MATH 321 Notes

Sushrut Tadwalkar

February 25, 2024

Quotes of the day: 01/08/2024 by Dr. Joshua Zahl

"Sometimes MVT stands for most valuable theorem."

"LaTeXis the language math is written in."

1 320 Review

Definition: Differentiable at point

Recall for some $f : [a, b] \to \mathbb{R}$, $c \in [a, b]$, we say that f is ***differentiable at*** c if $\lim_{x \to c} \dfrac{f(x) - f(c)}{x - c}$ exists (as a real number); we denote this by $f'(c)$.

This is a very elementary definition of what it means for something to be differentiable, but we look a bit deeper into what it means for the limit of a function. In particular, consider the case of the limit mentioned in the definition; what does it mean for this limit to exist?

- It satisfies the $\varepsilon - \delta$ definition of a limit.
- c is a limit point in $[a, b]$; in a metric space this means that any ball about the point c has a non empty intersection with the set $[a, b]$.
- $g(x) = \dfrac{f(x) - f(c)}{x - c}$ is a function with domain $[a, b] \backslash \{c\}$.

We might ask ourselves why go through all these layers of abstraction, when the high school definition of a limit works. Well, we have to make sure that the high school definition is consistent with what we have laid out so far: for any $c \in (a, b)$, the high school definition is just fine, but back then we had to separately check the end-points $c = a$ and $c = b$ with one sided limits, which we don't have to do when we satisfy one of the things laid out above. Hence, it is worth to delve into the abstraction.

Definition: Differentiable on a set

If $f : [a, b] \to \mathbb{R}$ is differentiable at *every* point $c \in [a, b]$, then we say f is differentiable on $[a, b]$, and this gives us a new function $f' : [a, b] \to \mathbb{R}$.

Furthermore, we can keep iterating this definition: if f' is differentiable at $c \in [a, b]$, we write $f''(c) = (f')'(c)$.

Notation 1. Some alternate notations for derivatives are:

- $f(c), f'(c), f''(c), \ldots$
- $f^{(0)}(c), f^{(1)}(c), f^{(2)}(c), \ldots, f^{(k)}(c)$.

Food for thought 1. Why have co-domain $\mathbb{R}$? Why not $\mathbb{C}$, or some arbitrary field F? Why not a general set/metric space?

Similarly, why make the domain a closed interval? Why not a more general subset of $\mathbb{R}$, or even $\mathbb{C}$? Why not a general set/metric space?

We cannot really have a notion of a derivative in a topological space, because in a TS we have no notion of a distance, only open and closed sets, so it does not really make sense to be talking about the rate of change of something as we get closer to a point. This is not a complete answer, but it's hard to give a better answer at this point in time. If we google a topological derivative, there might be some constructions that come close, but nothing that is a true generalization of a derivative using arbitrary topological spaces.

Theorem: Rolle's theorem

Let $f : [a, b] \to \mathbb{R}$ be differentiable with $f(a) = f(b)$. Then, there exists $c \in (a, b)$ such that $f'(c) = 0$.

We go on to showcase one of the more important theorems in differentiation:

Theorem: Taylor's theorem

Let $f : [a, b] \to \mathbb{R}$, and $n \geq 0$ be an integer. Suppose that f is $(n + 1)$ times differentiable on $[a, b]$. Let x_0 and x be points in $[a, b]$ with $x_0 \neq x$. Then, there exists a point c strictly between x_0 and x such that

$$f(x) = \underbrace{\left[\sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k\right]}_{P_n(x)} + \frac{f^{(n+1)}(c)}{(n + 1)!}(x - x_0)^{n+1}. \qquad (\dagger)$$

Call $P_n(x)$ the "degree n Taylor expansion of f around x_0".

Note (Choice of notation)**.** While choosing notation, we have many things competing for the "attention" of the notation; for example in case of $P_n(x)$, technically it is dependent on n, f, x_0, so it should be $P_n^{f,x_0}(x)$, but this is clunky. As we do more math, we get better with choosing what information notation should encode, and what can be omitted. In this particular case, we would generally know the f and x_0 and the more important part that needs to be encoded is the degree.

Food for thought 2. Let $f : \mathbb{R} \to \mathbb{R}$ which is infinitely differentiable. Suppose $f^{(k)}(0) = 0$ for all k; is it true that f must be the zero function?

Quotes of the day: 01/10/2024

No quotes today :(

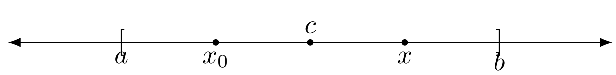

Figure 1: Visualization of points in Taylor's theorem.

Proof. We start by noting that for $n = 0$, *eq.* $(\dagger)$ says $f(x) = f(x_0) + f'(c)(x - x_0)$.

Define $A \in \mathbb{R}$ by

$$f(x) - P_n(x) = \frac{A}{(n + 1)!}(x - x_0)^{n+1}.$$

Our goal here is to show that there exists a c between x_0 and x such that $f^{(n+1)}(c) = A$.

Define $g(t) = f(t) - P_n(t) - \dfrac{A}{(n + 1)!}(t - x_0)^{n+1}$.

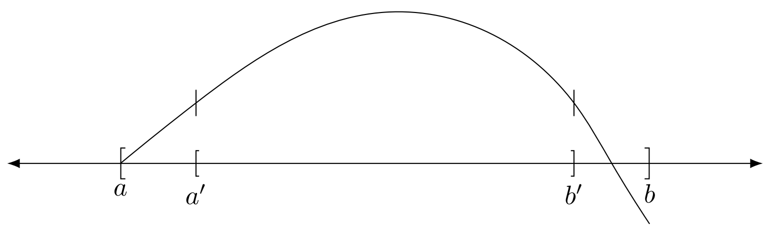

Figure 2: Visualization of how we shrink the interval to possibly apply Rolle's theorem.

Observe

$$\begin{aligned} g(x_0) =& f(x_0) - P_n(x_0) \\ =& f(x_0) - f(x_0) - 0 \\ =& 0, \end{aligned}$$

so $g(x) = 0$ by definition of A. Hence, for $j = 0, \ldots, n$,

$$\begin{aligned} g^{(j)}(x_0) =& f^{(j)}(x_0) - P_n^{(j)}(x_0) - \frac{d^j}{dt^j}\left\{\frac{A}{(n+1)!}(t-x_0)^{n+1}\right\}\Bigg|_{t=x_0} \\ =& f^{(j)}(x_0) - f^{(j)}(x_0) - 0 \\ =& 0, \end{aligned}$$

which tells us that $g^{(n+1)}(t) = f^{(n+1)}(t) - 0 - A$. Now, our goal is to find a c such that $g^{(n+1)}(c) = 0$.

Note that

$g(x_0) = 0, g(x) = 0$ by Rolle's thoerem, there exists c_1 between x_0 and x such that $g'(c_1) = 0$.
$g'(x_0) = 0, g'(x) = 0$ by Rolle's thoerem, there exists c_2 between x_0 and c such that $g''(c_2) = 0$.
$\vdots$
$g^{(n)}(x_0) = 0, g^{(n)}(x) = 0$ by Rolle's thoerem, there exists c_{n+1} between x_0 and c_n such that $g^{(n+1)}(c_{n+1}) = 0$.

Finally, set $c := c_{n+1}$ to conclude the proof. $\square$

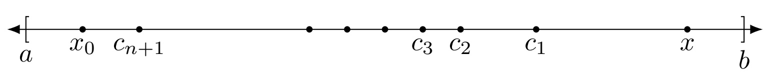

Figure 3: Visualization of the iterative process to find c_{n+1}.

Example 1. Why is Taylor's theorem so useful? We look at a few examples which illustrate this: set $x_0 = 0$,

1. f is a polynomial of degree D; $P_n(t)$ will be the first terms of f up to degree n.
2. If $f(t) = e^t$, we get
$$P_n(t) = \frac{1}{0!} + \frac{t}{1!} + \frac{t^2}{2!} + \frac{t^3}{3!} + \cdots + \frac{t^n}{n!}.$$
3. If $f(x) = \sin x$, we get
$$P_n(t) = 0 + t + 0 - \frac{t^3}{3!} + 0 + \frac{t^5}{5!} + \ldots$$

Quotes of the day: Dr. Joshua Zahl 01/12/2024

No quotes today :(

Recall food for thought 2: for a function $f : \mathbb{R} \to \mathbb{R}$ or $f : [-1, 1] \to \mathbb{R}$, given that $f(0) = 0$, and $f^{(k)}(0) = 0$ for all $k \in \mathbb{N}$, must it be true that $f(t) = 0$ for all t?

Solution. If we apply Taylor's theorem, we get

$$f(x) = \cancel{\sum_{k=0}^{n} \frac{f^{(k)}(0)}{k!}}x^n + \frac{f^{(n+1)}(c)}{(n+1)!}x^{n+1},$$

where the $P_n(x)$ term dies, but clearly the remainder term here could behave in unexpected ways (like blowing up), which would then be a function that fits our specification but is not identically zero. $\square$

Consider an example:

$$f(x) = \begin{cases} 0 & \text{if } x \leq 0 \\ e^{-1/x} & \text{if } x > 0 \end{cases}.$$

For $x > 0$, by applying chain rule, we get that $f^{(k)}(x) = Q(x)e^{-1/x}$, where $Q(x)$ is rational function in x. This function is in fact infinitely differentiable at the origin (good exercise); the intuition behind this is that the exponential function will always beat any rational function in decay at the origin, and the derivative at the origin will always be zero. However, just from Taylor's theorem, it would appear that the function is not zero at the origin, which in this case is not true. The point here is that while Taylor's theorem can aid in reconstructing a function by only using information about it at the origin, it can at times be misleading, and isn't as strong as it might seem.

2 The Riemann and Riemann-Stieltjes Integral

2.1 The Riemann integral

Definition: Partition

A ***partition*** of $[a, b]$ is a finite set $\mathcal{P} = \{x_0, x_1, \ldots, x_n\}$ with $a = x_0 < x_1 < \cdots < x_n = b$.

For $i = 1, \ldots, n$, let $\Delta x_i = x_i - x_{i-1}$. For $f : [a, b] \to \mathbb{R}$ bounded, define

$$\begin{aligned} M_i &= \sup\{f(x) : x \in [x_{i-1}, x_i]\} \\ m_i &= \inf\{f(x) : x \in [x_{i-1}, x_i]\}; \end{aligned}$$

also, define

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

$$L(\mathcal{P}, f) = \sum_{i=1}^{n} m_i \Delta x_i.$$

We define

$$\text{Upper Riemann integral} : \overline{\int_a^b} f\, dx = \inf_{\mathcal{P}} U(\mathcal{P}, f)$$

$$\text{Lower Riemann integral} : \underline{\int_a^b} f\, dx = \sup_{\mathcal{P}} L(\mathcal{P}, f);$$

the sup and inf are taken over all partitions of $[a,b]$.

Definition: Riemann integrable

We say $f:[a,b]\to\mathbb{R}$ is ***Riemann integrable*** if $\inf_{\mathcal{P}} U(\mathcal{P},f)=\sup_{\mathcal{P}} L(\mathcal{P},f)$, in which case we denote this number by $\displaystyle\int_a^b f\,dx$, and we say that $f\in\mathcal{R}[a,b]$: set of Riemann integrable functions on $[a,b]$.

A natural question that follows is what kinds of functions are Riemann integrable? We look at an example:

Example 2. Let $[a,b]=[0,1]$, $f(x)=x$. If $\mathcal{P}=\{x_0,\dots,x_n\}$ is a partition, $M_i=x_i$, $m_i=x_{i-1}$.

