Module 3 Notes for Math 104: Real Analysis

Sid Menon

Spring 2022

Contents

Chapter 1

Differentiation

1.1 Differentiation and the Mean Value Theorem

Let $f : [a, b] \to \mathbb{R}$ be a real valued function. Define $\forall x \in [a, b]$:

$$f'(x) = \lim_{t \to x} \left(\frac{f(t) - f(x)}{t - x} \right)$$

This limit may not exist for all points, and if it does, we say f is **differentiable** at x. We may formalize this notion by defining a function g_x, the **difference quotient** of f at x:

$$g_x(t) = \frac{f(t) - f(x)}{t - x}$$

Taking the limit of this function as t approaches x yields the derivative $f'(x)$.

Proposition 1

If $f : [a, b] \to \mathbb{R}$ is differentiable at $x_0 \in [a, b]$, then f is continuous at x_0.

Proof. We want to show that $\lim_{x \to x_0} f(x) - f(x_0) = 0$. We have the following:

$$f(x) - f(x_0) = \frac{f(x) - f(x_0)}{x - x_0} \cdot (x - x_0)$$

Taking limits, we see:

$$\lim_{x - x_0} f(x) - f(x_0) = \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} \cdot \lim_{x \to x_0} (x - x_0) = f'(0) \cdot 0 = 0$$

□

Remark: If $f(x)$ is differentiable at x_0, we don't necessarily maintain that f is continuous at points close to x_0. Consider:

$$f(x) = \begin{cases} x^2 & x \in \mathbb{Q} \\ -x^2 & x \in \mathbb{R} \setminus \mathbb{Q} \end{cases}$$

which is differentiable at 0 but is not continuous at any other point in $\mathbb{R}$.

Example 1: Given

$$f(x) = \begin{cases} x\sin\left(\frac{1}{x}\right) & x > 0 \\ 0 & x \leq 0 \end{cases}$$

Does $f'(0)$ exist? Let's construct left and right limits. If $x \geq 0$, then:

$$g_0(x) = \frac{f(x) - f(0)}{x} = \sin\left(\frac{1}{x}\right)$$

Observe that $\lim_{t \to 0} g_0(t)$ does not exist, and therefore f is not differentiable at 0.

Note that it's possible for a function to be differentiable on all $x \in \mathbb{R}$, but still be discontinuous at some points. An example of such a function is

$$f(x) = \begin{cases} x\sin\left(\frac{1}{x}\right) & x > 0 \\ 0 & x \leq 0 \end{cases}$$

After writing out the derivative, it's pretty simple to check that $f'(x)$ does not converge to $f'(0)$ as $x \to 0^+$.

Theorem 1: Properties of Derivatives

Let $f, g : [a, b] \to \mathbb{R}$ given that f, g are differentiable at $x_0 \in [a, b]$, then:

1. $\forall c \in \mathbb{R} \qquad (cf)'(x_0) = c(f'(0)$
2. $(f + g)'(x_0) = f'(x_0) + g'(x_0)$
3. $(fg)'(x_0) = f'(x_0)g(x_0) + f(x_0)g'(x_0)$
4. If $g(x_0) \neq 0$, then $\left(\frac{f}{g}\right)'(x_0) = \frac{f'(x_0)g(x_0) - f(x_0)g'(x_0)}{g(x_0)^2}$

Proof. Proofs 1,2 are simple to check using the definition of the derivative. 4 is simple given 3 and the chain rule (Theorem 2). We want to compute

$$\lim_{x \to x_0} \frac{(fg)(x) - (fg)(x_0)}{x - x_0}$$

We have that:

$$\begin{aligned} f(x)g(x) - f(x_0)g(x_0) &= [f(x) - f(x_0) + f(x_0)]\,[g(x) - g(x_0) + g(x_0)] \\ &= [f(x) - f(x_0)]\,[g(x) - g(x_0)] + [f(x) - f(x_0)]\,g(x) + f(x_0)\,[g(x) - g(x_0)] \end{aligned}$$

When we divide this entire expression by $x - x_0$, the first term goes to 0, since we multiply $f'(x_0)$ by 0. The second term goes to $f'(x_0)g(x_0)$ after factoring out $g(x_0)$, and the last term goes to $f(x_0)g'(x_0)$ after factoring out $f(x_0)$. We conclude that

$$\lim_{x \to x_0} \frac{(fg)(x) - (fg)(x_0)}{x - x_0} = f'(x_0)g(x_0) + f(x_0)g'(x_0)$$

□

Theorem 2: Chain Rule of Derivatives

Suppose $f : [a, b] \to \mathbb{R}$, and $g : I \to \mathbb{R}$, where $I \subseteq \mathbb{R}$. Suppose for some $x_0 \in [a, b]$, such that $f(x_0) = y_0$, and further suppose that $f([a, b]) \subseteq I$. Suppose that $f'(x_0), g'(y_0)$ exist. Then the composition $h = g \circ f : [a, b] \to \mathbb{R}$ is differentiable at x_0, and $h'(x_0) = g'(y_0)f'(x_0)$.

Proof. Since we have that f is differentiable at x_0, and g is differentiable at y_0, we have that:

$$f(x) - f(x_0) = (x - x_0)(f'(x_0) + U(x)) \tag{1.1}$$

$$g(y) - g(y_0) = (y - y_0)(g'(y_0) + V(y)) \tag{1.2}$$

Such that $\lim_{x \to x_0} U(x) = \lim_{y \to y_0} V(y) = 0$ We can write the following:

$$\begin{aligned} h(x) - h(x_0) &= g(f(x)) - g(f(x_0)) \\ &= (f(x) - f(x_0))(g'(f(x_0) + V(f(x))) \\ &= (x - x_0)(f'(x_0) + U(x))(g'(f(x_0)) + V(f(x))) \end{aligned}$$

We find that

$$\begin{aligned} \lim_{x \to x_0)} \frac{h(x) - h(x_0)}{x - x_0} &= \lim_{x \to x_0} (f'(x_0) + U(x))(g'(f(x_0)) + V(f(x))) \\ &= f'(x_0)g'(f(x_0)) = g'(y_0)f'(x_0) \end{aligned}$$

□

Let $f : [a, b] \to \mathbb{R}$. We say f has a **local maximum** at $p \in [a, b]$ if $\exists \epsilon > 0$ such that $\forall x \in [a, b] \cap B_\epsilon(p)$, we have $f(x) \leq f(p)$. Define a **local minimum** of f analogously.

Proposition 2

Let $f : [a, b] \to \mathbb{R}$. If f has a local max at $p \in (a, b)$, and if $f'(p)$ exists, then $f'(p) = 0$.

Before proving this, note that we only consider the open interval (a, b) because a function with maxima at it's endpoints may not have $f'(x) = 0$. Likewise, it's possible to have a function with a maximum at p but not be differentiable at p (sharp edges and stuff).

Proof. Since $p \in (a, b)$, $\exists \delta > 0$ such that $(p - \delta, p + \delta) \subseteq [a, b]$, and $f(p)$ is the max of the restriction $f\big|_{(p-\delta,p+\delta)}$. Consider difference quotient $g_p(x)$ for all $x \in [a, b] \setminus \{P\}$. The following two statements must hold

If $x \in (p - \delta, \delta)$, $g_p(x) \geq 0$ (numerator and denominator are negative), and therefore $lim_{x \to p^-} g_p(x) \geq 0$.

If $x \in (p, p + \delta)$, $g_p(x) \leq 0$ (negative numerator, positive denominator), and therefore $lim_{x \to p^+} g_p(x) \leq 0$.

Since f is differentiable at p, the limits must be equal, meaning that $lim_{x \to p} g_p(x) = 0$.

