Pergamon

Nonlinear Analysis, Theory, Methods & Applications, Vol. 24, No. 8, pp. 1163–1181, 1995

0362-546X/95 $9.50 + .00

0362-546X(94)00186-3

ON A THEOREM OF DANSKIN WITH AN APPLICATION TO A THEOREM OF VON NEUMANN–SION

PIERRE BERNHARD and ALAIN RAPAPORT

INRIA-Sophia-Antipolis, 2004 route des lucioles, B.P.93, 06902 Sophia-Antipolis Cedex, France

(*Received* 23 *November* 1992; *received in revised form* 17 *August* 1993; *received for publication* 17 *August* 1994)

Key words and phrases: Minimax, differentration, convex analysis, saddlepoint.

1. INTRODUCTION

1.1. *The problem considered and related work*

In a book published in 1967 Danskin [1], proves the following theorem.

Hypotheses. Let V be a compact toplogical space, and J a map from $\mathbb{R}^n \times V$ into $\mathbb{R}$, assumed to be jointly continuous, and C^1 w.r.t. the first variable. Let

$$\bar{J}(u) = \max_{v \in V} J(u, v),$$

and

$$\hat{V}(u) = \{v \in V \mid J(u, v) = \bar{J}(u)\}.$$

The theorem is as follows.

THEOREM 1 (Danskin). The function $\bar{J}$ has, for every u and h in $\mathbb{R}^n$ a directional derivative at u in the direction h given by

$$D\bar{J}(u; h) = \max_{v \in \hat{V}} \sum_{i=1}^{n} h_i J_i(u, v),$$

where J_i stands for the partial derivative w.r.t. the component u_i of u.

Let $D_1 J(u, v; h)$ represent the directional partial derivative of J w.r.t. its first variable in the direction h, so that the above formula can be written

$$D\bar{J}(u; h) = \max_{v \in \hat{V}} D_1 J(u, v; h). \tag{1.1}$$

Since 1967, much work has been devoted to improve this result, or to related ones. There have been two main directions of research, one in the domain of convex analysis, and the other for nonconvex nondifferentiable functions.

Early work in the first area is described in Valadier's contribution [2]. A recent account can be found in Aubin and Ekeland [3] or Aubin [4]. This last reference, for instance, contains in the theorem 4.4, p. 53, exactly our theorem C1 below. Most of the literature has concentrated on the infinite dimension, as we do. However, there has been little work carried out on the compacity assumption. Valadier's work is a notable exception, and although rarely quoted, supersedes many later accounts. We applied his formula to the Von Neumann–Sion theorem. While Valadier needed to look at the subdifferential at neighboring points, as well as the generalized subdifferentials work quoted in the next paragraph, we propose instead a set of hypotheses with more regularity (mainly *uniformity*). (Furthermore, the simple "convex–concave" result we obtained, somewhat ad hoc for the application to the problem of Von Neumann, does not seem to have been pointed out before.)

Although early work needed the differentiable hypothesis (a form can be found, for instance, in [5, lemma 15.1, p. 53]) later work concentrated on the use of generalized (sub)differentials, such as Clarke's. A typical case of such results can be found in Rockafellar [6]. These results are always in finite dimensional spaces, and lead to estimations of a generalized subdifferential (generally Clarke's subdifferential) of the upper envelope, i.e. supersets, while we concentrate on exact expressions, in infinite dimension spaces. More importantly, this work proves the existence of a generalized subdifferential while we have results giving the existence of ordinary directional derivatives. The price to be paid is that our results require more regularity, (again, uniformity), and in particular do not deal in detail with infinite slopes, and singular subdifferentials. Again, we give results without compactness, particularly useful in the context of infinite dimensional spaces, that do not seem to have been considered before.

Related work on the so-called sensitivity optimization function, where what is sought is the derivative of a constrained max (or min) with respect to a variable occurring in the constraint of the optimization problem, should also be mentioned. Typically, the derivative or subdifferential of $\hat{J}(u)$ is defined by

$$\tilde{V}(u) = \{v \in V \mid A(u,v) \leq 0\},$$

or, more specifically

$$\tilde{V}(u) = \{v \in V \mid A(v) \leq u\},$$

and

$$\hat{J}(u) = \max_{v \in \tilde{V}(u)} J(v).$$

The two problems are very closely connected: one simple way of observing this is to rewrite the latter as follows. Let

$$\tilde{J}(u,v) = J(v) - \chi_{\tilde{V}(u)}(v),$$

$$\hat{J}(u) = \max_{v \in V} \tilde{J}(u,v)$$

which is the form of problem considered. (Although this identification yields results in convex analysis it does not do so in differentiable analysis.)

Results on this problem most often fall into the second category above, typical examples being [7] and [8]. Let us quote also the earlier paper [9], which dealt with the more general problem where the variable u enters both the function to be maximized and the constraint. This paper falls into the category mentioned: finite dimension, and inclusions, very much in the spirit of those in the later papers quoted, which improve the results by weakening the regularity hypothesis, and by treating the singular subdifferential.

A somewhat different issue was taken up by Bonnans [10], who was able to relax all qualification hypotheses for the constraint, at the cost of having directional derivatives in certain directions only. This is still in finite dimension.

Finally, let us quote the notable exception of [11], which deals with the same general problem as [9], obtaining estimations of generalized subdifferentials, including the singular case, but in infinite dimension. However, his results depended on a hypothesis on the regularity of the solution of the optimization problem, and the only sufficient conditions we know to satisfy this hypothesis have been stated in a finite dimensional space. (Although we could not check one of his references (Dolecki, "to appear").)

Finally, the link we show between these results and the Von Neumann–Sion theorems seems to be new.

This paper is based on a previous internal report [12]. A specialized version of the "differentiable" theorem can be found in [13, Chapter 9].

1.2. *General framework*

The following framework holds throughout the paper, and will not be repeated in the sequel.

U and V are subsets of a Banach space $\mathscr{U}$ and a topological space $\mathscr{V}$, respectively. J is a mapping from $U \times V$ into $\mathbb{R}$. The directional derivative of $u \mapsto J(u, v)$ in a direction h of $\mathscr{U}$ is denoted by $D_1 J(u, v; h)$, and its subdifferential, in the case of a convex function, by $\partial_1 J(u, v)$. Let

$$\bar{J}(u) = \sup_{v \in V} J(u, v), \tag{1.2}$$

and, when it exists

$$\hat{V}(u) = \{v \in V \mid J(u, v) = \bar{J}(u)\}. \tag{1.3}$$

We will also need to consider

$$\mathscr{W}(u) = \left\{\{v_n\} \mid J(u, v_n) \to \bar{J}(u) \text{ as } n \to \infty\right\},$$

the set of maximizing sequences $\{v_n\}$ at u.

Our aim is to characterize the directional derivatives $D\bar{J}(u; h)$ of $\bar{J}$, or, when it is convex, its subdifferential $\partial \bar{J}(u)$.

2. THE DIFFERENTIABLE CASE

2.1. *V compact*

Firstly, a slightly improved version of theorem 1 above is stated.

Hypotheses D1.

D1.0. V is compact.

D1.1. $\forall v \in V$, the application $(t, v) \mapsto J(u + th, v)$ is upper semi-continuous (u.s.c.) at $(0, v)$.

