Arch. Math. 80 (2003) 630–646
0003–889X/03/060630–17
DOI 10.1007/s00013-003-0500-2

Archiv der Mathematik

A new version of the Hahn-Banach theorem

By

S. Simons

Abstract. We discuss a new version of the Hahn-Banach theorem, with applications to linear and nonlinear functional analysis, convex analysis, and the theory of monotone multifunctions. We show how our result can be used to prove a "localized" version of the Fenchel-Moreau formula — even when the classical Fenchel-Moreau formula is valid, the proof of it given here avoids the problem of the "vertical hyperplane". We give a short proof of Rockafellar's fundamental result on dual problems and Lagrangians — obtaining a necessary and sufficient condition instead of the more usual sufficient condition. We show how our result leads to a proof of the (well-known) result that if a monotone multifunction on a normed space has bounded range then it has full domain. We also show how our result leads to generalizations of an existence theorem with no *a priori* scalar bound that has proved very useful in the investigation of monotone multifunctions, and show how the estimates obtained can be applied to Rockafellar's surjectivity theorem for maximal monotone multifunctions in reflexive Banach spaces. Finally, we show how our result leads easily to a result on convex functions that can be used to establish a minimax theorem.

0. Introduction. In this paper, we discuss a new version of the Hahn-Banach theorem that has a number of applications in different fields of analysis. We shall give applications to linear and nonlinear functional analysis, convex analysis, and the theory of monotone multifunctions.

After a few preliminaries, the main result appears in Theorem 1.5, which uses the concept of "S-convexity" introduced in Definition 1.3. The full force of this concept will be used only in Theorem 2.4, a result on convex functions with applications to a minimax theorem. For all the other applications of Theorem 1.5 in this paper, the reader can substitute "affine" for "S-convex". This change shortens the proof of Lemma 1.4 by a few lines.

In Section 2, we sketch how Theorem 1.5 can be used to give the main existence theorems for linear functionals in functional analysis, and also how it gives the result referred to above that leads to a minimax theorem.

Section 3 contains two applications of Theorem 1.5 to convex analysis. The first, Theorem 3.4, is a "localized" version of the Fenchel-Moreau formula. Even in the situation when the classical Fenchel-Moreau formula is valid, the proof of it given here using

---

Mathematics Subject Classification (2000): 46A22, 46N10, 49J35, 47H05.

Theorem 1.5 allows us to avoid the problem of the "vertical hyperplane". The second application is a short proof of a fundamental result on dual problems and Lagrangians due to Rockafellar.

Theorem 1.5 has many applications to the theory of monotone multifunctions. In Section 4, we describe one of these, the (well-known) result that if a maximal monotone multifunction on a normed space has bounded range then it has full domain. We will describe another application to the theory of monotone multifunctions in Section 6.

Section 5 is also motivated by the theory of monotone multifunctions. Theorem 5.1 is an existence theorem without any a *priori* scalar bound that has proved very useful in the investigation of these multifunctions. There is a certain formal similarity between the statements of Theorem 5.1 and Theorem 1.5. The main result of Section 5 is Theorem 5.4, which generalizes and unifies Theorem 5.1 and Theorem 1.5.

In Section 6, we show how the estimates obtained in Section 5 can be applied to Rockafellar's surjectivity theorem for maximal monotone multifunctions in reflexive Banach spaces.

The author would like to express his thanks to Professors J. Borwein and R. T. Rockafellar for illuminating discussions on some of the results in this paper, and to Professor J. Zhu for finding a mistake in the original definition of "S-convex".

Part of the work for this paper was done while the author was visiting the *Centro de Modelamiento Matematico, Universidad de Chile* in Santiago. The author would like to thank the *Centro de Modelamiento Matematico* for its hospitality.

1. The main result. Theorem 1.5 contains the new version of the Hahn-Banach theorem that forms the main topic of this paper. Theorem 1.5 is proved using the technique of the "auxiliary sublinear functional" that has been used extensively by Prof H. König — most of the work for it is actually done in Lemma 1.4.

We start this section by recalling in Theorem 1.2 the classical Hahn-Banach theorem for sublinear functionals.

Definition 1.1. Let E be a nontrivial vector space. We say that $S\colon E \mapsto \mathbb{R}$ is *sublinear* if

$$x, y \in E \implies S(x+y) \leqq S(x) + S(y)$$

and

$$x \in E \text{ and } \lambda > 0 \implies S(\lambda x) = \lambda S(x).$$

Theorem 1.2. *Let E be a nontrivial vector space and $S\colon E \mapsto \mathbb{R}$ be sublinear. Then there exists a linear functional L on E such that $L \leqq S$ on E.*

Proof. See Kelly-Namioka, [5, 3.4, p. 21] for a proof using cones, Rudin, [11, Theorem 3.2, p. 56–57] for a proof using an extension by subspaces argument, and König, [6] and Simons, [12] for a proof using an ordering on sublinear functionals. □

Remark 2.5 contains some comments on the appropriateness of the various methods of proof for Theorem 1.2.

Definition 1.3. Let E be a nontrivial vector space and $S\colon E \mapsto \mathbb{R}$ be sublinear. Let A be a nonempty convex subset of a vector space and $g\colon A \mapsto E$. We say that g is *S-convex* if, for all $x \in E$,

$$a_1, a_2 \in A,\ \mu_1, \mu_2 > 0 \text{ and } \mu_1 + \mu_2 = 1$$
$$\Longrightarrow\ S\left(x + g\left(\sum_i \mu_i a_i\right)\right) \leqq S\left(x + \sum_i \mu_i g(a_i)\right).$$

Note that if we define an ordering "$\leqq_S$" on E by declaring that $y \leqq_S z$ if, for all $x \in E$, $S(x + y) \leqq S(x + z)$ then g is S-convex if, and only if,

$$a_1, a_2 \in A,\ \mu_1, \mu_2 > 0 \text{ and } \mu_1 + \mu_2 = 1$$
$$\Longrightarrow\ g\left(\sum_i \mu_i a_i\right) \leqq_S \sum_i \mu_i g(a_i).$$

An affine function is clearly S-convex. As observed in the introduction, apart from the application in Theorem 2.4, all the S-convex functions in this paper will, in fact, be affine.

Lemma 1.4. *Let E be a nontrivial vector space and $S\colon E \mapsto \mathbb{R}$ be sublinear. Let A be a nonempty convex subset of a vector space, $g\colon A \mapsto E$ be S-convex and $f\colon A \mapsto \mathbb{R}$ be convex. Let $\alpha := \inf_A [f + S \circ g] \in \mathbb{R}$. For all $x \in E$, let*

(1.4.1) $$T(x) := \inf_{a \in A,\ \lambda > 0} [S(x + \lambda g(a)) + \lambda f(a) - \lambda\alpha].$$

Then $T\colon E \mapsto \mathbb{R}$, T is sublinear, $T \leqq S$ on E and, for all $a \in A$, $-T(-g(a)) + f(a) \geqq \alpha$.

