F})$ is a lower compact pair. Then each mean l on $\mathbb{E}$ has a lower barycenter relative to $\mathbb{F}$.*

Proof. Let l be a mean on $\mathbb{E}$. For each $f \in \mathbb{F}$, let

$$A_f = \{x \in S : f(x) \le l(f)\}.$$

In the following, we consider on S the topology $\underline{\mathcal{T}}(\mathbb{F})$, hence S is compact and each $f \in \mathbb{F}$ is lower semicontinuous. For $f \in \mathbb{F}$, there is $y \in S$ such that $f(y) = \min_{x\in S} f(x)$. Then $f(y) \le l(f)$ and hence $A_f \ne \emptyset$. Let now $\{f_1, \ldots, f_n\}$ be any finite subset of $\mathbb{F}$. By Corollary 1,

$$\begin{aligned} \inf_{p\in P_{\mathrm{f}}(S)} \max_{1\le j\le n} \{f_j(p) - l(f_j)\} &= \max_{\lambda\in\Delta^{n-1}} \inf_{s\in S} \sum_{j=1}^{n} \lambda_j \{f_j(s) - l(f_j)\} \\ &= \max_{\lambda\in\Delta^{n-1}} \min_{s\in S} \Big\{ \sum_{j=1}^{n} \lambda_j f_j(s) - l\Big(\sum_{j=1}^{n} \lambda_j f_j \Big) \Big\} \\ &\le 0. \end{aligned} \tag{6}$$

Since $\mathbb{F}$ is jointly convex-like, for each $p \in P_{\mathrm{f}}(S)$, there is $s \in S$ such that

$$f_j(p) \ge f_j(s), \quad j = 1, \ldots, n,$$

hence

$$\max_{1\le j\le n} \{f_j(p) - l(f_j)\} \ge \max_{1\le j\le n} \{f_j(s) - l(f_j)\},$$

and consequently from (6) we have

$$0 \ge \inf_{p\in P_{\mathrm{f}}(S)} \max_{1\le j\le n} \{f_j(p) - l(f_j)\} \ge \min_{s\in S} \max_{1\le j\le n} \{f_j(s) - l(f_j)\}.$$

Thus there is $x \in S$ such that

$$f_j(x) \leq l(f_j), \quad j = 1, \ldots, n,$$

or $x \in \bigcap_{j=1}^{n} A_{f_j}$. We have shown that $\{A_f\}_{f \in \mathbb{F}}$ has the finite intersection property. Since S is compact, it follows that $\bigcap_{f \in \mathbb{F}} A_f \neq \emptyset$. Let $x \in \bigcap_{f \in \mathbb{F}} A_f$. Then x is a lower barycenter of l relative to $\mathbb{F}$. ∎

COROLLARY 2. *If X is a compact convex subset in a topological vector space, and $\mathbb{F}$ a family of bounded lower semicontinuous convex functions on X, then for each mean l on a $\vee$-subspace of $B(X)$ containing $\mathbb{F}$, there is $x_0 \in X$ such that*

$$f(x_0) \leq l(f) \quad \forall f \in \mathbb{F}.$$

In particular, if l is a mean on $C(X)$, then there is $x_0 \in X$ such that

$$f(x_0) = l(f)$$

for all continuous affine functions f on X.

The formulation and proof of Theorem 3 suggest the following question: under what condition on τ, the RHS of (1) in Theorem 2 can be replaced by $\inf_{s \in S} q(\tau(s))$? We now consider this question.

Let $\mathbb{E}$, q, S, and τ be as in Theorem 1. Then τ is called *almost q-convex* if for every $p \in P_{\mathrm{f}}(S)$ and $\varepsilon > 0$, there is $s \in S$ such that

$$l(\tau(s)) \leq l(\tau(p)) + \varepsilon \tag{7}$$

for all $l \in \mathbb{E}'$ with $l \leq q$. It is a routine matter to verify that τ is almost q-convex if for any s_1, s_2 in S and $\varepsilon > 0$, there is $s \in S$ such that

$$l(\tau(s)) \leq \tfrac{1}{2}\{l(\tau(s_1)) + l(\tau(s_2))\} + \varepsilon \tag{8}$$

for all $l \in \mathbb{E}'$ with $l \leq q$.

A set $S \subset \mathbb{E}$ is called *almost q-convex* if the identification map from S into $\mathbb{E}$ is almost q-convex, i.e. for any s_1, s_2 in S and $\varepsilon > 0$ there is $s \in S$ such that

$$l(s) \leq \tfrac{1}{2}\{l(s_1) + l(s_2)\} + \varepsilon \tag{9}$$

for all $l \in \mathbb{E}'$ with $l \leq q$. In particular, if $S \subset \mathbb{E}$ has the property that for any s_1, s_2 in S there is $s \in S$ such that

$$q\left(s - \frac{s_1 + s_2}{2}\right) \leq 0, \tag{10}$$

then S is almost q-convex. Sets $S \subset \mathbb{E}$ satisfying (10) have been first introduced in [7].

