VARIATIONAL METHODS IN CONVEX ANALYSIS

JONATHAN M. BORWEIN AND QIJI J. ZHU

Abstract. We use variational methods to provide a concise development of a number of basic results in convex and functional analysis. This illuminates the parallels between convex analysis and smooth subdifferential theory.

1. The purpose of this note is to give a concise and explicit account of the following folklore: several fundamental theorems in convex analysis such as the sandwich theorem and the Fenchel duality theorem may usefully be proven by variational arguments. Many important results in linear functional analysis can then be easily deduced as special cases. These are entirely parallel to the basic calculus of smooth subdifferential theory. Some of these relationships have already been discussed in [1, 2, 3, 4, 9, 14].

2. By a 'variational argument' we connote a proof with two main components: (a) an argument that an appropriate auxiliary function attains its minimum and (b) a 'decoupling' mechanism in a sense we make precise below.

It is well known that this methodology lies behind many basic results of smooth subdifferential theory [4, 16]. It is known, but not always made explicit, that this is equally so in convex analysis. Here we record in an organized fashion that this method also lies behind most of the important theorems in convex analysis.

In convex analysis the role of (a) is usually played by the following theorem attributed to Fenchel and Rockafellar (among others) for which some preliminaries are needed.

Let X be a real locally convex topological vector space. Recall that the *domain* of an extended valued convex function f on X (denoted $\mathrm{dom}\, f$) is the set of points with value less than $+\infty$. A subset T of X is *absorbing* if $X = \bigcup_{\lambda>0} \lambda T$ and a point s is in the *core* of a set $S \subset X$ (denoted by $s \in \mathrm{core}\, S$) provided that $S - s$ is absorbing. A symmetric, convex, closed and absorbing subset of X is called a *barrel*. We say X is *barrelled* if every barrel of X is a neighborhood of zero. All Baire—and hence all complete metrizable—locally convex spaces are barrelled, but not conversely.

Recall that $x^* \in X^*$ is a *subgradient* of $f : X \to (-\infty, +\infty]$ at $x \in \mathrm{dom}\, f$ provided that $f(y) - f(x) \geq \langle x^*, y - x\rangle$. The set of all subgradient of f at x is called the *subdifferential* of f at x and is denoted $\partial f(x)$. We use the standard convention that $\partial f(x) = \emptyset$ for $x \notin \mathrm{dom}\, f$.

Date: April 28, 2004.

Key words and phrases. Variational methods, convex analysis, sandwich theorem, Fenchel duality, functional analysis.

Research was supported by NSERC and by the Canada Research Chair Program.

Research was supported by National Science Foundation under grant DMS 0102496.

Theorem 1. (Fenchel-Rockafellar) *Let X be a barrelled locally convex topological vector space and let $f : X \to (-\infty, +\infty]$ be a convex function. Then for every x in* core dom$f$, $\partial f(x) \neq \emptyset$.

Combining this result with a decoupling argument we obtain the following lemma that can serve as a launching pad to develop many basic results in convex and in linear functional analysis.

Lemma 2. (Decoupling) *Let X and Y be locally convex topological vector spaces with Y barrelled. Suppose that both the functions $f : X \to (-\infty, +\infty]$ and $g : Y \to (-\infty, +\infty]$ are convex and that the map $A : X \to Y$ is linear. Let*

$$p = \inf\{f(x) + g(Ax)\}.$$

Suppose also that f, g and A satisfy the interiority condition

$$0 \in \operatorname{core}(\operatorname{dom} g - A \operatorname{dom} f). \tag{1}$$

Then, there is a $\phi \in Y^*$ such that, for any $x \in X$ and $y \in Y$,

$$p \leq [f(x) - \langle \phi, Ax\rangle] + [g(y) + \langle \phi, y\rangle]. \tag{2}$$