Proof. If $x \in E$, $a \in A$ and $\lambda > 0$ then

$$S(x + \lambda g(a)) + \lambda f(a) - \lambda\alpha$$
$$\geqq S(\lambda g(a)) + \lambda f(a) - \lambda\alpha - S(-x) \geqq -S(-x) > -\infty.$$

Taking the infimum over $a \in A$ and $\lambda > 0$, $T(x) \geqq -S(-x) > -\infty$. Thus $T\colon E \mapsto \mathbb{R}$. It is now easy to check that T is positively homogeneous, so to prove that T is sublinear it remains to show that T is subadditive. To this end, let $x_1,\ x_2 \in E$. Let $a_1,\ a_2 \in A$ and $\lambda_1,\ \lambda_2 > 0$ be arbitrary. Write $x := \sum_i x_i$, $\lambda := \sum_i \lambda_i$, $\mu_i := \lambda_i/\lambda$ and $a := \sum_i \mu_i a_i$. Then

$$\sum_i [S(x_i + \lambda_i g(a_i)) + \lambda_i f(a_i) - \lambda_i\alpha]$$
$$\geqq S\left(\sum_i x_i + \sum_i \lambda_i g(a_i)\right) + \sum_i \lambda_i f(a_i) - \sum_i \lambda_i\alpha$$

$$= S\left(x + \sum_i \lambda_i g(a_i)\right) + \sum_i \lambda_i f(a_i) - \lambda\alpha$$
$$= \lambda S\left(x/\lambda + \sum_i \mu_i g(a_i)\right) + \lambda \sum_i \mu_i f(a_i) - \lambda\alpha,$$

using the S-convexity of g and the convexity of f,

$$\geqq \lambda S(x/\lambda + g(a)) + \lambda f(a) - \lambda\alpha$$
$$= S(x + \lambda g(a)) + \lambda f(a) - \lambda\alpha$$
$$\geqq T(x) = T(x_1 + x_2).$$

Taking the infimum over a_1, a_2, λ_1 and λ_2 gives $T(x_1) + T(x_2) \geqq T(x_1 + x_2)$. Thus T is subadditive, and consequently, sublinear. Fix $a \in A$. Let x be an arbitrary element of E. Then, for all $\lambda > 0$, $T(x) \leqq S(x) + \lambda[S(g(a)) + f(a) - \alpha]$. Letting $\lambda \to 0$, $T(x) \leqq S(x)$. Thus $T \leqq S$ on E. Finally, let a be an arbitrary element of A. Then, taking $\lambda = 1$ in (1.4.1),

$$T(-g(a)) \leqq S(-g(a) + g(a)) + f(a) - \alpha = f(a) - \alpha,$$

hence $-T(-g(a)) + f(a) \geqq \alpha$, which completes the proof of Lemma 1.4. □

Theorem 1.5. *Let E be a nontrivial vector space and $S: E \mapsto \mathbb{R}$ be sublinear. Let A be a nonempty convex subset of a vector space, $g: A \mapsto E$ be S-convex and $f: A \mapsto \mathbb{R}$ be convex. Then there exists a linear functional L on E such that $L \leqq S$ on E and*

$$\inf_A [f + L \circ g] = \inf_A [f + S \circ g]. \tag{1.5.1}$$

Proof. Let $\alpha := \inf_A [f + S \circ g]$. If $\alpha = -\infty$, the result is immediate from Theorem 1.2 (take any linear functional L on E such that $L \leqq S$ on E). So we can suppose that $\alpha \in \mathbb{R}$. Define T as in Lemma 1.4. From Theorem 1.2, there exists a linear functional L on E such that $L \leqq T$ on E. Since $T \leqq S$ on E, $L \leqq S$ on E, as required. Let $a \in A$. Then

$$L(g(a)) + f(a) = -L(-g(a)) + f(a) \geqq -T(-g(a)) + f(a) \geqq \alpha.$$

Taking the infimum over $a \in A$,

$$\inf_A [f + L \circ g] \geqq \alpha = \inf_A [f + S \circ g].$$

On the other hand, since $L \leqq S$ on E, $\inf_A [f + L \circ g] \leqq \inf_A [f + S \circ g]$. □

Remark 1.6. It is worth pointing out that the definition of the auxiliary sublinear functional used to prove Theorem 1.5 is "forced" in the sense that if L is linear, $L \leqq S$ on E and (1.5.1) is satisfied with $\alpha := \inf_A [f + S \circ g] \in \mathbb{R}$ then, as the reader can easily verify, $L \leqq T$ on E.

2. Applications to functional analysis and minimax theorems. If E is a nontrivial vector space, let $\mathcal{PC}(E)$ stand for the set of all convex functions $h\colon E \mapsto (-\infty, \infty]$ such that $\operatorname{dom} h \neq \emptyset$, where $\operatorname{dom} h$, the *effective domain* of h, is defined by

$$\operatorname{dom} h := \{x \in E\colon h(x) \in \mathbb{R}\}.$$

(The "$\mathcal{P}$" stands for "proper", which is the adjective frequently used to denote the fact that a function is finite at at least one point.)

Theorem 2.1 is the *sandwich theorem* (see [6, Theorem 1.7, p. 112]). It follows immediately from Theorem 1.5 with $A := \operatorname{dom} h$, $g(a) := a$ and $f := h|_A$.

Theorem 2.1. *Let E be a nontrivial vector space, $S\colon E \mapsto \mathbb{R}$ be sublinear, $h \in \mathcal{PC}(E)$ and $-h \leqq S$ on E. Then there exists a linear functional L on E such that $-h \leqq L \leqq S$ on E.*

Theorem 2.1 implies in turn two other well known existence results: the *extension form of the Hahn-Banach theorem*, Corollary 2.2, (see [6, Corollary 1.8, p. 112]) and the *Mazur-Orlicz theorem*, Corollary 2.3, (see [6, Theorem 1.9, p. 112]). We leave to the reader the details of the substitutions that have to be made to derive them.

Corollary 2.2. *Let E be a nontrivial vector space, A be a linear subspace of E, $S\colon E \mapsto \mathbb{R}$ be sublinear, $M\colon A \mapsto \mathbb{R}$ be linear and $M \leqq S$ on A. Then there exists a linear functional L on E such that $L \leqq S$ on E and $L|_A = M$.*

Corollary 2.3. *Let E be a nontrivial vector space, $S\colon E \mapsto \mathbb{R}$ be sublinear and A be a nonempty convex subset of E. Then there exists a linear functional L on E such that $L \leqq S$ on E and $\inf_A L = \inf_A S$.*

Theorem 1.5 can also be used to give a very simple proof of Theorem 2.4 below, which was essentially proved by Fan-Glicksberg-Hoffman, (see [3, Theorem 1, p. 618]).

