ISRAEL JOURNAL OF MATHEMATICS, Vol. 54, No. 1, 1986

PARTIALLY MONOTONE OPERATORS AND THE GENERIC DIFFERENTIABILITY OF CONVEX–CONCAVE AND BICONVEX MAPPINGS†

BY
J. M. BORWEIN
Department of Mathematics, Dalhousie University, Halifax, Nova Scotia, Canada B3H 4H8

ABSTRACT
By studying partially monotone operators, we are able to show among other results that convex–concave and biconvex mappings defined on Asplund spaces or dually strictly convex spaces are respectively generically Fréchet or Gateaux differentiable.

Introduction

Recently Jouak and Thibault [9], [10] have undertaken a study of the continuity and differentiability of convex–concave and biconvex operators taking values in appropriate partially ordered vector spaces. This study was based on Rockafellar's work on saddle functions [18] and the present author's work on convex operators [3], [4]. In [10] Jouak and Thibault showed that continuous convex–concave operators defined on separable Fréchet spaces are Gateaux differentiable almost everywhere (in the sense of Christensen [5]). For convex operators a very satisfactory theory exists [3], based on the Baire category theorems of Mazur [14], Asplund [1] and others as described in [3]. This suggests that analogous results might exist for saddle functions and operators. This is indeed so as we now proceed to show. The results, which are modeled on adaptations of Kenderov's beautiful theorems [11], [12] on generic single-valuedness of monotone operators, appear new even in the scalar case.

† Research partially supported on a N.S.E.R.C. operating grant.
Received August 29, 1983 and in revised form October 18, 1985

Partial monotonicity

Throughout this paper all spaces are real Banach spaces unless otherwise specified. Let U be a Banach space with norm dual V. A relation $T: U \to V$ will be called a *partially monotone* operator if U can be represented as a finite product $U := \Pi_{i \in I} U_i$ and if the restriction of T to each coordinate space is monotone (increasing or decreasing). Explicitly, for each x and y in U with x' in $T(x)$ and y' in $T(y)$ we have

$$(1.1) \qquad \langle y' - x', y - x \rangle \geqq 0 \qquad (\leqq 0)$$

whenever $y - x$ lies in U_i. Let us consider $x = (x_i, \hat{x}_i)$ in the standard way with x_i in U_i and $\hat{x}_i$ in $U_i^{\perp}$ (the orthogonal complement). Also, let T_i represent the projection of T on $V_i := U_i^*$. Then it is easy to see that (1.1) is equivalent to saying that, for each i in I, $T_i(\cdot, \hat{x}_i)$ is monotone between U_i and V_i. The central case in which I is singleton yields the classical monotone operators of Minty [15] and others. The other important example is produced by considering the partial subgradients of partially convex functions. Specifically, we call an extended real valued function on U *partially convex* (*up* or *down*) if U may be factored as above and if, for each i in I, $f(\cdot, \hat{x}_i)$ is always convex or always concave. When I has cardinality two we will call f biconvex (up or down). Thus our biconvex functions may be convex–concave or convex–convex. We associate a partially monotone operator with f as follows. Assuming f is convex in the ith coordinate, we define the *ith partial subgradient* of f at x by

$$(1.2) \qquad \partial_i f(x) := \{x'_i \in V_i : \langle x'_i, y_i - x_i \rangle \leqq f(y_i, \hat{x}_i) - f(x_i, \hat{x}_i) \ \forall y \in U\}.$$

If f is concave in the ith coordinate we similarly define a *partial supergradient* and use the same notation. We now set

$$(1.3) \qquad T(f)(x) := \prod_{i \in I} \partial_i f(x)$$

and observe that $T(f)$ is partially monotone. We might well think of $T(f)$ as a generalized gradient for f. We do not so denote it to avoid notational conflicts with other concepts such as Rockafellar's [18]. We call $T(f)$ the *partial subgradient* of f.

