FUNCTION SPACES X
BANACH CENTER PUBLICATIONS, VOLUME 102
INSTITUTE OF MATHEMATICS
POLISH ACADEMY OF SCIENCES
WARSZAWA 2014

ON SIMONS' VERSION OF HAHN–BANACH–LAGRANGE THEOREM

JERZY GRZYBOWSKI, HUBERT PRZYBYCIEŃ and RYSZARD URBAŃSKI

Faculty of Mathematics and Computer Science, Adam Mickiewicz University
Umultowska 87, 61-614 Poznań, Poland
E-mail: jgrz@amu.edu.pl, hubert@amu.edu.pl, rich@amu.edu.pl

Abstract. In this paper we generalize in Theorem 12 some version of Hahn–Banach Theorem which was obtained by Simons. We also present short proofs of Mazur and Mazur–Orlicz Theorem (Theorems 2 and 3).

Simons, using the concept of p-convexity, proved a version of Hahn–Banach Theorem (Theorem 1.13 in [11]), which is a generalization of Hahn–Banach–Lagrange Theorem (Theorem 1.11 in [11]). Simons' theorem enabled him to present short proofs of a number of important and difficult theorems in functional analysis and to find applications in convex analysis and theory of monotone multifunctions (see [8, 9, 10, 11]).

In this paper we present short proofs of Mazur and Mazur–Orlicz Theorems (Theorems 2 and 3). Then we apply them to generalize Simons' theorem (Theorem 13) in our Theorem 12.

Throughout the paper by X we will denote a nontrivial vector space over the field of real numbers.

LEMMA 1. *Let $p : X \to \mathbb{R}$ be a convex function $y \in X$. For all $x \in X$, let*

$$p_y(x) := \inf_{\lambda>0} \frac{p(y+\lambda x) - p(y)}{\lambda}, \quad p'(x) := \inf_{\lambda>0} \frac{p(\lambda x)}{\lambda}.$$

Then:

(a) *$p_y : X \to \mathbb{R}$ is sublinear and $p(y) - p(2y) \le p_y(y) \le p(0) - p(y)$;*
(b) *if $p(0) \ge 0$ then p' is the greatest sublinear functional on X less than or equal to p;*
(c) *if p is sublinear then $p_y \le p$ and $p_y(-y) = -p(y)$.*

2010 *Mathematics Subject Classification*: Primary 46A22; Secondary 46N10.

Key words and phrases: Hahn–Banach theorem.

The paper is in final form and no version of it will be published elsewhere.

DOI: 10.4064/bc102-0-6

Proof. For $x, y \in X$ and $\lambda > 0$ we have
$$(1+\lambda)p(y) \le p(y+\lambda x) + \lambda p(y-x),$$
which implies
$$p(y) - p(y-x) \le \frac{p(y+\lambda x) - p(y)}{\lambda}\,.$$

Taking the infimum over $\lambda > 0$ we get $p_y(x) > -\infty$ and
$$p(y) - p(y-x) \le p_y(x) \le p(y+x) - p(y). \tag{1}$$
It is easy to observe that p_y is a positively homogeneous. Consider $x_1, x_2 \in X$ and arbitrary $\lambda_1, \lambda_2 > 0$. Since
$$p\Big(y + \frac{\lambda_1\lambda_2}{\lambda_1+\lambda_2}(x_1+x_2)\Big) \le \frac{\lambda_2}{\lambda_1+\lambda_2}\, p(y+\lambda_1 x_1) + \frac{\lambda_1}{\lambda_1+\lambda_2}\, p(y+\lambda_2 x_2), \tag{2}$$
p_y is subadditive, and we obtain (a).

Now let $p(0) \ge 0$. For $x \in X$ and $\lambda > 0$ from (1) we have
$$p(0) - p(-x) \le \frac{p(\lambda x) - p(0)}{\lambda} \le \frac{p(\lambda x)}{\lambda}\,.$$
Hence $p'(x) > -\infty$ and $p_0 \le p' \le p$. It is easy to observe that p' is positively homogeneous. Now from (2), p' is subadditive. Now let q be sublinear and $q \le p$ on X. Then $\lambda q(x) \le p(\lambda x)$ for every $\lambda > 0$. Hence $q \le p'$ on X, and we get (b). ∎

In [11] a short proof of the classical Hahn–Banach Theorem is given. Similarly, applying Lemma 1, we give a short proof of a basic version of classical Mazur Theorem.

