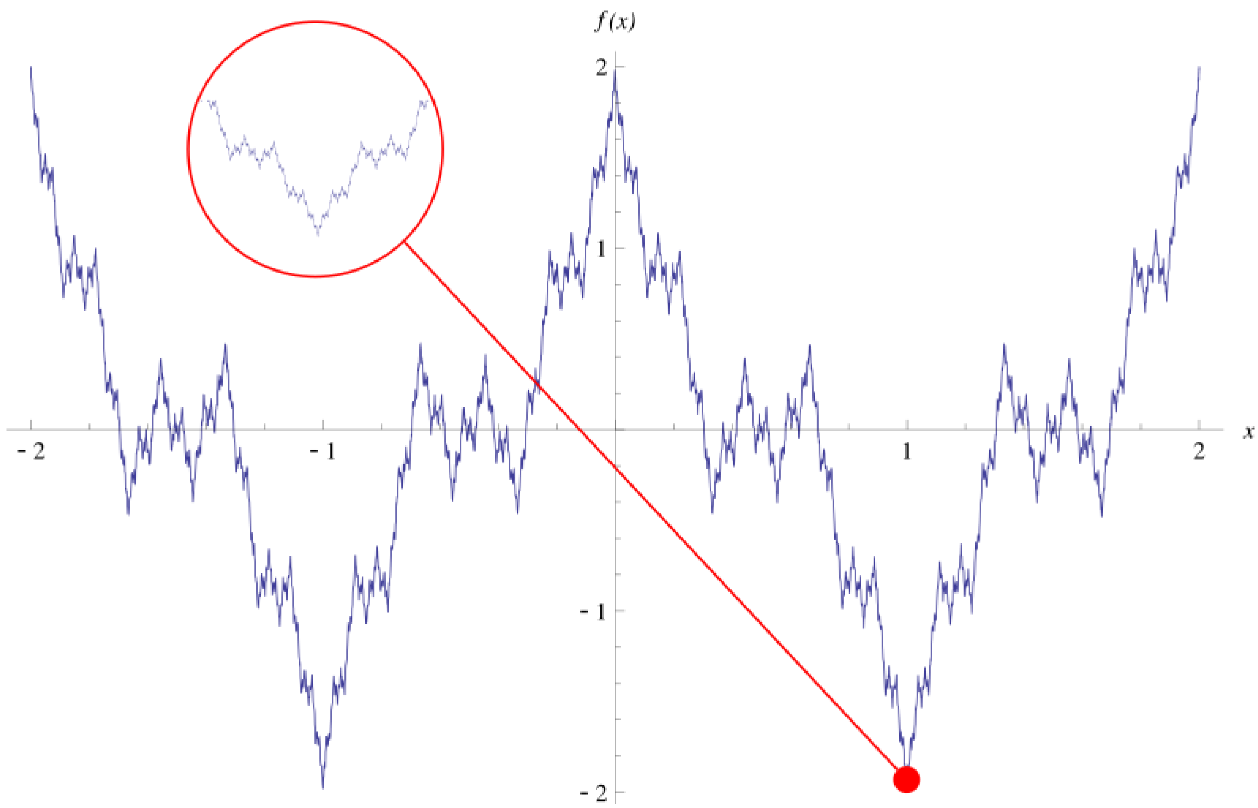
Math 140B Foundations of Real Analysis II

Jack Yang

Last Updated: March 19, 2021



$$f(x) = \sum_{n=0}^{\infty} a^n \cos(b^n \pi x) \text{ for } 0 < a < 1,\ b \text{ is a positive odd integer, and } ab > 1 + \frac{3}{2}\pi$$

Abstract

Differentiation. Riemann integral. Sequences and series of functions. Special functions. Fourier series. This corresponds to Chapters 5, 6, 7, and 8 in Walter Rudin's *Principle of Mathematical Analysis*.

Contents

1 Differentiation

Previously, we discussed sequences and functions in general metric spaces. Now, we will focus on concepts that are exclusive to functions defined on (the intervals in) $\mathbb{R}$.

Lecture 1
Monday
January 4

1.1 The Derivative of a Real Function

We can now begin the rigorous treatment of calculus in earnest, starting with the notion of a derivative. We can now define derivatives analytically, using limits.

Definition 1.1 (Derivative)**.** Let $f : E \to \mathbb{R}$ with $E \subseteq \mathbb{R}$. For $x \in E \cap E'$, the *derivative of f at x* is defined as

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

provided this limit exists.

More precisely, define $\phi : E \setminus \{x\} \to \mathbb{R}$ (*difference quotient*) by

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

and set $f' = \lim_{t \to x} \phi(t)$ provided the limit exists.

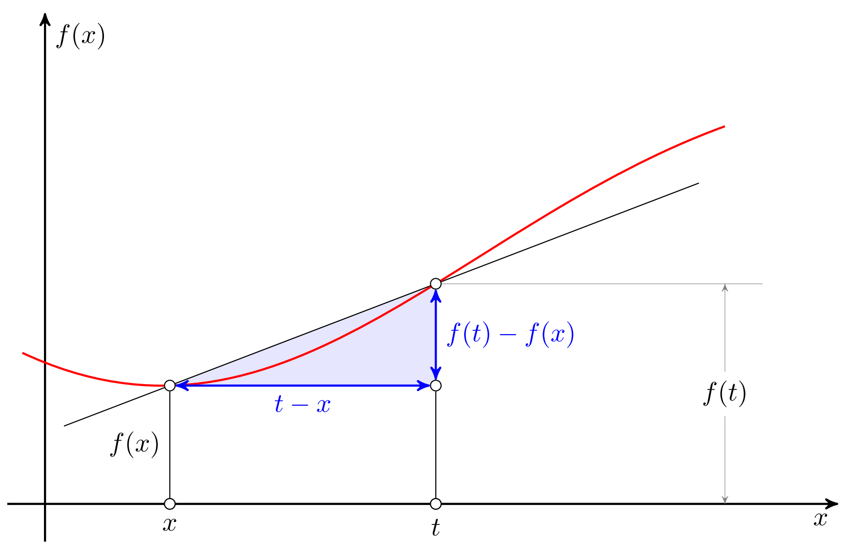


Figure 1.1: A Representation of Derivative.

Let $f : [a, b] \to \mathbb{R}$, rather than defined on an abstract domain E. In particular, at the point $x = a$, being an extreme of the interval $[a, b]$, we can only consider the right-hand limit. Otherwise, f is not defined on $x < a$.

To define the *derivative*, f needs to be *defined* and *continuous* at the point at which we are trying to differentiate. We say f is *differentiable* if it is differentiable on its *entire* domain.

Note that $\operatorname{dom}(f') \subseteq \operatorname{dom}(f)$ and often $\operatorname{dom}(f') \neq \operatorname{dom}(f)$. For example, if f is defined on an open set (a, b) and if $a < x < b$, then $f'(x)$ is defined as normal, but $f'(a)$ and $f'(b)$ are not defined in this case.

Theorem 1.2 (Differentiability Implies Continuity)**.** *Let $f : [a, b] \in \mathbb{R}$. If f is differentiable at $x \in [a, b]$, then f is continuous at x.*

Proof. As $t \to x$, we have

$$\begin{aligned}
\lim_{t\to x} f(t) &= \lim_{t\to x}\left(\frac{f(t)-f(x)}{t-x}\cdot(t-x)+f(x)\right)\\
&= \left(\lim_{t\to x}\frac{f(t)-f(x)}{t-x}\right)\cdot\left(\lim_{t\to x} t-x\right)+f(x)\\
&= f'\cdot 0 + f(x) = f(x).
\end{aligned}$$

Thus, we conclude that f is continuous at x.

□

The converse is not true. Consider $f(x) = |x|$ (in Figure 1.1. We note that f is continuous on $\mathbb{R}$ but is not differentiable at $x = 0$. There is a corner at $x = 0$.

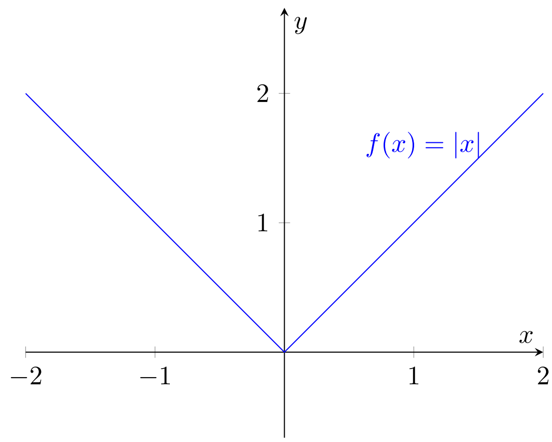


Figure 1.2: Counterexample for Converse of Theorem 1.2

In the future, we will construct a function (*Weierstrass function*) which is continuous on $\mathbb{R}$ without being differentiable at any $x \in \mathbb{R}$.

The following theorem comes from 140A first, but we will frequently use this in 140B as well.

Theorem 1.3 (Limit Rules)**.** *Let (x, d) be a metric space, $E \subseteq X$, and $p \in E'$. Let $f, g : E \to \mathbb{R}$ and $\lim_{x\to p} f(x) = A$ and $\lim_{x\to p} g(x) = B$. Then,*

(a) $\lim_{x\to p}(f+g)(x) = A + B$

(b) $\lim_{x\to p}(f\cdot g)(x) = A\cdot B$

(c) $\lim_{x\to p}\left(\frac{f}{g}\right)(x) = \frac{A}{B}$ *if* $B \neq 0$

Theorem 1.4. *Suppose f and g are real-valued functions differentiable at $x \in \mathbb{R}$. Then, $f+g$, $f\cdot g$, $\frac{f}{g}$ (when $g(x) \neq 0$) are differentiable at x and*

(a) $(f+g)'(x) = f'(x) + g'(x)$

(b) $(f\cdot g)'(x) = f'(x)\cdot g(x) + f(x)\cdot g'(x)$ *(Leibniz Rule)*

(c) $\left(\frac{f}{g}\right)'(x) = \frac{f'(x)\cdot g(x) - f(x)\cdot g'(x)}{g^2(x)}$

Proof.

(a) This is obvious by

$$\begin{aligned}(f+g)'(x) &= \lim_{t\to x}\frac{f(t)-f(x)+g(t)-g(x)}{t-x}\\ &= \lim_{t\to x}\frac{f(t)-f(x)}{t-x} + \lim_{t\to x}\frac{g(t)-g(x)}{t-x}\\ &= f'(x)+g'(x).\end{aligned}$$

(b) We note that

$$f(t)\cdot g(t) - f(x)\cdot g(x) = g(t)\cdot(f(t)-f(x)) + f(x)\cdot(g(t)-g(x)).$$

Thus,

$$\frac{f(t)\cdot g(t) - f(x)\cdot g(x)}{t-x} = g(t)\cdot\frac{f(t)-f(x)}{t-x} + f(x)\cdot\frac{g(t)-g(x)}{t-x}$$

has limit $g(x)\cdot f'(x) + f(x)\cdot g'(x)$ as $t\to x$.

(c) Let $h(x) = \dfrac{f(x)}{g(x)}$. Again, we note that

$$h(t)-h(x) = \frac{1}{g(t)\cdot g(x)}\left(g(x)\cdot(f(t)-f(x)) - f(x)\cdot(g(t)-g(x))\right).$$

Thus,

$$\frac{h(t)-h(x)}{t-x} = \frac{1}{g(t)\cdot g(x)}\left(\frac{f(t)-f(x)}{t-x}\cdot g(x) - f(x)\cdot\frac{g(t)-g(x)}{t-x}\right)$$

has limit $\dfrac{f'(x)\cdot g(x) - f(x)\cdot g'(x)}{g^2(x)}$ as $t\to x$.

□

Example 1.5 (Some Common Differentiable Functions)**.**

(a) Constant functions $f(x) = c$ have derivative $f'(x) = 0$.

(b) Let $f(x) = x$. Then, $f'(x) = 1$.

(c) Let $f(x) = x^m$. Then, $f'(x) = m\cdot x'$ by induction with *Leibniz Rule.*

(d) Polynomials are differentiable everywhere. Rational functions are differentiable on their entire domain.

The following theorem is known as the *Chain Rule* for differentiation. It deals with differentiation of composite functions.

Theorem 1.6 (Chain Rule)**.** *Let $f : [a,b] \to \mathbb{R}$, $f'(x)$ exists at some point $x \in [a,b]$, g is defined on an interval I that contains the range of f (that is, $f([a,b]) \subseteq I$), and g is differentiable at $f(x) \in I$. If $h(t) = g(f(t))$ for $a \le t \le b$, then h is differentiable at x, and*

$$h'(x) = g'(f(x))\cdot f'(x).$$

The intuition behind is that

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

has limit $g'(f(x))\cdot f'(x)$ as $t\to x$. However, there is a flaw. It is possible that there exists a $\delta>0$ such that $f(x)=f(t)$ for $t\in B_\delta(x)$. Also, this flaw does *not* matter if f is one-to-one.

Proof. Let $y=f(x)$. Then, we have

$$\begin{aligned}\frac{f(t)-f(x)}{t-x}&=f'(x)+u(t)\text{ for }\lim_{t\to x}u(t)=0\\ \frac{g(s)-g(y)}{s-y}&=g'(y)+v(s)\text{ for }\lim_{s\to y}v(s)=0\end{aligned}$$

where $t\in[a,b]$ and $s\in I$. We also define $u(x)=0$ and $v(y)=0$, thus u is continuous at x and v is continuous at y. Let $s=f(t)$, then

$$\begin{aligned}h(t)-h(x)&=g(f(t))-g(f(x))\\&=g(s)-g(y)\\&=(s-y)\cdot[g'(y)+v(s)]\\&=(f(t)-f(x))\cdot[g'(f(x))+v(s)]\\&=(t-x)\cdot[f'(x)+u(t)]\cdot[g'(f(x))+v(s)].\end{aligned}$$

With $t\neq x$, we have

$$\frac{h(t)-h(x)}{t-x}=[f'(x)+u(t)]\cdot[g'(f(x))+v(s)].$$

As $t\to x$, we note that $s=f(t)\to f(x)=y$ by continuity of f at x, then

$$\lim_{t\to x}\frac{h(t)-h(x)}{t-x}=f'(x)\cdot g'(f(x)).$$

□

In the proof above, we may get confused about the functions $u(x)$ and $v(y)$ initially. What do they mean? Let $f:[a,b]\to\mathbb{R}$ be differentiable at a point $t\in[a,b]$, then the derivative helps us find a suitable *linear approximation* for f; that is,

$$f(x)\approx f(t)+f'(t)\cdot(x-t).$$

Recall the definition of *derivative*

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

if we remove the notation lim, then this turns to an inequality. So, we need to use the *"error term"* $e(t)$ to make it up, that is,

$$\frac{f(t)-f(x)}{t-x}=f'(x)+e(t).$$

Lecture 2
Wednesday
January 6

Example 1.7. Define a function $f:\mathbb{R}\to\mathbb{R}$ as follows

$$f(x)=\begin{cases}x\sin\dfrac{1}{x} & x\neq 0\\ 0 & x=0.\end{cases}$$

First of all, note that f is continuous at all $x \neq 0$ since it is a composition of continuous functions. At $x = 0$, since

$$\left|\sin \frac{1}{x}\right| \leq 1 \text{ for all } x \in \mathbb{R},$$

so

$$|f(x)| \leq |x| \to 0 = f(0) \text{ as } x \to 0.$$

Thus, f is continuous on $\mathbb{R}$.

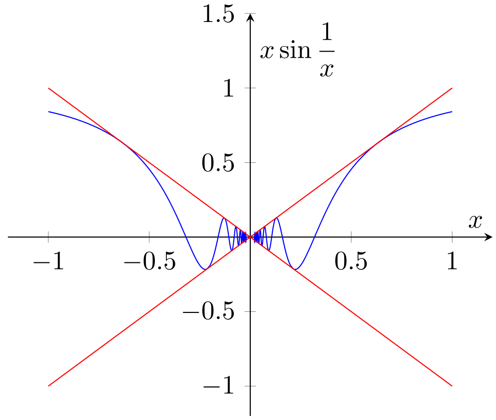


Figure 1.3: Graph of $x \sin \frac{1}{x}$

Assume we know

$$\frac{d}{dx} \sin x = \cos x \text{ and } \frac{d}{dx} \cos x = -\sin x,$$

then we can compute that f is differentiable at all points $x \neq 0$, and for such points, we have

$$f'(x) = \sin \frac{1}{x} + x \cos \frac{1}{x} \cdot \left(-\frac{1}{x^2}\right) = \sin \frac{1}{x} - \frac{1}{x} \cos \frac{1}{x}.$$

However, at $x = 0$, we define

$$\phi(t) = \frac{f(t) - f(0)}{t - 0} = \frac{f(t)}{t} = \sin \frac{1}{t}.$$

Does $\lim_{t \to 0} \phi(t)$ exist? Nope, since the function $\lim_{x \to 0} \sin \frac{1}{x}$ is oscillating between 0 and 1. So, f is not differentiable at $x = 0$, and there is no way to extend the function f' to be defined at 0 in such a way that it will be continuous. We conclude that f is defined and continuous on $\mathbb{R}$ but not differentiable at $x = 0$.

Example 1.8. Now consider the following function

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

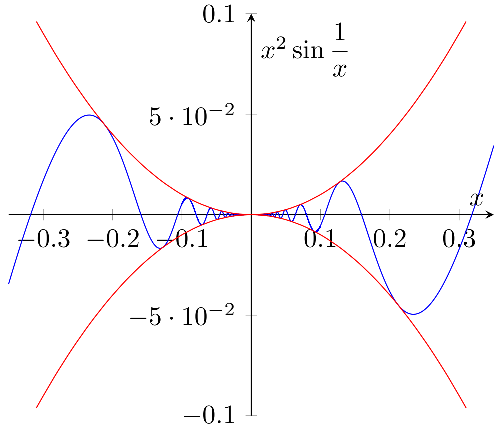


Figure 1.4: Graph of $x^2 \sin \frac{1}{x}$

Similarly, from Example 1.7, f is defined and continuous on $\mathbb{R}$ and differentiable everywhere except possibly at 0. Here, we have

$$f'(x) = 2x \sin \frac{1}{x} - \cos \frac{1}{x}.$$

Now, looking at $x = 0$ specifically, we define and note that

$$\phi(t) = \frac{f(t) - f(0)}{t - 0} = t \sin \frac{1}{t} \to 0 \text{ as } t \to 0.$$

This shows that f is actually differentiable on all $\mathbb{R}$, with $f'(0) = 0$.

Is $f'(x)$ continuous on $\mathbb{R}$? Again, $\lim_{x \to 0} 2x \sin \frac{1}{x} - \cos \frac{1}{x}$ oscillates between 0 and 1, so $f'(x)$ is not continuous at $x = 0$. In this case, $f'(x)$ has the *discontinuity of the second kind* at $x = 0$. We conclude that it is possible for a function to be continuous and differentiable everywhere, but for its derivative to have a discontinuity somewhere.

1.2 Mean Value Theorems

Recall from calculus, one of the main applications of derivatives is in the study of optimization. Let's define a familiar concept, *local maxima*, first.

Definition 1.9 (Local Maximum, Local Minimum)**.** Let (X, d) be a metric space, and let $f : X \to \mathbb{R}$. We say f has a *local maximum* at $p \in X$ if there exists a $\delta > 0$ such that for all $q \in B_\delta(p) = \{x \in X : d(x, p) < \delta\}$, we have $f(q) \leq f(p)$. *Local minimum* is defined similarly.

For a function $f : \mathbb{R} \to \mathbb{R}$, *local extrema* of f can be determined by locating the points x where $f'(x) = 0$.

Theorem 1.10. *Let $f : [a, b] \to \mathbb{R}$. If f has a local maximum at $x \in (a, b)$ and provided $f'(x)$ exists, then $f'(x) = 0$. Same statement holds for local minimum.*

Proof. Fix a $\delta > 0$ such that $f(t) \leq f(x)$ for all $t \in (x - \delta, x + \delta$. If $t \in (x - \delta, x)$, then

$$\frac{f(t) - f(x)}{t - x} \geq 0.$$

Therefore, $f'(x) \geq 0$.

Similarly, if $t \in (x, x + \delta)$, then

$$\frac{f(t) - f(x)}{t - x} \leq 0$$

and therefore $f'(x) \leq 0$. Hence, we conclude that $f'(x) = 0$.

□

Now, we get to a very important theorem in calculus, the *Mean Value Theorem*. It connects the values of a function and its derivative without using limits.

We will first look at the *Generalized Mean Value Theorem*, and then the traditional *Mean Value Theorem* immediately follows.

Theorem 1.11 (Cauchy Mean Value Theorem)**.** *If $f, g : [a, b] \to \mathbb{R}$ are continuous on $[a, b]$ and differentiable on (a, b), then there is $x \in (a, b)$ such that*

$$(f(b) - f(a))g'(x) = (g(b) - g(a))f'(x).$$

This theorem says that the functions $(f(b) - f(a))g(x)$ and $(g(b) - g(a))f(x)$ have the same net change on $[a, b]$.

Proof. Let $h(x) = (f(b) - f(a))g(x) - (g(b) - g(a))f(x)$. Then, h is continuous on $[a, b]$ and is differentiable on (a, b) and

$$h(a) = f(b)g(a) - g(b)f(a) = h(b).$$

It suffices to find $x \in (a, b)$ with $h'(x) = 0$.

The result is immediate if h is constant on $[a, b]$, then $h'(x)$ is constantly 0 on (a, b).

Now, assume h is not a constant. Without loss generality, assume that $h(t) > h(a) = h(b)$ for some $t \in [a, b]$. Since $[a, b]$ is compact, so h attains a maximum value on $[a, b]$ at some point x. Then, $h'(x) = 0$ by Theorem 1.10.

Similar argument works for the case $h(t) < h(a) = h(b)$ for some $t \in [a, b]$.

□

Theorem 1.12 (Lagrange Mean Value Theorem)**.** *If $f : [a, b] \to \mathbb{R}$ is continuous on $[a, b]$ and differentiable on (a, b), then there is $x \in (a, b)$ satisfying*

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

Proof. Apply Theorem 1.11 with $g(x) = x$ to obtain $x \in (a, b)$ such that

$$(b - a)f'(x) = f(b) - f(a).$$

□

Corollary 1.13 (Rolle's Theorem)**.** *Let $f : [a, b] \to \mathbb{R}$ be a continuous function which is differentiable on (a, b), and $f(a) = f(b)$. Then, there is a point $c \in (a, b)$ such that $f'(c) = 0$.*

This is obvious by Theorem 1.12.

The conditions of *Mean Value Theorem* are optimal and necessary. We can find some counterexamples if we modify some conditions.

Consider we change the interval of *continuity* to an open interval (a, b). We define a function

$$f(x) = \begin{cases} 1 & 0 < x \leq 1 \\ 0 & x = 1. \end{cases}$$

Indeed, f is continuous on $(0,1]$, not at $x=0$, and is differentiable on $(0,1)$ with $f'(x)=0$. However,

$$f(1)-f(0)=1\neq 0=f'(x)\cdot(1-x) \text{ for any } x\in(0,1).$$

So, the *Mean Value Theorem* fails.

Consider we change the interval of *differentiability* to a subset of (a,b), the original interval. We define a function

$$f(x)=\begin{cases}2x & x<\dfrac{1}{2}\\ 1 & x\geq\dfrac{1}{2}.\end{cases}$$

Indeed, f is continuous on $[0,1]$. Also,we have

$$f'(x)=\begin{cases}2 & x<\dfrac{1}{2}\\ 0 & x>\dfrac{1}{2}.\end{cases}$$

At $x=\dfrac{1}{2}$, $f'(x)$ does not exist since $f'(x-)\neq f'(x+)$. We have $f(1)-f(0)=1$, but

$$f'(x)\cdot(1-0)=\begin{cases}2 & x<\dfrac{1}{2}\\ 0 & x>\dfrac{1}{2}.\end{cases}$$

The *Mean Value Theorem* fails again.

Lecture 3
Friday
January 8

Corollary 1.14. *Suppose f is differentiable in (a,b).*

(a) If $f'(x)\geq 0$ for all $x\in(a,b)$, then f is monotonically increasing. This means $x_1>x_2$ implies $f(x_1)\geq f(x_2)$.

(b) If $f'(x)=0$ for all $x\in(a,b)$, then f is constant.

(c) If $f'(x)\leq 0$ for all $x\in(a,b)$, then f is monotonically decreasing. This means $x_1>x_2$ implies that $f(x_1)\leq f(x_2)$.

Proof. All conclusions are immediate from Mean Value Theorem. Suppose $a<x_1<x_2<b$. There exists $x\in(x_1,x_2)$ such that

$$f(x_2)-f(x_1)=f'(x)\cdot(x_2-x_1).$$

We always have $x_2-x_1>0$.

(a) If $f'(x)\geq 0$ for all $x\in(a,b)$, then $f(x_2)\geq f(x_1)$. We conclude that f is monotonically increasing.

(b) If $f'(x)=0$ for all $x\in(a,b)$, then $f(x_1)=f(x_2)$. We conclude that f is constant.

(c) If $f'(x)\leq 0$ for all $x\in(a,b)$, then $f(x_2)\leq f(x_1)$. We conclude that f is monotonically decreasing.

□

1.3 Intermediate Value Theorem

We have already seen that a function f may have a derivative f' which exists at every point, but is discontinuous at some point. However, not every function is a derivative. In particular, derivatives which exist at every point of an interval have one important property in common with functions which are continuous on an interval.

Theorem 1.15 (Intermediate Value Theorem for Derivatives)**.** *Suppose $f : [a, b] \to \mathbb{R}$ is differentiable on $[a, b]$ and $f'(a) < \lambda < f'(b)$. Then, there exists $x \in (a, b)$ such that $f'(x) = \lambda$.*

A similar result holds if $f'(a) > f'(b)$.

Proof. Let $g(t) = f(t) - \lambda \cdot t$. Then, $g'(t) = f'(t) - \lambda$. Thus, $g'(a) < 0$, so there exists $a < t_1 < b$ such that $g(t_1) < g(a)$. Also, $g'(b) > 0$, so that there exists $a < t_2 < b$ such that $g(t_2) < g(b)$.

Since $g(t_1) < g(a)$ and $g(t_2) < g(b)$ and g is continuous on $[a, b]$, a compact set, so g attains its minimum on $[a, b]$ and let $g(x)$ be the minimum value of g. Further,

$$g(x) = \min\{g(x) : x \in [a, b]\} \leq \min\{g(t_1), g(t_2)\} < \min\{g(a), g(b)\}.$$

So, $g(x)$ attains its minimum as $x \neq a$ and $x \neq b$. Then, there exists $x \in (a, b)$ such that $g(x)$ is minimum value of g. Then, $g'(x) = f'(x) - \lambda = 0$. We conclude that $f'(x) = \lambda$.

□

This theorem says that, although f' may not be continuous, the Intermediate Value Theorem holds for f'.

Remark 1.16. Not every function satisfies the Intermediate Value Theorem. We define a function

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

Clearly, $g(x)$ does not satisfy the Intermediate Value Theorem, since it is not continuous (this is simple discontinuity) at $x = 0$.

Is $g(x)$ a derivative of another function? No. Consider the function

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

In this case, we try to construct a continuous function such that $f'(x) = g(x)$, but we fail, since $f(x)$ is not differentiable at $x = 0$, while $g(x)$ is defined at $x = 0$. The derivative of f is given by

$$f'(x) = \begin{cases} 1 & x > 0 \\ 0 & x < 0. \end{cases}$$

Corollary 1.17. *If f is differentiable on $[a, b]$, then f' cannot have any simple discontinuities on $[a, b]$.*

Proof. Recall that f has a simple discontinuity at x if $f(x+) \neq f(x-)$ or $f(x+) = f(x-)$ but $f(x) \neq f(x+)$.

Towards a contradiction, assume $g(x) = f'(x)$ has a simple discontinuity at $c \in [a, b]$. Without loss of generality, we assume that $g(c+) \neq g(c)$. Further, assume that $g(c+) > g(c)$. Then, pick a λ such that $g(c) < \lambda < g(c+)$. Then, there exists a $\delta > 0$ such that

$$x - c < \delta \Rightarrow g(x) \in B_\varepsilon(f(c+)).$$

However, this is equivalent to

$$x - c < \delta \Rightarrow \lambda < g(x) < g(c+) \text{ for } x \in (c, c+\delta).$$

Then, consider the interval $\left[c, c+\dfrac{\delta}{2}\right]$, we have

$$g(c) < \lambda < g\left(c + \frac{\delta}{2}\right).$$

Hence, $g(x) \neq \lambda$ for all $x \in \left(c, c+\dfrac{\delta}{2}\right)$. However, this contradicts Intermediate Value Theorem, since it suggests that there must be $x \in \left(c, c+\dfrac{\delta}{2}\right)$ such that $f'(x) = \lambda$.

□

Of course, f' can have *discontinuity of the second kind.*

Example 1.18.

(a) Let $f(x) = x^2 \sin \dfrac{1}{x}$. Then, the derivative of f is given by

$$f'(x) = 2x \sin \frac{1}{x} - \cos \frac{1}{x}.$$

Indeed, $f'(x)$ is not continuous on $\mathbb{R}$. Since $f'(0)$ does not exist, so $f'(x)$ has *discontinuity of the second kind* at $x = 0$.

(b) Let $g(x) = x^{1/3}$. Then, the derivative of g is given by

$$g'(x) = \frac{1}{3}x^{-2/3}.$$

Also, in this case, $f'(0)$ is not defined, so $f'(x)$ has *discontinuity of the second kind* at $x = 0$.

1.4 L'Hospital's Rule

Starting from now, we are going to explore some interesting and important applications of Mean Value Theorem. One of the famous application is the *L'Hospital's Rule*, which helps us deal with limits of indeterminate forms.

Theorem 1.19 (L'Hospital's Rule)**.** *Let $-\infty \leq a < b \leq +\infty$. Let $f : (a,b) \to \mathbb{R}$ and $g : (a,b) \to \mathbb{R}$ be differentiable and $g'(x) \neq 0$ for all $x \in (a,b)$. Suppose $\displaystyle\lim_{x\to a} \frac{f'(x)}{g'(x)} = A \in \mathbb{R} \cup \{-\infty, +\infty\}$.*

(a) If $\displaystyle\lim_{x\to a} f(x) = \lim_{x\to a} g(x) = 0$, then $\displaystyle\lim_{x\to a} \frac{f(x)}{g(x)} = A$.

(b) If $\displaystyle\lim_{x\to a} g(x) = +\infty$, then $\displaystyle\lim_{x\to a} \frac{f(x)}{g(x)} = A$.

Similar statement is true if $x \to b$ or if $g(x) \to -\infty$.

Proof. If suffices to show that there exists $\displaystyle\lim_{x\to a} \frac{f(x)}{g(x)} = A$ such that

(a) for all $q > A$, there exists $c > a$ such that $\dfrac{f(x)}{g(x)} < q$ for all $a < x < c$, *and*

(b) for all $p < A$, there exists $d > a$ such that $\dfrac{f(x)}{g(x)} > p$ for all $a < x < d$.

In particular, if we consider the case $A = +\infty$, then we only need to show (b). If we consider the case $A = -\infty$, then we only need to show (a). So, we divide the proof into 2 parts.

We now prove (a). Assume $A \neq +\infty$ and let $q \in \mathbb{R}$ be such that $q > A$ (and q exists since $A < +\infty$). Then, pick $r \in \mathbb{R}$ such that $A < r < q$. By assumption, there exists $c \in (a, b)$ such that

$$\frac{f'(x)}{g'(x)} < r \text{ for all } a < x < c.$$

This follows from $\lim_{x\to a} \dfrac{f'(x)}{g'(x)} = A < r$. Then, if $a < x < y < c$, by Mean Value Theorem, there exists $t \in (x, y)$ such that

$$\frac{f(x) - f(y)}{g(x) - g(y)} = \frac{f'(t)}{g'(t)} < r. \tag{*}$$

(1) If $\lim_{t\to a} f(t) = \lim_{t\to a} g(t) = 0$, then let $x \to a$, we get

$$\frac{f(y)}{g(y)} = \frac{-f(y)}{-g(y)} \leq r < q \text{ for } a < y < c.$$

(2) If $\lim_{t\to a} g(t) = +\infty$, then for fixed y, pick $c_1 \in (a, y)$ such that $g(x) > g(y)$ for all $a < x < c_1$, by the fact that $g(t) \to +\infty$, and moreover $g(x) > 0$ for $a < x < c_1$. Multiply (*) by $\dfrac{g(x) - g(y)}{g(x)} > 0$ on (a, c_1), we get

$$\frac{f(x) - f(y)}{g(x)} < \left(1 - \frac{g(y)}{g(x)}\right) \cdot r.$$

Further,

$$\frac{f(x)}{g(x)} \leq \left(1 - \frac{g(y)}{g(x)}\right) \cdot r + \frac{f(y)}{g(x)}.$$

Let $x \to a$, then we have

$$\limsup_{x\to a} \frac{f(x)}{g(x)} \leq r < q.$$

So, we have $\limsup_{x\to a} \dfrac{f(x)}{g(x)} < q$ in both cases.

Proof of (b) is similar.

□

As *L'Hospital's Rule* states, we can deal two types of limits of indeterminate forms: $\dfrac{0}{0}$ and $\dfrac{C}{\infty}$ for a constant C. Lecture 4 Monday January 11

Example 1.20. Find $\lim_{x\to\infty} \left(1 + \dfrac{a}{x}\right)^x$. This is not in a form to directly apply *L'Hospital's Rule*. We note that

$$\ln\left(1 + \frac{a}{x}\right)^x = x \ln\left(1 + \frac{a}{x}\right) = \frac{\ln(1 + a/x)}{1/x} \to \frac{0}{0} \text{ as } x \to +\infty.$$

We have

$$\begin{aligned}
\lim_{x\to\infty} \ln\left(1+\frac{a}{x}\right)^x &= \lim_{x\to\infty} \frac{\ln(1+a/x)}{1/x} \\
&= \lim_{x\to\infty} \frac{\dfrac{1}{1+a/x}\cdot\left(-\dfrac{a}{x^2}\right)}{-\dfrac{1}{x^2}} \\
&= \lim_{x\to\infty} \frac{a}{1+a/x} \\
&= a.
\end{aligned}$$

Since $x \mapsto e^x$ is continuous, so

$$\lim_{x\to\infty}\left(1+\frac{a}{x}\right)^x = \lim_{x\to\infty} e^{\ln(1+a/x)^x} = e^a.$$

1.5 Taylor's Theorem

By Intermediate Value Theorem, f' must at least satisfy Intermediate Value Theorem, which means f' cannot be arbitrary, but f' certainly need not be continuous. In many circumstances, it is natural to assume that f' is continuous. In some cases, we use *second derivatives* to determine concavity, where we actually assume that f' is differentiable.