The *domain* of T, $D(T)$, is the set of points at which T has nonempty images. Also, T is said to be *USC* (*upper-semicontinuous*) with respect to the norm topology on U and a given topology on V if for each x in the interior of the domain of T and each neighbourhood N of $T(x)$ one can find a neighbourhood 0 of x with $T(0)$ contained in N. We will call an operator T_1 *weakly regular* if (locally) it lies inside a norm to weak-star USC partially monotone operator $T_2(T_1(x) \subset T_2(x))$. If, in addition, T is locally norm bounded we will call T *regular*. The justification for this is:

LEMMA 1.1. (a) *Every monotone operator is regular and partially monotone.*

(b) *The partial subgradient set of a locally Lipschitz partially convex function is regular and partially monotone.*

(c) *The partial subgradient set of a locally (separately) continuous biconvex function is regular and partially monotone.*

PROOF. (a) Every monotone operator T lies inside a maximal monotone operator M which is locally norm bounded throughout the interior of $D(M)$ [19]. Moreover, M has a (norm to weak-star) closed graph and has weak-star compact images. From these facts it follows that M is (norm to weak-star) USC and, hence, T is regular.

(b) Since f is continuous, $T(f)$ is norm to weak-star closed. Moreover, since f is locally Lipschitz, $T(f)$ is locally bounded. As in (a), $T(f)$ is USC.

(c) It is a consequence of the Baire category theorem that a separately continuous biconvex function is jointly continuous, and then is actually locally Lipschitz throughout the interior of its domain of finiteness. (The details can be found in [9].) The result now follows from (b). □

We will also have need for:

LEMMA 1.2. *Any weakly regular partially monotone operator lies inside a (norm to weak-star) USC partially monotone operator with weak-star closed convex images.*

PROOF. For T partially monotone and USC we define $C: U \to V$ by

$$C(x) := \text{cl-conv}\, T(x) \tag{1.4}$$

with closure taken in the weak-star topology. A little reflection on (1.1) shows that C is still partially monotone. Next, observe that C has weak-star compact images throughout $\operatorname{int} D(C)$. Indeed, each $T_i(\cdot, \hat{x}_i)$ is locally bounded; thus T has bounded images inside $\operatorname{int} D(T)$. A standard compactness argument shows that it suffices to verify that for each weak-star neighbourhood N of zero one can find a neighbourhood 0 of x with $C(0)$ contained in $C(x) + N$. Since N may be supposed convex and since $T(0)$ can be assumed to lie in $T(x) + N$ we are done. □

We can now establish our main results. Firstly:

THEOREM 1.3. *Suppose that U admits a strictly convex dual norm on V. Then every weakly regular partially monotone operator is single-valued at the points of a dense G_δ subset of the interior of its domain.*

PROOF. We may assume, by Lemma 1.2, that T is USC with compact convex images. We may also assume that T is actually partially isotone (monotone increasing) in each coordinate. For x in $\operatorname{int} D(T)$ we define m_i (for each i) by

$$m_i(x) := \min\{\| y_i \| : y \in T(x)\}. \tag{1.5}$$

Since the norm is weak-star lower-semicontinuous and T is USC, m_i is lower-semicontinuous; and the infimum is attained. Since $\operatorname{int} D(T)$ is a Baire space, m_i is actually continuous on a dense G_δ subset $G(i)$ of $\operatorname{int} D(T)$. We will show, as in [11], that $G(i)$ is comprised of points at which T_i is singleton. It follows that $G := \bigcap_{i\in I} G(i)$ is a suitable G_δ set. Since we may assume the norm to be strictly convex we must only show that $m_i(x) < \| y_i \|$ is impossible for y in $T(x)$ and x in $G(i)$. Select a unit vector x_i in U_i with $m_i(x) < \langle y_i, x_i \rangle$. By continuity of m_i we will have $m_i(x + tx_i) < \langle y_i, x_i \rangle$ for t small and positive. Choose $y(t)$ lying in $T(x + tx_i)$ with $\langle y_i(t), x_i \rangle < \langle y_i, x_i \rangle$. This contradicts the partial isotonicity of T. □

EXAMPLE. In the saddle function case we know there is a linear $1-1$ *surjective* mapping $L : X \to X$ such that L^*T is monotone; $L(x, y) := (x, -y)$ in fact [18]. In general this will not happen for partially convex functions (up or down). Indeed let $T(x, y) = (x^2y, yx^2)$ for $f(x, y) := \frac{1}{2}(xy)^2$. If L^*T is $(+)$ monotone we have

$$\left\langle (T(1,1) - T(1-x, 0)), L\begin{pmatrix} x \\ 1 \end{pmatrix} \right\rangle \geqq 0$$

for $x \in \mathbf{R}$. Then this says that if $L(x, y) := \langle ax + by, cx + dy \rangle$ we have

$$(a + c)x + (b + d) \geqq 0 \qquad \text{for all } x.$$

This is only possible if $a = -c$, symmetrically $b = -d$ and L is not $1-1$. Note that for single-valued results only injective L^* are any use if we wish to deduce that $T(x)$ is singleton when $L^*T(x)$ is. □

Our second result is the analogue result for norm to norm semicontinuity. Recall that a Banach space U is an *Asplund space* if every convex function on U is Fréchet differentiable on a dense G_δ subset of its points of continuity [16]. Such a set is said to be *generic*.

