MEAN VALUE INEQUALITIES

F. H. CLARKE AND YU. S. LEDYAEV

(Communicated by Andrew Bruckner)

ABSTRACT. We prove a new type of mean value theorem, one in which the functional differences are estimated in multiple directions simultaneously.

1. INTRODUCTION

The Mean Value Theorem of calculus in its inequality form asserts that if f is a smooth function on an interval $[x, y]$, then there is a point z in the interval such that

$$f'(z)(y - x) \geq f(y) - f(x).$$

The extension of this result to more than one dimension is straightforward; $[x, y]$ is then a line segment and we have the inequality

$$(1.1) \qquad \langle f'(z), y - x \rangle \geq f(y) - f(x).$$

However, a result such as this remains essentially one-dimensional, or unidirectional. Our purpose is to estimate the rate of growth in multiple directions simultaneously.

Let X and Y be compact subsets of $\mathbb{R}^n$ and f a smooth function on $\mathbb{R}^n$. Suppose for simplicity that f equals 1 on Y and 0 on X. Then for any $x \in X$, $y \in Y$, it follows from the inequality (1.1) cited above that for some point $z_{x,y}$ in the segment $[x, y]$ one has

$$(1.2) \qquad \langle f'(z_{x,y}), y - x \rangle \geq 1.$$

But we will show that when X and Y are convex, there is a *single* point z which fills the role of $z_{x,y}$ in (1.2) for *all* pairs (x, y); the point z lies in the convex hull of $X \cup Y$, which we denote $[X, Y]$.

We remark that the convexity of X and Y cannot simply be deleted in obtaining this conclusion. For example, if in $\mathbb{R}^n$ we take $X = \{0\}$ and $Y =$ the unit sphere, and if f is a function which vanishes at 0 and is equal to 1 on the sphere, the uniform inequality cited above becomes $\langle f'(z), y \rangle \geq 1$ for all unit vectors y, which is clearly impossible.

Theorem 2.1 of the following section does not require that f be constant on X and on Y. For example, when X and Y are compact it implies that for some z in the "interval" $[X, Y]$, one has

$$(1.3) \qquad \langle f'(z), y - x \rangle \geq \min_{Y} f - \max_{X} f \quad \text{for all } x \in X, y \in Y.$$

Further, X and Y need not be disjoint nor finite-dimensional. For example, let us set $X = Y = $ the closed unit ball B in a Banach space. By taking $y - x$ to be an appropriate negative multiple of $f'(z)$ in an extended form of (1.3), we obtain from Theorem 2.1 below the estimate

$$(1.4) \qquad \sup_{B} f - \inf_{B} f \geq 2 \inf_{B} |f'(z)|$$

for any continuously differentiable function f. We do not know whether (1.4) has been noted previously, but it serves to illustrate the deeper level of information contained in our theorem as compared to the classical unidirectional results.

The theorem is discussed and proved in the following sections, together with some variants. In a subsequent and more technical article we shall develop the applications in nonsmooth analysis and partial differential equations that were the original motivation for this work.

2. The Theorem

Throughout this article, X and Y are nonempty, closed convex bounded subsets of a Banach space E. We denote by $[X, Y]$ the "interval"

$$\{z \in E : z = \lambda x + (1 - \lambda)y \text{ for some } \lambda \in [0, 1], x \in X, y \in Y\}.$$

(This coincides with the convex hull of $X \cup Y$.) We are given a function f which is continuously Gâteaux differentiable on a neighborhood of $[X, Y]$.

Theorem 2.1. *Let $\epsilon > 0$ be given, and suppose that at least one of X and Y is compact. Then there exists a point z in $[X, Y]$ such that*

$$(2.1) \qquad \inf_{Y} f - \sup_{X} f < \langle f'(z), y - x \rangle + \epsilon \quad \forall x \in X, y \in Y$$

By switching the roles of X and Y in the theorem we obtain

Corollary 2.1. *There exists a point w in $[X, Y]$ such that*

$$(2.2) \qquad \sup_{Y} f - \inf_{X} f > \langle f'(w), y - x \rangle - \epsilon \quad \forall x \in X, y \in Y.$$

We now derive a result that is closer to the usual mean value theorem. To state it, we define two functions on E^*:

$$S(\zeta) := \max \{\langle \zeta, y - x \rangle : x \in X, y \in Y\},$$
$$s(\zeta) := \min \{\langle \zeta, y - x \rangle : x \in X, y \in Y\}.$$

Note that when X, Y are singletons $\{x\}, \{y\}$, then $s(\zeta) = S(\zeta) = \langle \zeta, x - y \rangle$. We nominate the following as "multidirectional mean value theorem".

