MTH5105 Differential and Integral Analysis

Lecture Notes 2010-2011

Thomas Prellberg

Version of March 31, 2011

Contents

0 Revision

Lecture 1: 10/01/11

Let $\mathcal{D} \subseteq \mathbb{R}$ be a domain (e.g. interval or all of $\mathbb{R}$).

Definition 0.1. *Let $f : \mathcal{D} \to \mathbb{R}$.*

(a) f is continuous at $a \in \mathcal{D}$ if

$$\forall \varepsilon > 0\ \exists \delta > 0\ \forall x \in \mathcal{D},\ |x - a| < \delta : |f(x) - f(a)| < \varepsilon\ .$$

(b) f is continuous if f is continuous at all $a \in \mathcal{D}$.

(c) $f(x)$ tends to the limit $L \in \mathbb{R}$ as x tends to $a \in \mathcal{D}$, $\lim_{x \to a} f(x) = L$, if

$$\forall \varepsilon > 0\ \exists \delta > 0\ \forall x \in \mathcal{D},\ 0 < |x - a| < \delta : |f(x) - L| < \varepsilon\ .$$

Remark. We use the short-hand notation $\lim_{x \to a} f(x) = f(a)$ to indicate that both (a) $\lim_{x \to a} f(x) = L$ exists and (b) $f(a) = L$.

Theorem 0.2. *Let $f : \mathcal{D} \to \mathbb{R}$. f is continuous at $a \in \mathcal{D}$ if and only if $\lim_{x \to a} f(x) = f(a)$.*

Proof. Let $f : \mathcal{D} \to \mathbb{R}$.

"$\Rightarrow$" Let f be continuous at $a \in \mathcal{D}$. Then

$$\forall \varepsilon > 0\ \exists \delta > 0\ \forall x \in \mathcal{D},\ |x - a| < \delta : |f(x) - f(a)| < \varepsilon\ .$$

If we set $L = f(a)$, then it follows that we can write

$$\forall \varepsilon > 0\ \exists \delta > 0\ \forall x \in \mathcal{D},\ 0 < |x - a| < \delta : |f(x) - L| < \varepsilon\ .$$

But this implies $\lim_{x \to a} f(x) = L$, so $\lim_{x \to a} f(x) = f(a)$ as needed.

"$\Leftarrow$" Let $\lim_{x \to a} f(x) = f(a)$. Then

$$\forall \varepsilon > 0\ \exists \delta > 0\ \forall x \in \mathcal{D},\ 0 < |x - a| < \delta : |f(x) - f(a)| < \varepsilon\ .$$

Additionally, for $x = a$, we have $|f(a) - f(x)| = 0 < \varepsilon$, so that

$$\forall \varepsilon > 0\ \exists \delta > 0\ \forall x \in \mathcal{D},\ |x - a| < \delta : |f(x) - f(a)| < \varepsilon\ .$$

This implies that f is continuous at $a \in \mathcal{D}$.

$\square$

Remark. If f is continuous, we are allowed to "exchange" lim and f, i.e.

$$\lim_{x\to a} f(x) = f\left(\lim_{x\to a} x\right) .$$

In other words, it does not matter whether we evaluate the function first and then take the limit or whether we first take the limit and then evaluate the function.

Theorem 0.3. *If $f : \mathcal{D} \to \mathbb{R}$ is continuous at $a \in D$ and $b = f(a) \neq 0$ then $f(x) \neq 0$ nearby, i.e.*

$$\exists \delta > 0\ \forall x \in \mathcal{D},\ |x - a| < \delta : f(x) \neq 0 .$$

Proof. f is continuous at a, and $b = f(a)$, so that

$$\forall \varepsilon > 0\ \exists \delta > 0\ \forall x \in \mathcal{D},\ 0 < |x - a| < \delta : |f(x) - b| < \varepsilon .$$

Now pick $\varepsilon = |b|$ so that $|f(x) - b| < |b|$. Then

$$|b| > |f(x) - b| \geq ||f(x)| - |b|| \geq |b| - |f(x)|$$

or, equivalently, $|f(x)| > 0$.

Therefore, by choosing ε as we did, we have shown

$$\exists \delta > 0\ \forall x \in \mathcal{D},\ |x - a| < \delta : f(x) \neq 0 .$$

$\square$

Reminder. Use the triangle inequality $|x + y| \leq |x| + |y|$ (Δ) to show

$$|x - y| \geq ||x| - |y|| .$$

Proof. We need to show both (a) $|x - y| \geq |x| - |y|$ and (b) $|x - y| \geq |y| - |x|$.

(a) is equivalent to $|x| \leq |x - y| + |y|$, and

$$|x| = |(x - y) + y| \leq |x - y| + |y| \quad \text{by } (\Delta).$$

(b) is equivalent to $|y| \leq |x - y| + |x|$, and

$$|y| = |(y - x) + x| \leq |y - x| + |x| \quad \text{by } (\Delta).$$

$\square$

1 Differentiation

Lecture 2: 13/01/11

Let $\mathcal{D} \subseteq \mathbb{R}$ be a domain without isolated points (to allow limits at all points of $\mathcal{D}$).

Definition 1.1. *Let $f : \mathcal{D} \to \mathbb{R}$.*

(a) f is differentiable at $a \in \mathcal{D}$ if the limit

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

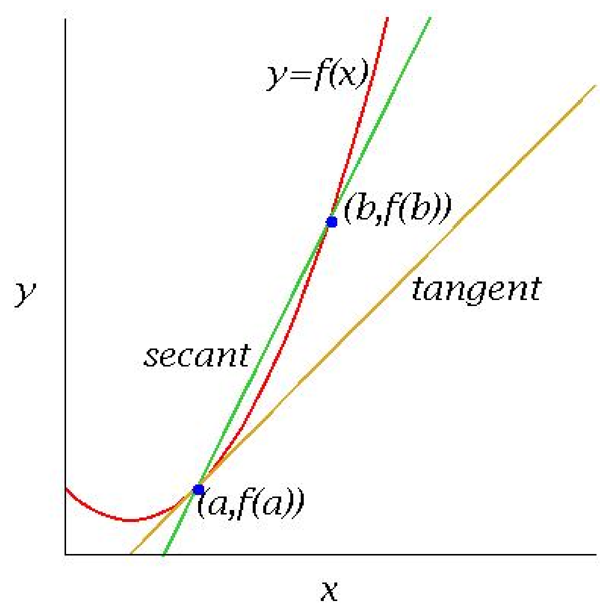
exists. $f'(a)$ is the derivative of f at a.

(b) f is differentiable if f is differentiable at all $a \in \mathcal{D}$. The function $f' : \mathcal{D} \to \mathbb{R}$ given by $x \mapsto f'(x)$ is the derivative of f.

Remark. Geometric interpretation: the difference quotient

$$\frac{f(b) - f(a)}{b - a}$$

is the slope of the secant line through the points $(a, f(a))$ and $(b, f(b))$, and the limit $f'(a)$ is the slope of the tangent line at $(a, f(a))$ of the graph of f.



Examples.

1) $f : \mathbb{R} \to \mathbb{R}$, $x \mapsto x^2$ is differentiable at every $a \in \mathbb{R}$:

   We have
   $$\frac{f(x) - f(a)}{x - a} = \frac{x^2 - a^2}{x - a} = x + a$$
   and
   $$\lim_{x \to a} \frac{f(x) - f(a)}{x - a} = \lim_{x \to a}(x + a) = 2a \ .$$
   The derivative is $f' : \mathbb{R} \to \mathbb{R}$, $x \mapsto 2x$.

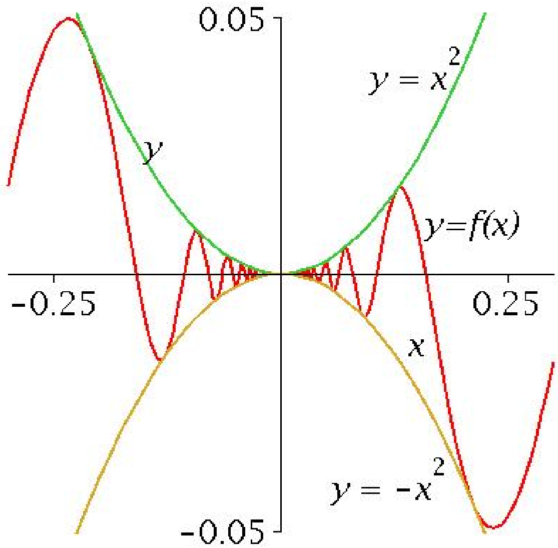
2) $f : \mathbb{R} \to \mathbb{R}$, $x \mapsto |x|$ is not differentiable at $a = 0$:

   We have
   $$\frac{f(x) - f(0)}{x - 0} = \frac{|x|}{x} = \begin{cases} -1 & x < 0 \\ 1 & x > 0 \end{cases}$$
   and $\lim_{x \to 0} \dfrac{f(x) - f(0)}{x - 0}$ does not exist.

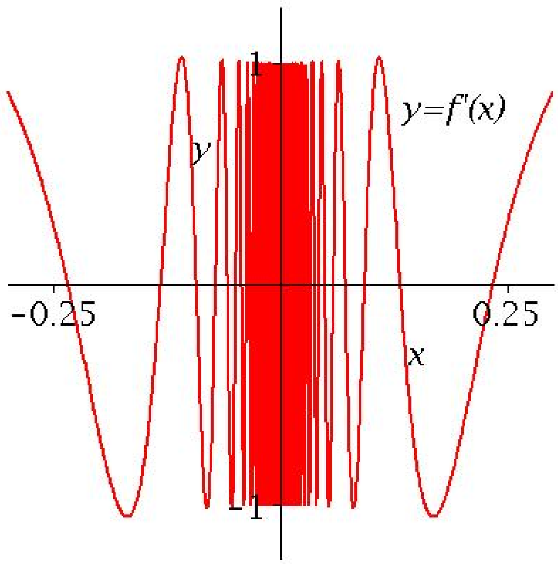
   If $a \neq 0$ then, by Theorem 0.3, x and a have the same sign when x is close to a, so that for $a > 0$
   $$\lim_{x \to a} \frac{f(x) - f(a)}{x - a} = \lim_{x \to a} \frac{x - a}{x - a} = 1 \ ,$$
   and for $a < 0$
   $$\lim_{x \to a} \frac{f(x) - f(a)}{x - a} = \lim_{x \to a} \frac{-x + a}{x - a} = -1 \ ,$$
   so that
   $$f'(x) = \begin{cases} 1 & x > 0 \\ \text{undefined} & x = 0 \\ -1 & x < 0 \ . \end{cases}$$

3) $f : \mathbb{R} \to \mathbb{R}$, $x \mapsto \begin{cases} x^2 \sin \dfrac{1}{x} & x \neq 0 \\ 0 & x = 0 \end{cases}$ is differentiable at $a = 0$:

   This is unclear from the graph of f, as f "wobbles" near zero.



Plotting the derivative doesn't help much, either:



We have

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

and noting that $\left|\sin \frac{1}{x}\right| \leq 1$, we have

$$\left|\lim_{x\to 0} x \sin \frac{1}{x}\right| = \lim_{x\to 0} \left|x \sin \frac{1}{x}\right| \leq \lim_{x\to 0} |x| = 0$$

and therefore $f'(0) = \lim_{x\to 0} x \sin \frac{1}{x} = 0$.

Lecture 3: 14/01/11

Lemma 1.2. *$f : \mathcal{D} \to \mathbb{R}$ is differentiable at a if and only if there exist $s, t \in \mathbb{R}$ and $r : \mathcal{D} \to \mathbb{R}$ such that*

(1) $f(x) = s + t(x - a) + r(x)(x - a)$ for all $x \in \mathcal{D}$, and

(2) $\lim_{x \to a} r(x) = 0$.

Remark. These properties say that $f(x)$ can be approximated by a linear function $y = s + t(x - a)$ for x close to a.

Proof. "$\Rightarrow$" Let f be differentiable at a. We define $r : \mathcal{D} \to \mathbb{R}$ by

$$r(x) = \begin{cases} \dfrac{f(x) - f(a)}{x - a} - f'(a) & x \neq a \\ 0 & x = a \end{cases} .$$

For $x \neq a$ it follows that

$$f(x) = f(a) + f'(a)(x - a) + r(x)(x - a) .$$

For $x = a$, this identity holds as well, as it reduces to $f(a) = f(a)$. Therefore (1) holds with $s = f(a)$ and $t = f'(a)$. To show (2) we compute

$$\lim_{x \to a} r(x) = f'(a) - f'(a) = 0 .$$

"$\Leftarrow$" Inserting $x = a$ into (1) gives $f(a) = s$, so that (1) gives

$$f(x) = f(a) + t(x - a) + r(x)(x - a) ,$$

and therefore

$$\frac{f(x) - f(a)}{x - a} = t + r(x) .$$

Now (2) implies that the limit

$$\lim_{x \to a} \frac{f(x) - f(a)}{x - a} = t + \lim_{x \to a} r(x) = t$$

exists.

$\square$

Remark. If $f(x) = s + t(x-a) + r(x)(x-a)$ with $\lim_{x\to a} r(x) = 0$, then f is differentiable at a with $s = f(a)$ and $t = f'(a)$. The equation of the tangent at a of the graph of f is therefore

$$y = f(a) + f'(a)(x-a) \ .$$

Theorem 1.3. *If $f : \mathcal{D} \to \mathbb{R}$ is differentiable at $a \in \mathcal{D}$ then f is continuous at a.*

Proof. By Lemma 1.2,

$$f(x) = s + t(x-a) + r(x)(x-a)$$

with $\lim_{x\to a} r(x) = 0$. Therefore $\lim_{x\to a} f(x) = s = f(a)$. $\square$

Remark. $f : \mathbb{R} \to \mathbb{R}$, $x \mapsto |x|$ is continuous at 0 but not differentiable. The converse of Theorem 1.3 is therefore not true.

Theorem 1.4. *Let $f, g : \mathcal{D} \to \mathbb{R}$ be differentiable at $a \in \mathcal{D}$ and let $c \in \mathbb{R}$. Then $f + g$, cf, fg, and f/g (if $g(a) \neq 0$) are differentiable at a. We have*

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

(b) $(cf)' = cf'$,

(c) $(fg)' = f'g + fg'$ *(product rule), and*

(d) $(f/g)' = (f'g - fg')/g^2$ *(quotient rule).*

Proof. (a) This is easy.

(b) This is a special case of (c) with the constant function $g(x) = c$.

(c) Write

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

As f and g are differentiable at a and g is continuous at a by Theorem 1.3,

$$(fg)'(a) = f'(a)g(a) + f(a)g'(a) \ .$$

(d) By Theorem 1.3, g is continuous at a. $g(a) \neq 0$, therefore $g(x) \neq 0$ nearby, i.e.

$$\exists \delta > 0 \, \forall x \in \mathcal{D}, \; |x-a| < \delta : g(x) \neq 0 \, .$$

Therefore $f(x)/g(x)$ is defined near a, and

$$\frac{\frac{f(x)}{g(x)} - \frac{f(a)}{g(a)}}{x-a} = \frac{1}{g(x)g(a)} \left(\frac{f(x)-f(a)}{x-a} g(x) - f(a) \frac{g(x)-g(a)}{x-a} \right) \, .$$

The limit as $x \to a$ exists on the right-hand-side, and therefore

$$\left(\frac{f}{g}\right)'(a) = \frac{1}{g(a)^2} \left(f'(a)g(a) - f(a)g'(a) \right) \, .$$

□

Example. Show that

$$\left(\frac{1}{f}\right)' = -\frac{f'}{f^2} \, :$$

(a) Use the quotient rule with constant function 1 in numerator:

$$\left(\frac{1}{f}\right)' = \frac{0 \cdot f - 1 \cdot f'}{f^2} = -\frac{f'}{f^2} \, .$$

(b) Use the product rule with $g = 1/f$, so that $fg = 1$, and differentiate this:

$$0 = (fg)' = f'g + fg' \qquad \text{and therefore} \qquad g' = -\frac{f'g}{f} = -\frac{f'}{f^2} \, .$$

Remark. All the derivatives from Calculus we shall assume as known. This is not cheating, as we can prove every single one in principle.

Lecture 4: 17/01/11

Theorem 1.5 (Chain Rule)**.** *let $f : \mathcal{D} \to \mathbb{R}$ be differentiable at $a \in D$, and let $g : f(\mathcal{D}) \to \mathbb{R}$ be differentiable at $b = f(a)$. Then $g \circ f : \mathcal{D} \to \mathbb{R}$ is differentiable at a and*

$$(g \circ f)'(a) = g'(f(a))f'(a) \, .$$

Remark. To get an idea for the formula, let us write

$$\frac{g \circ f(x) - g \circ f(a)}{x-a} = \frac{g \circ f(x) - g \circ f(a)}{f(x)-f(a)} \cdot \frac{f(x)-f(a)}{x-a} \, .$$

It looks like we can easily take the limit of $x \to a$ on the right-hand side. However, the problem is that $f(x) - f(a)$ might be zero for $x \neq a$, and we need to be more careful because of this.

Proof. By Lemma 1.2 we have

(1) $f(x) = f(a) + f'(a)(x-a) + r(x)(x-a)$, and

(2) $g(y) = g(b) + g'(b)(y-b) + s(y)(y-b)$

with $\lim_{x\to a} r(x) = 0$ and $\lim_{y\to b} s(y) = 0$. Define $s(b) = 0$ so that s is continuous at b. Let $y = f(x)$ to get

$$\begin{aligned} g\circ f(x) - g(b) &= \left(g'(b) + s(f(x))\right)\left(f(x) - b\right) \\ &= \left(g'(b) + s(f(x))\right)\left(f'(a) + r(x)\right)(x-a) \\ &= g'(b)f'(a)(x-a) + t(x)(x-a)\ , \end{aligned}$$

where $t(x) = s(f(x))f'(a) + g'(b)r(x) + s(f(x))r(x)$. Then

$$\begin{aligned} \lim_{x\to a} t(x) &= \lim_{x\to a}\left(s(f(x))f'(a) + g'(b)r(x) + s(f(x))r(x)\right) \\ &= \lim_{x\to a} s(f(x))f'(a) + g'(b)\lim_{x\to a} r(x) + \lim_{x\to a} s(f(x))\lim_{x\to a} r(x)\ . \end{aligned}$$

Now $\lim_{x\to a} r(x) = 0$, and also $\lim_{x\to a} s(f(x)) = 0$ (for the latter we crucially need that s is continuous at b), so that

$$\lim_{x\to a} t(x) = 0\ .$$

Thus $g\circ f$ is differentiable at a with $(g\circ f)'(a) = g'(b)f'(a) = g'(f(a))f'(a)$. $\square$

2 The Mean Value Theorem

Theorem 2.1. *If a function $f : [a,b] \to \mathbb{R}$ has a maximum (or minimum) at $c \in (a,b)$ and is differentiable at c, then $f'(c) = 0$.*

Proof. If f has a minumum at c then $-f$ has a maximum at c, so it suffices to consider the case of f having a maximum at c. By assumption f is differentiable at c, so

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

exists. Restricting to the one-sided limits, we have furthermore

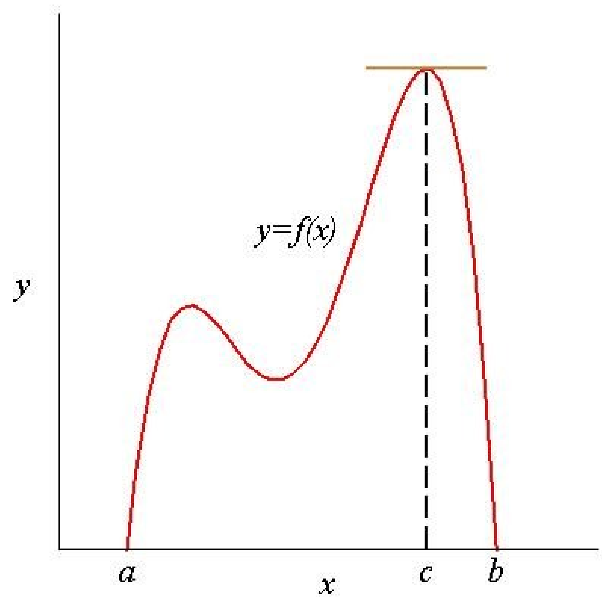
$$d = \lim_{x \to c^+} \frac{f(x) - f(c)}{x - c} \leq 0$$

and

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

Therefore $d = 0$. □

Theorem 2.2 (Rolle)**.** *Let $f : [a,b] \to \mathbb{R}$ be continuous on $[a,b]$ and differentiable on (a,b). If $f(a) = f(b) = 0$ then there exists a $c \in (a,b)$ such that $f'(c) = 0$.*



Proof. We consider three cases:

(1) $f(x) = 0$ for all $x \in (a, b)$. Then $f'(x) = 0$ for all $x \in (a, b)$.

