Mean value theorems and convexity: an example of cross-fertilization of two mathematical items

Jean-Baptiste Hiriart-Urruty

Abstract. With the help of two types of results, one for real-valued functions, the other one for vector-valued functions, we show how the classical mean value theorems (in an equality form) and the concept of convexity (for functions and for sets) are closely related.

Keywords. Mean value theorems, convex or concave functions, convex hull of a set.

Mathematics Subject Classification. 26A, 52A.

1. Introduction

The topic of mean value theorems for (real-valued or vector-valued) functions has been and still is one of my favorite ones in mathematics. During my career, I have written a lot on the subject : mean value theorems for convex or locally Lipschitz functions, witness the papers [3, 4] ; variants of the classical mean value theorems, like that of Cauchy, Pompeiu, Flett, etc. (see the first exercises in [8] for example).

As far as I remember, my first encounter with a mean value theorem goes back to my high school period. I remember a calculation integrated in the lesson itself : the first step was to prove Rolle's theorem, followed by the classical mean value theorem (also called Lagrange's theorem): For any $a < b$ in $\mathbb{R}$, there *exists* c in the open interval (a, b) such that

$$\frac{f(b) - f(a)}{b - a} = f'(c); \tag{1}$$

immediately followed the determination of such c for quadratic functions $f : x \mapsto f(x) = \alpha x^2 + \beta x + \gamma$, with $\alpha \neq 0$. It happens that finding out such c for quadratic functions is an easy calculation : a unique c pops up, it is $c = \frac{a+b}{2}$. One must confess that the result is somehow surprising for a beginner : for a, b close to 0 or not, for a, b far apart or not, the answer for c is always the midpoint of a and b. For a mathematician, a natural question which then arises is: what about the converse? In other words,

Q_1 : What are the functions for which the c in the mean value result (1) is always $\frac{a+b}{2}$?

A question akin to the one above is as follows. Consider $p > 0$ and $q > 0$ such that $p + q = 1$. We generalize (Q_1) with

Q_2 : What are the functions for which the (unique) c in the mean value result (1) is always $pa + qb$?

The above recalled LAGRANGE's mean value theorem is an *existence* result, it does not mention uniqueness or not of c. So, it is natural to ask the question

Q_3 : What are the functions for which the c in the mean value result (1) is *unique* for all a, b?

Answers to these three questions are more or less known, they are part of folklore in Calculus; we recall and prove them in the next section; we provide an original proof of the answer to the question (Q_3).

The main result in the first part of the present paper aims at identifying the functions for which the set of c satisfying (1) is always an *interval* (whatever a and b are); the broached question, generalizing (Q_3) therefore is

Q_4 : What are the functions for which the set of c satisfying the mean value result (1) is *an interval* for all a, b?

To the best of our knowledge, the result (Theorem 3 below) is new.

The second part of the paper deals with vector-valued functions $X : I \to \mathbb{R}^n$. Mean value theorems for such functions are usually derived in *inequality* forms, some authors like J. DIEUDONNÉ even claimed that they are the only possible[1]. This not true. We present a simple result, with its proof, showing how the mean value $\frac{X(b)-X(a)}{b-a}$ could be expressed as a *convex combination* of some values $X'(t_i)$ of the derivative of X at intermediate points $t_i \in (a, b)$. This result is not new, apparently not well-known, especially as no integral of any kind is called, only values of derivatives X' at points are used. Moreover, the kinematics interpretation of the result is very expressive.

2. The case of real-valued functions

Let $f : I \to \mathbb{R}$ be a differentiable function on the open interval I. There is no loss of generality in assuming that I is the whole of $\mathbb{R}$, which we do henceforth. For $a < b$ in $\mathbb{R}$, let $C_{a,b}$ denote the set of $c \in (a, b)$ for which $\frac{f(b)-f(a)}{b-a} = f'(c)$. The basic mean value theorem tells us that $C_{a,b}$ is nonempty for all a and b. In the next subsections, we intend to *characterize* functions f for which $C_{a,b}$ is the same fixed intermediate point between a and b, or always reduces to a single point between a and b, or always is an interval for all a, b.

[1] "The classical mean value theorem (for real-valued functions) is usually written as an *equality* $f(b)-f(a) = f'(c)(b-a)$. The trouble with that classical formulation is that there is nothing similar to it as soon as f has vector values... ". In J. Dieudonné, **Foundations of Modern Analysis,** Academic Press (1960), Section VIII.

