

# MEAN VALUE THEOREMS FOR VECTOR VALUED FUNCTIONS

by ROBERT M. McLEOD

(Received 28th April 1964)

## 1. Introduction

The object of this paper is to give a generalisation to vector valued functions of the classical mean value theorem of differential calculus. In that theorem we have

$$f(b) - f(a) = (b - a)f'(c) \quad (1)$$

for some  $c$  in the open interval  $]a, b[$  when  $f$  is a real valued function which is continuous on the closed interval  $[a, b]$  and differentiable on the open interval. The counterpart to (1) when  $f$  has values in an  $n$ -dimensional vector space turns out to be

$$f(b) - f(a) = (b - a) \sum_{k=1}^n \lambda_k f'(c_k) \quad (2)$$

where  $c_k \in ]a, b[$ ,  $0 \leq \lambda_k$ , and  $\sum_{k=1}^n \lambda_k = 1$ .

The key to the algebraic conclusion (2) lies in geometric ideas. Convex sets play a basic role in the statements and proofs of the mean value theorems developed in § 3. Let us recall a few definitions here.

**Definitions.** (a) Let  $x$  and  $y$  be elements of a vector space  $F$ . The segment joining  $x$  and  $y$  is the set of elements  $\lambda x + (1 - \lambda)y$  for all  $\lambda$  such that  $0 \leq \lambda \leq 1$ . (b) A subset  $A$  of  $F$  is convex provided the segment joining each pair of points in  $A$  is contained in  $A$ . (c) If  $A$  is any subset of  $F$ , the convex cover  $H_A$  of  $A$  is the minimal convex set containing  $A$ , that is, the intersection of all convex sets containing  $A$ . (d) When  $F$  is a topological vector space the closed convex cover  $K_A$  of  $A$  is the minimal closed convex set containing  $A$ .

The equation (2) is the end of a three step development. The first step is to prove that  $(f(b) - f(a))/(b - a)$  is in the closed convex cover of the set of values of the derivative. At this stage  $f$  may have values in any locally convex topological linear space. Then under additional hypotheses, for example that the space is finite dimensional, it is also true that  $(f(b) - f(a))/(b - a)$  is in the convex cover of the set of values of the derivative. The algebraic form of this conclusion is

$$f(b) - f(a) = (b - a) \sum_{k=1}^{n+1} \lambda_k f'(c_k) \quad (3)$$

when  $f$  has values in  $n$ -dimensional space. The final step is to show that the number of terms in the sum can be reduced from  $n + 1$  to  $n$ .

In the first two steps it is not necessary to require the existence of a derivative throughout  $]a, b[$ . A right-hand or left-hand derivative can take the place of the ordinary derivative and there may be a countable subset of  $]a, b[$  in which no derivative of any sort need exist. The transition from (3) to (2) is possible under two additional assumptions: (a) The derivative  $f'(t)$  exists for all  $t$  in  $]a, b[-M$  where  $M$  is a set having at most  $n - 1$  points. (b) When  $n \geq 2$  the derivative  $f'$  is continuous from the right at every point of  $]a, b[-M$  or continuous from the left at every point of  $]a, b[-M$ .

The condition (a) is essential for (2). In the case  $n = 1$ , let  $f$  be the real valued function given by  $f(t) = |t|$  for  $-1 \leq t \leq 1$ . Now (3) does hold. Indeed,  $f(1) - f(-1) = 0$  and  $0 = 2(\lambda_1 f'(c_1) + \lambda_2 f'(c_2))$  with  $\lambda_1 = \lambda_2 = \frac{1}{2}$ ,  $-1 < c_1 < 0$ , and  $0 < c_2 < 1$ . But (2) is impossible for all  $c$  in  $] -1, 1[-\{0\}$ . Similar examples can be given of functions having values in  $E_n$  for any  $n$ .

Mean value theorems of other kinds have already been given for vector valued functions. Bourbaki (3) and Dieudonné (4) give one in which the conclusion

$$\|f(b) - f(a)\| \leq g(b) - g(a)$$

is drawn from the hypothesis  $\|f'(t)\| \leq g'(t)$  for all  $t$  in  $]a, b[-M$  where  $M$  is a countable set. See Theorem B in the next section for a full statement of a theorem of this type. It will be convenient to refer to theorems of this kind as increment theorems.

Aziz and Díaz have given a third kind in (2). It is possible to generalise their result by using the version (3) of the mean value theorem for real valued functions. Details are given in Theorem 5.

§ 2 is preparatory. The main results are in § 3 and Theorem 1 is chief among them. The finite dimensional case comes in Theorem 3 and Theorem 4.

## 2. Increment theorems

It is usual to prove that a function with a positive derivative is increasing by using the mean value theorem. Here we shall reverse the procedure. From a sufficient condition for a function to be increasing we can obtain mean value theorems for vector valued functions. From the same starting point and by a very similar method it is also possible to prove increment theorems for vector valued functions. In fact, for economy of exposition it is advantageous to derive the mean value theorems from an increment theorem for vector valued functions.

The increasing function theorem that is suited to the present purpose, Theorem A below, is a special case of theorems given by Aumann (1, p. 222) and Gál (6, p. 310). The principal simplification is in replacing a semi-continuity condition on  $f$  by the assumption of continuity.