(2) $f(x) > 0$ for some $x \in (a, b)$. Then f is maximal at some $c \in [a, b]$ and $f(c) \geq f(x) > 0$. As $f(a) = f(b) = 0$, c must be different from a or b, so f is maximal at some $c \in (a, b)$. By Theorem 2.1 it follows that $f'(c) = 0$.

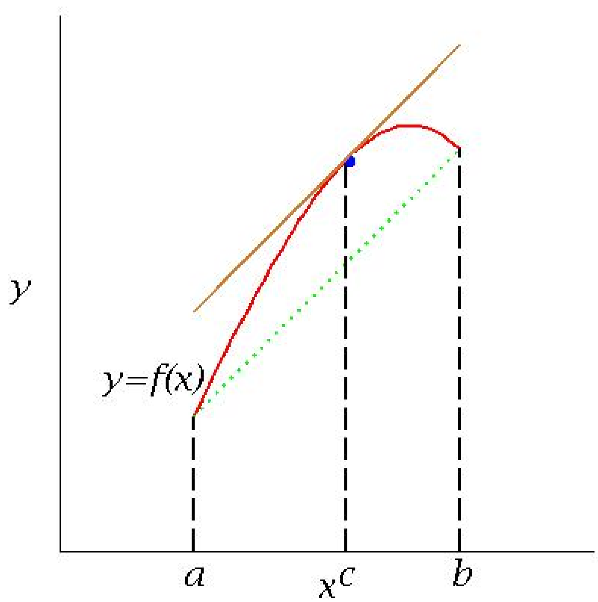
(2) $f(x) < 0$ for some $x \in (a, b)$. Then f is minimal at some $c \in [a, b]$ and $f(c) \leq f(x) < 0$. As $f(a) = f(b) = 0$, c must be different from a or b, so f is minimal at some $c \in (a, b)$. By Theorem 2.1 it follows that $f'(c) = 0$.

□

Lecture 5: 20/01/11

Theorem 2.3 (Mean Value Theorem)**.** *Let $f : [a, b] \to \mathbb{R}$ be continuous on $[a, b]$ and differentiable on (a, b). Then there exists a $c \in (a, b)$ such that*

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



Proof. The equation of the straight line through the points $(a, f(a))$ and $(b, f(b))$ is

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

Taking the difference between $y = f(x)$ and this equation, we define the auxiliary function

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

By construction, h is continuous on $[a, b]$ and differentiable on (a, b). Moreover

$$h(a) = 0 \qquad \text{and} \qquad h(b) = 0 \ ,$$

so that Rolle's Theorem applies: there exists a $c \in (a, b)$ such that $h'(c) = 0$. Now

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

so that $h'(c) = 0$ implies $f'(c) = \dfrac{f(b) - f(a)}{b - a}$ as claimed. □

Remark. Geometric interpretation: there exists a tangent to the graph of f which is parallel to the secant line through $(a, f(a))$ and $(b, f(b))$.

We continue with some applications of the Mean Value Theorem.

Theorem 2.4. *Let $f : [a, b] \to \mathbb{R}$ be continuous on $[a, b]$ and differentiable on (a, b).*

(a) If $f'(x) > 0$ for all $x \in (a, b)$, then f is strictly increasing on $[a, b]$, i.e. $x_1 < x_2$ implies $f(x_1) < f(x_2)$.

(b) If $f'(x) < 0$ for all $x \in (a, b)$, then f is strictly decreasing on $[a, b]$, i.e. $x_1 < x_2$ implies $f(x_1) > f(x_2)$.

Proof. (a) Let $x_1, x_2 \in [a, b]$ with $x_1 < x_2$. Applying the Mean Value Theorem to f on $[x_1, x_2]$, we have that there exists a $c \in (x_1, x_2)$ with

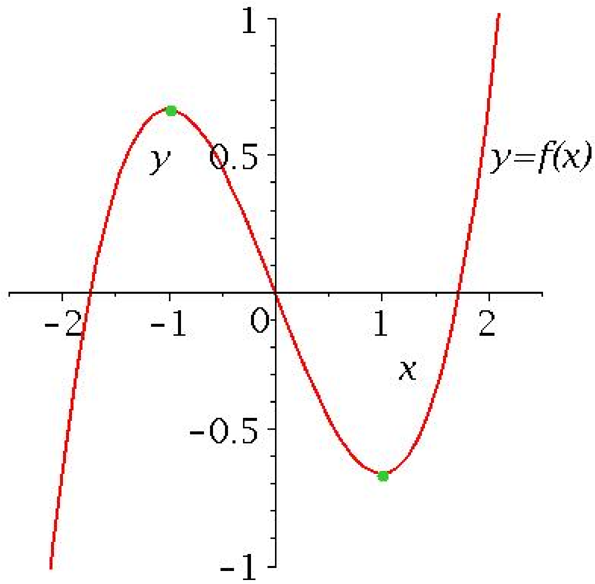
$$\frac{f(x_2) - f(x_1)}{x_2 - x_1} = f'(c) > 0 \ .$$

Therefore $f(x_2) - f(x_1) > 0$.

(b) Replace f by $-f$ and repeat.

□

Example. Find intervals on which $f : \mathbb{R} \to \mathbb{R}$, $x \mapsto \dfrac{x^3}{3} - x$ is strictly increasing or strictly decreasing.



As $f'(x) = x^2 - 1$, $f'(x) < 0$ on $(-1, 1)$ and $f'(x) > 0$ on $(-\infty, -1) \cup (1, \infty)$. Therefore f is strictly decreasing on $[-1, 1]$ and strictly increasing on $(-\infty, -1]$ and $[1, \infty)$.

Theorem 2.5. *Let $f : [a, b] \to \mathbb{R}$ be continuous on $[a, b]$ and differentiable on (a, b). If $f'(x) = 0$ for all $x \in (a, b)$, then f is constant on $[a, b]$, i.e. $f(x) = f(a)$ for all $x \in [a, b]$.*

Proof. Let $x \in (a, b]$ and apply the Mean Value Theorem to f on $[a, x]$: there exists a $c \in (a, x)$ such that $\dfrac{f(x) - f(a)}{x - a} = f'(c) = 0$. Therefore $f(x) = f(a)$. $\square$

Lecture 6: 21/01/11

We conclude this section with presenting an Intermediate Value Theorem for differentiable functions. First recall the Intermediate Value Theorem for continuous functions.

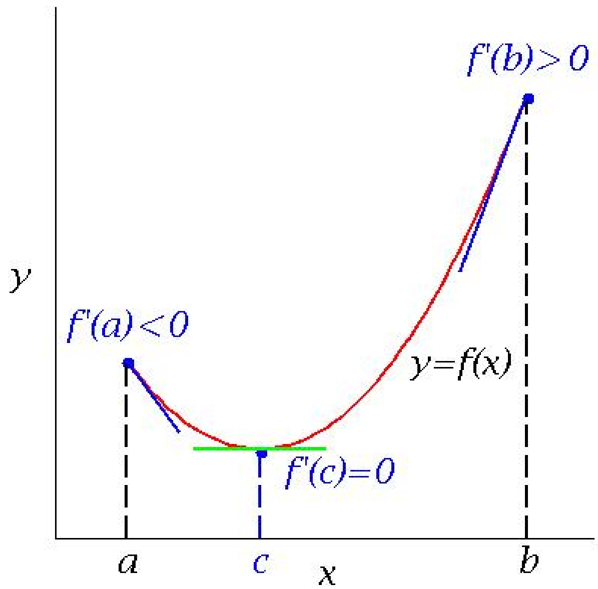
Theorem (Intermediate Value Theorem)**.** *Let $f : [a, b] \to \mathbb{R}$ be continuous and $f(a) < s < f(b)$. Then there exists a $c \in (a, b)$ such that $f(c) = s$.*

The following theorem looks very similar.

Theorem 2.6. *Let $f : [a, b] \to \mathbb{R}$ be differentiable and $f'(a) < s < f'(b)$. Then there exists a $c \in (a, b)$ such that $f'(c) = s$.*

Remark. This shows that the derivative of differentiable functions satisfies the intermediate value property. Note that the derivative doesn't have to be continuous, so this is different from the Intermediate Value Theorem for continuous functions.

Proof. Consider the case $s = 0$ first. We need to show that there exists a $c \in (a, b)$ such that $f'(c) = 0$:



As f is differentiable on $[a, b]$, f is continuous on $[a, b]$ and therefore attains its minimum on $[a, b]$. $f'(a) < 0$ implies that there exists an $a' > a$ with $(f(a') - f(a))/(a' - a) < 0$, thus there exists an $a' > a$ with $f(a') < f(a)$. Similarly, as $f'(b) > 0$, there exists a $b' < b$ with $f(b') < f(b)$. Therefore the minimum is not attained at the endpoints a or b, but at some point in (a, b). As f is differentiable at $c \in (a, b)$, $f'(c) = 0$ by Theorem 2.1. This concludes the proof for $s = 0$.

Now consider the general case of $s \neq 0$. We can reduce this to the case $s = 0$ by considering the function $g(x) = f(x) - sx$. g is differentiable on $[a, b]$ and $g'(x) = f'(x) - s$ implies $g'(a) = f'(a) - s < 0$ and $g'(b) = f'(b) - s > 0$. Therefore, $g'(c) = 0$ for some $c \in (a, b)$. This implies $f'(c) = s$.

□

3 The Exponential Function

Definition 3.1. *A differentiable function $f : \mathbb{R} \to \mathbb{R}$ with (a) $f'(x) = f(x)$ for all $x \in \mathbb{R}$, and (b) $f(0) = 1$ is called exponential function.*

Remark. We will show later that $f(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!}$ satisfies this definition. For now, we shall assume existence of such a function.

In items (A) to (J) we shall collect properties of an exponential function.

(A) $f(x)f(-x) = 1$.

Proof. Differentiate $h(x) = f(x)f(-x)$:

$$h'(x) = f'(x)f(-x) + f(x)f'(-x)(-1) = 0 ,$$

and by Theorem 2.5, h is constant and $h(0) = f(0)f(0) = 1$, so $h(x) = 1$. □

Lecture 7: 24/01/11

(B) $f(x) \neq 0$ for all $x \in \mathbb{R}$.

Proof. If $f(x) = 0$ for some $x \in \mathbb{R}$ then $0 = f(x)f(-x) = 1$, a contradiction. □

(C) Let $g : \mathbb{R} \to \mathbb{R}$ be differentiable and $g' = g$. Then there exists a $c \in \mathbb{R}$ such that $g = cf$.

Proof. Consider $h(x) = g(x)/f(x)$. By (B), h is defined on $\mathbb{R}$ and differentiable with

$$h'(x) = \frac{g'(x)f(x) - g(x)f'(x)}{f(x)^2} = 0 .$$

Therefore h is constant, $h(x) = c$, and $g(x) = cf(x)$. □

(D) Definition 3.1 determines f uniquely.

Proof. Assume g satisfies Definition 3.1. Then (C) implies $g = cf$. As $g(0) = 1 = f(0)$, we have $c = 1$, so $g = f$. □

Now that we have shown uniqueness, we will write $f(x) = \exp(x)$ for the function f defined by Definition 3.1.

Theorem 3.2. *For all $a, b \in \mathbb{R}$, $\exp(a+b) = \exp(a)\exp(b)$.*

Proof. Consider $g(x) = \exp(a+x)$. Then $g'(x) = \exp(a+x) = g(x)$, so $\exp(a+x) = c\exp(x)$ by (C). Letting $x = 0$, we find $\exp(a) = c$, so that $\exp(a+b) = c\exp(b) = \exp(a)\exp(b)$. □

Corollary. *For $a \in \mathbb{R}$ and $n \in \mathbb{N}$, $\exp(na) = (\exp(a))^n$.*

Proof. We use mathematical induction in n: For $n = 1$, we have

$$\exp(1a) = \exp(a) = (\exp(a))^1 .$$

Next, assuming that we have shown that $\exp(na) = (\exp(a))^n$ for some $n \in \mathbb{N}$, we deduce that

$$\exp((n+1)a) = \exp(na)\exp(a) = (\exp(a))^n \exp(a) = (\exp(a))^{n+1} .$$

□

(E) $\exp(x) > 0$ for all $x \in \mathbb{R}$.

Proof. The function exp is differentiable, therefore continuous. By (B), $\exp(x) \neq 0$ for all $x \in \mathbb{R}$, and $\exp(0) = 1 > 0$. Assume now that (E) is false, i.e. there exists an $x \in \mathbb{R}$ for which $\exp(x) < 0$. By the Intermediate Value Theorem it follows that there exists a $c \in \mathbb{R}$ such that $\exp(c) = 0$, a contradiction. □

(F) exp is strictly increasing.

Proof. $\exp'(x) = \exp(x) > 0$, and the claim follows from Theorem 2.4. □

Theorem 3.3. *For all $x \in \mathbb{R}$, $\exp(x) > x$.*

Proof. $x < 0$: $\exp(x) > 0 > x$.

$x = 0$: $\exp(x) = 1 > 0$.

$x > 0$: By the Mean Value Theorem applied to $[0, x]$, there exists a $c \in (0, x)$ such that

$$\frac{\exp(x) - \exp(0)}{x - 0} = \exp(c) .$$

Moreover, $\exp(c) > \exp(0) = 1$ by (F). Therefore $\exp(x) - 1 = x\exp(c) > x$, and thus $\exp(x) > x + 1 > x$. □

Lecture 8: 27/01/11

(G) $\exp(\mathbb{R}) = \mathbb{R}^+$ $(= \{x \in \mathbb{R} : x > 0\})$.

Proof. (E) implies that $\exp(\mathbb{R}) \subseteq \mathbb{R}^+$. We need to show that also $\mathbb{R}^+ \subseteq \exp(\mathbb{R})$, i.e.

$$\forall c > 0 \ \exists x \in \mathbb{R}, \ \exp(x) = c \ .$$

Case 1: $c \geq 1$.

We have $\exp(0) = 1 \leq c < \exp(c)$. By the Intermediate Value Theorem applied to $[0, c]$, there exists an $x \in (0, c)$ such that $\exp(x) = c$.

Case 2: $0 < c < 1$.

Now $1/c > 1$ and as in Case 1 we can deduce that there exists an $x \in (0, 1/c)$ such that $\exp(x) = 1/c$. As $\exp(x)\exp(-x) = 1$, we have $\exp(-x) = c$. $\square$

(H) $\exp(1) = e$, where $e = \lim_{n\to\infty} \left(1 + \frac{1}{n}\right)^n$.

Proof. Recall the Bernoulli inequality: $(1 + x)^n \geq 1 + nx$ for all $x \geq -1$ and for all $n \in \mathbb{N}_0$.

1) Show that $\lim_{n\to\infty} \left(1 + \frac{1}{n}\right)^n$ exists:

(a) $a_n = \left(1 + \frac{1}{n}\right)^n$ is increasing:

Using

$$\left(1 - \frac{1}{n^2}\right)\left(1 + \frac{1}{n-1}\right) = \frac{n^2 - 1}{n}\frac{n}{n-1} = 1 + \frac{1}{n} \ ,$$

it follows that

$$\begin{aligned} a_n = \left(1 + \frac{1}{n}\right)^n &= \left(1 - \frac{1}{n^2}\right)^n \left(1 + \frac{1}{n-1}\right)^n \\ &\geq \left(1 - \frac{1}{n}\right)\left(1 + \frac{1}{n-1}\right)\left(1 + \frac{1}{n-1}\right)^{n-1} = \left(1 + \frac{1}{n-1}\right)^{n-1} = a_{n-1} \ , \end{aligned}$$

where we have used the estimate $\left(1 - \frac{1}{n^2}\right)^n \geq 1 - \frac{1}{n}$ which follows from the Bernoulli inequality.

(b) $b_n = \left(1+\dfrac{1}{n}\right)^{n+1}$ is decreasing:

From the Bernoulli inequality it follows that

$$\left(1+\frac{1}{n^2-1}\right)^n \geq 1+\frac{n}{n^2-1} \geq 1+\frac{1}{n} \,.$$

Therefore

$$\begin{aligned} b_n &= \left(1+\frac{1}{n}\right)^n \left(1+\frac{1}{n}\right) \\ &\leq \left(1+\frac{1}{n}\right)^n \left(1-\frac{1}{n^2-1}\right)^n = \left(1+\frac{1}{n-1}\right)^n = b_{n-1} \,. \end{aligned}$$

(c) Each b_m is an upper bound for the sequence (a_n) and each a_m is a lower bound for the sequence (b_n). Therefore the limits $\lim\limits_{n\to\infty} a_n$ and $\lim\limits_{n\to\infty} b_n$ exist. We find

$$\lim_{n\to\infty} b_n = \lim_{n\to\infty} \left(a_n \left(1+\frac{1}{n}\right)\right) = \lim_{n\to\infty} a_n \,.$$

2) Show that

$$a_n = \left(1+\frac{1}{n}\right)^n \leq \exp(1) \leq \left(1+\frac{1}{n}\right)^{n+1} = b_n \ :$$

The Mean Value Theorem for exp on $[0, 1/n]$ implies that there exists a $c \in (0, 1/n)$ such that

$$\frac{\exp(1/n) - \exp(0)}{1/n - 0} = \exp(c)$$

so that $\exp(1/n) = 1 + \exp(c)/n$. As $1 \leq \exp(c) \leq \exp(1/n)$, we deduce that

$$1+\frac{1}{n} \leq \exp\left(\frac{1}{n}\right) \leq 1+\frac{1}{n}\exp\left(\frac{1}{n}\right) \,.$$

This implies firstly that

$$\left(1+\frac{1}{n}\right)^n \leq \left(\exp\left(\frac{1}{n}\right)\right)^n = \exp(1) \,.$$

Secondly, $(1-1/n)\exp(1/n) \leq 1$, so that $\exp(1/n) \leq n/(n-1)$ for $n \geq 2$. Shifting n by one, we deduce that $\exp(1/(n+1)) \leq (n+1)/n = 1+1/n$, so that

$$\left(1+\frac{1}{n}\right)^{n+1} \geq \left(\exp\left(\frac{1}{n+1}\right)\right)^{n+1} = \exp(1) \,.$$

$\square$

Lecture 9: 28/01/11

Corollary. $\exp(n) = e^n$ *for* $n \in \mathbb{Z}$.

Proof. $n \in \mathbb{N}$: $\exp(n) = (\exp(1))^n = e^n$.

$n = 0$: $\exp(0) = 1 = e^0$.

