INROADS

**Real Analysis Exchange**
Vol. 37(1), 2011/2012, pp. 189–202

Russell A. Gordon, Department of Mathematics, Whitman College, Walla Walla, WA 99362, U.S.A. email: gordon@whitman.edu

# LEAST SQUARES AND APPROXIMATE DIFFERENTIATION

**Abstract**

The least squares derivative and the approximate derivative are both generalizations of the ordinary derivative. The existence of either of these generalized derivatives does not guarantee the existence of the other and it is even possible for both generalized derivatives to exist at a point but have different values. Several examples of such functions are presented in this paper. In addition, conditions for which the existence of the approximate derivative implies the existence (and equality) of the least squares derivative are stated and proved. These conditions involve the notion of Hölder continuity and a stronger version of approximate differentiability.

The least squares derivative, a relatively recent concept that will be fully defined in a moment, is a generalization of the ordinary derivative. This means that every function with an ordinary derivative at a point has a least squares derivative at that point and the values are the same and, in addition, there are functions that have a least squares derivative at a point but do not have an ordinary derivative at that point. In this same sense, the approximate derivative is also a generalization of the ordinary derivative. Since these two generalizations are radically different, there appears to be little hope for any relationship between them. However, it turns out that some positive comments can be made on this subject. In this paper, we explore a few connections between the least squares derivative and the approximate derivative as well as present some examples to illustrate these two derivatives.

Although it is possible to develop the ideas presented here in a more general context, we restrict ourselves to continuous functions. The interested reader

Mathematical Reviews subject classification: Primary: 26A24, 26A16
Key words: approximate derivative, least squares derivative, Hölder continuity
Received by the editors September 18, 2007
Communicated by: Brian S. Thomson

may wish to consider the modifications that are needed when this restriction is removed. However, in what follows, we consider a function $f$ that is continuous on a neighborhood of a point $c$.

We begin with a definition of the least squares derivative. This derivative is motivated by taking a number of points on a curve and using linear regression to find the slope of the least squares line of best fit for these points. By taking a limit of this slope as all the points move toward a given fixed point, we obtain (assuming a limit exists) a method for determining the slope of the curve at the given point. The reader interested in the full details of this process can consult [1] or [2]. The result is a somewhat intriguing definition of a derivative that involves an integral. For the record, the two integral expressions that appear in the following definition are equivalent via a simple change of variables.

**Definition 1.** *Suppose that $f$ is a continuous function defined on a neighborhood of a point $c$. The least squares derivative of $f$ at $c$, denoted by $f'_\ell(c)$, is given by*

$$f'_\ell(c) = \lim_{h\to 0^+} \frac{3}{2h^3}\int_{-h}^{h} t\, f(c+t)\, dt \quad \text{or} \quad f'_\ell(c) = \lim_{h\to 0^+} \frac{3}{2h}\int_{-1}^{1} t\, f(c+ht)\, dt,$$

*provided that the limit exists.*

It is easy to verify that the least squares derivative satisfies the usual property of linearity. This version of a derivative has its roots in a textbook written by Lanczos [6]. Lanczos was interested in estimating derivatives of functions represented only as a table of values; such representations often occur as numerical data obtained from experiments. For practical reasons, he considered the case in which there were only a small number of data points, simply mentioning as an aside that an integral appears if the number of data points increases indefinitely, thus resulting in a somewhat ironic "differentiation by integration" process. Several authors (see [1], [4], and [5]) have recently done some work with this derivative, primarily with a focus on numerical analysis and statistics. For a study of the least squares derivative from the point of view of analysis, see [2].

For the reader interested in seeing how this derivative works, two helpful functions to consider are (i) the function $f$ defined by $f(x) = x^n$, where $n$ is a positive integer, and (ii) the function $f$ defined by $f(x) = ax$ for $x \le 0$ and $f(x) = bx$ for $x > 0$, where $a$ and $b$ are distinct real numbers. For the first function, any point $c$ will do; the second function has interesting behavior at 0. In general, it is not difficult to prove that the least squares derivative exists (and has the same value) whenever the symmetric derivative exists. The proof (a simple application of L'Hôpital's Rule and the Fundamental Theorem of

Calculus) is left to the reader. Examples of functions that have a least squares derivative but not a symmetric derivative are not too difficult to construct; an example presenting a class of such functions will be presented shortly.

