Topological Methods in Nonlinear Analysis
Journal of the Juliusz Schauder Center
Volume 2, 1993, 317–331

MEASURE SOLUTIONS OF SYSTEMS OF INEQUALITIES

Fon-Che Liu

(*Submitted by A. Granas*)

Dedicated to the memory of Juliusz Schauder

1 Introduction and a theorem of Mazur and Orlicz

Let X and Y be compact Hausdorff spaces, $\{f_i\}$ and $\{g_i\}$ two families of real-valued continuous functions defined on X and Y respectively and indexed by the same set I. We consider the question of existence of nonnegative Radon measures μ and ν on X and Y respectively and not both zero such that

$$\int_X f_i\,d\mu \le \int_Y g_i\,d\nu, \qquad i \in I.$$

Since $\mu = 0$ corresponds to the case that all f_i are zero functions, we may assume without loss of generality that $\mu \neq 0$, hence after normalization the problem becomes to find probability measures μ and ν on X and Y respectively and a nonnegative real number γ such that

$$\int_X f_i\,d\mu \le \gamma \int_Y g_i\,d\nu, \qquad i \in I, \tag{1.1}$$

where by measures we always mean Radon measures.

We consider first the case $\gamma = 1$ together with applications to minimax theorems and a duality theorem of Gale in the next section. The general case will then be treated in the third section by applying the result obtained for the

1991 *Mathematics Subject Classification.* Primary 46A22, 49N15; Secondary 49J35.

case $\gamma = 1$. An application to Lagrangian duality is then considered in the last section.

In the remaining part of this section we present the main tool of this work, which is a generalization (Theorem 1) of a theorem of Mazur and Orlicz [7, p. 174] (see also [8], [9]). The theorem of Mazur and Orlicz is a general form of the Hahn-Banach theorem. Our generalization is based on the following special case of the theorem of Mazur and Orlicz:

LEMMA 1. *Let τ be a map from a set S into a real vector space E and q a sublinear functional on E. Then the following two statements are equivalent:*

(I) *There is $\ell \in E^*$, the algebraic dual of E, with $\ell \leq q$ such that*

$$\ell(\tau(s)) \geq 0 \qquad \text{for all } \ s \in S.$$

(II) *$q(\sum_{i=1}^{n} \lambda_i \tau(s_i)) \geq 0$ for all n and all $s_1, \ldots, s_n$ in S, all $\lambda_1 \geq 0, \ldots, \lambda_n \geq 0$.*

For completeness we include a proof of Lemma 1 which relies on a simple geometric form of the Hahn-Banach theorem:

PROPOSITION 1. *Let E be a real vector space and C a linearly open convex cone not containing the zero element θ. Then there is a hyperplane H containing θ such that $H \cap C = \emptyset$.*

We recall that a set C in E is called *linearly open* if for every x in C and y in E, $x + ry$ is in C if $|r|$ is small. See [5] for a proof of Proposition 1 independent of the Hahn-Banach theorem.

Given a sublinear functional q on a real vector space E we shall use Q to denote the set $\{x \in E : q(x) < 0\}$.

PROPOSITION 2. *Let q be a sublinear functional on a real vector space E with $Q \neq \emptyset$, and let $\ell \in E^*$. Then the following two statements are equivalent;*

(A) *$\ell(x) \leq 0$ for all $x \in Q$.*

(B) *There is $\sigma > 0$ such that $\sigma\ell \leq q$ on E.*

PROOF. That (B) $\Rightarrow$ (A) is obvious. To show (A) $\Rightarrow$ (B) define a map t from E into $\mathbb{R}^2$ by

$$t(x) = (q(x), -\ell(x)), \qquad x \in E.$$

Let P be the convex hull of the image of E under the map t and $\mathbb{R}^2_-$ be the negative quadrant of $\mathbb{R}^2$, i.e. $\mathbb{R}^2_- = \{(r_1, r_2) \in \mathbb{R}^2 : r_1 < 0,\, r_2 < 0\}$. We claim that P and $\mathbb{R}^2_-$ are disjoint. Actually, if $v \in P$, then there are $x_1, \ldots, x_n$ in E and $\alpha_1 \geq 0, \ldots, \alpha_n \geq 0$ with $\sum_{i=1}^{n} \alpha_i = 1$ such that $v = (\sum_{i=1}^{n} \alpha_i q(x_i), -\ell(\sum_{i=1}^{n} \alpha_i x_i))$;

if $\sum_{i=1}^n \alpha_i q(x_i) < 0$, then $q(\sum_{i=1}^n \alpha_i x_i) \le \sum_{i=1}^n \alpha_i q(x_i) < 0 \Rightarrow \ell(\sum_{i=1}^n \alpha_i x_i) \le 0 \Rightarrow -\ell(\sum_{i=1}^n \alpha_i x_i) \ge 0 \Rightarrow v \notin \mathbb{R}^2_-$. This shows our claim. Hence by separation principle in $\mathbb{R}^2$ there is (α_1, α_2) in $\mathbb{R}^2$ with $\alpha_1^2 + \alpha_2^2 > 0$ such that

$$\alpha_1 r_1 + \alpha_2 r_2 \le \alpha_1 q(x) - \alpha_2 \ell(x) \qquad \forall (r_1, r_2) \in \mathbb{R}^2_- \text{ and } x \in E. \tag{1.2}$$

Since $\alpha_1 r_1 + \alpha_2 r_2$ is bounded from above for $(r_1, r_2) \in \mathbb{R}^2_-$, $\alpha_1 \ge 0$ and $\alpha_2 \ge 0$. Letting (r_1, r_2) tend to (0,0) in (1.2), we have

$$\alpha_2 \ell(x) \le \alpha_1 q(x) \qquad \forall x \in E,$$

from which since $Q \ne \emptyset$ we infer easily that both α_1 and α_2 are positive. We choose $\sigma = \alpha_2 \alpha_1^{-1}$ to finish the proof.