D1.2. $\forall v \in V$ and $\forall t$ in a right neighborhood of 0, there exists a bounded directional derivative

$$D_1(u+th,v;h) = \lim_{\tau \to 0^+} \frac{1}{\tau}[J(u+(t+\tau)h,v) - J(u+th,v)].$$

D1.3. Moreover, the map $(t,v) \mapsto D_1 J(u+th,v)$ is supper semi-continuous at $(0,v)$.

THEOREM D1. Under hypotheses D.1, the function $\bar{J}$ has a directional derivative at u in the direction h, given by the formula

$$D\bar{J}(u;h) = \max_{u \in \hat{V}(v)} D_1 J(u,v;h).$$

Proof. It should be noticed that by the assumption D1.3, the map $v \mapsto D_1 J(u,v;h)$ is u.s.c., so that, V being compact, the maximum is reached.

Let, for convenience,

$$\Delta(t) = \frac{1}{t}[\bar{J}(u+th) - \bar{J}(u)]. \tag{2.1}$$

PROPOSITION 1. One has

$$\liminf_{t \to 0} \Delta(t) \geq \max_{v \in \hat{V}(u)} D_1 J(u,v;h).$$

Proof of the proposition. Let $\hat{v} \in \hat{V}(u)$. By definition, $\bar{J}(u) = J(u,\hat{v})$, and $\bar{J}(u+th) \geq J(u+th,\hat{v})$. Thus

$$\Delta(t) \geq \frac{1}{t}[J(u+th,\hat{v}) - J(u,\hat{v})].$$

Taking the lim inf,

$$\liminf_{t \to 0} \Delta(t) \geq D_1 J(u,\hat{v}),$$

and since this holds for any $\hat{v}$ in $\hat{V}(u)$, the proposition is proved.

PROPOSITION 2. Let $\{t_n\}$ be a sequence of real positive numbers, going to zero, and for all $n, v_n \in \hat{V}(u+t_n h)$. Then

$$v_n \to \hat{V}(u), \qquad \text{and} \qquad J(u+t_n h, v_n) \to \bar{J}(u).$$

(The map $t \mapsto \hat{V}(u+th)$ is said to be u.s.c. at 0.)

Proof of the proposition. Proposition 1 implies that $\Delta(t_n)$ is bounded below: $\exists a$ such that, for n large enough, $\Delta(t_n) \geq a$. Thus also

$$\bar{J}(u+t_n h) \geq \bar{J}(u) + a t_n.$$

Therefore, $\liminf \bar{J}(u+t_n h) \geq \bar{J}(u)$. Now, V is compact. Let, therefore, $\bar{v}$ be a cluster point of the sequence $\{v_n\}$. One has

$$\bar{J}(u) \geq J(u,\bar{v}) \geq \limsup J(u+t_n h, v_n) \geq \liminf J(u+t_n h, v_n) \geq \bar{J}(u).$$

The first inequality follows from the definition of $\bar{J}$, the second one from the hypothesis D1.1, the last one from what has just been mentioned. Thus all inequalities are equalities, from which it can be concluded that $\bar{v} \in \hat{V}(u)$, and the existence of the limit $\lim J(u + t_n h, v_n) = \bar{J}(u)$.

PROPOSITION 3.

$$\limsup_{t \to 0} \Delta(t) \leq \max_{v \in \hat{V}(u)} D_1 J(u, v; h).$$

Proof of the proposition. With the same notations as in proposition 2, one has

$$\Delta(t_n) = \frac{1}{t_n}[J(u + t_n h, v_n) - J(u, v_n)] + \frac{1}{t_n}\left[J(u, v_n) - \bar{J}(u)\right].$$

By definition of $\bar{J}$, the second term is nonpositive, hence

$$\Delta(t_n) \leq \frac{1}{t_n}[J(u + t_n h, v_n) - J(u, v_n)].$$

The function $t \mapsto J(u + th, v_n)$ having for all $t \in [0, t_n]$ a bounded directional derivative, it is absolutely continuous, and there exists $t'_n \in [0, t_n]$ such that

$$D_1 J(u + t'_n h, v_n; h) \geq \frac{1}{t_n}[J(u + t_n h, v_n) - J(u, v_n)],$$

hence $\Delta(t_n) \leq D_1 J(u + t'_n h, v_n; h)$. Due to the hypothesis D1.3, taking the lim sup

$$\limsup \Delta(t_n) \leq D_1 J(u, \bar{v}; h),$$

where $\bar{v}$ is any cluster point of the sequence $\{v_n\}$. Using Proposition 2, $\bar{v} \in \hat{V}(u)$, and thus a fortiori the result claimed.

Finally, propositions 1 and 3 together prove the theorem. ■

COROLARY If $u \mapsto J(u, v)$ has a Gâteaux derivative J'_u, and if the max is unique: $\hat{V}(u) = \{\hat{v}\}$, then J has a Gâteaux derivative $J'(u)$ given by the simple formula

$$\bar{J}'(u) = J'_u(u, \hat{v}).$$

Proof. It follows from Theorem 1 that, since $D_1 J(u, v; h) = J'_u(u, v).h$, then

$$D\bar{J}(u; h) = J'_u(u, \hat{v}).h.$$

This equality proves the claim. ■

2.2. *Uniform case*

The compacity hypothesis on V can be traded for more regularity on J, for instance in the following way (u and h are as in hypothesis D1).

Hypotheses D2.

D2.1. The map $u \mapsto J(u, v)$ is uniformly directionally differentiable in the following sense

$$\forall \epsilon > 0, \exists \tau > 0: \forall t \in (0, \tau), \quad \forall v \in V,$$

$$\left|\frac{1}{t}[J(u + th, v) - J(u, v)] - D_1 J(u, v; h)\right| \leq \epsilon.$$

D2.2. The directional derivative $D_1J(u+th,v;h)$ is bounded in a right neighborhood of 0 in t, uniformly in $v \in V$.

D2.3. The map $t \mapsto D_1J(u+th,v;h)$ is u.s.c. at 0, uniformly in $v \in V$.

Remark. Hypothesis D2.1 and D2.3 may be lumped into any of the following two stronger hypotheses:

D2.a. The map $u \mapsto J(u,v)$ is uniformly directionally differentiable in the stronger following sense: for $\lambda > 0$, write $u_\lambda = u + \lambda h$. The hypothesis reads

$$\exists \theta > 0 : \forall \epsilon > 0, \exists \tau > 0 : \forall t \in (0,\tau), \forall \lambda < \theta, \forall v \in V,$$

$$\left| \frac{1}{t}[J(u_\lambda + th, v) - J(u_\lambda, v)] - D_1J(u_\lambda, v; h) \right| \le \epsilon.$$

D2.b At point u, J has a second directional derivative with respect to its first variable in the direction h, uniformly bounded in v.

THEOREM D2. Under hypotheses D2, for all t in a left neighborhood of 0, there exits $\bar{J}(u+th) < \infty$, and $\bar{J}$ has a directional derivative in the direction h, given by

$$D\bar{J}(u;h) = \sup_{\{v_k\} \in \mathscr{W}(u)} \limsup_{k \to \infty} D_1J(u, v_k; h).$$

Remark. It could be agreed, with no ambiguity, to simply write the r.h.s. above as

$$D\bar{J}(u;h) = \limsup_{\{v_k\} \in \mathscr{W}(u)} D_1J(u, v_k; h).$$

Proof. Let us call D the r.h.s. of the above equality, and let us define $\Delta(t)$ as in (1.1). In the sequel, it has been selected two sequences $\{t_n\}$ and $\{\epsilon_n\}$ of positive numbers such that $t_n \to 0$ and $\epsilon_n / t_n \to 0$ as $n \to \infty$. (say, e.g. $\epsilon_n = t_n^2$).

