Michael M. Neumann

On the Mazur-Orlicz theorem

Czechoslovak Mathematical Journal, Vol. 41 (1991), No. 1, 104–109

Persistent URL: http://dml.cz/dmlcz/102439

Czechoslovak Mathematical Journal, 41 (116) 1991, Praha

ON THE MAZUR-ORLICZ THEOREM

MICHAEL M. NEUMANN, Mississippi State

(Received February 19, 1990)

INTRODUCTION

The classical interpolation theorem due to Mazur and Orlicz [11] and several variants thereof have become a useful tool in different branches of analysis and its applications. In this note, we discuss a certain refinement of the Mazur-Orlicz theorem, where the usual convexity assumptions are considerably weakened. This sandwich type theorem dates back to König [6], [7] where a certain preliminary version turned out to be useful for applications to minimax theorems and to function algebras. The by now most general theorem of this type seems to be due to Fuchssteiner and König [1], p. 255 and can also be found in the monograph [8], p. 33. Unfortunately, the technicalities for the original proof of this theorem are somewhat involved and not convincingly simple. In this note, we present a completely different approach to this sandwich type theorem, which we believe to give some new insight into this theory. Moreover, our methods even work in the general context of vector-valued operators rather than real-valued functionals; therefore, the main results of this note also subsume and improve the Hahn-Banach type theorems recently found by Martellotti and Salvadori [10]. The basic idea of the present approach is to combine some elementary approximation arguments with Pták's beautiful idea of a suitable auxiliary mapping [14], which turned out to be the most natural tool for the proof of the classical version of the Mazur-Orlicz theorem.

THE GENERALIZED MAZUR-ORLICZ THEOREM

Throughout this note, let E be a real vector space, and let F denote a real vector lattice, which is assumed to be Dedekind complete in the sense that each upper bounded subset has a supremum. The positive cone of F will be denoted by F_+. We add a smallest element $-\infty$ to the space F and extend the algebraic operations from F to $F_* := F \cup \{-\infty\}$ in the usual way, i.e. we define $0 \cdot (-\infty) := 0$, $t \cdot (-\infty) := -\infty$ for all real $t > 0$, and $x + (-\infty) := -\infty$ for all $x \in F$. An

Research supported by NSF Grant DMS 90-96108.

operator $\vartheta: E \to F$ is called *sublinear* if $\vartheta(u + v) \leqq \vartheta(u) + \vartheta(v)$ and $\vartheta(tu) = t\,\vartheta(u)$ holds for all $u, v \in E$ and all real $t \geqq 0$. The most primitive version of the vector-valued Hahn-Banach theorem states that every sublinear operator $\vartheta: E \to F$ dominates at least one linear operator $\varphi: E \to F$ in the sense of $\varphi(u) \leqq \vartheta(u)$ for all $u \in E$; see for instance [13]. The following theorem is, of course, a much stronger version of this result.

Theorem. *Let $\vartheta: E \to F$ be a sublinear operator, and consider an arbitrary mapping $\varrho: K \to F$ on a nonempty subset K of E such that $\varrho \leqq \vartheta$ on K. Moreover, assume that for some pair of real numbers $\alpha, \beta > 0$ and some $u \in F_+$ the following condition is fulfilled*:

(1) *For all $x, y \in K$ and all $\varepsilon > 0$ there exists some $z \in K$ such that* $\vartheta(z - \alpha x - \beta y) \leqq \varrho(z) - \alpha\,\varrho(x) - \beta\,\varrho(y) + \varepsilon u$.

Then there exists a linear operator $\varphi: E \to F$ such that $\varrho \leqq \varphi$ on K and $\varphi \leqq \vartheta$ on E.