Theorem 2.4. *Let A be a nonempty convex subset of a vector space and $f_1, \dots, f_m$ be convex real functions on A. Then there exist $\lambda_1, \dots, \lambda_m \geqq 0$ such that $\lambda_1 + \dots + \lambda_m = 1$ and*

$$\inf_A [f_1 \vee \ldots \vee f_m] = \inf_A [\lambda_1 f_1 + \dots + \lambda_m f_m].$$

Proof. Let $\alpha := \inf_A [f_1 \vee \ldots \vee f_m]$. If $\alpha = -\infty$, the result is immediate with any $\lambda_1, \dots, \lambda_m$ satisfying the other conditions, so we can suppose that $\alpha \in \mathbb{R}$. The result follows from Theorem 1.5 with $E := \mathbb{R}^m$, $S(\mu_1, \dots, \mu_m) := \mu_1 \vee \ldots \vee \mu_m$, $g(a) := (f_1(a), \dots, f_m(a))$ and $f(a) := -\alpha$. $\square$

Theorem 2.4 leads in turn to a short proof of the minimax theorem proved by Fan in [2] (see [13, Theorem 3.1, p. 17] for details of this).

Remark 2.5. Since we have given Corollary 2.2 as (ultimately) a consequence of Theorem 1.2, in order to dispel any suspicion of circularity, it would seem better to avoid the "extension by subspaces" proof of Theorem 1.2, using instead the "cone" argument of Kelly-Namioka, [5], or the "minimal sublinear functional" argument outlined below, which is most in tune with the other analysis in this paper. It is easy to see from Zorn's lemma that if $U\colon E \mapsto \mathbb{R}$ is sublinear then there exists a sublinear functional S on E such that $S \leqq U$ on E and S is minimal with respect to the pointwise ordering of $\mathbb{R}^E$. Let a be an arbitrary element of E. If we define $A := \{a\}$, $g(a) := a$ and $f(a) := 0$ then Lemma 1.4 yields a sublinear functional T on E such that $T \leqq S$ on E and $-T(-a) \geqq S(a)$. The minimality of S now gives $T = S$, thus we have proved that $-S(-a) \geqq S(a)$. Since S is sublinear, it follows easily from this that S is linear. This gives us a proof of Theorem 1.2 that does not depend on an "extension by subspaces" argument. More details of this approach can be found in König, [6] and Simons, [12].

3. Applications to convex analysis. Let E be a nontrivial real Hausdorff locally convex space with dual E^*. $\mathcal{S}(E)$ stands for the family of all continuous seminorms on E. If $f \in \mathcal{PC}(E)$, the *Fenchel conjugate*, f^*, of f is the function from E^* into $(-\infty, \infty]$ defined by

$$f^*(x^*) := \sup_E (x^* - f).$$

It follows easily from the definitions above that, for all $y \in E$,

$$f(y) \geqq \sup_{E^*} (y - f^*). \tag{3.0.1}$$

One of the fundamental results in convex analysis is the *Fenchel-Moreau formula* that if $f \in \mathcal{PC}(E)$ is lower semicontinuous then *we always have equality in* (3.0.1). (See Moreau, [7], Section 5–6, p. 26–39.)

Now suppose that f is not necessarily lower semicontinuous. Let us say that $y \in E$ is a *Fenchel-Moreau* point of f if equality holds in (3.0.1). It is very tempting to speculate that every point of lower semicontinuity of f is a Fenchel-Moreau point of f. Remark 3.1 shows that this is false. However, we establish in Theorem 3.4 that every point of lower semicontinuity of f is a Fenchel-Moreau point provided that f is lower semicontinuous at at least one point of its effective domain. Our proof goes by way of Theorem 3.2, which contains a simple characterization of the Fenchel-Moreau points of f.

Remark 3.1. Let E be infinite-dimensional. Fix $x^* \in E^* \setminus \{0\}$ and a discontinuous linear functional L on E. Define

$$f(x) := \begin{cases} \infty, & \text{if } \langle x, x^* \rangle < 1; \\ L(x), & \text{if } \langle x, x^* \rangle \geqq 1. \end{cases}$$

Clearly, $f \in \mathcal{PC}(E)$ and f is lower semicontinuous at 0. Let y^* be an arbitrary element of E^*. Since x^* and $y^* - L$ are linearly independent, there exist $u, v \in E$ such that

$$\langle u, x^* \rangle = 1,\ \langle v, x^* \rangle = 0,\ (y^* - L)(u) = 0,\ \text{and } (y^* - L)(v) = 1.$$

Let $\lambda \in \mathbb{R}$, and set $x := u + \lambda v$. Then $\langle x, x^* \rangle = \langle u, x^* \rangle = 1$, and so $f(x) = L(x)$. Thus

$$f^*(y^*) \geqq \langle x, y^* \rangle - f(x) = (y^* - L)(x) = \lambda(y^* - L)(v) = \lambda.$$

Since this holds for all $\lambda \in \mathbb{R}$, $f^*(y^*) = \infty$. Thus we have

$$f(0) = \infty > -\infty = \sup_{E^*} (0 - f^*),$$

and so 0 is not a Fenchel-Moreau point of f. (This example can also be justified using the Moreau-Rockafellar theorem on the conjugate of the sum of two convex functions.)

We now come to the promised characterization of Fenchel-Moreau points.

Theorem 3.2. *Let $f \in \mathcal{PC}(E)$ and $y \in E$. Then $f(y) \leqq \sup_{E^*} (y - f^*)$ if, and only if,*

$$\text{(3.2.1)} \qquad \begin{cases} \textit{for all } \lambda < f(y), \textit{ there exists } S \in \mathcal{S}(E) \textit{ such that} \\ \qquad x \in \operatorname{dom} f \implies f(x) + S(x - y) \geqq \lambda. \end{cases}$$

Proof. Applying Theorem 1.5 with $A := \operatorname{dom} f$ and $g(x) := x - y$, (3.2.1) is equivalent to:

$$\begin{cases} \text{for all } \lambda < f(y), \text{ there exist } S \in \mathcal{S}(E) \\ \text{and a linear functional } L \text{ on } E \text{ such that} \\ \qquad L \leqq S \text{ on } E \quad \text{and} \quad x \in \operatorname{dom} f \implies f(x) + L(x - y) \geqq \lambda, \end{cases}$$

which is exactly equivalent to:

$$\begin{cases} \text{for all } \lambda < f(y), \text{ there exists } x^* \in E^* \text{ such that} \\ \qquad x \in \operatorname{dom} f \implies f(x) + \langle y - x, x^* \rangle \geqq \lambda, \end{cases}$$

which is, in turn, equivalent to

$$\text{for all } \lambda < f(y), \text{ there exists } x^* \in E^* \text{ such that } \langle y, x^* \rangle - f^*(x^*) \geqq \lambda.$$

This completes the proof of Theorem 3.2. □

Lemma 3.3 contains a positive result on Fenchel-Moreau points (which will be subsumed by Theorem 3.4).