Consider $\mathcal{P}=\left\{0,\frac{1}{n},\frac{2}{n},\dots,\frac{n}{n}\right\}$. In this case,

$$\begin{aligned}U(\mathcal{P},f)=&\sum_{i=1}^{n}\underbrace{\frac{i}{n}}_{M_i}\frac{1}{n}\\=&\frac{1}{n^2}\sum_{i=1}^{n}i\\=&\frac{1}{n^2}\frac{1}{2}n(n+1)\\=&\frac{1}{2}+\frac{1}{2n}.\end{aligned}$$

In particular,

$$\overline{\int_0^1} x\,dx\le\inf\left\{\frac{1}{2}+\frac{1}{2n}:n\in\mathbb{N}\right\}=\frac{1}{2}.$$

Similarly,

$$\begin{aligned}L(\mathcal{P},f)=&\sum_{i=1}^{n}\frac{i-1}{n}\frac{1}{n}\\=&\frac{1}{n^2}\frac{1}{2}n(n-1)\\=&\frac{1}{2}-\frac{1}{2n}.\end{aligned}$$

In particular,

$$\underline{\int_0^1} x\,dx\ge\frac{1}{2}.$$

At this point, we cannot really conclude that it is Riemann integrable, since we still need the inequality $U(\mathcal{P},f)\ge L(\mathcal{P},f)$, which we have not proved yet. However, rather than proving this, we will now define the Riemann-Stieltjes integral first, prove it for that, and we get it for the Riemann integral as a special case.

2.2 The Riemann-Stieltjes integral

Let $\alpha:[a,b]\to\mathbb{R}$ be (weakly) monotone increasing; let $\mathcal{P}=\{x_0,x_1,\dots,x_n\}$ be a partition of $[a,b]$.

For $i=1,\dots,n$, let $\Delta\alpha_i=\alpha(x_i)-\alpha(x_{i-1})$ (if $\alpha(x)=x$, then $\Delta\alpha_i=\Delta x_i$). For $f:[a,b]\to\mathbb{R}$ bounded, define

$$\begin{aligned}U(\mathcal{P},f,\alpha)=&\sum_{i=1}^{n}M_i\Delta\alpha_i\\L(\mathcal{P},f,\alpha)=&\sum_{i=1}^{n}m_i\Delta\alpha_i.\end{aligned}$$

Define

$$\text{Upper Riemann-Stieltjes integral} : \overline{\int_a^b} f\,dx = \inf_{\mathcal{P}} U(\mathcal{P}, f, \alpha)$$
$$\text{Lower Riemann-Stieltjes integral} : \underline{\int_a^b} f\,dx = \sup_{\mathcal{P}} L(\mathcal{P}, f, \alpha).$$

Definition: Riemann-Stieltjes integrable

We say $f : [a, b] \to \mathbb{R}$ is ***Riemann-Stieltjes integrable*** if $\inf_{\mathcal{P}} U(\mathcal{P}, f, \alpha) = \sup_{\mathcal{P}} L(\mathcal{P}, f, \alpha)$, in which case we denote this number by $\displaystyle\int_a^b f\,d\alpha$, and we say that $f \in \mathcal{R}_\alpha[a, b]$: set of Riemann-Stieltjes integrable functions on $[a, b]$.

Does $\alpha(x)$ always have to be continuous? We look at an example:

Example 3. Consider

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

what does the integral $\displaystyle\int_{-1}^{1} f\,d\alpha$ look like? It it literally just $f(0)$ (it is like the Dirac-δ "function"), and this showcases the power of the Riemann-Stieltjes integral, because $\alpha(x)$ does not have to be continuous.

Quotes of the day: Dr. Joshua Zahl 01/15/2024

No quotes today :(

Definition: Refinement and common refinement (Rudin 6.3)

Let $\mathcal{P}$ and $\mathcal{P}^*$ be partitions of $[a, b]$. We say $\mathcal{P}^*$ is a ***refinement*** of $\mathcal{P}$ is $\mathcal{P} \subseteq \mathcal{P}^*$.

If $\mathcal{P}_1$ and $\mathcal{P}_2$ are partitions of $[a, b]$, their ***common refinement*** is the partition $\mathcal{P}_1 \cup \mathcal{P}_2$.

Theorem: Baby Rudin 6.4

Let $\mathcal{P}^*$ is a refinement of $\mathcal{P}$. Then, $L(\mathcal{P}, f, \alpha) \leq L(\mathcal{P}^*, f, \alpha) \leq U(\mathcal{P}^*, f, \alpha) \leq U(\mathcal{P}, f, \alpha)$.

Proof. Middle inequality we have seen before (follows from the definition of inf and sup.) Proving the leftmost inequality is equivalent to proving the rightmost inequality, so we will just prove the leftmost one.

The only interesting case is when $\mathcal{P} \subsetneq \mathcal{P}^*$, since if they're the same set, we just get equality. So it suffices to prove the inequality when $\mathcal{P}^*$ has one additional point (the minimum for two sets to not be the same one; this can be extended to any number of points by induction.) Let the additional point be x^*, and let it be between two points x_i and x_{i+1} of $\mathcal{P}$.

We proceed by comparing the two lower sums $L(\mathcal{P}, f, \alpha)$ and $L(\mathcal{P}^*, f, \alpha)$:

$$\begin{aligned} L(\mathcal{P}, f, \alpha) &= \sum_{j=1}^{n} m_j \Delta\alpha_j \\ L(\mathcal{P}^*, f, \alpha) &= \sum_{j=1}^{i} m_j \Delta\alpha_j + (\inf\{f(x) : x \in [x_i, x^*]\})\,(\alpha(x^*) - \alpha(x_i)) \\ &\qquad + (\inf\{f(x) : x \in [x^*, x_{i+1}]\})\,(\alpha(x_{i+1}) - \alpha(x^*)) \\ &\qquad + \sum_{j=i+2}^{n} m_j \Delta\alpha_j. \end{aligned}$$

Hence,

$$\begin{aligned}
L(\mathcal{P}^*, f, \alpha) - L(\mathcal{P}, f, \alpha) &= \left(\inf_{x\in[x_i,x^*]} f(x)\right)(\alpha(x^*) - \alpha(x_i)) + \left(\inf_{x\in[x^*,x_{i+1}]} f(x)\right)(\alpha(x_{i+1}) - \alpha(x^*)) - m_{i+1}\Delta\alpha_{i+1} \\
&\geq \left(\inf_{x\in[x_i,x_{i+1}]} f(x)\right)(\alpha(x^*) - \alpha(x_i)) + \left(\inf_{x\in[x_i,x_{i+1}]} f(x)\right)(\alpha(x_{i+1}) - \alpha(x^*)) - m_{i+1}\Delta\alpha_{i+1} \\
&= m_{i+1}(\alpha(x^*) - \alpha(x_i) + \alpha(x_{i+1}) - \alpha(x^*)) - m_{i+1}\Delta\alpha_{i+1} \\
&= m_{i+1}\Delta\alpha_{i+1} - m_{i+1}\Delta\alpha_{i+1} = 0.
\end{aligned}$$

□

Theorem: Baby Rudin 6.5

Let $f : [a,b] \to \mathbb{R}$ be bounded, and $\alpha : [a,b] \to \mathbb{R}$ be monotone increasing. Then,

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

Proof. Let $\mathcal{P}_1$ and $\mathcal{P}_2$ be partitions of $[a,b]$; hence, let $\mathcal{P}^* = \mathcal{P}_1 \cup \mathcal{P}_2$. By theorem 6.4, $L(\mathcal{P}_1, f, \alpha) \leq U(\mathcal{P}_2, f, \alpha)$. Hence,

$$\underline{\int_a^b} f\,d\alpha = \sup_{\mathcal{P}_1} L(\mathcal{P}_1, f, \alpha) \leq U(\mathcal{P}_2, f, \alpha).$$

Since this is true for *every* $\mathcal{P}_2$,

$$\underline{\int_a^b} f\,d\alpha \leq \inf_{\mathcal{P}_2} U(\mathcal{P}_2, f, \alpha) = \overline{\int_a^b} f\,d\alpha.$$

□

Note. This was the missing piece that we required to show that $\displaystyle\int_0^1 x\,dx = \frac{1}{2}$.

Theorem: Baby Rudin 6.6

Let $f : [a,b] \to \mathbb{R}$ be bounded, and $\alpha : [a,b] \to \mathbb{R}$ be monotone increasing. Then $f \in \mathcal{R}_\alpha[a,b] \iff$ for all $\varepsilon > 0$, there exists $\mathcal{P}$ such that $U(\mathcal{P}, f, \alpha) - L(\mathcal{P}, f, \alpha) < \varepsilon$.

Proof. By hypothesis,

$$\sup_{\mathcal{P}} L(\mathcal{P}, f, \alpha) = \int_a^b f\,d\alpha = \inf_{\mathcal{P}} U(\mathcal{P}, f, \alpha).$$

Let $\varepsilon > 0$, then there exists a partition $\mathcal{P}_1$ such that

$$L(\mathcal{P}_1, f, \alpha) > \int_a^b f\,d\alpha - \frac{\varepsilon}{2},$$

and there exists $\mathcal{P}_2$ such that

$$U(\mathcal{P}_2, f, \alpha) < \frac{\varepsilon}{2} + \int_a^b f\,d\alpha.$$

Let $\mathcal{P} = \mathcal{P}_1 \cup \mathcal{P}_2$. By theorem 6.4,

$$L(\mathcal{P}_1, f, \alpha) \leq L(\mathcal{P}, f, \alpha) \leq U(\mathcal{P}, f, \alpha) \leq U(\mathcal{P}_2, f, \alpha).$$

Hence,

$$U(\mathcal{P}, f, \alpha) - L(\mathcal{P}, f, \alpha) < \varepsilon.$$

The other direction follows from definition. □

Quotes of the day: Dr. Joshua Zahl 01/17/2023

No quotes today :(

Theorem: Baby Rudin 6.8

Let $\alpha : [a,b] \to \mathbb{R}$ be monotone increasing, and let $f : [a,b] \to \mathbb{R}$ be continuous. Then $f \in \mathcal{R}_\alpha[a,b]$, i.e., $C([a,b]) \in \mathcal{R}_\alpha[a,b]$.

Proof. Given that f is continuous, since $[a,b]$ is compact, f is uniformly continuous. Hence, for all $\varepsilon_1 > 0$, there exists some $\delta > 0$ such that $|f(x) - f(y)| < \varepsilon$ for all x, y with $|x - y| < \delta$.

Thus, if $\mathcal{P}$ is a partition with $\Delta x_i < \delta$ for all i, then $M_i - m_i < \varepsilon_1$ for all i. Hence,

$$U(\mathcal{P}, f, \alpha) - L(\mathcal{P}, f, \alpha) \leq \varepsilon_1 \sum_{i=1}^{n} \Delta\alpha_i = \varepsilon\left(\alpha(b) - \alpha(a)\right).$$

Given $\varepsilon > 0$, select ε_1 sufficiently small, such that $\varepsilon_1\left(\alpha(b) - \alpha(a)\right) < \varepsilon$. Choose $\mathcal{P}$ as above for the corresponding ε_1. We have shown that for $\varepsilon > 0$, there exists a partition $\mathcal{P}$, such that $U(\mathcal{P}, f, \alpha) - L(\mathcal{P}, f, \alpha) < \varepsilon$. Therefore, by theorem 6.6, $f \in \mathcal{R}_\alpha[a,b]$. $\square$

Food for thought 3. Can we describe/characterize $\mathcal{R}_\alpha[a,b]$ or $\mathcal{R}[a,b]$?

Turns out there is a nice bi-conditional statement to characterize these sets, but we need to develop some more machinery before we can do so.

Theorem: Baby Rudin 6.9

Let $f : [a,b] \to \mathbb{R}$ be monotone (increasing or decreasing), $\alpha : [a,b] \to \mathbb{R}$ be monotone increasing, and continuous. Then $f \in \mathcal{R}_\alpha[a,b]$.