Definition 1.21 (Higher-Order Derivative)**.** Let f be a real-valued function. If $f'(x)$ is defined in a neighborhood of x_0 (that is an open interval containing x_0) and $f'(x)$ is differentiable at x_0, then we define

$$f''(x_0) = \lim_{y\to x_0} \frac{f'(y)-f'(x_0)}{y-x_0}.$$

The limit exists because $f'(x)$ is differentiable at x_0.

In the same way, we can define derivatives in any order, denoted by $f^{(n)}(x)$. Note that

$$\operatorname{dom}\left(f^{(n)}(x)\right) \subseteq \operatorname{dom}\left(f^{(n-1)}(x)\right) \subseteq \cdots$$

In particular, if $f^{(n)}(x)$ exists, then $f^{(n-1)}(x)$ must be defined in a neighborhood of x, therefore $f^{(n-2)}(x)$ must be differentiable.

Recall our motivation of differentiability. Continuity of f at a point x is that $f(x+h)-f(x)$ tends to 0 as $h\to 0$, but at what rate? If f is differentiable, the answer is that there is an error term with

$$f(x+h)-f(x)-f'(x)\cdot h = e(h).$$

Thus, differentiability implies that $f(x+h)$ is closer to $f(x)$ than any linear function. The difference goes to 0 faster than t as $t\to 0$. This leads us to ask what happens if there are higher-order derivatives? Is the difference even smaller, with the higher-order corrections? Yes. *Taylor's Theorem* gives us a polynomial that can actually approximate f very well.

Theorem 1.22 (Taylor)**.** *Suppose $f : [a,b] \to \mathbb{R}$ and let $n \in \mathbb{Z}_{>0}$. Suppose that $f^{(n-1)}(x)$ is continuous on $[a,b]$ and differentiable on (a,b) (that is, $f^{(n)}(t)$ exists for $t \in (a,b)$). Let $\alpha \neq \beta \in [a,b]$ and define*

$$P_{n-1}(t) = \sum_{k=0}^{n-1} \frac{f^{(k)}(\alpha)}{k!}(t-\alpha)^k.$$

Then there exists $x \in [\alpha,\beta]$ such that

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

Proof. For $n = 1$, this is just the Mean Value Theorem.

Let $M \in \mathbb{R}$ such that
$$f(\beta) = P_{n-1}(\beta) + M(\beta - \alpha)^n$$
and let $g(t) = f(t) - P_{n-1}(t) - M(t - \alpha)^n$ for $a \le t \le b$. We want to show that $n!M = f^{(n)}(x)$ for some $x \in (\alpha, \beta)$. By taking the derivative of g for n times, we find that
$$g^{(n)}(t) = f^{(n)}(t) - n!M \text{ for } a < t < b.$$
Hence, it suffices to show that $g^{(n)}(x) = 0$ for some $x \in (\alpha, \beta)$.

We note that $P^{(k)}_{n-1}(\alpha) = f^{(k)}(\alpha)$ for $0 \le k \le n - 1$, so
$$g^{(k)}(\alpha) = f^{(k)}(\alpha) - P^{(k)}_{n-1}(\alpha) - M \cdot n(n-1)\cdots(n-k+1)\cdot(\alpha - \alpha)^{n-k} = 0.$$
Also,
$$g(\beta) = f(\beta) - P_{n-1}(\beta) - \frac{f(\beta) - P_{n-1}(\beta)}{(\beta - \alpha)^n} \cdot (\beta - \alpha)^n = 0.$$
By Mean Value Theorem, there exists $x_1 \in (\alpha, \beta)$ such that $g'(x_1) = 0$. Similarly, we can apply Mean Value Theorem to g', then we can obtain $x_2 \in (\alpha, x_1)$ such that $g''(x_2) = 0$. Inductively repeating this process, after n steps, we find $x_n \in (\alpha, \beta)$ (in fact, $x_n \in (\alpha, x_{n-1})$) such that $g^{(n)}(x_n) = 0$.

□

Taylos's Theorem generallly says that if we know the bounds on $\left|f^{(n)}(x)\right|$, then we can approximate f by a polynomial of degree $n - 1$ and we can estimate the error.

Is that the case that every function which has infinitely many derivatives is locally (in a neighborhood of every point) a series?

Definition 1.23 (Real Analytic Function)**.** f is *real analytic* at x_0 if
$$\sum_{n=0}^{\infty} \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n \to f(x)$$
for every x in a neighborhood of x_0.

Is it the case that functions with infinitely many derivatives are analytic? No. Consider the following function
$$f(x) = \begin{cases} e^{-1/x} & x > 0 \\ 0 & x \le 0. \end{cases}$$
In fact, $f^{(n)}(0) = 0$ for every $n \in \mathbb{N}$. Also, $P_n(x) = 0$ is the case for all x in a neighborhood of 0, but of course $e^{-1/x}$ is not a *zero function*, that is almost everywhere zero.

Lecture 5
Wednesday
January 13

1.6 Differentiation of Vector-Valued Functions

The definition of derivative is not limited in real-valued functions. Further, we can apply the definition of derivative to complex-valued functions.

Definition 1.24 (Derivative of Complex-Valued Functions)**.** Let $f : [a, b] \to \mathbb{C}$, that is, f is complex valued. Then, f is *differentiable* at x if there exists
$$\lim_{y \to x} \frac{f(y) - f(x)}{y - x} = f'(x) \in \mathbb{C}.$$
Note that $f' : \operatorname{dom}(f') \to \mathbb{C}$ where $\operatorname{dom}(f') \subseteq \operatorname{dom}(f)$.

In particular, if $f(x) = f_1(x) + i \cdot f_2(x)$, where $f_1(x)$ is real part and $i \cdot f_2(x)$ is imaginary part, then f is differentiable at x if and only if f_1 and f_2 are differentiable and

$$f'(x) = f_1'(x) + i \cdot f_2'(x).$$

Proof Sketch.

$$\begin{aligned}
&\lim_{y\to x} \frac{f(y)-f(x)}{y-x} = f'(x) \\
\iff &\lim_{y\to x} \left| \frac{f(y)-f(x)}{y-x} - f'(x) \right| = 0 \\
\iff &\lim_{y\to x} \sqrt{\left| \frac{f_1(y)-f_1(x)}{y-x} - f_1'(x) \right|^2 + \left| \frac{f_2(y)-f_2(x)}{y-x} - f_2'(x) \right|^2} = 0 \\
\iff &\lim_{y\to x} \left| \frac{f_1(y)-f_1(x)}{y-x} - f_1'(x) \right| = 0 \text{ and } \lim_{y\to x} \left| \frac{f_2(y)-f_2(x)}{y-x} - f_2'(x) \right| = 0.
\end{aligned}$$

More generally, the definition of derivative can also be applied to vector-valued functions.

Definition 1.25 (Derivative of Vector-Valued Functions)**.** Let $f : [a,b] \to \mathbb{R}^n$, for $n \geq 1$. We say f is *differentiable* at $x \in [a,b]$ if there exists $p \in \mathbb{R}^n$ such that

$$\lim_{\substack{y\to x \\ y\neq x}} \frac{f(y)-f(x)}{y-x} = p = f'(x) \in \mathbb{R}^n.$$

This means that

$$\lim_{\substack{y\to x \\ y\neq x}} \left\| \frac{f(y)-f(x)}{y-x} - p \right\| = 0.$$

Note that $f' : \operatorname{dom}(f') \to \mathbb{R}^n$ where $\operatorname{dom}(f') \subseteq \operatorname{dom}(f)$.

Further, if $f'(x)$ is continuous, we say $f \in C_1$. More generally, for $k \in \mathbb{N}$, we say $f \in C^k$ if f has k continuous derivatives. If $f \in C^k$ for all k, we say $f \in C^\infty$, and call f *smooth*.

Proposition 1.26. *Let $f(x) = (f_1(x), f_2(x), \ldots, f_n(x)) : [a,b] \to \mathbb{R}^n$. Then, f is differentiable at x if and only if $f_i(x)$ is differentiable at x for $1 \leq i \leq n$ and*

$$f'(x) = (f_1'(x), f_2'(x), \ldots, f_n'(x)) \in \mathbb{R}^n.$$

More generally, f is C^k if and only its component functions are all C^k.

Theorem 1.27. *Let $f : [a,b] \to \mathbb{R}^n$ be differentiable at $x \in [a,b]$, then f is continuous at x.*

Theorem 1.28. *Let $f : [a,b] \to \mathbb{R}^n$ and $g : [a,b] \to \mathbb{R}^n$ be differentiable at $x \in [a,b]$. Then, $f+g$ and $f \cdot g$ (scalar product) and $\frac{f}{g}$ (in the complex case, since division is defined in the complex number system) are differentiable at x with the usual rules.*

(a) $(f+g)' = f' + g'$ *for addition.*

(b) $(f \cdot g)' = f' \cdot g + f \cdot g'$ *for product.*

(c) $\left(\frac{f}{g}\right)' = \frac{g \cdot f' - f \cdot g'}{g^2}$ *for quotient in the complex plane.*

However, *Mean Value Theorem* might fail in vector-valued functions.

Example 1.29 (Failure of Mean Value Theorem)**.** Let $f : \mathbb{R} \to \mathbb{C}$ with $f(x) = e^{ix} = \cos x + i \sin x$. We will define this function later. Then,

$$f(2\pi) - f(0) = 1 - 1 = 0.$$

The derivative of f is given by

$$f'(x) = ie^{ix} = i \cdot f(x).$$

Then,

$$|f'(x)| = \sin^2 x + \cos^2 x = 1$$

for all $x \in [0, 2\pi]$. Then,

$$f(2\pi) - f(0) = 0 \neq f'(x) \cdot 2\pi$$

for all $x \in [0, 2\pi]$. In this case, we conclude that f does not satisfy *Mean Value Theorem*.

Even so, *L'Hospital's Rule* fails, but not all the times.

Example 1.30 (Failure of L'Hospital's Rule)**.** Consider two complex-valued functions $f(x) = x$ and $g(x) = x + x^2 e^{i/x^2}$ on $(0, 1)$. By Example 1.29, we have $\left|e^{i/x^2}\right| = 1$ for $x \in \mathbb{R}$. Then,

$$\lim_{x\to 0} \frac{f(x)}{g(x)} = \lim_{x\to 0} \frac{x}{x + x^2 e^{i/x^2}} = \lim_{x\to 0} \frac{1}{1 + xe^{i/x^2}} = 1.$$

Also, $f'(x) = 1$ and

$$g'(x) = 1 + 2xe^{i/x^2} + x^2 e^{i/x^2} \cdot \left(-2 \cdot \frac{i}{x^3}\right) = 1 + \left(2x - \frac{2i}{x}\right) e^{i/x^2}.$$

Note that $\lim\limits_{x\to 0} f(x) = \lim\limits_{x\to 0} g(x) = 0$, so we can apply *L'Hospital's Rule*. However,

$$\lim_{x\to 0} \frac{f'(x)}{g'(x)} = \lim_{x\to 0} \frac{1}{1 + (2x - 2i/x)e^{i/x^2}} = 0 \neq \lim_{x\to 0} \frac{f(x)}{g(x)}.$$

In this case, we conclude that *L'Hospital's Rule* fails.

Will something weaker still hold? Oneo f the most useful consequence of *Mean Value Theorem* is that

$$|f(b) - f(a)| \leq (b - a) \cdot \sup_{x\in(a,b)} |f'(x)|.$$

If we try to prove $\sin x \leq x$ on $[0, 1]$, we get

$$\sin x = \sin x - \sin 0 \leq x \cdot \sup_{x\in(0,1)} |\cos x| \leq x.$$

Theorem 1.31. *Let $f : [a, b] \to \mathbb{R}^n$ be continuous on $[a, b]$ and differentiable on (a, b), then there exists $x \in (a, b)$ such that*

$$|f(b) - f(a)| \leq (b - a) \cdot |f'(x)|.$$

Proof. Let $z = f(b) - f(a) \in \mathbb{R}^n$ and define

$$\phi(t) = z \cdot f(t) : [a, b] \to \mathbb{R}.$$

So, $\phi(t)$ is a real-valued function and continuous on $[a, b]$ and differentiable on (a, b), so by Mean Value Theorem, there exists $x \in (a, b)$ such that

$$\phi(b) - \phi(a) = (b - a) \cdot \phi'(x) = (b - a) \cdot z \cdot f'(x).$$

On the other hand,

$$\phi(b) - \phi(a) = z \cdot f(b) - z \cdot f(a) = z \cdot (f(b) - f(a)) = z \cdot z = |z|^2.$$

Therefore, by *Cauchy Schwarz's Inequality*, we have

$$|z|^2 = \phi(b) - \phi(a) = (b - a) \cdot z \cdot f'(x) \leq (b - a) \cdot |z| \cdot |f'(x)|.$$

Then, $|z| \leq (b - a) \cdot f'(x)$. We conclude that

$$|f(b) - f(a)| \leq (b - a) \cdot |f'(x)|.$$

$\square$

2 Riemann-Stieltjes Integral

We have refined our rigorously theoretical understanding of *differentiation*, one of the two pillars of single variable calculus. The other pillar is, of course, *integration*. Previously, the *integral* was presented as a "limit of Riemann sums." However, this is not accurate. Further in 140C, we will discuss integration over sets other than intervals.

2.1 Definition and Existence of the Integral

Before carefully defining what is an *integral*, we first define *partition*. Also, it is assumed that we are working with a *bounded* function f on a closed interval $[a, b]$, which means there exists an $M > 0$ such that $|f(x)| \leq M$ for all $x \in [a, b]$.

Definition 2.1 (Partition)**.** Let $f : [a, b] \to \mathbb{R}$. A *partition* P of $[a, b]$ is a finite collection of points $P = \{x_0, x_1, \ldots, x_n\}$ where

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

We write $\Delta x_i = x_i - x_{i-1}$ for $i = 1, 2, \ldots, n$.

Definition 2.2 (Upper and Lower Sums)**.** Let $f : [a, b] \to \mathbb{R}$ be a bounded function. Then, given a partition P, we denote

$$M_i = \sup\{f(x) : x_{i-1} \leq x \leq x_i\}$$
$$m_i = \inf\{f(x) : x_{i-1} \leq x \leq x_i\}.$$

Further, we can define *upper and lower sums* of f on P:

$$U(P, f) = \sum_{i=1}^{|P|} M_i \cdot \Delta x_i \qquad \textit{upper Riemann sum}$$
$$L(P, f) = \sum_{i=1}^{|P|} m_i \cdot \Delta x_i \qquad \textit{lower Riemann sum}$$

Note that this only makes sense for *bounded* f; if f is not bounded, then on at least one partition interval $[x_{i-1}, x_i]$, f is not bounded, and so at least one of the terms in either $U(P, f)$ or $L(P, f)$ will be $\pm\infty$, but there could be more than one term that is $\pm\infty$, and so the sums may not even be defined. That's why we now restrict our attention to bounded functions for now.

Definition 2.3 (Upper and Lower Riemann Integrals)**.** Let $f : [a, b] \to \mathbb{R}$ be a bounded function. Then, we define the *upper and lower Riemann integrals* of f as

$$\overline{\int_a^b} f\,dx = \inf\{U(P, f) : P \text{ is a partition of } [a, b]\} \qquad \textit{upper Riemann integral}$$
$$\underline{\int_a^b} f\,dx = \sup\{L(P, f) : P \text{ is a partition of } [a, b]\} \qquad \textit{lower Riemann integral}$$

Definition 2.4 (Riemann Integrable)**.** If $\overline{\int_a^b} f\,dx = \underline{\int_a^b} f\,dx \in \mathbb{R}$, then we say that f is *Riemann integrable* and we write $f \in \mathscr{R}$ and

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

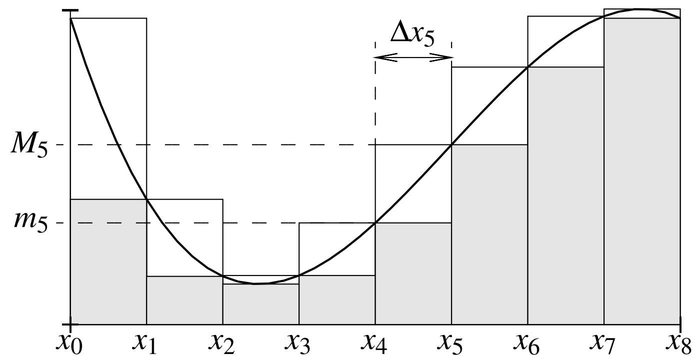


Figure 2.1: Riemann Sums

Lecture 6
Friday
January 15

Definition 2.5 (Riemann Stieltjes Integral)**.** Let α be a monotonically increasing function $\alpha : [a, b] \to \mathbb{R}$ such that $\alpha(a)$ and $\alpha(b)$ are finite (so α is bounded). Define

$$\Delta\alpha_i = \alpha(x_i) - \alpha(x_{i-1}) \geq 0.$$

For any real function f that is bounded on $[a, b]$, we define *upper and lower sums of Riemann-Stieltjes Integral* as

$$U(P, f, \alpha) = \sum_{i=1}^{|P|} M_i \cdot \Delta\alpha_i \quad \textit{upper sum}$$

$$L(P, f, \alpha) = \sum_{i=1}^{|P|} m_i \cdot \Delta\alpha_i \quad \textit{lower sum}$$

where M_i and m_i are defined exactly same as in Definition 2.2. Further, define *upper and lower Riemann-Stieltjes Integral*

$$\overline{\int_a^b} f\, d\alpha = \inf\{U(P, f, \alpha) : P \text{ is a partition of } [a, b]\} \quad \textit{upper integral}$$

$$\underline{\int_a^b} f\, d\alpha = \sup\{L(P, f, \alpha) : P \text{ is a partition of } [a, b]\} \quad \textit{lower integral.}$$

We say $f : [a, b] \to \mathbb{R}$ is *Riemann-Stieltjes integrable* if

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

and we write $f \in \mathscr{R}(\alpha)$ on $[a, b]$, in short, $f \in \mathscr{R}_a^b(\alpha)$.

When $\alpha(x) = x$, the above definition reduces to the standard *Riemann integral.* Thus, *Riemann-Stieltjes integration* is more general than *Riemann integration.*

Which functions are *Riemann integrable*? Suppose I have 2 partitions P_1 and P_2 of $[a, b]$ and I know something about $L(P_1, f, \alpha)$ and $U(P_2, f, \alpha)$, so how do I obtain a common information?

Definition 2.6 (Refinement)**.** We say that a partition P^* is a *refinement* of a partition P if $P \subseteq P^*$, that is, every point of P is a point of P^*.

Given 2 partitions, say P_1 and P_2, we say P^* is their *common refinement* if $P^* = P_1 \cup P_2$.

Lemma 2.7. *If P^* is a refinement of P, then*

(a) $L(P, f, \alpha) \leq L(P^*, f, \alpha)$.

(b) $U(P, f, \alpha) \geq U(P^*, f, \alpha)$.

Proof. We prove (a) (and (b) is similar). It suffices to consider that P^* contains just one more point than P, that is, let

$$P = \{x_0, x_1, \ldots, x_n\}$$
$$P^* = \{x_0, x_1, \ldots, x_{i-1}, x^*, x_i, \ldots, x_n\}.$$

Let

$$w_1 = \inf\{f(x) : x_{i-1} \leq x \leq x^*\} \text{ and } w_2 = \inf\{f(x) : x^* \leq x \leq x_i\}.$$

Recall that $m_i = \inf\{f(x) : x_{i-1} \leq x \leq x_i\}$, so by definition, we have $m_i \leq \min\{w_1, w_2\}$. Then,

$$\begin{aligned} L(P^*, f, \alpha) - L(P, f, \alpha) &= w_1 \left(\alpha(x^*) - \alpha(x_{i-1})\right) + w_2 \left(\alpha(x_i) - \alpha(x^*)\right) - m_i \left(\alpha(x_i) - \alpha(x_{i-1})\right) \\ &\geq \min\{w_1, w_2\} \cdot (\alpha(x^*) - \alpha(x_{i-1}) + \alpha(x_i) - \alpha(x^*)) - m_i(\alpha(x_i) - \alpha(x_{i-1})) \\ &= (\min\{w_1, w_2\} - m_i) \left(\alpha(x_i) - \alpha(x_{i-1})\right) \geq 0. \end{aligned}$$

If P^* has k more points than P, repeat the same proof k times.

□

Theorem 2.8. *Let f be a bounded function. Then,* $\underline{\int_a^b} f \, d\alpha \leq \overline{\int_a^b} f \, d\alpha$.

This is not trivial, since for any partition P, it is true that $L(P, f, \alpha) \leq U(P, f, \alpha)$ since $m_i \leq M_i$, which immediately implies

$$\sum_{i=1}^{|P|} m_i \cdot \Delta\alpha(x_i) \leq \sum_{i=1}^{|P|} M_i \cdot \Delta\alpha(x_i).$$

However, when we take inf and sup, does this still hold?

Proof. Let P_1 and P_2 be two partitions, then we want to compare $L(P_1, f, \alpha)$ and $U(P_2, f, \alpha)$. In order to do that, we cnosider a common refinement $P^* = P_1 \cup P_2$, then, by Lemma 2.7, we have

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

For any partition P_1 and P_2, $L(P_1, f, \alpha) \leq U(P_2, f, \alpha)$. Then,

$$\underline{\int_a^b} f \, d\alpha = \sup\{L(P_1, f, \alpha) : P_1 \text{ is a partition}\} \leq U(P_2, f, \alpha)$$

for every P_2 as a partition. Then,

$$\underline{\int_a^b} f \, d\alpha \leq \int \{U(P_2, f, \alpha) : P_2 \text{ is a partition}\} \leq \overline{\int_a^b} f \, d\alpha.$$

We conclude that

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

□

Therefore, for bounded function f, we have that

$$f \notin \mathscr{R}_a^b(\alpha) \iff \underline{\int_a^b} f \, d\alpha < \overline{\int_a^b} f \, d\alpha.$$

Theorem 2.9 (Criterion of Riemann-Stieltjes Integrability)**.** *Let $f : [a, b] \to \mathbb{R}$ be bounded. Then, $f \in \mathscr{R}_a^b(\alpha)$ if and only if for any $\varepsilon > 0$, there exists a partition P of $[a, b]$ such that*

$$0 \leq U(P, f, \alpha) - L(P, f, \alpha) < \varepsilon.$$

This is a much better criterion for *Riemann-Stieltjes integrability.*

Proof. Assume for any $\varepsilon > 0$, there exists a partition P such that

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

Since

$$\overline{\int_a^b} f \, d\alpha \leq U(P, f, \alpha) \text{ and } \underline{\int_a^b} f \, d\alpha \geq L(P, f, \alpha)$$

then

$$0 \leq \overline{\int_a^b} f \, d\alpha - \underline{\int_a^b} f \, d\alpha \leq U(P, f, \alpha) - L(P, f, \alpha).$$

Since this is true for all $\varepsilon > 0$, then let $\varepsilon \to 0$, we get

$$0 \leq \overline{\int_a^b} f \, d\alpha - \underline{\int_a^b} f \, d\alpha \leq 0$$

so

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

We conclude that $f \in \mathscr{R}_a^b(\alpha)$.

Assume that $f \in \mathscr{R}_a^b(\alpha)$, let $\varepsilon > 0$ be given. By definition of sup and inf, there exists two partitions P_1 and P_2 such that

$$U(P_2, f, \alpha) - \int_a^b f \, d\alpha = U(P_2, f, \alpha) - \overline{\int_a^b} f \, d\alpha < \frac{\varepsilon}{2}$$

and

$$\int_a^b f \, d\alpha - L(P_1, f, \alpha) = \underline{\int_a^b} f \, d\alpha - L(P_1, f, \alpha) < \frac{\varepsilon}{2}.$$

Let P be a common refinement of P_1 and P_2, that is, $P = P_1 \cup P_2$, then we have

$$\begin{aligned}
0 &\leq U(P, f, \alpha) - L(P, f, \alpha) \\
&= U(P_2, f, \alpha) - L(P_1, f, \alpha) \text{ by Lemma 2.7} \\
&= U(P_2, f, \alpha) - \int_a^b f \, d\alpha + \int_a^b f \, d\alpha - L(P_1, f, \alpha) \\
&< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.
\end{aligned}$$

$\square$

Indeed, this is a convenient criterion for *Riemann-Stieltjes integrability.* We state some closely related facts before we explore which types of functions are *Riemann-Stieltjes integrable.* Lecture 7 Wednesday January 20

Theorem 2.10. *If $U(P, f, \alpha) - L(P, f, \alpha) < \varepsilon$ holds for some partition P and some $\varepsilon > 0$, then*

(a) $U(P, f, \alpha) - L(P, f, \alpha) < \varepsilon$ holds for every refinement of P.

(b) If the partition P that works is $P = \{x_0, x_1, \ldots, x_n\}$ and if $s_i, t_i \in [x_{i-1}, x_i]$, then

$$\sum_{i=1}^{n} |f(s_i) - f(t_i)| \Delta\alpha_i < \varepsilon.$$

(c) If $f \in \mathscr{R}_a^b(\alpha)$ and $U(P, f, \alpha) - L(P, f, \alpha)$ holds for $P = \{x_0, x_1, \ldots, x_n\}$ and t_i is an arbitrary point in $[x_{i-1} - x_i]$ for $i = 1, 2, \ldots, n$, then

$$\left| \sum_{i=1}^{n} f(t_i) \cdot \Delta\alpha_i - \int_a^b f \, d\alpha \right| < \varepsilon.$$

Briefly, (a) says that if $U(P, f, \alpha) - L(P, f, \alpha)$ is true, refining the partition only makes it better, that is,

$$U(P, f, \alpha) - L(P, f, \alpha) \to 0 \text{ as refining } P.$$

(b) says that if $U(P, f, \alpha) - L(P, f, \alpha) - \varepsilon$ holds with m_i, M_i, then it also holds for $f(s_i), f(t_i)$ for $s_i, t_i \in [x_{i-1}, x_i]$. (c) says that if $f \in \mathscr{R}_a^b(\alpha)$ then you can approximate $\displaystyle\int_a^b f \, d\alpha$ with $\displaystyle\sum_i f(t_i) \cdot \Delta\alpha_i$ for any choice of $t_i \in [x_{i-1}, x_i]$.

Proof.

(a) Since

$$L(P, f, \alpha) \leq L(P^*, f, \alpha) \leq U(P^*, f, \alpha) \leq U(P, f, \alpha),$$

then the result is immediate.

(b) For any $s_i, t_i \in [x_{i-1}, x_i]$, by the definition of m_i and M_i, then $f(s_i), f(t_i) \in [m_i, M_i]$, so

$$|f(s_i) - f(t_i)| \leq M_i - m_i.$$

Then,

$$\begin{aligned} \sum_{i=1}^{n} |f(s_i) - f(t_i)| \cdot \Delta\alpha_i &\leq \sum_{i=1}^{n} (M_i - m_i) \cdot \Delta\alpha_i \\ &= \sum_{i=1}^{n} M_i \cdot \Delta\alpha_i - \sum_{i=1}^{n} m_i \cdot \Delta\alpha_i \\ &= U(P, f, \alpha) - L(P, f, \alpha) < \varepsilon. \end{aligned}$$

(c) For $t_i \in [x_{i-1}, x_i]$, by the definition of m_i and M_i, then we have $m_i \leq f(t_i) \leq M_i$ for $i = 1, 2, \ldots, n$. Therefore,

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

Further,

$$L(P, f, \alpha) \leq \int_a^b f \, d\alpha \leq U(P, f, \alpha),$$

so we have

$$\int_a^b f \, d\alpha - \sum_{i=1}^n f(t_i) \cdot \Delta\alpha_i \leq U(P, f, \alpha) - L(P, f, \alpha) < \varepsilon$$
$$\sum_{i=1}^n f(t_i) \cdot \Delta\alpha_i - \int_a^b f \, d\alpha \leq U(P, f, \alpha) - L(P, f, \alpha) < \varepsilon.$$

We conclude that

$$\left| \int_a^b f \, d\alpha - \sum_{i=1}^n f(t_i) \cdot \Delta\alpha_i \right| < \varepsilon.$$

□

Which properties of f and α are *sufficient* to guarantee that $f \in \mathscr{R}(\alpha)$? For a function f to be *Riemann integrable*, we will discuss the criterion in 140C, with the *Lebesgue integral*.

Theorem 2.11 (Continuity Implies Integrability)**.** *If f is continuous on $[a, b]$, then f is Riemann-Stieltjes integrable on $[a, b]$.*

To show a function is integrable, we want to show it satisfies the criterion

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

For every ε, we need to find a partition P such that $U(P, f, \alpha) - L(P, f, \alpha) < \varepsilon$. We want to find a partition such that $M_i - m_i$ is very small or $\Delta\alpha_i$ is very small.

Proof. Let $\varepsilon > 0$. Pick an $\eta > 0$ such that $(\alpha(b) - \alpha(a)) \cdot \eta < \varepsilon$. Since f is continuous on $[a, b]$, which is compact, then f is uniformly continuous on $[a, b]$. Then, there exists $\delta > 0$ (independent of $x \in [a, b]$) depends only on f and $[a, b]$, such that

$$|f(x) - f(t)| < \eta \text{ for } x, t \in [a, b] \text{ and } |x - t| < \delta.$$

Let P be a partition on $[a, b]$ such that $\Delta x_i < \delta$ for $1 \leq i \leq n$, then $M_i - m_i \leq \eta$ for $i = 1, 2, \ldots, n$ by uniform continuity. Then,

$$U(P, f, \alpha) - L(P, f, \alpha) = \sum_{i=1}^n (M_i - m_i) \cdot \Delta\alpha_i \leq \eta \sum_{i=1}^n \Delta\alpha_i = \eta \cdot (\alpha(b) - \alpha(a)) < \varepsilon$$

so we conclude that $f \in \mathscr{R}_a^b(\alpha)$.

□

So, now we know how to integrate continuous functions. In particular, taking $\alpha(x) = x$, this gives us the usual *Riemann integral* of continuous functions, which is the main object of study in integral calculus. However, the class of functions that can be integrated is not limited in continuous functions.

Theorem 2.12 (Monotonicity Implies Integrability)**.** *If f is monotonic on $[a, b]$ and α is continuous on $[a, b]$, then f is Riemann-Stieltjes integrable on $[a, b]$.*

Proof. Let $\varepsilon > 0$ be given and $n \in \mathbb{N}$. Since $\alpha : [a,b] \to \mathbb{R}$ is continuous and monotone increasing for $i = 1, 2, \ldots, n$, there exists $x_i \in [a,b]$ such that

$$x_i > x_{i-1} \text{ and } \alpha(x_i) = \alpha(a) + i \cdot \frac{\alpha(b) - \alpha(a)}{n}$$

so that $\Delta\alpha_i = \dfrac{\alpha(b) - \alpha(a)}{n}$ by Intermediate Value Theorem. (If $\alpha(x) = 0$, this is equivalent to divide $[a,b]$ in n subintervals of length $\dfrac{b-a}{n}$.) If $\alpha(x)$ is any monotone function, without loss of generality, assume f is increasing. Then,

$$M_i = f(x_i) \text{ and } m_i = f(x_{i-1})$$

for $1 \le i \le n$. Then,

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

Pick an $n \in \mathbb{N}$ sufficiently large (depends on ε), then we conclude that $f \in \mathscr{R}_a^b(\alpha)$.

□

From Theorem 2.11 and Theorem 2.12, we can see that to prove a function f is integrable, our target is to prove that

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

is small. If f is continuous, then we can guarantee that $M_i - m_i$ is small, hence the sum will be small. If f is monotonic and α is continuous, then $\Delta\alpha_i$ will be small and hence the sum will be small. In other words, two functions f and α support each other and in a proof of integrability, we will always want to restrict either f or α, and sometimes both.

Theorem 2.13. *Suppose f is bounded on $[a,b]$, f has only finitely many discontinuities on $[a,b]$, and α is continuous at every point at which f is discontinuous. Then f is Riemann-Stieltjes integrable on $[a,b]$.*

Lecture 8
Friday
January 22

Proof. Let $\varepsilon > 0$ be given. Let $M = \sup_{[a,b]} |f(x)|$ and let $E \subseteq [a,b]$ be the set of points of discontinuity of f. Since E is finite and α is continuous on E, then we can cover E with finitely many disjoint intervals $(u_j, v_j) \subseteq [a,b]$ such that

(1) $E \subseteq \bigcup_{j=1}^{N} (u_j, v_j)$,

(2) $[u_j, v_j] \cap [u_i, v_i] = \varnothing$ for $i \ne j$ (pairwise disjoint), and

(3) $\displaystyle\sum_{j=1}^{N} (\alpha(v_j) - \alpha(u_j)) < \varepsilon$ follows by the continuity of α on E.

The set $K = [a,b] \setminus \bigcup_{j=1}^{N} (u_j, v_j)$ is compact and f is continuous on K, so $f : K \to \mathbb{R}$ is uniformly continuous. Then, there exists a $\delta > 0$ such that for all $t, s \in K$, $|f(s) - f(t)| < \varepsilon$ if $|t - s| < \delta$. Let $P = \{x_0, x_1, \ldots, x_n\}$ be a partition of $[a,b]$ such that

(1) $u_j \in P$ and $v_j \in P$ for all $j = 1, \ldots, N$,

(2) $x \notin P$ if $x \in (u_j, v_j)$,

(3) if $x_{i-1} \neq u_j$ for $j = 1, \ldots, N$, then $\Delta x_i < \delta$.