1.2 Generalized Mean Value Theorem and L'Hopital's Rule

□

Proposition 3: Rolle's Theorem

Let $f : [a, b] \to \mathbb{R}$ be a continuous function, that is differentiable on (a, b). If $f(a) = f(b)$, then $\exists c \in (a, b)$ such that $f'(c) = 0$

Proof. If f is a constant function, then $f'(c) = 0$ for any $c \in (a, b)$ (by definition of the derivative). If f is nonconstant, then there is $t \in (a, b)$ such that $f(a) = f(b) \neq f(t)$. WLOG, suppose $f(t) > f(a)$. Let $x_0 \in (a, b)$ be such that $f(x_0)$ is the maximum value of $f([a, b])$, which exists since f is surjective onto it's image, and it's continuous.

Then $f(x_0) \geq f(t) > f(a)$, which means $x_0 \neq a \neq b$. By the earlier proposition $x_0 \in (a, b)$ is a local maximum and therefore has derivative 0, so take $c = x_0$. The proof works the same way for local mins (in which case replace $>$ with $<$, replace $\geq$ with $\leq$, etc). □

Theorem 3: Generalized Mean Value Theorem

Let $f, g : [a, b] \to \mathbb{R}$ be a continuous function that is differentiable on (a, b). Then $\exists c \in (a, b)$ such that:

$$[f(a) - f(b)]g'(c) = [g(a) - g(b)]f'(c)$$
$$\iff [f(b) - f(a)]g'(c) = [g(b) - g(a)]f'(c)$$

Proof. Let $h(x) := [f(a) - f(b)][g(x) - g(a)] - [f(x) - f(a)][g(a) - g(b)]$. Note where we use input x. We can observe that $h(a) = 0$ obviously, and $h(b) = 0$ since we're subtracting 2 equal things.Therefore, we can apply Rolle's theorem to $h(x)$, so there is $c \in (a, b)$ where $h'(c) = 0$.

We compute:

$$h'(x) = [f(a) - f(b)]g'(x) - f'(x)[g(a) - g(b)]$$
$$h'(c) = 0 \implies [f(a) - f(b)]g'(c) = f'(c)[g(a) - g(b)]$$

□

Remark: In the special case that $g(a) = f(a)$ and $g(b) = f(b)$, we're guaranteed a c such that $f'(c) = g'(c)$, which is achieved when $f(x) - g(x)$ has a local max/min.

Proposition 4

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

$$[f(b) - f(a)] = (b - a)f'(c))$$

Proof. Let $g(x) = x$ and apply generalized mean value theorem. □

A corollary of this is that if f is as above, and we have $|f'(x)| \leq M$ for some constant M, then f is uniformly continuous. This formalizes our earlier idea of uniformly continuous functions having a bounded slope.

Another corollary is that if we have $f : [a, b] \to \mathbb{R}$ that is continuous everywhere and differentiable on (a, b), if $f'(x) \geq 0$ $\forall x \in (a, b)$ then f is monotone increasing. Furthermore, if $f'(x) > 0$ $\forall x \in (a, b)$ then f is strictly increasing. The same argument applies to monotone and strictly decreasing functions.

Theorem 4: Intermediate Value Theorem for Derivatives

Assume that f is differentiable on $[a, b]$, with $f'(a) < f'(b)$. Then for each $M \in (f'(a), f'(b))$, there exists a $c(a, b)$ such that $M = f'(c)$.

Proof. Let $g(x) := f(x) - Mx$. Then $g'(a) = f'(a) - M < 0$, and $g'(b) = f'(b) - M > 0$, so neither are local minima (and thus not global minima). Let c be a global minimum of g on (a, b), then:

$$g'(c_0 \implies f'(c) = M$$

□

Now we get to L'Hopital's rule, in which given a quotient of functions where plugging in our limit value yields an indeterminate form, we want to find the limit value.

Theorem 5: L'Hopital's Rule

Assume that $f, g : (a, b) \to \mathbb{R}$ are differentiable, $g(x) \neq 0$ on a, b. If either:

1. $\lim_{x \to a} f(x) = 0$ and $\lim_{x \to a} g(x) = 0$
2. $\lim_{x \to a} f(x) = +\infty$

and if

$$\lim_{x \to a} \frac{f'(x)}{g'(x)} = A \in \mathbb{R} \cup \{-\infty, +\infty\}$$

then we have

$$\lim_{x \to a} \frac{f(x)}{g(x)} = A \in \mathbb{R} \cup \{-\infty, +\infty\}$$

Proof. The proof is very hard uwu. For any $\epsilon > 0$, $\exists \delta > 0$ such that $\forall x \in (a, a + \delta)$, we have that;

$$\left| \frac{f'(x)}{g'(x)} - A \right| < \epsilon \iff A - \epsilon < \frac{f'(x)}{g'(x)} < A + \epsilon$$

Now for all α, β such that $a < \alpha < \beta < a + \delta$, we have some $\gamma \in (\alpha, \beta)$ such that:

$$[f(\beta) - f(\alpha)]g'(\gamma) = [g(\beta) - g'(\alpha)]f'(\gamma)$$

by the mean value theorem. This means:

$$\frac{f(\beta) - f(\alpha)}{g(\beta) - g(\alpha)} = \frac{f'(\gamma)}{g'(\gamma)} \in (A - \epsilon, A + \epsilon)$$

Suppose we are in case 1 of the theorem. Then we know that $\lim_{\alpha \to a} f(\alpha) = 0$ and $\lim_{\alpha \to a} g(\alpha) = 0$. If we take the limit as $\alpha \to a$:

$$\lim_{\alpha \to a} \frac{f(\beta) - f(\alpha)}{g(\beta) - g(\alpha)} = \frac{f(\beta)}{g(\beta} \in (A - \epsilon, A + \epsilon)$$

If we like parse what we just did, for all $\epsilon > 0$ we have found a $\delta > 0$ such that $\forall \beta \in (a, a+\delta)$ such that $\frac{f(\beta)}{g(\beta)} \in (A-\epsilon, A+\epsilon)$, which is the definition that: $\lim_{\beta \to a} \frac{\beta}{g(\beta)} = A$.

Great so now what if we're in case 2 of the theorem, in which $\lim_{x\to a} f(x) = +\infty$. We pick $\beta \in (a, a+\delta$, pick $\alpha \in (a, \beta)$ such that α is close enough to a thatL

$$g(\alpha) > g(\beta) \implies \left(\frac{g(\alpha) - g(\beta)}{g(\alpha)}\right) > 0$$

$$A - \epsilon < \frac{f(\beta) - f(\alpha)}{g(\beta) - g(\alpha)} < A + \epsilon$$

$$\implies (A-\epsilon)\left(\frac{g(\alpha) - g(\beta)}{g(\alpha)}\right) < \left(\frac{f(\beta) - f(\alpha)}{g(\alpha)}\right) < (A+\epsilon)\left(\frac{g(\alpha) - g(\beta)}{g(\alpha)}\right)$$

If we take the limit as $\alpha \to a, \frac{g(\alpha)-g(\beta)}{g(\alpha)} \to 1$. Therefore, we have that:

$$A - \epsilon \leq \lim\ \text{inf}_{\alpha \to a}\left(\frac{f(\beta) - f(\alpha)}{g(\alpha)}\right) \leq \lim\ \text{sup}_{\alpha \to a}\left(\frac{f(\beta) - f(\alpha)}{g(\alpha)}\right) \leq A + \epsilon$$

Take the limit as $\epsilon \to 0$ to get that $\lim_{x\to a} \frac{f(x)}{g(x)} = A$

□

1.3 Higher Derivatives and Taylor Expansion

For continuous $f : \mathbb{R} \to \mathbb{R}$, we say that $f \in \mathcal{C}^0(\mathbb{R})$, and if $f'(x)$ exist for all $x \in \mathbb{R}$, and f' is continuous, we say that $f \in \mathcal{C}^1(\mathbb{R})$. If f' is also differentiable, we can get

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

We can denote this as $f''(x)$ or $f^{(2)}(x)$. If this exists for all x, and is continous, then $f \in \mathcal{C}^2(\mathbb{R})$. We may extend this definition to $\mathcal{C}^\infty(R)$, as functions that are in $\mathcal{C}^k(\mathbb{R})$ for all $k = 1, 2, ...$, which we may call "smooth" or infinitely differentiable functions.