We are now ready to prove Lemma 1.

PROOF OF LEMMA 1. That (I) $\Rightarrow$ (II) is obvious. We show that (II) $\Rightarrow$ (I). If $Q = \emptyset$, (II) holds always and (I) holds with ℓ being the zero functional. So we may assume that $Q \ne \emptyset$. Since q is sublinear, Q is a linearly open convex cone. Let P be the convex cone generated by the image of S under τ. If (II) holds, then $P \cap Q = \emptyset$, hence if we let $C = \{x - y : x \in P,\ y \in Q\}$, then $\theta \notin C$. It is clear that C is a linearly open convex cone. By Proposition 1, there is ℓ in E^* such that $\ell(z) > 0$, for $z \in C$, i.e. for $x \in P$ and $y \in Q$ we have

$$\ell(rx) > \ell(y),\ \ell(x) > \ell(ry) \qquad \text{for } r > 0,$$

from which by letting r tend to zero we obtain

$$\ell(y) \le 0 \quad \forall y \in Q \qquad \text{and} \qquad \ell(x) \ge 0 \quad \forall x \in P. \tag{1.3}$$

Since $\ell(y) \le 0$ for $y \in Q$, by Proposition 2, there is $\sigma > 0$ such that $\sigma\ell \le q$ on E and if we rename $\sigma\ell$ to be ℓ then (1.3) becomes

$$\ell \le q \quad \text{on } E \qquad \text{and} \qquad \ell(x) \ge 0 \quad \forall x \in P,$$

which establishes (I). The lemma is thus proved.

The following generalization of the Mazur-Orlicz theorem is our main preliminary result:

THEOREM 1. *For $i = 1, 2$, let q_i be a sublinear functional on a real vector space E_i and let τ_i be a map from a set S into E_i. Then the following two statements are equivalent:*

(I) *There are ℓ_1 and ℓ_2 in E_1^* and E_2^* respectively with $\ell_1 \le q_1$ and $\ell_2 \le q_2$ such that*

$$\ell_1(\tau_1(s)) \le \ell_2(\tau_2(s)) \quad \forall s \in S. \tag{1.4}$$

(II) *For any positive integer n we have*

$$-q_1\Big(-\sum_{i=1}^{n}\lambda_i\tau_1(s_i)\Big) \le q_2\Big(\sum_{i=1}^{n}\lambda_i\tau_2(s_i)\Big) \tag{1.5}$$

for all $s_1, \ldots, s_n$ in S and all $\lambda_1 \ge 0, \ldots, \lambda_n \ge 0$.

PROOF. (I) $\Rightarrow$ (II) is immediate. To show (II) $\Rightarrow$ (I) we shall use Lemma 1. Let E be the vector space product of E_1 and E_2. Elements of E will be denoted by $\{u, v\}$ with $u \in E_1$ and $v \in E_2$. Let q be the sublinear functional on E defined by

$$q(\{u,v\}) = q_1(u) + q_2(v), \qquad u \in E_1,\ v \in E_2;$$

finally, define $\tau : S \to E$ by

$$\tau(s) = \{-\tau_1(s), \tau_2(s)\}, \qquad s \in S.$$

It follows from (II) that for any positive integer n we have

$$q\Big(\sum_{i=1}^{n}\lambda_i\tau(s_i)\Big) \ge 0$$

for all $s_1, \ldots, s_n$ in S and all $\lambda_1 \ge 0, \ldots, \lambda_n \ge 0$. Consequently by Lemma 1 there is $\ell \in E^*$ satisfying $\ell \le q$ on E such that $\ell(\tau(s)) \ge 0$ for all $s \in S$. Let ℓ be expressed by $\{\ell_1, \ell_2\}$ with $\ell_1 \in E_1^*$ and $\ell_2 \in E_2^*$ through

$$\ell(\{u,v\}) = \ell_1(u) + \ell_2(v), \qquad u \in E_1,\ v \in E_2.$$

It is clear that $\ell \le q$ is equivalent to $\ell_1 \le q_1$ on E_1 and $\ell_2 \le q_2$ on E_2. Now $\ell(\tau(s)) \ge 0$ for all $s \in S$ implies (1.4). Thus (I) holds and the proof is complete.

If we choose $E_1 = \mathbb{R}$ and let $q_1(t) = t$ for $t \in \mathbb{R}$, then we obtain the theorem of Mazur and Orlicz.

2. The main theorem (special case) with applications

In this section X, Y, $\{f_i\}$ and $\{g_i\}$ will always be as introduced in the last section without further notice. We shall use $C(S)$ to denote the vector space of all continuous real-valued functions defined on a topological space S.

We now take up the question of the existence of probability measures μ and ν on X and Y respectively such that (1.1) holds with $\gamma = 1$. Applications to minimax theorems and a duality theorem of Gale will also be given.

THEOREM 2. *The following two statements are equivalent:*

($*$) *There exist probability measures μ and ν on X and Y respectively such that*

$$\int_X f_i \, d\mu \leq \int_Y g_i \, d\nu, \qquad i \in I. \tag{2.1}$$

($**$) *For any positive integer n we have*

$$\operatorname*{Min}_{x\in X} \sum_{k=1}^{n} \lambda_k f_{i_k}(x) \leq \operatorname*{Max}_{y\in Y} \sum_{k=1}^{n} \lambda_k g_{i_k}(y) \tag{2.2}$$

for all $i_1, \dots, i_n$ in I and all $\lambda_1 \geq 0, \dots, \lambda_n \geq 0$.