Corollary 2.2. *Let both* X *and* Y *be compact, and let* g *be any continuous function on* E^* *satisfying*

$$s(\zeta) \leq g(\zeta) \leq S(\zeta) \text{ for all } \zeta \in E^*.$$

Then there exists a point $u \in [X, Y]$ *such that*

(2.3) $$g(f'(u)) \in f(Y) - f(X).$$

Proof. By an evident limiting procedure in which the ϵ of the theorem goes to zero, there is a point z such that

$$\min(f(Y) - f(X)) = \min_Y f - \max_X f \leq s(f'(z)) \leq g(f'(z)).$$

Similarly, we obtain a point w satisfying

$$g(f'(w)) \leq S(f'(w)) \leq \max_Y f - \min_X f = \max(f(Y) - f(X)).$$

It follows from the intermediate value theorem that some point u between z and w satisfies (2.3). □

Example 1. As an illustration of the "multidirectional equality" embodied in the corollary, consider a C^1 function f on the following subset Ω of $\mathbb{R}^2$:

$$\Omega := \{(x, y) : 0 \leq x \leq 1, 0 \leq y \leq 1\},$$

and suppose that $f(0, \cdot) \equiv 0$, $f(1, \cdot) \equiv 1$. Then there is a point $w = (u, v)$ in Ω such that

$$f'_x(u, v) = 1 + |f'_y(u, v)|.$$

This follows from Corollary 2.2 by taking

$$g(u, v) = s(u, v) = u - |v|,$$
$$X = \{(u, v) \in \Omega : u = 0\}, \qquad Y = \{(u, v) \in \Omega : u = 1\}.$$

When X and Y are singletons $\{x\}$, $\{y\}$, it is evident that (2.3) reduces to the classical equality

$$f(y) - f(x) = \langle f'(z), y - x \rangle.$$

In contrast to this classical unidirectional case, it is not always possible, however, to affirm (2.3) for some z lying in the relative interior of $[X, Y]$. We illustrate by a counterexample.

Example 2. In two dimensions, we define

$$f(u, v) = \begin{cases} u + v^2(1-v)^2/2 - (u-v)^2/2 & \text{if } u \leq v, \\ u + v^2(1-v)^2/2 - v^2(u-v)^2/2 & \text{if } u \geq v, \end{cases}$$

$$X = \{(0, 0)\}, \qquad Y = \{(1, v) : 0 \leq v \leq 1\}.$$

Then f is continuously differentiable, equal to 0 on X, and to 1 on Y. Theorem 2.1 implies that for some (z, w) in the (closed) triangle formed by taking the convex hull of $X \cup Y$, one has

(2.4) $$\langle f'(z, w), (1, v) \rangle \geq 1 \text{ for all } v \in [0, 1].$$

We shall see that (z, w) cannot lie in the relative interior of the triangle. For if (2.4) holds, then necessarily $f'_u(z, w) \geq 1$ (take $v = 0$). But when $w < z$, we have

$$f'_u(z, w) = 1 - w^2(z - w),$$

which is strictly less than 1 in the relative interior. (We can go on to show that the points (z, w) satisfying (2.4) are those on the segment $(0, 0)$ to $(1, 0)$, and the point $(1, 1)$.)