Proof. Let $n \in \mathbb{N}$; by the intermediate value theorem, there exists a partition $\mathcal{P}$ such that $\Delta\alpha_i = \dfrac{\alpha(b) - \alpha(a)}{n}$ for all $i = 1, \ldots, n$. Note that

$$\begin{aligned}
U(\mathcal{P}, f, \alpha) - L(\mathcal{P}, f, \alpha) &= \sum_{i=1}^{n} (M_i - m_i)\Delta\alpha_i \\
&= \frac{\alpha(b) - \alpha(a)}{n} \sum_{i=1}^{n} (M_i - m_i).
\end{aligned}$$

Suppose, without loss of generality, f is monotone increasing; we have $M_i = f(x_i)$ and $m_i = f(x_{i-1})$, so

$$\begin{aligned}
U(\mathcal{P}, f, \alpha) - L(\mathcal{P}, f, \alpha) &= \frac{\alpha(b) - \alpha(a)}{n} \sum_{i=1}^{n} (f(x_i) - f(x_{i-1})) \\
&= \frac{\alpha(b) - \alpha(a)}{n} \left(\cancel{f(x_1)} - f(x_0) + \cancel{f(x_2)} - \cancel{f(x_1)} + \cancel{f(x_3)} - \cancel{f(x_2)} + \cdots + f(x_n) - \cancel{f(x_{n-1})}\right) \\
&= \left(\frac{\alpha(b) - \alpha(a)}{n}\right)(f(x_n) - f(x_0)) \\
&= \left(\frac{\alpha(b) - \alpha(a)}{n}\right)(f(a) - f(b)) \\
&= \frac{1}{n} \underbrace{\left(\alpha(b) - \alpha(a)\right)\left(f(b) - f(a)\right)}_{\in \mathbb{R}},
\end{aligned}$$

so given $\varepsilon > 0$, pick $n \in \mathbb{N}$ such that
$$\left| \frac{1}{n} \left(\alpha(b) - \alpha(a)\right) \left(f(b) - f(a)\right) \right| < \varepsilon.$$
For such a function, $|U(\mathcal{P}, f, \alpha) - L(\mathcal{P}, f, \alpha)| < \varepsilon$. By theorem 6.6, $f \in \mathcal{R}_\alpha[a, b]$. □

Note (f monotone decreasing)**.** In this case, the proof is pretty much the same; not tricky to work out the details.

Theorem: Baby Rudin 6.10

Let $f : [a, b] \to \mathbb{R}$ be bounded and continuous at all but finitely many points. Let $\alpha : [a, b] \to \mathbb{R}$ be monotone increasing, and continuous at every point where f is not continuous. Then, $f \in \mathcal{R}_\alpha[a, b]$.

Quotes of the day: Dr. Joshua Zahl 01/19/2024

No quotes today :(

We will now prove theorem 6.10:

Proof. Let $N := \sup_{x \in [a,b]} |f|$; this is finite since f is bounded. Let $\mathcal{E} := \{e_1, \ldots, e_k\}$ be the set of points where f is discontinuous.

Let $\varepsilon_1 > 0$. Since α is continuous at each $e_i \in \mathcal{E}$, we can pick $u_i < e_i < v_i$, where $u_i, v_i \in [a, b]$, such that $0 \le \alpha(v_i) - \alpha(u_i) < \varepsilon_1$. The inequalities can be equality if $e_i = a$ or $e_i = b$.

Let $\mathcal{K} := [a, b] \setminus \bigcup_{i=1}^{k} (u_i, v_i)$. Since $\mathcal{K}$ is closed and bounded, it is compact. Furthermore, since f is continuous on $\mathcal{K}$, it is uniformly continuous on $\mathcal{K}$: for all $x, y \in \mathcal{K}$ with $|x - y| < \delta$, we have $|f(x) - f(y)| < \varepsilon_1$.

Let $\{y_i\} \subseteq \mathcal{K}$ be a set of points such that for every $x \in \mathcal{K}$, there is an index i such that $y_i \le x \le y_{i+1}$, and $0 < y_{i+1} - y_i < \delta$. Also, let $\mathcal{P} := \{u_i, v_i\}_{i=1}^k \cup \{y_i\} \cup \{a, b\}$ (might have to re-order to pout these in increasing order). Hence,

$$\begin{aligned} 0 \le U(\mathcal{P}, f, \alpha) - L(\mathcal{P}, f, \alpha) &= \sum_{i=1}^{n} (M_i - m_i)\Delta\alpha_i \text{ (for interval } [x_{i-1}, x_i]) \\ &\le \underbrace{k(2N)\varepsilon_1}_{[u_i, v_i] \text{ intervals}} + \underbrace{\varepsilon_1 \left(\alpha(b) - \alpha(a)\right)}_{[y_{i-1}, y_i] \text{ intervals}}. \end{aligned}$$

Given $\varepsilon > 0$, choose ε_1 such that
$$k(2N)\varepsilon_1 + \varepsilon_1 \left(\alpha(b) - \alpha(a)\right) < \varepsilon;$$
we use the partition $\mathcal{P}$. Therefore, we have shown that $f \in \mathcal{R}_\alpha[a, b]$. □

Food for thought 4. What if f and α are both discontinuous at a common point? If $f \in \mathcal{R}_\alpha[a, b]$ always? Does it depend on f and α? Or is this never true?

Solution. Consider the case
$$f(x) = \alpha(x) = \begin{cases} 0 & x < 0 \\ 1 & x \ge 0 \end{cases}.$$
Let $\mathcal{P} = \{-1 = x_0, x_1, x_2, \ldots, x_n = 1\}$ be the partition on the interval $[-1, 1]$. There are two cases that we need to consider here: if we look at
$$U(\mathcal{P}, f, \alpha) - L(\mathcal{P}, f, \alpha) = \sum_{i=1}^{n} (M_i - m_i)\Delta\alpha_i,$$
the only interesting term is the one about the point of discontinuity (the origin); we can choose our partition to be such that the origin in between two points of the partition, but if we work this out, we get that $M_k - m_k = 1 - 0 = 1$, and $\alpha_k - \alpha_{k-1} = 1$, so

$f \notin \mathcal{R}_\alpha[a,b]$. In this case that zero is one of the partition points, we end up getting the same thing, so this function turns out to not be Riemann-Stieltjes integrable with this integrator.

However, while keeping f the same, if we slightly change α to be

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

and we let the origin be one of the partition points, we see that over the interval $[s, 0]$, where $s \in \mathcal{P}$,

$$U(\mathcal{P}, f, \alpha) - L(\mathcal{P}, f, \alpha) = \sum_{i=1}^{n} (M_i - m_i)\Delta\alpha_i = (1-0)(0-0) = 0,$$

so this in fact is now Riemann-Stieltjes integrable. It is still Riemann-Stieltjes integrable if we consider an interval of the form $[0, t]$ for $t \in \mathcal{P}$.

This is particularly interesting because if we compute the integral $\int_{-1}^{1} f\, d\alpha$, we get that it evaluates to zero, which means in this case, even though it is integrable, the integrator was unable to detect the step up in the function. So we conclude that if the function and the integrator share a point of discontinuity, then sometimes the function is still Riemann-Stieltjes integrable. However, funny things happen in such situations. □

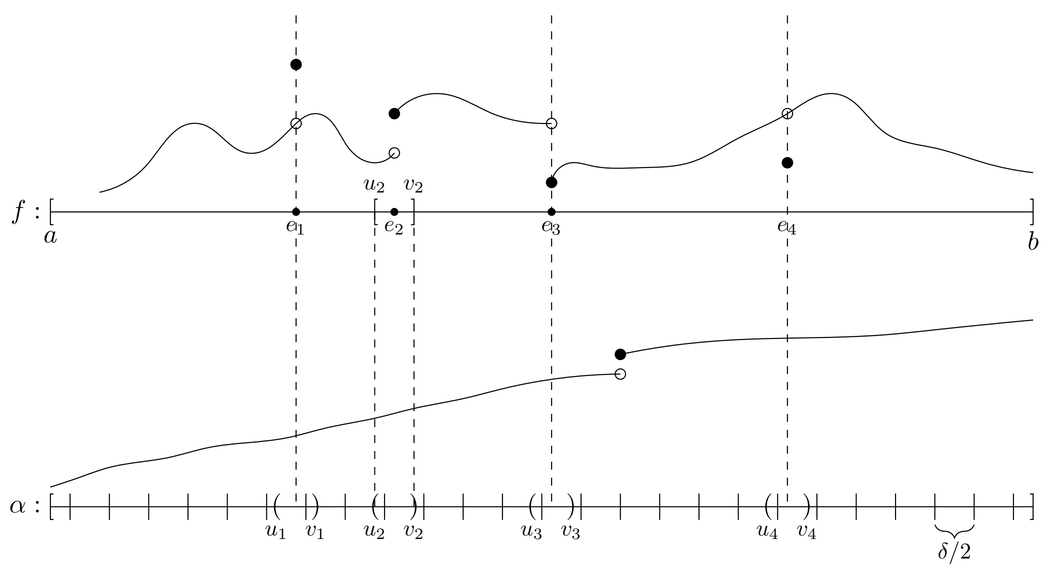

Figure 4: "Proof by picture" for the theorem.

Theorem: Baby Rudin 6.11

Let $f : [a,b] \to \mathbb{R}$ be bounded, and $\alpha : [a,b] \to \mathbb{R}$ be monotone increasing. Suppose $f \in \mathcal{R}_\alpha[a,b]$. Suppose $m \leq f(x) \leq M$ for all $x \in [a,b]$. Let $\varphi : [m, M] \to \mathbb{R}$ be continuous; then $\varphi \circ f \in \mathcal{R}_\alpha[a,b]$.

Proof. Given in Rudin. □

Quotes of the day: Dr. Joshua Zahl 01/22/2024

No quotes today :(

Theorem: Properties of the Riemann-Stieltjes integral (Baby Rudin 6.12)

Let $\alpha : [a, b] \to \mathbb{R}$ be monotonically increasing, and $f, f_1, f_2 : [a, b] \to \mathbb{R}$ be functions satisfying $f, f_1, f_2 \in \mathcal{R}_\alpha[a, b]$.

a) Linearity: $f_1 + f_2 \in \mathcal{R}_\alpha[a, b]$ and $\int_a^b (f_1 + f_2)\, d\alpha = \int_a^b f_1\, d\alpha + \int_a^b f_2\, d\alpha$. For $c \in \mathbb{R}$, $cf \in \mathcal{R}_\alpha[a, b]$ and $\int_a^b cf\, d\alpha = c\int_a^b f\, d\alpha$.

b) Weak positivity/non-negativity: If $f(x) \geq 0$ for all $x \in [a, b]$, then $\int_a^b f\, d\alpha \geq 0$.

If $f_1(x) \leq f_2(x)$ for all $x \in [a, b]$, then $\int_a^b f_1\, d\alpha \leq \int_a^b f_2\, d\alpha$.

c) For $c \in [a, b]$, $f \in \mathcal{R}_\alpha[a, c]$ and $f \in \mathcal{R}_\alpha[c, b]$, and

$$\int_a^c f\, d\alpha + \int_c^b f\, d\alpha = \int_a^b f\, d\alpha.$$

d) Boundedness: If $|f| \leq M$, then $\left|\int_a^b f\, d\alpha\right| \leq M\,(\alpha(b) - \alpha(a))$.

e) Let $\alpha_1, \alpha_2 : [a, b] \to \mathbb{R}$ br monotone increasing, and $f : [a, b] \to \mathbb{R}$ satisfying $f \in \mathcal{R}_{\alpha_1}[a, b]$ and $f \in \mathcal{R}_{\alpha_2}[a, b]$. Then, $f \in \mathcal{R}_{\alpha_1+\alpha_2}[a, b]$, 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 $c \in \mathbb{R}$, $f \in \mathcal{R}_{c\alpha_1}[a, b]$, and $\int_a^b f\, d(c\alpha_1) = c\int_a^b f\, d\alpha_1$.

Proof. The proof is given on page 128 of Baby Rudin; it's not very involved, so can be treated as an exercise as well. □

Recall $\mathcal{C}([a, b])$, the space of continuous functions $f : [a, b] \to \mathbb{R}$. Define $\|f\|_{\mathcal{C}([a,b])} = \sup_{x\in[a,b]} |f(x)|$. Hence, the metric is $d(f, g) = \|f - g\|_{\mathcal{C}([a,b])}$. We say that the pair $\left(\mathcal{C}([a, b]), \|\cdot\|_{\mathcal{C}([a,b])}\right)$ is a *normed vector space*.