Lemma 3.3. *If $f \in \mathcal{PC}(E)$ is lower semicontinuous at $y \in \operatorname{dom} f$ then y is a Fenchel-Moreau point of f.*

Proof. We will establish (3.2.1), and Theorem 3.2 will then give the desired result. Let $\lambda < f(y)$. Choose $S \in \mathcal{S}(E)$ such that

$$\text{(3.3.1)} \qquad S(u - y) \leqq f(y) - \lambda \implies f(u) > \lambda.$$

Let x be an arbitrary element of E. If $S(x-y) \leqq f(y)-\lambda$ then (using (3.3.1) with u replaced by x)

$$f(x)+S(x-y) \geqq f(x) \geqq \lambda.$$

If, on the other hand, $S(x-y) > f(y)-\lambda$, let $\gamma := (f(y)-\lambda)/S(x-y) \in (0,1)$ and put $u := \gamma x+(1-\gamma)y$. Then $S(u-y) = \gamma S(x-y) = f(y)-\lambda$ and so, from (3.3.1),

$$\lambda < f(u) = f(\gamma x+(1-\gamma)y) \leqq \gamma f(x)+(1-\gamma)f(y),$$

from which

$$\gamma f(x)+f(y)-\lambda > \gamma f(y) > \gamma\lambda.$$

Substituting in the formula for γ and clearing of fractions yields:

$$f(x)+S(x-y) > \lambda.$$

Thus (3.2.1) is satisfied, which completes the proof of Lemma 3.3. □

Theorem 3.4. *If $f \in \mathcal{PC}(E)$ is lower semicontinuous at $z \in \operatorname{dom} f$ and at $y \in E$ then y is a Fenchel-Moreau point of f.*

P r o o f. We will establish (3.2.1), and Theorem 3.2 will then give the desired result. Let $\lambda < f(y)$. Lemma 3.3 gives $T \in \mathcal{S}(E)$ such that

$$x \in E \implies f(x)+T(x-z) \geqq f(z)-1, \tag{3.4.1}$$

and, by making T larger if necessary, we can also suppose that

$$f(x) \leqq \lambda \implies T(x-y) \geqq 1. \tag{3.4.2}$$

Choose M so that $M \geqq 1$ and $M \geqq \lambda - f(z)+T(y-z)+2$, and define $S := MT \in \mathcal{S}(E)$. Let x be an arbitrary element of E. If $f(x) \geqq \lambda$ then obviously

$$f(x)+S(x-y) \geqq \lambda.$$

If, on the other hand, $f(x) < \lambda$, (3.4.2) gives $T(x-y) \geqq 1$ and so, using this and (3.4.1),

$$\begin{aligned} f(x)+S(x-y) &= f(x)+MT(x-y) \\ &= f(x)+T(x-y)+(M-1)T(x-y) \\ &\geqq f(x)+T(x-z)-T(y-z)+(M-1)T(x-y) \\ &\geqq f(z)-1-T(y-z)+(M-1) \geqq \lambda. \end{aligned}$$

Thus (3.2.1) is satisfied, which completes the proof of Theorem 3.4. □

Remark 3.5. Of course, Theorem 3.4 provides a proof of the original Fenchel-Moreau formula when f is lower semicontinuous on E. This is usually proved using the Eidelheit separation theorem in $E \times \mathbb{R}$. The advantage of the method of proof given here is that we do not have to deal with the elimination of the "vertical hyperplane".

In [10], Rockafellar develops a theory of dual problems and Lagrangians that gives a very large number of results in convex analysis. We will show in Theorem 3.6(c) how Theorem 1.5 can be used to give an efficient proof of [10, Theorem 17(a), p. 41], one of the main existence results in [10]. We note, in passing, that Theorem 3.6(b) gives a *necessary and sufficient* condition for (3.6.4). (Compare [10, Theorem 16, p. 40].)

Theorem 3.6. *Let X be a real vector space, U be a real locally convex space with dual U^* and $F\colon X \times U \mapsto [-\infty, \infty]$ be convex. (P) is the "primal problem" of finding the value of $\beta := \inf_{x \in X} F(x, 0)$, and (D) is the "dual problem" of finding $\sup_{u^* \in U^*} h(u^*)$, where $h\colon U^* \mapsto [-\infty, \infty]$ is defined by*

$$h(u^*) := \inf_{(x,u) \in X \times U} [F(x, u) + \langle u, u^* \rangle].$$

Finally, let $\alpha \in [-\infty, \infty]$. We consider the four conditions

(3.6.1) *there exists $u^* \in U^*$ such that* $\inf_{(x,u) \in X \times U} [F(x, u) + \langle u, u^* \rangle] \geqq \alpha$,

(3.6.2) *there exists $S \in \mathcal{S}(U)$ such that* $\inf_{(x,u) \in X \times U} [F(x, u) + S(u)] \geqq \alpha$,

(3.6.3) *there exists $S \in \mathcal{S}(U)$ such that* $\inf_{(x,u) \in X \times U} [F(x, u) + S(u)] \geqq \beta$,

and

(3.6.4) $\inf(P) = \sup(D)$ and there exists $u^* \in U^*$ *solving* (D).

Then

(a) (3.6.1) $\iff$ (3.6.2).
(b) (3.6.3) $\iff$ (3.6.4).
(c) *Suppose that the optimal value function $u \mapsto \inf_{x \in X} F(x, u)$ is bounded above on a neighborhood of* 0*. Then* (3.6.4) *is satisfied.*

Proof. (a)($\Longrightarrow$) This follows because any element of U^* is dominated by an element of $\mathcal{S}(U)$. ($\Longleftarrow$) Since (3.6.1) is automatic if $\alpha = -\infty$, we can and will suppose that $\alpha > -\infty$. (3.6.2) now implies that $F\colon X \times U \mapsto (-\infty, \infty]$. If F is identically $+\infty$ then, again, (3.6.1) is automatic, so we can suppose that $F \in \mathcal{PC}(X \times U)$. The result now follows from Theorem 1.5 with $A := \operatorname{dom} F$ and $g(x, u) := u$, since any linear functional on U dominated by an element of $\mathcal{S}(U)$ is continuous.

If $u^* \in U^*$ then $h(u^*) \leqq \inf_{x \in X} [F(x, 0) + \langle 0, u^* \rangle] = \beta$, and so (b) is immediate from (a).

The assumption in (c) is that there exist $T \in \mathcal{S}(U)$ and $M \in \mathbb{R}$ such that

(3.6.5) $\quad T(u) < 1 \implies$ there exists $z \in X$ such that $F(z, u) \leqq M$,

and we shall show that there exists $S \in \mathcal{S}(U)$ satisfying (3.6.3) — the result then follows from (b). It is clear by taking $u = 0$ in (3.6.5) that $\beta \leqq M$. If $\beta = -\infty$ then (3.6.3) is true with $S := 0$, so we can and will suppose that $\beta \in \mathbb{R}$. Let S be the continuous seminorm $(M - \beta)T$: we shall show that (3.6.3) is true for this value of S. So let $(x, u) \in X \times U$. Let $\lambda > T(u)$. Since $T(-u/\lambda) < 1$, (3.6.5) gives $z \in X$ such that $F(z, -u/\lambda) \leqq M$. But then, from the definition of β and the convexity of F,

$$\begin{aligned}\beta &\leqq F\left(\frac{x + \lambda z}{1 + \lambda}, 0\right) \leqq \frac{1}{1 + \lambda} F(x, u) + \frac{\lambda}{1 + \lambda} F\left(z, -\frac{u}{\lambda}\right) \\ &\leqq \frac{1}{1 + \lambda} F(x, u) + \frac{\lambda}{1 + \lambda} M.\end{aligned}$$