PROOF. ($*$) $\Rightarrow$ ($**$) is immediate. We now use Theorem 1 to show ($**$) $\Rightarrow$ ($*$). Let $E_1 = C(X)$ and $E_2 = C(Y)$, define sublinear functionals q_1 and q_2 by

$$q_1(f) = \operatorname*{Max}_{x\in X} f(x), \quad f \in E_1; \qquad q_2(g) = \operatorname*{Max}_{y\in Y} g(y), \quad g \in E_2;$$

and finally let $S = I$ and define $\tau_j : I \to E_j$, $j = 1, 2$, by

$$\tau_1(i) = f_i, \quad i \in I; \qquad \tau_2(i) = g_i, \quad i \in I.$$

($**$) implies that the statement (II) of Theorem 1 holds. Hence (I) of Theorem 1 holds. Thus there are $\ell_1 \in E_1^*$ and $\ell_2 \in E_2^*$ with $\ell_1 \leq q_1$ and $\ell_2 \leq q_2$ such that

$$\ell_1(f_i) \leq \ell_2(g_i) \qquad \forall i \in I. \tag{2.3}$$

But $\ell_1 \leq q_1$ and $\ell_2 \leq q_2$ means that ℓ_1 and ℓ_2 are probability Radon measures, say μ and ν on X and Y respectively. Now (2.3) becomes (2.1) and statement ($*$) holds. The proof is complete.

We consider now applications of Theorem 2 to minimax theorems and a duality theorem of Gale [3]. We remark first that in the statement of Theorem 2, $\lambda_1 \geq 0, \dots, \lambda_n \geq 0$ can be chosen with $\lambda_1 + \dots + \lambda_n = 1$, i.e. $\lambda = (\lambda_1, \dots, \lambda_n)$ is in the standard $(n-1)$-simplex σ^{n-1} of $\mathbb{R}^n$.

THEOREM 3. *Let $\mathbf{F}$ be a family of continuous functions defined on a compact Hausdorff space X and let $[\mathbf{F}]$ be the convex hull of $\mathbf{F}$. If we denote by $P(X)$ the set of all probability measures on X, then*

$$\operatorname*{Min}_{\mu\in P(X)} \operatorname*{Sup}_{f\in\mathbf{F}} \int_X f \, d\mu = \operatorname{Sup} \left\{ \operatorname*{Min}_{x\in X} g(x) : g \in [\mathbf{F}] \right\}. \tag{2.4}$$

PROOF. Since for each $\mu \in P(X)$ the inequality

$$\operatorname*{Sup}_{f\in\mathbf{F}} \int_X f \, d\mu \geq \operatorname{Sup} \left\{ \operatorname*{Min}_{x\in X} g(x) : g \in [\mathbf{F}] \right\} \tag{2.4}$$

clearly holds, it remains only to show that there is $\mu \in P(X)$ such that

$$\sup_{f\in \mathbf{F}} \int_X f\,d\mu \le \operatorname{Sup}\left\{\operatorname*{Min}_{x\in X} g(x) : g \in \mathbf{F}\right\}. \tag{2.5}$$

We may assume that the right hand side of (2.5) is a finite number L. By applying Theorem 2 with Y being any compact Hausdorff space and each g_i being the constant function with value L, we need only verify that for any positive integer n and all $(\lambda_1, \dots, \lambda_n)$ in σ^{n-1} the following inequality holds:

$$\operatorname*{Min}_{x\in X} \sum_{k=1}^{n} \lambda_k f_{i_k}(x) \le L$$

for all $i_1, \dots, i_n$ in I. But this is obvious from the definition of L and the fact that the function $\sum_{k=1}^{n} \lambda_k f_{i_k}$ is in $[\mathbf{F}]$. This proves (2.5) and completes the proof.

When $\mathbf{F}$ is a finite family and X is a finite set we obtain from Theorem 3 the following well-known minimax theorem of von Neumann for bilinear forms [10]:

THEOREM 4. *Let (a_{ij}), $i = 1, \dots, m$, $j = 1, \dots, n$, be an $m \times n$ matrix and denote by $\alpha = (\alpha_1, \dots, \alpha_m)$, and $\beta = (\beta_1, \dots, \beta_n)$ generic elements of σ^{m-1} and σ^{n-1} respectively. Then*

$$\operatorname*{Min}_{\alpha} \operatorname*{Max}_{\beta} \sum a_{ij}\alpha_i\beta_j = \operatorname*{Max}_{\beta} \operatorname*{Min}_{\alpha} \sum a_{ij}\alpha_i\beta_j.$$

Suppose now we have a finite number of functions $f_1, \dots, f_n$ defined on an arbitrary set S. Let $X = \{1, \dots, n\}$ and let $\mathbf{F}$ be the family $\{g_s\}$ of functions defined on X and indexed by S with $g_s(i) = f_i(s)$, for $s \in S$ and $i \in X$. Then the following corollary is a direct consequence of Theorem 3:

COROLLARY 1. *Let $f_1, \dots, f_n$ be real-valued functions defined on a set S, and let $\alpha = (\alpha_1, \dots, \alpha_n)$ be a generic element of the standard $(n-1)$-simplex in $\mathbb{R}^n$. Then*

$$\operatorname*{Min}_{\alpha} \operatorname*{Sup}_{s\in S} \sum_{j=1}^{n} \alpha_j f_j(s) = \operatorname{Sup} \operatorname*{Min}_{1\le j\le n} \sum_{i=1}^{m} \beta_i f_j(s_i),$$

where the supremum on the right hand side is taken over all possible finite sets $\{s_1, \dots, s_m\}$ and all possible $\beta = (\beta_1, \dots, \beta_m) \in \sigma^{m-1}$ with m running through all positive integers.