Proof. Define an *optimal value function* $h : Y \to [-\infty, +\infty]$ by

$$h(u) := \inf_{x \in X}\{f(x) + g(Ax + u)\}.$$

It is easy to check that h is convex and that $\operatorname{dom} h = \operatorname{dom} g - A \operatorname{dom} f$. Thus, the Fenchel-Rockafellar Theorem implies that $\partial h(0) \neq \emptyset$. Let $-\phi \in \partial h(0)$. Then, for all u in Y and x in X,

$$\begin{aligned} h(0) = p &\leq h(u) + \langle \phi, u\rangle \\ &\leq f(x) + g(Ax + u) + \langle \phi, u\rangle. \end{aligned} \tag{3}$$

For arbitrary $y \in Y$, setting $u = y - Ax$ in (3), we arrive at (2). **QED**

3. We now use Lemma 2 to recapture several basic theorems in convex analysis.

Theorem 3. (Sandwich) *Let X and Y be locally convex topological vector spaces with Y barrelled. Suppose that both the functions $f : X \to (-\infty, +\infty]$ and $g : Y \to (-\infty, +\infty]$ are convex and suppose that $A : X \to Y$ is a closed densely defined linear map (meaning that the adjoint A^* is well defined). Suppose also that*

$$f \geq -g \circ A$$

and f and g satisfy condition (1). Then, there is an affine function $\alpha : X \to R$ of the form

$$\alpha(x) = \langle A^*\phi, x\rangle + r,$$

for some ϕ in Y^*, satisfying

$$f \geq \alpha \geq -g \circ A.$$

Moreover, for any $\bar{x}$ satisfying $f(\bar{x}) = -g \circ A(\bar{x})$, one has $-\phi \in \partial g(A\bar{x})$.

Proof. By Lemma 2 there exists $\phi \in X^*$ such that, for any $x \in X$ and $y \in Y$,

$$0 \leq p \leq [f(x) - \langle \phi, Ax\rangle] + [g(y) + \langle \phi, y\rangle]. \tag{4}$$

For any $z \in X$ setting $y = Az$ in (4) we have

$$f(x) - \langle A^*\phi, x\rangle \geq -g(Az) - \langle A^*\phi, z\rangle.$$

Thus,

$$a := \inf_{x\in X}[f(x) - \langle A^*\phi, x\rangle] \geq b := \sup_{y\in Y}[-g(Ay) - \langle A^*\phi, y\rangle].$$

Picking any $r \in [a, b]$, and defining $\alpha(x) := \langle A^*\phi, x\rangle + r$ yields an affine function that separates f and $-g \circ A$.

Finally, when $f(\bar{x}) = -g \circ A(\bar{x})$, it follows from (4) that $-\phi \in \partial g(A\bar{x})$. **QED**

Theorem 4. (Fenchel duality) *Let X and Y be locally convex topological vector spaces with Y barrelled. Let $f : X \to (-\infty, +\infty]$ and $g : Y \to (-\infty, +\infty]$ be convex functions and let $A : X \to Y$ be a closed densely defined linear map. Suppose that f and g satisfy the condition*

$$0 \in \operatorname{core}(\operatorname{dom} g - A \operatorname{dom} f).$$

Define

$$p = \inf_{x\in X}\{f(x) + g(Ax)\} \tag{5}$$

$$d = \sup_{\phi\in Y^*}\{-f^*(A^*\phi) - g^*(-\phi)\}. \tag{6}$$

Then $p = d$, and the supremum in the dual problem (6) is attained whenever finite.

Proof. It follows from Fenchel's inequality

$$h(z) + h^*(\phi) \geq \langle \phi, z\rangle$$

for any function h, that $p \geq d$ always holds. This fact is usually referred to as *weak duality.*

If p is $-\infty$ there is nothing to prove, while if condition (1) holds and p is finite then by Lemma 2 there is a $\phi \in Y^*$ such that (2) holds. For any $u \in Y$, setting $y = Ax + u$ in (2) we have

$$\begin{aligned} p &\leq f(x) + g(Ax + u) + \langle \phi, u\rangle \\ &= \{f(x) - \langle A^*\phi, x\rangle\} + \{g(Ax + u) - \langle -\phi, Ax + u\rangle\}. \end{aligned}$$

Taking the infimum over all points u, and then over all points x, gives the inequalities

$$p \leq -f^*(A^*\phi) - g^*(-\phi) \leq d \leq p.$$

Thus ϕ attains the supremum in problem (6), and $p = d$. **QED**

Remark 5. *Let $\operatorname{cont} g$ denote the set of all continuity points of g. Note that the condition*

$$\operatorname{cont} g \cap A \operatorname{dom} f \neq \emptyset, \tag{7}$$

implies (1) and is convenient in many applications.

