

GENERALIZED VS. ORDINARY DIFFERENTIATION

J. MARSHALL ASH, STEFAN CATOIU, AND MARIANNA CSÖRNYEI

(Communicated by Alexander Iosevich)

ABSTRACT. There are difference quotients distinct from $\frac{f(x+h) - f(x)}{h}$ which characterize differentiability. We find all of them.

1. INTRODUCTION

1.1. Motivation. Taylor's basic Calculus theorem shows that if a real function f is n times differentiable at x , then it is well approximated about x by its n th Taylor polynomial,

$$(1.1) \quad f(x+h) = f_0(x) + f_1(x)h + \cdots + \frac{f_n(x)}{n!}h^n + o(h^n),$$

where $f_k(x) = f^{(k)}(x)$, for $k = 0, 1, \dots, n$. If the above equation holds for any function f at x , we say that f has n Peano derivatives at x , denoted $f_1(x), \dots, f_n(x)$. The n th Riemann difference of f is the difference

$$D_n f(x, h) = \sum_{k=0}^n (-1)^k \binom{n}{k} f(x + (n-k)h),$$

and the n th symmetric Riemann difference of f is

$$D_n^s f(x, h) = \sum_{k=0}^n (-1)^k \binom{n}{k} f\left(x + \left(\frac{n}{2} - k\right)h\right).$$

The function f is n times Riemann (resp. symmetric Riemann) differentiable at x if the limit $R_n f(x) = \lim_{h \rightarrow 0} \frac{D_n f(x, h)}{h^n}$ (resp. $R_n^s f(x) = \lim_{h \rightarrow 0} \frac{D_n^s f(x, h)}{h^n}$) exists as a finite number. It is well known that if f has n Peano derivatives at x , then it is n times Riemann and symmetric Riemann differentiable at x and $R_n f(x) = f_n(x)$.

Received by the editors March 17, 2016.

2010 *Mathematics Subject Classification.* Primary 26A24; Secondary 15A06, 26A27.

Key words and phrases. Derivatives, generalized derivatives, Peano derivatives, Riemann derivatives, \mathcal{A} -derivatives, generalized Riemann derivatives.

The second author's research was partially supported by a 2014 Summer Research Grant from the Faculty and Development Program at DePaul University, College of Science and Health.

Indeed, if $n = 2$, then

$$\begin{aligned} \lim_{h \rightarrow 0} \frac{D_2 f(x, h)}{h^2} &= \lim_{h \rightarrow 0} \frac{f(x) - 2f(x+h) + f(x+2h)}{h^2} \\ &= \lim_{h \rightarrow 0} \left[\frac{f_0(x) - 2(f_0(x) + f_1(x)h + \frac{f_2(x)}{2}h^2 + o(h^2))}{h^2} \right. \\ &\quad \left. + \frac{f_0(x) + f_1(x)2h + \frac{f_2(x)}{2}4h^2 + o(h^2)}{h^2} \right] = f_2(x) = f''(x). \end{aligned}$$

The converse of this is false in general. For example, the function $f(x) = |x|$ is symmetric Riemann differentiable of order one at $x = 0$, since

$$\lim_{h \rightarrow 0} \frac{f(\frac{h}{2}) - f(-\frac{h}{2})}{h} = \lim_{h \rightarrow 0} \frac{|\frac{h}{2}| - |-\frac{h}{2}|}{h} = 0,$$

while f is clearly not differentiable at zero. The major result in the area is due to [MZ]. It says that for differentiable functions f ,

$$nth \text{ Riemann} \implies nth \text{ Peano at a.e. } x.$$

Riemann derivatives were generalized in [As]. A generalized n th Riemann difference of a function f is a difference of the form

$$\Delta_{\mathcal{A}} f(x, h) = \sum_{i=1}^m A_i f(x + a_i h),$$

where $\mathcal{A} = \{A_1, \dots, A_m; a_1, \dots, a_m\}$ is a set of $2m$ parameters, the second half of which are distinct, that satisfy the Vandermonde conditions $\sum_{i=1}^m A_i a_i^j = n! \cdot \delta_{jn}$, for $j = 0, 1, \dots, n$. This implies that m must be at least $n + 1$. In other words, the excess number $e = m - (n + 1)$ is non-negative. Some interesting examples of \mathcal{A} -derivatives with positive excess appear in numerical analysis; see [AJ, AJJ]. The \mathcal{A} -derivative of f is defined by the limit

$$D_{\mathcal{A}}(x) = \lim_{h \rightarrow 0} \frac{\Delta_{\mathcal{A}} f(x, h)}{h^n}.$$

Most of the results for classical Riemann derivatives hold true for \mathcal{A} -derivatives of differentiable functions f . For example, the pointwise implication

$$(1.2) \quad nth \text{ Peano derivative exists at } x \implies nth \text{ } \mathcal{A}\text{-derivative exists at } x$$

has this generic converse

$$(1.3) \quad \begin{array}{ccc} nth \text{ } \mathcal{A}\text{-derivative exists at} & \implies & nth \text{ Peano derivative exists at} \\ \text{every } x \in E & & \text{a.e. } x \in E. \end{array}$$

This converse implication is, in general, not true pointwise. The example of $f(x) = |x|$ at $x = 0$ mentioned above shows this, since the first symmetric derivative is an \mathcal{A} -derivative with $\mathcal{A} = \{1, -1; \frac{1}{2}, -\frac{1}{2}\}$.