We now show that Corollary 1 implies the following minimax theorem of Ky Fan [1]:

THEOREM 5. *Let* $\mathbf{F}$ *be a concave family of lower semicontinuous convex functions on a compact convex subset* X *of a Hausdorff topological vector space. Then*

$$\operatorname*{Min}_{x\in X}\operatorname*{Sup}_{f\in\mathbf{F}} f(x) = \operatorname*{Sup}_{f\in\mathbf{F}}\operatorname*{Min}_{x\in X} f(x).$$

PROOF. Since $\operatorname{Min}_{x\in X}\operatorname{Sup}_{f\in\mathbf{F}} f(x) \geq \operatorname{Sup}_{f\in\mathbf{F}}\operatorname{Min}_{x\in X} f(x)$ holds always, we need only show

$$\operatorname*{Min}_{x\in X}\operatorname*{Sup}_{f\in\mathbf{F}} f(x) \leq \operatorname*{Sup}_{f\in\mathbf{F}}\operatorname*{Min}_{x\in X} f(x). \tag{2.6}$$

For this purpose we may assume that the right hand side of (2.6) is $< +\infty$. Let λ be any finite number greater than the right hand side of (2.6) and for each $f \in \mathbf{F}$ let

$$A_f = \{x \in X : f(x) \leq \lambda\}.$$

Consider an arbitrary finite subfamily $\{f_1, \dots, f_n\}$ of $\mathbf{F}$. Since $\mathbf{F}$ is concave and $\lambda - f_1, \dots, \lambda - f_n$ are concave upper semicontinuous, it follows from Corollary 1 that

$$0 \leq \operatorname*{Inf}_{f\in\mathbf{F}}\operatorname*{Max}_{x\in X}[\lambda - f(x)] \leq \operatorname*{Min}_{\alpha}\operatorname*{Max}_{x\in X}\left[\lambda - \sum_{j=1}^{n}\alpha_j f_j(x)\right]$$

$$= \operatorname*{Min}_{\alpha}\operatorname*{Max}_{x\in X}\sum_{j=1}^{n}\alpha_j[\lambda - f_j(x)] \leq \operatorname*{Max}_{x\in X}\operatorname*{Min}_{1\leq j\leq n}[\lambda - f_j(x)],$$

which means that $\bigcap_j A_{f_j}$ is non empty. As each A_f is compact, it follows that $\bigcap_{f\in\mathbf{F}} A_f$ is non empty. Hence $\operatorname{Min}_{x\in X}\operatorname{Sup}_{f\in\mathbf{F}} f(x) \leq \lambda$. Since λ is an arbitrary number greater than the right hand side of (2.6), (2.6) is proved. The proof is complete.

Now we consider an application of Theorem 2 to a duality theorem of Gale. Let M be a non empty subset of a normed vector space E and g a real-valued function defined on M. For $x \in M$, define a possibly infinite number $\sigma(x, g; M)$ by

$$\sigma(x, g; M) = \operatorname{Inf}\{s \geq 0 : g(y) - g(x) \leq s \cdot \|y - x\|,\ y \in M\},$$

where $\|\cdot\|$ denotes the norm on E as well as that on the topological dual E' of E. We note that $\sigma(x, g; M)$ is the larger of 0 and $\operatorname{Sup}_{y\in M, y\neq x}[g(y) - g(x)]/\|y - x\|$.

THEOREM 6 [3]. *Let* M *be a convex subset of a normed vector space* E *and* g *a concave function on* M*. Then for any* x *in* M *we have*

$$\sigma(x, g; M) = \operatorname{Min}\{\|\xi\| : \xi \in E',\ g(y) - g(x) \leq \xi(y - x),\ y \in M\}. \tag{2.7}$$

PROOF. If there is ξ in E' such that $g(y) - g(x) \leq \xi(y - x)$ for all y in M, then obviously the left hand side of (2.7) is $\leq \|\xi\|$. Hence we may assume that the left hand side of (2.7) is finite. It then remains to show that when $\sigma(x, g; M) > 0$

there is $\xi \in E'$ with $\|\xi\| \le \sigma(x, g; M)$ such that $g(y) - g(x) \le \xi(y - x)$ for all y in M. Let $K = \{\xi \in E' : \|\xi\| \le \sigma(x, g; M)\}$; then K is w^*-compact. Let $\{f_y\}_{y \in M}$ with $f_y = g(y) - g(x)$ be a family of constants and let $\{h_y\}_{y \in M}$ be a family of w^*-continuous functions on K, where each h_y is defined by

$$h_y(\xi) = \xi(y - x), \qquad \xi \in K.$$

Then for any subset $\{y_1, \dots, y_n\}$ of M and any $(\lambda_1, \dots, \lambda_n)$ in σ^{n-1} we have

$$\begin{aligned}\sum_{i=1}^{n} \lambda_i f_{y_i} &\le g(\sum_{i=1}^{n} \lambda_i y_i) - g(x) \le \sigma(x, g; M) \Big\| \sum_{i=1}^{n} \lambda_i y_i - x \Big\| \\ &= \sigma(x, g; M) \cdot \operatorname*{Max}_{\|\xi\|=1} \xi\Big(\sum_{i=1}^{n} \lambda_i y_i - x \Big) = \operatorname*{Max}_{\xi \in K} \sum_{i=1}^{n} \lambda_i h_{y_i}(\xi),\end{aligned}$$

from which we infer from Theorem 2 that there is a probability measure μ on K such that

$$g(y) - g(x) = f_y \le \int_K h_y \, d\mu = \int_K \xi(y - x) \, d\mu(\xi)$$

for all $y \in M$. Since $\eta(z) = \int_K \xi(z) \, d\mu(\xi)$, $z \in E$, is obviously linear and since

$$|\eta(z)| \le \|z\| \int_K \|\xi\| \, d\mu(\xi) \le \sigma(x, g; M) \|z\|,$$

we know that $\eta \in E'$ with $\|\eta\| \le \sigma(x, g; M)$ and $g(y) - g(x) \le \eta(y - x)$ for all $y \in M$. This completes the proof.