Several important convex subdifferential calculus rules follow immediately.

Theorem 6. (Convex subdifferential sum and composition rule) *Let X and Y be locally convex topological vector spaces with Y barrelled, let both $f : X \to (-\infty, +\infty]$ and $g : Y \to (-\infty, +\infty]$ be convex functions and let $A : X \to Y$ be a closed densely defined linear map. Then at any point x in X, the sum rule*

$$\partial(f + g \circ A)(x) \supset \partial f(x) + A^*\partial g(Ax) \tag{8}$$

holds, with equality if either condition (1) or (7) holds.

Proof. Inclusion (8) is easy. We prove the reverse inclusion under condition (1). Suppose $x^* \in \partial(f + g \circ A)(\bar{x})$. Since shifting by a constant does not change the subdifferential of a convex function, we may assume without loss of generality that

$$x \mapsto f(x) + g(Ax) - \langle x^*, x\rangle$$

attains its minimum 0 at $x = \bar{x}$. By the sandwich theorem of Theorem 3 there exists an affine function $\alpha(x) := \langle A^*\phi, x\rangle + r$ with $-\phi \in \partial g(A\bar{x})$ such that

$$f(x) - \langle x^*, x\rangle \geq \alpha(x) \geq -g(Ax).$$

Since equality is attained at $x = \bar{x}$, we have $x^* + A^*\phi \in \partial f(\bar{x})$. Therefore,

$$x^* = (x^* + A^*\phi) + A^*(-\phi) \in \partial f(\bar{x}) + A^*\partial g(A\bar{x}).$$

QED

Recall that the *convex normal cone* to C at x is defined to be

$$N_C(\bar{x}) := \{\phi \in X^* : \langle \phi, c - \bar{x}\rangle \leq 0, \forall c \in C\}.$$

With this notation, suppose $g := i_C$ where C is a closed convex subset of X and i_C denotes the *convex indicator function* of C, which is zero on C and $+\infty$ otherwise, and A is the identity mapping on X. Then we derive:

Theorem 7. (Pshenichnii–Rockafellar conditions [10]) *If the convex set C in a barrelled locally convex topological vector space X satisfies the condition that (i)* $\operatorname{cont} f \cap C \neq \emptyset$, *or the condition that (ii)* $\operatorname{dom} f \cap \operatorname{int} C \neq \emptyset$, *and if f is bounded below on C, then there is an affine function $\alpha \leq f$ with*

$$\inf_C f = \inf_C \alpha.$$

In addition, the point $\bar{x}$ minimizes f on C if and only if it satisfies

$$0 \in \partial f(\bar{x}) + N_C(\bar{x}).$$

Combining Theorems 6, 7 and Ekeland's variational principle [7]—applicable in the complete metrizable setting—we may next derive a convex version of the *multidirectional mean value theorem* [6, 9].

Theorem 8. (Convex multidirectional mean value inequality) *Let X be an arbitrary Banach space, let C be a nonempty, closed and convex subset of X. Fix x in X and let $f : X \to \mathbb{R}$ be a continuous convex function. Suppose that f is bounded below on $[x, C]$ and*

$$\inf_{y \in C} f(y) - f(x) > r.$$

Then, for any $\varepsilon > 0$, there exist $z \in [x, C]$ and $z^* \in \partial f(z)$such that

$$f(z) < \inf_{[x,C]} f + |r| + \varepsilon,$$

and

$$r < \langle z^*, y - x\rangle + \varepsilon\|y - x\|$$

for all y in C.

