

A MEAN VALUE THEOREM FOR GENERALIZED RIEMANN DERIVATIVES

H. FEJZIĆ, C. FREILING, AND D. RINNE

(Communicated by David Preiss)

ABSTRACT. Functional differences that lead to generalized Riemann derivatives were studied by Ash and Jones in (1987). They gave a partial answer as to when these differences satisfy an analog of the Mean Value Theorem. Here we give a complete classification.

1. INTRODUCTION

Throughout this article, Δ will denote the following functional difference:

$$\Delta f(x) = \sum_{i=1}^n a_i f(x + b_i h),$$

where a_1, \dots, a_n and $b_1 < b_2 < \dots < b_n$ are constants. We will also suppose that there is some positive integer d , called the *order* of Δ , such that

$$(1.1) \quad \|\Delta\|^j = \sum_{i=1}^n a_i b_i^j = 0 \text{ for } 0 \leq j < d,$$

$$(1.2) \quad \|\Delta\|^d = \sum_{i=1}^n a_i b_i^d = d! \text{ (normalized).}$$

A generalized Riemann derivative is created by

$$\lim_{h \rightarrow 0^+} \frac{\Delta f(x)}{h^d}$$

when this limit exists. It is a well-known fact that if the ordinary d th derivative, $f^{(d)}(x)$, exists, then so does the generalized Riemann derivative and the two must agree. In the case where $f^{(d)}$ is continuous at x this can be easily seen by d -applications of l'Hôpital's Rule. More generally, it follows from a variant of Taylor's Theorem first proved by Peano (see pp. 245–246 of [2] for a statement and history).

Definition 1.1. We will say that Δ possesses the *mean value property* if for any x , any $h > 0$, and any function $f(x)$ such that $f^{(d-1)}(x)$ is continuous on

Received by the editors March 28, 2006 and, in revised form, October 10, 2006.

2000 *Mathematics Subject Classification.* Primary 26A06, 26A24.

Key words and phrases. Mean value theorem, generalized derivatives.

The second author was supported in part by NSF.

$[x + b_1h, x + b_nh]$ and differentiable on $(x + b_1h, x + b_nh)$, we have

$$\frac{\Delta f(x)}{h^d} = f^{(d)}(z)$$

for some $z \in (x + b_1h, x + b_nh)$.

In [1] Ash and Jones gave a partial classification for the mean value property. First of all, it must be that $d \leq n - 1$ since if (1) holds for all $0 \leq j \leq n - 1$, then it holds for all j , rendering (2) impossible. The difference, $(n - 1) - d$ is called the *excess* of Δ .

In the case where $d = n - 1$, i.e. the excess is zero, the authors of [1] show that the values of a_i are uniquely determined from the values of b_i . The mean value property holds in this case, and this is a well-known numerical analysis fact (see e.g. [4]).

If $d < n - 1$, then Δ can be expressed naturally as a linear combination of differences whose excess is zero. These are the unique excess zero differences that come from the sets $\{b_1, \dots, b_{d+1}\}, \{b_2, \dots, b_{d+2}\}, \dots, \{b_{n-d}, \dots, b_n\}$. If this turns out to be a convex combination, then Δ satisfies the mean value property, and this is easy to show from the well-known Darboux property (or Intermediate Value Property) of derivatives. The authors then showed that the converse of this fact holds in the cases $d = 1, 2, n - 2$, and asked whether the converse holds in general. They also gave a specific order-three difference for which the mean value property was an open question:

$$-\frac{5}{8}f(x) + \frac{17}{8}f(x+h) - \frac{26}{8}f(x+2h) + \frac{26}{8}f(x+3h) - \frac{17}{8}f(x+4h) + \frac{5}{8}f(x+5h).$$

This difference could not be decided using their method since the corresponding linear combination had a negative coefficient.

In this note we will give a condition that is both necessary and sufficient for all $d > 0$. Using this classification, we will quickly determine that the particular difference given above does not satisfy the mean value property. We also examine a slight variation of this difference that retains its negative coefficient and yet it does satisfy the mean value property. This shows that the condition in [1] is not necessary for the case $d = 3, n = 6$.

