DUALITY INEQUALITIES AND SANDWICHED FUNCTIONS
December 1998
J. M. Borwein and S. P. Fitzpatrick

Abstract We establish a mixed convex–Lipschitz mean value inequality from which recent results of Clarke and Ledyaev and of Lewis and Ralph follow naturally. We also provide various refinements and extensions. Finally, we answer affirmatively several open questions on the existence of “squeeze” theorems for a finite number of Lipschitz functions.


Key words Nonconvex separation, sandwich theorem, mean–value inequality, Fenchel duality, Schauder fixed point theorem, Ekeland variational principle.

1. Introduction and Preliminaries
Motivated by Clarke and Ledyaev’s [2] striking multi-directional mean-value theorem and its elegant reformulation by Lewis and Ralph [8] as a non-smooth sandwich theorem, our intentions in this paper are two-fold.

Firstly, in section two we provide a self-contained proof of a general convex/Lipschitz inequality from which both results follow. In section three we exploit the underlying technique to obtain some more refined – and perhaps surprising – inequalities. These results rely on the Brouwer/Schauder fixed point theorem.

Secondly, in section four, we use variational methods to provide an affirmative answer to several open questions posed by Lewis and Lucchetti [7] on the existence of common subgradients for a finite family of Lipschitz functions. Section five provides various infinite dimensional and non-compact extensions of our results and poses several open questions.

The rest of this section is dedicated to three preparatory results.

If $E$ is a Banach space, possibly $\mathbb{R}^n$, we say $f : E \to \mathbb{R}_-$ is proper provided it is not identically $\infty$ and write $\text{dom}(f)$ for the set of points where $f$ is finite. We say $f$ is lower semicontinuous if its epigraph

$$\text{epi}(f) := \{(x, r) \in E \times \mathbb{R} : r \geq f(x)\}$$

is closed and convex if $\text{epi}(f)$ is convex. We say $h : E \to \mathbb{R} \cup \{-\infty\}$ is upper semicontinuous if $-h$ is lower semicontinuous, that is, the hypograph

$$\text{hyp}(h) := \{(x, r) \in E \times \mathbb{R}_- : r \leq h(x)\}$$

is closed, and concave if $\text{hyp}(h)$ is convex.

The indicator function of a closed nonempty convex subset $A$ of $E$, \[ I_A(x) := \begin{cases} 0 & \text{if } x \in A, \\ \infty & \text{otherwise} \end{cases} \]

is a proper convex lower semicontinuous function.

A real-valued function $g$ on a subset $X$ of $E$ is Lipschitz provided there is $M \geq 0$ such that $|g(x) - g(y)| \leq M\|x - y\|$ for all $x, y \in X$, and say $g$ is $M$-Lipschitz on $X$. We write $Lip(g)$ for the least such $M$. If $g$ is Lipschitz on some neighbourhood

of each point of $X$ we say it is locally Lipschitz. In particular the distance function $d(\cdot, A)$ to a closed nonempty subset $A$ of $E$ is 1-Lipschitz on $E$.

If $f$ is proper convex lower semicontinuous on $E$ the subdifferential $\partial f(x)$ is defined by $\partial f(x) := \{x^* \in E^* : \langle y - x, x^* \rangle \leq f(y) - f(x) \text{ for all } y\}$ while for $g$ locally Lipschitz we use the Clarke subdifferential $\partial g(x) := \{x^* \in E^* : \langle y, x^* \rangle \leq \limsup\limits_{z \to x, t \downarrow 0} \frac{g(z + ty) - g(z)}{t}, \forall y\}$

Finally for $f$ lower semicontinuous we use $\partial f(x) := \{x^* \in E^* : (x^*, -1) \in \mathbb{R}_+ \partial d(\cdot, \text{epi}(f))(x, f(x))\}$.

This agrees with the definitions above if $f$ is locally Lipschitz or convex.

The conjugate of a proper convex lower semicontinuous function $f$ on $E$ is defined by $f^*(x^*) := \sup\{\langle x, x^* \rangle - f(x) : x \in E\}$ for $x^* \in E^*$. We will use Fenchel’s equation \[
 f^*(x^*) + f(x) = \langle x, x^* \rangle \quad \text{if} \quad x^* \in \partial f(x) \quad (1.1)
\]
and the fact that if \( f \) is proper convex lower semicontinuous on \( \mathbb{R}^n \) and \( f^* \) is differentiable at \( x^* \) then \( \nabla f^*(x^*) \in \text{dom} f^* \), indeed \( x^* \in \partial f(\nabla f^*(x^*)) \).

We denote the convex hull of the union of two convex sets \( X \) and \( Y \) by \([X,Y]\).

To smoothen our functions we have two devices, given in the following two lemmas.

**Lemma 1.1.** If \( f \) is proper convex lower semicontinuous with \( \text{dom}(f) \) a bounded subset of \( \mathbb{R}^n \) and \( \epsilon > 0 \) then \((f + \epsilon \| \cdot \|^2)^* \) is continuously differentiable on \( \mathbb{R}^n \).

**Proof.** This follows from \( f + \epsilon \| \cdot \|^2 \) being strictly convex with bounded domain. QED

**Proposition 1.2.** If \( g \) is a Lipschitz real-valued function on \( B(x; \epsilon) \subset \mathbb{R}^n \) and \( \phi \) is a continuously differentiable nonnegative function with support contained in \( B(0; \epsilon) \) and integral equal to 1 let \( g_\phi := \phi \cdot g \). Then \( g_\phi \) is continuously differentiable, \( |g_\phi - g| \leq \epsilon \text{Lip}(g) \) and \( \nabla g_\phi(x) \in \text{conv} \partial g(\text{B}(x; \epsilon)) \).