1.2. Results. The goal of this article is to classify all n th generalized Riemann differences $\Delta_{\mathcal{A}}$ into those for which the pointwise implication

$$(1.4) \quad nth \text{ } \mathcal{A}\text{-derivative exists at } x \implies nth \text{ Peano derivative exists at } x$$

is always true and those for which it is sometimes false. We began this study to justify this conjecture made by one of us in a survey talk, [As1], in the fall of 2012.

“I have not worked through all the cases of the irreversibility of the implication (1.2), but I have no doubt that every one is irreversible.”

Obviously, this conjecture excluded the trivial $n = 1$ case of $\mathcal{A} = \{a^{-1}, a^{-1}; a, 0\}$ with $a \neq 0$, since these \mathcal{A} -derivatives, the ordinary first derivative, and the first Peano derivative all have the same definition. In fact,

$$\begin{aligned} D_{\mathcal{A}}(x) &= \lim_{h \rightarrow 0} \frac{a^{-1}f(x+ah) - a^{-1}f(x)}{h} = \lim_{h \rightarrow 0} \frac{f(x+ah) - f(x)}{ah} \\ &= \lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h} = f'(x) = f_1(x). \end{aligned}$$

One of the first cases we looked at was $\mathcal{A} = \{2, 1, -3; 1, -1, 0\}$. For concrete examples, we favor the more intuitive compact notation of writing out the difference quotient for the essentially general case of $x = 0$. So our assumption is that there is a number d so that as $h \rightarrow 0$,

$$(1.5) \quad \frac{2f(h) + f(-h) - 3f(0)}{h} \rightarrow d.$$

Replacing h by $-h$ yields

$$(1.6) \quad \frac{-2f(-h) - f(h) + 3f(0)}{h} \rightarrow d,$$

and addition of (1.5) multiplied by $2/3$ to (1.6) multiplied by $1/3$ produces

$$\begin{aligned} &\frac{\left(\frac{2}{3} \cdot 2 + \frac{1}{3} \cdot -1\right) f(h) + \left(\frac{2}{3} \cdot 1 + \frac{1}{3} \cdot -2\right) f(-h) + \left(\frac{2}{3} \cdot -3 + \frac{1}{3} \cdot 3\right) f(0)}{h} \\ &= \frac{f(h) - f(0)}{h} \rightarrow \frac{2}{3}d + \frac{1}{3}d = d. \end{aligned}$$

Thus this particular \mathcal{A} -derivative’s existence does imply ordinary classical differentiability! So the conjecture was false.

We have found the following complete classification of \mathcal{A} -derivatives.

Theorem 1. **A:** *The first order \mathcal{A} -derivatives which are dilates ($h \rightarrow sh$, for some $s \neq 0$) of*

$$(1.7) \quad \lim_{h \rightarrow 0} \frac{Af(x+rh) + Af(x-rh) + f(x+h) - f(x-h) - 2Af(x)}{2h},$$

where $Ar \neq 0$ are equivalent to ordinary differentiation.

B: *Given any other \mathcal{A} -derivative of any order $n = 1, 2, \dots$, there is a measurable function $f(x)$ such that $D_{\mathcal{A}}(0)$ exists, but the Peano derivative $f_n(0)$ does not.*

Notice that the \mathcal{A} -derivatives shown in equation (1.7) have parameters

$$(1.8) \quad \mathcal{A} = \left\{ \frac{A}{2}, \frac{A}{2}, \frac{1}{2}, \frac{-1}{2}, -A; r, -r, 1, -1, 0 \right\}.$$

Also notice that the first generalized derivative given by A and r is a five point derivative ($m = 5$) when $|r| \neq 1$, a three point derivative when $|r| = 1$ but $|A| \neq 1$, and ordinary differentiation (with $m = 2$) when $|r| = |A| = 1$. When $r = 1$,

changing from the parameter A to the parameter $\alpha = (A + 1) / 2$ gives the resultant subfamily of difference quotients the forms

$$(1.9) \quad \lim_{h \rightarrow 0} \frac{\alpha f(x+h) + (\alpha - 1) f(x-h) - (2\alpha - 1) f(x)}{h}.$$

In particular, the original example presented in (1.5) above can be considered either as the $m = 3$ case $\alpha = 2$, or as the degenerate $m = 5$ case $r = 1, A = 3$.

2. THREE EXAMPLES

Before proceeding with the proof of Theorem 1, we will give three very simple examples which together cover many cases of the theorem.

Example 1. Every \mathcal{A} -derivative with zero $f(x)$ -term, that is to say with none of $\{a_1, \dots, a_m\}$ equal to zero, differentiates the characteristic function of the non-zero real numbers, $\chi(x) = \begin{cases} 1 & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}$ at $x = 0$, since $\Delta_{\mathcal{A}}$ at $x = 0$ cannot distinguish χ from the constant function which is equal to 1 for every real x . However, χ has no Peano derivative at $x = 0$.

The second example has exactly the same property, but is a little less trivial. Let $\mathcal{S} = \mathbf{Q}(a_1, \dots, a_m)$ be the subfield of \mathbf{R} generated over \mathbf{Q} by the parameters a_1, \dots, a_m . This is a denumerable subset of \mathbf{R} .