PROPOSITION 1. $\liminf \Delta(t_n) \ge D$.

Proof of the proposition. Let δ be a positive integer. Choose N such that $\forall n > N$,

(i) $$\frac{\epsilon_n}{t_n} < \frac{\delta}{3},$$

and

(ii) $$\forall v \in V, \quad \frac{1}{t_n}[J(u+t_nh, v) - J(u,v)] \ge D_1J(u,v;h) - \frac{\delta}{3}.$$

This is possible due to hypothesis D2.1. Let also $\{v_k\} \in \mathscr{W}$ be a maximizing sequence at u,

$$\forall n, \exists K_n : \forall k > K_n, \qquad J(u, v_k) \ge \bar{J}(u) - \epsilon_n.$$

Hence, $\forall n > N, \forall K > K_n$,

$$\Delta(t_n) \ge \frac{1}{t_n}\left[\bar{J}(u+t_nh) - J(u,v_k)\right] - \frac{\epsilon_n}{t_n} \ge \frac{1}{t_n}[J(u+t_nh, v_k) - J(u,v_k)] - \frac{\delta}{3}.$$

By (ii), $\forall k > K_n$, $\Delta(t_n) \geq D_1 J(u, v_k) - 2\delta/3$. Let k go to infinity to conclude that

$$\Delta(t_n) \geq \limsup D_1 J(u, v_k; h) - 2\delta/3.$$

However, since $\{v_k\}$ is an arbitrary maximizing sequence, it may be chosen such that

$$\limsup D_1 J(u, v_k; h) \geq D - \delta/3.$$

This way it gives $\forall n > N$, $\Delta(t_n) > D - \delta$, and this proves the proposition given that δ was an arbitrary positive number.

PROPOSITION 2. Let $\{v_n\}$ be a sequence in V such that

$$\forall n > 0, \qquad J(u + t_n h, v_n) \geq \bar{J}(u + t_n h) - \epsilon_n. \tag{2.2}$$

Then $\{v_n\} \in \mathscr{W}(u)$.

Proof of the proposition. One has

$$\bar{J}(u) \geq J(u, v_n) \geq J(u + t_n h, v_n) - t_n D_1 J(u, v_n; h) - t_n \eta_n,$$

where $\eta_n \to 0$ by hypothesis D2.1. $D_1 J$ being bounded by the hypothesis D2.2, it also gives

$$\bar{J}(u) \geq J(u, v_n) \geq J(u + t_n h, v_n) - \delta_n, \qquad \delta_n \to 0.$$

Making use of the definition (1.2) of the sequence $\{v_n\}$, and of proposition 1 that implies that $\bar{J}(u + t_n h) \geq \bar{J}(u) + \gamma_n$ where $\gamma_n \to 0$, finally gives

$$\bar{J}(u) \geq J(u, v_n) \geq \bar{J}(u) - \delta_n - \epsilon_n + \gamma_n \to \bar{J}(u),$$

which proves the proposition.

PROPOSITION 3. $\limsup \Delta(t_n) \leq D$.

Proof of the proposition. The sequence $\{v_n\}$ is still as in (2.2). By definition, one has

$$\Delta(t_n) \leq \frac{1}{t_n} \left[J(u + t_n h, v_n) - \bar{J}(u) \right] + \frac{\epsilon_n}{t_n} \leq \frac{1}{t_n} [J(u + t_n h, v_n) - J(u, v_n)] + \frac{\epsilon_n}{t_n}. \tag{2.3}$$

As in the proof of theorem D1, by hypothesis D2.2, there exists $t'_n \in [0, t_n]$ such that

$$D_1 J(u - t'_n h, v_n; h) \geq \frac{1}{t_n} [J(u + t_n h, v_n) - J(u, v_n)].$$

Moreover, making use of hypothesis D2.3, for n large enough

$$D_1 J(u + t'_n h, v_n; h) \geq D_1 J(u, v_n; h) + \eta_n, \qquad \eta_n \to 0,$$

so that, making further use of (2.3)

$$\Delta(t_n) \leq D_1 J(u, v_n; h) + \frac{\epsilon_n}{t_n} + \eta_n,$$

and taking a lim sup

$$\limsup \Delta(t_n) \leq \limsup D_1 J(u, v_n; h) \leq D.$$

This proves the proposition, because due to proposition 2, $\{v_n\} \in \mathscr{W}(u)$.

Finally, propositions 1 and 3 together prove the theorem. ■

3. THE CONVEX CASE

Versions in convex analysis of the preceding two theorems are now given. They are closely connected to them by the remark that for a convex function f, the map $h \mapsto Df(u; h)$ is the support function of its subdifferential $\partial f(u)$. Thus the two theorems with compacity have identical conclusions. However, slight differences in the regularity requirements seem to prevent the "convex" theorems from being strict corollaries of the "differentiable" ones.

As it has been pointed out, the first theorem below is not new, (see [2, 4]). The proof given below is not as elegant as in these references. It has been chosen on the one hand to parallel the proofs in the differentiable case, and on the other hand to prepare the stage for the proof of the theorem without compacity, which seems to be original.

3.1. *V compact*

Hypotheses C.

C0. V is (sequentially) compact in a topology for which $\forall u \in U$, the map $v \mapsto J(u, v)$ is u.s.c.

C1. U is convex and $\forall v \in V$, the function $u \mapsto J(u, v)$ is convex. Let us denote as $\partial_1 J(u, v)$ its subdifferential.

C2. There exists $u_0 \in U$, a neighborhood $\tilde{U}$ of u_0 and a real number a such that $\forall (u, v) \in \tilde{U} \times V$, $J(u, v) \leq a$.

Let notice that hypothesis C2 implies that, $\forall \tilde{u} \in \tilde{U}$, $\bar{J}(\tilde{u}) \leq a$. Therefore, it will be introduced in the following definition.

Definition. Let U_0 be the interior of the subset of U, where $\bar{J}$ is finite.

LEMMA 1. In the presence of the hypothesis C1, the hypothesis C2 is equivalent to the following hypothesis C2a.

C2a Let $u_0 \in U$. There exists a (bounded) neighborhood $\tilde{U}$ of u_0 and a real number b such that

$$\forall \tilde{u} \in \tilde{U}, \forall v \in V, \exists \tilde{p} \in \partial_1 J(\tilde{u}, v) \qquad \text{with } \|\tilde{p}\| \leq b.$$

Proof of the lemma. Let us show that the hypothesis C2a and $\bar{J}(u_0) \leq \infty$ imply C2. Let η be such that $\tilde{u} \in \tilde{U}$ imply $\|\tilde{u} - u\| \leq \eta$. One has, $\forall (\tilde{u}, v) \in U \times V$, and with $\tilde{p} \in \partial_1 J(\tilde{u}, v)$, chosen such that $\|\tilde{p}\| \leq b$,

$$\bar{J}(u_0) \geq J(u_0, v) \geq J(\tilde{u}, v) - (\tilde{p}, \tilde{u} - u_0)$$

hence

$$J(\tilde{u}, v) \leq \bar{J}(u_0) + (\tilde{p}, \tilde{u} - u_0) \leq \bar{J}(u_0) + b\eta.$$

The converse is elementary, taking $\tilde{U}$ in C2a strictly included in $\tilde{U}$ in C2.

