Real Analysis:
Single Variable Calculus.

Edward Chernysh

Last updated: August 8, 2017.

Contents

In these notes we study calculus in a single variable from a rigorous point of view: that of mathematical analysis. In almost all calculus courses available, the subject is treated from an "engineering perspective" in which the student is only shown the methods and not the reasoning or abstractions. We do the opposite here. Instead of providing countless examples of tedious and unenlightening calculations, we shall prove the results covered in standard calculus courses, leaving no doubt as to their truth.

This first text in *real analysis* covers standard topics: differentiation, Riemann integration and more diverse topics such as null sets and the Lebesgue criterion for Riemann integrability. The first part covers the derivative and its applications, including proofs of the "differentiation rules" the reader is likely familiar with.

The second part introduces the notion of a *Riemann integral* in a proper setting, without the ambiguity oft found in calculus texts.

In this text we assume the reader has distinguished countable infinity from uncountable infinities, and has understood the rigorous $\varepsilon - \delta$ or sequential formulations of the limit.

1 Differentiation

This section is devoted to the topic of differentiation. Most readers should be aware of the motivation behind the derivative. Loosely speaking, given a function $f : [a, b] \to \mathbb{R}$ we wish to determine the conditions for which we can "measure" the "instantaneous rate of change" at a point $c \in (a, b)$. This is made easier when one understands the $\varepsilon - \delta$ formulation of the limit.

Despite being a prerequisite for this text, we shall state this definition below for the sake of completeness.

Definition 1. Let $I \subseteq \mathbb{R}$ be an open set and $f : I \longrightarrow \mathbb{R}$. If $c \in I$ we write

$$\lim_{x \to c} f(x) = L,$$

for a complex number L, if for all $\varepsilon > 0$ there exists $\delta > 0$ such that $x \in I$ with $0 < |x - c| < \delta$ implies

$$|f(x) - L| < \varepsilon.$$

1.1 Monotonicity

We briefly consider the *continuity* of monotone functions. We shall see that monotonic functions are "almost continuous".

Theorem 1.1 (Injectivity and Monotonicity). Let $f : I \longrightarrow \mathbb{R}$ be monotonic and injective. Then f^{-1} is monotone.

Proof. Without loss of generality we suppose that f is increasing on I. We now show f^{-1} is also increasing on $f(I)$. We argue by contradiction, so suppose not. Then there exists $y_1, y_2 \in f(I)$ with $y_1 < y_2$ but $f^{-1}(y_1) > f^{-1}(y_2)$. Thus, taking $x_1 = f^{-1}(y_1)$ and $x_2 = f^{-1}(y_2)$ in I we have $x_1 > x_2$. By hypothesis on f this implies that $f(x_1) \geq f(x_2)$, and $y_1 \geq y_2$. Contradiction. $\square$

Theorem 1.2. Let $I \subseteq \mathbb{R}$ be an interval and $f : I \to \mathbb{R}$ be monotonic. Denote by $\mathfrak{D}$ the set of all discontinuity points of f in I. Then, $\mathfrak{D}$ is countable.

Proof. Let $c \in \mathfrak{D}$. Without harm to the proof let us suppose that f is increasing. Thus, since f is discontinuous at c, we necessarily have (recall that f is monotone increasing)

$$\lim_{x \to c+} f(x) > \lim_{x \to c-} f(x).$$

Now, for any $c \in \mathfrak{D}$ we may define a non-empty open interval:

$$\Delta_c := \left(\lim_{x \to c-} f(x), \lim_{x \to c+} f(x) \right).$$

Moreover, for $c, d \in \mathfrak{D}$ with $c \neq d$ we have $\Delta_c \cap \Delta_d = \varnothing$. To see this, suppose not. Then there exists $z \in \Delta_c$ and $z \in \Delta_d$. Assume without loss of generality that $c < d$. Now, $z \in \Delta_c$ implies $z < \lim_{x \to c+} f(x)$, but $z \in \Delta_d$ implies $z > \lim_{x \to d-} f(x)$. However,

$$\lim_{x \to c+} f(x) \leq \lim_{x \to d-} f(x)$$

since $c < d$. Contradiction, so $\Delta_c \cap \Delta_d = \varnothing$.

Therefore, to each $c \in \mathcal{D}$ we may associate an open interval Δ_c. By the density of $\mathbb{Q}$ in $\mathbb{R}$ we select some $q_c \in \Delta_c$. Since each interval is disjoint, each selected rational is unique. Define now a map:

$$\Psi : \mathcal{D} \to \mathbb{Q}, \quad c \mapsto q_c$$

Clearly, this map is injective and hence the set $\mathcal{D}$ is countable (since[1] $\#\mathbb{Q} = \#\mathbb{N}$ and $\#\mathfrak{D} \leq \#\mathbb{Q}$). $\square$

1.2 The Derivative, Continuity and Local Extrema

We are now ready to define the derivative of a function.

Definition 2. Let $f : [a, b] \longrightarrow \mathbb{R}$ and $c \in (a, b)$. We say f is differentiable at c if the limit below exists, if this is the case, we call

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

the derivative of f at c.

[1] The notation $\#(\cdot)$ represents the cardinality of a set.

Theorem 1.3. Let $f : [a,b] \subseteq \mathbb{R} \longrightarrow \mathbb{R}$ be differentiable at a point $c \in (a,b)$. Then, f is continuous at c.

Proof. For $x \neq c$ we may write

$$(f(x) - f(c)) = \left(\frac{f(x) - f(c)}{x - c}\right)(x - c).$$

Note now that,

$$\lim_{x \to c}\left(\frac{f(x) - f(c)}{x - c}\right)(x - c) = \lim_{x \to c}\left(\frac{f(x) - f(c)}{x - c}\right)(x - c) \cdot \lim_{x \to c}(x - c) = f'(c) \cdot 0 = 0.$$

Thus, $\lim_{x \to c}(f(x) - f(c)) = 0$, whence $\lim_{x \to c} f(x) = f(c)$ thereby establishing continuity. □

Just as in calculus courses, we consider briefly the notion of an *extremum*. Assume $I \subseteq \mathbb{R}$ is an interval of the form (a,b), where $a = -\infty$ and $b = \infty$ are allowed. If $c \in I$ we say that f achieves a **local maximum** at c provided there exists $\delta > 0$ such that

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

whenever $x \in (c - \delta, c + \delta)$. We may often write $V_\delta(c) := (c - \delta, c + \delta)$ as well.

Theorem 1.4. Let I be an open interval and assume $f : I \longrightarrow \mathbb{R}$ is differentiable at a point $c \in I$. If c is a local extremum, then $f'(c) = 0$.

Proof. If $f(c)$ is a local maximum, then $-f'(c)$ is a local minimum, and $f'(c) = 0$ if and only if $-f'(c) = 0$; thus we may assume without harm that $f(c)$ is a local maximum.

Thus, there exists $\delta > 0$ such that $x \in V_\delta(c)$ implies $f(x) \leq f(c)$. Consider the two disjoint intervals $A := (c - \delta, c)$ and $B := (c, c + \delta)$. For all $x \in A$ one has:

$$\frac{f(x) - f(c)}{x - c} \geq 0 \implies \lim_{x \to c^-} \frac{f(x) - f(c)}{x - c} \geq 0 \text{ (if it exists)}$$

And for all $x \in B$:

$$\frac{f(x) - f(c)}{x - c} \leq 0 \implies \lim_{x \to c^+} \frac{f(x) - f(c)}{x - c} \leq 0 \text{ (if it exists)}$$

Since the derivative exists at c, both sided limits must be equal. Hence, $f'(c) \geq 0$ and $f'(c) \leq 0$: whence $f'(c) = 0$. □

1.3 The Chain Rule

There are many proofs of the chain rule, several of which are horribly tedious. We offer here an elegant proof, one that requires a preliminary criterion.

Lemma 1.5. Let $I \subseteq \mathbb{R}$ be an open interval and $f : I \to \mathbb{R}$ with $c \in I$. The following statements are equivalent:

1. f is differentiable at c.
2. There exists a function $\varphi : I \to \mathbb{R}$ on I such that $\varphi(c) = f'(c)$.

Proof. (1) $\implies$ (2). Let f be differentiable at c. Define a function

$$\varphi(x) := \begin{cases} \frac{f(x)-f(c)}{x-c} & x \neq c \\ f'(c) & x = c \end{cases}$$

This function is clearly continuous on $I \setminus \{c\}$, furthermore continuity at c follows from the fact that the derivative at c exists, certainly:

$$\lim_{x \to c} \phi(x) = \lim_{x \to c} \frac{f(x) - f(c)}{x - c} = f'(c).$$

(2) $\implies$ (1). Suppose such a function exists. Then, for all $x \in I \setminus \{c\}$ one has:

$$f(x) - f(c) = \varphi(x)(x - c).$$

Thence, for all $x \in I$ with $x \neq c$,

$$\frac{f(x) - f(c)}{x - c} = \varphi(x).$$

Taking the limit as $x \to c$ it follows by continuity of φ,

$$\lim_{x \to c} \frac{f(x) - f(c)}{x - c} = \lim_{x \to c} \varphi(x) = \varphi(c).$$

So, $f'(c)$ exists at c and $f'(c) = \varphi(c)$. $\square$

The Chain Rule. Let I, J be open intervals in $\mathbb{R}$, $f : I \longrightarrow \mathbb{R}$ and $g : J \longrightarrow \mathbb{R}$. Suppose that $f(I) \subseteq J$ If f is differentiable at $c \in I$ and g is differentiable at $d = f(c)$ then $(g \circ f)'(c) = g'(f(c)) \cdot f'(c)$.