Example 2: Polynomials are smooth functions. $f^{(k)}$ exists and is also a polynomial. Therefore an inductive argument shows that $f \in \mathcal{C}^\infty(\mathbb{R})$.

Example 3: Consider the piecewise defined function:

$$f(x) = \begin{cases} 0 & x \leq 0 \\ x^2 & x > 0 \end{cases}$$

This function looks smooth when graphed, but if after taking the first derivative, we get a sharp edge at $x = 0$ in which $f''(0)$ jumps from 0 to 1.

Suppose we want to approximate a function at a point by creating a function with the same derivative. Consider

$$p(x) = a_0 + \frac{a_1}{1}x + \frac{a_2}{1\cdot 2}x^2 + \cdots + \frac{a_n}{n!}x^n$$
$$p'(x) = 0 + a_1 + \frac{a_2}{1}x + \frac{a_3}{1\cdot 2}x^2 + \cdots + \frac{a_n}{(n-1)!}x^{n-1}$$

We note here that $p(0) = a_0$, $p'(0) = a_1$, and $p^{(k)}(0) = a_k$. So there is a nice function such that it's value and k^{th} derivative at $x = 0$ can be specified (note that $x = 0$ isn't special, that takes values at $x = a$)

Define the **n-th order Taylor expansion centered at a point** as follows. Assume $f : \mathbb{R} \to \mathbb{R}$ is a $\mathcal{C}^n$ function. Then we can use $f(x_0), f'(x_0), ..., f^n(x_0)$ to create a polynomial:

$$P_{x_0}(x) = f(x_0) + f'(x_0)\frac{x}{1} + \cdots + f^n(x_0)\frac{(x-x_0)^n}{n!}$$

Note that $P_{x_0}^{(k)}(x) = f^{(k)}(x_0)$.

Theorem 6: Taylor's Theorem

Suppose $f : \mathbb{R} \to \mathbb{R} \in \mathcal{C}^n(\mathbb{R})$, and $f^{(n)}(x)$ exists (but may not be continuous). Let $P_{x_0}(x)$ be the n-th order Taylor approximation of f at x_0.

Then $\forall x \in \mathbb{R}$, $\exists \theta \in [0, 1]$ such that if we define $x_\theta = x_0(1-\theta) + x_0 \cdot \theta$, we have:

$$f(x) - P_{x_0}(x) = f^{(n+1)}(x_\theta) \cdot \frac{(x-x_0)^{n+1}}{(n+1)!}$$

Sanity Check: if $n = 0$, then $P_{x_0}(x) = f(x_0)$, then $\exists x_\theta$ such that:

$$f(x) - f(x_0) = f'(x_\theta)\frac{(x-x_0)}{1}$$

which looks exactly like the mean value theorem. The general case isn't that different than this.

Proof. Fix x_0, and $x_1 \in \mathbb{R}$. We're trying to find x_θ such that $f(x_1) - P_{x_0}(x_1) = f^{(n+1)}(x_\theta)\frac{(x_1-x_0)^{n+1}}{(n+1)!}$. Define $M \in R$ such that $f(x_1) - P_{x_0}(x_1) = (x_1 - x_0) \cdot M$. Let

$$g(x) = f(x) - P_{x+0}(x) - M(x-x_0)^{n+1}$$

Note that $g(x_0) = f(x_0) - P_{x_0}(x_0) = 0$ and $g(x_1)$ is 0 by definition of M. At x_0 we have perfect approximation, and therefore $g^{(k)}(x_0) = f^{(k)}(x_0) - P_{x_0}^{(k)}(x_0) = 0$ for $0 \leq k \leq n$. We will iterate the mean value theorem:

- Use $g(x_0) = 0$, $g(x_1) = 0$, so $\exists a_1 \in (x_0, x_1)$ such that $g'(a_1) = 0$.
- Use $g'(x_0) = 0$, $g'(a_1) = 0$ to get $a_2 \in (x_0, a_1)$ such that $g''(a_2) = 0$.
- Continue this process until we get $a_{n+1} \in (x_0, a_n)$ such that $g^{(n+1)}(a_{n+1}) = 0$.

Therefore we have that:

$$0 = g^{(n+1)}(a_{n+1}) = f^{(n+1)}(a_{n+1}) - 0 - M(n+1)!$$

Rearranging the equation from when we defined M, we realize that:

$$f(x_1) - P(x_1) = (x_1 - x_0)^{n+1}\frac{f^{(n+1)}(a_{n+1})}{(n+1)!}$$

We see that this is what we set out to prove, where a_{n+1} is the x_θ we're looking for. □

1.4 Taylor's Theorem, Differentiation, Integration

Let $f : \mathbb{R} \to \mathbb{R} \in \mathcal{C}^\infty(\mathbb{R})$. Let $x_0 \in \mathbb{R}$, and let N be a positive integer. The N-th order Taylor expansion of f, centered at x_0, is the polynomial $P(x)$ such that $P^{(k)}(x_0) = f^{(k)}(x_0)$ for all natural numbers k. Also, the degree of P is at most N. Concretely,

$$P_{x_0}(x) = \sum_{k=0}^{n} f^{(k)}(x_0)\frac{(x-x_0)^k}{k!}$$

The remainder $f(x) - P(x) = R(x)$ has the property that $R^{(k)}(x_0) = 0$ for all $0 \le k \le N$.

We say a smooth function $f : \mathbb{R} \to \mathbb{R}$ is **analytic** at point $x_0 \in \mathbb{R}$ if $\exists R > 0$ such that $f(x) = \sum_{n=0}^{\infty} a_n(x-x_0)^n$ for all $|x - x_0| < R$. If f is analytic at x_0, then we can obtain the coefficients of the expansion from the derivative: $a_n = \frac{f^{(n)}(x_0)}{n!}$.

Remark: There are smooth functions f such that $f(0) = 0$, $f'(0) = 0, \ldots, f^{(n)}(0) = 0$, but $f(x)$ is not identically 0. For example, the function defined as

$$f(x) = \begin{cases} 0 & x \le 0 \\ e^{-\frac{1}{x}} & x > 0 \end{cases}$$

Proposition 5: example of non-analytic function

$$\lim_{x \to 0^+} \frac{e^{-\frac{1}{x}}}{x^n} = 0$$

Proof. Let $u = \frac{1}{x}$, in which case (*) is equivalent to:

$$\lim_{u \to \infty} \frac{e^{-u}}{\left(\frac{1}{u}\right)^n} = \lim_{u \to \infty} u^n e^{-u} = \lim_{u \to \infty} \frac{n!}{e^u} = 0$$

where the last equality follows from repeatedly applying L'Hopital's Rule. This is an example of a smooth function that is not analytic at 0. □

To determine whether $f(x) = \frac{1}{1+x}$ is analytic at $x = 0$ we need to study the Taylor expansion around $x = 0$.We find that:

$$f^{(n)}(x) = \frac{(-1)...(-n)}{(1+x)^{n+1}}$$

and at $x = 0$, we get Taylor expansion:

$$\sum_{n=0}^{\infty} (-1)^n x^n$$

A necessary and sufficient condition for this to converge is $|x| < 1$. We know that this geometric series converges when $|x| < 1$, or when $x \in B_1(0)$. 1 is the **radius of convergence**. In general, you may ask how can I cook up an analytic function? The answer is that you just need to provide a convergent power series. You just need to specify the coefficients.

