The Electronic Journal of Mathematics and Technology, Volume 4, Number 3, ISSN 1933-2823

Making Teaching Calculus Accessible

Qun LIN
Institute of Computational Mathematics,
Chinese Academy of Sciences
P.O. Box 2719, Beijing 100190, CHINA
E-mail: linq@lsec.cc.ac.cn

Wei-Chi YANG
Department of Mathematics
and Statistics
Radford University
Radford, Virginia 24060, USA
E-mail: wyang@radford.edu

Abstract

We give an intuitive definition for derivative of a function at a point without using the epsilon and delta concept, which leads to the concept of integration and the outcome of the Fundamental Theorem of Calculus (FTC). We include the use of graphical representations to strengthen the understandings of our definitions on differentiation and integration. All of our constructions do not depend on axiom of completeness of real numbers. This paper is intended for two purposes: First is to makes the concepts of differentiation and integration more accessible to beginners who start learning Calculus; second is to give those who have learned Calculus and beyond to ponder if approaches described in this article can be adapted in teaching Calculus in their respective countries.

1 The Derivative and Slope Difference

Consider a 'smooth function' f in an open interval containing a point at $x = x_0$ below in Figure 1:

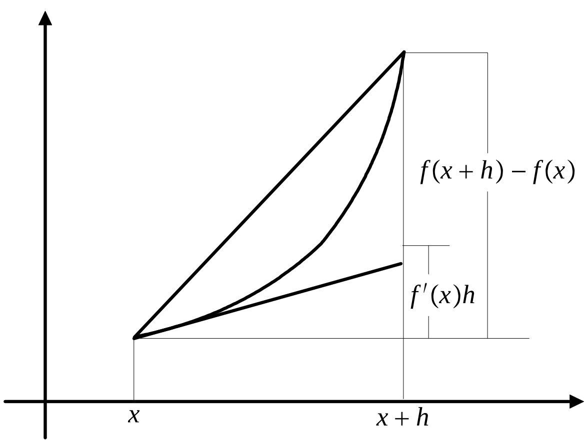


Figure 1. The height of secant

The difference between $f(x)$ and $f(x+h)$ is the height between the secant line connecting the fixed point $(x, f(x))$ and its neighboring point $(x+h, f(x+h))$. In other words, the height of the secant is $f(x+h)-f(x)$, where the base h is a variable. If we fix the point $P=(x, f(x))$, and make the point $Q=(x+h, f(x+h))$ get closer and closer to P (or make $|h|$ smaller and smaller), will the slope of the secant reach a number when $|h| \rightarrow 0$? If the answer is affirmative, then the slope of the secant PQ will get closer and closer to the slope of the tangent line at P, and we call such number the derivative of f at x or simply $f'(x)$. If f' exists at every point throughout an open interval (a, b), then the function f is called smooth or differentiable over the interval (a, b). This is what most students encounter in their first year Calculus course. We would like to motivate the idea of $f'(x)$ without using limit. We first note the slope of the secant line PQ is $\frac{f(x+h)-f(x)}{h}$. If the slope of the tangent line at $P=(x, f(x))$ exists, let's say $A(x)$, then the difference between $\frac{f(x+h)-f(x)}{h}$ and $A(x)$ should be small. In other words, we would like to see how we can control the quantity $\left|\frac{f(x+h)-f(x)}{h}-A(x)\right|$ and make it small. We start with a polynomial say $f(x)=x^n$ for some positive integer n. We observe that

$$\begin{aligned}\frac{f(x+h)-f(x)}{h} &= \frac{\binom{n}{1}x^{n-1}h+\binom{n}{2}x^{n-2}h^2+\ldots+\binom{n}{n}h^n}{h} \\ &= nx^{n-1}+h\left(\binom{n}{2}x^{n-2}h+\ldots+\binom{n}{n}h^{n-1}\right).\end{aligned}$$

If we can make the quantity $\left|h\left(\binom{n}{2}x^{n-2}h+\ldots\binom{n}{n}h^{n-1}\right)\right|$ small when h is small, then the quantity $\left|\frac{f(x+h)-f(x)}{h}-nx^{n-1}\right|$ should be small too. Alternatively, if we write

$$\frac{f(t)-f(x)}{t-x}=\frac{(t-x)\left(t^{n-1}+t^{n-2}x+t^{n-3}x^2+\cdots+tx^{n-2}+x^{n-1}\right)}{t-x}$$

and note that $\frac{f(t)-f(x)}{t-x}=\left(t^{n-1}+t^{n-2}x+t^{n-3}x^2+\cdots+tx^{n-2}+x^{n-1}\right)$ if $t \neq x$. If we let t get close to x, we see the expression $\left|\frac{f(t)-f(x)}{t-x}-nx^{n-1}\right|$ is small. Let us explore the following examples from algebraic point of view as follows.

Example 1 *For polynomials such as $f(x)=x^2$, the slope of the secant in the neighboring of the fixed point x is*

$$\frac{(x+h)^2-x^2}{h}=2x+h.$$

By making $|h|$ small but $h \neq 0$ in $2x+h$ above, the right hand side contains the constant dominant term, is called the derivative $f'(x)=2x$ (which is independent of h).