Proof. Using the auxiliary function $F(x,t) := f(x) - rt$ we can convert the general case to the special case when $r = 0$. So we will only prove this special case. Let $\tilde{f} := f + i_{[x,C]}$. Then $\tilde{f}$ is bounded below on X. By taking a smaller $\varepsilon > 0$ if necessary, we may assume that

$$\varepsilon < \inf_{y \in C} f(y) - f(x).$$

Applying Ekeland's variational principle [7] we conclude that there exists z such that

$$\tilde{f}(z) < \inf \tilde{f} + \varepsilon \tag{9}$$

and

$$\tilde{f}(z) \leq \tilde{f}(u) + \varepsilon\|u - z\|, \ \forall u \in X. \tag{10}$$

That is to say

$$u \to f(u) + i_{[x,C]}(u) + \varepsilon\|u - z\|$$

attains a minimum at z. By (9) $\tilde{f}(z) < +\infty$ hence $z \in [x, C]$.

The sum rule for convex subdifferentials given in Theorem 6 (with A being the identity mapping) implies that there exists $z^* \in \partial f(z)$ such that $0 \leq \langle z^*, w - z\rangle + \varepsilon\|w - z\|$, $\forall w \in [x, C]$. Using a smaller ε to begin with if necessary we have, for $w \neq z$,

$$0 < \langle z^*, w - z\rangle + \varepsilon\|w - z\|, \quad \forall w \in [x, C]\backslash\{z\}. \tag{11}$$

Moreover by inequality (9) we have $f(z) = \tilde{f}(z) \leq f(x) + \varepsilon < \inf_C f$, so $z \notin C$. Thus we can write $z = x + \bar{t}(\bar{y} - x)$ where $\bar{t} \in [0, 1)$. For any $y \in C$ set $w = y + \bar{t}(\bar{y} - y) \neq z$ in (11) yields

$$0 < \langle z^*, y - x\rangle + \varepsilon\|y - x\|, \quad \forall y \in C. \tag{12}$$

QED

Note that in the proof of this result besides using the subdifferential sum rule (which we have seen is a consequence of the decoupling lemma) we centrally used Ekeland's variational principle to locate the mean value point z.

The multidirectional mean value inequality can be used to prove a quite general *open mapping* theorem [9]. Recall that a multifunction $F : X \to 2^Y$ is a *closed convex multifunction* if the *graph* of F ($\{(x, y) : y \in F(x)\}$) is a closed convex set.

Theorem 9. (Open mapping) *Let X and Y be Banach spaces. Let $F : X \to 2^Y$ be a closed convex multifunction. Suppose that*

$$y_0 \in \text{core } F(X).$$

Then F is open at y_0; that is, for any $x_0 \in F^{-1}(y_0)$ and any $\eta > 0$,

$$y_0 \in \text{int } F(x_0 + \eta B_X).$$

Proof. Let $T : X \times Y \to Y$ be a linear operator defined by $T(x, y) := y$ and let $G :=$ Graph F. It is plain that we need only to show that $T|_A$ is open at (x_0, y_0). Since

$$0 \in \text{core } T(G - (x_0, y_0)) = F(X) - y_0$$

and G is convex, a standard Baire category argument implies that there exists $\varepsilon > 0$ such that

$$\varepsilon B_Y \subset \text{cl } T((G - (x_0, y_0)) \cap B_{X\times Y}). \tag{13}$$

We need to remove the closure above and so to show that

$$T(x_0, y_0) + (\varepsilon\eta/2)B_Y \subset T(((x_0, y_0) + \eta B_{X\times Y}) \cap G).$$

Let $z \in T(x_0, y_0) + (\varepsilon\eta/2)B_Y$ and set $h(x, y) := \|T(x, y) - z\|$.

