INROADS

Real Analysis Exchange
Vol. 35(1), 2009/2010, pp. 205–228

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

A LEAST SQUARES APPROACH TO DIFFERENTIATION

Abstract

We consider an approach to differentiation that involves least squares lines of best fit rather than the traditional secant lines and use elementary techniques to show how this leads to the Lanczos derivative. A number of examples are presented to illustrate this concept and to show that the product, quotient, and chain rules fail for the Lanczos derivative. Several results giving conditions for which these rules do hold are discussed and proved. A brief introduction to higher order Lanczos derivatives is included.

We consider an approach to differentiation that involves least squares lines of best fit rather than the traditional secant lines. This rather intriguing view of the derivative leads to a generalization of both the ordinary derivative and the symmetric derivative. We carefully explain the motivation behind the first order least squares derivative and present a number of examples to illustrate the new concept. Interestingly, the product rule, the quotient rule, and the chain rule all fail for this least squares derivative. However, several results giving conditions for which these rules do hold are discussed and proved. By viewing this new derivative in the lens of inner product spaces, we show how to extend least squares differentiation to higher order derivatives. It turns out that the nth order least squares derivative is a generalization of the nth order Peano derivative.

Although it is possible to develop the ideas presented here in the context of measurable functions, we restrict ourselves to continuous functions so as to avoid trivial examples that arise when sets of measure zero are ignored and to keep the material accessible to a wider audience. Hence, throughout this

Mathematical Reviews subject classification: Primary: 26A24, 26A16
Key words: Lanczos derivative, Hölder continuity, Peano derivative
Received by the editors September 30, 2008
Communicated by: Paul D. Humke

paper, we consider a function f that is continuous on a neighborhood of a point c.

For the sake of completeness, we begin with two familiar definitions. The derivative of f at c and the symmetric derivative of f at c are defined by

$$f'(c) = \lim_{h\to 0} \frac{f(c+h) - f(c)}{h}, \qquad f'_s(c) = \lim_{h\to 0^+} \frac{f(c+h) - f(c-h)}{2h},$$

respectively. It is easy to verify that $f'_s(c)$ exists whenever $f'(c)$ exists. However, the absolute value function (at the origin) shows that $f'_s(c)$ can exist even when $f'(c)$ does not exist. A little more generally, the symmetric derivative exists whenever f has left and right derivatives at c and the value of the symmetric derivative is the average of these two one-sided derivatives. The converse is false since the function S defined by $S(x) = x\sin(1/x)$ for $x \neq 0$ and $S(0) = 0$ has a symmetric derivative at 0 but does not have one-sided derivatives at 0.

The difference quotients that appear in the definitions of the derivative and the symmetric derivative are attempts to determine the slope of a curve at a point. They do this in a very natural way; take two points on the curve and use them to determine a slope. To get the slope of the curve at c, we then take a limit as the points move closer to c. The only difference between the methods is the choice of the two points. Since a curve is full of points, however, it is possible to use more than two of them at a time. Given a number of points, the most common way of using them to determine the slope of a line is the method of least squares. Suppose we take several points on the curve and find the least squares line of best fit for those points, then take the limit of the slope of this line as all the points converge to $(c, f(c))$. In order to simplify matters, we always assume that the x-coordinates of our points are symmetrically placed about the point c. The case in which three points are used is illustrated in Figure 1.

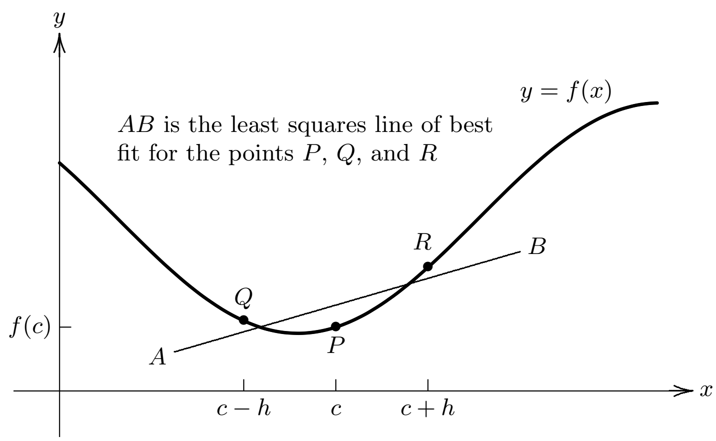

Figure 1: A least squares line using three points

For the ordinary derivative, we consider the slope of either of the lines PQ or PR, while for the symmetric derivative, we consider the slope of the line QR. In each case, we then let $h \to 0$ and determine whether or not the slopes have a limit. For the least squares line of best fit approach, more effort is required to find the slope of the desired line. We need to find a line of the form $y = m(x - c) + b$, where the constants m and b are the best least squares approximation to the linear system

$$\begin{aligned} b - hm &= f(c-h); \\ b \qquad &= f(c); \\ b + hm &= f(c+h). \end{aligned}$$

Using standard techniques (illustrated briefly in a moment), the least squares solution is given by

$$b = \frac{f(c-h) + f(c) + f(c+h)}{3}, \qquad m = \frac{f(c+h) - f(c-h)}{2h}.$$

Note that the least squares slope is the same as the difference quotient that appears in the definition of the symmetric derivative, that is, the line AB is parallel to the line QR in Figure 1. Hence, the symmetric derivative can be interpreted as a type of least squares derivative.

The ideas are the same when we consider more points. To illustrate this explicitly and to show how the least squares solutions are found, consider the situation in which there are five symmetrically placed points with x-coordinates

$$c-h,\ c-\tfrac{1}{2}h,\ c,\ c+\tfrac{1}{2}h,\ c+h.$$

The corresponding points on the graph are represented by bullets in Figure 2.