Example 2. Suppose $a_i \neq 0$, for all $i = 1, 2, \dots, m$, and let f be the Lebesgue measurable function

$$f(t) = \begin{cases} 1 & \text{if } t \notin \mathcal{S}, \\ 0 & \text{if } t \in \mathcal{S}. \end{cases}$$

Clearly, \mathcal{S} denumerable and containing \mathbf{Q} implies f discontinuous at 0, hence it has no Peano derivatives at zero of any order. On the other hand, for each i , we have $h \in \mathcal{S}$ if and only if $a_i h \in \mathcal{S}$, hence

$$\Delta_{\mathcal{A}} f(0, h) := \sum_{i=1}^m A_i f(a_i h) = \begin{cases} \sum_{i=1}^m A_i \cdot 1 & \text{if } h \notin \mathcal{S}, \\ \sum_{i=1}^m A_i \cdot 0 & \text{if } h \in \mathcal{S}. \end{cases}$$

If $n \geq 1$, then the Vandermonde condition of \mathcal{A} for $j = 0$, namely $\sum A_i = \sum A_i a_i^0 = 0$, makes $\Delta_{\mathcal{A}} f(0, h)$ identically equal to zero. Thus f is \mathcal{A} -differentiable at zero, for each \mathcal{A} (of order $n \geq 1$), and $D_{\mathcal{A}}(0) = 0$.

The third example is \mathcal{A} -differentiable for every \mathcal{A} of order $n \geq 2$, but has no Peano derivative of any order ≥ 1 .

Example 3. Here we allow some $a_i = 0$. Let g be the function $g(t) = tf(t)$

$$g(t) = tf(t) = \begin{cases} t & \text{if } t \notin \mathcal{S}, \\ 0 & \text{if } t \in \mathcal{S}, \end{cases}$$

for the same f as in Example 2. Then g is continuous at zero, and has no Peano derivatives of order ≥ 1 . Moreover,

$$\Delta_{\mathcal{A}} g(0, h) = \sum_{i=1}^m A_i g(a_i h) = \begin{cases} \sum_{i=1}^m A_i a_i h & \text{if } h \notin \mathcal{S}, \\ \sum_{i=1}^m A_i 0 & \text{if } h \in \mathcal{S}. \end{cases}$$

This is identically zero by the Vandermonde condition of \mathcal{A} for $j = 1$, provided $n \geq 2$. Then g is \mathcal{A} -differentiable at zero for any \mathcal{A} with $n \geq 2$, and $D_{\mathcal{A}}(0) = 0$.

3. PROOF OF PART A OF THEOREM 1

The proof of the (unmotivated) Part A of Theorem 1 is just about immediate. The motivation will come later.

Proof of Part A of Theorem 1. Let non-zero A and r be given and suppose that f has a derivative of the form (1.7) equal to d at $x = 0$. Let $g(h) = f(h) - f(0)$ and notice that as $h \rightarrow 0$,

$$\frac{Ag(rh) + Ag(-rh) + g(h) - g(-h)}{2h} \rightarrow d.$$

Successively replace h by $-h$, $r^{-1}h$ and $-r^{-1}h$ to see that additionally

$$\frac{-Ag(-rh) - Ag(rh) - g(-h) + g(h)}{2h} \rightarrow d,$$

$$\frac{Ag(h) + Ag(-h) + g(r^{-1}h) - g(-r^{-1}h)}{2r^{-1}h} \rightarrow d,$$

and

$$\frac{-Ag(-h) - Ag(h) - g(-r^{-1}h) + g(r^{-1}h)}{2r^{-1}h} \rightarrow d.$$

Now multiply these relations respectively by $\frac{1}{2}$, $\frac{1}{2}$, $\frac{1}{2Ar}$, and $\frac{-1}{2Ar}$ and add. We get

$$\begin{aligned} & \left(\frac{1}{2} \frac{A}{2} + \frac{1}{2} \frac{-A}{2} \right) \frac{g(rh)}{h} \\ & + \left(\frac{1}{2} \frac{A}{2} + \frac{1}{2} \frac{-A}{2} \right) \frac{g(-rh)}{h} \\ & + \left(\frac{1}{2} \frac{1}{2} + \frac{1}{2} \frac{1}{2} + \frac{1}{2Ar} \frac{A}{2r^{-1}} + \frac{-1}{2Ar} \frac{-A}{2r^{-1}} \right) \frac{g(h)}{h} \\ & + \left(\frac{1}{2} \frac{-1}{2} + \frac{1}{2} \frac{-1}{2} + \frac{1}{2Ar} \frac{A}{2r^{-1}} + \frac{-1}{2Ar} \frac{-A}{2r^{-1}} \right) \frac{g(-h)}{h} \\ & + \left(\frac{1}{2Ar} \frac{1}{2r^{-1}} + \frac{-1}{2Ar} \frac{1}{2r^{-1}} \right) \frac{g(r^{-1}h)}{h} \\ & + \left(\frac{1}{2Ar} \frac{-1}{2r^{-1}} + \frac{-1}{2Ar} \frac{-1}{2r^{-1}} \right) \frac{g(-r^{-1}h)}{h} \\ \rightarrow & \left(\frac{1}{2} + \frac{1}{2} + \frac{1}{2Ar} + \frac{-1}{2Ar} \right) d. \end{aligned}$$

Thus $\frac{f(0+h)-f(0)}{h} = \frac{g(h)}{h} \rightarrow d$ so that Part A of Theorem 1 is proved. □

4. RESTRICTED PROOF OF PART B OF THEOREM 1

The “restricted” Part B of Theorem 1 is the same result under an additional linear algebra hypothesis or restriction. This is the subject of Lemma 2. Before dealing with the restricted result, we simplify notation twice and provide the linear algebra result that made the restricted proof possible.