3. The main theorem (general case)

We now consider the general problem introduced in the first section. We denote by $\mathbf{F}$ and $\mathbf{G}$ the families of functions $\{f_i\}$ and $\{g_i\}$ respectively and by $[\mathbf{F}]$, $[\mathbf{G}]$ their convex hulls. We also define $\mathbf{G}^-$ and $\mathbf{G}^+$ as follows:

$$\mathbf{G}^- = \Big\{ g \in [\mathbf{G}] : \operatorname*{Max}_{y \in Y} g(y) < 0 \Big\};$$

$$\mathbf{G}^+ = \Big\{ g \in [\mathbf{G}] : \operatorname*{Max}_{y \in Y} g(y) > 0 \Big\}.$$

For notational convenience, for $g \in [\mathbf{G}]$ with $g = \sum_{k=1}^{n} \lambda_k g_{i_k}$, $(\lambda_1, \dots, \lambda_n) \in \sigma^{n-1}$ and $\{i_1, \dots, i_n\} \subset I$, we denote by f_g the function $\sum_{k=1}^{n} \lambda_k f_{i_k} \in [\mathbf{F}]$. In this way different expressions of convex combinations of functions in $\mathbf{G}$ of the same function will be considered as giving different functions; in other words, we mean by $g \in [\mathbf{G}]$ a function with a given expression of the form $\sum_{k=1}^{n} \lambda_k g_{i_k}$. It follows from Theorem 2 that the existence of probability measures μ, ν on X and Y respectively and $\gamma \ge 0$ such that (1.1) holds is equivalent to the following condition:

(3.1) $\qquad$ For all $g \in [\mathbf{G}]$, $\qquad \operatorname*{Min}_{x \in X} f_g(x) \le \gamma \operatorname*{Max}_{y \in Y} g(y),$

from which it follows in particular that

$$\text{(3.2)} \qquad \operatorname*{Min}_{x\in X} f_g(x) \le 0 \text{ whenever } \operatorname*{Max}_{y\in Y} g(y) \le 0, \text{ if } g \in [\mathbf{G}].$$

If we let

$$\gamma_- = \operatorname*{Inf}_{g\in\mathbf{G}^-} \operatorname{Max} f_g/\operatorname{Max}\ g\}; \qquad \gamma_+ = \operatorname*{Sup}_{g\in\mathbf{G}^+}\{\operatorname{Min} f_g/\operatorname{Max}\ g\},$$

we observe readily from (3.1) that

$$\text{(3.3)} \qquad \gamma_+ \le \gamma \le \gamma_-.$$

In the foregoing definitions of γ_- and γ_+, we have compressed $\operatorname{Min}_{x\in X} f_g(x)$ and $\operatorname{Max}_{y\in Y} g(y)$ to Min f_g and Max g respectively and adopted the usual convention that the infimum over an empty set is $+\infty$, while the supremum over an empty set is $-\infty$. It is clear that if (3.2) and (3.3) hold with γ being a finite real number not necessarily nonnegative, then (3.1) holds. Hence a first natural question is the validity of the inequality $\gamma_+ \le \gamma_-$. Actually we have:

LEMMA 2. *If* (3.2) *holds, then* $\gamma_- \ge 0$ *and* $\gamma_+ \le \gamma_-$.

PROOF. That $\gamma_- \ge 0$ follows directly from (3.2). If either $\mathbf{G}^-$ or $\mathbf{G}^+$ is empty, $\gamma_+ \le \gamma_-$ holds trivially. We may assume therefore that both $\mathbf{G}^-$ and $\mathbf{G}^+$ are non empty. Let Z be the disjoint union of X and Y and let $\mathbf{H} = \{h_i\}$, $i \in I$, be the family of continuous functions on Z defined by $h_i(z) = f_i(z)$ or $-g_i(z)$ according as z is in X or in Y. Then (3.2) implies that Min $h \le 0$ for all $h \in [\mathbf{H}]$, where $[\mathbf{H}]$ is the convex hull of $\mathbf{H}$. We infer from Theorem 2 that there is a probability measure μ on Z such that

$$\int_Z h_i\, d\mu \le 0 \qquad \text{for all } i \in I.$$

We claim that $\mu(X) > 0$. Otherwise, $\mu(Y) = 1$ and

$$\int_Y g_i\, d\mu \ge 0 \qquad \text{for all } i \in I,$$

which means by Theorem 2 that Max $g \ge 0$ for all $g \in [\mathbf{G}]$, contradicting the assumption that $\mathbf{G}^- \neq \emptyset$. Similarly, $\mu(Y) > 0$. Thus we may write $\int_Z h_i\, d\mu \le 0$, $i \in I$, as

$$\int_X f_i\, d\mu/\mu(X) \le [\mu(Y)/\mu(X)] \int_Y g_i\, d\mu/\mu(Y), \qquad i \in I,$$

which implies again by Theorem 2 that

$$\operatorname{Min} f_g \le [\mu(Y)/\mu(X)] \operatorname{Max} g$$

for all $g \in [\mathbf{G}]$, from which we infer that $\gamma_+ \le \mu(Y)/\mu(X) \le \gamma_-$. This proves the lemma.