As an aside (and with thanks to one of the referees), we make an observation that links the least squares derivative and the symmetric derivative. With $f$ and $c$ as above, for some sufficiently small $\delta > 0$, define functions $F$ and $G$ on $[0, \delta]$ by

$$F(x) = \int_{-x}^{x} t\, f(c+t)\, dt \quad \text{and} \quad G(x) = \frac{2}{3}\, x^3.$$

For each $h \in (0, \delta)$, the Cauchy Mean Value Theorem guarantees the existence of a point $v_h \in (0, h)$ such that

$$\frac{F(h) - F(0)}{G(h) - G(0)} = \frac{F'(v_h)}{G'(v_h)}.$$

Performing the elementary computations yields

$$\frac{3}{2h^3} \int_{-h}^{h} t\, f(c+t)\, dt = \frac{f(c+v_h) - f(c-v_h)}{2v_h}.$$

In particular, letting $w_n$ be the point corresponding to $h = 1/n$, we find that

$$f'_\ell(c) = \lim_{n\to\infty} \frac{f(c+w_n) - f(c-w_n)}{2w_n},$$

assuming that $f$ has a least squares derivative at $c$. It may prove interesting to examine the size and properties of the set $\{v_h : 0 < h < \delta\}$ as well as to consider what properties of the symmetric derivative (see [7] and some of its references) carry over to the least squares derivative.

We next present a brief review of the approximate derivative. Let $E$ be a measurable set and let $c$ be a real number. The density of $E$ at $c$ is defined by (using $\mu(E)$ to denote the measure of $E$)

$$d_c E = \lim_{h\to 0^+} \frac{\mu(E \cap (c-h, c+h))}{2h},$$

provided the limit exists. It is clear that $0 \le d_c E \le 1$ when it exists. The point $c$ is a point of density of $E$ if $d_c E = 1$ and a point of dispersion of $E$ if $d_c E = 0$. Some simple properties of this concept include the following (the set $E$ is assumed to be measurable and $\mathcal{C}E$ is its complement).

1. A point $c$ is a point of density of $E$ if and only if $c$ is a point of dispersion of $\mathcal{C}E$.
2. Almost all the points of $E$ are points of density of $E$.
3. If $c$ is a point of dispersion of both $E$ and $H$, then $c$ is a point of dispersion of $E \cup H$.
4. If $c$ is a point of density of both $E$ and $H$, then $c$ is a point of density of $E \cap H$.
5. If $c$ is a point of density of $E$ and $\mu(H) = 0$, then $c$ is a point of density of $E \setminus H$.

Using this concept, we can define the approximate derivative of a function. (The assumption that $f$ be continuous is not required for this definition.)

**Definition 2.** *Suppose that $f$ is a continuous function defined on a neighborhood of a point $c$. The function $f$ is approximately differentiable at $c$ if there exists a measurable set $E$ such that $0$ is a point of density of $E$ and the limit*

$$\lim_{\substack{x \to 0 \\ x \in E}} \frac{f(c+x) - f(c)}{x}$$

*exists. The approximate derivative of $f$ at $c$ will be denoted by $f'_{ap}(c)$.*

Property (4) of points of density indicates that the number $f'_{ap}(c)$ is unique. It is immediately obvious that the approximate derivative is a generalization of the ordinary derivative. The basic idea is that the approximate derivative ignores some of the difference quotients. In spite of this reduced number of difference quotients, approximate derivatives share many of the properties of ordinary derivatives. For instance, if a function $f$ has an approximate derivative on an interval $I$, then $f'_{ap}$ is a Darboux Baire class one function on $I$; the textbook [3] is one of several sources for proofs of these facts. A literature search will reveal many other facts about the approximate derivative and some of its generalizations.

It is easy to see that the existence of the least squares derivative does not imply the existence of the approximate derivative; the absolute value function at 0 is a simple example. There are also functions that have an approximate derivative at a point but do not have a least squares derivative at that point. Furthermore, even when both derivatives exist at a point, the values are not necessarily equal. We turn to examples of these last two cases now.

Before presenting these examples, we make two computational observations, leaving their proofs to the reader. Verification of the first observation is related to the proof of the integral test for series, while the second observation follows most easily from Simpson's Rule (which gives the exact value of an integral for quadratic functions) using $n = 4$.

A. If $\lambda > 1$, then $\displaystyle\lim_{n\to\infty} n^{\lambda-1}\sum_{k=n}^{\infty}\frac{1}{k^{\lambda}} = \frac{1}{\lambda-1}$.