4.1. First simplification. Because of Example 3 given above we only need to study \mathcal{A} -derivatives of order $n = 1$. Because of the other two examples, we may as well assume that $a_m = 0$ and $A_m \neq 0$. Let f be \mathcal{A} -differentiable of order 1 at x . Without loss, we may assume $x = 0$ and write

$$(4.1) \quad D_{\mathcal{A}}f(0) = \lim_{h \rightarrow 0} \frac{\Delta_{\mathcal{A}}f(h)}{h} = d,$$

where

$$\Delta_{\mathcal{A}}f(h) = \sum_{i=1}^m A_i f(a_i h)$$

and the data \mathcal{A} satisfies the Vandermonde conditions

$$(4.2) \quad \sum_{i=1}^m A_i = 0 \quad \text{and} \quad \sum_{i=1}^m A_i a_i = 1.$$

4.2. Second simplification. Since $a_m = 0$, without loss of generality we may replace $f(x) - f(0)$ by $f(x)$. The Vandermonde conditions are now updated to

$$(4.3) \quad \sum_{i=1}^{m-1} A_i \neq 0 \quad \text{and} \quad \sum_{i=1}^{m-1} A_i a_i = 1.$$

In this reduced notation, the ordinary first derivative has $m = 2$, but, because of our reduction, is written as simply $\lim_{h \rightarrow 0} \frac{f(h)}{h}$.

4.3. A basic linear algebra result. The following lemma on infinite dimensional linear algebra will be used during the restricted proof of Part B of Theorem 1.

Lemma 1. *An infinite system of linear equations over a field has a solution if every finite subsystem has a solution.*

Proof. Let Σ be an infinite system of linear equations, all of whose finite subsystems have solutions. Say that the pair $(S, \{c_\alpha\})$ is a solved subsystem of Σ if $S \subseteq \Sigma$ and $\{c_\alpha\}$ are the field elements assigned to the variables $\{x_\alpha\}$ appearing in the equations of S . Say that $(S, \{c_\alpha\}) \preceq (T, \{d_\alpha\})$ if $c_\alpha = d_\alpha$ whenever x_α is a variable appearing in an equation of S . If $\{(S_\iota, \{c_\alpha\}_\iota)\}_{\iota \in I}$ is a chain with respect to the partially ordering defined by \preceq , then $(S, \{d_\alpha\})$ is an upper bound where $S = \bigcup_{\iota \in I} S_\iota$ and the $\{d_\alpha\}$ are defined by assigning to any equation in S the variable evaluations arising from the membership of some S_ι to which it belongs. This evaluation is well defined. By Zorn's Lemma, there is a maximal element $(M, \{e_\alpha\}_{\alpha \in J})$. Assume $M \neq \Sigma$, and let $e : \sum_{\alpha \in K} r_\alpha x_\alpha = d$ be an equation in $\Sigma \setminus M$. By maximality, $M^* = M \cup \{e\}$ has no solution. If $K \not\subseteq J$ letting $x_\alpha = e_\alpha$ for $\alpha \in J$ and selecting the remaining x_β appropriately would solve M^* . So $K \subseteq J$. If $\alpha \in K$ is such that the participation of x_α in the finite subsets of M^* forces x_α to have the unique value e_α , then x_α will take on the value e_α in evaluating e also. Finally, if x_α is not forced to have a unique value despite the participation of x_α in the finite subsets of M^* , then, as is shown in [CE], one can change the value of x_α to 0 without spoiling the solving of any equation in M^* . This solving of M^* is a contradiction. \square

Lemma 1 is proved in references [Ab] and [CE]. The proof in [Ab] cleverly shows the lemma to be a corollary of the well known fact that any linearly independent subset of a vector space can be extended to a basis. The proof in [CE] is by transfinite induction and makes use of the argument mentioned at the end of our proof.

4.4. The restricted Part B of Theorem 1 and its proof. In Theorem 1, Part B, we are given an \mathcal{A} -derivative $\{A_1, \dots, A_m; a_1, \dots, a_m\}$ and our goal is to create a function $f : \mathbb{R} \rightarrow \mathbb{R}$ whose values $\{f(h)\}$ must satisfy an infinite system of equations S ; one of these equations is $\sum_{i=1}^m A_i f(a_i) = 0$. These equations will be so numerous that for every real number h , $f(h)$ will appear in at least one of the equations. To create f , we will first assign to each real number h a variable x_h , and then think of the set $\{x_h\}_{h \in \mathbb{R}}$ as a subset of an infinite dimensional real vector space. We will next substitute x_h for $f(h)$ throughout the system S creating an associated system of linear equations S' ; one of these linear equations will be $\sum_{i=1}^m A_i x_{a_i} = 0$. Next we will show that the hypothesis of Theorem 1, Part B guarantees that there is a solution for the linear system S' . Denote one such solution as $\{x_h = c_h\}_{h \in \mathbb{R}}$, where each c_h is a real number; in particular, $\sum_{i=1}^m A_i c_{a_i} = 0$. Finally, we will create the required function f by setting $f(h) = c_h$ for every real number h .