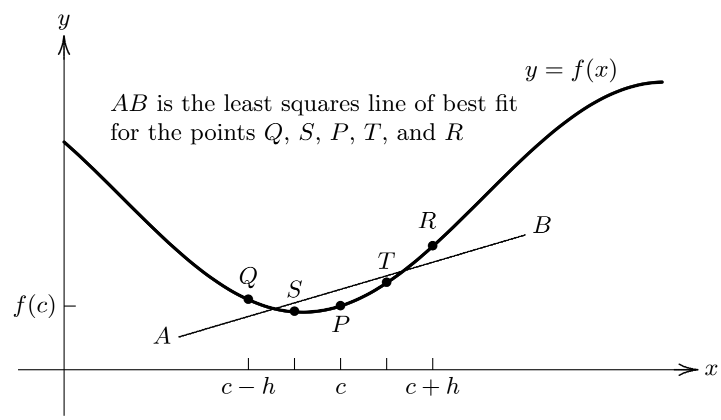

Figure 2: A least squares line using five points

In this case, the constants m and b for the line $y = m(x-c)+b$ of best fit are the least squares solution to the matrix equation

$$\begin{bmatrix} 1 & -h \\ 1 & -\dfrac{h}{2} \\ 1 & 0 \\ 1 & \dfrac{h}{2} \\ 1 & h \end{bmatrix} \begin{bmatrix} b \\ m \end{bmatrix} = \begin{bmatrix} f(c-h) \\ f\left(c-\dfrac{h}{2}\right) \\ f(c) \\ f\left(c+\dfrac{h}{2}\right) \\ f(c+h) \end{bmatrix} \equiv \begin{bmatrix} \mathbf{c} & \mathbf{d} \end{bmatrix} \begin{bmatrix} b \\ m \end{bmatrix} = \mathbf{f},$$

where the definitions of the vectors $\mathbf{c}$, $\mathbf{d}$, and $\mathbf{f}$ should be clear. Since the columns of the matrix (that is, the vectors $\mathbf{c}$ and $\mathbf{d}$) are orthogonal (this is guaranteed by the symmetry condition), the value of m is simply

$$m = \frac{\mathbf{f}\cdot\mathbf{d}}{\mathbf{d}\cdot\mathbf{d}} = \frac{f(c+h) + \frac{1}{2}f\left(c+\frac{h}{2}\right) - \frac{1}{2}f\left(c-\frac{h}{2}\right) - f(c-h)}{\frac{5}{2}h}.$$

In general, if we fix a positive number h and a positive integer n and use $2n+1$ symmetrically placed data points with x-coordinates

$$c-h,\ c-\frac{n-1}{n}\,h,\ c-\frac{n-2}{n}\,h,\ \ldots\ ,\ c,\ \ldots\ ,\ c+\frac{n-2}{n}\,h,\ c+\frac{n-1}{n}\,h,\ c+h,$$

then the slope $m(h,n)$ of the least squares line of best fit is given by

$$m(h,n)=\frac{\displaystyle\sum_{k=-n}^{n}\frac{k}{n}\,h\,f\Big(c+\frac{k}{n}\,h\Big)}{\displaystyle\sum_{k=-n}^{n}\Big(\frac{k}{n}\,h\Big)^2}.$$

For each fixed positive integer n, the expression $\lim\limits_{h\to 0^+} m(h,n)$ is a least squares derivative. (Each of these derivatives is a generalized Riemann derivative, see [1]. I suspect, but have not taken the time to construct examples, that this generates a sequence of proper extensions of the ordinary derivative.) However, we are interested in the least squares derivative that arises when the number of data points increases indefinitely, that is, we want to look at

$$\lim_{h\to 0^+}\lim_{n\to\infty} m(h,n).$$

In order to find an expression for $\lim\limits_{n\to\infty} m(h,n)$, we begin by writing

$$m(h,n)=\frac{\displaystyle\Big(\frac{h}{n}\Big)^2\sum_{k=-n}^{n}k\,f\Big(c+\frac{k}{n}\,h\Big)}{\displaystyle\Big(\frac{h}{n}\Big)^3\sum_{k=-n}^{n}k^2}.$$

It is easy to verify that the denominator has a limit of $2h^3/3$ as n goes to infinity. The numerator can be interpreted as a Riemann sum using right endpoints to generate the value of the function, but there is one extra term that disappears in the limit. It turns out that

$$\lim_{n\to\infty}\Big(\frac{h}{n}\Big)^2\sum_{k=-n}^{n}k\,f\Big(c+\frac{k}{n}\,h\Big)=\int_{-h}^{h}t\,f(c+t)\,dt;$$

we leave the elementary, but somewhat tedious, details to the reader. Putting the two pieces together, we find that (with a simple change of variable for the second integral)

$$\lim_{n\to\infty} m(h,n) = \frac{3}{2h^3}\int_{-h}^{h} t\,f(c+t)\,dt = \frac{3}{2h}\int_{-1}^{1} t\,f(c+ht)\,dt.$$

Either of these equivalent integral expressions forms the basis for the least squares derivative that we will consider. Depending on the circumstances, one of these expressions is often more useful than the other.

Definition 1. Suppose that f is a continuous function defined on a neighborhood of a point c. The Lanczos derivative of f at 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.

We note in passing that the Lanczos derivative is a linear operator. As indicated by its name, this derivative is due to Lanczos, see [9]. Lanczos was interested in approximating derivatives of functions represented only as a table of values; such representations often occur as numerical data obtained from experiments. He considered the case in which n is a small positive integer and was also interested in higher order derivatives such as second derivatives and third derivatives. For the first derivative, he simply mentions 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 [3], [6], and [8]) have recently done some work with this derivative, primarily with a focus on numerical analysis and statistics.

To illustrate the definition, we begin with three elementary examples.

Example 2. Suppose that f is defined by $f(x) = x^n$, where n is a positive integer. Using properties of even and odd functions as well as the Binomial Theorem, we find that

$$\begin{aligned}
f'_\ell(c) &= \lim_{h\to 0^+} \frac{3}{2h}\int_{-1}^{1} t\,(c+ht)^n\,dt \\
&= \lim_{h\to 0^+} \frac{3}{2h}\int_{-1}^{1} t\sum_{k=0}^{n}\binom{n}{k}(ht)^k c^{n-k}\,dt \\
&= \lim_{h\to 0^+} \frac{3}{2h}\int_{-1}^{1}\Big(c^n t + nhc^{n-1}t^2 + h^2\sum_{k=2}^{n}\binom{n}{k}h^{k-2}c^{n-k}t^{k+1}\Big)\,dt \\
&= nc^{n-1}.
\end{aligned}$$

Combining this fact with the linearity of the Lanczos derivative, we see that the Lanczos derivative of any polynomial is the same as the ordinary derivative of that polynomial.

Example 3. Let a and b be distinct real numbers and define a function f by $f(x) = ax$ for $x \leq 0$ and $f(x) = bx$ for $x > 0$. Then

$$f'_\ell(0) = \lim_{h \to 0^+} \frac{3}{2h^3}\Big(\int_{-h}^{0} at^2\,dt + \int_0^h bt^2\,dt\Big) = \frac{a+b}{2}.$$

This example shows that the Lanczos derivative can exist even when the ordinary derivative does not exist. However, the function f does have a symmetric derivative at 0 and $f'_s(0) = f'_\ell(0)$ in this case.

Example 4. For this example, rather than write an equation for a function f, we will let a graph speak for itself; see Figure 3.