B. Let $c$ be the midpoint of the interval $[a,b]$ and let $v$ be a real number. Define a function $f$ on $[a,b]$ by setting $f(a) = 0 = f(b)$ and $f(c) = v$, then making $f$ linear on the intervals $[a,c]$ and $[c,b]$. To be specific, the function $f$ is defined by

$$f(x) = v\Big(1 - \frac{|x-c|}{c-a}\Big)$$

on the interval $[a,b]$. Then $\displaystyle\int_a^b tf(t)\,dt = \frac{v}{4}(b^2 - a^2)$.

**Example 3.** *There exist functions that are approximately differentiable but not least squares differentiable at a given point and there are functions which are differentiable in both senses but the values of the derivatives are unequal.*

For each positive integer $n$, let

$$a_n = \frac{1}{n^2}, \quad b_n = \frac{1}{n^2} - \frac{1}{4n^4}, \quad \text{and} \quad c_n = \frac{a_n + b_n}{2}.$$

It is easy to verify that $a_{n+1} < b_n < a_n$ for all $n$. Let $A$ be the set defined by $A = \bigcup_{k=1}^{\infty}(b_k, a_k)$ and let $E = [-1,1] \setminus A$. We claim that 0 is a point of dispersion of $A$ and thus a point of density of $E$. To see this, take any $h$ that satisfies $a_{n+1} < h \le a_n$ and note that

$$\begin{aligned}\frac{\mu\big(A\cap(-h,h)\big)}{2h} &\le \frac{\mu\big(A\cap(-a_n,a_n)\big)}{2a_{n+1}}\\ &= \frac{(n+1)^2}{2}\sum_{k=n}^{\infty}(a_k - b_k)\\ &= \frac{(n+1)^2}{8n^3}\Big(n^3\sum_{k=n}^{\infty}\frac{1}{k^4}\Big).\end{aligned}$$

Since $n$ goes to infinity as $h \to 0^+$, observation (A) tells us that the limit of the last expression in the displayed equation is $0 \cdot (1/3) = 0$, as desired.

Define a function $f$ on $[-1,1]$ by setting $f(x)=0$ for $x\in E$ and letting the graph of $f$ on each of the intervals $[b_n,a_n]$ consist of a triangular spike with peak at the ordered pair $(c_n,16c_n^r)$, where $r$ is a fixed positive number. A portion of the graph of $f$ is shown (not to scale) in Figure 1.

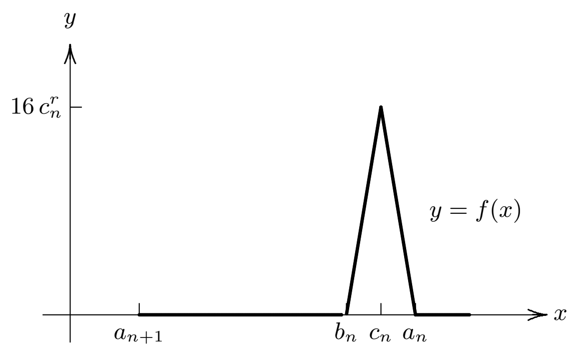


**Figure 1:** A portion of the graph of $f$.

Since 0 is a point of density of $E$, it is clear that $f'_{ap}(0)=0$, regardless of the value of $r$. For the record, note that $f$ is not differentiable at 0 (and also not symmetrically differentiable at 0) for any $r$ that satisfies $0<r\le 1$.

To compute the least squares derivative of $f$ at 0, we need to examine the limit

$$\lim_{h\to 0^+}\frac{3}{2h^3}\int_{-h}^{h} t\,f(0+t)\,dt=\lim_{h\to 0^+}\frac{3}{2h^3}\int_0^h t\,f(t)\,dt.$$

For $a_{n+1}<h\le a_n$, over and under estimates show that

$$\frac{3}{2h^3}\int_0^h tf(t)\,dt\le\frac{3}{2a_{n+1}^3}\sum_{k=n}^{\infty}\int_{b_k}^{a_k} tf(t)\,dt=\frac{3a_n^3}{2a_{n+1}^3}\Big(\frac{1}{a_n^3}\sum_{k=n}^{\infty}\int_{b_k}^{a_k} tf(t)\,dt\Big);$$

$$\frac{3}{2h^3}\int_0^h tf(t)\,dt\ge\frac{3}{2a_n^3}\sum_{k=n+1}^{\infty}\int_{b_k}^{a_k} tf(t)\,dt=\frac{3a_{n+1}^3}{2a_n^3}\Big(\frac{1}{a_{n+1}^3}\sum_{k=n+1}^{\infty}\int_{b_k}^{a_k} tf(t)\,dt\Big).$$