Theorem 2 (Mazur). *Let $p : X \to \mathbb{R}$ be a convex functional, $p(0) \ge 0$. Then there exists a linear functional l on X such that $l \le p$.*

Proof. By $\mathcal{C}_p(X)$ denote the set of all convex functionals q on X such that $p \ge q$ and $q(0) \ge 0$. Since for every $q \in \mathcal{C}_p(X)$ and $x \in X$, $q(x) \ge -q(-x) + 2q(0) \ge -p(-x)$, by using Kuratowski–Zorn Lemma, there exists a minimal element l in $\mathcal{C}_p(X)$. Now, by Lemma 1, a functional l' is sublinear and $l' \le l$. Hence $l' = l$ and l is sublinear. Since $l_y \le l$, $l_y = l$. Again, from Lemma 1, we have $l(-y) = l_y(-y) = -l_y(y) = -l(y)$ for $y \in X$. Thus l is linear. ∎

In 1953 Mazur and Orlicz [5] proved some generalization of Hahn–Banach Theorem. We present a version of Mazur–Orlicz Theorem [1, 6] for convex functionals. Our short proof of Mazur–Orlicz Theorem is based on the idea of Pták [6] and Mazur Theorem (Theorem 2).

Theorem 3 (Mazur–Orlicz). *Let $p : X \to \mathbb{R}$ be a convex functional. Moreover, let $g : A \to X$ and $f : A \to \mathbb{R}$ be functions defined on a nonempty subset A of X. Then the following statements are equivalent:*

(a) *there exists a linear functional l on X such that $l \le p$ on X and $f \le l \circ g$ on A;*
(b) *for every finite sequence $a_1, \dots, a_n \in A$,*
$$\sum_{i=1}^{n} \lambda_i f(a_i) \le p\Big(\sum_{i=1}^{n} \lambda_i g(a_i)\Big)$$
for all non-negative real numbers $\lambda_1, \dots, \lambda_n$.

Proof. Obviously, the condition (a) implies (b). Suppose that the condition (b) holds and consider a functional

$$p_1(x) := \inf\Big\{p\Big(x + \sum_{i=1}^{n} \lambda_i g(a_i)\Big) - \sum_{i=1}^{n} \lambda_i f(a_i) \;\Big|\; a_i \in A,\ \lambda_i \geq 0\Big\}$$

for all $x \in X$. Then $p_1 : X \to \mathbb{R}$ is convex, $p_1(0) \geq 0$ and $p_1 \leq p$. By Mazur Theorem (Theorem 2) there exists a functional l on X such that $l \leq p_1$. Since $l(-ng(a)) \leq p(0) - nf(a)$ for all $a \in A$, $n \in \mathbb{N}$, we obtain $f \leq l \circ g$ on A. ∎

All three theorems: Simons' version of Mazur–Orlicz Theorem (Lemma 1.6 in [11]), classical Mazur–Orlicz Theorem [5] and Mazur Theorem [1, 4] which is a generalization of Hahn–Banach Theorem [2, 7] follow from Theorem 3.

The following lemma will be applied in our proof of Theorem 12. The proof is based on Theorems 2 and 3.

LEMMA 4. *Let $p : X \to \mathbb{R}$ be a convex functional. Moreover, let $g : A \to X$ and $f : A \to \mathbb{R}$ be functions defined on a nonempty subset A of X. Then the following statements are equivalent*:

(a) *there exists a linear functional l on X such that $l \leq p$ and*

$$\inf_A [f + l \circ g] = \inf_A [f + p \circ g]$$

(b) *for every finite sequence $a_1, \ldots, a_n \in A$ and for arbitrary non-negative real numbers $\lambda_1, \ldots, \lambda_n$ the following inequality holds*

$$\sum_{i=1}^{n} \lambda_i(\alpha - f(a_i)) \leq p\Big(\sum_{i=1}^{n} \lambda_i g(a_i)\Big), \tag{3}$$

where $\alpha = \inf_A [f + p \circ g]$.