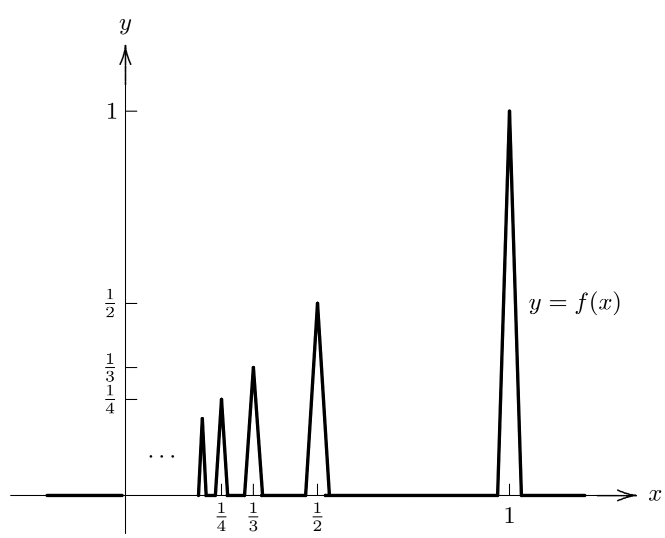

Figure 3: A function f for which $f'_\ell(0) = 0$ but $f'_s(0)$ does not exist

The function f assumes the value 0 except at points in small intervals centered at each of the points with x-coordinates $1/k$, where k is a positive integer. On these intervals, the graph of f forms a triangle with upper vertex on the line $y = x$, that is, $f(1/k) = 1/k$, and the lengths of the intervals are defined so that the area of the triangle surrounding the x-coordinate $1/k$ is $1/2^{k+4}$.

It is then clear that the symmetric derivative of f at 0 does not exist. For $1/(n+1) < h \le 1/n$, where n is any positive integer,

$$\begin{aligned} 0 < \frac{3}{2h^3}\int_{-h}^{h} t\,f(t)\,dt &\le \frac{3}{2}\,(n+1)^3 \int_0^{1/n} f(t)\,dt \\ &\le \frac{3}{2}\,(n+1)^3 \sum_{k=n}^{\infty} \frac{1}{2^{k+4}} = \frac{3}{2}\cdot\frac{(n+1)^3}{2^{n+3}}. \end{aligned}$$

As h goes to 0^+, we know that n goes to infinity. The inequality thus shows that $f'_\ell(0)$ exists and has a value of 0.

Examples 3 and 4 show that there are continuous functions which have Lanczos derivatives but do not have ordinary derivatives or symmetric derivatives. The next theorem shows that a function has a Lanczos derivative whenever it has a symmetric derivative. Hence, the Lanczos derivative is a proper extension of the symmetric derivative (which in turn is a proper extension of the ordinary derivative).

Theorem 5. *Suppose that f is a continuous function defined on a neighborhood of a point c. If $f'_s(c)$ exists, then $f'_\ell(c)$ exists and the values are the same.*

PROOF. Using L'Hôpital's Rule and the Fundamental Theorem of Calculus, we find that

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

Note the use of the continuity of f when differentiating the integral. □

The next two examples are more complicated than our earlier examples. They illustrate other facets of the Lanczos derivative and show that it is not as well-behaved as ordinary and symmetric derivatives.

Example 6. In this example, we construct a continuous function f for which $f'_\ell(0)$ exists and is nonzero but $f'_s(0)$ does not exist. Of course, a simple way to do this is to add a linear term mx to the function in Example 4. However, we want to illustrate a different type of function. Once again, a graph should be sufficient to show what such a function looks like. Fix a positive number m. For each positive integer k, let

$$a_k = \frac{1}{k}, \quad c_k = \frac{a_k + a_{k+1}}{2}.$$

The function f has a value of 0 at each a_k, a value of mc_k at each c_k, and is piecewise linear; see Figure 4.

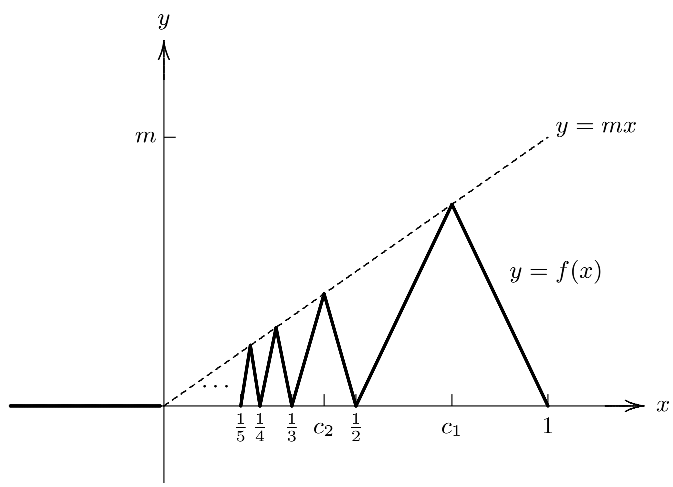

Figure 4: A function f for which $f'_\ell(0) \neq 0$ and $f'_s(0)$ does not exist

Once again, it is clear that f does not have a symmetric derivative at 0. To determine the Lanczos derivative of f at 0, we begin with a computation of the integral of $tf(t)$ over the interval $[a_{k+1}, a_k]$. On each half of this interval, the function $tf(t)$ is quadratic so we can use the formula (which forms the basis for Simpson's Rule)

$$\int_a^b Q(t)\,dt = \frac{Q(a) + 4Q((a+b)/2) + Q(b)}{6}\,(b-a),$$

where Q is any quadratic (or even cubic) polynomial. The relevant table of values for our function f is

x	a_{k+1}	$\frac{3a_{k+1}+a_k}{4}$	$\frac{a_{k+1}+a_k}{2}$	$\frac{a_{k+1}+3a_k}{4}$	a_k
$f(x)$	0	$\frac{m(a_{k+1}+a_k)}{4}$	$\frac{m(a_{k+1}+a_k)}{2}$	$\frac{m(a_{k+1}+a_k)}{4}$	0

and we find that

$$\begin{aligned}\int_{a_{k+1}}^{a_k} t\,f(t)\,dt &= \frac{a_k - a_{k+1}}{12}\Big(1\cdot a_{k+1}\cdot 0 + 4\cdot\frac{3a_{k+1}+a_k}{4}\cdot\frac{m(a_{k+1}+a_k)}{4}\\ &\qquad + 2\cdot\frac{a_{k+1}+a_k}{2}\cdot\frac{m(a_{k+1}+a_k)}{2}\\ &\qquad + 4\cdot\frac{a_{k+1}+3a_k}{4}\cdot\frac{m(a_{k+1}+a_k)}{4} + 1\cdot a_k\cdot 0\Big)\\ &= \frac{m}{8}(a_k - a_{k+1})(a_k + a_{k+1})^2\\ &= \frac{m}{8}\cdot\frac{(2k+1)^2}{k^3(k+1)^3}.\end{aligned}$$