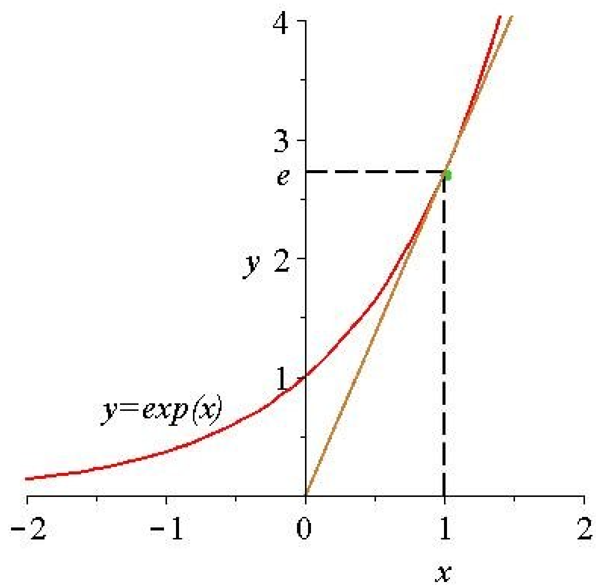
$-n \in \mathbb{N}$: $\exp(-n) = (\exp(n))^{-1} = e^{-n}$. $\square$

We also have $(\exp(n/m))^m = \exp(n) = e^n$, so that $\exp(n/m) = e^{n/m}$. Summarising we have proved the following result.

Theorem 3.4. *(1)* $\exp$ *is strictly increasing,*

(2) $\exp(\mathbb{R}) = \mathbb{R}^+$, *and*

(3) $\exp(x) = e^x$ *for all* $x \in \mathbb{Q}$.



4 Inverse Functions

Definition 4.1. *Let $f : \mathcal{D} \longrightarrow \mathbb{R}$, and let $\mathcal{E} = f(\mathcal{D})$ be the image of f. Then f is <u>invertible</u> if there exists $g : \mathcal{E} \longrightarrow \mathbb{R}$ such that*

$$g \circ f(x) = x \text{ for all } x \in \mathcal{D} \text{ and } f \circ g(x) = x \text{ for all } x \in \mathcal{E}.$$

g is an <u>inverse</u> of f.

Properties of the inverse:

1) The inverse is uniquely defined.

Proof. Let $\mathcal{E} = f(\mathcal{D})$ and $g_1, g_2 : \mathcal{E} \longrightarrow \mathbb{R}$ be inverses of f. Let $y \in \mathcal{E}$. There exists an $x \in \mathcal{D}$ with $y = f(x)$ and

$$g_1(y) = g_1 \circ f(x) = x = g_2 \circ f(x) = g_2(y) ,$$

so $g_1 = g_2$. $\square$

As the inverse is uniquely defined, we can write $g = f^{-1}$.

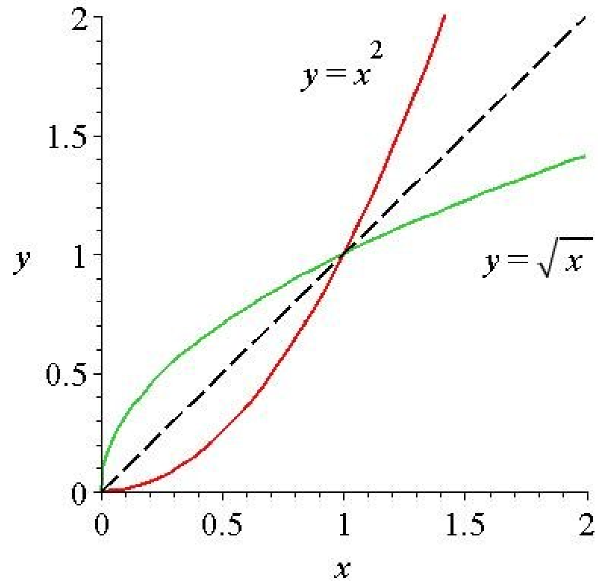
2) If f is invertible, then f^{-1} is invertible as well, and $(f^{-1})^{-1} = f$.

3) The graphs of f and f^{-1} are mirror images with respect to the straight line $y = x$.

Proof. $\text{Graph}(f) = \{(x, f(x)) : x \in \mathcal{D}\}$ and $\text{Graph}(f^{-1}) = \{(y, f^{-1}(y)) : y \in \mathcal{E}\} = \{(f(x), f^{-1} \circ f(x)) : x \in \mathcal{D}\} = \{(f(x), x) : x \in \mathcal{D}\}$ is its mirror image. $\square$

Example.

$$f : \mathbb{R}_0^+ \to \mathbb{R} \qquad f(x) = x^2 \qquad f(\mathbb{R}_0^+) = \mathbb{R}_0^+$$
$$f^{-1} : \mathbb{R}_0^+ \to \mathbb{R} \qquad f^{-1}(x) = \sqrt{x} \qquad f^{-1}(\mathbb{R}_0^+) = \mathbb{R}_0^+$$



Lecture 10: 31/01/11

Theorem 4.2. *$f : \mathcal{D} \to \mathbb{R}$ is invertible if and only if it is injective (one-to-one).*

Proof. "⇒" Let f be invertible and $f(x_1) = f(x_2)$. Then $x_1 = f^{-1} \circ f(x_1) = f^{-1} \circ f(x_2) = x_2$.

"⇐" Let f be injective and let $\mathcal{E} = f(\mathcal{D})$. Then for each $y \in \mathcal{E}$ there is a unique $x \in \mathcal{D}$ such that $y = f(x)$. Define $g : \mathcal{E} \to \mathbb{R}$ via $g(y) = x$. Then

$$g \circ f(x) = g(y) = x \quad \forall x \in \mathcal{D} \text{ and}$$
$$f \circ g(y) = f(x) = y \quad \forall y \in \mathcal{E} \; .$$

□

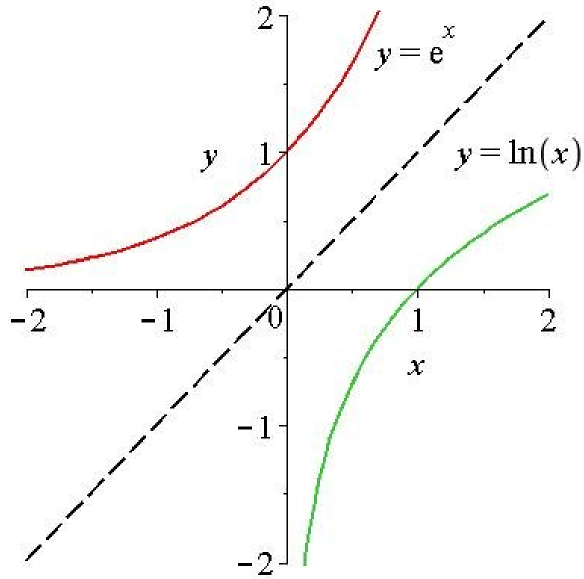
Corollary. *If $f : \mathcal{D} \to \mathbb{R}$ is strictly increasing (or decreasing) then f is invertible.*

Proof.

$$\left.\begin{aligned} x_1 < x_2 \Rightarrow f(x_1) < f(x_2) \\ x_1 > x_2 \Rightarrow f(x_1) > f(x_1) \end{aligned}\right\} \qquad \text{implies } x_1 \neq x_2 \Rightarrow f(x_1) \neq f(x_2).$$

□

Example. $\exp : \mathbb{R} \to \mathbb{R}$ is strictly increasing, therefore invertible.

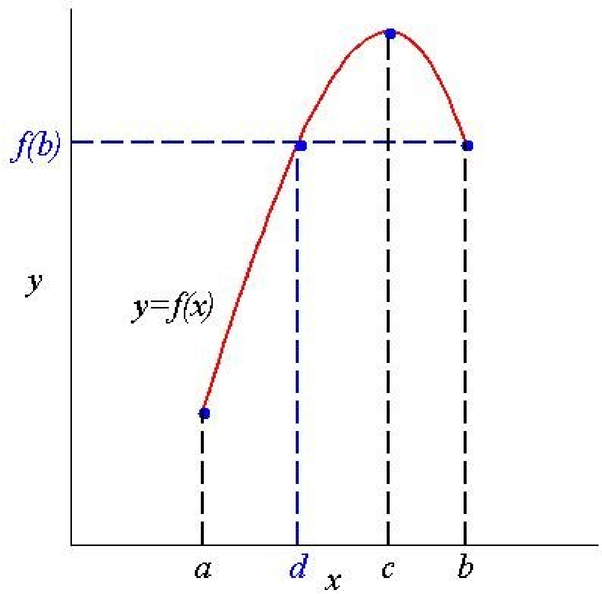


$$\exp(\mathbb{R}) = \mathbb{R}^+ \qquad\qquad \exp^{-1} = \log : \ \mathbb{R}^+ \to \mathbb{R} \ .$$

Let I be an interval ($a, b \in I$, $a \leq c \leq b \Rightarrow c \in I$). If $f : I \to R$ is continuous then $f(I)$ is an interval (by the Intermediate Value Theorem).

Theorem 4.3. *Let $f : [a, b] \to \mathbb{R}$ be continuous and injective. Then f attains its minimum and maximum at a or b.*

Proof. Without loss of generality, let $f(a) \leq f(b)$. f is continuous, therefore f attains its maximum at $c \in [a, b]$.



If $c < b$ then $f(a) \leq f(b) \leq f(c)$ and by the Intermediate Value Theorem there exists a $d \in [a, c]$ such that $f(d) = f(b)$. Now $d \leq c < b$ implies $d \neq b$, a contradiction to injectivity. Thus $c = b$ and f is maximal at b. An analogous argument shows that f is minimal at a. □

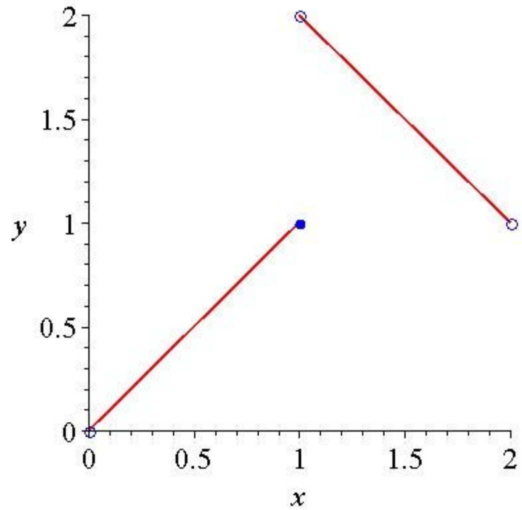
Theorem 4.4. *Let I be an interval and $f : I \to \mathbb{R}$ be continuous and injective. Then f is strictly increasing or decreasing.*

Proof. (1) Consider $I = [a, b]$ and assume without loss of generality that $f(a) < f(b)$. Let $x, y \in I$ with $x < y$. Then, by Theorem 4.3, f attains its maximum in b and therefore $f(x) \leq f(b)$. Restricted to the interval $[x, b]$, the minimum of f is attained at x, and thus $f(x) \leq f(y)$. As f is injective, in fact $f(x) < f(y)$.
(2) Consider now an arbitrary interval I. f is continuous and injective when restricted to any closed and bounded interval $[a, b] \subseteq I$, therefore by (1) it is strictly increasing or decreasing on $[a, b]$.

Now pick $u, v \in I$ with $u < v$ and assume without loss of generality that $f(u) < f(v)$. Let $x, y \in I$ with $x < y$, and choose a closed interval $[a, b] \subseteq I$ containing x, y, u, v. f is strictly increasing or decreasing on $[a, b]$, so $f(x) < f(y)$. □

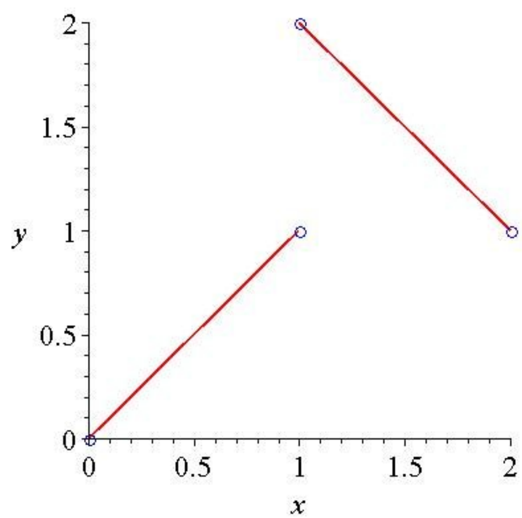
Examples.

1) $f:(0,2)\to\mathbb{R},\ f(x)=\begin{cases} x & x\in(0,1] \\ 3-x & x\in(1,2) \end{cases}.$



f is injective, but not strictly increasing or decreasing (it is not continuous).

2) $f:(0,1)\cup(1,2)\to\mathbb{R},\ f(x)=\begin{cases} x & x\in(0,1) \\ 3-x & x\in(1,2) \end{cases}.$



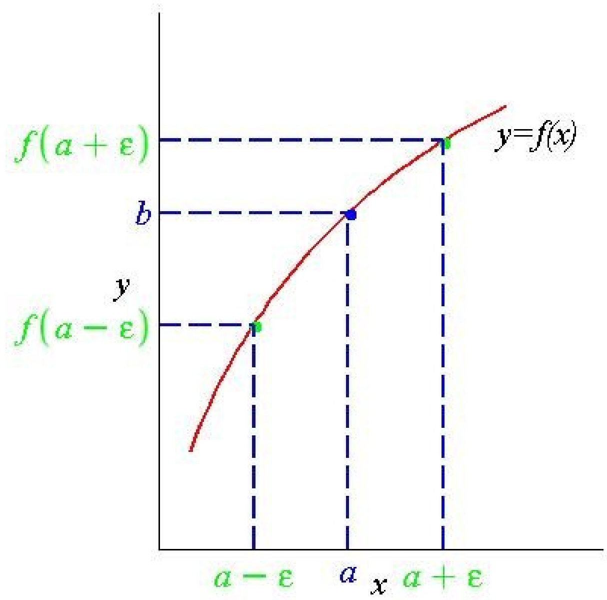
f is injective, continuous, but not strictly increasing or decreasing ($(0,1)\cup(1,2)$ is not an interval).

Lecture 11: 03/02/11

Theorem 4.5. *Let I be an interval and $f : I \to \mathbb{R}$ be continuous and injective. Then $f^{-1} : f(I) \to \mathbb{R}$ is continuous.*

Proof. Theorem 4.4 inplies that f is strictly increasing or decreasing. Consider the case of strictly increasing f. Let $a \in I$. Then $b = f(a) \in f(I)$, and we need to show that f^{-1} is continuous at b:

Fix $\varepsilon > 0$. If $y = f(x) \in f(I)$ satisfies $f(a - \varepsilon) < y < f(a + \varepsilon)$ then $a - \varepsilon < x < a + \varepsilon$.



Choose now

$$\delta := \min\{f(a + \varepsilon) - b, b - f(a - \varepsilon)\} .$$

Then $|y - b| < \delta$ implies $|x - a| < \varepsilon$, so f^{-1} is continuous at b. $\square$

Theorem 4.6. *Let I be an interval and $f : I \to \mathbb{R}$ be continuous and injective. Let f be differentiable at $a \in I$ and $b = f(a)$.*

(a) If $f'(a) = 0$ then f^{-1} is not differentiable at b.

(b) If $f'(a) \neq 0$ then f^{-1} is differentiable at b and

$$(f^{-1})'(b) = \frac{1}{f'(a)} = \frac{1}{f'(f^{-1}(b))} .$$

Proof. (a) Let $f'(a) = 0$ and assume f^{-1} is differentiable at $b = f(a)$. Then differentiating $x = f^{-1}(f(x))$ gives a contradiction:

$$1 = (f^{-1})'(f(a))f'(a) = 0 .$$

(b) Let $f'(a) \neq 0$. Define the difference quotient

$$A(y) = \frac{f^{-1}(y) - f^{-1}(b)}{y - b} \quad \text{for } y \neq b .$$

We need to show that $(f^{-1})'(b) = \lim_{y \to b} A(y)$ exists. Define now

$$B(x) = \begin{cases} \dfrac{f(x) - f(a)}{x - a} & x \neq a , \\ f'(a) & x = a . \end{cases}$$

Note that $\lim_{x \to a} B(x) = f'(a) = B(a)$, so B is continuous at a, and therefore continuous on I.

f^{-1} is continuous on $f(I)$, and so $B \circ f^{-1}$ is also continuous on $f(I)$. We compute

$$B \circ f^{-1}(y) = \begin{cases} \dfrac{y - b}{f^{-1}(y) - f^{-1}(b)} & y \neq b . \\ f'(a) & y = b . \end{cases}$$

Therefore $B \circ f^{-1}(y) = 1/A(y)$ for $y \neq b$ and

$$\lim_{y \to b} \frac{1}{A(y)} = B \circ f^{-1}(b) = f'(a) ,$$

so $(f^{-1})'(b)$ exists and equals $1/f'(a)$.

$\square$

Examples.

1) Consider $f : \mathbb{R} \to \mathbb{R}$, $x \mapsto x^3$. f is differentiable, and $f'(x) = 3x^2$. Moreover, $f(\mathbb{R}) = \mathbb{R}$ (and f is continuous by Theorem 1.3).

   $f' > 0$ on $(-\infty, 0)$ and on $(0, \infty)$, so by Theorem 2.4 f is strictly increasing on both $(-\infty, 0]$ and $[0, \infty)$, hence on all of $\mathbb{R}$.

   By the corollary to Theorem 4.2, f is invertible. (The inverse $f^{-1} : \mathbb{R} \to \mathbb{R}$ it is given by $x \mapsto x^{1/3}$).

   From Theorem 4.5 it follows that f^{-1} is continuous.

   From Theorem 4.6 it follows that f^{-1} is not differentiable at $x = 0$, and differentiable for all $x \neq 0$ with derivative

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

2) Consider $f : \mathbb{R} \to \mathbb{R}$, $x \mapsto \exp(x)$. $f(\mathbb{R}) = \mathbb{R}^+$, f is differentiable, and $f'(x) = \exp(x) > 0$.

   Therefore $f^{-1} : \mathbb{R}^+ \to \mathbb{R}$, $x \mapsto \log(x)$ is differentiable, and

$$(f^{-1})'(x) = \frac{1}{\exp(\log(x))} = \frac{1}{x} .$$

Lecture 12: 04/02/11

General powers, exponentials, and logarithms

For $a \in \mathbb{R}$ and $b \in \mathbb{R}^+$, we define

$$b^a = \exp(a \log(b)) .$$

We have $x^a = \exp(a \log(x))$ for $a \in \mathbb{R}$ and $x \in \mathbb{R}^+$, and differentiating using the chain rule gives

$$(x^a)' = \exp(a \log(x)) \frac{a}{x} = a x^{a-1} .$$

We have $b^x = \exp(x \log(b))$ for $b \in \mathbb{R}^+$ and $x \in \mathbb{R}$, and differentiating using the chain rule gives

$$(b^x)' = \exp(x \log(b)) \log(b) = \log(b) b^x .$$

For $a \in \mathbb{R}^+$ and $b \in \mathbb{R}^+ \setminus \{1\}$ we define

$$\log_b(a) = \frac{\log(a)}{\log(b)} .$$

Considering the function $\log_b : \mathbb{R}^+ \to \mathbb{R}$, $x \mapsto \dfrac{\log x}{\log b}$, we find that for $x \in \mathbb{R}^+$

$$b^{\log_b(x)} = \exp\left(\log(b)\frac{\log(x)}{\log(b)}\right) = \exp(\log(x)) = x$$

and that for $x \in \mathbb{R}$

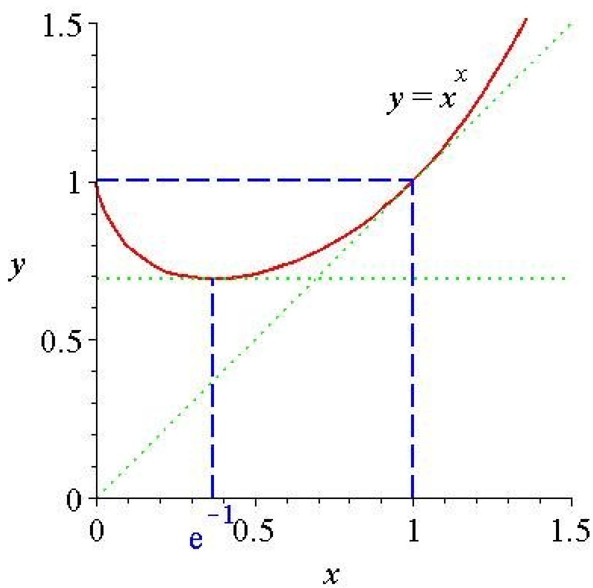
$$\log_b(b^x) = \frac{1}{\log(b)}\log(\exp(\log(b)x)) = \frac{1}{\log(b)}\log(b)x = x ,$$

so that $\log_b$ is the inverse of the function $x \mapsto b^x$.