Example 2 *For rational polynomials such as $f(x)=\frac{1}{x}$ for $x \neq 0$, the slope of the secant in the neighboring of the fixed point x is*

$$\frac{\frac{1}{x+h}-\frac{1}{x}}{h}=\frac{-1}{x(x+h)}=\frac{-1}{x^2}+\frac{1}{x^2(x+h)}h,$$

Notice that the absolute value of the ending term, $\left|\frac{\frac{1}{x+h}-\frac{1}{x}}{h}+\frac{1}{x^2}\right| = \left|\frac{1}{x^2(x+h)}h\right| \leq C|h|$ *(where* $C = \frac{2}{|x|^3}$ *with* $|h| \leq \frac{|x|}{2}$*). By making* $|h|$ *small in the tail term,* $\frac{1}{x^2(x+h)}h$*, the constant dominant term* $\frac{-1}{x^2}$ *is called the derivative* $f'(x) = \frac{-1}{x^2}$ *(which is independent of* h*).*

Example 3 *For the radicals such as* $f(x) = \sqrt{x}$ *(*$x > 0$ *or* $x \in (a, \infty)$*, with* $a > 0$*), the slope of the secant in the neighboring of* x *is*

$$\frac{\sqrt{x+h}-\sqrt{x}}{h} = \frac{1}{\sqrt{x+h}+\sqrt{x}} = \frac{1}{2\sqrt{x}} + \frac{-h}{2\sqrt{x}(\sqrt{(x+h)}+\sqrt{x})^2}.$$

We note the tail

$$\left|\frac{\sqrt{x+h}-\sqrt{x}}{h} - \frac{1}{2\sqrt{x}}\right| = \left|\frac{-h}{2\sqrt{x}(\sqrt{(x+h)}+\sqrt{x})^2}\right| \leq \frac{|h|}{2x\sqrt{x}} = C\,|h|\,,$$

where $C = \frac{1}{2x\sqrt{x}}$ *or* $\frac{1}{2a\sqrt{a}}$ *when* $|h| \leq a$*. By making* $|h|$ *small into* $\frac{-h}{2\sqrt{x}(\sqrt{(x+h)}+\sqrt{x})^2}$*, the constant dominant term* $\frac{1}{2\sqrt{x}}$ *is called the derivative* $f'(x) = \frac{1}{2\sqrt{x}}$ *(which is independent of* h*). Alternatively, we may write*

$$\frac{\sqrt{x}-\sqrt{a}}{x-a} = \frac{\sqrt{x}-\sqrt{a}}{(\sqrt{x}-\sqrt{a})\,(\sqrt{x}+\sqrt{a})} = \frac{1}{\sqrt{x}+\sqrt{a}},$$

and if we substitute $x = a$ *in the last expression above we get the derivative of* $\sqrt{x}$ *at* $x = a$ *to be* $\frac{1}{2\sqrt{a}}$*.*

Example 4 *For the trigonometric functions such as* $f(x) = \sin x$*, we consider*

$$\begin{aligned}\frac{\sin(x+h)-\sin x}{h} &= \frac{\sin x\cos h + \sin h\cos x - \sin x}{h} \\ &= \frac{\sin x(\cos h - 1)}{h} + \frac{\sin h \cos x}{h}.\end{aligned}$$

We observe that $1-\cos h = 2\sin^2\frac{h}{2} \leq \frac{h^2}{2}$*, and* $\left|\frac{\cos(h)-1}{h}\right| \leq \frac{|h|}{2}$*. Furthermore, since* $\cos(h) < \frac{\sin(h)}{h} < 1$*,* $\frac{\sin(h)}{h}$ *is close to* 1 *when* h *is small. Therefore,*

$$\left|\frac{\sin(x+h)-\sin x}{h} - \frac{\sin h\cos x}{h}\right| = \left|\frac{\sin x(\cos h-1)}{h}\right| \leq \frac{|h|}{2}.$$

By making $|h|$ *small, we call the derivative of* $\sin x$ *at* x *to be* $\cos x$*.*

We see from Examples 1 to 4 that, in general, for most *elementary functions* (such as polynomials, rational functions, radical functions, and trigonometric functions) we can use algebraic identities to figure out their respective derivatives (the largest constant) of respective functions; although sometimes we need to elaborate on finding proper inequalities. It is expected that we need more differentiation rules to generate the derivatives of arbitrary elementary functions from the those of basic elementary functions, which we will skip in this article.

Indeed, the Examples 1 through 4 above give us motivation of how we should formulate our general definition of the derivative for a function at a point. The derivative at a fixed point x is obtained from the identity expanded by a dominant term, $A(x)$, and a tail term, $r(x, h)$. In other words, we can write

$$\frac{f(x+h)-f(x)}{h} = A(x) + r(x,h) \text{ or} \tag{1}$$

$$f(x+h) - f(x) - A(x)h = r(x,h)\cdot h. \tag{2}$$

Alternatively, we may write

$$\frac{f(t)-f(x)}{t-x} = A(x) + r(x,t) \text{ or} \tag{3}$$

$$f(x) - f(t) - A(x)(x-t) = r(x,t)\cdot(x-t). \tag{4}$$

Notice that the tail function $r(x,h) = \frac{f(x+h)-f(x)}{h} - f'(x)$ has the obvious geometry sense, it is the difference between the slope of the secant line (between two points) and that of the tangent at a given point. We shall call the expression $r(x, h)$ as *the slope difference*. When $|h|$ gets smaller, the slope difference $r(x, h)$ gets closer to 0, which is consistent with the the concept using epsilon and delta.