To compute the Lanczos derivative of f at 0, we need to evaluate the limit

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

For $1/(n+1) \le h < 1/n$, we can use over and under estimates to obtain

$$\begin{aligned}\frac{3}{2h^3}\int_0^h t\,f(t)\,dt &\le \frac{3}{2}\,(n+1)^3\sum_{k=n}^{\infty}\int_{a_{k+1}}^{a_k} t\,f(t)\,dt\\ &= \frac{3m}{16}\cdot\frac{(n+1)^3}{n^3}\cdot n^3\sum_{k=n}^{\infty}\frac{(2k+1)^2}{k^3(k+1)^3}\end{aligned}$$

and

$$\begin{aligned}\frac{3}{2h^3}\int_0^h t\,f(t)\,dt &\ge \frac{3}{2}\,n^3\sum_{k=n+1}^{\infty}\int_{a_{k+1}}^{a_k} t\,f(t)\,dt\\ &= \frac{3m}{16}\cdot\frac{n^3}{(n+1)^3}\cdot(n+1)^3\sum_{k=n+1}^{\infty}\frac{(2k+1)^2}{k^3(k+1)^3}.\end{aligned}$$

By comparing the two expressions

$$n^3\sum_{k=n}^{\infty}\frac{(2k+1)^2}{k^3(k+1)^3}, \qquad n^3\sum_{k=n}^{\infty}\Big(\frac{1}{k^3} - \frac{1}{(k+1)^3}\Big),$$

it is not difficult to show that as n tends to infinity, both the over and under estimates have limits of $m/4$. The Squeeze Theorem can then be invoked to reveal that $f_\ell'(0) = m/4$.

Although the Lanczos derivative is a linear operator, it turns out that some of the other familiar differentiation rules for ordinary derivatives fail for the Lanczos derivative. In particular, the product rule, the quotient rule, and the chain rule do not hold. Providing an example to illustrate these facts is our next goal.

Example 7. Let f be the function defined by

$$f(x) = \begin{cases} 0, & \text{if } x \le 0; \\ \sqrt{x}\sin(\pi/x), & \text{if } x > 0; \end{cases}$$

It is easy to verify that f does not have a symmetric derivative at 0. We will show that f has a Lanczos derivative of 0 at 0 and that f^2 has a nonzero Lanczos derivative at 0. It then follows that the product rule (for the product $f \cdot f$) and the chain rule (for the composition f^2) fail for the Lanczos derivative.

The following equations determine $f'_\ell(0)$. The substitutions $x = 1/h$ and $s = 1/t$ convert the limit and the integral into a form in which integration by parts plays a key role.

$$\begin{aligned} f'_\ell(0) &= \lim_{h\to 0^+} \frac{3}{2h^3}\int_0^h t^{3/2}\sin(\pi/t)\,dt \\ &= \lim_{x\to\infty} \frac{3x^3}{2}\int_x^\infty s^{-7/2}\sin(\pi s)\,ds \\ &= \lim_{x\to\infty} \frac{3x^3}{2\pi}\Big(-s^{-7/2}\cos(\pi s)\Big|_x^\infty - \frac{7}{2}\int_x^\infty s^{-9/2}\cos(\pi s)\,ds\Big) \\ &= \lim_{x\to\infty} \frac{3x^3}{2\pi}\Big(\frac{\cos(\pi\, x)}{x^{7/2}} - \frac{7}{2}\int_x^\infty s^{-9/2}\cos(\pi s)\,ds\Big) \\ &= 0. \end{aligned}$$

The value of the limit follows from the fact that the first term clearly has a limit of 0 as $x \to \infty$ and the second does as well because

$$\Big|\frac{7}{2}\,x^3\int_x^\infty s^{-9/2}\cos(\pi s)\,ds\Big| \le x^3\int_x^\infty \frac{7}{2}\,s^{-9/2}\,ds = \frac{1}{\sqrt{x}}.$$

We conclude that $f'_\ell(0) = 0$.

A similar computation determines $(f^2)'_\ell(0)$; this time the half-angle identity for sine plays a role.

$$\begin{aligned}
(f^2)'_\ell(0) &= \lim_{h\to 0^+} \frac{3}{2h^3}\int_0^h t^2\sin^2(\pi/t)\,dt \\
&= \lim_{x\to\infty} \frac{3x^3}{2}\int_x^\infty s^{-4}\sin^2(\pi s)\,ds \\
&= \lim_{x\to\infty} \frac{3x^3}{4}\int_x^\infty s^{-4}\big(1-\cos(2\pi s)\big)\,ds \\
&= \lim_{x\to\infty} \frac{3x^3}{4}\Big(\frac{1}{3x^3} - \int_x^\infty s^{-4}\cos(2\pi s)\,ds\Big) \\
&= \frac{1}{4},
\end{aligned}$$

since the remaining integral term has a limit of 0 by repeating the integration by parts argument used in the previous computation. We conclude that $(f^2)'_\ell(0) = 1/4$.

For the record, if $\sqrt{x}$ is replaced with $\sqrt[3]{x}$ in the definition of the function f in Example 7, then the new function f has a Lanczos derivative at 0 but f^2 does not have a Lanczos derivative at 0. This illustrates that an even more dramatic breakdown of the product rule and the chain rule can occur. The details (just slight modifications of those given) are left to the reader.

The failure of the quotient rule for the Lanczos derivative follows from the fact that the validity of the quotient rule implies the validity of the product rule. To see this, write

$$fg = \frac{f}{1/g} \quad \text{or} \quad fg = \frac{f}{1/(1+g)} - f,$$

the latter being necessary if $g(c) = 0$. Using the linearity of the Lanczos derivative and the assumed quotient rule, we obtain the product rule. Once again, the elementary details are left to the reader. Thus the failure of the product rule implies the failure of the quotient rule.

We should mention that the product rule and the quotient rule are valid for symmetric derivatives (when the functions are assumed to be continuous as we are doing in this paper) and the proofs of these results are quite similar to the usual proofs for ordinary derivatives. However, the chain rule fails for symmetric derivatives. For instance, if $f(x) = |x|$ and

$$g(x) = \begin{cases} 0, & \text{if } x \le 0; \\ 2x, & \text{if } x > 0; \end{cases}$$

then $f \circ g = g$. But since $f_s'(0) = 0$ and $g_s'(0) = 1$, the chain rule formula fails to hold at 0.

In summary, the product rule, the quotient rule, and the chain rule all fail for the Lanczos derivative. Although these differentiation rules fail in general, there are conditions on the functions that guarantee that these differentiation rules do hold. Our next goal is to explore some of these conditions. We begin with a version of the product rule.