**Definitions.** (a) Let  $f$  be a continuous function from an interval  $[a, b]$  into a topological linear space  $F$  over the real field  $R$ . Let  $t \in [a, b[$ . An

element  $y$  of  $F$  is a right-hand derivative value of  $f$  at  $t$  if there exists a sequence  $t_1, t_2, t_3, \dots$  decreasing to  $t$  such that

$$y = \lim_{n \rightarrow \infty} \frac{f(t_n) - f(t)}{t_n - t}.$$

(b) For a real valued function  $f$  let  $D^R f$  be the upper right derivative, i.e.

$$D^R f(t) = \limsup_{s \rightarrow t^+} \frac{f(s) - f(t)}{s - t}.$$

( $D^R f$  may take on the values  $-\infty$  and  $+\infty$  but it will be assumed throughout that right-hand derivative values are finite when  $f$  is real valued.)

**Theorem A.** Let  $f$  be a continuous function from  $[a, b]$  into  $R$ . Let  $M$  be a countable subset of  $]a, b[$  and  $RDf$  a real valued function such that  $RDf(t)$  is a right-hand derivative value of  $f$  for all  $t$  in  $]a, b[ - M$ . If there is a set  $N$  of Lebesgue measure zero such that  $0 \leq D^R f(t)$  for all  $t$  in  $]a, b[ - N$ , then  $f$  is an increasing function on  $[a, b]$ . Moreover, if  $0 < D^R f(t)$  for at least one  $t$  in  $]a, b[$ , then  $f(a) < f(b)$ .

**Definition.** A sublinear functional on a linear space  $F$  is a function  $p$  from  $F$  into  $R$  such that

$$p(\lambda x) = \lambda p(x) \text{ and } p(x + y) \leq p(x) + p(y)$$

for all  $\lambda \geq 0$  and all  $x$  and  $y$  in  $F$ . ( $p$  is a linear functional when the first relation holds for all  $\lambda$  in  $R$  and equality holds for all  $x$  and  $y$  in the second relation.)

The following theorem differs in detail from those stated by Bourbaki and Dieudonné. The proof is the one suggested in (4), p. 156, problem 6.

**Theorem B.** Let  $F$  be a topological linear space. Let  $p$  be a continuous sublinear functional on  $F$ . Let  $f$  and  $g$  be continuous functions from  $[a, b]$  into  $F$  and  $R$  respectively. Suppose there is a countable subset  $M$  of  $]a, b[$  such that, for each  $t$  in  $]a, b[ - M$ ,  $f$  and  $g$  have right-hand derivative values  $RDf(t)$  and  $RDg(t)$  associated with the same sequence decreasing to  $t$ . Suppose also there is a set  $N$  of Lebesgue measure zero such that  $M \subseteq N \subseteq ]a, b[$  and

$$p(RDf(t)) \leq RDg(t) \text{ for all } t \in ]a, b[ - N. \tag{4}$$

Then

$$p(f(d) - f(c)) \leq g(d) - g(c)$$

for all  $[c, d]$  contained in  $[a, b]$ . Moreover, if inequality holds in (4) for at least one value of  $t$ , then  $p(f(b) - f(a)) < g(b) - g(a)$ .

**Proof.** (a) Given  $x_0 \in F$ , let  $F_0 = \{\lambda x_0 : \lambda \in R\}$ . On the subspace  $F_0$  define  $u_0$  by  $u_0(\lambda x_0) = \lambda p(x_0)$ . Then  $u_0(x) \leq p(x)$  for all  $x$  in  $F_0$ . The Hahn-Banach theorem (7, p. 42) provides an extension  $u$  of  $u_0$  to all of  $F$  such that  $u$  is a linear functional and  $u(x) \leq p(x)$  for all  $x$  in  $F$ . Moreover  $u$  is continuous at 0 since  $p$  is continuous and

$$|u(x)| \leq \max \{p(x), p(-x)\}$$

for all  $x$  in  $F$ . Therefore  $u$  is continuous on  $F$ .

(b) Take  $x_0 = f(b) - f(a)$ . Set  $h(t) = g(t) - u(f(t))$ . Since the right-hand derivative values of  $f$  and  $g$  are determined by the same sequence decreasing to  $t$ ,  $RDg(t) - u(RDf(t))$  is a right-hand derivative value of  $h$ . Call it  $RDh(t)$ . Then

$$0 \leq RDg(t) - p(RDf(t)) \leq RDh(t) \quad \text{for all } t \in ]a, b[-N.$$

Since  $RDh(t) \leq D^R h(t)$  Theorem A is applicable. Then  $h(a) \leq h(b)$  or

$$u(f(b)) - u(f(a)) \leq g(b) - g(a).$$

But

$$u(f(b)) - u(f(a)) = u(f(b) - f(a)) = p(f(b) - f(a))$$

and the first conclusion is established for  $c = a$  and  $d = b$ . The restriction of  $f$  to  $[c, d] \subseteq [a, b]$  satisfies the same hypotheses. The first conclusion follows. If inequality holds in (4) then  $0 < RDh(t) \leq D^R h(t)$  and the final conclusion follows from the last assertion of Theorem A.

Note that existence of a right-hand derivative for one of  $f$  and  $g$  and existence of a right-hand derivative value for the other suffices for the existence of right-hand derivative values on a common sequence.

One important case of Theorem B occurs when  $p$  is a norm on  $F$ . But for application to the proofs of mean value theorems it is important that  $p$  can be a linear functional also.

### 3. Mean value theorems

A few more basic ideas now come into use. (a) Subsets  $S$  and  $U$  of a topological linear space  $F$  are a closed half-space and a closed hyperplane, respectively, when

$$S = \{x \in F : u(x) \leq \alpha\} \text{ and } U = \{x \in F : u(x) = \alpha\}$$

for some non-zero continuous linear functional  $u$  and some real  $\alpha$ . (b)  $F$  is a locally convex topological linear space when each neighbourhood of 0 contains a convex neighbourhood of 0.

