REAL ANALYSIS
TOPIC 31 - CALCULUS REVIEW

PAUL L. BAILEY

ABSTRACT. We review the three main definitions from Calculus: limits, derivatives, and integrals, as presented in Thomas' Calculus. We will focus on real-valued functions of a real variable, but encourage the reader to consider potential generalizations to $\mathbb{R}^n$, metric spaces, or topological spaces.

1. Topology of $\mathbb{R}$

Let us restate the basic topological concepts we need for Calculus.

Let $U \subset \mathbb{R}$. We say that U is *open* if, for every $x \in U$, there exits $\epsilon > 0$ such that $(x - \epsilon, x + \epsilon) \subset U$.

Let $F \subset \mathbb{R}$. We say that F is *closed* if its complement, $F^c = \mathbb{R} \smallsetminus F$, is open.

Let $p \in \mathbb{R}$.

A *neighborhood* of p is a subset of $\mathbb{R}$ which contains an open set which contains p. That is, a set $N \subset \mathbb{R}$ is a neighborhood of p if there exists $\epsilon > 0$ such that $(p - \epsilon, p + \epsilon) \subset N$.

A *deleted neighborhood* of p is a set of the form $N \smallsetminus \{p\}$, where N is a neighborhood of p.

Let $A \subset \mathbb{R}$.

We say that p is a *closure point* of A if every neighborhood of p intersects A.

We say that p is an *interior point* of A if there exists a neighborhood of p which is contained in A.

We say that p is a *boundary point* of A if every neighborhood of p intersects A and A^c.

We say that p is an *accumulation point* of A if every deleted neighborhood of p intersects A.

We say that p is an *isolated point* of A if $p \in A$ and there exists a deleted neighborhood of p which is disjoint from A. Thus if $a \in A$, then a is either an accumulation point of A, or a is an isolated point of A, but not both.

We will use the definitions and results regarding compactness and connectedness which we have previously explored, in the context of the real numbers. In particular, we will use the following results:

- **(a)** A compact subset of $\mathbb{R}$ is closed.
- **(b)** A connected subset of $\mathbb{R}$ is an interval.
- **(c)** A subset of $\mathbb{R}$ is compact if and only if it is closed and bounded.
- **(d)** A compact, connected subset of $\mathbb{R}$ is a bounded closed interval.

Date: February 26, 2019.

2. Continuity

A *real-valued function of a real variable* is a function of the form $f : D \to \mathbb{R}$, where $D \subset \mathbb{R}$. We allow D to be any subset; it may have accumulation points which are not in D, and it may have isolated points. We will become particularly interested in situations where D is an unusual set. We review the basic facts of continuity we have already studied, in this context.

Definition 1. Let $D \subset \mathbb{R}$ and let $f : D \to \mathbb{R}$. We say that f is *continuous* at $x_0 \in D$ if

$$\forall \epsilon > 0 \, \exists \delta > 0 \, \forall x \in D \, : \, |x - x_0| < \delta \, \Rightarrow \, |f(x) - f(x_0)| < \epsilon.$$

We say that f is *continuous* if f is continuous at every point in D.

Proposition 1. *Let $f : D \to \mathbb{R}$ be any function. If $x_0 \in D$ is an isolated point of D, then f is continuous at x_0.*

Proof. Exercise. □

We have seen the following consequences of the definition of continuity:

(a) The continuous image of a connected set is connected.
(b) The continuous image of a compact set is compact.

Theorem 1 (Intermediate Value Theorem (IVT))**.** *Let $f : [a, b] \to \mathbb{R}$ be a continuous function. If $f(a)f(b) < 0$, then there exists $c \in (a, b)$ such that $f(c) = 0$.*

Proof. We have seen that the connected, compact subsets of $\mathbb{R}$ are closed intervals of finite length. We have also seen that the continuous image of a compact set is compact, and that the continuous image of a connected set is connected. Thus, since the domain $[a, b]$ is compact and connected, so is its image; thus, $f([a, b]) = [y, z]$ for some $y, z \in \mathbb{R}$. Since $yz < 0$, either $y < 0 < z$, so $0 \in [y, z]$. □

Theorem 2 (Extreme Value Theorem (EVT))**.** *Let $f : [a, b] \to \mathbb{R}$ be a continuous function. Then f admits a minimum and maximum value on $[a, b]$; that is, there exist $c, d \in \mathbb{R}$ such that, for every $x \in [a, b]$,*

$$f(c) \leq f(x) \leq f(d).$$

Proof. The image of $[a, b]$ is a closed interval, so $f([a, b]) = [y, z]$ for some $y, z \in \mathbb{R}$. Thus y and z are in the image of f, so there exist $c, d \in [a, b]$ such that $f(c) = y$ and $f(d) = z$. The result follows. □

3. Limits

3.1. **Definition of Limit.** Even though we have previous thoroughly explored continuity, we nevertheless now require the notion of the limit of a function. This is because the notion of derivative of f at a point $a \in \text{dom}(f)$ uses the difference quotient of f, which we may denote by $\widehat{f}$; we have

$$\widehat{f}(x) = \frac{f(x) - f(a)}{x - a}.$$

But this function is not (yet) defined at $x = a$, and we wish to understand its behavior close to a; the idea of limit precedes continuity in this case.