Example.

The function $f : \mathbb{R}^+ \to \mathbb{R}$, $x \mapsto x^x$ is differentiable, and

$$f'(x) = (x^x)' = (\exp(x\log(x)))' = \exp(x\log x)\left(\log(x) + \frac{x}{x}\right) = (1 + \log x)x^x .$$



5 Higher Order Derivatives

Theorem 5.1 (Second Mean Value Theorem)**.** *Let $f, g : [a, b] \to \mathbb{R}$ be continuous on $[a, b]$ and differentiable on (a, b). Then there exists a $c \in (a, b)$ such that*

$$g'(c)(f(b) - f(a)) = f'(c)(g(b) - g(a)) \,.$$

Proof. Consider the auxiliary function $h : [a, b] \to \mathbb{R}$ given by

$$h(x) = f(x)(g(b) - g(a)) - g(x)(f(b) - f(a)) \,.$$

h is continuous on $[a, b]$ and differentiable on (a, b). By the Mean Value Theorem there exists a $c \in (a, b)$ such that

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

and inserting the definition of h, we find

$$\begin{aligned} & f'(c)(g(b) - g(a)) - g'(c)(f(b) - f(a)) \\ &= \frac{1}{b-a}\Big(f(b)(g(b)-g(a))-g(b)(f(b)-f(a))-f(a)(g(b)-g(a))+g(a)(f(b)-f(a))\Big) = 0 \,. \end{aligned}$$

$\square$

Remark. For $g(x) = x$, this reduces to the Mean Value Theorem.

If the derivative of a function $f : \mathcal{D} \to \mathbb{R}$ is again differentiable, we can consider the second derivative $f'' = (f')'$. We generalise this to higher order derivatives.

Definition 5.2. *Let $f : \mathcal{D} \to \mathbb{R}$ be n times differentiable at $a \in \mathcal{D}$ for some $n \in \mathbb{N}_0$. We call $f^{(n)}$ the n-th derivative of f. It is given by*

$$f^{(0)}(a) = f(a) \quad \text{and} \quad f^{k+1}(a) = (f^{(k)})'(a) \quad \text{for } 0 \leq k < n.$$

We say a function is n times continuously differentiable at $a \in \mathcal{D}$ if $f^{(n)}$ is continuous at a.

Remark. Conventionally, n-th derivatives are denoted by repeating dashes, i.e.

$$f = f^{(0)} \,, \quad f' = f^{(1)} \,, \quad f'' = f^{(2)} \,, \quad f''' = f^{(3)} \,, \quad f'''' = f^{(4)} \,,$$

but this becomes cumbersome for large n.

Example. For $n \in \mathbb{N}$, let $f : \mathbb{R} \to \mathbb{R}$, $x \mapsto |x|x^n$.

Lecture 13: 07/02/11

(a) If $n \geq 1$ then $f'(x) = (n+1)|x|x^{n-1}$:

Consider three cases:

$x > 0$: $f(x) = x^{n+1}$, $f'(x) = (n+1)x^n$

$x < 0$: $f(x) = -x^{n+1}$, $f'(x) = -(n+1)x^n$

$x = 0$: $f'(0) = \lim_{x \to 0} \frac{f(x) - f(0)}{x - 0} = \lim_{x \to 0} |x|x^{n-1} = 0$

(b) For $0 \leq k < n$, $f^{(k)}(x) = \left(\prod_{i=0}^{k-1} (n+1-i) \right) |x|x^{n-k}$:

Use mathematical induction in k:

$k = 0$:

$$f^{(0)}(x) = \left(\prod_{i=0}^{-1} (n+1-i) \right) |x|x^n = |x|x^n \ .$$

$k \to k+1$: For $k < n$,

$$\begin{aligned} f^{(k+1)}(x) = (f^{(k)})'(x) &= \left(\prod_{i=0}^{k-1} (n+1-i) \right) (|x|x^{n-k})' \\ &= \left(\prod_{i=0}^{k-1} (n+1-i) \right) (n+1-k)|x|x^{n-k-1} \\ &= \left(\prod_{i=0}^{k} (n+1-i) \right) |x|x^{n-k} \end{aligned}$$

So f is precisely n times differentiable.

Theorem 5.3 (Taylor's Theorem)**.** *Let $f : [a,x] \to \mathbb{R}$ be n times continuously differentiable (i.e. $f^{(n)}$ exists and is continuous) on $[a,x]$ and $(n+1)$ times differentiable on (a,x). Then there exists a $c \in (a,x)$ such that*

$$f(x) = f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \ldots + \frac{f^{(n)}(a)}{n!}(x-a)^n + \frac{f^{(n+1)}(c)}{(n+1)!}(x-a)^{n+1} \ .$$

Remark. A similar statement holds for $x < a$ (replace $[a,x]$ by $[x,a]$ and (a,x) by (x,a)).

Proof. Let

$$\begin{aligned} F(t) =& f(t) + \frac{f'(t)}{1!}(x-t) + \frac{f''(t)}{2!}(x-t)^2 + \ldots + \frac{f^{(n)}(t)}{n!}(x-t)^n \\ =& \sum_{k=0}^{n} \frac{f^{(k)}(t)}{k!}(x-t)^k \ . \end{aligned}$$

Then F is continuous on $[a, x]$ and differentiable on (a, x), and

$$
\begin{aligned}
F'(t) &= \sum_{k=0}^{n} \frac{f^{(k+1)}(t)}{k!}(x-t)^k - \sum_{k=1}^{n} \frac{f^{(k)}(t)}{(k-1)!}(x-t)^{k-1} \\
&= \frac{f^{(n+1)}(t)}{n!}(x-t)^n \; .
\end{aligned}
$$

Applying Theorem 5.1 to $F(t)$ and $g(t) = (x-t)^{n+1}$ on $[a, x]$ shows that there exists a $c \in (a, x)$ such that $F'(c)(g(x) - g(a)) = g'(c)(F(x) - F(a))$. As $F(x) = f(x)$ and $g(x) = 0$, we find that

$$
\frac{f^{(n+1)}(c)}{n!}(x-c)^n \left(0 - (x-a)^{n+1}\right) = -(n+1)(x-c)^n \left(f(x) - F(a)\right) \; ,
$$

so that

$$
f(x) = F(a) + \frac{f^{(n+1)}(c)}{(n+1)!}(x-a)^{n+1} \; .
$$

□

Remark. We call

$$
T_{n,a}(x) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k
$$

the n-th degree Taylor polynomial of f at a and

$$
R_n = \frac{f^{(n+1)}(c)}{(n+1)!}(x-a)^{n+1}
$$

the Lagrange form of the remainder term. The equation

$$
f(x) = T_{n,a}(x) + R_n
$$

is also called Taylor's formula, and

$$
\sum_{k=0}^{\infty} \frac{f^{(k)}(a)}{k!}(x-a)^k
$$

is called the Taylor series of f at a.

Examples.

Lecture 14: 10/02/11

1) Estimate $e = \exp(1)$ using Taylor's formula:

For $f(x) = \exp(x)$, we have $f^{(k)}(x) = \exp(x)$, and thus

$$T_{n,0}(x) = \sum_{k=0}^{n} \frac{\exp(0)}{k!}(x-0)^k = \sum_{k=0}^{n} \frac{x^k}{k!}$$

and

$$R_n = \frac{\exp(c)}{(n+1)!}x^{n+1} .$$

Taylor's Theorem applied to $f = \exp$ on $[0,1]$ says that there exists a $c \in (0,1)$ such that

$$e = \exp(1) = \sum_{k=0}^{n} \frac{1}{k!} + \frac{\exp(c)}{(n+1)!} .$$

Using that $\exp(c) < \exp(1) < (1+1/1)^2 = 4$, we find

$$\sum_{k=0}^{n} \frac{1}{k!} < e < \sum_{k=0}^{n+1} \frac{1}{k!} + \frac{3}{(n+1)!} .$$

Evaluating this chain of inequalities for $n = 10$ gives the bounds

$$2.718281826 < e < 2.718281901 .$$

Moreover, as

$$\left| e - \sum_{k=0}^{n} \frac{1}{k!} \right| < \frac{4}{(n+1)!} ,$$

we find

$$e = \sum_{k=0}^{\infty} \frac{1}{k!} .$$

2) Show that $\exp(x) = \sum_{k=0}^{\infty} \frac{x^k}{k!}$ for all $x \in \mathbb{R}$:

Taylor's Theorem applied to $f = \exp$ on $[0,x]$ for $x > 0$, or on $[x,0]$ for $x < 0$, says that there exists a c with $|c| < |x|$ such that

$$|\exp(x) - T_{n,0}(x)| = |R_n| = \left| \frac{\exp(c)}{(n+1)!}x^{n+1} \right| .$$

Now $\lim_{n\to\infty} \frac{x^n}{n!} = 0$, so that $R_n \to 0$ as $n \to \infty$.

3) Show that $\log(x) = \sum\limits_{k=1}^{\infty} \dfrac{(-1)^{k-1}}{k}(x-1)^k$ for $1 < x \leq 2$:

For $f(x) = \log(x)$, we have $f'(x) = 1/x$, $f''(x) = -1/x^2$, $f''' = 2/x^3$, ... From this we conjecture that for $k \geq 1$

$$f^{(k)}(x) = \frac{(-1)^k (k-1)!}{x^k} \ .$$

holds and prove this via mathematical induction (this is a standard argument which we omit here). We choose $a = 1$ and get

$$T_{n,1}(x) = \sum_{k=0}^{n} \frac{f^{(k)}(1)}{k!}(x-1)^k = \sum_{k=1}^{n} \frac{(-1)^{k-1}}{k}(x-1)^k$$

and

$$R_n = \frac{f^{(n+1)}(c)}{(n+1)!}(x-1)^{n+1} = \frac{(-1)^n}{n+1}\left(\frac{x-1}{c}\right)^{n+1} \ .$$

Taylor's Theorem applied to $f = \log$ on $[1, x]$ for $1 < x \leq 2$ says that there exists a c with $c \in (1, x) \subseteq (1, 2)$ such that

$$|\log(x) - T_{n,1}(x)| = |R_n| \leq \frac{1}{n+1}\left|\frac{x-1}{c}\right|^{n+1} \ .$$

Now $0 < x - 1 \leq 1$ and $1 < c < x \leq 2$, so that $\left|\dfrac{x-1}{c}\right| < 1$. Therefore $R_n \to 0$ as $n \to \infty$.

(It can be shown that this result holds not only for $1 < x \leq 2$ but for $0 < x < 2$, or, equivalently, for $|x - 1| < 1$.)

We return now to our discussion of the exponential function.

(I) $\exp(x) = \sum\limits_{n=0}^{\infty} \dfrac{x^n}{n!}$.

Proof. From Example 2) above. □

(J) $\lim\limits_{x\to\infty} x^n \exp(-x) = 0$ for all $n \in \mathbb{N}_0$.

Proof. From (I) it follows that $\exp(x) > \dfrac{x^{n+1}}{(n+1)!}$ for $x > 0$ and $n \in \mathbb{N}_0$. Therefore

$$0 < x^n \exp(-x) < \frac{(n+1)!}{x} \ ,$$

and, taking the limit of $x \to \infty$,

$$0 \leq \lim_{x\to\infty} x^n \exp(-x) \leq \lim_{x\to\infty} \frac{(n+1)!}{x} = 0 \, .$$

$\square$

Lecture 15: 11/02/11

Theorem 5.4. *Let $f : \mathbb{R} \to \mathbb{R}$ be given by*

$$f(x) = \begin{cases} \exp(-1/x) & x > 0 \, , \\ 0 & x \leq 0 \, . \end{cases}$$

Then

$$f^{(k)}(x) = \begin{cases} P_k(1/x)\exp(-1/x) & x > 0 \, , \\ 0 & x \leq 0 \, , \end{cases}$$

where P_k is a polynomial of degree at most $2k$.

Corollary. *The n-th degree Taylor polynomial of f at zero is $T_{n,0}(x) = 0$.*

Remark. While the Taylor polynomial can be a good approximation to a function, it need not be. In this case all Taylor polynomials are zero, so $f(x) = R_n$ and the remainder does not get small.

When looking for the cause of this, one finds that close to zero the derivatives of f become arbitrarily large. From the Lagrange form of the remainder we know that for each $n \in \mathbb{N}$ there exists a $c_n \in (0, x)$ such that

$$\exp(-1/x) = R_{n-1} = \frac{f^{(n)}(c_n)}{n!} x^n \, .$$

This implies that for x fixed,

$$f^{(n)}(c_n) = \frac{n!}{x^n} \exp(-1/x) \to \infty \quad \text{as } n \to \infty.$$

In other words, no matter how close x is to zero, there exists a sequence (c_n) with $c_n \in (0, x)$ such that $\lim_{n\to\infty} f^{(n)}(c_n) = \infty$.

Proof (Theorem 5.4). We use mathematical induction in k. In the case $k = 0$ we only need to choose $P_0(1/x) = 1$. For the inductive step from k to $k+1$, we need to compute the derivative of

$$f^{(k)}(x) = \begin{cases} P_k(1/x)\exp(-1/x) & x > 0\ , \\ 0 & x \leq 0\ . \end{cases}$$

For $x < 0$ we find $f^{(k+1)}(x) = 0$, and for $x > 0$ we compute

$$\begin{aligned} f^{(k+1)}(x) =& P_k'(1/x)(-1/x^2)\exp(-1/x) + P_k(1/x)\exp(-1/x)(1/x^2) \\ =& (1/x^2)\left(P_k(1/x) - P_k'(1/x)\right)\exp(-1/x) \\ =& P_{k+1}(1/x)\exp(-1/x)\ , \end{aligned}$$

where $P_{k+1}(t) = t^2(P_k(t) - P_k'(t))$ is a polynomial of degree at most $2k+2$. For $x = 0$ we compute the left and right limits of the difference quotient separately. We have $\lim\limits_{x\to 0^-} \dfrac{f^{(k)}(x) - f^{(k)}(0)}{x - 0} = 0$ and find

$$\begin{aligned} \lim_{x\to 0^+} \frac{f^{(k)}(x) - f^{(k)}(0)}{x - 0} &= \lim_{x\to 0^-} (1/x)P_k(1/x)\exp(-1/x) \\ &= \lim_{t\to\infty} tP_k(t)\exp(-t) = 0 \end{aligned}$$

by (J). This concludes the inductive step. □

Theorem 5.5 (L'Hospital's Rule)**.** *Let $f, g : \mathcal{D} \to \mathbb{R}$ be differentiable for $|x - a| < \varepsilon$ and let $g'(x) \neq 0$ for $0 < |x - a| < \varepsilon$. If $\lim\limits_{x\to a} f(x) = \lim\limits_{x\to a} g(x) = 0$ and if $\lim\limits_{x\to a} \dfrac{f'(x)}{g'(x)}$ exists, then $\lim\limits_{x\to a} \dfrac{f(x)}{g(x)}$ exists and*

$$\lim_{x\to a} \frac{f(x)}{g(x)} = \lim_{x\to a} \frac{f'(x)}{g'(x)}\ .$$

Proof. We first show that $g(x) \neq 0$ for $0 < |x - a| < \varepsilon$. By assumption $g(a) = 0$. If $g(b) = 0$ for some b with $0 < |x - b| < \varepsilon$, then we apply Rolle's Theorem to g and find that there exists a c between a and b such that $g'(c) = 0$, but this contradicts the assumption that $g'(x) \neq 0$ for $0 < |x - a| < \varepsilon$.

Next, by the Second Mean Value Theorem applied to f and g, there exists a c between a and x such that

$$g'(c)(f(x) - f(a)) = f'(c)(g(x) - g(a)) \; .$$

By assumption $f(a) = g(a) = 0$, and as $g(x) \neq 0$ as well as $g'(c) \neq 0$, we can write

$$\frac{f(x)}{g(x)} = \frac{f'(c)}{g'(c)} \; .$$

Finally, when $x \to a$ then necessarily $c \to a$, so that

$$\lim_{x \to a} \frac{f(x)}{g(x)} = \lim_{c \to a} \frac{f'(c)}{g'(c)} \; .$$

$\square$

Lecture 16: 14/02/11

Examples.

1) Apply l'Hospital's rule:

$$\lim_{x \to 0} \frac{\sqrt{1+2x} - \sqrt{1+x}}{x} = \lim_{x \to 0} \frac{1/\sqrt{1+2x} - 1/2\sqrt{1+x}}{1} = 1 - \frac{1}{2} = \frac{1}{2} \; .$$

2) Apply l'Hospital's rule twice:

$$\lim_{x \to 0} \frac{\exp(x) - 1 - x}{x^2} = \lim_{x \to 0} \frac{\exp(x) - 1}{2x} = \lim_{x \to 0} \frac{\exp(x)}{2} = \frac{1}{2} \; .$$

The rule also holds if $f(x), g(x) \to \infty$:

3)

$$\lim_{x \to 0} x \log(|x|) = \lim_{x \to 0} \frac{\log(|x|)}{1/x} = \lim_{x \to 0} \frac{1/x}{-1/x^2} = 0 \; .$$

4)

$$\lim_{x \to 0} |x|^x = \lim_{x \to 0} \exp(x \log(|x|)) = \exp(\lim_{x \to 0} x \log(|x|)) = \exp(0) = 1 \; .$$

6 Definition of the Riemann Integral

Let $I = [a, b]$ for $a < b$ be an interval. Given

$$a = x_0 < x_1 < x_2 < \ldots < x_{n-1} < x_n = b\,,$$

we call

$$P = \{x_0, x_1, x_2, \ldots, x_{n-1}, x_n\}$$

a partition of I. We denote the set of all partitions of I by $\mathcal{P}$.

We denote $I_i = [x_{i-1}, x_i]$ and $\Delta x_i = x_i - x_{i-1}$ for $i = 1, 2, \ldots, n$. A partition is called equidistant, if all I_i have equal length Δx_i.

P_2 is called a refinement of P_1 if $P_1 \subseteq P_2$. Two partitions P_1 and P_2 have a common refinement, for example $P = P_1 \cup P_2$ is such a refinement. The notion of refinement defines a partial order on $\mathcal{P}$.

$\sigma(P) = \max\{\Delta x_i : i = 1, 2, \ldots, n\}$ is called the mesh of P. $P_1 \subseteq P_2$ implies $\sigma(P_1) \geq \sigma(P_2)$, i.e. a refinement has a smaller mesh.

Examples.

1) $P = \left\{a, a + \frac{b-a}{n}, a + 2\frac{b-a}{n}, \ldots, a + n\frac{b-a}{n} = b\right\}$ is an equidistant partition of $[a, b]$ with $\sigma(P) = \frac{b-a}{n}$.

2) $P_2 = \left\{0, \frac{1}{2n}, \frac{2}{2n}, \ldots, \frac{2n}{2n}\right\}$ is a refinement of $P_1 = \left\{0, \frac{1}{n}, \frac{2}{n}, \ldots, \frac{n}{n}\right\}$. $\sigma(P_2) = \frac{1}{2n} < \sigma(P_1) = \frac{1}{n}$. Note that $P_3 = \left\{0, \frac{1}{n+1}, \frac{2}{n+1}, \ldots, \frac{n+1}{n+1}\right\}$ is not a refinement of P_1.