THEOREM 4. *Let $\mathbb{E}$, q, and τ be as in Theorem* 1 *and assume that τ is almost q-convex. Then*

$$\max_{\substack{l\in\mathbb{E}'\\ l\le q}} \inf_{s\in S} l(\tau(s)) = \inf_{s\in S} q(\tau(s)). \tag{11}$$

Proof. By Theorem 2 and the fact that $\inf_{s\in S} q(\tau(s)) \ge \inf_{p\in P_{\mathrm{f}}(S)} q(\tau(p))$, it is sufficient to show that for each $p \in P_{\mathrm{f}}(S)$ and $\varepsilon > 0$ there is $s \in S$ such that

$$q(\tau(s)) \le q(\tau(p)) + \varepsilon. \tag{12}$$

Since τ is almost q-convex, there is $s \in S$ such that

$$l(\tau(s)) \le l(\tau(p)) + \varepsilon$$

for all $l \in \mathbb{E}'$ with $l \le q$. For this s, there is $\widehat{l} \in \mathbb{E}'$ with $\widehat{l} \le q$ such that $\widehat{l}(\tau(s)) = q(\tau(s))$ (see Remark 1), hence

$$q(\tau(s)) = \widehat{l}(\tau(s)) \le \widehat{l}(\tau(p)) + \varepsilon \le q(\tau(p)) + \varepsilon,$$

thus (12) holds. ∎

COROLLARY 3. *If $S \subset \mathbb{E}$ is almost q-convex, then there is $l \in \mathbb{E}'$ with $l \le q$ such that*

$$\inf_{s\in S} l(s) = \inf_{s\in S} q(s).$$

REMARK 3. If $S \subset \mathbb{E}$ satisfies (10), then S is almost q-convex and Corollary 3 holds. This special case is proved in [7] together with many applications.

REMARK 4. Corollary 3 contains as special case a strict separation principle: if $\inf_{s\in S} q(s) = \sigma > 0$, then there is $l \in \mathbb{E}'$ with $l \le q$ such that $l(s) \ge \sigma$ for all $s \in S$.

3. Further applications. In this section we consider two further applications of Theorem 2. The first application is a generalization of the Lax–Milgram theorem in reflexive Banach spaces; the second application is a simplified proof of a special case of a theorem of Hartman and Stampacchia [6] on a variational inequality as given by Dugundji and Granas [2]. This special case is itself a fairly general form of variational inequality.

Let $\mathbb{E}$ be a reflexive Banach space and let $(x, x^*) \mapsto B(x, x^*)$ be bilinear on $\mathbb{E} \times \mathbb{E}^*$. Then B is called *bounded* if there is $C \ge 0$ such that

$$|B(x, x^*)| \le C\|x\|\,\|x^*\|, \qquad x \in \mathbb{E},\ x^* \in \mathbb{E}^*,$$

and *positive definite* if there is $\sigma > 0$ such that

$$\max_{x^*\in J_x} B(x, x^*) \ge \sigma\|x\|^2 \quad \forall x \in \mathbb{E}, \tag{13}$$

and for $x \neq 0$ we have

$$B(x, x^*) > 0 \quad \forall x^* \in J_x, \tag{14}$$

where for $x \in \mathbb{E}$,

$$J_x = \{x^* \in \mathbb{E}^* : \|x^*\| = \|x\| \text{ and } \langle x^*, x\rangle = \|x\|^2\}.$$

It is known from the Hahn–Banach theorem that $J_x \neq \emptyset$ for $x \in \mathbb{E}$. The set-valued map J is called the *duality map* of $\mathbb{E}$.

Remark 5. If $\mathbb{E}$ is a Hilbert space and $\mathbb{E}^*$ is identified with $\mathbb{E}$ through the Riesz representation theorem, then $J_x = \{x\}$ for $x \in \mathbb{E}$. Then positive definiteness defined above coincides with the positive definiteness of bilinear forms on Hilbert space.

The following theorem is a generalization of the Lax–Milgram theorem (see Remark 5):

Theorem 5. *Let $\mathbb{E}$ be a reflexive Banach space and B a bounded and positive definite bilinear form on $\mathbb{E} \times \mathbb{E}^*$. Then for each $l \in \mathbb{E}^*$, there is a unique $y_0^* \in \mathbb{E}^*$ such that*

$$l(x) = B(x, y_0^*), \qquad x \in \mathbb{E}.$$

Furthermore, y_0^* satisfies $\|y_0^*\| \le \|l\|/\sigma$, where σ is the constant in (13).