Definition 2. Let $D \subset \mathbb{R}$ and let $x_0 \in D$ be an accumulation point of D. Let $f : D \to \mathbb{R}$. We say that the *limit* of f at x_0 is L if

$$\forall \epsilon > 0 \, \exists \delta > 0 \, \forall x \in D \, : \, 0 < |x - x_0| < \delta \, \Rightarrow \, |f(x) - f(x_0)| < \epsilon.$$

We write $\lim\limits_{x \to x_0} f(x) = L$ to mean that

- x_0 is an accumulation point of $\text{dom}(f)$, and
- the limit of f at x_0 is L.

We observe that f is continuous at x_0 if and only if $\lim\limits_{x \to x_0} f(x) = L$.

Question 1. We have already discussed and kind of understand continuity; why isn't it enough? We do we need this new definition?

3.2. **Arithmetic of Limits.** Next, we prove the standard arithmetic limit laws, summarized as follows:

- The limit of a constant is that constant.
- The limit of a sum is the sum of the limits.
- The limit of a product is the product of the limits.
- The limit of a quotient is the quotient of the limits.
- The limit operator commutes with continuous functions.

Proposition 2. *Let $D \subset \mathbb{R}$ and let $x_0 \in D$ be an accumulation point of D. Let $C \in \mathbb{R}$ be constant. Define $f : D \to \mathbb{R}$ by $f(x) = C$ for all $x \in D$. Then*

$$\lim_{x \to x_0} f(x) = C.$$

Proof. Let $\epsilon > 0$ and let $\delta > 0$. Then if $x \in D$ and $0 < |x - x_0| < \delta$, we have $|f(x) - C| = |C - C| = 0 < \epsilon$. $\square$

Proposition 3. *Let $D \subset \mathbb{R}$ and let $x_0 \in D$ be an accumulation point of D. Let $f : D \to \mathbb{R}$ and $g : D \to \mathbb{R}$. Suppose that $\lim\limits_{x \to x_0} f(x) = L$ and $\lim\limits_{x \to x_0} g(x) = M$. Then*

$$\lim_{x \to x_0} (f + g)(x) = L + M.$$

Proof. Let $\epsilon > 0$.

Let $\delta_1 > 0$ be so small that $0 < |x - x_0| < \delta$ implies $|f(x) - L| < \dfrac{\epsilon}{2}$.

Let $\delta_2 > 0$ be so small that $0 < |x - x_0| < \delta$ implies $|g(x) - M| < \dfrac{\epsilon}{2}$.

Let $\delta = \min\{\delta_1, \delta_2\}$. Then, for $x \in D$ with $0 < |x - x_0| < \delta$, we have

$$|(f(x) + g(x)) - (L + M)| \leq |f(x) - L| + |g(x) - M| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

$\square$

Proposition 4. *Let $D \subset \mathbb{R}$ and let $x_0 \in D$ be an accumulation point of D. Let $f : D \to \mathbb{R}$ and $g : D \to \mathbb{R}$. Suppose that $\lim_{x \to x_0} f(x) = L$ and $\lim_{x \to x_0} g(x) = M$. Then*

$$\lim_{x \to x_0} (fg)(x) = LM.$$

Proof. Let $\epsilon > 0$.

Let $\delta_0 > 0$ be so small that $0 < |x - x_0| < \delta_0$ implies $|g(x) - M| < 1$. Suppose that $0 < |x - x_0| < \delta_0$. Then $-1 + M < g(x) < 1 + M$, and also, $-1 - M < -g(x) < 1 - M$. Now let $B = \max\{|1 - M|, |1 + M|\}$; therefore $|f(x)| < B$.

Let $\delta_2 > 0$ be so small that $0 < |x - x_0| < \delta_2$ implies $|g(x) - M| < \dfrac{\epsilon}{2(|L| + 1)}$. We require the plus one in case $L = 0$.

Let $\delta = \min\{\delta_0, \delta_1, \delta_2\}$. Then, for $x \in D$ with $0 < |x - x_0| < \delta$, we have

$$\begin{aligned}
|f(x)g(x) - LM| &\leq |f(x)g(x) - Lg(x)| + |Lg(x) - LM| \\
&\leq |f(x) - L||g(x)| + |L||g(x) - M| \\
&< \frac{\epsilon}{2B} B + |L| \frac{\epsilon}{2(|L| + 1)} \\
&< \epsilon.
\end{aligned}$$

□

Proposition 5. *Let $D \subset \mathbb{R}$ and let $x_0 \in D$ be an accumulation point of D. Let $f : D \to \mathbb{R}$ and $g : D \to \mathbb{R}$. Suppose that $\lim_{x \to x_0} f(x) = L$ and $\lim_{x \to x_0} g(x) = M$, and that $M \neq 0$. Then*

$$\lim_{x \to x_0} \left(\frac{f}{g}\right)(x) = \frac{L}{M}.$$

Proof. It suffices to show that $\lim_{x \to x_0} \left(\dfrac{1}{g}\right)(x) = \dfrac{1}{M}$, as the stated result simply combines this with Proposition 4.