Since $M = \sup_{[a,b]} |f(x)|$, then $M_i - m_i \leq 2M$ for $1 \leq i \leq n$, and if $x_{i-1} \neq u_j$, then $\Delta x_i < \delta$, so $M_i - m_i < \varepsilon$ (by the continuity of f on K). Therefore,

$$\begin{aligned}
U(P, f, \alpha) - L(P, f, \alpha) &= \sum_{i=1}^{|P|} (M_i - m_i) \cdot \Delta\alpha_i \\
&= \sum_{\substack{i=1 \\ x_{i-1} \neq u_j}}^{|P|} (M_i - m_i) \cdot \Delta\alpha_i + \sum_{\substack{i=1 \\ x_{i-1} = u_j}}^{|P|} (M_i - m_i) \cdot \Delta\alpha_i \\
&\leq \varepsilon \sum_{\substack{i=1 \\ x_{i-1} \neq u_j}}^{|P|} \Delta\alpha_i + 2M \sum_{\substack{i=1 \\ x_{i-1} = u_j}}^{|P|} \Delta\alpha_i \\
&\leq \varepsilon(\alpha(b) - \alpha(a)) + 2M \cdot \varepsilon.
\end{aligned}$$

Since ε is arbitrary, then we can conclude that $f \in \mathscr{R}_a^b(\alpha)$.

$\square$

If f is *Riemann-Stieltjes integrable*, how about f^2? How about more general compositions of functions?

Theorem 2.14 (Integrability of Composition)**.** *Suppose f is Riemann-Stieltjes integrable on $[a, b]$ and $m \leq f \leq M$. Let $\phi : [m, M] \to \mathbb{R}$ be continuous. Then, $h(x) = \phi(f(x))$ is Riemann-Stieltjes integrable on $[a, b]$.*

Proof. Fix an $\varepsilon > 0$. Since ϕ is continuous on $[m, M]$, which is compact, so ϕ is uniformly continuous on $[m, M]$ and so there exists $0 < \delta < \varepsilon$ such that $|\phi(s) - \phi(t)| < \varepsilon$ if $|s - t| < \delta$ for $s, t \in [m, M]$. Since $f \in \mathscr{R}_a^b(\alpha)$, there exists a partition $P = \{x_0, x_1, \ldots, x_n\}$ on $[a, b]$ such that

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

Let

$$\begin{aligned}
m_i &= \inf\{f(x) : x_{i-1} \leq x \leq x_i\}, \\
m_i^* &= \inf\{h(x) : x_{i-1} \leq x \leq x_i\}, \\
M_i &= \sup\{f(x) : x_{i-1} \leq x \leq x_i\}, \\
M_i^* &= \sup\{h(x) : x_{i-1} \leq x \leq x_i\}.
\end{aligned}$$

Let $A, B \subseteq \{1, 2, \ldots, n\}$ be defined as follows: $i \in A$ if $M_i - m_i < \delta$ and $i \in B$ if $M_i - m_i \geq \delta$. Then, $A \cup B = \{1, 2, \ldots, n\}$ and $A \cap B = \varnothing$. If $i \in A$, then $M_i^* - m_i^* < \varepsilon$ since $h = \phi \circ f$ and ϕ is uniformly continuous and $f([x_{i-1}, x_i]) \subseteq [m, M]$ and $M_i - m_i < \delta$. If $i \in B$, then $M_i^* - m_i^* \leq 2 \sup_{[m,M]} |\phi|$. Note that

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

so $\sum_{i\in B} \Delta\alpha_i < \delta < \varepsilon$. Therefore, we can conclude

$$\begin{aligned} U(P,h,\alpha) - L(P,h,\alpha) &= \sum_{i\in A}(M_i^* - m_i^*)\cdot\Delta\alpha_i + \sum_{i\in B}(M_i^* - m_i^*)\cdot\Delta\alpha_i \\ &\le \varepsilon\cdot\sum_{i\in A}\Delta\alpha_i + 2\sup_{[m,M]}|\phi|\cdot\sum_{i\in B}\Delta\alpha_i \\ &\le \varepsilon[\alpha(b)-\alpha(a) + 2\sup_{[m,M]}|\phi|] \end{aligned}$$

so $h \in \mathscr{R}_a^b(\alpha)$.

□

2.2 Properties of Integrals

Proved in Lecture 9

We now turn to the basic properties of the *Riemann-Stieltjes integral*, all of which reflect its nature as a kind of limit sum. We state them as a sequence of theorems. Let $f, f_1, f_2 : [a,b] \to \mathbb{R}$ are bounded functions, and $\alpha, \alpha_1, \alpha_2 : [a,b] \to \mathbb{R}$ are monotone increasing functions.

Theorem 2.15. *If $f_1, f_2 \in \mathscr{R}_a^b(\alpha)$, then $f_1 + f_2 \in \mathscr{R}_a^b(\alpha)$. If $c \in \mathbb{R}$ and $f \in \mathscr{R}_a^b(\alpha)$, then $c\cdot f \in \mathscr{R}_a^b(\alpha)$. Further,*

$$\int_a^b (f_1+f_2)\,d\alpha = \int_a^b f_1\,d\alpha + \int_a^b f_2\,d\alpha$$
$$\int_a^b (c\cdot f)\,d\alpha = c\cdot\int_a^b f\,d\alpha.$$

Proof. Let $f = f_1 + f_2$ and P be a partition of $[a,b]$, then

$$\inf_{[x_{i-1},x_i]} f_1 + \inf_{[x_{i-1},x_i]} f_2 \le \inf_{[x_{i-1},x_i]} f \le \sup_{[x_{i-1},x_i]} f \le \sup_{[x_{i-1},x_i]} f_1 + \sup_{[x_{i-1},x_i]} f_2$$

fir $1 \le i \le |P|$. Then,

$$L(P,f_1,\alpha) + L(P,f_2,\alpha) \le L(P,f,\alpha) \le U(P,f,\alpha) \le U(P,f_1,\alpha) + U(P,f_2,\alpha).$$

Let $\varepsilon > 0$ be given. Since $f_j \in \mathscr{R}(\alpha)$ for $j = 1,2$, then there are partitions P_j of $[a,b]$ such that $U(P_j,f_j,\alpha) - L(P_j,f_j,\alpha) < \frac{\varepsilon}{2}$. Let P be a common refinement of P_1 and P_2, then we have

$$U(P,f,\alpha) - L(P,f,\alpha) \le U(P,f_1,\alpha) - L(P,f_1,\alpha) + U(P,f_2,\alpha) - L(P,f_2,\alpha) < \varepsilon.$$

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

Moreover,

$$U(P,f_j,\alpha) < \int_a^b f_j\,d\alpha + \frac{\varepsilon}{2} \text{ for } j = 1,2.$$

then

$$\int_a^b f\,d\alpha \le U(P,f,\alpha) \le U(P,f_1,\alpha) + U(P,f_2,\alpha) < \int_a^b f_1\,d\alpha + \int_a^b f_2\,d\alpha + \varepsilon.$$

Let $\varepsilon \to 0$, we have

$$\int_a^b f\,d\alpha \le \int_a^b f_1\,d\alpha + \int_a^b f_2\,d\alpha.$$

Using the fact that $L(P, f_j, \alpha) - \frac{\varepsilon}{2} \leq \int_a^b f_j \, d\alpha$, in the same way, we can get

$$\int_a^b f_1 \, d\alpha + \int_a^b f_2 \, d\alpha \leq \int_a^b f \, d\alpha.$$

Therefore, we can conclude that

$$\int_a^b f \, d\alpha = \int_a^b f_1 \, d\alpha + \int_a^b f_2 \, d\alpha.$$

The proof for $c \cdot f$ is similar, so left as an exercise.

□

Theorem 2.16. *If $f_1, f_2 \in \mathscr{R}_a^b(\alpha)$ and $f_1(x) \leq f_2(x)$ for all $x \in [a, b]$, then*

$$\int_a^b f_1 \, d\alpha \leq \int_a^b f_2 \, d\alpha.$$

Proof. Suppose that $f \geq 0$ and $f \in \mathscr{R}(\alpha)$ on $[a, b]$, then let $P = \{a, b\}$, we have

$$\int_a^b f \, d\alpha \geq L(P, f, \alpha) \geq \inf_{[a,b]} f \cdot (\alpha(b) - \alpha(a)) \geq 0.$$

Pick f_1 and f_2 as in the statement, let $g = f_2 - f_1 \geq 0$ on $[a, b]$ by assumption. Then, $\int_a^b g \, d\alpha \geq 0$. Therefore, we conclude that

$$\int_a^b f_2 \, d\alpha \geq \int_a^b f_1 \, d\alpha.$$

□

Theorem 2.17. *Let $c \in [a, b]$. Then, $f \in \mathscr{R}_a^b(\alpha)$ if and only if $f \in \mathscr{R}_a^c(\alpha)$ and $f \in \mathscr{R}_c^b(\alpha)$ and*

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

Proof. It is obvious if $c = a$ or $c = b$, so we fix $c \in (a, b)$. Suppose $f \in \mathscr{R}(\alpha)$ on $[a, b]$. Then, given $\varepsilon > 0$, there exists a partition P of $[a, b]$ such that

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

Let $P' = P \cup \{c\}$ be a refinement of P (might be P itself) and let $Q = P' \cap [a, c]$ and $R = P' \cap [c, b]$. Then,

(a) $P' = Q \cup R$

(b) Q is a part of $[a, c]$, R is a part of $[c, b]$.

(c) $L(P', f, \alpha) = L(Q, f, \alpha) + L(R, f, \alpha)$, $U(P', f, \alpha) = U(Q, f, \alpha) + L(R, f, \alpha)$.

Therefore,

$$\begin{aligned} U(Q, f, \alpha) - L(Q, f, \alpha) &= (U(P', f, \alpha) - L(P', f, \alpha)) - (U(R, f, \alpha) - L(R, f, \alpha)) \\ &\leq U(P', f, \alpha) - L(P', f, \alpha) \\ &\leq U(P, f, \alpha) - L(P, f, \alpha) < \varepsilon. \end{aligned}$$

So, $f \in \mathscr{R}(\alpha)$ on $[a, c]$. In the same way, $f \in \mathscr{R}(\alpha)$ on $[a, b]$.

Vice versa, if $f \in \mathscr{R}(\alpha)$ on $[a,c]$ and $[c,b]$, then there exists partitions Q and R of $[a,c]$ and $[c,b]$ respectively such that

$$U(Q,f,\alpha) - L(Q,f,\alpha) < \frac{\varepsilon}{2} \text{ and } U(R,f,\alpha) - L(R,f,\alpha) < \frac{\varepsilon}{2}.$$

Let $P = Q \cup R$, then

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

so $f \in \mathscr{R}_a^b(\alpha)$. The rest part of the proof is left as an exercise.

$\square$

Theorem 2.18. *If $f \in \mathscr{R}_a^b(\alpha)$ and $\sup\limits_{[a,b]} |f| \leq M$, then*

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

Proof. By assumption, we have $-M \leq -|f| \leq f \leq |f| \leq M$, so by Theorem 2.16

$$-M \cdot (\alpha(b) - \alpha(a)) \leq \int_a^b f \, d\alpha \leq M \cdot (\alpha(b) - \alpha(a)).$$

Therefore,

$$\left| \int_a^b f \, d\alpha \right| \leq M \cdot (\alpha(b) - \alpha(a)).$$

$\square$

Theorem 2.19. *If $f \in \mathscr{R}_a^b(\alpha_1)$ and $f \in \mathscr{R}_a^b(\alpha_2)$, then $f \in \mathscr{R}_a^b(\alpha_1 + \alpha_2)$ 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 $f \in \mathscr{R}_a^b(\alpha)$ and $c > 0$, then $f \in \mathscr{R}_a^b(c \cdot \alpha)$ and

$$\int_a^b f \, d(c\alpha) = c \int_a^b f \, d\alpha.$$

Proof. We note that

$$\begin{aligned}
\Delta(\alpha_1 + \alpha_2)_i &= (\alpha_1 + \alpha_2)(t_i) - (\alpha_1 + \alpha_2)(t_{i-1}) \\
&= \alpha_1(t_i) - \alpha_1(t_{i-1}) + \alpha_2(t_i) - \alpha_2(t_{i-1}) \\
&= \Delta(\alpha_1)_i + \Delta(\alpha_2)_i
\end{aligned}$$

for each $1 \leq i \leq n$. Then, for partition P, we have

$$U(P, f, \alpha_1 + \alpha_2) = U(P, f, \alpha_1) + U(P, f, \alpha_2).$$

Taking inf on both sides yields

$$\int_a^b f \, d(\alpha_1 + \alpha_2) = \int_a^b f \, d\alpha_1 + \int_a^b f \, d\alpha_2.$$

We can handle lower sums in a similar way. Further, the proof for $c \cdot f$ is similar.

$\square$

Lecture 9
Monday
January 25

Theorem 2.20. *Let $f, g \in \mathscr{R}(\alpha)$ on $[a, b]$. Then,*

(a) $f \cdot g \in \mathscr{R}(\alpha)$ *on* $[a, b]$.

(b) $|f| \in \mathscr{R}(\alpha)$ *on* $[a, b]$ *and* $\left| \int_a^b f \, d\alpha \right| \leq \int_a^b |f| \, d\alpha.$

Recall the integrability of composition functions. Suppose $f \in \mathscr{R}_a^b(\alpha)$, $m \leq f \leq M$, and $\phi : [m, M] \to \mathbb{R}$ is continuous. Then, $h(x) = \phi \circ f$ is *Riemann-Stieltjes integrable* on $[a, b]$.

Proof.

(a) We note that

$$f \cdot g = \frac{1}{4} \left[(f+g)^2 - (f-g)^2 \right].$$

Since $f, g \in \mathscr{R}_a^b(\alpha)$, then $f + g, f - g \in \mathscr{R}_a^b(\alpha)$ as well. If $h \in \mathscr{R}(\alpha)$, then $h^2 \in \mathscr{R}(\alpha)$. Thus, $(f+g)^2, (f-g)^2 \in \mathscr{R}(\alpha)$. Therefore,

$$f \cdot g = \frac{1}{4} \left[(f+g)^2 - (f-g)^2 \right] \in \mathscr{R}(\alpha).$$

(b) Let $\phi(t) = |t|$, which is continuous. So, $|f| \in \mathscr{R}(\alpha)$. Moreover, let $c = \pm 1$ be chosen, so that

$$c \cdot \int_a^b f \, d\alpha \geq 0 \qquad \text{for } c = \operatorname{sign}\left(\int_a^b f \, d\alpha \right).$$

Then, $c \cdot f \leq |c \cdot f| = |c| \cdot |f| = |f|$ gives

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

□

For now, we know that we can integrate a reasonably large class of functions, against arbitrary monotone increasing integrators. What benefit have we gained by including more general integrator weights α?

Definition 2.21 (Step Function)**.** The *unit step function* I is defined by

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

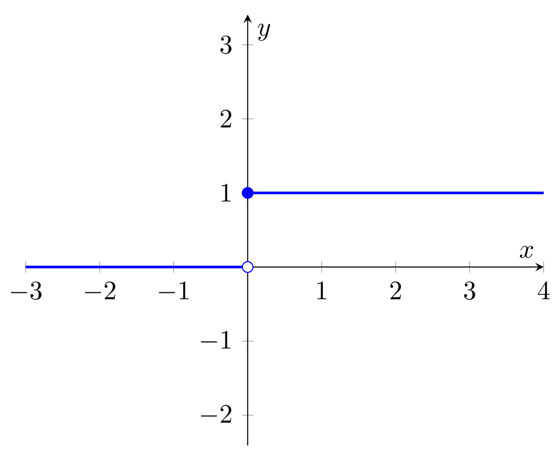


Figure 2.2: Graph of the *unit step function* $I(x)$

Theorem 2.22. *If $a < s < b$, f is bounded on $[a,b]$, f is continuous at s, and $\alpha(x) = I(x-s)$, then*

$$\int_a^b f\,d\alpha = f(s).$$

Proof. For any partition $P = \{x_0, x_1, \ldots, x_n\}$. Suppose $s \in (x_{i-1}, x_i)$, then

$$U(P,f,\alpha) = \sum_{j=1}^{|P|} M_j \cdot \Delta\alpha_j = M_i$$

since $\Delta\alpha_j = 0$ for $j \neq i$ and $\Delta\alpha_j = 1$ for $j = i$. Similarly,

$$L(P,f,\alpha) = \sum_{j=1}^{|P|} m_j \cdot \Delta\alpha_j = m_i.$$

On the other hand, for $\varepsilon > 0$, there exists a $\delta > 0$ such that $|x_{i-1} - x_i| < \delta$ implies

$$M_i - f(s) < \varepsilon \text{ and } f(s) - m_i < \varepsilon$$

since f is continuous. Thus,

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

Therefore, by taking sup adn inf for two inequality, we get

$$\left| \int_a^b f\,d\alpha - f(s) \right| < \varepsilon.$$

Since ε is arbitrary, so we conclude that

$$\int_a^b f\,d\alpha = f(s).$$

$\square$

The graph below helps us understand the proof easier and visualize the partition with the change of α.

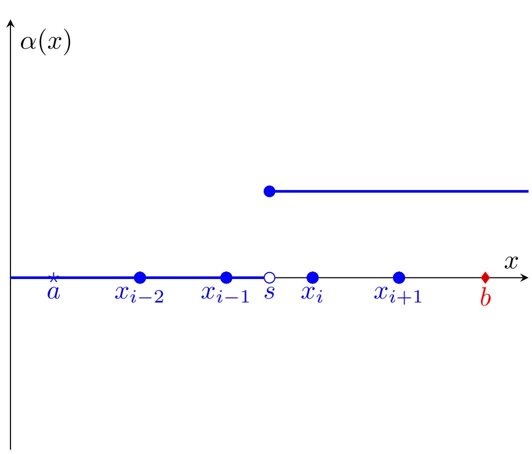


Figure 2.3: Graph of $I(x-s)$

The backend idea for this proof is the following. Fix an $\varepsilon > 0$,

$$\int_a^b f\, d\alpha \geq L(P,f,\alpha) \Longrightarrow f(s) - \int_a^b f\, d\alpha \leq f(s) - L(P,f,\alpha) < \varepsilon$$
$$\int_a^b f\, d\alpha \leq U(P,f,\alpha) \Longrightarrow \int_a^b f\, d\alpha - f(s) \leq U(P,f,\alpha) - f(s) < \varepsilon.$$

Lecture 11 Friday January 29

We can use the additivity of the integral to generalize Theorem 2.22, and extend this result to discrete infinite sums. If (s_n) is an increasing (possibly finite) sequence in (a,b), and if (c_n) is a nonnegative sequence such that $\sum_{n=1}^{\infty} c_n < \infty$, we can define

$$\alpha(t) = \sum_{n=1}^{\infty} c_n \cdot I(t - s_n).$$

This is a well defined and monotone increasing function on $[a,b]$. It is constant on $[a, a+s_1]$, taking value c_1; it is constant on $(a+s_1, a+s_2]$, taking value c_2; and so forth.

Theorem 2.23. *Suppose $c_n \geq 0$ for $n \geq 1$ such that $\sum_{n=1}^{\infty} c_n$ converges. Let $(s_n)_n$ be a sequence of distinct points in (a,b) and*

$$\alpha(x) = \sum_{n=1}^{\infty} c_n \cdot I(x - s_n).$$

Let f be continuous on $[a,b]$. Then

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

Proof. Note that $\alpha(a) = 0$ and $\alpha(b) = \sum_{n=1}^{\infty} c_n$ and α is monotone increasing since $c_n \geq 0$.

Let $\varepsilon > 0$ be fixed and pick $N \in \mathbb{N}$ such that $\sum_{n=N+1}^{\infty} c_n < \varepsilon$ (such an N exists since $\sum_{n=1}^{\infty} c_n$ converges). Let

$$\alpha_1(x) = \sum_{n=1}^{N} c_n \cdot I(x - s_n) \text{ and } \alpha_2(x) = \sum_{n=N+1}^{\infty} c_n \cdot I(x - s_n).$$

Then, we have

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

As $a < s_n < b$, so $I(a - s_n) = 0$ adn $I(b - s_n) = 1$ for every n. Since $\sum_{n=N+1}^{\infty} c_n < \varepsilon$, so

$$\alpha_2(b) - \alpha_2(a) = \sum_{n=N+1}^{\infty} c_n < \varepsilon.$$

Then, set $M = \sup_{[a,b]} |f|$ (this exists since f is continuous on compact set $[a,b]$),

$$\left| \int_a^b f\, d\alpha_2 \right| \leq \int_a^b |f| d\, \alpha_2 \leq U(\{a,b\}, |f|, \alpha_2) < M \cdot \varepsilon.$$

To conclude, we have

$$\left|\int_a^b f\,d\alpha - \sum_{n=1}^{N} c_n \cdot f(s_n)\right| = \left|\int_a^b f\,d\alpha_1 + \int_a^b f\,d\alpha_2 - \sum_{n=1}^{N} c_n \cdot f(s_n)\right| = \left|\int_a^b f\,d\alpha_2\right| < M \cdot \varepsilon.$$

Since this is true for all $\varepsilon > 0$, let $N \to \infty$, so

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

$\square$

Thus, if we use a step function as our integrator, we unify the theory of infinite series and integration. On the flip side, what will happen if α is differentiable? In that case, we will see that integration with respect to α actually involved the derivative α'. This will motivate the connection afterward to the *Fundamental Theorem of Calculus.*

Theorem 2.24. *Let $\alpha : [a, b] \to \mathbb{R}$ be monotone increasing, and suppose that α is differentiable on $[a, b]$ with $\alpha' \in \mathscr{R}$. Let $f : [a, b] \to \mathbb{R}$ be bounded. Then, $f \in \mathscr{R}(\alpha)$ if and only if $f \cdot \alpha' \in \mathscr{R}$, and*

$$\int_a^b f\,d\alpha = \int_a^b f(x) \cdot \alpha'(x)\,dx.$$

This theorem is often summarized as $d\alpha(t) = \alpha'(t)\,dt$. For example,

$$\int_0^1 f(x)\,dx^2 = \int_0^1 f(x) \cdot 2x\,dx.$$

Proof. Let $\varepsilon > 0$ be given. As $\alpha' \in \mathscr{R}$ on $[a, b]$, so there exists a partition P of $[a, b]$ such that

$$U(P, \alpha') - L(P, \alpha') < \varepsilon.$$

By the Mean Value Theorem, for $1 \le i \le n$, there exists $t_i \in [x_{i-1}, x_i]$ such that

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

If $s_i \in [x_{i-1}, x_i]$, then, as $t_i \in [x_{i-1}, x_i]$ also, so we conclude that

$$\sum_{i=1}^{n} |\alpha'(t_i) - \alpha'(s_i)| \cdot \Delta x_i \le U(P, \alpha') - L(P, \alpha') < \varepsilon.$$

Let $M = \sup\limits_{[a,b]} |f|$. Since

$$\left|\sum_{i=1}^{n} f(s_i) \cdot \Delta\alpha_i - \sum_{i=1}^{n} f(s_i) \cdot \alpha'(s_i) \cdot \Delta x_i\right| \le \sum_{i=1}^{n} |f(s_i)| \cdot |\alpha'(s_i) - \alpha'(t_i)| \cdot \Delta x_i < M \cdot \varepsilon,$$

so, with $f(s_i) \cdot \alpha'(s_i) \le \sup\limits_{[x_{i-1}, x_i]} f \cdot \alpha'$ for $1 \le i \le n$, we have

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

Then, $U(P, f, \alpha) \le U(P, f \cdot \alpha') + M \cdot \varepsilon$. With the same argument, we get

$$U(P, f \cdot \alpha') \le U(P, f, \alpha) + M \cdot \varepsilon.$$

Thus,
$$|U(P,f,\alpha) - U(P,f\cdot\alpha')| \le M\cdot\varepsilon.$$
Now, if $\widetilde{P}$ be any partition of $[a,b]$ such that $P \subseteq \widetilde{P}$, then
$$L(P,\alpha') \le L(\widetilde{P},\alpha') \le U(\widetilde{P},\alpha') \le U(P,\alpha'),$$
which means
$$U(\widetilde{P},\alpha') - L(\widetilde{P},\alpha') \le U(P,\alpha') - L(P,\alpha').$$
Thus, if $\widetilde{P}$ is any refinement of P, then
$$\left|U(\widetilde{P},f,\alpha) - U(\widetilde{P},f\cdot\alpha')\right| \le M\cdot\varepsilon.$$
Towards a contradiction, assume
$$\overline{\int_a^b} f\,d\alpha - \overline{\int_a^b} f\cdot\alpha'\,dx > M\cdot\varepsilon.$$
Since the upper integral is evaluated as the infimum of the corresponding upper sums, we can pick a partition $\hat{P}$ such that
$$\overline{\int_a^b} f\,d\alpha - U(\hat{P},f\cdot\alpha') > M\cdot\varepsilon.$$
Then,
$$U(P,f,\alpha) - U(\hat{P},f\cdot\alpha') \ge \overline{\int_a^b} f\,d\alpha - U(\hat{P},f\cdot\alpha') > M\cdot\varepsilon.$$
Now, $P' = P\cup\hat{P}$ is a refinement of P, and also
$$U(P',f,\alpha) - U(\hat{P},f\cdot\alpha') \ge \overline{\int_a^b} f\,d\alpha - U(\hat{P},f\cdot\alpha') > M\cdot\varepsilon.$$
Since $U(P',f\cdot\alpha') \le U(\hat{P},f\cdot\alpha')$, so
$$U(P',f,\alpha) - U(P,f\cdot\alpha') \ge U(P',f,\alpha) - U(\hat{P},f\cdot\alpha') > M\cdot\varepsilon,$$
which contradicts to the conclusion that
$$\left|U(\widetilde{P},f,\alpha) - U(\widetilde{P},f,\cdot\alpha')\right| \le M\cdot\varepsilon$$
if $\widetilde{P}$ is a refinement of P. This means that we can make the difference between $\overline{\int_a^b} f\,d\alpha$ and $\overline{\int_a^b} f\cdot\alpha'\,dx$ arbitrarily small.

Since ε is arbitrary, so we conclude that
$$\overline{\int_a^b} f\,d\alpha = \overline{\int_a^b} f\cdot\alpha'\,dx.$$
An analogous statement gives
$$\underline{\int_a^b} f\,d\alpha = \underline{\int_a^b} f\cdot\alpha'\,dx.$$
Therefore, we conclude that f is *Riemann-Stieltjes integrable* if and only if $f\cdot\alpha'$ is *Riemann integrable*, and further
$$\int_a^b f\,d\alpha = \int_a^b f(x)\cdot\alpha'(x)\,dx.$$
$\square$

The two proceding theorems set up the generality and flexibility which are inherent in the Riemann-Stieltjes integration.

(1) If $\alpha = \displaystyle\sum_{n=1}^{\infty} c_n \cdot I(x - s_n)$ and f is continuous, then the integral reduces to a finite or infitite series as

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

(2) If α is differentiable and $f \in \mathscr{R}(\alpha)$, then the integral reduces to an ordinary Riemann integral as

$$\int_a^b f \, d\alpha = \int_a^b f \cdot \alpha'(x) \, dx.$$

This brings us to one of the most important tool for actually *computing* integrals.

Theorem 2.25 (The Change of Variable Formula)**.** *Let $\varphi : [c, d] \to [a, b]$ be a strictly increasing, continuous, and surjective function, in particular, φ is injective. Let $\alpha : [a, b] \to \mathbb{R}$ be a monotone increasing function and $f \in \mathscr{R}(\alpha)$ on $[a, b]$. Define $\beta, g : [c, d] \to \mathbb{R}$ by*

$$\beta(y) = \alpha(\varphi(y)) \text{ and } g(y) = f(\varphi(y)).$$

Then, $g \in \mathscr{R}(\beta)$ and

$$\int_a^b d\alpha = \int_c^d g \, d\beta.$$

The relation of these functions is presented in the following diagram.

$$\begin{tikzcd}
 & & \mathbb{R} \\
{[c,d]} \arrow[urr, "g = f \circ \varphi"] \arrow[r, "\varphi"] \arrow[drr, "\beta = \alpha \circ \varphi"'] & {[a,b]} \arrow[ur, "f"'] \arrow[dr, "\alpha"] & \\
 & & \mathbb{R}
\end{tikzcd}$$

Lecture 12
Monday
February 1

Proof. Since φ is strictly increasing and surjective, so there exists an inverse map $\varphi^{-1} : [a, b] \to [c, d]$. This gives a bijection between partition $P = \{x_0, x_1, \ldots, x_n\}$ of $[a, b]$ and partition $\widetilde{P} = \{y_0, y_1, \ldots, y_n\}$ of $[c, d]$ where $c = \varphi^{-1}(x_0) = y_0, y_1 = \varphi^{-1}(x_1), \ldots, y_n = \varphi^{-1}(x_n) = d$. So, the correspondence is $t \in P$ if and only if $\varphi^{-1}(t) \in \widetilde{P}$. Also, $x_i > x_{i-1}$ if and only if $y_i > y_{i-1}$. Then, values taken by f on $[x_{i-1}, x_i]$ with α are equivalent to values taken by g on $[y_{i-1}, y_i]$ with β. Therefore, we have

$$U(\widetilde{P}, g, \beta) = U(P, f, \alpha) \text{ and } L(\widetilde{P}, g, \beta) = L(P, f, \alpha).$$

So, finding a partition P so that $U(P, f, \alpha) - L(P, f, \alpha) < \varepsilon$ is equivalent to finding a partition $\widetilde{P} = \varphi^{-1}(P)$ for which $U(\widetilde{P}, g, \beta) - L(\widetilde{P}, g, \beta) < \varepsilon$, and thus $g \in \mathscr{R}(\beta)$. Since $P \to \widetilde{P}$ is a bijection of partitions, so

$$\int_a^b f \, d\alpha = \inf_P U(P, f, \alpha) = \inf_{\widetilde{P}} U(\widetilde{P}, g, \beta) = \int_c^d g \, d\beta.$$

$\square$

This theorem may not look like the change of variable formula we learn from calculus, but it is actually the generalization of it to the *Riemann-Stieltjes integral.* The following corollary restores the theorem that we recall exactly, with some special conditions.

Corollary 2.26. *If $\alpha(x) = x$, $\beta = \varphi$, and φ is differentiable with $\varphi' \in \mathscr{R}$ on $[c, d]$, then*

$$\int_a^b f(x)\, dx = \int_c^d f(\varphi(y))\, df(y) = \int_c^d f(\varphi(y)) \cdot \varphi'(y)\, dy.$$

2.3 Integration and Differentiation

We now arrive the central result of calculus, the *Fundamental Theorem of Calculus*: the Riemann integral is (more or less) the inverse of the derivative.

Theorem 2.27 (Fundamental Theorem of Calculus I)**.** *Let $f \in \mathscr{R}$ on $[a, b]$. For $a \le x \le b$, define*

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

Then F is continuous on $[a, b]$. Further, if f is continuous at $x_0 \in [a, b]$, then F is differentiable at x_0 and

$$F'(x_0) = f(x_0).$$

Proof. Since $f \in \mathscr{R}$ is bounded on $[a, b]$, assume $|f(x)| \le M$ for $x \in [a, b]$. Then, for $a \le x < y \le b$,

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

So, F is M-Lipschitz.

Fix $\varepsilon > 0$, if $|y - x| < \dfrac{\varepsilon}{M}$, then $|F(x) - F(y)| < \varepsilon$, so F is uniformly continuous. In particular, if f is continuous at $x_0 \in [a, b]$, then there is $\delta > 0$ such that $|f(t) - f(x_0)| < \varepsilon$ if $|t - x_0| < \delta$ for $a \le t \le b$. If $x_0 - \delta < s \le x_0 \le t < x_0 + \delta$ and $a \le s < t \le b$, then

$$\left| \frac{F(t) - F(s)}{t - s} - f(x_0) \right| = \left| \frac{1}{t - s} \int_s^t f(x)\, dx - \frac{1}{t - s} \int_s^t f(x_0)\, dx \right| \le \frac{1}{t - s} \int_s^t |f(x) - f(x_0)|\, dx < \varepsilon.$$

Since ε is arbitrary, we conclude

$$F'(x_0) = \lim_{t, s \to x_0} \frac{F(t) - F(s)}{t - s} = f(x_0).$$

□

Theorem 2.28 (Fundamental Theorem of Calculus II)**.** *Let $f \in \mathscr{R}$ on $[a, b]$. If there is a differentiable function $F : [a, b] \to \mathbb{R}$ such that $F'(x) = f(x)$ on $[a, b]$, then*

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

Proof. Fix $\varepsilon > 0$. Since $f \in \mathscr{R}$ on $[a, b]$, so there is a partition $P = \{x_0, x_1, \ldots, x_n\}$ of $[a, b]$ such that

$$U(P, f) - L(P, f) < \varepsilon.$$

By Mean Value Theorem, for $1 \le i \le n$, there exists $t_i \in [x_{i-1}, x_i]$ such that

$$F(x_i) - F(x_{i-1}) = F'(t_i) \cdot \Delta x_i = f(t_i) \cdot \Delta t_i,$$

so

$$\sum_{i=1}^{n} f(t_i) \cdot \Delta x_i = \sum_{i=1}^{n} (F(x_i) - F(x_{i-1})) = F(b) - F(a).$$

By combining

$$L(P, f) \leq \sum_{i=1}^{n} f(t_i) \cdot \Delta x_i \leq U(P, f) \text{ and } L(P, f) \leq \int_a^b f\, dx \leq U(P, f),$$

we get

$$\left| \int_a^b f\, dx - \sum_{i=1}^{n} f(t_i) \cdot \Delta x_i \right| \leq U(P, f) - L(P, f) < \varepsilon.$$

By sending ε to 0, we conclude

$$\int_a^b f\, dx = \sum_{i=1}^{n} f(t_i) \cdot \Delta x_i = F(b) - F(a).$$

□

Let us now use the *Fundamental Theorem of Calculus* to turn the product rule into a powerful computational and theoretical tool for Riemann integration.

Theorem 2.29 (Integration by Parts). *Let $f, g : [a, b] \to \mathbb{R}$ be differentiable functions with $f', g' \in \mathscr{R}$. Then, $f \cdot g', f'g \in \mathscr{R}$, and*

$$\int_a^b f(t) \cdot g(t)\, dt = f(b) \cdot g(b) - f(a) \cdot g(a) - \int_a^b f'(t) \cdot g(t)\, dt.$$

Proof. The chain rule gives

$$(f \cdot g)' = f \cdot g' + f' \cdot g.$$

By assumption, $f', g' \in \mathscr{R}$, and so are f adn g since they are differentiable (hence continuous). Thus, $f \cdot g' \in \mathscr{R}$ and $f' \cdot g \in \mathscr{R}$. By the *Fundamental Theorem of Calculus*, we have

$$f(b) \cdot g(b) - f(a) \cdot g(a) = \int_a^b (f \cdot g)'\, dt = \int_a^b f(t) \cdot g'(t)\, dt + \int_a^b f'(t) \cdot g(t)\, dt.$$

□

2.4 Vector-Valued Functions and Rectifiable Curves

To conclude *Riemann-Stieltjes integral*, we consider an application to curves. Before introducing the rectifiable curves, we first need to briefly extend the integral to curves (more general, vector-valued functions of a real variable).