Theorem 8. *Let f and g be continuous functions defined on a neighborhood of a point c. If f has a Lanczos derivative at c and g has a derivative at c, then fg has a Lanczos derivative at c and $(fg)_\ell'(c) = f(c)g'(c) + f_\ell'(c)g(c)$.*

PROOF. Since g is differentiable at c, there exists a function η that is continuous in a neighborhood of 0 for which $\lim\limits_{h\to 0} \eta(h) = 0$ and

$$g(c+h) = g(c) + g'(c)\,h + \eta(h)\,h$$

for all values of h that are sufficiently close to 0. The following set of equations gives the gist of the proof:

$$\begin{aligned}
(fg)_\ell'(c) &= \lim_{h\to 0^+} \frac{3}{2h}\int_{-1}^{1} t\,f(c+ht)g(c+ht)\,dt \\
&= \lim_{h\to 0^+} \frac{3}{2h}\int_{-1}^{1} t\,f(c+ht)\big(g(c) + g'(c)\,ht + \eta(ht)\,ht\big)\,dt \\
&= \lim_{h\to 0^+} \Big(g(c)\cdot\frac{3}{2h}\int_{-1}^{1} t\,f(c+ht)\,dt + g'(c)\cdot\frac{3}{2}\int_{-1}^{1} t^2\,f(c+ht)\,dt \\
&\qquad\qquad + \frac{3}{2}\int_{-1}^{1} t^2\,\eta(ht)f(c+ht)\,dt\Big) \\
&= g(c)f_\ell'(c) + g'(c)f(c).
\end{aligned}$$

We need to show that the indicated values of the limits of the three terms are correct. The first limit follows easily from the definition of the Lanczos derivative of f at c. For the third term, observing that

$$\Big|\int_{-1}^{1} t^2\,\eta(ht)f(c+ht)\,dt\Big| \le \int_{-1}^{1} \big|\eta(ht)f(c+ht)\big|\,dt,$$

and recalling that f is continuous (and thus bounded) and that $\lim\limits_{h\to 0}\eta(h) = 0$ gives the desired result. Finally, for the middle term, we compute

$$
\begin{aligned}
\lim_{h\to 0^+} g'(c)\cdot\frac{3}{2}\int_{-1}^{1} t^2\, f(c+ht)\,dt
&= g'(c)\lim_{h\to 0^+}\Big(\frac{3}{2}\int_{-1}^{1} t^2\, f(c)\,dt + \frac{3}{2}\int_{-1}^{1} t^2\big(f(c+ht)-f(c)\big)\,dt\Big)\\
&= g'(c)f(c),
\end{aligned}
$$

where the continuity of f guarantees that the limit of the last integral is 0. This completes the proof. $\square$

It is not possible to weaken the hypotheses of Theorem 8 by assuming only that g is symmetrically differentiable at c. To see why, define functions f and g by

$$
f(x) = \begin{cases} 0, & \text{if } x \le 0;\\ \sqrt{x}\,\sin(\pi/x), & \text{if } x > 0; \end{cases} \qquad g(x) = \begin{cases} 0, & \text{if } x = 0;\\ \sqrt{|x|}\,\sin(\pi/|x|), & \text{if } x \ne 0; \end{cases}
$$

The function f is the function considered in Example 7 and the function g has a symmetric derivative of 0 at 0 since it is an even function. Since $fg = f^2$, it follows easily from Example 7 that the product rule formula is not valid for this pair of functions.

We turn now to the chain rule and alert the reader to the fact that the level of abstraction increases at this point in the paper. As a start, we consider the definitions of two concepts. The first is well-known, but the second has been created for the purpose of this paper.

Definition 9. Let α be a nonnegative number. A function f is α-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. Similarly, a function f is α-Hölder continuous on an interval I if there exists a constant M such that the inequality $|f(y)-f(x)| \le M|y-x|^\alpha$ is valid for all values of x and y in I.

If f is α-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 α-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 α-Hölder continuous at c if f is β-Hölder continuous at c. A simple collection of functions that illustrate the full range of α-Hölder continuity at 0 are functions of the form $f(x) = x^\alpha \sin(\pi/x)$ for $x \ne 0$ and $f(0) = 0$. Finally, note that a function with a bounded derivative on an interval I is 1-Hölder continuous on I.

Definition 10. Let $\beta \geq 1$ be a fixed real number and let f be a function that is differentiable at c. The function f has a β-strong derivative at c if there exists a function ϵ such that $\lim\limits_{h\to 0} \epsilon(h) = 0$ and

$$f(c+h) = f(c) + f'(c)\,h + \epsilon(h)\,h^{\beta}$$

for all values of h in some neighborhood of 0. □

If f is differentiable at c, then f has a 1-strong derivative at c; this is a simple consequence of the definition of the derivative. If f is twice differentiable at c, then f has a β-strong derivative at c for each β in the interval $[1,2)$. This follows from the fact that such an f satisfies

$$f(c+h) = f(c) + f'(c)\,h + \frac{f''(c)}{2}\,h^2 + \eta(h)\,h^2$$

for h near 0, where $\lim\limits_{h\to 0} \eta(h) = 0$. For some particular examples, the function $f(x) = x^5$ has a β-strong derivative at 0 for each β in the interval $[1,5)$. The function f defined by $f(x) = x^{4/3}\sin(1/x)$ for $x \neq 0$ and $f(0) = 0$ has a β-strong derivative at 0 for each β in the interval $[1,4/3)$. (This concept is somewhat related to the notion of Peano bounded as defined in [1], but the focus here is on the magnitude of the real number β.)

With these two concepts at hand, we can tackle the chain rule problem. Two versions are given, one in which the outside function is differentiable and one in which the inside function is differentiable.

Theorem 11. *Let f and g be continuous functions and let c be a point for which g is defined on a neighborhood of c and f is defined on a neighborhood of $g(c)$. Suppose that g is α-Hölder continuous at c and Lanczos differentiable at c and that f has a β-strong derivative at $g(c)$. If $\alpha\beta \geq 1$, then $f \circ g$ is Lanczos differentiable at c and $(f \circ g)'_{\ell}(c) = f'(g(c))g'_{\ell}(c)$.*