Applying the convex multidirectional mean value inequality of Theorem 8 to function h, set $Y := ((x_0, y_0) + \eta B_{X\times Y}) \cap G$ and point (x_0, y_0) yields that there exist $u \in ((x_0, y_0) + \eta B_{X\times Y}) \cap A$ and $u^* \in \partial h(u)$ such that

$$\inf_Y h - h(x_0, y_0) - \varepsilon\eta/4 \le \langle u^*, (x, y) - (x_0, y_0)\rangle, \ \forall x \in Y. \tag{14}$$

If $h(u) = 0$ then $T(u) = z$ and we are done. Otherwise $u^* = T^*y^*$ with $y^* \in \partial\| \cdot \|(T(u) - z)$ being a unit vector. Then we can rewrite (14) as

$$\begin{aligned} 0 \ &\le \ \inf_Y h \le h(x_0, y_0) + \varepsilon\eta/4 + \langle y^*, T((x, y) - (x_0, y_0))\rangle \\ &\le \ \varepsilon\eta/2 + \varepsilon\eta/4 + \langle y^*, T((x, y) - (x_0, y_0))\rangle, \ \forall (x, y) \in ((x_0, y_0) + \eta B_{X\times Y}) \cap G. \end{aligned}$$

Observe that $\eta\varepsilon B_Y \subset \mathrm{cl}\, T((G - (x_0, y_0)) \cap \eta B_{X\times Y})$ the infimum of the right hand side of the above inequality is $-\varepsilon\eta/4$, a contradiction. **QED**

As an easy corollary we have the following boundedness result for convex functions, which holds somewhat more generally in Baire or barrelled normed spaces.

Theorem 10. (Boundedness of convex functions) *Let X be a Banach space and let $f : X \to \bar{R}$ be a lower semicontinuous convex function. Then f is continuous at every point in the core (equivalently interior) of its domain.*

In particular, f is everywhere continuous if and only if f is everywhere finite.

Proof. We need only prove the first assertion. Consider

$$F(x) := f(x) + [0, +\infty).$$

Then F and F^{-1} are closed convex multifunctions because graph $F :=$ epi f is a closed convex set. Let $x \in$ core (dom f) $=$ core $F^{-1}(R)$. By the Open Mapping Theorem 9, F^{-1} is open at x. Now, consider any open interval (a, b) that contains $f(x)$. The lower semicontinuity of f implies that $\{x : f(x) \le a\}$ is closed. Thus, x is in the open set

$$f^{-1}((a, b)) = F^{-1}((a, b))\backslash\{x : f(x) \le a\}.$$

Therefore, f is continuous at x. **QED**

4. Much of linear functional analysis can be viewed as a special case of convex analysis. Below we recall how to derive the basic results of linear functional analysis from the results of the previous section.

Theorem 11. (Hahn–Banach extension) *Let X be a barrelled locally convex topological vector space. Suppose the function $f : X \to \mathbb{R}$ is everywhere finite and sublinear, and suppose for some linear subspace L of X the function $h : L \to \mathbb{R}$ is linear and* dominated *by f, that is, $f \ge h$ on L.*

Then there is a linear function $\alpha : X \to \mathbb{R}$, dominated by f, which agrees with h on L.

Proof. Let $X = Y$, let A be the identity mapping of X, let $g = -h + i_L$ and apply the sandwich result of Theorem 3. **QED**

Theorem 12. (Hahn–Banach separation) *Let X be a barrelled locally convex topological vector space and let C_1 and C_2 be two convex subsets of X. Suppose that $\operatorname{int} C_1 \neq \emptyset$ but that $C_2 \cap \operatorname{int} C_1 = \emptyset$.*

Then there exists an affine function α on X such that

$$\sup_{c_1 \in C_1} \alpha(c_1) \leq \inf_{c_2 \in C_2} \alpha(c_2).$$

Proof. Consider the *gauge* function of C_1 defined by

$$\gamma(x) := \inf\{r : x \in rC_1\}.$$

Then γ is convex and $\operatorname{dom}\gamma = X$. Moreover, $\operatorname{int} C_1 = \{x \in X : \gamma(x) < 1\}$ and, consequently $C_1 \subset \{x \in X : \gamma(x) \leq 1\}$. Applying the sandwich Theorem 3 with $f = i_{C_2}$, A is the identity mapping of X and $g = \gamma - 1$ we have there exists an affine function α on X such that $f \geq \alpha \geq -g$. Now for any $c_1 \in C_1$, $\alpha(c_1) \geq 1 - \gamma(c_1) \geq 0$ and for any $c_2 \in C_2$, $\alpha(c_2) \leq i_{C_2}(c_2) = 0$. **QED**

The following classical open mapping theorem for linear mappings is a direct corollary of Theorem 9 in which $F(x) = \{Ax\}$ and, as usual, a linear mapping A from X to Y is said to be *open* when it maps open sets in X to open sets in Y.