Definition 2.30 (Integral of Vector-Valued Functions). Let $f_1, \ldots, f_k : [a, b] \to \mathbb{R}$ and let

$$f = (f_1, \ldots, f_k) : [a, b] \to \mathbb{R}^k$$

for $k \in \mathbb{N}$. Let α be monotonically increasing on $[a, b]$. We say $f \in \mathscr{R}(\alpha)$ if $f_j \in \mathscr{R}(\alpha)$ for $j = 1, \ldots, k$.

If this is the case, we define

$$\int_a^b f\, d\alpha = \left(\int_a^b f_1\, d\alpha, \ldots, \int_a^b f_k\, d\alpha \right) \in \mathbb{R}^k.$$

As with derivatives, most of the theorems about integrals of scalar-valued functions extends immediately to vector-valued functions, so long as it applies separately to the components.

Theorem 2.31 (Fundamental Theorem of Calculus II in Vector-Valued Functions)**.** *If $f, F : [a, b] \to \mathbb{R}^k$ such that $f \in \mathscr{R}$ on $[a, b]$ and $F' = f$, then*

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

Proof. We deal with each component of the integrals of the vector-valued function.

$$\begin{aligned}\int_a^b f\,dx &= \left(\int_a^b f_1\,dx, \dots, \int_a^b f_k\,dx\right)\\ &= (F_1(b) - F_1(a), \dots, F_k(b) - F_k(a))\\ &= (F_1(b), \dots, F_k(b)) - (F_1(a), \dots, F_k(a))\\ &= F(b) - F(a).\end{aligned}$$

$\square$ Lecture 13 Wednesday February 3

One result about integrals that does not obviously carry over to the vector-valued case is Theorem 2.20: if $|f| \in \mathscr{R}(\alpha)$ on $[a, b]$, then

$$\left|\int_a^b f\,d\alpha\right| \le \int_a^b |f|\,d\alpha.$$

If we try to apply this componentwise, we will get

$$\left|\int_a^b f\,d\alpha\right| = \sqrt{\sum_{i=1}^n \left|\int_a^b f_i\,d\alpha\right|^2} \le \sqrt{\sum_{i=1}^n \left(\int_a^b |f_i|\,d\alpha\right)^2},$$

but this is not related in any clear way to

$$\int_a^b |f|\,d\alpha = \int_a^b \sqrt{\sum_{i=1}^n |f_i|^2}\,d\alpha.$$

Nonetheless, these two are comparable in precisely the same manner.

Proposition 2.32. *If $f : [a, b] \to \mathbb{R}^k$, $f \in \mathscr{R}(\alpha)$ on $[a, b]$, then $|f| \in \mathscr{R}(\alpha)$ (defined as $|f| = \sqrt{f_1^2 + \cdots + f_k^2}$) and*

$$\left|\int_a^b f\,d\alpha\right| \le \int_a^b |f|\,d\alpha.$$

Proof. Since $f \in \mathscr{R}(\alpha)$ on $[a, b]$ if and only if $f_i \in \mathscr{R}(\alpha)$ on $[a, b]$ for $i = 1, \dots, k$, so $f_i^2 \in \mathscr{R}(\alpha)$ on $[a, b]$ for $i = 1, \dots, k$. By linearity, $f_1^2 + \cdots + f_k^2 \in \mathscr{R}(\alpha)$ on $[a, b]$. Since $\phi(x) = \sqrt{x}$ is continuous on $[0, +\infty)$, so $|f| = \sqrt{f_1^2 + \cdots + f_k^2} \in \mathscr{R}(\alpha)$ on $[a, b]$.

Let $y_i = \displaystyle\int_a^b f_i\,d\alpha$ and $y = \displaystyle\int_a^b f\,d\alpha$. Then,

$$|y|^2 = \sum_{i=1}^k y_i^2 = \sum_{i=1}^k y_i \cdot \int_a^b f_i\,d\alpha = \sum_{i=1}^k \int_a^b y_i \cdot f_i\,d\alpha = \int_a^b \sum_{i=1}^k y_i \cdot f_i\,d\alpha = \int_a^b y \cdot f\,d\alpha.$$

By Cauchy-Schwarz inquality,

$$\int_a^b y \cdot f \, d\alpha \leq \int_a^b |y| \cdot |f| \, d\alpha = |y| \cdot \int_a^b |f| \, d\alpha,$$

so we conclude that

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

$\square$

Now, we formally start a topic of geometric interest which provides an application of some of the preceding theory.

Definition 2.33 (Curve)**.** A continuous map $\gamma : [a, b] \to \mathbb{R}^k$ is called a *curve* in $\mathbb{R}^k$ on $[a, b]$. If γ is one-to-one, then it is called an *arc.* If $\gamma(a) = \gamma(b)$, then it is called a *closed curve.*

We define a curve to be a mapping, instead of a point set, so different curves might have the same image.

Example 2.34. The curve $\gamma(t) = (\cos t, \sin t) : [0, 2\pi] \to \mathbb{R}^2$ describes a unit circle. Actually, the curve $\gamma'(t) = (\cos 2t, \sin 2t) : [0, \pi] \to \mathbb{R}^2$ also gives a unit circle as a result of the map.

Our present goal is to measure the *length* of a curve. To that target, we begin by approximating the curve by pieces of line segments. Let $P = \{x_0, x_1, \ldots, x_n\}$ be a partition of $[a, b]$, and $\gamma : [a, b] \to \mathbb{R}^k$ be a curve. We replace γ by the path which passes through the points $\gamma(x_0), \gamma(x_1), \ldots, \gamma(x_n)$, and is a straight line segment between successive points. The length of each line segment is just the Euclidean length of the vector $\gamma(x_i) - \gamma(x_{i-1})$. We define

$$\Lambda(P, \gamma) = \sum_{i=1}^{n} |\gamma(x_i) - \gamma(x_{i-1})| \, .$$

Definition 2.35 (Rectifiable Curve)**.** Let $\gamma : [a, b] \to \mathbb{R}^k$ be a curve. The *length* of γ is

$$\Lambda(\gamma) = \sup\{\Lambda(P, \gamma) : P \text{ is a partition of } [a, b]\}.$$

If $\Lambda(\gamma)$ is finite, then we say γ is a *rectifiable curve.*

The $\sup \Lambda(P, \gamma)$ may well be infinite, which means the length of a curve is infinite. In this case, the graph might have vertical asymptote, which the curve is infinite and never approaches to the end. However, we cannot parametrize this curve to a *closed interval.*

Example 2.36. Consider the curve $\gamma : [0, 1] \to \mathbb{R}^2$ which traces out the graph of the function

$$f(x) = \begin{cases} x \cdot \sin \dfrac{1}{x} & x \neq 0 \\ 0 & x = 0. \end{cases}$$

Since f is continuous on $[0, 1]$, it follows that γ is also continuous on $[0, 1]$, and at least γ is a curve. For each n, consider the partition

$$P = \{1 = x_0 > x_1 > \cdots > x_n = 0\}$$

where for $0 < i < n$, $t_i = \dfrac{1}{(n-1)\pi + \pi/2}$. So, we have

$$\frac{1}{x_1} = \frac{\pi}{2}, \frac{1}{x_2} = \frac{3}{2}\pi, \ldots, \frac{1}{x_{n-1}} = \frac{2n-3}{2}\pi.$$

At all of these points, $\sin\dfrac{1}{x_{i-1}} = \pm 1$, with the sign changing from one term to the next. Hence,

$$|\gamma(x_i) - \gamma(x_{i-1})| = \sqrt{(x_i - x_{i-1})^2 + (x_i + x_{i-1})^2} > x_i + x_{i-1} > x_i.$$

Now, $x_i = \dfrac{2}{(2i-1)\pi} > \dfrac{1}{4(i-1)}$ for $0 < i < n$, so for $2 \le i \le n-1$, we have

$$\Lambda(P,\gamma) > \sum_{i=2}^{n-1} |\gamma(x_i) - \gamma(x_{i-1})| > \frac{1}{4}\sum_{i=2}^{n-1} \frac{1}{i-1} = \frac{1}{4}\sum_{i=1}^{n-2}\frac{1}{i} = \infty$$

by Harmonic series. Therefore, $\Lambda(\gamma) = \sup \Lambda(P,\gamma) = \infty$.

In certain cases, $\Lambda(\gamma)$ is given by a Riemann integral. With the facts about integration of vector-valued functions, we can verify that the *continuously differentiable* curve is rectifiable.

Theorem 2.37. *Let γ be a continuously differentiable curve on $[a,b]$, which means γ' is continuous on $[a,b]$. Then, γ is rectifiable, and*

$$\Lambda(\gamma) = \int_a^b |\gamma'(t)|\,dt.$$

Consider the curve $\gamma(t) = (\cos t, \sin t)$ in the prior examples. Then, we have

$$\gamma'(t) = (-\sin t, \cos t) \text{ and } |\gamma'(t)| = \sqrt{\sin^2 t + \cos^2 t} = 1$$

and indeed, it is the circumference of a unit circle

$$\Lambda(\gamma) = \int_0^{2\pi} |\gamma'(t)|\,dt = \int_0^{2\pi} 1\,dt = 2\pi.$$

Proof. If $a \le x_{i-1} < x_i \le b$, then

$$|\gamma(x_i) - \gamma(x_{i-1})| = \left|\int_{x_{i-1}}^{x_i} \gamma'(t)\,dt\right| \le \int_{x_{i-1}}^{x_i} |\gamma'(t)|\,dt.$$

Then,

$$\Lambda(P,\gamma) \le \sum_{i=1}^{n} \int_{x_{i-1}}^{x_i} |\gamma'(t)|\,dt = \int_a^b |\gamma'(t)|\,dt.$$

Taking sup over P, we have

$$\Lambda(\gamma) \le \int_a^b |\gamma'(t)|\,dt.$$

To prove the other direction, fix an $\varepsilon > 0$, since $\gamma'(t)$ is continuous on $[a,b]$, so $\gamma'(t)$ is also uniformly continuosu on $[a,b]$, and there is $\delta > 0$ such that

$$|\gamma'(s) - \gamma'(t)| < \varepsilon \text{ if } |s-t| < \delta.$$

Let $P = \{x_0, x_1, \ldots, x_n\}$ be a partition of $[a,b]$ such that $\Delta x_i < \delta$ for $i = 1, \ldots, n$, then for $x_{i-1} \le t \le x_i$, $|\gamma'(t) - \gamma'(x_i)| < \varepsilon$, so

$$|\gamma'(t)| \le |\gamma'(x_i)| + \varepsilon.$$

Thus,

$$\begin{aligned}
\int_{x_{i-1}}^{x_i} |\gamma'(t)|\,dt &\le (|\gamma'(x_i)| + \varepsilon)\cdot \Delta x_i \\
&= \left|\int_{x_{i-1}}^{x_i} \gamma'(x_i)\,dx\right| + \varepsilon \cdot \Delta x_i \\
&= \left|\int_{x_{i-1}}^{x_i} (\gamma'(t) + \gamma'(x_i) - \gamma'(t))\,dx\right| + \varepsilon \cdot \Delta x_i \\
&\le \left|\int_{x_{i-1}}^{x_i} \gamma'(t)\,dx\right| + \left|\int_{x_{i-1}}^{x_i} (\gamma'(x_i) - \gamma'(t))\,dx\right| + \varepsilon \cdot \Delta x_i \\
&\le |\gamma(x_i) - \gamma(x_{i-1})| + 2\varepsilon \cdot \Delta x_i.
\end{aligned}$$

Summation over i gives

$$\int_a^b |\gamma'(t)|\,dt \le \Lambda(P,\gamma) + 2\varepsilon\cdot(b-a) \le \Lambda(\gamma) + 2\varepsilon\cdot(b-a).$$

Since ε is arbitrary, so

$$\int_a^b |\gamma'(t)|\,dt \le \Lambda(\gamma).$$

Therefore, we conclude that

$$\Lambda(\gamma) = \int_a^b |\gamma'(t)|\,dt.$$

□

3 Sequences and Series of Functions

Although we used many of the familiar functions from calculus (like cos and sin) and assumed properties of them, we have yet to formally define and develop these functions rigorously. To do so, we will investigate convergence of sequences and series of functions. The results will be extensions of numerical sequences and series.

In particular, we shall be concerned with the following problem. Let $(f_n)_{n\in\mathbb{N}}$ be a sequence of functions. If each element of (f_n) has a particular property, when does the limit function f have the same property? For example, we will ask whether the limit function f is continuous if each function f_n is continuous. We begin by defining convergence of a sequence of functions.

Lecture 14
Friday
February 5

3.1 Pointwise Convergence is Problematic

For each $n \in \mathbb{N}$, suppose f_n be a real-valued function, defined on the set E which does not vary with n. Then, we can investigate the *pointwise* limit of these functions (and the existence).

Definition 3.1 (Pointwise Convergence)**.** Let (X, d) be a metric space and $E \subseteq X$. Let $(f_n)_{n\in\mathbb{N}}$ be a sequence of real-valued functions $f_n : E \to \mathbb{R}$. Suppose that the sequence of real numbers $(f_n(x))_n$ converges for all $x \in E$. Then, we define

$$f(x) = \lim_{n\to\infty} f_n(x) \in \mathbb{R}$$

for all $x \in E$. In the case, we say that f is the *pointwise limit* of $(f_n)_n$ and that (f_n) *converges pointwise* on E to f.

Example 3.2. Let $f_n(x) = x^n$ for $0 \le x \le 1$.

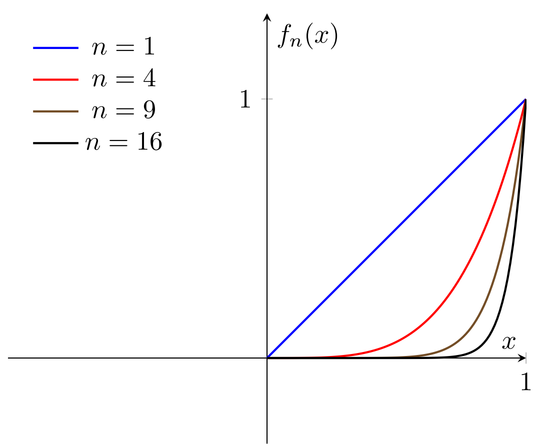


Figure 3.1: The Function $f_n(x) = x^n$ for $n = 1, 4, 9, 16$

For $0 \le x < 1$, we have

$$\lim_{n\to\infty} f_n(x) = 0.$$

For $x = 1$, then $\lim_{n\to\infty} f_n(x) = 1$. So, (f_n) converges pointwise to

$$f(x) = \begin{cases} 0 & 0 \le x < 1 \\ 1 & x = 1 \end{cases}$$

on $[0,1]$. For all $n \in \mathbb{N}$, the function $f_n(x) = x^n$ is continuous on $[0,1]$, but the pointwise limit function $f(x)$ is *not* continuous on $[0,1]$. Thus, continuity is not in general preserved under taking pointwise limits.

Does our notion of convergence preserve any property of the sequence all the way to the limit? For properties, we will discuss continuity, integrability, differentiability, summability, etc. In short, can we exchange the order in which we take limits? The answer will be *no* without further assumptions.

Recall that f is continuous at x if and only if

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

Suppose each f_n is continuous at x and f is the pointwise limit of (f_n). Is f continuous? Is it true that

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

For continuity, it is a NO. When $f_n(x) = x^n$, we have

$$\lim_{n \to \infty} \lim_{t \to 1} f_n(t) = 1,$$

but

$$\lim_{t \to 1} \lim_{n \to \infty} f_n(t) = 0.$$

It is always dangerous to change the order of limits. We say $\lim\limits_{x \to p} f(x) = q$ if for all $\varepsilon > 0$, there exists a $\delta > 0$, such that $d(x,p) < \delta$ implies $d(f(x), q) < \varepsilon$. The definition is rigorous, and contains some quantifiers. Changing the order of quantifiers can really change what we intend to mean. Therefore, the order matters a lot. Here is another straightforward example.

Example 3.3. Consider the double sequence $a_{m,n} = \dfrac{m}{m+n}$ for $m \geq 1$ and $n \geq 1$. For fixed m, we have

$$\lim_{n \to \infty} a_{m,n} = \lim_{n \to \infty} \frac{m}{m+n} = 0.$$

For fixed n, we have

$$\lim_{m \to \infty} a_{m,n} = \lim_{m \to \infty} \frac{m}{m+n} = 1.$$

Thus, we note that

$$\lim_{m \to \infty} \lim_{n \to \infty} a_{m,n} = 0$$

but

$$\lim_{n \to \infty} \lim_{m \to \infty} a_{m,n} = 1.$$

Lecture 15
Monday
February 8

3.2 Uniform Convergence

To preserve continuity, integrability, differentiability, or other properties, we need a stronger form of convergence. One of the most useful notions is that of *uniform convergence.*

Definition 3.4 (Uniform Convergence)**.** Let (X,d) be a metric space and $E \subseteq X$. Let $(f_n)_{n \in \mathbb{N}}$ be a sequence of functions defined on E. Let f be a function defined on E as well. We say that (f_n) *converges uniformly* to f on E if for all $\varepsilon > 0$, there exists an $N \in \mathbb{N}$ such that for all $n \geq N$,

$$|f_n(x) - f(x)| \leq \varepsilon$$

for all $x \in E$.

The crucial part of *uniform convergence* is that $|f_n(x) - f(x)| \le \varepsilon$ for every $x \in E$ if $n \ge N$. That is, N depends *only* on ε, but *not* on x. For *pointwise convergence*, N can depend on x and ε, that is, $N = N(\varepsilon, x)$. So, it is clear that if a sequence of functions (f_n) converges uniformly, it also converges pointwise.

Remark 3.5. Like comparing *continuity* to *uniform continuity*, *uniform convergence* differs from *pointwise convergence* by a reordering of quantifiers. For *uniform convergence*, we must be able to choose N independently of x, so we have $\forall \varepsilon > 0 \, \exists N \, \forall n \ge N \, \forall x$, as opposed to the *pointwise convergence* which beings with $\forall x \, \forall \varepsilon > 0 \, \exists N \, \forall n \ge N$.

Now, we can formally verify the convergence of a sequence of functions $(f_n)_{n\in\mathbb{N}}$. With this stronger form of convergence, we can get some nice properties.

Example 3.6 (Continuation of Example 3.2)**.** Let $f_n(x) = x^n$ for $0 \le x \le 1$. Then, (f_n) converges pointwise to

$$f(x) = \begin{cases} 0 & 0 \le x < 1 \\ 1 & x = 1. \end{cases}$$

Towards a contradiction, assume that $f_n \to f$ uniformly. Then, for $\varepsilon > 0$, there exists $N_\varepsilon \in \mathbb{N}$ such that for $n \ge N_\varepsilon$, $|f_n(x) - f(x)| \le \varepsilon$ for $x \in [0,1]$. Let $\varepsilon = \frac{1}{2}$. For all $n \ge N_{1/2}$, we have $|f_n(x) - f(x)| \le \frac{1}{2}$ for all $x \in [0,1]$. Pick $x = \left(\frac{3}{4}\right)^{1/N} \in [0,1)$, then

$$|f_N(x) - f(x)| = |f_N(x)| = \frac{3}{4} > \frac{1}{2}.$$

So, (f_n) cannot converge uniformly to f on $[0,1]$.

The continuity is not preserved in Example 3.2 since $f_n(x) = x^n$ does not converge uniformly. Does uniform convergence of a sequence of functions (f_n) guarantee the preservation of continuity? Before we arrive at this step, we first derive some criteria for uniform convergence.

Definition 3.7. Let (X, d) be a metric space and $E \subseteq X$. Let $(f_n)_{n\in\mathbb{N}}$ be a sequence of functions defined on E. We say $\sum_{n=1}^{\infty} f_n(x)$ *converges uniformly* on E if the sequence of partial sums $s_N(x) = \sum_{n=1}^{N} f_n(x)$ *converges uniformly* on E, that is, for all $\varepsilon > 0$, there exists an $N \in \mathbb{N}$ such that for all $n \ge N$, we have

$$|s_N(x) - f(x)| = \left| \sum_{n=1}^{N} f_n(x) - f(x) \right| < \varepsilon$$

for all $x \in E$.

Proposition 3.8 (Uniform Cauchy Criterion)**.** *Let $(f_n)_{n\in\mathbb{N}}$ be a sequence of real-valued functions on E. We say (f_n) converges uniformly on E if and only if for all $\varepsilon > 0$, there exists an $N \in \mathbb{N}$ such that, for all $n, m \ge N$, $|f_n(x) - f_m(x)| < \varepsilon$ for all $x \in E$.*

Proof. Assume that there exists a function f on E such that $f_n \to f$ uniformly. Pick an $N \in \mathbb{N}$ such that for all $n \ge N$ and $x \in E$, we have

$$|f_n(x) - f(x)| \le \frac{\varepsilon}{2}.$$

So, for $n, m \ge N$ and $x \in E$,

$$|f_n(x) - f_m(x)| \le |f_n(x) - f(x)| + |f(x) - f_m(x)| \le \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Hence, (f_n) is uniformly Cauchy.

Now, fix $x \in E$, then $|f_n(x) - f_m(x)| \leq \varepsilon$, so $(f_n(x))$ is a Cauchy sequence in $\mathbb{R}$. Since $\mathbb{R}$ is a complete metric space, so there is a real number $f(x)$ such that $f_n(x) \to f(x)$. Thus, we find a pointwise limit for (f_n). Fix an $\varepsilon > 0$. Pick an N such that $|f_n(x) - f_m(x)| \leq \dfrac{\varepsilon}{2}$ for all $n, m \geq N$. Taking the limit as $m \to \infty$, we obtain that if $n \geq N$ and $x \in E$, then $|f_n(x) - f(x)| \leq \varepsilon$. We conclude that $f_n \to f$ uniformly on E.

□

The following corollaries are quite easy to prove just by using the definition of uniform convergence.

Corollary 3.9. *Let $(f_n)_{n\in\mathbb{N}}$ be a sequence of real-valued functions on E and f be a function defined on E. Suppose (f_n) converges pointwise to f on E. Let*

$$M_n = \sup_{x\in E} |f_n(x) - f(x)|.$$

Then, (f_n) converges uniformly to f on E if and only if $\lim_{n\to\infty} M_n = 0$.

Example 3.10 (Continuation of Example 3.2)**.** Let $f_n(x) = x^n$ for $0 \leq x \leq 1$. Then, (f_n) converges pointwise to

$$f(x) = \begin{cases} 0 & 0 \leq x < 1 \\ 1 & x = 1. \end{cases}$$

We have $\sup_{x\in[0,1]} |f_n(x) - f(x)| = 1$ for every $n \in \mathbb{N}$, so

$$\lim_{n\to\infty} M_n = 1 \neq 0.$$

So, we get the same result as previous: (f_n) does not converge uniformly to f.

The previous example demonstrates an important and simple technique, which is the following corollary.

Corollary 3.11 (Weierstrass M-test)**.** *Let $(f_n)_{n\in\mathbb{N}}$ be a sequence of real-valued functions on E. Suppose $|f_n(x)| \leq M_n$ for all $x \in E$. Then $\sum_{n=1}^{\infty} f_n(x)$ converges uniformly if $\sum_{n=1}^{\infty} M_n$ converges.*

Proof. Let $s_n(x) = \sum_{k=1}^{n} f_k(x)$. For any $m > n$, we have

$$|s_m(x) - s_n(x)| = \left|\sum_{k=n+1}^{m} f_k(x)\right| \leq \sum_{k=n+1}^{m} |f_k(x)| \leq \sum_{k=n+1}^{m} M_k.$$

Since the sequence $S_n = \sum_{k=1}^{n} M_k$ converges, so S_n is Cauchy. For any $\varepsilon > 0$, there exists an $N \in \mathbb{N}$ such that for all $m > n \geq N$, we have $|S_m - S_n| < \varepsilon$. With $S_m - S_n = \sum_{k=n+1}^{m} M_k$, the sequence of functions (s_n) is uniformly Cauchy. Therefore, it converges uniformly.

□

3.3 Uniform Convergence and Continuity

Previously, for $f_n(x) = x^n$ on $[0, 1]$, we find that

$$\lim_{t\to 1}\lim_{n\to\infty} f_n(t) = 0 \quad \text{but} \quad \lim_{n\to\infty}\lim_{t\to 1} f_n(t) = 1.$$

So, pointwise limits of functions do note preserve limits in general. The following theorem shows that uniform convergence *does* preserve limits, and hence continuity.

Theorem 3.12. *Let (X, d) be a metric space and $E \subseteq X$. Let $(f_n)_{n\in\mathbb{N}}$ be a sequence of functions on E and f be a function on E such that $f_n \to f$ uniformly on E. Let x be a limit point of E and*

$$\lim_{t\to x} f_n(x) = a_n \in \mathbb{R}$$

for $n \in \mathbb{N}$. Then, $(a_n)_n$ converges and

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

In other words,

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

Proof. Fix an $\varepsilon > 0$. Since $f_n \to f$ uniformly on E, there exists $N \in \mathbb{N}$ such that for all $n, m \geq N$ and $t \in E$, we have

$$|f_n(t) - f_m(t)| \leq \varepsilon.$$

Let $t \to x$, we get $|a_n - a_m| \leq \varepsilon$. Then, $(a_n)_{n\in\mathbb{N}}$ is Cauchy, so (a_n) converges and set $\lim_{n\to\infty} a_n = a$. By triangle inequality,

$$|f(t) - a| \leq |f(t) - f_n(t)| + |f_n(t) - a_n| + |a_n - a|.$$

Since $f_n \to f$ uniformly and $a_n \to a$, so there exists $N \in \mathbb{N}$ such that for $n \geq N$, we have $|f_n(t) - f(t)| \leq \dfrac{\varepsilon}{3}$ for $t \in E$ and $|a_n - a| \leq \dfrac{\varepsilon}{3}$. Since $\lim_{t\to x} f_n(t) = a_n$, we can choose some $\delta > 0$ such that for all $t \in B_\delta(x) \setminus \{x\}$, we have $|f_n(t) - a_n| < \dfrac{\varepsilon}{3}$. Therefore, we note that for $0 < |t - x| < \delta$

$$|f(t) - a| \leq \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon.$$

We conclude that $\lim_{t\to x} f(t) = a$.

$\square$

Corollary 3.13. *Let (X, d) be a metric space and $E \subseteq X$. If $(f_n)_{n\in\mathbb{N}}$ is a sequence of continuous functions on E which converges uniformly to a function f on E, then f is continuous on E.*

This is immediate by letting $a_n = f_n(x)$ in Theorem 3.12.

Lecture 16
Wednesday
February 10

Definition 3.14 (Space of Continuous Functions)**.** Let (X, d) be a metric space. We define $\mathscr{C}^0(X)$ as a set of real-valued bounded continuous functions on X, that is,

$$\mathscr{C}^0(X) = \{f : f \text{ is continuous, } |f(x)| \leq M \text{ for all } x \in X \text{ for some } M\}.$$

The *supremum norm* of $f \in \mathscr{C}^0(X)$ is defined as

$$\|f\|_\infty = \sup_{x\in X} |f(x)|.$$

To check $\|f\|_\infty$ is a *norm*, we need to check (1) $\|f\|_\infty = 0$ if and only if $f = 0$, (2) $\|f\|_\infty < \infty$, and (3) $\|f+g\|_\infty \leq \|f\|_\infty + \|g\|_\infty$. So $(\mathscr{C}^0(X), d_\infty)$, where

$$d_\infty(f, g) = \|f - g\|_\infty,$$

is a metric space. With Proposition 3.8 and Theorem 3.12, we can get the *completeness* of the metric space $(\mathscr{C}^0(X), d_\infty)$.

Theorem 3.15. *$(\mathscr{C}^0(X), d_\infty)$ is a complete metric space.*

Proof. Let $(f_n)_{n\in\mathbb{N}}$ be a Cauchy sequence in the metric space $(\mathscr{C}^0(X), d_\infty)$. For $\varepsilon > 0$, there exists an $N \in \mathbb{N}$ such that

$$d_\infty(f_n, f_m) < \varepsilon$$

for all $n, m \geq N$. For every $x \in X$, we have

$$|f_n(x) - f_m(x)| \leq \sup_{x\in X} |f_n(x) - f_m(x)| < \varepsilon,$$

so $(f_n)_{n\in\mathbb{N}}$ is uniformly Cauchy, which means (f_n) converges uniformly. Let $f(x) = \lim_{n\to\infty} f_n(x)$. So, f is continuous by Theorem 3.12. Let $m \to \infty$, we have

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

for $x \in X$ and $n \geq N$. Let $n = N$, we have

$$|f(x)| \leq |f(x) - f_N(x)| + |f_N(x)| \leq \varepsilon + d_\infty(f_N, 0).$$

Therefore, f is bounded. We conclude that $(\mathscr{C}^0(X), d_\infty)$ is a complete metric space.

□

From this proof, we see that the completeness of $\mathscr{C}^0(X)$ is inherited from the completeness of $\mathbb{R}$, so the underlying space X does not play any role in this aspect.

3.4 Integration and Differentiation of Uniformly Convergent Sequences

As with continuity, in general we require uniform convergence for integrals to play nicely with limits. First, we provide an example about the problem with pointwise convegence.

Example 3.16. Let $f_n(x) = n^2 x(1-x^2)^n$ for $x \in [0, 1]$.

We note that $f_n(0) = 0$ for all $n \in \mathbb{N}$. For $0 \leq x \leq 1$, the sequence $x(1-x^2)^n$ converges to 0 exponentially fast, and therefore even multiplying by n, we still have $f_n(x) \to 0$ as $n \to \infty$. Thus, the pointwise limit of (f_n) is the constant function 0 on $[0, 1]$. Nevertheless, by substitution $u = 1 - x^2$, we compute that

$$\int_0^1 f_n(x)\, dx = n \cdot \int_0^1 x(1-x^2)^n\, dx = n \cdot \int_0^1 \frac{1}{2} u^n\, du = \frac{n}{2(n+1)}.$$

Thus,

$$\lim_{n\to\infty} \int_0^1 f_n(x)\, dx = \frac{1}{2}$$

but

$$\int_0^1 \lim_{n\to\infty} f_n(x)\, dx = 0.$$

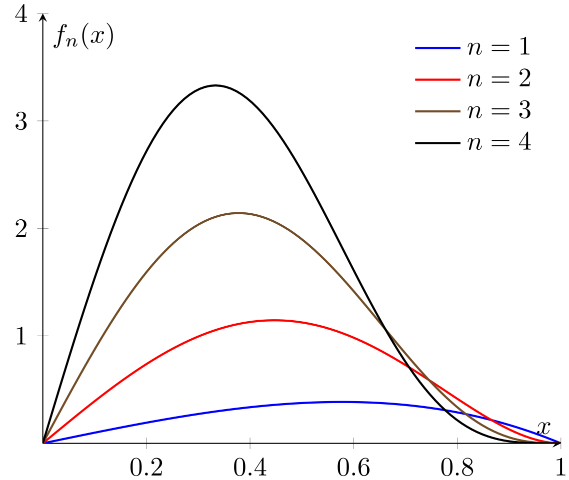


Figure 3.2: The Function $f_n(x) = n^2 x(1 - x^2)^n$ for $n = 1, 2, 3, 4$

So although the functions f_n are all integrable, as is their pointwise limit function, the limit of the integrals is not the integral of the limit.

Theorem 3.17. *Let $\alpha : [a, b] \to \mathbb{R}$ be a monotone increasing function. Suppose $f_n \in \mathscr{R}(\alpha)$ on $[a, b]$ for each $n \in \mathbb{N}$. If $f_n \to f$ uniformly on $[a, b]$, then $f \in \mathscr{R}(\alpha)$, and*

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

Proof. Let $\varepsilon_n = \sup_{x\in[a,b]} |f_n(x) - f(x)|$. Then, for $x \in [a, b]$, we have

$$f_n - \varepsilon_n \le f \le f_n + \varepsilon_n,$$

so

$$\int_a^b (f_n - \varepsilon_n) \, d\alpha \le \underline{\int_a^b} f \, d\alpha \le \overline{\int_a^b} f \, d\alpha \le \int_a^b (f_n + \varepsilon_n) \, d\alpha.$$

Then,

$$0 \le \overline{\int_a^b} f \, d\alpha - \underline{\int_a^b} f \, d\alpha \le \int_a^b (f_n + \varepsilon_n) \, d\alpha - \int_a^b (f_n - \varepsilon_n) \, d\alpha \le 2\varepsilon_n(\alpha(b) - \alpha(a)).$$

Since $\varepsilon_n \to 0$ as $n \to \infty$, so

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

thus $f \in \mathscr{R}(\alpha)$. Further,

$$\left| \int_a^b f \, d\alpha - \int_a^b f_n \, d\alpha \right| \le \varepsilon_n(\alpha(b) - \alpha(a)) \to 0$$

as $n \to \infty$.

$\square$

This theorem further means that the metric distance $d(f_n, f) \to 0$, provided that $f_n \to f$ uniformly, that is, uniform convergence is a stronger notion.

Corollary 3.18. *Let $(f_n)_{n\in\mathbb{N}}$ be a sequence of functions defined on $[a,b]$. If $f_n \in \mathscr{R}(\alpha)$ on $[a,b]$ and for $a \le x \le b$,*

$$f(x) = \sum_{n=1}^{\infty} f_n(x),$$

the series converges uniformly on $[a,b]$, then $f \in \mathscr{R}(\alpha)$ and

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

In other words, the series may be integrated term by term.