Lemma 2. *Suppose that no finite linear combination of dilates of $\sum_{i=1}^{m-1} A_i x_{a_i}$ sums to x_h for any h . Then there is a measurable function f such that $D_{\mathcal{A}}(0) = 0$, but $f'(0)$ does not exist.*

Proof. Let $s_n \rightarrow 0$, where $\{s_n\}$ is an algebraically independent set of numbers over the field $K = \mathbb{Q}(a_1, \dots, a_{m-1})$. For each $n = 1, 2, \dots$, let $S_n = \{p s_n : p \in K \setminus \{0\}\}$. Decompose the uncountable system of equations

$$(4.4) \quad \begin{aligned} \sum_{i=1}^{m-1} A_i x_{a_i h} &= 0, \quad h \in \mathbb{R}, \\ x_{s_n} &= 1, \quad n = 1, 2, \dots, \end{aligned}$$

into the following countable set of subsystems.

$$\begin{aligned} \sum_{i=1}^{m-1} A_i x_{a_i h} &= 0, \quad h \in \mathbb{R} \setminus \bigcup_{n=1}^{\infty} S_n, \\ (*)_n &= \begin{cases} \sum_{i=1}^{m-1} A_i x_{a_i h} = 0, & h \in S_n \\ x_{s_n} = 1 \end{cases}, \quad n = 1, 2, \dots \end{aligned}$$

Solve the first homogeneous subsystem with the trivial solution. In other words, set $x_{a_i h} = 0$ for every $h \notin \bigcup_{n=1}^{\infty} S_n$. Fix a positive integer n . The system $(*)_n$ has a solution: By Lemma 1, it is enough to prove that every finite subsystem has a solution. Take a finite subsystem. Without loss of generality, assume that the subsystem includes the equation $x_{s_n} = 1$. The only possible impediment to a solution is if there can be deduced a contradiction of the form $0 = c$ where c is a non-zero constant. But this could only happen if, for some n , x_{s_n} were a finite linear combination of the homogeneous equations, thereby providing $x_{s_n} = 0$, contrary to the hypothesis of this lemma.

Let $f(h)$ be the function that was created by solving the system (4.4). Then $\{k : f(k) \neq 0\}$ has Lebesgue measure 0 because it is a subset of the countable set $\bigcup_{i=1}^{m-1} \bigcup_{n=1}^{\infty} a_i S_n$ so that f is a measurable function. In particular, because of the non-equality part of condition (4.3), $\sum A_i x_{a_i} 0 = (\sum A_i) x_0 = 0$ implies $f(0) = x_0 = 0$. On the one hand, $D_{\mathcal{A}}(0) = \lim_{h \rightarrow 0} \frac{\Delta_{\mathcal{A}} f(h)}{h} = \lim_{h \rightarrow 0} \frac{0}{h} = 0$; on the other hand, $\lim_{n \rightarrow \infty} \frac{f(s_n)}{s_n} = \lim_{n \rightarrow \infty} \frac{1}{s_n}$ is not zero, so $f'(0)$ does not exist. \square

5. UNRESTRICTED PROOF OF PART B OF THEOREM 1

Look at the proof of Part A of Theorem 1 above. Each \mathcal{A} -derivative listed there was shown to imply ordinary differentiation in just about the simplest imaginable way: a linear combination of four of its dilates was shown to add to ordinary differentiation. It seems to be logically possible to prove the implication true for some other \mathcal{A} -derivative in another, quite possibly very indirect, way. The meaning of Lemma 2 is that this is not the case. Because of Lemma 2, as soon as an \mathcal{A} -derivative is known to not have x_h in the span of its dilates, that \mathcal{A} -derivative's existence at $x = 0$ is known not to imply the existence of the ordinary derivative. So the plan for our proof is to assume that the hypothesis of Lemma 2 is not satisfied, that is, to assume x_h is a linear combination

$$(5.1) \quad \sum_{j=1}^n L_j \left(\sum_{i=1}^{m-1} A_i x_{a_i r_j h} \right) = x_h$$

of r_j -dilates of $\Delta_{\mathcal{A}} = \sum_{i=1}^{m-1} A_i x_{a_i h}$, and to infer that $\Delta_{\mathcal{A}}$ has the form shown in relation (1.7) above, namely for an appropriate dilation s and numbers A and r with $Ar \neq 0$,

$$\frac{\Delta_{\mathcal{A}}(sh)}{sh} = \frac{Af(rh) + Af(-rh) + f(h) - f(-h)}{2h}.$$

Replacing $\Delta_{\mathcal{A}}$ by one of the dilates in equation (5.1) for which some $a_i r_j = 1$ if necessary, we may assume that some $a_i = 1$.

5.1. Symmetrization of the linear system. We adapt a more symmetrical notation for $\Delta_{\mathcal{A}}$. Write $\sum A_i x_{a_i h}$ as

$$\sum_{i=1}^k A_i x_{t_i} + B_i y_{t_i}$$

where for convenience we take $h > 0$ and set $t_i = |a_i| h$, $x_{t_i} = f(t_i)$, $y_{t_i} = f(-t_i)$, and for each i , at least one of A_i and B_i is non-zero. Then $0 < t_1 < t_2 < \dots < t_k$ and some $t_i = h$. The Vandermonde conditions (4.3) are now

$$(5.2) \quad \sum_{i=1}^k A_i + B_i \neq 0 \quad \text{and} \quad \sum_{i=1}^k (A_i - B_i) a_i = 1.$$