As with the proof in [1], our approach will be based on the Darboux property of derivatives. As for the numerical analysis fact stated above, rather than using it as a lemma, we will derive it as a corollary. Our basic technique is to first represent the difference as an integral and then apply several integrations by parts. All of our integrals will be Denjoy integrals, and we summarize some of their relevant properties in the next section.

2. DENJOY INTEGRALS

Although it would be possible to carry out all of the arguments in this paper using the Riemann integral, the proofs are simplified by making use of the Denjoy integral. The utility of the Denjoy integral comes from the fact that it is powerful enough to integrate any derivative. More generally, it gives us the following integration by parts formula.

Theorem 2.1 (see e.g. [3], Theorem 12.6). *Let F be a continuous function on $[a, b]$ that is differentiable on (a, b) with $F'(x) = f(x)$, and let $g(x)$ be Lebesgue integrable*

on $[a, b]$ with $G(x) = \int_a^x g(t)dt + c$. Then Fg is Lebesgue integrable, fG is Denjoy integrable and

$$\int_a^b Fgdx = F(x)G(x)]_a^b - \int_a^b fGdx.$$

Setting $g(x) = 0$ and $G(x) = 1$ in the above theorem results in the Fundamental Theorem of Calculus,

$$\int_a^b f(x)dx = F(x)]_a^b$$

for any differentiable function $F(x)$.

Other expected properties of integration that hold for the Lebesgue integral continue to hold for the Denjoy integral. For example, the Denjoy integral is both linear and monotone:

Theorem 2.2 (see e.g. [3], Theorem 7.4). *If $f(x)$ and $g(x)$ are two Denjoy integrable functions on $[a, b]$ and k is a constant, then $kf(x) + g(x)$ is also Denjoy integrable on $[a, b]$ and*

$$\int_a^b kf(x) + g(x)dx = k \int_a^b f(x)dx + \int_a^b g(x)dx.$$

Furthermore, if $f(x) \leq g(x)$, then

$$\int_a^b f(x)dx \leq \int_a^b g(x)dx.$$

Also, if f is Denjoy integrable on an interval, then it is Denjoy integrable on every subinterval (see e.g. [3], Theorem 7.4).

Finally, for many functions, the two notions of integration agree:

Theorem 2.3 (see e.g. [3], Theorem 7.7). *If $f(x)$ is Lebesgue integrable on $[a, b]$, then it is also Denjoy integrable on $[a, b]$ and the two integrals agree. If $f(x)$ is Denjoy integrable and f is bounded from below, then $f(x)$ is Lebesgue integrable.*

3. DIFFERENCES AS INTEGRALS

To represent the difference Δ as an integral, we will use the Dirac delta function $\delta(x)$. Although technically not a function, $\delta(x)$ acts just like a function when used inside an integral. There are slight variations of $\delta(x)$ in the literature, depending on the exact property needed. For us, the relevant property is that for any function $f(t)$ continuous from above at 0, we have

$$(3.1) \quad \int_0^\infty \delta(t)f(t)dt = f(0).$$

The antiderivative of $\delta(x)$ is the Heaviside unit step function, which is an actual function,

$$\mathcal{H}(x) = \int_{-\infty}^x \delta(t)dt = \begin{cases} 0 & \text{if } x \leq 0, \\ 1 & \text{if } x > 0. \end{cases}$$

Using $\mathcal{H}(x)$ we may write (3.1) more generally as

$$(3.2) \quad \int_u^v \delta(t)f(t)dt = f(0)[\mathcal{H}(v) - \mathcal{H}(u)].$$

One of the function-like properties that $\delta(x)$ enjoys is integration by parts. Thus, if $f(x)$ is differentiable on (u, v) and continuous on $[u, v]$, then

$$\int_u^v \delta(t)f(t)dt = f(t)\mathcal{H}(t)]_u^v - \int_u^v f'(t)\mathcal{H}(t)dt,$$

a fact that follows easily from (3.2).