Proof. Since f is differentiable at c the previous lemma ensures that there exists a function $\varphi : I \longrightarrow \mathbb{R}$ such that $\varphi(c) = f'(c)$ and

$$f(x) - f(c) = \varphi(x)(x - c) \quad \text{for all } x \in I \text{ with } x \neq c.$$

Similarly, by differentiability of g at d there exists $\phi : J \longrightarrow \mathbb{R}$ such that $\phi(d) = g'(d)$ and

$$g(y) - g(d) = \phi(y)(y - d) \quad \text{for } y \in J \text{ with } y \neq d.$$

Consider for $y = f(x) \neq f(c)$:

$$g(y) - g(d) = g(f(x)) - g(f(c)) = \phi(f(x))(f(x) - f(c)).$$

Or, rather,

$$g(f(x)) - g(f(c)) = \phi(f(x))\varphi(x)(x - c)$$

whence

$$\phi(f(x))\varphi(x) = \frac{g(f(x)) - g(f(c))}{x - c}, \quad \text{for all } x \in I \text{ with } x \neq c.$$

Passing to the limit as $x \to c$ we conclude,

$$\lim_{x \to c} \psi(f(x))\phi(x) = \psi(f(c))\phi(c) = \lim_{x \to c} \frac{g(f(x)) - g(f(c))}{x - c} = (g \circ f)'(c)$$

Since $\phi(f(c)) = g'(f(c))$ and $\varphi(c) = f'(c)$ the result follows. □

1.4 The Mean Value Theorem

Unfortunately, because of how modern calculus courses are taught, the *mean value theorem* is usually the most forgetful result. In reality, the mean value theorem is possibly the most powerful and important result arising from differential calculus. If there is one theorem one should remember, it is the mean value theorem!

Rolle's Theorem. Suppose $f : [a, b] \longrightarrow \mathbb{R}$ is continuous on $[a, b]$ and differentiable on (a, b). Assume further that $f(a) = f(b) = 0$, there exists a point $c \in (a, b)$ such that $f'(c) = 0$.

Proof. By continuity on a closed and bounded (i.e. compact) interval, the function achieves absolute maximum on $[a, b]$. Let $m, M \in \mathbb{R}$ be the minimum and maximum values respectively. We distinguish two possibilities:

1. $M = 0 = m$. In this case $f \equiv 0$ which implies that $f'(x) = 0$ for all $x \in (a, b)$. Hence any $c \in (a, b)$ is admissible.

2. One of $m, M \neq 0$. That is, it achieves some local extrema at a point c in the open interval (a, b). Since it is a local extrema, a previous theorem implies $f'(c) = 0$.

□

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

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

Proof. Consider the function

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

By composition this function is continuous on I and differentiable on (a, b). Moreover, $g(a) = 0 = g(b)$, thus by Rolle's Theorem there exists a point $c \in (a, b)$ such that $g'(c) = 0$. On the other-hand, it is obvious that

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

Thus,

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

□

Proposition 1.6. Let $f : [a, b] \longrightarrow \mathbb{R}$ be continuous[2] and differentiable on (a, b). If $f'(x) = 0$ for all $x \in (a, b)$ then f is a constant function.

Proof. We prove $f(a) = f(x)$ for all $x \in (a, b]$. Let $x \in (a, b)$ and consider the interval $[a, x]$. We know f is continuous on $[a, x]$ and differentiable on (a, x). Thus, by the mean value theorem there is a point $c \in (a, x)$ such that

$$0 = f'(c) = \frac{f(x) - f(a)}{x - a}.$$

Therefore, $f(x) = f(a)$. Since $x \in (a, b]$ was arbitrary, the proof is concluded. □

Proposition 1.7. In the context of the previous theorem suppose $f'(x) = g'(x)$ for all $x \in (a, b)$. Then $f(x) = g(x) + C$ for some constant C.

Proof. Set $h(x) = f(x) - g(x)$. Then $h'(x) = 0$ everywhere which implies $h(x)$ is constant, ie $h(x) = C$ and hence $f(x) = g(x) + C$. □

Theorem 1.8. Let $f : I = [a, b] \longrightarrow \mathbb{R}$ be continuous on $[a, b]$ and differentiable in (a, b). Then, f is monotone increasing on I if and only if $f'(x) \geq 0$ on (a, b).

[2] We mean continuous on $[a, b]$.]

Proof. Suppose f is increasing on I. That is, given any $x, y \in (a, b)$ with $x < y$ we have $f(x) \leq f(y)$. Then,

$$f'(y) = \lim_{x \to y^-} \frac{f(x) - f(y)}{x - y} \geq 0.$$

Since f is differentiable the two-sided limits must be equal, so $f'(y) \geq 0$.

Conversely, suppose $f'(x) \geq 0$ for all $x \in (a, b)$. Pick any $x_1, x_2 \in [a, b]$ with $x_1 < x_2$. Then, by the *mean value theorem* there exists $z \in (x_1, x_2)$ such that

$$f'(z) = \frac{f(x_2) - f(x_1)}{x_2 - x_1} \geq 0.$$

Since $x_2 > x_1$ we must have $f(x_2) - f(x_1) \geq 0$, and $f(x_2) \geq f(x_1)$. □

Theorem 1.9. Let $f : (a, b) \longrightarrow \mathbb{R}$ be differentiable on (a, b). If there is a neighborhood $(c - \delta, c + \delta) \subseteq (a, b)$ such that $x \in (c - \delta, c)$ implies $f'(x) \geq 0$ and for each $x \in (c, c + \delta)$ one $f'(x) \leq 0$ there exists a local maximum at c.

Proof. Let $x \in (c - \delta, c)$, by the *mean value theorem* there exists $c = c_x \in (x, c)$ such that

$$\frac{f(c) - f(x)}{c - x} = f'(c_x) \geq 0 \implies f(x) \leq f(c).$$

Choose $y \in (c, c + \delta)$, there exists $c = c_y \in (c, y)$ such that

$$\frac{f(y) - f(c)}{y - c} \leq 0 \implies f(y) \leq f(c).$$

□

1.5 The Intermediate Properties of the Derivative

Lemma 1.10. If $f'(c) > 0$ there exists $\delta > 0$ such that $f(x) > f(c)$ for $x \in (c, c + \delta)$.

Proof. Choose $\varepsilon := \frac{f(c)}{2} > 0$. Then by definition of limit there exists $\delta > 0$ such that $0 < |x - c| < \delta$ implies

$$\left| \frac{f(x) - f(c)}{x - c} - f'(c) \right| < \frac{f'(c)}{2}.$$

Thence for all such x:

$$0 < \frac{f'(c)}{2} < \frac{f(x) - f(c)}{x - c}.$$

This implies, $f(x) > f(c)$ whenever $x \in (c, c + \delta)$. □

Darboux Theorem. If f is differentiable on an open interval containing $[a, b]$ and $k \in (f'(a), f'(b))$ there exists $c \in (a, b)$ such that $f'(c) = k$.

Proof. Let $g(x) := kx - f(x)$. By continuity this function achieves a maximum on $[a, b]$. Observe that

$$g'(a) = k - f'(a) > 0.$$

It follows by the previous lemma that there is a point, say, $c > a$ where $g(c) > a$. Therefore, g does not achieve its maximum at a. Similarly, $g'(b) < 0$ and an argument nearly identical to that used in the previous lemma shows there is a point $c < b$ where $g(c) > g(b)$. Hence, g achieves its maximum at a point $x_\star \in (a, b)$; thus $x_\star$ is a *local maximum*. A previous result then ensures that,

$$g'(x_\star) = k - f'(x_\star) = 0$$

since $g'(x_\star) = 0$. Take $c := x_\star$. $\square$

2 Riemann Integration

In this section we turn to the notion of integration. There are many different *theories of integration*; the most commonly known is that due to Riemann, which we explore in this section.

As in the previous section, our goal here is to develop the Riemann integral with rigor and precision. Although very useful, it is important to realize that this is a "primitive" integral. There are more modern theories of integration, the most prominent of which is Lebesgue integration. The readers who go on to study more advanced analysis shall likely see this constructed; we cannot do this here as it requires more advanced notions such as measure theory. What is important to understand are the benefits of Lebesgue integration over the theory that we shall unveil here: the Lebesgue integral is much more *forgiving*, i.e. there are many functions that are not Riemann integrable yet have a Lebesgue integral.

Nevertheless, most "nice" functions have a Riemann integral and for many cases this theory of integration is sufficient.

Throughout this section, we denote by I an interval of the form $[a, b]$ in $\mathbb{R}$.

2.1 Partitions, Tags and the Riemann Integral

Definition 3 (Partition). Let $I \subseteq \mathbb{R}$ be an interval, a set denoted $\mathcal{P} := \{[x_{i-1}, x_i]\}_{i=1}^{n}$ is called a partition of I provided that x_0, x_n are the end points of I, $x_{i-1} \leq x_i$ for each i and

$$\bigcup_{i=1}^{n}[x_{i-1}, x_i] = I.$$

A collection $\dot{\mathcal{P}} := \left\{[x_[i-1, x_i], u_i\right\}_{i=1}^{n}$ is called a tagged partition of I provided the intervals $[x_{i-1}, x_i]$ form a partition of I and each $u_i \in [x_{i-1}, x_i]$.