Property a) of theorem 6.12 says: If $\alpha : [a, b] \to \mathbb{R}$ is monotone increasing, then the function $T(f) = \int_a^b f\, d\alpha$ is a linear function from the vector space $\mathcal{C}([a, b])$ to $\mathbb{R}$. Hence,

$$T(f + g) = T(f) + T(g)$$
$$T(cf) = cT(f).$$

Property d) says that T is bounded, i.e., $|T(f)| \leq (\alpha(b) - \alpha(a))\, \|f\|_{\mathcal{C}([a,b])}$.

Notation 2. People sometimes write Tf instead of $T(f)$, however it's the same thing. For example, in linear algebra, we write Mv where M is a matrix and v is a vector, but this is technically $M(v)$.

Property b) says that T is non-negative, i.e., if $f \in \mathcal{C}([a,b])$ with $f(x) \geq 0$ for all $x \in [a,b]$. Then $Tf \geq 0$.

In functional analysis (MATH 421), and more generally in Physics, we want to study linear functions whose domain is $\mathcal{C}([a,b])$ (or more general), and whose co-domain is $\mathbb{R}$ (or more often $\mathbb{C}$). Functions of this type are called "linear operators" or "linear functionals".

Theorem: Riesz Representation Theorem 1.0

Let $T : \mathcal{C}([a,b]) \to \mathbb{R}$ be linear, bounded, and non-negative. Then, there exists a unique monotone increasing $\alpha : [a,b] \to \mathbb{R}$, such that $Tf = \displaystyle\int_a^b f\, d\alpha$.

We want to find a better version of the theorem where we can drop the non-negative hypothesis:

Theorem: Riesz Representation Theorem 2.0

Let $T : \mathcal{C}([a,b]) \to \mathbb{R}$ be linear and bounded. Then, there exist two monotone increasing functions $\alpha, \beta : [a,b] \to \mathbb{R}$ such that

$$T(f) = \int_a^b f\, d\alpha - \int_a^b f\, d\beta = \int_a^b f\, d(\alpha - \beta).$$

Note (Extension of the definition of the Riemann-Stieltjes integral)**.** Note that for monotone increasing α, β, $\alpha - \beta$ is not necessarily monotonically increasing, so we would have to change the definition of the Riemann-Stieltjes integral from monotonically increasing α to α that is the difference of monotonically increasing functions. However, we don't really need to get into that since we can just write it as the first equality shown above.

Quotes of the day: Dr. Joshua Zahl 01/24/2024

"I love nitpicking, because math is meant to be precise."

Theorem: Baby Rudin 6.13

Let $f, g \in \mathcal{R}_\alpha[a,b]$. Then

(a) Then $fg \in \mathcal{R}_\alpha[a,b]$.

(b) Then $|f| \in \mathcal{R}_\alpha[a,b]$, and $\left|\displaystyle\int_a^b f\, d\alpha\right| \leq \displaystyle\int_a^b |f|\, d\alpha$.

(a) *Proof.* By theorem 6.11, $\alpha(x) = x^2$, $(f+g)^2, (f-g)^2 \in \mathcal{R}_\alpha[a,b]$. By theorem 6.12(a), $(f+g)^2 - (f-g)^2 = 4fg \in \mathcal{R}_\alpha[a,b]$. Finally, by theorem 6.12(a), for $c = \frac{1}{4}$, $fg \in \mathcal{R}_\alpha[a,b]$. □

(b) *Proof.* By theorem 6.11, $\alpha(x) = |x|$, $|f| \in \mathcal{R}_\alpha[a,b]$. Let $c = \operatorname{sgn} \displaystyle\int_a^b f\, d\alpha$, so

$$\left|\int_a^b f\, d\alpha\right| = c\int_a^b f\, d\alpha = \int_a^b cf\, d\alpha \leq \int_a^b |f|\, d\alpha.$$

(by theorem 6.12(a)). □

Theorem: Baby Rudin 6.15

Let $f : [a, b] \to \mathbb{R}$ be bounded. Let $s \in (a, b)$ and suppose f is continuous at s. Let

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

Then, $f \in \mathcal{R}_\alpha[a, b]$, and $\int_a^b f\, d\alpha = f(s)$.

Proof. Let $\mathcal{P} = \{x_0, x_1, x_2, x_3\}$ where $a = x_0,\ s = x_1,\ b = x_3$. Then, $U(\mathcal{P}, f, \alpha) = \sum_{i=1}^{3} M_i \Delta\alpha_i = M_2 = \sup_{x \in [x_1, x_2]} f(x)$ and $L(\mathcal{P}, f, \alpha) = \sum_{i=1}^{3} m_i \Delta\alpha_i = m_2 = \inf_{x \in [x_1, x_2]} f(x)$. Since f is continuous at s, for all $\varepsilon > 0$, there exists δ such that if $|x - y| < \delta$, then $|f(x) - f(y)| < \frac{\varepsilon}{2}$, so let $x_2 \in (x_1, x_1 + \delta)$. Then,

$$\begin{aligned} \sup_{x \in [x_1, x_2]} f(x) \leq& f(x_1) + \frac{\varepsilon}{2} \Rightarrow M_2 \leq f(x_1) + \frac{\varepsilon}{2} \\ \inf_{x \in [x_1, x_2]} f(x) \geq& f(x_1) - \frac{\varepsilon}{2} \Rightarrow m_2 \geq f(x_1) - \frac{\varepsilon}{2}. \end{aligned}$$

Hence, $M_2 - m_2 \leq \varepsilon$. $\square$

Food for thought 5. How would we change the proof if $f(x) = 1$ or $x \geq s$? We would get the same result, but would need to change the roles of x_1, x_2. If defined at neither, then probably $s \in [x_1, x_2]$. This step function is quite important in electrical engineering; it even has a special name:

Definition: Heavyside step function

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

Theorem: Baby Rudin 6.16

Let $\{c_n\}_{n=1}^{\infty}$ be positive real numbers with $\sum_{n=1}^{\infty} c_n < \infty$. Let $[a, b]$ be an interval, and let $\{s_n\}_{n=1}^{\infty} \subseteq (a, b)$ be distinct points. Let $\alpha(x) = \sum_{n=1}^{\infty} x_n I(x - s_n)$ (a bunch of steps at s_n by c_n). Let $f : [a, b] \to \mathbb{R}$ be continuous, so $f \in \mathcal{R}_\alpha[a, b]$ (always true if f is continuous and α is monotone increasing). Then

$$\int_a^b f\, d\alpha = \sum_{n=1}^{\infty} c f(s_n).$$

Note. We know that both the integral and sum exist because the sum converges, and f is bounded.

Proof. Let $R_N = \int_a^b f\, d\alpha - \sum_{n=1}^{N} c_n f(s_n)$. Our goal is to show that for all $\varepsilon > 0$, there exists N_0 such that for all $n \geq N_0$, $|R_N| < \varepsilon$. So, fix N, let $\alpha_1 = \sum_{n=1}^{N} c_n I(x - s_n)$, $\alpha_2 = \sum_{n=N+1}^{\infty} c_n I(x - s_n)$. By theorem 6.12(c), $\int_a^b f\, d\alpha = \int_a^b f\, d\alpha_1 + \int_a^b f\, d\alpha_2$, and

$f \in \mathcal{R}_{\alpha_1}[a,b]$, $f \in \mathcal{R}_{\alpha_2}[a,b]$ since f is continuous. Hence,

$$\int_a^b f\,d\alpha_1 = \sum_{n=1}^{N} \int_a^b f(x)\,d[c_n I(x-s_n)] = \sum_{n=1}^{N} c_n f(s_n),$$

by theorem 6.15. Therefore, $R_N = \displaystyle\int_a^b f\,d\alpha_2$. Let $K = \displaystyle\sup_{x\in[a,b]} |f|$. By theorem 6.12(b),

$$\int_a^b f\,d\alpha_2 \le K \int_a^b 1\,d\alpha_2 = K[\alpha_2(b) - \alpha_2(a)] = K \sum_{n=N+1}^{\infty} c_n \to 0 \text{ as } N \to \infty.$$

In summary, we wrote the difference as the sum of a main term and a tail term, and showed that the tail term goes to zero. We could make this more formal using ε's, but we were out of time. $\square$

Theorem: Baby Rudin 6.17

Let $f : [a,b] \to \mathbb{R}$ be bounded, and $\alpha : [a,b] \to \mathbb{R}$ be differentiable and monotone increasing. Suppose $\alpha' \in \mathcal{R}[a,b]$. Then $f \in \mathcal{R}_\alpha[a,b] \iff f\alpha' \in \mathcal{R}[a,b]$, and if so

$$\int_a^b f\,d\alpha = \int_a^b f\alpha'\,dx.$$

Proof. **Step-1**: It suffices to show that

$$\overline{\int_a^b} f\,d\alpha = \overline{\int_a^b} f\alpha'\,dx \quad \text{and} \quad \underline{\int_a^b} f\,d\alpha = \underline{\int_a^b} f\alpha'\,dx.$$

We will prove the first equality, and the second one is left as an exercise.

Step-2: Since $\alpha' \in \mathcal{R}[a,b]$ for all $\varepsilon > 0$, there exists $\mathcal{P}$ (of $[a,b]$) such that $U(\mathcal{P},\alpha') - L(\mathcal{P},\alpha') < \varepsilon$. This inequality continues to hold for every refinement $\mathcal{P}'$ of $\mathcal{P}$.

We have $U(\mathcal{P},\alpha') - L(\mathcal{P},\alpha') = \displaystyle\sum_{i=1}^{n}(A_i - a_i)\Delta x_i$, where $A_i := \sup\{\alpha'(x) : x \in [x_{i-1},x_i]\}$, and $a_i := \inf\{\alpha'(x) : x \in [x_{i-1},x_i]\}$. By the Mean Value Theorem for each $i = 1,\dots,n$, there exists $t_i \in [x_{i-1},x_i]$ such that $\Delta\alpha_i = \alpha'(t_i)\Delta x_i$. Now, this suggests $\displaystyle\overline{\int_a^b} f\,d\alpha = \overline{\int_a^b} f\alpha'\,dx$, but we need to be careful:

For every $s_i \in [x_{i-1},x_i]$, we have $|\alpha'(s_i) - \alpha'(t_i)| \le A_i - a_i$, so $\displaystyle\sum_{i=1}^{n} |\alpha'(s_i) - \alpha'(t_i)|\Delta x_i \le \sum_{i=1}^{n}(A_i - a_i)\Delta x_i$ for *every* choice of $s_i \in [x_{i-1},x_i]$, $1 \le i \le n$. Let $K := \displaystyle\sup_{x\in[a,b]} |f|$, then

$$\sum_{i=1}^{n} |f(s_i)\alpha'(s_i)\Delta x_i - f(s_i)\alpha'(t_i)\Delta x_i| \le K\varepsilon. \tag{1}$$

Hence,

$$\sum_{i=1}^{n} f(s_i)\Delta\alpha_i \le \sum_{i=1}^{n} f(s_i)\alpha'(s_i)\Delta x_i + K\varepsilon \le U(\mathcal{P}, f\alpha') + K\varepsilon.$$

Recall that if $|a+b| \le c$, then $a \le b + c$ and $b \le a + c$.

Note. Whenever we have an inequality, we might wonder whether it is "sharp" or "tight", meaning it is equality or the closest to equality as possible.

Taking the supremum of $s_i \in [x_{i-1}, x_i]$, $i = 1, \ldots, n$, we conclude

$$U(\mathcal{P}, f, \alpha) = \sum_{i=1}^{n} M_i \Delta\alpha_i \leq U(\mathcal{P}, f\alpha') + K\varepsilon.$$

Hence,

$$\overline{\int_a^b} f\, d\alpha \leq U(\mathcal{P}, f, \alpha) \leq U(\mathcal{P}, f\alpha') + K\varepsilon,$$

i.e., for every ε, we found some partition $\mathcal{P}$ that makes the inequality hold. Recall that it also holds for every refinement $\mathcal{P}'$ of $\mathcal{P}$.