Given a functional difference Δ we define an associated distribution

$$(3.3) \quad D(t) = \sum_{i=1}^{n-1} a_i \delta(t - b_i h) + a_n \delta(b_n h - t).$$

The reason for choosing a different δ -function for $b_n h$ is because we want to work with a function f that is continuous on the closed interval $[x + b_1 h, x + b_n h]$, so that it is continuous from above at $x + b_i h$ as long as $i < n$, but it is continuous from below at $x + b_n h$. Then, assuming f is such a function, we may use (3.2) to write

$$(3.4) \quad \Delta f(x) = \sum_{i=1}^n a_i f(x + b_i h) = \int_{b_1 h}^{b_n h} D(t)f(t + x)dt,$$

which gives us the desired integral for Δ .

We wish to integrate (3.4) several times, so we let $D^{[j]}(x)$ denote the j th-order antiderivative of $D(x)$ defined inductively by $D^{[0]} = D$ and $D^{[k+1]}(x) = \int_{-\infty}^x D^{[k]}(t)dt$. The next lemma tells us how to quickly compute the values of $D^{[j]}$ at the endpoints $b_1 h$ and $b_n h$.

Lemma 3.1. *If Δ is a normalized difference of order $d > 0$, then $D^{[j]}(b_1 h) = 0$ for all j , $D^{[j]}(b_n h) = 0$ for $j = 0, \dots, d$ and $D^{[d+1]}(b_n h) = (-1)^d h^d$.*

Proof. Let $u = b_1 h$ and $v = b_n h$. The values at u are immediate from the definition, since D is identically zero on $(-\infty, u)$. At the other endpoint, for $j = 0$ this says $D(v) = 0$, which follows by the definition of D and the fact that $d > 0$. We proceed by induction. Assume the lemma holds for $j = 0, 1, \dots, k \leq d$. Then using integration by parts,

$$\begin{aligned} D^{[k+1]}(v) &= \int_u^v D^{[k]}(t)dt \\ &= D^{[k]}(t)t]_u^v - D^{[k-1]}(t) \left. \frac{t^2}{2} \right]_u^v + \dots + (-1)^{k-1} D^{[1]}(t) \left. \frac{t^k}{k!} \right]_u^v \\ &+ (-1)^k \int_u^v D(t) \frac{t^k}{k!} dt. \end{aligned}$$

By the induction hypothesis, all of the terms on the right are zero except possibly the last one, which is

$$\begin{aligned} (-1)^k \int_u^v D(t) \frac{t^k}{k!} dt &= (-1)^k \sum_{i=1}^n a_i \frac{(b_i h)^k}{k!} \\ &= (-1)^k \frac{\|\Delta\|^k}{k!} h^k. \end{aligned}$$

Since Δ has order d , this is zero for $k < d$. Since Δ is normalized, this is $(-1)^d h^d$ when $k = d$. □

Lemma 3.1 allows us to do many integration by parts very easily, as in the next lemma.

Lemma 3.2. *Let Δ be a normalized difference of order $d > 0$ and $j \leq d$. If $f^{(j-1)}$ is continuous on $[x + u, x + v] = [x + b_1h, x + b_nh]$ and differentiable on $(x + u, x + v)$, then*

$$\Delta f(x) = (-1)^j \int_u^v D^{[j]}(t) f^{(j)}(t + x) dt .$$

Proof. Using integration by parts, we compute

$$\begin{aligned} \Delta f(x) &= \int_u^v D(t) f(t + x) dt \\ &= D^{[1]}(t) f(t + x) \Big|_u^v - D^{[2]}(t) f'(t + x) \Big|_u^v + \dots \\ &\quad + (-1)^{j-1} D^{[j]}(t) f^{(j-1)}(t + x) \Big|_u^v + (-1)^j \int_u^v D^{[j]}(t) f^{(j)}(t + x) dt . \end{aligned}$$

But by Lemma 3.1, all of the terms on the right except the last one are zero. □

We now summarize our results so far.

Lemma 3.3. *If Δ is a normalized difference of order $d > 0$ and $f^{(d-1)}$ is continuous on $[x + u, x + v] = [x + b_1h, x + b_nh]$ and differentiable on $(x + u, x + v)$, then*

$$\begin{aligned} h^d &= (-1)^d \int_u^v D^{[d]}(t) dt , \\ \Delta f(x) &= (-1)^d \int_u^v D^{[d]}(t) f^{(d)}(t + x) dt . \end{aligned}$$

4. MAIN RESULTS

Theorem 4.1. *If Δ is a normalized difference of order $d > 0$. Then the generalized Riemann derivative*

$$\lim_{h \rightarrow 0^+} \frac{\Delta f(x)}{h^d}$$

possesses the mean value property if and only if $D^{[d]}$ does not change sign.