Proof. Let $s_n(x) = \displaystyle\sum_{k=1}^{n} f_k(x)$. Then, $s_n \to f$ uniformly on $[a,b]$, so $s_n(x) \in \mathscr{R}(\alpha)$ for $N \in \mathbb{N}$ on $[a,b]$. By Theorem 3.17, $f \in \mathscr{R}(\alpha)$, and

$$\lim_{n\to\infty} \int_a^b s_n(x)\,d\alpha = \int_a^b f\,d\alpha,$$

so

$$\lim_{n\to\infty} \int_a^b s_n\,d\alpha = \lim_{n\to\infty} \int_a^b \left(\sum_{k=1}^{n} f_k\right) d\alpha = \lim_{n\to\infty} \sum_{k=1}^{n} \int_a^b f_i\,d\alpha = \sum_{n=1}^{\infty} \int_a^b f_n\,d\alpha.$$

□

Now, we shift our focus to uniform convergence and differentiation. Let's start with an example about the problem of pointwise convergence.

Example 3.19. Assume that sin and cos are differentiable functions with $\sin' x = \cos x$. Let

$$f_n(x) = \frac{\sin nx}{\sqrt{n}}$$

for $x \in \mathbb{R}$ and $n \in \mathbb{N}$.

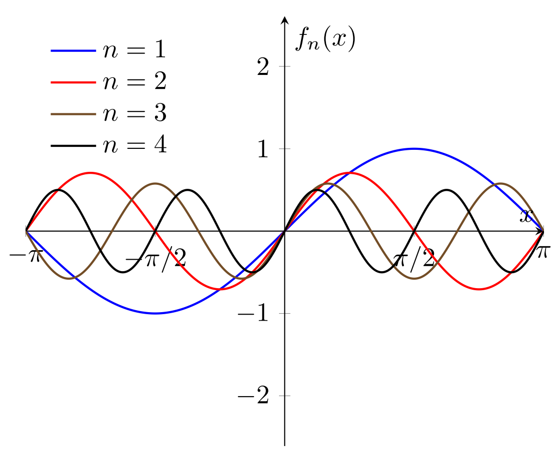


Figure 3.3: The Function $f_n(x) = \dfrac{\sin nx}{\sqrt{n}}$ for $n = 1, 2, 3, 4$

Since $|\sin nx| \le 1$, so $\lim_{n\to\infty} f_n(x) = 0$ for all x. Also, we note that

$$f_n'(x) = n \cdot \frac{\cos nx}{\sqrt{n}} = \sqrt{n}\cos nx.$$

As $n \to \infty$, we have $f_n' \to \infty$. So, while f_n is differentiable and $f_n \to 0$ pointwise, the derivatives of f_n do not converge pointwise.

Further, we note that $f_n(x) = \dfrac{\sin nx}{\sqrt{n}}$ converges *uniformly* to 0. Since

$$|f_n(x)| \le \frac{1}{\sqrt{n}},$$

so we can make $|f_n(x)| < \varepsilon$ for all $x \in \mathbb{R}$ by picking $N > \dfrac{1}{\varepsilon^2}$. Unfortunately, the uniform convergence of (f_n) implies nothing about the sequence (f_n'). Thus stronger hypotheses are required for this assertion that $f_n' \to f'$ if $f_n \to f$.

Theorem 3.20. *Let $(f_n)_{n\in\mathbb{N}}$ be a sequence of differentiable functions on $[a,b]$. Suppose there exists at least one point $x_0 \in [a,b]$ such that $(f_n(x_0))_{n\in\mathbb{N}}$ converges. If $(f_n')_{n\in\mathbb{N}}$ converges uniformly on $[a,b]$, then $f_n \to f$ uniformly on $[a,b]$, and f is differentiable on $[a,b]$, and*

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

Proof. Fix $\varepsilon > 0$. Since $(f_n(x_0))_{n\in\mathbb{N}}$ converges and $(f_n')_{n\in\mathbb{N}}$ converges uniformly, so $(f_n(x_0))$ is Cauchy and (f_n') is uniformly Cauchy, that is, there exists $N \in \mathbb{N}$ such that for all $n, m \ge N$, we have

$$|f_n(x_0) - f_m(x_0)| \le \frac{\varepsilon}{2}$$

and for every $x \in [a,b]$,

$$|f_n'(x) - f_m'(x)| < \frac{\varepsilon}{2(b-a)}.$$

Let $g(x) = f_n(x) - f_m(x)$ that is differentiable on $[a,b]$. By the Mean Value Theorem,

$$\begin{aligned}
|f_n(x) - f_m(x) - f_n(t) + f_m(t)| &= |g(x) - g(t)| \\
&\le \sup_{[a,b]} |g'(s)| \cdot |x-t| \\
&\le \sup_{[a,b]} |f_n'(s) - f_m'(s)| \cdot |x-t| \\
&\le |x-t| \cdot \frac{\varepsilon}{2(b-a)} \\
&\le \frac{\varepsilon}{2}
\end{aligned}$$

for $x, t \in [a,b]$ and $n, m \ge N$. Therefore,

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

So, there exists $N \in \mathbb{N}$ such that for all $n, m \ge N$ and $x \in E$,

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

Thus, the sequence (f_n) is uniformly Cauchy, so it converges uniformly on $[a,b]$.

Now, we define

$$\phi_n(t) = \frac{f_n(t) - f_n(x)}{t - x} \text{ and } \phi(t) = \frac{f(t) - f(x)}{t - x}$$

for $a \le t \le b$ and $t \ne x$. By definition of derivative,

$$\lim_{t\to x} \phi_n(t) = f_n'(x).$$

For $n, m \geq N$, $t \in [a,b]$, and $t \neq x$, then

$$|\phi_n(t) - \phi_m(t)| \leq \frac{|g(t) - g(x)|}{|t-x|} \leq \sup_{[a,b]} |g'(s)| \leq |f_n'(s) - f_m'(s)| < \frac{\varepsilon}{2(b-a)},$$

so $(\phi_n)_n$ is uniformly Cauchy on $[a,b]$, so (ϕ_n) converges uniformly on $[a,b]$. If $t \neq x$, then

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

pointwise, so

$$\lim_{t \to x} \phi_n(t) = f_n'(x).$$

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

$\square$

Corollary 3.21. *Let $(f_n)_{n \in \mathbb{N}}$ be a sequence of continuously differentiable functions on $[a,b]$. Suppose there exists at least one point $x_0 \in [a,b]$ such that $(f_n(x_0))_{n \in \mathbb{N}}$ converges. If $(f_n')_{n \in \mathbb{N}}$ converges uniformly on $[a,b]$, then $f_n \to f$ uniformly on $[a,b]$, and f is continuously differentiable on $[a,b]$, and*

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

This is a weaker conclusion of Theorem 3.20. The proof of this corollary is based on Theorem 3.17 and the Fundamental Theorem of Calculus.

Lecture 17
Friday
February 12

Theorem 3.22. *Let $\varphi : \mathbb{R} \to \mathbb{R}$ be a continuous function defined by $\varphi(x) = |x|$ for $-1 \leq x \leq 1$, and extended periodically to satisfy $\varphi(x+2) = \varphi(x)$ for all $x \in \mathbb{R}$. Then, we define*

$$f(x) = \sum_{n=0}^{\infty} \left(\frac{3}{4}\right)^n \varphi(4^n x).$$

The function f is continuous on $\mathbb{R}$, but nowhere differentiable.

Proof. Note that

$$|\varphi(s) - \varphi(t)| < |s - t|$$

for all $s, t \in \mathbb{R}$, so φ is 1-Lipchitz, thus φ is continuous on $\mathbb{R}$. Let $f_n(x) = \left(\frac{3}{4}\right)^n \varphi(4^n \cdot x)$. With $\varphi(4^n \cdot x) \leq 1$, we have

$$|f_n(x)| \leq \left(\frac{3}{4}\right)^n |\varphi(4^n \cdot x)| \leq \left(\frac{3}{4}\right)^n$$

for $x \in \mathbb{R}$. Since $\sum_{n=0}^{\infty} \left(\frac{3}{4}\right)^n$ converges, so $f(x)$ converges uniformly on $\mathbb{R}$, and hence f is continuous.

Now, we want to show that f is differentiable at x, for any given $x \in \mathbb{R}$. To do so, we will construct a sequence $\delta_m \to 0$ such that

$$\lim_{m \to \infty} \left| \frac{f(x + \delta_m) - f(x)}{\delta_m} \right| = \infty.$$

For any $\delta \in \mathbb{R}$, we have

$$\left| \frac{f(x+\delta) - f(x)}{\delta} \right| = \left| \sum_{n=0}^{\infty} \left(\frac{3}{4}\right)^n \cdot \frac{\varphi(4^n(x+\delta)) - \varphi(4^n x)}{\delta} \right|.$$

Define

$$\gamma_n(\delta) = \frac{\varphi(4^n(x+\delta)) - \varphi(4^n x)}{\delta}.$$

If $4^n \cdot \delta$ is an even integer, then since φ is 2-periodic, we will have $\gamma_n(\delta) = 0$. Let $x \in \mathbb{R}$ and $m \in \mathbb{N}$, set

$$\delta_m = \pm\frac{1}{2} \cdot \frac{1}{4^m}$$

where $\delta_m > 0$ if $\mathbb{Z} \cap (4^m x, 4^m(x + \delta_m)) = \varnothing$ and $\delta_m < 0$ if $\mathbb{Z} \cap (4^m(x - \delta_m), 4^m x) = \varnothing$. We can always make this choice since $|4^m \delta_m| = \dfrac{1}{2}$. Then, $|4^n \cdot \delta_m| = 2^{2(n-m)-1}$, so $4^n \cdot \delta_m$ is an even integer for all $n > m$, and so $\gamma_n(\delta_m) = 0$ for $n > m$. This means that

$$\left|\frac{f(x+\delta_m) - f(x)}{\delta_m}\right| = \left|\sum_{n=0}^{\infty} \left(\frac{3}{4}\right)^n \gamma_n(\delta_m)\right| = \left|\sum_{n=0}^{m} \left(\frac{3}{4}\right)^n \gamma_n(\delta_m)\right|.$$

By reverse triangle inequality, we get

$$\left|\frac{f(x+\delta_m) - f(x)}{\delta_m}\right| \geq \left(\frac{3}{4}\right)^n |\gamma_m(\delta_m)| - \sum_{n=0}^{m-1} \left(\frac{3}{4}\right)^n |\gamma_n(\delta_m)|.$$

For $0 \leq n < m$, we have

$$|\gamma_n(\delta_m)| \leq \left|\frac{4^n(x+\delta_m) - 4^n x}{\delta_m}\right| \leq 4^n.$$

For $n = m$, we have

$$|\gamma_m(\delta_m)| = \frac{|\varphi(4^m x \pm 1/2) - \varphi(4^m x)|}{|\delta_m|}.$$

Since the interval $\left[4^m x - \dfrac{1}{2}, 4^m x + \dfrac{1}{2}\right]$ has length 1, it contains at most one integer in its *interior*. By the definition of δ_m, in each case, either there is no integer in $\left(4^m x - \dfrac{1}{2}, 4^m x\right)$ or $\left(4^m x, 4^m x + \dfrac{1}{2}\right)$, so we have

$$|\gamma_m(\delta_m)| = \frac{|\varphi(4^m x \pm 1/2) - \varphi(4^m x)|}{|\delta_m|} = \left|\frac{\pm 1/2}{\delta_m}\right| = \left|\frac{1}{2} \cdot 2 \cdot 4^m\right| = 4^m$$

since φ is linear with slope ± 1 on any interval between two integers. Thus,

$$\left|\frac{f(x+\delta_m) - f(x)}{\delta_m}\right| \geq \left(\frac{3}{4}\right)^m \cdot 4^m - \sum_{n=0}^{m-1} \left(\frac{3}{4}\right)^n \cdot 4^n = \frac{3^m + 1}{2}.$$

Therefore, there exists a sequence $\delta \to 0$ such that

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

We conclude that $f'(x)$ does not exist.

□

3.5 Equicontinuous Families of Functions

Let (X, d) be a metric space. We define $\mathscr{C}^0(X)$ to be the set of continuous and bounded real-valued functions defined on X. With the *supremum metric*

$$d_\infty(f, g) = \sup_{x \in X} |f(x) - g(x)|,$$

$(\mathscr{C}^0(X), d_\infty)$ is a *complete* metric space. Thus, $(f_n)_{n \in \mathbb{N}}$ is a Cauchy sequence in $(\mathscr{C}^0(X), d_\infty)$ and $f_n \to f \in \mathscr{C}^0(X)$ uniformly.

Every sequence in a compact metric space admits a convergent subsequence. What are the compact subsets of $(\mathscr{C}^0(X), d_\infty)$? Which sequences admit convergent subsequences? Is there an analog for sequences of functions?

A sequence of functions $(f_n)_n$ is bounded in $(\mathscr{C}^0(X), d_\infty)$ if and only if there exists $M > 0$ such that

$$d_\infty(f_n, 0) = \sup_{x \in X} |f_n(x)| \leq M$$

for any $n \in N$. This happens if and only if there exists $M > 0$ such that $|f_n(x)| \leq M$ for all $n \in \mathbb{N}$ and $x \in X$.

Definition 3.23 (Uniformly Bounded)**.** Let $(f_n)_{n\in\mathbb{N}}$ be a sequence of functions defined on a set X. We say that (f_n) is *uniformly bounded* if there exists an $M \in \mathbb{R}$ such that

$$|f_n(x)| \leq M$$

for all $x \in X$, for all $n \in \mathbb{N}$.

Definition 3.24 (Pointwise Bounded)**.** Let $(f_n)_{n\in\mathbb{N}}$ be a sequence of functions defined on a set X. We say that (f_n) is *pointwise bounded* if for all $x \in X$, there exists $\phi(x) \in \mathbb{R}$ such that

$$|f_n(x)| \leq \phi(x)$$

for $n \in \mathbb{N}$ (that is, the sequence of real numbers $(f_n(x))_n$ is bounded).

In particular, $(f_n)_{n\in\mathbb{N}}$ is uniformly bounded, then (f_n) has a uniformly convergent subsequence. In $\mathbb{R}$, this is always true.

Example 3.25. Let $f_n(x) = \sin nx$ for $0 \leq x \leq 2\pi$. Assume we know properties of sin, then $f_n \in \mathscr{C}^0([0, 2\pi])$ and $|f_n(x)| \leq 1$ for all $n \in \mathbb{N}$ and all $x \in [0, 2\pi]$, but $(f_n)_n$ has no uniformly convergent subsequence.

Towards a contradiction, (f_n) has a uniformly convergent subsequence $(f_{n_k})_k$, then

$$\lim_{k\to\infty} (\sin n_k x - \sin n_{k+1} x) = 0$$

uniformly. Also,

$$\lim_{k\to\infty} (\sin n_k x - \sin n_{k+1} x)^2 = 0$$

uniformly. So, (by Lebesgue's Theorem concerning integration of boundedly convergent sequences in 140C)

$$\lim_{k\to\infty} \int_0^{2\pi} (\sin n_k x - \sin n_{k+1} x)^2 \, dx = 0$$

but

$$\int_0^{2\pi} (\sin n_k x - \sin n_{k+1} x)^2 \, dx = 2\pi.$$

Therefore, $(\sin nx)_n$ does not admit uniformly convergent subsequence.

Recall the Heine-Borel Theorem from 140A:

Theorem 3.26 (Heine-Borel)**.** *Let $E \subseteq \mathbb{R}^k$. The following two statements are equivalent:*

(a) E is closed and bounded.

(b) E is compact.

In short, boundedness and closedness imply compactness in $\mathbb{R}$. However, in $(\mathscr{C}^0(X), d_\infty)$, this implication is not true. We need an additional property.

Definition 3.27 (Equicontinuity)**.** Let $\mathscr{F} \subseteq \mathscr{C}^0(X)$ be a collection of functions. Let (X, d) be a metric space and $E \subseteq X$. We say $\mathscr{F}$ is *equicontinuous* on E if for every $\varepsilon > 0$, there exists a $\delta > 0$ such that for all $x, y \in E$ and all $f \in \mathscr{F}$,

$$d(x, y) < \delta \Longrightarrow |f(x) - f(y)| < \varepsilon.$$

We upgrade the definition of *continuity* of a function to *uniform continuity* by insisting that δ depends only on ε without considering the point x. The new notion of *equicontinuity* says that, not only that, we must able to pick the same δ for all functions $f \in \mathscr{F}$ simultaneously. Now, it is clear that every member of an equicontinuous family is uniformly continuous.

Lecture 18
Wednesday
February 17

We want to study the compact subsets of $\mathscr{C}^0(X)$.

Theorem 3.28 (Arzelà–Ascoli)**.** *Let K is a compact subset of a metric space (X, d). Let $(f_n)_{n\in\mathbb{N}}$ be a sequence of functions such that $f_n \in \mathscr{C}(K)$ for every $n \in \mathbb{N}$. If (f_n) is uniformly bounded and equicontinuous on K, then (f_n) contains a uniformly convergent subsequence on K.*

The compact subset of $(\mathscr{C}([a, b]), d_\infty)$ are the subsets of equicontinuous, uniformly bounded, and closed functions on $[a, b]$.

Lemma 3.29 (Cantor Diagonalization Technique)**.** *Let $(f_n)_{n\in\mathbb{N}}$ be a pointwise bounded sequence of functions on a countable set E. Then (f_n) has a subsequence $(f_{n_k})_k$ such that $(f_{n_k}(x))_k$ converges for all $x \in E$, that is, $(f_{n_k})_k$ converges pointwise on E.*

Proof. Since E is countable, so let $E = \{x_i : i \in \mathbb{N}\}$. Since $(f_n(x_i))_n$ is a bounded sequence in $\mathbb{R}$, so, by Bolzano Weierstrass, there exists a subsequence $(f_{1,k})_k$ such that $(f_{1,k}(x_1))_k$ converges as $k \to \infty$. Also, $(f_{1,k}(x_2))_k$ is a bounded sequence in $\mathbb{R}$, so there is a subsequence $(f_{2,k}(x_2))_k$ converges as $k \to \infty$. Note that the sequence of functions $(f_{2,k})_k$, since it is a subsequence of $(f_{1,k})_k$, converges at both x_1 and x_2. Proceeding in the same way, we can obtain a countable collection of subsequences of our original sequence

$$\begin{array}{ccccc}
S_1 : & f_{1,1} & f_{1,2} & f_{1,3} & \cdots \\
S_2 : & f_{2,1} & f_{2,2} & f_{2,3} & \cdots \\
S_3 : & f_{3,1} & f_{3,2} & f_{3,3} & \cdots \\
\vdots & \vdots & \vdots & \vdots & \\
S_n : & f_{n,1} & f_{n,2} & f_{n,3} & \cdots
\end{array}$$

where the sequence in the nth row converges at the points $x_1, \ldots, x_n$, and each row is a subsequence of the one above it.

Consider the diagonal sequence $(f_{k,k})_k$. For each $i \in \mathbb{N}$, the sequence $(f_{k,k})$ is a subsequence of $(f_{i,k})$. Since $(f_{i,k}(x_i))$ converges, then $(f_{k,k}(x_i))$ converges. Therefore, (f_n) admits a convergent subsequence $(f_{k,k})_k$ such that $(f_{k,k}(x))$ converges for all $x \in E$.

□

Lemma 3.30. *Let K is a compact subset of a metric space (X, d). Let $(f_n)_{n\in\mathbb{N}}$ be a sequence of functions such that $f_n \in \mathscr{C}(K)$ for $n \in \mathbb{N}$. If (f_n) is pointwise bounded and equicontinuous on K, then* *(a) (f_n) is uniformly bounded on K.*

(b) (f_n) contains a uniformly convergent subsequence.

Proof.

(a) Let $\varepsilon > 0$ be given. Pick a $\delta > 0$ such that for all $x, y \in K$ and $n \in \mathbb{N}$,

$$d(x,y) < \delta \Longrightarrow |f_n(x) - f_n(y)| < \varepsilon.$$

Let $\bigcup_{x \in K} B_\delta(x)$ be an open cover of K. Since K is compact, then there is $p_1, p_2, \ldots, p_r \in K$ such that $K \subseteq \bigcup_{i=1}^r B_\delta(p_i)$, that is, for $x \in K$, there is $i \in \{1, 2, \ldots, r\}$ such that $d(x, p_i) < \delta$. Since (f_n) is pointwise bounded, then for all $1 \le i \le r$, there exists $M_i \in \mathbb{R}$ such that $|f_n(x_i)| \le M_i$ for all $n \in \mathbb{N}$. Let $M = \max\{M_1, \ldots, M_r\}$. Then, for all $x \in K$,

$$|f_n(x)| \le |f_n(x) - f_n(p_i)| + |f_n(p_i)|$$

where i satisfies that $d(x, p_i) < \delta$, so for all $n \in \mathbb{N}$ and $x \in K$,

$$|f_n(x)| \le \varepsilon + M.$$

We conclude that (f_n) is uniformly bounded.

(b) Let E be a countable dense subset of K. Then, by Lemma 3.30, (f_n) has a subsequence $(f_{n_i})_i$ that converges pointwise on E. Let $f_{n_i} = g_i$ for ease of notation. Fix an $\varepsilon > 0$. Pick a $\delta > 0$ such that for all $x, y \in K$ and $n \in \mathbb{N}$,

$$d(x,y) < \delta \Longrightarrow |f_n(x) - f_n(y)| < \varepsilon.$$

Since E is a dense subset of K, we have $K \subseteq \bigcup_{x \in E} B_\delta(x)$ as an open cover. Since K is compact, there exists $x_1, \ldots, x_m \in E$ such that $K \subseteq \bigcup_{i=1}^m B_\delta(x_i)$. Since $(g_i(x))_i$ converges for all $x \in E$, then $(g_i(x_s))_i$ converges for $i = 1, \ldots, m$. Therefore, there exists $N_s \in \mathbb{N}$ such that

$$|g_i(x_s) - g_j(x_s)| < \varepsilon$$

for any $i, j \ge N_s$. Let $N = \max\{N_1, \ldots, N_m\}$, then for $1 \le s \le m$ and $i, j \ge N$,

$$|g_i(x_s) - g_j(x_s)| < \varepsilon.$$

Since $K \subseteq \bigcup_{i=1}^m B_\delta(x_i)$, then $x \in K$ implies that $x \in B_\delta(x_s)$ for some $s \in \{1, \ldots, m\}$. Then, by equicontinuity, we have

$$g_i(x) - g_i(x_s)| < \varepsilon$$

for all $i \in \mathbb{N}$. Therefore, for $x \in K$ and $i, j \ge N$, then

$$\begin{aligned}|g_i(x) - g_j(x)| &\le |g_i(x) - g_i(x_s)| + |g_i(x_s) - g_j(x_s)| + |g_j(x_s) - g_j(x)| \\ &\le \varepsilon + \varepsilon + \varepsilon \\ &= 3\varepsilon.\end{aligned}$$

So, (g_i) is uniformly Cauchy. We conclude that (g_i) converges uniformly.

$\square$

With these two lemmas, we can prove our characterization of compact subsets in $\mathscr{C}^0(X)$.

Proof of Theorem 3.28. Since (f_n) is uniformly bounded by assumption, then (f_n) is pointwise bounded. Therefore, it follows directly from (b) in the prior lemma. □

In fact, equicontinuity is necessary for uniform convergence.

Theorem 3.31. *Let K be a compact subset of a metric space (X,d). Let $(f_n)_{n\in\mathbb{N}}$ be a sequence of functions such that $f_n \in \mathscr{C}(K)$ for all $n \in \mathbb{N}$. If (f_n) converges uniformly to f on K, then (f_n) is equicontinuous on K.*

Proof. Fix an $\varepsilon > 0$. Since $f_n \to f$ uniformly on K, so (f_n) is uniformly Cauchy and there exists $N \in \mathbb{N}$ such that for all $x \in K$ and $n \geq N$,

$$|f_n(x) - f_N(x)| \leq \varepsilon.$$

Since f_n is continuous on a compact set K, so f_n is uniformly continuous on K, so for any $n \in \mathbb{N}$, there is $\delta_n > 0$ such that

$$d(x,y) < \delta_n \Longrightarrow d(f_n(x), f_n(y)) < \varepsilon.$$

Pick $\delta = \min\{\delta_1, \ldots, \delta_N\} > 0$. Then,

$$d(x,y), \delta \Longrightarrow |f_i(x) - f_i(y)| < \varepsilon$$

for $i = 1, 2, \ldots, N$. If $n > N$, then $d(x,y) < \delta$ implies

$$|f_n(x) - f_n(y)| \leq |f_n(x) - f_N(x)| + |f_N(x) - f_N(y)| + |f_N(y) - f_n(y)| \leq 3\varepsilon.$$

Since $\varepsilon > 0$ is arbitrary, so we conclude that (f_n) is equicontinuous on K. □

Lecture 19
Friday
February 19

3.6 The Stone-Weierstrass Approximation Theorem via Convolution

We investigate the compact subsets of the space of continuous functions via Arzelà–Ascoli Theorem. Let K be compact. The compact subset of $(\mathscr{C}(K), d_\infty)$ are the subsets of equicontinuous, unifomrly bounded, closed functions on K.

Now, we want to study the dense subsets of the metric space $(\mathscr{C}([a,b]), d_\infty)$. Recall that a set E in a metric space (X,d) is *dense* if the closure of E is exactly the space X. The following theorem shows that the continuous real-valued functions on a compact interval can be uniformly approximated by polynomials. In other words, the polynomials are dense in $\mathscr{C}([a,b])$ with respect to the supremum norm.

Theorem 3.32 (Stone-Weierstrass Approximation Theorem)**.** *Let $f : [0,1] \to \mathbb{R}$ be a continuous function. Then, for each $\varepsilon > 0$, there exists a polynomial P such that for all $x \in [0,1]$, $|f(x) - P(x)| < \varepsilon$. In other words, for any such $f \in \mathscr{C}([0,1])$, there exists a sequence of polynomials $(P_n)_{n\in\mathbb{N}}$ such that $P_n \to f$ uniformly on $[0,1]$.*

Before giving a proof directly, we first dive into the main tool lying in the proof, *convolution*. This plays an important role in Harmonic Analysis, Partial Differential Equations, and Probability Theory. In particular, the convolution with an *approximate identity* is the important case on which this theorem is based.

Definition 3.33 (Convolution)**.** Let $f, g : \mathbb{R} \to \mathbb{R}$ be functions such that they are Riemann integrable on all compact integrals. Then, for each fixed $t \in \mathbb{R}$, the function $t \mapsto f(x-t)$ and $t \mapsto g(x-t)$ are also Riemann integrable on all compact intervals. Thus, the function $t \mapsto f(x-t) \cdot g(t)$ is Riemann integrable on all compact intervals. Since one of f and g is 0 outside some compact set, the same is true of the function $t \mapsto f(x-t) \cdot g(t)$. If the integrand is 0 outside $[a, b]$, the *convolution* $f * g : \mathbb{R} \to \mathbb{R}$ is defined by

$$(f * g)(x) = \int_a^b f(x-t) \cdot g(t) \, dt.$$

Provided that $t \mapsto f(x-t) \cdot g(t)$ is 0 outside $[a, b]$, the value of $(f * g)(x)$ does not depend on which $a < b$ are chosen.

The (closure of) the set where a function is nonzero is called its *support*. We are just integrating over the support of $t \mapsto f(x-t) \cdot g(t)$; if we integrate over a larger interval, it makes no difference to the value.

One property of convolution from the definition is that it does not matter what order you convolve the functions.

Lemma 3.34 (Commutativity)**.** *Let $f, g : \mathbb{R} \to \mathbb{R}$ be functions such that they are Riemann integrable on compact intervals. For any $x \in \mathbb{R}$,*

$$(f * g)(x) = (g * f)(x).$$

Proof. Fix an $x \in \mathbb{R}$, and pick $a < b$ such that $f(x-t) \cdot g(t) = 0$ for $t \notin [a, b]$. Let $u = x - t$. Then, $du = -dt$, and we have

$$(f * g)(x) = \int_a^b f(x-t) \cdot g(t) \, dt = \int_{x-a}^{x-b} f(u) \cdot g(u-x) \, (-du) = \int_{x-b}^{x-a} g(u-x) \cdot g(u) \, du.$$

Since $f(x-t) \cdot g(t) = 0$ when $t \notin [a, b]$, it follows that $g(u-x) \cdot f(u) = 0$ when $u \notin [x-b, x-a]$. Therefore, we conclude that

$$(g * f)(x) = \int_{x-b}^{x-a} g(u-x) \cdot g(u) \, du = (f * g)(x).$$

□

Also, by the definition of *convolution*, this operator is associative. Another identity that the convolution operator holds is the *approximate identity*, the sequence of functions (ψ_n) such that for each continuous function f, $f * g_n \to f$ in some sense.

Definition 3.35 (Approximate Identity)**.** An *approximate identity* (ψ_n) is a sequence of functions with the following properties:

(a) $\displaystyle\int_{\mathbb{R}} \psi_n(y) \, dy = 1$;

(b) $\displaystyle\int_{\mathbb{R}} |\psi_n(y)| \, dy < M$ for some $M \in \mathbb{R}$ that depends on $n \in \mathbb{N}$;

(c) for all $\delta > 0$,

$$\lim_{n \to \infty} \int_{|y| > \delta} \psi_n(y) \, dy = 0.$$

Lemma 3.36. *Let $(\psi_n)_{n \in \mathbb{N}}$ be the approximate identity. If f is a continuous function on $\mathbb{R}$, then $\psi_n * f$ converges uniformly to f on compact subsets of $\mathbb{R}$.*

Proof. Let $r > 0$ be sufficiently large enough that all the supports of ϕ_n are contained in $[-r, r]$. Since $\displaystyle\int_{-r}^{r} \psi_n(y)\, dy = 1$, we can write

$$f(x) = \int_{-r}^{r} f(x) \cdot \psi_n(y)\, dy$$

for any $n \in \mathbb{N}$. Then,

$$|(\psi_n * f)(x) - f(x)| = \left| \int_{-r}^{r} \psi_n(y) \cdot (f(x-y) - f(x))\, dy \right| \leq \int_{-r}^{r} \psi_n(y) \cdot |f(x-y) - f(x)|\, dy.$$

Now, let $K \subseteq \mathbb{R}$ be compact. Then, $K \subseteq [-s, s]$ for some $s > 0$. For all $x \in K$, $x - y$ is contained in $[-r-s, r+s]$ for all $|y| \leq r$. Since f is continuous on $\mathbb{R}$, it is uniformly continuous on $[-r-s, r+s]$. Fix an $\varepsilon > 0$. Pick a $\delta > 0$ such that $|f(s) - f(t)| < \dfrac{\varepsilon}{2}$ for $|s - t| < \delta$ with $s, t \in [-r-s, r+s]$. Then,

$$\int_{-r}^{r} \psi_n(y) \cdot |f(x-y) - f(x)|\, dy = \int_{-\delta}^{\delta} \psi_n(y) \cdot |f(x-y) - f(x)|\, dy + \int_{\delta < |y| \leq r} \psi_n(y) \cdot |f(x-y) - f(x)|\, dy.$$

For $-\delta \leq y \leq \delta$, and $x - y, x \in [-r-s, r+s]$, we have $|f(x-y) - f(x)| < \dfrac{\varepsilon}{2}$. For $\delta < |y| \leq r$, let $M = \displaystyle\sup_{|t| \leq r+s} f(t)$, we have $|f(x-y) - f(x)| \leq 2M$. Therefore,

$$\sup_{x \in K} |(\psi_n * f)(x) - f(x)| \leq \int_{-\delta}^{\delta} \psi_n(y) \cdot \frac{\varepsilon}{2}\, dy + \int_{\delta < |y| \leq r} \psi_n(y) \cdot 2M\, dy \leq \frac{\varepsilon}{2} + 2M \int_{|y| > \delta} \psi_n(y)\, dy.$$

There exists $N \in \mathbb{N}$ such that $\displaystyle\int_{|y| > \delta} \psi_n(y)\, dy$ for all $n \geq N$. Hence, $\psi_n * f \to f$ uniformly on K.

$\square$

If the approximate identity sequence (ψ_n) consists of all smooth functions (say $\mathscr{C}^\infty$), then the functions $\psi_n * f$ will be $\mathscr{C}^\infty$ as well, no matter how rough f is. This concludes that it is always possible to approximate any continuous function f uniformly by smooth functions.

Lemma 3.37. *If f is Riemann integrable and supported in $[0, 1]$, and if $P : [-1, 1] \to \mathbb{R}$ be a polynomial on this domain, then $P * f$ is a polynomial on $[0, 1]$.*

Proof. Since P is a polynomial on $[-1, 1]$, with Binomial Theorem, for $x, y \in [-1, 1]$, $(x, y) \mapsto P(x - y)$ is also a polynomial. In particular,

$$P(x-y) = \sum_{k=0}^{d} a_k(y) \cdot x^k$$

where d is the degree of polynomial P, and a_k is polynomial of y. For $x, y \in [0, 1]$, the difference $x - y$ is in $[-1, 1]$. Then,

$$(p * f)(x) = \int_0^1 P(x-y) \cdot f(y)\, dy = \int_0^1 \sum_{k=0}^{d} a_k(y) \cdot x^k \cdot f(y)\, dy = \sum_{k=0}^{d} x^k \int_0^1 a_k(y) \cdot f(y)\, dy.$$

The polynomial $a_k(\cdot)$ is Riemann integrable on $[0, 1]$, as is f, so $\displaystyle\int_0^1 a_k(y) \cdot f(y)\, dy$ is finite for all $0 \leq k \leq d$. Therefore, for $x \in [0, 1]$, $x \mapsto (P * f)(x)$ is a polynomial.

$\square$

Combining Lemma 3.36 and Lemma 3.37 allows us th prove the Stone-Weierstrass Approximation Theorem.

Proof of Theorem 3.32. Define

$$g(x) = f(x) - f(0) - x \cdot (f(1) - f(0))$$

for $0 \leq x \leq 1$. Here g is continuous on $[0,1]$ and $g(0) = g(1) = 0$. Hence, we can extend g continuously to $\mathbb{R}$ by setting $g(x) = 0$ for all $x \notin [0,1]$. It suffices to find a sequence of polynomials (P_n) uniformly approximates g. By Lemma 3.36, we want to show that there exists an approximate identity sequence consisting of polynomials on $[-1,1]$.