2.1 Case where $C_{a,b}$ is the same fixed intermediate point between a and b

Theorem 1. *Let $p > 0$ and $q > 0$ such that $p + q = 1$. Suppose that $C_{a,b} = \{pa + qb\}$ for all a and b. Then :*

(i) If $p = \frac{1}{2}$, the function is necessarily quadratic, that is to say $f : x \mapsto f(x) = \alpha x^2 + \beta x + \gamma$, with $\alpha \neq 0$.

(ii) If $p \neq \frac{1}{2}$, there is no function f with the required property on $C_{a,b}$.

Proof. Written in another form, the assumption made on f writes: There exists $p \in (0, 1)$ such that

$$f(x + h) = f(x) + hf'(x + qh) \text{ for all } x \text{ and } h \text{ in } \mathbb{R}. \tag{2}$$

First point. Due to the functional relationship (2), it is easy to derive that f is twice differentiable, even of class C^∞.

Second point. We differentiate the relationship (2) with respect to h, so that we get at:

$$f'(x + h) = f'(x + qh) + hqf''(x + qh) \text{ for all } x \text{ and } h \text{ in } \mathbb{R}. \tag{3}$$

We therefore have: For all x and $h \neq 0$ in $\mathbb{R}$,

$$\begin{aligned} qf''(x + qh) &= \frac{f'(x + h) - f'(x + qh)}{h} \\ &= \frac{f'(x + h) - f'(x)}{h} - q\frac{f'(x + qh) - f'(x)}{qh}. \end{aligned}$$

Passing to the limit $h \to 0$, since f'' is continuous, we get:

$$\begin{aligned} qf''(x) &= f''(x) - qf''(x) \\ &\text{or} \\ (1 - 2q)f''(x) &= 0 \text{ for all } x \text{ in } \mathbb{R}.4 \end{aligned} \tag{1}$$

We here examine two situations.

Situation (ii): q (or, equivalently, p) is different from $\frac{1}{2}$. Then it comes from (4) that $f''(x) = 0$ for all x in $\mathbb{R}$. Consequently, f is affine,

$$f(x) = \beta x + \gamma \text{ for all } x \text{ in } \mathbb{R}.$$

But, in that case, we would have $C_{a,b} = (a, b)$ for all a and b, which contradicts the assumption made on $C_{a,b}$.

Situation (i): q (or, equivalently, p) equals $\frac{1}{2}$. In such a case, (3) rewrites as:

$$f'(x+h) = f'(x+\frac{h}{2}) + \frac{h}{2}f''(x+\frac{h}{2}) \text{ for all } x \text{ and } h \text{ in } \mathbb{R}. \tag{5}$$

Changing into the new variables $u = x + \frac{h}{2}, r = \frac{h}{2}$, we get from (5):

$$f'(u+r) = f'(u) + rf''(u) \text{ for all } u \text{ and } r \text{ in } \mathbb{R}. \tag{6}$$

We take the derivative with respect to the variable r in (6), so that:

$$f''(u+r) = f''(u) \text{ for all } u \text{ in } \mathbb{R}.$$

Consequently, f'' is constant on $\mathbb{R}$, therefore f is a quadratic function. Here again, since $C_{a,b}$ is assumed to reduce to one point $c = \frac{a+b}{2}$, affine functions are excluded. $\boxdot$

Remarks. We indeed have proved a little more than what is stated in Theorem 1, namely:

"$\frac{a+b}{2} \in C_{a,b}$ for all a, b" happens only in two cases:

- for affine functions, in which case $C_{a,b} = (a, b)$ for all a, b;
- for quadratic functions, in which case $C_{a,b} = \left\{\frac{a+b}{2}\right\}$ for all a, b.

Given $p > 0$, $p \neq \frac{1}{2}$, and $q > 0$ such that $p + q = 1$, "$pa + qb \in C_{a,b}$ for all a, b" happens only in one specific situation:

- for affine functions, in which case $C_{a,b} = (a, b)$ for all a, b.