Definition 4. Let $I \subseteq \mathbb{R}$ and $\mathcal{P}$ be a partition of I. We define the **mesh** of a partition, denoted $\|\cdot\|$, by:

$$\left\|\dot{\mathcal{P}}\right\| := \max_{\substack{1 \leq i \leq n \\ i \in \mathbb{N}}} |x_i - x_{i-1}|$$

Definition 5 (Riemann Sum). Let $f : I \longrightarrow \mathbb{R}$ and $\dot{\mathcal{P}}$ be a tagged partition of I. The Riemann Sum of f over $\dot{\mathcal{P}}$ is

$$S(f; \dot{\mathcal{P}}) := \sum_{i=1}^{n} f(u_i)\Delta x_i$$

where $\Delta x_i := (x_i - x_{i-1})$.

Definition 6 (Riemann Integral)**.** Let $f : I \longrightarrow \mathbb{R}$, we say f is **Riemann Integrable** on I provided that there exists $\mathcal{I} \in \mathbb{R}$ such that for all $\varepsilon > 0$ there is $\delta > 0$ so that for any $\dot{\mathcal{P}}$ with $\left\|\dot{\mathcal{P}}\right\| < \delta$ one has

$$\left|S(f;\dot{\mathcal{P}}) - \mathcal{I}\right| < \varepsilon. \tag{1}$$

If this is the case, we call $\mathcal{I}$ the **Riemann Integral** of f on I which we denote $\int_I f$.

From hereon, given an interval $[a,b] \subset \mathbb{R}$, the **space of Riemann integrable functions** $f : [a,b] \longrightarrow \mathbb{R}$ will be denoted by $\mathcal{R}([a,b])$.

2.2 Uniqueness and Elementary Alegbraic Proprties

Theorem 2.1. Let $f \in \mathcal{R}([a,b])$, then the Riemann integral of f is unique.

Proof. We prove this by contradiction. Suppose not, then there exists at least two $L, L' \in \mathbb{R}$ satisfying (1) with $L' \neq L$. Then, for any $\varepsilon > 0$ there exists $\delta_1 > 0$ such that $\dot{\mathcal{P}}$ with $\left\|\dot{\mathcal{P}}\right\| < \delta_1$ implies $\left|S(f;\dot{\mathcal{P}}) - L\right| < \varepsilon/2$. Similarly, we may find δ_2 such that $\left\|\dot{\mathcal{P}}\right\| < \delta_2$ implies $\left|S(f;\dot{\mathcal{P}}) - L'\right| < \varepsilon/2$. Especially, setting $\delta := \min\{\delta_1, \delta_2\}$ we see that for any $\dot{\mathcal{Q}}$ with $\left\|\dot{\mathcal{Q}}\right\| < \delta$ one has,

$$\left|L - L'\right| \leq \left|S(f;\dot{\mathcal{Q}}) - L\right| + \left|S(f;\dot{\mathcal{Q}}) - L'\right| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

Hence for all $\varepsilon > 0$, $|L - L'| < \varepsilon$ which implies $L = L'$. Contradiction. □

Example. Any constant function $f : I \longrightarrow \mathbb{R}$, defined by $f(x) = k$, is Riemann Integrable with integral $\int_I f = k(b-a)$. To see this, let $\varepsilon > 0$ and take $\delta = 1 > 0$, for any tagged partition $\dot{\mathcal{P}} = \{[x_{i-1}, x_i], u_i\}$ we have

$$\begin{aligned}\left|S(f;\dot{\mathcal{P}}) - k(b-a)\right| = \left|\sum_{i=1}^{n} f(u_i)\Delta x_i - k(b-a)\right| &= k\left|\sum_{i=1}^{n}(x_i - x_{i-1}) - (b-a)\right| \\ &= k\left|(b-a) - (b-a)\right| \\ &= 0 < \varepsilon\end{aligned} \tag{$\mathfrak{A}$}$$

Theorem 2.2 (Properties)**.** Let $f, g \in \mathcal{R}([a,b])$.

(1) Fix $k \in \mathbb{R}$, then kf is Riemann Integrable on I with integral $k\int_I f$.

(2) $(f+g)(x) \in \mathcal{R}([a,b])$ with integral $\int_I f + \int_I g$.

(3) If $f(x) \leq g(x)$ on I then $\int_I f \leq \int_I g$.

Proof.

1. Fix $k \neq 0$ and take $\varepsilon > 0$, by hypothesis $f \in \mathcal{R}([a,b])$ which implies there exists $\delta > 0$ such for any $\dot{\mathcal{P}}$ with $\left\|\dot{\mathcal{P}}\right\| < \delta$ implies $\left|S(f;\dot{\mathcal{P}}) - \int_I f\right| < \varepsilon/k$. Indeed, for such δ it easily follows

$$\left|S(kf;\dot{\mathcal{P}}) - k\int_I f\right| = \left|\sum_{i=1}^{n} kf(u_i)\Delta x_i - k\int_I f\right| = \left|S(f;\dot{\mathcal{P}}) - \int_I f\right| < k \cdot \frac{\varepsilon}{k} = \varepsilon$$

   Since this holds trivially for $k = 0$ by ($\mathfrak{A}$) we are done.

2. Let $\varepsilon > 0$ be given, there exists $\delta > 0$ (take the minimum of two δ's) such that for any $\dot{\mathcal{P}}$ with $\left\|\dot{\mathcal{P}}\right\| < \delta$ implies both $\left|S(f;\dot{\mathcal{P}}) - \int_I f\right| < \varepsilon/2$ and $\left|S(g;\dot{\mathcal{P}}) - \int_I g\right| < \varepsilon/2$. Whence,

$$\begin{aligned}
\left|S(f+g;\dot{\mathcal{P}}) - \left(\int_I f + \int_I g\right)\right| &= \left|\sum_{i=1}^{n}(f(u_i) + g(u_i)\Delta x_i - \int_I f - \int_I g\right| \\
&= \left|\sum_{i=1}^{n} f(u_i)\Delta x_i - \int_I f + \sum_{i=1}^{n} g(u_i)\Delta x_i - \int_I g\right| \\
&= \left|S(f;\dot{\mathcal{P}}) - \int_I f + S(g;\dot{\mathcal{P}}) - \int_I g\right| \\
&\leq \left|S(f;\dot{\mathcal{P}}) - \int_I f\right| + \left|S(g;\dot{\mathcal{P}}) - \int_I g\right| \\
&< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.
\end{aligned}$$

3. Let $\varepsilon > 0$ be given; by Riemann Integrability of both function we have $\delta > 0$ such that for any $\dot{\mathcal{P}}$ with $\left\|\dot{\mathcal{P}}\right\| < \delta$

$$\left|S(f;\dot{\mathcal{P}}) - \int_I f\right| < \frac{\varepsilon}{2} \text{ and } \left|S(g;\dot{\mathcal{P}}) - \int_I g\right| < \frac{\varepsilon}{2}$$

   From each we derive:

$$\int_I f - \frac{\varepsilon}{2} < S(f;\dot{\mathcal{P}}) \text{ and } S(g;\dot{\mathcal{P}}) < \int_I g + \frac{\varepsilon}{2}$$

   But $f(x) \leq g(x)$ on I implies $S(f;\dot{\mathcal{P}}) \leq S(g;\dot{\mathcal{P}})$ whence

$$\int_I f - \frac{\varepsilon}{2} < \int_I g + \frac{\varepsilon}{2}$$

Or equivalently,
$$\int_I f < \int_I g + \varepsilon, \quad \text{for each } \varepsilon > 0.$$
Passing to the limit as $\varepsilon \to 0$ we have
$$\int_I f \leq \int_I g.$$

□

2.3 Riemann Integrable Functions are Bounded

Theorem 2.3 (Boundedness)**.** Every Riemann Integrable function is bounded.

Proof. By way of contradiction let $f \in \mathcal{R}([a,b])$ be an unbounded function on I. By hypothesis f is Riemann Integrable, which implies given $\varepsilon = 1$ there exsts $\delta > 0$ such that for any $\dot{\mathcal{P}}$ of I with mesh $\left\|\dot{\mathcal{P}}\right\| < \delta$ implies $\left|S(f;\dot{\mathcal{P}}) - \int_I f\right| < 1$. Recall the triangle inequality in the form $|a| - |b| \leq |a - b|$. From which we derive from the above inequality
$$\left|S(f;\dot{\mathcal{P}})\right| < 1 + \underbrace{\left|\int_I f\right|}_{:=L} = 1 + L$$
However since f is unbounded on I it is unbounded on some subinterval of $\dot{\mathcal{P}}$, whence we can pick some $u_k \in [x_{k-1}, x_k] \subset \dot{\mathcal{P}}$ such that
$$|f(u_k)\Delta x_k| > 1 + L + \left|\sum_{i \neq k} f(u_i)\Delta x_i\right|$$
However,
$$\begin{aligned}\left|S(f;\dot{\mathcal{P}})\right| = \left|\sum_{i=1}^{n} f(u_i)\Delta x_i\right| &= \left|f(u_k)\Delta x_k + \sum_{i \neq k} f(u_i)\Delta x_i\right| \\ &\geq |f(u_k)\Delta x_k| - \left|\sum_{i \neq k} f(u_i)\Delta x_i\right| \\ &> 1 + L.\end{aligned}$$
Contradiction.