Proof. Suppose first that $D^{[d]}$ does not change sign and let $f(x)$ be a function such that $f^{(d-1)}$ is continuous on $[x + u, x + v] = [x + b_1h, x + b_nh]$ and differentiable on $(x + u, x + v)$. From Lemma 3.3 we have

$$\Delta f(x) = (-1)^d \int_u^v D^{[d]}(t) f^{(d)}(t + x) dt .$$

Let m and M be the infimum and supremum, respectively, of $f^{(d)}$ on $(x + u, x + v)$. Since $D^{[d]}$ does not change sign, assume $(-1)^d D^{[d]}(t) \geq 0$; in the other case, the following inequalities will be reversed. Then by monotonicity,

$$m(-1)^d \int_u^v D^{[d]}(t) dt \leq \Delta f(x) \leq M(-1)^d \int_u^v D^{[d]}(t) dt$$

or, using Lemma 3.3, we may write

$$m \leq \frac{\Delta f(x)}{h^d} \leq M.$$

If both of these inequalities are strict, then $f^{(d)}$ attains values on each side of $\frac{\Delta f(x)}{h^d}$. The theorem follows since $f^{(d)}$ possesses the Darboux property. If one of the inequalities is not strict, say if $m = \frac{\Delta f(x)}{h^d}$, then m is finite. Using Theorem 2.3 almost everywhere on $(x + u, x + v)$ we have either $f^{(d)} = m$ or $D^{[d]}(t) = 0$. By Lemma 3.3, $D^{[d]}(t)$ cannot be zero almost everywhere, so the theorem follows.

Now suppose that $D^{[d]}$ has at least one sign change. We set $x = 0$ and $h = 1$. We will find a function $f(x)$ that is d times differentiable, but for which the mean value property fails. We may assume that $D^{[d]}(t)$ is positive on an open interval $I \subseteq (b_1, b_n)$, and negative on an open interval $J \subseteq (b_1, b_n)$. If d is odd, we pick g to be a continuous function that is negative on I and 0 elsewhere and set $f^{(d)} = g$. Using Lemma 3.3, we get $\Delta f(0) = -\int_{b_1}^{b_n} D^{[d]}(t)f^{(d)}(t)dt > 0$ while $f^{(d)} \leq 0$. If d is even, we pick g to be a continuous function that is negative on J and 0 elsewhere and set $f^{(d)} = g$. Thus $\Delta f(0) = \int_{b_1}^{b_n} D^{[d]}(t)f^{(d)}(t)dt > 0$ while $f^{(d)} \leq 0$. In either case the mean value property fails. \square

Although the number of sign changes in $D^{[d]}$ is certainly computable, it may be tedious to determine by hand. The next theorem gives a quick but rough estimate.

Theorem 4.2. *Let Δ be a normalized difference of degree d . Let c_k be the number of sign changes in $D^{[k]}$, $1 \leq k \leq d$. Then*

- the sequence c_1, c_2, \dots, c_d is strictly decreasing,
- c_1 is the number of sign changes in the partial sums of $\sum_{i=1}^n a_i = 0$,
- $c_k \leq n - k - 1$.

Proof. For each $1 \leq k \leq d$ we have from Lemma 3.1 that $D^{[k]}(b_1h) = D^{[k]}(b_nh) = 0$. Since D is zero outside the interval (b_1h, b_nh) all sign changes must occur in this interval. Between any two sign changes in $D^{[k]}(x) = \int_{-\infty}^x D^{[k-1]}(t)dt$ there must be a sign change in $D^{[k-1]}$. To satisfy the requirements $D^{[k]}(b_1h) = D^{[k]}(b_nh) = 0$, there must also be a sign change in $D^{[k-1]}$ before the first and after the last sign change in $D^{[k]}$. Therefore, $c_{k-1} \geq c_k + 1$. The second part follows directly from the definition $D^{[1]} = \sum_{i=1}^{n-1} a_i \mathcal{H}(t - b_ih) - a_n \mathcal{H}(b_nh - t)$. The third part follows directly from the first two parts, observing that the partial sums of $\sum_{i=1}^n a_i = 0$ can have at most $n - 2$ sign changes, and so $c_1 \leq n - 2$. \square

If $d = n - 1$, then Theorem 4.2 says that $c_d = 0$. Combining this with Theorem 4.1 we get the following well-known result.