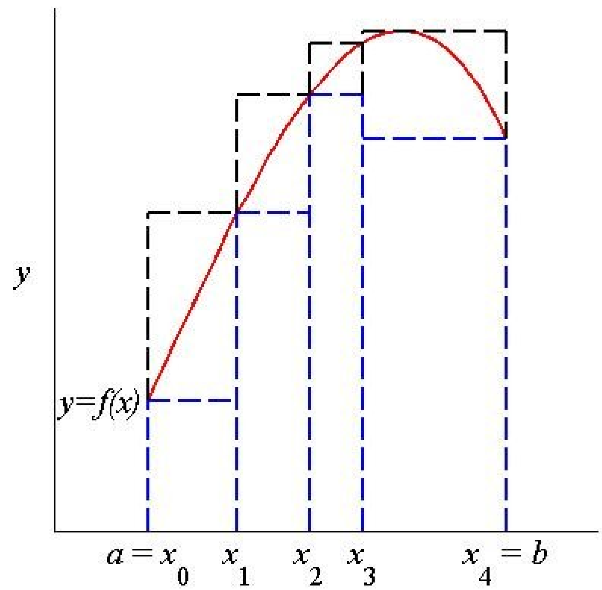
Definition 6.1. *Let $f : [a, b] \to \mathbb{R}$ be bounded and $P = \{x_0, x_1, \ldots, x_n\}$ be a partition of $[a, b]$. We define the upper sum of f with respect to P*

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

and the lower sum of f with respect to P

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

where $M_i = \sup\{f(x) : x \in I_i\}$ and $m_i = \inf\{f(x) : x \in I_i\}$.



Remark: Geometrically, if f is positive then the area A between the x-axis and the graph of $f(x)$ from a to b should satisfy

$$L(f,P) \leq A \leq U(f,P) \ .$$

Lecture 17: 17/02/11

Example.

Given $f : [-2,1] \to \mathbb{R}$, $x \mapsto x^2 - x$, consider the partition $P = \{-2,-1,1\}$. Then $I_1 = [-2,-1]$ and $I_2 = [-1,1]$. We find (and make sure you understand why!)

$$M_1 = 6 \ , \qquad m_1 = 2 \ ,$$
$$M_2 = 2 \ , \qquad m_2 = -1/4 \ ,$$

and this together with $\Delta x_1 = 1$ and $\Delta x_2 = 2$ implies

$$U(f,P) = 6 \cdot 1 + 2 \cdot 2 = 10 \ ,$$
$$L(f,P) = 2 \cdot 1 + (-1/4) \cdot 2 = 3/2 \ .$$

Theorem 6.2. *Let $f : [a,b] \to \mathbb{R}$ be bounded. If P_2 is a refinement of the partition P_1 then*

(1) $U(f,P_2) \leq U(f,P_1)$, and

(2) $L(f,P_2) \geq L(f,P_1)$.

Proof. Let $P_1 = \{x_0, x_1, \ldots, x_n\}$ and $P_2 = P_1 \cup \{y\}$. If $x_{i-1} < y < x_i$ then

$$\begin{aligned} M' =& \sup\{f(x) : x \in [x_{i-1}, y]\} \leq M_i \quad \text{and} \\ M'' =& \sup\{f(x) : x \in [y, x_i]\} \leq M_i \ . \end{aligned}$$

Therefore $M_i \Delta x_i = M_i(y - x_{i-1}) + M_i(x_i - y) \geq M'(y - x_{i-1}) + M''(x_i - y)$, so that

$$\begin{aligned} U(f, P_1) =& \sum_{\substack{j=1 \\ j \neq i}}^{n} M_j \Delta x_j + M_i \Delta x_i \\ \geq& \sum_{\substack{j=1 \\ j \neq i}}^{n} M_j \Delta x_j + M'(y - x_{i-1}) + M''(x_i - y) \\ =& U(f, P_2) \ . \end{aligned}$$

Now let P_2 be an arbitrary refinement of P_1. Then P_2 is obtained from P_1 by adding a finite number of points y_j, creating a chain of partitions

$$P_1 = Q_0 \subseteq Q_1 \subseteq \ldots \subseteq Q_r = P_2$$

and

$$U(f, Q_0) \geq U(f, Q_1) \geq \ldots \geq U(f, Q_r) \ .$$

A similar argument leads to $L(f, P_2) \geq L(f, P_1)$. $\square$

Corollary. *Let P_1, P_2 be partitions of $[a, b]$. Then*

$$L(f, P_1) \leq U(f, P_2) \ .$$

Proof. Let $P = P_1 \cup P_2$ be a common refinement of P_1 and P_2. Then

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

$\square$

Corollary. *$\{U(f, P) : P \in \mathcal{P}\}$ is bounded below and $\{L(f, P) : P \in \mathcal{P}\}$ is bounded above.*

Definition 6.3. *Let $f : [a, b] \to \mathbb{R}$ be bounded. We call*

$$\int_a^{*b} f(x)\,dx = \inf\{U(f, P) : P \in \mathcal{P}\}$$

the upper integral of f and

$$\int_{*a}^{b} f(x)\,dx = \sup\{L(f, P) : P \in \mathcal{P}\}$$

the lower integral of f.

Remark. Clearly,

$$\int_a^{*b} f(x)\,dx \geq \int_{*a}^{b} f(x)\,dx \,.$$

Definition 6.4. *A bounded function $f : [a, b] \to \mathbb{R}$ is Riemann integrable if the upper and lower integral of f agree. The quantity*

$$\int_a^b f(x)\,dx = \int_a^{*b} f(x)\,dx = \int_{*a}^{b} f(x)\,dx$$

is called the Riemann integral of f over $[a, b]$.

Lecture 18: 18/02/11

Theorem 6.5. *Let $f : [a, b] \to \mathbb{R}$ be bounded. f is Riemann integrable if and only if*

$$\forall \varepsilon > 0\, \exists P \in \mathcal{P} : U(f, P) - L(f, P) < \varepsilon \,.$$

Proof. "$\Rightarrow$" Let f be Riemann integrable and

$$A = \sup\{L(f, P) : P \in \mathcal{P}\} = \inf\{U(f, P) : P \in \mathcal{P}\} \,.$$

Then for a given $\varepsilon > 0$ there exist $P_1, P_2 \in \mathcal{P}$ such that

$$A - \frac{\varepsilon}{2} < L(f, P_1) \quad \text{and} \quad U(f, P_2) < A + \frac{\varepsilon}{2} \,.$$

For $P = P_1 \cup P_2$ we have

$$U(f, P) - L(f, P) \leq U(f, P_2) - L(f, P_1) < A + \frac{\varepsilon}{2} - \left(A - \frac{\varepsilon}{2}\right) = \varepsilon \,.$$

"$\Leftarrow$" If for any $\varepsilon > 0$ there is a $P \in \mathcal{P}$ such that

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

then

$$\int_a^{*b} f(x)\,dx - \int_{*a}^b f(x)\,dx \le U(f,P) - L(f,P) < \varepsilon \ .$$

As $\varepsilon > 0$ can be arbitrarily small,

$$\int_a^{*b} f(x)\,dx = \int_{*a}^b f(x)\,dx \ ,$$

so f is Riemann integrable.

$\square$

Examples.

1) Let $f : [a,b] \longrightarrow \mathbb{R}$, $x \mapsto c$ be the constant function.

   For $P = \{x_0, x_1, \ldots, x_n\}$ we find $m_i = M_i = c$ and thus

$$U(f,P) = \sum_{i=1}^n M_i \Delta x_i = c \sum_{i=1}^n \Delta x_i = c(b-a)$$

   and

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

   Therefore f is Riemann integrable with

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

2) Let $f : [a,b] \longrightarrow \mathbb{R}$, $x \mapsto \begin{cases} 1 & x \in \mathbb{Q} \ , \\ 0 & x \notin \mathbb{Q} \ . \end{cases}$

   For $P = \{x_0, x_1, \ldots, x_n\}$ we find $m_i = 0$ and $M_i = 1$ and thus

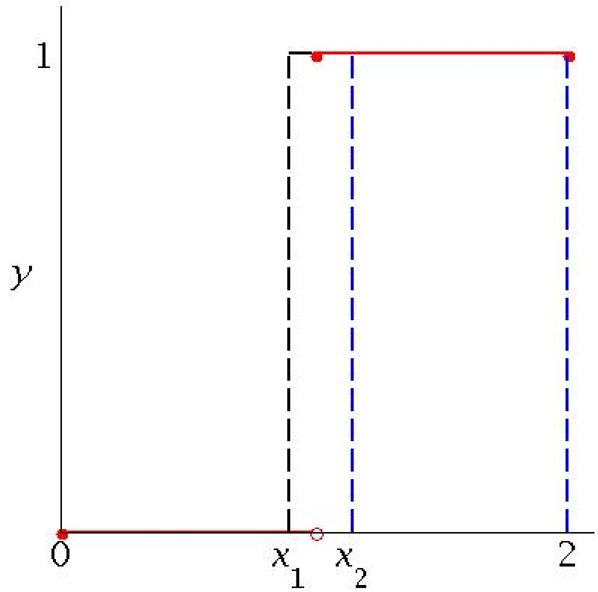
$$U(f,P) = \sum_{i=1}^n M_i \Delta x_i = \sum_{i=1}^n \Delta x_i = (b-a)$$

   and

$$L(f,P) = \sum_{i=1}^n m_i \Delta x_i = 0 \ .$$

   Therefore f is <u>not</u> Riemann integrable.

3) Let $f : [0,2] \to \mathbb{R}$, $x \mapsto \begin{cases} 0 & x \in [0,1) \ , \\ 1 & x \in [1,2] \ . \end{cases}$



Choose $0 < x_1 < 1 < x_2 < 2$ with $x_2 - x_1 < \varepsilon$ and $P = \{0, x_1, x_2, 2\}$. Then

$$M_1 = m_1 = 0 \ , \quad M_2 = 1 \ , \quad m_2 = 0 \ , \quad M_3 = m_3 = 1 \ ,$$

and thus

$$U(f,P) = 0 \cdot (x_1 - 0) + 1 \cdot (x_2 - x_1) + 1 \cdot (2 - x_2)$$

and

$$L(f,P) = 0 \cdot (x_1 - 0) + 0 \cdot (x_2 - x_1) + 1 \cdot (2 - x_2) \ ,$$

so that

$$U(f,P) - L(f,P) = x_2 - x_1 < \varepsilon \ .$$

Therefore f is Riemann integrable with

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

Theorem 6.6. *Every increasing or decreasing function $f : [a, b] \to \mathbb{R}$ is Riemann integrable.*

Proof. Assume without loss of generality that f is increasing. Then $f(a) \leq f(x) \leq f(b)$ for $x \in [a, b]$, so f is bounded.

Let $\varepsilon > 0$. Choose a partition P with a mesh

$$\sigma(P) \leq \frac{\varepsilon}{f(b) - f(a) + 1} .$$

As f is increasing, $M_i = f(x_i)$ and $m_i = f(x_{i-1})$, so that

$$\begin{aligned} U(f, P) - L(f, P) =& \sum_{i=1}^{n} (M_i - m_i)\Delta x_i = \sum_{i=1}^{n} (f(x_i) - f(x_{i-1}))\Delta x_i \\ \leq& \sum_{i=1}^{n} (f(x_i) - f(x_{i-1}))\sigma(P) = (f(b) - f(a)\sigma(P) \\ \leq& (f(b) - f(a)) \frac{\varepsilon}{1 + f(b) - f(a)} < \varepsilon . \end{aligned}$$

By Theorem 6.5, f is Riemann integrable. $\square$

Lecture 19: 28/02/11

Definition 6.7. *A function $f : \mathcal{D} \to \mathbb{R}$ is uniformly continuous if*

$$\forall \varepsilon > 0 \; \exists \delta > 0 \; \forall c \in \mathcal{D} \; \forall x \in \mathcal{D}, \; |x - c| < \delta : |f(x) - f(c)| < \varepsilon .$$

Remark. This means that δ is chosen independently of c. The statement that a function $f : \mathcal{D} \to \mathbb{R}$ is merely *continuous* is equivalent to

$$\forall c \in \mathcal{D} \; \forall \varepsilon > 0 \; \exists \delta > 0 \; \forall x \in \mathcal{D}, \; |x - c| < \delta : |f(x) - f(c)| < \varepsilon .$$

Note how the statement "$\forall c \in \mathcal{D}$" has moved places. Clearly a uniformly continuous function is continuous, but a continuous function need not be uniformly continuous.

Example.

$f : \mathbb{R} \to \mathbb{R}$, $x \mapsto x^2$ is continuous, but not uniformly continuous:

To show this, assume that f is uniformly continuous. Then for $\varepsilon = 1$, say, there exists a $\delta > 0$ such that $|x - c| < \delta \Rightarrow |x^2 - c^2| < \varepsilon = 1$ for all $x, c \in \mathbb{R}$. As δ is independent of c, this should be true for all c, for example if $c = 1/\delta$. But then, for $x = c + \delta/2$, we find $|x - c| = \delta/2 < \delta$ and

$$|x^2 - c^2| = |(c + \delta/2)^2 - c^2| = |c\delta + \delta^2/4| = 1 + \delta^2/4 > 1$$

which is a contradiction.

This example works because the domain is not closed and bounded. Continuous functions on closed and bounded domains are in fact uniformly continuous. We shall see below that this is an important ingredient in proving Riemann integrability of continuous functions.

Theorem 6.8. *Let $f : [a, b] \to \mathbb{R}$ be continuous. Then f is uniformly continuous.*

Proof. Suppose f is continuous on $[a, b]$ but *not* uniformly continuous. Then

$$\exists \varepsilon > 0\ \forall \delta > 0\ \exists c \in \mathcal{D}\ \exists x \in \mathcal{D},\ |x - c| < \delta : |f(x) - f(c)| \geq \varepsilon .$$

So there exists $\varepsilon > 0$ such that for $\delta = 1/n$ there exist $c_n, x_n \in \mathcal{D}$ with

$$|x_n - c_n| < \delta \quad \text{but} \quad |f(x_n) - f(c_n)| \geq \varepsilon .$$

Now (and this is the key step!) using Bolzano-Weierstraß, (c_n) contains a convergent subsequence. Therefore there exist $(n_r)_{r\in\mathbb{N}}$ such that

(a) $\lim\limits_{r\to\infty} c_{n_r} = d$ for some $d \in [a, b]$,

(b) $\lim\limits_{r\to\infty} x_{n_r} = d$ (as $|x_{n_r} - d| \leq |x_{n_r} - c_{n_r}| + |c_{n_r} - d|$), and

(c) $\lim\limits_{r\to\infty} f(c_{n_r}) = f(d)$ and $\lim\limits_{r\to\infty} f(x_{n_r}) = f(d)$.

But by assumption for all n, $|f(x_n) - f(c_n)| \geq \varepsilon$, which is a contradiction. □

Theorem 6.9. *Every continuous function $f : [a, b] \to \mathbb{R}$ is Riemann integrable.*

Lecture 20: 03/03/11

Proof. By Theorem 6.8, f is uniformly continuous on $[a, b]$, so that

$$\forall \varepsilon > 0\ \exists \delta > 0\ \forall c, c' \in [a, b],\ |c - c'| < \delta : |f(c) - f(c')| < \frac{\varepsilon}{b - a} .$$

Now choose a partition P with $\sigma(P) < \delta$. Then on each interval I_i, f assumes its minimum m_i at some c_i and its maximum M_i at some c'_i, so that $m_i = f(c_i)$ and $M_i = f(c'_i)$. As $|c_i - c'_i| \leq \sigma(P) < \delta$,

$$M_i - m_i = |f(c'_i) - f(c_i| < \frac{\varepsilon}{b - a} .$$

Therefore

$$U(f,P) - L(f,P) = \sum_{i=0}^{n}(M_i - m_i)\Delta x_i < \frac{\varepsilon}{b-a}\sum_{i=1}^{n}\Delta x_i = \varepsilon\ .$$

By Theorem 6.5, f is Riemann integrable. $\square$

Examples.

1) $f:[a,b] \longrightarrow \mathbb{R}$, $f(x) = x$:

   f is increasing, therefore Riemann integrable. To compute the Riemann integral, choose

$$P_n = \{a, a+\Delta, a+2\Delta, \ldots, a+n\Delta = b\}$$

   where $\Delta = \dfrac{b-a}{n}$. The mesh of the partition is given by $\sigma(P_n) = \Delta = \dfrac{b-a}{n}$. We find

$$m_i = a + (i-1)\Delta\ , \quad \text{and} \quad M_i = a + i\Delta\ .$$

   Therefore

$$\begin{aligned} L(f,P_n) =& \sum_{i=1}^{n}(a + (i-1)\Delta)\Delta \\ =& an\Delta + \frac{n(n-1)}{2}\Delta^2 \\ =& a(b-a) + \frac{1}{2}(b-a)^2\left(1 - \frac{1}{n}\right)\ . \end{aligned}$$

   Therefore

$$\int_{*a}^{b} f(x)\,dx = \lim_{n\to\infty} L(f,P_n) = a(b-a) + \frac{1}{2}(b-a)^2 = \frac{b^2}{2} - \frac{a^2}{2}\ .$$

   As we already know that f is Riemann integrable, we now conclude that

$$\int_a^b f(x)\,dx = \int_{*a}^{b} f(x)\,dx = \frac{b^2}{2} - \frac{a^2}{2}\ .$$

   If we didn't know that f was Riemann integrable, a computation of the upper sums shows that

$$U(f,P_n) = a(b-a) + \frac{1}{2}(b-a)^2\left(1 + \frac{1}{n}\right)\ .$$

Just as we should, we find that $U(f,P_n)-L(f,P_n) = (b-a)^2\frac{1}{n} \to 0$ as $n \to \infty$, and that

$$\int_a^{*b} f(x)\,dx = \frac{b^2}{2} - \frac{a^2}{2} = \int_{*a}^{b} f(x)\,dx\ .$$

2) $f : [1,a] \to \mathbb{R}$, $f(x) = 1/x$:

f is decreasing, therefore Riemann integrable. To compute the Riemann integral, choose

$$P_n = \{1 = q^0, q^1, q^2, \ldots, q^n = a\}$$

where $q = \sqrt[n]{a}$. We find

$$\Delta x_i = q^i - q^{i-1} = (q-1)q^{i-1}\ ,$$

so that the mesh of the partition is given by $\sigma(P_n) = (q-1)q^{n-1}$. We find

$$m_i = \frac{1}{q^i}\ , \quad \text{and} \quad M_i = \frac{1}{q^{i-1}}\ .$$

Therefore

$$\begin{aligned} L(f,P_n) &= \sum_{i=1}^{n} \frac{1}{q^i}(q-1)q^{i-1} \\ &= \sum_{i=1}^{n} \frac{1}{q}(q-1) = n\left(1-\frac{1}{q}\right) = n\left(1-\frac{1}{\sqrt[n]{a}}\right)\ . \end{aligned}$$

Therefore

$$\begin{aligned} \int_{*1}^{a} f(x)\,dx &= \lim_{n\to\infty} L(f,P_n) = \lim_{n\to\infty} n\left(1-a^{-1/n}\right) \\ &= \lim_{n\to\infty} n\left(1-\exp\left(-\frac{1}{n}\log(a)\right)\right) \\ &= \lim_{t\to 0} \frac{1-\exp(-t\log(a))}{t} \\ &= \lim_{t\to 0} \frac{\log(a)\exp(-t\log(a))}{1} = \log(a)\ . \end{aligned}$$

As we already know that f is Riemann integrable, we now conclude that

$$\int_1^a f(x)\,dx = \int_{*1}^{a} f(x)\,dx = \log(a)\ .$$

If we didn't know that f was Riemann integrable, a computation of the upper sums shows that

$$U(f, P_n) = n(q-1)\ .$$

Just as we should, we find that $U(f, P_n) - L(f, P_n) = n(q-1)^2/q \to 0$ as $n \to \infty$, and that

$$\int_1^{*a} f(x)\, dx = \log(a) = \int_{*1}^{a} f(x)\, dx\ .$$

7 Properties of the Riemann Integral

Lecture 21: 04/03/11

Theorem 7.1. *Let $f : [a, b] \to \mathbb{R}$ be Riemann integrable. If $[c, d] \subseteq [a, b]$ then f is Riemann integrable on $[c, d]$.*