Clearing of fractions, we obtain $F(x, u) + (M - \beta)\lambda \geqq \beta$. If we now let $\lambda \to T(u)+$ in this, we derive $F(x, u) + (M - \beta)T(u) \geqq \beta$, that is to say $F(x, u) + S(u) \geqq \beta$. Since this holds for all $(x, u) \in X \times U$, we obtain (3.6.3), completing the proof of (c). $\square$

4. An application to monotone multifunctions. We showed in [13] how the minimax theorem of Fan referred to after Theorem 2.4 can be used to obtain a large number of results on (or related to) monotone multifunctions on a Banach space. In some of these cases, these results can be obtained using Theorem 1.5 instead. (Specifically, Lemma 11.1, p. 41, Lemma 18.1, p. 65–66, Lemma 20.1, p. 77–78, Corollary 29.2, p. 114, Lemma 36.1, p. 141–142, Theorem 38.2, p. 146–147 and Theorem 38.3, p. 147–149 of [13] fall into this category.) Since Theorem 1.5 uses the sublinear functional (nearly always, a scalar multiple of the norm) directly, this alternative method of proof is not only shorter, but also avoids the use of the Banach-Alaoglu theorem. As an illustration, we give in Theorem 4.1 a proof of the least technical of the above results ([13, Lemma 11.1, p. 41]) using Theorem 1.5. Theorem 4.1 can also be established using the Debrunner-Flor extension theorem (which depends on Brouwer's fixed-point theorem, see Phelps, [8, Lemma 1.7, p. 4] and the comments preceding), or the Farkas Lemma (see Fitzpatrick-Phelps, [4, Lemma 2.4, p. 580–581]). In words, Theorem 4.1 says that if a maximal monotone multifunction on a normed space has bounded range then it has full domain.

Theorem 4.1. *Let E be a non-trivial normed space with dual E^* and S: $E \to 2^{E^*}$ be a maximal monotone multifunction with graph G. Suppose that there exists M such that, for all $(x, x^*) \in G$, $\|x^*\| \leqq M$. Then, for all $x \in E$, $Sx \neq \emptyset$.*

Proof. It follows from the monotonicity of S that there exist a convex subset A of a (large) vector space V and maps $\delta: G \mapsto A$, $p: A \mapsto E$, $q: A \mapsto E^*$ and $r: A \mapsto \mathbb{R}$ such that

(4.1.1) $\quad A$ is the convex hull of $\delta(G)$,

(4.1.2) $\quad p, q$ and r are affine,

$$(4.1.3) \qquad \begin{cases} (x, x^*) \in G \implies \\ p \circ \delta(x, x^*) = x, \ q \circ \delta(x, x^*) = x^* \\ \text{and } r \circ \delta(x, x^*) = \langle x, x^* \rangle, \end{cases}$$

and

$$(4.1.4) \qquad a \in A \implies r(a) \geqq \langle p(a), q(a) \rangle.$$

(We can take V to be the direct sum of $E \times E^*$ copies of $\mathbb{R}$, for $(x, x^*) \in G$, $\delta(x, x^*) \in V$ defined by

$$\delta(x, x^*)(s, s^*) := \begin{cases} 1 & ((s, s^*) = (x, x^*)) \\ 0 & ((s, s^*) \neq (x, x^*)), \end{cases}$$

and p, q and r, defined by $p(\mu) := \sum_{(s,s^*)\in E\times E^*} \mu(s, s^*)s$, $q(\mu) := \sum_{(s,s^*)\in E\times E^*} \mu(s, s^*)s^*$ and $r(\mu) := \sum_{(s,s^*)\in E\times E^*} \mu(s, s^*)\langle s, s^*\rangle$ — see [13, Section 9, p. 32–33].) Let x be an arbitrary element of E. Then, using (4.1.1), (4.1.2) and (4.1.3), for all $a \in A$, $\|q(a)\| \leqq M$ and so, using (4.1.4),

$$\begin{aligned} r(a) - \langle x, q(a)\rangle + M\|x - p(a)\| &\geqq r(a) - \langle x, q(a)\rangle + \langle x - p(a), q(a)\rangle \\ &= r(a) - \langle p(a), q(a)\rangle \geqq 0. \end{aligned}$$

From Theorem 1.5 with $S := M\|\ \|$, $f(a) := r(a) - \langle x, q(a)\rangle$ and $g(a) := x - p(a)$, there exists $x^* \in E^*$ such that ($\|x^*\| \leqq M$ and)

$$a \in A \implies r(a) - \langle x, q(a)\rangle + \langle x - p(a), x^*\rangle \geqq 0.$$

In particular, setting $a = \delta(s, s^*)$, it follows from (4.1.3) that

$$\begin{aligned} (s, s^*) \in G &\implies \langle s, s^*\rangle - \langle x, s^*\rangle + \langle x - s, x^*\rangle \geqq 0 \\ &\iff \langle x - s, x^* - s^*\rangle \geqq 0. \end{aligned}$$

The maximal monotonicity of S now implies that $(x, x^*) \in G$, which gives the required result. □

5. Existence theorems without a priori scalar bounds. We first state as Theorem 5.1 an existence theorem for linear functionals which appeared in [13, Theorem 7.2, p. 27–28], and was used in [13, Section 10] to obtain a number of criteria for a monotone multifunction on a reflexive Banach space to be maximal monotone, in [13, Sections 20–23] to obtain conditions for the sum of maximal monotone multifunctions on a reflexive Banach space to be maximal monotone, and in [13, Section 27] to obtain results on the closure of the range of a maximal monotone multifunction of Gossez's type (D) on an arbitrary Banach space.

The proof of Theorem 5.1 given in [13, Theorem 7.2] was quite nonconstructive, so it is not clear from the conditions assumed what would be an appropriate estimate for the value of $\|x^*\|$. One of the by-products of the analysis given here (see Remark 5.7) is an estimate for $\|x^*\|$ which (happily) coincides with the estimate obtained by working backwards from the conclusion of the theorem. One issue that might be worth pursuing is whether this new information gives any further insight into the theory of maximal monotone multifunctions.

Theorem 5.2. *Let A be a nonempty convex subset of a vector space, E be a real normed space, $g\colon A \mapsto E$ be affine and $f\colon A \mapsto \mathbb{R}$ be convex. Then* (5.1.1) $\Longleftrightarrow$ (5.1.2).

(5.1.1) $\quad a \in A \implies f(a) + \|g(a)\|^2 \geqq 0.$

(5.1.2) $\quad$ *There exists $x^* \in E^*$ such that* $\quad a \in A \implies f(a) - 2\langle g(a), x^*\rangle \geqq \|x^*\|^2.$

Note that (5.1.1) can be written $\inf_A [f + \psi \circ S \circ g] \geqq 0$, where $\psi\colon \mathbb{R} \mapsto \mathbb{R}$ is defined by $\psi(\lambda) := \lambda^2$ and $S := \|\ \|$, and $\inf_A [f + S \circ g]$ in (1.5.1) can be written $\inf_A [f + \psi \circ S \circ g]$ where $\psi\colon \mathbb{R} \mapsto \mathbb{R}$ is defined by $\psi(\lambda) := \lambda$. Thus it is natural to ask whether there is a result that simultaneously generalizes Theorem 1.5 and Theorem 5.1. Indeed, there is such a result, which we will state as Theorem 5.4. The work for Theorem 5.4 will actually be done in Lemma 5.3, which depends on Theorem 1.5.