If we dilate by $r_j > 0$, the inner summand in equation (5.1) becomes

$$\sum_{i=1}^k A_i x_{t_i r_j} + \sum_{i=1}^k B_i y_{t_i r_j}$$

whereas if we dilate by $-r_j$, the inner summand becomes

$$\sum_{i=1}^k A_i y_{t_i r_j} + \sum_{i=1}^k B_i x_{t_i r_j}.$$

These two summands involve the same variables, so to look at all possible ways to attain x_h in equation (5.1) we must add, say a λ_j -multiple of the former and a

μ_j -multiple of the latter. Equation (5.1) will be written as

$$(5.3) \quad \sum_{j=1}^n \sum_{i=1}^k ((\lambda_j A_i + \mu_j B_i) x_{t_i r_j} + (\lambda_j B_i + \mu_j A_i) y_{t_i r_j}) = x_h$$

and we now may assume that all r_j are positive and ordered as $0 < r_1 < \dots < r_n$.

5.2. Reduction to five cases. Call an i *small* (or *large*) if there is an r_j such that the product $t_i r_j$ is smaller (or larger) than h , respectively. Similarly, call a j small (or large) if there is a t_i such that $t_i r_j$ is smaller (or larger) than h . We have to consider several cases:

Case 1. There is a j which is neither small nor large.

Case 2. There is an i which is neither small nor large.

After treating Cases 1-2, for Cases 3-5 we assume (*) each i is either small or large and each j is either small or large.

Case 3. All i and j are small, or, all i and j are large.

After treating Cases 1-3, for Cases 4-5 we assume (*) and also the negation of Case 3: (**) there is an i or a j that is not small and there is an i or a j that is not large. This can be rephrased as “there are some small i, j and there are some large i, j ”.

Case 4. There is a j (or an i) which is both small and large.

After treating Cases 1-4, we assume (*), (**), and the negation of Case 4. What remains is Case 5.

Case 5. Each i , and each j is either small or large, but not both. There are some small i, j and there are some large i, j .

5.3. Proofs of the five cases. We shall see that Cases 1 and 5 lead to particular versions of the derivative (1.7), while Cases 2, 3 and 4 are impossible.

Case 1. There is a j which is neither small nor large.

Proof. The hypothesis implies that for every $i, t_i r_j = h$. So there is only one t_i and $\Delta_{\mathcal{A}} = \alpha f(h) + \beta f(-h)$. By (4.3), $\alpha + \beta \neq 0$ and $\alpha - \beta = 1$ and we have arrived at a special case of derivative (1.7), namely derivative (1.9) where $\beta = \alpha - 1$. \square

Case 2. There is an i which is neither small nor large.

Proof. For every j , we have $t_i r_j = h$. In particular, there is only one r_j . Then x_h must be a linear combination of a single $\Delta_{\mathcal{A}}$ at h and the same $\Delta_{\mathcal{A}}$ at $-h$. Say

$$\sum_{i=1}^k ((\lambda A_i + \mu B_i) x_{t_i} + ((\lambda B_i + \mu A_i)) y_{t_i}) = x_h.$$

If $k = 1$, this is exactly the same situation as in Case 1. But we must have $k = 1$, for if $k \geq 2$, there is an i so that $t_i = h$ yielding the equations

$$\begin{cases} \lambda A_i + \mu B_i = 1 \\ \lambda B_i + \mu A_i = 0 \end{cases}, \text{ or } \begin{pmatrix} \lambda & \mu \\ \mu & \lambda \end{pmatrix} \begin{pmatrix} A_i \\ B_i \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix},$$

so that $\lambda^2 - \mu^2 \neq 0$; as well as an $i' \neq i$, yielding

$$\left\{ \begin{array}{l} \lambda A_{i'} + \mu B_{i'} = 0 \\ \lambda B_{i'} + \mu A_{i'} = 0 \end{array} \right. \text{ or } \begin{pmatrix} \lambda & \mu \\ \mu & \lambda \end{pmatrix} \begin{pmatrix} A_{i'} \\ B_{i'} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

so that $\lambda^2 - \mu^2 = 0$, a contradiction. □

Case 3. All i and j are small, or, all i and j are large.

Proof. We will suppose that all i and j are small, the other case being similar. Since $i = k$ is small, there is a j so that $r_j t_k < h$. Similarly, there is an i so that $r_n t_i < h$. Because $r_1 t_1 \leq r_j t_k < h$, from equation (5.3) we have

$$\begin{pmatrix} \lambda_1 & \mu_1 \\ \mu_1 & \lambda_1 \end{pmatrix} \begin{pmatrix} A_1 \\ B_1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

so that $\lambda_1^2 - \mu_1^2 = 0$ and $\lambda_1 = \pm \mu_1$. If $\lambda_1 = -\mu_1$, then $\lambda_1(A_1 - B_1) = 0$, so $A_1 = B_1$.

Claim. There holds $\lambda_j = -\mu_j$ for all j and $A_i = B_i$ for all i .

The equation (5.3) can be written as

$$\sum_{\eta} \left\{ \sum_{t_i r_j = \eta} (\lambda_j A_i + \mu_j B_i) x_{t_i r_j} + (\lambda_j B_i + \mu_j A_i) y_{t_i r_j} \right\} = x_h$$

where the inner sums are all 0 when $\eta \neq h$. When $\eta = h$ we must have the inner sum equal to x_h . However we will soon see that the inner sums being 0 whenever $\eta < h$ will validate the claim and the claim will contradict the inner sum being non-zero when $\eta = h$. This means that Case 3 cannot occur.