Proof. Let $\varepsilon > 0$. Then there exists a partition P of $[a, b]$ such that $U(f, P) - L(f, P) < \varepsilon$. If we let

$$P' = P \cup \{c, d\} = \{x_0, x_1, \ldots, x_k = c, x_{k+1}, \ldots, x_{k+r} = d, x_{k+r+1}, \ldots, x_n\}$$

then

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

. Now let

$$P'' = \{x_k, x_{k+1}, \ldots, x_{k+r}\} .$$

This is a partition of $[c, d]$ with

$$\begin{aligned} U(f, P'') - L(f, P'') &= \sum_{i=k+1}^{k+r} (M_i - m_i)\Delta x_i \\ &\leq \sum_{i=1}^{n} (M_i - m_i)\Delta x_i \\ &= U(f, P') - L(f, P') < \varepsilon . \end{aligned}$$

Thus f is Riemann integrable on $[c, d]$. ☐

Theorem 7.2. *Let $f : [a, b] \to \mathbb{R}$ be Riemann integrable on $[a, c]$ and $[c, b]$ where $a < c < b$. Then f is Riemann integrable on $[a, b]$ and*

$$\int_a^b f(x)\, dx = \int_a^c f(x)\, dx + \int_c^b f(x)\, dx .$$

Proof. Let $\varepsilon > 0$ and let P_1 and P_2 be partitions of $[a, c]$ and $[c, b]$, respectively, with

$$U(f, P_1) - L(f, P_1) < \frac{\varepsilon}{2} \quad \text{and} U(f, P_2) - L(f, P_2) < \frac{\varepsilon}{2} .$$

Then $P = P_1 \cup P_2$ is a partition of $[a, b]$ with

$$U(f, P) - L(f, P) = U(f, P_1) + U(f, P_2) - L(f, P_1) - L(f, P_2) < \varepsilon$$

and hence f is Riemann integrable on $[a, b]$. Moreover, as

$$L(f, P_1) \leq \int_a^c f(x)\,dx \leq U(f, P_1) \quad \text{and} \quad L(f, P_2) \leq \int_c^b f(x)\,dx \leq U(f, P_2)$$

we have

$$L(f, P) \leq \int_a^c f(x)\,dx + \int_c^b f(x)\,dx \leq U(f, P)\ .$$

Clearly we also have

$$L(f, P) \leq \int_a^b f(x)\,dx \leq U(f, P)\ ,$$

and taking differences leads to

$$L(f, P) - U(f, P) \leq \int_a^c f(x)\,dx + \int_c^b f(x)\,dx - \int_a^b f(x)\,dx \leq U(f, P) - L(f, P)$$

or, equivalently,

$$\left|\int_a^c f(x)\,dx + \int_c^b f(x)\,dx - \int_a^b f(x)\,dx\right| \leq U(f, P) - L(f, P)\ .$$

Therefore. we have shown that for all $\varepsilon > 0$

$$\left|\int_a^c f(x)\,dx + \int_c^b f(x)\,dx - \int_a^b f(x)\,dx\right| < \varepsilon$$

so that

$$\int_a^c f(x)\,dx + \int_c^b f(x)\,dx = \int_a^b f(x)\,dx\ .$$

□

Remark. Because of Theorem 7.2 it makes sense to define for $a > b$

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

Then, if f is Riemann integrable on a closed and bounded interval I, and $a, b, c \in I$, we have

$$\int_a^c f(x)\,dx + \int_c^b f(x)\,dx = \int_a^b f(x)\,dx\ .$$

Theorem 7.3. *Let $f, g : [a,b] \to \mathbb{R}$ be bounded and P be a partition of $[a,b]$. Then*

(a) $U(f+g,P) \le U(f,P) + U(g,P)$, and

(b) $L(f+g,P) \ge L(f,P) + L(g,P)$.

Proof. For a subinterval I_i of the partition P, we write $M_i(h) = \sup\{h(x) : x \in I_i\}$ and $m_i(h) = \inf\{h(x) : x \in I_i\}$.

(a) On a subinterval I_i of the partition P we have

$$\begin{aligned} M_i(f+g) &= \sup\{f(x)+g(x) : x \in I_i\} \\ &\le \sup\{f(x) : x \in I_i\} + \sup\{g(x) : x \in I_i\} = M_i(f) + M_i(g)\ . \end{aligned}$$

Thus

$$\begin{aligned} U(f+g,P) &= \sum_{i=1}^{n} M_i(f+g)\Delta x_i \\ &\le \sum_{i=1}^{n} M_i(f)\Delta x_i + \sum_{i=1}^{n} M_i(g)\Delta x_i = U(f,P) + U(g,P)\ . \end{aligned}$$

(b) Similarly,

$$\begin{aligned} L(f+g,P) &= \sum_{i=1}^{n} m_i(f+g)\Delta x_i \\ &\ge \sum_{i=1}^{n} m_i(f)\Delta x_i + \sum_{i=1}^{n} m_i(g)\Delta x_i = L(f,P) + L(g,P)\ . \end{aligned}$$

$\square$

Lecture 22: 07/03/11

Theorem 7.4. *Let $f, g : [a,b] \to \mathbb{R}$ be Riemann integrable and $c \in \mathbb{R}$. Then $f+g$ and cf are Riemann integrable, and*

$$\int_a^b f(x) + g(x)\, dx = \int_a^b f(x)\, dx + \int_a^b g(x)\, dx \quad \text{and}$$
$$\int_a^b cf(x)\, dx = c\int_a^b f(x)\, dx\ .$$

Proof. (a) Let $\varepsilon > 0$. Then there exist partitions P_1 and P_2 of $[a, b]$ such that

$$U(f, P_1) - L(f, P_1) < \frac{\varepsilon}{2} \quad \text{and} \quad U(g, P_2) - L(g, P_2) < \frac{\varepsilon}{2} \, .$$

Let $P = P_1 \cup P_2$. Then

$$\begin{aligned} U(f, P) - L(f, P) &\le U(f, P_1) - L(f, P_1) < \frac{\varepsilon}{2} \quad \text{and} \\ U(g, P) - L(g, P) &\le U(g, P_2) - L(g, P_2) < \frac{\varepsilon}{2} \, . \end{aligned}$$

By Theorem 7.3 it follows that

$$U(f + g, P) - L(f + g, P) \le U(f, P) + U(g, P) - L(f, P) - L(g, P) < \varepsilon \, ,$$

so $f + g$ is Riemann integrable on $[a, b]$.

We proceed now as in the proof of Theorem 7.2. As

$$L(f, P) \le \int_a^b f(x)\, dx \le U(f, P) \quad \text{and} \quad L(g, P) \le \int_a^b g(x)\, dx \le U(g, P)$$

we have

$$L(f, P) + L(g, P) \le \int_a^b f(x)\, dx + \int_a^b g(x)\, dx \le U(f, P) + U(g, P) \, .$$

Clearly we also have

$$\begin{aligned} L(f, P) + L(g, P) \le L(f + g, P) \le \int_a^b f(x) + g(x)\, dx \\ \le U(f + g, P) \le U(f, P) + U(g, P) \, , \end{aligned}$$

and taking differences leads to

$$\begin{aligned} \left| \int_a^b f(x)\, dx + \int_a^b g(x)\, dx - \int_a^b f(x) + g(x)\, dx \right| \\ \le U(f, P) + U(g, P) - L(f, P) - L(g, P) \, . \end{aligned}$$

Therefore we have shown that for all $\varepsilon > 0$

$$\left| \int_a^b f(x) + g(x)\, dx - \int_a^b f(x)\, dx - \int_a^b g(x)\, dx \right| < \varepsilon \, ,$$

so that

$$\int_a^b f(x) + g(x)\, dx = \int_a^b f(x)\, dx + \int_a^b g(x)\, dx \, .$$

(b) This is an exercise. The key step is to show that

$$U(cf, P) - L(cf, P) \le |c|(U(f, P) - L(f, P)) .$$

$\square$

Theorem 7.5. *Let $f : [a, b] \to \mathbb{R}$ be Riemann integrable. If $g : [a, b] \to \mathbb{R}$ differs from f at finitely many points then g is also Riemann integrable, and*

$$\int_a^b g(x)\,dx = \int_a^b f(x)\,dx .$$

Proof. For $c \in [a, b]$, define

$$\chi_c(x) = \begin{cases} 1 & x = c , \\ 0 & x \ne c . \end{cases}$$

If g differs from f at $\{c_1, c_2, \ldots, c_n\}$, then

$$g(x) = f(x) + \sum_{i=1}^{n} (g(c_i) - f(c_i))\chi_{c_i}(x) ,$$

and it suffices to show that $\chi_c(x)$ is Riemann integrable with $\int_a^b \chi_c(x)\,dx = 0$. We shall show this by choosing suitable partitions.

If $a < c < b$, choose $P = \{a, x_1, x_2, b\}$ with $a < x_1 < x_2 < b$ and $x_2 - x_1 < \varepsilon$. It follows that

$$0 = L(\chi_c, P) < U(\chi_c, P) < \varepsilon .$$

If $c = a$, choose $P = \{a, x_1, b\}$ with $a < x_1 < b$ and $x_1 - a < \varepsilon$. It follows that

$$0 = L(\chi_a, P) < U(\chi_a, P) < \varepsilon .$$

If $c = b$, choose $P = \{a, x_1, b\}$ with $a < x_1 < b$ and $b - x_1 < \varepsilon$. It follows that

$$0 = L(\chi_b, P) < U(\chi_b, P) < \varepsilon .$$

Thus, for all $\varepsilon > 0$ there exists a partition P with $U(\chi_c, P) - L(\chi_c, P) < \varepsilon$. Therefore χ_c is Riemann integrable. As $L(\chi_c, P) = 0$ for any partition P,

$$\int_a^b \chi_c(x)\,dx = 0 .$$

$\square$

Lecture 23: 10/03/11

Theorem 7.6. *Let $f, g : [a,b] \to \mathbb{R}$ be Riemann integrable. If $f(x) \leq g(x)$ for all $x \in [a,b]$ then*

$$\int_a^b f(x)\,dx \leq \int_a^b g(x)\,dx\ .$$

Proof. As $g(x) - f(x) \geq 0$, we find

$$0 \leq L(g-f, P) \leq \int_a^b g(x) - f(x)\,dx = \int_a^b g(x)\,dx - \int_a^b f(x)\,dx\ .$$

□

Theorem 7.7. *If $f : [a,b] \to \mathbb{R}$ is Riemann integrable, then $|f|$ is Riemann integrable, and*

$$\left|\int_a^b f(x)\,dx\right| \leq \int_a^b |f(x)| dx\ .$$

Proof. For a partition P of $[a,b]$, we define

$$M_i = \sup\{f(x) : x \in I_i\}\ , \qquad M_i^* = \sup\{|f(x)| : x \in I_i\}\ ,$$

$$m_i = \inf\{f(x) : x \in I_i\}\ , \qquad m_i^* = \inf\{|f(x)| : x \in I_i\}\ .$$

Starting with

$$||f(x)| - |f(y)|| \leq |f(x) - f(y)|$$

we can show (exercise problem) that

$$M_i^* - m_i^* \leq M_i - m_i\ .$$

Therefore

$$\begin{aligned} U(|f|, P) - L(|f|, P) &= \sum_{i=1}^n (M_i^* - m_i^*)\Delta x_i \\ &\leq \sum_{i=1}^n (M_i - m_i)\Delta x_i = U(f, P) - L(f, P)\ . \end{aligned}$$

As f is Riemann integrable, it follows that $|f|$ is Riemann integrable. Furthermore,

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

implies by Theorem 7.6 that

$$-\int_a^b |f(x)|dx \le \int_a^b f(x)\,dx \le \int_a^b |f(x)|dx \;.$$

$\square$

Theorem 7.8. *If $f : [a,b] \to \mathbb{R}$ is Riemann integrable then f^2 is Riemann integrable.*

Proof. As f is bounded on $[a,b]$, there exists an $M \in \mathbb{R}$ such that $|f(x)| \le M$ for all $x \in [a,b]$. Given a partition P of $[a,b]$, we have

$$M_i(f^2) = (M_i(|f|))^2 \quad \text{and} \quad m_i(f^2) = (m_i(|f|))^2 \;.$$

Therefore

$$M_i(f^2) - m_i(f^2) = (M_i(|f|) + m_i(|f|))(M_i(|f|) - m_i(|f|)) \le 2M(M_i(|f|) - m_i(|f|)) \;.$$

Thus

$$U(f^2, P) - L(f^2, P) \le 2M(U(|f|, P) - L(|f|, P)) \;,$$

and hence f^2 is Riemann integrable. $\square$

Remark. The above proof shows also that

$$\int_a^b f^2(x)\,dx \le 2M \int_a^b |f(x)|dx \;.$$

Theorem 7.9. *If $f, g : [a,b] \to \mathbb{R}$ are Riemann integrable then fg is Riemann integrable.*

Proof. We write

$$f(x)g(x) = \frac{1}{4}\left((f(x)+g(x))^2 - (f(x)-g(x))^2\right) \;.$$

Now $f+g$ and $f-g$ are Riemann integrable by Theorem 7.4, and thus $(f+g)^2$ and $(f-g)^2$ are Riemann integrable by Theorem 7.8. By Theorem 7.4 it follows that $fg = \frac{1}{4}((f+g)^2 - (f-g)^2)$ is Riemann integrable. $\square$

8 The Fundamental Theorem of Calculus

Lecture 24: 11/03/11

Definition 8.1. *Let I be an interval and let $f : I \to \mathbb{R}$. A differentiable function $F : I \to \mathbb{R}$ is called an <u>antiderivative of f</u> if $F'(x) = f(x)$ for all $x \in I$.*

Theorem 8.2. *If F and G are antiderivatives of f, then $G = F + c$ for some $c \in \mathbb{R}$. Also, $F + c$ is an antiderivative of f for all $c \in \mathbb{R}$.*

Proof. $(G - F)' = G' - F' = f - f = 0$, so $G - F$ is constant. Also $(F + c)' = F' = f$ for all $c \in \mathbb{R}$. □

Theorem 8.3 (The Fundamental Theorem of Calculus)**.** *Let $f : [a, b] \to \mathbb{R}$ be Riemann-integrable. If F is an antiderivative of f then*

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

Proof. Let P be a partition of $[a, b]$. Applying the Mean Value Theorem to F on I_i, there exists a $c_i \in (x_{i-1}, x_i)$ such that

$$F(x_i) - F(x_{i-1}) = F'(c_i)(x_i - x_{i-1}) = f(c_i)\Delta x_i\ .$$

As

$$m_i = \inf\{f(x) : x \in I_i\} \le f(c_i) \le \sup\{f(x) : x \in I_i\} = M_i\ ,$$

it follows that

$$L(f, P) \le \sum_{i=1}^{n} (F(x_i) - F(x_{i-1})) \le U(f, P)\ .$$

Therefore

$$\int_{*a}^{b} f(x)\,dx \le F(b) - F(a) \le \int_a^{*b} f(x)\,dx\ ,$$

and as f is Riemann integrable, it follows that

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

□

Example. An antiderivative of $f(x) = 1/x$ is $F(x) = \log(x)$, as $F'(x) = f(x)$. We use this to compute

$$\int_1^a \frac{dx}{x} = \log(x)\big|_1^a = \log(a) - \log(1) = \log(a)\ .$$

For further examples, see Calculus I.

Theorem 8.4. *Let $f : [a, b] \to \mathbb{R}$ be Riemann integrable and define $F : [a, b] \to \mathbb{R}$ by*

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

Then

(a) F is continuous on $[a, b]$.

(b) If f is continuous at $c \in [a, b]$ then F is differentiable at c and $F'(c) = f(c)$.

Proof. (a) f is Riemann integrable, hence bounded, i.e. there exists an $M \in \mathbb{R}$ such that $|f(x)| \le M$ for all $x \in [a, b]$.

Given $t, t_0 \in [a, b]$, we have

$$|F(t) - F(t_0)| = \left| \int_a^t f(x)\,dx - \int_a^{t_0} f(x)\,dx \right| = \left| \int_{t_0}^t f(x)\,dx \right| \le M|t - t_0| \ .$$

If $|t - t_0| < \delta = \dfrac{\varepsilon}{M}$ then $|F(t) - F(t_0)| < \varepsilon$, implying continuity of F.

(b) Let f be continuous at c, i.e. $\forall \varepsilon > 0\ \exists \delta > 0\ \forall x \in [a, b], |x - c| < \delta$: $|f(x) - f(c)| < \varepsilon$. Hence, if $0 < |t - c| < \delta$ then

$$\left| \frac{F(t) - F(c)}{t - c} - f(c) \right| = \left| \frac{\int_c^t f(x)\,dx - \int_c^t f(c)\,dx}{t - c} \right| \le \left| \frac{\int_c^t |f(x) - f(c)|dx}{t - c} \right| < \varepsilon \ .$$

Thus $F'(c) = \lim_{t \to c} \dfrac{F(t) - F(c)}{t - c}$ exists and $F'(c) = f(c)$.

□

Lecture 25: 14/03/11

Example. Let $f : [-1, 1] \to \mathbb{R}$ be given by

$$f(x) = \begin{cases} 0 & x \in [-1, 0] \ , \\ 1 & x \in (0, 1] \ . \end{cases}$$

Then

$$F(t) = \int_{-1}^t f(x)\,dx = \begin{cases} 0 & t \in [-1, 0] \ , \\ t & t \in (0, 1] \ . \end{cases}$$

The function F is continuous on $[-1, 1]$ and differentiable on $[-1, 0) \cup (0, 1]$, but not differentiable at $t = 0$.

Corollary. *Every continuous function $f : [a, b] \to \mathbb{R}$ has an antiderivative.*

Proof. By Theorem 8.4, $F(t) = \int_a^t f(t)\,dt$ is an antiderivative of f. $\square$

Definition 8.5. *If F is an antiderivative of f, we define*

$$\int f(x)\,dx = F(x) + c\,,$$

the indefinite integral of f.

Theorem 8.6. *If f and g have antiderivatives on I, then so do $f + g$ and cf for $c \in \mathbb{R}$. Moreover,*

$$\int f(x) + g(x)\,dx = \int f(x)\,dx + \int g(x)\,dx \quad \text{and} \quad \int cf(x)\,dx = c\int f(x)\,dx\,.$$

Proof. $F' = f$ and $G' = g$ imply $(F + G)' = F' + G' = f + g$. Therefore

$$\int f(x) + g(x)\,dx = F(x) + G(x) = \int f(x)\,dx + \int g(x)\,dx\,.$$

Similarly, $(cF)' = cF'$, so that

$$\int cf(x)\,dx = cF(x) = c\int f(x)\,dx\,.$$

$\square$

Theorem 8.7. *Let $f, g : I \to \mathbb{R}$ be differentiable. If fg' has an antiderivative, then so does $f'g$, and*

$$\int f'(x)g(x)\,dx = f(x)g(x) - \int f(x)g'(x)\,dx;.$$

Proof. Let H be the antiderivative of $h = fg'$, i.e. $H' = h = fg'$. Then $(fg)' = f'g + fg'$ implies that

$$f'g = (fg)' - fg' = (fg)' - H' = (fg - H)'\,.$$

Therefore $fg - H$ is an antiderivative of $f'g$, and

$$\int f'(x)g(x)\,dx = f(x)g(x) - H(x) = f(x)g(x) - \int f(x)g'(x)\,dx\,.$$

$\square$

Theorem 8.8. *Let $g : I \to \mathbb{R}$ be differentiable and let F be an antiderivative of $f : g(I) \to \mathbb{R}$. Then $F \circ g$ is an antiderivative of $(f \circ g)g'$, i.e.*

$$F(g(x)) = \int f(g(x))g'(x)\,dx\ .$$

Proof. We verify that $(F \circ g)'(x) = F'(g(x))g'(x) = f(g(x))g'(x)$. $\square$

Corollary. *Let $g : [a, b] \to \mathbb{R}$ be continuously differentiable and let $f : g([a, b]) \to \mathbb{R}$ be continuous. Then*

$$\int_a^b f(g(x))g'(x)\,dx = \int_{g(a)}^{g(b)} f(u)du\ .$$

Proof. f and $(f \circ g)g'$ are both continuous on $[a, b]$, hence Riemann integrable. As f is continuous, it has an antiderivative, F. By Theorem 8.8, $F \circ g$ is an antiderivative of $(f \circ g)g'$, and

$$\int f(g(x))g'(x) = F(g(x))\ .$$

By the Fundamental Theorem of Calculus,

$$\int_a^b f(g(x))g'(x)\,dx = F(g(b)) - F(g(a)) = \int_{g(a)}^{g(b)} f(u)\,du\ .$$

$\square$

9 Sequences and Series of Functions

Let $\mathcal{D} \subseteq \mathbb{R}$ be a domain. Unless stated otherwise, in this section all functions map $\mathcal{D} \to \mathbb{R}$.