2.2 Case where $C_{a,b}$ is a singleton for all a and b

We consider in this subsection the case where $C_{a,b}$ is a singleton for all a and b, *i.e.*, $C_{a,b} = \{c_{a,b}\}$ for all a, b. It clearly covers the case of quadratic functions seen in the previous subsection ($c_{a,b} = \frac{a+b}{2}$ for all a, b). However, in the considered present case, $c_{a,b}$ is not "rigidified" via a formula, but varies with a, b. The answer to the question "What are the functions for which the c in the mean value result is *unique* for all a, b?" is known ; it consists of strictly convex functions or strictly concave functions ; this is even a characterization of such functions. The result is mentioned as early as in BOURBAKI's text (1958, [1, page 54]), where it is proposed as an exercise (without proof). One proof that we know, at the first year of Calculus level, consists in proving that the derivative f' is monotone (either increasing or decreasing). For that, knowing that "a derivative function does not create any hole", *i. e.,* DARBOUX' theorem stating that the image of an interval by f' is again an interval, helps a lot. Other proofs start by contradiction : "Suppose that f is not convex and f is not concave", or "Suppose that f' is not increasing and f' is not decreasing", but the sequel of reasonings is "laborious" or "tortuous" (several cases and subcases to treat) and even flawed. We propose below an alternate proof, based on an argument from a more advanced level in Analysis.

Theorem 2. *The following statements are equivalent:*
(i) $C_{a,b}$ is a singleton for all a, b.
(ii) There are not three (distinct) points aligned on the graph of f.
(iii) Either f is strictly convex or f is strictly concave.

Proof. [*(i)* $\Rightarrow$ *(ii)*]. Suppose that there are three points aligned on the graph of f, namely $(x_1, f(x_1)), (x_2, f(x_2)), (x_3, f(x_3))$, with $x_1 < x_2 < x_3$. Acording to the mean value theorem, applied on the line-segments $[x_1, x_2]$ and $[x_2, x_3]$, there exists $c_1 \in (x_1, x_2)$ and $c_2 \in (x_2, x_3)$ such that

$$\frac{f(x_2) - f(x_1)}{x_2 - x_1} = f'(c_1) \text{ and } \frac{f(x_3) - f(x_2)}{x_3 - x_2} = f'(c_2). \tag{7}$$

But, since the three points $(x_1, f(x_1)), (x_2, f(x_2)), (x_3, f(x_3))$ are on the same line,

$$\frac{f(x_3) - f(x_1)}{x_3 - x_1} = \frac{f(x_3) - f(x_2)}{x_3 - x_2} = \frac{f(x_2) - f(x_1)}{x_2 - x_1}.$$

In view of (7), that would induce that there are two different points c_1 and c_2 such that

$$\frac{f(x_3) - f(x_1)}{x_3 - x_1} = f'(c_1) = f'(c_2).$$

That would mean that C_{x_1,x_3} contains at least two points, therefore contradicting the assumption *(i)*.

[*(ii)* $\Rightarrow$ *(iii)*]. Consider the following open set in $\mathbb{R}^3$

$$\Omega = (0, 1) \times \left\{(x, y) \in \mathbb{R}^2 \text{ with } x < y\right\}$$

and the next function

$$F : (\lambda, x, y) \in \Omega \mapsto F(\lambda, x, y) = f\left[\lambda x + (1 - \lambda)\, y\right] - \left[\lambda f(x) + (1 - \lambda)\, f(y)\right].$$

Clearly, F measures the "default of convexity" or "default of concavity" of the function f. Here are two clear properties: Ω is an open connected (even convex) set; F is a continuous function. Hence, the image $F(\Omega)$ is a connected set, that is to say an interval of $\mathbb{R}$. But, according to the assumption *(ii)*, $F(\Omega)$ does not contain 0. We therefore have only two possibilities:

- either $F(\Omega) \subset (-\infty, 0)$, which amounts to having f strictly convex,

- or $F(\Omega) \subset (0, +\infty)$, which amounts to having f strictly concave.

$[(iii) \Rightarrow (i)]$. Under the assumption *(iii)*, the derivative f' is strictly increasing or decreasing; whence the equation $\frac{f(b)-f(a)}{b-a} = f'(c)$ has only one solution c for any a, b. $\boxdot$

2.3 Case where $C_{a,b}$ is an interval for all a and b

This subsection contains the newest part of the Section 2. The theorem that we present below generalizes Theorem 2, although the way of proving it is not the same.

Theorem 3. *The following statements are equivalent:*
(i) $C_{a,b}$ is an interval for all a, b.
(ii) The level-sets of f' are intervals.
(iii) f' is monotone (either increasing or decreasing).
(iv) Either f is convex or f is concave.