We first consider the conditions on the function ψ that we need for Theorem 5.4.

Definition 5.2. Let $S\colon E \mapsto \mathbb{R}$ and $g\colon A \to E$. We shall say that ψ is S, g–*compatible* if $\psi \in \mathcal{PC}(\mathbb{R})$ and $\operatorname{dom}\psi \cap \bigcup_{a\in A} (S \circ g(a), \infty) \neq \emptyset$. If $\psi\colon \mathbb{R} \mapsto \mathbb{R}$ is convex (as is the case with the two examples mentioned above) then ψ is clearly S, g-compatible whatever the values of S and g.

Lemma 5.3. *Let E be a nontrivial vector space and $S\colon E \mapsto \mathbb{R}$ be sublinear. Let A be a nonempty convex subset of a vector space, $g\colon A \mapsto E$ be S-convex and $f\colon A \mapsto \mathbb{R}$ be convex. Suppose, further, that ψ is S, g-compatible and*

(5.3.1) $\quad a \in A$ *and* $S \circ g(a) \leqq \alpha \implies f(a) + \psi(\alpha) \geqq 0.$

Let

$$M := \inf\left\{\frac{f(b) + \psi(\beta)}{\beta - S \circ g(b)} \colon b \in A,\ \beta \in \operatorname{dom}\psi,\ \beta > S \circ g(b)\right\}.$$

Then there exist $\gamma \in [0, M]$ and a linear functional L on E such that

(5.3.2) $\quad L \leqq \gamma S$ *on* E $\quad$ *and* $\quad f + L \circ g \geqq \psi^*(\gamma)$ *on* A.

Proof. We first observe from (5.3.1) and the S, g–compatibility of ψ that $M \in [0, \infty)$. We now show that

(5.3.3) $\quad a \in A$ and $\alpha \in \operatorname{dom}\psi \implies f(a) + \psi(\alpha) + M(S \circ g(a) - \alpha)^+ \geqq 0.$

Let a be an arbitrary element of A, and α be an arbitrary element of $\operatorname{dom}\psi$. If $S \circ g(a) \leqq \alpha$ then, from (5.3.1),

$$f(a) + \psi(\alpha) + M(S \circ g(a) - \alpha)^+ = f(a) + \psi(\alpha) \geqq 0.$$

Suppose, on the other hand, that $S \circ g(a) > \alpha$. Let b be an arbitrary element of A, and β be an arbitrary element of $\operatorname{dom}\psi$ subject to $\beta > S \circ g(b)$. Write

$$\mu := \frac{S \circ g(a) - \alpha}{\beta - S \circ g(b)} > 0.$$

The sublinearity of S and the S-convexity of g now imply that

$$\begin{aligned}\frac{\alpha + \mu\beta}{1+\mu} &= \frac{S \circ g(a) + \mu S \circ g(b)}{1+\mu} \\ &\geqq S\left(\frac{1}{1+\mu} g(a) + \frac{\mu}{1+\mu} g(b)\right) \geqq S \circ g\left(\frac{a+\mu b}{1+\mu}\right)\end{aligned}$$

thus, using (5.3.1) for the third time (with a replaced by $(a+\mu b)/(1+\mu)$ and α replaced by $(\alpha+\mu\beta)/(1+\mu)$),

$$0 \leqq f\Big(\frac{a+\mu b}{1+\mu}\Big) + \psi\Big(\frac{\alpha+\mu\beta}{1+\mu}\Big) \leqq \frac{f(a) + \psi(\alpha) + \mu f(b) + \mu\psi(\beta)}{1+\mu}.$$

Consequently,

$$\begin{aligned}0 &\leqq f(a) + \psi(\alpha) + \mu(f(b) + \psi(\beta)) \\ &= f(a) + \psi(\alpha) + \frac{f(b) + \psi(\beta)}{\beta - S \circ g(b)}(S \circ g(a) - \alpha).\end{aligned}$$

Taking the infimum over b and β gives

$$0 \leqq f(a) + \psi(\alpha) + M(S \circ g(a) - \alpha),$$

which completes the proof of (5.3.3).

Now let $\widetilde{E} := E \times \mathbb{R}$, $\widetilde{S}\colon \widetilde{E} \mapsto \mathbb{R}$ be defined by $\widetilde{S}(x, \lambda) := M(S(x) - \lambda)^+$, $\widetilde{A} := A \times \operatorname{dom}\psi$, $\widetilde{g}\colon \widetilde{A} \mapsto \widetilde{E}$ be defined by $\widetilde{g}(a, \alpha) := (g(a), \alpha)$ and $\widetilde{f}\colon \widetilde{A} \mapsto \mathbb{R}$ be defined by $\widetilde{f}(a, \alpha) := f(a) + \psi(\alpha)$. It is easy to check that $\widetilde{S}$ is sublinear on $\widetilde{E}$, and routine but tedious to check that $\widetilde{g}$ is $\widetilde{S}$–convex. (5.3.3) can now be written

$$\widetilde{f} + \widetilde{S} \circ \widetilde{g} \geqq 0 \text{ on } \widetilde{A},$$

and so it follows from Theorem 1.5 that there exists a linear functional $\widetilde{L}$ on $\widetilde{E}$ such that $\widetilde{L} \leqq \widetilde{S}$ on $\widetilde{E}$ and

(5.3.4) $$\widetilde{f} + \widetilde{L} \circ \widetilde{g} \geqq 0 \text{ on } \widetilde{A}.$$

Let L be a linear function on E and $\gamma \in \mathbb{R}$ be chosen so that for all $(x, \lambda) \in \widetilde{E}$, $\widetilde{L}(x, \lambda) = L(x) - \gamma\lambda$. Now $-\gamma = \widetilde{L}(0, 1) \leqq \widetilde{S}(0, 1) = 0$ and $\gamma = \widetilde{L}(0, -1) \leqq \widetilde{S}(0, -1) = M$, and so $\gamma \in [0, M]$. If x is an arbitrary element of E then $L(x) - \gamma S(x) = \widetilde{L}(x, S(x)) \leqq \widetilde{S}(x, S(x)) = 0$. Thus we have proved that $L \leqq \gamma S$ on E. Finally, substituting the formulas for $\widetilde{L}$, $\widetilde{g}$ and $\widetilde{f}$ into (5.3.4) gives us that

$$a \in A \text{ and } \alpha \in \operatorname{dom}\psi \implies f(a) + \psi(\alpha) + L \circ g(a) - \gamma\alpha \geqq 0,$$

which leads rapidly to (5.3.2), and thus completes the proof of Lemma 5.3. □

Finally, we give the promised result that unifies Theorems 5.1 and 1.5. We will show how these generalizations work in Remarks 5.5 and 5.6.