THEOREM C1. Under the hypothesis C, the function $\bar{J}$ is convex continuous over U_0, and its subdifferential at $u \in U_0$ is given by the formula

$$\partial \bar{J}(u) = \overline{co} \bigcup_{v \in \hat{V}(u)} \partial_1 J(u,v).$$

Proof. Let us first notice that being the upper envelope of a family of convex functions, $\bar{J}$ is itself convex. According to C2, it is bounded in a neighborhood of u_0, and thus also over U_0, providing a uniform upperbound of $J(u,v)$ in the neighborhood of every point of U_0. Thus $\partial \bar{J}$ and $\partial_1 J$ exist over that set, and by compacity of V, $\hat{V}(u)$ exists, so that the above formula has a meaning.

Notice also that the classical proof of the continuity of a locally bounded convex function also proves the uniformity in v of the continuity of $u \mapsto J(u,v)$, since the upper bound is uniform. Then it easily follows, making use of C0, that the map $(u,v) \mapsto J(u,v)$ is u.s.c.

PROPOSITION 1. One has

$$\partial \bar{J}(u) \supset \overline{co} \bigcup_{v \in \hat{V}(u)} \partial_1 J(u,v).$$

Proof of proposition 1. Let $\hat{v} \in \hat{V}(u)$, and $p \in \partial_1 J(u,\hat{v})$. Then

$$\forall w \in U, \qquad J(w,\hat{v}) \geq J(u,\hat{v}) + (p, w-u) = \bar{J}(u) + (p, w-u),$$

and thus

$$\forall w \in U, \qquad \bar{J}(w) \geq \bar{J}(u) + (p, w-u),$$

i.e. $p \in \partial \bar{J}(u)$. Since $\hat{v}$ was arbitrary in $\hat{V}$, and $p \in \partial_1 J(u,\hat{v})$, we infer that $\partial \bar{J}(u)$ includes the union of the subdifferentials $\partial_1 J$. Finally, a subdifferential being convex, this proves the proposition.

PROPOSITION 2. Let $h \in U - u$ and $t_n \to 0^+$ (or $t_n \searrow 0$) when $n \to \infty$, and $v_n \in \hat{V}(u + t_n h)$. Then $v_n \to \hat{V}(u)$.

Proof of proposition 2. Since V is compact, the sequence v_n has at least one cluster point $\bar{v}$. Let $\hat{v} \in \hat{V}(u)$. This gives

$$J(u,\bar{v}) \geq \limsup J(u + t_n h, v_n) \geq \liminf J(u + t_n h, \hat{v}) = J(u,\hat{v}) = \bar{J}(u).$$

The first inequality because of the semicontinuity of J, the second by the definition of v_n. The continuity of J in u and the definition of $\hat{v}$ give the two equalities. Therefore, $J(u,\bar{v}) = \bar{J}(u)$, and this proves the proposition.

PROPOSITION 3. Let h, t_n and v_n be as above, and $p_n \in \partial_1 J(u + t_n h, v_n)$. There exists $\hat{v} \in \hat{V}(u)$ such that

$$\limsup (p_n, h) \leq \sup_{p \in \partial_1 J(u,\hat{v})} (p,h) = D_1 J(u,\hat{v};h).$$

Proof of the proposition. Let $L = \lim\sup(p_n, h)$, and p_m be a subsequence such that $(p_m, h) \to L$. Let also $v_m \in \hat{V}(u + t_m h)$, and again a subsequence with index k such that $v_k \to \hat{v} \in \hat{V}(u)$. Let us write $D = D_1 J(u, \hat{v}; h)$.

Let $\epsilon > 0$ be fixed. The slope $[J(u + th, \hat{v}) - J(u, \hat{v})]/t$ being, for a convex function, decreasing as t decreases to 0, this function has a directional derivative, and

$$\exists \tau > 0 : \forall t < \tau, \qquad J(u + th, \hat{v}) < J(u, \hat{v}) + t(D + \epsilon). \tag{3.1}$$

On the other hand, one always has

$$\forall t, \qquad J(u + t_k h + th, v_k) \geq J(u + t_k h, v_k) + t(p_k, h) \geq J(u + t_k h, \hat{v}) + t(p_k, h).$$

Taking the lim sup, and taking into account the fact that J is u.s.c.,

$$J(u + th, \hat{v}) \geq J(u, \hat{v}) + t \lim\sup(p_k, h) = J(u, \hat{v}) + tL.$$

Comparing this last inequality with (3.1), for $t < \tau$, it gives $L < D + \epsilon$, and this proves the proposition.

PROPOSITION 4.

$$\sup_{\bar{p} \in \partial \bar{J}(u)} (\bar{p}, h) \leq \sup_{\substack{v \in \hat{V}(u) \\ p \in \partial_1 J(u, v)}} (p, h).$$

Proof of the proposition. The subdifferential is a monotonous operator

$$\forall \bar{p}_n \in \partial \bar{J}(u + t_n h), \qquad \forall \bar{p} \in \partial \bar{J}(u), \qquad (\bar{p}_n, h) \geq (\bar{p}, h).$$

Specifically,

$$\inf_{\bar{p}_n \in \partial \bar{J}(u + t_n h)} (\bar{p}_n, h) \geq \sup_{\bar{p} \in \partial \bar{J}(u)} (\bar{p}, h).$$

Moreover, making use of proposition 1 at $u + t_n h$, it gives, with the same notations as above

$$\inf_{\bar{p}_n \in \partial \bar{J}(u + t_n h)} (\bar{p}_n, h) \leq \inf_{p_n \in \partial_1 J(u + t_n h, v_n)} (p_n, h).$$

Regroup the two inequalities to obtain

$$\sup_{\bar{p} \in \partial \bar{J}(u)} (\bar{p}, h) \leq (p_n, h) \qquad \forall p_n \in \partial_1 J(u + t_n h, v_n).$$

Making use of proposition 3, it can be infered that there exists $\hat{v} \in \hat{V}(u)$ such that

$$\sup_{\bar{p} \in \partial \bar{J}(u)} (\bar{p}, h) \leq \sup_{p \in \partial_1 J(u, \hat{v})} (p, h),$$

and a fortiori the inequality claimed.

Now, proposition 4 implies the inclusion opposite to that proved in proposition 1, and the two together prove the equality claimed in the theorem. ■

3.2. *The case without compacity*

Let us adopt hypotheses C1, C2, and D2.1, where it is recalled that the directional derivative can be seen as the support of the subdifferential. It will be seen further what can be said without hypothesis D2.1, which is not very natural in this context.

The lemma 1 still holds, with the following precisions.

LEMMA 2. Hypothesis C2 is implied by C2a and the hypothesis that there exists $\bar{J}(u_0) \leq \infty$. Moreover, the hypotheses C1 and C2 imply that, if $u_n \to u \in U_0$, $\{v_n\} \in \mathscr{W}(u)$ and $p_n \in \partial_1 J(u_n, v_n)$, then $\limsup \|p_n\| \leq \infty$.