□

2.4 Cauchy's Criteria

In practice it is quite difficult to determine whether or not a function $f : I \longrightarrow \mathbb{R}$ is Riemann integrable. Luckily, there are several equivalent formulations of Riemann integrability that are easier to use in practice; these are presented below.

2.4.1 Cauchy's First Criterion

Theorem 2.4 (Cauchy's First Criterion). A function $f : I \longrightarrow \mathbb{R}$ is Riemann Integrable on I if and only if for all $\varepsilon > 0$ there exists $\delta > 0$ such that for any two $\dot{\mathcal{P}}, \dot{\mathcal{Q}}$ of I with mesh less than δ one has

$$\left|S(f;\dot{\mathcal{P}}) - S(f;\dot{\mathcal{Q}})\right| < \varepsilon.$$

Proof. Suppose now that $f \in \mathcal{R}([a,b])$, then for all $\varepsilon > 0$ there exists $\delta > 0$ such that for any tagged partition $\dot{\mathcal{W}}$ with $\left\|\dot{\mathcal{W}}\right\| < \delta$ implies

$$\left|S(f;\dot{\mathcal{W}}) - \int_I f\right| < \varepsilon/2.$$

Hence, for any two tagged partitions $\dot{\mathcal{P}}, \dot{\mathcal{Q}}$ that satisfy this we have:

$$\left|S(f;\dot{\mathcal{Q}}) - S(f;\dot{\mathcal{P}})\right| \leq \left|S(f;\dot{\mathcal{Q}}) - \int_I f\right| + \left|S(f;\dot{\mathcal{P}}) - \int_I f\right| < \varepsilon/2 + \varepsilon/2 = \varepsilon.$$

Conversely, suppose for all $\varepsilon > 0$ there exists $\delta > 0$ such that for any two $\dot{\mathcal{P}}, \dot{\mathcal{Q}}$ of I with mesh less than δ implies $\left|S(f;\dot{\mathcal{P}}) - S(f;\dot{\mathcal{Q}})\right| < \varepsilon$. Taking $\varepsilon = 1/n$ for $n \in \mathbb{N}$ we construct a sequence $(\delta_n) \subset \mathbb{R}$ such that for any $\dot{\mathcal{P}}_n$, $\dot{\mathcal{Q}}_n$ with $\left\|\dot{\mathcal{P}}_n\right\| < \delta_n$ and $\left\|\dot{\mathcal{Q}}_n\right\| < \delta_n$ implies

$$\left|S(f;\dot{\mathcal{P}}_n) - S(f;\dot{\mathcal{Q}}_n)\right| < \frac{1}{n}.$$

There is no harm in choosing $\delta_n \geq \delta_{n+1}$. Hence, given $\varepsilon > 0$ we may find $N \in \mathbb{N}$ such that $1/N < \varepsilon$, whence for $n, m \geq N$ we have $\delta_n, \delta_m \leq \delta_N$ as well as,

$$\left|S(f;\dot{\mathcal{P}}_n) - S(f;\dot{\mathcal{P}}_m)\right| < \varepsilon.$$

In other words, $(S(f;\dot{\mathcal{P}}_n))_{n=1}^{\infty}$ is a Cauchy sequence in $\mathbb{R}$. By completeness, this sequence converges to $L \in \mathbb{R}$. So, given $\varepsilon > 0$ we may choose $N_1 \in \mathbb{N}$ such that $n \geq N_1$ implies $\left|S(f;\dot{\mathcal{P}}_n) - L\right| < \varepsilon/2$. Similarly, for this same $\varepsilon > 0$ there exists $N_2 \in \mathbb{N}$ such that $1/N_2 < \varepsilon/2$, whence for any $n, m \geq N_2$ we have

$$\left|S(f;\dot{\mathcal{P}}_n) - S(f;\dot{\mathcal{P}}_m)\right| < 1/N_2 < \varepsilon/2.$$

We claim $L = \int_I f$, to see this consider $N := \max\{N_1, N_2\}$ and let $\dot{\mathcal{P}}$ be any tagged partition of I with $\left\|\dot{\mathcal{P}}\right\| < \delta_N$. Note that since $\left\|\dot{\mathcal{P}}\right\| < \delta_N \leq \delta_n, \delta_m$:

$$\left|S(f;\dot{\mathcal{P}}) - L\right| \leq \left|S(f;\dot{\mathcal{P}}) - S(f;\dot{\mathcal{P}}_N)\right| + \left|S(f;\dot{\mathcal{P}}_N) - L\right| < \frac{1}{N} + \frac{\varepsilon}{2} < \varepsilon.$$

Thus, $f \in \mathcal{R}([a,b])$. □

2.4.2 Cauchy's Second Criterion

Again, $I \subseteq \mathbb{R}$ is a closed and bounded interval.

Theorem 2.5 (Cauchy's Second Criterion). A function $f : I \longrightarrow \mathbb{R}$ is Riemann Integrable on I if and only if given $\varepsilon > 0$ there exist maps $\alpha, \beta : I \longrightarrow \mathbb{R}$, both Riemann Integrable on I, such that $\alpha(x) \leq f(x) \leq \beta(x)$ for all $x \in I$ and

$$\int_I (\beta - \alpha) < \varepsilon.$$

Proof. Suppose $f \in \mathcal{R}([a,b])$, then take $\alpha(x) := \beta(x) := f(x)$, whence the " $\Longrightarrow$ " direction follows immediately.

Conversely, let $\varepsilon > 0$ be given, there exists $\delta_\alpha, \delta_\beta > 0$ such that for any partition $\dot{\mathcal{P}}$ with $\left\|\dot{\mathcal{P}}\right\| < \delta_\alpha$ and $\left\|\dot{\mathcal{P}}\right\| < \delta_\beta$ implies both

$$\left|S(f;\alpha) - \int_I \alpha\right| < \varepsilon/3 \quad \text{and} \quad \left|S(f;\beta) - \int_I \beta\right| < \varepsilon/3.$$

By hypothesis, we know

$$\int_I (\beta - \alpha) < \varepsilon/3.$$

Fix $\delta := \min\{\delta_\alpha, \delta_\beta\}$, thereby securing the inequalities above. In particular for $\dot{\mathcal{P}}$ with $\left\|\dot{\mathcal{P}}\right\| < \delta$, these imply:

$$\int_I \alpha - \frac{\varepsilon}{3} < S(\alpha, \dot{\mathcal{P}}) \tag{2}$$

as well as,

$$S(\beta;\dot{\mathcal{P}}) < \int_I \beta + \varepsilon/3. \tag{3}$$

It is clear that $S(\alpha;\dot{\mathcal{P}}) \leq S(f;\dot{\mathcal{P}}) \leq S(\beta;\dot{\mathcal{P}})$, whence by (2)-(3) we find,

$$\int_I \alpha - \frac{\varepsilon}{3} < S(f;\dot{\mathcal{P}}) < \int_I \beta + \frac{\varepsilon}{3}. \tag{4}$$

Furthermore, for any tagged partition $\dot{\mathcal{Q}}$ of I with $\left\|\dot{\mathcal{Q}}\right\| < \delta$ the same holds, i.e:

$$\int_I \alpha - \frac{\varepsilon}{3} < S(f;\dot{\mathcal{Q}}) < \int_I \beta + \frac{\varepsilon}{3} \tag{5}$$

By subtracting equations (4) and (5) we have

$$\left|S(f;\dot{\mathcal{P}}) - S(f;\dot{\mathcal{Q}})\right| < \int_I (\beta - \alpha) + \frac{2\varepsilon}{3} < \varepsilon$$

Ergo, by Theorem 2.4 $f \in \mathcal{R}([a,b])$. □

2.5 Riemann Interability of Important Functions

In this section we shall consider many "elementary" functions, and show that they are Riemann integrable on some interval $I = [a,b]$. Of course, this integral would be more or less useless if our nicer functions were not integrable.

2.5.1 Elementary Functions

Definition 7. Let $J := [c,d] \subseteq [a,b]$ be an interval. We define a function $\chi_J : [a,b] \to \mathbb{R}$ as follows:

$$\chi_J(x) := \begin{cases} 1 & x \in J \\ 0 & x \notin J \end{cases}$$

Lemma 2.6. Let $J = [c,d] \subset I$ and $\chi_J(x)$ on J as in Definition 7, then $\chi_J \in \mathcal{R}([a,b])$ and moreover

$$\int_I \chi_J = (d-c).$$

Proof. Let $\varepsilon > 0$, take $\delta \leq \varepsilon/4 > 0$ such that $c + \delta < d - \delta$.

Let $\dot{\mathcal{P}}$ be a tagged partition of I such that $\left\|\dot{\mathcal{P}}\right\| < \delta \leq \varepsilon/4$. It is not difficult to see that:

$$[c+\delta, d-\delta] \subseteq \bigcup_{\substack{i \\ u_i \in J}} [x_{i-1}, x_i] \subseteq [c-\delta, d+\delta]$$

Whence, by monotonicity,

$$\left|S(\chi_J, \dot{\mathcal{P}}) - (d-c)\right| < 2\delta < \varepsilon.$$

□

Theorem 2.7. Any step function is Riemann Integrable.