Say that i is associated with η if there is a j so that $t_i r_j = \eta$ and that j is associated with η if there is an i with $t_i r_j = \eta$. We induct on η to prove that for all $\eta < h$, (*) whenever i (resp. j) is associated with η , $A_i = B_i$ (resp. $\lambda_j = -\mu_j$). For the smallest $\eta = r_1 t_1$, (*) is true for both $j = 1$ and $i = 1$. Assume (*) has been proved for all $\eta' < \eta$. The coefficients of x_{η} and y_{η} are

$$\sum_{t_i r_j = \eta} \lambda_j A_i + \mu_j B_i \quad \text{and} \quad \sum_{t_i r_j = \eta} \lambda_j B_i + \mu_j A_i.$$

The corresponding equations are

$$(5.4) \quad \sum_{t_i r_j = \eta} \lambda_j A_i + \mu_j B_i = 0 \quad \text{and} \quad \sum_{t_i r_j = \eta} \lambda_j B_i + \mu_j A_i = 0.$$

At most one r_j that has not appeared in lower indices may appear here multiplied by t_1 . Then by the inductive hypothesis, we have

$$\begin{aligned} (\lambda_J + \mu_J) A_1 + \sum_{t_i r_j = \eta, j \neq J} \lambda_j A_i + \mu_j B_i &= 0, \\ (\lambda_J + \mu_J) A_1 - \sum_{t_i r_j = \eta, j \neq J} \mu_j B_i + \lambda_j A_i &= 0. \end{aligned}$$

Add these two equations and then note that $(A_1, B_1) \neq (0, 0)$ and $A_1 = B_1$ implies $A_1 \neq 0$, so

$$2(\lambda_J + \mu_J) A_1 = 0 \quad \text{and} \quad \lambda_J = -\mu_J.$$

Also, at most one new t_I may appear and if it does, it will be multiplied by r_1 . In this case the equations (5.4) become

$$\begin{aligned} \lambda_1(A_I - B_I) + \sum_{t_i r_j = \eta, i \neq I} \lambda_j A_i + \mu_j B_i &= 0, \\ \lambda_1(B_I - A_I) + \sum_{t_i r_j = \eta, i \neq I} \lambda_j A_i + \mu_j B_i &= 0. \end{aligned}$$

This time subtract to get

$$2\lambda_1(A_I - B_I) = 0 \quad \text{and} \quad A_I = B_I.$$

This proves (*). The claim follows, since every i is small means that every i is associated with some $\eta < h$, and every j small means that also every j is associated with some $\eta < h$. Finally, because of the claim, every term in curly brackets is 0, even when $\eta = h$, when that term is supposed to be x_h .

Similarly, the case when $\lambda_1 = \mu_1$ and consequently $A_1 = -B_1$ leads to the fact that then there holds $\lambda_j = \mu_j$ for all j and $A_i = -B_i$ for all i . This, in turn, leads to the same contradiction as in the first case. □

Case 4. There is a j (or an i) which is both small and large.

Proof. Suppose there is a j_0 which is both small and large; the other subcase is similar. Then there are t' and t'' so that $r_1 t_1 \leq r_{j_0} t' < h < r_{j_0} t'' \leq r_n t_k$. As in the proof of Case 3, we must have either

- (1_{low}): $\lambda_j = -\mu_j$ and $A_i = B_i$ for all (j, i) such that $r_j t_i < h$ or
- (2_{low}): $\lambda_j = \mu_j$ and $A_i = -B_i$ for all (j, i) such that $r_j t_i < h$.

Also we must have

- (1_{high}): $\lambda_j = -\mu_j$ and $A_i = B_i$ for all (j, i) such that $r_j t_i > h$ or
- (2_{high}): $\lambda_j = \mu_j$ and $A_i = -B_i$ for all (j, i) such that $r_j t_i > h$.

Conditions (1_{low}) and (2_{high}) are incompatible, since $r_{j_0} t' < h$ implies $\lambda_{j_0} = -\mu_{j_0}$, while $h < r_{j_0} t''$ implies $\lambda_{j_0} = \mu_{j_0}$, a contradiction since $(\lambda_j, \mu_j) \neq (0, 0)$ for all j . Similarly, conditions (1_{high}) and (2_{low}) are incompatible. So we may assume either

- (1): $\lambda_j = -\mu_j$ and $A_i = B_i$ for all (j, i) such that $r_j t_i \neq h$ or
- (2): $\lambda_j = \mu_j$ and $A_i = -B_i$ for all (j, i) such that $r_j t_i \neq h$.

Now suppose $r_j t_I = h$. Since there are at least two t_i , there is at least one t for which $r_j t \neq h$; similarly there is at least one r such that $r t_I \neq h$. This shows that we may assume either

- (1'): $\lambda_j = -\mu_j$ and $A_i = B_i$ for all (j, i) or
- (2'): $\lambda_j = \mu_j$ and $A_i = -B_i$ for all (j, i) .

Just as in the proof of Case 3, condition (1') leads to a contradiction; and, similarly, condition (2') leads to a contradiction. These contradictions show that Case 4 is also impossible. □

Case 5. Each i and each j is either small or large, but not both. There are some small i, j and there are some large i, j .