How can we make this inequality be as close to equality as possible (how much strength can we squeeze out of the inequality?) Taking the infimum over all refinements $\mathcal{P}'$ of $\mathcal{P}$, we have

$$\overline{\int_a^b} f\, d\alpha \leq \int_{\mathcal{P}'} U(\mathcal{P}', f\alpha') + K\varepsilon.$$

Since we are considering a more strict set of partitions, will this give us the infimum we want? The answer is yes: for any non-refinement partition, we union it with $\mathcal{P}$ to get a refinement, i.e., for all $\varepsilon > 0$, $\overline{\int_a^b} f\, d\alpha \leq \overline{\int_a^b} f\alpha'\, dx + K\varepsilon \Rightarrow \overline{\int_a^b} f\, d\alpha \leq \overline{\int_a^b} f\alpha'\, dx$. The other three inequalities can be done as an exercise after the following note:

Note. For the inequality

$$\overline{\int_a^b} f\, d\alpha \geq \overline{\int_a^b} f\alpha'\, dx,$$

from eq. (1), we get

$$\sum_{i=1}^{n} f(s_i)\alpha'(s_i)\Delta x_i \leq \sum_{i=1}^{n} f(s_i)\Delta\alpha_i + K\varepsilon,$$

and so on.

□

Theorem: Baby Rudin 6.19

Let $\varphi : [A, B] \to [a, b]$ be a strictly increasing, surjective, and continuous function.

Let $\alpha : [a, b] \to \mathbb{R}$ be our monotone increasing integrator, and let $f \in \mathcal{R}_\alpha[a, b]$.

Define $g := f \circ \varphi : [A, B] \to \mathbb{R}$, and $\beta := \alpha \circ \varphi : [A, B] \to \mathbb{R}$. Hence, $g \in \mathcal{R}_\beta[A, B]$ and

$$\int_A^B g\, d\beta = \int_a^b f\, d\alpha.$$

Example 4. Let $\alpha(x) = x$, and φ is differentiable. Then $d\beta = \varphi'(x)\, dx$, i.e.,

$$\int_a^b f\, d\alpha = \int_{\varphi^{-1}(a)}^{\varphi^{-1}(b)} f(x)\, \varphi'(x) dx$$

Proof. Partitions $\mathcal{P} := \{a = x_0, x_1, \ldots, x_n = b\}$ of $[a, b]$, and partitions $\mathcal{Q}$ of $[A, B]$ are in 1-1 correspondence via $x_i = \varphi(y_i)$.

We have $\alpha(x_i) = \alpha \circ \varphi(y_i) = \beta(y_i)$, and

$$\{f(x) : x \in [x_{i-1}, x_i]\} = \{g(y) : y \in [y_{i-1}, y_i]\}.$$

Hence, $U(\mathcal{P}, f, \alpha) = U(\mathcal{Q}, g, \beta)$ and $L(\mathcal{P}, f, \alpha) = L(\mathcal{Q}, g, \beta)$. For all $\varepsilon > 0$, since $f \in \mathcal{R}_\alpha[a, b]$ there exists a partition $\mathcal{P}$ of $[a, b]$ such that $U(\mathcal{P}, f, \alpha) - L(\mathcal{P}, f, \alpha) < \varepsilon$. Therefore, $U(\mathcal{Q}, f, \alpha) - L(\mathcal{Q}, f, \alpha) < \varepsilon$, and $g \in \mathcal{R}_\beta[A, B]$.

Finally,

$$\int_A^B g\,d\beta = \inf_{\mathcal{Q}} U(\mathcal{Q}, g, \beta) = \inf_{\mathcal{P}} U(\mathcal{P}, f, \alpha) = \int_a^b f\,d\alpha.$$

$\square$

Note (About the properties of φ)**.** Without φ being strictly increasing, surjective, and continuous in the theorem hypothesis, we won't get a 1-1 correspondence between the partition $\mathcal{P}$ and the partition $\mathcal{Q}$.

Theorem: Baby Rudin 6.20

Let $f \in \mathcal{R}[a, b]$. For $x \in [a, b]$, define $F(x) := \int_a^x f(t)\,dt$, $F(a) := 0$. Then, F is continuous on $[a, b]$. If $c \in [a, b]$, and f is continuous at c, then F is differentiable at c, and the derivative of $F'(c) = f(c)$.

Proof. *Continuity*: Let $K = \sup_{t\in[a,b]} |f(t)|$. By theorem 6.12(c), for $a \le x \le y \le b$,

$$F(y) - F(x) = \int_a^y f(t)\,dt - \int_a^x f(t)\,dt = \int_x^y f(t)\,dt.$$

Thus,

$$|F(y) - F(x)| = \left|\int_x^y f(t)\,dt\right| \le \int_x^y |f(t)|\,dt \le \int_x^y K\,dt = K(y - x).$$

Hence, for every $\varepsilon > 0$, select $\delta = \frac{\varepsilon}{K}$ (or $\delta = \varepsilon$ if $K = 0$); if $|x - y| < \delta$, $|f(x) - f(y)| < \varepsilon$.

Differentiability at c: Suppose $c \neq b$, i.e., $c \in [a, b)$. Let us compute $\lim_{h\searrow 0} \frac{F(c+h) - F(c)}{h}$.

For $h > 0$, we have

$$\left|\frac{1}{h}[F(c+h) - F(c)] - f(c)\right| = \left|\frac{1}{h}\int_c^{c+h} f(t)\,dt - f(c)\right|.$$

Here we exploit a trick, where we write $f(c) = \frac{1}{h}\int_c^{c+h} f(c)\,dt$. Hence, we have

$$\left|\frac{1}{h}[F(c+h) - F(c)] - f(c)\right| = \left|\frac{1}{h}\int_c^{c+h} [f(t) - f(c)]\,dt\right| \le \frac{1}{h}\int_c^{c+h} |f(t) - f(c)|\,dt.$$

Since f is continuous at c, for all $\varepsilon > 0$, there exists $\delta > 0$ such that for all $y \in [a, b]$ with $|y - c| < \delta$, we have $|f(c) - f(y)| < \varepsilon$. Hence, for $h < \delta$, we have

$$\frac{1}{h}\int_c^{c+h} |f(t) - f(c)|\,dt < \frac{1}{h}\int_c^{c+h} \varepsilon\,dt = \varepsilon,$$

i.e., for all $\varepsilon > 0$, there exists $\delta > 0$ such that for all $0 < h < \delta$,

$$\left|\frac{1}{h}\left(F(c+h) - F(c)\right) - f(c)\right| < \varepsilon \Rightarrow \lim_{h\searrow 0}\frac{F(c+h) - F(c)}{h} = f(c).$$

If $c \neq a$, i.e., if $c \in (a, b]$, an identical argument shows $\lim_{h\nearrow 0}\frac{F(c+h) - F(c)}{h} = f(c)$. $\square$

Food for thought 6. If f is *not* continuous at c, is F

(a) *Never* differentiable at c.

(b) Maybe differentiable (depends on f and c).

(c) Always differentiable.

The answer to this should be (b) Maybe differentiable, since we could have a removable discontinuity, which the Riemann integral cannot see, so it will be just fine: If $f(x) = g(x)$ *except* at one point, then $\int_a^b f(x)\,dx = \int_a^b g(x)\,dx$. In contrast if it was even a jump discontinuity, f fails to be continuous, and hence it does not work.

Quotes of the day: Dr. Joshua Zahl 01/31/2024

No quotes today :(

We showed last time that if $f : [a, b] \to \mathbb{R}$ continuous, and $F(x) = \int_a^b f(t)\,dt$, then $F'(x) = f(x)$ for all $x \in [a, b]$.

Theorem: Fundamental theorem of Calculus (Baby Rudin 6.21)

Let $f \in \mathcal{R}[a, b]$, let $F : [a, b] \to \mathbb{R}$ be differentiable and suppose $F'(x) = f(x)$ for $x \in [a, b]$. Then

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

Proof. By MV, for any partition $P = \{x_0, \dots, x_n\}$ there are numbers $t_i \in [x_{i-1}, x_i]$ for $1 \leq i \leq n$ such that $F'(t_i) = (F(x_i) - F(x_{i-1}))/\Delta x_i$. So,

$$\begin{aligned} F(b) - F(a) &= \sum_{i=1}^{n} (F(x_i) - F(x_{i-1})) \\ &= \sum_{i=1}^{n} F'(t_i)\Delta x_i. \end{aligned}$$

Then,

$$\left| \int_a^b f\,dx - (F(b) - F(a)) \right| \leq U(P, f) - L(P, f).$$

Since $f \in \mathcal{R}[a, b]$, then for all $\varepsilon > 0$, there exists a partition P such that $U(P, f) - L(P, f) < \varepsilon$, and therefore,

$$\left| \int_a^b f\,dx - (F(b) - F(a)) \right| < \varepsilon.$$

□

This sets us up for proving things we know to be true about integration. We start by integration parts:

Theorem: Integration by parts (Baby Rudin 6.22)

Let $F, G : [a, b] \to \mathbb{R}$ be differentiable. Let $f = F'$, $g = G'$, and suppose $f, g \in \mathcal{R}[a, b]$. Then

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

Proof. Let $H(x) = F(x)G(x)$. Then $H'(x) = f(x)G(x) + F(x)g(x) \in \mathcal{R}[a, b]$. Apply Theorem 6.21 to H, then

$$H(b) - H(a) = \int_a^b H'(x)\,dx$$

i.e.,

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

□

In both os these results, we have this hypothesis that $f, g \in \mathcal{R}[a,b]$.

Food for thought 7. If $F; [a,b] \to \mathbb{R}$ is differentiable and $F' = f$, do we need repeat $f \in \mathcal{R}[a,b]$, or does this hold automatically, i.e., is $F' \in \mathcal{R}[a,b]$ for every $F : [a,b] \to \mathbb{R}$ differentiable?

If we ask that there exists $F : [a,b] \to \mathbb{R}$ differentiable, so that F' is discontinuous at every $x \in [a,b]$? The professor noted that "we've replaced a hard question with a harder question." We won't be doing this in class, but the answer to this question is *no*.

It is an interesting question: which sets can be the set of discontinuities of a derivative? We get that $S \subseteq [0,1]$, so can we find an F' that is discontinuous at S (where $F : [0,1] \to \mathbb{R}$ is differentiable). These are called $F - \delta$ sets.

Perhaps we wish for the derivative to blow up, but then it isn't Riemann integrable; here is a function that is worth remembering:

$$F(x) = \begin{cases} 0 & x = 0 \\ x^2 \sin \frac{1}{x^2} & x \neq 0 \end{cases}.$$

This function is differentiable, but its derivative is unbounded. On $x_n = \dfrac{1}{\sqrt{\pi n}}$, $F'(x_n)$ blows up.

Another type of counter-example is: F' is bounded, but F' is discontinuous at so many places that it is not Riemann integrable. Uncountable is not enough in this case: they might still be Riemann integrable. The condition is that it is discontinuous at points with positive Lebesgue measure: we try to cover all the discontinuities with open intervals, the smallest we can make the intervals will always add up to a positive value. However, this is a MATH 420 topic.

We will explore some definitions:

Definition: Absolute convergence of an integral

If $f : [a,\infty) \to \mathbb{R}$ satisfies $f \in \mathcal{R}[a,b]$ for all $b > a$, then we define

$$\int_a^\infty f(x)\,dx = \lim_{b\to\infty} \int_a^b f(x)\,dx.$$

If the limit exists (as a real number), we say that the $\displaystyle\int_a^\infty |f|\,dx$ exists (as a real number), then we say that $\displaystyle\int_a^\infty f(x)\,dx$ ***converges absolutely***.

Note. This is the same idea as conditional/absolute convergence of a sequence. We can make an equivalent definition for $\displaystyle\int_{-\infty}^b f(x)\,dx$.

If $f : \mathbb{R} \to \mathbb{R}$ and both $\displaystyle\int_0^\infty f(x)\,dx$ and $\displaystyle\int_{-\infty}^0 f(x)\,dx$ converges (absolutely), we define

$$\int_{-\infty}^\infty f(x)\,dx := \int_{-\infty}^0 f(x)\,dx + \int_0^\infty f(x)\,dx,$$

and we say $\displaystyle\int_{-\infty}^\infty f(x)\,dx$ converges (absolutely).

Food for thought 8. Can we construct a function that converges absolutely?