**Proof.** We can represent \( g_\phi \) as an integral functional in two ways:

\[
g_\phi(y) = \int g(z)\phi(y - z)d\lambda^n(z) = \int g(y - z)\phi(z)d\lambda^n(z)
\]

and apply [1], Theorem 2.7.2, to get in the respective cases,

\[
\nabla g_\phi(y) = \{\nabla g_\phi(y)\} = \int g(z)\nabla \phi(y - z)d\lambda^n(z)
\]

\[
\subset \int \partial g(y - z)\phi(z)d\lambda^n(z)
\]

so that \( \nabla g_\phi(x) \in \text{conv} \partial g(B(x; \epsilon)) \) and since

\[
g(x) - g_\phi(x) = \int (g(x) - g(x - z))\phi(z)d\lambda^n(z)
\]

we have \( |g(x) - g_\phi(x)| \leq \epsilon \text{Lip}(g) \). QED

**Remark.** The continuous differentiability of \( g_\phi \) is a standard result. We have not seen the precise subdifferential inclusion stated explicitly, but it can be proved (not quite as easily) without the use of Clarke’s integral theorem.

\[\]

\[\]

\[\geq -9M\|x - u\|,\]

as replacing the general \( b \) by a particular \( z \) in the \( \inf \) moves the inequality in the right way for us. But that follows from \( f(w) - f(w) \geq -M\|b - w\|, \|y\|^2 - \|w\|^2 \geq -M\|y - w\| \) and \( \|y - w\| = \|x - u\| \).

To calculate the subdifferential, we can represent

\[q(x) = 1 \wedge \inf_z \{f(x + z) - h(z) + \|x + z\|^2 + \|z\|^2\}
\]

so that \( \partial q(0) \subset \partial(f + \| \cdot \|^2)(0) = \partial f(0) \) using [1], Theorem 2.8.2, as \( f(z) - h(z) + 2\|z\|^2 > 0 \) if \( z \neq 0 \).

Similarly

\[q(x) = 1 \wedge \inf_y \{f(y) - h(y - x) + \|y\|^2 + \|y - x\|^2\}
\]

so that \( \partial q(0) \subset \partial(-h(-) + \| \cdot \|^2)(0) = \partial h(0) \). QED

**Remark.** The use of Clarke’s general pointwise maximum theorem could be avoided at the expense of an intricate direct calculation. This lemma also holds with \( \mathbb{R}^n \) replaced by a general Banach space. The minimum with 1 is used only to assure the global Lipschitzness of \( q \), without it \( q \) would grow like \( \| \cdot \|^2 \).
2. Three Functions

We now establish our fundamental inequality in finite dimensions. Our proof shows that the result is really no easier for smooth functions and seemingly relies unavoidably on topological fixed point theory. In section five we will extend the result to the general Banach space setting.

**Theorem 2.1.** Let $C$ be a nonempty compact convex subset of $\mathbb{R}^n$ and $f$ and $h$ proper convex lower semicontinuous functions with $f^*$ and $h^*$ continuously differentiable and $\text{dom}(f) \cup \text{dom}(h) \subset C$. For any continuously differentiable $g : C \to \mathbb{R}$ there is $z \in C$ such that

$$\max(g-f) + \max(-g-h) \geq f^*(\nabla g(z)) + h^*(\nabla g(z))$$

**Proof.** Let $M := 2\sup\{\|c\| : c \in C\}$ and $W := \{x : [0,1] \to C : \text{Lip}(x) \leq M\}$. Then $W$ is compact in the uniform norm topology, by the Arzela-Ascoli theorem, [4]. For $x \in W$ define

$$T x(t) := \int_0^t \nabla f^* \circ \nabla g \circ x + \int_t^1 \nabla h^* \circ (\nabla g) \circ x$$

and note that $T : W \to W$ is continuous. Since $W$ is compact and convex, the Schauder fixed point theorem (see [3] page 60) shows that there is $x \in W$ such that $x = Tx$. That is, $x(t) = \int_0^t \nabla f^* \circ \nabla g \circ x + \int_t^1 \nabla h^* \circ (\nabla g) \circ x$.

Now we have

$$g(x(1)) - g(x(0)) = \int_0^1 \langle \nabla f^* \circ \nabla g \circ x - \nabla h^* \circ (\nabla g) \circ x, \nabla g \circ x \rangle$$

$$= \int_0^1 \langle \nabla f^* \circ \nabla g \circ x, \nabla g \circ x \rangle$$

$$+ \int_0^1 \langle \nabla h^* \circ (\nabla g) \circ x, -\nabla g \circ x \rangle$$

$$= \int_0^1 \langle f^* \circ \nabla g \circ x + f \circ \nabla f^* \circ \nabla g \circ x \rangle$$

$$+ \int_0^1 \langle h^* \circ (\nabla g) \circ x + h \circ \nabla h^* \circ (\nabla g) \circ x \rangle$$

by Fenchel’s equation (1.1), so that by the integral form of Jensen’s inequality,

$$g(x(1)) - g(x(0)) \geq \int_0^1 (f^* \circ \nabla g \circ x + h^* \circ (\nabla g) \circ x)$$

$$+ f \left( \int_0^1 \nabla f^* \circ \nabla g \circ x + h \left( \int_0^1 \nabla h^* \circ (\nabla g) \circ x \right) \right)$$

$$= \int_0^1 (f^* \circ \nabla g \circ x + h^* \circ (\nabla g) \circ x) + f(x(1) + h(x(0)))$$

so that for some $z = x(t) \in C$ we have

$$g(x(1)) - g(x(0)) - f(x(1)) - h(x(0)) \geq f^*(\nabla g(z)) + h^*(\nabla g(z)).$$

Thus $x(1) \in C$ and $x(0) \in C$ will give the required inequality. QED

We may now establish a hybrid inequality mixing features of the Fenchel duality theorem and of (non-smooth) mean-value theorems. Recall that the classic formulation of Fenchel duality asserts that

$$\min(f + h) = \max(-f^* + h^*(-))$$

when $0 \in \text{relint}(\text{dom}(f) - \text{dom}(h))$.

