Northwestern University

Real Analysis

MATH 321-2

uniform continuity go brrrr

Author:
Elliott Yoon

28 March 2023

1 The Riemann Stieljis Integral

Remark 1.1. For this section, let there be a standing assumption that f is bounded.

Definition 1.2. Let $[a,b]$ be a given interval. A **partition** P of $[a,b]$ is a finite set of points $\{x_0, x_1, \ldots, x_n\}$ where

$$a = x_0 \leq x_1 \leq \cdots \leq x_n = b.$$

We will adopt the following notation: $\Delta x_i = x_i - x_{i-1}$. Now, let P be any partition of $[a,b]$. We put

1. $M_i = \sup f(x) \quad (x_{i-1} \leq x \leq x_i)$,
2. $m_i = \inf f(x) \quad (x_{i-1} \leq x \leq x_i)$,
3. $U(P,f) = \sum_{i=1}^n M_i \Delta x_i$,
4. $L(P,f) = \sum_{i=1}^n m_i \Delta x_i$,

and finally obtain the *upper* and *lower Riemann integrals* of f over $[a,b]$:

1. $\overline{\int_a^b} f\,dx = \inf_{P \in \mathscr{P}} U(P,f)$,
2. $\underline{\int_a^b} f\,dx = \sup_{P \in \mathscr{P}} L(P,f)$

where $\mathscr{P}$ is the set of all partitions P of $[a,b]$.

Lemma 1.3

The set $\{U(P,f) \mid P \in \mathscr{P}\}$ is bounded below.

Proof. Since f is bounded, $f(x) \geq m$ for all $x \in [a,b]$. Notice that

$$U(P,f) = \sum_{i=1}^n \sup_{x \in [x_{i-1}, x_i]} f(x)\,\Delta x_i \geq \sum_{i=1}^n m \cdot \Delta x_i = m(b-a).$$

□

Definition 1.4. We say that f is **Riemann-integrable** and write $f \in \mathscr{R}([a,b])$ if

$$\overline{\int_a^b} f\,dx = \underline{\int_a^b} f\,dx.$$

Remark 1.5. Notice that $L(P,f)$ and $U(P,f)$ are bounded by $m(b-a)$ and $M(b-a)$, where $m \leq f(x) \leq M$ for all $x \in [a,b]$. In other words, the upper and lower integrals exist for *every* bounded real function.

Definition 1.6. Let α be nondecreasing (monotonically increasing) function on $[a,b]$. We write $\Delta\alpha_i = \alpha(x_i) - \alpha(x_{i-1})$. (Clearly, $\Delta\alpha_i \geq 0$. We put

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

where M_i, m_i have the same meaning as in Definition 1.2 and we define

$$\overline{\int_a^b} f\,d\alpha = \inf_{P\in\mathscr{P}} U(P,f,\alpha) \tag{1}$$

and

$$\underline{\int_a^b} f\,d\alpha = \sup_{P\in\mathscr{P}} L(P,f,\alpha). \tag{2}$$

If (1) and (2) are equal, then we say f is *integrable with respect to α over* $[a,b]$, written $f \in \mathscr{R}(\alpha)$, and notate their common value, known as the **Riemann-Stieltjes integral** as

$$\int_a^b f\,d\alpha.$$

Question 1.7. When is $f \in \mathscr{R}(\alpha)$

It may be helpful to rephrase the question to ask when f is *not* in $\mathscr{R}(\alpha)$.

- Nonexample: The function
$$f(x) = \begin{cases} 1 & x \in \mathbb{Q} \\ 0 & x \notin \mathbb{Q} \end{cases}$$
is *not* in $\mathscr{R}([a,b])$. Notice that for any partition P, $U(P,f) = 1 \neq 0 = L(P,f)$.

So whenever $\inf_{P\in\mathscr{P}} U(P,f,\alpha)$ is *strictly* greater than $\sup_{P\in\mathscr{P}} L(P,f,\alpha)$, we know $f \notin \mathscr{R}(\alpha)$.

Definition 1.8. For partitions $P, Q \in \mathscr{P}$,

1. If $Q \supset P$, we say Q is a **refinement** of P.
2. We call $P^* = P \cup Q$ a **common refinement**.

Lemma 1.9

If $Q \supset P$, then $U(Q,f,\alpha) \le U(P,f,\alpha)$ and $L(Q,f,\alpha) \ge L(P,f,\alpha)$.

Proof. Let $Q = P \cup \{x_0, \ldots, x_k\}$. If $k = 0$, the conclusion obviously holds. Now, suppose $k \in \mathbb{N}$ and $U(Q,f,\alpha) \le U(P,f,\alpha)$, and let P^* contain just one more point than P, x^*, where $x_{i-1} < x^* < x_i$. Write $w_1 = \sup_{x_{i-1}\le x\le x^*} f(x)$ and $w_2 = \sup_{x^*\le x\le x_i} f(x)$. Notice $w_1, w_2 \le M_i$ where $M_i = \sup_{x_{i-1}\le x\le x_i} f(x)$. Then

$$\begin{aligned} U(P,f,\alpha) - U(P^*,f,\alpha) &= M_i[\alpha(x_i) - \alpha(x_{i-1}) - w_1[\alpha(x^*) - \alpha(x_{i-1})] - w_2[\alpha(x_i) - \alpha(x^*)] \\ &= (M_i - w_1)[\alpha(x^*) - \alpha(x_{i-1})] + (M_i - w_2)[\alpha(x_i) - \alpha(x^*)] \\ &\ge 0. \end{aligned}$$

The proof for the lower integrals is the same. ∎

Remark 1.10. Notice that for any partitions P_1, P_2, it follows with the common refinement $P^* = P_1 \cup P_2$ that

$$L(P_1,f,\alpha) \le L(P^*,f,\alpha) \le U(P^*,f,\alpha) \le U(P_2,f,\alpha).$$

Corollary 1.11

$$\underline{\int_a^b} f\, d\alpha \leq \overline{\int_a^b} f\, d\alpha.$$

We now arrive at a useful lemma relating integrability to being able to find partitions that allow the distance between upper and lower integrals to be arbitrarily small:

Lemma 1.12

$f \in \mathscr{R}(\alpha)$ if, and only if, $\forall \epsilon > 0$, there exists a partition P such that $U(P, f, \alpha) - L(P, f, \alpha) < \epsilon$.

Proof. Let $f \in \mathscr{R}$. Then there exists a partition P_1 such that $0 \leq U(P_1, f, \alpha) - \overline{\int_a^b} f\, d\alpha < \epsilon/2$. Similarly, there exists P_2 such that $0 \leq \underline{\int_a^b} f\, d\alpha - L(P_2, f, \alpha) < \epsilon/2$. (Notice that $f \in \mathscr{R}(\alpha)$, so $\overline{\int} f\, d\alpha = \underline{\int} f\, d\alpha = \int f\, d\alpha$.) Let $P = P_1 \cup P_2$ be the common refinement of P_1 and P_2. Then

$$U(P, f, \alpha) \leq U(P_1, f, \alpha) < \int f\, d\alpha + \epsilon/2 < L(P_2, f, \alpha) + \epsilon \leq L(P, f, \alpha) + \epsilon.$$

Now assume the converse. Recall that $\overline{\int} f\, d\alpha \leq U(P, f, \alpha)$ and $\underline{\int} f\, d\alpha \geq L(P, f, \alpha)$ for any partition P. Let $\epsilon > 0$. Then there exists a partition P such that

$$0 \leq \overline{\int} f\, d\alpha - \underline{\int} f\, d\alpha \leq U(P, f, \alpha) - L(P, f, \alpha) < \epsilon.$$

□

Now, let's introduce a bit of notation to make our lives easier. We can write that f is continuous on a metric space X as $f \in \mathscr{C}(X)$. Furthermore, we can improve upon our notation of integrability to write $f \in \mathscr{R}(\alpha, S)$ to mean that f is integrable on with respect to α over S.

Theorem 1.13

Let $f \in \mathscr{C}([a, b])$. Then $f \in \mathscr{R}(\alpha, [a, b])$.

Proof. Notice $[a, b]$ is compact. Thus f is uniformly continuous on $[a, b]$, so for $\epsilon > 0$, there exists $\delta > 0$ such that $|f(x) - f(y)| < \frac{\epsilon}{\alpha(b) - \alpha(a)}$ whenever $|x - y| < \delta$. Now, pick a partition P (with n elements) such that $\Delta x_j < \delta$ for all j. Then

$$\begin{aligned}
\overline{\int} f\, d\alpha - \underline{\int} f\, d\alpha \leq U(P, f, \alpha) - L(P, f, \alpha) &= \sum_{j=1}^{n} \sup_{I_j} f\, \Delta x_j - \sum_{j=1}^{n} \inf_{I_j} f\, \Delta x_j \\
&= \sum_{j=1}^{n} \left(\sup_{I_j} f - \inf_{I_j} f \right) \Delta x_j \\
&< \sum_{j=1}^{n} \frac{\epsilon}{\alpha(b) - \beta(a)} \Delta x_j \\
&= \frac{\epsilon}{\alpha(b) - \alpha(a)} (\alpha(b) - \alpha(a)) = \epsilon.
\end{aligned}$$

Since ϵ is arbitrary, the proof is complete. □

Remark 1.14. It turns out, we need not require that f is continuous on the entire interval; it suffices for f to be continuous *except at finitely many points*, with α continuous where f is not! Zoo wee mama!

Remark 1.15. In fact, the **Lebesgue Criterion for Riemann Integrability** states that

$$f \in \mathscr{R} \iff f \text{ is discontinuous on a set of } \textit{measure zero}!$$

(As a reminder, a set E has measure zero if for $\epsilon > 0$, there exists a collection of intervals $\{I_n\} \supset E$ such that $\sum_n \operatorname{diam}(I_n) < \epsilon$.)

Theorem 1.16 (The cooler Daniel)
If f is continuous at except finitely many points and α is continuous at the points of f's discontinuity, then $f \in \mathscr{R}(\alpha)$.

Proof. Let f be continuous except at finitely many points, say $\{x_0, \dots, x_n\}$. Because f is continuous at except *finitely* many points, we can let $M = |f|$.

Since the set of discontinuities $S = \{x_0, \dots, x_n\}$ is finite, α is uniformly continuous on S. It follows from the triangle inequality and the monotone increasing property of α that for all $\epsilon > 0$ there exists a $\delta > 0$ such that $\alpha(x_j + \delta) - \alpha(x_j - \delta) < \epsilon$.

We can always choose δ to be smaller, so without loss of generality, assume the set of $[x_j - \delta, x_j + \delta]$ is disjoint. Let $F = [a,b] \setminus \bigcup_{i=1}^n (x_j - \delta, x_j + \delta)$. F is compact, so for all $\epsilon > 0$, there exists $\delta' > 0$ such that $|f(u) - f(v)| < \epsilon$ for all $u, v \in F$ where $|u - v| < \delta'$.

We can now partition F into intervals I_j with $\Delta x_j < \delta'$. Let $J_i = [x_i - \delta, x_i + \delta]$. We can now partition $[a,b]$ into a partition P consisting of the I_i's and J_i's. Then

$$\begin{aligned}
\overline{\int} f\,d\alpha - \underline{\int} f\,d\alpha &\le \sum_j \sup_{I_j} f\,\Delta x_j - \sum_j \inf_{I_j} f\,\Delta x_j + \sum_j \sup_{J_j} f\,\Delta x_j - \sum_j \inf_{J_j} f\,d\alpha \\
&= \sum_j \left(\sup_{I_j} f - \inf_{I_j} f\right) \Delta x_j + \sum_j \left(\sup_{J_j} f - \inf_{J_j} f\right), \Delta x_j \\
&\le \epsilon \sum_j \Delta x_j + \sum_j 2M\epsilon \\
&= K\epsilon,
\end{aligned}$$

where $K \in \mathbb{R}$. □

Remark 1.17. What if we want to compose functions? Will their composition be integrable? Well it turns out that if the inner function is integrable, then the outer function being continuous on the range of the inner function is sufficient for integrability of the composition.

Theorem 1.18 (Integrability of composition of functions)
If f takes values in $[m, M]$ on $[a,b]$, $f \in \mathscr{R}(\alpha, [a,b])$, and ϕ continuous on $[m, M]$, then $\phi \circ f \in \mathscr{R}(\alpha, [a,b])$.

The proof for this theorem is pretty funny, so hang on.

Proof. ϕ is uniformly continuous on $[m, M]$ (why?) so for some $\epsilon > 0$ there exists a $\delta < \epsilon$ such that $|\phi(u) - \phi(v)| < \epsilon$ whenever $|u - v| < \delta$. Note that if we find a sufficiently small δ, then any value less than δ

also works so we can restrict ourselves to only working with $\delta < \epsilon$. It turns out, this restriction will become very useful later on!

Since $f \in \mathscr{R}(\alpha)$, it follows from **Lemma 1.12** that there exists a partition P such that

$$U(P, f, \alpha) - L(P, f, \alpha) < \delta^2.$$

For each $j = 1, \ldots, n$ (where $n = |P|$), if $\sup_{I_j} f - \inf_{I_j} f < \delta$, place $j \in A$. Otherwise place $j \in B$.

1. If $j \in A$, then $|\phi(f(x)) - \phi(f(y))| < \epsilon$, $x, y \in I_j$.
2. If $j \in B$, then $\sup_{I_j}(\phi \circ f) - \inf_{I_j}(\phi \circ f) \leq 2 \sup_{[m,M]} |\phi|$, and let's notate $K = \sup_{[m,M]} |\phi|$.

   But $U(P, f, \alpha) - L(P, f, \alpha) < \delta^2$, so

$$\sum_{j \in B} \delta \Delta\alpha_j \leq \sum \left(\sup_{I_j} f - \inf_{I_j} f \right) \Delta x_j \leq \delta^2.$$

   Dividing both sides by δ, we get $\sum_{j \in B} \Delta\alpha_j < \delta$.

Thus

$$\begin{aligned}
\overline{\int} \phi \circ f \, d\alpha - \underline{\int} \phi \circ f \, d\alpha &\leq U(P, \phi \circ f, \alpha) - L(P, \phi \circ f, \alpha) \\
&\leq \sum_{j=1}^{n} \epsilon \Delta\alpha_j + \sum_{j \in B} 2K \Delta\alpha_j \\
&\leq \epsilon(\alpha(b) - \alpha(a)) + 2K\delta \\
&< \epsilon(\alpha(b) - \alpha(a) + 2K).
\end{aligned}$$

□

> **Remark 1.19.** Note that we (stupidly, in the words of Jared Wunsch,) overcount in the third-to-last line of the extended equation; summing over *all* j instead of just $j \in A$.