Since norms are sublinear, the balls  $\{x \in F : \|x\| < \alpha\}$  are convex for all  $\alpha > 0$ . Thus every normed space is locally convex. In particular the Euclidean space  $E_n$  of  $n$ -tuples  $(x_1, x_2, \dots, x_n)$  of real numbers with

$$\|x\| = \left( \sum_{k=1}^n (x_k)^2 \right)^{\frac{1}{2}}$$

is locally convex.

**Theorem 1.** *Let  $F$  be a locally convex topological linear space. Let  $f$  be a continuous function from  $[a, b]$  into  $F$ . Let  $M$  be a countable subset of  $]a, b[$  and  $RDf$  a right-hand derivative value function for  $f$  on  $]a, b[-M$ . Let  $N$  be a set of Lebesgue measure zero such that  $M \subseteq N \subseteq ]a, b[$ . Set*

$$D = \{RDf(t) : t \in ]a, b[-N\}$$

and

$$Q = \{(f(d) - f(c)) / (d - c) : a \leq c < d \leq b\}.$$

Let  $K_D$  and  $K_Q$  be the closed convex covers of  $D$  and  $Q$ . Then  $K_D = K_Q$ .

**Proof.**  $K_D \subseteq K_Q$  is elementary.  $RDf(t)$  is a limit point of  $Q$  for all  $t$  in  $]a, b[-N$ . Thus  $D \subseteq \text{closure of } Q \subseteq K_Q$ . Now  $K_D \subseteq K_Q$  follows from the definition of the closed convex cover.

If  $K_D = F$  then trivially  $K_Q \subseteq K_D$ . Suppose  $K_D \neq F$ . In a locally convex space each closed convex set is the intersection of all the closed half-spaces containing it. (See (7) p. 151, problem 5.) Let  $K_D \subseteq S$  where

$$S = \{x \in F : u(x) \leq \alpha\}.$$

Set  $g(t) = \alpha t$ . Then Theorem B is applicable with  $u$  in place of  $p$ . The conclusion is  $u(f(d) - f(c)) \leq \alpha d - \alpha c$  for all  $[c, d]$  contained in  $[a, b]$ . Thus  $u(q) \leq \alpha$  where  $q = (f(d) - f(c))/(d - c)$ . This is equivalent to  $Q \subseteq S$ . Since  $S$  is an arbitrary half-space containing  $K_D$  the inclusion  $Q \subseteq K_D$  follows. Therefore  $K_Q \subseteq K_D$  and the proof is complete.

Since there may be more than one right-hand derivative value for each  $t$  the function  $RDf$  is not unique. Also many choices may be made for  $N$ . Nevertheless the equality  $K_D = K_Q$  guarantees the uniqueness of  $K_D$  since  $Q$  depends neither on  $RDf$  nor on  $N$ .

In some circumstances it is possible to conclude further that

$$(f(b) - f(a))/(b - a) \in H_D.$$

The proof turns on  $H_D$  having an interior point. A convex subset of  $F$  may fail to have interior points merely because it is contained in a hyperplane and should be seen in a different context. The next theorem is formulated so as to take care of this difficulty.

**Definition.** A linear variety in  $F$  is a translation  $x_0 + F_0$  of a subspace  $F_0$  of  $F$  by an element  $x_0$  of  $F$ .

$x_0 + F_0$  is closed if and only if  $F_0$  is closed. The intersection of any collection of closed linear varieties with a common point is again a closed linear variety.

**Theorem 2.** Let  $F$  and  $f$  be as in Theorem 1. Let  $H_D$  be the convex cover of  $D$ . Let  $V$  be the minimal closed linear variety containing  $H_D$ , i.e. the intersection of all closed linear varieties containing  $H_D$ . If  $H_D$  has an interior point in the relative topology of  $V$ , then  $(f(b) - f(a))/(b - a)$  is an interior point of  $H_D$  relative to  $V$ .

**Proof.** There is no loss in supposing  $0 \in D$  since  $f$  can be replaced by the function  $t \rightarrow f(t) - tRDf(t_0)$  for some  $t_0$  in  $]a, b[-M$  if necessary. With this agreement  $V$  becomes a closed subspace of  $F$ . Any continuous linear functional on  $V$  can be extended to a continuous linear functional on all of  $F$ . (See (7) p. 148.) Therefore the minimality of  $V$  implies that  $H_D$  is not contained in any closed hyperplane in  $V$ .

In any topological linear space a convex set having an interior point has the same interior as its closure. (See (7) p. 133, problem 2.) Since  $V$  is closed the closure of  $H_D$  in  $V$  is  $K_D$ . By Theorem 1,  $(f(b) - f(a))/(b - a) \in K_D$ . It is enough to eliminate the possibility that  $(f(b) - f(a))/(b - a)$  is a boundary point of  $K_D$  relative to  $V$  to complete the proof.