If now (3.2) holds and $\gamma_+ < +\infty$, then for all finite γ such that $\gamma_+ \leq \gamma \leq \gamma_-$ the condition (3.1) holds. Thus from Theorem 2 results the main theorem of this note:

THEOREM 7. *If* (3.2) *holds, then the following two statements are equivalent:*

(A) *There are probability measures μ, ν on X, Y respectively and a real number γ such that*

$$\int_X f_i \, d\mu \leq \gamma \int_Y g_i \, d\nu$$

for all $i \in I$.

(B) $\gamma_+ < +\infty$.

Furthermore, if (B) *holds, then* (A) *holds for γ being any real number between γ_+ and γ_- and with μ and ν depending on γ.*

REMARKS. (i) Condition (3.2) is a consequence of (1.1) for $\gamma \geq 0$, but under condition (3.2), if $\gamma_+ < 0$ we can also settle our problem for requiring γ to be negative. (ii) Since (3.2) implies $\gamma_- \geq 0$, if (3.2) holds and $\gamma_+ < +\infty$, our general problem always has solutions when restricted to $\gamma \geq 0$. (iii) In the proof of Lemma 2, we have shown that $\gamma_+ < +\infty$ if $\mathbf{G}^- \neq \emptyset$. Hence (A) holds if $\mathbf{G}^- \neq \emptyset$. One recognizes readily that $\mathbf{G}^- \neq \emptyset$ is a generalized form of constraint qualification condition of Slater in optimization theory (see [6] and [4]).

4. Lagrangian duality

We now consider an application of Theorem 7 to Lagrangian duality. We refer to [4] and the references there for general information on dualities in mathematical programming. Let X be a nonempty set and let Y and Z be compact Hausdorff spaces. Consider functions f and g defined on $X \times Y$ and $X \times Z$ respectively and satisfying the condition that the functions $f(x, \cdot)$ and $g(x, \cdot)$ are continuous on Y and Z respectively for each $x \in X$. We consider the following primary problem:

$$\text{(PP)} \qquad \operatorname*{Min}_{y \in Y} f(x, y) \to \text{Sup}, \qquad \text{for } x \in N,$$

where $N = \{x \in X : g(x, z) \leq 0 \ \forall z \in Z\}$. If $N \neq \emptyset$, the problem (PP) is called *feasible*. Let

$$L(\mu, \nu, \gamma; x) = \int_Y f(x, y) \, d\mu(y) - \gamma \int_Z g(x, z) \, d\nu(z), \tag{4.1}$$

where $\mu \in P(Y)$, $\nu \in P(Z)$ and $\gamma \geq 0$. $L(\mu, \nu, \gamma; x)$ is the Lagrangian function of our problem (PP). The following problem is called the *Lagrangian dual problem* of (PP):

(LDP) $$\operatorname*{Sup}_{x\in X} L(\mu,\nu,\gamma;x) \to \text{Inf},$$

where the infimum is taken over all $\mu \in P(Y)$, $\nu \in P(Z)$ and $\gamma \geq 0$.

Let u, v be the values respectively of problems (PP) and (LDP), i.e.

$$u = \operatorname*{Sup}_{x\in N} \operatorname*{Min}_{y\in Y} f(x,y);$$

$$v = \operatorname{Inf}\left\{\operatorname*{Sup}_{x\in X} L(\mu,\nu.\gamma;x) : \mu \in P(Y),\ \nu \in P(Z),\ \gamma \geq 0\right\};$$

then it is obvious that

$$u \leq v. \tag{4.2}$$

We now introduce another problem which is closely related to problem (PP). For this purpose let us denote by $P_f(X)$ the space of all probability measures supported on finite subsets of X, i.e. if $p \in P_f(X)$, then there is a finite subset $\{x_1,\ldots,x_n\}$ of X such that $p(x_i) \equiv p(\{x_i\}) > 0$, $i = 1,\ldots,n$, and $\sum_{i=1}^n p(x_i) = 1$. Since each element x of X can be identified with the probability measure supported at x, X is imbedded in $P_f(X)$. For a function h on X we shall denote the integral $\int h dp$ of h with respect to $p \in P_f(X)$ by $h(p)$. If $p \in P_f(X)$ is supported at an element x of X, then $h(p) = h(x)$; hence this convenient notation is consistent with the usual notation for function values and does not cause ambiguity. The following problem is called the *extended problem* of (PP):

(EPP) $$\operatorname*{Min}_{y\in Y} f(p,y) \to \text{Sup}, \qquad p \in N^*,$$

where $N^* = \{p \in P_f(X) : g(p,z) \leq 0\ \forall z \in Z\}$; the value of (EPP) is denoted by u^*, i.e. $u^* = \operatorname{Sup}_{p\in N^*} \operatorname{Min}_{y\in Y} f(p,y)$. Since X is imbedded in $P_f(X)$, we have $N \subset N^*$ and hence

$$u \leq u^*. \tag{4.3}$$

We also have the following obvious lemma for equality of u and u^*.