Taking inspiration from series, we can take a step function of $\frac{(-1)^n}{n}$; this converges conditionally, but not absolutely.

Quotes of the day: Dr. Joshua Zahl 02/02/2024

No quotes today :(

Definition: Riemann-Stieltjes integrability of complex valued functions

Let $f : [a, b] \to \mathbb{C}$, and $\alpha : [a, b] \to \mathbb{R}$ be monotone increasing. We say that $f \in \mathcal{R}_\alpha[a, b]$ if $\operatorname{Re} f \in \mathcal{R}_\alpha[a, b]$ and $\operatorname{Im} f \in \mathcal{R}_\alpha[a, b]$, and if so we define

$$\int_a^b f\, d\alpha := \int_a^b \operatorname{Re} f\, d\alpha + i \int_a^b \operatorname{Im} f\, d\alpha.$$

This definition can be extended for $\int_a^\infty f\, d\alpha$ and $\int_{-\infty}^\infty f\, d\alpha$ as well, and also for $f : [a, b] \to F^n$, where F is a field (generally just $\mathbb{R}$ or $\mathbb{C}$).

Food for thought 9. What happens if $\alpha : [a, b] \to \mathbb{C}$? We cannot make these "monotone increasing" because there is no order in the complex plane.

Answer. If α is differentiable, we can still try to use the Riemann integral, i.e.,

$$\int f\, d\alpha = \int f\alpha'\, dx.$$

We might also try to define what a "complex function of bounded variation" looks like. However, in practice, this is not really used since by the time we get around to thinking about complex integrators, we are usually working with Lebesgue integration.

3 Sequences and series of functions

3.1 The setup

Let $\mathcal{E}$ be a set ($\mathcal{E} = [a, b]$, $\mathcal{E} = \mathbb{R}$). Let $\{f_n\}_{n\in\mathbb{N}}$ be a sequence of functions $f_n : \mathcal{E} \to \mathcal{S}$ – usually $\mathcal{S} = \mathbb{R}$, $\mathcal{S} = \mathbb{C}$, or some metric space $(\mathcal{M}, d)$ – and let $f : \mathcal{E} \to \mathcal{S}$.

What does it mean for f_n to converge to f, i.e., $f_n \to f$, and the functions f_n all have some property P (e.g. all continuous, all integrable, etc.), must it be true that f also has P?

3.2 Point-wise convergence

Definition: Pointwise convergence

Let $\{f_n\}_{n\in\mathbb{N}}$ be a sequence of functions $F_n : \mathcal{E} \to \mathcal{M}$. If the sequence $\{f_n(x)\}_{n\in\mathbb{N}}$ converges for every $x \in \mathcal{E}$, then we say $\{f_n\}_{n\in\mathbb{N}}$ ***converges point-wise*** (on $\mathcal{E}$).

Note that there are two definitions that are used for convergence of sequence of functions:

- **Uniform convergence**: For all $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that for all $n > N$ and for all $x \in \mathcal{E}$, $d(f_n(x), f(x)) < \varepsilon$.
- **Point-wise convergence**: For all $x \in \mathcal{E}$, given $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that for all $n > N$, $d(f_n(x), f(x)) < \varepsilon$.

Uniform convergence is something we will cover later on in the course.

For point-wise convergence, what properties hold?

(a) Continuity: $\mathcal{E} = \mathbb{R}$. If each f_n is continuous and $f_n \to f$ point-wise, must it be true that f is continuous?

This is not necessarily true; consider

$$f_n = e^{-nx^2}, \quad \text{and} \quad f = \begin{cases} 1, & x = 0 \\ 0, & x \neq 0 \end{cases},$$

but $f_n \to f$. Also,

$$f_n = x^n, \quad \text{and} \quad f = \begin{cases} 1, & x = 1 \\ 0, & x \neq 1 \end{cases},$$

where $\mathcal{E} = [0, 1]$, but $f_n \to f$. In fact, we can make our points of discontinuity arbitrary where we want.

(b) Boundedness: If each f_n is bounded and $f_n \to f$ point-wise, must f be bounded?

This is not necessarily true; consider for $\mathcal{E} = (0, 1)$, $f_n = \dfrac{1}{x - a_n}$ such that $a_n \to 0^-$. Also, for $\mathcal{E} = \mathbb{R}$,

$$f_n = \begin{cases} x, & x \in [-n, n] \\ n, & x \in [n, \infty) \\ -n, & x \in (-\infty, n] \end{cases},$$

where $f_n \to f(x) = x$, but f_n is bounded for each $n \in \mathbb{N}$, whereas $f(x) = x$ is unbounded.

(c) Quantitative boundedness: If each f_n is bounded by 1, i.e., $|f_n(x)| \leq 1$, for all $x \in \mathcal{E}$, and $f_n \to f$ point-wise, must it be true that $|f(x)| \leq 1$ for all $x \in \mathcal{E}$?

This is in fact true: for the sake of contradiction, assume $|f(x')| > 1$ for some $x' \in \mathcal{E}$. Hence, $|f(x')| = 1 + \varepsilon > 1$, there must be some $f_n(x') > 1$.

If we change the condition in the hypothesis to be strictly less than 1, this is no longer true; we can get $|f(x)| = 1$.

(d) Riemann integrability: Given $f_n \in \mathcal{R}[0, 1]$, $f_n \to f$ point-wise, must it be true that $f \in \mathcal{R}[0, 1]$?

This is not necessarily true; we don't know many functions that fail to be integrable, but one example is

$$f(x) = \begin{cases} 1, & x \in \mathbb{Q} \\ 0, & x \notin \mathbb{Q} \end{cases}, \quad \text{and} \quad f_n(x) = \begin{cases} 1, & x \in \{q_1, q_2, \ldots, q_n\} \\ 0, & \text{otherwise} \end{cases},$$

where $\{q_n\}_{n \in \mathbb{N}}$ is an enumeration of the rationals.

Food for thought 10. Assuming f_n and f are both Riemann integrable, must $\displaystyle\int_a^b f_n(x)\,dx$ converge to $\displaystyle\int_a^b f(x)\,dx$?

Quotes of the day: Dr. Joshua Zahl 02/06/2024

No quotes today :(

We continue with solution of the problem from last time.

Solution. Essentially, we are asking that if $f_n - f \to 0$ point-wise, does it mean that $\displaystyle\lim_{n\to\infty} \int_0^1 g_n\,dx = 0$? This is in fact not necessarily true: consider

$$g_n(x) = \begin{cases} n, & 0 < x < \frac{1}{n} \\ 0, & \text{otherwise} \end{cases}.$$

We get $\int_0^1 g_n\,dx = 1$. An even more extreme counter-example

$$g_n(x) = \begin{cases} n^2, & 0 < x < \frac{1}{n} \\ 0, & \text{otherwise} \end{cases}, \quad \text{then} \quad \int_0^1 g_n\,dx = n.$$

□

Food for thought 11. We have seen examples of $f_n \to f$ point-wise, where f_n are continuous or integrable and the limit is not. Why does this happen?

Suppose $f_n \to f$, $f_n : \mathbb{R} \to \mathbb{R}$ and f_n are continuous at $c \in \mathbb{R}$. Is f continuous at c?

Solution. f is continuous at c if $\lim_{x\to c} f(x) = f(c)$, so f is continuous at c if

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

Similarly, $f \in \mathcal{R}[0,1]$ if $\lim_{m\to\infty}[U(P_m, f) - L(P_m, f)] = 0$.

If $f_n \to f$ point-wise, is it true that

$$\begin{aligned} \lim_{m\to\infty} U\left(P_m, \lim_{n\to\infty} f_n(x)\right) &= \lim_{n\to\infty}\lim_{m\to\infty} U(P_m, f_n) \\ \lim_{m\to\infty} L\left(P_m, \lim_{n\to\infty} f_n(x)\right) &= \lim_{n\to\infty}\lim_{m\to\infty} L(P_m, f_n)? \end{aligned}$$

This is true for a sequence of Lipschitz continuous functions, however the limit does not have to be Lipschitz continuous (due to failure to interchange limits).

Our final example showing that we *cannot* (in general) interchange limits. Consider $a_{n,m} = \dfrac{n}{n+m}$, where $n, m \in \mathbb{N}$; we get that

$$\begin{aligned} \lim_{n\to\infty} a_{n,m} = 1 &\Rightarrow \lim_{m\to\infty}\lim_{n\to\infty} a_{n,m} = 1. \\ \lim_{m\to\infty} a_{n,m} = 0 &\Rightarrow \lim_{n\to\infty}\lim_{m\to\infty} a_{n,m} = 0. \end{aligned}$$

□

Definition: Uniform convergence of a sequence of functions

Let $\mathcal{E}$ be a set. For a metric space $(\mathcal{M}, d)$ (i.e., $\mathbb{R}$ or $\mathbb{C}$), let $\{f_n\}$ be a sequence of functions $f_n : \mathcal{E} \to \mathcal{M}$ and $f : \mathcal{E} \to \mathcal{M}$. We say $f_n \to f$ ***uniformly*** is:

$$\text{For all } \varepsilon > 0, \text{ there exists } N \in \mathbb{N} \text{ such that, for all } n > N, x \in \mathcal{E}, d\left(f_n(x), f(x)\right) < \varepsilon.$$

Theorem: Cauchy criteria for uniform convergence (Baby Rudin 7.8)

Let $\mathcal{E}$ be a set, $(\mathcal{M}, d)$ a *complete* metric space. Then $\{f_n\}$ converges uniformly (to same $f : \mathcal{E} \to \mathcal{M}$) iff

$$\text{For all } \varepsilon > 0, \text{ there exists } N \in \mathbb{N} \text{ such that, for all } m, n > N, \text{ for all } x \in \mathcal{E}, d\left(f_n(x), f_m(x)\right) < \varepsilon,$$

which is the Cauchy criterion for uniform convergence.

Proof. ($\Rightarrow$) Suppose $f_n \to f$ uniformly. Let $\varepsilon > 0$; there exists $N \in \mathbb{N}$ such that for all $n > N$, for all $x \in \mathcal{E}$, $d(f_n(x), f(x)) < \dfrac{\varepsilon}{2}$. So for all $m, n > N$, for all $x \in \mathcal{E}$,

$$d(f_n(x), f_m(x)) \le d(f_n(x), f(x)) + d(f_m(x), f_n(x)) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

($\Leftarrow$) Suppose $\{f_n(x)\}$ satisfies the Cauchy criteria. For each $x \in \mathcal{E}$, $\{f_n(x)\}$ is a Cauchy sequence in $\mathcal{M}$, and $(\mathcal{M}, d)$ is complete, so $\{f_n(x)\}$ converges, i.e., $\lim_{n\to\infty} f_n(x)$ exists.

Define $f(x) = \lim_{n\to\infty} f_n(x)$. Let $\varepsilon > 0$; since $\{f_n(x)\}$ satisfies Cauchy criteria, there exists $N \in \mathbb{N}$ such that for all $m, n > N$, for all $x \in \mathcal{E}$, we have

$$d(f_n(x), f_m(x)) < \frac{\varepsilon}{2}.$$

Hence,

$$d(f_n(x), f(x)) \leq d(f_n(x), f_m(x)) + d(f_m(x), f(x)) < \frac{\varepsilon}{2} + d(f_m(x), f(x)).$$

Since $f_m(x) \to f(x)$, we can select $m > N$ such that $d(f_m(x), f(x)) < \frac{\varepsilon}{2}$. Therefore,

$$d(f_n(x), f(x)) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

□

Quotes of the day: Dr. Joshua Zahl 02/08/2024

No quotes today :(

Theorem: Baby Rudin 7.11

Let $(\mathcal{M}_1, d_1)$ and $(\mathcal{M}_2, d_2)$ be metric spaces with $(\mathcal{M}_2, d_2)$ complete, i.e., $\mathbb{R}$ or $\mathbb{C}$. Let $\mathcal{E} \subseteq \mathcal{M}_1$, and let $\{f_n\}$ be a sequence of functions $f_n : \mathcal{E} \to \mathcal{M}_2$, and suppose $f_n \to f$ uniformly on $\mathcal{E}$.