Since the ratio $a_n/a_{n+1}$ converges to 1 as $n$ goes to infinity, the problem reduces to evaluating

$$\lim_{n\to\infty}\frac{3}{2a_n^3}\sum_{k=n}^{\infty}\int_{b_k}^{a_k} tf(t)\,dt.$$

By observation (B), we find that

$$\frac{3}{2a_n^3}\sum_{k=n}^{\infty}\int_{b_k}^{a_k} tf(t)\,dt = \frac{3}{2}\,n^6\sum_{k=n}^{\infty}\frac{16\,c_k^r}{4}(a_k^2-b_k^2)$$
$$= 6n^6\sum_{k=n}^{\infty}c_k^r(a_k+b_k)(a_k-b_k)$$
$$= 3n^6\sum_{k=n}^{\infty}\frac{c_k^{r+1}}{k^4}.$$

From the definition of $c_k$, we know that

$$c_k = \frac{1}{k^2}-\frac{1}{8k^4} = \frac{1}{k^2}\Big(1-\frac{1}{8k^2}\Big).$$

Since the graph of $y = (1+x)^{1+r}$ is concave up on the interval $[-1,1]$, the graph lies above its tangent lines. In particular, using the tangent line at the point $(0,1)$,

$$(1+x)^{1+r} \geq 1+(1+r)\,x$$

for all $x \in [-1,1]$. It follows that

$$\Big(\frac{1}{k^2}\Big)^{1+r}\Big(1-\frac{1+r}{8k^2}\Big) \leq c_k^{1+r} < \Big(\frac{1}{k^2}\Big)^{1+r}$$

for each positive integer $k$ and hence

$$3n^6\sum_{k=n}^{\infty}\Big(\frac{1}{k^{6+2r}}-\frac{1+r}{8k^{8+2r}}\Big) \leq 3n^6\sum_{k=n}^{\infty}\frac{c_k^{r+1}}{k^4} \leq 3n^6\sum_{k=n}^{\infty}\frac{1}{k^{6+2r}}.$$

Noting that

$$3n^6\sum_{k=n}^{\infty}\frac{1+r}{8k^{8+2r}} \leq \frac{3(1+r)}{8n}\Big(n^7\sum_{k=n}^{\infty}\frac{1}{k^8}\Big),$$

which goes to 0 as $n$ goes to infinity (observation (A)), we see that

$$f_\ell'(0) = \lim_{n\to\infty}3n^6\sum_{k=n}^{\infty}\frac{1}{k^{6+2r}} = \lim_{n\to\infty}3n^{1-2r}\Big(n^{5+2r}\sum_{k=n}^{\infty}\frac{1}{k^{6+2r}}\Big).$$

Appealing to observation (A) again, the term in parentheses has a nonzero limit (namely, $(5+2r)^{-1}$) as $n$ goes to infinity. Therefore,

$$f_\ell'(0) = \begin{cases} 0, & \text{if } r > 1/2; \\ 1/2, & \text{if } r = 1/2; \\ \text{does not exist}, & \text{if } 0 < r < 1/2. \end{cases}$$

In other words, it is possible for the least squares derivative to fail to exist even when the approximate derivative exists (which is not a real surprise) and, more interestingly, it is possible for both derivatives to exist but have different values (the $r = 1/2$ case). □

Functions such as the absolute value function and those presented in Example 3 indicate that, in general, there is no relationship between the least squares derivative and the approximate derivative. However, we next present a result that gives conditions for the existence of the approximate derivative to imply the existence (and equality) of the least squares derivative. To do so, we need to introduce two concepts, one familiar and one not so familiar.