**Theorem 2.2.** Let $C$ be a nonempty compact convex subset of $\mathbb{R}^n$ and $f$ and $h$ proper convex lower semicontinuous functions with $\text{dom}(f) \cup \text{dom}(h) \subset C$. For any Lipschitz $g : C \to \mathbb{R}$ there is $z^* \in \partial g(C)$ such that

$$\min(f - g) + \min(h + g) \leq -f^*(z^*) - h^*(-z^*)$$

$$\leq \min(f + h).$$

**Proof.** The right-hand inequality follows easily from the definition of the Fenchel conjugate. To obtain the other inequality, let $x > 0$ and $\phi$ as in Proposition 1.2. Then we apply Theorem 2.1 to $f + x \phi$ and $h \phi$. Since $C$ is compact, in the limit as $x \to 0$ we get the desired conclusion. QED

**Corollary 2.3.** (The Lewis-Ralph sandwich theorem, [8]) Let $C$ be a nonempty compact convex subset of $\mathbb{R}^n$ and $f$ and $h$ proper convex lower semicontinuous functions with $\text{dom}(f) \cup \text{dom}(h) \subset C$. For any Lipschitz $g : C \to \mathbb{R}$ such that $f \geq g \geq -h$ there is $z^* \in \partial g(C)$

$$0 \geq f^*(z^*) + h^*(-z^*).$$

Note that the inequality is potentially useful when only one of $f \geq g$ or $g \geq -h$ holds. On considering $\mathcal{F} := f - \min(f - g)$ and $\mathcal{H} := h - \min(h + g)$ we can actually recover Theorem 2.2 from Corollary 2.3.
Corollary 2.4. (The Clarke-Ledyaev two set mean-value theorem, [2]) Let \( X \) and \( Y \) be compact convex subsets of \( \mathbb{R}^n \) and \( g : [X,Y] \to \mathbb{R} \) Lipschitz. Then there is \( z^* \in \partial g([X,Y]) \) such that
\[
\langle x - y, z^* \rangle \leq \max_x g - \min_y g
\]
for all \( x \in X \) and \( y \in Y \).

Proof. In Theorem 2.2 we take \( f := I_X \) and \( h := I_Y \) and \( C := [X,Y] \). Since \( f^*(z^*) = \max_x z^* \) and \( h^*(z^*) = \max_y z^* \) we get the desired inequality. QED

Remark. We observe that the key in Theorem 2.1 is to show that \( T \) has a fixed point, the rest of the argument follows much more easily. When \( f^* \) or \( h^* \) is affine one may deduce the existence of a fixed point from the Banach contraction principle and so from Ekeland’s principle (see [1], [5]).

3. Two Functions

Revisiting our technique in the case when one of \( f \) and \( h \) is removed (relatively when \( f = h(\cdot - \cdot) \)), we can quite substantially improve our conclusions.

Theorem 3.1. Let \( C \) be a nonempty compact convex subset of \( \mathbb{R}^n \) and \( f \) a proper convex lower semi-continuous function with \( \text{dom}(f) \subset C \) and \( f^* \) continuously differentiable. If \( \alpha \neq 1 \) and \( g : [C,\alpha C] \to \mathbb{R} \) is continuously differentiable then there are \( z \in [C,\alpha C] \) and \( a \in C \) such that
\[
(g(a\alpha) - g(a))/(\alpha - 1) - f(a) \geq f^*(\nabla g(z)).
\]

Proof. Let \( M := (1 + |\alpha|)\sup \{ ||c|| : c \in C \} \) and \( W := \{ x : [0,1] \to [C,\alpha C] : \text{Lip}(x) \leq M \} \). Then \( W \) is compact in the uniform norm topology, again by the Arzela-Ascoli theorem. For \( x \in W \) define
\[
Tx(t) := \alpha \int_0^t \nabla f^* \circ \nabla g \circ x + \int_0^t \nabla f^* \circ \nabla g \circ x
\]
and note that \( T : W \to W \) is continuous. Since \( W \) is compact and convex, once more the Schauder fixed point theorem shows that there is \( x \in W \) such that \( x = Tx \). That is,
\[
x(t) = \alpha \int_0^t \nabla f^* \circ \nabla g \circ x + \int_1^t \nabla f^* \circ \nabla g \circ x.
\]
Now we have
\[
g(x(1)) - g(x(0)) = (\alpha - 1) \int_0^1 \langle \nabla f^* \circ \nabla g \circ x, \nabla g \circ x \rangle
\]
\[
= (\alpha - 1) \int_0^1 \left( \nabla f^* \circ \nabla g \circ x + f \circ \nabla f^* \circ \nabla g \circ x \right)
\]
by Fenchel’s equation, so that by Jensen’s inequality,
\[
\left( g(x(1)) - g(x(0)) \right)/(\alpha - 1) \\
\geq \int_0^1 \nabla f^* \circ \nabla g \circ x + f \left( \int_0^1 \nabla f^* \circ \nabla g \circ x \right)
\]
\[
= \int_0^1 \nabla f^* \circ \nabla g \circ x + f(x(0))
\]
and letting \( a := x(0) \) we see there is some \( z = x(t) \in [C,\alpha C] \) such that
\[
(g(\alpha a) - g(a))/(\alpha - 1) - f(a) \geq f^*(\nabla g(z)).
\]
since \( x(1) = \alpha x(0) = \alpha a \). QED