Theorem 13. (Open mapping theorem for linear mappings) *Let X and Y be Banach spaces and let A be a closed linear mapping from X to Y such that $A(X) = Y$.*

Then A is an open mapping.

Next, we recall how directly to deduce the *principle of uniform boundedness* of linear functional analysis from Theorem 10.

Theorem 14. (Principle of uniform boundedness) *Let X and Y be Banach spaces. Let Γ be a set of bounded linear operators from X to Y such that for each $x \in X$,*

$$\sup\{\|Ax\| : A \in \Gamma\} < +\infty.$$

Then

$$\sup\{\|A\| : A \in \Gamma\} < +\infty.$$

Proof. Define

$$f(x) := \sup\{\|Ax\| : A \in \Gamma\}.$$

Then it is easy to verify that f is a lower semicontinuous convex function, as a supremum of convex continuous functions. Since, by assumption, $f(x) < +\infty$ for all $x \in X$, it follows by Theorem 10 that f is continuous. In particular, there exists a constant $\eta > 0$ such that $\sup\{f(x) : x \in \eta B_X\} < \infty$. Then

$$\begin{aligned} & \sup\{\|A\| : A \in \Gamma\} = \sup\{\|Ax\| : A \in \Gamma, x \in B_X\} \\ = \quad & \frac{1}{\eta}\sup\{\|Ax\| : A \in \Gamma, x \in \eta B_X\} = \frac{1}{\eta}\sup\{f(x) : x \in \eta B_X\} < +\infty. \end{aligned}$$

QED

5. Finally we use the Fitzpatric function to give variational proofs of Rockafellar's results on the range of maximal monotone multifunctions and Kirszbraun-Valentine extension theorem on nonexpansive mappings. Throughout this section, $(X, \|\cdot\|)$ is a reflexive Banach space with dual X^* and $T \colon X \to 2^{X^*}$ is maximal monotone. The

Fitzpatrick function F_T [8], associated with T, is the proper closed convex function defined on $X \times X^*$ by

$$\begin{aligned} F_T(x,x^*) &:= \sup_{y^* \in Ty} [\langle y^*, x\rangle + \langle x^*, y\rangle - \langle y^*, y\rangle] \\ &= \langle x^*, x\rangle + \sup_{y^* \in Ty} \langle x^* - y^*, y - x\rangle. \end{aligned}$$

Since T is maximal monotone

$$\sup_{y^* \in Ty} \langle x^* - y^*, y - x\rangle \geq 0$$

and the equality holds if and only if $x^* \in Tx$. It follows that

$$F_T(x,x^*) \geq \langle x^*, x\rangle \tag{15}$$

with equality holds if and only if $x^* \in Tx$. Thus, we capture much of a maximal monotone function via a convex associate.

Using only the Fitzpatrick function and the sandwich theorem we can prove the following fundamental result remarkably easily [15].

Theorem 15. (Rockafellar) *Let* X *be a reflexive Banach space and let* $T\colon X \to 2^{X^*}$ *be a maximal monotone operator. Then* $R(T+J) = X^*$. *Here* J *is the* duality map *defined by* $J(x) := \partial\|x\|^2/2$.