**Definition 4.** *Suppose that $f$ is defined on a neighborhood of a point $c$ and let $\alpha$ be a nonnegative number. The function $f$ is $\alpha$-Hölder continuous at $c$ if there exists a constant $M$ such that $|f(c+x) - f(c)| \le M|x|^\alpha$ for all values of $x$ in some neighborhood of 0.*

If $f$ is $\alpha$-Hölder continuous at $c$ for any $\alpha > 0$, then $f$ is continuous at $c$, while $f$ is merely bounded in a neighborhood of $c$ when $f$ is 0-Hölder continuous at $c$. If $f$ is 1-Hölder continuous at $c$, then $f$ has bounded difference quotients at $c$. If $f$ is $\alpha$-Hölder continuous at $c$ for some $\alpha > 1$, then $f$ is differentiable at $c$ with $f'(c) = 0$. When $0 \le \alpha < \beta$, it is easy to see that $f$ is $\alpha$-Hölder continuous at $c$ if $f$ is $\beta$-Hölder continuous at $c$. A simple collection of functions that illustrate the full range of $\alpha$-Hölder continuity at 0 are functions of the form $f(x) = x^\alpha \sin(\pi/x)$ for $x \ne 0$ and $f(0) = 0$. (For the record, a function that is continuous at zero and satisfies $f(1/n^n) = 1/n$ for each positive integer $n$ is not $\alpha$-Hölder continuous at 0 for any $\alpha > 0$.)

**Definition 5.** *Suppose that $f$ is a continuous function defined on a neighborhood of a point $c$ and that $f$ has an approximate derivative at $c$. Let $\beta \ge 1$ be a real number. The function $f$ has a $\beta$-thick approximate derivative at $c$ if there exists a measurable set $E$ such that*

$$f'_{ap}(c) = \lim_{\substack{x\to 0\\ x\in E}} \frac{f(c+x) - f(c)}{x} \qquad \text{and} \qquad \lim_{h\to 0^+} \frac{\mu(\mathcal{C}E \cap (-h,h))}{2h^\beta} = 0.$$

*The usual approximate derivative corresponds to the case $\beta = 1$.*

The essential idea behind this concept is that more difference quotients are considered because the set $E$ is "bigger" in some specific sense. In other words, the measure of the set where the difference quotients are ill-behaved must be smaller (when $\beta > 1$) than in the case of the approximate derivative. A simple example of this concept is presented below; the interested reader can create further examples.

**Example 6.** *There exist functions that do not have an ordinary derivative at a point but do have a $\beta$-thick approximate derivative at that point for various values of $\beta$.*

Let $\lambda > 2$ be a fixed real number. For each positive integer $n$, let

$$a_n = \frac{1}{n} \quad \text{and} \quad b_n = \frac{1}{n} - \frac{w}{n^\lambda},$$

where $w$ is chosen so that $a_{n+1} < b_n < a_n$ for all $n$. (Note that any $w$ that satisfies $0 < w < 1/2$ will do the job.) Let $A$ be the set defined by

$$A = \bigcup_{k=1}^{\infty} (-a_k, -b_k) \cup \bigcup_{k=1}^{\infty} (b_k, a_k)$$

and let $E = [-1, 1] \setminus A$ (so that $A = \mathcal{C}E$). For $a_{n+1} < h \leq a_n$, we find that

$$\begin{aligned} \frac{\mu\big(A \cap (-h, h)\big)}{2h^\beta} &\leq \frac{\mu\big(A \cap (-a_n, a_n)\big)}{2a_{n+1}^\beta} \\ &= (n+1)^\beta \sum_{k=n}^{\infty} (a_k - b_k) \\ &= \frac{w(n+1)^\beta}{n^{\lambda-1}} \Big( n^{\lambda-1} \sum_{k=n}^{\infty} \frac{1}{k^\lambda} \Big). \end{aligned}$$

By observation (A), the limit as $n$ goes to infinity (which occurs as $h \to 0^+$) of the last expression is 0 for any $\beta$ that satisfies the inequality $1 \leq \beta < \lambda - 1$. Suppose that $f$ is defined on $[-1, 1]$ by setting $f(x) = 0$ for $x \in E$ and letting the graph of $f$ on the intervals $[b_n, a_n]$ and $[-a_n, -b_n]$ consist of triangular spikes in such a way that $f$ is continuous at 0 but not differentiable at 0. Then the function $f$ has a $\beta$-thick approximate derivative at the point 0 for each $1 \leq \beta < \lambda - 1$, but $f$ does not have an ordinary derivative at 0. For the record, if $b_n = (1/n) - (w/2^n)$, then it is possible to construct a function $f$ such that $f$ has a $\beta$-thick approximate derivative at 0 for any $\beta \geq 1$ even though $f$ does not have an ordinary derivative at 0. $\square$

With these two concepts at our disposal, we can state and prove conditions for which the existence of the approximate derivative implies the existence of the least squares derivative. These same conditions also imply that the two derivatives are equal.