Corollary 4.3. *If $d = n - 1$, then Δ possesses the mean value property.*

5. EXAMPLES

Example 5.1. In [1] the authors asked whether the following degree-difference three possesses the mean value property. For $i = 0, 1, 2$, let

$$\Delta_i f(x) = -f(x + ih) + 3f(x + (i + 1)h) - 3f(x + (i + 2)h) + f(x + (i + 3)h)$$

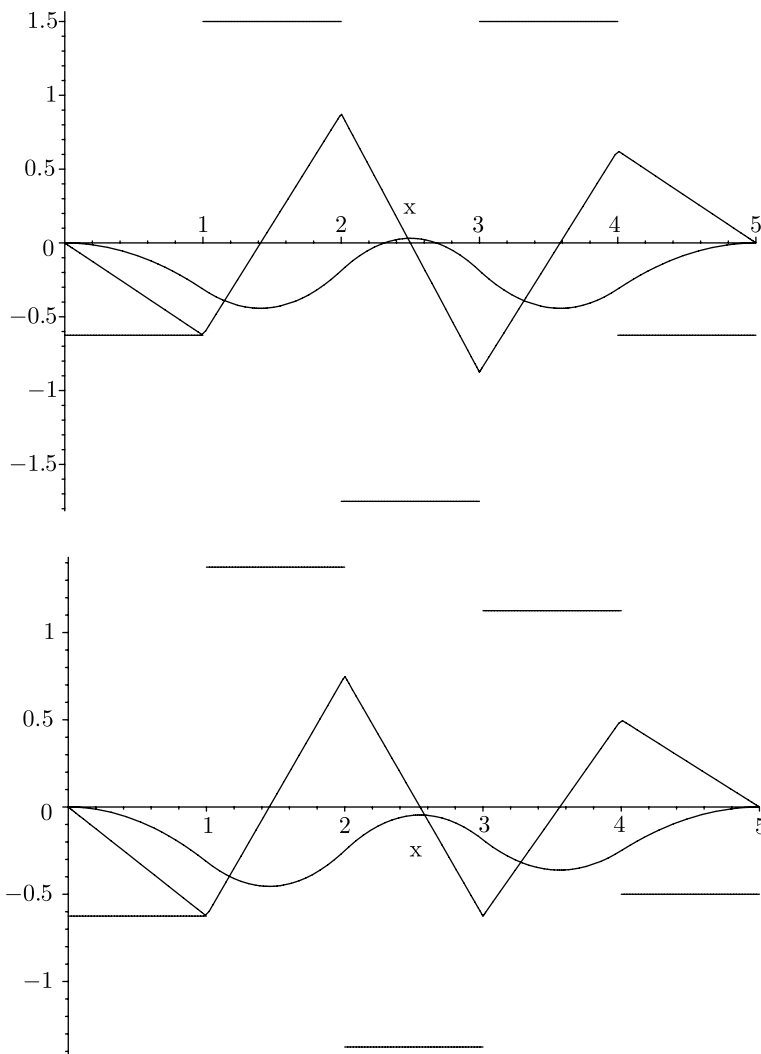


FIGURE 1. The first three antiderivatives of D for Examples 5.1 and 5.2. The mean value property fails in the first case and succeeds in the second.

and define

$$\begin{aligned} \Delta f(x) &= \frac{5}{8}\Delta_0 f(x) - \frac{1}{4}\Delta_1 f(x) + \frac{5}{8}\Delta_2 f(x) \\ &= -\frac{5}{8}f(x) + \frac{17}{8}f(x+h) - \frac{26}{8}f(x+2h) \\ &\quad + \frac{26}{8}f(x+3h) - \frac{17}{8}f(x+4h) + \frac{5}{8}f(x+5h). \end{aligned}$$