3. PROOF OF THE THEOREM

Let ϵ' belong to $(0, \epsilon)$, and let M denote the quantity

$$\sup_{z \in [X, Y]} \inf_{x \in X, y \in Y} \langle f'(z), y - x \rangle.$$

We shall establish the existence of a point $(\tilde{x}, \tilde{y})$ in $X \times Y$ such that

$$f(\tilde{y}) - f(\tilde{x}) \leq M + \epsilon',$$

which clearly implies (2.1). Note that M belongs to $(-\infty, \infty]$; we may limit ourselves to the case in which M is finite, for otherwise the existence of z satisfying (2.1) is trivial.

We define a multifunction Γ on $[X, Y]$ with values in $X \times Y$ as follows:

$$\Gamma(z) := \{(x, y) \in X \times Y : \langle f'(z), y - x \rangle \leq M + \epsilon'\},$$

where $f' : E \to E^*$ denotes the Gâteaux derivative.

Lemma 3.1. *There exists a locally Lipschitz function* $\gamma : E \to E$ *such that*

(3.1) $\qquad\qquad\qquad \gamma(z) \in \Gamma(z)$ *for all* $z \in [X, Y]$,

(3.2) $\qquad\qquad\qquad \gamma(z) \in X \times Y$ *for all* $z \in E$.

Proof. We begin by constructing an open covering of E. Let z belong to $[X, Y]$. By definition of M, there exists a point $(\overline{x}_z, \overline{y}_z)$ in $X \times Y$ such that

$$\langle f'(z), \overline{y}_z - \overline{x}_z \rangle \leq M + \epsilon'/2.$$

Since f' is continuous at z, there is an open neighborhood $U(z)$ of z such that

(3.3) $\qquad\qquad \langle f'(z'), \overline{y}_z - \overline{x}_z \rangle \leq M + \epsilon' \quad \forall z' \in U(z).$

Fix any point $(\overline{x}, \overline{y}) \in X \times Y$. For $z \notin [X, Y]$ (which is a closed set), let $U(z)$ be an open neighborhood of z such that $U(z) \cap [X, Y] = \varnothing$, and define $(\overline{x}_z, \overline{y}_z) = (\overline{x}, \overline{y})$.

We have an open covering $\{U(z)\}_{z \in E}$ of E. Since a Banach space is paracompact, there is a locally finite refinement $\{V_\alpha\}_{\alpha \in A}$ of this covering (see for example [3]). Thus for each α there exists a point z_α such that V_α is contained in $U(z_\alpha)$. Also, for each z, the set of α for which $z \in V_\alpha$ is finite. We now use a classical construction to define a certain partition of unity.

Let $p_\alpha(z)$ denote the distance from z to $E \backslash V_\alpha$; then $p_\alpha(z) = 0$ iff $z \notin V_\alpha$. We define

$$\varphi_\alpha(z) := p_\alpha(z) \bigg/ \left(\sum_{\alpha' \in A} p_{\alpha'}(z) \right)$$

(note that for each z, only finitely many terms in the sum are nonzero and that the sum is positive). The functions $z \to p_\alpha(z)$ are globally Lipschitz with Lipschitz constant 1; it follows that each φ_α is locally Lipschitz. We have that

$$(3.4) \qquad \sum_{\alpha \in A} \varphi_\alpha(z) = 1 \quad \forall z \in E.$$

Thus $\{\varphi_\alpha(\cdot)\}_A$ is a locally Lipschitz partition of unity.

We proceed to construct a function γ as follows:

$$\gamma(z) := \sum_{\alpha \in A} \left(\overline{x}_{z_\alpha}, \overline{y}_{z_\alpha} \right) \varphi_\alpha(z)$$

where the function $z \to (\overline{x}_z, \overline{y}_z)$ was defined earlier.

Again, only finitely many terms in this sum are nonzero, and it follows that γ is locally Lipschitz. From (3.4), because X and Y are convex, it follows that γ has values in $X \times Y$, which establishes (3.2). We need only establish (3.1).