Proof. We can express a step function $\psi : [a, b] \to \mathbb{R}$ by breaking it into elementary functions χ_k over sub-intervals J_k. Hence for some constants C_k on each J_k:

$$\begin{aligned}\psi(x) = \sum_{k=1}^{n} C_k \chi_k &\implies \int_I \psi(x) = \int_I \sum_{k=1}^{n} C_k \chi_k \\ &\implies \int_I \psi = \sum_{k=1}^{n} C_k (d_k - c_k)\end{aligned}$$

□

2.5.2 Upgrading To Uniform Continuity

Theorem 2.8 (Uniform Continuity of Continuous Functions)**.** Let $I = [a, b]$ be a closed and bounded interval and $f : I \to \mathbb{R}$ a continuous mapping. Then f is uniformly continuous on $[a, b]$. That is,

$$(\forall \varepsilon > 0)(\exists \delta > 0)(\forall x, y \in I)(|x - y| < \delta \implies |f(x) - f(y)| < \varepsilon)$$

Proof. We argue by contradiction, instead suppose f is not uniformly continuous on I, hence by negation we have

$$\begin{aligned}&\neg(\forall \varepsilon > 0)(\exists \delta > 0)(\forall x, y \in I)(|x - y| < \delta \implies |f(x) - f(y)| < \varepsilon) \\ &\equiv (\exists \varepsilon > 0)(\forall \delta > 0)(\exists x, y \in I)(|x - y| < \delta \implies |f(x) - f(y)| \geq \varepsilon) \qquad (6)\end{aligned}$$

Fix such $\varepsilon > 0$. Taking $\delta = 1/n$ for all $n \in \mathbb{N}$ we construct a sequence $(\delta_n) \subset \mathbb{R}$, moreover for each d_n we can associate a pair $x_n, y_n \in I$ such that $|x_n - y_n| < 1/n$ and $|f(x_n) - f(y_n)| \geq \varepsilon$ for all n. Indeed, (x_n) is a sequence in a closed and bounded interval, by Bolzano-Weierstrass, this sequence has a subsequence (x_{n_k}) that converges to some $x \in I$. We claim (y_{n_k}) converges to the same point. To see this, fix $\varepsilon > 0$, indeed there exists $N_1 \in \mathbb{N}$ such that $n_k \geq N$ implies $|x_{n_k} - x| < \varepsilon/2$. Take $N_2 \in \mathbb{N}$ such that $1/N_2 < \varepsilon/2$. Then, for $N = \max\{N_1, N_2\}$ we have for all $n_k \geq N$:

$$|y_{n_k} - x| \leq |y_{n_k} - x_{n_k}| + |x_{n_k} - x| < \frac{1}{n_k} + \frac{\varepsilon}{2} < \varepsilon$$

Thus, $\lim y_{n_k} = x$. By continuity of f and the sequential criterion for continuity, both $\lim f(x_{n_k}) = f(x) = \lim f(y_{n_k})$. Whence by the same argument given $\varepsilon > 0$ we may choose $N \in \mathbb{N}$ such that for all $n_k \geq N$ we have

$$|f(x_{n_k}) - f(x)| < \frac{\varepsilon}{2} \text{ and } |f(y_{n_k}) - f(x)| < \frac{\varepsilon}{2}$$

But, for $n_k \geq N$ we would then have

$$|f(x_{n_k}) - f(y_{n_k})| \leq |f(x_{n_k}) - f(x)| + |f(y_{n_k}) - f(x)| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

This contradicts the assumption that f is not-uniformly continuous in (6) and hence we conclude f is uniformly continuous. $\square$

2.5.3 Continuous and Monotonic Functions

Theorem 2.9 (Riemann Integrability of Continuous Functions)**.** Let $f : I \longrightarrow \mathbb{R}$ be a continuous function. Then f is Riemann Integrable on I.

Proof. By Theorem 2.8 f is uniformly continuous on $[a, b]$. Hence for all $\varepsilon > 0$ there exists $\delta > 0$ such that for any $x, y \in I$ $|x - y| < \delta \implies |f(x) - f(y)| < \varepsilon/(b - a)$. Fix such $\delta > 0$ and let $\mathcal{P} = \{x_{i-1}, x_i\}_{i=1}^n$ be any partition with $\|\mathcal{P}\| < \delta$. Since f is continuous on $[a, b]$ it must be continuous on each $[x_{i-1}, x_i] \subset [a, b]$. Thus, on each sub-interval it achieves a absolute max and min at some points v_i and u_i respectively. We further define the step functions $\alpha, \beta : [a, b] \longrightarrow \mathbb{R}$ as follows:

$$\alpha(x) := \begin{cases} f(u_i) & \text{on } [x_{i-1}, x_i) \text{ for } 1 \leq i \leq n-1 \\ f(u_n) & \text{on } [x_{n-1}, x_n] \end{cases}$$

$$\beta(x) := \begin{cases} f(v_i) & \text{on } [x_{i-1}, x_i) \text{ for } 1 \leq i \leq n-1 \\ f(v_n) & \text{on } [x_{n-1}, x_n] \end{cases}.$$

By choice, we have $\alpha(x) \leq f(x) \leq \beta(x)$ for all $x \in [a, b]$. Moreover, since f is uniformly continuous on $[a, b]$ and the points are chosen such that $|v_i - u_i| < \delta$ we have:

$$|f(v_i) - f(u_i)| < \frac{\varepsilon}{b - a}$$

Thus,

$$\int_a^b \beta - \alpha = \sum_{i=1}^n (f(v_i) - f(u_i))(x_i - x_{i-1}) < \frac{\varepsilon}{b - a} \sum_{i=1}^n \Delta x_i = \varepsilon$$

Hence f is Riemann integrable on $[a, b]$ by Theorem 2.5. $\square$

Theorem 2.10. Every monotonic function on a closed and bounded interval is Riemann Integrable.

Proof. Let $f : [a, b] \longrightarrow \mathbb{R}$ be monotonic. Without loss of generality we may suppose f is increasing, since $-f$ is decreasing and hence would also be Riemann integrable. Let $\varepsilon > 0$ be given, fix $n \in \mathbb{N}$ be such that

$$\frac{b - a}{n} (f(b) - f(a)) < \varepsilon \tag{7}$$

We will establish Theorem (2.5) to show $f \in \mathcal{R}([a,b])$. Let $\mathcal{P} = \{[x_{i-1}, x_i]\}_{i=1}^n$ be a partition of $[a,b]$ such that $\Delta x_i = (b-a)/n$. Further define

$$A_i := \begin{cases} [x_{i-1}, x_i] & 1 \leq i < n \\ [x_{n-1}, x_n] & i = n \end{cases}$$

On any A_i, since f is increasing for $x \in A_i$ we have that $f(x_{i-1}) \leq f(x) \leq f(x_i)$, thus defining

$$\alpha(x) := \sum_{i=1}^n f(x_{i-1})\chi_{A_i} \text{ and } \beta(x) := \sum_{i=1}^n f(x_i)\chi_{A_i}$$

We define two step functions that are hence Riemann interable satisying:

$$\alpha(x) \leq f(x) \leq \beta(x) \quad \forall x \in [a,b]$$

Especially, recalling our choice of $n \in \mathbb{N}$ in (7)

$$\int_a^b \beta - \alpha = \sum_{i=1}^n (f(x_i), f(x_{i-1})\Delta x_i = \frac{b-a}{n}(f(b) - f(a)) < \varepsilon$$

□

3 The Fundamental Theorem of Calculus

3.1 Form 1

Theorem 3.1 (Fundamental Theorem of Calculus Form 1)**.** Let $F, f : [a,b] \to \mathbb{R}$ and suppose that $F'(x) = f(x)$ for all $x \in (a,b)$. Suppose further that $f \in \mathcal{R}([a,b])$. Then,

$$\int_{[a,b]} f = F(b) - F(a). \tag{8}$$

Proof. Let $\varepsilon > 0$ be given, since f is Riemann integrable on $[a,b]$ there exists $\delta > 0$ such that for any tagged partition of $[a,b]$ given by $\dot{\mathcal{P}}$ with $\left\|\dot{\mathcal{P}}\right\| < \delta$ implies

$$\left| \sum_{i=1}^n f(x_i^\star)(x_i - x_{i-1}) - \int_{[a,b]} f \right| < \varepsilon. \tag{9}$$