> **Remark 1.20.** You've probably caught on to the style of proving a function is integrable: find a partition such that the difference $U - L$ is bounded above by an arbitrary ϵ.

We will now explore the properties of the integral, which pretty much agree with the intuition of someone who studied linear algebra and multivariate calculus with Aaron Peterson in MATH 291 @ Northwestern University:

1. The integral is *linear* over $\mathbb{R}$;
2. If a function bounds another from above, then the integral of the first will bound the integral of the second from above;
3. We can split integrals by an intermediate bound;
4. If the magnitude of a function is bounded by a finite number M, then the magnitude of the integral of that function will by bounded by the product of M and the width of the integral's bounds.
5. The sum of functions integrable with respect to different "clock speeds" is integrable with respect to the sum of their individual clock speeds. (Really pushing the metaphore here..)

Theorem 1.21 (Rudin 6.12)

1. If $f_1, f_2 \in \mathscr{R}(\alpha)$ then $f_1 + f_2 \in \mathscr{R}(\alpha)$, $cf \in \mathscr{R}(\alpha)$ for every $c \in \mathbb{R}$, and
$$\int_a^b (f_1 + f_2)\, d\alpha = \int_a^b f_1\, d\alpha + \int_a^b f_2\, d\alpha \quad \text{and} \quad \int_a^b cf\, d\alpha = c\int_a^b f\, d\alpha.$$
2. If $f_1(x) \le f_2(x)$ on $[a, b]$, then
$$\int_a^b f_1\, d\alpha \le \int_a^b f_2\, d\alpha.$$
3. If $f \in \mathscr{R}(\alpha)$ on $[a, b]$ and $a < c < b$, then $f \in \mathscr{R}(\alpha)$ on $[a, c]$ and $[c, b]$, and
$$\int_a^c f\, d\alpha + \int_c^b f\, d\alpha = \int_a^b f\, d\alpha.$$
4. If $f \in \mathscr{R}(\alpha)$ on $[a, b]$ and if $|f(x)| \le M$ on $[a, b]$, then
$$\left|\int_a^b f\, d\alpha\right| \le M[\alpha(b) - \alpha(a)];$$
5. If $f \in \mathscr{R}(\alpha_1)$ and $f \in \mathscr{R}(\alpha_2)$, then $f \in \mathscr{R}(\alpha_1 + \alpha_2)$ and
$$\int_a^b f\, d(\alpha_1 + \alpha_2) = \int_a^b f\, d\alpha_1 + \int_a^b f\, d\alpha_2;$$
If $f \in \mathscr{R}(\alpha)$ and $c \in \mathbb{R}^+$, then $f \in \mathscr{R}(c\alpha)$ and
$$\int_a^b f\, d(c\alpha) = c\int_a^b f\, d\alpha.$$

Proof. The proofs for each part are very similar, so we will only prove (1). However Wunsch messed up here so we'll skip this for now. A proof is in Rudin if you really want to read it. □

The previous theorem (**Rudin 6.12**) gives us a lot of power to determine the integrability of functions; we just need to be adept at manipulating expressions into sums and compositions of continuous functions. Thankfully, $x \to x^2$ is continuous and we have a useful identity to translate multiplication into addition:

A useful identity: $xy = \frac{1}{4}\left((x + y)^2 - (x - y)^2\right)$**.**

Theorem 1.22

Let $f, g \in \mathscr{R}(\alpha)$. Then

1. $fg \in \mathscr{R}(\alpha)$,
2. $|f| \in \mathscr{R}(\alpha)$, and
3. $\left|\int_a^b f\, d\alpha\right| \le \int_a^b |f|\, d\alpha$.

Proof. Notice that $f \pm g \in \mathscr{R}(\alpha)$, so $(f \pm g)^2 \in \mathscr{R}(\alpha)$. Then
$$fg = \frac{1}{4}\left((f + g)^2 - (f - g)^2\right) \in \mathscr{R}(\alpha).$$

Since $u \to u^2$ is continuous, $|f| \in \mathscr{R}(\alpha)$. Finally, there exists a $c = \pm 1$ where

$$\left| \int f \, d\alpha \right| = c \int f \, d\alpha = \int cf \, d\alpha \leq \int |f| \, d\alpha.$$

□

Example 1.23 (Heaviside Function)
We define the **Heaviside Function** as

$$H(x) = \begin{cases} 0 & x \leq 0 \\ 1 & x > 1 \end{cases}$$

If $a < 0 < b$ and f is continuous at $x = 0$, then $f \in \mathscr{R}([a,b], H)$ and $\int_a^b f \, dH = f(0)$.

Proof. Again, we choose a funny partition that will result in some clean shit: Let $P = \{x_0, x_1, x_2, x_3\}$, where $x_0 = a, x_1 = 0, x_3 = b$, and $x_2 \in (0, b)$. Then

$$\begin{aligned} U(P, f, H) &= \sup_{[a,0]} f \cdot (H(0) - H(a)) + \sup_{[0,x_2]} f \cdot (H(x_2) - H(0)) + \sup_{[x_2,b]} f \cdot (H(b) - H(x_2)) \\ &= \sup_{[a,0]} f \cdot (0 - 0) + \sup_{[0,x_2]} f \cdot (1 - 0) + \sup_{[x_2,b]} f \cdot (1 - 1) \\ &= \sup_{[0,x_2]} f. \end{aligned}$$

Similarly, $L(P, f, H) = \inf_{[0,x_2]} f$. Letting x_2 approach 0 from the right, notice that $U(P, f, H) \to f(0)^+$ and $L(P, f, H) \to f(0)^-$. So $\int_a^b f \, dH = 0$. □

Corollary 1.24 (Basically Heaviside, with **linearity!**)
Let $\alpha = \sum_{j=1}^{N} c_j H(x - s_j)$, $s \in [a,b]$, and $f \in \mathscr{C}([a,b])$. Then

$$\int_a^b f \, d\alpha = \sum_{i=1}^{N} c_j f(s_j)$$

Proof. Immediate by Theorem 1.20. □

Remark 1.25. Rudin extends α to be an infinite sum, but we don't need to get that crazy here...

Theorem 1.26
Say α' exists for all $x \in [a,b]$, α' is bounded, and f is *Riemann*-integrable (i.e. $f \in \mathscr{R}([a,b], x)$. Then $f \in \mathscr{R}(\alpha)$. If $\alpha' \in \mathscr{R}([a,b], x)$, then $\int_a^b f \, d\alpha = \int_a^b f(x)\alpha'(x) \, dx$.

Proof. Let α' be bounded for all $x \in [a,b]$ and $f \in \mathscr{R}([a,b], x)$. For $\epsilon > 0$, there exists a partition P of $[a,b]$ such that $U(P, f) - L(P, f) < \epsilon / K$, where $K = \sup_{[a,b]} \alpha'$. Now, note that for any $j \in P$, the Mean Value

Theorem implies there exists some $x_j^* \in [x_{j-1}, x_j]$ such that $\alpha(x_j) - \alpha(x_{j-1}) = \alpha'(x_j^*)\Delta x_j$. With this, notice

$$\begin{aligned} U(P,f,\alpha) - L(P,f,\alpha) &= \sum_{j\in P}\left(\sup_{I_j} f - \inf_{I_j} f\right)(\alpha(x_j) - \alpha(x_{j-1})) \\ &= \sum_{j\in P}(M_j - m_j)\alpha'(x_j^*)\Delta x_j \\ &\leq K\sum_{j\in P}(M_j - m_j)\Delta x_j \\ &< K\cdot \epsilon/K = \epsilon. \end{aligned}$$

Thus $f \in \mathscr{R}(\alpha)$. Now, let $\alpha' \in \mathscr{R}([a,b],x)$. Then for $\epsilon > 0$, there exists a partition P such that

1. $U(P,f) - L(P,f) < \epsilon$,
2. $U(P,f,\alpha') - L(P,f,\alpha') < \epsilon$,
3. $\sum_{j\in P}\left(\sup_{I_j}\alpha' - \inf_{I_j}\alpha'\right)\Delta x_j < \epsilon$, or $U(P,\alpha') - L(P,\alpha') < \epsilon$, and
4. $U(P,f,\alpha) - L(P,f,\alpha) < \epsilon$, shown earlier in the proof.

Now, for each $j \in P$ pick any $u_j \in I_j$. By the Mean Value Theorem, there exists some $x_j^* \in I_j$ such that

$$\sum_{j\in P} f(u_j)\Delta\alpha_j = \sum_{j\in P} f(u_j)\alpha'(x_j^*)\Delta x_j = \left(\sum_{j\in P} f(u_j)\alpha'(u_j)\Delta x_j\right) + \left(\sum_{j\in P} f(u_j)(\alpha'(x_j^*) - \alpha'(u_j))\Delta x_j\right).$$

Define, for sake of brevity, $A = \sum_{j\in P} f(u_j)\alpha'(u_j)\Delta x_j$ and $B = \sum_{j\in P} f(u_j)(\alpha'(x_j^*) - \alpha'(u_j))\Delta x_j$. (These are the last two sums in the previous equation.) Letting $M = \sup|f|$, we can bound B as follows:

$$|B| \leq \sum_{j\in P}\sup|f|\cdot\left(\sup_{I_j}\alpha' - \inf_{I_j}\alpha'\right)\Delta x_j \leq M\epsilon.$$

Since $L(P,f,\alpha) \leq A + B \leq U(P,f,\alpha)$, we have $L(P,f,\alpha) - M\epsilon \leq A \leq U(P,f,\alpha) + M\epsilon$, and thus

$$A - \int_a^b f\,d\alpha < \epsilon + M\epsilon.$$

Since A is a Riemann Sum, we also have

$$\left|A - \int_a^b f\alpha'\,dx\right| < \epsilon.$$

Combining all our Pokémon card inequality cards collected throughout the proof, we finally get

$$\left|\int_a^b f\,d\alpha - \int_a^b f\alpha'\,dx\right| < \epsilon + M\epsilon + \epsilon.$$

□

2 Integration and Differentiation

We will explore the dynamics between integration and differentiation, and as expected, the two act as quasi-inverse functions.

Theorem 2.1 (Fundamental Theorem of Calculus 1)
Let $f \in \mathscr{R}([a,b])$ and f be continuous at a point $x_0 \in [a,b]$. Then

$$f(x_0) = \frac{d}{dx}\int_a^x f(s)\,ds\bigg|_{x=x_0}.$$

Proof. Differentiating our funny integral, we have that

$$\frac{d}{dx}\int_a^x f(s)\,ds = \lim_{h\to 0}\frac{\int_0^{x_0+h} f(s)\,ds - \int_a^{x_0} f(s)\,ds}{h}.$$

We now have to inspect both right and left hand limits, but as the proofs for each case are analogous, we'll just look at the right hand limit: $h \to 0^+$. Since f is continuous at x_0, for $\epsilon > 0$, there exists $\delta > 0$ such that if $0 \le |y - x_0| < \delta$, then $|f(y) - f(x_0)| < \epsilon$. Since we're taking the limit as h approaches 0, we can limit our choice of h to only those with $h < \delta$. Pick any of them. Then $|f(s) - f(x_0)| < \epsilon|$ for all $s \in (x_0, x_0 + h)$.

We'll now employ a slick trick: since $f(x_0)$ is constant, we can write $f(x_0) = \frac{1}{h}\int_{x_0}^{x_0+h} ds$.

Then

$$\begin{aligned}
\left|\frac{1}{h}\int_{x_0}^{x_0+h} f(s)\,ds - f(x_0)\right| &= \left|\frac{1}{h}\int_{x_0}^{x_0+h} f(s)\,ds - \frac{1}{h}\int_{x_0}^{x_0+h} f(x_0)\,ds\right| \\
&= \left|\frac{1}{h}\int_{x_0}^{x_0+h} f(s) - f(x_0)\,ds\right| \\
&\le \frac{1}{h}\int_{x_0}^{x_0+h} |f(s) - f(x_0)|\,ds \\
&< \frac{1}{h}\int_{x_0}^{x_0+h} \epsilon\,ds \\
&= \epsilon.
\end{aligned}$$

□

Remark 2.2. Notice that $F(x) = \int_a^x f(s)\,ds$ is continuous on $[a,b]$.

Proof. By continuity, if $x < y$ then

$$|F(x) - F(y)| = \left|\int_x^y f(s)\,ds\right| \le \int_x^y |f(s)|\,ds \le \sup|f|(y-x).$$

So for $\epsilon > 0$, take $\delta = \epsilon/\sup|f|$. □

Continuous things have antiderivatives!

Theorem 2.3 (Fundamental Theorem of Calculus 2 (le célèbre))
Let $f \in \mathscr{R}([a,b])$, and there exist a differentiable F such that $F' = f$ on $[a,b]$. Then

$$\int_a^b f(s)\,ds = F(b) - F(a).$$

> **Remark 2.4.** Recall that integrable functions need not be continuous. (What's an example of a finitely discontinuous function that is is Riemann-integrable?) However, the large majority of commonly used integrable functions are continuous, so we'll prove this theorem for continuous functions first, and then weaken our hypothesis for the *real* proof.

Proof. *(naive)* Let f be continuous, and set $G(x) = \int_c^x f(s)\,ds$. By FTC1,

$$\frac{d}{dx}G(x) = f(x) = F'(x),$$

so $G(x) = F(x) + C$, where C is constant. Thus

$$\int_a^b f(s)\,ds = \int_c^b f(s)\,ds - \int_c^a f(s)\,ds = G(b) - G(a) = F(b) - F(a).$$

□

Proof. *(The real one..)* You know the drill: For $\epsilon > 0$, there exists a partition $P = \{x_0, x_1, \ldots, x_n\}$ such that $U(P, f) - L(P, f) < \epsilon$. By the Mean Value Theorem, there exists a $x_j^* \in [x_{j-1}, x_j]$ such that $F(x_j) - F(x_{j-1}) = f(x_j^*)\Delta x_j$. Then

$$F(b) - F(a) = \sum_{i=1}^{n} F(x_j) - F(x_{j-1}) = \sum_{i=1}^{n} f(x_j^*)\Delta x_j.$$

Thus,

$$\left| F(b) - F(a) - \int_a^b f(s)\,ds \right| < \epsilon.$$

□

We now approach the topic of **integration by parts**, which is often thought of as a computational integration tool by calculus students. As it turns out, it also carries much importance in analysis by showing that one can move derivatives around inside the integrand at the cost of a negative sign:

Theorem 2.5 (Integration by Parts)

Say F, G are differentiable functions, $F' = f$, $G' = g$, and $f, g \in \mathscr{R}$. Then

$$\int_a^b Fg\,dx = FG\Big|_a^b - \int_a^b fG\,dx.$$

Proof. By the chain rule, $(FG)' = Fg + fG$. Rearrange to isolate Fg and apply FTC2. □

Corollary 2.6

If $G = 0$ and a, b, then $\int_a^b FG'\,dx = -\int_a^b F'G\,dx$.