Proof of the lemma. The first claim has been proved in lemma 1. Let now u be fixed in U. Let us recall that by C2, J is continuous at u, uniformly in v. Let $\rho > 0$ be such that the ball $B(u, 2\rho)$ be included in $\tilde{U}$. Let $h \in U$, with $\|h\| = \rho$. For n large enough, $u_n + h \in \tilde{U}$, and thus

$$a \geq J(u_n + h, v_n) \geq J(u_n, v_n) + (p_n, h) \geq J(u, v_n) - \epsilon_n + (p_n, h),$$

where ϵ_n goes to zero independently of v_n due to the uniform continuity in u. Then, taking into account the fact that, by hypothesis, $\{v_n\} \in \mathscr{W}(u)$,

$$a \geq \bar{J}(u) - \eta_n - \epsilon_n + (p_n, h),$$

where again, $\eta_n \to 0$. Therefore,

$$(p_n, h) \leq a - \bar{J}(u) + \epsilon_n + \eta_n,$$

whence,

$$\limsup \|p_n\| \leq \frac{a - \bar{J}(u)}{\rho}.$$

In fact, otherwise, it could be chosen $\delta > 0$ such that $\limsup \|p_n\| > (a - \bar{J}(u))/\rho + 2\delta$, N such that for $n > N$, $\epsilon_n + \eta_n < \delta\rho$, and $k > N$ such that $\|p_k\| > (a - \bar{J}(u))/\rho + \delta$. Let then $\bar{u}_k \in U$ of unit norm such that $(p_k, \bar{u}_k) = \|p_k\|$, taking $h = \rho\bar{u}_k$, one obtains a contradiction with the above inequality.

In order to simplify the statement of the next theorem, let us introduce the following natural definition.

Definition. Let $\mathscr{U}'$ be the topological dual space of $\mathscr{U}$, and $\{\mathscr{D}_n\}$ a sequence of subsets of $\mathscr{U}'$. Let define $\limsup \mathscr{D}_n$ as the set of all limits in the weak-star topology of sequences $\{d_n\}$ of elements of $\mathscr{D}_n$

$$\limsup_{n \to \infty} \mathscr{D}_n = \left\{ d \mid \exists d_n \in \mathscr{D}_n : d_n \xrightarrow{*} d \right\} = \bigcap_{k=1}^{\infty} \overline{\bigcup_{n=k}^{\infty} \partial_1 J(u, v_n)}.$$

(The closure operator in the last expression being in the sense of the weak-star topology.)

It can now be stated the next theorem.

THEOREM C2. Under hypothesis C1 and C2, the function $\bar{J}$ has at every $u \in U_0$ a subdifferential given by the following formula

$$\partial \bar{J}(u) = \overline{co} \bigcup_{\{v_n\} \in \mathscr{W}(u)} \limsup_{n \to \infty} \partial_1 J(u, v_n)$$

or, equivalently,

$$\partial \bar{J}(u) = \overline{co} \bigcup_{\{v_n\} \in \mathscr{W}(u)} \bigcap_{k=1}^{\infty} \overline{\bigcup_{n=k}^{\infty} \partial_1 J(u, v_n)}.$$

(See another formulation after the proof.)

Proof. As in theorem C1, $\bar{J}$ is convex continuous over U. Let us use the notation

$$\mathscr{D} = \bigcup_{\{v_n\}\in\mathscr{W}(u)} \limsup_{n\to\infty} \partial_1 J(u, v_n).$$

According to lemma 2, $\mathscr{D}$ is bounded.

PROPOSITION 1. $\partial\bar{J}(u) \supset \overline{co}\mathscr{D}$.

Proof of the proposition. Let $\tilde{p} \in \mathscr{D}$. By definition, there exist a maximizing sequence $\{v_k\}$ and a sequence $p_k \in \partial_1 J(u, v_k)$ such that $p_k \overset{*}{\rightharpoonup} \tilde{p}$. It gives, $\forall h$

$$\bar{J}(u+h) \geq J(u+h, v_k) \geq J(u, v_k) + (p_k, h) = \bar{J}(u) - \epsilon_k + (p_k, h),$$

where $\epsilon_k \to 0$, whence, taking the limit

$$\bar{J}(u+h) \geq \bar{J}(u) + (\tilde{p}, h),$$

thus $\mathscr{D} \subset \partial\bar{J}(u)$, but as the latter is convex, the proposition is proved.

PROPOSITION 2. Let $t_n \searrow 0$, $\epsilon_n \searrow 0$, and v_n such that $J(u + t_n h, v_n) \geq \bar{J}(u + t_n h) - \epsilon_n$ and, finally, $p_n \in \partial_1 J(u + t_n h, v_n)$. Then $\{v_n\} \in \mathscr{W}(u)$.

Proof of the proposition. One has

$$\bar{J}(u) \geq J(u, v_n) \geq J(u + t_n h, v_n) - t_n(p_n, h) \geq \bar{J}(u + t_n h) - \epsilon_n - t_n(p_n, h).$$

Let $p \in \partial\bar{J}(u)$. (It has been seen that it is not empty.) Using it to upperbound the last occurrence of J above, it can be easily obtained

$$\bar{J}(u) \geq J(u, v_n) \geq \bar{J}(u) - t_n(p, h) - \epsilon_n - t_n(p_n, h).$$

By the lemma 2, p_n is bounded, hence the proposition.

PROPOSITION 3. Let t_n, ϵ_n, v_n be as in proposition 2. Let, furthermore,

$$D = \sup_{\tilde{p}\in\mathscr{D}} (\tilde{p}, h) \quad \text{and} \quad D_n = \sup_{p_n \in \partial_1 J(u, v_n)} (p_n, h).$$

Then $\limsup D_n \leq D$.

Proof of the proposition. For all n, one can choose $\hat{p}_n \in \partial_1 J(u, v_n)$ such that $D_n \geq (\hat{p}_n, h) \geq D_n - \epsilon_n$. Thus

$$\limsup D_n = \limsup(\hat{p}_n, h).$$

Extracting a subsequence $\{\hat{p}_k\}$ of $\{\hat{p}_n\}$ such that $(\hat{p}_k, h) \to \limsup(\hat{p}_n, h)$, and again a weak-star convergent subsequence converging to, say, $\tilde{p} \in \mathscr{D}$, it gives

$$\limsup(\hat{p}_n, h) = (\tilde{p}, h) \geq D.$$

which proves the proposition.