Theorem 3.2. Let \( C \) be a nonempty compact convex subset of \( \mathbb{R}^n \) and \( f \) a proper convex lower semi-continuous function with \( \text{dom}(f) \subset C \). If \( \alpha \neq 1 \) and \( g : [C,\alpha C] \to \mathbb{R} \) is Lipschitz then there are \( z^* \in \partial g([C,\alpha C]) \) and \( a \in C \) such that
\[
(g(a\alpha) - g(a))/(\alpha - 1) - f(a) \geq f^*(z^*).
\]

Proof. This follows from Theorem 3.1 in the same way that Theorem 2.2 followed from Theorem 2.1. QED

Two pleasant specialisations follow.

Corollary 3.3. Let \( C \) be a nonempty compact convex subset of \( \mathbb{R}^n \) and \( f \) a proper convex lower semi-continuous function with \( \text{dom}(f) \subset C \). If \( g : [C, \alpha C] \to \mathbb{R} \) is Lipschitz then there are \( z^* \in \partial g([C, \alpha C]) \) and \( a \in C \) such that
\[
(g(a) - g(-a))/(2 - f(a)) \geq f^*(z^*)
\]
which becomes \( f^*(z^*) \leq 0 \) if \( f \) dominates the odd part of \( g \) on \( C \).

The comparison of \( f \) to the odd part of \( g \) reinforces the suggestion that fixed point theory is central to these results.
Corollary 3.4. Let $C$ be nonempty, compact and convex in $\mathbb{R}^n$ and $f$ a proper convex lower semicontinuous function with $\text{dom}(f) \subset C$. If $g : [C, 0] \to \mathbb{R}$ is Lipschitz then there are $z^* \in \partial g([C, 0])$ and $a \in C$ such that $f(a) + f^*(z^*) \leq g(a) - g(0)$. It becomes $f^*(z^*) \leq 0$ if $f$ dominates $g - g(0)$ on $C$.

Corollary 3.5. Let $B$ be the closed unit ball in $\mathbb{R}^n$ and $g : B \to \mathbb{R}$ a Lipschitz function. Then for $\alpha \in [-1, 1]$ there is $x^* \in \partial g(B)$ such that

$$\|x^*\| \leq \max_{a \in B}(g(\alpha a) - g(a))/(\alpha - 1).$$

When $g$ is smooth, a more careful analysis shows that when $B$ is strictly convex the fixed point arc, $x$ either lies entirely interior to the ball or is a straight line with endpoints on the sphere. In any event we may assert that $x^* \in \partial g(\text{int} B)$. Example 2 in [2] illustrates the need for strict convexity.

Corollary 3.6. Let $B$ be the closed unit ball in $\mathbb{R}^n$ and $g : B \to \mathbb{R}$ a Lipschitz function. Then there is $x^* \in \partial g(B)$ such that

$$\|x^*\| \leq \max_{a \in B}(g(a) - g(-a))/2.$$

which is impossible as $\alpha + \epsilon < t$. Thus $u \in \partial X$. Now putting $y := x$ we obtain $h(u) \leq h(x) + \|u - x\|$, that is, $\alpha\|u - x\| - g(u) \leq \alpha\|x - x\| - g(x) + \|u - x\|$. Therefore

$$\sup_{u \in \partial X} (g(u) - (\alpha - \epsilon)\|u - x\|) \geq g(x).$$

Now let $\alpha \to t$ and $\epsilon \to 0$ to get our inequality. QED

Note that Theorem 3.8 holds for any generalised gradient with a reasonable calculus such as that of Moreau or that of Michel-Penot. As an immediate application we recover a non-smooth version of Rolle's theorem asserting when $\sup_{u \in \partial X} g(u) = 0 = \inf_{u \in \partial X} g(u)$ that $g$ has an approximate critical point in int $X$.

Corollary 3.9. Let $B$ be the closed unit ball in a Banach space $E$ and $g : B \to \mathbb{R}$ Lipschitz. Then

$$\inf\{\|z\| : z^* \in \partial g(z), \|z\| < 1\} \leq (\sup_{\partial B} g - g(0)) \vee 0$$

Proof. Put $x := 0$ and $X := B$ in Theorem 3.8. QED

We note that in [6] an example is given of a globally Lipschitz $C^1$ function on $l_2$ which vanishes on the unit sphere and has no critical point in the closed unit ball.

Remark. In the last two sections, all our results rely on uniform approximation by smooth functions. They thus remain valid if $g$ is merely assumed continuous and Warga's derivative containers [11] are used. In the locally Lipschitz case, Clarke's subdifferential is the smallest such convex container. It is instructive to examine Example 5.10 from this perspective.
4. Squeezed Functions

Lewis and Lucchetti [7] observed that when, in the setting of Corollary 2.3, \( f + h = 0 \) and is attained at \( x_0 \), one deduces that \( f, g \) and \( h \) share a common subgradient at the point. They asked if this somewhat unexpected result had a variational proof and investigated as to what was true with more than three functions. Theorems 4.1, 4.2 and 4.3 provide variational proofs of appropriate generalisations.