Let $x \in \mathcal{M}_1$ be a limit point of $\mathcal{E}$. Suppose $\lim_{t\to x} f_n(t) = y_n$ exists for each n; $\{y_n\}$ is a convergent sequence, i.e., $y_n \to y \in \mathcal{M}_2$, and $\lim_{t\to x} f(t) = y$, i.e.,

$$\lim_{t\to x} \underbrace{\lim_{n\to\infty} f_n(t)}_{f(t)} = \lim_{n\to\infty} \underbrace{\lim_{t\to x} f_n(t)}_{y_n}$$

Proof. **Step-1:** Show that $\{y_n\}$ converges.

It suffices to show that $\{y_n\}$ is Cauchy. Let $\varepsilon > 0$ be given. Choose N such that for all $m, n > N$, for all $t \in \mathcal{E}$, $d_2(f_n(t), f_m(t)) < \frac{\varepsilon}{3}$, and thus

$$\begin{aligned} d_2(y_n, y_m) \leq & d_2(y_n, f_n(t)) + d_2(f_n(t), y_m) \\ \leq & d_2(y_n, f_n(t)) + d_2(f_n(t), f_m(t)) + d_2(f_m(t), y_m). \end{aligned}$$

We can choose t such that the above is at most $< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon$; call this the "$\frac{\varepsilon}{3}$ trick".

In conclusion, for all $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that for all $m, n > N$, $d_2(y_n, y_m) < \varepsilon$, i.e., $\{y_n\}$ is Cauchy, and by completeness of $(\mathcal{M}_2, d_2)$, hence convergent.

Step-2: Prove that $f(t) \to y$ as $t \to x$.

For all $t \in \mathcal{E}$ and n,

$$d_2(f(t), y) \leq d_2(f(t), f_n(t)) + d_2(f_n(t), y_n) + d_2(y_n, y). \tag{$\star$}$$

Let $\varepsilon > 0$; since $f_n \to f$ uniformly, there exists $N_1 \in \mathbb{N}$ such that for all $n > N_1$, for all $t \in \mathcal{E}$,

$$d_2(f(t), f_n(t)) < \frac{\varepsilon}{3}.$$

Since $y_n \to y$, there exists $N_2 \in \mathbb{N}$ such that for all $n > N_2$, $d_2(y_n, y) < \dfrac{\varepsilon}{3}$. Let $N := \max\{N_1, N_2\}$. Applying eq. ($\star$) with this choice of N, we have

$$d_2(f(t), y) \leq \frac{\varepsilon}{3} + d_2(f_N(t), y_N) + \frac{\varepsilon}{3}.$$

Since $\lim_{t \to x} f_N(t) = y_N$, there exists $\delta > 0$ such that for all $t \in \mathcal{E}$, $d_1(t, x) < \delta$, we have $d_2(f_N(t), y_N) < \dfrac{\varepsilon}{3}$. Hence, for all $t \in \mathcal{E}$, for all x obeying $d_1(t, x) < \delta$, we have

$$d_2(f(t), y) < \varepsilon.$$

□

Corollary: Baby Rudin 7.12

Let $(\mathcal{M}_1, d_1)$, $(\mathcal{M}_2, d_2)$, $\{f_n\}$, f, and $\mathcal{E}$ be as before. If each f_n is continuous on $\mathcal{E}$, and $f_n \to f$ uniformly, then f is continuous on $\mathcal{E}$.

Effectively, "the uniform limit of continuous functions is continuous."

Proof. f is always continuous at isolated points, so we only need to consider limit points, $x \in \mathcal{E} \cap \mathcal{E}'$, For every such x, theorem 7.11 implies

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

□

3.3 Series of functions

Definition: Convergence of a series of functions to a function

Let $\mathcal{E}$ be a set, let $\{f_n\}$ be a sequence of functions, $f_n : \mathcal{E} \to \mathbb{R}$ or $\mathcal{E} \to \mathbb{C}$, and let $g : \mathcal{E} \to \mathbb{R}$ or $\mathcal{E} \to \mathbb{C}$. We say $\displaystyle\sum_{n \in \mathbb{N}} f_n$ converges point-wise (uniformly) to g is the sequence $S_n := \sum_{i=1}^{n} f_i$ converges point-wise (uniformly) to g.

Example 5. The series $1 + \displaystyle\sum_{n \in \mathbb{N}} \frac{x^n}{n!}$ converges to $g(x) = e^x$

- point-wise on $\mathbb{R}$.
- uniformly on any bounded set $\mathcal{E} \subseteq \mathbb{R}$, or any compact set $\mathcal{K} \subseteq \mathbb{R}$.

Theorem: Weierstraß M-test

Let $\mathcal{E}$ be a set, $f_n : \mathcal{E} \to \mathbb{R}$ or $\mathcal{E} \to \mathbb{C}$. If $|f_n(x)| \leq M$ for all $n > N_0 \in \mathbb{N}$, for all $x \in \mathcal{E}$, and if $\displaystyle\sum_{n=N_0}^{\infty} M_n < \infty$, then $\displaystyle\sum_{n \in \mathbb{N}} f_n$ converges uniformly.

Quotes of the day: Dr. Joshua Zahl 02/09/2024

No quotes today :(

Theorem: Dini's uniform convergence theorem (Baby Rudin 7.13)

Let $(\mathcal{M}, d)$ be a compact metric space (i.e., $[a, b]$), $\{f_n\}$ a sequence of functions, $f_n : \mathcal{M} \to \mathbb{R}$. Suppose that

(a) Each f_n is continuous.

(b) f_n converges *point-wise* to some continuous $f : \mathcal{M} \to \mathbb{R}$.

(c) $f_{n+1}(x) \geq f_n(x)$ for each $x \in \mathcal{M}$, $n \in \mathbb{N}$.

Then, $f_n \to f$ *uniformly* on $\mathcal{M}$.

Proof. Let $g_n = f - f_n$. Then, (a) g_n is continuous, (b) $g_n \to 0$ point-wise, (c) $g_n(x) \geq g_{n+1}(x) \geq 0$ for all $n \in \mathbb{N}$.

Goal: Prove $g_n \to 0$ uniformly, i.e.,

$$\text{For all } \varepsilon > 0, \text{ there exists } N \in \mathbb{N} \text{ such that for all } n > N, x \in \mathcal{M}, 0 \leq g_n(x) < \varepsilon.$$

Since g_n is monotonically decreasing, it is sufficient to show for all $x \in \mathcal{M}$, $g_n(x) < \varepsilon$.

Let $\mathcal{K}_n = g_n^{-1}([\varepsilon, \infty))$, $\mathcal{K}_n$ is closed, hence compact ($\mathcal{M}$ compact). Since $\{g_n\}$ is decreasing, $\mathcal{K}_n$ are nested, i.e., $\mathcal{K}_{n+1} \subseteq \mathcal{K}$. Since $g_n \to 0$ point-wise, for each $x \in \mathcal{M}$, there exists n such that $g_n(x) < \varepsilon \Rightarrow x \notin \mathcal{K}_n$. Since x was arbitrary, $\bigcap\limits_{n=1}^{\infty} \mathcal{K}_n = \emptyset$. By theorem 2.36, there exists $N \in \mathbb{N}$ such that $\mathcal{K}_N = \emptyset$, i.e.,

$$\begin{aligned} & g_N(x) < \varepsilon \quad \text{for all } x \in \mathcal{M} \\ \Rightarrow & g_n(x) < \varepsilon \quad \text{for all } x \in \mathcal{M}, n \geq N \\ \Rightarrow & |g_n(x)| < \varepsilon \quad \text{for all } x \in \mathcal{M}, n \geq N. \end{aligned}$$

□

Definition: Supremum norm

Let $(\mathcal{X}, d)$ be a non-empty metric space. Define

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

For each $f \in \mathscr{C}(\mathcal{X})$, define the "supremum norm"

$$\|f\| = \sup_{x \in \mathcal{X}} |f(x)|, \text{ for } f \in \mathscr{C}(\mathcal{X}), \|f\| < \infty.$$

Note. If $\mathcal{X}$ is compact in the above definition, f being bounded is superfluous.

Notation 3 (Alternate notation)**.** Some other notation for the supremum norm is: $\|f\|_{\mathscr{C}(\mathcal{X})}$, $\|f\|_{\mathscr{C}^0(\mathcal{X})}$, $\|f\|_\infty$, where the first one is probably the best one.

Note that $\mathscr{C}(\mathcal{X})$ is a vector space over $\mathbb{C}$, with $\|\cdot\|$ as the norm. For this, we have

1. $\|f\| \geq 0$, $\|f\| = 0$ iff $f(x) = 0$ for all $x \in \mathcal{X}$, i.e., $f = 0$.
2. For $\lambda \in \mathbb{C}$, $\|\lambda f\| = \lambda \|f\|$.
3. $|f(x) + g(x)| \leq |f(x)| + |g(x)| \leq \|f\| + \|g\| \Rightarrow \|f + g\| \leq \|f\| + \|g\|$.

Thus, if we define $\varrho(f, g) = \|f - g\|$, then ϱ is a metric, and $(\mathscr{C}(\mathcal{X}), \varrho)$ is a metric space. Therefore,

$$\begin{aligned} f_n \to f \text{ uniformly} & \iff \text{for all } \varepsilon > 0, \text{ there exists } N \in \mathbb{N} \text{ such that for all } x \in \mathcal{X}, \text{ for all } n > N, |f_n(x) - f(x)| < \varepsilon \\ & \iff \text{for all } \varepsilon > 0, \text{ there exists } N \in \mathbb{N} \text{ such that for all } n > N, \|f - f_n\| < \varepsilon \\ & \iff f_n \to f \text{ in the metric space } \mathscr{C}(\mathcal{X}). \end{aligned}$$

Theorem: Baby Rudin 7.15

$\mathscr{C}(\mathcal{X})$ is complete.

Proof. Let $\{f_n\}$ be a Cauchy sequence (in $\mathscr{C}(\mathcal{X})$). By theorem 7.8 (Cauchy's criteria), $f_n \to f$ uniformly for some $f : \mathcal{X} \to \mathbb{C}$. by corollary 7.12, f is continuous, since it is the uniform limit of a continuous function. Finally, f is bounded, and $f_n \to f$ uniformly, so there exists $N \in \mathbb{N}$ such that $|f(x) - f_N(x)| < 1$ for all $x \in \mathcal{X}$, so

$$\begin{aligned} &|f(x)| < |f_N(x)| + 1 \leq \|f_N\| + 1 \\ \Rightarrow &\|f\| < \|f_N\| + 1 < \infty, \end{aligned}$$

so f is bounded, and hence $f \in \mathscr{C}(\mathcal{X})$. □

Quotes of the day: Dr. Joshua Zahl 02/12/2024

No quotes today :(

We continue the talk about $\mathscr{C}(\mathcal{X})$. What about the metric space $\mathscr{C}(\mathcal{X}, \mathcal{Y})$, the bounded continuous functions $f : \mathcal{X} \to \mathcal{Y}$ ($\mathcal{X}, \mathcal{Y}$ are metric spaces) (bounded here means $f(\mathcal{X})$ is contained in some r-ball in $\mathcal{Y}$). Our metric is $d(f, g) = \sup\limits_{x \in \mathcal{X}} d_{\mathcal{Y}}(f(x), g(x))$ (check this is a metric). Is $\mathscr{C}(\mathcal{X}, \mathcal{Y})$ complete? Yes, this is true if and only if $\mathcal{Y}$ is complete; this has the same proof as before: consider constant function of Cauchy sequence that doesn't converge.

Theorem: Baby Rudin 7.16

Let $\alpha : [a, b] \to \mathbb{R}$ be a monotone increasing function. Let $\{f_n\}$ be a sequence $f_n \in \mathcal{R}_\alpha[a, b]$. Let $f : [a, b] \to \mathbb{R}$ and suppose $f_n \to f$ uniformly. Then $f \in \mathcal{R}_\alpha[a, b]$, and $\lim\limits_{n\to\infty} \int_a^b f \, d\alpha = \int_a^b f \, d\alpha$.