Theorem 7

Let $\sum_{n=0}^{\infty} a_n(x-x_0)^n$ be a **power series** at x_0. Then, we can define $\alpha = \limsup_{n\to\infty} |a_n|^{\frac{1}{n}}$, and let $R = \frac{1}{\alpha}$. Then,

- If $|x - x_0| < R$, the series converges.
- If $|x - x_0| > R$, the series diverges.
- If $|x - x_0| = R$, then it depends.
- If $\alpha = 0$, so $R = \infty$, then the series is always convergent.

As an example, consider

$$\sum \frac{1}{n^2} x^n$$

We have

$$\alpha = \lim_{n\to\infty} \sup \left(\frac{1}{n^2}\right)^{1/n} = 1$$

so $R = 1$.

- if $|x| < R = 1$ it converges
- if $|x| > R = 1$ it diverges
- if $|x| = 1$ it still converges in this case because $\sum n^{-2} = \frac{\pi^2}{6} < \infty$. (situational)

Remark: Taylor expansion is just one way to approximate a function. The advantage is that it accurately reproduces the derivatives around a point. For a period function, with period T, i.e. $f(x+T) = f(x)$, we can use Fourier series;

$$f(x) \approx \sum_{n=0}^{\infty} a_n \sin\left(2\pi n \frac{x}{T}\right) + b_n \cos\left(2\pi n \frac{x}{T}\right)$$

We can also use polynomial interpolation, where we want a polynomial $P(x)$ such that $P(x_n) = f(x_n)$ for finitely many distinct $x_i \in \mathbb{R}$.

Theorem 8: Weierstrauss Approximation Theorem

Given a continuous function f on $[0,1]$, and given any $\epsilon > 0$, there exists a polynomial $P(x)$ such that $\forall x \in [0,1]$, $|f(x) - P(x)| < \epsilon$.

Remark: This theorem can work for any closed interval by rescaling.

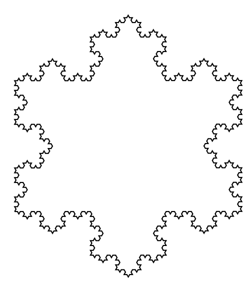
Chapter 2

Integration

2.1 Introduction to Riemann Integration

What is integration? $\int_0^1 f(x)\,\mathrm{d}x$ is the area under the graph $f(x)$, accounting for the fact that if $f(x)$ is below the x axis, it's area is counted as negative (**signed area** under the graph). One way to count area is to cut the piece into more regular pieces. For example if we cut the figure into squares we can count the number of squares strictly inside the figure and the number of squares on the boundary. By making the square size smaller and smaller we can get a more accurate approximation of the area of the figure. That is the plan, but will the plan succeed? Subscribe to find out!

A **fractal** is a curve on $\mathbb{R}^2$ for which we can try to measure the length by straight line approximations on the boundary, but where taking the limit as the length of each segment approximation fails to converge (pictured below is the Koch fractal).



Riemann integral: Let $f : [a, b] \to \mathbb{R}$ be a bounded (not necessarily continuous) function. Let P be a partition of $[a, b]$:

$$P = \{a = x_0 \leq x_1 \leq \cdots \leq x_N = b\}$$

Let $\Delta x_i = x_i - x_{i-1}$ be the length of the i-th segment. We can also define $M_i = \sup_{[x_{i-1},x_i]} f(x)$ and $m_i = \inf_{[x_{i-1},x_i]} f(x)$.

For a given partition P, define:

$$U(P, f) = \sum_{i=1}^{N} M_i \cdot \Delta x_i$$

$$L(P, f) = \sum_{i=1}^{N} m_i \cdot \Delta x_i$$

We say a partition Q **refines** P if "as a set of cut points" $Q \supsetneq P$.

Proposition 6

If Q refines P, then:
$$L(Q, f) \geq L(P, f)$$
$$U(Q, f) \leq U(P, f)$$

Proof. The proof is clear from the fact that Δx_i becomes smaller, but the sup cannot increase and the inf cannot decrease.

Informally, the limit as P gets more refined:

$$L(f) := \lim_{P} L(P, f)$$

The expression P gets more and more refined is not that well defined. So instead we can take the supremum over all possible partitions.

$$\underline{\int} = L(f) = \sup_{P \text{ any partition}} L(P, f)$$

Similarly, we can do the same for the U and the inf.

$$\overline{\int} = U(f) = \inf_{P \text{ any partition}} U(P, f)$$

□

We say f is **Riemann integrable** if $\underline{\int_a^b} f \, \mathrm{d}x = \overline{\int_a^b} f \, \mathrm{d}x$.

Theorem 9: Continuous Functions are Riemann integrable

If $f : [a, b] \to \mathbb{R}$ is continuous (hence bounded and uniformly continuous since $[a, b]$ compact), then f is Riemann integrable

Proof. We want to show that $\forall \epsilon > 0$, $\exists$ partition P such that:

$$\overline{\int_a^b} f \mathrm{d}x - \underline{\int_a^b} f \mathrm{d}x < \epsilon$$

Let $\tilde{\epsilon} = \frac{\epsilon}{b-a}$, and by unif. continuity, $\exists \delta > 0$ such that $|x - y| < \delta \implies |f(x) - f(y)| < \tilde{\epsilon}$. Choose any partition P such that $\Delta x_i < \delta$ (for example let $N = \lceil \frac{b-a}{\delta} \rceil$). Then $M_i = \sup_{[x_{i-1}, x_i]} f(x) = f(s_i)$ for some $s_i \in [x_{i-1}, x_i]$. Similarly, $m_i = \inf_{[x_{i-1}, x_i]} f(x) = f(t_i)$ for some $t_i \in [x_{i-1}, x_i]$. Therefore, $|M_i - m_i| = |f(s_i) - f(t_i)| < \tilde{\epsilon}$. Thus,

$$\begin{aligned} U(P, f) - L(P, f) &= \sum (M_i - m_i) \Delta x_i \leq \sum \tilde{\epsilon} \Delta x_i \\ &= \tilde{\epsilon} \sum \Delta x_i = \tilde{\epsilon} \cdot (b - a) = \epsilon \end{aligned}$$

□

Example 1: Consider $f : [0,1] \to \mathbb{R}$ bounded, given by

$$f(x) = \begin{cases} 0 & x = 0 \\ \sin\left(\frac{1}{x}\right) & x \in (0,1] \end{cases}$$

Proof. This function is integrable. $\forall \epsilon > 0$, consider a partition of the form $\left[0, \frac{\epsilon}{4}\right]$, and some other partition of $\left[\frac{\epsilon}{4}, 1\right]$. Let P' be a partition of $\left[\frac{\epsilon}{4}, 1\right]$ such that:

$$U\left(P', f, \left[\frac{\epsilon}{4}, 1\right]\right) - L\left(P', f, \left[\frac{\epsilon}{4}, 1\right]\right) < \frac{\epsilon}{2}$$

Then, $P = \left[0, \frac{\epsilon}{4}\right] \cup P'$. Then

$$U\left(P, f, [0,1]\right) - L\left(P, f, [0,1]\right) \leq (1 - (-1)) \cdot \left(\frac{\epsilon}{4}\right) + \frac{\epsilon}{2} \leq \epsilon$$

□

Proposition 7

If $f : [a,b] \to \mathbb{R}$ is a bounded function with finitely many discontinuities, then f is integrable.

Proof. Repeat the above argument for all (finitely many) discontinuities □

Theorem 10: Bounded monotone functions are integrable

If f is a monotone function over $[a,b]$, then f is integrable.