Proof. Since f has a β-strong derivative at $g(c)$, there exists a function ϵ such that $\lim\limits_{h\to 0} \epsilon(h) = 0$ and

$$f(g(c)+h) = f(g(c)) + f'(g(c))\,h + \epsilon(h)\,h^{\beta}$$

for all values of h in a sufficiently small neighborhood of 0. Since g is α-Hölder continuous at c, there exists $M > 0$ such that $|g(c+h) - g(c)| \leq M|h|^{\alpha}$ for all values of h in a sufficiently small neighborhood of 0. We first note that

$$
\begin{aligned}
&\int_{-1}^{1} t\, f(g(c+ht))\, dt \\
&= \int_{-1}^{1} t\,\big[f\big(g(c)+g(c+ht)-g(c)\big)-f(g(c))\big]\, dt \\
&= \int_{-1}^{1} t\big[f'(g(c))\big(g(c+ht)-g(c)\big)+\epsilon(g(c+ht)-g(c))(g(c+ht)-g(c))^{\beta}\big]\, dt \\
&= f'(g(c))\int_{-1}^{1} t\, g(c+ht)\, dt+\int_{-1}^{1} t\,\epsilon(g(c+ht)-g(c))(g(c+ht)-g(c))^{\beta}\, dt,
\end{aligned}
$$

as long as h is close to 0. Since

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

we just need to prove that

$$\lim_{h\to 0^+}\frac{3}{2h}\int_{-1}^{1} t\,\epsilon(g(c+ht)-g(c))(g(c+ht)-g(c))^{\beta}\, dt = 0.$$

The α-Hölder continuity of g at c implies that

$$
\begin{aligned}
&\Big|\frac{3}{2h}\int_{-1}^{1} t\,\epsilon(g(c+ht)-g(c))(g(c+ht)-g(c))^{\beta}\, dt\Big| \\
&\qquad\qquad \le \frac{3}{2|h|}\int_{-1}^{1} |t|\,|\epsilon(g(c+ht)-g(c))|\,\big(M|ht|^{\alpha}\big)^{\beta}\, dt \\
&\qquad\qquad \le \frac{3M^{\beta}}{2}|h|^{\alpha\beta-1}\int_{-1}^{1} |\epsilon(g(c+ht)-g(c))|\, dt.
\end{aligned}
$$

The final integral term goes to 0 as h goes to 0 due to the continuity of g and the properties of ϵ. Combining this fact with the conditions on α and β, we see that the entire quantity goes to 0 with h. □

Corollary 12. *Let f and g be continuous functions and let c be a point for which g is defined on a neighborhood of c and f is defined on a neighborhood of $g(c)$. Suppose that g is α-Hölder continuous at c and Lanczos differentiable at c and that f is twice differentiable at $g(c)$. If $\alpha > 1/2$, then $f\circ g$ is Lanczos differentiable at c and $(f\circ g)'_{\ell}(c) = f'(g(c))g'_{\ell}(c)$.*

To be consistent with the notation in Theorem 11 and Corollary 12, let g be the function (that was called f) in Example 7 and let $f(x) = x^2$. We know that the chain rule fails for $f \circ g$. Note that g is α-Hölder continuous at 0 for any α that satisfies $0 \le \alpha \le 1/2$ and that f has a β-strong derivative at $g(0) = 0$ for any β that satisfies $1 \le \beta < 2$. It follows that neither Corollary 12 nor Theorem 11 apply, although $\alpha\beta$ can be made as close to 1 as possible. This example shows that, in a certain sense, these are the best results possible for this type of chain rule.

We make the simple observation that a change of variables shows that

$$\frac{3}{2(-h)} \int_{-1}^{1} t\, f(c - ht)\, dt = \frac{3}{2h} \int_{-1}^{1} t\, f(c + ht)\, dt$$

for all (sufficiently small) nonzero values of h. It follows that

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

when f has a Lanczos derivative at c; that is, the limit can be assumed to be two-sided. This fact is used in the proof of the next theorem.

Theorem 13. *Let f and g be continuous functions and let c be a point for which g is defined on a neighborhood of c and f is defined on a neighborhood of $g(c)$. Suppose that f is α-Hölder continuous on a neighborhood of $g(c)$ and Lanczos differentiable at $g(c)$ and that g has a β-strong derivative at c. If $\alpha\beta \ge 1$, then $f \circ g$ is Lanczos differentiable at c and $(f \circ g)'_\ell(c) = f'_\ell(g(c))g'(c)$.*

Proof. The proof is similar to the proof of Theorem 11 but different enough to warrant its inclusion. Since g has a β-strong derivative at c, there exists a function ϵ such that $\lim_{h \to 0} \epsilon(h) = 0$ and

$$g(c + h) = g(c) + g'(c)\, h + \epsilon(h)\, h^\beta$$

for all values of h in a sufficiently small neighborhood of 0. Since f is α-Hölder continuous on a neighborhood of $g(c)$, there exists a positive number M such that

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

for all values of x and y in a sufficiently small neighborhood of 0. Once again, we begin by noting that

$$\begin{aligned}
&\int_{-1}^{1} t\, f(g(c+ht))\, dt \\
&\quad = \int_{-1}^{1} t\, f\big(g(c) + g'(c)\, ht + \epsilon(ht)\,(ht)^{\beta}\big)\, dt \\
&\quad = \int_{-1}^{1} t\, f\big(g(c) + g'(c)\, ht\big)\, dt \\
&\qquad\qquad + \int_{-1}^{1} t\,\big[f\big(g(c) + g'(c)\, ht + \epsilon(ht)\,(ht)^{\beta}\big) - f\big(g(c) + g'(c)\, ht\big)\big]\, dt
\end{aligned}$$

when h is near 0. Assuming that $g'(c) \neq 0$ (the $g'(c) = 0$ case is much easier), we see that

$$\begin{aligned}
&\lim_{h\to 0} \frac{3}{2h} \int_{-1}^{1} t\, f\big(g(c) + g'(c)\, ht\big)\, dt \\
&\qquad\qquad = g'(c) \cdot \lim_{h\to 0} \frac{3}{2[g'(c)\, h]} \int_{-1}^{1} t\, f\big(g(c) + [g'(c)\, h]t\big)\, dt \\
&\qquad\qquad = g'(c) f'_{\ell}(g(c)).
\end{aligned}$$

(This is where the observation made before the theorem is used.) To complete the proof, we need to confront the term

$$\frac{3}{2h} \int_{-1}^{1} t\,\big[f\big(g(c) + g'(c)\, ht + \epsilon(ht)\,(ht)^{\beta}\big) - f\big(g(c) + g'(c)\, ht\big)\big]\, dt.$$

The α-Hölder continuity of f on a neighborhood of $g(c)$ shows that the magnitude of this term is bounded above by

$$\frac{3}{2|h|} \int_{-1}^{1} |t|\, M \big|\epsilon(ht)(ht)^{\beta}\big|^{\alpha}\, dt \leq \frac{3M}{2} |h|^{\alpha\beta - 1} \int_{-1}^{1} |\epsilon(ht)|^{\alpha}\, dt.$$

As in the proof of Theorem 11, the conditions on α and β imply that this quantity goes to 0 with h. □

Since the condition that f be α-Hölder continuous on an interval is quite a bit stronger than f being α-Hölder continuous at a point, it is worth noting that the hypotheses of Theorem 13 can be weakened to f being α-Hölder continuous at $g(c)$ in the special case in which $g'(c) = 0$. A careful reading of the proof reveals the validity of this fact. It should also be noted that the hypotheses of Theorem 13 are not necessary for the conclusion to hold. The relevant inequalities in the proof involve moving an absolute value into an integral; this is somewhat analogous to working with absolutely convergent series rather than conditionally convergent series. Referring once again to the function f from Example 7, we note that $f(x^2)$ is Lanczos differentiable at 0 even though the functions do not satisfy the hypotheses of Theorem 13 (the product $\alpha\beta$ is less than 1). The details, which are quite similar to those of Example 7, are left to the reader.