Now, F is continuous on $[a,b]$ and is thus continuous on any sub-interval of $[a,b]$. Moreover, since $F(x)$ is differentiable on $[a,b] \setminus \{a,b\} = (a,b)$ it is differentiable (and hence continuous) on any open sub-interval of $\dot{\mathcal{P}}$, so, for $1 \le i \le$ by the Mean Value Theorem there exists $z_i \in (x_{i-1}, x_i)$ such that

$$F'(z_i) = f(z_i) = \frac{F(x_i) - F(x_{i-1})}{x_i - x_{i-1}} \tag{10}$$

Note that for our definition of the Riemann Integral it does not matter which tag we choose in the sub-intervals, and since $z_i \in [x_{i-1}, x_i]$ for any $1 \le i \le n$ each z_i is a perfectly valid tag. Moreover, these tags preserve the norm of the tagged partition and hence preserve inequality (10). Now, by choosing each $x_i^\star = z_i$ we see that

$$\begin{aligned}
\left| \sum_{i=1}^{n} f(x_i^\star)(x_i - x_{i-1}) - \int_{[a,b]} f \right| &= \left| \sum_{i=1}^{n} f(z_i)(x_i - x_{i-1}) - \int_{[a,b]} f \right| \\
&= \left| \sum_{i=1}^{n} \frac{F(x_i) - F(x_{i-1})}{x_i - x_{i-1}}(x_i - x_{i-1}) - \int_{[a,b]} f \right| \\
&= \left| \sum_{i=1}^{n} [F(x_i) - F(x_{i-1})] - \int_{[a,b]} f \right| \\
&= \left| F(x_n) - F(0) - \int_{[a,b]} f \right| \\
&= \left| F(b) - F(a) - \int_{[a,b]} f \right|.
\end{aligned}$$

However, by (9) we find for $\varepsilon > 0$ and the choice of $\delta > 0$ that $\left\| \dot{\mathcal{P}} \right\| < \delta$ implies

$$\left| \sum_{i=1}^{n} f(x_i^\star)(x_i - x_{i-1}) - \int_{[a,b]} f \right| < \varepsilon.$$

By equality, it follows

$$\left| F(b) - F(a) - \int_{[a,b]} f \right| < \varepsilon.$$

Hence for all $\varepsilon > 0$ one has $\left| F(b) - F(a) - \int_{[a,b]} f \right| < \varepsilon$ which implies

$$\left| F(b) - F(a) - \int_{[a,b]} f \right| = 0$$

Ergo,

$$\boxed{F(b) - F(a) = \int_{[a,b]} f}$$

□

Theorem 3.2. Let $f : [a, b] \to \mathbb{R}$ with $E \subset [a, b]$ a finite set $\{c_0, ..., c_n\}$. If

- F is continuous on $[a, b]$.
- $F'(x) = f(x)$ for all $x \in [a, b] \setminus E$.
- $f \in \mathcal{R}([a, b])$.

then

$$\int_a^b f = F(b) - F(a)$$

Proof. We break $[a, b]$ into subintervals $[c_0, c_1], [c_1, c_2], ..., [c_{n-1}, c_n]$. Indeed, on each $[c_{k-1}, c_k]$ the hypothesis of the theorem are satisfied and the previous theorem applies. That is,

$$\int_{c_{k-1}}^{c_k} f = F(c_k) - F(c_{k-1})$$

But,

$$\int_a^b f = \sum_{k=1}^{n} \int_{c_{k-1}}^{c_k} f = \sum_{k=1}^{n} (F(c_k) - F(C_{k-1})) = F(b) - F(a)$$

□

3.2 Form 2

Definition 8 (Indefinite Integral). If $f \in \mathcal{R}([a, b])$ then the function defined by

$$F(x) := \int_a^x f \quad x \in [a, b] \tag{11}$$

is called the indefinite integral of f with basepoint a.

Theorem 3.3. The indefinite integral F defined by (8) is continuous on $[a, b]$. Moreover, if f is bounded above, then F is Lipschitz on $[a, b]$.

Proof. Let $z, w \in [a, b]$ with $w \leq z$ By additivity,

$$F(z) = \int_a^z f = \int_a^w f + \int_w^z f = F(w) + \int_w^z f$$

Or,

$$F(z) - F(w) = \int_w^z f \tag{12}$$

Now if $|f(x)| \leq M \iff -M \leq f(x) \leq M$ on $[a, b]$ by monotonicity and equality,

$$-M(z - w) \leq \int_w^z f \leq M(z - w)$$

$$\implies -M(z - w) \leq F(z) - F(w) \leq M(z - w) \iff |F(z) - F(w)| \leq M\, |z - w|$$

□

Theorem 3.4 (Fundamental Theorem of Calculus Form 2)**.** Let $f \in \mathcal{R}([a,b])$ and let f be continuous at $c \in [a,b]$. Then the indefinite integral is differentiable at c and $F'(c) = f(c)$.

Proof. We first consider the case where $c \in [a,b)$. Let $\varepsilon > 0$ be given, since f is continuous at c, there exists $\delta > 0$ such that $x \in V_\delta(c) \cap [a,b]$ implies

$$f(c) - \varepsilon < f(x) < f(c) + \varepsilon \tag{13}$$

We may choose $\delta > 0$ small enough such that $[c, c+\delta] \subseteq [a,b]$. Obviously, this inequality holds for $x \in [c, c+\delta]$. Let $h > 0$ be such that $[c, c+h] \subseteq [c, c+\delta]$. Then for $x \in [c, c+h]$ inequality (13) holds. Moreover, since f is integrable on $[a,b]$ it is integrable on $[c, c+h] \subseteq [c, c+\delta]$, thus

$$F(c+h) - F(c) = \int_c^{c+h} f$$

However, on $[c, c+h]$ we have by inequality (13):

$$(f(c) - \varepsilon)h \leq \int_c^{c+h} f \leq (f(c) + \varepsilon)h$$

But by Theorem (3.1) we must have:

$$(f(c) - \varepsilon)h \leq F(c+h) - F(c) \leq (f(c) + \varepsilon)h$$

Whence, since $h > 0$

$$f(c) - \varepsilon \leq \frac{F(c+h) - F(c)}{h} \leq f(c) + \varepsilon$$

Thus where $h > 0$ is small,

$$\left| \frac{F(c+h) - F(c)}{h} - f(c) \right| \leq \varepsilon$$

Since $\varepsilon > 0$ is arbitrary we conclude that:

$$\lim_{h \to 0+} \frac{F(c+h) - F(c)}{h} = f(c) \tag{14}$$

It remains to show the left handed limit for $c \in (a,b]$. We proceed by a similar argument. Since f is continuous at c there exists $\eta > 0$ such that for all $x \in [a,b] \cap V_\eta(c)$ one has inquality (13). Let $h < 0$ be such that both

1. $|h| < \eta$

2. $[c+h, c] \subseteq [a,b]$

In this case, $x \in [c+h, c]$ implies $x \in [a,b]$ and $|x-c| \leq |h| < \eta$ and thus that inequality (13) holds. Since $[c+h, c] \subseteq [a,b]$ it is Riemann integrable on $[c+h, c]$ by the Additivity theorem and hence

$$\int_{c+h}^{c} f = F(c) - F(c+h)$$

By monotonicity of integration together with inequality (13) it is easy to see that

$$(f(c) - \varepsilon)(-h) \leq \int_{c+h}^{c} f \leq (f(c) + \varepsilon)(-h)$$

Whence by the previous two inequalities we infer

$$\begin{aligned}
(f(c) - \varepsilon)(-h) \leq F(c) - F(c+h) \leq (f(c) + \varepsilon)(-h) \\
\implies f(c) - \varepsilon \leq \frac{F(c+h) - F(c)}{h} \leq f(c) + \varepsilon \\
\implies \left| \frac{F(c+h) - F(c)}{h} - f(c) \right| \leq \varepsilon
\end{aligned}$$

Since $h < 0$ and $\varepsilon > 0$ is arbitrary we conclude that

$$\lim_{h \to 0-} \frac{F(c+h) - F(c)}{h} = f(c) \tag{15}$$

Thus, for $c \in \{a, b\}$ the respective single sided derivatives exist, and for $c \in (a,b)$ by (14) and (15) both sided limits exist and are equal to $f(c)$. We then infer that $F'(c) = f(c)$ wherever f is continuous. □

4 Advanced Topics

4.1 Null Sets

Definition 9. Let $\Omega \subseteq \mathbb{R}$. We call Ω a null set if for any $\varepsilon > 0$ there exists countably many open intervals $\{(a_k, b_k)\}_{k=1}^{\infty}$ such that

$$\Omega \subseteq \bigcup_{k \in \mathbb{N}} (a_k, b_k)$$

and

$$\sum_{k=0}^{\infty} (b_k - a_k) \leq \varepsilon$$

Alternatively,

$$(\forall \varepsilon > 0)\, (\exists \{\omega_k\}_{k \in \mathbb{N}}))\left(\Omega \subseteq \bigcup_{k \in \mathbb{N}} \omega_n \;\wedge\; \sum_{k \in \mathbb{N}} |\omega_k| \leq \varepsilon \right)$$

Remark: Some authors will say a set Ω is null if it has measure 0, or $\ell(\Omega) = 0$.

Theorem 4.1 (Null Properties of Countable Sets)**.** Any countable subset of $\mathbb{R}$ is null.