**Theorem 7.** *Let $f$ be a continuous function defined on a neighborhood of a point $c$. Suppose that $f$ is $\alpha$-Hölder continuous at $c$ and that $f$ has a $\beta$-thick approximate derivative at $c$. If $\alpha+\beta \ge 2$, then $f$ is least squares differentiable at $c$ and $f'_\ell(c) = f'_{ap}(c)$.*

PROOF. Since $f$ is $\alpha$-Hölder continuous at $c$, there exists $M > 0$ so that

$$|f(c+x) - f(c)| \le M|x|^\alpha$$

for all values of $x$ in a neighborhood of 0. Since $f$ has a $\beta$-thick approximate derivative at $c$, there exists a measurable set $E$ such that

$$f'_{ap}(c) = \lim_{\substack{x\to 0\\ x\in E}} \frac{f(c+x)-f(c)}{x} \qquad \text{and} \qquad \lim_{h\to 0^+} \frac{\mu(\mathcal{C}E \cap (-h,h))}{2h^\beta} = 0.$$

For ease of writing, let $E_h = E \cap (-h,h)$ and $A_h = (-h,h) \setminus E$ for each $h > 0$, and let $v = f'_{ap}(c)$. Since $f$ has an approximate derivative at $c$, there exists a function $\epsilon$ such that

$$\lim_{\substack{x\to 0\\ x\in E}} \epsilon(x) = 0 \quad \text{and} \quad f(c+x) = f(c) + vx + \epsilon(x)\, x$$

for all $x \in E$. In order to show that

$$\lim_{h\to 0^+} \frac{3}{2h^3} \int_{-h}^{h} t\, f(c+t)\, dt = v,$$

we begin by noting that

$$\begin{aligned}
\frac{3}{2h^3} \int_{-h}^{h} t\, f(c+t)\, dt &= \frac{3}{2h^3} \int_{-h}^{h} t\big(f(c+t) - f(c)\big)\, dt \\
&= \frac{3}{2h^3} \int_{E_h} t\big(f(c+t) - f(c)\big)\, dt + \frac{3}{2h^3} \int_{A_h} t\big(f(c+t) - f(c)\big)\, dt \\
&= \frac{3}{2h^3} \int_{E_h} vt^2\, dt + \frac{3}{2h^3} \int_{E_h} \epsilon(t)\, t^2\, dt + \frac{3}{2h^3} \int_{A_h} t\big(f(c+t) - f(c)\big)\, dt.
\end{aligned}$$

Referring to the last line of the previous displayed equation, we will show that the limit of the first integral expression is $v$ and that the limits of the other two are 0.

For the first integral expression, we note that

$$\frac{3}{2h^3} \int_{E_h} vt^2\, dt = \frac{3}{2h^3} \int_{-h}^{h} vt^2\, dt - \frac{3}{2h^3} \int_{A_h} vt^2\, dt = v - \frac{3}{2h^3} \int_{A_h} vt^2\, dt,$$

where

$$\left|\frac{3}{2h^3}\int_{A_h} vt^2\,dt\right| \le 3|v|\cdot\frac{\mu(A_h)}{2h}.$$

Since 0 is a point of density of $E$ (recall that $\beta \ge 1$), the limit of $\mu(A_h)/h$ is 0 as $h$ tends to 0, and it follows that

$$\lim_{h\to 0^+}\frac{3}{2h^3}\int_{E_h} vt^2\,dt = v.$$

Estimating the magnitude of the second integral expression reveals that

$$\left|\frac{3}{2h^3}\int_{E_h}\epsilon(t)\,t^2\,dt\right| \le \frac{3}{2h}\int_{E_h}|\epsilon(t)|\,dt \le 3M_\epsilon(h)\cdot\frac{\mu(E_h)}{2h} \le 3M_\epsilon(h),$$

where $M_\epsilon(h) = \sup\{|\epsilon(t)| : t \in E_h\}$. The properties of $\epsilon$ then imply that this expression goes to 0 with $h$. It is the third integral expression that is the most sensitive. We find that

$$\begin{aligned}\left|\frac{3}{2h^3}\int_{A_h} t\big(f(c+t)-f(c)\big)\,dt\right| &\le \frac{3}{2h^3}\int_{A_h}|t|\big|f(c+t)-f(c)\big|\,dt\\ &\le \frac{3}{2h^3}\int_{A_h} M|t|^{1+\alpha}\,dt\\ &\le \frac{3}{2h^3}M|h|^{1+\alpha}\mu(A_h)\\ &= 3M\,|h|^{\alpha+\beta-2}\cdot\frac{\mu(A_h)}{2h^\beta}.\end{aligned}$$

