MTH5105 Differential and Integral Analysis

Lecture Notes 2012-2013

Andrew Treglown

Version of March 28, 2013

Contents

0 Revision

Lecture 1: 7/01/13

Let $\mathcal{D} \subseteq \mathbb{R}$. (Most commonly we take $\mathcal{D}$ to be either an interval, or all of $\mathbb{R}$.)

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

(a) f is <u>continuous at $a \in \mathcal{D}$</u> if

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

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

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

Often it is helpful to consider 'one-sided limits'.

Definition 0.2. *Let $f : \mathcal{D} \to \mathbb{R}$.*

(a) (Left-hand limit) Given $a \in \mathcal{D}$ and $L \in \mathbb{R}$, we write $\lim_{x \nearrow a} f(x) = L$, if

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

(b) (Right-hand limit) Given $a \in \mathcal{D}$ and $L \in \mathbb{R}$, we write $\lim_{x \searrow a} f(x) = L$, if

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

Remark. Clearly $\lim_{x \to a} f(x) = L$ if and only if both $\lim_{x \searrow a} f(x) = L$ and $\lim_{x \nearrow a} f(x) = L$.

Theorem 0.3. *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 $f : \mathcal{D} \to \mathbb{R}$.

"⇒" 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.

"⇐" 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 $a \in \mathcal{D}$.

□

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 0.4. *If $f : \mathcal{D} \to \mathbb{R}$ is continuous at $a \in \mathcal{D}$ 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. Pick $\varepsilon = |b|$. Since f is continuous at a, and $b = f(a)$, by definition

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

Then, for such x, we have

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

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

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

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

□

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

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

□

1 Differentiation

Lecture 2: 10/01/13

Let $\mathcal{D} \subseteq \mathbb{R}$ be a set 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*

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

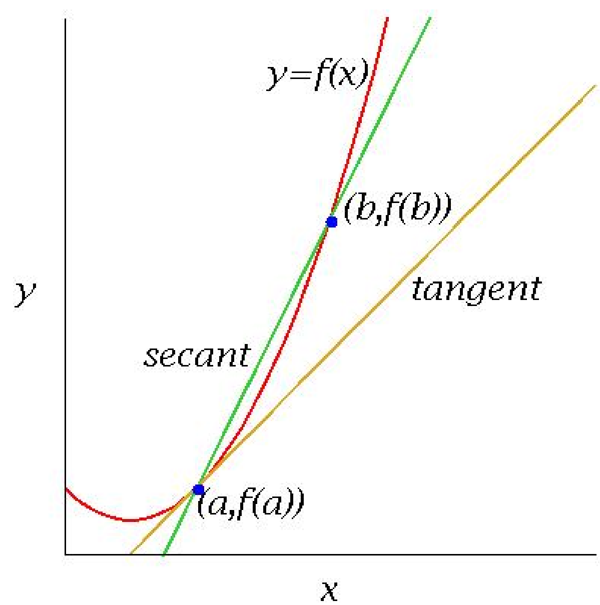
exists. The value $f'(a)$ is the derivative of f at a.

(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
$$\lim_{x \to a} \frac{f(x) - f(a)}{x - a} = \lim_{x \to a}(x + a) = 2a\ ,$$
so
$$f'(a) = 2a \quad \text{for all } a \in \mathbb{R}\,.$$
The derivative is $f' : \mathbb{R} \to \mathbb{R}$, $x \mapsto 2x$.

2) Consider $f : \mathbb{R} \to \mathbb{R}$, $x \mapsto |x|$

(i) f is not differentiable at $a = 0$:

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

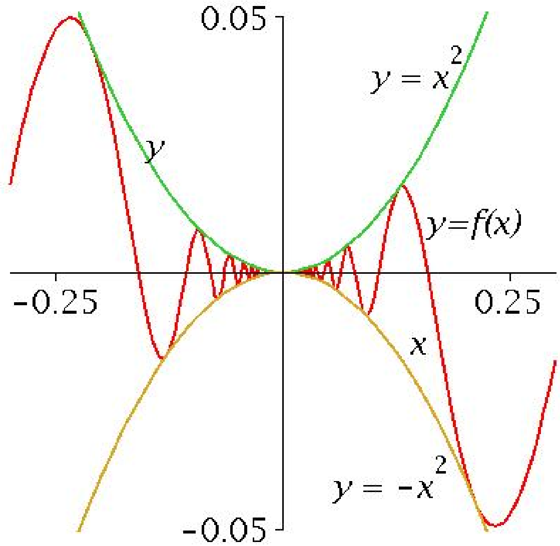
(ii) If $a \neq 0$ then f is differentiable at a:

If $a \neq 0$ then x and a have the same sign when x is sufficiently close to a.