A closed convex set with interior points has a supporting hyperplane at each of its boundary points. (See (7) p. 142, problem 2.) That is, if  $K_D$  is the set and  $y$  is a boundary point of  $K_D$ , there is a closed half-space  $S$  bounded by a hyperplane  $U$  such that  $K_D \subseteq S$  and  $y \in U$ . Let  $S = \{x \in V : u(x) \leq \alpha\}$  and  $U = \{x \in V : u(x) = \alpha\}$  where  $u$  is a continuous linear functional on  $V$ . Since  $H_D$  is not contained in any closed hyperplane in  $V$  there is at least one  $t$  in  $]a, b[ - N$  such that  $u(RDf(t)) < \alpha$ . Apply Theorem B with  $u$  in place of  $p$ . From the final assertion of Theorem B,

$$u(f(b) - f(a)) < \alpha b - \alpha a.$$

Therefore  $(f(b) - f(a))/(b - a)$  is distinct from the boundary point  $y$  of  $K_D$  lying in  $U$ . This completes the proof since  $y$  may be any boundary point of  $K_D$ .

**Corollary.** *Under the hypotheses of Theorem 2 there is an integer  $m$  and numbers  $c_k$  and  $\lambda_k, k = 1, 2, \dots, m$ , such that*

$$f(b) - f(a) = (b - a) \sum_{k=1}^m \lambda_k RDf(c_k),$$

$c_k \in ]a, b[ - N, 0 \leq \lambda_k$ , and  $\sum_{k=1}^m \lambda_k = 1$ .

**Proof.** The convex cover of any set  $A$  consists of all sums of the form  $\sum_{k=1}^m \lambda_k x_k$  where  $x_k \in A, 0 \leq \lambda_k, \sum_{k=1}^m \lambda_k = 1$ , and  $m$  ranges over all the positive integers. (See (7) p. 131.) The conclusion follows. (For convenience call such a sum a convex combination.)

More can be said about  $m$  when  $F$  is finite dimensional. Also the hypothesis that  $H_D$  has an interior point relative to  $V$  becomes superfluous.

If  $A$  is a subset of  $E_n$ , each element of  $H_A$  can be expressed as a convex combination of elements of  $A$  having  $n + 1$  terms. Moreover, if  $A$  has at most  $n$  connected components, the convex combinations having  $n$  terms yield all of  $H_A$ . (See Eggleston (5) p. 35.) Simple examples show no further reduction of  $m$  is possible.

**Theorem 3.** *Replace  $F$  by  $E_n$  in the conditions of Theorem 1. Then  $Q \subseteq H_D$  and  $(f(b) - f(a))/(b - a)$  is an interior point of  $H_D$  relative to the minimal linear variety  $V$  containing  $H_D$ . Moreover,*

$$f(b) - f(a) = (b - a) \sum_{k=1}^{n+1} \lambda_k RDf(c_k) \tag{5}$$

where  $c_k \in ]a, b[ - N, 0 \leq \lambda_k$ , and  $\sum_{k=1}^{n+1} \lambda_k = 1$ . Finally, if  $D$  has at most  $n$  connected components the sum in (5) may be replaced by a sum having  $n$  terms.

**Proof.** Every convex set in  $E_n$  has an interior point relative to the minimal variety containing the set (5, p. 16). Moreover, every variety in  $E_n$  is a closed variety. Thus Theorem 2 can be applied to the restriction of  $f$  to each sub-interval  $[c, d]$ . Then, with obvious notation,

$$(f(d) - f(c))/(d - c) \in H_D[c, d] \subseteq H_D$$

and thus  $Q \subseteq H_D$ . The remaining conclusions follow directly from Theorem 2, its corollary, and the facts about convex sets in  $E_n$  set down in the paragraph preceding Theorem 3.

It does not follow that  $Q$  is contained in the interior of  $H_D$  relative to  $V$ . This is because the minimal variety  $V[c, d]$  associated with the restriction of  $f$  to  $[c, d]$  need not be  $V$ . Let  $f$  be the function from  $[-1, 2]$  into  $R$  given by  $f(t) = |t|$ ,  $-1 \leq t \leq 2$ . In this example  $H_D = [-1, 1]$  and  $V = R$ , but  $H_D[0, 2] = \{1\}$  and  $V[0, 2] = \{1\}$ . Now  $(f(2) - f(0))/2 = 1$  and 1 is not an interior point of  $[-1, 1]$  relative to  $R$ . On the other hand 1 is an interior point of  $\{1\}$  relative to  $\{1\}$ .

**Theorem 4.** *Let  $f$  be a continuous function from  $[a, b]$  into  $E_n$ . Suppose there is a finite set  $M$  having at most  $n - 1$  points such that  $f'(t)$  exists for all  $t$  in  $]a, b[-M$ . When  $n \geq 2$  suppose also that  $f'$  is continuous from the right on  $]a, b[-M$  or continuous from the left on  $]a, b[-M$ . Then  $(f(b) - f(a))/(b - a)$  is equal to a convex combination of  $n$  values of  $f'$ .*

**Proof.** In the case  $n = 1$  the derivative  $f'$  exists throughout  $]a, b[$ . The theorem is just the classical mean value theorem in this case.

Suppose now that  $n \geq 2$ . There is no loss of generality in supposing that  $f'$  is right-hand continuous on  $]a, b[-M$  and in supposing  $f(a) = f(b)$ . From

Theorem 3,  $0 = (f(b) - f(a))/(b - a) = \sum_{k=1}^{n+1} \lambda_k f'(c_k)$  with  $c_k \in ]a, b[-M, 0 \leq \lambda_k$ ,

and  $\sum_{k=1}^{n+1} \lambda_k = 1$ . If any proper subset of  $\{f'(c_1), \dots, f'(c_{n+1})\}$  has a convex combination equal to 0 the proof is complete. Suppose not. Then, in particular,  $\lambda_k f'(c_k) \neq 0$  for  $k = 1, 2, \dots, n + 1$ . There are at most  $n$  open intervals forming  $]a, b[-M$ . Therefore some interval contains two or more of the numbers  $c_k$ . Suppose the notation chosen so that  $c_1$  and  $c_2$  are in the same interval with  $c_1 < c_2$ .