Let

$$\psi_n(x) = \frac{1}{c_n}(1-x^2)^n \text{ for } -1 \leq x \leq 1,$$

where

$$c_n = \int_{-1}^{1} (1-x^2)^n \, dx.$$

By Lemma 3.37, $\psi_n * f$ is a polynomial on $[0,1]$ for all $n \in \mathbb{N}$. It is immediate that $\psi_n(x) \geq 0$ for all $n \in \mathbb{N}$ and $x \in [-1,1]$, and $\displaystyle\int_{-1}^{1} \psi_n(y) \, dy = 1$. Now, we want to show that for all $\delta > 0$,

$$\lim_{n \to \infty} \int_{|y|>\delta} \psi_n(y) \, dy = 0.$$

Fix a $\delta \in (0,1)$. We note that

$$\int_0^1 (1-x^2)^n \, dx \geq \int_0^\delta (1-x^2)^n \, dx \geq \int_0^\delta \frac{x}{\delta}(1-x^2)^n \, dx$$

for $0 \leq x \leq \delta$. Similarly, we have

$$\int_\delta^1 (1-x^2)^n \, dx \leq \int_\delta^1 \frac{x}{\delta}(1-x^2)^n \, dx.$$

Since $\psi_n(x) = \psi_n(-x)$, so

$$\int_{|x|>\delta} \psi_n(x) \, dx = 2\int_\delta^1 \frac{1}{c_n}(1-x^2)^n \, dx \leq \frac{2}{\delta \cdot c_n} \int_\delta^1 x(1-x^2)^n \, dx.$$

Similarly,

$$c_n = \int_{-1}^1 (1-x^2)^n \, dx = 2\int_0^1 (1-x^2)^n \, dx \geq 2\int_0^\delta \frac{x}{\delta}(1-x^2)^n \, dx.$$

Therefore,

$$\int_{|x|>\delta} \psi_n(x) \, dx \leq \frac{\displaystyle\int_\delta^1 x(1-x^2)^n \, dx}{\displaystyle\int_0^\delta x(1-x^2)^n \, dx}.$$

Let $u = 1 - x^2$, then $du = -2x \, dx$. So,

$$\frac{\displaystyle\int_\delta^1 x(1-x^2)^n \, dx}{\displaystyle\int_0^\delta x(1-x^2)^n \, dx} = \frac{\displaystyle\int_0^{1-\delta^2} u^n \, du}{\displaystyle\int_{1-\delta^2}^1 u^n \, du} = \frac{(1-\delta^2)^{n+1}}{1-(1-\delta^2)^{n+1}}.$$

As $n \to \infty$, we have $(1-\delta^2)^{n+1} \to 0$ and $1-(1-\delta^2)^{n+1} \to 1$, so

$$\lim_{n\to\infty} \int_{|x|>\delta} \psi_n(x)\, dx = 0.$$

This means that (ψ_n) is an approximate identity sequence. It follows that $P_n = \psi_n * f$ converges to f uniformly on $[0,1]$, by Lemma 3.36. P_n is a polynomial on $[0,1]$ by Lemma 3.37. □

To generalize from $[0,1]$ to arbitrary compact interval $[a,b]$, we can translate and dilate back.

Corollary 3.38. *For every interval $[-a,a]$, there is a sequence of real polynomials $(P_n)_{n\in\mathbb{N}}$ such that $P_n(0) = 0$ and $P_n \to |x|$ uniformly on $[-a,a]$.*

Proof. By Stone-Weierstrass, there exists a sequence of polynomials $(P_n^*)_{n\in\mathbb{N}}$ which converges to $|x|$ uniformly on $[-a,a]$. In particular, $P_n^*(0) \to 0$ as $n \to \infty$, so $(P_n^*(0))_{n\in\mathbb{N}}$ is a sequence converges to 0. Then, the polynomial

$$P_n(x) = P_n^*(x) - P_n^*(0)$$

has the property $P_n(0) = 0$. We conclude that $P_n(x) \to |x|$ uniformly on $[-a,a]$. □

Lecture 20
Monday
February 22

3.7 The Stone-Weierstrass Theorem for Compact Set X

What is so special about polynomials that allows them to uniformly approximate continuous functions? Can we find other families of nice functions that uniformly approximate continuous functions? We now wish to generalize Theorem 3.32 to general metric spaces and to generalize to collections other than polynomials.

Definition 3.39 (Algebra)**.** Let $\mathscr{A}$ be a collection of real valued functions all defined on some set E. We say $\mathscr{A}$ is an *algebra* if it is closed under pointwise addition, scalar multiplication, and pointwise multiplication, that is, given $f, g \in \mathscr{A}$ and $\lambda \in \mathbb{R}$, then $f+g \in \mathscr{A}$, $\lambda \cdot f \in \mathscr{A}$, and $f \cdot g \in \mathscr{A}$.

Definition 3.40. Let $\mathscr{A}$ be a collection of real valued functions all defined on some set E. We say $\mathscr{A}$ is *uniformly closed* if $f \in \mathscr{A}$ whenever $f_n \in \mathscr{A}$ for $n \in \mathbb{N}$, and $f_n \to f$ uniformly on E.

Given this collection $\mathscr{A}$, we define the *uniform closure* $\mathscr{B}$ of $\mathscr{A}$ as

$$\mathscr{B} = \{f : f_n \in \mathscr{A} \text{ for all } n \in \mathbb{N} \text{ such that } f_n \to f \text{ uniformly}\},$$

that is, $\mathscr{B}$ is the collection of uniform limit points of sequences in $\mathscr{A}$.

Let A be the set of polynomials ddefined on $[a,b]$. By Stone-Weierstrass, the *uniform closure* of A is $\mathscr{C}([a,b])$. This means

(a) If there exists a sequence of polynomials $(P_n)_{n\in\mathbb{N}}$ such that $P_n \to f$ uniformly on $[a,b]$, then f is continuous on $[a,b]$.

(b) For $f \in \mathscr{C}([a,b])$, there exists a sequence of polynomials $(P_n)_{n\in\mathbb{N}}$ in A such that $P_n \to f$ uniformly on $[a,b]$.

Theorem 3.41. *Let $\mathscr{A}$ be an algebra of bounded real-valued functions on a compact set X and $\mathscr{B}$ be the uniform closure of $\mathscr{A}$. Then, $\mathscr{B}$ is a uniformly closed algebra. In other words, the uniform closure of an algebra is still an algebra.*

Proof. Let $f, g \in \mathscr{B}$ and $\lambda \in \mathbb{R}$. Then, there are uniformly convergent sequences $(f_n)_{n\in\mathbb{N}}$ and $(g_n)_{n\in\mathbb{N}}$ in $\mathscr{A}$ such that $f_n \to f$ and $g_n \in g$ uniformly. Since $\mathscr{A}$ is an algebra, so $f_n + \lambda \cdot g_n \in \mathscr{A}$ for $n \in \mathbb{N}$. Now,

$$\begin{aligned} d_\infty(f_n + \lambda \cdot g_n, f + \lambda \cdot g) &= \sup_{x\in X} |(f_n(x) + \lambda \cdot g_n(x)) - (f(x) + \lambda \cdot g(x))| \\ &\leq \sup_{x\in X} |f_n(x) - f(x)| + \lambda \cdot \sup_{x\in X} |g_n(x) - g(x)| \\ &= d_\infty(f_n, f) + \lambda \cdot d_\infty(g_n, g). \end{aligned}$$

Since $d_\infty(f_n, f) \to 0$ and $d_\infty(g_n, g) \to 0$, then

$$d_\infty(f_n + \lambda \cdot g_n, f + \lambda \cdot g) \to 0.$$

Therefore, $f + \lambda \cdot g \in \mathscr{B}$, which means that $\mathscr{B}$ is closed under addition and scalar multiplication.

Since f and g are defined on X, which is compact, so there exists an $M_1, M_2 \in \mathbb{R}$ such that $|f_n(x)| < M_1$ and $|g_n(x)| < M_2$ for all $n \in \mathbb{N}$ and $x \in X$. Fix $\varepsilon > 0$. Since $f_n \to f$ and $g_n \to g$ uniformly, so there exists $N_1, N_2 \in \mathbb{N}$ such that $|f_n(x) - f(x)| < \dfrac{\varepsilon}{2M_2}$ for all $n \geq N_1$ and $|g_n(x) - g(x)| < \dfrac{\varepsilon}{2M_1}$ for all $n \geq N_2$. Let $N = \max\{N_1, N_2\}$. Then, for $n \geq N$ and $x \in X$, we have

$$\begin{aligned} |f_n(x) \cdot g_n(x) - f(x) \cdot g(x)| &= |f_n(x) \cdot g_n(x) - f_n(x) \cdot g(x) + f_n(x) \cdot g(x) - f(x) \cdot g(x)| \\ &\leq |f_n(x)| \cdot |g_n(x) - g(x)| + |g(x)| \cdot |f_n(x) - f(x)| \\ &< M_1 \cdot \frac{\varepsilon}{2M_1} + M_2 \cdot \frac{\varepsilon}{2M_2} = \varepsilon. \end{aligned}$$

So, $f_n \cdot g_n \to f \cdot g$ uniformly on X. As $\mathscr{A}$ is an algebra, so $f_n \cdot g_n \in \mathscr{A}$ for $n \in \mathbb{N}$, and this proves that the product $f \cdot g$ is in the uniform closure of $\mathscr{A}$. Thus, $\mathscr{B}$ is closed under pointwise multiplication. So, $\mathscr{B}$ is an algebra.

Let $(f_n)_{n\in\mathbb{N}}$ be an uniformly convergent sequence of functions defined on X from $\mathscr{B}$. There is a sequence of functions $(g_n)_{n\in\mathbb{N}}$ such that $|f_n(x) - g_n(x)| < \dfrac{1}{n}$. If $f_n \to f$, then $g_n \to f$ as well, so by the definition of uniform closure $\mathscr{B}$, so $f \in \mathscr{B}$ and $\mathscr{B}$ is uniformly closed.

□

Definition 3.42. Let $\mathscr{A}$ be a family of functions on a set E. We say that $\mathscr{A}$ *separates points on E* if for every $x_1, x_2 \in E$ such that $x_1 \neq x_2$, there exists $f \in \mathscr{A}$ such that $f(x_1) \neq f(x_2)$, and $\mathscr{A}$ *vanishes at no point of E* if for every $x \in E$, there exists $g \in \mathscr{A}$ such that $g(x) \neq 0$.

Example 3.43 (Polynomials)**.**

(a) The algebra $\mathscr{A} = \{P : \text{ polynomials on } \mathbb{R}\}$ is separating and vanishing nowhere.

(b) The algebra $\mathscr{A}_1 = \{P : \text{ polynomials on } \mathbb{R} \text{ such that } P(x) = P(-x) \text{ for } x \in \mathbb{R}\}$ is not separating. Consider $x_1 = 1$ and $x_2 = -1$, for every $P \in \mathscr{A}_1$, we have $P(1) = P(-1)$.

(c) The algebra $\mathscr{A}_2 = \{P : \text{ polynomials on } \mathbb{R} \text{ such that } P(-x) = -P(x) \text{ for } x \in \mathbb{R}\}$ is vanishing somewhere. For every $P \in \mathscr{A}_2$, we have $P(x) + P(-x) = 0$, so $P(0) = 0$, which means $\mathscr{A}_2$ vanishes at $x = 0$.

Let X be a compact metric space. By the Extreme Value Theorem, the function $f : X \to \mathbb{R}$ is bounded. Thus, the supremum norm $d_\infty(f, g) = \sup_{x \in X} |f(x) - g(x)|$ is a genuine metric on $\mathscr{C}(X)$. Let $(f_n)_{n\in\mathbb{N}}$ be a sequence of functions in $\mathscr{C}(X)$. Then, we say (f_n) uniformly approximates a function $f \in \mathscr{C}(X)$ means that $d_\infty(f_n, f) \to 0$ as $n \to \infty$. In this setting, the Stone-Weierstrass Approximation Theorem says that every function $f \in \mathscr{C}([a, b])$ is the uniform limit of some sequence (p_n) of polynomials on $[a, b]$, that is, the set of polynomials is *dense*.

Theorem 3.44 (Stone-Weierstrass)**.** *Let $\mathscr{A}$ be an algebra of functions $f : K \to \mathbb{R}$ where K is compact. If $\mathscr{A}$ separates points on K and vanishes at no point of K, then the uniform closure $\mathscr{B}$ of $\mathscr{A}$ is the set of continuous functions on K, that is, $\mathscr{C}(K)$.*

The real part of Theorem 3.44 is the statement that a real sub-algebra of $\mathscr{C}(K)$ that separates points and vanishes nowhere is dense in $\mathscr{C}(K)$. In light of Theorem 3.41, we aim to prove that *if $\mathscr{A} \subseteq \mathscr{C}(K)$ is a uniformly closed algebra that separates points and vanishes nowhere, then $\mathscr{A} = \mathscr{C}(K)$.* To do that, we need two more lemmas first.

Lemma 3.45. *Suppose $\mathscr{A}$ is an algebra of functions on E that separates points on E and vanishes at no point of E. Suppose $x_1, x_2 \in E$ but $x_1 \neq x_2$ and $c_1, c_2 \in \mathbb{R}$. Then, there exists a function $f \in \mathscr{A}$ such that $f(x_1) = c_1$ and $f(x_2) = c_2$.*

Proof. By assumption, there exists $g, h, k : E \to \mathbb{R}$ and $g, h, k \in \mathscr{A}$ such that $g(x_1) \neq g(x_2)$, $h(x_1) \neq 0$, and $k(x_2) \neq 0$.

Let

$$u = g \cdot k - g(x_1) \cdot k \text{ and } v = g \cdot h - g(x_2) \cdot h$$

and $u, v \in \mathscr{A}$. Note that $u(x_1) = 0 = v(x_2)$. Then,

$$u(x_2) = (g(x_2) - g(x_1)) \cdot k(x_2) \neq 0$$

and $v(x_1) \neq 0$. Now, define

$$f = \frac{c_1 \cdot v}{v(x_1)} + \frac{c_2 \cdot u}{u(x_2)}.$$

Then, $f \in \mathscr{A}$, and we conclude that $f(x_1) = c_1$ and $f(x_2) = c_2$.

□

Lemma 3.46. *Let K be a compact metric space, and let $\mathscr{A} \subseteq \mathscr{C}(K)$ be a uniformly closed algebra. Then, for each $f \in \mathscr{A}$, $|f| \in \mathscr{A}$ as well. Consequently, if $f, g \in \mathscr{A}$, then $\max\{f, g\}$ and $\min\{f, g\}$ are in $\mathscr{A}$ as well.*

Proof. Let $M = \sup_{x \in K} |f(x)|$. Fix an $\varepsilon > 0$. By Stone-Weierstrass Approximation Theorem, there exists a sequence of polynomials $(P_n)_{n\in\mathbb{N}}$ such that $P_n \to |y|$ uniformly on $[-M, M]$. In particular, for $\varepsilon > 0$, there exists $c_1, \ldots, c_n \in \mathbb{R}$ such that

$$\left| \sum_{i=1}^{n} c_i \cdot y^i - |y| \right| < \varepsilon$$

for $-M \leq y \leq M$. Since $\mathscr{A}$ is an algebra, then the function $g = \sum_{i=1}^{n} c_i \cdot f^i$ is in $\mathscr{A}$. Then,

$$|g(x) - |f(x)|| = \left| \sum_{i=1}^{n} c_i \cdot (f(x))^i - |f(x)| \right| < \varepsilon$$

for $x \in K$. Consider a sequence (ε_n) such that $\varepsilon_n \to 0$, and let $g_n(x) \in \mathscr{A}$ be the corresponding function such that $|g_n - |f|| < \varepsilon$, then $g_n \to |f|$ uniformly on K and $g_n \in \mathscr{A}$ for $n \in \mathbb{N}$. Since $\mathscr{A}$ is uniformly closed, so $|f| \in \mathscr{A}$.

We note that

$$\max\{f,g\} = \frac{f+g}{2} + \frac{|f-g|}{2} \text{ and } \min\{f,g\} = \frac{f+g}{2} - \frac{|f-g|}{2}.$$

By the definition of algebra and prior paragraph, we conclude that $\max\{f,g\} \in \mathscr{A}$ and $\min\{f,g\} \in \mathscr{A}$ as well. Further, by induction, for $f_1, \dots, f_n \in \mathscr{A}$, then $\max_{1\leq i\leq n} f_i, \min_{1\leq i\leq n} f_i \in \mathscr{A}$.

$\square$

Lecture 21
Wednesday
February 24

We are finally in a position to complete the proof for a more general version of Stone-Weierstrass Theorem (in a compactness setting).

Proof of Theorem 3.44. Fix a function $f \in \mathscr{C}(K)$. Fix a point $x \in K$. Since $\mathscr{A} \subseteq \mathscr{B}$ and $\mathscr{A}$ separates points and vanishes at no points of K, then for any point $y \in K$, by Lemma 3.45, there exists a function h_y such that

$$h_y(x) = f(x) \text{ and } h_y(y) = f(y).$$

Fix an $\varepsilon > 0$. By continuity of $h_y \in \mathscr{B}$, there exists an open neighborhood $V_y \in K$ such that $y \in V_y$ and $h_y(t) > f(y) - \frac{\varepsilon}{2}$ and $f(t) < f(y) + \frac{\varepsilon}{2}$ for all $t \in V_y$. In particular, $h_y(t) > f(t) - \varepsilon$ for all $t \in V_y$. Now, K is compact, and $\{V_y : y \in K\}$ covers K, so there exists $y_1, \dots, y_n$ such that $K \subseteq \bigcup_{i=1}^n V_{y_i}$. Define

$$g_x = \max\{h_{y_1}, h_{y_2}, \dots, h_{y_n}\}.$$

By (induction on) Lemma 3.46, $g_x \in \mathscr{B}$. By contruction, $g_x(t) \geq h_{y_i}(t) > f(t) - \varepsilon$ for all $t \in K$. The function h_y satisfies $h_y(x) = f(x)$, so $g_x(x) = f(x)$ for $x \in K$.

Since g_x and f are continuous, there is an open neighborhood V_x such that $x \in V_x$ and $g_x(u) < f(x) + \varepsilon$ for all $u \in V_x$. (Consider $r = g_x - f$, then $r(x) = 0 < \varepsilon$. Then, there is a neighborhood V_x of x such that $r(t) < \varepsilon$ for $t \in V_x$, so $g_x(t) - f(t) < \varepsilon$ for $t \in V_x$. Thus, $g_x(t) < f(t) + \varepsilon$ for all $t \in V_x$.) Let $\{V_x : x \in K\}$ covers K. By compactness of K, there exists $x_1, \dots, x_m \in K$ such that $K \subseteq \bigcup_{i=1}^m V_{x_i}$. Now, define

$$h = \min\{g_{x_1}, g_{x_2}, \dots, g_{x_m}\}.$$

Again by (induction on) Lemma 3.46, $h \in \mathscr{B}$. Thus, $h(t) > f(t) - \varepsilon$ for $t \in K$. By construction, since $K \subseteq \bigcup_{i=1}^m V_{x_i}$, there is $1 \leq i \leq m$ such that $t \in V_{x_i}$, so

$$h(t) \leq g_{x_i}(t) < f(t) + \varepsilon.$$

Therefore, $f(t) - \varepsilon < h(t) < f(t) + \varepsilon$ for $t \in K$. We conclude that for every $\varepsilon > 0$, there is a function $h \in \mathscr{B}$ such that $|f(t) - h(t)| < \varepsilon$ for all $t \in K$.

Let $\varepsilon_n = \frac{1}{n}$. We can construct a sequence of functions $f_n \in \mathscr{B}$ such that

$$|f_n(t) - f(t)| < \varepsilon_n = \frac{1}{n}$$

for $t \in K$. So, $f_n \to f$ uniformly on K. Since $\mathscr{B}$ is uniformly closed, so $f \in \mathscr{B}$, which means $\mathscr{C}(K) \subseteq \mathscr{B}$. By assumption, $\mathscr{B}$ is the uniform closure of a family of continuous functions on compact set K, then $\mathscr{B} \subseteq \mathscr{C}(K)$ since the uniform limit of continuous function is continuous. Hence, $\mathscr{C}(K) = \mathscr{B}$.

$\square$

3.8 Applications to Power Series

We discussed *real-analytic* functions, that is, functions that can be written as

$$f(x) = \sum_{n=0}^{\infty} c_n \cdot (x-a)^n$$

for $x \in (a-R, a+R)$ where R is the radius of convergence of a power series.

In 140A, we proved that $\sum_{n=0}^{\infty} a_n \cdot (x-a)^n$ converges absolutely if $|x-a| < R$ and diverges if $|x-t| > R$. Now, we will use the results of series of functions to prove that a power series may be differentiated or integrated term by term for $|x-a| < R$.

Theorem 3.47. *Let $\sum_{n=0}^{\infty} c_n \cdot x^n$ be the power series that converges for $|x| < R$. Define*

$$f(x) = \sum_{n=0}^{\infty} c_n \cdot x^n$$

for $x \in (-R, R)$. Then, $\sum_{n=0}^{\infty} c_n \cdot x^n$ converges uniformly on $[-R+\varepsilon, R-\varepsilon]$ for any $\varepsilon > 0$, and f is continuous and differentiable on $(-R, R)$ with

$$f'(x) = \sum_{n=1}^{\infty} n \cdot c_n \cdot x^{n-1}$$

for $x \in (-R, R)$.

Proof. Let $\varepsilon > 0$ be given. For $|x| \leq R - \varepsilon$, we have

$$|x|^n \leq (R-\varepsilon)^n$$

and $\sum_{n=0}^{\infty} c_n \cdot (R-\varepsilon)^n$ converges, so the partial sums $s_N(x) \leq \sum_{n=0}^{N} c_n \cdot (R-\varepsilon)^n$ for $|x| \leq R-\varepsilon$. By Weierstrass M-test, $\sum_{n=0}^{\infty} c_n \cdot x^n$ converges uniformly on $[-R+\varepsilon, R-\varepsilon]$. Therefore, f is continuous, since f is the uniform limit of $s_N(x)$.

Since $\lim_{n\to\infty} \sqrt[n]{n} = 1$, we have

$$\limsup_{n\to\infty} \sqrt[n]{n \cdot |c_n|} = \limsup_{n\to\infty} \sqrt[n]{|c_n|} = R$$

by the definition of radius of convergence and root test. Then, $\sum_{n=0}^{\infty} c_n \cdot x^n$ and $\sum_{n=1}^{\infty} n \cdot c_n \cdot x^n$ have the same radius of convergence, that is, R. Reasoning as above, $\sum_{n=1}^{\infty} n \cdot c_n \cdot x^n$ converges uniformly on $|x| < R - \varepsilon$. Therefore,

$$f'(x) = \sum_{n=1}^{\infty} n \cdot c_n \cdot x^n$$

on $[-R+\varepsilon, R-\varepsilon]$. Since $\varepsilon > 0$ is arbitrary, so we are done.

$\square$

Corollary 3.48 (Link between Power Series and Taylor Series)**.** *Let* $\sum_{n=0}^{\infty} c_n \cdot x^n$ *be the power series that converges for* $|x| < R$*. Define*

$$f(x) = \sum_{n=0}^{\infty} c_n \cdot x^n$$

for $x \in (-R, R)$*. Then,* $f \in \mathscr{C}^\infty((-R, R))$*, that is,* f *is infinitely differentiable, and*

$$f^{(k)}(x) = \sum_{n=k}^{\infty} n(n-1)\cdots(n-k+1) \cdot c_n \cdot x^{n-k}.$$

In particular, $f^{(k)}(0) = k! \cdot c_k$*.*

Proof. Using Theorem 3.47 k times gives

$$f^{(k)}(x) = \sum_{n=k}^{\infty} n(n-1)\cdots(n-k+1) \cdot c_n \cdot x^{n-k}.$$

Plug in $x = 0$, with $0^0 = 1$, we have

$$f^{(k)}(0) = \sum_{n=k}^{\infty} n(n-1)\cdots(n-k+1) \cdot c_n \cdot 0^{n-k} = k! \cdot c_k.$$

□

Further, we get $c_k = \dfrac{f^{(k)}(0)}{k!}$ and

$$f(x) = \sum_{n=0}^{\infty} c_n \cdot x^n = \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} \cdot x^n$$

on $(-R, R)$, which is the Taylor series at 0 for f.

Corollary 3.49 (Uniqueness of Power Series)**.** *Suppose* $f(x) = \sum_{n=0}^{\infty} a_n \cdot x^n$ *and* $g(x) = \sum_{n=0}^{\infty} b_n \cdot x^n$ *on* $(-R, R)$*. If* $f(x) = g(x)$ *for all* $x \in (-R, R)$*, then* $a_n = b_n$ *for* $n \in \mathbb{N}$*.*

This is immediate by noting that

$$a_n = \frac{f^{(n)}(0)}{n!} = \frac{g^{(n)}(0)}{n!} = b_n$$

for all $n \in \mathbb{N}$.

Lecture 23
Monday
March 1

We should remark that if the power series $\sum_{n=0}^{\infty} a_n x^n$ converges conditionally at $x = R$, then it is possible for it to diverge when $x = -R$. The series $\sum_{n=1}^{\infty} \frac{(-1)^n \cdot x^n}{n}$ with $R = 1$ is an example.

Theorem 3.50 (Abel's Test for Uniform Convergence)**.** *Let* $(u_n)_{n\in\mathbb{N}}$ *be a sequence of functions on a set* X *such that* $\sum_{n=1}^{\infty} u_n$ *converges uniformly* X*. Let* $(v_n)_{n\in\mathbb{N}}$ *be a uniformly bounded sequence on* X *such that* $v_{n+1}(x) \le v_n(x)$ *for all* $x \in X$ *and* $n \in \mathbb{N}$*. Then,* $\sum_{n=1}^{\infty} u_n v_n$ *converges uniformly on* X*.*

Proof. We will use an analogue of "summation by parts" technique. Let $(a_n)_{n\in\mathbb{N}}$ and $(b_n)_{n\in\mathbb{N}}$ be sequences of real numbers. Let

$$S_{j,k} = \sum_{n=j}^{k} a_n \text{ for } j \le k.$$

Similarly, we can derive

$$\sum_{k=m+1}^{n} a_k b_k = \sum_{k=m+1}^{n} S_{m+1,k}(b_k - b_{k+1}) + S_{m+1,n} b_{n+1}$$

for $m < n$.

Let $x \in X$. Let $a_n = u_n(x)$ and $b_n = v_n(x)$. Since (v_n) is uniformly bounded, so there is an $M \in \mathbb{R}$ such that $|v_n(x)| < M$ for every $x \in X$ and $n \in \mathbb{N}$. Fix an $\varepsilon > 0$. There exists $N \in \mathbb{N}$ such that for $n > m \ge N$, we have

$$\left| \sum_{k=m+1}^{n} u_k(x) \right| < \frac{\varepsilon}{3M}$$

Then, for $n > m \ge N$,

$$\begin{aligned}
\left| \sum_{k=m+1}^{n} u_k(x) \cdot v_k(x) \right| &= \left| \sum_{k=m+1}^{n} S_{m+1,k}(x) \cdot (v_k(x) - v_{k+1}(x)) + S_{m+1,n}(x) \cdot v_{n+1}(x) \right| \\
&\le \sum_{k=m+1}^{n} |S_{m+1,k}(x)| \cdot |v_k(x) - v_{k+1}(x)| + |S_{m+1,n}(x)| \cdot |v_{n+1}(x)| \\
&\le \frac{\varepsilon}{3M} \sum_{k=m+1}^{n} (v_k(x) - v_{k+1}(x)) + \frac{\varepsilon}{3M} \cdot M \\
&= \frac{\varepsilon}{3M}(v_{m+1}(x) - v_{n+1}(x)) + \frac{\varepsilon}{3} \\
&< \frac{\varepsilon}{3M}(M + M) + \frac{\varepsilon}{3} = \varepsilon.
\end{aligned}$$

Therefore, $\sum_{n=1}^{\infty} u_n v_n$ converges uniformly on X.

□

Corollary 3.51 (Abel's Theorem). *If $\sum_{n=0}^{\infty} a_n$ is convergent, then the power series $\sum_{n=0}^{\infty} a_n x^n$ converges uniformly on $(-1, 1)$.*

This is immediate by setting $u_n(x) = a_n$ and $v_n(x) = x^n$ for $x \in [0, 1]$ and applying Theorem 3.50.

Theorem 3.52 (Abel's Limit Theorem). *Let $\sum_{n=0}^{\infty} a_n$ be convergent. Let $f(x) = \sum_{n=0}^{\infty} a_n x^n$ for $-1 < x < 1$. Then, $\lim_{x\to 1} f(x) = f(1) = \sum_{n=0}^{\infty} a_n$.*

Proof. By Theorem 3.50 and Theorem 3.12, f is continuous on $x = 1$.

□

We will need a theorem on swapping the limits of series.

Theorem 3.53 (Fubini's Theorem for Sums)**.** *Let* $(a_{ij})_{i,j\in\mathbb{Z}_+}$ *be a double sequence* ($a : \mathbb{Z}_+ \times \mathbb{Z}_+ \to \mathbb{R}$) *and assume that* $\sum_{j=1}^{\infty} |a_{ij}| = b_i < \infty$ *and* $\sum_{i=1}^{\infty} b_i$ *converges. Then,*

$$\sum_{i=1}^{\infty}\sum_{j=1}^{\infty} a_{ij} = \sum_{j=1}^{\infty}\sum_{i=1}^{\infty} a_{ij}.$$

Proof. We note that $\sum_{i=1}^{\infty}\sum_{j=1}^{\infty} a_{ij}$ converges since it is bounded in absolute value by $\sum_{i=1}^{\infty}\sum_{j=1}^{\infty} |a_{ij}| = \sum_{i=1}^{\infty} b_i < \infty$. Also, $\sum_{i=1}^{\infty} a_{ij}$ converges for every $j \in \mathbb{Z}_+$ since $|a_{ij}| \le b_i$ and $\sum_{i=1}^{\infty} b_i$ converges.

Let $\varepsilon > 0$. Pick $M \in \mathbb{Z}_+$ such that $\sum_{i=M+1}^{\infty} b_i < \frac{\varepsilon}{2}$. Then, pick $N \in \mathbb{Z}_+$ such that

$$\sum_{j=N+1}^{\infty} |a_{ij}| < \frac{\varepsilon}{2M} \text{ for } 1 \le i \le M.$$

Then, for $n \ge N$, we have

$$\begin{aligned}
\left|\sum_{j=1}^{n}\sum_{i=1}^{\infty} a_{ij} - \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} a_{ij}\right| &= \left|\sum_{i=1}^{\infty}\sum_{j=1}^{n} a_{ij} - \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} a_{ij}\right| \\
&= \left|\sum_{i=1}^{\infty}\left(\sum_{j=1}^{n} a_{ij} - \sum_{j=1}^{\infty} a_{ij}\right)\right| \\
&\le \sum_{i=1}^{\infty}\sum_{j=n+1}^{\infty} |a_{ij}| \\
&\le \sum_{i=1}^{M}\sum_{j=n+1}^{\infty} |a_{ij}| + \sum_{i=M+1}^{\infty}\sum_{j=1}^{\infty} |a_{ij}| \\
&< \frac{\varepsilon}{2} + \sum_{i=M+1}^{\infty} b_i < \varepsilon.
\end{aligned}$$

We can conclude that

$$\sum_{j=1}^{\infty}\sum_{i=1}^{\infty} a_{ij} = \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} a_{ij}.$$

□

Up to now, we are too restrictive towards the power series. Everything above is for power series centered at 0. We can just as well discuss power series like $\sum_{n=0}^{\infty} a_n(x-x_0)^n$ centered at any point $x_0 \in \mathbb{R}$, which means we can "re-center" the power series.

Theorem 3.54 (Taylor's Theorem for Real-Analytic Functions)**.** *Let* $f(x) = \sum_{n=0}^{\infty} a_n x^n$ *on* $|x| < R$*. If* $a \in (-R, R)$*, then*

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n$$

for all $|x - a| < R - |a|$*.*

Proof. We play the trick $x = (x-a)+a$. Then, we have

$$f(x) = \sum_{n=0}^{\infty} c_n x^n = \sum_{n=0}^{\infty} c_n((x-a)+a)^n = \sum_{n=0}^{\infty} c_n \sum_{k=0}^{n} \binom{n}{k} \cdot (x-a)^k \cdot a^{n-k}.$$

Applying Theorem 3.47 k times to the power series $\sum_{n=0}^{\infty} c_n x^n$, we get

$$f^{(k)}(x) = \sum_{n=k}^{\infty} n(n-1)\cdots(n-k+1) \cdot c_n \cdot x^{n-k}$$

for $|x-a| < R$. Then,

$$f^{(k)}(a) = \sum_{n=k}^{\infty} n(n-1)\cdots(n-k+1) \cdot c_n \cdot a^{n-k} = \sum_{n=k}^{\infty} \binom{n}{k} \cdot k! \cdot c_n \cdot a^{n-k}.$$

We note that

$$\sum_{n=0}^{\infty}\sum_{k=0}^{n} \left| c_n \cdot \binom{n}{k} \cdot a^{n-k} \cdot (x-a)^k \right| = \sum_{n=0}^{\infty} |c_n| \sum_{k=0}^{n} \binom{n}{k} \cdot |a|^{n-k} \cdot |x-a|^k = \sum_{n=0}^{\infty} |c_n| \cdot (|x-a| + |a|)^n$$

converges as long as $|x-a| + |a| < R$. By Theorem 3.53, we can swap the order of summation as

$$\sum_{n=0}^{\infty} c_n \sum_{k=0}^{n} \binom{n}{k} \cdot a^{n-k} \cdot (x-a)^k = \sum_{k=0}^{\infty}\sum_{n=k}^{\infty} \binom{n}{k} \cdot c_n \cdot a^{n-k} \cdot (x-a)^k.$$

Then, with

$$\frac{f^{(k)}(a)}{k!} = \sum_{n=k}^{\infty} \binom{n}{k} \cdot c_n \cdot a^{n-k},$$

we conclude that

$$f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(a)}{k!} \cdot (x-a)^k$$

for $|x-a| < R - a$.

□

If two power series converge to same function in $(-R,R)$, Corollary 3.49 shows that the two series must be equal, that is, they must have the same coefficients. Actually, we can arrive at the same conclusion with a much weaker hypotheses.

Theorem 3.55. *Suppose the series $\sum_{n=0}^{\infty} a_n x^n$ and $\sum_{n=0}^{\infty} b_n x^n$ converge on $S = (-R,R)$ where R is the maximum of the radius of convergence of the series. Let E be the set of all $x \in S$ at which*

$$\sum_{n=0}^{\infty} a_n x^n = \sum_{n=0}^{\infty} b_n x^n.$$

If E has a limit point in S, then $a_n = b_n$ for $n \in \mathbb{N}$. Hence, the equality holds for all $x \in S$.