Let $\epsilon > 0$. Let $B = \dfrac{|M|}{2}$.

Let δ_0 be so small that $0 < |x - x_0| < \delta_0$ implies that $|g(x) - M| < \dfrac{|M|}{2}$. If $M > 0$, we have $\dfrac{M}{2} < g(x) - M$, so $0 < \dfrac{|M|}{2} = \dfrac{M}{2} < g(x) = |g(x)|$. If $M < 0$, we have $g(x) - M < -\dfrac{M}{2}$, so $g(x) < \dfrac{M}{2}$, so $|g(x)| = -g(x) > -\dfrac{M}{2} = \dfrac{|M|}{2}$. In either case, $|g(x)| > \dfrac{|M|}{2}$.

Let δ_1 be so small that $0 < |x - x_0| < \delta_1$ implies that $|g(x) - M| < \dfrac{\epsilon |M|^2}{2}$

Let $\delta = \min\{\delta_0, \delta_1\}$.

Then, for $x \in D$ with $0 < |x - x_0| < \delta$, we have

$$\begin{aligned}
\left|\frac{1}{g(x)} - \frac{1}{M}\right| &= \left|\frac{M - g(x)}{g(x)M}\right| \\
&= |g(x) - M| \cdot \frac{1}{|g(x)||M|} \\
&< \frac{\epsilon |M|^2}{2} \cdot \frac{2}{|M|^2} \\
&= \epsilon.
\end{aligned}$$

□

Observation 1. Let $D \subset \mathbb{R}$ and let $x_0 \in D$ be an accumulation point of D. Let $f : D \to \mathbb{R}$ and $g : D \to \mathbb{R}$. Let $c \in \mathbb{R}$. Suppose that $\lim_{x \to x_0} f(x) = L$ and $\lim_{x \to x_0} g(x) = M$. Then

(a) $\lim_{x \to x_0} (cf)(x) = cf$;
(b) $\lim_{x \to x_0} (-f)(x) = -f$;
(c) $\lim_{x \to x_0} (1/f)(x) = 1/L$, when $L \neq 0$.

Proof. We use the fact that the limit of a constant function is that constant.
Part (a) follows from Proposition 4 by changing f to c and g to f.
Part (b) follows from Part (a) by changing c to -1
Part (c) follows from Proposition 5 by changing f to 1 and g to f. □

3.3. **Theory of Limits.** We give two additional important theorems with respect to limits.

Proposition 6. *Let $f : D \to \mathbb{R}$ and let $g : E \to \mathbb{R}$. Let x_0 be an accumulation point of D and y_0 be an accumulation point of $E \cap f(D)$. Suppose that $\lim_{x \to x_0} f(x) = y_0$ and $\lim_{y \to y_0} g(y) = z_0$. Additionally, suppose that either*

(a) *$g(y_0) = z_0$, so that g is continuous at y_0, or*
(b) *there exists $\beta > 0$ such that $0 < |x - x_0| < \beta$ implies $f(x) \neq y_0$.*

Then $\lim_{x \to x_0} (g \circ f)(x) = z_0$.

Proof. Let $\epsilon > 0$.
Let $\delta > 0$ be so small that $0 < |y - y_0| < \delta$ implies $|g(y) - z_0| < \epsilon$.
Let $\gamma > 0$ be so small that $0 < |x - x_0| < \gamma$ implies $|f(x) - y_0| < \delta$.
Suppose **(a)** holds, and let $x \in D$ with $0 < |x - x_0| < \gamma$. Then $|f(x) - y_0| < \delta$. If $f(x) = y_0$, then $g(f(x)) = g(y_0) = z_0$, so $|g(f(x)) - z_0| = 0 < \epsilon$. If $f(x) \neq y_0$, then $0 < |f(x) - y_0| < \delta$, so $|g(f(x)) - z_0| < \epsilon$.
Otherwise, **(b)** holds. Let $\alpha = \min\{\beta, \gamma\}$. Then for $x \in D$, if $0 < |x - x_0| < \alpha$, then $f(x) - y_0 \neq 0$, so $0 < |f(x) - y_0| < \delta$, which implies that $|g(f(x)) - z_0| < \epsilon$. □

The next proposition relates our previous work with the definition currently under consideration. Is is in fact possible to prove some of the theorems regarding the arithmetic of series using the next definition, but we thought it informative to proceed as we did.

Proposition 7. *Let $D \subset \mathbb{R}$ and let $x_0 \in D$ be an accumulation point of D. Let $f : D \to \mathbb{R}$. Then $\lim_{x \to x_0} f(x) = L$ if and only if for every sequence (x_n) in D which converges to L, the sequence $(f(x_n))$ converges to $f(L)$.*

Proof. We prove both directions of the implication.

($\Rightarrow$) Suppose that $\lim_{x \to x_0} f(x) = L$, and let (x_n) be a sequence in D which converges to x_0. We wish to show that $(f(x_n))$ converges to L.