Proof. We recall that a monotone function can have at most countably many discontinuities. Remember how we prove that, if we have a discontinuity, there will be a gap in the range. But this means that there is a rational number inside that gap, so therefore the number of gaps is less than or equal to the number of rational numbers, which is countable. Also, for all discontinuities p, the left and right limits exist at p.

Fix $\epsilon > 0$. $\forall n \in \mathbb{N}, n > 0$, consider an equipartition P_n with n segments such that each segment has length $\frac{b-a}{n} = \delta$. Then:

$$\begin{aligned} U(P_n, f) - L(P_n, f) &= \sum (M_i - m_i) \cdot \delta \\ &= \sum_{i=1}^{n} (f(x_i) - f(x_{i-1})) \cdot \delta \\ &= \delta \cdot \sum_{i=1}^{n} (f(x_i) - f(x_{i-1}) \\ &= \delta \cdot [f(b) - f(a)] \end{aligned}$$

By making n large enough, we can make δ small enough that:

$$\delta \cdot [f(b) - f(a)] < \epsilon$$

□

2.2 The Riemann-Stieltjes Integral

Now we can define the **Riemann-Stieltjes integral**, which accounts for density.

• ——— •

For a rod of uniform density, the center of mass will be the midpoint. For this uniform rod, we find that the center of mass will be :

$$\frac{\int_0^L x \, \mathrm{d}x}{\int_0^L \mathrm{d}x} = \frac{1}{2}L$$

Now consider a rod of non-uniform density, in which we have density $\rho(x)$ at x. Then $\rho(x)\mathrm{d}x$ denotes the mass of a small segment. We see that the center of mass is :

$$\frac{\int_0^L x\rho(x) \, \mathrm{d}x}{\int_0^L \rho(x) \, \mathrm{d}x} = M$$

Suppose we want our density function to be able to encode things like jumps and point discontinuities. One general way to replace $\rho(x)\mathrm{d}x$, is by $\mathrm{d}\alpha(x)$, where α is the cumulative mass function. $\alpha(x)$ is the the mass on the interval $[a, x]$:

$$\alpha(x) = \int_a^x \rho(x) \, \mathrm{d}x$$

We want this function to be monotone increasing, because we should have non-negative density, so the more volume we include, the heavier our object should get. This means that $\alpha(x)$ is monotone increasing, which by the previous theorem is integrable. For example:

$$\alpha(x) = \begin{cases} 0 & 0 \leq x \leq \frac{1}{2} \\ 3 & \frac{1}{2} \leq x \leq 1 \end{cases}$$

in which case $\mathrm{d}(\alpha(x)) = 3\delta(x - \frac{1}{2}) \cdot \, \mathrm{d}x$.

Suppose we want to compute the center of mass of 2 points. We expect this to be:

$$\frac{x_1 m_1 + x_2 m_2}{m_1 + m_2}$$

We just generalize this to an interval using integration:

$$\frac{\int_0^1 x \, \mathrm{d}\alpha(x)}{\int_0^1 \mathrm{d}\alpha(x)}$$

Let $\alpha : [a, b] \to \mathbb{R}$ be monotone increasing. Let $f : [a, b] \to \mathbb{R}$ be bounded. Let P be a partition of $[a, b]$ $a = x_0 \leq x_1 < \cdots < x_N = b$, and let $I_i = (x_{i-1}, x_i]$. Also, define $\Delta\alpha(I_i) = \alpha(x_i) - \alpha(x_{i-1})$, and recall $M_i = \sup_{I_i} f, m_i = \inf_{I_i} f$ as in the definition of the Riemann integral.

$$U(P, f, \alpha) = \sum M_i \cdot \Delta\alpha(I_i)$$

$$L(P, f, \alpha) = \sum m_i \cdot \Delta\alpha(I_i)$$

If we let:

$$U(f, \alpha) = \inf_P U(P, f, \alpha)$$

$$L(f, \alpha) = \sup_P L(P, f, \alpha)$$

if $U(f, \alpha) = L(f, \alpha)$, we say that f is **Riemann integrable with respect to** α.

Theorem 11: Density analog of Thm 9

If f is continuous and α is monotone, then $\int_a^b f \, \mathrm{d}\alpha(x)$ exists.

Remark: if we have:

$$f(x) = \begin{cases} 0 & x \in \left[0, \frac{1}{2}\right) \\ 1 & x \in \left[\frac{1}{2}, 1\right] \end{cases}$$

and

$$\alpha(x) = \begin{cases} 0 & x \in \left[0, \frac{1}{2}\right) \\ 1 & x \in \left[\frac{1}{2}, 1\right] \end{cases}$$

Then $U(P, f, \alpha) - L(P, f, \alpha) = 1$. Any partition with a bin containing the jump (i.e. there is a segment with $\frac{1}{2}$ in the interior), then the mass of the segment will be 1, so $U - L = (1 - 0) \cdot 1 = 1$. Alternatively, what if we have $\frac{1}{2}$ as a cut point of the partition, then $U - L = 1$ if $1/2$ is in the lower partition.

Theorem 12: Density analog of Thm 10

Suppose f is monotone, α is continuous and monotone. Then $\int_a^b f \, \mathrm{d}\alpha(x)$ exists.

Proof. Since α is continous and monotone on $[a, b]$, and for each $n \in \mathbb{N}, n > 0$, we can let let $y_0, y_1, ..., y_N$ be an equipartition of $[\alpha(a), \alpha(b)]$. Then choose corresponding x_i in the pre-image of y_i under α. Then $\alpha(x_i) - \alpha(x_{i-1}) = y_i - y_{i-1} = \frac{\alpha(b)-\alpha(a)}{n}$. Then:

$$\begin{aligned} U(P, f, \alpha) - L(P, f, \alpha) &= \sum [f(x_i) - f(x_{i-1})] \cdot \delta \\ &\leq \sum_i [f(x_i) - f(x_{i-1})] \cdot \delta \\ &= (f(b) - f(a)) \cdot \delta \end{aligned}$$

and again we can make n arbitrarily large to make the whole quantity less than ϵ.

□

Example 2: If $\alpha(x)$ is smooth.

- $[a, b] = [0, 1]$
- $\alpha(x) = 2 + 3x$
- $f(x) = 1$

Then

$$\int_0^1 f(x) \, \mathrm{d}\alpha(x) = \lim_{P \text{ a partition}} \sum f(x_i) \Delta\alpha_i = \alpha(1) - \alpha(0) = 3$$

If α is a smooth function (at least differentiable, say $\alpha'(x) = \rho(x)$, then $\mathrm{d}\alpha(x) = \rho(x)\mathrm{d}x$, and we can just do a normal Riemann integral. If α has finitely many jumps, we can still do the integration. We can deal with the jumps with some special treatment.

$$\alpha(x) = \begin{cases} x & x \in [0, 1] \\ 1 + x & x \in (1, 2] \\ 2 + x & x \in (2, 3] \end{cases}$$

$$\int_0^3 1 \, d\alpha(x) = \int_{0^+}^{1^-} d\alpha(x) + \int_{1^+}^{2^-} d\alpha(x) + \int_{2^+}^{3^-} d\alpha(x) + \sum_{P \text{ are jumps of } \alpha} (\alpha(P_+) - \alpha(P_-))$$
$$= 1 + 1 + 1 + (1 + 1) = 5$$

Theorem 13

If $f : [a, b] \to \mathbb{R}$ is bounded, and has finitely many discontinuities, and if α is continuous when f is discontinuous, then $\int_a^b f \, d\alpha$ exists.

Remark: If $\alpha(x) = x$, the usual Riemann integral, we show this by creating partitions around the jumps/discontinuities of f.