The proof combines the contributions of the author and that proposed in [10] as an answer to a posed question by the author. As usual in dealing with a derivative function, DARBOUX' theorem stating that the image of an interval by f' is again an interval is instrumental.

The equivalence $[(ii) \Leftrightarrow (iv)]$ makes echo to another property of f': following ROWE's theorem (1926) (see [6]), f' is continuous if and only if its level-sets are closed.

In the proof of Theorem 3, appears, in a hidden form, an equivalence with a further statement,

(v) If there are three (distinct) points aligned on the graph of f, for example $(x_1, f(x_1)), (x_2, f(x_2)), (x_3, f(x_3))$ with $x_1 < x_2 < x_3$, then f is affine on $[x_1, x_3]$,

in the same vein as *(ii)* in Theorem 2.

Proof of Theorem 3. $[(i) \Rightarrow (ii)]$. Let $u < v$ in the level-set $\Gamma_r = \{x : f'(x) = r\}$; we intend to prove that the whole line-segment $[u, v]$ is contained in Γ_r, that is to say f' is constant on $[u, v]$.

Like in the proof of the mean value theorem from ROLLE's one, we call the auxiliary function

$$g : x \mapsto g(x) = f(x) - \frac{f(v) - f(u)}{v - u} x.$$

We have $g(u) = g(v)$, as also $g'(u) = g'(v)$ (since $f'(u) = f'(v)$). Changing g into $-g$ if necessary, we can assume that $g'(u) \geqslant 0$. Since g is a "tilted" version of f by a linear function, it is clear that property *(ii)* in Theorem 3 transfers to g: the level-sets of g' are intervals.

Let $J = \{x : g'(x) = 0\}$. It is nonempty according to ROLLE's theorem, and an interval since it is the level-set of g' at level 0. Define $\underline{u} = \inf J$ and $\overline{v} = \sup J$. By construction, g' does not vanish neither on $[u, \underline{u})$ nor on $(\overline{v}, v]$. DARBOUX' property of the derivative g' allows us to deduce that g' is of a constant sign on $[u, \underline{u})$ and on $(\overline{v}, v]$. But $g'(u) = g'(v) \geqslant 0$, hence g is increasing on $[u, \underline{u})$ and on $(\overline{v}, v]$. Consequently, g is constant on $[u, v]$, $g' = 0$ on $[u, v]$, so that f' is constant on $[u, v]$.

$[(ii) \Rightarrow (iii)]$. The reasoning is by contradiction. Changing f' into $-f'$ if necessary, one may suppose that there exists $x_1 < x_2 < x_3$ such that $f'(x_1) < f'(x_2)$ and $f'(x_2) > f'(x_3)$. Let us choose a real value α such that:

$$\begin{aligned} f'(x_1) &< \alpha < f'(x_2); \\ f'(x_2) &> \alpha > f'(x_3). \end{aligned}$$

Since f' transforms intervals into intervals (DARBOUX' property), there exists y_1 in (x_1, x_2) and y_2 in (x_2, x_3) such that:

$$f'(y_1) = f'(y_2) = \alpha.$$

Then, because the level-sets of f' are intervals (assumption (ii)), we have:

$$f'(x) = \alpha \text{ for all } x \in [y_1, y_2].$$

Consequently, for the "intermediate" point x_2, we get $f'(x_2) = \alpha$. We therefore have gotten at a contradiction since $f'(x_2) > \alpha$.

$[(iii) \Rightarrow (iv)]$ is classical.

$[(iv) \Rightarrow (i)]$ or even $[(iv) \Rightarrow (iii) \Rightarrow (ii) \Rightarrow (i)]$ does not offer any difficulty. $\boxdot$

To summarize this Section 2, here is a tableau offering *characterizations* of classes of convex (or concave) functions via the properties of the sets $C_{a,b}$.

$$\begin{array}{ccc}
\left[C_{a,b} = \left\{\frac{a+b}{2}\right\} \text{ for all } a \text{ and } b\right] & \Leftrightarrow & [\text{quadratic functions}] \\[2ex]
\left[\begin{array}{c} \text{For some } \frac{1}{2} \neq p > 0 \text{ and } q > 0 \\ \text{such that } p + q = 1, \\ pa + qb \in C_{a,b} \text{ for all } a \text{ and } b \end{array}\right] & \Leftrightarrow & [\text{affine functions}] \\
[C_{a,b} \text{ is a singleton for all } a \text{ and } b] & \Leftrightarrow & \left[\begin{array}{c} \text{strictly convex or} \\ \text{strictly concave functions} \end{array}\right] \\
[C_{a,b} \text{ is an interval for all } a \text{ and } b] & \Leftrightarrow & [\text{convex or concave functions}].
\end{array}$$