Let $\epsilon > 0$. Let $\delta > 0$ be so small that for $x \in D$ such that $0 < |x - x_0| < \delta$, we have $|f(x) - L| < \epsilon$. Let $N \in \mathbb{N}$ be so large that $n \geq N$ implies $|x_n - x_0| < \delta$. Then, for $n \geq N$, we have $0 < |x_n - x_0|$, so $|f(x_n) - L| < \epsilon$. Thus $(f(x_n))$ converges to L.

($\Leftarrow$) Here we prove the contrapositive. That is, assume that the limit of f at x_0 is not L, and construct a sequence (x_n) which converges to x_0 such that $(f(x_n))$ avoids L.

Since the limit of f at x_0 is not L, there exists $\epsilon > 0$ such that for every $n \in \mathbb{N}$ there exists $x_n \in (x_0 - \frac{1}{n}, x_0 + \frac{1}{n})$ we that $|(|f(x_n) - L) > \epsilon$. As is clear to all those who observe the same, the sequence (x_n) converges to x_0, but $(f(x_n))$ does not converge to L. □

4. Differentiability

4.1. **Definition of Derivative.** Let $D \subset \mathbb{R}$ and let $f : D \to \mathbb{R}$. Let a be an accumulation point of D; it is possible to define derivatives in this case.

Definition 3. Let $D \subset \mathbb{R}$ and let $f : D \to \mathbb{R}$. Let $a \in D$ be an accumulation point of D. We say that f is *differentiable at a* if

$$\lim_{x \to a} \frac{f(x) - f(a)}{x - a}$$

exists. In this case, set

$$f'(a) = \lim_{x \to a} \frac{f(x) - f(a)}{x - a};$$

we call $f'(a)$ the *derivative* of f at a.

We wish to vary the point at which we are taking the derivative, and for notational clarity, we to call it x. We rephrase the definition by replacing $x - a$ with h, and then replacing x_0 with x. It is clear, then, that

$$f'(x) = \lim_{h \to 0} \frac{f(x + h) - f(x)}{h}.$$

Let $D' \subset D$ be the set of all points in D at which f is differentiable. We obtain a function

$$f : D' \to \mathbb{R} \quad \text{given by } f'(x) = \lim_{h \to 0} \frac{f(x + h) - f(x)}{h}.$$

Proposition 8. *Let $D \subset \mathbb{R}$ and let $a \in D$ be an accumulation point of D. Let $f : D \to \mathbb{R}$. If f is differentiable at a, then f is continuous at a.*

Proof. We use the fact that f is continuous at a if and only if $\underset{x \to a}{f}(x) = f(a)$.

Suppose that f is differentiable at a. Note that

$$f(a+h) = f(a) + f(a+h) - f(a) = f(a) + \frac{f(a+h) - f(a)}{h} \cdot h.$$

Take the limit as $h \to 0$ of both sides; the limit laws we have previously derived show that

$$\lim_{h \to 0} f(a+h) = \lim_{h \to 0} f(a) + \lim_{h \to 0} \frac{f(a+h) - f(a)}{h} \cdot \lim_{h \to 0} h = f(a) + f'(a) \cdot 0 = f(a).$$

Thus, f is continuous at a. □

4.2. **Arithmetic of Derivatives.** We now prove the appropriate rules for taking derivatives of sums, products, and quotients of differentiable functions.

Proposition 9 (Sum Rule)**.** *Let $D \subset \mathbb{R}$ and let $x \in D$ be an accumulation point of D. Let $f : D \to \mathbb{R}$ and $g : D \to \mathbb{R}$ be differentiable at x. Then $f+g$ is differentiable at x, and $(f+g)'(x) = f'(x) + g'(x)$.*

Proof. Since the necessary limits exist, we have

$$\begin{aligned}
(f+g)'(x) &= \lim_{h \to 0} \frac{(f+g)(x+h) - (f+g)(x)}{h} \\
&= \lim_{h \to 0} \frac{f(x+h) - f(a) + g(x+h) - g(x)}{h} \\
&= \lim_{h \to 0} \frac{f(x+h) - f(x)}{h} + \lim_{h \to 0} \frac{g(x+h) - g(x)}{h} \\
&= f'(x) + g'(x).
\end{aligned}$$

□

Proposition 10 (Product Rule)**.** *Let $D \subset \mathbb{R}$ and let $x \in D$ be an accumulation point of D. Let $f : D \to \mathbb{R}$ and $g : D \to \mathbb{R}$ be differentiable at x. Then fg is differentiable at x, and $(fg)'(x) = f'(x)g(x) + f(x)g'(x)$.*

Proof. Compute

$$\begin{aligned}
(fg)'(x) &= \lim_{h \to 0} \frac{(fg)(x+h) - (fg)(x)}{h} \\
&= \lim_{h \to 0} \frac{f(x+h)g(x+h) - f(x)g(x+h) + f(x)g(x+h) - f(x)g(x)}{h} \\
&= \lim_{h \to 0} \frac{f(x+h) - f(x)}{h} \cdot g(x) + \lim_{h \to 0} f(x) \cdot \frac{g(x+h) - g(x)}{h} \\
&= f'(x)g(x) + f(x)g'(x).
\end{aligned}$$