Let $z \in [X, Y]$. Let α be any index in A for which $z \in V_\alpha$. Since V_α is a subset of $U(z_\alpha)$, from (3.3) we deduce

$$\langle f'(z), \overline{y}_{z_\alpha} - \overline{x}_{z_\alpha} \rangle \leq M + \epsilon',$$

which gives

$$(\overline{x}_{z_\alpha}, \overline{y}_{z_\alpha}) \in \Gamma(z).$$

But $\gamma(z)$ is a finite convex combination of such points $(\overline{x}_{z_\alpha}, \overline{y}_{z_\alpha})$, and $\Gamma(z)$ is convex. This establishes (3.1) and completes the proof of the lemma. $\quad \square$

We now suppose that X is compact; the case in which Y is compact is handled by simply switching the roles of X and Y in what follows. We pause to prove an apparently novel "geometric controllability" result that may be of independent interest.

Lemma 3.2. *Let u and v be locally Lipschitz mappings from E to X and Y respectively, where X is compact. Then there exists a solution z on $[0, 1]$ of the differential equation*

$$(3.5) \qquad \dot{z}(t) = v(z(t)) - u(z(t))$$

such that for each $t \in [0, 1]$,

$$z(t) \in t\,Y + (1 - t)\,X.$$

Proof. Since u and v are bounded, it is well known [2] that for any x, there is a unique solution z defined on $[0, \infty)$ of (3.5) with initial condition $z(0) = x$, in the sense that there is a Bochner integrable function $\dot{z}$ for which (3.5) holds, and such that for all $t \geq 0$,

$$z(t) = x + \int_0^t \dot{z}(s)\,ds.$$

Further, the solution $z(t, x)$ depends continuously upon $t \in [0, 1]$ and x.

Consider now the function F on X defined by

$$F(x) = \int_0^1 u(z(s, x))\, ds.$$

Then F is a continuous map from X to X, so by the Schauder fixed point theorem [4, Theorem 2.3.7], there is a point $\overline{x}$ in X such that $F(\overline{x}) = \overline{x}$. Let $\overline{z}$ be the solution of (3.5) satisfying $\overline{z}(0) = \overline{x}$. Then

$$\overline{z}(t) = \overline{x} + \int_0^t v(\overline{z}(s))\, ds - \int_0^t u(\overline{z}(s))\, ds$$

$$= \int_0^t v(\overline{z}(s))\, ds + \int_t^1 u(\overline{z}(s))\, ds$$

$$\in t\, Y + (1 - t)\, X,$$

completing the proof of the lemma. □

We proceed now to apply Lemma 3.2 to the function $\gamma = (u, v)$ produced by Lemma 3.1. The function $z(t)$ has values in $[X, Y]$, so $\gamma(z(t))$ belongs to $\Gamma(z(t))$, which gives

$$\langle f'(z(t)), \dot{z}(t)\rangle = \langle f'(z(t)), v(z(t)) - u(z(t))\rangle \le M + \epsilon'.$$

The left side of this inequality is the derivative of the function $t \to f(z(t))$, which, as a locally Lipschitz function, is absolutely continuous. Integrating the inequality over $[0, 1]$ therefore yields

$$f(\tilde{y}) - f(\tilde{x}) \le M + \epsilon',$$

where $\tilde{y} := z(1)$, $\tilde{x} := z(0)$. This implies the existence of a point z satisfying (2.1). □

Remark 3.1. An analysis of the proof shows that the only requirement on f is the following: on a neighborhood of $[X, Y]$, f is defined and Gâteaux differentiable; the Gâteaux derivative f' is continuous at each point of $[X, Y]$.

Remark 3.2. The proof technique applies to vector-valued functions. As an illustration, suppose that f has values in $\mathbb{R}^m$, and let $\le$ denote the usual component-wise partial ordering in that space. Let r be any vector in $\mathbb{R}^m$ satisfying

$$r \le f(Y) - f(X)$$

(i.e., $r \le f(y) - f(x)$ for all $(x, y) \in X \times Y$), and let Δ be a vector such that for any $z \in [X, Y]$, for some $(x, y) \in X \times Y$ (depending on z), one has

$$\langle f'(z), y - x\rangle < \Delta$$

(i.e., strict inequality in each component). We deduce $r < \Delta$, by an obvious adaptation of the proof. This rather indirect statement is equivalent to that of the theorem when $m = 1$.

4. Two variants of the Theorem

We now present two extensions of the theorem which weaken in turn the smoothness and convexity requirements of Theorem 2.1. For the first, we recall two notions from nonsmooth analysis [1].