Proof. Let $\alpha = -\infty$. Then obviously (a) implies (b). From the condition (b), by Mazur Theorem there exists a linear functional l on X such that $l \leq p - p(0)$. Thus (a) holds.

Now let $\alpha \in \mathbb{R}$ and $l \leq p$ on X then by (a) we get $l \circ g \geq \alpha - f$ on A and by Mazur–Orlicz Theorem we get (3). Conversely if (3) is satisfied then there exists a linear functional l on X such that $l \leq p$ and $f + l \circ g \geq \alpha$ on A. ∎

In order to present the main result (Theorem 12) and Simons' theorem (Theorem 13) we need to give definitions of p-convex, p_f^2-convex and p_f-convex functions.

DEFINITION 5. Let $p : X \to \mathbb{R}$ be a sublinear functional. A function $g : A \to X$ defined on a nonempty convex subset A of X is said to be *p-convex* if

$$p\big(x + g(\lambda_1 a_1 + \lambda_2 a_2)\big) \leq p\big(x + \lambda_1 g(a_1) + \lambda_2 g(a_2)\big)$$

for all $x \in X$, $a_1, a_2 \in A$ and $\lambda_1, \lambda_2 > 0$ such that $\lambda_1 + \lambda_2 = 1$.

REMARK 6. Let us note that the function g is *p-convex* if and only if $p\big(g(\lambda_1 a_1 + \lambda_2 a_2) - \lambda_1 g(a_1) - \lambda_2 g(a_2)\big) \leq 0$ for all $a_1, a_2 \in A$ and $\lambda_1, \lambda_2 > 0$ such that $\lambda_1 + \lambda_2 = 1$. In fact, p-convexity depends only on the cone $\{x \in X \,|\, p(x) \leq 0\}$.

Definition 7. Let $f : A \to (-\infty, \infty]$. The set $\operatorname{dom} f := \{x \in A \,|\, f(x) \in \mathbb{R}\}$ is called the *effective domain* of f. We say that f is *proper* if $\operatorname{dom} f \neq \emptyset$. By $\mathcal{PC}(A)$ we will denote the set of all proper convex functions from A into $(-\infty, \infty]$.

In [11] Simons generalized p-convexity as follows.

Definition 8. Let $p : X \to \mathbb{R}$ be a sublinear functional, A be a nonempty subset of X and $f : A \to \mathbb{R}$. A function $g : A \to X$ is said to be p^2_f*-convex* if for all $a_1, a_2 \in A$, there exists $a \in A$ such that

$$p\Big(g(a) - \Big(\frac{1}{2}\, g(a_1) + \frac{1}{2}\, g(a_2)\Big)\Big) \le 0 \quad \text{and} \quad f(a) \le \frac{1}{2}\, f(a_1) + \frac{1}{2}\, f(a_2).$$

Now we introduce a broader class of p_f-convex functions.

Definition 9. Let $p : X \to \mathbb{R}$, A be a nonempty subset of X and $f : A \to \mathbb{R}$. A function $g : A \to X$ is said to be p_f*-convex* if for every $b \in \operatorname{conv} g(A)$, $b = \sum_{i=1}^n \lambda_i g(a_i)$, $\epsilon > 0$ there exists $a \in A$ such that for every $\lambda \ge 0$

$$\lambda p \circ g(a) \le p(\lambda b) + \epsilon \quad \text{and} \quad f(a) \le \sum_{i=1}^n \lambda_i f(a_i) + \epsilon.$$

The following lemma shows the connection between p^2_f-convexity and p_f-convexity.