□

Proposition 11 (Quotient Rule)**.** *Let $D \subset \mathbb{R}$ and let $x \in D$ be an accumulation point of D. Let $f : D \to \mathbb{R}$ and $g : D \to \mathbb{R}$ be differentiable at x. If $g(x) \neq 0$, then $\frac{f}{g}$ is differentiable at x, and $\left(\frac{f}{g}\right)'(x) = \frac{f'(x)g(x) - f(x)g'(x)}{(g(x))^2}$.*

Proof. Compute

$$\begin{aligned}
\left(\frac{f}{g}\right)'(x) &= \lim_{h\to 0} \frac{\frac{f}{g}(x+h) - \frac{f}{g}(x)}{h} \\
&= \lim_{h\to 0} \frac{f(x+h)g(x) - f(x)g(x+h)}{hg(x)g(x+h)} \\
&= \lim_{h\to 0} \frac{f(x+h)g(x) - f(x)g(x) + f(x)g(x) - f(x)g(x+h)}{hg(x)g(x+h)} \\
&= \lim_{h\to 0} \frac{(f(x+h) - f(x))g(x) - f(x)(g(x+h) - g(x))}{hg(x)g(x+h)} \\
&= \left(\lim_{h\to 0} \frac{f(x+h) - f(x)}{h} \cdot g(x) - \lim_{h\to 0} f(x) \cdot \frac{g(x+h) - g(x)}{h}\right) \cdot \lim_{h\to 0} \frac{1}{g(x)g(x+h)} \\
&= \frac{f'(x)g(x) - f(x)g'(x)}{(g(x))^2}.
\end{aligned}$$

□

4.3. **Leibnitz Notation.** Let $f : D \to \mathbb{R}$ with x_0 an accumulation point of D.

Let $\Delta x = x - x_0$; viewing x_0 as fixed, this is implicitly a function of x. Let $\Delta f = f(x) - f(x_0)$; viewing f as fixed, this is also a function of x.

Now x goes to x_0, we see that Δx goes to 0. Thus

$$\lim_{x\to 0} \frac{\Delta f}{\Delta x} = \lim_{x\to x_0} \frac{f(x) - f(x_0)}{x - x_0}.$$

Thus we may define the derivative to be

$$\frac{df}{dx} = \lim_{x\to 0} \frac{\Delta f}{\Delta x}.$$

Moreover, in Leibnitz notation, it is traditional to start with a function whose name is y instead of f, so this becomes

$$\frac{dy}{dx} = \lim_{x\to 0} \frac{\Delta y}{\Delta x}.$$

4.4. **Chain Rule.** If we have two lines $y = m_1 z + b_1$ and $z = m_2 x + b_2$ and compose them, we obtain

$$y = m_1 m_2 x + (m_1 b_2 + b_1),$$

a line with slope $m_1 m_2$. Since we view a differentiable function as a function which is approximately a line whose slope is the derivative, we guess that the derivative of a composition is the product of the derivatives.

Suppose that y is a function of u and u is a function of x. Then we may attempt to right

$$\frac{\Delta y}{\Delta x} = \frac{\Delta y}{\Delta u}\frac{\Delta u}{\Delta x}.$$

Then $\Delta u \to 0$ as $\Delta x \to 0$, so taking the limit of both sides we would arrive at

$$\frac{dy}{dx} = \frac{dy}{du}\frac{du}{dx}.$$

The problem with this reasoning is that Δu may be zero even when Δx is nonzero. We have to get around this problem.

To adopt some standard notation, we let $\widehat{f}$ be the difference quotient which we are to take the limit of. That is, we fix $a \in D$ and set

$$\widehat{f}(x) = \begin{cases} \dfrac{f(x) - f(a)}{x - a} & \text{if } x \neq a \ ; \\ f'(a) & \text{if } f'(a) \text{ exists} \ ; \\ \text{undefined} & \text{otherwise} \ . \end{cases}$$

In what follows, this plays the role of $\dfrac{\Delta y}{\Delta x}$.

Proposition 12 (Chain Rule)**.** *Let $X, Y \subset \mathbb{R}$ with $x_0 \in D$ an accumulation point of X and $y_0 \in Y$ and accumulation point of Y. Let $f : X \to \mathbb{R}$ and $g : Y \to \mathbb{R}$ with $f(X) \subset Y$ and $f(x_0) = y_0$. If f is differentiable at x_0 and g is differentiable at y_0, then $g \circ f$ is differentiable at x_0 and*

$$(g \circ f)'(x_0) = g'(y_0) f'(x_0).$$

Proof. Let $h : X \to \mathbb{R}$ be given by $h = g \circ f$.

Define a function $\widehat{h} : X \to \mathbb{R}$ by $\widehat{h}(x) = \frac{g(f(x)) - g(f(x_0))}{x - x_0}$; we wish to show that $\widehat{h}(x)$ has a limit at $x = x_0$, and that $\lim_{x \to x_0} \widehat{h}(x) = g'(y_0) f'(x_0)$.