The proof of this theorem will be given in the next section. We first discuss the main assumption (1) and some variants thereof. Condition (1) may be considered as a weak convexity assumption: if the set K is convex and the function $\varrho: K \to F$ is concave, then this condition is obviously fulfilled with an arbitrary choice of real numbers $\alpha, \beta > 0$ satisfying $\alpha + \beta = 1$ and even with the trivial choice $u = 0$. Actually, in this particular case the preceding theorem is easily seen to be an equivalent version of the vector-valued Mazur-Orlicz theorem, which will be reproduced below as lemma 1; see [11], p. 147 and [13], p. 79. Moreover, in the classical case $F = \mathbb{R}$, condition (1) is obviously equivalent to the following König type convexity condition [1], [6], [7], [8]: there exists a pair of real numbers $\alpha, \beta > 0$ such that

(2) $$\inf\{\vartheta(z - \alpha x - \beta y) - \varrho(z) + \alpha\,\varrho(x) + \beta\,\varrho(y) : z \in K\} \leqq 0 \quad \text{for all } x, y \in K.$$

Actually, even in the vector-valued setting condition (1) always implies (2), since it is well-known and easily seen [3] that the order of a Dedekind complete vector lattice is Archimedean. On the other hand, condition (1) is obviously fulfilled, if for some pair of real numbers $\alpha, \beta > 0$ we have:

(3) For all $x, y \in K$ there is some $z \in K$ such that $\vartheta(z - \alpha x - \beta y) \leqq \varrho(z) - \alpha\,\varrho(x) - \beta\,\varrho(y)$.

To illustrate the role of the vector $u \in F_+$ in the general version of condition (1), let us assume that F is a Dedekind complete Banach lattice with respect to the order unit norm given by some order unit $u \in F_+$. Then, by Kakutani's theorem, F can be represented as the space $C(S)$ of all real-valued continuous functions on some extremally disconnected compact Hausdorff space S; see [3], p. 164. Since the order unit u corresponds to the constant function 1, condition (1) for this choice of u

means precisely that for some pair of real numbers $\alpha, \beta > 0$ the following uniform estimates are fulfilled:

$$\text{(4)} \qquad \inf\{\sup_{s \in S} [\vartheta(z - \alpha x - \beta y) - \varrho(z) + \alpha\, \varrho(x) + \beta\, \varrho(y)]\,(s) : z \in K\} \leqq 0$$
$$\text{for all} \quad x, y \in K\,.$$

Let us note that there are numerous situations where the weak form of the convexity conditions (1)–(4) turns out to be crucial: for instance in the general theory of minimax theorems [1], [6], [8], [10]; in potential theory [6]; in the theory of function algebras [7]; in Choquet theory [7]; in the theory of monotone operators [7], [8]; and in optimization theory [12]. We include the following useful consequences of the preceding theorem, which improve the corresponding results from [1] and [10] and have similar applications:

Corollary 1. *Let $\vartheta\colon E \to F$ be a sublinear operator, and let K be a nonempty subset of E such that $\vartheta \geqq 0$ on K. Moreover, assume that for some pair of real numbers $\alpha, \beta > 0$ and some $u \in F_+$ the following condition is fulfilled:*

For all $x, y \in K$ and $\varepsilon > 0$ there is some $z \in K$ such that $\vartheta(z - \alpha x - \beta y) \leqq \varepsilon u$. Then there exists a linear operator $\varphi\colon E \to F$ such that $\varphi \geqq 0$ on K and $\varphi \leqq \vartheta$ on E.

Corollary 2. *Let $\vartheta\colon E \to F$ be a sublinear operator, and consider a nonempty subset K of E such that for some $0 < \lambda < 1$ and some $u \in F_+$ the following condition is fulfilled:*

For all $x, y \in K$ and $\varepsilon > 0$ there is some $z \in K$ such that $\vartheta(z - \lambda x - (1 - \lambda)\, y) \leqq \varepsilon u$. Then there exists a linear operator $\varphi\colon E \to F$ such that $\varphi \leqq \vartheta$ on E and

$$\inf\{\varphi(x)\colon x \in K\} = \inf\{\vartheta(x)\colon x \in K\}\,.$$

Of course, corollary 1 follows immediately from our theorem with the choice $\varrho = 0$. To prove corollary 2, let $\varrho := \inf\{\vartheta(x)\colon x \in K\} \in F_*$. If $\varrho = -\infty$ then every linear operator $\varphi\colon E \to F$ with $\varphi \leqq \vartheta$ on E has the desired property, whereas in the case $\varrho \in F$ the assertion is clear from the preceding theorem.