Proof. We may assume that $l \neq 0$ and let $\beta = \|l\|/\sigma > 0$. Let K be the closed unit ball in $\mathbb{E}^*$. Since $\mathbb{E}$ is reflexive, K is compact in the weak topology of $\mathbb{E}^*$ and hence so is $S = \beta K$. In the following, the topology on S is the topology inherited from $\mathbb{E}^*$ with weak topology. Consider now the map $\tau : \mathbb{E} \to C(S)$ defined by

$$\tau(x) = B(x, \cdot) - l(x), \qquad x \in \mathbb{E}.$$

Obviously τ is a linear map, and hence if $p \in P_{\mathrm{f}}(\mathbb{E})$, then $\tau(p) = B(x_0, \cdot) - l(x_0) = \tau(x_0)$ for some $x_0 \in \mathbb{E}$. Hence

$$\inf_{p \in P_{\mathrm{f}}(\mathbb{E})} \sup \tau(p) = \inf_{x \in \mathbb{E}} \sup \tau(x),$$

where the sublinear functional sup on $C(S) \subset B(S)$ is defined in Section 1. Now for $x \in \mathbb{E}$, $x \neq 0$,

$$\begin{aligned}
\sup \tau(x) &= \max_{y^* \in S} \{B(x, y^*) - l(x)\} \\
&\ge \max_{y^* \in J_x} \left\{ B\left(x, \frac{\beta}{\|x\|}\, y^*\right) - l(x) \right\} \\
&= \frac{\beta}{\|x\|} \max_{y^* \in Jx} B(x, y^*) - l(x) \\
&\ge \frac{\beta}{\|x\|}\, \sigma \|x\|^2 - l(x) = \|l\| \cdot \|x\| - l(x) \ge 0,
\end{aligned}$$

and consequently

$$\inf_{p\in P_{\mathrm{f}}(\mathbb{E})} \sup \tau(p) \geq 0.$$

From Theorem 2, it follows that there is a mean μ on $C(S)$ such that

$$\mu(\tau(x)) \geq 0 \quad \forall x \in \mathbb{E}.$$

Since τ is linear, $\mu(\tau(x)) = 0$ for all $x \in \mathbb{E}$. But by Corollary 2, there is $y_0^* \in S$ such that

$$\tau(x)(y_0^*) = \mu(\tau(x)) = 0$$

for all $x \in \mathbb{E}$, or

$$l(x) = B(x, y_0^*) \quad \forall x \in \mathbb{E}. \tag{15}$$

Now we claim that y_0^* is uniquely determined by (15). Suppose y_1^* satisfies (15) with y_0^* replaced by y_1^* and $y_1^* \neq y_0^*$. Then, since $\mathbb{E}$ is reflexive, there is x_0 in $\mathbb{E}$ with $\|x_0\| = 1$ such that

$$\|y_0^* - y_1^*\| = \langle y_0^* - y_1^*, x_0\rangle,$$

which implies $\frac{y_0^*-y_1^*}{\|y_0^*-y_1^*\|} \in Jx_0$ and hence

$$B\left(x_0, \frac{y_0^* - y_1^*}{\|y_0^* - y_1^*\|}\right) > 0,$$

consequently $B(x_0, y_0^* - y_1^*) > 0$, or

$$l(x_0) = B(x_0, y_0^*) > B(x_0, y_1^*) = l(x_0),$$

which is absurd. So y_0^* is uniquely determined. Since $y_0^* \in S = \beta K$, we have $\|y_0^*\| \leq \beta = \|l\|/\sigma$. ∎

REMARK 6. From the proof of Theorem 5, it is clear that the existence of y_0^* still holds even if we do not assume that condition (14) holds for B. But then uniqueness of y_0^* is not guaranteed.