**Notation.** Given \( f \) and \( h \) on \( \mathbb{R}^m \) we denote \((f \# h)(x) := \min(1, \inf \{ f(y) - h(z) + \|y\|^2 + \|z\|^2 : x = y - z \})\)
as appeared in Lemma 1.3. Our basic result, normalised to \( x_0 = 0 \), is:

**Theorem 4.1.** Let \( g_1 \geq g_2 \geq \ldots \geq g_k \) be Lipschitz on \( \mathbb{R}^n \) with \( g_1(0) = g_2(0) = \ldots = g_k(0) = 0 \). Then

\[ \partial g_1(0) \cap \partial g_2(0) \cap \ldots \cap \partial g_k(0) \neq \emptyset. \]

**Proof.** For \( k = 2 \) this is true by Lemma 1.3 as \( \partial (g_1 \# g_2)(0) \) is nonempty. If it is true for \( k = m \) then given \( g_1 \geq g_2 \geq \ldots \geq g_{m+1} \) Lipschitz functions

on \( \mathbb{R}^n \) with \( g_1(0) = g_2(0) = \ldots = g_{m+1}(0) = 0 \) we define \( h_i := g_i \# g_{m+1} \) for \( 1 \leq i \leq m \). By Lemma 1.3, \( h_i \) are Lipschitz, \( \partial h_i(0) \subset \partial g_i(0) \cap \partial g_{m+1}(0) \) and \( h_i(0) = 0 \). Clearly \( h_1 \geq h_2 \geq \ldots \geq h_m \) so we have

\[ \partial h_1(0) \cap \partial h_2(0) \cap \ldots \cap \partial h_m(0) \neq \emptyset \]

from which it follows that

\[ \partial g_1(0) \cap \partial g_2(0) \cap \ldots \cap \partial g_{m+1}(0) \neq \emptyset. \]

QED

Since the conclusion is local to the point of attainment, we may equally well assume, here and below, that the functions are only locally Lipschitz. The Lipschitz condition may also be relaxed dramatically as follows:

**Theorem 4.2.** Let \( f_1 \geq f_2 \geq \ldots \geq f_j \) be lower semicontinuous, \( h_1 \geq h_2 \geq \ldots \geq h_j \) upper semicontinuous and \( f_j \geq g \geq h_1 \) for some Lipschitz \( g \) on \( \mathbb{R}^n \). If \( 0 = g(0) = h_1(0) = f_j(0) \) for \( 1 \leq i \leq j \) then there is \( x^* \in \partial g(0) \) such that \( x^* \in \partial f_i(0) \) and \( x^* \in -\partial(-h_i)(0) \) for \( 1 \leq i \leq j \).

**Proof.** Choose \( K > 1 + \text{Lip}(g) \) and define \( p_{j+1}(x, r) := g(x) - r \) and

\[ p_i(x, r) := Kd((x, r), \text{epi}(f_i)), \]

\[ p_{i+j+1}(x, r) := -Kd((x, r), h_i) \]

for \( 1 \leq i \leq j \). Then \( p_1 \geq p_2 \geq \ldots \geq p_{2j+1} \) are Lipschitz on \( \mathbb{R}^n \times \mathbb{R} \). Indeed Lipschitz constant \( K \) will work for all of them, while \( p_i \geq p_{i+1} \) for \( 1 \leq i < j \) as \( \text{epi}(f_i) \subset \text{epi}(f_{i+1}) \) and similarly \( p_{i+j+1} \geq p_{i+j+2} \) for \( 1 \leq i < j \). So we need to show \( p_j \geq p_{j+1} \geq p_{j+2} \).

If \( r \geq g(x) \) then \( p_{j+1}(x, r) \leq 0 \leq p_j(x, r) \). Otherwise

\[ p_j(x, r) = Kd((x, r), \text{epi}(f_j)) \geq Kd((x, r), \text{epi}(g)) \geq Kd((x, r), g(x) - r) \]

\[ = K(g(x) - r)/\sqrt{1 + (\text{Lip}(g))^2} > g(x) - r = p_{j+1}(x, r) \]

where the distance calculation follows from the formula for the distance from a point to a line in the plane. Similarly we obtain \( p_{j+1} \geq p_{j+2} \).

Also it is immediate that \( p_i(0, 0) = 0 \) for \( 1 \leq i \leq 2j+1 \). Thus we can apply Theorem 4.1 to find \((x^*, r^*) \)
in \( \mathbb{R}^n \times \mathbb{R} \) such that \((x^*, r^*) \in \partial p_i(0, 0) \) for \( 1 \leq i \leq 2j+1 \). But it is clear setting \( i = j+1 \) that \( r^* = -1 \) and \( x^* \in \partial g(0) \). Now it follows from the definition that \((x^*, r^*) \in \partial f_i(0) \) and \((x^*, r^*) \in -\partial(-h_i)(0) \) for \( 1 \leq i \leq j \). QED

Note that on repeating functions, there is no loss in assuming an equal number of \( h_i \) and \( f_i \). We may move to infinite dimensions painlessly:

**Theorem 4.3.** Let \( f_1 \geq f_2 \geq \ldots \geq f_j \) be lower semicontinuous, \( h_1 \geq h_2 \geq \ldots \geq h_j \) upper semicontinuous and \( f_j \geq g \geq h_1 \) for some Lipschitz \( g \) on a Banach space \( E \). If \( 0 = g(0) = h_1(0) = f_j(0) \) for \( 1 \leq i \leq j \) then there is \( x^* \in \partial g(0) \) such that \( x^* \in \partial f_i(0) \) and \( x^* \in -\partial(-h_i)(0) \) for \( 1 \leq i \leq j \).