Finally, we introduce machinery that will facilitate changing the bounds of integration. In doing so, we must account for the "stretch" factor when stretching or shrinking the region of integration.

Theorem 2.7 (Change of Variables)
Let $\phi : [a,b] \to [A,B]$ be strictly increasing, where $\phi(a) = A$ and $\phi(b) = B$. Let ϕ be differentiable, with $\phi' \in \mathscr{R}$, and $f : [A,B] \to \mathbb{R}$ be continuous. Then

$$\int_a^b f(\phi(x))\phi'(x)\,dx = \int_A^B f(y)\,dy.$$

Proof. Set $F(x) = \int_A^x f(s)\,ds$. By FTC1, $F' = f$. By the chain rule, $\frac{d}{dx}F(\phi(x)) = f(\phi(x))\phi'(x)$. By FTC2, we have

$$\int_a^b f(\phi(x))\phi'(x)\,dx = \frac{d}{dx}\int_a^b F(\phi(x)) = F(\phi(x))\Big|_a^b = \int_A^B f(s)\,ds.$$

□

2.1 Appendix

There are some arguments utilized throughout the section worth having in writing for posterity:

Theorem 2.8
Recall that $F \in \mathscr{R}(\alpha)$ on $[a,b]$ if and only if for every $\epsilon > 0$ there exists a partition P such that

$$U(P,f,\alpha) - L(P,f,\alpha) < \epsilon. \tag{3}$$

It is now the case that

1. If (3) holds for some P and ϵ, then (3) holds (with the same ϵ) for *every* refinement of P.
2. If (3) holds for $P = \{x_0, \dots, x_n\}$ and if s_i, t_i are arbitrary points in $[x_{i-1}, x_i]$, then
$$\sum_{i=1}^n |f(s_i) - f(t_i)|\Delta\alpha_i < \epsilon.$$
3. If $f \in \mathscr{R}(\alpha)$ and the hypothesis of (2) hold, then
$$\left|\sum_{i=1}^n f(t_i)\Delta\alpha_i - \int_a^b f\,d\alpha\right| < \epsilon.$$

Proof. Immediate after noting $\sum_{i=1}^n |f(s_i) - f(t_i)|\Delta x_i \leq U(P,f,\alpha) - L(P,f,\alpha)$,

$$L(P,f,\alpha) \leq \sum f(t_i)\Delta\alpha_i \leq U(P,f,\alpha) \quad \text{and} L(P,f,\alpha) \leq \int f\,d\alpha \leq U(P,f,\alpha).$$

□

Theorem 2.9
If f is monotonic on $[a,b]$ and α is continuous on $[a,b]$, then $f \in \mathscr{R}(\alpha)$.

Proof. Monotonic functions are discontinuous at most countably many times. Countable subsets have measure zero, so we're done. Thanks, Lebesgue! □

3 Sequences and Series of Functions

Say $f_n : E \to \mathbb{C}$ are functions.

Definition 3.1. The sequence $\{f_n\}$ **converges pointwise** on E (to $f(x)$) if for all $x \in E$,

$$\lim_{n\to\infty} f_n(x) = f(x).$$

Question 3.2. What good properties of f_n might f inherit?

LMAO none. Anyway, an example:

Example 3.3

Let $f_n = \arctan(nx) \subset \mathbb{R}$. f_n converges pointwise to

$$f(x) = \begin{cases} -\pi/2 & x < 0 \\ 0 & x = 0 \\ \pi/2 & x > 0. \end{cases}$$

So f_n is infinitely differentiable, but $\lim_{n\to\infty} f(n)$ is not even continuous!

Definition 3.4. f_n **converges uniformly** on E if for all $\epsilon > 0$, there exists an $N \in \mathbb{N}$ such that if $n \geq N$ and $x \in E$, then

$$|f_n - f(x)| < \epsilon.$$

Remark 3.5. N is *independent* of x! (What else does this independence remind you of?)

Example 3.6 (3.3, revisited.)

Pick $\epsilon = \frac{\pi}{4}$. Given N, there exists $x > 0$ such that $\arctan(Nx) < \frac{\pi}{4}$ (since $\lim_{n\to\infty} \arctan(Nx) = 0$). Then

$$\left| f_N(x) - f(x)| < \frac{\pi}{4} \right|.$$

Theorem 3.7 (Cauchy Criterion for sequences of functions, kinda.)

$f_n \to f$ uniformly on E if, and only if, for every $\epsilon > 0$, there exists an $N \in \mathbb{N}$ such that for $m, n \geq N$, for all $x \in E$,

$$|f_m(x) - f_n(x)| < \epsilon.$$

Proof. Say $f_n \to f$ uniformly. Then for $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that for all $n \geq N$, $|f_n(x) - f(x)| < \frac{\epsilon}{2}$ for all x. Then for $m \geq N$, we obtain the same inequality and the proof follows directly from the triangle inequality. Conversely, suppose for every $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that for $m, n \geq N$, for all $x \in E$, $|f_m(x) - f_n(x)| < \frac{\epsilon}{2}$. Then for all x, if we fix m, $\{f_n(x)\}$ is Cauchy in $\mathbb{C}$. By completeness of $\mathbb{C}$, there exists a $f(x)$ such that $\lim_{n\to\infty} f_n(x) = f(x)$. By uniform convergence, for all $m \geq N$

$$\lim_{n\to\infty} |f_m(x) - f_n(x)| = |f_m(x) - f(x)| \leq \frac{\epsilon}{2}.$$

□

Theorem 3.8

If f_n are continuous functions on X, a metric space, and $f_n \to f$ uniformly on X, then f is continuous.

Proof. Fix $y \in X$. Then for $\epsilon > 0$:

1. There exists $N \in \mathbb{N}$ such that $|f_n(x) - f(x)| < \frac{\epsilon}{3}$ for all $n \geq N$ and $x \in X$.
2. If f_N continuous, there exists $\delta > 0$ such that $|f_N(x) - f_N(y)| < \frac{\epsilon}{3}$ whenever $d(x, y) < \delta$.

Now for all x such that $d(x, y) < \delta$, (1) and (2) give

$$|f(x) - f(y)| \leq |f(x) - f_N(x)| + |f_N(x) - f_N(y)| + |f_N(y) - f(y)| < \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3}.$$

□

We will now introduce an important notion of distance between functions, which will nicely lead to the notion of a metric space of continuous functions!

Definition 3.9. If $f : X \to \mathbb{C}$ is bounded, set $\|f\| = \sup |f|$, where X is a nonempty metric space. Let

$$\mathscr{C}(X) = \{f : X \to \mathbb{C} \mid f \text{ is continuous and bounded}\}.$$

For $f, g \in \mathscr{C}(X)$, we define $d_{\mathscr{C}}(f, g) = \|f - g\|$.

Of course, we wouldn't be defining a distance function if we didn't think we could use it as a metric...

Lemma 3.10

$\|f + g\| \leq \|f\| + \|g\|$.

Proof. As one would intuitively expect,

$$\|f + g\| = \sup |f(x) + g(x)| \leq \sup(|f(x)| + |g(x)|) \leq \sup |f| + \sup |g| = \|f\| + \|g\|.$$

□

Proposition 3.11

$d_{\mathscr{C}}$ is a metric on $\mathscr{C}(X)$.

Proof. Using the lemma,

- $d_{\mathscr{C}}$ is symmetric since $|f - g| = |g - f|$.
- $d_{\mathscr{C}}(f, g) = 0 \iff \sup |f - g| = 0 \iff |f(x) - g(x)| = 0 \quad \forall x \iff f = g$.
- $d(f, h) = \|f - h\| = \|f - g + g - h\| \leq \|f - g\| + \|g - h\| = d(f, g) + d(g, h)$.

□

Sick, so $(\mathscr{C}(X), d_{\mathscr{C}})$ is a metric space.

But what's the point of going through all this work to verify this fact?

Proposition 3.12 (Convergence in $(\mathscr{C}(X), d_{\mathscr{C}})$ is analogous to uniform convergence of functions.)
$f_n \to f$ in $\mathscr{C}(X)$ if, and only if, $f_n \to f$ uniformly.

Proof. Let $f_n \to f$ in $\mathscr{C}(X)$. Then for $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that

$$\sup |f(x) - f_n(x)| = \|f_n - f\| < \epsilon$$

whenever $n \geq N$. Thus, $|f(x) - f_n(x)| < \epsilon$ for all x. Conversely, let $f_n \to f$ uniformly. Then for $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that if $n \geq N$, $x \in X$, then

$$|f_n(x) - f(x)| < \frac{\epsilon}{2},$$

so

$$\|f_n - f\| = \sup |f_n(x) - f(x)| \leq \frac{\epsilon}{2}.$$

□

Theorem 3.13 (this seems important)
$\mathscr{C}(X)$ is complete.

Proof. Say $\{f_n\}$ is Cauchy in $\mathscr{C}(X)$. Then for $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that if $m, n \geq N$, then

$$\|f_n - f_m\| = \sup |f_n - f_m| < \epsilon.$$

Thus for all $x \in X$, $|f_n(x) - f_m(x)| < \epsilon$ and thus $\{f_n\}$ is uniformly convergent to some $f(x)$. By the previous proposition, f is continuous. Then there exists $N \in \mathbb{N}$ such that if $n \geq N$, $|f_n - f(x)| < 1$ for all x, and thus $|f(x)| < 1 + |f_n(x)|$, so f is bounded. □

Some notation: If $E \subset \mathbb{R}$, we write $\mathscr{C}^k(E) = \{f : E \to \mathbb{C} \mid f, f', \ldots, f^{(k)} \in \mathscr{C}(E)\}$. (Notice $\mathscr{C}^0(E) = \mathscr{C}(E)$.)

Theorem 3.14
Let α be nondecreasing on $[a, b] \subset \mathbb{R}$, $f_n \in \mathscr{R}(\alpha)$ for all n, and assume $f_n \to f$ uniformly. Then

1. $f \in \mathscr{R}(\alpha)$, and
2. $\lim_{n\to\infty} \int_a^b f_n \, d\alpha = \int_a^b f \, d\alpha$.

Proof. TBD. □

3.1 I missed class

wip: need to catch up on the lecture I missed (1/13/2023)

Definition 3.15. Two notions of **boundedness**:

1. A sequence of functions $\{f_n\} \in \mathbb{C}(X)$ is said to be **pointwise bounded** if for all $x \in X$, there exists $C(x)$ such that $|f_n(x)| \leq C(x)$ for all x.

2. A sequence of functions $\{f_n\} \in \mathbb{C}(X)$ is said to be **uniformly bounded** if for all $x \in X$, there exists a constant M such that $|f_n(x)| \le M$ for all x.

Definition 3.16. A family $\mathscr{F}$ of complex functions f defined on a set E of a metric space X is said to be **equicontinuous** on E if for every $\epsilon > 0$, there exists a $\delta > 0$ such that

$$|f(x) - f(y)| < \epsilon$$

whenever $d(x, y) < \delta$, where $x, y \in E$, $f \in \mathscr{F}$.

Indeed, it is the case that uniform convergence of sequences of functions and this notion of equicontinuity are related to one another.

Theorem 3.17

If K is compact set, $f_n \in \mathscr{C}(K)$ and $f_n \to f$ uniformly on K for $n \in \mathbb{N}$, then $\{f_n\}_{n\in\mathbb{N}}$ is equicontinuous.

Proof. For $\epsilon > 0$, there exists an $N \in \mathbb{N}$ such that for any $x \in K$,

$$|f_n(x) - f(x)| < \frac{\epsilon}{3}$$

whenever $n \ge N$. Furthermore, since $f_n \to f$ uniformly, f is continuous; because K is compact, we even have f uniformly continuous. So there exists a $\delta' > 0$ such that if $d(x, y) < \delta'$, then

$$|f(x) - f(y)| < \frac{\epsilon}{3}.$$

So, if $d(x, y) < \delta'$ and $m, n \ge N$, we have

$$|f_n(x) - f_n(y)| \le |f_n(x) - f(x)| + |f(x) - f(y)| + |f(y) - f_n(y)| < \epsilon.$$

Finally, $f_1, \ldots, f_N$ are continuous on a compact metric space, so they are uniformly continuous. Thus for each $1 \le j \le N$, there exists a $\delta_j > 0$ such that

$$|f_j(x) - f_j(y)| < \epsilon \quad (j = 1, \ldots, N).$$

Set $\delta = \min(\delta', \delta_1, \ldots, \delta_N)$, and we're done. □

Lemma 3.18

If K is compact, then there exists countable dense subset of K.

Proof. We can cover K with $\{B(x, \frac{1}{n}) \mid x \in K\}_{n\in\mathbb{N}}$. (Here, we shall abuse the notation $B(a, b)$ to represent the neighborhood of radius b centered at a, and x_j^n to represent the element x_j with corresponding neighborhood $B(x_j, \frac{1}{n})$.) For any $n \in \mathbb{N}$, there exists a finite subcover $\{B(x_j^n, \frac{1}{n}\}$ of balls of fixed radius $\frac{1}{n}$. Then for all $n \in \mathbb{N}$ and $y \in K$, there exists x_j^n such that $d(y, x_j^n) < \frac{1}{n}$. Take

$$S = \{x_j^n \mid n \in \mathbb{N}, j = 1, \ldots, N\}.$$

□

Lemma 3.19

Given countable S and uniformly bounded sequence of functions f_n, there exits subsequence converging to every element of S

Proof. Let $S = \{x_1, \dots\}$. Since the sequence $f_n(x_1)$ is bounded in $\mathbb{C}$ (i.e $\sup|f_n| \leq M$ for all $n \in \mathbb{N}$), there exists a subsequence $f_{n_j^1}$ such that $f_{n_j^1}(x_1)$ converges. (Abusing more notation, let the sequence of subindices n_j^n only consist of the sequence of indices n_j^m if $m < n$.) Similary, because $f_{n_j^1}(x_2) \leq M$ for all j, there exists a subsequence n_j^2 of n_j^1 such that $f_{n_j^2}(x_2)$ converges. Since $f_{n_j^2}$ is a subsequence of $f_{n_j^1}$, it also converges at x_1. Then $f_{n_j^3}$ converges at x_3 (and thus at x_1 and x_2); by induction, it can be seen that $f_{n_j^k}$ converges at $x_1, \dots, x_k$. Now, to obtain an explicit subsequence, we shall **diagonalize** (recall Cantor's diagonalization argument from proving the countable union of countable sets is countable!), by setting

$$g_j = f_{n_j^j}.$$

For all k, if $j \geq k$, then $f_{n_j^j}$ is a subsequence of $f_{n_j^k}$, so $g_j(x_1), \dots, g_j(x_j)$ converge as $j \to \infty$. □

Theorem 3.20 (Arzelà–Ascoli)

K compact, $f_n \in \mathscr{C}(K)$ for $n = 1, 2, \dots$. Assume $\{f_n\}_{n\in\mathbb{N}}$ bounded in $\mathscr{C}(K)$ (i.e. uniformly bounded) and equicontinuous. Then there exists a convergent subsequence in $\mathscr{C}(K)$ (i.e. uniformly convergent).