THEOREM 6 (Hartman–Stampacchia). *Let X be a reflexive Banach space, and C a closed, bounded, and convex subset of X. Suppose $f : C \to X^*$ is monotone, i.e. $\langle f(x)-f(y), x-y\rangle \geq 0$ for any x and y in C. Assume further that $f|_{L\cap C}$ is continuous for each line L in X. Then there is $y_0 \in C$ such that $\langle f(y_0), y_0 - x\rangle \leq 0$ for all x in C.*

Proof. We claim first that there is $y_0 \in C$ such that $\langle f(x), y_0 - x\rangle \leq 0$ for all $x \in C$. For this purpose, let $\mathbb{E}$ be the Banach space of all continuous functions on C with sup-norm, where C is equipped with the weak topology and therefore is compact by the reflexivity of X, and let $\tau : C \to \mathbb{E}$ be defined by

$$\tau(s)(x) = \langle f(s), s - x\rangle$$

for s and x in C. Let $p \in P_{\mathrm{f}}(C)$, say $p = \left\langle {s_1,\dots,s_n \atop \lambda_1,\dots,\lambda_n} \right\rangle$. Then

$$\begin{aligned}\tau(p)(x) &= \sum_{i=1}^{n} \lambda_i \langle f(s_i), s_i - x\rangle \\ &= \sum_{i=1}^{n} \lambda_i \{\langle f(s_i) - f(x), s_i - x\rangle + \langle f(x), s_i - x\} \\ &\ge \sum_{i=1}^{n} \lambda_i \langle f(x), s_i - x\rangle = \Big\langle f(x), \sum_{i=1}^{n} \lambda_i s_i - x \Big\rangle,\end{aligned}$$

and hence

$$\sup \tau(p) \ge \sup_{x\in C} \Big\langle f(x), \sum_{i=1}^{n} \lambda_i s_i - x \Big\rangle \ge \Big\langle f\Big(\sum_{i=1}^{n} \lambda_i s_i\Big), 0 \Big\rangle = 0.$$

Now apply Theorem 2 with $S = C$ and $q = \sup$ to infer that there is a mean l on $\mathbb{E}$ such that

$$l(\tau(s)) \ge 0, \qquad s \in C.$$

Since each $\tau(s)$ is a continuous affine function on C, it follows from Corollary 2 that there is $y_0 \in C$ such that $l(\tau(s)) = \tau(s)(y_0) = \langle f(s), s - y_0\rangle \ge 0$ for all s in C. Thus we have shown that there is y_0 in C such that

$$\langle f(x), y_0 - x\rangle \le 0, \qquad x \in C. \tag{16}$$

Now we use the same argument as in [2] to show that

$$\langle f(y_0), y_0 - x\rangle \le 0, \qquad x \in C. \tag{17}$$

For any x in C and $t \in [0,1]$, let $z_t = tx + (1-t)y_0$. Then $z_t \in C$ and hence $\langle f(z_t), y_0 - z_t\rangle \le 0$. But $\langle f(z_t), y_0 - z_t\rangle = \langle f(z_t), t(y_0 - x)\rangle = t\langle f(z_t), y_0 - x\rangle$ implies $\langle f(z_t), y_0 - x\rangle \le 0$. Let now $t \to 0$. By our assumption of the continuity of $f|_{L\cap C}$ for each line L in X we have

$$\langle f(y_0), y_0 - x\rangle = \lim_{t\to 0} \langle f(z_t), y_0 - x\rangle \le 0,$$

which completes our proof. ∎

References

[1] M. Altman, *A general separation theorem for mappings, saddle points, duality and conjugate functions*, Studia Math. 36 (1970), 131–167.

[2] J. Dugundji and A. Granas, *KKM maps and variational inequalities*, Ann. Scuola Norm. Sup. Pisa 5 (1978), 679–682.

[3] K. Fan, *Minimax theorems*, Proc. Nat. Acad. Sci. U.S.A. 39 (1953), 42–47.

[4] A. Granas and J. Dugundji, *Fixed Point Theory*, Springer, 2003.

[5] A. Granas et M. Lassonde, *Sur un principe géométrique en analyse convexe*, Studia Math. 101 (1991), 1–18.

[6] P. Hartman and G. Stampacchia, *On some nonlinear elliptic differential equations*, Acta Math. 115 (1966), 271–310.

[7] H. König, *Über das von Neumannsche Minimax-Theorem*, Arch. Math. (Basel) 19 (1968), 482–487.

[8] M. Landsberg and W. Schirotzek, *Mazur–Orlicz type theorems with applications*, Math. Nachr. 79 (1977), 331-341.

[9] F. C. Liu, *Measure solutions of systems of inequalities*, Topol. Methods Nonlinear Anal. 2 (1993), 317–331.

[10] S. Mazur et W. Orlicz, *Sur les espaces métriques linéaires (II)*, Studia Math. 13 (1953), 137–179.

[11] V. Pták, *On a theorem of Mazur and Orlicz*, ibid. 15 (1956), 365–366.

[12] R. Sikorski, *On a theorem of Mazur and Orlicz*, ibid. 13 (1953), 180–182.

Department of Mathematics
Tamkang University
Tamshui, Taiwan
E-mail: maliufc@math.sinica.edu.tw

Received October 17, 2007
Revised version May 20, 2008 (6231)