**Proof.** For each finite dimensional subspace \( F \) of \( E \) we can apply Theorem 4.2 to get \( x^*_F \in F^* \) such that \( x^*_F \in \partial \theta(g|_F) (0) \), \( x^*_F \in \partial \theta(f_i|_F) (0) \) and \( x^*_F \in -\partial(-h_i|_F) (0) \) for \( 1 \leq i \leq j \). Also there is \( y^*_F \in \partial g(0) \) such that \( x^*_F = y^*_F \). By the Hahn-Banach theorem.

Now ordering the finite dimensional subspaces of \( E \) by inclusion we have a net \( (y^*_p) \) and since \( g \) is Lipschitz it is a bounded net. Now let \( x^* \) be the weak*
limit of any weak* convergent subnet of $(y_i^*)$. Then $x^* \in \partial g(0)$ and $x^* \in \partial f_i(0)$ and $x^* \in -\partial(-h_i)(0)$ for $1 \leq i \leq j$.

**Alternative proof.** Since Lemma 1.3 works for a general Banach space, so do Theorems 4.1 and 4.2 with no change to the structure of the proofs. QED

We note that an application of the finite intersection property means that Theorem 4.3 remains valid generally for infinite monotone collections of functions, \(\{f_i\}_{i \in I}\) and \(\{g_j\}_{j \in J}\).

Applying Theorem 4.3 in a product space leads to a different generalisation of Lewis and Lucchetti’s squeeze theorem.

**Corollary 4.4. (Parallel squeeze theorem) Let \(E\) be a Banach space. Consider \(f_k : E \rightarrow (-\infty, \infty]\) proper, convex and lower semicontinuous and \(g_k : E \rightarrow R\) locally Lipschitz. Suppose \(\sum_{k=1}^{N} f_k(0) = 0\), and \(f_k(x) \geq g_k(x)\) for \(x\) near zero and \(k = 1, \ldots, N\). Suppose also that \(\sum_{k=1}^{N} g_k \geq 0\).

Then there are \(\lambda_k, k = 1, \ldots, N\) such that

\[
\lambda_k \in \partial f_k(0) \cap \partial g_k(0)
\]

while

\[
\sum_{k=1}^{N} \lambda_k = 0.
\]

**Proof.** (\(N > 1\)) Apply Theorem 4.3 for \(x \in E^N\) to

\[
f(x) := \sum_{k=1}^{N} f_k(x_k), \quad g(x) := \sum_{k=1}^{N} g_k(x_k), h := -I_N
\]

where \(I_N\) is the diagonal in \(E^N\), QED.

The original result of Lewis and Lucchetti is recovered by setting \(g_1 := g, g_2 := -g, f_1 := f, f_2 := -h\).

**5. Extensions and Examples.**

The most immediate extension to infinite dimensions occurs when \(g\) is globally Lipschitz.

**Theorem 5.1.** Let \(E\) be a Banach space and \(g : E \rightarrow R\) globally Lipschitz and \(f\) and \(h\) proper convex lower semicontinuous on \(E\). Then there is \(z^*\) in the weak* closure of the range of \(\partial g\) such that

\[
\sup(g - f) + \sup(-g - h) \geq f^*(z^*) + h^*(-z^*).
\]

**Proof.** For \(F \subset E\) finite-dimensional and \(m\) a large enough positive integer we can take \(g|_F, f|_F + I_{mB_E}\) and \(h|_F + I_{mB_E}\) in Theorem 2.2 to get \(z^*_m|_F \in \partial (g|_F) (z_m, F)\) such that

\[
\max_{mB_E} (g - f) + \max_{mB_E} (-g - h) \geq (f|_F + I_{mB_E})^* (z^*_m|_F) + (h|_F + I_{mB_E})^* (-z^*_m|_F).
\]

Now there is \(z^*_m|_F \in \partial g(z_m, F)\) such that \(z^*_m|_F = z^*_m|_F\).

Since \(g\) is Lipschitz, the net \((z^*_m|_F)\), \((m, F\) increasing), is bounded and the weak* limit \(z^*\) of any weak* convergent subnet \((x^*_{m,F})\) is in the weak* closure of the range of \(\partial g\).

Now for any \(x\) and \(y\) we have \(\beta\) so that \(\alpha \geq \beta\) implies \(m_{\alpha} \geq \|x\|, m_{\alpha} \geq \|y\|\) and \((x, y) \subset F_{\alpha}\). Thus

\[
\sup(g - f) + \sup(-g - h) \geq (x, z^*_{x_m, F_{\alpha}}) - f(x) - (y, x^*_{x_m, F_{\alpha}}) - h(y) \rightarrow (x, z^*) - f(x) - (y, z^*) - h(y)
\]

and taking the sup over \(x\) and \(y\) gives our result. QED.

The following example shows that even on \(R\) one cannot generally avoid taking the closure.

**Example 5.2.** If \(f := 2, h := 2\) and \(g := \arctan\) then \(g\) strongly separates \(f\) and \(-h\) on \(R\) but there is no \(z \in R\) such that \(f^*(\nabla g(z)) + h^*(-\nabla g(z)) \leq 0\).