We first introduce the notation of $a << 1$ if a is a real number and a is smaller than any arbitrary positive small number. Obviously, this is to avoid the language of using epsilon and delta, which is also a convenient way for us to say that our following Definition 5 is of no difference than the regular epsilon delta concept. More importantly, we would like to explore how to control the tail function $r(x, h)$ using Definition 6 below. As expected, we have the following

Definition 5 *We say that a function f is differentiable at x if there exists a function $A(x)$, and a neighborhood $(x-\delta, x+\delta)$ of x, with $\delta > 0$, and*

$$|r(x,h)| = \left|\frac{f(x+h)-f(x)}{h} - A(x)\right| << 1 \tag{5}$$

for all $|h| < \delta$, or

$$|r(x,t)| = \left|\frac{f(x)-f(t)}{x-t} - A(x)\right| << 1 \tag{6}$$

for all $t \in (x-\delta, x+\delta)$.

We will examine what it means by $|r(x, h)| << 1$ in more details; the treatment of $|r(x, t)| << 1$ can be done analogously, which we will skip here. For those elementary functions demonstrated in Examples 1 through 4, we have seen that in order to make $|r(x, h)|$ as small as we wish, it will be easier if $|r(x, h)|$ can be bounded above by $C|h|$, where C is a constant. More precisely, we consider

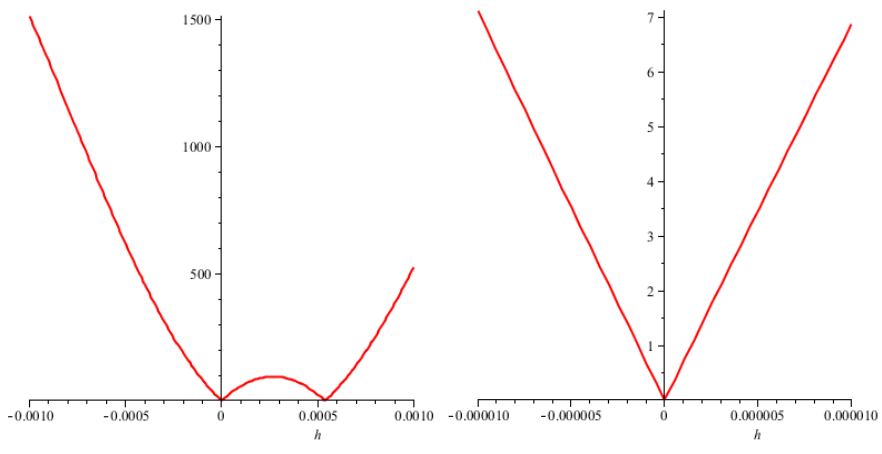
Definition 6 *Let $x \in (a.b)$, and we say that a function f is differentiable at x if there exists a function $A(x), C > 0$, and $\delta > 0$ such that*

$$|r(x,h)| = \left|\frac{f(x+h)-f(x)}{h} - A(x)\right| \le C|h|, \tag{7}$$

for each h for which $|h| < \delta$ and $(x - |h|, x + |h|) \subset (x - \delta, x + \delta) \subset (a, b)$. In other words, by shortening $|h|$, we call the ending term $A(x)$ in (7) the derivative of f at x, which we denote it by $f'(x) = A(x)$. It is important to note that the derivative is unique whenever it exists (see [1] on this page).

1.1 Comparing the Suggested Definition with the Standard Definition

1. We note that our Definition 6 is more restrictive than Definition 5 or the ordinary one which uses epsilon and delta notations. Noticeably, we need to find the extra constant $C > 0$ in the Definition 6. We will see Definition 6 is actually useful and sufficient for many elementary functions. To lighten up the burden of working out the constant C algebraically, we demonstrate here the meaning and the importance of selecting proper values for C and h in order for the inequality $|r(x_0, h)| \leq C|h|$ to be valid for a fixed x_0. Intuitively, if a function f is differentiable, by zooming in the graph of $|r(x_0, h)|$, it should look like that of an absolute function of $C|h|$ with proper values of C and h respectively. In other words, if a function f is differentiable at a point $x = x_0$ we expect the graph of $|r(x_0, h)|$ to be similar to an absolute function $y = C|h|$ when proper values of C and range of h are found. It is not difficult to conjecture that the required respective values for C and h will be different for each different x. For example, we consider $f(x) = \sin(x^2)$ and choose x_0 to be a large number, say $x_0 = 1,000$, then the slope difference is $r(x_0, h) = \frac{f(x_0+h)-f(x_0)}{h} - 2x\cos x^2$. Then the plot of $|r(1000, h)|$ when $h \in [-0.001, 0.001]$ can be seen in Figure 2(a) below. We see that the graph of $|r(1000, h)|$ does not look like that of an absolute function of h yet when $h \in [-0.001, 0.001]$. On the other hand, if we choose $h \in [-0.00001, 0.00001]$, we see from Figure 2(b) that the function $|r(1000, h)|$ is similar to the graph of an absolute function (combination of two lines with the proper slope values C and $-C$ respectively). In such case, we have found an desired neighborhood and a constant C. We leave it to the readers to verify that the required C and h respectively will be more stringent when x gets larger, which is intuitive because the graph of $y = \sin(x^2)$ gets denser when x gets larger.



Figures 2(a) and 2(b) $r(1000, h)$ with various $h's$.

[1] If there are two tangents with respective slopes A_1 and A_2, we set $|A_1 - A_2| = d > 0$, then there exists h such that $|r_1| < \frac{d}{2}$, $|r_2| < \frac{d}{2}$, $|A_1 - A_2| = |r_1 - r_2| < d$, which is a contradiction.

2. We use the slope difference when discussing the differentiability of f' at one point, which is to consider if the inequality (5), (6) or (7) holds. The advantage is we can extend this idea to Section 3 when discussing the condition of integrability of f' over an interval, which will be again involving the slope difference, see (11).