Let  $x_k = -f'(c_k)$ ,  $k = 1, 2, \dots, n + 1$ , and  $A = \{x_2, \dots, x_{n+1}\}$ . Let  $C$  be the cone subtended at the origin by the convex cover of  $A$ . Then  $C$  consists of all points  $\sum_{k=2}^{n+1} \mu_k x_k$  with  $\mu_k \geq 0$  for  $2 \leq k \leq n + 1$ . Since  $\sum_{k=2}^{n+1} \lambda_k x_k = 0$  and  $\lambda_1 x_1 \neq 0$  we see that  $-x_1 \in C$ . We have assumed that 0 does not belong to the convex cover of any proper subset of  $\{x_1, x_2, \dots, x_{n+1}\}$ . Consequently  $-x_2 \notin C$ .

$A$  is a linearly independent set. Suppose  $\sum_{k=2}^{n+1} v_k x_k = 0$  with at least one  $v_k$  not zero. Then

$$\lambda_1 x_1 + \sum_{k=2}^{n+1} (\lambda_k + t v_k) x_k = 0$$

for all real  $t$ . Let  $E$  be the set of all real  $t$  such that  $\lambda_k + t v_k \geq 0$  for  $k = 2, \dots, n + 1$ . Then  $0 \in E$  since every  $\lambda_k$  is positive. There is no loss in generality in supposing some  $v_k$  is negative. Then  $E$  is bounded above. Let  $t_0 = \sup E$ .

By continuity  $t_0 \in E$  and  $\lambda_k + t_0 v_k = 0$  for at least one integer  $k$ . Let

$$v = \lambda_1 + \sum_{k=2}^{n+1} (\lambda_k + t_0 v_k).$$

Then  $v \geq \lambda_1 > 0$  and  $v^{-1} \lambda_1 x_1 + \sum_{k=2}^{n+1} v^{-1} (\lambda_k + t_0 v_k) x_k$  is a convex combination of a proper subset of  $\{x_1, \dots, x_{n+1}\}$  which is equal to zero. This contradicts our assumption.

The linearly independent set  $A$  contains  $n$  elements. Thus any element  $y$  in  $E_n$  has one and only one representation  $\sum_{k=2}^{n+1} \mu_k x_k$ . Furthermore  $y$  is a boundary point of  $C$  if and only if  $\mu_k \geq 0$  for  $2 \leq k \leq n+1$  and  $\mu_k = 0$  for at least one  $k$ .

Let  $d = \sup \{t : f'([c_1, t]) \subseteq C\}$ . Since  $f'(c_1) \in C$  and  $f'(c_2) \notin C$  it follows that  $c_1 \leq d \leq c_2$ . Clearly  $f'([c_1, d]) \subseteq C$ .  $C$  is a closed convex set. From Theorem 1,  $(f(t) - f(d))/(t - d) \in C$  for all  $t$  in  $[c_1, d[$ . Thus  $f'(d) \in C$  and  $d < c_2$ . On the other hand there exists a sequence  $t_1, t_2, t_3, \dots$  decreasing to  $d$  such that  $f'(t_k) \notin C$ . Since  $f'$  is right-hand continuous at  $d$  it is true that  $f'(d)$  belongs to the closure of the complement of  $C$ . Therefore  $f'(d) = \sum_{k=2}^{n+1} \mu_k x_k$  with  $\mu_k \geq 0$  for all  $k$  and  $\mu_k = 0$  for at least one  $k$ . This equation can be rewritten as

$$0 = f'(d) + \sum_{k=2}^{n+1} \mu_k f'(c_k).$$

Multiplication by  $1 / \left(1 + \sum_{k=2}^{n+1} \mu_k\right)$  produces the desired convex combination of  $n$  values of  $f'$  equal to 0.

The condition for reduction of the number of terms in the convex combination given in Theorem 4 is not a consequence of the one given in Theorem 3 when  $n > 1$ . To establish this it suffices to give an example of a function  $f$  satisfying the hypotheses of Theorem 4 for which  $f'([a, b[-M)$  has more than  $n$  components. The function  $f$  whose construction follows is continuous on  $[0, 1/48]$ , differentiable on  $]0, 1/48[$ , and its derivative is right-hand continuous on  $]0, 1/48[$ . It has its values in  $E_2$ . Since  $E_2$  is a linear subspace of  $E_n$  for  $n > 2$  we can also consider  $f$  as a function having values in  $E_n$  for all  $n \geq 2$ . The components of  $f'([0, 1/48[$  are the same whether this set is considered as a subset of  $E_2$  or as a subset of  $E_n$  for  $n > 2$ . The collection of components is countably infinite. Thus this one example serves for every  $n$  greater than 1.