THEOREM 1.5. (a) *Every partially monotone operator T defined on a Banach space is single-valued and norm to norm USC exactly at the points of a G_δ subset, G, of* $\operatorname{int} D(T)$.

(b) *If the Banach space is an Asplund space and T is regular then G is dense in* $\operatorname{int} D(T)$.

PROOF. (a) For each i in I and n in $\mathbf{N}^x$, consider

(1.6) $G(i,n) := \{x \in \operatorname{int} D(T) : \operatorname{diam}(T_i(x + (1/m)B)) < 1/n \text{ for some } m \text{ in } \mathbf{N}^x\}$.

Then $G := \bigcap_{i \in I, n \in \mathbf{N}^x} G(i,n)$ exhibits the points at which T is singleton and norm to norm USC as a G_δ set. (Here B is the unit ball in U.)

(b) Since T is regular we may assume that T is actually USC norm to weak-star and locally bounded. Again, following Kenderov [12], we argue as follows. Fix x in $\operatorname{int} D(T)$. Let 0 be an open neighbourhood of x, small enough so that $T_i(0)$ is bounded. Since U is Asplund, $A := \operatorname{cl}\operatorname{conv} T_i(0)$ contains a weak-star strongly exposed point and one can find a point x_i in U_i which weak-star dents $T_i(0)$ [7], [16]. Let n in $\mathbf{N}$ be given and select $r > 0$ such that the nonempty set

$$S(r) := \{y_i \in T_i(0) : \sup\{\langle y_i, x_i\rangle \mid y_i \in T_i(0)\} - r < \langle y_i, x_i\rangle\} \tag{1.7}$$

has diameter less than $1/n$. Let z in 0 and w in $T(z)$ be chosen with w_i in $S(r)$. Set $x(t) := z + tx_i$, with t positive and small enough so that $x(t)$ is in 0. For each $w(t)$ in $T(x(t))$ we have, by partial monotonicity,

$$\langle w_i(t), x_i\rangle = \langle w_i, x_i\rangle + (1/t)\langle w(t) - w, x(t) - z\rangle \geqq \langle w_i, x_i\rangle \tag{1.8}$$

so that $T_i(x(t))$ lies in $S(r)$. Since T is USC norm to weak-star, there is a neighbourhood of $x(t)$, W, lying in 0 with $T_i(W)$ in $S(r)$. Thus diam $T_i(W)$ is less than $1/n$ and $x(t)$ lies in $G(i,n)$. Thus each $G(i,n)$ is dense and the proof is complete. □

REMARK 1.6. (a) An examination of the proof shows that we actually only need each coordinate space to be Asplund. This is equivalent [15]; our argument gives another proof of this.

(b) We have no example of a weakly regular partially monotone mapping which is not regular.

(c) It is possible to simplify the proof if one only wishes to deal directly with the partially convex case.

(d) Part (b) of Theorem 1.5 can also be deduced as a consequence of an elegant recent result of Christensen and Kenderov [6].

Applications to differentiability

An immediate application of the previous results is that the set of points at which all the partial derivatives of a locally Lipschitz partially convex function exist can be shown to be generic. We can do better in the case of biconvex functions and operators.

We now consider mappings taking values in a real locally convex space Y endowed with a partial order induced by a closed convex cone S. (See [3], [4], [9] for details.) We avoid infinity by considering mappings defined on an open convex subset C of U. Since all our results are local this is no restriction. A mapping $f: C \to Y$ is said to be *convex* (*with respect to* S) if for each x and y in C and $0 < t < 1$ we have

$$tf(x)+(1-t)f(y)-f(tx+(1-t)y)\in S. \tag{2.1}$$

We may then define *partially convex* and *biconvex operators with respect to* S, exactly as in Section 1. We presume throughout that the cone S is *normal*: there is a base at zero for the topology on Y such that $(S-W)\cap(W-S)\subset W$ for each W in the base; such neighbourhoods are said to be *full.*

PROPOSITION 2.1. *Let $f: C\to Y$ be* (*separately*) *continuous and biconvex* (*up or down*).