Proof. Using Lemma 3.18, pick a countable and dense $S \subset K$. Using Lemma 3.19, pick a subsequence g_j of f_n converging on x for all $x \in S$. If we show g_j is uniformly convergent on K, then we're done:

We can do so by showing g_j is uniformly Cauchy. It follows from equicontinuity that for all $\epsilon > 0$, there exists $\delta > 0$ such that if $d(x, y) < \delta$, then for all j,

$$|g_j(x) - g_j(y)| < \frac{\epsilon}{3}. \tag{4}$$

By Lemma 3.18, we can obtain, from the open cover $\{B(x_j, \delta) \mid x_j \in S\}$, a finite subcover $\{B(x, \delta) \mid x \in S_\delta\}$, where S_δ is finite. Now, since g_j converges to every element of S, there exists $N \in \mathbb{N}$ such that for all $x \in S_\delta$

$$|g_n(x) - g_m(x)| < \frac{\epsilon}{3} \tag{5}$$

whenever $m, n \geq N$. (Take $N = \max\{N_x \mid x \in S_\delta\}$.) Now, for all $i, j \geq N$, $y \in K$, there exists $x \in S_\delta$ such that $d(y, x) < \delta$, so (4) and (5) give

$$\begin{aligned} |g_i(y) - g_j(y)| &\leq |g_i(y) - g_i(x)| + |g_i(x) - g_j(x)| + |g_j(x) - g_j(y)| \\ &< \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} \\ &= \epsilon. \end{aligned}$$

□

Remark 3.21. Rudin displays the Arzelà–Ascoli theorem with slightly weaker conditions: he requires $\{f_n\}$ to be a pointwise bounded sequence of complex functions on a countable set. However, pointwise boundedness is almost never used, so we'll choose not to think about it.

> **Remark 3.22.** TODO: bounds on derivatives (or difference quotients) give equicontinuity. (MVT probably comes into play.....)

4 A Special Function

Definition 4.1. For $z \in \mathbb{C}$, we define a **power series** to be the infinite series

$$\sum_{n=0}^{\infty} c_n z^n = \lim_{n\to\infty} \sum_{n}^{N} c_n z^n. \tag{6}$$

Lemma 4.2 (Weierstrauss M-Test)

Consider the series of functions $\sum_{j=0}^{\infty} f_j(x)$. If there exists M_j such that $\sup|f(x)| \le M_j$ and $\sum M_j < \infty$, then $\sum_{j=0}^{\infty} f_j(x)$ converges uniformly.

Proof. Let $s_n = \sum_{j=0}^{n} f_j(x)$. For $m < n$,

$$|(s_n - s_m)(x)| \le \sum_{j=m+1}^{n} |f_j(x)| \le \sum_{j=m+1}^{n} M_j,$$

and if $\sum M_j$ converges, then for all $\epsilon > 0$, there exists an $N \in \mathbb{N}$ such that $\sum_{j=m+1}^{n} M_j < \epsilon$ if $m, n \ge N$. □

Theorem 4.3

There exists an $R \in [0, +\infty)$ such that

1. (6) converges absolutely for all $z \in \mathbb{C}$ with $|z| < R$, and converges uniformly on $\{z \mid |z| \le R'\}$ for all $0 \le R' < R$.
2. (6) diverges for $|z| > R$, with no information on R.

Proof. Recall that $\sum a_n$ converges if $\limsup |a_n|^{\frac{1}{n}} =: \alpha < 1$, and diverges if $\alpha > 1$. Now, notice that

$$\limsup |c_n z^n|^{\frac{1}{n}} = |z| \limsup |c_n|^{\frac{1}{n}} = \frac{|z|}{R}.$$

If $\frac{|z|}{R} < 1$, we get absolute convergence (and divergence if $\frac{|z|}{R} > 1$). We'll now check uniform convergence on the closure of $B(0, R')$: If $R' < R$, then $\frac{1}{R} < \frac{1}{R'}$. Pick s to be sandwiched such that $\frac{1}{R} < s < \frac{1}{R'}$. Now, $\limsup |c_n|^{\frac{1}{n}} = \frac{1}{R}$, so there exists $N \in \mathbb{N}$ such that if $n \ge N$, then $|c_n|^{\frac{1}{n}} \le s$. So, for $n \ge N$, if $|z| < R'$, observe

$$|c_n z^n| \le s_n (R')^n = (sR')^n < 1.$$

Defining $\beta := sR'$, we have $|c_n z^n| < \beta^n$, so $\beta < 1$ and we get uniform convergence for $|z| \le R'$ by the M-test! □

Theorem 4.4

Fix a series $\sum_n^\infty c_n z^n$ with radius of convergence R. For $|z| < R$, let

$$f(z) = \sum_{n=0}^{\infty} c_n z^n.$$

Then the derivative evaluated at real $z \in \mathbb{R}$

$$f'(z) = \sum_{n=1}^{\infty} c_n n z^{n-1} = \sum_{n=0}^{\infty} c_{k+1}(k+1)z^k$$

has the same radius of convergence R.

Proof. It suffices to verify that both $f(z)$ and $f'(z)$ have the same radius of convergence. Notice that $\limsup_{n\to\infty} |n \cdot c_n|^{\frac{1}{n}} = \limsup n^{\frac{1}{n}} |c_n|^{\frac{1}{n}} = \frac{1}{R_{f'}}$, so

$$\frac{1}{R_{f'}} = \lim_{n\to\infty} n^{\frac{1}{n}} \cdot \limsup_{n\to\infty} |c_n|^{\frac{1}{n}} = 1 \cdot \frac{1}{R}.$$

So for any $R' < R$, $\sum n c_n z^{n-1} = f'(z)$ converges uniformly if $|z| \le R'$. □

Corollary 4.5

For all k, $f^{(k)}(x) = \sum_{n=k}^{\infty} c_n n(n+1)\cdots(n-k+1)z^{n-k}$ for $|x| < R$.

Corollary 4.6

$f^{(k)}(0) = k! c_k$.

Remark 4.7. By the last corollary, we get that the infinite series we've been working with were Taylor series for f.

What about the converse? Can we represent every function by some Taylor series?

Example 4.8

Let

$$f(x) = \begin{cases} 0 & x \le 0 \\ e^{-\frac{1}{x^2}} & x > 0 \end{cases}.$$

Then $f \in \mathscr{C}^\infty$, but $f^{(j)}(0) = 0$ for all j, so f is obviously not equal to its Taylor series.

Then what functions work?

Definition 4.9. A function that can be represented by a series $\sum c_n z^n$ is said to be an **analytic function**.

Example 4.10

The function $\frac{1}{1-z} = \sum_{n=0}^{\infty} z^n$ converges if $|z| < 1$ and diverges if $|z| > 1$.

4.1 The exponential function

Example 4.11

Define
$$E(z)=\sum_{n=0}^{\infty}\frac{z^n}{n!}.$$

By the ratio test, $\lim_{n\to\infty}\left|\frac{z^{n+1}}{(n+1)!}\right|\left|\frac{z^n}{n!}\right| = \lim_{n\to\infty}\frac{|z|}{n+1} = 0$ for all $z \in \mathbb{C}$. Thus $E(z)$ converges uniformly on any disc. Furthermore, notice that

$$E'(x)=\sum_{n=0}^{\infty}\frac{nz^{n-1}}{n!}=\sum_{n=1}^{\infty}\frac{z^n}{n!}=E(x),$$

and $E(0) = 1$. Taking $z, w \in \mathbb{C}$, we get from some sad algebraic manipulation that

$$\begin{aligned}E(z)E(w) &= \sum_{n=0}^{\infty}\frac{z^n}{n!}\sum_{k=0}^{\infty}\frac{w^k}{k!}\\ &= \sum_{m=0}^{\infty}\sum_{n+k=m}\frac{z^n w^k}{n!k!}\\ &= \sum_{m=0}^{\infty}\left(\sum_{k=0}^{m}\frac{z^{m-k}w^k}{(m-k)!k!}\right)\\ &= \sum_{m=0}^{\infty}\frac{1}{m!}\sum_{k=0}^{m}\frac{m!}{k!(m-k)!}z^{m-k}w^k\\ &= \sum_{m=0}^{k}\frac{1}{m!}(z+w)^m\\ &= E(z+w).\end{aligned}$$

Corollary 4.12

$E(-x) = \frac{1}{E(x)}$, $x \in \mathbb{C}$. Thus $E(z) \neq 0$ for all $z \in \mathbb{C}$.

Definition 4.13.

$$e = E(1) = \sum_{n=0}^{\infty}\frac{1}{n!}.$$

Notice

1. $E(w) = E(1+\cdots+1) = E(1)^n = e^n$ for all $n \in \mathbb{N}$,
2. $E\left(\frac{p}{q}\right)^q = E\left(\frac{p}{q}+\cdots+\frac{p}{q}\right) = E(p) = e^p$, so $E\left(\frac{p}{q}\right) = e^{p/q}$.

Proposition 4.14

$E(x)$ is strictly increasing on $\mathbb{R}$.

Proof. $E'(x) = E(x) > 0$ for all x. □

Remark 4.15. e^x was also defined (at some point) as $\sup_{p/q<x} E\left(\frac{p}{q}\right) = E(x)$.

Example 4.16

Note that e^x is the unique function characterized by satisfying the following ordinary differential equation:

$$\begin{cases} E'(x) = E(x) & \forall x \in \mathbb{R} \\ E(0) = 1 \end{cases}$$

We can further derive the aforementioned properties of e^x with this formulation, e.g. both $E(z+tw)$ and $E(z)E(tw)$ satisfy $\frac{d}{dt}g = wg$:

$$\frac{d}{dt}E(z+tw) = wE(z+tw)$$
$$\frac{d}{dt}E(z)E(tw) = E(z) + wE'(tw)$$

, $E(z+w) = E(z)E(tw)$.

4.2 The Natural Logarithm

Since $e > 1$, it follows that $e^n \to \infty$ as $n \to \infty$ and $e^{-x} \to 0$ as $x \to \infty$. Moreover, $x \mapsto E(x)$ maps $\mathbb{R} \to (0,\infty)$, and is a one-to-one and onto. Thus, there exists an inverse function $L : (0,\infty) \to \mathbb{R}$, $E \circ L = Id$. By properties of the derivative,

$$L'(y) = \frac{1}{E(L(y))} = \frac{1}{L^{-1}(L(y))} = \frac{1}{y}$$

so by the Fundamental Theorem of Calculus, $L(y) = \int_1^y \frac{1}{s}\,ds + C$. Since $E(0) = 1$, $L(1) = 0$ and thus $C = 0$.

Definition 4.17. We define the natural logarithm function to be inverse of e^x:

$$\log(y) = L(y) = \int_1^y \frac{1}{s}\,ds.$$

4.3 feat. sine & cosine

Recall that $E(z) = \sum_{j=0}^{\infty} \frac{z^j}{j!}$, so

$$E(ix) = \sum_{j=0}^{\infty} \frac{(ix)^j}{j!} = \sum_{j=0}^{\infty} \frac{i^j x^j}{j!}. \tag{7}$$

Definition 4.18. Notice $C'(x) = -S(x)$ and $C(0) = 1$; $S'(x) = C(x)$ and $S(0) = 0$. As you've probably guessed, $C(x) = \cos(x)$ and $S(x) = \sin(x)$.

Theorem 4.19 (Sum and Difference Rule)

aka SACB, CASB or *SINE COSINE COSINE SINE, COSINE, COSINE, SINE SINE. SINE THE SAME, COSINE CHANGE, nanananananana* -Paul Zaclin, Fall 2018, Calculus 1.

$$\sin(x \pm y) = \sin(x)\cos(y) \pm \cos(x)\sin(y)$$
$$\cos(x \pm y)\cos(x)\cos(y) \mp \sin(x)\sin(y)$$

Proof. Let $E(ix) = C(x) + iS(x)$, where $C(x), S(x) \in \mathbb{R}$. Then, (7) implies that

$$C'(x) + iS'(x) = i(C(x) + iS(x)) = iC(x) - S(x).$$

Thus

$$E(i(x+y)) = E(ix)E(iy) = C(x+y) + iS(x+y) = (C(x) + iS(x))(C(y) + iS(y)),$$

so

$$C(x+y) = (C(x)C(y) - S(x)S(y))$$
$$S(x+y) = (S(x)C(y) + C(x)S(y))$$

□

Theorem 4.20 (Pythagorean Identity)

$$\cos^2(x) + \sin^2(x) = 1.$$

Proof. Notice

$$\overline{e^{ix}} = \overline{\sum \frac{(ix)^j}{j!}} = \sum \frac{(-ix)^j}{j!} = e^{-ix} = \frac{1}{e^{ix}}.$$

Thus $1 = e^{ix}e^{-ix} = e^{ix}\overline{e^{-ix}} = |e^{ix}|^2 = \cos^2(x) + \sin^2(x)$. □

5 Fourier Analysis

The year is 1807. France just had a revolution or something. Enter Joseph Fourier. This guy is living the life: Napolean likes him and makes him do math n stuff. Of specific interest to us is Fourier's study of heat equations, which led him to an extremely important proposition:

Proposition 5.1

Any nice* 2π-periodic function can be written as the **Fourier Series**

$$f(x) = \sum_{n=-\infty}^{\infty} a_n e^{inx}$$

for some coefficients $a_n \in \mathbb{C}$. (Barring exponentials, he would've written the series as

$$f(x) = \sum_{j=0}^{\infty} a_j \cos(jx) + \sum_{j=1}^{\infty} b_j \sin(jx)$$

but the bookkeeping with this representation is a wee bit more annoying than the first.)

Remark 5.2. Notice that $e^{in(x+2\pi)} = e^{inx}$ for all x so $\sum a_n e^{inx}$ is 2π-periodic (if the series converges).

Definition 5.3. We shall define, for convenience, the function

$$e_n(x) = \frac{1}{\sqrt{2\pi}} e^{inx}.$$

Say $f(\alpha) = \sum_{-\infty}^{\infty} a_n e_n(x)$. We say convergence of f is **uniform** if

$$\lim_{N\to\infty} \sum_{n=-N}^{n=N} a_n e_n(x) = f(x)$$

with uniform convergence.