Proof. Let $\Lambda \subset \mathbb{R}$ be a countable set. We can thus enumerate $\Lambda = \{\lambda_1, ..., \lambda_n, ...\}$. Let $\varepsilon > 0$ be given, for any $k \in \mathbb{N}$ we consider the open interval

$$\Delta_k := \left(\lambda_k - \frac{\varepsilon}{2^{k+1}}, \lambda_k + \frac{\varepsilon}{2^{k+1}}\right) \ni \lambda_k$$

Obviously each Δ_k contains λ_k and a simple calculation reveals that for each $k \in \mathbb{N}$ the length $|\cdot|$ of Δ_k corresponds to

$$|\Delta_k| = \frac{\varepsilon}{2^k}$$

Indeed, since $\lambda_k \in \Delta_k$ for each $k \in \mathbb{N}$ and Λ is countable by hypothesis it follows:

$$\Lambda \subseteq \bigcup_{k \in \mathbb{N}} \Delta_k$$

Moreover,

$$\sum_{k \in \mathbb{N}} |\Delta_k| = \sum_{k=1}^{\infty} \frac{\varepsilon}{2^k} = \varepsilon$$

Ergo, Λ is a null set. □

Corollary 4.1.1. Let $f : [a, b] \to \mathbb{R}$ be monotonic and let $\mathcal{D}$ be the set of all points in $[a, b]$ at which f is discontinuous. $\mathcal{D}$ is null.

Proof. By Theorem 1.2 $\mathcal{D}$ is countable. By Theorem 4.1 $\mathcal{D}$ is null. □

4.2 Riemann's Criterion for Riemann Integrability

Theorem 4.2 (Riemann's Theorem)**.** Let $f : [a, b] \to \mathbb{R}$ be a bounded function. The following are equivalent.

1. f is Riemann integrable on $[a, b]$.

2. For all $\varepsilon > 0$ there exists a partition $\mathcal{W} = \{I_i\}_{i=1}^{n}$ of $[a, b]$ such that for any tagged partitions $\dot{\mathcal{Q}}, \dot{\mathcal{P}}$ with the same sub-intervals as $\mathcal{W}$ one has

$$\left| S(f; \dot{\mathcal{P}}) - S(f; \dot{\mathcal{Q}}) \right| < \varepsilon$$

3. For all $\varepsilon > 0$ there exists $\mathcal{P} = \{[x_{i-1}, x_i]\}_{i=1}^{n}$ with

$$m_i = \inf\{f(x) \mid x \in [x_{i-1}, x_i]\} \text{ and } M_i = \sup\{f(x) \mid x \in [x_{i-1}, x_i]\}$$

one has

$$\sum_{i=1}^{n} (M_i - m_i) \Delta x_i < \varepsilon$$

Proof. (1 $\implies$ 2). Let $\varepsilon > 0$, by Theorem 2.4 there exists $\delta > 0$ such that for any two tagged partition $\dot{\mathcal{Q}}, \dot{\mathcal{P}}$ with $\left\|\dot{\mathcal{Q}}\right\| < \delta$, $\left\|\dot{\mathcal{P}}\right\| < \delta$ implies

$$\left|S(f;\dot{\mathcal{P}}) - S(f;\dot{\mathcal{Q}})\right| < \varepsilon$$

Let $\mathcal{W}$ be any partition of $[a,b]$ with $\|W\| < \delta$. Then, for any tagged partition $\dot{\mathcal{P}}, \dot{\mathcal{Q}}$ with the same endpoints as $\mathcal{W}$ we must have $\left\|\dot{\mathcal{Q}}\right\| < \delta$ and $\left\|\dot{\mathcal{P}}\right\| < \delta$ whence by Cauchy's First Criterion it follows

$$\left|S(f;\dot{\mathcal{P}}) - S(f;\dot{\mathcal{Q}})\right| < \varepsilon$$

(2 $\implies$ 3). Let $\varepsilon > 0$ be given, by our hypothesis in (2) we can select a partition $\mathcal{W}$ of $[a,b]$ such that for any two tagged partitions $\dot{\mathcal{P}}, \dot{\mathcal{Q}}$ with the same end points as $\mathcal{W} = \{[x_{i-1}, x_i]\}_{i=1}^n$ one has

$$\left|S(f;\dot{\mathcal{P}}) - S(f;\dot{\mathcal{Q}})\right| < \frac{\varepsilon}{2} \tag{16}$$

Indeed, f is bounded on $[a,b]$ and hence on any $[x_{i-1}, x_i]$, thus f achieves infimum and supremum on each sub-interval. We may thus set m_i and M_i as in (3). Thus, for a given $\varepsilon > 0$ and i we can find $u_i, v_i \in [x_{i-1}, x_i]$ satisfying

$$f(u_i) < m_i + \frac{\varepsilon}{4(b-a)} \text{ and } f(v_i) > M_i - \frac{\varepsilon}{4(b-a)}$$

$$\iff -f(u_i) > -m_i - \frac{\varepsilon}{4(b-a)} \text{ and } f(v_i) > M_i - \frac{\varepsilon}{4(b-a)}$$

Whence we infer for $1 \leq i \leq n$:

$$f(v_i) - f(u_i) + \frac{\varepsilon}{2(b-a)} > (M_i - m_i)$$

$$\implies \left(f(v_i) - f(u_i) + \frac{\varepsilon}{2(b-a)}\right)(x_i - x_{i-1}) > (M_i - m_i)(x_i - x_{i-1})$$

Since i is arbitrary within $[1,n] \cap \mathbb{N}$ we may sum over such i preserving the inequality:

$$\sum_{i=1}^n (M_i - m_i)(x_i - x_{i-1}) < \sum_{i=1}^n \left(f(v_i) - f(u_i) + \frac{\varepsilon}{2(b-a)}\right)(x_i - x_{i-1})$$

$$= \sum_{i=1}^n \left(f(v_i) - f(u_i)\right)(x_i - x_{i-1}) + \frac{\varepsilon}{2} = S(f;\dot{\mathcal{P}}) - S(f;\dot{\mathcal{Q}}) + \frac{\varepsilon}{2}$$

Where $\dot{\mathcal{P}} = \{[x_{i-1}, x_i], v_i\}$ and $\dot{\mathcal{Q}} = \{[x_{i-1}, x_i], u_i\}$. However, both of these are tagged partitions with the same endpoints as $\mathcal{W}$, thus by hypothesis

$$\sum_{i=1}^n (M_i - m_i)(x_i - x_{i-1}) < S(f;\dot{\mathcal{P}}) - S(f;\dot{\mathcal{Q}}) + \frac{\varepsilon}{2} < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

$(3 \implies 1)$. Fix $\varepsilon > 0$ and let $\mathcal{P} = \{[x_{i-1}, x_i]\}$ be a partition satisfying our hypothesis in (3). Let $A_i := [x_{i-1}, x_i)$ for $1 \leq i \leq n-1$ and $A_n := [x_{n_1}, x_n]$. On each A_i set $\alpha := m_i$ and $\omega := M_i$. Hence, on $[a, b]$ we have that $\alpha(x) \leq f(x) \leq \omega(x)$. Moreover, since α, ω are step functions we can express them as follows:

$$\alpha(x) = \sum_{i=1}^{n} m_i \chi_{A_i} \quad \wedge \quad \omega(x) = \sum_{i=1}^{n} M_i \chi_{A_i}$$

Indeed, these step function are Riemann Integrable and by previous results

$$\int_a^b \alpha = \sum_{i=1}^{n} m_i (x_i - x_{i-1})$$

$$\int_a^b \omega = \sum_{i=1}^{n} M_i (x_i - x_{i-1})$$

Moreover, by hypothesis in (3)

$$\int_a^b \omega - \alpha = \sum_{i=1}^{n} (M_i - m_i)(x_i - x_{i-1}) < \varepsilon$$

Thus, by Theorem (2.5) $f \in \mathcal{R}([a, b])$. □

4.3 Lebesgue's Criterion for Riemann Integrability

Definition 10. Let $A \subseteq \mathbb{R}$ and $f : A \rightarrowtail \mathbb{R}$ a bounded function. Given $S \subseteq A$ we define the **oscillation of f over S** as

$$W(f; S) := \sup_{x,y \in S} |f(x) - f(y)|$$

We consider the following propositions, which are of great importance in the proof of Lebesgue's Criterion for Riemann Integrability:

Proposition 4.3. In the context of the previous definition

$$W(f; S) \leq 2 \sup\{|f(x)| \ : \ x \in S\}$$

Proposition 4.4.

$$W(f; S) = \sup_{x \in S} |f(x)| - \inf_{x \in S} |f(x)|$$

Proposition 4.5. If $S \subseteq T \subseteq A$ we have

$$W(f; S) \leq W(f; T)$$

Lemma 4.6. f is continuous at c if and only if $w(f;c) = 0$.

Proof. Suppose f is continuous at c, and let $\varepsilon > 0$, then there exists $\delta > 0$ such that $x \in A$ with $x \in V_\delta(c)$ implies

$$|f(x) - f(c)| < \varepsilon/2$$

Now take $x, y \in V_\delta(c)$, thus $|f(x) - f(y)| < \varepsilon$. Thus, $W(f; V_\delta(c)) = \sup\{|f(x) - f(y)| \mid x, y \in V_\delta(c)\}$ implies

$$W(f; V_\delta(c)) \leq \varepsilon$$

But since $w(f;c) = \inf\{W(f; V_r(c)) \mid r > 0\}$ we have that $w(f;c) \leq \varepsilon$. But $\varepsilon > 0$ is arbitary.

Suppose now $w(f;c) = \inf\{W(f; V_r(c)) : r > 0\} = 0$. Then, by properties of the infimum for $\varepsilon > 0$ there exists $r > 0$ such that

$$0 < W(f; V_r(c)) < 0 + \varepsilon$$

Thus, $W(f; V_r(c)) = \sup\{|f(x) - f(y)| \mid x, y \in V_r(c)\} < \varepsilon$ Fix $y = c$, then we have for all $x \in V_r(c)$ that

$$|f(x) - f(c)| < \varepsilon$$

□

Theorem 4.7. Let $\delta(x)$ be a gauge on $[a, b]$. Then there exists a δ-fine partition of $[a, b]$.