PROPOSITION 4. Let $t_n \searrow 0$, and for each n, $\{v_n^k\}_k \in \mathscr{W}(u + t_n h)$. Let

$$\mathscr{D}_n = \limsup_{k\to\infty} \partial_1 J(u + t_n h, v_n^k).$$

Then, if $\tilde{p}_n \in \mathscr{D}_n$, one has

$$\limsup(\tilde{p}_n, h) \le D = \sup_{\tilde{p}\in\mathscr{D}} (\tilde{p}, h).$$

Proof of the proposition. Let $p_n \in \mathscr{D}_n$. Let us choose $\{v_n^k\}_k \in \mathscr{W}(u + t_n h)$ and $p_n^k \in \partial_1 J(u + t_n h, v_n^k)$ such that $p_n^k \overset{*}{\rightharpoonup} \tilde{p}_n$. Let us also choose k_n such that, for a fixed sequence $\epsilon_n \searrow 0$, with the notations $v_n^{k_n} = v_n$ and $p_n^{k_n} = p_n$, the following holds

$$J(u + t_n h, v_n) \ge \bar{J}(u + t_n h) - \epsilon_n \qquad \text{and} \qquad |(p_n - \tilde{p}_n, h)| \le \epsilon_n.$$

The sequence $\{v_n\}$ is as in proposition 2, and in particular is in $\mathscr{W}(u)$. Moreover, for all $\alpha > 0$,

$$J(u + t_n h + \alpha h, v_n) \ge J(u + t_n h, v_n) + \alpha(p_n, h) \ge \bar{J}(u + t_n h) + \alpha(\tilde{p}_n, h) - 2\epsilon_n,$$

holds, i.e. for $\tilde{p} \in \mathscr{D}$, hence $\tilde{p} \in \partial\bar{J}(u)$ according to proposition 1,

$$J(u + t_n h + \alpha h, v_n) \ge \bar{J}(u) + t_n(\tilde{p}, h) - 2\epsilon_n + \alpha(\tilde{p}_n, h).$$

On the other hand, let us set

$$D_n = \sup_{\hat{p}_n \in \partial_1 J(u, v_n)} (\hat{p}_n, h).$$

For every positive η, there exists a positive α_0 such that, for every positive α smaller or equal to α_0, there holds

$$J(u + \alpha h, v_n) < J(u, v_n) + \alpha(D_n + \eta) \le \bar{J}(u) + \alpha(D_n + \eta).$$

Moreover, due to hypothesis D2.1, it may be picked α_0 independently of v_n. (i.e. fixed as $n \to \infty$.) Since J is continuous in u, uniformly in v, for n large enough, one has

$$J(u + t_n h + \alpha h, v_n) \le J(u + \alpha h, v_n) + \epsilon_n,$$

whence, regrouping the last three inequalities,

$$\exists \alpha_0 > 0 : \forall \alpha \in [0, \alpha_0], \qquad \alpha(D_n + \eta) \ge \alpha(\tilde{p}_n, h) - 3\epsilon_n.$$

Taking the limit, using proposition 3, to derive

$$D + \eta \ge \limsup(\tilde{p}_n, h)$$

which proves the proposition, since η was arbitrary.

PROPOSITION 5. $\partial\bar{J}(u) \subset \overline{co\mathscr{D}}$

Proof of the proposition. Let $\bar{p} \in \partial\bar{J}(u)$. Since $\partial\bar{J}$ is a monotonous operator,

$$\forall \bar{p}_n \in \partial\bar{J}(u + t_n h), \qquad (\bar{p}, h) \le (\bar{p}_n, h).$$

Therefore, making use of proposition 1

$$(\bar{p}, h) \le \inf_{\bar{p}_n \in \partial\bar{J}(u + t_n h)} (\bar{p}_n, h) \le \inf_{\tilde{p}_n \in \mathscr{D}_n} (\tilde{p}_n, h).$$

Finally, taking the lim sup and making use of proposition 4,

$$\forall \bar{p} \in \partial \bar{J}(u), \qquad (\bar{p}, h) \leq \sup_{\tilde{p} \in \mathscr{D}} (\tilde{p}, h).$$

Thus, $\bar{p} \in \overline{co}\mathscr{D}$, which proves the proposition.

Finally, propositions 1 and 5 together prove the theorem. ∎

It is useful, at this point, to give an alternate form of the formula of theorem C2. Define the *level sets at* u, V_ϵ, in the following way.

Definition. Let ϵ be a positive number, define

$$V_\epsilon(u) = \{v \in V \mid J(u,v) \geq \bar{J}(u) + \epsilon\}.$$

They are convex sets, increasing with ϵ. When it exists, $\hat{V}(u)$ is just $V_0(u)$. In terms of these sets, the formula of theorem C2 may be rewritten as follows

$$\partial \bar{J}(u) = \bigcap_{\epsilon > 0} \overline{co} \bigcup_{v \in V_\epsilon} \partial_1 J(u,v).$$

The above formulation is the natural one to state the result without the uniformity hypothesis D2.1. This theorem is proved in [2].

THEOREM C3 (Valdier). Under hypothesis C1 and C2, one has

$$\partial \bar{J}(u) = \bigcap_{\substack{\epsilon > 0 \\ \Omega}} \overline{co} \bigcup_{\substack{v \in V_\epsilon \\ u \in \Omega}} \partial_1 J(u,v),$$

where Ω ranges over a complete set of neighborhoods of u.

4. THE CONVEX CONCAVE CASE

The following additional hypothesis will be made.

Hypothesis CC. V is a convex subset of a Banach space $\mathscr{V}$, and $\forall u \in U$, $v \mapsto J(u,v)$ is concave.

Remark. In this case, if furthermore $\mathscr{V}$ is reflexive, in hypothesis C0 the compacity of V may be replaced by V closed and bounded. Owing to $v \mapsto J(u,v)$ being concave, its being u.s.c. is preserved in the weak topology.

The previous two theorems can be simplified in the following way.

THEOREM CC1. Under hypotheses C and CC, the subdifferential of $\bar{J}$ is given at any point u in U_0 by the formula

$$\partial \bar{J}(u) = \bigcup_{v \in \hat{V}(u)} \partial_1 J(u,v).$$

Proof. According to theorem C1, it suffices to prove the following proposition.

PROPOSITION. $\mathscr{D} = \bigcup_{v \in \hat{V}(u)} \partial_1 J(u,v)$ is convex.

Proof. For $i = 1,2$, let $v_i \in \hat{V}(u)$, and $p_i \in \partial_1 J(u, v_i)$. We know that $\hat{V}(u)$ is convex, and thus $\forall \lambda \in [0,1]$, $w = \lambda v_1 + (1-\lambda) v_2 \in \hat{V}(u)$. Let us also set $q = \lambda p_1 + (1-\lambda) p_2$. Let $h \in U - u$. Making use of hypothesis CC,

$$J(u+h, w) \geq \lambda J(u+h, v_1) + (1-\lambda) J(u+h, v_2) \geq \lambda J(u, v_1) + (1-\lambda) J(u, v_2) + (q, h).$$

In addition since, by definition, $J(u, v_i) = \bar{J}(u)$,

$$J(u+h, w) \geq \bar{J}(u) + (q,h) = J(u,w) + (q,h).$$

Thus $q \in \partial_1 J(u,w)$, where $w \in \hat{V}(u)$. ∎

THEOREM CC2. Under hypotheses C1, C2, and CC, the differential of $\bar{J}$ is given at any point u in U_0 by the formula

$$\partial \bar{J}(u) = \bigcup_{\{v_n\} \in \mathscr{W}(u)} \limsup_{n \to \infty} \partial_1 J(u, v_n).$$

Proof. Again, it suffices to prove that

$$\mathscr{D} = \bigcup_{\{v_n\} \in \mathscr{W}(u)} \limsup_{n \to \infty} \partial_1 J(u, v_n)$$

is convex.