Let $f : E \to \mathbb{R}$ be a locally Lipschitz function defined on a Banach space E. The generalized directional derivative at x in direction v is given by

$$f^0(x\,;\,v) := \limsup_{\substack{x' \to x \\ \lambda \downarrow 0}} \frac{f(x' + \lambda v) - f(x')}{\lambda},$$

and the generalized gradient (or subdifferential) of f at x is the nonempty set defined by

$$\partial f(x) := \left\{ \zeta \in E^* : f^0(x\,;\,v) \geq \langle v, \zeta \rangle \ \forall v \in E \right\}.$$

Then $\partial f(x)$ reduces to $\{f'(x)\}$ if f is strictly differentiable (or continuously differentiable) at x [1, Proposition 2.2.4].

We assume now that X and Y are as in Theorem 2.1, but we weaken the regularity of f to the following: f is locally Lipschitz on $[X, Y]$; i.e., there corresponds to each z in $[X, Y]$ an open neighborhood of z in E in which f satisfies a Lipschitz condition.

Theorem 4.1. *Under these hypotheses, for any $\epsilon > 0$, there exist a point z in $[X, Y]$ and an element ζ of $\partial f(z)$ such that*

$$\inf_Y f - \sup_X f < \langle \zeta, y - x \rangle + \epsilon \quad \forall x \in X,\ y \in Y.$$

Proof. We merely indicate how to adapt the proof of Theorem 2.1 to the present setting. The multifunction Γ is now defined as follows:

$$\Gamma(z) := \{(x, y) \in X \times Y : \langle \zeta, y - x \rangle \leq M + \epsilon' \quad \forall \zeta \in \partial f(z)\}$$
$$= \left\{(x, y) \in X \times Y : f^0(z\,;\,y - x) \leq M + \epsilon'\right\},$$

where M is defined as before but with ∂f for f'.

Because $f^0(z\,;\,v)$ is upper semicontinuous as a function of z [1, Proposition 2.1.1], the proof of Lemma 3.1 requires only trivial modification; Lemma 3.2 requires none at all. When the final step of the proof is reached, we have a function $z(t)$ satisfying

$$\langle \zeta, \dot{z}(t) \rangle \leq M + \epsilon' \quad \forall \zeta \in \partial f(z(t)).$$

The function $t \to f(z(t))$ is Lipschitz (and hence absolutely continuous), and its derivative at t, when it exists, lies in the set

$$\{\langle \zeta, \dot{z}(t) \rangle : \zeta \in \partial f(z(t))\},$$

as follows readily from a direct argument, or from a chain rule of nonsmooth calculus [1, Theorem 2.3.10]. The preceding inequality therefore yields

$$f(z(1)) - f(z(0)) \leq M + \epsilon',$$

which, as before, completes the proof. $\square$

The analogous corollaries to those of Section 2 can be derived in the nonsmooth setting above; we limit ourselves here to stating just one.

Corollary 4.1. *In addition to the hypotheses of the theorem, let both X and Y be compact. Then there exist z in $[X, Y]$ and $\zeta \in \partial f(z)$ such that*

$$\min_Y f - \max_X f \leq \langle \zeta, y - x \rangle \quad \forall x \in X, \, y \in Y.$$

The second extension of the theorem involves a mapping $\varphi : E \to E$ which admits, at each point z in a neighborhood of $[X, Y]$, a Fréchet derivative $D\varphi(z)$ (thus, a continuous linear functional from E to itself). We assume that at each such z, the map $z' \to D\varphi(z')$ is continuous for the natural induced norm topologies (i.e., $\|D\varphi(z)\| := \sup_{\|u\| \leq 1} \|D\varphi(z)(u)\|$).

The hypotheses on X, Y are those of Theorem 2.1 and we assume that f is continuously differentiable on a neighborhood of $\varphi([X, Y])$. ($*$ denotes the adjoint.)