3. For those readers who have knowledge beyond Calculus. We know that if a function is uniformly differentiable or f continuously differentiable over a fixed interval, then we have the following stronger inequality

$$|r(x,h)| \leq r(h) << 1, \tag{8}$$

where $r(h)$ does not depend on x.

4. The Definition 6 is a special case of Livshits [10], where author defines the derivative without using limits by requiring the estimates

$$\left| \frac{f(x+h) - f(x)}{h} - f'(x) \right| \leq m(|h|)$$

for a suitable function m (a modulus of continuity), such as $K|h|$ or $K|h|^{\alpha}$. We note that Zhang [19] also presents an elementary method (so called Zhang's inequality).

2 The Fundamental Theorem of Calculus

2.1 Motivations

The slope of the tangent (if exists) deals with the curve locally only at a point. We now turn to studying a curve globally. We first divide the curve into many short curves (there is no need to subdivide the domain evenly) and then each short curve is replaced by the tangent at a node (or at any interior point), where we assume that the tangent exists everywhere or we have |slope difference| $\ll 1$ on each subinterval.

We note from Figure 3 that the height between two points on the curve can be approximated by the total heights of all the tangents at a given point in each subinterval.

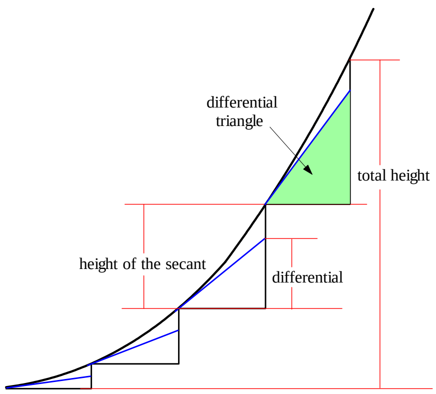


Figure 3. Using differentials to approximate the total height

Let f be a differentiable function and dx be any given real number, and dy be defined in terms of dx. We call dy simply the *differential* (when there is no confusion with dx). Consequently, the height of a tangent at a point x to a curve is the differential or dy at the point x. In other words, we have

$$\text{differential} = \text{the height of tangent at } x \text{ or}$$
$$dy = f'(x)dx = f'(x)h.$$

The associated triangle with respect to the tangent is called the *differential triangle*-see the shaded triangle in Figure 4 below. Then the total height of the curve between two points $(a, f(a))$ and $(b, f(b))$, can be approximated by sum of the height of corresponding differential triangles. Equivalently, we can use the sum of differentials to estimate the height between these two points, $f(b) - f(a)$.

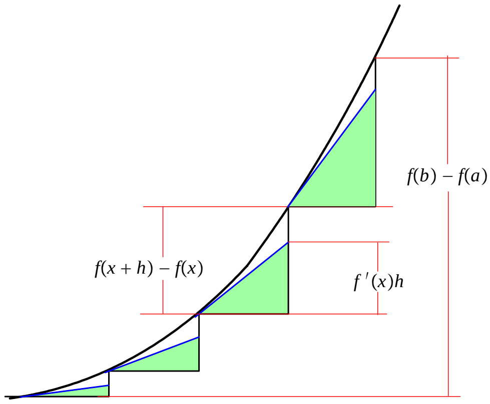


Figure 4. Differential Triangles

In other words, it follows from Figure 4 that the total height, $f(b) - f(a)$, can be approximated by the sum of the heights of small triangles, $\sum_{i=1}^{n} f'(x_i)h_i$. First, let's consider the difference of these two quantities as follows:

$$\begin{aligned}
&\text{Total height} - \text{Sum of differentials} \\
&= (f(b) - f(a)) - \left(\sum_{i=1}^{n} f'(x_i)h_i\right) \\
&= \sum_{i=1}^{n} [(f(x_i + h_i) - f(x_i) - f'(x_i)h_i] \\
&= \sum_{i=1}^{n} r(x_i, h_i)h_i \qquad (9)
\end{aligned}$$

We note the equation (9) is an identity which involves the slope difference again. We investigate (9) further and make the following observation

$$\begin{aligned}
f(b) - f(a) - \sum_{i=1}^{n} f'(x_i)h_i &= \sum_{i=1}^{n} [(f(x_i + h_i) - f(x_i) - f'(x_i)h_i] \\
&= (b-a)\sum_{i=1}^{n}\left(\frac{h_i}{b-a}\right)\left(\frac{f(x_i+h_i) - f(x_i)}{h_i} - f'(x_i)\right), \qquad (10)
\end{aligned}$$

where the coefficients $\dfrac{h_i}{\sum_{i=1}^n h_i}, i = 1, 2, ...n$ are all positive and sum up to be 1, so we call $\sum_{i=1}^n \left(\dfrac{h_i}{b-a}\right)\left(\dfrac{f(x_i+h_i)-f(x_i)}{h_i} - f'(x_i)\right)$ **the average of the slope differences** of $\left\{\dfrac{f(x_1+h_1)-f(x_1)}{h_1} - f'(x_1), \dfrac{f(x_2+h_2)-f(x_2)}{h_2} - f'(x_2), ..., \dfrac{f(x_n+h_n)-f(x_n)}{h_n} - f'(x_n)\right\}$.