Recall that a sequence (a_n) of real numbers converges to a limit a if

$$\forall \epsilon > 0 \; \exists n_0 \in \mathbb{N} \; \forall n \geq n_0 : |a_n - a| < \varepsilon \; .$$

Similarly, for a sequence of functions (f_n) we can discuss convergence of this sequence to a limiting function. This leads to the consideration of the convergence of the sequence (a_n) where $a_n = f_n(x)$ for $x \in \mathcal{D}$. Keeping the point x fixed, this leads to the notion of pointwise convergence, while allowing x to vary within the domain $\mathcal{D}$ leads to the notion of uniform convergence. The next definition makes this idea more precise.

Definition 9.1. *Let (f_n) be a sequence of functions.*

(1) f_n converges pointwise to a function f if

$$\forall x \in \mathcal{D} \; \forall \epsilon > 0 \; \exists n_0 \in \mathbb{N} \; \forall n \geq n_0 : |f_n(x) - f(x)| < \varepsilon \; .$$

(2) f_n converges uniformly to a function f if

$$\forall \epsilon > 0 \; \exists n_0 \in \mathbb{N} \; \forall n \geq n_0 \; \forall x \in \mathcal{D} : |f_n(x) - f(x)| < \varepsilon \; .$$

Remark. In (1) n_0 depends on x and ε, whereas in (2) n_0 depends on ε, but not on x. In both cases, we can write Lecture 26: 17/03/11

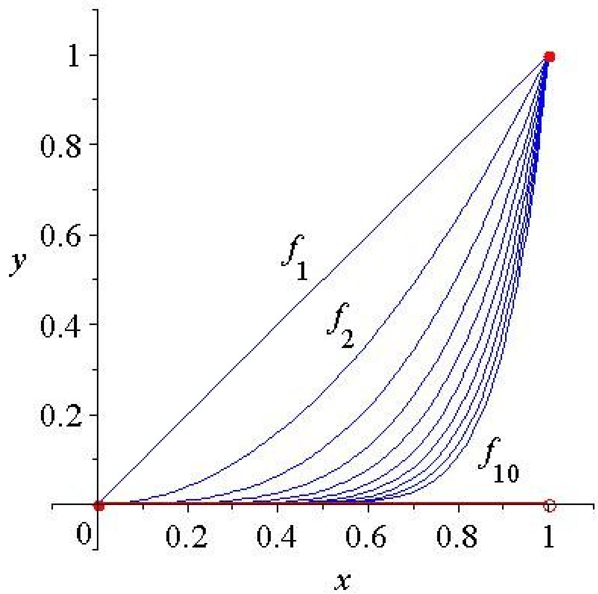
$$f(x) = \lim_{n \to \infty} f_n(x) \; .$$

Note that the limit notation does not indicate whether the convergence is uniform or pointwise.

Clearly uniform convergence implies pointwise convergence, but the converse is not true.

Examples.

(1) $f_n : [0,1] \to \mathbb{R}$, $x \mapsto x^n$.



We find

$$\lim_{n\to\infty} f_n(x) = \lim_{n\to\infty} x^n = \begin{cases} 0 & 0 \le x < 1 \ , \\ 1 & x = 1 \ . \end{cases}$$

Thus f_n converges pointwise to the discontinuous function

$$f : [0,1] \to \mathbb{R} \ , \quad x \mapsto \begin{cases} 0 & 0 \le x < 1 \ , \\ 1 & x = 1 \ . \end{cases}$$

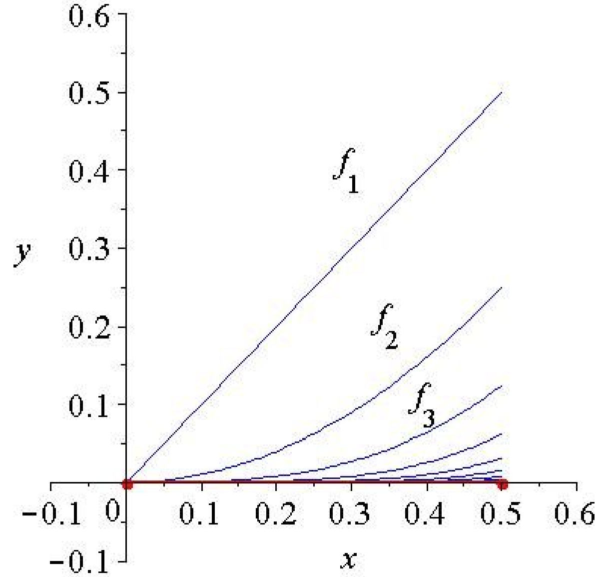
This convergence is <u>not</u> uniform: we need to show

$$\exists \varepsilon > 0 \ \forall n_0 \in \mathbb{N} \ \exists n \ge n_0 \ \exists x \in [0,1] : |f_n(x) - f(x)| \ge \varepsilon \ .$$

Take $\varepsilon = 1/2$ and consider $x = 2^{-1/n}$:

$$|f_n(2^{-1/n}) - f(2^{-1/n})| = |(2^{-1/n})^n - 0| = \frac{1}{2} \ge \varepsilon \ .$$

(2) $f_n : [0, 1/2] \to \mathbb{R}$, $x \mapsto x^n$.



For $0 \le x \le 1/2$ we find $\lim_{n\to\infty} f_n(x) = \lim_{n\to\infty} x^n = 0$. Thus f_n converges to

$$f : [0, 1/2] \to \mathbb{R}\ , \quad x \mapsto 0\ .$$

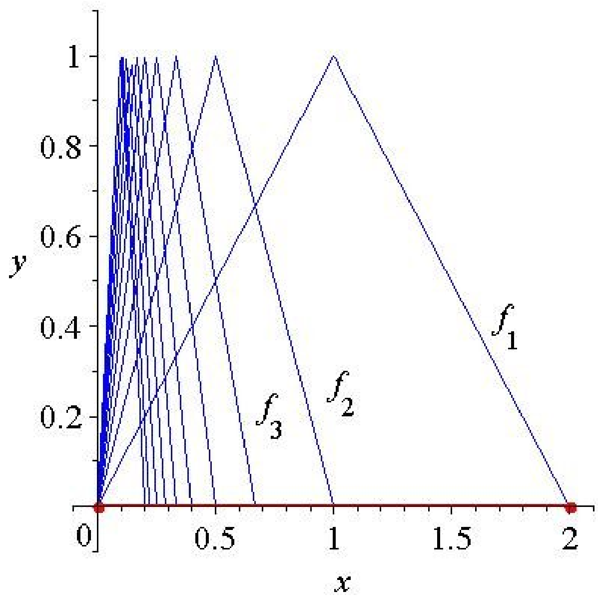
This convergence is uniform:

The difference between $f_n(x)$ and $f(x)$ is largest at $x = 1/2$. Therefore, if we pick an integer n_0 such that $n_0 > -\log(\varepsilon)/\log(2)$ to ensure $(1/2)^{n_0} < \varepsilon$, then for all $n \ge n_0$,

$$|f_n(x) - f(x)| = |x^n - 0| \le (1/2)^n \le (1/2)^{n_0} < \varepsilon\ .$$

(3) $f_n : [0,2] \to \mathbb{R}$,

$$x \mapsto \begin{cases} nx & 0 \le x \le 1/n \ , \\ 2 - nx & 1/n < x \le 2/n \ , \\ 0 & 2/n < x \le 2 \ . \end{cases}$$



$f_n(0) = 0$, and if $0 < x \le 2$ then $f_n(x) = 0$ if $n \ge 2/x$, so that

$$\lim_{n\to\infty} f_n(x) = 0 \quad \text{for all } 0 \le x \le 2.$$

Thus f_n converges to

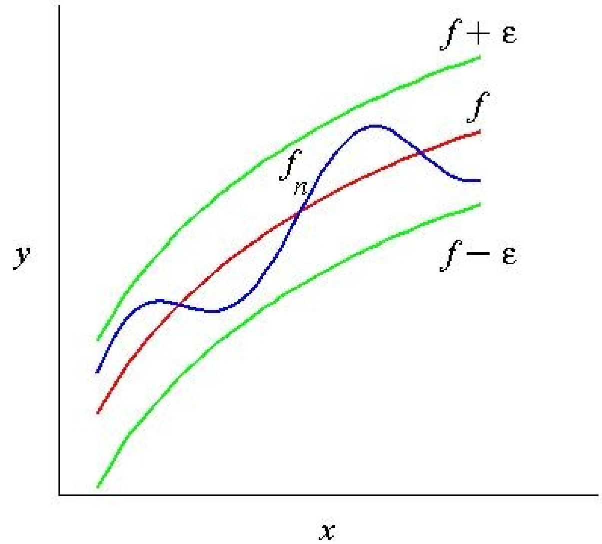
$$f : [0,2] \to \mathbb{R} \ , \quad x \mapsto 0 \ .$$

This convergence is <u>not</u> uniform: take $\varepsilon = 1$ and consider $x = 1/n$:

$$|f_n(1/n) - f(1/n)| = |1 - 0| = 1 \ge \varepsilon \ .$$

Remark. The following figures indicate the idea of an "ε-tube" around the limiting function f.

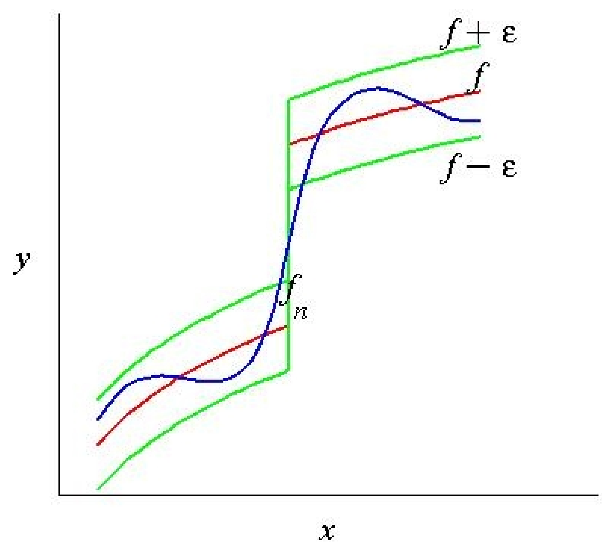
Lecture 27: 18/03/11



ε − *tube of uniform convergence*

In the case of uniform convergence, given $\varepsilon > 0$, the graph of $y = f_n(x)$ must lie entirely within the ε-tube of f for all sufficiently large n.

When the limiting function f is discontinuous, the ε-tube is "broken".



ε − *tube of a discont. function is broken*

If f is a limit of continuous f_n, no f_n can lie entirely within the ε-tube of f if ε is sufficiently small.

Theorem 9.2. *Let $f_n : \mathcal{D} \to \mathbb{R}$ converge uniformly to $f : \mathcal{D} \to \mathbb{R}$. If f_n are continuous at $a \in \mathcal{D}$ then f is continuous at a.*

Proof. We need to show

$$\forall \varepsilon > 0 \; \exists \delta > 0 \; \forall x \in \mathcal{D}, |x-a| < \delta : |f(x) - f(a)| < \varepsilon \, .$$

By assumption we have

(a) $\forall \varepsilon' > 0 \; \exists n_0 \in \mathbb{N} \; \forall n \geq n_0 \; \forall x \in \mathcal{D} : |f(x) - f_n(x)| < \varepsilon'$, and

(b) $\forall \varepsilon'' > 0 \; \exists \delta > 0 \; \forall x \in \mathcal{D}, |x-a| < \delta : |f_n(x) - f_n(a)| < \varepsilon''$.

We start estimating the distance between $f(x)$ and $f(a)$ by splitting $|f(x) - f(a)|$ into three parts:

$$|f(x) - f(a)| \leq |f(x) - f_n(x)| + |f_n(x) - f_n(a)| + |f_n(a) - f(a)| \, .$$

First, given $\varepsilon > 0$, we choose $\varepsilon' = \varepsilon/3$. By (a) there is an n_0 such that for all $n \geq n_0$ and for all $x \in \mathcal{D}$:

$$|f(x) - f_n(x)| < \varepsilon/3$$

(so that clearly also $|f(a) - f_n(a)| < \varepsilon/3$). Next, fix an $n > n_0$ and choose $\varepsilon'' = \varepsilon/3$. By (b) there exists a $\delta > 0$ such that for all $x \in \mathcal{D}$,

$$|x - a| < \delta : |f_n(x) - f_n(a)| < \varepsilon/3 \, .$$

Thus, given $\varepsilon > 0$ we have shown that there is a $\delta > 0$ such that

$$|f(x) - f(a)| < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon$$

for $|x - a| < \delta$. □

Remark. This theorem implies that *under the assumption of uniform convergence of the functions* we can exchange limits as follows:

$$\lim_{x\to a} \underbrace{\lim_{n\to\infty} f_n(x)}_{f(x)} = \lim_{n\to\infty} \underbrace{\lim_{x\to a} f_n(x)}_{f_n(a)} \, .$$

If the convergence of f_n to f is not uniform, this is generally not correct. For example $\lim_{x\to 1^-} \lim_{n\to\infty} x^n = 0$ but $\lim_{n\to\infty} \lim_{x\to 1^-} x^n = 1$ (see example (1) above).

An immediate consequence of Theorem 9.2 is the next theorem.

Theorem 9.3. *If a sequence of continuous functions converges uniformly, then the limiting function is continuous.*

Remark. If the limiting function of a sequence of continuous functions is discontinuous, the convergence cannot be uniform.

Examples (continued).

(1) f_n are continuous, the limiting function is not continuous. Therefore the convergence of f_n to f cannot be uniform.

(2) f_n are continuous, and the convergence is uniform. Therefore the limiting function is continuous.

(3) f_n are continuous, the limiting function is continuous. However, this does not imply uniform convergence.

Theorem 9.4. *Let $f_n : [a,b] \to \mathbb{R}$ be Riemann integrable. If f_n converges uniformly to $f : [a,b] \to \mathbb{R}$ then f is Riemann integrable and*

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

Remark. This theorem implies that *under the assumption of uniform convergence of the functions* we can exchange limits as follows:

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

Lecture 28: 21/03/11

Proof. Let $\varepsilon > 0$. We want to show that there exists a partition P such that $U(f,P) - L(f,P) < \varepsilon$. We shall do this in three steps.

(a) We know that f_n converges uniformly to f:

$$\exists n \in \mathbb{N}\ \forall x \in [a,b] : |f(x) - f_n(x)| < \frac{\varepsilon}{3(b-a)}\ .$$

(b) Once n is chosen, we use Riemann integrability for f_n:

$$\exists P : U(f_n,P) - L(f_n,P) < \frac{\varepsilon}{3}\ .$$

(c) Now we constrain upper and lower sums $U(f,P)$ and $L(f,P)$: f_n is bounded, and (a) implies that $f - f_n$ is bounded, so that

$$\begin{aligned}
M_i = \sup\{f(x) : x \in I_i\} &\leq \sup\{f_n(x) : x \in I_i\} + \sup\{f(x) - f_n(x) : x \in I_i\} \\
&\leq M_i^{(n)} + \frac{\varepsilon}{3(b-a)} \text{ , and} \\
m_i = \inf\{f(x) : x \in I_i\} &\geq \inf\{f_n(x) : x \in I_i\} + \inf\{f(x) - f_n(x) : x \in I_i\} \\
&\geq m_i^{(n)} - \frac{\varepsilon}{3(b-a)} \text{ .}
\end{aligned}$$

Therefore

$$\begin{aligned}
U(f,P) - U(f_n,P) &\leq \sum_{i=1}^{n} (M_i - M_i^{(n)}) \Delta x_i \leq \frac{\varepsilon}{3(b-a)} \sum_{i=1}^{n} \Delta x_i = \frac{\varepsilon}{3} \text{ , and} \\
L(f,P) - L(f_n,P) &\geq \sum_{i=1}^{n} (m_i - m_i^{(n)}) \Delta x_i \geq -\frac{\varepsilon}{3(b-a)} \sum_{i=1}^{n} \Delta x_i = -\frac{\varepsilon}{3} \text{ .}
\end{aligned}$$

Thus

$$\begin{aligned}
&U(f,P) - L(f,P) = \\
&\quad (U(f,P) - U(f_n,P)) + (U(f_n,P) - L(f_n,P)) + (L(f_n,P) - L(f,P)) \\
&\qquad \leq \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon \text{ .}
\end{aligned}$$

Therefore f is Riemann integrable.

Moreover

$$\begin{aligned}
\left| \int_a^b f(x)\,dx - \int_a^b f_n(x)\,dx \right| &= \left| \int_a^b f(x) - f_n(x)\,dx \right| \\
&\leq \int_a^b |f(x) - f_n(x)|\,dx \leq (b-a) \sup\{|f(x) - f_n(x)| : x \in [a,b]\} < \frac{\varepsilon}{3} \text{ ,}
\end{aligned}$$

so

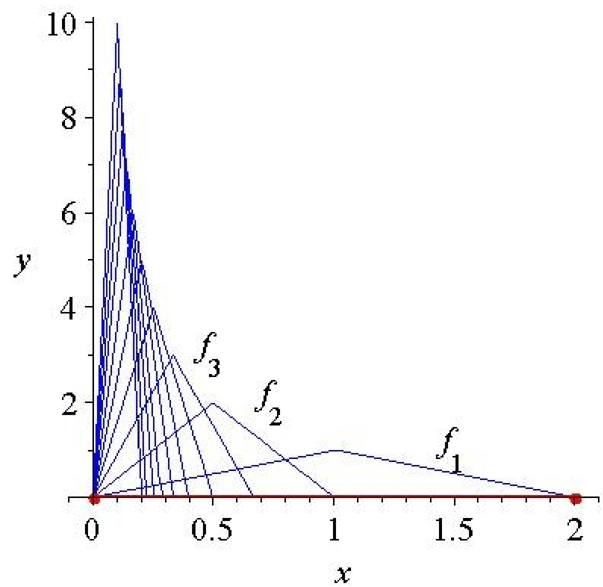
$$\lim_{n\to\infty} \int_a^b f_n(x)\,dx = \int_a^b f(x)\,dx \text{ .}$$

□

Example.

(4) Consider

$$f_n : [0,2] \to \mathbb{R} \; , \quad x \mapsto \begin{cases} n^2 x & 0 \le x \le 1/n \; , \\ 2n - n^2 x & 1/n < x \le 2/n \; , \\ 0 & 2/n < x \le 2 \; . \end{cases}$$



As in Example (3), as $n \to \infty$, $f_n(x) \to f(x) = 0$ pointwise, but not uniformly.

We compute

$$\int_0^2 f_n(x)\, dx = \int_0^{1/n} n^2 x \, dx + \int_{1/n}^{2/n} (2n - n^2 x)\, dx = 1$$

which is not equal to

$$\int_0^2 f(x)\, dx = 0 \; .$$

Theorem 9.5. *Let $f_n : [a,b] \to \mathbb{R}$ be continuously differentiable. If f_n converges pointwise to $f : [a,b] \to \mathbb{R}$ and f_n' converges uniformly to $g : [a,b] \to \mathbb{R}$, then f is differentiable and $f' = g$.*

Remark.