Theorem 4.2. *For any $\epsilon > 0$, there exists a point z in $[X, Y]$ such that*

$$\inf_{\varphi(Y)} f - \sup_{\varphi(X)} f \leq \langle D\varphi(z)^* f'(\varphi(z)), y - x \rangle + \epsilon \quad \forall x \in X, \, y \in Y.$$

Proof. We adapt the proof of Theorem 2.1 given in Section 3 to this extended setting. The definition of the multifunction Γ is modified as follows:

$$\Gamma(z) := \{(x, y) \in X \times Y : \langle D\varphi(z)^* f'(\varphi(z)), y - x \rangle \leq M + \epsilon'\},$$

where M now signifies the quantity

$$\sup_{z \in [X, Y]} \inf_{x \in X, y \in Y} \langle D\varphi(z)^* f'(\varphi(z)), y - x \rangle,$$

which as before can be supposed finite.

The proof then proceeds along the same lines as before. In the final phase, we arrive at a smooth function $z(t)$ such that

$$\langle D\varphi(z(t))^* f'(\varphi(z(t))), \dot{z}(t) \rangle \leq M + \epsilon'.$$

Integration and the chain rule give

$$f(\varphi(z(1))) - f(\varphi(z(0))) \leq M + \epsilon',$$

which yields the conclusion of Theorem 4.2. $\square$

Remark. The proof technique adapts easily to the case in which $\varphi(t, z)$ depends on the real parameter t. In this nonautonomous setting, the conclusion of the theorem becomes

$$\inf_{\varphi(1, Y)} f - \sup_{\varphi(0, X)} f \leq \langle f'(\varphi(t, z)), \varphi_t'(t, z) \rangle$$

$$+ \langle D\varphi(t, z)^* f'(\varphi(t, z)), y - x \rangle + \epsilon \quad \forall x \in X, \, y \in Y,$$

where φ_t' denotes the t-derivative. We omit the details.

It is possible to combine simultaneously the two extensions of this section; here is an example whose proof is omitted.

Corollary 4.2. *Let X and Y be compact, let φ be as above, and suppose f is locally Lipschitz in a neighborhood of $\varphi([X, Y])$. Then there exists z in $[X, Y]$ and $\zeta \in \partial f(\varphi(z))$ such that*

$$\min_{\varphi(Y)} f - \max_{\varphi(X)} f \leq \langle D\varphi(z)^* \zeta, y - x \rangle \quad \forall x \in X, \, y \in Y.$$

We obtain the next result by applying Corollary 4.1 to the (nonsmooth, globally Lipschitz) function $f(z) := d_{\varphi(X)}(z)$ (distance to the set $\varphi(X)$) and upon noting that any element of ∂f has norm at most 1.

Corollary 4.3. *Let* X *and* Y *be compact, let* φ *be as above, and suppose the Hausdorff distance between* $\varphi(X)$ *and* $\varphi(Y)$ *equals* $\Delta > 0$. *Then there exist a point* z *in* $[X, Y]$ *and an element* ζ *of* E^* *of norm at most* 1 *such that*

$$\Delta \le \langle D\varphi(z)^* \zeta, y - x \rangle \quad \forall x \in X, y \in Y.$$

The above can be viewed as a "linearized separation theorem" for the generally nonconvex sets $\varphi(X), \varphi(Y)$.

REFERENCES

1. F. H. Clarke, *Optimization and nonsmooth analysis*, Classics Appl. Math., vol. 5, SIAM, Philadelphia, 1990 (originally published by Wiley, New York, 1983).
2. K. Deimling, *Ordinary differential equations in Banach spaces*, Lecture Notes in Math., vol. 596, Springer-Verlag, New York, 1977.
3. J. Dugundji, *Topology*, Allyn and Bacon, Boston, 1966.
4. D. R. Smart, *Fixed point theorems*, Cambridge Univ. Press, Cambridge, 1974.

CENTRE DE RECHERCHES MATHÉMATIQUES, UNIVERSITÉ DE MONTRÉAL, C.P. 6128–A, MONTRÉAL, QUÉBEC, CANADA H3C 3J7

STEKLOV MATHEMATICAL INSTITUTE, VAVILOV STR. 42, MOSCOW 117966, RUSSIA