Question 5.4. What values a_n are nice?

This will motivate further exploration into this silly goofy world.

Definition 5.5. Let $f, g \in \mathscr{R}([0, 2\pi])$. We define the following the notation:

$$< f, g >= \int_0^{2\pi} f(x)\overline{g(x)}\, dx.$$

As an analogy, imagine we go from the continuous world of $[0, 2\pi]$ to the discrete world of n points: $x_1, \ldots, x_n$. Then $\sum_{i=1}^n f(x_i)\overline{g(x_i)}$ is the dot (or inner) product on $\mathbb{C}$. Thus we can think of $< \vec{v}, \vec{w} >$ to be an infinitessimal notion of an inner product! Moreover, recall the $\mathscr{L}^2$-norm is defined

$$\|f\|_2^2 = \left\{ \int |f(x)|^2\, dx \right\}^{1/2}$$

Turns, out $\|f\|^2 = \int_0^{2\pi} |f(x)|^2\, dx =< f, f >$. Hmmmm. Okay, so that doesn't really mean anything to me. But let's keep digging: For $n \in \mathbb{N}$,

$$< e_n(x), e_j(x) > \quad = \frac{1}{2\pi} \int_0^{2\pi} e^{inx} e^{-ijx}\, dx = 1 \cdot \delta_{nj},$$

where $d_{nj} = \begin{cases} 1 & j = n \\ 0 & j \neq n \end{cases}$. So

$$< f, e_j > \quad = \quad < \sum_{-\infty}^{\infty} a_n e_n(x), e_j(x) > \quad = \sum_{-\infty}^{\infty} a_n < e_n(x), e_j(x) > \quad = a_j.$$

Proposition 5.6 (a conjecture, really)
For nice** f,

$$f(x) = \sum_{n=-\infty}^{\infty} \hat{f}(n) e^{inx}$$

with

$$\hat{f}(n) =< f, e_n >= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-inx}\, dx.$$

****Sadly, it turns out bounded and uniformly continuous does not suffice. Well, at least it almost does..*

Remark 5.7. The bounds of integration for $\hat{f}(n)$ are often 0 to 2π instead. Since f is periodic, it doesn't really matter.

5.1 Priodicity of sine and cosine

Before proceeding, it is first imperative that we establish the **periodicity of sine and cosine**. Recall that $E(ix) = C(x) + iS(x)$, where C is cosine, S is sine, and $C(0) = 1, S(0) = 0$.

Lemma 5.8

$C(x) > 1 - x$ for $x > 0$.

Proof. Let $f(x) = C(x) + x - 1$. (We want to show that $f(x) > 0$ if $x > 0$.) Notice $f'(x) = -S(x) + 1$, so $f'(0) = 1$ and $f(x) \geq 0$ for all x. Thus $f(x) > 0$ on $x \in (0, \epsilon)$ for $\epsilon > 0$. Thus, for sufficientlys small h, $\lim_{h\to 0} \frac{f(h)}{h} = 1$ and so $f(x) \geq \frac{1}{2}h$. For sake of contradiction, suppose there exists $a > 0$ with $f(a) \leq 0$. Then, by the Mean Value Theorem, there exists $b \in [\epsilon, a]$ with $f'(b) < 0$, a contradiction since $S(x) \leq 1$. □

Set $T(x) = \frac{S(x)}{C(X)}$ on $[0,1]$, since $C(X) > 0$ on $[0,1]$. Notice $T'(X) = \frac{1}{C(X)^2} = \frac{S(X)^2 + C(X)^2}{C(X)^2} = T(X)^2 + 1 \geq 1$. Thus $T(1) \geq 1$, so there exists $a \in [0,1]$ such that $T(a) = 1$.

Definition 5.9. Define $\pi = 4a$.

Then $T\left(\frac{\pi}{4}\right) = 1$, so $C\left(\frac{\pi}{4}\right) = S\left(\frac{\pi}{4}\right) > 0$ since $C(X) > 0$ on $[0,1]$. Since $C^2 + S^2 = 1$, $C\left(\frac{\pi}{4}\right) = S\left(\frac{\pi}{4}\right) = \frac{\sqrt{2}}{2}$,and thus

$$e^{i\frac{\pi}{4}} = \left(\frac{\sqrt{2}}{2} + i\frac{\sqrt{2}}{2}\right) \Rightarrow e^{i\pi} = -1.$$

So for all $x \in \mathbb{R}$,

$$e^{i(2\pi + x)} = e^{ix}e^{2\pi i} = e^{ix},$$

and thus

$$C(X + 2\pi) = C(X), \quad S(X + 2\pi) = S(X).$$

So sine and cosine are periodic! Whew, that was a lot.

5.2 The Direchlet and Fejer Kernel

Okay, so we've defined the notation of $e_n = \frac{1}{\sqrt{2\pi}} e^{inx}$, $n \in \mathbb{Z}$, the operation $< f, g >= \int_0^{2\pi} f(x)\overline{g(x)}\, dx$, and the set

$$\mathscr{C}^k(S^1) = \{f \in \mathscr{C}^k(\mathbb{R}) \mid f(x + 2\pi) = f(x) \quad \forall x\}.$$

It would be really nice if for all $f \in \mathscr{C}^k(S^1)$, $f(x) = \sum_{n=-\infty}^{\infty} a_n c_n(x)$. If all goes well, then

$$< f, e_j >= \sum_{n=-\infty}^{\infty} a_n < e_n, e_j >= a_j.$$

We want $a_j =< f, e_j >= \hat{f}(j)$. In other words, we want

$$f = \sum_{j=\infty}^{\infty} \hat{f}(j) e_j = \sum_{-\infty}^{\infty} < f, e_j > e_j = \lim_{N\to\infty} \sum_{j=-N}^{N} < f, e_j > e_j.$$

Definition 5.10. We define the **Direchlet Kernel** to be

$$D_N(x) = \frac{1}{2\pi} \sum_{j=-N}^{N} e^{ijx}.$$

> **Remark 5.11.** Here, "kernel" has no relationship to the notion of kerenels of a vector space, or the like. In this case, a kernel is something that is synechdoche with its integral.

Then

$$\begin{aligned} f &= \lim_{N\to\infty} \sum_{j=-N}^{N} < f, e_j > e_j \\ &= \lim_{N\to\infty} \sum_{j=-N}^{N} \left(\int_0^{2\pi} f(y)\overline{e_j(y)} \right) e_j(x) \\ &= \lim_{N\to\infty} \sum_{j=-N}^{N} \left(\int_0^{2\pi} f(y) \frac{1}{2\pi} e^{ij(x-y)}\, dy \right) \\ &= \lim_{N\to\infty} \int_0^{2\pi} \frac{1}{2\pi} f(y) \sum_{j=-N}^{N} e^{ij(x-y)}\, dy \\ &= \lim_{N\to\infty} \int_0^{2\pi} f(y) D_N(x-y)\, dy. \end{aligned}$$

Definition 5.12. Let $s_N(x) = \sum_{j=-N}^{N} \hat{f}(y) e_j(x)$.

We've just shown that $s_N(x) = \int_0^{2\pi} f(y) D_N(x-y)\, dy$, which if you've taken probability might recognize a convolution buried in notational junk. Wait, a what?

Definition 5.13. Let $f, g \in \mathscr{C}(S^1)$. The **convolution** of f and g is defined as $(f \star g)(x) = \int f(y) g(x-y)\, dy$

Well, this is cool, but how can it be applied?

Question 5.14. Let $f, g \in \mathscr{C}(S^1)$ and $y = s - z$. Show

$$\int f(y) g(x-y)\, dy = \int f(z-x) g(z)\, dz.$$

(In other words, show $(f \star g)(x) = (g \star f)(x)$.)

> **Remark 5.15.** If we consider the convolution $f \star D_N(x) = \int f(x-y) D_N(y)\, dy$, notice that it is a quasi-"weighted average" function, that smooths out f. Furthermore, the N^{th} partial sum of a Fourier series:
>
> $$s_N(x) = (f \star D_N)(x),$$
>
> where as a reminder, $D_N(x) = \frac{1}{2\pi} \sum_{j=-N}^{N} e^{ijx} = \frac{1}{2\pi} \sum_{j=-N}^{N} (e^{ix^j})$.

Notice that

$$e^{ix} D_N(x) = D_N(x) + \frac{1}{2\pi} e^{i(N+1)x} - \frac{1}{2\pi} e^{-iNx},$$

so

$$D_N(x)(e^{ix} - 1) = \frac{1}{2\pi} (e^{i(N+1)x} - e^{-i(Nx)}),$$

and thus

$$\begin{aligned} D_N(x) &= \frac{1}{2\pi}\frac{e^{i(N+1)x} - e^{-iNx}}{e^{ix}-1} \\ &= \frac{1}{2\pi}\frac{e^{i(N+\frac{1}{2})x} - e^{-i(N+\frac{1}{2})x}}{e^{i\frac{x}{2}} - e^{-i\frac{x}{2}}} \\ &= \frac{1}{2\pi}\frac{2i\sin(N+\frac{1}{2}x)}{2i\sin(\frac{1}{2}x)}) \\ &= \frac{1}{2\pi}\frac{\sin(N+\frac{1}{2}x)}{\sin(\frac{1}{2}x)} \end{aligned}$$

Definition 5.16. We say $s_n(x)$ converges Cesaro if $\sigma_N(x) = \frac{1}{N}(s_0 + s_1 + \cdots + s_{N-1})(x) \to f(x)$.

Theorem 5.17 (Fejér)
If $f \in \mathscr{C}^0(S^1)$, then $\sigma_N(x) \to f(x)$ uniformly.

Proof. It's coming, don't you worry. We just need some more machinery first... □

Definition 5.18. We define the **Fejér Kernel** to be $F_N(x) = \frac{1}{N}(D_0 + \cdots + D_{N-1})$.

Remark 5.19. Notice that we can compute the following equality:

$$\begin{aligned} \sigma_N(x) &= \frac{1}{N}(s_0 + s_1 + \cdots + s_{N-1})(x) \\ &= \frac{1}{N}\left(\int_0^{2\pi} f(y)D_0(x-y)\,dy + \cdots + \int_0^{2\pi} f(y)D_{N-1}(x-y)\,dy\right) \\ &= \int_0^{2\pi} f(y)\left(\frac{1}{N}(D_0 + \cdots + D_{N-1})(x-y)\right)dy \\ &= \int_0^{2\pi} f(y)F_N(x-y)\,dy. \end{aligned}$$

Next, recall that $e^{ia} - e^{-ia} = 2\pi\sin(a)$, so

$$\begin{aligned} F_N(x) &= \frac{1}{2\pi N}\sum_{j=0}^{N-1}\frac{\sin\left(j+\frac{1}{2}\right)x}{\sin\left(\frac{1}{2}x\right)} \\ &= \frac{1}{2\pi N\sin\left(\frac{1}{2}x\right)}\Im\sum_{j=0}^{N-1} e^{i(j+\frac{1}{2})x} \\ &= \frac{1}{2\pi N\sin\left(\frac{1}{2}x\right)}\Im\left\{e^{i\frac{x}{2}}\sum_{j=0}^{N-1}(e^{ix})^j\right\} \\ &= \frac{1}{2\pi N\sin\left(\frac{1}{2}x\right)}\Im\left\{e^{i\frac{x}{2}}\left(\frac{e^{iNx}-1}{e^{ix}-1}\right)\right\} \\ &= \frac{1}{2\pi N\sin\left(\frac{1}{2}x\right)}\Im\left\{\frac{e^{i\frac{N}{2}x}\left(e^{i\frac{N}{2}x} - e^{-i\frac{N}{2}x}\right)}{e^{i\frac{x}{2}} - e^{-i\frac{x}{2}}}\right\} \\ &= \frac{1}{2\pi N\sin\left(\frac{1}{2}x\right)}\Im\left(\cos\frac{N}{2} + i\sin\frac{N}{2}x\right)\left(\frac{\sin\left(\frac{N}{2}x\right)}{\sin(\frac{1}{2}x)}\right) = \frac{1}{2\pi N}\frac{\sin^2\left(\frac{N}{2}x\right)}{\sin^2\left(\frac{1}{2}x\right)} \end{aligned}$$

and thus $F_N(x) \to 0$ as N grows to infinity. On the other hand, $D_N(x)$ oscillates forever.

Lemma 5.20
The following hold:

1. $\int_0^{2\pi} D_N(x)\,dx = 1$ for all N,
2. $\int_0^{2\pi} F_N(x)\,dx = 1$ for all N, and
3. $F_N(x) \geq 0$ for all x. For $\delta > 0$, $F_N(x) \to 0$ uniformly as $N \to \infty$ on $x \in [-\pi, \pi] \setminus (-\delta, \delta)$.