Therefore, we suggest consider the following

Definition 7 *If f is continuous over $[a, b]$ and differentiable over (a, b). We say that f' is Riemann integrable over $[a, b]$ if*

$$\left|\sum_{i=1}^n \left(\frac{h_i}{b-a}\right)\left(\frac{f(x_i+h_i)-f(x_i)}{h_i} - f'(x_i)\right)\right| \ll 1, \tag{11}$$

for any partition $a = x_1 < x_2 < ... < x_n = b$, with $x_{i+1} - x_i = h_i$,where $i = 1, 2, ...n-1$, and we write

$$\int_a^b f' = f(b) - f(a). \tag{12}$$

In short, the integrability for f' is requiring the average of the slope differences to be arbitrary small.

Remarks:

1. It is easy to see that if f' is bounded in a closed interval and satisfying 11, then f is Riemann integrable.

2. The advantage of the Definition 7 above is that the fundamental theorem of calculus is being imbedded in the definition.

3. In view of the Definition 7 above, if we can find one partition in $[a, b]$ such that

$$\left(\frac{h_i}{b-a}\right)\left(\frac{f(x_i+h_i)-f(x_i)}{h_i} - f'(x_i)\right) \tag{13}$$

   can not be made as small as possible for some $1 \le i \le n$, then f' is not Riemann integrable over $[a, b]$, see Example 13 when considering $f(x) = \sqrt{x}$ over $[0, 1]$.

We now turn our attention to describing the condition for the integrability for an elementary function. If f is an elementary function, f is differentiable at x, and we have

$$\left|\frac{f(x+h)-f(x)}{h} - f'(x)\right| = |r(x, h)| \le |r(h)| \le C_1 \cdot h,$$

where C_1 is a constant and is independent of x. For any partition $a = x_1 < x_2 < ... < x_n = b$, with $x_{i+1} - x_i = h_i$, we have

$$|\text{the maximum of } \{|r(h_i)|\}_{i=1}^n| \le C \cdot (\max\{|h_i|\}_{i=1}^n),$$

where C is a constant and is independent of x, and hence we have

$$|\text{the average of the slope differences}| \leq C \cdot \left(\max\left\{|h_i|\right\}_{i=1}^{n}\right) \text{ or}$$
$$\left|\sum_{i=1}^{n}\left(\frac{h_i}{b-a}\right)\left(\frac{f(x_i+h_i)-f(x_i)}{h_i}-f'(x_i)\right)\right| \leq C \cdot \left(\max\left\{|h_i|\right\}_{i=1}^{n}\right).$$

Therefore

$$\begin{aligned}&\left|f(b)-f(a)-\sum_{i=1}^{n}f'(x_i)h_i\right| \\ &= \left|(b-a)\sum_{i=1}^{n}\left(\frac{h_i}{b-a}\right)\left(\frac{f(x_i+h_i)-f(x_i)}{h_i}-f'(x_i)\right)\right| \\ &\leq \left(C\cdot\left(\max\left\{h_i\right\}_{i=1}^{n}\right)\right)\cdot(b-a). \end{aligned} \tag{14}$$

Hence, if f is an elementary function over an interval $[a, b]$, it is clear that the total error $|f(b) - f(a) - \sum_{i=1}^{n} f'(x_i)h_i|$ is small, f' is Riemann integrable and we have $\int_a^b f' = f(b) - f(a)$.

Remark: For those readers who have knowledge beyond Calculus. It is clear that finding the maximum slope difference described in (14) over a closed interval is equivalent to requiring the uniformity for the slope difference. This is equivalent of requiring a function f to be uniformly differentiable over an interval (which implies that f' is uniformly continuous there) or demanding f to have continuous derivative in a closed interval. We also remind readers that we do allow a function f to be discontinuous at finitely many points over an interval $[a, b]$, so long as f' is bounded over $[a, b]$. The following example shows that f' is continuous in $[0, 1]$ except at $x = 0$, f' is bounded over $[0, 1]$, and hence f' satisfies condition (11) and f' is integrable over $[0, 1]$.

Example 8 *We define*

$$f(x) = \begin{cases} x + 2x^2\sin(1/x), & x \neq 0 \\ 0 & 0. \end{cases} \tag{15}$$

Then $\left|\dfrac{f(h)-f(0)}{h} - 1\right| = |2h\sin(1/h)| \leq 2|h|$ *so* $f'(0) = 1$, *and we leave it to reader to verify that*

$$f'(x) = \begin{cases} 1 + 4x\sin(1/x) - 2\cos(1/x), & x \neq 0 \\ 1 & 0. \end{cases}$$

Notice that f' is bounded and continuous in $(0, 1]$ (f' is continuous in $[0, 1]$ except at $x = 0$.) We can prove that f'satisfies condition (11) or Definition 7. [See [6] page 73]. Consequently, f' is integrable over $[0, 1]$ and $\int_0^1 f' = f(1) - f(0) = 1 + 2\sin(1)$.

2.2 Experiment Integrability Numerically

We briefly recall the definition of a definite integral of a function over an interval $[a, b]$ in a regular Calculus textbook is to examine if the Riemann sum is convergent or not, which is hard to examine or compute by hand sometimes but it is manageable if we have a computational tool. Similarly, the theoretical integrability condition for f' in (11) sometimes is neither simple nor direct. We will use some computational techniques here to make this concept clearer. Hongtao Chen et al from the Institute of Computational Mathematics, use Matlab to check validity of condition (11) as follows:

Example 9 *Consider* $f(x) = x^2$ *in* $[0, 1]$*, the slope difference in each sub-interval is*

$$\frac{(x_i + h_i)^2 - x_i^2}{h_i} - 2\delta_i \leq 2x_i + h_i - 2\delta_i \leq h_i \leq h, \tag{16}$$

where $\delta_i \in [x_i, x_{i+1}]$*, and* $h = \max\{h_1, \ldots, h_n\}$ *is called the scale of division. From equation (16) we note that the total error decreases to* 0 *if the scale of division decreases to* 0*. For example, we choose the scale of division* $h = 0.1, 0.09, ...,$ *and etc., and we graph the total error of the scale of division in Figure 5 below, and thus we conjecture that the condition (11) is met.*