This theorem implies that *under the assumption of uniform convergence of the derivative of the functions* we can exchange limits as follows:

$$\left(\lim_{n\to\infty} f_n\right)' = \lim_{n\to\infty} (f_n') \ .$$

Lecture 29: 24/03/11

Proof. Consider $g_n = f_n'$. By assumption, g_n converges uniformly to g on $[a,b]$. Hence, by Theorem 9.3, g is continuous.

Moreover, g_n is Riemann integrable on $[a,b]$. Restricting to the interval $[a,x]$ for $a < x \leq b$, we apply Theorem 9.4 to g on $[a,x]$. It follows that g is Riemann integrable on $[a,x]$ and that

$$\int_a^x g(t)\,dt = \lim_{n\to\infty} \int_a^x g_n(t)\,dt \ .$$

Now $f_n(x) = f_n(a) + \int_a^x g_n(t)\,dt$ is an antiderivative of $g_n = f_n'$, and as f_n converges pointwise to f, we compute

$$\begin{aligned} f(x) = \lim_{n\to\infty} f_n(x) &= \lim_{n\to\infty} \left(f_n(a) + \int_a^x g_n(t)\,dt \right) \\ &= \lim_{n\to\infty} f_n(a) + \lim_{n\to\infty} \int_a^x g_n(t)\,dt = f(a) + \int_a^x g(t)\,dt \ . \end{aligned}$$

As g is continuous, by Theorem 8.4 f is differentiable. This implies that f is an antiderivative of g and, hence, that $f' = g$.

□

Remarks.

(1) We only need convergence of f_n to f at *one* point x_0. Moreover, it follows that f_n converges uniformly to f.

Proof. By the Mean Value Theorem, $(f_n-f)(x)=(f_n-f)(x_0)+(x-x_0)(f_n'-f')(c_n)$ for some $c_n\in(a,b)$. Hence

$$|f_n(x)-f(x)|\le|f_n(x_0)-f(x_0)|+(b-a)|f_n'(c_n)-f'(c_n)|\ .$$

The first term tends to zero because $f_n(x_0)$ converges to $f(x_0)$, and the second term tends to zero because f_n' converges to f' uniformly. $\square$

(2) It suffices for f_n to be differentiable, i.e. f_n' need not be continous (without proof).

(3) Even if f_n is differentiable and $f_n\to f$ uniformly, the limiting function need not be differentiable.

Definition 9.6. *(a)* $\sum\limits_{n=1}^{\infty} f_n(x)$ *converges pointwise if*

$$s_k(x)=\sum_{n=1}^{k} f_n(x)$$

converges pointwise as $k\to\infty$*.*

(b) $\sum\limits_{n=1}^{\infty} f_n(x)$ *converges uniformly if*

$$s_k(x)=\sum_{n=1}^{k} f_n(x)$$

converges uniformly as $k\to\infty$*.*

Example. $\sum\limits_{n=1}^{\infty}\dfrac{1}{(2+x^2)^n}$ converges uniformly: we compute

$$s_k(x)=\sum_{n=1}^{k}\frac{1}{(2+x^2)^n}=\frac{1}{2+x^2}\cdot\frac{1-\dfrac{1}{(2+x^2)^k}}{1-\dfrac{1}{2+x^2}}=\frac{1}{1+x^2}\left(1-\frac{1}{(2+x^2)^k}\right)\ .$$

As $\dfrac{1}{2+x^2}\le\dfrac{1}{2}$ for all $x\in\mathbb{R}$, $\dfrac{1}{(2+x^2)^k}\to 0$ as $k\to\infty$, which implies (pointwise) convergence

$$\sum_{n=1}^{\infty}\frac{1}{(2+x^2)^n}=\frac{1}{1+x^2}\ .$$

We estimate

$$\left|\frac{1}{1+x^2} - s_k(x)\right| = \frac{1}{1+x^2} \cdot \frac{1}{(2+x^2)^k} \leq \frac{1}{2^k} \, .$$

The bound $1/2^k$ tends to zero as $k \to \infty$ *independently* of x, so convergence is uniform. Lecture 30: 25/03/11

Theorem 9.7 (Weierstraß M-Test)**.** *Let* $\sum\limits_{n=1}^{\infty} a_n$ *be convergent. If* $|f_n(x)| \leq a_n$ *for all* $x \in \mathcal{D}$ *then* $\sum\limits_{n=1}^{\infty} f_n(x)$ *converges uniformly on* $\mathcal{D}$*.*

Proof. We estimate

$$\left|\sum_{n=1}^{\infty} f_n(x) - \sum_{n=1}^{k} f_n(x)\right| = \left|\sum_{n=k+1}^{\infty} f_n(x)\right| \leq \sum_{n=k+1}^{\infty} |f_n(x)| \leq \sum_{n=k+1}^{\infty} a_n \, .$$

As $\sum\limits_{n=1}^{\infty} a_n$ converges, the bound $\sum\limits_{n=k+1}^{\infty} a_n \to 0$ as $k \to \infty$ independently of $x \in \mathcal{D}$. □

Example (continued). For $f_n(x) = \dfrac{1}{(2+x^2)^n}$ we estimate

$$|f_n(x)| \leq \frac{1}{2^n} = a_n \, ,$$

and as $\sum\limits_{n=1}^{\infty} a_n = \sum\limits_{n=1}^{\infty} \dfrac{1}{2^n} = 1$ converges, by the Weierstraß M-Test $\sum\limits_{n=1}^{\infty} f_n(x)$ converges uniformly for $x \in \mathbb{R}$.

Theorem 9.8. *(a) Let* f_n *be continuous. If* $\sum\limits_{n=1}^{\infty} f_n$ *is uniformly convergent then* $f = \sum\limits_{n=1}^{\infty} f_n$ *is continuous.*

(b) Let f_n *be continuously differentiable. If* $\sum\limits_{n=1}^{\infty} f_n$ *is convergent and* $\sum\limits_{n=1}^{\infty} f_n'$ *is uniformly convergent then* $f = \sum\limits_{n=1}^{\infty} f_n$ *is differentiable and* $f' = \sum\limits_{n=1}^{\infty} f_n'$*.*

(c) Let f_n *be Riemann integrable on* $[a,b]$*. If* $\sum\limits_{n=1}^{\infty} f_n$ *is uniformly convergent then* $f = \sum\limits_{n=1}^{\infty} f_n$ *is Riemann integrable and* $\int\limits_a^b f(x)\,dx = \sum\limits_{n=1}^{\infty} \int_a^b f_n(x)\,dx$*.*

Proof. This is an immediate consequence of Theorems 9.3, 9.4, and 9.5. □

10 Power Series

Definition 10.1. $\sum_{n=0}^{\infty} a_n x^n$ *with* $a_n \in \mathbb{R}$ *is called a* *power series.*
Its *radius of convergence* r *is given by*

$$r = \sup \left\{ |x| : \sum_{n=0}^{\infty} a_n x^n \text{ converges} \right\} .$$

(a finite r *may not exist if* $\sum_{n=0}^{\infty} a_n x^n$ *converges for all* $x \in \mathbb{R}$*.)*

Theorem 10.2. *(a) If* $\sum_{n=0}^{\infty} a_n x^n$ *converges for* $x = c$*, then* $\sum_{n=0}^{\infty} a_n x^n$ *converges absolutely for all* $x \in \mathbb{R}$ *with* $|x| < |c|$*.*

(b) If $\sum_{n=0}^{\infty} a_n x^n$ *diverges for* $x = c$*, then* $\sum_{n=0}^{\infty} a_n x^n$ *diverges for all* $x \in \mathbb{R}$ *with* $|x| > |c|$*.*

Proof. (a) Convergence of $\sum_{n=0}^{\infty} a_n c^n$ implies that $\lim_{n\to\infty} a_n c^n = 0$. Thus for $|x| < |c|$ there exists a $n_0 \in \mathbb{N}$ such that

$$|a_n x^n| = |a_n c^n| \cdot \left|\frac{x}{c}\right|^n \leq \left|\frac{x}{c}\right|^n \text{ for } n \geq n_0.$$

Therefore $\sum_{n=n_0}^{\infty} |a_n x^n|$ is majorised by $\sum_{n=n_0}^{\infty} \left|\frac{x}{c}\right|^n$, which converges absolutely.

(b) If $\sum_{n=0}^{\infty} a_n x^n$ converged for some x with $|x| > |c|$, then by (a) $\sum_{n=0}^{\infty} a_n y^n$ would converge for all y with $|y| < |x|$, in particular for $y = c$, which is a contradiction.

□

Corollary. $\sum_{n=0}^{\infty} a_n x^n$ *converges absolutely for all* $x \in \mathbb{R}$ *with* $|x| < r$ *and diverges for all* $x \in \mathbb{R}$ *with* $|x| > r$*, where* r *is the radius of convergence of* $\sum_{n=0}^{\infty} a_n x^n$*.*

Remark. Convergence for $x = \pm r$ must be considered separately.

Theorem 10.3. *Let $r > 0$ be the radius of convergence of $\sum_{n=0}^{\infty} a_n x^n$ and let $0 < \rho < r$. Then $\sum_{n=0}^{\infty} a_n x^n$ converges uniformly on $\mathcal{D} = \{x \in \mathbb{R} : |x| \le \rho\}$.*

Proof. As $\rho < r$, $\sum_{n=0}^{\infty} a_n x^n$ converges absolutely. As $|a_n x^n| \le |a_n \rho^n|$ for $x \in \mathcal{D}$, the Weierstraß M-Test implies uniform convergence of $\sum_{n=0}^{\infty} a_n x^n$ on $\mathcal{D}$. □

Lecture 31: 28/03/11

Theorem 10.4. *Let $r > 0$ be the radius of convergence of $f(x) = \sum_{n=0}^{\infty} a_n x^n$. Then for all $x \in \mathbb{R}$ such that $|x| < r$,*

$$\int_0^x f(t)\,dt = \sum_{n=0}^{\infty} a_n \frac{x^{n+1}}{n+1} .$$

Proof. Choose $\rho \in \mathbb{R}$ such that $0 < \rho < r$. Then $\sum_{n=0}^{\infty} a_n x^n$ converges uniformly on $\mathcal{D} = \{x \in \mathbb{R} : |x| \le \rho\}$. As $f_n(x) = a_n x^n$ is Riemann integrable, Theorem 9.8(c) implies that $f(x) = \sum_{n=0}^{\infty} a_n x^n$ is Riemann integrable on $\mathcal{D}$ and that

$$\int_0^x f(t)\,dt = \sum_{n=0}^{\infty} \int_0^x a_n t^n\,dt = \sum_{n=0}^{\infty} a_n \frac{x^{n+1}}{n+1} .$$

□

Theorem 10.5. *Let $r > 0$ be the radius of convergence of $f(x) = \sum_{n=0}^{\infty} a_n x^n$. Then for all $x \in \mathbb{R}$ such that $|x| < r$,*

$$f'(x) = \sum_{n=1}^{\infty} n a_n x^{n-1} .$$

Proof. Choose $\rho \in \mathbb{R}$ such that $0 < \rho < r$. Then $\sum_{n=0}^{\infty} a_n x^n$ converges uniformly on $\mathcal{D} = \{x \in \mathbb{R} : |x| \le \rho\}$. To apply Theorem 9.8(b), we need to show that $\sum_{n=0}^{\infty} n a_n x^n$ also converges uniformly on $\mathcal{D}$. Once this is established, it follows that f is differentiable on $\mathcal{D}$ and that $f'(x) = \sum_{n=0}^{\infty} n a_n x^{n-1}$.

Now pick ρ' such that $\rho < \rho' < r$. Then $\sum_{n=0}^{\infty} a_n \rho'^n$ converges absolutely, and

$$|n a_n x^n| \le |n a_n \rho^n| = |a_n \rho'^n| \underbrace{\left| n \left(\frac{\rho}{\rho'}\right)^n \right|}_{\le 1 \text{ for } n \ge n_0} \le |a_n \rho'^n|$$

implies by the Weierstraß M-Test uniform convergence of $\sum_{n=0}^{\infty} n a_n x^n$ for $|x| \le \rho$, as needed. $\square$

Corollary. *$f(x) = \sum_{n=0}^{\infty} a_n x^n$ is for $|x| < r$ infinitely often differentiable, and $f^{(k)}(x) = \sum_{n=k}^{\infty} n(n-1)\dots(n-k+1) a_n x^{n-k}$.*

Remark. We find $f^{(k)}(0) = k! a_k$, so that $f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} x^n$, the Taylor series of f about zero.

Examples.

(1) For $|x| < 1$ we have

$$\frac{1}{1+x} = 1 - x + x^2 - x^3 + \dots = \sum_{n=0}^{\infty} (-1)^n x^n ,$$

and integration gives by Theorem 10.4

$$\log(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \dots = \sum_{n=0}^{\infty} (-1)^n \frac{x^{n+1}}{n+1}$$

for $|x| < 1$ (we had only proved this earlier for $0 \le x < 1$).

Note that for $x = 1$ the first sum diverges $(1 - 1 + 1 - 1 + \dots)$ but the second sum converges $(1 - 1/2 + 1/3 - 1/4 + \dots)$, whereas for $x = -1$ both sums diverge.

Lecture 32: 31/03/11

(2) For $|x| < 1$ we have

$$\frac{1}{1-x^2} = \sum_{n=0}^{\infty} x^{2n} .$$

As $\frac{1}{1-x^2} = \frac{1}{2}\left(\frac{1}{1-x} + \frac{1}{1+x}\right)$, we have for $|x| < 1$

$$\frac{1}{2} \log \frac{1+x}{1-x} = \int_0^x \frac{dx}{1-x^2} = \sum_{n=0}^{\infty} \frac{x^{2n+1}}{2n+1} .$$

Thus, for example, $x = 1/2$ gives

$$\log 3 = 2\left(\frac{1}{2} + \frac{1}{3 \cdot 2^3} + \frac{1}{5 \cdot 2^5} + \dots\right) .$$

(3) $\exp(-x^2) = \sum\limits_{n=0}^{\infty} \dfrac{(-1)^n x^{2n}}{n!}$ for all $x \in \mathbb{R}$, so that

$$\int_0^x \exp(-t^2)\,dt = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{n!(2n+1)} \text{ for all } x \in \mathbb{R}.$$

(4) $\dfrac{1}{1+x^2} = \sum\limits_{n=0}^{\infty} (-1)^n x^{2n}$ for $|x| < 1$, so that

$$\arctan x = \int_0^x \frac{dt}{1+t^2} = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{2n+1} \text{ for } |x| < 1.$$

We shall now connect power series to Taylor series. We note that

$$f(x) = \sum_{n=0}^{\infty} a_n (x-a)^n$$

converges for $|x-a| < r$, where $r > 0$ is the radius of convergence of $\sum\limits_{n=0}^{\infty} a_n x^n$. We identify $f^{(k)}(a) = k! a_k$, so that

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n \ ,$$

which is just the Taylor series of f about a.

Theorem 10.6 (Taylor's Theorem with Integral Form of the Remainder)**.** *Let $f : [a,x] \to \mathbb{R}$ be n times continuously differentiable on $[a,x]$ and $(n+1)$ times differentiable on (a,x). Then*

$$f(x) = T_{n,a}(x) + \int_a^x \frac{f^{(n+1)}(t)}{n!}(x-t)^n\,dt \ .$$

Proof. As in the proof of Taylor's Theorem (Theorem 5.3), we write

$$F(t) = T_{n,t}(x) = \sum_{k=0}^{n} \frac{f^{(k)}(t)}{k!}(x-t)^k$$

and compute

$$F'(t) = \frac{f^{(n+1)}(t)}{n!}(x-t)^n \ .$$

Therefore by the Fundamental Theorem of Calculus

$$F(x) - F(a) = \int_a^x F'(t)\,dt = \int_a^x \frac{f^{(n+1)}(t)}{n!}(x-t)^n\,dt \ ,$$

and with $F(x) = T_{n,x}(x) = f(x)$ and $F(a) = T_{n,a}(x)$ we have

$$f(x) = T_{n,a}(x) + \int_a^x \frac{f^{(n+1)}(t)}{n!}(x-t)^n \, dt \, .$$

$\square$

Remark. An analogous result holds if $[a, x]$ is replaced by $[x, a]$ for $x < a$.

Theorem 10.7. *For $\alpha \in \mathbb{R}$ we have*

$$(1+x)^\alpha = \sum_{k=0}^{\infty} \binom{\alpha}{k} x^k \text{ for } |x| < 1,$$

where $\binom{\alpha}{k} = \dfrac{\alpha(\alpha-1)\dots(\alpha-k+1)}{k!}$.

Proof. We need only consider $x \neq 0$. We apply Theorem 10.6 to $f(x) = (1+x)^\alpha$. From

$$f^{(k)}(x) = \alpha(\alpha-1)\dots(\alpha-k+1)(1+x)^{\alpha-k}$$

we see that $f^{(k)}(0) = \alpha(\alpha-1)\dots(\alpha-k+1)$. Therefore

$$(1+x)^\alpha = \sum_{k=0}^{n} \binom{\alpha}{k} x^k + \int_0^x \frac{\alpha(\alpha-1)\dots(\alpha-n)}{n!}(1+t)^{\alpha-n-1}(x-t)^n \, dt \, .$$

We need to estimate the remainder term

$$\int_0^x \frac{\alpha(\alpha-1)\dots(\alpha-n)}{n!}(1+t)^{\alpha-n-1}(x-t)^n \, dt = \alpha\binom{\alpha-1}{n} \int_0^x (1+t)^{\alpha-1} \left(\frac{x-t}{1+t}\right)^n \, dt$$

If $x > 0$ we have $0 \leq t \leq x < 1$, so that

$$0 \leq \frac{x-t}{1+t} = x - t\frac{1+x}{1+t} \leq x \, .$$

Similarly, if $x < 0$ we have $0 \geq t \geq x > -1$, so that

$$0 \geq \frac{x-t}{1+t} = x - t\frac{1+x}{1+t} \geq x \, .$$

Taken together, we conclude that inside the integral we can estimate

$$\left|\frac{x-t}{1+t}\right| \leq |x| \, .$$

Moreover, for $|x| < 1$, $M = \max\{|1+t|^{\alpha-1} : |t| \le |x|\}$ is finite. Putting this together, we arrive at

$$\left|\alpha\binom{\alpha-1}{n}\int_0^x (1+t)^{\alpha-1}\left(\frac{x-t}{1+t}\right)^n dt\right| \le M\left|\alpha\binom{\alpha-1}{n}\right||x|^n .$$

Applying the quotient test, we find that

$$\frac{M\left|\alpha\binom{\alpha-1}{n+1}\right||x|^{n+1}}{M\left|\alpha\binom{\alpha-1}{n}\right||x|^n} = \left|1-\frac{\alpha}{n+1}\right||x| \to |x| < 1 \text{ as } n\to\infty,$$

and thus $M\left|\alpha\binom{\alpha-1}{n}\right||x|^n \to 0$ as $n\to\infty$. This proves that

$$\int_0^x \frac{\alpha(\alpha-1)\dots(\alpha-n)}{n!}(1+t)^{\alpha-n-1}(x-t)^n\, dt \to 0$$

as $n\to\infty$, as required. □

Examples. For $|x| < 1$,

$$\frac{1}{\sqrt{1+x}} = \sum_{k=0}^{\infty}\binom{-1/2}{k}x^k ,$$

so that (also for $|x| < 1$)

$$\frac{1}{\sqrt{1-x^2}} = \sum_{k=0}^{\infty}\binom{-1/2}{k}(-1)^k x^{2k} .$$

Term-by-term intergration gives

$$\begin{aligned}\arcsin(x) = \int_0^x \frac{dt}{\sqrt{1-t^2}} &= \sum_{k=0}^{\infty}\binom{-1/2}{k}\frac{(-1)^k}{2k+1}x^{2k+1} \\ &= x + \frac{1}{2}\frac{x^3}{3} + \frac{1\cdot 3}{2\cdot 4}\frac{x^5}{5} + \frac{1\cdot 3\cdot 5}{2\cdot 4\cdot 6}\frac{x^7}{7} + \dots .\end{aligned}$$