Lemma 10. *Let $p : X \to \mathbb{R}$ be a sublinear functional, A be a nonempty subset of X and $f : A \to \mathbb{R}$. If $g : A \to X$ is p^2_f-convex, then g is p_f-convex.*

Proof. Let $x_1 = g(b_1) - (\frac{1}{2} g(a_1) + \frac{1}{2} g(a_2))$, $x_2 = g(b_2) - (\frac{1}{2} g(a_3) + \frac{1}{2} g(a_4))$ and $x_3 = g(a) - (\frac{1}{2} g(b_1) + \frac{1}{2} g(b_2))$ for some $a, b_1, b_2, a_1, a_2, a_3, a_4 \in A$. Assume that $p(x_1) \le 0$, $p(x_2) \le 0$ and $p(x_3) \le 0$. Since p is subadditive and positively homogenous, $p\big(g(a) - (\frac{1}{4} g(a_1) + \frac{1}{4} g(a_2) + \frac{1}{4} g(a_3) + \frac{1}{4} g(a_4))\big) = p\big(x_3 + \frac{1}{2} x_1 + \frac{1}{2} x_2\big) \le p(x_3) + \frac{1}{2} p(x_1) + \frac{1}{2} p(x_2) \le 0$. Therefore, for every $b = \sum_{i=1}^n \lambda_i g(a_i)$, where λ_i are binary rational (i.e. of the form $\frac{m}{2^k}$ where $m, k \in \mathbb{Z}$) and $a_i \in A$ there exists $a \in A$ such that

$$p \circ g(a) \le p(b) \quad \text{and} \quad f(a) \le \sum_{i=1}^n \lambda_i f(a_i). \tag{4}$$

Let us fix $b \in \operatorname{conv} g(A)$, $b = \sum_{i=1}^n \lambda_i g(a_i)$, $\lambda \ge 0$, $\epsilon > 0$. Since p is sublinear, p is continuous on $\operatorname{conv}\{g(a_1), \ldots, g(a_n)\}$. Hence for some binary rational $\lambda'_i \ge 0$, $i = 1, \ldots, n$, which are sufficiently close to λ_i, $i = 1, \ldots, n$, and such that $\sum_{i=1}^n \lambda'_i = 1$ we have $\lambda p(\sum_{i=1}^n \lambda'_i g(a_i)) \le \lambda p(b) + \varepsilon$ and $\sum_{i=1}^n \lambda'_i f(a_i) \le \sum_{i=1}^n \lambda_i f(a_i) + \varepsilon$. Now we can find $a \in A$ for $b' = \sum_{i=1}^n \lambda'_i g(a_i)$ and apply (4). ∎

The class of p_f-convex functions is substantially broader than the class of p^2_f-convex functions. In order to show it we give the following example.

Example 11. Let $X = \mathbb{R}^2$, $a_1 = (0,0)$, $a_2 = (0,1)$, $A = \{a_1, a_2\}$, $f(a_1) = f(a_2) = 0$, $g = \mathrm{Id}_A$ and $p(x) = p(x_1, x_2) = \sqrt{x_1^2 + x_2^2} + x_1$. Since

$$p(g(a_1) - g(a_2)) = p(g(a_2) - g(a_1)) = 1 > 0,$$

the function g is not p^2_f-convex. On the other hand, if $b \in \operatorname{conv} g(A) = \{0\} \times [0,1]$ then $b = \alpha g(a_1) + \beta g(a_2) = \beta a_2$, where α, $\beta \ge 0$ and $\alpha + \beta = 1$. Let $a = a_1$ and $\lambda \ge 0$.

We have an obvious equality $f(a) = 0 = \alpha f(a_1) + \beta f(a_2)$. The function g is p_f-convex because of the inequality

$$\lambda p(g(a)) = 0 \le \lambda\beta = \lambda p(b) = p(\lambda b).$$

The example shows that the following theorem is an essential generalization of Simons' theorem (Theorem 13).