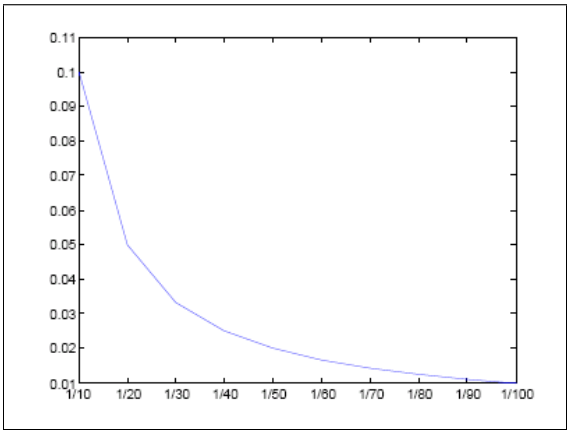


Figure 5.

Example 10 *Consider* $f(x) = 1/x$ *in* $[1, 2]$*, the slope difference in each sub-interval is*

$$\left| \frac{1/(x_i + h_i) - 1/x_i}{h_i} + \frac{1}{\delta_i^2} \right| \leq h_i \leq h, \tag{17}$$

where $\delta_i \in [x_i, x_{i+1}]$*. It is clear that the total error decreases to* 0 *when the scale of the division decreases to* 0*. If we choose the scale of division* $h = 0.1, 0.09, ...$ *and etc., we have the same figure for the total error of the scale of division as shown in Figure 5, and thus condition (11) is met and* f' *is Riemann integrable over the interval* $[1, 2]$*.*

Example 11 *Consider* $f(x) = \sin x$ *in* $[0, 1]$*, the slope difference in each sub-interval is*

$$\left| \frac{\sin(x_i + h_i) - \sin x_i}{h_i} - \cos \delta_i \right| \leq h_i/2 \leq h, \tag{18}$$

where $\delta_i \in [x_i, x_{i+1}]$*, and from the formula (18) we find the total error decreases to* 0 *when the scale of division decreases to* 0*. We choose the scale of division* $h = 0.1, 0.09, ...,$ *and etc. We see from*

Figure 6 below that the scale of division decreases to 0, so does the total error. Thus we conjecture that the condition (11) is met.

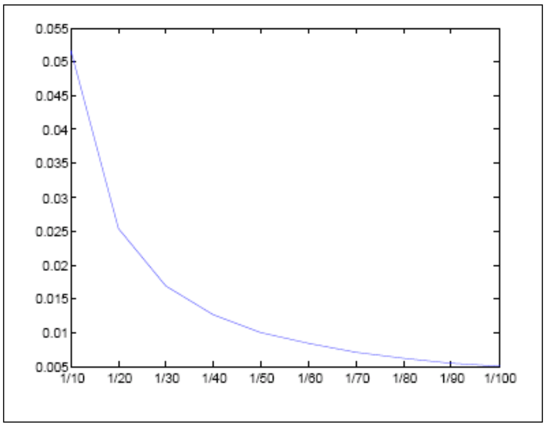


Figure 6.

Example 12 *Consider* $f(x) = \sqrt{x}$ *in* $[1, 2]$*, the slope difference in each sub-interval is*

$$\left| \frac{\sqrt{x_i + h_i} - \sqrt{x_i}}{h_i} - \frac{1}{2\sqrt{\delta_i}} \right| \leq h_i/2 \leq h, \tag{19}$$

where $\delta_i \in [x_i, x_{i+1}]$*, and it follows from (19) that the scale of division decreases to* 0*, so does the total error, see Figure 7 below.*

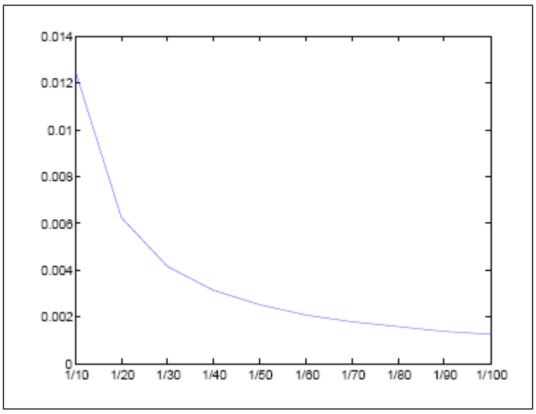


Figure 7.

On the other hand, if we consider $f(x) = \sqrt{x}$ *in* $[0, 1]$ *and divide* $[0, 1]$ *into* n *subintervals. For the first sub-interval* $[0, 1/n]$*, we calculate the slope difference at the point* $x = n^{-1}$*, with* $h = -n^{-1}$*, then consider*

$$\left| \left(\frac{\sqrt{0} - \sqrt{1/n}}{-1/n} - \frac{1}{2\sqrt{n^{-1}}} \right) \right| = \frac{\sqrt{n}}{2},$$

we see that the slope difference at $x = n^{-1}$ *cannot be made as small as we want when* n *gets large. It follows from the Figure 8 that as the scale of division decreases to* 0*, the error stays near* 0.45 *(see*