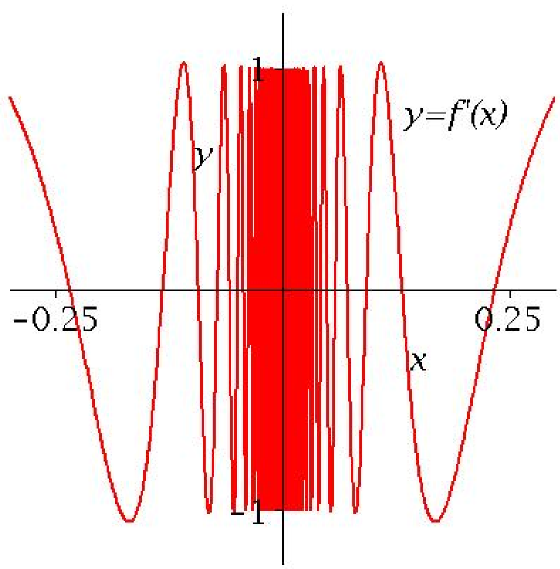
Thus, if $a > 0$,
$$\lim_{x \to a} \frac{f(x) - f(a)}{x - a} = \lim_{x \to a} \frac{|x| - |a|}{x - a} = \lim_{x \to a} \frac{x - a}{x - a} = 1\ .$$
If $a < 0$,
$$\lim_{x \to a} \frac{f(x) - f(a)}{x - a} = \lim_{x \to a} \frac{|x| - |a|}{x - a} = \lim_{x \to a} \frac{-x + a}{x - a} = -1\ .$$
In summary,
$$f'(x) = \begin{cases} 1 & x > 0 \\ \text{undefined} & x = 0 \\ -1 & x < 0\ . \end{cases}$$

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 claim that

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

To see this, note that for any $\varepsilon > 0$ we may choose $\delta = \varepsilon$, so that if $0 < |x| < \delta$ then

$$\left| \frac{f(x) - f(0)}{x - 0} - 0 \right| = \left| x \sin \frac{1}{x} \right| = |x| \left| \sin \frac{1}{x} \right| \leq |x| < \delta = \varepsilon \,,$$

as required (note we used that $|\sin \frac{1}{x}| \leq 1$ for all $x \neq 0$).

The following result gives us some properties about limits that we often use implicitly/explicitly in proofs. Lecture 3: 11/01/13

Theorem 1.2 (Algebra of limits at a point)**.** *Consider* $f : \mathcal{D} \to \mathbb{R}$ *and* $g : \mathcal{D} \to \mathbb{R}$ *and some* $a \in \mathcal{D}$*. Suppose that* $\lim_{x\to a} f(x) = L$ *and* $\lim_{x\to a} g(x) = M$ *for some* $L, M \in \mathbb{R}$*. Then the following conditions hold:*

- $\lim_{x\to a}(f(x) + g(x)) = L + M$;
- $\lim_{x\to a}(f(x)g(x)) = LM$;
- *If* $M \neq 0$ *then* $\lim_{x\to a}(\frac{f(x)}{g(x)}) = \frac{L}{M}$.

Proof. Omitted. □

Lemma 1.3. $f : \mathcal{D} \to \mathbb{R}$ *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. "⇒" 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, so f is differentiable (see Definition 1.1).

□

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 line to the graph of f, at the point $(a, f(a))$, is therefore

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

Theorem 1.4. *If $f : \mathcal{D} \to \mathbb{R}$ is differentiable at $a \in \mathcal{D}$ then f is continuous at a.*

Proof. By Lemma 1.3,

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

with $\lim_{x \to a} r(x) = 0$, $s = f(a)$ and $t = f'(a)$. Therefore $\lim_{x \to a} f(x) = s = f(a)$, so f is continuous at a, by Theorem 0.3. □

Remark. $f : \mathbb{R} \to \mathbb{R}$, $x \mapsto |x|$ is continuous at 0 but not differentiable. The converse of Theorem 1.4 is therefore not true.

Theorem 1.5. *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) Follows from the Algebra of limits. (Check it!)