Proof. Let $c_n = a_n - b_n$ and $f(x) = \sum_{n=0}^{\infty} c_n x^n$ for $x \in S$. Then, $f(x) = 0$ on E.

Let A be the set of all limit points of E in S. Then, $A \subseteq S$, $A \neq \varnothing$ by assumption, and A is closed (since the collection of limit points E is precisely the closure of E). Pick $x_0 \in A$, by Theorem 3.54, we can write

$$f(x) = \sum_{n=0}^{\infty} d_n (x - x_0)^n$$

for $|x - x_0| < R - |x_0|$ and $d_n = \dfrac{f^{(n)}(x_0)}{n!}$. We want to show that $d_n = 0$ for all $n \in \mathbb{N}$. Towards a contradiction, assume there exists some $n \in \mathbb{N}$ such that $d_n \neq 0$. Let $k > 0$ be the smallest positive integer such that $d_k \neq 0$. We define

$$g(x) = \sum_{m=0}^{\infty} d_{k+m} (x - x_0)^m$$

for $|x - x_0| < R - |x_0|$. Then, on $|x - x_0| < R - |x_0|$, we can write

$$f(x) = (x - x_0)^k \cdot g(x) = \sum_{n=k}^{\infty} d_n (x - x_0)^n.$$

We note that g is continuous at x_0 and $g(x_0) = d_k \neq 0$, so there exists a $\delta > 0$ such that $g(x) \neq 0$ if $|x - x_0| < \delta$. Then,

$$f(x) = (x - x_0)^k \cdot g(x) \neq 0$$

for $x \neq x_0$ such that $|x - x_0| < \delta$. Therefore, x_0 is not a limit point of points of E, that is, $x_0 \notin A$. This contradicts to our assumption $x_0 \in A$. Thus, A is open.

Since S is connected without doubt, with $A \subseteq S$ open, closed, and nonempty, so $A = S$. For every $x \in S$, there is a sequence of points (x_n) in E such that $x_n \to x \in A = S$. By continuity of f, we have

$$f(x) = \lim_{n \to \infty} f(x_n) = 0$$

for all $x \in S$.

$\square$

4 Transcendental Functions

Recall that functions that are given by a convergent power series are called *analytic*.

Definition 4.1 (Real Analytic Functions)**.** Let E be an open subset of $\mathbb{R}$, and let $f : E \to \mathbb{R}$ be a function. We say f is *real analytic* on E if for any $x_0 \in E$, we can write $f(x) = \sum_{n=0}^{\infty} a_n (x - x_0)^n$ where $a_n \in \mathbb{R}$ for $n \in \mathbb{N}$, and the series converges to f for x in a neighborhood of x_0.

Previously, we prove that analytic functions are $\mathscr{C}^\infty$, infinitely differentiable. The converse is generally false. There are many $\mathscr{C}^\infty$ functions that are not analytic.

Corollary 4.2 (Taylor)**.** *Let f be analytic in a neighborhood of 0. Then f can be represented by the Taylor series*

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} x^n.$$

In particular, analytic functions have unique power series expansions.

This is immediate by Theorem 3.54. Combining with Taylor's Theorem, we can get a sense of what should be true in order for a infinitely differentiable function to be analytic.

Proposition 4.3. *Let $f \in \mathscr{C}^\infty$ over $(-R, R)$. Suppose that for each $\alpha \in (0, R)$ and $k \in \mathbb{N}$,*

$$\frac{1}{k!} \sup_{|x|,|\xi| \le \alpha} \left| f^{(k)}(\xi) \cdot x^k \right| \to 0$$

as $k \to \infty$. Then, f is analytic over $(-R, R)$.

Proof. By Taylor's Theorem, for each $x \in [-\alpha, \alpha]$ and $k \in \mathbb{N}$, there exists $0 \le \xi(x, k) \le x$ such that

$$f(x) = \sum_{n=0}^{k-1} \frac{f^{(n)}(0)}{n!} \cdot x^n + \frac{1}{k!} \cdot f^{(k)}(\xi) \cdot x^k.$$

Then,

$$\sup_{|x| \le \alpha} \left| f(x) - \sum_{n=0}^{k-1} \frac{f^{(n)}(0)}{n!} \cdot x^n \right| \le \frac{1}{k!} \cdot \sup_{|x| \le \alpha} \left| f^{(k)}(\xi(x,k)) \cdot x^k \right| \le \frac{1}{k!} \sup_{|x|,|\xi| \le \alpha} \left| f^{(k)}(\xi) \cdot x^k \right| \to 0.$$

Thus, the Taylor polynomials converge uniformly to f on $[-\alpha, \alpha]$ for every $\alpha < R$. Therefore, f is given by the convergent Taylor series on $(-R, R)$.

□

Further, functions can be classified into two broad groups. Polynomial functions are called *algebraic*, as are functions obtained from them by addition, multiplication, division, or taking powers and roots. Functions that are not algebraic are called *transcendental*. The exponential, logarithmic, and trigonometric functions are transcendental, as are their inverses.

Lecture 24
Wednesday
March 3

4.1 The Exponential Function

We can now use the machinery developed in the power series to develop a rigorous foundation for many functions. Before rigorously defining the exponential function, we start with an example for motivation.

Example 4.4. Suppose that $E : (-R, R) \to \mathbb{R}$ is a function which satisfies

$$E(x + y) = E(x) \cdot E(y)$$

for all $x, y \in (-R, R)$ such that $x + y \in (-R, R)$. We note that for any $x \in (-R, R)$,

$$E(x) = E(x + 0) = E(x) \cdot E(0),$$

so $E(x) = 0$ for all x or $E(0) = 1$. It is boring to discussion $E(x) = 0$ for $x \in (-R, R)$. In the other case $E(0) = 1$, we assume that E is differentiable at 0 additionally. Then,

$$E'(0) = \lim_{t \to 0} \frac{E(t) - 1}{t}$$

exists. Then, we can compute for any $x \in (-R, R)$. Pick t sufficiently small such that $x + t \in (-R, R)$,

$$E'(x) = \lim_{t \to 0} \frac{E(x + t) - E(x)}{t} = \lim_{t \to 0} \frac{E(x) \cdot E(t) - E(x)}{t} = E(x) \cdot \lim_{t \to 0} \frac{E(t) - 1}{t} = E'(0) \cdot E(x).$$

Therefore, E is differentiable at all points in its domain. Let $\lambda = E'(0)$, then E satisfies the differential equation $E'(x) = \lambda \cdot E(x)$ for all x. By induction, we have $E^{(k)}(x) = \lambda^k \cdot E(x)$. Since E is defined on $(-R, R)$, this shows that $E \in \mathscr{C}^k(-R, R)$ for al $k \in \mathbb{N}$, so E is infinitely differentiable on $(-R, R)$. For $\alpha < R$, we have

$$\frac{1}{k!} \sup_{|x|, |\xi| \leq \alpha} \left| E^{(k)}(\xi) \cdot x^k \right| = \frac{|\lambda|^k}{k!} \sup_{|x|, |\xi| \leq \alpha} \left| E(\xi) \cdot x^k \right| = \frac{(|\lambda| \cdot \alpha)^k}{k!} \sup_{|\xi| \leq \alpha} |E(\xi)|.$$

Since E is infinitely differentiable over $(-R, R)$, it is continuous on $[-\alpha, \alpha]$, which is compact, so E is bounded over $[-\alpha, \alpha]$. In addition, as $k \to \infty$,

$$\frac{(|\lambda| \cdot \alpha)^k}{k!} \to 0.$$

By Proposition 4.3, E is analytic in $(-R, R)$.

We can now compute the Taylor expansion of E. We have $E^{(n)}(0) = \lambda^n \cdot E(0) = \lambda^n$, so

$$E(x) = \sum_{n=0}^{\infty} \frac{\lambda^n}{n!} \cdot x^n.$$

Let $a_n = \dfrac{\lambda^n}{n!}$ for $n \in \mathbb{N}$. Then,

$$\frac{a_{n+1}}{a_n} = \frac{\lambda^{n+1}}{(n + 1)!} \cdot \frac{n!}{\lambda^n} = \frac{\lambda}{n + 1} \to 0.$$

By ratio test, the series converges. Then, by Theorem 3.47, the power series expansion for E actually converges uniformly on $\mathbb{R}$ (that is, the radius of convergence can be $R = \infty$).

Definition 4.5 (Exponential Function)**.** For $x \in \mathbb{R}$, with the special case $\lambda = 1$, we define the *exponential function* e^x to be the real value

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}.$$

This functions is analytic on $\mathbb{R}$. Let $x = 1$. It comes with the famous *Euler's number*, that is,

$$e = \sum_{n=0}^{\infty} \frac{1}{n!} \approx 2.71828183...$$

Theorem 4.6 (Properties of Exponential Function)**.** *Let* $e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}$*. Then*

(a) $e^x : \mathbb{R} \to (0, \infty)$ *is continuous and differentiable at every order.*

(b) For $x \in \mathbb{R}$, $(e^x)' = e^x$*.*

(c) e^x *is strictly monotone increasing. In other words, for* $x, y \in \mathbb{R}$*, we have* $e^y > e^x$ *if and only if* $y > x$*.*

(d) For $x, y \in \mathbb{R}$, $e^{x+y} = e^x \cdot e^y$*.*

(e) $e^x \to \infty$ *as* $x \to \infty$ *and* $e^x \to 0$ *as* $x \to -\infty$*.*

(f) $\lim_{x\to\infty} x^n e^{-x} = 0$ *for every* $n \in \mathbb{N}$*. In other words,* e^x *converges to* 0 *faster than any polynomial.*

Proof.

(a) Let $a_n = \frac{x^n}{n!}$. We have

$$\frac{a_{n+1}}{a_n} = \frac{x}{n+1} \to 0$$

as $n \to \infty$. By the ratio test, $\sum_{n=0}^{\infty} a_n$ converges on $\mathbb{R}$. Continuity and differentiability follows from Theorem 3.47.

(b) This is a direct computation.

$$(e^x)' = \sum_{n=0}^{\infty} \left(\frac{x^n}{n!}\right)' = \sum_{n=1}^{\infty} n \cdot \frac{x^{n-1}}{n!} = \sum_{n=1}^{\infty} \frac{x^{n-1}}{(n-1)!} = \sum_{k=0}^{\infty} \frac{x^k}{k!} = e^x.$$

(c) By (b), we have $(e^x)' = e^x > 0$ for $x \in \mathbb{R}$, so e^x is strictly monotone increasing.

(d) This is a direct computation.

$$\begin{aligned}
e^{x+y} &= \sum_{n=0}^{\infty} \frac{(x+y)^n}{n!} \\
&= \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{k=0}^{n} \binom{n}{k} \cdot x^k \cdot y^{n-k} \\
&= \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{k=0}^{n} \frac{n!}{k! \cdot (n-k)!} \cdot x^k \cdot y^{n-k} \\
&= \sum_{n=0}^{\infty} \sum_{k=0}^{\infty} \frac{x^k}{k!} \cdot \frac{y^{n-k}}{(n-k)!} \\
&= \sum_{n=0}^{\infty} \frac{x^n}{n!} \cdot \sum_{n=0}^{\infty} \frac{y^n}{n!} = e^x \cdot e^y.
\end{aligned}$$

(e) Consider the function $f(x) = e^x - x$. Then, by (b), we have

$$f'(x) = e^x - 1.$$

For $x \geq 0$, $e^x \geq 1$, so $f'(x) \geq 0$. With $f(0) = 1$, f is monotone increasing, and $f(x) \geq 1 > 0$. Then, $e^x \geq x$ for $x \geq 0$. Therefore,

$$\lim_{x\to\infty} e^x \geq \lim_{x\to\infty} x = \infty.$$

We note that $e^x \cdot e^{-x} = 1$, so

$$\lim_{x\to\infty} e^{-x} = \lim_{x\to\infty} \frac{1}{e^x} = 0.$$

(f) Fix $n \in \mathbb{N}$. By the definition of exponential funciton, we have

$$e^x > \frac{x^{n+1}}{(n+1)!}$$

for $x > 0$. Then,

$$x^n \cdot e^{-x} < \frac{(n+1)!}{x} \to 0$$

as $x \to \infty$. We conclude that $\lim\limits_{x\to\infty} x^n \cdot e^{-x} = 0$.

$\square$

4.2 The Logarithmic Function

Since $e^x : \mathbb{R} \to (0,\infty)$ is strictly increasing and differentiable on $\mathbb{R}$, then it has an inverse function from $(0,\infty)$ onto $\mathbb{R}$.

Definition 4.7 (Logarithm Function)**.** We define the *natural logarithm function* $\log : (0,\infty) \to \mathbb{R}$ to be the inverse of the exponential function. Thus, $\log e^x = x$ for $x \in \mathbb{R}$ and $e^{\log y} = y$ for $y > 0$.

Theorem 4.8 (Properties of Logarithmic Function)**.**

(a) $\log' x = \dfrac{1}{x}$ *for* $x \in (0,\infty)$.

(b) $\log 1 = 0$ *and* $\log e = 1$.

(c) $\log x = \displaystyle\int_1^x \frac{dt}{t}$.

(d) $\log xy = \log x + \log y$ *for* $x, y \in (0,\infty)$.

(e) $\lim\limits_{x\to\infty} \log x = \infty$ *and* $\lim\limits_{x\to 0} \log x = -\infty$.

(f) For $p \in \mathbb{Q}$ *and* $x > 0$, $x^p = e^{p\log x}$. *For* $\alpha \in \mathbb{R}$ *and* $x > 0$, $x^\alpha = e^{\alpha \log x}$.

(g) $\lim\limits_{x\to\infty} x^{-n} \log x = 0$ *for every positive integer* n. *In other words,* $\log x$ *converges to infinity slower than any polynomial.*

Proof.

(a) By the definition of inverse function, we have $\log e^x = x$ for $x \in \mathbb{R}$. Differentiation gives

$$\log' e^x \cdot e^x = 1.$$

Let $y = e^x > 0$. Then, $\log' y \cdot y = 1$. Therefore,

$$\log' y = \frac{1}{y}.$$

(b) Since $\log x$ is the inverse of e^x and $e^0 = 1$ and $e^1 = e$, so $\log 1 = 0$ and $\log e = 1$.

(c) By (a) and the Fundamental Theorem of Calculus, we have

$$\int_1^x \frac{dt}{t} = \log x - \log 1 = \log x.$$

(d) Let $x = e^u$ and $y = e^v$ and $u, v \in \mathbb{R}$ indeed exist. Then,

$$\log(x \cdot y) = \log(e^u \cdot e^v) = \ln e^{u+v} = u + v = \log x + \log y.$$

(e) By Theorem 4.6 (e), it is obvious that $\log x \to \infty$ as $x \to \infty$ and $\log x \to -\infty$ as $x \to 0$.

(f) Let $p = \dfrac{n}{m}$ where n and m are coprime. Then, $e^{p \log x} > 0$ and

$$(e^{p \log x})^m = e^{m \cdot p \log x} = e^{n \log x} = (e^{\log x})^n = x^n.$$

If $x \geq 1$, then $\log x \geq 0$, so

$$e^{\alpha \log x} = \sup_{p \in \mathbb{Q}, p < \alpha} e^{p \log x} = \sup_{p \in \mathbb{Q}, p < \alpha} x^p = x^\alpha.$$

If $0 < x < 1$, then there exists $y \in \mathbb{R}$ such that $xy = 1$. Therefore,

$$e^{\alpha \log x} = e^{\alpha \log 1/y} = e^{-\alpha \log y} = y^{-\alpha} = x^\alpha.$$

(g) Fix $0 < \varepsilon < \alpha$ and $x > 1$. Then, by (c),

$$x^{-\alpha} \ln x = x^{-\alpha} \int_1^x \frac{dy}{y} < x^{-\alpha} \int_1^x y^{\varepsilon - 1}\, dy = x^{-\alpha} \cdot \frac{x^\varepsilon - 1}{\varepsilon} < \frac{x^{\varepsilon - \alpha}}{\varepsilon}.$$

This converges to 0 as $\alpha \to \varepsilon$.

$\square$

Lecture 25
Friday
March 5

Definition 4.9 (Complex Exponential)**.** Let $z \in \mathbb{C}$. We define the *complex exponential* by

$$e^z = \sum_{n=0}^{\infty} \frac{z^n}{n!}.$$

Proposition 4.10 (Properties of Complex Exponential)**.** *Let $z, w \in \mathbb{C}$. Then*

(a) $e^z \cdot e^w = e^{z+w}$.

(b) $e^z \cdot e^{-z} = 1$.

(c) $e^{\overline{z}} = \overline{e^z}$.

We can immediately modify our proof from real exponential to complex exponential for (a) and (b). For (c), we note that $\overline{z + w} = \overline{z} + \overline{w}$ and $\overline{z \cdot w} = \overline{z} \cdot \overline{w}$. Then,

$$e^{\overline{z}} = \sum_{n=0}^{\infty} \frac{\overline{z}^n}{n!} = \sum_{n=0}^{\infty} \frac{\overline{z^n}}{n!} = \overline{\sum_{n=0}^{\infty} \frac{z^n}{n!}} = \overline{e^z}.$$

4.3 The Trigonometric Functions

Trigonometric functions are often defined using the geometric concepts. However, we will define them in a more analytic way, in particular, the complex exponential function.

Definition 4.11 (Trigonometric Functions)**.** Let $x \in \mathbb{C}$. Then, we define

$$\cos x = \frac{e^{ix} + e^{-ix}}{2}$$

and

$$\sin x = \frac{e^{ix} - e^{-ix}}{2i}.$$

We refer to cos and sin as the *cosine* and *sine* functions respectively.

From the power series definition of the exponential function, we have

$$e^{ix} = 1 + ix - \frac{x^2}{2!} - \frac{ix^3}{3!} + \frac{x^4}{4!} + \cdots$$

and

$$e^{-ix} = 1 - ix - \frac{x^2}{2!} + \frac{ix^3}{3!} + \frac{x^4}{4!} - \cdots ,$$

so

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \cdots = \sum_{n=0}^{\infty} (-1)^n \cdot \sum_{x^{2n}} (2n)!$$

and

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots = \sum_{n=0}^{\infty} (-1)^n \cdot \frac{x^{2n+1}}{(2n+1)!}.$$

In particular, $\cos x$ and $\sin x$ are always real whenever x is real. By the ratio test, two power series definitions of cos and sin are absolutely convergent for $x \in \mathbb{R}$, thus $\cos x$ and $\sin x$ are real analytic at 0 with an infinite radius of convergence. Actually, $\cos x$ and $\sin x$ are real analytic at $x \in \mathbb{R}$.

Theorem 4.12 (Properties of Trigonometric Functions)**.** *Let* $x \in \mathbb{R}$.

(a) $e^{ix} = \cos x + i \cdot \sin x$. *This is Euler's formula.*

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

(c) $\cos 0 = 1$ *and* $\sin 0 = 0$.

(d) $\cos' x = -\sin x$ *and* $\sin' x = \cos x$.

Proof.

(a) By the definition of cos and sin, we immediately get

$$\cos x + i \cdot \sin x = \frac{e^{ix} + e^{-ix}}{2} + i \cdot \frac{e^{ix} - e^{-ix}}{2} = e^{ix}.$$

(b) We note that

$$|e^{ix}|^2 = e^{ix} \cdot \overline{e^{ix}} = e^{ix} \cdot e^{-ix} = 1.$$

Therefore, $|e^{ix}| = 1$ for $x \in \mathbb{R}$. Then, we conclude

$$\cos^2 x + \sin^2 x = |e^{ix}|^2 = 1.$$

(c) By the complex exponential definition of cos and sin, we get $\cos 0 = 1$ and $\sin 0 = 0$.

(d) By the power series definition of cos and sin, we have

$$\cos' x = \sum_{n=1}^{\infty}(-1)^n \cdot 2n \cdot \frac{x^{2n-1}}{(2n)!} = \sum_{n=1}^{\infty}(-1)^n \cdot \frac{x^{2n-1}}{(2n-1)!} = -\sum_{n=0}^{\infty}(-1)^n \cdot \frac{x^{2n+1}}{(2n+1)!} = -\sin x$$

and

$$\sin' x = \sum_{n=0}^{\infty}(-1)^n \cdot (2n+1) \cdot \frac{x^{2n}}{(2n+1)!} = \sum_{n=0}^{|} infty(-1)^n \cdot \frac{x^{2n}}{(2n)!} = \cos x.$$

□

Now, we describe some other nice properties of cos and sin.

Theorem 4.13. *There exists a positive real number x_0 such that $\cos x_0 = 0$.*

Proof. Towards a contradiction, with $\cos 0 = 1$ and continuity of cos, assume that $\cos x > 0$ for all $x > 0$. On the other hand, since $\sin' x = \cos x > 0$ for $x > 0$, so $\sin x$ is strictly increasing. Since $\sin 0 = 0$, so $\sin x > 0$ for all $x > 0$. Also, $\cos' x = -\sin x < 0$ for $x > 0$, so $\cos x$ is strictly decreasing for $x > 0$. However, for $0 < x < y$, we have

$$\sin x \cdot (y - x) < \int_x^y \sin t \, dt = \cos x - \cos y < \cos x < 1. \tag{*}$$

Since $\sin x > 0$, then Equation * cannot hold for $y \to \infty$. Thus, we attain a contradiction. We conclude that there exists $x_0 > 0$ such that $\cos x_0 = 0$.

□

Let $E = \{x > 0 : \cos x = 0\}$, that is, E is the set of zeroes of cos on $(0, \infty)$. By Theorem 4.13, E is non-empty. Since cos is continuous on $(0, \infty)$ and $\{0\}$ is closed, so E is closed. Then, E has the greatest lower bound.

Definition 4.14. We define π to be the number

$$\pi = 2 \cdot \inf\{x > 0 : \cos x = 0\}.$$

Now, we conclude that

$$\cos\frac{\pi}{2} = 0 \text{ and } \sin\frac{\pi}{2} = 1.$$

Another set of special values happens at $x = \pi$. We can also define $\pi = \inf\{x > 0 : \sin x = 0\}$. Then, $\sin \pi = 0$ and $\cos \pi = -1$ by the properties of functions over $[0, \pi]$. How do cos and sin behave outside the interval $[0, \pi]$?

Theorem 4.15 (Periodicity of Trigonometric Functions)**.** *Let $z \in \mathbb{C}$.*

(a) e^z is $2\pi i$-periodic. In other words, $e^{z+2\pi i} = e^z$.

(b) cos and sin are 2π-periodic. In other words, $\cos(z + 2\pi) = \cos z$ and $\sin(z + 2\pi) = \sin z$.

(c) If $0 < t < 2\pi$, then $e^{it} \neq 1$.

(d) If $|z| = 1$, then there exists a unique $t \in [0, 2\pi)$ such that $z = e^{it}$.

Proof.

(a) By definition of π, we have $\cos\dfrac{\pi}{2} = 0$. Then, $\sin^2\dfrac{\pi}{2} = 1$, which implies $\sin\dfrac{\pi}{2} = \pm 1$. On the other hand, $\sin 0 = 0$ and $\sin' z = \cos z > 0$ for $0 < z < \dfrac{\pi}{2}$, so $\sin z$ is increasing over $0 < z < \dfrac{\pi}{2}$, and $\sin\dfrac{\pi}{2} = 1$. Now, we have

$$e^{i\cdot\pi/2} = \cos\frac{\pi}{2} + i\cdot\sin\frac{\pi}{2} = i.$$

Therefore,

$$e^{i\cdot 2\pi} = (e^{i\cdot\pi/2})^4 = i^4 = 1.$$

We conclude that $e^{z+2\pi i} = e^z \cdot e^{2\pi i} = e^z$.

(b) This is immediate from (a).

(c) Suppose $0 < t < \dfrac{\pi}{2}$ and $e^{it} = x + i\cdot y$ with $x, y \in \mathbb{R}$. By Euler's formula, we can let $x = \cos t$ and $y = \sin t$. For $0 < t < \dfrac{\pi}{2}$, $0 < \cos t < 1$ and $0 < \sin t < 1$. Then, $0 < x < 1$ and $0 < y < 1$. We note that

$$e^{4it} = (x + i\cdot y)^4 = x^4 - 6x^2y^2 + y^4 + 4ixy(x^2 - y^2).$$

Towards a contradiction, assume there is $0 < t < \dfrac{\pi}{2}$ such that $e^{4it} = 1$. This leads to $x^2 = y^2$. Since $x^2 + y^2 = 1$, then $x^2 = y^2 = \dfrac{1}{2}$, so $e^{4it} = -1$, which contradicts to our assumption $e^{4it} = 1$. Hence, $e^{4it} \neq 1$ for $0 < t < \dfrac{\pi}{2}$, which means $e^{it} \neq 1$ for $0 < t < \dfrac{\pi}{2}$. For the rest of cases, similar arguments give the same result.

(d) Towards a contradiction, there exists $0 \leq t_1 < t_2 < 2\pi$ such that $z = e^{it_1} = e^{it_2}$. Then,

$$|z|^2 = z\cdot\overline{z} = e^{it_1}\cdot e^{-it_2} = e^{i(t_1 - t_2)} \neq 1,$$

which contradicts to our assumption $|z| = 1$. This establishes the uniquess.

For existence, fix $z \in \mathbb{C}$ such that $|z| = 1$, and we can write $z = x + i\cdot y$ for $x, y \in \mathbb{R}$. Suppose $x \geq 0$ and $y \geq 0$. Then, for $0 < x < \dfrac{\pi}{2}$, $\cos x$ decreases from 1 to 0. Since $|z|^2 = x^2 + y^2$, so $0 \leq x \leq 1$. Then, there exists $0 \leq t \leq \dfrac{\pi}{2}$ such that $\cos t = x$ by Intermediate Value Theorem on $\cos t$ with $0 < t < \dfrac{\pi}{2}$. Since $\cos^2 t + \sin^2 t = 1 = x^2 + y^2$, so $\sin t = \pm y$. By assumption, $y > 0$, so $\sin t > 0$ on $0 < t < \dfrac{\pi}{2}$. So, $\sin t = y$. Therefore, there exists $0 \leq t \leq \dfrac{\pi}{2}$ such that $z = e^{it}$. For the other cases, (x, y) in other quadrants, that is, $x < 0$ and $y \geq 0$, $x < 0$ and $y < 0$, or $x > 0$ and $y < 0$, similar argument holds. Hence, we conclude that there exists a unique $t \in [0, 2\pi]$ such that $z = e^{it}$.

$\square$

Remark 4.16.

(a) The curve $\gamma(t) = e^{it}$ for $0 \leq t \leq 2\pi$ describes a unit circle.

(b) The length of a unit circle is given by

$$\Lambda(\gamma) = \int_0^{2\pi} \left|\cos^2 t + \sin^2 t\right|\, dt = 2\pi.$$

This is of course the expected result for the circumference of a circle of radius 1. This shows that π has the geometric significance, other than analytic significance in Definition 4.14.

(c) The functions cos and sin are defined geometrically in the followingn way. For $t \in [0, 2\pi)$, $(\cos t, \sin t)$ is the point on the unit circle such that the arclength of the circle curve from $(0, 0)$ to $(\cos t, \sin t)$ is t. This is how angles are defined: angle t means the arclength t on a unit circle.

4.4 The Algebraic Completeness of the Complex Field

Consider the function $f(x) = x^2 + 1$. In the real field $\mathbb{R}$, it does not have zeroes, which means there is no $x_0 \in \mathbb{R}$ such that $x_0^2 + 1 = 0$. So, $\mathbb{R}$ is not *algebraic complete*, that is to say, every nonconstant polynomial with real coefficients has a real root. In the complex field $\mathbb{C}$, indeed, we have $i^2 = -1$. Now we are going to prove $\mathbb{C}$ is *algebraic complete*.

Theorem 4.17 (Fundamental Theorem of Algebra)**.** *Let $a_0, a_1, \ldots, a_n \in \mathbb{C}$, $n \geq 1$, and $a_n \neq 0$. We define*

$$P(z) = \sum_{i=0}^{n} a_i \cdot z^i.$$

Then, there exists a $z_0 \in \mathbb{C}$ such that $P(z_0) = 0$.

Lecture 26 Monday March 8

Proof. Without loss of generality, we can assume that $a_n = 1$; otherwise, divide $P(z)$ by a_n. Let $\mu = \inf |P(z)|$. If $|z| = R$, then, by triangle inequality, we have

$$|P(z)| \geq R^n \cdot \left(1 - \frac{|a_{n-1}|}{R} - \frac{|a_{n-2}|}{R^2} - \cdots - \frac{|a_0|}{R^n}\right).$$

$|P(z)|$ tends to ∞ as $R \to \infty$. There exists $R_0 > 0$ such that $|P(z)| > \mu$ for $|z| > R_0$. Since $P(z)$ is continuous on

$$\overline{B_r(x_0)} = \{z \in \mathbb{C} : |z| \leq R_0\},$$

which is compact, so there exists $z_0 \in \mathbb{C}$ such that

$$|P(z_0)| = \mu = \min |P(z)|.$$

Now, it suffices to show $\mu = 0$, which implies $|P(z_0)| = 0$. Up to now, we have not used any property about $\mathbb{C}$. Towards a contradiction, assume $\mu \neq 0$. We define

$$\tilde{P}(z) = \frac{P(z + z_0)}{P(z_0)}.$$

Then, $\tilde{P}$ is a nonconstant polynomial, with $\tilde{P}(0) = 1$ and $|\tilde{P}(z)| \geq 1$ for all $z \in \mathbb{C}$. There exists a smallest integer $1 \leq k \leq n$ such that

$$\tilde{P}(z) = 1 + b_k z^k + \cdots + b_n z^n = \sum_{i=0}^{n} b_k z^k$$

where $\tilde{P}(0) = 1$ so $b_0 = 1$ and such k always exists since Q is nonconstant. By Theorem 4.12 (d), there exists $\theta \in \mathbb{R}$ such that

$$e^{ik\theta} = \frac{b_k}{|b_k|}.$$

Consider z of the form

$$z = r \cdot |b_k|^{-1/k} \cdot e^{i(\pi - \theta)/k},$$

with $r > 0$. For z of this form, we have

$$\tilde{P}(z) = 1 - r^k + r^{k+1} \cdot h(r),$$

where h is a polynomial. Then, for $r < 1$, by the triangle inequality, we have

$$|\tilde{P}| \leq 1 - r^k - r^{k+1} \cdot |h(r)|.$$

For $r > 0$ sufficiently small, we have $r \cdot |h(r)| < 1$, by the continuity of the function $r \cdot h(r)$ and the fact that it vanishes as $r = 0$. Hence,

$$|\tilde{P}(z)| \leq 1 - r^k \cdot (1 - r \cdot |h(r)|) < 1,$$

for some z in that particular form with $r \in (0, r_0)$ and $r_0 > 0$ sufficiently small, but then the minimum of the function $|\tilde{P}| : \mathbb{C} \to \mathbb{C}$ cannot possibly achieve 1, which contradicts to our assumption $|\tilde{P}(z)| \geq 1$ for all $z \in \mathbb{C}$. We conclude that $\mu = 0$ and $P(z_0) = 0$.

□

Corollary 4.18. *Let $a_0, a_1, \ldots, a_n \in \mathbb{C}$, $a_n \neq 0$, and $n \geq 1$. We define*

$$f(z) = \sum_{i=0}^{n} a_i \cdot z^i.$$

Then

(a) Given any complex number $w \in \mathbb{C}$, we have that $f(w) = 0$ if and only if there exists a polynomial $g : \mathbb{C} \to \mathbb{C}$ of degree $n - 1$ such that

$$f(z) = (z - w) \cdot g(z)$$

for $z \in \mathbb{C}$.

(b) There are at most n distinct complex numbers $w \in \mathbb{C}$ for which $f(w) = 0$. In other words, f has at most n distinct roots.

(c) (Restatement of the Fundamental Theorem of Algebra) There exists exactly $n + 1$ complex numbers $w_0, w_1, \ldots, w_n \in \mathbb{C}$ (not necessarily distinct) such that

$$f(z) = w_0 \cdot (z - w_1) \cdots (z - w_n)$$

for $z \in \mathbb{C}$. In other words, every polynomial function with coefficients over $\mathbb{C}$ can be factored into linear factors over $\mathbb{C}$.

Proof.

(a) Pick a $w \in \mathbb{C}$. Assume that $f(w) = 0$. Then, in particular, we have

$$a_n w^n + \cdots + a_1 w + a_0 = 0.$$

Given $z \in \mathbb{C}$, it follows that

$$\begin{aligned} f(z) &= a_n z^n + \cdots + a_1 z + a_0 - (a_n w^n + \cdots + a_1 w + a_0) \\ &= a_n(z^n - w^n) + a_{n-1}(z^{n-1} - w^{n-1}) + \cdots + a_1(z - w) \\ &= a_n(z - w) \sum_{k=0}^{n-1} z^k w^{n-1-k} + a_{n-2} \sum_{k=0}^{n-2} z^k w^{n-2-k} + \cdots + a_1(z - w) \\ &= (z - w) \sum_{i=1}^{n} a_i \sum_{k=0}^{i-1} z^k w^{i-k}. \end{aligned}$$

Thus, we define

$$g(z) = \sum_{i=1}^{n} a_i \sum_{k=0}^{i-1} z^k w^{i-k}$$

for $z \in \mathbb{C}$. We indeed construct a polynomial g of degree $n-1$ such that $f(z) = (z-w) \cdot g(z)$ for $z \in \mathbb{C}$.

Now, assume that there exists a polynomial $g : \mathbb{C} \to \mathbb{C}$ of degree $n-1$ such that

$$f(z) = (z-w) \cdot g(z)$$

for $z \in \mathbb{C}$. Then it follows that $f(w) = 0$ immediately.

(b) We prove by induction on the degree n of f.

For $n = 1$, then $f(z) = a_1 z + a_0$ is linear. The unique solution of $a_1 z + a_0 = 0$ is $z = -\dfrac{a_0}{a_1}$.

Assume that the result holds for $n-1$. In other words, assume that every polynomial of degree $n-1$ has at most $n-1$ roots. By Theorem 4.17, there exists a complex number $w \in \mathbb{C}$ such that $f(w) = 0$. Moreover, there exists a polynomial g of degree $n-1$ such that $f(z) = (z-w) \cdot g(z)$ for $z \in \mathbb{C}$. Then, it follows by the induction hypothesis that g has at most $n-1$ distinct roots, and so f has at most n distinct roots.