Sandwich theorems of this type have proven to be much more efficient than the usual version of the Hahn-Banach theorem on the extension of continuous linear functionals, which is, of course, contained in the classical Mazur-Orlicz theorem [11] as a special case. Also, the sandwich theory can most effectively replace the separation arguments from traditional functional analysis; for further information we refer to the monographs [2] and [8]. Simplified proofs of the classical Mazur-Orlicz theorem have been given by Sikorski [15], Pták [14], Simons [16], and others. There are also interesting variants and extensions of the classical result to the case of additive functionals on abelian semigroups [2], [4], [9] and also to the case of submodular set functions [5].

We will deduce our main result from an elementary approximation argument and the following version of the Mazur-Orlicz theorem. For completeness, we include a short proof based on Pták [14].

Lemma 1. *Let $\vartheta\colon E \to F$ be a sublinear operator, and consider an arbitrary mapping $\varrho\colon K \to F$ on a nonempty subset K of E. Then the following assertions are equivalent*:

(5) *There exists a linear operator* $\varphi\colon E \to F$ *such that* $\varrho \leqq \varphi$ *on* K *and* $\varphi \leqq \vartheta$ *on* E.

(6) $\sum_{k=1}^{r} t_k\, \varrho(x_k) \leqq \vartheta\big(\sum_{k=1}^{r} t_k x_k\big)$ *for all* $x_1, \ldots, x_r \in K$ *and* $t_1, \ldots, t_r \geqq 0$.

Proof. Obviously, condition (5) implies (6). Conversely, assume that (6) holds and consider the operator $\psi\colon E \to F$ given by

$$\psi(x) := \inf\{\vartheta(x + \sum_{k=1}^{r} t_k x_k) - \sum_{k=1}^{r} t_k\, \varrho(x_k)\colon\ x_1, \ldots, x_r \in K \text{ and } t_1, \ldots, t_r \geqq 0\}$$

for all $x \in E$. Since $\psi\colon E \to F$ is sublinear, by the basic version of the vector-valued Hahn-Banach theorem, there exists a linear operator $\varphi\colon E \to F$ which is dominated by ψ on E; see [3] or [13]. It is easily seen that such an operator φ satisfies condition (5).

Lemma 2. *Let X be an arbitrary nonempty set, and consider a family T of upper bounded mappings $f\colon X \to F_*$. Moreover, assume that for some real number $0 < \lambda < 1$ the following condition is fulfilled*:

(7) *For all* $f, g \in T$ *there exists some* $h \in T$ *such that* $h \leqq \lambda f + (1 - \lambda)\, g$ *on* X.

Then $\inf\{\sup(f)\colon f \in T\} = \inf\{\sup(f)\colon f \in \operatorname{co} T\}$, *where* $\operatorname{co} T$ *denotes the convex hull of T.*

Proof. Define $A_1 := \{0, 1\}$ and $A_{k+1} := \lambda A_k + (1 - \lambda)\, A_k$ for all $k \in \mathbb{N}$. Then an obvious inductive argument shows that T satisfies condition (7) for all $\sigma \in A_k$ and all $k \in \mathbb{N}$. We next claim that for every collection of finitely many $f_1, \ldots, f_r \in T$ and real numbers $0 < s_1, \ldots, s_r < 1$ with $s_1 + \ldots + s_r < 1$ there exists some $g \in T$ such that

$$g \leqq \sum_{k=1}^{r} s_k f_k + \big(1 - \sum_{k=1}^{r} s_k\big) \max(f_1, \ldots, f_r) \quad \text{on} \quad X,$$

where the maximum is taken in the pointwise sense. The proof of this claim is by induction on r. The case $r = 1$ being trivial with the choice $g = f_1$, let us assume that the claim holds for a given $r \in \mathbb{N}$, and consider the situation for the case $r + 1$. Since the union of the sets A_k over all $k \in \mathbb{N}$ is easily seen to be dense in the interval $[0, 1]$, there exists a $\sigma \in A_k$ for some suitable $k \in \mathbb{N}$ such that $s_1 + \ldots + s_r < \sigma <$