Corollary 14. *Let f and g be continuous functions and let c be a point for which g is defined on a neighborhood of c and f is defined on a neighborhood of $g(c)$. Suppose that g is differentiable at c and that f is Lanczos differentiable at $g(c)$. If f has bounded difference quotients on a neighborhood of $g(c)$, then $f \circ g$ is Lanczos differentiable at c and $(f \circ g)'_\ell(c) = f'_\ell(g(c))g'(c)$.*

The condition on f in the corollary is equivalent to f being 1-Hölder continuous on a neighborhood of $g(c)$; the statement of the corollary just uses more familiar terms. It should be mentioned that a differentiable function need not have bounded difference quotients on an interval. As an example, let $f(x) = x^2 \sin(\pi/x^2)$ for $x \neq 0$ and $f(0) = 0$. Then f is differentiable on $[-1, 1]$ but f does not have bounded difference quotients in any neighborhood of 0. Note also that f' is not continuous at 0; it is this fact that opens the door to such examples. An example of a function f that satisfies the hypotheses of Corollary 14 is $|x|$; the reader should have no trouble finding similar examples. However, an open question that arises in the context of this corollary is whether or not the following statement is valid:

> If f is Lanczos differentiable at c and if f has bounded difference quotients on a neighborhood of c, then f has a symmetric derivative at c.

Both a search for a counterexample and an attempt at a proof of this statement have been fruitless.

Since the chain rule is related to the product rule (by interpreting the product $f \cdot f$ as the composite function f^2) and the quotient rule (via the composite function $1/f$), we can obtain versions of these rules from the chain rules we have proved. In particular, Corollary 12 implies the following variations on the product rule and the quotient rule for Lanczos derivatives.

Theorem 15. *Let f and g be continuous functions defined on a neighborhood of a point c and suppose that both f and g are α-Hölder continuous at c for some $\alpha > 1/2$. If f and g are Lanczos differentiable at c, then fg is Lanczos differentiable at c and $(fg)'_\ell(c) = f(c)g'_\ell(c) + f'_\ell(c)g(c)$.*

Proof. The functions $f+g$ and $f-g$ are both α-Hölder continuous at c. By Corollary 12 and the linearity of the Lanczos derivative, the functions $(f+g)^2$ and $(f-g)^2$ are Lanczos differentiable at c. The theorem follows by writing

$$fg = \frac{1}{4}\,(f+g)^2 - \frac{1}{4}\,(f-g)^2$$

and once again using the linearity of the Lanczos derivative. □

Theorem 16. *Let f and g be continuous functions defined on a neighborhood of a point c for which $g(c) \neq 0$ and suppose that both f and g are α-Hölder continuous at c for some $\alpha > 1/2$. If f and g are Lanczos differentiable at c, then f/g is Lanczos differentiable at c and*

$$(f/g)'_\ell(c) = \big(g(c)f'_\ell(c) - f(c)g'_\ell(c)\big)/g(c)^2.$$

Proof. We first show that $1/g$ is α-Hölder continuous at c. To do so, let $M > 0$ be an appropriate constant for the α-Hölder continuity of g at c, then choose a neighborhood of 0 for which $|g(c+h) - g(c)| \le M|h|^\alpha$ and $|g(c+h)| > |g(c)|/2$ for all h in this neighborhood. For these values of h, we then have

$$\Big|\frac{1}{g(c+h)} - \frac{1}{g(c)}\Big| = \Big|\frac{g(c+h)-g(c)}{g(c)g(c+h)}\Big| \le \frac{2M}{|g(c)|^2}\,|h|^\alpha.$$

This shows that $1/g$ is α-Hölder continuous at c. The reciprocal rule (the derivative formula for $1/g$) then holds for g as a consequence of Corollary 12. Finally, the product rule (Theorem 15) can be applied to the functions f and $1/g$ to complete the proof. □

As mentioned in the introduction of the paper, we have limited ourselves to a consideration of continuous functions. The interested reader may want to see what happens if this condition is dropped. In addition, we have not considered questions such as the following:

> Suppose that f has a Lanczos derivative at each point of an interval I. What can be said about the existence of the symmetric derivative of f? In particular, does it follow that f must have a symmetric derivative at almost all of the points of I?

Since results such as this are known to hold for other derivatives, there may be some interesting possibilities here for further study.

These last results and comments bring to a close our discussion of the first order Lanczos derivative. We now return to the motivating idea behind the Lanczos derivative and put it into a different perspective with the hope of determining how to define higher order Lanczos derivatives. Since our purpose is to indicate how higher order Lanczos derivatives can be defined rather than present an in-depth study of them, some of the details will be streamlined. (See [10] and [2] for similar approaches, each with a different perspective.) Recall that an inner product on the space $\mathcal{C}([-1,1])$ of continuous functions defined on the interval $[-1,1]$ can be defined by

$$\langle f, g\rangle = \int_{-1}^{1} f(t)g(t)\,dt.$$

Suppose, as we have done throughout this paper, that f is a continuous function defined on a neighborhood of a point c. For each value of h that is sufficiently close to 0, define a function g_h by $g_h(t) = f(c+ht)$. These functions are defined and continuous on $[-1,1]$ when h is small. Note that

$$\frac{3}{2}\int_{-1}^{1} t\,g_h(t)\,dt = \frac{\int_{-1}^{1} t\,g_h(t)\,dt}{\int_{-1}^{1} t^2\,dt} = \frac{\langle g_h(t), t\rangle}{\langle t, t\rangle} = \frac{\langle g_h, P_1\rangle}{\langle P_1, P_1\rangle},$$

where P_1 is the Legendre polynomial of degree 1. (We will review these polynomials in just a moment.) It follows that

$$f_\ell'(c) = \lim_{h\to 0^+} \frac{3}{2h}\int_{-1}^{1} t\,f(c+ht)\,dt = \lim_{h\to 0^+} \frac{1}{h}\cdot\frac{\langle g_h, P_1\rangle}{\langle P_1, P_1\rangle}.$$

What happens in this formula if we use the nth degree Legendre polynomial P_n rather than P_1? Is it possible to obtain an nth order derivative in this way? An affirmative answer depends on the properties of the Legendre polynomials.