LEMMA 3. *Let X be a convex set in a real vector space and suppose that for each y in Y and each z in Z, $x \to f(x,y)$ is concave and $x \to g(x,z)$ is convex on X. Then $u = u^*$.*

We recapitulate now the results of the previous section by replacing I, X, and Y there by X, Y, and Z of this section. For a given finite number r we consider the following conditions:

(4.4) $\exists \mu \in P(Y)$, $\nu \in P(Z)$, and $\gamma \geq 0$ such that $L(\mu,\nu,\gamma;x) \leq r\ \forall x \in X$;

$$\operatorname*{Max}_{y\in Y} f(p,y) \le r \quad \text{for } p \in N^*. \tag{4.5}$$

If we choose $\mathbf{F}$ and $\mathbf{G}$ of Section 2 to be the families $\{f(x,\cdot)-r\}_{x\in X}$ and $\{g(x,\cdot)\}_{x\in X}$, then conditions (4.4) and (4.5) are the conditions (1.1) and (3.2) respectively, hence (4.4) $\Rightarrow$ (4.5).

LEMMA 4. (i) *If* $r < u^*$, *then* $\operatorname{Sup}_{x\in X} L(\mu,\nu,\gamma;x) > r \ \forall \mu \in P(Y)$, $\nu \in P(Z)$, *and* $\gamma \ge 0$; (ii) $u^* \le v$.

PROOF. Since (i) $\Rightarrow$ (ii), we need only prove (i). If $r < u^*$, (4.5) does not hold and hence (4.4) does not hold either, and thus (i) follows.

COROLLARY 2. *If* $u^* = +\infty$, *then any* $\mu \in P(Y)$, $\nu \in P(Z)$, *and* $\gamma \ge 0$ *solve the problem* (LDP). *In particular, if* $u = +\infty$, *then* (LDP) *has all* $\mu \in P(Y)$, $\nu \in P(Z)$, *and* $\gamma \ge 0$ *as solutions.*

We now assume that u^* is a finite number and let $r = u^*$. Then (4.5) holds and it follows from Theorem 7 that there exist μ in $P(Y)$, ν in $P(Z)$ and $\gamma \ge 0$ such that

$$\operatorname*{Sup}_{x\in X} L(\mu,\nu,\gamma;x) \le u^* \tag{4.6}$$

if and only if

$$\operatorname*{Sup}_{p\in CN^*} [\operatorname*{Min}_{y\in Y} f(p,y) - u^*]/[\operatorname*{Max}_{z\in Z} g(p,z)] < +\infty, \tag{4.7}$$

where $CN^* = P_f(X)\backslash N^*$.

But if (4.6) holds, then it follows from Lemma 4
(i) that

$$\operatorname*{Sup}_{x\in X} L(\mu,\nu,\gamma;x) = u^*.$$

We have thus proved the main theorem of this section:

THEOREM 8. *If the problem* (PP) *is feasible and* $u^* < +\infty$, *then* (LDP) *has a solution and* $v = u^*$ *if and only if* (4.7) *holds.*

REMARK. If $CN^* = \emptyset$, then (4.7) holds trivially and (LDP) has a solution with $\gamma = 0$ if (PP) is feasible and $u^* < +\infty$, i.e. there is $\mu_0 \in P(Y)$ such that

$$\operatorname*{Sup}_{p\in N^*} \operatorname*{Min}_{y\in Y} f(p,y) = \operatorname*{Sup}_{x\in X} \int_Y f(x,y)\, d\mu_0(y).$$

This is exactly formula (2.4), hence, if we choose $g(x,z)$ to be the zero function, we recover Theorem 3. Thus the form we choose for problem (PP) allows the minimax principles to be included in our framework.

There is one useful situation for which (4.7) holds, namely, when the following generalized Slater condition holds (see Remark (iii) following Theorem 7):

(GSC) There exists $p_0 \in P_f(X)$ such that $g(p_0, z) < 0 \ \forall z \in Z$.

Thus the following theorem holds:

THEOREM 9. *If* $-\infty < u^* < +\infty$, *and* (GSC) *holds, then there are* $\mu_0 \in P(Y)$, $\nu_0 \in P(Z)$, *and* $\gamma_0 \geq 0$ *such that*

$$\operatorname*{Sup}_{x \in X} L(\mu_0, \nu_0, \gamma_0; x) = u^*.$$

COROLLARY 3. *Assume that X is a convex set in a real vector space and functions f and g satisfy the conditions of Lemma* 3. *Suppose that* (PP) *is feasible, $u < +\infty$, and that the Slater condition holds, i.e. there is x_0 in X such that $g(x_0, z) < 0$ for $z \in Z$. Then there are $\mu_0 \in P(Y)$, $\nu_0 \in P(Z)$, and $\gamma_0 \geq 0$ such that*

$$u = \operatorname*{Sup}_{x \in X} L(\mu_0, \nu_0, \gamma_0; x).$$

To look more closely at the condition (4.7) we introduce perturbed problems of (EPP) in the following way. For $\varepsilon > 0$ let $N^*(\varepsilon) = \{p \in P_f(X) : g(p, z) \leq \varepsilon \ \forall z \in Z\}$ and define $u^*(\varepsilon)$ by

$$u^*(\varepsilon) = \operatorname*{Sup}_{p \in N^*(\varepsilon)} \operatorname*{Min}_{y \in Y} f(p, y).$$

If u^* is a finite number and $\operatorname{Sup}_{\varepsilon > 0} \varepsilon^{-1}[u^*(\varepsilon) - u^*] < +\infty$, the problem (EPP) will be said to be *stable*. For the role of stability in mathematical programming we refer to [4]. Since the left hand side of (4.7) is bounded by $\operatorname{Sup}_{\varepsilon > 0} \varepsilon^{-1}[u^*(\varepsilon) - u^*]$, we have the following corollary of Theorem 8:

COROLLARY 4. *Suppose that the problem* (PP) *is feasible and $u^* < +\infty$. If* (EPP) *is stable, then* (LDP) *has a solution and $v = u^*$.*