$< 1 - s_{r+1}$. Then, by our inductive hypothesis, we obtain some $f \in T$ such that

$$\sigma f \leqq \sum_{k=1}^{r} s_k f_k + \left(\sigma - \sum_{k=1}^{r} s_k\right) \max(f_1, \ldots, f_r) \quad \text{on} \quad X .$$

And from condition (7) for $\sigma \in A_k$ we obtain some $g \in T$ such that $g \leqq \sigma f + (1 - \sigma) f_{r+1}$ on X. An obvious combination of these inequalities shows that g has the desired property for the case $r + 1$. Now, to complete the proof of the lemma, let $a := \inf\{\sup(f): f \in T\} \in F_*$ and $b := \inf\{\sup(f): f \in \operatorname{co} T\} \in F_*$. To show that $a \leqq b$, we may assume that $a \in F$. Given an arbitrary $f \in \operatorname{co} T$, let us choose $f_1, \ldots, f_r \in T$ and $t_1, \ldots, t_r > 0$ with $t_1 + \ldots + t_r = 1$ such that $f = t_1 f_1 + \ldots + t_r f_r$. Then, for each $n \in \mathbb{N}$ with $n \geqq 2$, we may apply our claim with the choice $s_k := (1 - 1/n)\, t_k$ to obtain some $g_n \in T$ such that

$$g_n \leqq \left(1 - \frac{1}{n}\right) \sum_{k=1}^{r} t_k f_k + \frac{1}{n} \max\{\sup(f_k): k = 1, \ldots, r\} \quad \text{on} \quad X .$$

Note that the maximum on the right-hand side exists as an element of F, since $a \in F$ and since $f_1, \ldots, f_r$ are bounded above. Taking the supremum over X, we arrive at

$$a \leqq \sup(g_n) \leqq \left(1 - \frac{1}{n}\right) \sup(f) + \frac{1}{n} c$$

and hence $n(a - \sup(f)) \leqq c - \sup(f)$ for some $c \in F$ and all $n \in \mathbb{N}$. But this implies that $a \leqq \sup(f)$, since the order of a Dedekind complete vector lattice is Archimedean [3]. Thus $a \leqq b$, which completes the proof.

Proof of the theorem. Let X be the set of all linear operators $\varphi: E \to F$ with $\varphi \leqq \vartheta$ pointwise on E, and for each $a \in E$ let $\hat{a}: X \to F$ denote the corresponding vector-valued Gelfand transform given by $\hat{a}(\varphi) := \varphi(a)$ for all $\varphi \in X$. From an obvious application of lemma 1 to singletons we obtain $\vartheta(a) = \sup(\hat{a})$ for all $a \in E$. Now consider the family $S := \{\hat{a} - \varrho(a) + \varepsilon u: a \in K \text{ and } \varepsilon > 0\}$ of bounded functions from X into F and observe that the convexity assumption (1) implies that for all $f, g \in S$ there exists some $h \in S$ such that $h \leqq \alpha f + \beta g$ holds on X. In particular, taking $f = g$ and iterating this property, we obtain for every $f \in S$ and every $k \in \mathbb{N}$ some $h \in S$ such that $h \leqq (\alpha + \beta)^k f$ on X. From this observation it follows that condition (7) of lemma 2 is satisfied for the choice

$$T := \bigcup_{n=0}^{\infty} (\alpha + \beta)^{-n} S \quad \text{and} \quad \lambda := \frac{\alpha}{\alpha + \beta} \in (0, 1) .$$

Moreover, if a given $f \in T$ is represented in the form $f = (\alpha + \beta)^{-n} (\hat{a} - \varrho(a) + \varepsilon u)$ for some $n \geqq 0$, $a \in K$, and $\varepsilon > 0$, we obtain the estimate

$$\begin{aligned} \sup(f) &= (\alpha + \beta)^{-n} (\sup(\hat{a}) - \varrho(a) + \varepsilon u) = \\ &= (\alpha + \beta)^{-n} (\vartheta(a) - \varrho(a) + \varepsilon u) \geqq 0 , \end{aligned}$$