Let  $a$  be a real constant. Set  $g(0, a) = (0, 0)$  and  $g(t, a) = t^2(\cos a/t, \sin a/t)$  for  $t \neq 0$ . The map  $t \rightarrow g'(t, a)$  carries each interval  $[r, 0]$  onto  $\{(0, 0)\} \cup S$  where  $S$  is a spiral which approaches the circle centred at  $(0, 0)$  with radius  $a$  from the outside. Set

$$h(t) = g(t, a) - g(r, a) + (t - r)[y - g'(r, a)] + x$$



where  $x$  and  $y$  are elements of  $E_2$ . Then  $h'(t) = g'(t, a) + y - g'(r, a)$  and the image of  $[r, 0]$  under  $h'$  is  $\{y - g'(r, a)\} \cup T$  where  $T$  is the translate of  $S$  which begins at  $y$ . Since  $\|g'(t, a)\|^2 = 4t^2 + a^2$  the diameter of  $S$ , and of  $T$ , is less than  $2(4r^2 + a^2)^{\frac{1}{2}}$ . We want to choose a sequence of functions  $h_k$  so that the spiral  $T_{k+1}$  starts at the centre of the limiting circle of  $T_k$  and stays inside that circle. This will be true if  $y_{k+1} = y_k - g'(r_k, a_k)$  and  $2(4r_{k+1}^2 + a_{k+1}^2)^{\frac{1}{2}} < a_k$ . We may choose  $a_k = 4^{-k}$ ,  $r_k = -4^{-k-2}$ , and  $y_1 = (0, 0)$ . To complete the determination of  $h_k$  we require that  $h_{k+1}(r_{k+1}) = h_k(0)$ . Equivalently

$$x_{k+1} = -g(r_k, a_k) - r_k y_{k+1} + x_k.$$

The choice  $x_1 = (0, 0)$  completes the definition of all the functions.

Now let  $t_k - t_{k-1} = -r_k$  and  $t_0 = 0$ . Then

$$\lim_{k \rightarrow \infty} t_k = \sum_{k=1}^{\infty} 4^{-k-2} = 1/48.$$

Define  $f$  on  $[0, 1/48]$  as follows:

$$f(t) = h_k(t - t_k) \text{ for } t_{k-1} \leq t < t_k,$$

$$f(1/48) = \lim_{k \rightarrow \infty} x_k.$$

Then  $f$  satisfies the hypotheses of Theorem 4. But  $f'$  maps  $]0, 1/48[$  onto an infinite union of spirals and each spiral is a connected component of the union.

In the case  $n = 1$  Theorem 4 is a consequence of Theorem 3. Indeed,  $f'(t)$  exists throughout  $]a, b[$  hence, by the Theorem of Darboux,  $f'$  has the intermediate value property. Thus  $f'(]a, b[)$  is an interval and the number of components is 1.

We used right-hand derivative values in Theorems 1, 2 and 3. Left-hand derivative values can be used equally well. In fact the corresponding theorems are corollaries of the ones given. Indeed, if  $f$  has left-hand derivative values  $LDf(t)$  for all  $t$  in  $]a, b[-M$  and

$$g(t) = -f(-t), \quad -b \leq t \leq -a,$$

then  $g$  has right-hand derivative values such that  $RDg(t) = LDf(-t)$ . Moreover  $(g(d) - g(c))/(d - c) = (f(-d) - f(-c))/(-d + c)$ . Thus application of Theorems 1, 2 and 3 to  $g$  yields corresponding theorems for  $f$  in which  $D$ ,  $H_D$  and  $K_D$  are formed from the values of  $LDf$ . Furthermore, if  $f$  has both a left-hand and a right-hand derivative value function, the two sets  $K_D$  formed from  $LDf$  and  $RDf$  are equal since both are equal to  $K_Q$ .

#### 4. Applications

Many of the theorems of elementary calculus which are usually derived from the classical mean value theorem can just as well be got from an increment theorem like Theorem B. Dieudonné has given an excellent account of their extensions to Banach spaces in (4) pp. 156 ff. We will avoid duplicating them

here and concentrate on some applications which come more readily from mean value theorems.

The first is a generalisation of the theorem of Aziz and Diaz (2, p. 261).

**Theorem 5.** *Let  $F$  be a topological linear space. Let  $f$  be a continuous function from  $[a, b]$  into  $F$  having a right-hand derivative value function  $RDf$  on  $]a, b[ - M$ , where  $M$  is countable. Let  $p$  be a continuous sublinear functional on  $F$ . If  $N$  is any set of Lebesgue measure zero such that  $M \subseteq N \subseteq ]a, b[$ , there exists  $c \in ]a, b[ - N$  such that*

$$p(f(b) - f(a)) \leq p(RDf(c)).$$

**Proof.** From part (a) of the proof of Theorem B there is a continuous linear functional  $u$  on  $F$  such that  $p(f(b) - f(a)) = u(f(b) - f(a))$  and  $u(x) \leq p(x)$  for all  $x$  in  $F$ . Consider the real valued function  $h$  given by  $h(t) = u(f(t))$  for all  $t$  in  $[a, b]$ . Since  $u(RDf(t))$  is a right-hand derivative value for  $h$ , Theorem 3 is applicable. Thus

$$u(f(b)) - u(f(a)) = \lambda_1 u(RDf(c_1)) + \lambda_2 u(RDf(c_2))$$

with  $0 \leq \lambda_1, 0 \leq \lambda_2$ , and  $\lambda_1 + \lambda_2 = 1$ . Let  $c$  be that one of  $c_1$  and  $c_2$  giving the larger value of  $u(RDf(t))$  for  $t = c_1$  and  $t = c_2$ . The left-hand side of the equation is  $p(f(b) - f(a))$  and the right-hand side is dominated by  $u(RDf(c))$  and hence by  $p(RDf(c))$ . This completes the proof.