However if the range of \(\partial g\) is weak* closed then that is not an issue. Here is one case where that is guaranteed.
Theorem 5.3. Let $E$ be a Banach space and $g : E \to \mathbb{R}$ locally Lipschitz and $f$ and $h$ proper convex lower semicontinuous on $E$. If $\text{dom}(f)$ and $\text{dom}(h)$ are relatively compact then there is $z^*$ in the range of $\partial g$ such that

$$\sup (g - f) + \sup (-g - h) \geq f^*(z^*) + h^*(-z^*).$$

Proof. Consider $C := \text{dom}(f) \cap \text{dom}(h)$. Then $C$ is compact so $\partial g(C)$ is weak* closed. Since $g$ is Lipschitz on $C$ the result follows from Theorem 5.1. QED

We define the core of a convex set $C$ to be $\text{core}(C) := \{ x \in C : \forall y \in E \exists x + ty \in C \text{ for some } t > 0 \}.$

We note the following useful form of the Baire category principle for convex functions (Simons [10]).

Lemma 5.4. (The dom-dom lemma) Let $f$ and $h$ be proper convex lower semicontinuous on $E$ and $0 \in \text{core}(\text{dom}(f) - \text{dom}(h))$. Then there are $n > 0$ and $\epsilon > 0$ such that if $\| x \| < \epsilon$ then $x = y - z$ for some $y$ and $z$ in $B[0 ; n]$ such that $f(y) \leq n$ and $h(z) \leq n$.

and $h(z) \leq n$. Now if $K \subset F$ and $m \geq n$, for $x \in F$ with $\| x \| < \epsilon$ we have $y \in F$ as $z \in K$.

So

$$\langle x, z^*_m \rangle = \langle y - z, z^*_m \rangle \leq \sup (g - f) + \sup (-g - h) + f(y) + h(z) \leq \sup (g - f) + \sup (-g - h) + 2n$$

so that $\| z^*_m \| = \| z^*_m \| \leq (\sup (g - f) + \sup (-g - h) + 2n) / \epsilon$. Thus $\langle z^*_m \rangle (m \geq n, K \subset F)$ is a bounded net in $E^*$ and we take $z^*$ to be the weak* limit of any weak* convergent subnet $(z^*_{m_i, n_i})$. As in Theorem 5.1, we have $\sup (g - f) + \sup (-g - h) \geq f^*(z^*) + h^*(-z^*)$.

Finally, to see $z^*$ is in the weak* closure of the range of $\partial g$, let $x_1, \ldots, x_k \in E$ and $\delta > 0$. Then there is $\beta$ such that for $\alpha \geq \beta$ we have $\| x_{i+1} \| - \| x_i \| < \delta$ and $x_i \in F_n$ for $i = 1, \ldots, k$. Now for any $v^*_m \in \partial g(z^*_m, n_m)$ which agrees with $g^*_m$ on $F_n$ we have $\| v^*_m \| = \| v^*_m \| < \delta$ for $i = 1, \ldots, k$. Thus a weak* neighbourhood of $z^*$ cannot avoid the range of $\partial g$. QED

Suppose $f$ is everywhere continuous. By taking $g := f$ we recover a reasonably general version of the classical Fenchel duality theorem.

Theorem 5.5. Let $E$ be a Banach space and $g : E \to \mathbb{R}$ locally Lipschitz and $f$ and $h$ proper convex lower semicontinuous on $E$. If

$$0 \in \text{core}(\text{dom}(f) - \text{dom}(h + I_K))$$

for some finite dimensional subspace $K$ of $E$ then there is $z^*$ in the weak* closure of the range of $\partial g$ such that

$$\sup (g - f) + \sup (-g - h) \geq f^*(z^*) + h^*(-z^*).$$

Proof. We proceed as in the proof of Theorem 5.1 to find $z^*_m \in \partial (g^*_m) \cap \text{dom}(f)$ such that

$$\max (g - f)|_{B \rho} + \sup (-g - h) \geq (f|_{B \rho} + I_{B \rho})^*(z^*_m) + (h|_{B \rho} + I_{B \rho})^*(-z^*_m).$$

Now we take $x^*_m \in E^*$ which agree with $z^*_m$ on $F$ and $\| x^*_m \| = \| z^*_m \|$ by the Hahn-Banach theorem.

Next we apply Lemma 5.4 to $f$ and $h + I_K$ to get $n > 0$ and $\epsilon > 0$ such that if $\| y \| < \epsilon$ then $x = y - z$ for some $y$ and $z$ in $B[0 ; n]$ such that $f(y) \leq n$, $z \in K$

Remark. Considering the infimal convolution $f_M := f \square M : \| \| \text{ for } M > \text{ Lip}(g)$ we may rederive Theorem 5.1 from Theorem 5.5 applied to $f_M$ rather than $f$. In finite dimensions we may replace $0 \in \text{core}(\text{dom}(f) - \text{dom}(h + I_K))$ by $0 \in \text{relin}(\text{dom}(f) - \text{dom}(h))$.

We also observe that $g$ need only be assumed locally Lipschitz on $\text{cl}(\text{dom}(f), \text{dom}(h))$.

Example 5.6. We note that if $f$ and $h$ are separated by a locally Lipschitz function and their domains are disjoint but distance zero apart while $f - h \geq \epsilon > 0$ then there can be no uniformly continuous let alone globally Lipschitz or affine separator. Such an example [7] is afforded by $g : (x, y) := -xy, f(x, y) := I_P$ where $P := \{(x, y) : |xy| \geq 1, x \geq 0 \}$ and $h := I_{(x \geq 0)}$. Relatedly, the function $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) := \sin(\exp(x)) / (1 + x^2)$ is locally Lipschitz, uniformly continuous and not globally Lipschitz.