Proof. Fix $\epsilon > 0$. Let $M = \sup f(x)$. Let $E = \{c_1, \ldots, c_m\}$ be the points at which f is discontinuous. The steps are given below:

1. Choose small enough intervals $[u_j, v_j]$ centered around c_j, such that $\sum \alpha(v_j) - \alpha(u_j) < \epsilon$ and these intervals are disjoint. We can let v_j, u_j get arbitrarily close to c_j and therefore shrink the bound (since α is continuous)

2. Let $K = [a, b] \setminus \bigcup_{j=1}^m (u_j, v_j)$. This is still a compact set. There exists a partition P of K that's fine enough such that:
$$U(P, f, \alpha) - L(P, f, \alpha) < \epsilon$$
Then, let $\tilde{P} = P \cup \bigcup_{i=1}^m [u_i, v_i]$. We conclude
$$U(\tilde{P} - L(\tilde{P}) < \epsilon + \sum_{i=1}^m (M - (-M)) \cdot \Delta\alpha_i)$$
$$< \epsilon + 2m\epsilon = (1 + 2m)\epsilon$$
We can make this arbitrarily small with ϵ, and therefore the integral exists.

□

Theorem 14: Rudin 6.11

Let $f : [a, b] \to \mathbb{R}$ be integrable with respect to weight α, and assume the range $f([a, b]) \subseteq [m, M]$. If $\phi : [m, M] \to R$ is continuous, then $h(x) = \phi(f(x))$ is integrable with respect to $\alpha(x)$

Example 3:

- $\alpha(x) = x$
- $f(x)$ is some monotone function
- ϕ is some smooth function like exponential $e^{|x|}$

Then $\int_a^b f \, dx = \int_a^b e^{|f(x)|} \, dx$ exists.

Proof. Fix $\epsilon > 0$. Since ϕ is continuous on $[m, M]$, it is uniformly continuous (domain is compact), so $\exists\delta$ such that if $|y_2 - y_1| < \delta$ then $|\phi(y_2) - \phi(y_1)| < \epsilon$.

Since f is integrable, $\exists P$ a partition of $[a, b]$ such that $U(P, f, \alpha) - L(P, f, \alpha) < \delta^2$. For interval $I_i = [x_{i-1}, x_i]$, let $M_i = \sup_{I_i} f$ and $m_i = \inf_{I_i} f$. Also let

$$M_i^* = \sup_{x \in I_i} \phi(f(x)) \qquad m_i^* = \inf_{x \in I_i} \phi(f(x))$$

we say I_i is of **short type** if $M_i - m_i < \delta$. We claim that then

$$M_i^* - m_i^* < \epsilon$$

$M_i^* - m_i^* = \sup x_1, x_2 \in I_i |h(x_1) - h_(x_2)|$, because $x_1, x_2 \in I_i$, then $f(x_1), f(x_2) \in [m_i, M_i]$, which has length $< \delta$, so we use uniform continuity to verify our claim. Otherwise, I_i is of **"long" type**, where we write the indices into 2 sets $A \sqcup B$ where I_i is a short type iff $i \in A$. Then:

$$M_i^* - m_i^* \leq 2 \sup |h| = 2K$$

We just take the peak of the function K, and realize that the entire function is contained in $[-K, K]$. Then:

$$\delta \cdot \sum_{i \in B} \leq \sum_{i \in B} (M_i - m_i) \cdot \Delta\alpha_i \leq U(P, f, \alpha) - L(P, f, \alpha) \leq d^2$$

so we get that:

$$\sum_{i \in B} \Delta\alpha_i < \delta$$

and thus

$$\begin{aligned} U(P, h, \alpha) - L(P, h, \alpha) = \sum_{i=1}^{n} (M_i^* - m_i^*) \cdot \Delta\alpha_i &= \sum_{i \in A}^{n} (M_i^* - m_i^*) \cdot \Delta\alpha_i + \sum_{i \in B}^{n} (M_i^* - m_i^*) \cdot \Delta\alpha_i \\ &\leq \sum_{i \in A}^{n} (M_i^* - m_i^*) \cdot \Delta\alpha_i + \sum_{i \in B} 2K\Delta\alpha_i \\ &\leq \epsilon[\alpha(b) - \alpha(a)] + 2k \cdot \delta \end{aligned}$$

we can use the same trick where we make δ super small, and use ϵ' or something to make that small too. ☐

Theorem 15: Rudin 6.12

Informally, $\int f \, d\alpha$ is linear in f and α. Formally:

1. If f, g are integrable with respect to α, then:

   $$\int c \cdot f \, d\alpha$$

   exists for all constants c and is equal to $c \cdot \int f \, d\alpha$, and:

   $$\int f + g \, d\alpha$$

   also exists and is equal to $\int f \, d\alpha + \int g \, d\alpha$

2. If f is integrable with respect to α_1, α_2, then f is integrable with respect to $c\alpha_1$ for positive c, and f is integrable with respect to $\alpha_1 + \alpha_2$

Theorem 16: Rudin 6.13

1. If f and g are integrable with respect to α then $f \cdot g$ is integrable.
2. If f is integrable, then $|f|$ is integrable.

The latter follows from setting $\phi(y) = |y|$ and applying our result about composition with continuous functions.

Recall our setup for integration. We have $f : [a, b] \to \mathbb{R}$ bounded, $\alpha : [a, b] \to \mathbb{R}$ monotone increasing, and P a partition of $[a, b]$ into closed intervals $I_i = [x_{i-1}, x_i]$, where $\Delta\alpha_i = \alpha(x_i) - \alpha(x_{i-1})$. We also defined $M_i = \sup_{I_i} f, m_i = \inf_{I_i} f$, so that:

$$U(P, f, \alpha) = \sum_{i=1}^{n} M_i \cdot \Delta\alpha_i$$

$$L(P, f, \alpha) = \sum_{i=1}^{n} m_i \cdot \Delta\alpha_i$$

We said that f is integrable if $\lim_P U(P, f, \alpha) - L(P, f, \alpha) = 0$ as P gets more refined.

Theorem 17: Sampling Lemma

- $\forall i = 1, \ldots, n$, pick $s_i \in I_i$, then:

$$L(P, f, \alpha) \leq \sum f(s_i)\Delta \cdot \alpha_i \leq U(P, f, \alpha)$$

- If $U - L \leq \epsilon$, then for any $s_i, t_i \in I_i$:

$$\sum |f(s_i) - f(t_i)| \cdot \Delta\alpha_i \leq \sum (M_i - m_i) \cdot \Delta\alpha_i = U - L \leq E$$

Last time, we said that if there's a smooth positive density function$\rho(x)$ such that $\alpha'(x) = \rho(x)$, $\int f\rho(x)\ \mathrm{d}x = \int f\ \mathrm{d}\alpha$, so now we only require that we have integrable α.

If $\int f\ \mathrm{d}x$ exists, we say that f is Riemann integrable, and we say that $f \in \mathcal{R}$, **the set of Riemann integrable functions**. If $\int f\ \mathrm{d}\alpha$ exists, then $f \in \mathcal{R}(\alpha)$.