Proof. We first show that $f \in \mathcal{R}_\alpha[a, b]$. We will show $\underline{\int_a^b} f \, d\alpha = \overline{\int_a^b} f \, d\alpha$. Note that we can always assume that these upper and lower integrals exist: we just need to show that they are bounded, so take $\varepsilon = 1$, and there exists $N \in \mathbb{N}$ such that $|f(x) - f_N(x)| \leq 1$ for all x, and $|f_N(x)| < K$ for some $K \in \mathbb{R}$ since $f_N(x)$ is Riemann integrable. Hence. we have

$$|f(x)| \leq |f(x) - f_N(x)| + |f_N(x)| < K + 1.$$

For all $\varepsilon > 0$, since $f_n \to f$ uniformly, there exists $N \in \mathbb{N}$ such that for all $n > N$, for all $x \in [a, b]$, we have $|f(x) - f_n(x)| \leq \varepsilon$. Note that $f_n(x) - \varepsilon \leq f(x) \leq f_n(x) + \varepsilon$, so

$$\int_a^b (f_n - \varepsilon) \, d\alpha \leq \underline{\int_a^b} f \, d\alpha \leq \overline{\int_a^b} f \, d\alpha \leq \int_a^b (f_n + \varepsilon) \, d\alpha;$$

since each m_i is the greatest lower bound, it is greater than $f_n - \varepsilon$ on each interval for any partition, so $\underline{\int_a^b} (f_n - \varepsilon) \, d\alpha \leq \underline{\int_a^b} f \, d\alpha$, but the upper and lower integrals converge because it is Riemann integrable. This is only the first inequality, but the others follow in the same manner. Rearranging, we get

$$\overline{\int_a^b} f \, d\alpha - \underline{\int_a^b} f \, d\alpha \leq \int_a^b (f_n + \varepsilon) \, d\alpha - \int_a^b (f_n - \varepsilon) \, d\alpha = \int_a A^b 2\varepsilon \, d\alpha = 2\varepsilon[\alpha(b) - \alpha(a)].$$

Since $\varepsilon > 0$ is arbitrary, $\underline{\int_a^b} (f_n + \varepsilon) \, d\alpha = \overline{\int_a^b} (f_n + \varepsilon) \, d\alpha$.

Now, we need to show that the integral agrees with the limit of the integrals of f_n, which is not something we get to assert for free as we could in the point-wise case. Re-arranging our inequalities , we get

$$\int_a^b f_n \, d\alpha - \int_a^b \varepsilon \, d\alpha \leq \int -a^b f \, d\alpha \leq \int_a^b f_n \, d\alpha + \int_a^b \varepsilon \, d\alpha,$$

and thus,

$$\left| \int_a^b f \, d\alpha - \int_a^b f_n \, d\alpha \right| < \varepsilon[\alpha(b) - \alpha(a)].$$

□

Corollary

If $f_n \in \mathcal{R}_\alpha[a, b]$, and $\sum_{i=1}^{\infty} f_i$ converges uniformly on $[a, b]$ to f, then $f \in \mathcal{R}_\alpha[a, b]$, and

$$\int_a^b f \, d\alpha = \sum_{n=1}^{\infty} \int_a^b f_n \, d\alpha.$$

Proof sketch. Let $g_n := \sum_{i=1}^{n} f_i$, and $g_n \to f$ uniformly. Apply theorem 7.16, the rest of the proof is left as an exercise. The idea is to move an infinite sum inside an integral, which we cannot normally do, but the point of theorem 7.16 is to characterize when we can do this. □

The next theorem we look at is a bit trick to prove. Note that we can have a series of differentiable functions that converge uniformly to some function which is not differentiable. The functions $f_0 = \sin(x)$, $f_1 = \frac{1}{2}\sin(4x)$, $f_n = \frac{1}{2^n}\sin(4^n x)$are all differentiable, but $\sum_{i=1}^{\infty} f_i$, which converges uniformly – by the WeierstraßM-test – is not differentiable (in fact this is not differentiable anywhere, but this is hard to show.) The derivative of f_n is $\frac{4^{n^2}}{2^n}\cos(4^n x)$, and basically show that the cosine terms cancel out in such a way that it cannot happen.

If we assume a lot, however, we can ay some things about differentiability.

Theorem: Baby Rudin 7.17

Let $\{f_n\}$ be a sequence of functions $f_n : [a, b] \to \mathbb{R}$. Suppose

(a) Each $\{f_n\}$ is differentiable on $[a, b]$.

(b) There exists $x_0 \in [a, b]$ such that $\{f_n(x)\}$ converges.

(c) f_n' converge uniformly on $[a, b]$.

Then there exists f such that $f_n \to f$ uniformly on $[a, b]$, and $f'(x)$ exists for all $x \in [a, b]$ and $f'(x) = \lim_{n\to\infty} f_n'(x)$, i.e., $f_n' \to f'$ uniformly.

We need the last hypothesis, because our example above fails it. The second is also important, as $f_n = n$ is differentiable, has derivatives, but it does not converge.

We won't do the proof in this lecture, but we will get started with an important estimate:

Step-1: Show that $f_n \to f$ uniformly.

Let $\varepsilon > 0$; let $N \in \mathbb{N}$ be large enough such that $|f_m(x_0) - f_n(x_0)| < \varepsilon$ for all $m, n > N$ (convergent sequences are Cauchy) and $|f'_m(x_0) - f'_n(x_0)| < \varepsilon$ for all $m, n > N$, for all $x \in [a, b]$. Here is the crucial idea of this proof: we apply MVT to the difference $f_m - f_n$. For $x, t \in [a, b]$, $x \neq t$, there exists $c \in [x, t]$ such that

$$|[f_m(x) - f_n(x)] - [f_n(t) - f_m(t)]| = |(f'_n f'_m)(c)||x - t| \leq \varepsilon|x - t|.$$

How do we use this inequality? We have two consequences:

1. $|[f_n(x) - f_m(x)] - [f_n(t) - f_m(t)]| < \varepsilon|b - a|$; this is useful for uniform convergence.
2. $\dfrac{|[f_n(x) - f_m(x)] - [f_n(t) - f_m(t)]|}{|x - t|} < \varepsilon$; this is useful for differentiability.

Quotes of the day: Dr. Joshua Zahl 02/16/2024

"It's like looking for hay in a haystack, but we never get any hay." - On trying to write down a function that is continuous everywhere, but differentiable nowhere.

We pick up the proof of theorem 7.17 from last time:

Proof. Till now, we have fixed an $\varepsilon > 0$, and found an $N \in \mathbb{N}$ such that for all $m, n > N$,

$$|f_n(x_0) - f_m(x_0)| < \varepsilon \quad \text{and} \quad |f'_n(x) - f'_m(x)| < \varepsilon, \text{ for all } x \in [a, b].$$

Furthermore, using MVT, we get that for $x, t \in [x, b]$ ($x \neq t$):

1. $|[f_n(x) - f_m(x)] - [f_n(t) - f_m(t)]| < \varepsilon|b - a|$; this is useful for uniform convergence.
2. $\dfrac{|[f_n(x) - f_m(x)] - [f_n(t) - f_m(t)]|}{|x - t|} < \varepsilon$; this is useful for differentiability.

The goal here is to prove this in steps:

1. Show that there exists some f to which f_n uniformly converges.
2. Prove the statement of the theorem for the derivative.

By (1), for $t = x_0$, consider $m, n > N$, $x \in [a, b]$, which gives us

$$|f_n(x) - f_m(x)| \leq |(f_n(x) - f_m(x)) - (f_n(x_0) - f_m(x_0))| + |f_n(x_0) - f_m(x_0)| \leq \varepsilon_1(b - a) + \varepsilon_1.$$

Hence, let $\varepsilon := \varepsilon_1(b - a + 1)$, and $\{f_n\}$ satisfies the Cauchy criterion for uniform convergence.

We move on to step-2: showing that $f' = \lim\limits_{n\to\infty} f'_n$. Fix $x \in [a, b]$; we first need to show that f is differentiable at x. Since the derivative involves limits, this will involve a careful interchange of limits. The derivative is

$$\varphi_n(t) = \frac{f_n(t) - f_n(x)}{t - x}, \quad \varphi(t) = \frac{f(t) - f(x)}{t - x}.$$

Hence, $f'_n = \lim\limits_{n\to\infty} \varphi_n(t)$, and similarly we get f', if it exists. Inequality (2) says that for $m, n > N$,

$$|\varphi_n(t) - \varphi_m(t)| \leq \varepsilon,$$

i.e., $\{\varphi_n\}$ satisfies Cauchy criterion for uniform convergence, which implies φ_n converges uniformly on the domain $[a, b]\backslash\{x\}$; here x is a limit point, which is something that comes up later. This says *something* about φ_n, but we don't know if it converges to $\varphi(t)$ at any fixed t. Showing point-wise is enough: fix $t \in [a, b]\backslash\{x\}$. Hence,

$$\varphi_n(t) - \varphi(t) = \left|\frac{(f_n(t) - f_n(x)) - (f(t) - f(x))}{t - x}\right| \leq \underbrace{\left|\frac{f_n(t) - f(t)}{t - x}\right| + \left|\frac{f_n(x) - f(x)}{t - x}\right|}_{\to 0 \text{ as } n\to\infty}.$$

In conclusion, $\varphi_n \to \varphi$ point-wise, hence uniformly on $[a, b]\backslash\{x\}$. Finally, since x is a limit point of $[a, b]\backslash\{x\}$, and $\varphi_n \to \varphi$ uniformly on $[a, b]\backslash\{x\}$, we apply theorem 7.11 to conclude that

$$f'(x) = \lim_{t\to x} \varphi(t) = \lim_{t\to x} \lim_{n\to\infty} \varphi_n(t) = \lim_{n\to\infty} \lim_{t\to x} \varphi_n(t) = \lim_{n\to\infty} f'_n(x)$$

when the limit exists, which it does in this case. $\square$

3.4 The Weierstraß function

Theorem: Baby Rudin 7.18

The exists a continuous function $f : \mathbb{R} \to \mathbb{R}$ such that $f'(x)$ does not exist for any $x \in \mathbb{R}$.

One way to do this is construct such a function. Note that if we pick a function from $\mathcal{C}(\mathbb{R})$ at random, it will almost surely be a function that is not differentiable anywhere. However, if we try to write a function down, it is usually differentiable (this could be seen as an analogue to the fact that most of the numbers that exist are transcendental, but if we think of a real number, it will almost surely be algebraic.) This happens often in math: we can prove that almost always functions have a certain property, but it's significantly harder (if even feasible) to write an example down.

Proof. Consider the periodization of $|x|$ on $[-1, 1]$ to the whole real line; call this $\varphi(x)$. This function is continuous, in fact it is Lipschitz continuous with Lipschitz constant 1 (also called 1-Lipschitz): $|\varphi(x) - \varphi(y)| \leq 1|x - y|$. Note that being Lipschitz is not preserved under point-wise limits, but being k-Lipschitz, for a fixed k, is.

Let $f(x) := \left(\frac{3}{4}\right)^n \varphi(4^n x)$. This series converges absolutely by the Weierstraß M-test. Each of these terms are continuous, and since absolutely convergent series of continuous functions converges to a continuous function, f is continuous.

For large n, note that $f(x)$ is very small, but very spiky. However, note that the 4^n is getting large faster than $\left(\frac{3}{4}\right)^n$ is getting small, i.e., if we multiply them, we get 3^n, which still blows up. Hence, we get that $(f(x)$ is 3^n-Lipschitz.

Fix $x \in \mathbb{R}$. We wish to show that $f(x)$ is not differentiable at x. It suffices to find $\delta_m \searrow 0$ such that

$$\left| \frac{f(x + \delta_m) - f(x)}{\delta_m} \right| \nearrow \infty$$

as $m \to \infty$. The trick here is finding δ_m such that bigger m cancel out (equal spots on the period), an for the smaller values of m, the Lipschitz constant is too small to make a difference: even if all the other peaks were working against it, $3^j - \sum_{n=0}^{j-1} 3^j = \frac{1}{6} 3^j$ (or something along those lines).

The full proof is given in Baby Rudin, or can be treated as an exercise. $\square$