THEOREM 12. *Let $p : X \to \mathbb{R}$ be a convex functional and let A be a nonempty convex subset of X. If $f : A \to \mathbb{R}$ and $g : A \to X$ is p_f-convex, then there exists a linear functional l on X such that $l \le p$ and*

$$\inf_A [f + l \circ g] = \inf_A [f + p \circ g].$$

Proof. Let $\alpha = \inf_A [f + p \circ g]$. Then $\alpha - f \le p \circ g$ on A. For arbitrary non-negative real numbers $\lambda_1, \dots, \lambda_n$, let us put $\lambda = \lambda_1 + \ldots + \lambda_n$. Without loss of generality, we may assume that $\lambda > 0$. Then, for any $a_1, \dots, a_n \in A$, $\epsilon > 0$ there exists $a \in A$ such that

$$\sum_{i=1}^n \lambda_i(\alpha - f(a_i)) \le \lambda(\alpha - f(a)) + \epsilon \le \lambda p \circ g(a) + \epsilon \le p\Big(\sum_{i=1}^n \lambda_i g(a_i)\Big) + 2\epsilon.$$

Hence the condition (b) of Lemma 4 is satisfied, so there exists a linear functional l on X such that $l \le p$ and $\alpha = \inf_A [f + l \circ g]$. ∎

Theorem of Simons (Theorem 1.13 in [11]) is a simple corollary of Theorem 12 and Lemma 4:

THEOREM 13 (Simons). *Let $p : X \to \mathbb{R}$ be a sublinear functional and let A be a nonempty convex subset of X. If $f \in \mathcal{PC}(A)$ and $g : A \to X$ is p_f^2-convex, then there exists a linear functional l on X such that $l \le p$ and*

$$\inf_A [f + l \circ g] = \inf_A [f + p \circ g].$$

In [8, 9, 10] Simons proved for p-convex functions some version of Hahn–Banach Theorem which he calls Hahn–Banach–Lagrange Theorem (Theorem 1.11 in [11]). Theorem 13 is a generalization of Theorem 14.

THEOREM 14 (Hahn–Banach–Lagrange). *Let $p : X \to \mathbb{R}$ be a sublinear functional and let A be a nonempty convex subset of X. If $f \in \mathcal{PC}(A)$ and $g : A \to X$ is p-convex, then there exists a linear functional l on X such that $l \le p$ and*

$$\inf_A [f + l \circ g] = \inf_A [f + p \circ g].$$

REMARK 15. If p is a sublinear functional then, by Lemma 10, every p_f^2-convex function is p_f-convex. Hence Theorem 13 follows from Theorem 12.

REMARK 16. If $p : X \to \mathbb{R}$ is convex we can reformulate the definition of p_f-convexity. The function g is *p_f-convex* if and only if $p(0) \ge 0$ and for every $b \in \operatorname{conv} g(A)$, $b = \sum_{i=1}^n \lambda_i g(a_i)$, $\epsilon > 0$ there exists $a \in A$ such that

$$p \circ g(a) \le p'(b) + \epsilon \quad \text{and} \quad f(a) \le \sum_{i=1}^n \lambda_i f(a_i) + \epsilon,$$

where p' is given in Lemma 1.

References

[1] A. Alexiewicz, *Functional Analysis*, PWN, Warszawa 1969 (in Polish).
[2] S. Banach, *Sur les fonctionnelles linéaires II*, Studia Math. 1 (1929), 223–239.
[3] H. König, *On certain applications of the Hahn–Banach and minimax theorems*, Arch Math. (Basel) 21 (1970), 583–591.
[4] S. Mazur, *Über konvexe Mengen in linearen normierten Räumen*, Studia Math. 4 (1933), 70–84.
[5] S. Mazur, W. Orlicz, *Sur les espaces métriques linéaires II*, Studia Math. 13 (1953), 137–179.
[6] V. Pták, *On a theorem of Mazur and Orlicz*, Studia Math. 15 (1956), 365–366.
[7] W. Rudin, *Functional Analysis*, McGraw-Hill Ser. in Higher Math., McGraw-Hill, New York 1973.
[8] S. Simons, *A new version of the Hahn–Banach theorem*, Arch. Math. (Basel) 80 (2003), 630–646.
[9] S. Simons, *Hahn–Banach theorems and maximal monotonicity*, in: Variational Analysis and Applications, Nonconvex Optim. Appl. 79, Springer, New York 2005, 1049–1083.
[10] S. Simons, *The Hahn–Banach–Lagrange theorem*, Optimization 56 (2007), 149–169.
[11] S. Simons, *From Hahn–Banach to Monotonicity*, Lecture Notes in Math. 1693, Springer, Berlin 2008.