This question was posed in [1] since writing $\Delta f(x)$ as a linear combination of Δ_i involves negative coefficients. In this example $D^{[1]}$ takes on the values $-\frac{5}{8}, \frac{12}{8}, -\frac{14}{8}, \frac{12}{8}, -\frac{5}{8}$ on the intervals $(i, i+1), 0 \leq i \leq 4$. It is then fairly easy to determine that

$D^{[3]}$ has two sign changes, so Theorem 4.1 answers the question in the negative. Figure 1 shows the functions $D^{[1]}$, $D^{[2]}$, and $D^{[3]}$ with $h = 1$. Notice how the graph of $D^{[3]}$ climbs slightly above the x -axis.

Example 5.2. Changing the coefficients of the Δ_i slightly can change the answer to yes. Consider

$$\begin{aligned}\Delta f(x) &= \frac{5}{8}\Delta_0 f(x) - \frac{1}{8}\Delta_1 f(x) + \frac{4}{8}\Delta_2 f(x) \\ &= -\frac{5}{8}f(x) + \frac{16}{8}f(x+h) - \frac{22}{8}f(x+2h) \\ &\quad + \frac{20}{8}f(x+3h) - \frac{13}{8}f(x+4h) + \frac{4}{8}f(x+5h).\end{aligned}$$

Now $D^{[1]}$ takes on the values $-\frac{5}{8}, \frac{11}{8}, -\frac{11}{8}, \frac{9}{8}, -\frac{4}{8}$ on the intervals $(i, i+1)$, $0 \leq i \leq 4$. It is then easy to determine that $D^{[3]}$ does not change sign, so Theorem 4.1 answers the question in the positive. Again, we graph the functions $D^{[1]}$, $D^{[2]}$, and $D^{[3]}$ with $h = 1$ in Figure 1, but this time the graph of $D^{[3]}$ stays below the x -axis.

Example 5.3. This example illustrates the use of Theorem 4.2. It is easy to check that the difference

$$\Delta f(x) = \frac{4}{15}(-f(x) + f(x+h) + f(x+\frac{3}{2}h) - f(x+\frac{7}{2}h) - f(x+4h) + f(x+5h))$$

has degree three and is normalized. Also, the sequence of partial sums of $\sum a_i$ is $-\frac{4}{15}, 0, \frac{4}{15}, 0, -\frac{4}{15}, 0$, which has two sign changes, so $c_1 = 2$. But $c_1 > c_2 > c_3$, so $c_3 = 0$, and Δ possesses the mean value property by Theorem 4.1.

REFERENCES

1. J.M. Ash and R.L. Jones, *Mean Value Theorems for Generalized Riemann Derivatives*, Proc. Amer. Math. Soc., vol. 101, no. 2, October, 1987. MR902539 (88i:26011)
2. J.M. Ash, A. E. Gatto, and S. Vági, *A multidimensional Taylor's Theorem with minimal hypotheses*, Colloq. Math., **60-61** (1990), 245-252. MR1096374 (92b:26001)
3. Russell A. Gordon, *The Integrals of Lebesgue, Denjoy, Perron, and Henstock*, Graduate Studies in Mathematics, vol. 4, American Mathematical Society, 1994. MR1288751 (95m:26010)
4. E. Isaacson and H. B. Keller, *Analysis of Numerical Methods*, Wiley, New York, 1966. MR0201039 (34:924)

DEPARTMENT OF MATHEMATICS, CALIFORNIA STATE UNIVERSITY, SAN BERNARDINO, CALIFORNIA 92407

E-mail address: hfejzic@csusb.edu

DEPARTMENT OF MATHEMATICS, CALIFORNIA STATE UNIVERSITY, SAN BERNARDINO, CALIFORNIA 92407

E-mail address: cfreilin@csusb.edu

DEPARTMENT OF MATHEMATICS, CALIFORNIA STATE UNIVERSITY, SAN BERNARDINO, CALIFORNIA 92407

E-mail address: drinne@csusb.edu