For the sequel in this proof, the following notations will be used for $i = 1,2$, let $p^i \in \mathscr{D}$. There exist $\{v_k^i\}_k \in \mathscr{W}(u)$ and $p_k^i \in \partial_1 J(u, v_k^i)$ such that $p_k^i \overset{*}{\rightharpoonup} p^i$. For $\lambda \in [0,1]$, let

$$w_k = \lambda v_k^1 + (1-\lambda) v_k^2, \qquad q = \lambda p^1 + (1-\lambda) p^2, \qquad q_k = \lambda p_k^1 + (1-\lambda) p_k^2.$$

PROPOSITION 1. $\{w_k\}_k \in \mathscr{W}(u)$.

Proof of the proposition. By concavity, one has

$$J(u, w_k) \geq \lambda J(u, v_k^1) + (1-\lambda) J(u, v_k^2) \to \bar{J}(u).$$

In addition since, by definition of $\bar{J}$, $J(u, w_k) \leq \bar{J}(u)$, $J(u, w_k) \to \bar{J}(u)$, which is the definition of $\{w_k\}_k \in \mathscr{W}(u)$.

PROPOSITION 2. Let $\mathscr{C}_k$ be a sequence of convex subsets of $\mathscr{U}'$, and $\mathscr{C} = \limsup \mathscr{C}_k$. Let $D_k = \sup(p_k, h)$ and $D = \sup(p, h)$, for $p_k \in \mathscr{C}_k$ and $p \in \mathscr{C}$. Then:

(i) $\mathscr{C}_n$ is convex;

(ii) $\limsup D_k \leq D$.

Proof of the proposition. The first item is elementary. The second one is the proposition 3 of the proof of theorem C2.

PROPOSITION 3. With the notations introduced for this proof, let $\mathscr{C}_k = \partial_1 J(u, w_k)$, and $\mathscr{C} = \limsup \mathscr{C}_k$. Then $q \in \mathscr{C}$.

Proof of the proposition. As in the previous theorem, $\forall \alpha > 0$,

$$J(u+\alpha h, w_k) \geq \lambda J(u+\alpha h, v_k^1) + (1-\lambda) J(u+\alpha h, v_k^2),$$

i.e.

$$\forall \alpha, \qquad J(u+\alpha h, w_k) \geq \lambda J(u, v_k^1) + (1-\lambda) J(u, v_k^2) + \alpha (q_k, h).$$

Making use of the uniform continuity of J, it can be infered that

$$\forall \alpha, \qquad J(u+\alpha h, w_k) \geq \bar{J}(u) - \epsilon_k + \alpha (q_k, h),$$

where $\{\epsilon_k\}$ is a sequence decreasing to zero independently of α and h.

On the other hand, $\forall h, \forall \eta > 0$, $\exists \alpha_0 : \forall \alpha \in (0, \alpha_0)$,

$$J(u+\alpha h, w_k) \leq J(u, w_k) + \alpha (D_k + \eta) \leq \bar{J}(u) + \alpha (D_k + \eta).$$

Whence, comparing the last two inequalities, for $\alpha \leq \alpha_0$,

$$(q_k, h) \leq D_k + \eta + \frac{\epsilon_k}{\alpha}$$

and making use of proposition 2,

$$(q, h) = \lim (q_k, h) \leq D + \eta.$$

Since η was arbitrary, it can be concluded that $(q, h) \leq D$ and, since, according to proposition 2, $\mathscr{C}$ is convex, the proposition is proved.

Finally, since $\mathscr{C} \subset \mathscr{D}$, $q = \lambda p_1 + (1-\lambda) p_2 \in \mathscr{D}$, and the theorem is proved. ■

5. APPLICATION TO THE VON NEUMANN–SION THEOREM

It is shown here that classical theorems of the existence of a saddle point, or at least of a value, (inf sup = sup inf), to a convex–concave function, are simple consequences of the above theorems.

The first theorem below is often called "Von Neumann's theorem", although Von Neumann [16] himself only treated the case needed for matrix games, i.e. where U and V are simplices in Euclidean space, and J is linear. Sion [14] credits Shiffman for a more general form. The second theorem below is often called "Sion's theorem", although Sion credits Kneser and Fan for it. In [14], Sion gives a rather complete, and more general treatment of that question. An elegant theory can be found in [4].

Our hypotheses are similar to those of the previous section. We state them anew adapted to the present aim.

Hypotheses VN.

VN1. U is convex compact, contained in an open subset $\tilde{U} \subset \mathscr{U}$, and $\forall v \in V$, the function $u \mapsto J(u, v)$ is convex l.s.c. from $\tilde{U}$ into $\mathbb{R}$. Furthermore, J is bounded above, uniformly in v, in a neighborhood of any point of U in $\tilde{U}$.

VN2. V is convex, and $\forall u \in \tilde{U}$, the function $v \mapsto J(u, v)$ is concave.

VN3. V is (sequentially) compact, and $\forall u \in \tilde{U}$, the function $v \mapsto J(u, v)$ is u.s.c.

THEOREM VN1. Under hypotheses VN1 to VN3, the function J has a saddle point over $U \times V$, i.e. there exist $\hat{u} \in U$ and $\hat{v} \in V$ such that

$$\forall (u,v) \in U \times V, \qquad J(\hat{u},v) \le J(\hat{u},\hat{v}) \le J(u,\hat{v}).$$

Remark. The existence of a saddle point implies that

$$\min_{u \in U} \max_{v \in V} J(u,v) = \max_{v \in V} \min_{u \in U} J(u,v) = J(\hat{u},\hat{v}).$$

Proof. Theorem CC1 applies. In particular, hypothesis VN1 insures that $U \subset U_0$, where $\bar{J}$ is continuous, therefore, l.s.c. (even in the weak topology if necessary). It reaches its minimum at a point $\hat{u} \in U$. There exists thus $\hat{p} \in \partial \bar{J}(\hat{u})$ such that

$$\forall u \in U, \qquad (\hat{p}, u - \hat{u}) \ge 0.$$

Making use of theorem CC1, there exists $\hat{v} \in \hat{V}(\hat{u})$ such that $\hat{p} \in \partial_1 J(\hat{u},\hat{v})$. Whence

$$J(\hat{u},\hat{v}) \le J(u,\hat{v}) - (\hat{p}, u - \hat{u}).$$

Remembering that $\hat{v} \in \hat{V}(\hat{u})$, one can take the left hand inequality of the saddlepoint, and with the above two inequalities, the right hand one. ■

THEOREM VN2. Under hypotheses VN1 and VN2, there exists $\hat{u} \in U$ such that

$$\sup_{v \in V} J(\hat{u},v) = \min_{u \in U} \sup_{v \in V} J(u,v) = \sup_{v \in V} \min_{u \in U} J(u,v).$$

Proof. The proof makes use of Valadier's formula.