Proof. Recall that $D_N = \frac{1}{\sqrt{2\pi}} \sum_{j=-N}^{N} e_j(x) = \sum_{j=-N}^{N} \frac{e^{ijx}}{2\pi}$, so

$$\int_0^{2\pi} D_N(x)\,dx = \sum_{j=-N}^{N} \frac{1}{2\pi} \int_0^{2\pi} e^{ijx} = 1.$$

Furthermore,

$$\int_0^{2\pi} F_N(x)\,dx = \frac{1}{N} \int_0^{2\pi} \sum_{j=0}^{N-1} D_j(x)\,dx = \frac{1}{N} \cdot N = 1.$$

$F_N(x) \geq 0$ since $F_N(x) = \left(\frac{1}{\sqrt{2\pi}} \frac{\sin(\frac{N}{2}x)}{\sin(\frac{1}{2}x)} \right)^2$. On $[\delta, \pi)$, notice $\sin \frac{1}{2}x$ is increasing (take the derivative!), so for sufficiently small $\delta > 0$,

$$\left| \frac{\sin \frac{N}{2}x}{\sin \frac{1}{2}x} \right| \leq \frac{|\sin \frac{N}{2}x|}{|\sin \frac{\delta}{2}|} \leq \frac{1}{\sin \frac{\delta}{2}}.$$

So on $[\delta, \pi]$, $F_N(x) \leq \frac{1}{N} \frac{1}{\sin^2 \frac{\delta}{2}}$. Likewise, on $[-\pi, -\delta]$, since F_N is even, $0 \leq F_N(x) \leq \frac{1}{N} \frac{1}{\sin^2 \frac{\delta}{2}}$ on $[-\pi, pi] \setminus (\delta, \delta)$, and this tends to 0 as N goes to infinity. □

Lemma 5.21
Let the conditions for the Fejer Theorem hold. Then for $\epsilon > 0$, there exists a $\delta >$ such that there exists an $N_0 \in \mathbb{N}$ where if $N \geq N_0$, then

$$|f(x) - \sigma_N(x)| \leq \epsilon.$$

Proof. For $\epsilon > 0$, continuity of f gives $|f(x) - f(y)| < \frac{\epsilon}{2}$ whenever $|x - y| < \delta$. Given this δ, there exist an N_0 such that if $N \geq N_0$, then

$$0 \leq F_N(x) < \frac{\epsilon}{4 \sup |f| 2\pi}$$

on $[-\pi, \pi] \setminus (-\delta, \delta)$. Then for all $N \geq N_0$,

$$\begin{aligned}
|f(x) - \sigma_N(x)| &\leq \int_{-\pi}^{\pi} |f(x) - f(y)| F_N(x - y)\,dy \\
&= \int_{x-\delta}^{x+\delta} |f(x) - f(y)| F_N(x - y)\,dy + \int_{[x-\pi, x+\pi] \setminus (x-\delta, x+\delta)} |f(x) - f(y)| F_N(x - y)\,dy \\
&\leq \int_{x-\delta}^{x+\delta} \frac{\epsilon}{2} F_N(x - y)\,dy + \int_{[x-\pi, x+\pi] \setminus (x-\delta, x+\delta)} 2 \sup |f| \frac{\epsilon}{4 \sup |f| 2\pi}\,dy \\
&\leq \int_{-\pi}^{\pi} \frac{\epsilon}{2} F_N(x - y)\,dy + 2\pi \cdot 2 \sup |f| \cdot \frac{\epsilon}{4 \sup |f| 2\pi} \\
&= \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.
\end{aligned}$$

□

Theorem 5.22 ((Fejer, again))

If $f \in \mathscr{C}^0(S^1)$, then $\int_0^{2\pi} f(y)F_N(x-y)\,dy = \sigma_N(x) \to f(x)$ uniformly on $\mathbb{R}$. (Indeed, if $f \in \mathscr{R}([0,2\pi])$ is continuous at $a \in \mathbb{R}$, then $\sigma_N(a) \to f(a)$.)

Proof. We can do the clever trick of multiplying by a funny representation of 1: $f(x) = \int_0^{2\pi} f(x)F_N(x-y)\,dy$. Then

$$\begin{aligned} |f(x) - \sigma_N(x)| &= \left| f(x) - \int_0^{2\pi} f(y)F_N(x-y)\,dy \right| \\ &= \left| \int_0^{2\pi} (f(x) - f(y))F_N(x-y)\,dy \right| \\ &\le \int_0^{2\pi} |f(x) - f(y)|F_N(x-y)\,dy \\ &< \epsilon. \end{aligned}$$

□

Corollary 5.23

Trigonometric polynomials are dense in $\mathscr{C}^0(S^1)$. (Trigonometric polyomials are finite Fourier series of the form $\sum_{j=-N}^{N} b_j e^{ijx} = \sum_{j=-N}^{N} b_j(\cos(x) + i\sin(x))^j$ for some N, $b_{-N_j}, \ldots, b_{N_j}$.)

Corollary 5.24 (Weierstrass Approximation Theorem)

Let $f : [0,1] \to \mathbb{R}$ be continuous. Then there is a sequence of polynomials $P_n(x)$ such that

$$\lim_{n\to\infty} \sup_{x\in[0,1]} |P_n(x) - f(x)| = 0$$

Remark 5.25. This holds not just for $[0,1]$, but also any closed interval $[a,b]$.

Lemma 5.26 (This is important!)

If $f \in \mathscr{C}^k(S^1)$, then $|\hat{f}(n)| \le \frac{C_f}{|n|^k}$, $n \neq 0$, and $\widehat{D^k f}(n) = n^k \hat{f}(n)$.

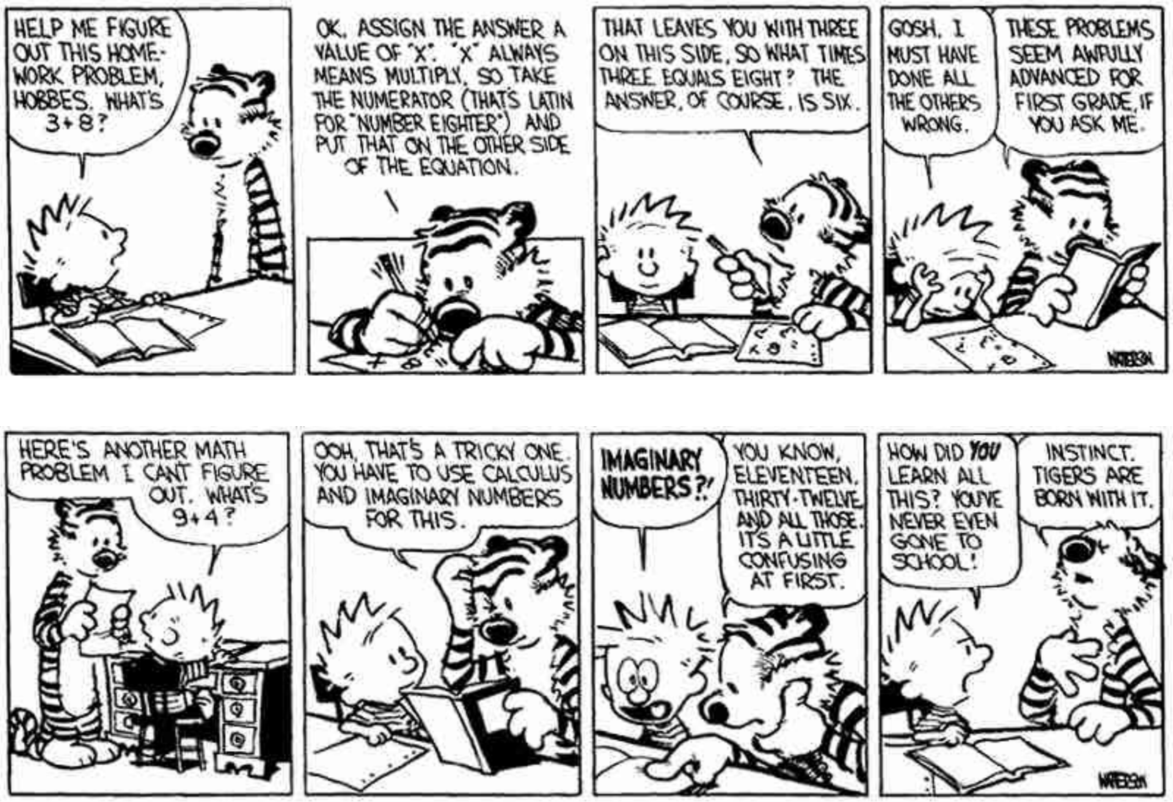
Proof. Since $f \in \mathscr{R}$, $|\hat{f}(n)| = |\int_0^{2\pi} f(x)\frac{1}{\sqrt{2\pi}}e^{inx}\,dx| \le \sup|f|\sqrt{2\pi}$. Since $f \in \mathscr{C}^1$, we can use integration by parts:

$$\begin{aligned} \widehat{Df}(n) &= \int_0^{2\pi} \frac{1}{i} f'(x) \frac{1}{\sqrt{2\pi}} e^{inx}\,dx \\ &= \int_0^{2\pi} \frac{1}{i\sqrt{2\pi}}(-1)(-in)f(x)e^{inx}\,dx + \frac{1}{i} f(x)\frac{1}{\sqrt{2\pi}}e^{inx}\Big|_0^{2\pi} \\ &= n\int_0^{2\pi} f(x)\overline{e_n(x)}\,dx \\ &= n\hat{f}(n). \end{aligned}$$

It follows from an induction argument that if $f \in \mathscr{C}^k(S^1)$, then $\widehat{D^k f}(n) = n^k \hat{f}(n)$ since $\widehat{D^k f}(n)$ is bounded. □

> **Remark 5.27.** In other words, taking Fourier coefficients connects differentiation to multiplication! Moreover, there is a correlation between how smooth (i.e. how many times differentiable) a function is and how fast its Fourier series decays.

Addendum: didn't have time to finish typing these up so the rest of the notes are hastily handwritten:


HELP ME FIGURE OUT THIS HOME-WORK PROBLEM, HOBBES. WHAT'S 3+8?
OK, ASSIGN THE ANSWER A VALUE OF "X". "X" ALWAYS MEANS MULTIPLY, SO TAKE THE NUMERATOR (THAT'S LATIN FOR "NUMBER EIGHTER") AND PUT THAT ON THE OTHER SIDE OF THE EQUATION.
THAT LEAVES YOU WITH THREE ON THIS SIDE, SO WHAT TIMES THREE EQUALS EIGHT? THE ANSWER, OF COURSE, IS SIX.
GOSH, I MUST HAVE DONE ALL THE OTHERS WRONG.
THESE PROBLEMS SEEM AWFULLY ADVANCED FOR FIRST GRADE, IF YOU ASK ME.
HERE'S ANOTHER MATH PROBLEM I CAN'T FIGURE OUT. WHAT'S 9+4?
OOH, THAT'S A TRICKY ONE. YOU HAVE TO USE CALCULUS AND IMAGINARY NUMBERS FOR THIS.
IMAGINARY NUMBERS?!
YOU KNOW, ELEVENTEEN, THIRTY-TWELVE AND ALL THOSE. IT'S A LITTLE CONFUSING AT FIRST.
HOW DID YOU LEARN ALL THIS? YOU'VE NEVER EVEN GONE TO SCHOOL!
INSTINCT. TIGERS ARE BORN WITH IT.

A Teaser

Parseval's Theorem:

(kinda): the sum of the square of a function is equal to the sum of the square of its transformation.

Let $A(x), B(x) : [0, 2\pi) \to \mathbb{C}$ be 2π-periodically square-integrable, with $A(x) = \sum_{-\infty}^{\infty} a_n e^{inx}$, $B(x) = \sum_{-\infty}^{\infty} b_n e^{inx}$. Then

$$\sum_{-\infty}^{\infty} a_n \overline{b_n} = \frac{1}{2\pi} \int_{-\infty}^{\infty} A(x) \overline{B(x)} \, dx$$

An Introduction to Fourier Coefficients

Notation: $e_n(x) = \frac{1}{2\pi} e^{inx}$, $n \in \mathbb{N}$

If we can write $f(x) = \sum_{n=-\infty}^{\infty} a_n e_n(x)$, we obtain

$$\int_0^{2\pi} f(x) e_{-m}(x) \, dx = \int_0^{2\pi} \left(\sum_{n=-\infty}^{\infty} \frac{1}{2\pi} a_n e^{inx} \right) \frac{1}{2\pi} e^{-imx} \, dx$$

$$= \frac{1}{(2\pi)^2} \sum_{n=-\infty}^{\infty} a_n \int_0^{2\pi} e^{inx} e^{-imx} \, dx$$

$$\longrightarrow \frac{1}{in - im} e^{ix(m-n)} \Big]_0^{4\pi} = 0 \text{ if } m \neq n, \quad 1 \text{ if } m = n$$

$$= \frac{1}{2\pi} \sum_{-\infty}^{\infty} a_n \delta_{m,n}$$

$$= \frac{1}{2\pi} a_n$$

Conjecture: If $f(x)$ is a reasonable periodic function on $\mathbb{R}$ with period 2π, we can write

$$f(x) = \sum_{n=-\infty}^{\infty} \hat{f}(n) e_n(x) \qquad (*)$$

where

$$\hat{f}(n) = \int_0^{2\pi} f(x) e_{-n}(x) \, dx$$

$(*) \Rightarrow \forall x \in \mathbb{R}$, $S_N = \sum_{n=-N}^{N} \hat{f}(n) e_n(x) \longrightarrow f(x)$.

Note: What do we mean by "reasonable"? An intuitive guess would be $f \in \mathcal{C}^0(S^1)$, but unfortunately there are some awful continuous functions out there, so that turns out to be incorrect.

Conjecture (nicer) A reasonable function on S^1 is determined uniquely by its Fourier coefficients.

↓

Theorem Let $f \in \mathcal{C}^1(S^1)$. Then $f(x) = \lim_{N \to \infty} S_N(x) = \sum_{N=-\infty}^{\infty} \hat{f}(n) e_n(x)$ $\forall x \in S^1$. In fact, the convergence of $S_N(x)$ is uniform!

Convolution Interlude

Definition Let g_1, g_2 be defined on S^1. The convolution of g_1 & g_2 is $g_1 * g_2(x) = \int_0^{2\pi} g_1(x-y) g_2(y)\, dy$.

(commutativity is immediate from change of variables)

WHAT WE WANT for appropriately nice functions: $\widehat{g_1 * g_2}(n) = \sqrt{2\pi}\, \hat{g}_1(n) \hat{g}_2(n)$

Lemma Let $g_1(x) = \sum_{n=-N}^{N} a_n e_n(x)$ and $g_2(x) \in \mathcal{C}^0(S^1)$. Then $g_1 * g_2(x) = \sqrt{2\pi} \sum_{n=-N}^{N} a_n \hat{g}_2(n) e_n(x)$

Pf $g_1 * g_2(x) = \int_0^{2\pi} g_1(x-y) g_2(y)\, dy = \int_0^{2\pi} \left(\sum_{n=-N}^{N} a_n e_n(x-y) \right) g_2(y)\, dy = \int_0^{2\pi} \sum_{n=-N}^{N} a_n \underbrace{e_n(x-y)}_{\sqrt{2\pi} e_n(x) e_{-n}(y)} g_2(y)\, dy = \sqrt{2\pi} \sum_{n=-N}^{N} a_n e_n(x) \int_0^{2\pi} g_2(y) e_{-n}(y)\, dy$ ☒

→ as $n \to \infty$ $a_n = \hat{g}_1(x)$ $\Rightarrow g_1 * g_2(x) = \sqrt{2\pi} \sum_{-\infty}^{\infty} \hat{g}_1(n) \hat{g}_2(n) e_n(x)$

Bessel's Inequality

Lemma Let $f \in \mathcal{C}^0(S^1)$. Then $\sum_{n=-\infty}^{\infty} |\hat{f}(n)|^2 \leq \int_0^{2\pi} |f(x)|^2 dx$

Pf Notice $\underbrace{\int_0^{2\pi} \left| f(x) - \sum_{n=-N}^{N} \hat{f}(n) e_n(x) \right|^2 dx}\geq 0$.

$\int_0^{2\pi} |f(x)|^2 dx - \sum_{n=-N}^{N} \overline{\hat{f}(n)} \int_0^{2\pi} f(x) e_n(x)\, dx - \sum_{n=-N}^{N} \hat{f}(n) \int_0^{2\pi} \overline{f(x) e_n(x)}\, dx + \sum_{m,n=-N}^{N} \hat{f}(n) \overline{\hat{f}(m)} \int e_n(x) e_m(x)\, dx$ ☒

Convergence of Fourier Series for Continuous Functions

Theorem (Fejer) If $f \in \mathcal{C}^0(S^1)$, the sequence $A_N = \frac{1}{N}(S_0 + \dots + S_{N-1})$ converges to f uniformly as $N \to \infty$.