Theorem 18: Rudin 6.17

Suppose that f is bounded and α is increasing. Further suppose α' is exists and is integrable. Then:

1. We have
$$f \in \mathcal{R}(\alpha) \iff f\alpha' \in \mathcal{R}$$
2. We have
$$f \in \mathcal{R}(\alpha) \implies \int_a^b f\ \mathrm{d}\alpha = \int_a^b f\alpha'\ \mathrm{d}x$$

Recall that $U(f, \alpha) = \lim_P U(P, f, \alpha) = \overline{\int}_a^b f\ \mathrm{d}\alpha$ and that $L(f, \alpha) = \lim_P L(P, f, \alpha) = \underline{\int}_a^b f\ \mathrm{d}\alpha$. The idea is that we want to prove that:

$$\overline{\int_a^b} f \ \mathrm{d}\alpha = \overline{\int_a^b} f\alpha' \ \mathrm{d}x$$

$$\underline{\int_a^b} f \ \mathrm{d}\alpha = \underline{\int_a^b} f\alpha' \ \mathrm{d}x$$

Proof. We're going to show that for any $\epsilon > 0$, there's a partition such that:

$$|U(P, f, \alpha) - U(P, f, \alpha')| < \epsilon$$

which holds for any refinement of partition P. Since α' is integrable, there's a partition P such that:

$$U(P, f, \alpha') - L(P, f, \alpha') < \epsilon$$

By the mean value theorem, since we know that $\exists t_i \in I_i$ such that

$$\Delta\alpha_i = \alpha(x_i) - \alpha(x_{i-1}) = \alpha'(t_i)\Delta x_i$$

and by the sampling lemma, $\forall s_i \in I_i$, we have:

$$\sum |\alpha'(s_i) - \alpha'(t_i)| \ \Delta x_i < \epsilon$$

Thus:

$$\left|\sum_i f(s_i) \ \Delta\alpha_i - f(s_i)\alpha(s_i)\Delta x_i\right| = \left|\sum_i f(s_i)\alpha'(t_i)\Delta x_i - f(s_i)\alpha'(s_i)\Delta x_i\right|$$

for some sample point $s_i \in I_i$. By our boundedness assumption, we can let $M = \sup_{[a,b]} |f|$. Then the entire quantity above is less than or equal to:

$$\sum_i |f(s_i)||\alpha'(t_i) - \alpha'(s_i)| \ \Delta x_i \leq M \sum_i |\alpha'(t_i) - \alpha'(s_i)| \ \Delta x_i \leq M \cdot \epsilon$$

for all $s_i \in I_i$. This implies that:

$$U(P, f, \alpha) \leq U(P, f\alpha') + M\epsilon$$

where the new α is just uniform density as in the normal Riemann integral. Similarly,, we can use the absoute value in the other direction to get that:

$$U(P, f\alpha') \leq U(P, f, \alpha) + M\epsilon$$

we conclude that the limit of their difference is less than ϵ, and since ϵ is arbitrarily sall, we have that the limit :

$$|U(P, f\alpha') - U(P, f, \alpha)| = 0$$

The exact same argument applies to the lower bounds, so:

$$|L(P, f\alpha') - L(P, f, \alpha)| = 0$$

and we're done. □

Theorem 19: Rudin 6.19 - Change of variables

Let α be increasing on $[a, b]$, and $f \in \mathcal{R}(\alpha)$. Let $\phi : [A, B] \to [a, b]$ be a strictly increasing function. Define $g : [A, B] \to \mathbb{R}$ such that $g(y) = f(\phi(y))$. Similarly define $\beta(y) = \alpha(\phi(y))$. Then:

$$\int_a^b f \ \mathrm{d}\alpha = \int_A^B g \ \mathrm{d}\beta$$

2.3 The Fundamental Theorem of Calculus

For this section, we only consider the Riemann integral.

Theorem 20: 6.20 Fundamental Theorem of Calculus I

Let $f \in \mathcal{R}$ on $[a, b]$. For any $a \leq x \leq b$, define:

$$F(x) := \int_a^x f(t)\ \mathrm{d}t$$

Then:

1. $F(x)$ is a continous function
2. if $f(x)$ is continuous at a point $x_0 \in [a, b]$ then $F(x)$ is differentiable at x_0 and $F'(x_0) = f(x_0)$

Remark: Suppose that f has a jump discontinuity, so for example:

$$f(x) = \begin{cases} 0 & x \in \left[0, \frac{1}{2}\right] \\ 1 & x \in \left(\frac{1}{2}, 1\right] \end{cases}$$

then F will not be differentiable at $x = \frac{1}{2}$.

Proof. Let $M = \sup_{a,b} |f(x)|$. Then for any $a \leq x < y \leq b$ we have:

$$|F(y) - F(x)| = \left| \int_a^y f(t)\ \mathrm{d}t - \int_a^x f(t)\ \mathrm{d}t \right| = \left| \int_x^y f(t)\ \mathrm{d}t \right|$$

$$\leq \int_x^y |f(t)|\ \mathrm{d}t \leq \int_x^y M\ \mathrm{d}t \leq M|y - x|$$

Thus F is **Lipschitz continuous with constant M**. This is like the definition of a Lipschitz function. Lipschitz continuity $\implies$ continuity.

Now suppose f is continuous at x_0, then $\forall \epsilon > 0$, $\exists \delta > 0$ s.t.:

$$|x - x_0| < \delta \implies |f(x) - f(x_0)| < \epsilon$$

Then for any $s, t \in [a, b]$, such that:

$$x_0 - \delta < s < x_0 < t < x_0 + \delta$$

Then:

$$\frac{F(t) - F(s)}{t - s} = \frac{1}{t - s} \int_s^t f(u)\ \mathrm{d}u$$

We realize that we can write $f(x_0) = \frac{1}{t-s} \int_s^t f(x_0)\ \mathrm{d}u$ since $f(x_0)$ is just a constant.

$$\left| \frac{F(t) - F(s)}{t - s} - f(x_0) \right| = \left| \frac{1}{t - s} \int_s^t \int_s^t (f(u) - f(x_0))\ \mathrm{d}u \right|$$

$$\leq \frac{1}{t - s} \int_s^t |f(u) - f(x_0)|\ \mathrm{d}u \leq \frac{1}{t - s} \int_s^t \epsilon\ \mathrm{d}u = \frac{1}{t - s}(t - s)\epsilon = \epsilon$$

Therefore $\lim_{n \to 0} \frac{F(x+n) - F(x_0)}{n} = f(x_0)$ as desired. ☐

Theorem 21: Fundamental Theorem of Calculus II

Let F be a differentiable function on $[a, b]$ with $F'(x) = f(x)$. If $f(x)$ is integrable, then:

$$\int_a^b f(x)\ \mathrm{d}x = F(b) - F(a)$$

Note that f might not always be integrable, we may consider the *Volterra function* as an example.

Proof. Fix $\epsilon > 0$. We have that $\exists P$ a partition of $[a, b]$ where $U(P, f) - L(P, f) < \epsilon$ and:

$$\begin{aligned}
F(b) - F(a) &= \sum_{i=1}^{n} F(x_i) - F(x_{i-1}) \\
&= \sum_{1}^{n} F'(s_i)\ \Delta x_i \\
&= \sum_{1}^{n} f(s_i)\ \Delta x_i \in [L(P, f), U(P, f)]
\end{aligned}$$

Therefore $|F(b) - F(a) - \int_a^b f\ \mathrm{d}x| < U - L < \epsilon$, so the left hand side goes to 0. ☐

The canonical example is integration by parts.

$$\int_a^b f\ \mathrm{d}g = \int_a^b \mathrm{d}(fg) - f\mathrm{d}g$$

if we differentiate both sides using the fundamental theorem of calculus, we see that we get a restatement of the chain rule. For a more rigorous treatment:

Theorem 22: Integration by Parts

Suppose f, g are differentiable and f', g' are integrable functions. Then:

$$\int_a^b fg'\ \mathrm{d}x = [fg]_a^b - \int_a^b gf'\ \mathrm{d}x$$

2.4 Integrability, Differentiability, and Uniform Convergence

Recall that a sequence of functions converges uniformly if:

$$\sup_{x \in [a,b]} |f_n(x) - f(x)| = 0$$

Recall that if $f_n \to f$ uniformly and $\{f_n\}$ are continuous then f is continuous. Further, if $f_n(x)$ are integrable with respect to weight function $\alpha(x)$, what can we say about whether or not $f \in \mathcal{R}(\alpha)$? Integrability is preserved under uniform convergence, but differentiability is not always.