3. The case of vector-valued functions

In this section, we consider vector-valued differentiable functions $X : I \to \mathbb{R}^n$; here again there is no loss of generality in assuming that $I = \mathbb{R}$. The first thing we learn concerning mean value theorems for such functions X is that a result like "$\frac{X(b)-X(a)}{b-a} = X'(c)$ for some $c \in (a,b)$" is hopeless... Usually, a mean value theorem for vector-valued X is presented in the following form ([2]):

$$\left\| \frac{X(b)-X(a)}{b-a} \right\| \leqslant \sup_{c\in(a,b)} \|X'(c)\|, \tag{8}$$

a way of bounding from above the norm of the mean value $\frac{X(b)-X(a)}{b-a}$ which can be fairly weak. However, it is indeed possible to exactly express $\frac{X(b)-X(a)}{b-a}$ in terms of convex combinations of derivatives $X'(c_i)$, $c_i \in (a,b)$. Even when the image space of X are normed vector spaces, there are powerful results on the subject, by McLeod ([9]) for example. Here, with X taking values in the finite-dimensional space $\mathbb{R}^n$, we give a short proof of a mean value theorem for X, in an *equality* form. In doing that, convexity enters into the picture very naturally. Our arguments are merely based on simple techniques from convex analysis.

Given a nonempty set $S \subset \mathbb{R}^n$, we denote by:

coS (resp. $\overline{co}S$) the convex hull of S (resp. the closed convex hull of S);
$d \in \mathbb{R}^n \mapsto \sigma_S(d) = \sup_{s\in S} \langle s, d\rangle$ the support function of S[2].

When C is convex, there is a unique smallest affine set containing C; this set is called the affine hull of C. The relative interior of C, which is denoted by riC, is defined as the interior which results when C is regarded as a subset of its affine hull. The properties of support functions, of relative interiors of convex sets, are expounded in Rockafellar's book ([11]) or in our texbook ([7]).

Theorem 4. *Let $a < b$. Then:*

$$\frac{X(b)-X(a)}{b-a} \in ri\ co\{X'(c) : c \in (a,b)\}. \tag{9}$$

As a consequence, there are $n+1$ elements $c_i \in (a,b)$, $n+1$ nonnegative coefficients λ_i summing up to 1, such that:

$$\frac{X(b)-X(a)}{b-a} = \sum_{i=1}^{n+1} \lambda_i X'(c_i). \tag{10}$$

[2] $\langle . \rangle$ denotes the usual inner product in $\mathbb{R}^n$.

Proof. It is divided into three steps.

Step1. We intend to prove that the mean value $\frac{X(b)-X(a)}{b-a}$ belongs to the closed convex hull of the image-set $\{X'(c) : c \in (a,b)\}$.

For $d \in \mathbb{R}^n$, let f_d be the "scalarized version of X in the d direction", that is

$$f_d : t \in \mathbb{R} \mapsto f_d(x) = \langle X(t), d \rangle .$$

The function f_d is real-valued and, according to the classical mean value theorem for such functions, there exists $c_d \in (a,b)$ such that

$$\left\langle \frac{X(b)-X(a)}{b-a}, d \right\rangle = \langle X'(c_d), d \rangle .$$

The difficulty comes here from the fact that the intermediate point c_d depends on the d direction. Never mind. We deduce from above

$$\left\langle \frac{X(b)-X(a)}{b-a}, d \right\rangle \leqslant \sup_{c \in (a,b)} \langle X'(c), d \rangle . \tag{11}$$

We recognize in the right-hand side of (11) the support function of the image-set $\{X'(c) : c \in (a,b)\}$ or, which amounts to the same, of its closed convex hull $\overline{co}\,\{X'(c) : c \in (a,b)\}$. The left-hand side in (11) is a linear form, "directed" by the vector $\frac{X(b)-X(a)}{b-a}$. A consequence on sets of the inequality (11) on support functions is that ([7, Theorem 2.2.2 in p. 137])

$$\frac{X(b)-X(a)}{b-a} \in \overline{co}\,\{X'(c) : c \in (a,b)\} . \tag{12}$$

Step 2. For the sake of simplicity, we denote:

$$\frac{X(b)-X(a)}{b-a} = X_m \ ; \ co\,\{X'(c) : c \in (a,b)\} = C.$$

By the result of Step 1, X_m belongs to $\overline{C}$. We intend to prove here that X_m cannot be on the relative boundary $rbd\ C = \overline{C} \setminus ri\ C$ of C.