(c) This follows from an inductive argument on n that is almost identical to (b).

□

5 A Glimpse of Fourier Analysis

Fourier Analysis provides a collection of techniques for resolving general functions into sums or integrals of simple functions or functions with certain special properties. This is a powerful tool for solving various differential equations of interest in science and engineering.

The most fundamental of these applications is the expansion of periodic functions. We will devote our time to the study of periodic functions, such as the trigonometric functions. In many respects, it is simpler and neater to work with the complex exponential function $e^{i\theta}$ instead of the trigonometric functions $\cos\theta$ and $\sin\theta$. We recall these funcitons are related by

$$\cos\theta = \frac{e^{i\theta}+e^{-i\theta}}{2} \text{ and } \sin\theta = \frac{e^{i\theta}-e^{-i\theta}}{2i}$$

and the famous Euler's formula $e^{i\theta} = \cos\theta + i\sin\theta$.

5.1 The Fourier Series of a Periodic Function

Previously, we define 2π-*periodic function* as the function defined on $\mathbb{R}$ such that $f(x+2\pi) = f(x)$ for all x. We assume that f is *Riemann integrable* on every *bounded* interval. This will be the case if f is bounded and is continuous except perhaps at finitely many points in each bounded interval. We wish to know if f can be expanded in the form of trigonometric series defined as below.

Definition 5.1 (Trigonometric Polynomial)**.** A *trigonometric polynomial* is a finite sum of the form

$$T_N(x) = a_0 + \sum_{n=1}^{N}(a_n\cos nx + b_n\sin nx)$$

for $x \in \mathbb{R}$ and $a_0, a_1, \ldots, a_N, b_1, \ldots, b_N \in \mathbb{C}$.

By the definition of sin and cos, we get

$$\begin{aligned}
f(x) &= a_0 + \sum_{n=1}^{N}\left(a_n \cdot \frac{e^{inx}+e^{-inx}}{2} + b_n \cdot \frac{e^{inx}-e^{-inx}}{2i}\right) \\
&= a_0 + \sum_{n=1}^{N}\left(\frac{i\cdot a_n + b_n}{2i}\cdot e^{inx} + \frac{i\cdot a_n - b_n}{2i}\cdot e^{-inx}\right) \\
&= a_0 + \sum_{n=1}^{N}\left(\frac{a_n - i\cdot b_n}{2}\cdot e^{inx} + \frac{a_n + i\cdot b_n}{2}\cdot e^{-inx}\right).
\end{aligned}$$

Then, we can rewritten the trigonometric polynomial $T_N(x)$ as

$$T_N(x) = \sum_{n=-N}^{N} c_n \cdot e^{inx}$$

with $c_0 = a_0$, $c_n = \dfrac{a_n - i\cdot b_n}{2}$, and $c_{-n} = \dfrac{a_n + i\cdot b_n}{2}$ for $n \geq 1$. Alternatively, if we start out with

$$f(x) = \sum_{n=-N}^{N} c_n \cdot e^{inx},$$

then we can verify $a_0 = c_0$, $a_n = c_n + c_{-n}$, and $b_n = i\cdot(c_n - c_{-n})$ for $n \geq 1$. Further, it is clear that the trigonometric polynomial is periodic, with period 2π.

Proposition 5.2. *A trigonometric polynomial* $T_N(x) = \sum_{n=-N}^{N} c_n e^{inx}$ *is real-valued for* $x \in \mathbb{R}$ *if and only if* $c_{-m} = \overline{c_m}$ *for* $m = -N, \ldots, N$. *In this case, we can write*

$$T_N(x) = a_0 + \sum_{n=1}^{N} (a_n \cos nx + b_n \sin nx)$$

where $a_n = 2\operatorname{Re}(c_n)$ *and* $b_n = 2\operatorname{Im}(c_n)$.

Definition 5.3 (Trigonometric Series)**.** The *trigonometric series* is a series in the form

$$f(x) = a_0 + \sum_{n=1}^{\infty} (a_n \cdot \cos nx + b_n \cdot \sin nx) \tag{5.1}$$

or

$$f(x) = \sum_{n=-\infty}^{\infty} c_n \cdot e^{inx} \tag{5.2}$$

for $x \in \mathbb{R}$.

We have seen that analytic functions can be represented as power series

$$f(x) = \sum_{n=0}^{\infty} a_n x^n,$$

and every continuous function can be uniformly approximated with sequences of polynomials. Can we express any continuous function f as a trigonometric series? Which functions can be expressed as trigonometric series?

Lemma 5.4. *If* $f(x) = \sum_{n=-N}^{N} c_n \cdot e^{inx}$, *then*

$$c_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) \cdot e^{-inx}\, dx$$

for every $-N \le n \le N$.

Proof. We multiply both sides of $f(x) = \sum_{n=-N}^{N} c_n \cdot e^{inx}$ by e^{-ikx} and integrate $-\pi$ to π. Then, we have

$$\int_{-\pi}^{\pi} f(x) \cdot e^{-ikx}\, dx = \sum_{n=-N}^{N} \int_{-\pi}^{\pi} c_n \cdot e^{i(n-k)x}\, dx = \sum_{n=-N}^{N} c_n \int_{-\pi}^{\pi} e^{i(n-k)x}\, dx.$$

Recall that e^{inx} is 2π-periodic, so if $n \neq k$,

$$\int_{-\pi}^{\pi} e^{i(n-k)x}\, dx = \frac{1}{i(n-k)} \cdot e^{i(n-k)x}\Big|_{-\pi}^{\pi} = 0.$$

Otherwise, if $n = k$, we have

$$\int_{-\pi}^{\pi} e^{i(n-k)x}\, dx = \int_{-\pi}^{\pi} 1\, dx = 2\pi.$$

Hence the only term in the series that survives the integration is the term with $n = k$. Therefore,

$$\int_{-\pi}^{\pi} f(x) \cdot e^{-ikx}\, dx = 2\pi \cdot c_k.$$

In other words, relabeling the integer k as n, we get the desired formula for c_n, that is,

$$c_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) \cdot e^{-inx}\, dx.$$

$\square$

If f has a series expansion of the form Equation 5.1 or Equation 5.2, and the series converges decently so that term-by-term integration is permissible, then the coefficients a_n and b_n (or c_n) are given by

$$a_0 = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x)\, dx$$

and for $n \geq 1$,

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cdot \cos nx\, dx \text{ and } b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cdot \sin nx\, dx \tag{5.3}$$

(or in Lemma 5.4). Now, if f is any Riemann integrable periodic function, the integrals to define a_n, b_n, and c_n make sense.

Definition 5.5 (Fourier Series)**.** If $f : [-\pi, \pi] \to \mathbb{R}$ is Riemann integrable, then we define the *Fourier coefficients* of f by

$$c_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) \cdot e^{-inx}\, dx$$

for $n \in \mathbb{Z}$. We define the *Fourier series* of f by

$$\sum_{n=-\infty}^{\infty} c_n \cdot e^{inx}.$$

Does $\displaystyle\sum_{n=-\infty}^{\infty} c_n e^{inx}$ converge? Does it converge to f? Lecture 27 Wednesday March 10

5.2 Orthogonal and Orthonormal Sets in Inner Product Spaces

In the Euclidean space $\mathbb{R}^k$, we encounters the *inner product* or (*dot product*) defined by

$$(x, y) = \sum_{i=1}^{k} x_i \cdot y_i$$

where $x = (x_1, \ldots, x_k) \in \mathbb{R}^k$ and $y = (y_1, \ldots, y_k) \in \mathbb{R}^k$. The *norm* of a vector $x \in \mathbb{R}^k$ is related to the inner product by the equation

$$\|x\| = \sqrt{(x, x)}.$$

Definition 5.6 (Inner Product Space)**.** Let V be a vector space. An *inner product* for V is a function $(\ ,\ ) : V \times V \to \mathbb{R}$ which satisfies the following

(a) $(cx + y, z) = c(x, z) + (y, z)$ for all $c \in \mathbb{R}$ and all $x, y, z \in V$.

(b) $(x, y) = (y, x)$ for all $x, y \in V$.

(c) $(x, x) \geq 0$ for all $x \in V$.

(d) If $(x, x) = 0$, then $x = 0$.

The ordered pair $(V, (\ ,\ ))$ (usually denoted by just V) is called an *inner product space.*

Theorem 5.7 (Examples of Inner Product Space)**.**

(a) Let $x = (x_n)_{n\in\mathbb{N}}, y = (y_n)_{n\in\mathbb{N}} \in L^2$. The equation

$$(x, y) = \sum_{n=1}^{\infty} x_n \cdot y_n$$

defines an inner product of L^2.

(b) Let $f, g \in \mathscr{R}$ on $[a, b]$. The equation

$$(f, g) = \int_a^b f \cdot g \, dx$$

defines an inner product on $\mathscr{R}_a^b$ with the exception that property (d) in Definition 5.6 is replaced by $f = 0$ almost everywhere (defined in 140C) in $[a, b]$ if $(f, f) = 0$.

Now, we study a more general series for a function. One fundamental property of the exponentials that make Fourier series what it is that the exponentials are a so-called *orthonormal system.*

Definition 5.8 (Orthogonal and Orthonormal Systems)**.** Let $(\phi_n)_{n\in\mathbb{N}}$ be a sequence of complex functions on $[a, b]$ such that

$$\int_a^b \phi_n(x) \cdot \overline{\phi_m(x)} \, dx = 0$$

for $n \neq m$. Then (ϕ_n) is called an *orthogonal system.* If, moreover,

$$\int_a^b |\phi_n|^2 \, dx = \int_a^b \phi_n \cdot \overline{\phi_n} \, dx = 1,$$

then it is called an *orthonormal system.*

Example 5.9 (Examples of Orthonormal Sets)**.**

(a) In $\mathbb{R}^n$, we define the standard basis e_i where the ith entry is 1 and all others are 0. Then, $\{e_1, \ldots, e_n\}$ form an orthonormal base of $\mathbb{R}^n$. Indeed, for $1 \leq i, j \leq n$, we have

$$(e_i, e_j) = e_i \cdot e_j = \begin{cases} 0 & i \neq j \\ 1 & i = j. \end{cases}$$

(b) Let

$$\phi_0(x) = \frac{1}{\sqrt{2\pi}}, \qquad \phi_{2n-1}(x) = \frac{\cos nx}{\sqrt{\pi}}, \qquad \phi_{2n}(x) = \frac{\sin nx}{\sqrt{\pi}}$$

for $x \in \mathbb{R}$ and $n \in \mathbb{N}$. Then, $\{\phi_0, \phi_1, \ldots\}$ form an orthogonal system in $[t, t + 2\pi]$ for $t \in \mathbb{R}$.

We verify the special case $t = 0$. First, we need to show that any two distinct vectors in the system are orthogonal. We note that

$$(\phi_0, \phi_{2n-1}) = \int_0^{2\pi} \frac{\cos nx}{\sqrt{2\pi}} \, dx = \frac{\sin nx}{n \cdot \sqrt{2\pi}} \Big|_0^{2\pi} = 0.$$

Similarly, $(\phi_0, \phi_{2n}) = 0$. If $n = m$, we have

$$(\phi_{2n-1}, \phi_{2m}) = \int_0^{2\pi} \frac{\cos nx \cdot \sin nx}{\pi} \, dx = -\frac{1}{2\pi} \frac{\cos 2nx}{2n} \Big|_0^{2\pi} = 0.$$

If $n \neq m$, we compute

$$(\phi_{2n-1}, \phi_{2m}) = \int_0^{2\pi} \frac{\cos nx \cdot \sin mx}{\pi} \, dx = -\frac{1}{2\pi} \cdot \left[\frac{\cos(m+n)x}{m+n} - \frac{\cos(m-n)x}{m-n} \right]_0^{2\pi} = 0.$$

Similarly, we have

$$(\phi_{2n-1}, \phi_{2m-1}) = (\phi_{2n}, \phi_{2m}) = 0$$

for $n \neq m$. Now, for the additional part, we have $\|\phi_0\|^2 = \int_0^{2\pi} \left(\frac{1}{\sqrt{2\pi}} \right)^2 \, dx = 1$ immediately and

$$\|\phi_{2n-1}\|^2 = \int_0^{2\pi} \frac{\cos^2 nx}{\pi} \, dx = \frac{1}{\pi} \int_0^{2\pi} \frac{1 + \cos 2nx}{2} \, dx = \frac{1}{\pi} \cdot \left[\frac{1}{2}x + \frac{\sin nx \cdot \cos nx}{2n} \right]_0^{2\pi} = 1.$$

Similarly, we have $\|\phi_{2n}\|^2 = 1$ as well.

Having an orthonormal system (ϕ_n) on $[a, b]$ and a Riemann integrable function f on $[a, b]$, we can write a Fourier series relative to (ϕ_n). We define

$$c_n = (f, \phi_n) = \int_a^b f(t) \cdot \overline{\phi_n(t)} \, dt$$

and write $f(x) \sim \sum_{n=1}^{\infty} c_n \phi_n$. In other words, the series is $\sum_{n=1}^{\infty} (f, \phi_n) \cdot \phi_n(x)$.

Theorem 5.10. *Let (ϕ_n) be an orthonormal system on $[a, b]$. Suppose f is a Riemann integrable function on $[a, b]$ and*

$$f(x) = \sum_{n=1}^{\infty} c_n \cdot \phi_n(x).$$

Let

$$s_n(x) = \sum_{k=1}^{n} c_k \cdot \phi_k(x)$$

be the nth partial sum of the Fourier series of f and

$$t_n(x) = \sum_{k=1}^{n} \gamma_k \cdot \phi_k(x)$$

for some other sequence (γ_k). Then

$$\int_a^b |f - s_n|^2 \, dx \leq \int_a^b |f - t_n|^2 \, dx$$

with equality holds if and only if $c_k = \gamma_k$ for $k = 1, 2, \ldots, n$.

In other words, the partial sums of the Fourier series are the best approximation with respect to the L^2 norm.

Proof. We have

$$\int_a^b |f - t_n|^2 \, dx = \int_a^b |f|^2 \, dx - \int_a^b f \cdot \overline{t_n} \, dx - \int_a^b \overline{f} \cdot t_n \, dx + \int_a^b |t_n|^2 \, dx.$$

Now, we compute

$$\int_a^b f\cdot\overline{t_n}\,dx=\int_a^b f\cdot\sum_{k=1}^{n}\overline{\gamma_k}\cdot\overline{\phi_k}\,dx=\sum_{k=1}^{n}\overline{\gamma_k}\int_a^b f\cdot\overline{\phi_k}\,dx=\sum_{k=1}^{n}\overline{\gamma_k}\cdot c_k,$$

$$\int_a^b |t_n|^2\,dx=\int_a^b\sum_{k=1}^{n}\gamma_k\cdot\phi_k\cdot\sum_{k=1}^{n}\overline{\gamma_k}\cdot\overline{\phi_k}=\sum_{k=1}^{n}\gamma_k\cdot\overline{\gamma_k}=\sum_{k=1}^{n}|\gamma_k|^2,$$

and

$$\int_a^b f\cdot\overline{s_n}\,dx=\sum_{k=1}^{n}|c_k|^2=\int_a^b|s_n|^2\,dx.$$

Since (ϕ_n) is orthonormal, so

$$\begin{aligned}\int_a^b|f-t_n|^2\,dx&=\int_a^b|f|^2\,dx+\int_a^b|t_n|^2\,dx-\int_a^b f\cdot\overline{t_n}\,dx-\int_a^b\overline{f}\cdot t_n\,dx\\&=\int_a^b|f|^2\,dx+\sum_{k=1}^{n}|\gamma_k|^2-\sum_{k=1}^{n}\gamma_k\cdot\overline{c_k}-\sum_{k=1}^{n}\overline{\gamma_k}\cdot c_k\\&=\int_a^b|f|^2\,dx-\sum_{k=1}^{n}|c_k|^2+\sum_{k=1}^{n}|c_k-\gamma_k|^2\end{aligned}$$

and

$$\int_a^b|f-s_n|^2\,dx=\int_a^b|f|^2\,dx-\sum_{k=1}^{n}|c_k|^2+\sum_{k=1}^{n}|c_n-c_n|^2=\int_a^b|f|^2\,dx-\sum_{k=1}^{n}|c_k|^2.$$

Therefore,

$$\int_a^b|f-s_n|^2\,dx\le\int_a^b|f-t_n|^2\,dx$$

and the equality holds if and only if $\sum_{k=1}^{n}|c_k-\gamma_k|^2=0$ if and only if $c_k=\gamma_k$ for $1\le k\le n$.

□

Theorem 5.11 (Bessel's Inequality). *Suppose f is a Riemann integrable function on $[a,b]$. Let (ϕ_n) be an orthonormal system on $[a,b]$ and suppose $f(x)\sim\sum_{n=1}^{\infty}c_n\cdot\phi_n(x)$ where $c_n=\int_a^b f\cdot\overline{\phi_n}\,dx$. Then*

$$\sum_{n=1}^{\infty}|c_n|^2\le\int_a^b|f|^2\,dx.$$

In particular, $\lim_{n\to\infty}c_n=0$.

Proof. From Theorem 5.10, we have

$$\sum_{k=1}^{n}|c_k|^2=\int_a^b|s_n|^2\,dx=\int_a^b|f|^2\,dx-\int_a^b|f-s_n|^2\,dx\le\int_a^b|f|^2\,dx$$

for $n\in\mathbb{N}$. The sequence $\left(\sum_{k=1}^{n}|c_k|^2\right)_n$ is monotone increasing and bounded above, so there exists $\lim_{n\to\infty}\sum_{k=1}^{n}|c_n|^2$ bounded by $\int_a^b|f|^2\,dx$. So, it is immediate that $|c_n|^2\to 0$ and $c_n\to 0$ as $n\to\infty$.

□

Bessel's inequality can also be stated in terms of the coefficients a_n and b_n defined by Equation 5.3. Then, we have

$$\begin{aligned} |a_n|^2 + |b_n|^2 &= a_n \cdot \overline{a_n} + b_n \cdot \overline{b_n} \\ &= (c_n + c_{-n}) \cdot (\overline{c_n} + \overline{c_{-n}}) + i(c_n - c_{-n}) \cdot (-i)(\overline{c_n} - \overline{c_n}) \\ &= 2c_n\overline{c_n} + 2c_{-n}\overline{c_{-n}} \end{aligned}$$

so that $|a_0|^2 = 4|c_0|^2$ and $|a_n|^2 + |b_n|^2 = 2\left(|c_n|^2 + |c_{-n}|^2\right)$ for $n \geq 1$. Therefore,

$$\frac{1}{4}|a_0|^2 + \frac{1}{2}\sum_{n=1}^{\infty}\left(|a_n|^2 + |b_n|^2\right) = \sum_{n=-\infty}^{\infty} |c_n|^2 \leq \frac{1}{2\pi}\int_{-\pi}^{\pi} |f(x)|^2 \, dx.$$

Theorem 5.12 (Riemann-Lebesgue Lemma)**.** *Let $f : [-\pi, \pi] \to \mathbb{R}$ be Riemann integrable with period 2π. Then*

$$\lim_{n\to\infty} \int_{-\pi}^{\pi} f(x) \cdot \sin nx \, dx = \lim_{n\to\infty} \int_{-\pi}^{\pi} f(x) \cdot \cos nx \, dx = 0.$$

In other words, this says that the Fourier coefficients a_n, b_n, and c_n all tend to 0 as $n \to \infty$ (and as $n \to -\infty$ in the case of c_n).

Lecture 28
Friday
March 12

5.3 Pointwise Convergence of Fourier Series

Let V be the inner product space of Riemann integrable functions on $[-\pi, \pi]$. Now, we start to specialize the general theory of Fourier series to the inner product space V by considering the problem of when the Fourier series of a function f in V converges *pointwise* to f. To show that the Fourier series of a function in V converges at x, we must show that the limit of the Nth partial sum of the Fourier series of f at x,

$$s_N(x) = \sum_{n=-N}^{N} e^{inx}$$

exists. By Bessel's inequality,

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} |s_N(x)|^2 \, dx = \sum_{n=-N}^{N} |c_n|^2 \leq \frac{1}{2\pi}\int_{-\pi}^{\pi} |f(x)|^2 \, dx.$$

Definition 5.13 (Dirichlet Kernel)**.** The *Dirichlet kernel* D_n is defined by

$$D_N(x) = \sum_{n=-N}^{N} e^{inx}. \tag{5.4}$$

We derive a more explicit formula for the Dirichlet kernel and its properties.

Corollary 5.14. $D_N(x) = \dfrac{\sin(N + 1/2)x}{\sin(x/2)}$.

Proof. We multiply both sides of Equation 5.4 by $e^{ix} - 1$

$$(e^{ix} - 1) \cdot D_N(x) = \sum_{n=-N}^{N} (e^{i(n+1)x} - e^{inx}).$$

Then, we multiply both sides by $e^{-ix/2}$

$$(e^{ix/2} - e^{-ix/2}) \cdot D_N(x) = \sum_{n=-N}^{N} (e^{i(n+1/2)x} - e^{i(n-1/2)x}).$$

Further, we multiply both sides by $\frac{1}{2i}$

$$\frac{1}{2i}(e^{ix/2} - e^{-ix/2}) \cdot D_N(x) = \frac{1}{2i} \sum_{n=-N}^{N} (e^{i(n+1/2)x} - e^{i(n-1/2)x}) = \frac{1}{2i}(e^{i(N+1/2)x} - e^{-i(N+1/2)x}).$$

Recall the complex exponential definition of sin, we conclude that

$$D_N(x) = \frac{\sin(N+1/2)x}{\sin(x/2)}.$$

□

Corollary 5.15. *For any $N \in \mathbb{N}$,*

$$\int_{-\pi}^{0} D_N(x)\,dx = \int_{0}^{\pi} D_N(x)\,dx = \pi.$$

Proof. By the definition of Dirichlet kernel, we have

$$D_N(x) = \sum_{n=-N}^{N} e^{inx} = 1 + \sum_{n=1}^{N} \cos nx,$$

so that

$$\int_0^\pi D_N(x)\,dx = \left[x + \sum_{n=1}^{N} \frac{1}{n} \cdot \sin nx\right]_0^\pi = \pi$$

and likewise for the integral from $-\pi$ to 0. Finally, we have

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} D_N(x)\,dx = 1.$$

□

Lemma 5.16. *Let $f : [-\pi, \pi] \to \mathbb{R}$ be Riemann integrable with period 2π. Again, let s_N be the Nth partial sum of the Fourier series of f. Then*

$$s_N(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x-t) \cdot D_N(t)\,dt.$$

Proof. By Lemma 5.4, we have

$$\begin{aligned}
s_N(x) &= \sum_{n=-N}^{N} c_n e^{inx} \\
&= \sum_{n=-N}^{N} \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t) \cdot e^{-int}\,dx \cdot e^{inx} \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t) \cdot \sum_{n=-N}^{N} e^{in(x-t)}\,dt \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x) \cdot D_N(x-t)\,dt \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x-t) \cdot D_N(t)\,dt.
\end{aligned}$$

□

Now, we can derive the sufficient condition for pointwise convergence of Fourier series.

Theorem 5.17. *Let f be continuous on $[-\pi, \pi]$ with period 2π. For $x \in (-\pi, \pi)$, if there exists a constant $\delta > 0$ and $M < \infty$ such that*

$$|f(x+t) - f(x)| < M|t|$$

for all $t \in (-\delta, \delta)$. Then,

$$\lim_{N\to\infty} s_N(x) = f(x).$$

In other words, if f satisfies a Lipschitz condition at x, then the Fourier series of f converges to f at x.

Proof. We have

$$\begin{aligned}
s_N(x) - f(x) &= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x-t)\cdot D_N(t)\,dt - f(x) \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x-t)\cdot D_N(t)\,dt - f(x)\cdot\frac{1}{2\pi}\int_{-\pi}^{\pi} D_N(t)\,dt \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} (f(x-t) - f(x))\cdot D_N(t)\,dt \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{f(x-t) - f(x)}{\sin(t/2)}\cdot\sin\left(N + \frac{1}{2}\right)t\,dt.
\end{aligned}$$

Define $g(t) = \dfrac{f(x-t) - f(x)}{\sin(t/2)}$ for $0 < |t| \le \pi$ and $g(0) = 0$. Then, with the identity

$$\sin(u+v) = \sin u\cdot\cos v + \cos u\cdot\sin v,$$

we have

$$\begin{aligned}
s_N(x) - f(x) &= \frac{1}{2\pi}\int_{-\pi}^{\pi} g(t)\cdot\left(\cos\frac{t}{2}\cdot\sin Nt + \cos Nt\cdot\sin\frac{t}{2}\right)\,dt \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} g(t)\cdot\cos\frac{t}{2}\cdot\sin Nt\,dt + \frac{1}{2\pi}\int_{-\pi}^{\pi} g(t)\cdot\sin\frac{t}{2}\cdot\cos Nt\,dt.
\end{aligned}$$

Now, it suffices to show that as $N \to \infty$,

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} g(t)\cdot\cos\frac{t}{2}\cdot\sin Nt\,dt \to 0 \text{ and } \frac{1}{2\pi}\int_{-\pi}^{\pi} g(t)\cdot\sin\frac{t}{2}\cdot\cos Nt\,dt \to 0.$$

We define

$$a_N = \frac{1}{2\pi}\int_{-\pi}^{\pi} g(t)\cdot\cos\frac{t}{2}\cdot\sin Nt\,dt \text{ and } b_N = \frac{1}{2\pi}\int_{-\pi}^{\pi} g(t)\cdot\sin\frac{t}{2}\cdot\cos Nt\,dt.$$

By Bessel's inequality, to show that $\lim_{N\to\infty} a_N = 0 = \lim_{N\to\infty} b_N$, it suffices to show that $g(t)\cdot\cos\frac{t}{2}$ and $g(t)\cdot\sin\frac{t}{2}$ are Riemann integrable on $[-\pi, \pi]$. By Taylor's Theorem, with sufficiently small δ, for $0 < |t| < \delta$, we have

$$\left|\sin\frac{t}{2}\right| \ge \frac{1}{4}|t|.$$

We note that

$$|g(t)| = \left|\frac{f(x-t) - f(x)}{\sin(t/2)}\right| = \left|\frac{f(x-t) - f(x)}{t}\cdot\frac{t}{\sin(t/2)}\right| \le 4M,$$

so g is bounded on $[-\pi, \pi]$ and continuous except at 0. This concludes Riemann integrability of $g(t)\cdot\cos\frac{t}{2}$ and $g(t)\cdot\sin\frac{t}{2}$.

□

Corollary 5.18. *Let f be continuous on $[-\pi, \pi]$. For $x \in (-\pi, \pi)$, if f is differentiable at x, then the Fourier series of f converges to f at x.*

This follows from the fact that f is Lipschitz at x if f is differentiable at x and Theorem 5.17.

Corollary 5.19 (Riemann Localization Theorem)**.** *Let f and g be a 2π-periodic and continuous function on $[-\pi, \pi]$. If $f(t) = 0$ for all t in some neighborhood of x, then $\lim_{N\to\infty} s_N(x) = 0$. In particular, if $f(t) = g(t)$ for all t in the neighborhood of x, then*

$$s_N(x; f) - s_N(x; g) = s_N(x; f - g) \to 0$$

as $N \to \infty$.

In other words, convergence of Fourier series of f at x only depends on the values of the function near x.

Remark 5.20. Both Fourier series and Taylor series (or power series) can represent a function in a different form, but there exists a huge difference. To make sure a function f has a Fourier series, by Theorem 5.17, f needs to be Lipschitz. For f to admit a representation of power series, f requires to be *smooth*, that is infinitely differentiable, which is a much stronger and strict requirement. For example, $f(x) = |x|$ is Lipschitz but not analytic, so $f(x) = |x|$ admits a representation in Fourier series, that is, for $x \in (-\pi, \pi)$,

$$|x| \sim \frac{\pi}{2} - \frac{4}{\pi} \sum_{n=1}^{\infty} \frac{\cos(2n-1)x}{(2n-1)^2}.$$

Theorem 5.21. *If $f : \mathbb{R} \to \mathbb{C}$ is continuous and 2π-periodic, then for every $\varepsilon > 0$, there exists a trigonometric polynomial P such that*

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

for all $x \in \mathbb{R}$.

In other words, given any continuous, periodic, complex-valued function f, there exists a sequence of trigonometric polynomials that converges uniformly to f, and note that P_n is not the Fourier series of f. This is a direct application of the Stone-Weierstrass Theorem for complex-valued functions.

Proof. We define $T = \{z \in \mathbb{C} : |z| = 1\}$, which is compact. For $z \in T$, we pick an $x \in \mathbb{R}$ such that $e^{ix} = z$ and let $g(z) = f(x)$ (note that if $e^{ix} = z = e^{it}$, then $x - t = k \cdot 2\pi$ for $k \in \mathbb{Z}$ and hence $f(x) = f(t)$).

Let $\mathscr{A}$ be the set of all functions

$$\sum_{n=-N}^{N} c_n \cdot z^n$$

on T. Then, $\mathscr{A}$ is a self-adjoint algebra, that is, for every trinometric polynomial $T_N(z) \in \mathscr{A}$, there is a trigonometric polynomial $\overline{T_N(z)} \in \mathscr{A}$, that separates points and vanishes nowhere.

By Stone-Weierstrass Theorem, there exists $P \in \mathscr{A}$ such that

$$|P(z) - g(z)| < \varepsilon$$

for every $z \in T$. Now, note that $P(e^{ix})$ is a trigonometric polynomial and for any $x \in \mathbb{R}$, setting $z = e^{ix}$, we have

$$|P(e^{ix}) - f(x)| = |P(e^{ix}) - g(e^{ix})| < \varepsilon.$$

Therefore, the family of trigonometric polynomials is dense in the space of the complex exponentials.

□

Finally, convergence always happen in the L^2 space and that is the reason why we introduce and focus on the L^2 space and orthonormal systems.

Theorem 5.22. *Let f and g be Riemann integrable functions on $[-\pi, \pi]$ with period 2π. Suppose*

$$s(x; f) = \sum_{n=-\infty}^{\infty} c_n \cdot e^{inx} \text{ for } c_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) \cdot e^{-inx}\, dx$$

and

$$s(x; g) = \sum_{n=-\infty}^{\infty} \gamma_n \cdot e^{inx} \text{ for } \gamma_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} g(x) \cdot e^{-inx}\, dx$$

Let $s_N(x)$ be the Nth partial sum of the Fourier series of f. Then

(a) $\displaystyle\lim_{N\to\infty} \int_{-\pi}^{\pi} |f(x) - s_N(x; f)|^2\, dx = 0.$

(b) $\displaystyle\frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) \cdot \overline{g(x)}\, dx = \sum_{n=-\infty}^{\infty} c_n \cdot \overline{\gamma_n}.$

(c) (Parseval) $\displaystyle\frac{1}{2\pi} \int_{-\pi}^{\pi} |f(x)|^2\, dx = \sum_{n=-\infty}^{\infty} |c_n|^2.$

Proof.

(a) For Riemann integrable f over $[a, b]$, we define

$$\|f\|_2 = \sqrt{\frac{1}{2\pi} \int_{-\pi}^{\pi} |f(x)|^2\, dx}.$$

Fix $\varepsilon > 0$. Since f is Riemann integrable and 2π-periodic, there is a continuous and 2π-periodic function h such that

$$\|f - h\|_2 < \varepsilon.$$

By Theorem 5.21, there is a trigonometric polynomial P such that

$$|h(x) - P(x)| < \varepsilon$$

for $x \in [-\pi, \pi]$. In particular,

$$\|h - P\|_2 < \sqrt{\frac{1}{2\pi} \int_{-\pi}^{\pi} \varepsilon^2\, dx} = \varepsilon.$$

Let N_0 be the degree of P, then by Theorem 5.10,

$$\|h - s_N(x; h)\|_2 \le \|h - P\|_2 < \varepsilon$$

for all $N \ge N_0$. Also, by Theorem 5.10 and Theorem 5.11,

$$\|s_N(x; h) - s_N(x)\|_2 = \|s_N(x; h - f)\|_2 = \|h - f\|_2 < \varepsilon.$$

By triangle inequality for $\|\cdot\|_2$, for $N \ge N_0$,

$$\|f - s_N(x; f)\|_2 \le \|f - h\|_2 + \|h - s_N(x; h)\|_2 + \|s_N(x; h) - s_N(x)\|_2 < 3\varepsilon.$$

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

$$\lim_{N\to\infty} \frac{1}{2\pi} \int_{-\pi}^{\pi} |f - s_N(x)|^2\, dx = \lim_{N\to\infty} \frac{1}{2\pi} \int_{-\pi}^{\pi} \|f - s_N(x)\|_2^2\, dx = 0.$$

(b) We note that

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} s_N(x;f)\cdot\overline{g(x)}\,dx = -\sum_{n=-N}^{N} c_n\cdot\frac{1}{2\pi}\int_{-\pi}^{\pi} e^{inx}\cdot\overline{g(x)}\,dx = \sum_{n=-N}^{N} c_n\cdot\overline{\gamma_n}. \tag{5.5}$$

Moreover,

$$\begin{aligned}\left|\int_{-\pi}^{\pi} f(x)\cdot\overline{g(x)}\,dx - \int_{-\pi}^{\pi} s_N(x)\cdot\overline{g(x)}\,dx\right| &\leq \int_{-\pi}^{\pi}|f-s_N(x)|\cdot|g(x)|\,dx\\ &\leq\sqrt{\int_{-\pi}^{\pi}|f-s_N(x)|^2\,dx\cdot\int_{-\pi}^{\pi}|g(x)|^2\,dx}.\end{aligned}$$

This goes to 0 as $N\to\infty$. Therefore,

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} s_N(x)\cdot\overline{g(x)}\,dx \to \int_{-\pi}^{\pi} f(x)\cdot\overline{g(x)}\,dx$$

as $N\to\infty$. By Equation 5.5, we conclude that

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)\cdot\overline{g(x)}\,dx \to \sum_{n=-N}^{N} c_n\cdot\overline{\gamma_n}$$

as $N\to\infty$.

(c) This is immediate by setting $f=g$ in (b).

□

This is the end of 140B.