**Theorem 6.** *Let  $f$  and  $F$  be as in Theorem 1. Let  $F_0$  be a closed subspace of  $F$ . The set of values of  $f$  is a subset of the variety  $x_0 + F_0$  if and only if  $D \subseteq F_0$  and  $f(t) \in x_0 + F_0$  for some  $t$  in  $[a, b]$ .*

**Proof.** Suppose the values of  $f$  all lie in  $x_0 + F_0$ . Clearly  $Q \subseteq F_0$ . Since  $F_0$  is closed,  $D \subseteq \text{closure of } Q \subseteq F_0$ . Conversely, suppose  $D \subseteq F_0$  and  $f(t_0) \in x_0 + F_0$ . Since  $F_0$  is closed and convex,  $K_D \subseteq F_0$ . By Theorem 1,  $Q \subseteq K_D$ . Thus

$$f(t) = f(t_0) + (t - t_0)y, \quad y \in F_0,$$

for each  $t$  in  $[a, b]$ . But  $f(t_0) = x_0 + y_0$  with  $y_0 \in F_0$ . Since  $F_0$  is a subspace,  $y_0 + (t - t_0)y \in F_0$ . Thus  $f(t) \in x_0 + F_0$  for all  $t$  in  $[a, b]$ .

When  $F = E_3$  and  $F_0$  is a plane through the origin the geometric content of Theorem 6 is that  $f$  is represented by a plane curve in a plane parallel to  $F_0$  if and only if the values of the function  $RDf$  lie in the plane  $F_0$ . Note that it is not merely a matter of the values of  $RDf$  lying in a plane. For example, let

$$f(t) = (\cos t, \sin t, t), \quad 0 \leq t \leq 2\pi.$$

The curve which represents the derivative  $f'$  is a plane curve, a circle in the plane  $x_3 = 1$ , and  $f$  is represented by a spiral which lies in no plane.

We continue with a geometrical point of view and consider closed curves in  $E_2$ . It seems intuitively clear that a continuously turning tangent vector on a closed curve must turn through at least  $\pi$  radians in one circuit of the curve. It is possible to make Theorem 4 yield precise information to this effect.

**Theorem 7.** *Suppose  $f$  is a continuous function from  $[a, b]$  into  $E_2$  having a derivative  $f'(t)$  for all  $t$  in  $]a, b[-\{c\}$  for some  $c$  in  $]a, b[$ . Suppose  $f'$  is right-hand continuous on  $]a, b[-\{c\}$  or left-hand continuous on  $]a, b[-\{c\}$ . Suppose also  $f'(t) \neq 0$  for all  $t$  in  $]a, b[-\{c\}$ . If  $f(a) = f(b)$  there are numbers  $c_1$  and  $c_2$  in  $]a, b[-\{c\}$  and a positive number  $\mu$  such that  $f'(c_1) = -\mu f'(c_2)$ . (In geometrical language, when the curve is closed there is at least one pair of oppositely directed forward tangent vectors.)*

**Proof.** By Theorem 4,

$$f(b) - f(a) = (b - a)(\lambda_1 f'(c_1) + \lambda_2 f'(c_2))$$

with  $0 \leq \lambda_1, 0 \leq \lambda_2$  and  $\lambda_1 + \lambda_2 = 1$ . Since  $f(b) = f(a)$  the convex combination is 0. Since neither of the values of the derivative is zero neither  $\lambda_1$  nor  $\lambda_2$  is zero. Thus  $0 < \lambda_1$  and  $0 < \lambda_2$ . The conclusion follows with  $\mu = \lambda_2/\lambda_1$ .

If  $f$  maps  $[a, a_1], [a_1, a_2]$  and  $[a_2, b]$  linearly onto the three successive sides of a triangle it fails to satisfy the hypotheses of Theorem 7 only in having no derivative at two points in  $]a, b[$  instead of at one point. The conclusion fails also.

We turn now to some applications of the mean value theorems to functions from one normed space to another. The definition of the derivative, or total differential, is needed in order to state them.

**Definition.** Let  $f$  be a function from an open set  $G$  in a normed space  $E$  into a normed space  $F$ .  $f$  is said to be differentiable at  $y$  in  $G$  if there is a continuous linear mapping  $u$  from  $E$  into  $F$  such that

$$\lim_{x \rightarrow y} \|f(x) - f(y) - u(x - y)\| / \|x - y\| = 0.$$

The linear mapping  $u$  is called the derivative of  $f$  at  $y$  and is denoted by  $f'(y)$ . Its value at any  $x$  in  $E$  is denoted by  $f'(y) \cdot x$ .

If  $g$  is a function from  $[a, b]$  into  $G$  having a derivative value  $RDg(t)$  the chain rule for the composite function  $fg$  reads

$$RDfg(t) = f'(g(t)) \cdot RDg(t).$$

**Definitions.** (a) Let  $x$  and  $y$  belong to the connected open set  $G$ . Let  $P(x, y)$  denote the collection of all functions  $g$  from  $[0, 1]$  into  $G$  such that  $g(0) = x, g(1) = y, g$  is continuous on  $[0, 1]$ , and there exists a countable set  $M$  such that  $g$  has a right-hand derivative value function  $RDg$  on  $]0, 1[-M$ . (b) Suppose  $f$  is a differentiable function from  $G$  into  $F$ . Let  $K(x, y; g)$  denote the closed convex cover of  $D(x, y; g)$  where, for fixed  $g$  in  $P(x, y)$ ,

$$D(x, y; g) = \{f'(g(t)) \cdot RDg(t) : t \in ]0, 1[-M\}.$$

Also let

$$K(x, y) = \bigcap_{g \in P(x, y)} K(x, y; g).$$

**Theorem 8.** *Let  $f$  be a differentiable function from the connected open set  $G$  in the normed space  $E$  into the normed space  $F$ . Then*

$$f(y) - f(x) \in K(x, y)$$

for all  $x$  and  $y$  in  $G$ . Moreover, if  $F = E_n$ ,  $f(y) - f(x)$  is an interior point of  $K(x, y; g)$  relative to the minimal variety containing it for each  $g$  in  $P(x, y)$ . Also  $f(y) - f(x)$  is a convex combination of  $n + 1$  elements of  $D(x, y; g)$ .