Suppose that $X_m \in rbd\ C$. There then exists a non-trivial supporting hyperplane to $\overline{C}$ at X_m, that is one which does not contain C itself ([11, p. 100]). Written with the help of support functions σ_C of C (see [11, Section 13] or [7, Section 4 of Chapter A]), there exists $d \in \mathbb{R}^n$ such that:

$$\begin{aligned} \langle x, d \rangle &\leqslant \sigma_C(d) \text{ for all } x \in \overline{C}; 13 && (2) \\ \langle X_m, d \rangle &= \sigma_C(d); 14 && (3) \\ \langle x, d \rangle &< \sigma_C(d) \text{ for all } x \in ri\ C.15 && (4) \end{aligned}$$

Let now $g_d : \mathbb{R} \to \mathbb{R}$ be defined by

$$g_d(t) = \langle X(t), d \rangle - \langle X_m, d \rangle \, t.$$

According to (13) and (14), remembering that $\{X'(t) : t \in (a,b)\} \subset C$, we have: $g'_d(t) = \langle X'(t), d \rangle - \langle X_m, d \rangle \leqslant 0$ for all $t \in (a,b)$. Moreover, according to (15) and the mere definition of C, there is at least one $t^* \in (a,b)$ for which $g'_d(t^*) < 0$. As a consequence,

$$g_d(b) - g_d(a) = \langle X(b) - X(a), d \rangle - \langle X_m, d \rangle \, (b-a) < 0,$$

a contradiction. We therefore have proved (9).

Step 3. According to the classical CARATHEODORY's theorem in convex analysis, each element in $S \subset \mathbb{R}^n$ can be expressed as a convex combination of $n+1$ terms in the image-set $\{X'(c) : c \in (a,b)\}$. Hence (10) is proved. $\boxdot$

At the first glance, Theorem 4 does not seem to retrieve the classical mean value theorem for real-valued functions, that is when $n = 1$. But it does. Indeed, according to DARBOUX' theorem, $\{f'(c) : c \in (a,b)\}$ is an interval for real-valued f; therefore, (9) infers that

$$\frac{f(b) - f(a)}{b-a} \in ri \ \{f'(c) : c \in (a,b)\} . \tag{16}$$

This subsumes two situations : either f is affine on (a,b), in which case $\{f'(c) : c \in (a,b)\} = ri \ \{f'(c) : c \in (a,b)\} = \{s\}$ (a singleton), or

$$\frac{f(b) - f(a)}{b-a} \in int \ \{f'(c) : c \in (a,b)\} . \tag{17}$$

There is a slight improvement of Theorem 4 when, for example, X is continuously differentiable. This relies on a very fine result in convex analysis, called FENCHEL-BUNT theorem, which states the following ([7, page 30]) : if a set $S \subset \mathbb{R}^n$ has no more than n connected components (in particular, if S is connected), then any $x \in coS$ can be expressed as a convex combination of n elements in S.

Corollary 5. *Suppose that the image-set $S = \{X'(t) : t \in (a,b)\}$ of the derivative of X has at most n connected components, which is the case if X is continuously differentiable (in that case, S is a connected set in $\mathbb{R}^n$). Then, there are n elements $c_i \in (a,b)$, n nonnegative coefficients λ_i summing up to 1, such that:*

$$\frac{X(b) - X(a)}{b-a} = \sum_{i=1}^{n} \lambda_i X'(c_i). \tag{18}$$

We end by giving a very simple example which illustrates Theorem 4 and Theorem 5.

Let $X : t \in \mathbb{R} \to \mathbb{R}^3$ be defined by $X(t) = (t^2, t^2, t^3)$ and choose $a = 0, b = 1$. Then, $S = \{X'(t) : t \in (0,1)\} = \left\{(u, u, \frac{3u^2}{4}) : u \in (0,2)\right\}$ is a connected curve whose affine hull is two-dimensional. Consequently, the mean value $(1,1,1)$ of X on $[0,1]$ can be expressed as a convex combination of two elements lying on the curve S.

Final remarks.