Pf "show the appropriate kernel is an approximate identity" (In this case, the Fejer kernel: $F_N(x) = \frac{1}{N} \sum_{n=0}^{N-1} D_n(x)$ ⟶ $\frac{1}{2\pi} \sum_{n=-N}^{N} e_n(x)$

Theorem (Weierstrass) Let $f \in \mathcal{C}^0$ on $[0, 2\pi]$. Then $\forall \varepsilon > 0$ $\exists$ polynomial P s.t. $\sup_{x \in [0,2\pi]} |f(x) - P(x)| < \varepsilon$.

Pf Given any $f \in \mathcal{C}^0([0,\pi])$, extend f to be a continuous periodic function on $[0, 2\pi]$, so $f \in \mathcal{C}^0[0, 2\pi]$

By Fejer, for $\varepsilon > 0$ $\exists N \in \mathbb{N}$ s.t. $\sup |S_N(x) - f(x)| < \varepsilon/2$. Notice S_N is a polynomial Q in e^{ix}. Moreover, we can choose $M \in \mathbb{Z}$ s.t. $\sup_{[0,2\pi]} \left| Q(e^{ix}) - Q\left(\sum_{m=0}^{M} \frac{(ix)^m}{m!} \right) \right| < \frac{\varepsilon}{2}$ for $x \in [0, 2\pi]$

$\Rightarrow \left| f(x) - Q\left(\sum_{m=0}^{M} \frac{(ix)^m}{m!} \right) \right| \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$ ☒

Mean-Square Convergence of Plancherel + Failure of the Riemann Integral.

We can view $\{f \in \mathcal{C}^0(S^1)\}$ as a vector space, with inner product $(f,g) = \int f(x)\bar{g}(x)\,dx$.

- $(e_m, e_n) = \delta_{m,n}$ (i.e. complex exponentials are orthonormal)
- $\hat{f}(n) = (f, e_n)$.

$\Rightarrow$ The Fourier series for f is $\sum_{n=-\infty}^{\infty} (f,e_n)\, e_n$

- We assign norm $\|f\|_2 = (f,f)^{1/2} = \left(\int_0^{2\pi} |f(x)|^2\,dx\right)^{1/2}$

$\rightarrow$ Cauchy Schwartz (Revisited): $(f,g) \le \|f\|\,\|g\| \rightsquigarrow \int f(x)\bar{g}(x)\,dx \le \left(\int |f(x)|^2 dx\right)^{1/2}\left(\int |g(x)|^2 dx\right)^{1/2}$

- Triangle inequality is also gucci :)

From all of the above mentioned junk, we should try to get L^2 convergence of Fourier Series: $\int |S_N(x) - f(x)|^2 dx \to 0$

In fact, we can see S_N as the result of projecting f onto the subspace spanned by the orthonormal vectors $e_{-N}, \dots, e_0, \dots, e_N$, which we'd like to say is an orthonormal basis of S^1.

Turns out, for once, all of this is true!

Theorem For $f \in \mathcal{C}^0(S^1)$, $\|f - S_N(f)\| \to 0$ as $N \to \infty$

Pf By Fejer, $\forall \varepsilon > 0$ $\exists$ trig polynomial $h(x) = P(e^{ix})$ s.t. $|h(x) - f(x)| < \varepsilon/\sqrt{2\pi}$

$\rightarrow \|h - f\| = \left(\int_0^{2\pi} |h(x) - f(x)|^2\,dx\right)^{1/2} < \varepsilon$.

$\forall x \in [0, 2\pi]$. Let $S_N(h)$, $S_N(f)$ be the partial sums of Fourier series of h & f. For sufficiently large N, $S_N(h) = h$.

$\Rightarrow \|f - S_N(f)\| \le \|f-h\| + \|h - S_N(h)\| + \|S_N(h) - S_N(f)\| \le \varepsilon + \|S_N(h) - S_N(f)\|$

$= \varepsilon + \|S_N(h-f)\|$

Setting $g = h-f$, we get

$\|S_N(g)\|^2 = \int \left|\sum_{-N}^{N} \hat{g}(n) e_n(x)\right|^2 dx = \sum_{-N}^{N} |\hat{g}(n)|^2 \le \int |g(x)|^2 dx < \varepsilon^2$ ☒

Plancherel pt 1.

Theorem If $f \in \mathcal{C}^0(S^1)$, then $\|f\|_2^2 = \sum_{n=-\infty}^{\infty} |\hat{f}(n)|^2$

Pf Note that $\|S_N\|^2 = \sum_{-N}^{N} |\hat{f}(n)|^2$. But $\big|\|S_N\| - \|f\|\big| \le \|S_N - f\| \le \varepsilon$ ☒

Plancherel pt 2.

Theorem Let $f, g \in \mathcal{C}^0(S^1)$. Then $(f,g) = \sum_{-\infty}^{\infty} \hat{f}(n)\overline{\hat{g}(n)}$

Pf Apply previous Theorem to $f+g$ & $f-g$. $\Rightarrow (f+g, f+g) - (f-g, f-g) = \sum_{-\infty}^{\infty} |\widehat{f+g}(n)|^2 - |\widehat{f-g}(n)|^2$

$= \sum_{-\infty}^{\infty} |\hat{f}(n) + \hat{g}(n)|^2 - |\hat{f}(n) - \hat{g}(n)|^2$ ☒

Linear Transformations (aka the Basics®)

Theorem Let $r \in \mathbb{N}$. If a vector space X is spanned by a set of r vectors, $\dim X \geq r$.

$\Rightarrow \dim \mathbb{R}^n = n$.

Theorem Suppose X is a vector space and $\dim X = n$.

a) A set E of n vectors in X spans $X \Leftrightarrow E$ is independent.

b) X has a basis, and every basis consists of n vectors.

c) If $1 \leq r < n$ and $\{y_1, \ldots, y_r\}$ is independent in X, X has a basis $\supset \{y_1, \ldots, y_r\}$.

Definition A mapping $A : X \to Y$, X, Y v.s.'s, is a linear transformation if

$$A(x_1 + x_2) = Ax_1 + Ax_2, \quad A(cx_1) = cAx_1$$

$\forall x_1, x_2 \in X$, $c \in \mathbb{R}$.

- If $\{x_1, \ldots, x_n\}$ is a basis of X, every $x \in X$ has a unique representation of the form

$$x = \sum_{i=1}^{n} c_i x_i$$

$$\Rightarrow Ax = \sum_{i=1}^{n} c_i A x_i$$

- If A is an injective AND surjective linear operator on X, A is invertible.

$\Rightarrow \exists A^{-1}$ on X s.t. $A^{-1}Ax = x \quad \forall x \in X$.

(A^{-1} ↑ linear!)

Theorem A linear operator A on a finite-dimensional vector space X is injective $\Leftrightarrow$ range$(A) = X$.

Definition

a) Let $L(X,Y) = \{$ linear transformation $: X \to Y\}$. We write $L(X) := L(X,X)$

b) Let X, Y, Z be vector spaces, $A \in L(X,Y)$, $B \in L(Y,Z)$, we define the product BA as

$$(BA)x = B(Ax), \quad x \in X.$$

$\Rightarrow BA \in L(X,Z)$.

c) For $A \in L(\mathbb{R}^n, \mathbb{R}^m)$, define the norm $\|A\| = \sup_{x \in \mathbb{R}^n, |x| \leq 1} |Ax|$

- Notice $|Ax| \leq \|A\| |x| \quad \forall x \in \mathbb{R}^n$ and $|Ax| \leq \lambda |x| \Rightarrow \|A\| \leq \lambda$

Theorem

a) $A \in L(\mathbb{R}^n, \mathbb{R}^m) \Rightarrow \|A\| < \infty$, A is uniformly continuous.

b) $\|A+B\| \leq \|A\| + \|B\|$, $\|cA\| = |c|\,\|A\|$

c) $A \in L(\mathbb{R}^n, \mathbb{R}^m)$, $B \in L(\mathbb{R}^m, \mathbb{R}^k) \Rightarrow \|BA\| \leq \|B\|\|A\|$

Theorem (Inverses of Linear Operators)

Let: $\Omega = \{ A \in L(\mathbb{R}^n) \mid A \text{ invertible} \}$

a) If $A \in \Omega$, $B \in L(\mathbb{R}^n)$, and $\|B-A\| \cdot \|A^{-1}\| < 1$,
$\Rightarrow B \in \Omega$.

b) Ω is an open subset of $L(\mathbb{R}^n)$; $A \to A^{-1}$ is continuous.

Differentiation

Intuition In $\mathbb{R}$, if $f:(a,b) \to \mathbb{R}$ ($(a,b) \subset \mathbb{R}$), then $x \in (a,b) \Rightarrow f'(x) := \lim_{h\to 0} \frac{f(x+h)-f(x)}{h}$ if the limit exists.

$\Rightarrow f(x+h) - f(x) = f'(x)h + r(h)$

small bc: $\left(\lim_{h\to 0} \frac{r(h)}{h} = 0 \right)$

sum of linear function $h \to f'(x)h$ (w/ small remainder)

$\to$ Now, if $f:(a,b) \to \mathbb{R}^m$ ($(a,b) \subset \mathbb{R}$), then $f'(x)$ is defined to be the vector $y \in \mathbb{R}^n$ for which $\lim_{h\to 0} \left\{ \frac{f(x+h)-f(x)}{h} - y \right\} = 0$.

$\Rightarrow f(x+h) - f(x) = yh + r(h)$ (again)

Note: every $y \in \mathbb{R}^n$ induces a linear transformation: $\mathbb{R}^1 \to \mathbb{R}^n$ by associating to each $h \in \mathbb{R}$ the vector $hy \in \mathbb{R}^n$.

So we can regard $f'(x)$ as a member of $L(\mathbb{R}^1, \mathbb{R}^n)$.

Thus, if f is a differentiable mapping: $(a,b) \to \mathbb{R}^n$ ($(a,b) \subset \mathbb{R}$), $x \in (a,b) \Rightarrow f'(x)$ is the linear transformation $\mathbb{R}^1 \to \mathbb{R}^n$ with

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

or

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

Definition Let $E \subset \mathbb{R}^n$ be open, $f: E \to \mathbb{R}^m$, $x \in E$. If $\exists$ linear transformation $A: \mathbb{R}^n \to \mathbb{R}^m$ s.t.

$$\lim_{h \to 0} \frac{|f(x+h) - f(x) - Ah|}{|h|} = 0,$$

we say f is differentiable at x, and write $f'(x) = A$. If f is differentiable at every $x \in E$, we say f is differentiable in E.

Theorem Suppose the above definition holds w/ $A = A_1$ and $A = A_2$. Then $A_1 = A_2$.

Theorem (Chain Rule, kinda)

Let $E \subset \mathbb{R}^n$ be open. $f: E \to \mathbb{R}^m$, f differentiable @ $x_0 \in E$, g maps an open set containing $f(E)$ into $\mathbb{R}^k$, and g differentiable @ $f(x_0)$. Then $F: E \to \mathbb{R}^k$ defined by $F(x) = g(f(x))$ is differentiable @ x_0, and

$$F'(x_0) = g'(f(x_0))\, f'(x_0)$$

Pf Let $y_0 = f(x_0)$, $A = f'(x_0)$, $B = g'(y_0)$, and define $u(h) = f(x_0 + h) - f(x_0) - Ah$, $v(k) = g(y_0 + k) - g(y_0) - Bk$ $\forall$ viable $h \in \mathbb{R}^n$, $k \in \mathbb{R}^m$.

Then $|u(h)| = \varepsilon(h)|h|$, $|v(k)| = \mu(k)|k|$, where $\varepsilon(h), \mu(k) \to 0$ as $h, k \to 0$.

Given h, put $k = f(x_0 + h) - f(x_0)$. Then

$$|k| = |Ah + u(h)| \le [\|A\| + \varepsilon(h)]\,|h|$$

and $F(x_0 + h) - F(x_0) - BAh = g(y_0 + k) - g(y_0) - BAh = B(k - Ah) + v(k) = Bu(h) + v(k)$. Then for $h \neq 0$,

$$\frac{|F(x_0+h) - F(x_0) - BAh|}{|h|} \le \|B\|\,\varepsilon(h) + [\|A\| + \varepsilon(h)]\,\mu(k)$$

Let $h \to 0$. Then $\varepsilon(h) \to 0$. also, $k \to 0$ so $\mu(k) \to 0$. It follows that $F'(x_0) = BA$. ∎

Consider a function f that maps an open set $E \subset \mathbb{R}^n$ into $\mathbb{R}^m$. Let $\{e_1, \dots, e_n\}$ and $\{u_1, \dots, u_m\}$ be the standard bases of $\mathbb{R}^n$ and $\mathbb{R}^m$. The components of f are the real functions $f_1, \dots, f_m$ defined by

$$f(x) = \sum_{i=1}^{m} f_i(x) u_i \qquad (x \in E)$$

(or $f_i(x) = f(x) \cdot u_i$, $1 \le i \le m$). We write the partial derivative $D_j f_i$, or $\frac{\partial f_i}{\partial x_j}$, is the derivative of f_i w/ respect to x_j.

Note Even for continuous functions, existence of partial derivatives $\not\Rightarrow$ existence of the derivative. However,

existence of the derivative $\Rightarrow$ existence of partial derivatives.

Theorem Suppose $f: E \to \mathbb{R}^m$ ($E \subset \mathbb{R}^n$ open), f is differentiable at a point $x \in E$. Then the partial derivatives $(D_j f_i)(x)$ exist, and

$$f'(x)e_j = \sum_{i=1}^{m} (D_j f_i)(x)\, u_i \qquad (1 \le j \le n)$$

$\Rightarrow$ Let $[f'(x)]$ be the matrix that represents $f'(x)$ w/ respect to standard bases.

Then

$$[f'(x)] = \begin{bmatrix} (D_1 f_1)(x) & \cdots & (D_n f_1)(x) \\ \vdots & \ddots & \vdots \\ (D_1 f_m)(x) & \cdots & (D_n f_m)(x) \end{bmatrix}$$

Fun stuff: gradient $\nabla f(x) = \sum_{i=1}^{n} D_i f(x) e_i$, Directional derivative $D_u(x) = \lim_{t \to 0} \frac{f(x+tu) - f(x)}{t} = \nabla f(x) \cdot u$

$$= \sum_{i=1}^{n} (D_i f)(x) \cdot u_i$$

Theorem (Budget MVT) Suppose f maps a convex open set $E \subset \mathbb{R}^n$ into $\mathbb{R}^m$, f is differentiable in E, and $\exists M \in \mathbb{R}$ s.t.

$$\| f'(x) \| \le M$$

$\forall x \in E$. Then

$$|f(b) - f(a)| \le M |b - a|$$

$\forall a, b \in E$.