Figure 8 below),

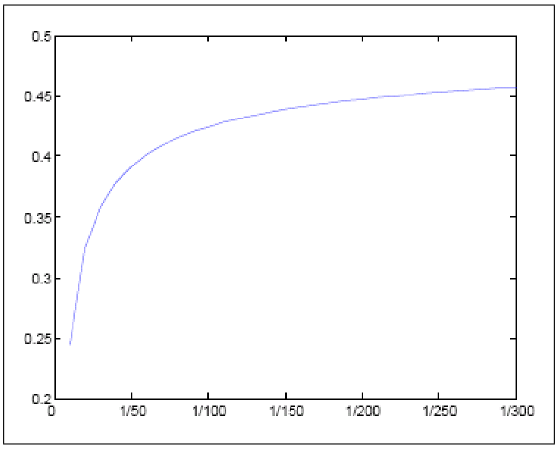


Figure 8.

which suggests that the condition (11) is not met, and thus $\int_0^1 f'(x)dx = \int_0^1 \frac{1}{2\sqrt{x}}dx$ *does not exist in the Riemann sense. That is why we define the improper integral of* $\int_0^1 \frac{1}{2\sqrt{x}}dx$ *to be* $f(1) - f(0)$ *in the regular Calculus text books.*

The next example is intended for those readers who have knowledge about the Lebesgue integration. We show that the slope difference $<< 1$ at every point in the interval does not guarantee $|$the average of slope differences$|$ $<< 1$ over the whole interval.

Example 13 *We define*

$$g(x) = \begin{cases} x^2\cos(\frac{\pi}{x^2}), & x \neq 0 \\ 0 & 0. \end{cases} \tag{20}$$

Then we can verify that

$$g'(x) = \begin{cases} 2x\cos(\pi x^{-2}) + 2\pi x^{-1}\sin(\pi x^{-2}), & x \neq 0 \\ 0 & 0. \end{cases} \tag{21}$$

We note that g' *is not continuous at* $x = 0$ *and* g' *is not bounded on any interval containing* $x = 0$. *We first divide the interval* $[0, 1]$ *into* n *equal sized subintervals. Consider the first subinterval* $[0, 1/n]$, *we choose the tangent at the specific point* $x = 1/n$, *and choose* $h = -1/n$. *Then*

$$\begin{aligned} &\frac{g(x+h) - g(x)}{h} - g'(x) \\ &= \frac{g(0) - g(1/n)}{-1/n} - g'(1/n) \\ &= \frac{(-1/n^2)\cos\pi n^2}{-1/n} - \left[(2/n)\cos\pi n^2 + 2\pi n\sin\pi n^2\right] \\ &= (1/n)\cos\pi n^2 - \left[(2/n)\cos\pi n^2 + 2\pi n\sin\pi n^2\right] \\ &= (-1/n)\cos\pi n^2 - 2\pi n\sin\pi n^2. \end{aligned} \tag{22}$$

It is clear that we need $h\left(\frac{g(x+h)-g(x)}{h}-g'(x)\right)$ *with* $h=-1/n$ *to be made as small as possible before saying that* g' *is Riemann integrable over* $[0,1]$. *However,*

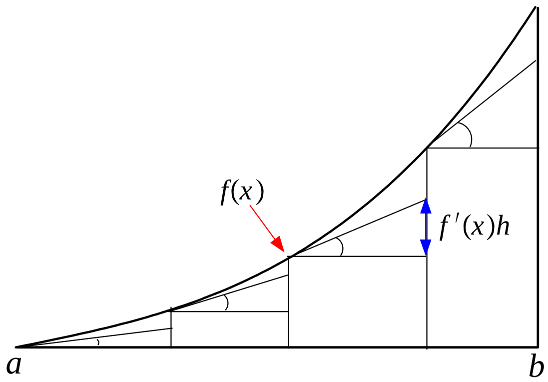
$$h\left(\frac{g(x+h)-g(x)}{h}-g'(x)\right)$$
$$=(1/n^2)\cos\pi n^2+2\pi\sin\pi n^2, \tag{23}$$

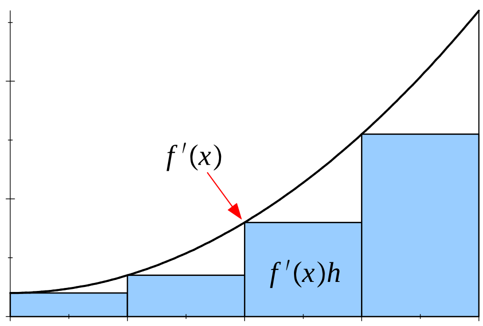
which can not be made as small as we wish. Therefore g' *is not Riemann integrable over* $[0,1]$. *On the other hand, since* g *is differentiable over the interval* $[0,1]$, *the slope difference* $<< 1$ *at every point in* $[0,1]$.

We note that it can be shown that g' is not even Lebesgue integrable over the interval $[0,1]$ (see [12], page 29). However, since g is differentiable over $[0,1]$, it is reasonable to expect from the Fundamental Theorem of Calculus that $\int_0^1 g' = g(1)-g(0)$, the integral $\int_0^1 g'$ should depend on only the value of $g(1)-g(0)$ regardless if g' is integrable. This is indeed the deficiency of the Lebesgue theory and motivation of the modern theory of Henstock and Kurzweil integration theory, see [2, 3, 12, 17].

2.3 The Area Calculation

The formula (12) can be used to calculate the area of f'. This is because the 'sum of differentials' can be interpreted not only as the total height of f between two points on the curve but also as the area of f' over a proper interval, see the illustrations in Figures 9(a) and (b) respectively below.





Figures 9 (a) and (b).