Proof. The Cauchy inequality and (15) implies that, for all x, x^*,

$$F_T(x,x^*) + \frac{\|x\|^2 + \|x^*\|^2}{2} \geq 0. \tag{16}$$

Applying the sandwich Theorem 3 to (16) we conclude that there exist points $(w^*, w) \in X^* \times X$ and $\alpha \in R$ such that

$$\begin{aligned} F_T(x,x^*) &\geq \langle w^*, x\rangle + \langle x^*, w\rangle + \alpha \qquad (17)\\ &\geq -\frac{\|x\|^2 + \|x^*\|^2}{2}. \qquad (18) \end{aligned}$$

Choose $x \in -Jw^*$ and $x^* \in -Jw$ in inequality (18) we have

$$\alpha \geq \frac{\|w\|^2 + \|w^*\|^2}{2}. \tag{19}$$

Next, for any $x^* \in Tx$ in (17) we have

$$\langle x^*, x\rangle \geq \langle w^*, x\rangle + \langle x^*, w\rangle + \frac{\|w\|^2 + \|w^*\|^2}{2}.$$

Adding $\langle w^*, w\rangle - \langle w^*, x\rangle - \langle x^*, w\rangle$ to both sides of the above inequality we obtain

$$\begin{aligned} \langle x^* - w^*, x - w\rangle &= \langle x^*, x\rangle - \langle w^*, x\rangle - \langle x^*, w\rangle + \langle w^*, w\rangle \\ &\geq \frac{\|w\|^2 + \|w^*\|^2}{2} + \langle w^*, w\rangle \geq 0. \qquad (20) \end{aligned}$$

Since (20) holds for all $x^* \in Tx$ and T is maximal we must have $w^* \in Tw$. Now setting $x^* = w^*$ and $x = w$ in (20) yields

$$\frac{\|w\|^2 + \|w^*\|^2}{2} + \langle w^*, w\rangle = 0,$$

which implies $-w^* \in Jw$. Thus, $0 \in (T+J)w$. Since the argument applies equally well to all translations of T, we have $R(T+J) = X^*$ as required. **QED**

The original proofs [13] were very extended and quite sophisticated—they used tools such as Brouwer's fixed point theorem and Banach space renorming theory. As with our proof of local boundedness, ultimately the result is reduced to much more accessible geometric convex analysis. The short proof here is a rework of that of [15]. It well illustrates the techniques of variational analysis: a properly constructed auxiliary function— the Fitzpatrick function—the variational principle with decoupling in the form of a sandwich theorem, followed by an appropriate decoding of the information.

There is a tight relationship between nonexpansive mappings and monotone operators in Hilbert spaces as stated in the next lemma.

Lemma 16. *Let H be a Hilbert space. Suppose that P and T are two multifunctions from subsets of H to 2^H whose graphs are related by $(x, y) \in$ graph P if and only if $(v, w) \in$ graph T where $x = w + v$ and $y = w - v$. Then*

(i) *P is nonexpansive (and single-valued) if and only if T is monotone.*
(ii) *$D(P) = R(T + I)$.*

Proof. Consider $v_n \in Tw_n, n = 1, 2$. Then $y_n \in Px_n$ where $x_n = w_n + v_n$ and $y_n = w_n - v_n$. Direct computation yields

$$\begin{aligned}\langle v_1 - v_2, w_1 - w_2\rangle &= \langle\frac{x_1 - x_2 - (y_1 - y_2)}{2}, \frac{x_1 - x_2 + (y_1 - y_2)}{2}\rangle \\ &= \frac{\|x_1 - x_2\|^2 - \|y_1 - y_2\|^2}{4}.\end{aligned}$$

It is easy to see that P is nonexpansive if and only if T is monotone.

To prove (ii), note that if $x \in D(P)$ and $y = Px$ then $\frac{x+y}{2} \in T(\frac{x-y}{2})$ by definition so that $x = \frac{x+y}{2} + \frac{x-y}{2} \in R(T + I)$. Conversely if $w \in R(T + I)$ then there exists v such that $w \in (T + I)v$ or $w - v \in Tv$ which implies that $w = (w - v) + v \in D(P)$. **QED**

This very easily leads to the Kirszbraun-Valentine theorem on the existence of nonexpansive extensions to all of Hilbert space of nonexpansive mappings on subsets of Hilbert space.