(a) *If f is partially Gateaux differentiable at x then f is Gateaux differentiable at x.*

(b) *Similarly, if Y is normed and f is partially Fréchet differentiable at x then f is Fréchet differentiable at x.*

PROOF. (a) This is a consequence of Propositions 3.1 and 3.2 of [10].

(b) We establish only the convex–concave case. The convex–convex case follows with an adjustment as in [10, Proposition 3.2]. Let $D_1f(x)$ and $D_2f(x)$ represent the two partial derivatives (singleton sub- and super-gradient sets). By an extension of Corollaries 2.9 and 3.6 of of [9] we can show that f is Lipschitz in a neighbourhood of x. Let W be an arbitrary full open neighbourhood of zero. Fix $\varepsilon>0$ so that for all unit vectors h in U and $0<t<\varepsilon$ we have

$$(1/t)[f(x+th_1)-f(x)]-D_1f(x)h_1\in W, \tag{2.2}$$

and

$$(1/t)[f(x+th_2)-f(x)]-D_2f(x)h_2\in W. \tag{2.3}$$

We also take ε small enough so that f is Lipschitz on a ball of radius ε around x.

Now select $0 < \delta < \varepsilon$ such that for $0 < s < \delta$ and all unit vectors h we have

$$(2.4) \qquad (1/\varepsilon)[f(x_1 + \varepsilon h_1, x_2 + sh_2) - f(x_1, x_2 + sh_2)] - D_1 f(x)h_1 \in W,$$

and

$$(2.5) \qquad (1/\varepsilon)[f(x_1 + sh_1, x_2 + \varepsilon h_2) - f(x_1 + sh_1, x_2)] - D_2 f(x)h_2 \in W.$$

Now (2.4) is possible as the functions

$$g(h, \cdot) := (1/\varepsilon)[f(x_1 + \varepsilon h_1, \cdot) - f(x_1, \cdot)]$$

are equi-Lipschitz at x_2; (2.5) is similar. Now by the convexity–concavity of f, we have

$$(2.6) \qquad (1/s)[f(x_1 + sh_1, x_2 + sh_2) - f(x_1, x_2 + sh_2)] - D_1 f(x)h_1 \in W - S$$

and

$$(2.7) \qquad (1/s)[f(x_1 + sh_1, x_2 + sh_2) - f(x_1 + sh_1, x_2)] - D_2 f(x)h_2 \in W + S$$

for $0 < s < \delta$. If we now add (2.2) with $t = s$ to (2.7) and similarly add (2.3) to (2.6) we deduce that for $0 < s < \delta$ and all unit vectors h

$$(1/s)[f(x + sh) - f(x)] - [D_1 f(x)h + D_2 f(x)h] \in (W - S) \cap (W + S)$$

which, since W is full, completes (b) and shows that

$$(2.8) \qquad Df(x)h = D_1 f(x)h_1 + D_2 f(x)h_2. \qquad \square$$

By applying this to scalar functions we derive:

COROLLARY 2.2. (a) *If U is an Asplund space then every biconvex (up or down) function is generically Fréchet differentiable (on its domain of continuity).*

(b) *If U admits a strictly convex dual norm every such function is generically Gateaux differentiable.*

PROOF. (a) Theorem 1.5 shows that $T(f)$ is generically singleton and norm to norm USC. It follows that f is partially Fréchet at these points [3], and Proposition 2.1 now applies.

(b) follows similarly from Theorem 1.3. $\square$

Thus the generic Fréchet differentiability of saddle-functions actually characterizes Asplund spaces. Asplund spaces include those with equivalent Fréchet norms and, hence, all reflexive spaces [15]. All weakly compactly generated spaces admit strictly convex dual norms [7], [13].

Finally we give the analogue results for operators. We first recall some terminology. A cone S is said to be *countably Daniell* [3], [1] if every decreasing sequence in S has an infimum which is also the topological limit. A continuous linear function ϕ on Y is *strictly positive* if $\phi(s) > 0$ for all nonzero s in S. When Y is metrizable, a cone with a bounded complete base is normal, countably Daniell and admits a strictly positive function [3]. We now combine results, much as in [9].