Proof. We must have $r_1 t_1 < h$ and $r_n t_k > h$. Now $r_1 t_k < h$ would imply t_k small and large and $r_1 t_k > h$ would imply r_1 small and large, so $r_1 t_k = h$. Similarly, trichotomy forces $r_n t_1 = h$. If $n \geq 3$, then $r_2 t_1 < r_n t_1 = h$ and $r_2 t_k > r_1 t_k = h$ making r_2 small and large. So n must be 2. Similarly, $k \geq 3$ is impossible. Thus, we have only two t_i 's and two r_j 's and they satisfy $r_1 t_2 = r_2 t_1 = h$. This means that $\{t_i\}$ consists of h and one other number which we will denote as rh with $r \neq 1$. The only possible pair of distinct positive dilates that can lead to a solvable system of four equations is $\{1, r^{-1}\}$. With $t = r^{-1}h$, our set of differences becomes:

$$\begin{aligned} Af(rh) + Bf(-rh) + Cf(h) + Df(-h), \\ Bf(rh) + Af(-rh) + Df(h) + Cf(-h), \\ Af(h) + Bf(-h) + C(t) + Df(-t), \\ Bf(h) + Af(-h) + Df(t) + Cf(-t). \end{aligned}$$

We must determine all possible coefficient choices of $\{A, B, C, D\}$ so that $f(h)$ is a linear combination of these differences. Since the $f(rh)$ and $f(-rh)$ terms must both sum to 0, $A = \pm B$. Since the $f(t)$ and $f(-t)$ terms must both sum to 0, $C = \pm D$. But $(A, C) = (-B, -D)$ contradicts the first part of (4.3), so either $A = B$ or $C = D$ must hold.

If $A = B$, then the second part of (4.3) applied to the first difference forces $C = -D = 1/2$ and we have uncovered the derivatives (1.7) of Part A of Theorem 1. Similarly, if $C = D$, apply (4.3) to the third difference to deduce $A = -B = 1/2$ and we reach the same derivatives again. \square

Notice that our proof of Part B of Theorem 1 has discovered the small family of generalized derivatives that are equivalent to ordinary differentiation in a natural way.

NOTE ADDED IN PROOF

This paper was originally submitted to the PAMS on 7/10/14. Since then there has been (1) a simple application of our method that produces new characterizations of continuity ([AAC]), (2) an extension of this paper to real functions ([ACC]), and (3) an extension of this paper to the complex domain ([ACC1]). Paper [AAC] solves a related problem. The proof of the main theorem there gives insight into the proof of the main theorem here.

Say that two generalized Riemann derivatives are equivalent if the existence of either one at a point implies the existence of the other at that point. Thus this paper determines the equivalence class of the ordinary first derivative. Paper [ACC] determines for each positive integer n , all equivalence classes for all generalized Riemann derivatives of order n when the functions are real valued functions of a real variable. Paper [ACC1] does the same thing for complex valued functions of a complex variable. Here the equivalence classes are much different.

REFERENCES

- [Ab] Alexander Abian, *Solvability of infinite systems of linear equations*, Arch. Math. (Brno) **12** (1976), no. 1, 43–44. MR0424845
- [AAC] Arlene Ash, J. Marshall Ash and Stefan Catoiu, *New definitions of continuity*, Real Anal. Exchange **40** (2014/15), no. 2, 403–420. MR3499773
- [As] J. Marshall Ash, *Generalizations of the Riemann derivative*, Trans. Amer. Math. Soc. **126** (1967), 181–199. MR0204583

- [As1] J. Marshall Ash, *Remarks on various generalized derivatives*, Special functions, partial differential equations, and harmonic analysis, Springer Proc. Math. Stat., vol. 108, Springer, Cham, 2014, pp. 25–39, DOI 10.1007/978-3-319-10545-1.5. MR3297652
- [ACC] J. Marshall Ash, Stefan Catoiu and William Chin, *The classification of generalized Riemann derivatives*, preprint.
- [ACC1] J. Marshall Ash, Stefan Catoiu and William Chin, *The classification of complex generalized Riemann derivatives*, in preparation.
- [AJ] J. Marshall Ash and Roger L. Jones, *Optimal numerical differentiation using three function evaluations*, Math. Comp. **37** (1981), no. 155, 159–167, DOI 10.2307/2007508. MR616368
- [AJJ] J. Marshall Ash, S. Janson, and R. L. Jones, *Optimal numerical differentiation using N function evaluations*, Calcolo **21** (1984), no. 2, 151–169, DOI 10.1007/BF02575911. MR799618
- [CE] Robert Cowen and William Emerson, *A compactness theorem for linear equations*, Studia Logica **57** (1996), no. 2-3, 355–357, DOI 10.1007/BF00370839. MR1416740
- [MZ] J. Marcinkiewicz and A. Zygmund, *On the differentiability of functions and summability of trigonometric series*, Fund. Math. **26** (1936), 1–43.

DEPARTMENT OF MATHEMATICS, DEPAUL UNIVERSITY, CHICAGO, ILLINOIS 60614

E-mail address: mash@depaul.edu

URL: <http://www.depaul.edu/~mash/>

DEPARTMENT OF MATHEMATICS, DEPAUL UNIVERSITY, CHICAGO, ILLINOIS 60614

E-mail address: scatoiu@depaul.edu

URL: <http://www.depaul.edu/~scatoiu/>

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CHICAGO, CHICAGO, ILLINOIS 60637

E-mail address: csornyei@math.uchicago.edu