Theorem 17. (Kirszbraun-Valentine) *Let H be a Hilbert space and let D be a nonempty subset of H. Suppose that $P : D \to H$ is a nonexpansive mapping. Then there exists a nonexpansive mapping $\widehat{P} : H \to H$ defined on all of H such that $\widehat{P}|_D = P$.*

Proof. Associate P to a monotone multifunction T as in Lemma 16. Extend T to a maximal monotone multifunction $\widehat{T}$. Define $\widehat{P}$ from $\widehat{T}$ using Lemma 16 again. Then Rockafellar's Theorem 15 to we have $D(\widehat{P}) = R(\widehat{P} + I) = H$. It is easy to check that $\widehat{P}$ is indeed an extension of P. **QED**

Alternatively [11], one may directly associate a convex Fitzpatrick function F_P with a non-expansive mapping P, and thereby derive the Kirszbraun-Valentine theorem.

6. We have seen that in the proofs of all the results discussed here the decoupling Lemma is either explicitly or implicitly involved. Thus, a variational argument is indeed a common thread behind many of the fundamental results in convex and functional analysis. Such matters are also discussed in [5] where additional examples are to be found. In particular, similar proofs are given the local boundedness of maximal monotone operators throughout the core of their domains, and of the surjectivity of coercive maximal monotone operators in reflexive space.

This is by no means a claim of the intrinsic superiority of the treatment herein, which in part follows [3]. For example, recently S. Simons showed an elegant and different way of explaining a similar basket of results starting from a generalized Hahn-Banach extension theorem [14]. Indeed [14] was in part what stimulated us to record the present perspective.

References

[1] Borwein J. M., Convex relations in analysis and optimization, in S. Schaible and W. T. Ziemba eds. *Generalized Concavity in Optimization and Economics*, Academic Press, New York, 1981.

[2] Borwein J. M., Stability and regular points of inequality systems, *J. Optimization Theory and Applications*, **48** (1986), 9–52.

[3] Borwein J. M. & Lewis A. S., *Convex Analysis and Nonlinear Optimization: Theory and Examples*, CMS Springer-Verlag Books, Springer-Verlag, New York, 2000.

[4] Borwein J. M. & Zhu Q. J., A survey of subdifferential calculus with applications, *Nonlinear Analysis, TMA,* **38** (1999), 687–773.

[5] Borwein J. M. & Zhu Q. J., *Techniques of Variational Analysis*, CMS Springer-Verlag Books, Springer-Verlag, New York, 2004.

[6] Clarke F. H. & Ledyaev Yu. S., Mean value inequalities, *Proc. Amer. Math. Soc.*, **122** (1994), 1075–1083.

[7] Ekeland I., On the variational principle, *J. Math. Anal. Appl.*, **47** (1974), 324–353.

[8] Fitzpatrick, S., Representing monotone operators by convex functions. *Workshop/Miniconference on Functional Analysis and Optimization* (Canberra, 1988), 59–65, Proc. Centre Math. Anal. Austral. Nat. Univ., 20, Austral. Nat. Univ., Canberra, 1988.

[9] Ledyaev Yu. S. & Zhu Q. J. Implicit multifunction theorems, *Set-Valued Analysis*, **7** (1999), 209–238.

[10] Pshenichnii B., *Necessary Conditions for an Extremum*, Marcel Dekker, New York, 1971.

[11] Reich S. and Simons S., Fenchel duality, Fitzpatrick functions and the Kirszbraun-Valentine extension theorem, preprint, 2004.

[12] Rockafellar R. T., *Convex Analysis*, Princeton University Press, Princeton, NJ, 1970.

[13] Rockafellar R. T., On the maximality of sums of nonlinear monotone operators. *Trans. Amer. Math. Soc.* **149** (1970), 75–88.

[14] Simons S. A new version of the Hahn-Banach theorem, preprint.

[15] Simons S. and Zalinescu C., A New Proof for Rockafellar's Characterization of Maximal Monotonicity. *Proc Amer. Math. Soc.*, in press.

[16] Zhu Q. J. The equivalence of several basic theorems for subdifferentials, *Set-Valued Analysis*, **6** (1998), 171-185.

Faculty of Computer Science, Dalhousie University, Halifax, N. S. B3H 1W5, Canada, e-mail: jborwein@cs.dal.ca,

Department of Mathematics, Western Michigan University, Kalamazoo, MI 49008, USA, e-mail: zhu@wmich.edu,