Define $\widehat{g} : Y \to \mathbb{R}$ by

$$\widehat{g}(y) = \begin{cases} \frac{g(y) - g(y_0)}{y - y_0} & \text{if } y \neq y_0; \\ g'(y_0) & \text{if } y = y_0. \end{cases}$$

Since g is differentiable at y_0, we have $\lim_{y \to y_0} \widehat{g}(y) = g'(y_0) = \widehat{g}(y_0)$, so $\widehat{g}$ is continuous at y_0. Since f is differentiable at x_0, it is continuous at x_0, and since $f(x_0) = y_0$, then $\widehat{g} \circ f$ is continuous at x_0.

Set $\widehat{f}(x) = \frac{f(x) - f(x_0)}{x - x_0}$. We claim that for $x \in D \setminus \{x_0\}$, we have $\widehat{h}(x) = \widehat{g}(f(x)) \cdot \widehat{f}(x)$. If $f(x) = f(x_0)$, then $g(f(x)) = g(f(x_0)) = g(y_0)$. In this case, $\widehat{h}(x) = 0$ and $\widehat{g}(f(x)) \cdot \widehat{f}(x) = 0$. Otherwise, $\widehat{h}(x) = \frac{g(f(x)) - g(f(x_0))}{f(x) - f(x_0)} \frac{f(x) - f(x_0)}{x - x_0} = \widehat{g}(f(x)) \cdot \widehat{f}(x)$.

Now take the limit to see that

$$\lim_{x \to x_0} \widehat{h}(x) = \lim_{x \to x_0} \widehat{g}(f(x)) \lim_{x \to x_0} \widehat{f}(x) = g'(y_0) f'(x_0).$$

$\square$

5. Consequences of Differentiability

5.1. **Extreme Value Theorem.** We have previously stated the Extreme Value Theorem on a closed interval; we revisit this here, in order to distinguish global and local extreme values.

Definition 4. Let $f : D \to \mathbb{R}$, where $D \subset \mathbb{R}$, and let $x_0 \in D$.

We say that f has a *global minimum* at x_0 if $f(x) \geq f(x_0)$ for all $x \in [a, b]$.

We say that f has a *global maximum* at x_0 if $f(x) \leq f(x_0)$ for all $x \in [a, b]$.

We say that f has a *global extremum* at x_0 if f has a global minimum or global maximum at x_0.

Proposition 13. *Let $f : [a, b] \to \mathbb{R}$ be continuous. Then f has a global minimum and a global maximum on $[a, b]$.*

Proof. This is Theorem 2. □

Definition 5. Let $f : D \to \mathbb{R}$, where $D \subset \mathbb{R}$, and let $x_0 \in D$ be an accumulation point of D.

We say that f has a *local minimum* at x_0 if there exists $\delta > 0$ such that $f(x) \geq f(x_0)$ for all $x \in (x_0 - \delta, x_0 + \delta) \cap D$.

We say that f has a *local maximum* at x_0 if there exists $\delta > 0$ such that $f(x) \leq f(x_0)$ for all $x \in (x_0 - \delta, x_0 + \delta) \cap D$.

We say that f has a *local extremum* at x_0 if f has a local minimum or local maximum at x_0.

Proposition 14. *Let $f : [a, b] \to \mathbb{R}$ and let $x_0 \in (a, b)$ be a local extremum of f. If f is differentiable at x_0, then $f'(x_0) = 0$.*

Proof. We may assume that f has a local maximum at x_0, the proof of a local minimum being analogous. Then there exists $\delta > 0$ such that $f(x) \leq f(x_0)$ for all x satisfying $|x - x_0| < \delta$.

Set $\widehat{f}(x) = \frac{f(x) - f(x_0)}{x - x_0}$ for $x \in D \setminus \{x_0\}$. Since f is differentiable at x_0, $\lim_{n\to\infty} \widehat{f}(x_n) = f'(x_0)$ for every sequence from $(x - \delta, x + \delta)$ which converges to x_0.

Note that the numerator of $\widehat{f}(x)$ is negative for x near x_0, but that the denominator is positive on one side of x_0 and negative on the other. For $x_n = x - \frac{\delta}{n}$, we see that $\widehat{f}(x_n) \geq 0$, so $f'(x_0) \geq 0$. However, for $x_n = x + \frac{\delta}{n}$, we have $\widehat{f}(x_n) \leq 0$, so $f'(x_0) \leq 0$. This shows that $f'(x_0) = 0$. □

5.2. **Mean Value Theorem.**

Lemma 1 (Rolle's Theorem)**.** *Let $f : [a, b] \to \mathbb{R}$ be continuous on $[a, b]$ and differentiable on (a, b). Then if $f(a) = f(b)$, there exists $c \in (a, b)$ such that $f'(c) = 0$.*

Proof. Since $[a, b]$ is compact, there exists $x_1, x_2 \in [a, b]$ such that $f(x_1)$ is a global minimum and $f(x_2)$ is a global maximum. If $f(x_1) = f(x_2)$, then f is constant, and $f'(x) = 0$ for every $x \in [a, b]$. Otherwise, either $x_1 \neq f(a)$ or $x_2 \neq f(a)$. Therefore either $x_1 \in (a, b)$ or $x_2 \in (a, b)$. If $x_1 \in (a, b)$, then x_1 is a local minimum, and $f'(x_1) = 0$. If $x_2 \in (a, b)$, then x_2 is a local maximum, and $f'(x_2) = 0$. □