since by assumption $\varrho \leqq \vartheta$ on K and $u \geqq 0$. Thus $\sup(f) \geqq 0$ for all $f \in T$ and consequently, by lemma 2, even for all $f \in \operatorname{co} T$. Now let $a_1, \ldots, a_r \in K$ and $t_1, \ldots, t_r \geqq 0$

with $t_1 + \ldots + t_r = 1$ be arbitrarily given. Then for each $\varepsilon > 0$ we have

$$f_\varepsilon := \sum_{k=1}^{r} t_k(\hat{a}_k - \varrho(a_k) + \varepsilon u) \in \operatorname{co} T$$

and hence $\sup(f_\varepsilon) \geqq 0$, which implies that

$$\vartheta(\sum_{k=1}^{r} t_k a_k) - \sum_{k=1}^{r} t_k\, \varrho(a_k) = \sup(\sum_{k=1}^{r} t_k(\hat{a}_k - \varrho(a_k)) \geqq -\varepsilon u \quad \text{for all} \quad \varepsilon > 0\,.$$

Since the order of F is Archimedean, we arrive at

$$\vartheta(\sum_{k=1}^{r} t_k a_k) - \sum_{k=1}^{r} t_k\, \varrho(a_k) \geqq 0 \quad \text{for all} \quad a_1, \ldots, a_r \in K \text{ and } \quad t_1, \ldots, t_r \geqq 0\,.$$

Consequently, condition (6) is fulfilled so that an application of lemma 1 completes the proof of the theorem.

References

[1] *Fuchssteiner, B., König, H.:* New versions of the Hahn-Banach theorem. In: General Inequalities 2, E. F. Beckenbach (ed.). ISNM 47, pp. 255—266. Basel. Birkhäuser Verlag 1980.

[2] *Fuchssteiner, B., Lusky, W.:* Convex Cones. North-Holland Math. Studies Vol. 56. Amsterdam–New York–Oxford. North-Holland Publishing Company 1981.

[3] *Jameson, G.:* Ordered Linear Spaces. Springer Lecture Notes in Math. Vol. 141. Berlin–Heidelberg–New York. Springer 1970.

[4] *Kaufman, R.:* Interpolation of additive functionals. Studia Math. *27*, 269—272 (1966).

[5] *Kindler, J.:* A Mazur-Orlicz type theorem for submodular set functions. J. Math. Anal. Appl. *120*, 533—546 (1986).

[6] *König, H.:* Über das von Neumannsche Minimax-Theorem. Arch. Math. *19*, 482—487 (1968).

[7] *König, H.:* On certain applications of the Hahn-Banach and minimax theorems. Arch. Math. *21*, 583—591 (1970).

[8] *König, H., Neumann, M. M.:* Mathematische Wirtschaftstheorie. Math. Systems in Economics Vol. 100. Königstein/Taunus. Hain Verlag bei Athenäum 1986.

[9] *Kranz, P.:* Additive functionals on abelian semigroups. Ann. Soc. Math. Pol. *16*, 239—246 (1972).

[10] *Martellotti, A., Salvadori, A.:* A minimax theorem for functions taking values in a Riesz space. J. Math. Anal. Appl. *133*, 1—13 (1988).

[11] *Mazur, S., Orlicz, W.:* Sur les espaces métriques linéaires II. Studia Math. *13*, 137—179 (1953).

[12] *Oettli, W.:* On a general formulation of the Hahn-Banach principle with application to optimization theory. In: Optimization — Theory and Algorithms, J. B. Hiriart-Urruty, W. Oettli, and J. Stoer (eds.), pp. 91—101. New York–Basel. Marcel Dekker 1983.

[13] *Peressini, A. L.:* Ordered Topological Vector Spaces. New York–Evanston–London. Harper, and Row 1967.

[14] *Pták, V.:* On a theorem of Mazur and Orlicz. Studia Math. *15*, 365—366 (1956).

[15] *Sikorski, R.:* On a theorem of Mazur and Orlicz. Studia Math. *13*, 180—182 (1953).

[16] *Simons, S.:* Extended and sandwich versions of the Hahn-Banach theorem. J. Math. Anal. Appl. *21*, 112—122 (1968).

Author's address: Department of Mathematics and Statistics, Mississippi State University, Mississippi State, MS 39762, USA.