We now consider an application of Corollary 4. Let X be a real vector space, $\xi_1, \ldots, \xi_n, \zeta_1, \ldots, \zeta_m$ be linear functionals on X, and let $b_1, \ldots, b_n, c_1, \ldots, c_m$ be real constants. Then the following duality holds:

THEOREM 10. *Assume that* $N = \{x \in X : \xi_i(x) \leq b_i,\ i = 1, \ldots, n\} \neq \emptyset$. *Then*

$$\operatorname*{Sup}_{x \in N} \operatorname*{Min}_{1 \leq j \leq m} [\zeta_j(x) + c_j] = \operatorname{Min} \left\{ \sum_{j=1}^{m} \alpha_j c_j + \sum_{i=1}^{n} \lambda_i b_i \right\}, \tag{4.8}$$

where the minimum is taken over all $\alpha = (\alpha_1, \dots, \alpha_m)$ *in* σ^{m-1} *and all* $\lambda_1 \geq 0, \dots, \lambda_n \geq 0$ *satisfying* $\sum_{j=1}^m \alpha_j \zeta_j = \sum_{i=1}^n \lambda_i \xi_i$.

PROOF. Let $Y = \{1, \dots, m\}$, $Z = \{1, \dots, n\}$ and define f and g by

$$f(x, j) = \zeta_j(x) + c_j, \quad x \in X,\ 1 \leq j \leq m;$$

$$g(x, i) = \xi_i(x) - b_i, \quad x \in X,\ 1 \leq i \leq n.$$

By Lemma 3, (EPP) is the same as (PP). If the left hand side of (4.8) is $+\infty$, by Corollary 2,

$$\sup_{x \in X} \left\{ \sum_{j=1}^m \alpha_j [\zeta_j(x) + c_j] - \sum_{i=1}^n \lambda_i [\xi_i(x) - b_i] \right\} = +\infty \tag{4.9}$$

for all $\alpha = (\alpha_1, \dots, \alpha_m)$ in σ^{m-1} and all $\lambda_1 \geq 0, \dots, \lambda_n \geq 0$. Since X is a vector space, (4.9) holds if and only if the left hand side of (4.9) is not a constant, i.e. $\sum_{j=1}^m \alpha_j \zeta_j \neq \sum_{i=1}^n \lambda_i \xi_i$, hence the minimum on the right hand side of (4.8) is taken over empty set and is $+\infty$. We may assume then that the left hand side of (4.8) is a finite number. In this situation it is clear that the problem (PP) is stable. Thus by Corollary 4 there are $\alpha = (\alpha_1, \dots, \alpha_m)$ in σ^{m-1} and $\lambda_1 \geq 0, \dots, \lambda_n \geq 0$ such that

$$\sup_{x \in N} \min_{1 \leq j \leq m} [\zeta_j(x) + c_j] = \sup_{x \in X} \left\{ \sum_{j=1}^m \alpha_j [\zeta_j(x) + c_j] - \sum_{i=1}^n \lambda_i [\xi_i(x) - b_i] \right\}. \tag{4.10}$$

But for the right hand side of (4.10) to be finite it is necessary that

$$\sum_{j=1}^m \alpha_j [\zeta_j(x) + c_j] - \sum_{i=1}^n \lambda_i [\xi_i(x) - b_i] = \text{constant},$$

i.e. $\sum_{j=1}^m \alpha_j \zeta_j = \sum_{i=1}^n \lambda_i \xi_i$; this proves the theorem.

We remark here that when $m = 1$, Theorem 10 is the duality theorem in linear programming; while if $b_1 = \dots = b_n = c_1 = \dots = c_m = 0$, Theorem 10 leads to a generalization of a lemma of Farkas [2]:

COROLLARY 5 (Generalized Farkas lemma). *Let* $\xi_1, \dots, \xi_n, \zeta_1, \dots, \zeta_m$ *be linear functionals on a real vector space* X. *If for* $x \in X$, $\xi_i(x) \leq 0$, $1 \leq i \leq n$ *imply* $\operatorname{Min}_{1 \leq j \leq m} \zeta_j(x) \leq 0$, *then the cone generated by* $\xi_1, \dots, \xi_n$ *contains a convex combination of* $\zeta_1, \dots, \zeta_m$.

Acknowledgements. I want to thank Andrzej Granas for discussions that lead to the final form of this note.

REFERENCES

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

[2] J. FARKAS, *Theorie der einfachen Ungleichungen*, J. Reine Angew. Math. **124** (1902), 1–27.

[3] D. GALE, *A geometric duality theorem with economic applications*, Rev. Economic Studies **34** (1967), 19–24.

[4] A. M. GEOFFRION, *Duality in nonlinear programming: a simplified applications-oriented development*, SIAM Rev: **13** (1971), 1–37.

[5] P. J. HIGGINS, *An Introduction to Topological Groups*, Cambridge University Press, 1974.

[6] O. L. MANGASARIAN, *Nonlinear Programming*, McGraw-Hill, New York, 1969.

[7] S. MAZUR ET W. ORLICZ, *Sur les espaces métriques linéaires* (II), Studia Math. **13** (1953), 137–179.

[8] V. PTÁK, *On a theorem of Mazur and Orlicz*, Studia Math. **15** (1956), 365–366.

[9] R. SIKORSKI, *On a theorem of Mazur and Orlicz*, Studia Math. **13** (1953), 180–182.

[10] J. VON NEUMANN AND O. MORGENSTERN, *The Theory of Games and Economic Behavior*, Princeton University Press, 1944.

Manuscript received October 10, 1992

FON-CHE LIU
Institute of Mathematics
Academia Sinica, Taipei
Taipei, R.O.C

E-mail address: maliufc@ ccvax.sinica.edu.tw