1. A mean value result in an equality form like (10) immediately induces an inequality like (8).

2. Mean value theorems indeed have nice kinematic interpretations. Think of $X(t)$ as the position of a moving bicycle at time t, while $X'(t)$ represents its instantaneous velocity at time t. We suppose that the cyclist leaves a starting point at $t_0 = 0$ and goes back there at $t = t_f$.

- If the cyclist moves on a straight road, there necessarily is a time $t^* \in (t_0, t_f)$ at which its instantaneous velocity is null, $X'(t^*) = 0$.

- If the cyclist moves on a plane, he could make his whole trip with a non-null velocity $X'(t)$ all the time. However, at least assuming that X is continuously differentiable, it comes from (18) that there are two times $t_1^*, t_2^* \in (t_0, t_f)$ and some $r > 0$ for which

$$X'(t_2^*) = -rX'(t_1^*),$$

i.e., the velocity vectors are in opposite directions.

3. There are some extensions of Theorem 4 and Corollary 5 to the case where X is not differentiable, but not for any nondifferentiable (or nonsmooth) function X. Here is an example. Suppose that $X : \mathbb{R} \to \mathbb{R}^n$ is locally Lipschitz, that is to say satisfying a Lipschitz property on each bounded interval of $\mathbb{R}$. According to an old theorem by Rademacher (1919), such functions are differentiable almost everywhere, *i.e.*, except on a set of Lebesgue measure zero. Moreover, on the other points, the behavior of the derivative $X'(t)$ can be "controlled". For such functions, a mean value result analogous to (9) is as follows ([4, Theorem 7]):

$$\frac{X(b) - X(a)}{b - a} \in ri\ co\{X'(c) : c \in (a,b)\backslash\Lambda\}, \tag{19}$$

where Λ is the set of points in (a,b) where the derivative fails to exist. Beware however that the image-set $\{X'(c) : c \in (a,b)\backslash\Lambda\}$ may have as many connected components as desired.

4. Conclusion

"Taking a mean value" is a mathematical operation which immediately makes echo to convexity; there therefore is no surprise that the two concepts mix harmoniously in statements of results as well as in proofs. In this note we have shown, on two different contexts (Theorem 3 for real-valued functions, Theorem 4 for vector-valued functions), how they cross-fertilize each other.

References

[1] N. Bourbaki, *Fonctions d'une variable réelle. Chapitres 1, 2, 3 (Dérivées, Primitives et intégrales, Fonctions élémentaires).* Hermann (1958).

[2] T. M. Flett, *Differential analysis.* Cambridge University Press (1980).

[3] J.-B. Hiriart-Urruty, *A note on the mean value theorem for convex functions.* Bolletino U. M. I. (5) 17-B (1980), 765 – 775.

[4] J.-B. Hiriart-Urruty, *Mean value theorems in nonsmooth analysis.* Numer. Funct. Anal. Optim. 2 (1980), 1 – 30.

[5] J.-B. Hiriart-Urruty, *Théorèmes de valeur moyenne sous forme d'égalité pour les fonctions à valeurs vectorielles.* Revue de mathématiques spéciales (RMS), N°7 (1983), 287 – 293.

[6] J.-B. Hiriart-Urruty, *Que manque t-il à une fonction vérifiant la propriété des valeurs intermédiaires pour être continue ?* Revue de mathématiques spéciales (RMS), N°9 (1984), 370 – 371.

[7] J.-B. Hiriart-Urruty and C. Lemaréchal, *Fundamentals of convex analysis.* Grundlehren Text Editions, Springer (2001).

[8] J.-B. Hiriart-Urruty, *Mathematical tapas. Vol. 2 (From undergraduate to graduate level).* Springer Undergraduate Mathematics Series (2017).

[9] R. M. McLeod, *Mean value theorems for vector valued functions.* Proc. Edinburgh Math. Soc., 14, Serie II (1964 – 1965), 197 – 209.

[10] RMS-Revue de la filière mathématiques, Vol. 128-n°3 (2017 – 2018); answer to Question 886 of J.-B. Hiriart-Urruty by C. Madeline-Dérou.

[11] R.T. Rockafellar, *Convex analysis.* Princeton University Press (1970).

Jean-Baptiste Hiriart-Urruty
Institut de Mathématiques
118, route de Narbonne
31062 Toulouse Cedex 9, France
www.math.univ-toulouse.fr/∽jbhu/