Let us first notice that again, hypothesis VN1 insures the existence of the minima in u. In particular, $\bar{J}$ has a minimum at a point $\hat{u} \in U$. Let

$$\hat{J} = \bar{J}(\hat{u}) = \min_{u \in U} \sup_{v \in V} J(u,v).$$

$\bar{J}$ being convex, there exists $\hat{p} \in \partial \bar{J}(\hat{u})$ such that

$$\forall u \in U, (\hat{p}, u - \hat{u}) \ge 0.$$

It will be exhibited a sequence $\{w_k\} \in V$ such that

$$\forall \epsilon > 0, \qquad \exists N, \qquad \min_{u \in U} J(u, w_N) \ge \hat{J} - \epsilon.$$

Then, one can conclude

$$\sup_{u \in V} \min_{u \in U} J(u,v) \ge \hat{J}.$$

Let ϵ_k be a decreasing sequence of positive numbers, and consider $\Omega_k \subset U$ and $V_k \subset V$ such that

$$\forall u \in \Omega_k, \qquad \forall v \in V, \qquad J(u,v) \ge J(\hat{u},v) - \epsilon_k,$$

$$V_k = V_{\epsilon_k}(\hat{u}) = \left\{ v \in V / J(\hat{u},v) \ge \hat{J} - \epsilon_k \right\}.$$

Such Ω_ks exist because $u \mapsto J(u,v)$ is l.s.c. uniformly in v, and any sequence $\{u_k\} \in \Omega_k$ converges to $\hat{u}$. Due to the definition of the level sets, any sequence $\{v_k\} \in V_k$ belongs to $\mathscr{W}(\hat{u})$. Let us also define the sequence of sets of subgradients

$$P_k = \bigcup_{\substack{u \in \Omega_k \\ v \in V_k}} \partial_1 J(u,v).$$

According to Valadier's formula, there exists a sequence of finite barycenters over P_k (see [15 T. 2, XIX, 2; 2] for finitude even in infinite dimension) such that

$$q_k = \sum_{i=0}^{n_k} \lambda_{k_i} p_{k_i} \overset{*}{\rightharpoonup} \hat{p}$$

and, of course,

$$\forall k \geq 0, \qquad \lambda_{k_i} \geq 0, \qquad \sum_{i=1}^{n_k} \lambda_{k_i} = 1.$$

Then, for each k, by definition of P_k, we can define two maps: $u_k: P_k \mapsto \Omega_k$ and $v_k: P_k \mapsto V_k$ such that

$$\forall p \in P_k, \qquad p \in \partial_1 J(u_k(p), v_k(p)).$$

For all $u \in U$, one has

$$J(u, v_k(p)) \geq J(u_k(p), v_k(p)) + (p, u - u_k(p)).$$

Since $u_k(p) \in \Omega_k$,

$$\forall u \in U, \qquad J(u, v_k(p)) \geq J(\hat{u}, v_k(p)) + (p, u - u_k(p)) - \epsilon_k.$$

Due to the concavity of $v \mapsto J(u,v)$ and the convexity of V, one can take the convex combination of all the inequalities in p_{k_i}, and obtain

$$\forall u \in U, \qquad J(u, w_k) \geq \sum_{i=0}^{n_k} \lambda_{k_i} \{ J(\hat{u}, v_k(p_{k_i})) + (p_{k_i}, u - u_k(p_{k_i})) \} - \epsilon_k,$$

where

$$w_k = \sum_{i=1}^{n_k} \lambda_{k_i} v_k(p_{k_i}) \in V.$$

It has been seen that $\{v_k(p), p \in P_k\} \in \mathscr{W}(\hat{u})$, so

$$\exists K_1, \qquad \forall k \geq K_1, \qquad \sum_{i=0}^{n_k} \lambda_{k_i} J(\hat{u}, v_k(p_{k_i})) \geq \hat{J} - \frac{\epsilon}{4}.$$

Due to $u_k(p) \to \hat{u}, p \in P_k$ and lemma 2 which provides the fact that the elements of P_k are bounded,

$$\exists K_2, \qquad \forall u \in U, \qquad \forall k \geq K_2, \qquad \forall p \in P_k, \qquad (p, u - u_k(p)) \geq (p, u - \hat{u}) - \frac{\epsilon}{4}.$$

For all u in U, there exists an open neighborhood $\mathscr{O}(u)$ and an integer n such that

$$\forall \tilde{u} \in \mathscr{O}(u), \qquad \forall k \geq n, \qquad (q_k, \tilde{u} - \hat{u}) \geq (q_k, u - \hat{u}) - \frac{\epsilon}{8} \geq (\hat{p}, u - \hat{u}) - 2\frac{\epsilon}{8} \geq -\frac{\epsilon}{4}.$$

As U is compact, a finite covering from the $\mathscr{O}(u)$ can be extracted. Let K_3 be the maximum of the corresponding ns. Then

$$\forall u \in U, \qquad \forall k \geq K_3, \qquad (q_k, u - \hat{u}) \geq -\frac{\epsilon}{4},$$

for u belongs to one of the $\mathscr{O}(u)$ selected in the finite covering.

Thus,

$$\forall u \in U, \qquad \forall k \geq \max(K_1, K_2, K_3), \qquad J(u, w_k) \geq \hat{J} - 3\frac{\epsilon}{4} - \epsilon_k.$$

Then, a $N \geq \max(K_1, K_2, K_3)$ can be chosen such that $\epsilon_N \leq \epsilon/4$, and the claim is proved. ∎

A trivial, but may be useful corollary is as follows.

COROLLARY. Under the hypotheses VN1 and VN2, if $\hat{U} \subset U$,

$$\sup_{v \in V} \inf_{u \in \hat{V}} J(u,v) = \inf_{u \in \hat{U}} \sup_{v \in V} J(u,v).$$

Proof. It suffices to make use of the continuity of J and $\tilde{J}$ and to apply the previous theorem to the closure $\bar{U}$ of $\hat{U}$. ∎

REFERENCES

1. DANSKIN J. M., *The Theory of Max Min*. Springer, Berlin, (1967).
2. VALADIER M., Contribution à l'analyse convexe, Thèse d'état, Paris (1970).
3. AUBIN J. P. & EKELAND I., *Applied Nonlinear Analysis*. Wiley-Interscience, New York (1984).
4. AUBIN J. P., *L'analyse non linéaire et ses motivations économiques*. Masson, Paris (1984).
5. FLEMING W. H. & RISHEL R. W., *Deterministic and Stochastic Optimal Control*. Springer, New York (1975).
6. ROCKAFELLAR R. T., Extensions of subgradients calculus with applications to optimization, *Nonlinear Analysis* **9**, 665–698 (1985).
7. CLARKE F. H., *Optimization and Nonsmooth Analysis*. Wiley-Interscience, New York (1983).
8. GAUVIN J., The generalized gradient of a marginal function in mathematical programming, *Math. Operat. Res.* **4**, 458–463 (1979).
9. HIRIART-URRUTY J. B., Gradients généralisés de fonctions marginales, *SIAM J. Control Optim.* **16**, 301–316 (1978).
10. BONNANS J. F., Directional derivatives of optimal solutions in smooth nonlinear programming, Rapport de recherche INRIA 1006, Rocquencourt, France (1989).
11. THIBAULT L., On subdifferentials of optimal value functions, *SIAM J. Control Optim.* **29**, 1019–1036 (1991).
12. BERHARD P., Variations sur un thème de Danskin avec une coda sur un thème de Von Neumann, Rapport de recherche INRIA 1238. Sophia-Antipolis, France (1990).
13. BAŞAR T. & BERHARD P., *H_∞-Optimal Control and Related Minimax Design Problems, a Dynamic Game Approach*. Birkhaüser, Boston (1991).
14. SION M., On General Minimax Theorems, *Pacif. J. Math.* **8**, 171–176 (1958).
15. SCHWARTZ L., *Analyse: topologie générale et analyse fonctionnelle*. Hermann, Paris (1970).
16. VON NEUMANN J. & MORGENSTERN O., *Theory of Games and Economic Behaviour*. Princeton University Press, Princeton (1947).