Theorem 3 (Mean Value Theorem)**.** *Let $f : [a, b] \to \mathbb{R}$ be continuous on $[a, b]$ and differentiable on (a, b). Then there exists $c \in (a, b)$ such that $f'(c) = \frac{f(b)-f(b)}{b-a}$.*

Proof. Define $g : [a, b] \to \mathbb{R}$ by $g(x) = f(x) - \frac{x-a}{b-a}(f(b) - f(a))$. There g is continuous on $[a, b]$ and differentiable on (a, b), and we compute that

$$g'(x) = f'(x) - \frac{f(b) - f(a)}{b - a}.$$

However, $g(a) = f(a)$ and $g(b) = f(a)$. By Rolle's Theorem, there exists $c \in [a, b]$ such that $g'(c) = 0$, so $f'(c) = \frac{f(b)-f(a)}{b-a}$. □

Corollary 1. *Let $f : [a, b] \to \mathbb{R}$ be continuous on $[a, b]$ and differentiable on (a, b). Suppose that $f'(x) = 0$ for all $x \in (a, b)$. Then $f(x) = C$ for all $x \in [a, b]$ and some fixed constant $C \in \mathbb{R}$.*

Proof. Let $C = f(a)$, and let $x \in (a, b]$; it suffices to show that $f(x) = f(a)$. We note that f is continuous on $[a, x]$, and that f is differentiable on (a, x). Hence, by Theorem 3, there exists $c \in (a, x)$ such that

$$f'(c) = \frac{f(x) - f(a)}{x - a};$$

but $f'(c) = 0$, so $f(x) - f(a) = 0$, whence $f(x) = f(a) = C$. □

Corollary 2. *Let $f, g : [a, b] \to \mathbb{R}$ be continuous on $[a, b]$ and differentiable on (a, b). Suppose that $g'(x) = f'(x)$ for all $x \in (a, b)$. Then $g(x) = f(x) + C$, for all $x \in [a, b]$ and some fixed constant $C \in \mathbb{R}$.*

Proof. Let $h : [a, b] \in \mathbb{R}$ be given by $h(x) = g(x) - f(x)$. Then h is continuous on $[a, b]$ and differentiable on (a, b).

Moreover, $h'(x) = g'(x) - f'(x) = 0$, for all $x \in (a, b)$. Thus, by Corollary 1, we have $h(x) = C$, for all $x \in [a, b]$ and some fixed constant $C \in \mathbb{R}$. Thus $g(x) - f(x) = C$, so $g(x) = f(x) + C$, for all $x \in [a, b]$. □

5.3. **Inverse Function Theorem.**

Proposition 15. *Let $f : [a,b] \to \mathbb{R}$ be continuous on $[a,b]$ and differentiable on (a,b). If $f'(x) \neq 0$ for $x \in [a,b]$, then f is injective, and the inverse of f is differentiable on $f((a,b))$, with*

$$(f^{-1})'(f(x)) = \frac{1}{f'(x)}$$

for every $x \in (a,b)$.

Proof. Suppose that f is not injective. Then there exist $x_1, x_2 \in [a,b]$ such that $f(x_1) = f(x_2)$. By Rolle's Theorem, there exists $c \in [x_1, x_2]$ such that $f'(c) = 0$; this violates the hypothesis. Thus f is injective.

We have seen that a continuous bijective function on a compact set has a continuous inverse; since $[a,b]$ is compact, $f^{-1} : f([a,b]) \to [a,b]$ is continuous.

Now let $y_0 \in f((a,b))$, and let (y_n) be an arbitrary sequence from $f((a,b)) \setminus \{y_0\}$ which converges to y_0. Set $x_0 = f^{-1}(y_0)$ and $x_n = f^{-1}(y_n)$. It suffices to show that $\lim_{n\to\infty} \frac{f^{-1}(y_n) - f^{-1}(y_0)}{y_n - y_0} = \frac{1}{f'(x_0)}$.

Since f^{-1} is continuous, we see that $\lim_{n\to\infty} f^{-1}(y_n) = f^{-1}(y_0)$, that is, $\lim_{n\to\infty} x_n = x_0$. Thus, since f is differentiable at x_0, we have

$$\lim_{n\to\infty} \frac{f(x_n) - f(x_0)}{x_n - x_0} = f'(x_0).$$

Since f is injective, $f(x_n) - f(x_0) \neq 0$ unless $x_n = x_0$, so by a property of limits of sequences we have

$$\frac{1}{f'(x_0)} = \lim_{n\to\infty} \frac{x_n - x_0}{f(x_n) - f(x_0)} = \lim_{n\to\infty} \frac{f^{-1}(y_n) - f^{-1}(y_0)}{y_n - y_0}.$$

$\square$

6. Integration

The definition of the Riemann Integral, as presented in Thomas' Calculus.

Definition 6. Let $a, b \in \mathbb{R}$ with $a < b$.