**Proof.** This is a direct application of Theorems 1 and 3 using the chain rule noted above.

A global sufficient condition for  $f$  to be one-to-one on  $G$  comes out of Theorem 8 at once.

**Theorem 9.** Under the conditions of Theorem 8,  $f$  is one-to-one on  $G$  if  $0 \notin K(x, y)$  for all  $x$  and  $y$  in  $G$ . When  $F = E_n$  it is also sufficient that for each  $x$  and  $y$  in  $G$  there exists  $g$  in  $P(x, y)$  such that  $0$  is not a relative interior point of  $K(x, y; g)$ .

(A local sufficient condition for  $f$  to be one-to-one can be obtained in a useful form from Theorem B. See (4) p. 156 and p. 268.)

Now let  $f$  be a holomorphic complex valued function defined on a connected open set  $G$  in the complex plane. There are two concepts of derivative for  $f$ , both denoted by  $f'(y)$ , which are related so that  $f'(y) \cdot x$  can be read correctly as the value of the linear mapping  $f'(y)$  at  $x$  or as the product of the complex number  $f'(y)$  and the complex number  $x$ . Consequently, Theorems 8 and 9 are applicable to holomorphic functions. Since  $f'$  is continuous we can use Theorem 4 to get the following fact also.

**Theorem 10.** Suppose  $f$  is a holomorphic function defined on a connected open set  $G$  in the complex plane. If  $z_1$  and  $z_2$  are points in  $G$  such that the segment joining them is also in  $G$  then

$$f(z_2) - f(z_1) = (z_2 - z_1)(\lambda_1 f'(w_1) + \lambda_2 f'(w_2))$$

for some  $w_1$  and  $w_2$  on the segment joining  $z_1$  and  $z_2$  and some  $\lambda_1$  and  $\lambda_2$  such that  $0 \leq \lambda_1, 0 \leq \lambda_2$  and  $\lambda_1 + \lambda_2 = 1$ .

The convex combination in this mean value equation cannot be replaced by a value  $f'(w)$  in general. The example  $f(z) = e^z$  with  $z_2 = z_1 + 2\pi i$  shows this since  $f(z_2) - f(z_1) = 0$  but  $(z_2 - z_1) f'(w) = 2\pi i e^w \neq 0$  for all  $w$ .

These facts throw light on why the Jacobian condition  $|f'(t)| \neq 0$  is a global sufficient condition for a real valued function of a real variable to be one-to-one on an interval while the counterpart  $|f'(z)| \neq 0$  is only a local sufficient condition in the case of holomorphic functions.

### 5. Examples

There are standard examples which show that the countable set  $M$  cannot be replaced by a set of measure zero in Theorem A and its consequences.

Let  $T$  be the Cantor ternary set in  $[0, 1]$ . Let  $k$  be the continuous function on  $[0, 1]$  such that  $k(0) = 1, k(1) = 0, k(t) = \frac{1}{2}$  for  $\frac{1}{3} < t < \frac{2}{3}, k(t) = \frac{3}{4}$  for  $\frac{1}{9} < t < \frac{2}{9}, k(t) = \frac{1}{4}$  for  $\frac{7}{9} < t < \frac{8}{9}$ , etc. Then  $k'(t) = 0$  for all  $t$  in  $[0, 1] - T$ . Since  $T$  is a set of Lebesgue measure zero this example shows that it is not possible

to replace the countable set  $M$  by a set of Lebesgue measure zero in Theorem A and Theorem 1.

Let  $k_2(t) = (t, k(t))$  for all  $t$  in  $[0, 1]$  and  $k_3(t) = (t, k(t), 1)$  for  $0 \leq t \leq 1$ ,  $k_3(t) = (t, 0, k(t-1))$  for  $1 < t \leq 2$ . In both of these examples  $D$  is contained in a one dimensional subspace if we use only values of  $t$  outside the Cantor sets. On the other hand the values of  $k_2$  span  $E_2$  and the values of  $k_3$  span  $E_3$ . These examples contrast strongly with Theorem 6.

## REFERENCES

- (1) GEORG AUMANN, *Reelle Funktionen* (Berlin, 1954).
- (2) A. K. AZIZ and J. B. DIAZ, On a mean value theorem of the differential calculus of vector-valued functions, and uniqueness theorems for ordinary differential equations in a linear-normed space, *Contributions to Differential Equations*, 1 (1963), 251-269.
- (3) N. BOURBAKI, *Éléments de Mathématique*, Livre IV, *Fonctions d'une variable réelle*, Chaps. 1, 2, 3 (2nd ed., Paris, 1958).
- (4) J. DIEUDONNÉ, *Foundations of Modern Analysis* (New York, 1960).
- (5) H. G. EGGLESTON, *Convexity* (Cambridge, 1958).
- (6) ISTVÁN S. GÁL, On the fundamental theorems of the calculus, *Trans. Amer. Math. Soc.* 86 (1957), 309-320.
- (7) ANGUS E. TAYLOR, *Introduction to Functional Analysis* (New York, 1958).

AMERICAN UNIVERSITY OF BEIRUT  
BEIRUT, LEBANON