Theorem 5.4. *Let E be a nontrivial vector space and $S\colon\ E \mapsto \mathbb{R}$ be sublinear. Let A be a nonempty convex subset of a vector space, $g\colon\ A \mapsto\ E$ be S-convex and $f\colon\ A \mapsto \mathbb{R}$ be convex.* (*So far, these are the conditions of Theorem* 1.5). *Suppose, further, that ψ is S, g-compatible and*

(5.4.1) $\qquad \psi$ *is nondecreasing on* $\bigcup_{a\in A}[S \circ g(a), \infty)$.

Then (5.4.2) $\Longleftrightarrow$ (5.4.3).

(5.4.2) $\qquad f + \psi \circ S \circ g \geqq 0$ on A.

(5.4.3) $\qquad \begin{cases} \text{There exist } \ \gamma \geqq 0 \text{ and a linear functional } L \text{ on } E \text{ such that} \\ \qquad\qquad\quad L \leqq \gamma S \text{ on } E \quad\text{and}\quad f + L \circ g \geqq \psi^*(\gamma) \text{ on } A. \end{cases}$

Proof. ($\Longrightarrow$) If $a \in A$ and $S \circ g(a) \leqq \alpha$ then, from (5.4.1), $f(a) + \psi(\alpha) \geqq f(a) + \psi \circ S \circ g(a)$ and so (5.3.1) follows from (5.4.2). The result now follows from Lemma 5.3. □

($\Longleftarrow$) This is immediate since, for all $a \in A$, $L \circ g(a) \leqq \gamma S \circ g(a) \leqq \psi(S \circ g(a)) + \psi^*(\gamma)$. (We note that (5.4.3) implies that $\psi^*(\gamma) \in \mathbb{R}$.)

Remark 5.5. Let us first show that Theorem 5.4 implies Theorem 5.1. We consider the implication (5.1.1) $\Longrightarrow$ (5.1.2), since the opposite implication is not difficult. So we suppose that $S = \|\ \|$ and $\psi(\lambda) := \lambda^2$. Since

$$\bigcup_{a\in A}\big[S \circ g(a), \infty\big) = \bigcup_{a\in A}\big[\|g(a)\|, \infty\big) \subset [0, \infty),$$

(5.4.1) is clearly satisfied. So if (5.1.1) (or equivalently, (5.4.2)) is satisfied then Theorem 5.4 gives $\gamma > 0$ and $L \in E^*$ such that $\|L\| \leqq \gamma$ and $f + L \circ g \geqq \gamma^2/4$ on A. Then $f + L \circ g \geqq \|L\|^2/4$ on A, and so (5.1.2) is satisfied with $x^* := -L/2$.

Remark 5.6. We next show that Theorem 5.4 implies Theorem 1.5. So we suppose that $\psi(\lambda) := \lambda$, and (5.4.1) is again satisfied. As in the proof of Theorem 1.5, we can and will suppose that $\alpha := \inf_A \big[f + S \circ g\big] \in \mathbb{R}$. Then, replacing f by $f - \alpha$, we can suppose that (5.4.2) is satisfied. Now $\psi^*(\gamma) = 0$ if $\gamma = 1$ and $\psi^*(\gamma) = \infty$ otherwise. Theorem 5.4 now gives (5.4.3), which forces $\psi^*(\gamma) < \infty$, from which $\gamma = 1$, and so $\psi^*(\gamma) = 0$. Now (5.4.3) implies (1.5.1).

Remark 5.7. Suppose, in the context of Theorem 5.4, that $\gamma \geqq 0$ and L is a linear functional on E such that $L \leqq \gamma S$ on E and $f + L \circ g \geqq \psi^*(\gamma)$ on A. Let $b \in A$ and $\beta > S \circ g(b)$. Then

$$f(b) + \gamma S \circ g(b) \geqq f(b) + L \circ g(b) \geqq \psi^*(\gamma) \geqq \gamma\beta - \psi(\beta),$$

from which it follows easily that

$$\gamma \leqq \inf\left\{\frac{f(b)+\psi(\beta)}{\beta - S\circ g(b)} \colon b\in A,\ \beta\in\operatorname{dom}\psi,\ \beta > S\circ g(b)\right\}.$$

So the value of M defined in Lemma 5.3 coincides with the a *posteriori* bound for γ obtained by working backwards from (5.4.3). In the situation of Theorem 5.1 and Remark 5.5, after some elementary calculus, we obtain the corresponding upper bound for $\|x^*\|$ to be

$$\inf_{b\in A}\left[\|g(b)\| + \sqrt{f(b)+\|g(b)\|^2}\right].$$

6. Rockafellar's surjectivity theorem. Rockafellar proved in [9, Proposition 1, p. 77–78] that if E is a non-trivial reflexive Banach space with dual E^* and duality map $J\colon E\to 2^{E^*}$, J and J^{-1} are single-valued and $T\colon E\to 2^{E^*}$ is a monotone multifunction then T is maximal monotone $\iff$ $T+J$ is surjective. Now ($\Longleftarrow$) of the above statement fails if J or J^{-1} is not single-valued (see [13, Remark 10.8, p. 39] for a discussion of this), while ($\Longrightarrow$) remains true (see [13, Theorem 10.7, p. 38]). It follows from a simple translation argument that, in order to prove that $T+J$ is surjective, it suffices to prove that there exists $y\in E$ such that $Ty+Jy\ni 0$. In Theorem 6.1 below, we give a proof of this result with an upper bound on $\|y\|$ obtained from Remark 5.7. (See also [13, Theorem 10.3, Corollary 10.4 and Theorem 10.6 p. 36–37] for characterizations of maximal monotonicity that are valid in general reflexive spaces with no restriction on J.) We mention parenthetically that the result of Rockafellar mentioned above depends on results of Browder, [1], which depend, in turn, on Brouwer's fixed-point theorem.

Theorem 6.1. *Let E be a non–trivial reflexive Banach space with dual E^* and $T\colon E\to 2^{E^*}$ be a maximal monotone multifunction. Let A, $p\colon A\mapsto E$, $q\colon A\mapsto E^*$ and $r\colon A\mapsto\mathbb{R}$ be as in the proof of Theorem* 4.1. *Then there exists $y\in E$ such that*

(6.1.1) $$\|y\| \leqq \frac{1}{2}\inf_{a\in A}\left[\|p(a)\|+\|q(a)\|+\sqrt{4r(a)+(\|p(a)\|+\|q(a)\|)^2}\right].$$

and

(6.1.2) $$Ty+Jy\ni 0.$$

P r o o f. From (4.1.4), for all $a\in A$,

$$\begin{aligned}4r(a)+(\|p(a)\|+\|q(a)\|)^2 &\geqq (\|p(a)\|+\|q(a)\|)^2+4\langle p(a),q(a)\rangle\\ &\geqq (\|p(a)\|+\|q(a)\|)^2-4\|p(a)\|\|q(a)\|\\ &= (\|p(a)\|-\|q(a)\|)^2 \geqq 0.\end{aligned}$$