Theorem 23

If $f_n \to f$ uniformly, and $f_n \in \mathcal{R}(\alpha)$, then $f \in \mathcal{R}(\alpha)$, given that $f_n, f : [a, b] \to \mathbb{R}$ and $\alpha : [a, b] \to \mathbb{R}$ is monotone increasing.

Proof. We'll use the fact that for uniformly convergent functions, given $\epsilon > 0$, we can find an N such that:

$$n \geq N \implies f_n - \epsilon < f < f_n + \epsilon$$

Therefore for any P partition, we have

$$L(P, f_n - \epsilon, \alpha) - \epsilon(\alpha(b) - \alpha(a)) \leq L(P, f, \alpha) \leq U(P, f, \alpha) \leq U(P, f_n + \epsilon, \alpha) + \epsilon(\alpha(b) - \alpha(a))$$

We can control both sources of error, since we can arbitrarily refine the partition, and we can make our function approximation be better (by making n larger). Fix an $n > N$, we may choose partition P such that;

$$U(P, f_n, \alpha) - L(P, f_n, \alpha) \leq \epsilon(\alpha(b) - \alpha(a))$$

Thus

$$\begin{aligned} U(P, f, \alpha) - L(P, f, \alpha) &\leq U(P, f_n, \alpha) - L(P, f_n, \alpha) + 2\epsilon(\alpha(b) - \alpha(a)) \\ &= 3\epsilon(\alpha(b) - \alpha(a)) \end{aligned}$$

Therefore $\forall \epsilon > 0$, there exists a partition P to make $U(P, f, \alpha) - L(P, f, \alpha) < \epsilon$ work. Hence f is in $\mathcal{R}(\alpha)$. □

Corollary: Let $f_n(x) \in \mathcal{R}(\alpha)$ over $[a, b]$ and assume that $F(x) = \sum_{n=1}^{\infty} f_n(x)$ is a uniformly convergent series, so the partial sums converge uniformly. Then:

$$\int_a^b F(x) \, \mathrm{d}x = \sum_{n=0}^{\infty} \int_a^b f_n(x) \, \mathrm{d}x$$

Proof. Define the partial sum $F_N(x) = \sum_{n=1}^{N} f_n(x)$. This is the finite sum of $\alpha-$integrable sums so it is $\alpha-$integrable as well. Hence $F_N(x) \in \mathcal{R}(\alpha)$. By the previous theorem, since $F_N \to F$ uniformly and F_N integrable, then $F(x) \in \mathcal{R}(\alpha)$.

By our previous theorem,

$$\begin{aligned} \int_a^b F(x) \, \mathrm{d}x &= \lim_{N\to\infty} \int_a^b F_N(x) \, \mathrm{d}x \\ &= \lim_{N\to\infty} \int_a^b \left(\sum_{n=1}^{N} f_n(x) \right) \mathrm{d}x \\ &= \lim_{N\to\infty} \sum_{n=1}^{N} \int_a^b f_n(x) \, \mathrm{d}x \\ &= \sum_{n=1}^{\infty} \int_a^b f_n(x) \, \mathrm{d}x \end{aligned}$$

□

Now we can consider uniform convergence with respect to differentiation.

Example: $f_n \to 0$, f_n' exists and is continuous, but f_n' doesn't converge to 0:

$$f_n(x) = \frac{1}{n} \sin(n^2 x)$$

$$f_n' = n \cos(n^2 x)$$

Theorem 24

Suppose f_n is a sequence of differentiable functions such that the derivative converges uniformly:

$$f_n' \to g$$

and further, $\exists x_0 \in [a,b]$ such that $f_n(x_0) \to c$. Then:

1. $\exists f$ such that $f_n \to f$ uniformly
2. f is differentiable and $f'(x) = g(x) = \lim f_n'(x)$

Proof. Fix $\epsilon > 0$. Choose N large enough such that:

1. $\forall n, m > N$ we have $|f_n(x_0) - f_m(x_0)| < \frac{\epsilon}{2}$
2. $d_\infty\left(f_n', f_m'\right) < \frac{\epsilon}{2}\frac{1}{b-a}$

We can apply the mean value theorem to $f_n - f_m$ on interval $[x,t]$:

$$|f_n(x) - f_m(x) - f_n(t) + f_n(t)| = | = |f_n'(s) - f_m'(s)| \cdot |t-x| \leq \frac{\epsilon}{2}\frac{1}{b-a}(b-a) \leq \frac{\epsilon}{2}$$

We have that $\forall x \in [a,b]$:

$$\begin{aligned} |f_n(x) - f_m(x)| &\leq |f_n(x) - f_m(x) - (f_n(x_0) - f_m(x_0))| + |f_n(x_0) - f_m(x_0)| \\ &\leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon \end{aligned}$$

So since f_n is uniformly Cauchy, it's uniformly convergent. We still need to prove that the limit function is differentiable. We can use the definition to show that we can always compute the difference quotient everywhere in $[a,b]$. Fix a point $x \in [a,b]$ and define;

$$\phi(t) = \frac{f(t) - f(x)}{t - x}$$

We want to show that $\lim_{t\to x} \phi(t) = g(x)$. We will define an approximation:

$$\phi_n(t) = \frac{f_n(t) - f_n(x)}{t - x}$$

and since $\lim_{t\to x} \phi_n(t) = f_n'(x)$, and we have that $\lim_{n\to\infty} \phi_n(t) = \phi(t)$, it suffices to show that $\lim_{t\to x} \lim_{n\to\infty} \phi_n(t) = \lim_{n\to\infty} \lim_{t\to x} \phi_n(t)$. This is true if ϕ_n converges to g uniformly.

Recall that:

$$|f_n(x) - f_m(x) - (f_n(t) - f_m(t))| \leq \frac{\epsilon}{2(b-a)} \cdot |t - x|$$

Divide both sides by $|t-x|$ which implies that;

$$|\phi_n(t) - \phi_m(t)| \leq \frac{\epsilon}{2(b-a)}$$

so ϕ_n is uniformly Cauchy, so it's uniformly convergent on the punctured interval $[a,b] \backslash x$. Therefore we conclude that the limits commute, and therefore:

$$\lim_{t \to x} \phi(t) = g(x)$$

□

Example 1: $\phi_n(t) = t^n$ for $t \in (0,1)$. Then $\lim_{n\to\infty} \phi_n(t) = 0$, but for each fixed n, we have that $lim_{t\to 1}\phi_n(t) = 1$, so the limits will not commute since this sequence doesn't uniformly converge.

Example 2:

$$lim_{n\to\infty}\left(\frac{1}{n+1} + ... + \frac{1}{3n}\right)$$

How do we use integration to solve this limit? Let $f(x) = \frac{1}{1+x}$ We have:

$$\lim_{n\to\infty}\sum_{k=1}^{2n}\frac{1}{n+k} = \int_0^2 \frac{1}{x}\ \mathrm{d}x$$

this is because we let $x = \frac{k}{n}$.

Example 3: We can try and find;

$$\lim_{n\to\infty} n^2\left(\frac{1}{n^3+1} + ...\frac{1}{2n^3}\right)$$

We realize:

$$\frac{n^2}{n^3+k} = \frac{1}{1+\frac{k^3}{n^3}}\frac{1}{n}$$

This goes from $k = 1$ to $k = n$. Therefore this is the integral:

$$\int_0^1 \frac{1}{1+x^3}\ \mathrm{d}x$$

Exercise: Find:

$$\lim_{n\to\infty}\left(\frac{\sin\left(\frac{\pi}{n+1}\right)}{1} + \cdots + \frac{\sin\left(\frac{n\pi}{n+1}\right)}{n}\right)$$

which is

$$\lim_{n\to\infty}\sum_{k=1}^{n}\frac{1}{k}\sin\left(\frac{k\pi}{n+1}\right)$$