The Legendre polynomials form an orthogonal set in the inner product space $\mathcal{C}([-1,1])$. There are several ways to obtain these polynomials, each approach exhibiting a different feature. For our purposes, it is sufficient to note that Rodrigues' formula shows that the Legendre polynomial P_n is defined by

$$P_n(x) = \frac{1}{2^n n!}\frac{d^n}{dx^n}(x^2-1)^n.$$

The first few Legendre polynomials are

$$P_0(x) = 1, \quad P_1(x) = x, \quad P_2(x) = \frac{1}{2}\left(3x^2-1\right), \quad P_3(x) = \frac{1}{2}\left(5x^3-3x\right).$$

We make note of the following properties of Legendre polynomials.

1. P_n is a polynomial of degree n with leading coefficient $\dfrac{(2n)!}{2^n(n!)^2}$.

2. P_n is an even function when n is even and an odd function when n is odd.

3. $P_n(1) = 1$ and $P_n(-1) = (-1)^n$ for all nonnegative integers n.

4. $\displaystyle\int_{-1}^{1} w(t)P_n(t)\,dt = \frac{(-1)^n}{2^n n!}\int_{-1}^{1} w^{(n)}(t)(t^2-1)^n\,dt$, assuming that w has a continuous nth derivative.

5. $\displaystyle\int_{-1}^{1} t^m P_n(t)\,dt = 0$ whenever m is a nonnegative integer less than n.

6. $\displaystyle\int_{-1}^{1} t^n P_n(t)\,dt = \frac{2^{n+1}(n!)^2}{(2n+1)!}$.

7. $\displaystyle\int_{-1}^{1} P_n(t)^2\,dt = \frac{2}{2n+1}$.

These properties of Legendre polynomials are not difficult to prove and are listed in an order for which later properties depend on previous properties. Property (4) is simply repeated integration by parts. The most difficult property to prove is property (6), but a change of variables reduces the problem to the evaluation of a well-known integral involving integral powers of the sine function. If you get stuck verifying any of these properties, there are many resources available for this particular collection of orthogonal polynomials.

With this material as background, we make the following definition.

Definition 17. Suppose that f is a continuous function defined on a neighborhood of a point c. The nth order Lanczos derivative of f at c is defined by

$$\begin{aligned} f_\ell^{(n)}(c) &= \frac{(2n)!}{2^n n!}\lim_{h\to 0^+}\frac{1}{h^n}\cdot\frac{\langle g_h, P_n\rangle}{\langle P_n, P_n\rangle} \\ &= \frac{(2n+1)!}{2^{n+1}n!}\lim_{h\to 0^+}\frac{1}{h^n}\int_{-1}^{1} f(c+ht)\,P_n(t)\,dt, \end{aligned}$$

provided that the limit exists.

To obtain a sense for how this higher order derivative operates, suppose that for h near 0, the function f can be expressed as

$$f(c+h) = a_0 + a_1\,h + \frac{a_2}{2}\,h^2 + \cdots + \frac{a_{n-1}}{(n-1)!}\,h^{n-1} + \frac{a_n}{n!}\,h^n + \epsilon(h)\,h^n,$$

where $\lim_{h\to 0}\epsilon(h) = 0$. This sort of expression is certainly valid if f happens to have an nth derivative at c; it is Taylor's Theorem if f is $n+1$ times differentiable on a neighborhood of c. (A reader who has experience with other derivatives may recognize that f can be written in this form if it has an nth order Peano derivative at c, see [4].) In any event, we find that

$$\begin{aligned}
&\frac{1}{h^n}\int_{-1}^{1} f(c+ht)\,P_n(t)\,dt \\
&\qquad = \frac{1}{h^n}\Big(\sum_{k=0}^{n}\frac{a_k}{k!}\,h^k\int_{-1}^{1} t^k\,P_n(t)\,dt + h^n\int_{-1}^{1}\epsilon(ht)\,t^n P_n(t)\,dt\Big) \\
&\qquad = \frac{1}{h^n}\Big(\frac{a_n}{n!}\,h^n\int_{-1}^{1} t^n\,P_n(t)\,dt + h^n\int_{-1}^{1}\epsilon(ht)\,t^n P_n(t)\,dt\Big) \\
&\qquad = \frac{a_n}{n!}\cdot\frac{2^{n+1}(n!)^2}{(2n+1)!} + \int_{-1}^{1}\epsilon(ht)\,t^n P_n(t)\,dt \\
&\qquad \to \frac{2^{n+1}n!}{(2n+1)!}\,a_n
\end{aligned}$$

as $h\to 0^+$. Thus, if f has an nth order derivative (or even an nth order Peano derivative) at c, then f has an nth order Lanczos derivative at c and the values are the same. As we have seen in our earlier work, the converse is false even for the $n=1$ case.

This approach to higher order differentiation was first mentioned in an article by Haslam-Jones [7]. Haslam-Jones formulates these ideas in a very general setting, but his work appears to have caught the attention of very few people. In addition to defining a number of higher order derivatives that all include the Peano derivative, he was able to obtain a similar expression that is actually equivalent to the Peano derivative. In particular, it can be shown that the nth order Peano derivative of f at c is given by

$$\frac{(2n-1)!}{2^{n-1}(n-1)!}\lim_{h\to 0}\frac{1}{h^n}\left(f(c+h) - \frac{1}{2}\int_{-1}^{1} f(c+ht)\big(P_n'(t) + P_{n-1}'(t)\big)\,dt\right).$$

If f has an nth order Peano derivative at c, then it is not difficult to show that the above limit yields the same value. However, a proof of the converse (see [7]) is nontrivial. A more accessible proof of the same result can be found in [5]. We leave a detailed study of these higher order Lanczos derivatives to the interested reader.

References

[1] J. Marshall Ash, *Generalizations of the Riemann derivative*, Trans. Amer. Math. Soc. **126** (1967), 181–199.

[2] N. Burch, P. Fishback, *Orthogonal polynomials and regression-based symmetric derivatives*, Real Anal. Exchange **32** (2007), 597–607.

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

[4] M. Evans, C. Weil, *Peano derivatives: a survey*, Real Anal. Exchange **7** (1982), 5–23.

[5] R. Gordon, *Peano differentiation via integration*, Real Anal. Exchange, **34** (2009), 147–156.

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

[7] U. S. Haslam-Jones, *On a generalized derivative*, Quart. J. Math, Oxford Ser. (2) **4** (1953), 190–197.

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

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

[10] S. K. Rangarajan, S. P. Purushothaman, *Lanczos' generalized derivative for higher orders*, J. Comput. Appl. Math. **177** (2005), 461–465.