When $f$ and $h$ grow super-linearly, we may recapture an extension of one of Lewis and Lucchetti’s results.
Corollary 5.7. (Coercivity) Suppose in the setting of Theorem 5.1 or Theorem 5.5 that \(f\) and \(h\) are coercive with domains that are boundedly relatively compact. Then the conclusion obtains for some \(z^*\) in the range of \(\partial g\).

Proof. Since \(C_m := C \cap mB_{R \epsilon}\) is compact we may apply Theorem 2.2 to \(C_m\) for \(m > 0\). An inspection of the proof of Theorem 5.1 or Theorem 5.5 reveals that the dual vectors \(z_m^*\) remain norm bounded as \(m \to \infty\). Let \(f_m := f + I_{C_m}, \ h_m := h + I_{C_m}\). For sufficiently large \(m\), the coercivity of \(f\) and \(h\) implies that

\[
\sup(g-f) + \sup(-g-h) \geq \sup(g-f_m) + \sup(-g-h_m)
\]

\[
\geq f^*_m(z_m^*) + h^*_m(-z_m^*) = f^*(z_m^*) + h^*(z_m^*).
\]

QED

Remark. We observe that Clarke and Ledyaev give an approximate version of Corollary 2.4, valid in Banach space when one of \(X\) and \(Y\) is compact and the other is merely bounded. Theorem 2.1 may be similarly adapted to the case when one of \(\text{dom}(f)\) and \(\text{dom}(h)\) is relatively compact and the other bounded.

33

Corollary 5.7 has an easier counterpart in the two function setting. We need only assume that \(g\) is globally Lipschitz or that \(f\) is (relatively) continuous at zero. We illustrate with one application.

Example 5.8. Suppose that \(E\) is finite dimensional and that \(g\) is a locally Lipschitz function whose odd part is majorised by \(\|\cdot\|^p/p + K\) for some \(1 < p < \infty\) and \(K \geq 0\). Then \(g\) has a subgradient whose dual norm is no greater than \((Kq)^{1/p}\) where \(1/q + 1/p = 1\). The inequality \(x \leq |x|^p/p + 1/q\) shows that the estimate may be best possible.

Example 5.9. We record that the inequalities of the last few sections may well be strict. For example, \(f := h := I_{[-1,1]}\) and \(g\) the identity shows that Theorem 2.2 may be strict. Also, easy examples show that even in one dimension Corollaries 3.6 and 3.7 are incomparable and that in Corollary 3.5 the infimum over \(\alpha \in [0,1]\) may be greater than the maximum value of a subgradient.

Indeed let \(g(x) := x \exp(-|x|)\) on \([-1,1]\). Then as \(t\) increases to \(1\)

\[
\max\{tg(x) - g(x)/(t - 1), |x| \leq 1\} \to \gamma^3 \exp(-\gamma^2)
\]

\[= 0.16112\ldots\]

And we finish with the following:

Open Problems.

(i) In Theorem 5.1, suppose \(g\) is merely locally Lipschitz but that an affine separator is assumed to exist. Must the conclusion of the theorem obtain?

(ii) Suppose, in Banach space that the hypotheses of Theorem 2.2 hold, but \(g\) is only assumed locally Lipschitz and \(C\) bounded. Construct an example with

\[
\sup(g-f) + \sup(-g-h) < \inf f^*(z^*) + h^*(-z^*)
\]
REFERENCES


APPENDIX

Proof of the Remark after Lemma 1.3.

For x near 0 we have

\[ f(x + u) - h(u) + \|x + u\|^2 + \|u\|^2 \]
\[ = f(x + u) - f(u) + f(u) - h(u) \]
\[ + \|x + u\|^2 + \|u\|^2 \]
\[ \geq -M\|x\| + 0 + \|x + u\|^2 + \|u\|^2 \]
\[ \geq 2\|u\|^2 - (M + 2\|u\|)\|x\| \]

so that we may suppose the inf, which is bounded by \(9M\|x\|\), occurs when \(\|u\|^2 \leq \|x\|\).

Now as \(x \to 0\) and \(t \to 0+\), for each \(y\) in \(E\) we have

\( (q(x + ty) - q(x))/t = \)
\( \inf_{z} (f(x + ty + z) - h(z) + \|x + ty + z\|^2 + \|z\|^2)/t) \)
\( - \inf_{u} (f(x + u) - h(u) + \|x + u\|^2 + \|u\|^2)/t \)
\( \leq \sup_{\|u\|^2 \leq \|x\|} (f(x + u + ty) - f(x + u) \)
\( + \|x + u + ty\|^2 - \|x + u\|^2)/t \)

and as \(x + u \to 0\) and \(t \to 0+\) we have \(q^*(0; y) \leq f^*(0; y) + 0\) whence \(\partial q(0) \subset \partial f(0)\).

Jonathan Borwein (jborwein@ccm.sfu.ca)  
Department of Mathematics and Statistics  
Simon Fraser University  
Burnaby BC Canada V5A 1S6  
Research supported by NSERC and by the Shrum Chair Endowment at Simon Fraser University.

Simon Fitzpatrick (fitzpatr@maths.uwa.edu.au)  
Department of Mathematics  
The University of Western Australia  
Nedlands WA 6907 Australia.

Work on this paper was commenced while both authors were visiting the University of Newcastle, Australia and was completed while the second author was visiting Simon Fraser University. The authors wish to thank John Giles, Brailey Sims and Xianfu Wang for many useful discussions.