In other words, the formula (12) can be interpreted as saying that the height of f is equal to the area under f'. So, if we want to calculate the area of f' we make use of the primitive function f (or the antiderivative of f). Finding a primitive function or an antiderivative for a function f can be done for many elementary functions if we have table for differentiations ahead of time. For example, we see from Section 1 that

$$(\sin x)' = \cos x$$
$$(x^{n+1})' = (n+1)x^n,$$

and etc., from which we get the corresponding table for integration immediately, e.g.

$$\int \cos x dx = \sin x + C$$

$$\int x^n dx = \frac{x^{n+1}}{n+1} + C \quad (n \neq -1),$$

where C is a constant. In addition, we may apply more integration techniques such as integration by substitution, integration by parts and etc. on more elementary functions to obtain more primitives for more elementary functions. We recall that there are elementary functions, for example for $\sqrt{1+\cos^2 x}$, while calculating the arc length of $\sin x$ on $[0, 2\pi]$), whose primitive is not an elementary function. Moreover, we do not expect to find the primitive for any function all the time. It is suggested in ([16], page 91) that we may fill this gap by using a computational tool through the process of computing the Riemann sum to construct the primitive for a function f.

3 Conclusion

Many articles using the slope difference and the average slope differences when defining the differentiability and integrability for a function have been published in Chinese publications. It is clear in order to avoid using the concept of epsilon and delta, one sometimes needs to work on algebraic equations extensively to find an upper bound for the tail function or upper bound for the quantity of $|(f(b) - f(a)) - (\sum_{i=1}^{n} f'(x_i)h_i)|$. The demands on algebraic manipulation skills seems to be expected and accepted in Chinese school systems; however, this may not be always the case for students in the US or other parts of the world. To bridge this gap, authors use graphical tools to make the concepts of differentiation and integration more accessible. The fundamental theorem has been publicized as a graph in two Chinese newspapers, Guangming Daily (27th Jun., 1997) and Renmin Daily (6th Aug., 1997) (see [4]) and later this graph was incorporated into some mathematics textbooks for college (see, e.g. [5, 8–10, 13, 14, 18]). It was also used as a cover page for the textbook for senior high school (see [1, 7, 11]).

Acknowledgement Author would like to thank Jerry Bona, Alain Leger, Jonathan Lewin, Michael Livshits, Michal Krizek, and Liyeng Sung for communicating with authors and make many invaluable suggestions to this article.

References

[1] K.-S Cai (2008), *Practical Calculus*, Suzhou University Press, Suzhou China, ISBN 978-7-81137-059-1.

[2] P. Y. Lee (1989), *Lanzhou Lectures on Henstock Integration*, World Scientific, Singapore.

[3] P. Y. Lee & R. Výborný (2000), *Integral: An Easy Approach after Kurzweil and Henstock*, Cambridge University Press, Cambridge UK, ISBN 0 521 77968 5.

[4] Q. Lin (1997), *Calculus: a graph.* Guangming Daily, June 27 & People's Daily, August 6.

[5] Q. Lin (1999), *Calculus Cartoon*, Guangxi Normal University Press, Nanning China, ISBN 7563327851.

[6] Q. Lin (2009), *Learning School Calculus in China*, Mathematics Bulletin, ISSN 0583-1458, Special issue, 44-52.

[7] Q. Lin & S.-X. Liu (2005, 2007, 2009), *Mathematics textbooks for high school*, People's Education Press, Beijing China, ISBN 978-7-107-19170-1. (http://www.pep.com.cn/czsx/jszx/jnjsc/9shangnew/).

[8] Y. Liu & X.-L Chen (2010), *Mathematic Teaching Practice*, Peking Unviersity Press, Beijing China, ISBN 9787301143858.

[9] Q. Lin & C.-X Wu (2002), *Liberal Arts Mathematics*, Hebe University Press, Baoding China, ISBN 9787810289061.

[10] M. Livshits (2009), *You Could Simplify Calculus*, http://www.mathfoolery.com/Article/simpcalc-v1.pdf.

[11] S.-H. Meng, W.-H. Zou, Z. Gan & C.-L. Zou (2005), *High School Mathematic Textbook 2-2 for Science*, Hunan Education Press, Changsha China, ISBN 987-7-5355-4613-5.

[12] W.F. Pfeffer (1993), *The Riemann Approach to Integration: Local Geometric Theory*, Cambridge University Press, ISBN 0-521-44035-1.

[13] M. Ryan (2003), *Calculus for Dummies*, Wiley Publishing Inc, Indianapolis, ISBN 0-7654-2498-4.

[14] L.-X Tian (2008) *Advanced mathematics*, Jiangsu Press, Nanjing China, ISBN 9787811300000.

[15] J.-Q. Yao (2010), Book Reviews for *Quick Meal of Calculus*, Scientific Times.

[16] W.-C. Yang & J. Lewin (1998), *Exploring Mathemaics with Scientific Notebook*, page 91, Springer Verlag, ISBN 981-3083-88-3.

[17] W.-C. Yang, P.-Y. Lee & X. Ding (2009), *Numerical Integration On Some Special Henstock-Kurzweil Integrals*, the Electronic Journal of Mathematics and Technology (eJMT), Issue 3, Vol. 3, ISSN 1933-2823.

[18] C.-H Yu (2003), *Mathematical Analysis*, Higher Education Press, Beijing China, ISBN 7-04-013852-2.

[19] J.-Z. Zhang & Y. Feng (2009), *New perspective on calculus*, Science China Series A, Issue 2, Vol.39, ISSN 247-256.