A *partition* of the closed interval $[a, b]$ is a finite set

$$P = \{x_0, x_1, x_2, \ldots, x_n\}$$

with the property that

$$a = x_0 < x_1 < \cdots < x_{n-1} < x_n = b.$$

Let $P = \{x_0, x_1, x_2, \ldots, x_n\}$ be a partition of $[a, b]$. We view P as indicating a way of breaking the interval $[a, b]$ into n subintervals. The width of the i^{th} subinterval is $\Delta x_i = x_i - x_{i-1}$, for $i = 1, \ldots, n$.

The *norm* of the partition P is

$$||P|| = \max\{\Delta x_i \mid i = 1, \ldots, n\}.$$

A *choice set* for P is a finite set

$$C = \{c_1, c_2, \ldots, c_n\}$$

such that $c_i \in [x_{i-1}, x_i]$, for $i = 1, \ldots, n$. Note that this implies

$$c_1 < c_2 < \cdots < c_n.$$

Let $f : [a, b] \to \mathbb{R}$. The *Riemann sum* associated to a partition P and a choice set C for P is

$$R(f, P, C) = \sum_{i=1}^{n} f(c_i)\Delta x_i.$$

We say that f is *Riemann integrable with integral I* if there exists a real number $I \in \mathbb{R}$ such that, for every positive real number $\epsilon > 0$, there exists a real number $\delta > 0$ such that for every partition P and choice set C of P,

$$\|P\| < \delta \quad \Rightarrow \quad |R(f, P, C) - I| < \epsilon.$$

If f is Riemann integrable with integral I, we write

$$\int_a^b f(x)\, dx.$$

This is read, "the integral from a to b of $f(x)\, dx$".

The Riemann integral represents the area between the graph of f and the x-axis. Note that this is *signed area*; that is, area below the x-axis is counted as negative.

Remark 1 (Properties of the Riemann Integral)**.** Let $f : [a, b] \to \mathbb{R}$ and $g : [a, b] \to \mathbb{R}$ be integrable.
Let $c, d \in [a, b]$ with $a \le c \le d \le b$. Let $k \in \mathbb{R}$.

(a) f is integrable on $[c, d]$
(b) $\int_a^a f(x)\, dx = 0$
(c) $\int_b^a f(x)\, dx = -\int_a^b f(x)\, dx$
(d) $\int_a^c f(x)\, dx + \int_c^b f(x)\, dx = \int_a^b f(x)\, dx$
(e) $\int_a^b kf(x)\, dx = k\int_a^b f(x)\, dx$
(f) $\int_a^b f(x) + g(x)\, dx = \int_a^b f(x)\, dx + \int_a^b g(x)\, dx$
(g) if $f(x) \le g(x)$ for all $x \in [a, b]$, then $\int_a^b f(x)\, dx \le \int_a^b f(x)\, dx$

Theorem 4 (Fundamental Theorem of Calculus, Part I). **(FTC I)**
Let $f : [a, b] \to \mathbb{R}$ be integrable. Define a function

$$F : [a, b] \to \mathbb{R} \quad \text{by} \quad F(x) = \int_a^x f(t)\, dt.$$

Then F is differentiable at x for $x \in (a, b)$, and $F'(x) = f(x)$.

Reason. Consider

$$F(x + h) - F(x) = \int_a^{x+h} f(t)\, dt - \int_a^x f(t)\, dt = \int_x^{x+h} f(t)\, dt.$$

Now $\int_x^{x+h} f(t)\, dt$ is the area under the graph of f from x to $x + h$. Since f is continuous, it is clear that, for very small h, this area is approximately the area of the rectangle whose height is $f(x)$ and whose width is h; that is,

$$\int_x^{x+h} f(t)\, dt \approx f(x)h.$$

Thus, for very small h,

$$F'(x) \approx \frac{F(x + h) - F(x)}{h} = \frac{\int_x^{x+h} f(t)\, dt}{h} \approx \frac{f(x)h}{h} = f(x).$$

These approximations become precise as h approaches zero, so

$$F'(x) = \lim_{h\to 0} \frac{F(x + h) - F(x)}{h} = f(x).$$

□

Theorem 5 (Fundamental Theorem of Calculus, Part II). **(FTC II)**
Let $f : [a, b] \to \mathbb{R}$ and suppose that F is an antiderivative for f on (a, b). Then

$$\int_a^b f(t)\, dt = F(b) - F(a).$$

Proof. Let $G(x) = \int_a^x f(t)\, dt$. Then by FTC I, G is differentiable on (a, b), and $G'(x) = F'(x) = f(x)$. Since F and G have the same derivative, they differ by a constant. Thus there exists a constant $C \in \mathbb{R}$ such that

$$G(x) = F(x) + C \quad \text{for all } x \in [a, b].$$

Plugging in $x = a$, we have $G(a) = F(a) + C$. But $G(a) = \int_a^a f(x)\, dx = 0$, so $F(a) = -C$, so

$$G(x) = F(x) - F(a).$$

Finally, plug in $x = b$ to get $G(b) = F(b) - F(a)$, so

$$\int_a^b f(x)\, dx = F(b) - F(a).$$

□

Department of Mathematics and CSci, BASIS Scottsdale
Email address: paul.bailey@basised.com