$\Rightarrow$ If also, $f'(x) = 0 \quad \forall x \in E$, then f is constant.

Definition A differentiable mapping f of an open set $E \subset \mathbb{R}^n$ into $\mathbb{R}^m$ is continuously differentiable in E if f' is a continuous mapping of E into $L(\mathbb{R}^n, \mathbb{R}^m)$. More explicitly,

$$\forall x \in E \ \forall \varepsilon > 0 \ \exists \delta > 0 \text{ s.t. } \| f'(y) - f'(x) \| < \varepsilon \text{ if } y \in E \text{ and } |x - y| < \delta.$$

If so, say f is a $\mathscr{C}'$-mapping, or $f \in \mathscr{C}'(E)$.

Theorem Suppose f maps an open set $E \subset \mathbb{R}^n$ into $\mathbb{R}^m$. Then $f \in \mathscr{C}'(E) \iff$ the partial derivatives $D_j f_i$ exist and are continuous on E.

Pf ($\Rightarrow$) Assume $f \in \mathscr{C}'(E)$. Then $D_j f_i(x) = f'(x) e_j \cdot u_i \quad \forall i, j \ \forall x \in E$. Hence

$$D_j f_i(y) - D_j f_i(x) = \{ [f'(y) - f'(x)] e_j \} \cdot u_i$$

Since $|u_i| = |e_j| = 1$,

$$| D_j f_i(y) - D_j f_i(x) | \le | [f'(y) - f'(x)] e_j | \le \| f'(y) - f'(x) \| \checkmark$$

($\Leftarrow$) Annoying convexity argument

The Contraction Principle

↳ a fixed point theorem valid in arbitrary complete metric spaces.

Definition Let (X, d) be a metric space. If $\varphi : X \to X$ and $\exists c < 1$ such that $d(\varphi(x), \varphi(y)) \leq c\, d(x,y)$ $\forall x, y \in X$. Then φ is a contraction of X into X.

Theorem If X is a complete metric space, φ is a contraction of X into X. Then $\exists!\, x \in X$ s.t. $\varphi(x) = x$.
(φ has a singular fixed point) — ONE, and ONLY ONE

↳ uniqueness is trivial: $\varphi(y) = y$, $\varphi(x) = x \Rightarrow d(x,y) = d(\varphi(x), \varphi(y)) \leq c\, d(x,y) \Rightarrow \Leftarrow$.

The Inverse Function Theorem

↳ roughly speaking, a continuously differentiable mapping f is invertible in a neighborhood of any point x at which the linear transformation $f'(x)$ is invertible.

Theorem Suppose f is a $\mathscr{C}'$ mapping of an open set $E \subset \mathbb{R}^n$ into $\mathbb{R}^n$, $f'(a)$ is invertible for some $a \in E$, $b = f(a)$.
Then
a) $\exists$ open sets $U, V \subset \mathbb{R}^n$ such that $a \in U$, $b \in V$, f is one-to-one on U, and $f(U) = V$.
b) If g is the inverse of f (which exists by (a)) defined in V by

$$g(f(x)) = x \qquad (x \in U).$$

then $g \in \mathscr{C}'(V)$.

↳ Writing $y = f(x)$ in component form, we arrive at the following interpretation of the conclusion:
The system of n equations $y_i = f_i(x_1, \dots, x_n)$ $(1 \leq i \leq n)$
can be solved for $x_1, \dots, x_n$ in terms of $y_1, \dots, y_n$; if we restrict x and y to small enough neighborhoods of a and b, the solutions are unique and continuously differentiable.

⇒ **Theorem** If f is a $\mathscr{C}'$-mapping of an open set $E \subset \mathbb{R}^n$ into $\mathbb{R}^n$ and if $f'(x)$ is invertible $\forall x \in E$, then $f(W)$ is an open subset of $\mathbb{R}^n$ $\forall$ open $W \subset \mathbb{R}^n$, i.e. f is an open mapping of E into $\mathbb{R}^n$

The Implicit Function Theorem

Notation: · If $x = (x_1, \dots, x_n) \in \mathbb{R}^n$, $y = (x_1, \dots, x_m) \in \mathbb{R}^m$, write

$(x, y) = (x_1, \dots, x_n, y_1, \dots, y_m) \in \mathbb{R}^{n+m}$

· If $A \in L(\mathbb{R}^{n+m}, \mathbb{R}^n)$ can be split into two linear transformations A_x, A_y defined by

$A_x h = A(h, 0)$, $A_y k = A(0, k)$

$\forall h \in \mathbb{R}^n, k \in \mathbb{R}^m$. Then $A_x \in L(\mathbb{R}^n)$, $A_y \in L(\mathbb{R}^m, \mathbb{R}^n)$, and $A(h,k) = A_x h + A_y k$.

↓

Theorem (Linear Implicit Function Theorem)

Let $A \in L(\mathbb{R}^{n+m}, \mathbb{R}^n)$ and A_x be invertible. Then $\forall k \in \mathbb{R}^m$ $\exists!$ $h \in \mathbb{R}^n$ such that $A(h,k) = 0$.

$h = -(A_x)^{-1} A_y k$.

Pf $A(h,k) = 0 \iff A_x h + A_y k = 0$. ∎

Theorem (Implicit Function Theorem)

Let f be a $\mathscr{C}'$ mapping of an set $E \subset \mathbb{R}^{n+m}$ into $\mathbb{R}^n$, such that $f(a,b) = 0$ for some point $(a,b) \in E$. Put $A = f'(a,b)$ and assume A_x is invertible. Then $\exists$ open sets $U \subset \mathbb{R}^{n+m}$, $W \subset \mathbb{R}^m$, with $(a,b) \in U$ and $b \in W$, s.t.

- $\forall y \in W$ $\exists! x$ s.t. $(x,y) \in U$ and $f(x,y) = 0$

  If this $x := g(y)$, then ① g is a $\mathscr{C}'$ mapping of W into $\mathbb{R}^n$

  ② $g(b) = a$

  ③ $f(g(y), y) = 0$ $(y \in W)$

  ④ $g'(b) = -(A_x)^{-1} A_y$

Note: The function g is "implicitly" defined, hence the name.

The equation $f(x,y) = 0$ can be written as a system of n equations in $n+m$ variables:

$$(*) \quad \begin{cases} f_1(x_1, \dots, x_n, y_1, \dots, y_m) = 0 \\ \vdots \\ f_n(x_1, \dots, x_n, y_1, \dots, y_m) = 0 \end{cases}$$

The assumption of A_x being invertible means $\begin{bmatrix} D_1 f_1 & \cdots & D_n f_1 \\ \vdots & \ddots & \vdots \\ D_1 f_n & \cdots & D_n f_n \end{bmatrix}$ evaluated @ (a,b) is invertible, and thus has non-zero determinant.

If $(*)$ holds when $x = a$, $y = b$, then by the theorem, $(*)$ can be solved for $x_1, \dots, x_n$ in terms of $y_1, \dots, y_m$ $\forall y$ near b, and these solutions are continuously differentiable functions of y.

Definition If $f : E \to \mathbb{R}^n$ ($E \subset \mathbb{R}^n$ open) is differentiable at a point $x \in E$, the determinant of $f'(x)$ is called the **Jacobian** of f at x:

$$J_f(x) := \det f'(x)$$
$$:= \frac{\partial(y_1,\dots,y_n)}{\partial(x_1,\dots,x_n)} \quad \text{if } (y_1,\dots,y_n) = f(x_1,\dots,x_n)$$

- The crucial hypothesis of the inverse function theorem can be restated as follows: $J_f(a) \neq 0$.
- In the implicit function theorem, the crucial assumption is that $\frac{\partial(f_1,\dots,f_n)}{\partial(x_1,\dots,x_n)} \neq 0$.

Derivatives of Higher Order

Theorem Suppose f is defined in an open set $E \subset \mathbb{R}^2$, $D_1 f$, $D_{21} f$, and $D_2 f$ exist at every point of E, and $D_{21} f$ is continuous at some point $(a,b) \in E$. Then $D_{12} f$ exists at (a,b) and

$$D_{12} f(a,b) = D_{21} f(a,b).$$

↳ **Corollary**: $D_{21} f = D_{12} f$ if $f \in \mathscr{C}''(E)$

Differentiation of Integrals

Motivating Question: under what conditions on φ can one prove the equation $\frac{d}{dt}\int_a^b \varphi(x,t)\,dx = \int_a^b \frac{\partial \varphi}{\partial t}(x,t)\,dx$?

Theorem Suppose (a) $\varphi(x,t)$ is defined for $a \le x \le b$, $c \le t \le d$, (b) α is an increasing function on $[a,b]$

(c) $\varphi^t \in \mathscr{R}(\alpha)$ $\forall t \in [c,d]$

(d) $c < s < d$, and to every $\varepsilon > 0$ $\exists \delta > 0$ s.t.

$$\left| (D_2\varphi)(x,t) - (D_2\varphi)(x,s) \right| < \varepsilon$$

$\forall x \in [a,b]$, $\forall t \in (s-\delta, s+\delta)$.

(holds whenever $D_2\varphi$ continuous on rectangle φ is defined.)

Define $f(t) = \int_a^b \varphi(x,t)\,d\alpha(x)$ $\quad (c \le t \le d)$

Then $(D_2\varphi)^s \in \mathscr{R}(\alpha)$, $f'(s)$ exists, and $f'(s) = \int_a^b (D_2\varphi)(x,s)\,d\alpha(x)$

Theorems from Spivak Integration

Lemma 3-1

Suppose P' be a refinement of P. Then $L(f,P) \le L(f,P')$ and $U(f,P') \le U(f,P)$.

↳ **Corollary 3-2** If P, Q are any two partitions, $L(f,P) \le U(f,Q)$. Pf: Let R be a common refinement of P & Q. ☒

A function $f: A \to \mathbb{R}^n$ is *integrable* on A if f is ① bounded, and ② $\sup_P \{L(f,P)\} = \inf_P \{U(f,P)\}$.

Theorem 3-3 (Integrability Criterion)

A bounded function $f: A \to \mathbb{R}^n$ is integrable $\iff$ $\forall \varepsilon > 0$ $\exists$ partition P with $U(f,P) - L(f,P) < \varepsilon$.

Pf ($\Rightarrow$) $\sup\{L(f,P)\} = \inf\{U(f,P)\}$. So for $\varepsilon > 0$ $\exists P, P'$ such that $U(f,P) - L(f,P') < \varepsilon$. Take P'' to be a refinement of $P \cup P'$. Then $U(f,P'') - L(f,P'') \le U(f,P) - L(f,P') < \varepsilon$. ✓ ($\Leftarrow$) Immediate ☒

Theorem 3-4

The countable union of sets of measure zero also has measure zero, i.e. $A = \bigcup_{i=1}^{\infty} A_i$ has measure zero if each A_i does.

Pf use Cantor's diagonalization argument to obtain a series $< \sum_{i=1}^{\infty} \frac{\varepsilon}{2^i} = \varepsilon$ ☒

An important distinction: A has measure zero if $\forall \varepsilon > 0$ $\exists$ cover of closed rectangles $\{V_i\}_{i=1}^{\infty}$ such that $\sum_{i=1}^{\infty} V_i < \varepsilon$.

A has content zero if $\forall \varepsilon > 0$ $\exists$ finite cover of closed rectangles $\{U_i\}_{i=1}^{\infty}$ such that $\sum_{i=1}^{\infty} U_i < \varepsilon$.

Theorem 3-5

① $[a,b] \subset \mathbb{R}$, $a < b$ does not have content 0. ② Moreover, if $\{U_1, \dots, U_n\}$ is a finite cover of $[a,b]$ by closed intervals, then $\sum_{i=1}^{n} v(U_i) \ge b - a$

Pf Obviously, each $U_i \subset [a,b]$. Let $\{t_i\}_0^k$, $a = t_0 < t_1 < \cdots < t_k$ be all the endpoints of all U_i. $\sum_{i=1}^{n} v(U_i) \ge \sum_{j=1}^{k} (t_j - t_{j-1}) = b - a$. ☒

Theorem 3-6

If A is compact and has measure zero, then A has content zero.

Pf Immediate by compactness. ☒

Integrable Functions

$\lim_{\delta \to 0} \sup\{f(x) : |x-a|<\delta\} - \inf\{f(x) : |x-a|<\delta\}$

Lemma 3-7

Let A be a closed rectangle, $f: A \to \mathbb{R}^n$ be bounded such that $o(f,a) < \varepsilon$ $\forall a \in A$. Then $\exists$ partition P of A with $U(f,P) - L(f,P) < \varepsilon \cdot v(A)$.

Theorem 3-8

Let A be a closed rectangle, $f: A \to \mathbb{R}^n$ be bounded. Let $B = \{x \mid f \text{ discontinuous @ } x\}$. Then

f integrable $\iff$ B has measure zero.

$\begin{cases} 0 & x \notin C \\ 1 & x \in C \end{cases}$

Theorem 3-9 The function $\chi_C : A \to \mathbb{R}$ is integrable $\iff$ ∂C has measure 0 (content zero since ∂C closed).

- A bounded set C whose boundary has measure 0 is called Jordan-measurable.
- The integral $\int_C 1$ is called the content/volume of C.

  NOTE: Even if C is open and f is continuous, C might not be Jordan-measurable so $\int_C f$ is not necessarily defined.

Theorem 3-10 (Fubini)

Let $A \subset \mathbb{R}^n$, $B \subset \mathbb{R}^m$ be closed rectangles, $f: A \times B \to \mathbb{R}$ be integrable. For $x \in A$, let $g_x : B \to \mathbb{R}$ be defined by $g_x(y) = f(x,y)$ and let

$$\mathcal{L}(x) = \underline{\int_B} g_x = \underline{\int_B} f(x,y)\,dy$$

$$\mathcal{U}(x) = \overline{\int_B} g_x = \overline{\int_B} f(x,y)\,dy$$

Then $\mathcal{L}(x), \mathcal{U}(x) \in \mathcal{R}(A)$, and

$$\int_{A\times B} f = \int_A \mathcal{L} = \int_A \left(\underline{\int_B} f(x,y)\,dy\right)dx \quad \& \quad \int_{A\times B} f = \int_A \mathcal{U} = \int_A \left(\overline{\int_B} f(x,y)\,dy\right)dx.$$