Now write $F := E\times E^*$ with $\|(x,x^*)\| := \|x\|+\|x^*\|$ and, for all $a\in A$,

$$f(a) := 4r(a) \quad\text{and}\quad g(a) := (p(a),q(a)).$$

It follows from Theorem 5.1 and Remark 5.7 (with E replaced by F) that there exists $x^* \in F^*$ such that

$$(6.1.3) \qquad \|x^*\| \leqq \inf_{a \in A} \left[\|p(a)\| + \|q(a)\| + \sqrt{4r(a) + (\|p(a)\| + \|q(a)\|)^2} \right]$$

and

$$(6.1.4) \qquad a \in A \implies 4r(a) - 2\langle g(a), x^* \rangle \geqq \|x^*\|^2.$$

Now we can write $x^* = (2y^*, 2\widehat{y})$ for some $(y, y^*) \in E \times E^*$, and $\|x^*\| = 2\|y\| \vee 2\|y^*\|$. (6.1.1) follows immediately from (6.1.3). Dividing (6.1.4) by 4, we obtain

$$a \in A \implies$$
$$r(a) - \langle p(a), y^* \rangle - \langle y, q(a) \rangle \geqq \|y\|^2 \vee \|y^*\|^2 \geqq (\|y\|^2 + \|y^*\|^2)/2.$$

Now let G and $\delta: G \mapsto A$ be as in Theorem 4.1. If now $(s, s^*) \in G$ and we substitute $a = \delta(s, s^*)$, we obtain from the above and (4.1.3) that

$$(s, s^*) \in G \implies \langle s, s^* \rangle - \langle s, y^* \rangle - \langle y, s^* \rangle \geqq (\|y\|^2 + \|y^*\|^2)/2,$$

from which

$$(6.1.5) \qquad (s, s^*) \in G \implies \langle s - y, s^* - y^* \rangle \geqq (\|y\|^2 + \|y^*\|^2 + 2\langle y, y^* \rangle)/2.$$

Since $\|y\|^2 + \|y^*\|^2 + 2\langle y, y^* \rangle \geqq \|y\|^2 + \|y^*\|^2 - 2\|y\|\|y^*\| = (\|y\| - \|y^*\|)^2 \geqq 0$, (6.1.5) and the maximal monotonicity of T imply that $(y, y^*) \in G$, that is to say, $y^* \in Ty$. Substituting $(s, s^*) = (y, y^*)$ in (6.1.5) yields $\|y\|^2 + \|y^*\|^2 + 2\langle y, y^* \rangle \leqq 0$, from which it follows easily that $-y^* \in Jy$. Since $0 = y^* + (-y^*)$, (6.1.2) is now immediate. □

Remark 6.2. In this remark, we show how (6.1.2) implies (6.1.1), in other words, the bound for $\|y\|$ obtained in (6.1.1) coincides with the *a posteriori* bound. So suppose that $y \in E$ and (6.1.2) holds. Then there exists $y^* \in Ty$ such that $\|y\|^2 + \|y^*\|^2 + 2\langle y, y^* \rangle = 0$. Reversing the steps of Theorem 6.1, we find that

$$(s, s^*) \in G \implies \langle s, s^* \rangle - \langle s, y^* \rangle - \langle y, s^* \rangle \geqq (\|y\|^2 + \|y^*\|^2)/2,$$

hence, using (4.1.3),

$$(s, s^*) \in G \implies$$
$$r(\delta(s, s^*)) - \langle p(\delta(s, s^*)), y^* \rangle - \langle y, q(\delta(s, s^*)) \rangle \geqq (\|y\|^2 + \|y^*\|^2)/2.$$

Since p, q and r are affine on A, it follows from (4.1.1) that

$$a \in A \implies r(a) - \langle p(a), y^* \rangle - \langle y, q(a) \rangle \geqq (\|y\|^2 + \|y^*\|^2)/2.$$

Now $\|y^*\| = \|y\|$ and so, on completing the square, we obtain that

$$a \in A$$
$$\implies r(a) + (\|p(a)\| + \|q(a)\|)\|y\| \geqq \|y\|^2$$
$$\implies \|y\| \leqq \frac{1}{2}\left[\|p(a)\| + \|q(a)\| + \sqrt{4r(a) + (\|p(a)\| + \|q(a)\|)^2} \right].$$

Of course, (6.1.1) is an immediate consequence of this. We emphasize that Theorem 6.1 gives the *existence* of a solution y of (6.1.2) satisfying (6.1.1). The argument above shows that *all* solutions of (6.1.2) satisfy (6.1.1).

One can perform a similar extension to the computations of [13, Lemma 10.1, p. 34–35] using the technique of Theorem 6.1, and one obtains the bound:

$$\|y\| \leqq \frac{1}{\sqrt{2}} \inf_{a \in A} \left[\sqrt{\|p(a)\|^2 + \|q(a)\|^2} + \sqrt{2r(a) + \|p(a)\|^2 + \|q(a)\|^2} \right].$$

We leave it to the reader to show that the bound obtained in (6.1.1) is tighter.

References

[1] F. E. BROWDER, Nonlinear maximal monotone operators in Banach spaces. Math. Ann. **175**, 89–113 (1968).

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

[3] K. FAN, I. GLICKSBERG and A. J. HOFFMAN, Systems of inequalities involving convex functions. Proc. Amer. Math. Soc. **8**, 617–622 (1957).

[4] S. P. FITZPATRICK and R. R. PHELPS, Bounded approximants to monotone operators on Banach spaces. Ann. Inst. H. Poincaré Anal. non linéaire **9**, 573–595 (1992).

[5] J. L. KELLEY, I. NAMIOKA and co-authors, Linear Topological Spaces. Princeton-Toronto-London-Melbourne 1963.

[6] H. KÖNIG, Some Basic Theorems in Convex Analysis. In Optimization and operations research, B. Korte, ed. Amsterdam 1982.

[7] J. J. MOREAU, Fonctionelles convexes. In: Séminaire sur les équations aux derivées partielles, Lecture notes, Collège de France, Paris 1966.

[8] R. R. PHELPS, Lectures on Maximal Monotone Operators. Extracta Math. **12**, 193–230 (1997).

[9] R. T. ROCKAFELLAR, On the Maximality of Sums of Nonlinear Monotone Operators. Trans. Amer. Math. Soc. **149**, 75–88 (1970).

[10] R. T. ROCKAFELLAR, Conjugate duality and optimization. HCBMS-NSF Regional Conf. Sci. in Appl. Math. **16**, (1974).

[11] W. RUDIN, Functional analysis. New York 1973.

[12] S. SIMONS, Minimal sublinear functionals. Studia Math. **37**, 37–56 (1970).

[13] S. SIMONS, Minimax and monotonicity. LNM **1693** Berlin-Heidelberg-New York 1998.

Received: 18 October 2001; revised manuscript accepted: 25 January 2001

S. Simons
Department of Mathematics
University of California
Santa Barbara
CA 93106-3080
USA
simons@math.ucsb.edu