Proof. Let E be the collection of all points in $[a, b]$ such that there exists a δ-fine partition of $[a, x]$. We note that $E \neq \varnothing$, since for any point $x \in [a, b]$ with $a \leq a \leq a + \delta(a)$ we can define $\dot{\mathcal{P}} := \{[a, x], a\}$ as our tagged partition. Moreover, E is bounded above and below. Thus, $u = \sup\{E\}$ exists in $\mathbb{R}$. We will show two things. That is, $u \in E$ and $u = b$. Let us do this

1. We claim $u \in E$. Suppose not, and consider $u - \delta(u)$, since $\delta(u) > 0$ there exists $v \in E$ such that
$$u - \delta(u) < v < u$$
Since $v \in E$ there exists a δ-fine partition of $[a, v]$. We denote this partition by $\dot{\mathcal{P}}$ and consider the partition given by $\dot{\mathcal{P}} \cup \{[v, u], u\}$. Whence, $[v, u] \subseteq [u - \delta(u), u + \delta(u)]$. This shows we have a parition of $[a, u]$ that is δ-fine by choice of v and thus we have $u \in E$. Contradiction.

2. We claim that $u = b$. Suppose not then $u < b$. Considering $u + \delta(u)$ we chose $w \in [a, b]$ such that
$$u < w < u + \delta(u)$$
Now, $u \in E$ so there exists a δ fine partition of $[a, u]$ denoted $\dot{\mathcal{P}}$. Whereby
$$\dot{\mathcal{P}} \cup \{[u, w], u\}, \implies [u, w] \subseteq [u - \delta(u), u + \delta(u)]$$
Thus $w \in E$. But, then we must have $w \leq u$. Contradiction.

$\square$

Theorem 4.8. Let $f : [a, b] \to \mathbb{R}$ be a bounded function and Z be the set of all discontinuity points of f. TFAE

1. $f \in \mathcal{R}([a, b])$
2. Z is a null set.

Proof. (1 implies 2). By our previous lemma we can express $Z = \{x \in [a, b] \mid w(f; x) > 0\}$. Define further

$$Z = \bigcup_{n=1}^{\infty} H_k \quad H_k := \{x \in [a, b] \mid w(f; x) \geq 1/2^k\}$$

Fix $\varepsilon > 0$, by the Riemann Criterion there exists a parition $\mathcal{P} = \{[x_{i-1}, x_i]\}_{i=1}^{n}$ such that

$$\sum_{i=1}^{n} (M_i - m_i)(x_i - x_{i-1}) < \frac{\varepsilon}{4^k}$$

If $H_k \cap (x_{i-1}, x_i) \neq \varnothing$ choose x in this intersection. Then there exists $r > 0$ such that $(x_{i-1}, x_i) \supseteq V_r(x)$. Since $x \in H_k$ it follows that

$$\frac{1}{2^k} \leq w(f; x) \leq W(f; V_r(x)) \leq W(f; [a, b]) = M_i - m_i$$

Thus,

$$\frac{1}{2^k} \leq M_i - m_i$$

Set $S = \{i \mid H_k \cap (x_{i-1}, x_i) \neq \varnothing\}$. Then,

$$\frac{1}{2^k} \sum_{i \in S} (x_i - x_{i-1}) \leq \sum_{i \in S} (M_i - m_i)(x_i - x_{i-1}) \leq \sum_{i=1}^{n} (M_i - m_i)(x_i - x_{i-1}) < \frac{\varepsilon}{4^k}$$

Whence we have

$$\sum_{i \in S} (x_i - x_{i-1}) < \frac{\varepsilon}{2^k}$$

Now,

$$H_k \subseteq \bigcup_{i \in S} ((x_{i-1}, x_i) \cup \{x_{i-1}, x_i\})$$

$$\subseteq \bigcup_{i \in S} ((x_{i-1}, x_i)) \cup \bigcup_{j=1}^{n} (x_j - \frac{\varepsilon}{4(n+1)}, x_j + \frac{\varepsilon}{4(n+1)})$$

$$= A + B$$

Now the length of A is less than $\varepsilon/2$ as shown above. Similarly, the length of $B < \varepsilon/2$ as well. Thus Z is null.

(2 implies 1). Let Z be null, take

$$\varepsilon' = \frac{\varepsilon}{b-a+1}, \ \wedge |f(x)| \leq M$$

Since Z is a null set we define a countable collection of disjoint open intervals $\{(a_n, b_n)\}_{n\in\mathbb{N}} = \{J_n\}_{n\in\mathbb{N}}$ such that

$$Z \subseteq \bigcup k \in \mathbb{N} J_k \ \ \sum_{k\in\mathbb{N}} \ell(J_k) < \frac{\varepsilon'}{2M}$$

We now construct a gauge δ on $[a, b]$ as follows,

1. If $x \notin Z$ we know f is continuous at x. Thus, there exists some $\delta(x) > 0$ such that for all $y \in V_{\delta(x)}(x)$ one has $|f(x) - f(y)| < \varepsilon'/2$. Thus, for

$$M_x := \sup_{y\in V_{\delta(x)(x)}} \{|f(y)|\} \ \ m_x := \inf\{y \in V_{\delta(x)(x)}\}\{|f(y)|\}$$

   We have $M_x - m_x \leq \varepsilon'$ since $|f(y') - f(y)| < \varepsilon'$ for $y, y' \in V_{\delta(x)}(x)$.

2. If $x \in Z$ then since $\{J_k\}_{k=1}^{\infty}$ covers Z and are disjoint $x \in J_k$ for a unique $k \in \mathbb{N}$. But J_k is open, thus there exists $\delta(x) > 0$ such that $(x - \delta(x), x + \delta(x)) \subseteq J_k$

Taking such values of $\delta(x) > 0$ constructs a gauge on $[a, b]$. Especially, $\delta(x)/2$ is also a gauge on $[a, b]$. We take this as our gauge.

Let $\dot{\mathcal{P}}\{[x-i-1, x_i]u_i\}_{i=1}^{n}$ be δ-fine. Then

$$[x_{i-1}, x_i] \subseteq \left[u_i - \frac{\delta(u_i)}{2}, u_i + \frac{\delta(u_i)}{2}\right] \bigcap [a, b] = V_{\delta(u_i)}(u_i)$$

Let $\Lambda := \{i \mid u_i \text{ is a tag of } \dot{\mathcal{P}}\}$. We may write this set as two disjoint sets:

$$S = \{i \mid u_i \notin Z\} \ \ S' = \{i \mid u_i \in Z\}$$

Now we distinguish two cases:

1. $i \in S$. Then $[x_{i-1}, x_i] \subseteq V_{\delta(u_i)}(u_i)$. So, for M_i, m_i as in (3) of the Riemann Criterion we have $M_i - m_i \leq M_{t_i} - m_{t_i} \leq \varepsilon'$

2. $i \in S'$. Then $[x_{i-1}, x_i] \subseteq (t_i - \delta(t_i), t_i + \delta(t_i) \cap [a, b] = V_{\delta(u_i)}(u_i) \subseteq J_k$ for a unique k. This implies $(x_i - x_{i-1}) \leq \ell(J_k)$.

We are now ready to prove $f \in \mathcal{R}([a,b])$. Note that

$$\sum_{i=1}^{n}(M_i - m_i)\Delta x_i = \sum_{i \in S}(M_i - m_i)\Delta x_i + \sum_{i \in S'}(M_i - m_i)\Delta x_i$$

$$\leq \sum_{i \in S}\varepsilon' \Delta x_i + 2M\sum_{i \in S'}\Delta x_i \leq \varepsilon'(b-a) + 2M\sum_{k=1}^{\infty}\ell(J_k)$$

$$< \varepsilon'(b-a) + 2M\left(\frac{\varepsilon'}{2M}\right) = \varepsilon$$

□

Proposition 4.9. Let $f : [a,b] \to \mathbb{R}$ be Riemann integrable and $\phi : [c,d] \to \mathbb{R}$ be continuous. Suppose further that $f([a,b]) \subseteq [c,d]$. Then $\phi \circ f$ is Riemann integrable on $[a,b]$.

Proof. Let $Z := \{x \in [a,b] \mid f \text{ is discontinuous}\}$ and $Z' := \{y \in [c,d] \mid \phi \circ f \text{ is discontinuous}\}$. We claim $Z \supseteq Z'$, and we argue as follows. Let $x \notin Z$, i.e $x \in Z^c$. Then f is continuous at x and by continuity of ϕ on $[c,d] \supseteq f([a,b])$ the composition $\phi \circ f$ is continuous at c. Thus, $x \in Z'$. Symbolically,

$$Z^c \subseteq (Z')^c \iff Z' \subseteq Z$$

Since Z is null, we conclude Z' is null. By Lebesgue's criterion we conclude $\phi \circ f$ is Riemann integrable. □

References

1. Bartle, Robert Gardner, and Donald R. Sherbert. Introduction to Real Analysis. Hoboken, NJ: Wiley, 2011. Print.