Given the properties of the set $\mathcal{C}E$ (note that $A_h = \mathcal{C}E \cap (-h,h)$) and the fact that $\alpha+\beta-2 \ge 0$, we see that the third integral expression also goes to 0 with $h$. Putting all the pieces together, we find that $f$ is least squares differentiable at $c$ and $f'_\ell(c) = v$, that is, $f'_\ell(c) = f'_{ap}(c)$. □

To get a sense for this theorem, let's consider the function $f$ from Example 3 and adopt all of the notation from that example. First of all, it is easy to see that $f$ is $r$-Hölder continuous at 0. We claim that $f$ has a $\beta$-thick approximate derivative at 0 for any $\beta$ that satisfies $1 \le \beta < 3/2$. To verify this, take any $h$ that satisfies $a_{n+1} < h \le a_n$ and compute

$$\begin{aligned}\frac{\mu\big(A\cap(-h,h)\big)}{2h^{\beta}} &\le \frac{\mu\big(A\cap(-a_n,a_n)\big)}{2a_{n+1}^{\beta}}\\ &= \frac{(n+1)^{2\beta}}{2}\sum_{k=n}^{\infty}(a_k-b_k)\\ &= \frac{(n+1)^{2\beta}}{8n^3}\Big(n^3\sum_{k=n}^{\infty}\frac{1}{k^4}\Big).\end{aligned}$$

Using observation (A), we see that

$$\lim_{h\to 0^+}\frac{\mu\big(A\cap(-h,h)\big)}{2h^{\beta}}=0$$

as long as $2\beta<3$. This establishes the claim. Theorem 7 then says that $f$ is least squares differentiable at 0 with $f'_{\ell}(0)=f'_{ap}(0)$ as long as $r>1/2$. The case in which $r=1/2$ shows that the inequality $\alpha+\beta\ge 2$ in Theorem 7 is sharp.

After working with these two derivatives, it is tempting to try to define a least squares approximate derivative that includes both the least squares derivative and the approximate derivative. The fact that these two derivatives can both exist but have different values makes such a definition impossible. However, we close this paper by stating the following result that indicates how to obtain approximate derivatives with integration. As the proof uses ideas very similar to those in the proof of Theorem 7, it will be left for the interested reader. Note also that the restriction to continuous functions can certainly be weakened for results of this type.

**Theorem 8.** *Suppose that $f$ is a continuous function defined on a neighborhood of a point $c$. If $f$ has an approximate derivative at $c$, then there exists a measurable set $E$ such that* 0 *is a point of density of $E$ and*

$$f'_{ap}(c)=\lim_{h\to 0^+}\frac{3}{2h^3}\int_{E_h} t\big(f(c+t)-f(c)\big)\,dt,$$

*where $E_h=E\cap(-h,h)$.*

As a final note, it is possible to construct a measurable set $E$ having $c$ as a point of density and a continuous function $f$ for which the limit given in the conclusion of Theorem 8 exists, but $f$ has neither a least squares derivative nor an approximate derivative at $c$. One way to construct such a function is to redefine the function $f$ from Example 3 on the set $E$.

## References

[1] N. Burch, P. Fishback, R. Gordon, *The least squares property of the Lanczos derivative*, Math. Magazine **75** (2005) 368–378.

[2] R. Gordon, *A least squares approach to differentiation*, Real Anal. Exchange **35** (2010) 205-228.

[3] R. Gordon, *The integrals of Lebesgue, Denjoy, Perron, and Henstock*, Graduate Studies in Mathematics, Vol. 4, American Mathematical Society, Providence, 1994.

[4] C. Groetsch, *Lanczos' generalized derivative*, Amer. Math. Monthly **105** (1998) 320–326.

[5] D. Hicks and L. Liebrock, *Lanczos' generalized derivative: insights and applications*, Appl. Math. and Comp. **112** (2000) 63-73.

[6] C. Lanczos, *Applied Analysis*, Prentice Hall, 1960.

[7] B. Thomson, *Differentiation*, Handbook of measure theory, Vol. I, II, 179–247, North-Holland, Amsterdam, 2002.