THEOREM 2.3. *Let $f : C \to Y$ be continuous and biconvex.*

(a) *If U is an Asplund space, Y is normed and S has a bounded complete base then f is generically Fréchet differentiable.*

(b) *If U admits a strictly convex dual norm and S is normal, countably Daniell and admits a strictly positive functional then f is generically Gateaux differentiable.*

PROOF. (a) Let ϕ be strictly positive. Then Corollary 2.2(a) shows ϕf to be generically Fréchet differentiable. *A fortiori*, $\phi f(\cdot, x_2)$ and $\phi f(x_1, \cdot)$ are Fréchet differentiable and are convex or concave. Since S has a bounded base, Theorem 5.2 of [3] shows f to be partially Fréchet differentiable. Now Proposition 2.1(b) completes the proof.

(b) This follows similarly, using Proposition 4.2 of [3] and Proposition 2.1(a).

□

REMARK 2.4. (a) If either U or Y is a separable normed space we may drop the hypothesis in Theorem 2.3(b) that strictly positive functionals exist, by arguing as in [4, Theorem 3.1]. Part (b) also holds if U is assumed to be an Asplund space.

(b) If we combine the previous result with [10, Proposition 4.6] we may deduce that when U is separable the generic set is actually *Haar-full* measure.

(c) As in [10] the conclusion of Theorem 2.3(b) can be strengthened to show that f is *Michel–Bastiani differentiable.*

(d) The limiting convex examples given in [3] show that the hypotheses on U and Y cannot be substantially weakened, unless more restrictions are imposed on f.

REFERENCES

1. E. Asplund, *Frechet-differentiability of convex functions*, Acta Math. **121** (1968), 31–47.

2. E. Asplund and R. T. Rockafellar, *Gradients of convex functions*, Trans. Amer. Math. Soc. **139** (1969), 443–467.

3. J. M. Borwein, *Continuity and differentiability properties of convex operators*, Proc. London Math. Soc. **44** (1982), 420–444.

4. J. M. Borwein, *Subgradients of convex operators*, Math. Operationforschung **15** (1984), 179–191.

5. J. P. R. Christensen, *Topological and Borel Structure*, North-Holland, New York, 1974.

6. J. P. Christensen and P. S. Kenderov, *Dense strong continuity of mappings and the Radon–Nykodym property*, preprint.

7. J. Diestel, *Geometry of Banach Spaces*, Lecture Notes in Mathematics **485**, Springer, New York, 1975.

8. S. P. Fitzpatrick, *Continuity of nonlinear monotone operators*, Proc. Amer. Math. Soc. **62** (1977), 111–116.

9. M. Jouak and L. Thibault, *Equicontinuity of families of convex and concave–convex operators*, Canad. Math. J. **34** (1984), 883–893.

10. M. Jouak and L. Thibault, *Directional derivatives and almost everywhere differentiability of biconvex and concave–convex operators*, preprint.

11. P. Kenderov, *Multivalued monotone mappings are almost everywhere single-valued*, Studia Math. **56** (1976), 199–203.

12. P. Kenderov, *Monotone operators in Asplund spaces*, Compte Rendu Bulg. Sci. **30** (1977), 936–964.

13. D. G. Larman and R. Phelps, *Gateaux differentiability of convex functions on Banach spaces*, J. London Math. Soc. **20** (1979), 115–127.

14. S. Mazur, *Über konvexe Mengen in linearen normierten Raumen*, Studia Math. **4** (1933), 70–84.

15. G. J. Minty, *Monotone (nonlinear) operators in Hilbert space*, Duke Math. J. **29** (1962), 341–346.

16. I. Namioka and R. Phelps, *Banach spaces which are Asplund spaces*, Duke Math J. **42** (1975), 735–750.

17. J.-P. Penot, *Calcul sous-differential et optimisation*, J. Funct. Anal. **27** (1978), 248–276.

18. R. T. Rockafellar, *Convex Analysis*, Princeton University Press, 1970.

19. R. T. Rockafellar, *Local boundedness of nonlinear, monotone operators*, Michigan Math. J. **16** (1969), 397–407.