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

(c) We have

$$\lim_{x\to a}\left(\frac{f(x)g(x) - f(a)g(a)}{x-a}\right) = \lim_{x\to a}\left(\frac{f(x) - f(a)}{x-a}g(x) + f(a)\frac{g(x) - g(a)}{x-a}\right) .$$

As f and g are differentiable at a and g is continuous at a by Theorem 1.4, we may apply the Algebra of limits (Theorem 1.2) to get:

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

(d) By Theorem 1.4, g is continuous at a. Now $g(a) \neq 0$, therefore by Theorem 0.4, $g(x) \neq 0$ 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

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

Using Theorem 1.2 we see that the limit as $x \to a$ exists on the right-hand-side, and

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

□

Example. Show that

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

Here there are two possible solutions:

(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 $1 = fg$, and differentiate this:

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

Remark. Often, the derivatives of 'standard' functions from Calculus will be assumed as known (e.g. $\sin' = \cos$, etc). If we wanted to, we could rigorously justify each of these using Definition 1.1.

Lecture 4: 14/01/13

Theorem 1.6 (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 1.3 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

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

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, by Lemma 1.3, $g \circ f$ is differentiable at a with $(g \circ f)'(a) = g'(b)f'(a) = g'(f(a))f'(a)$. □

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 minimum 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

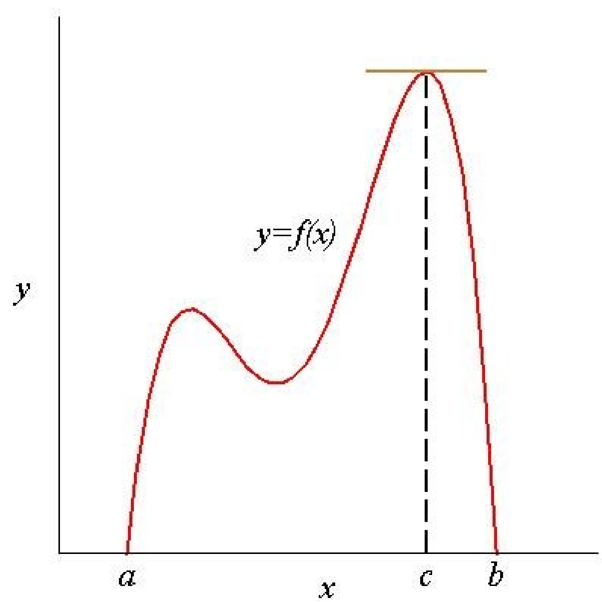
$$d = \lim_{x \searrow c} \frac{f(x) - f(c)}{x - c} \leq 0$$

and

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

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 $c \in (a, b)$ 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 $x \in (a, b)$.

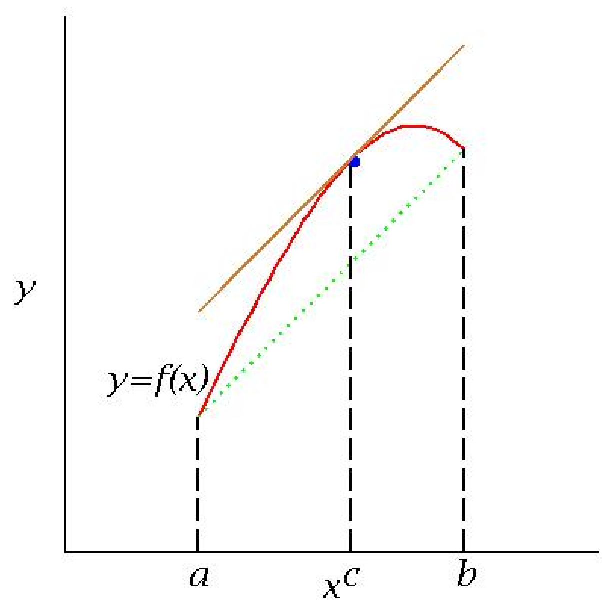
(2) $f(x) > 0$ for some $x \in (a, b)$. Then f attains its maximum on $[a, b]$ at some $c \in [a, b]$ and $f(c) \geq f(x) > 0$. Now $f(a) = f(b) = 0$, so f does not attain its maximum at either a or b, so c does not equal a or b, therefore f attains its maximum at some $c \in (a, b)$. By Theorem 2.1 it follows that $f'(c) = 0$.

(2) $f(x) < 0$ for some $x \in (a, b)$. Then f attains its minimum on $[a, b]$ 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 attains its minimum at some $c \in (a, b)$. By Theorem 2.1 it follows that $f'(c) = 0$.

$\square$

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 $c \in (a, b)$ such that*

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



Lecture 5: 17/01/13

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 $c \in (a, b)$ such that $h'(c) = 0$. But

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

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

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 $c \in (x_1, x_2)$ with

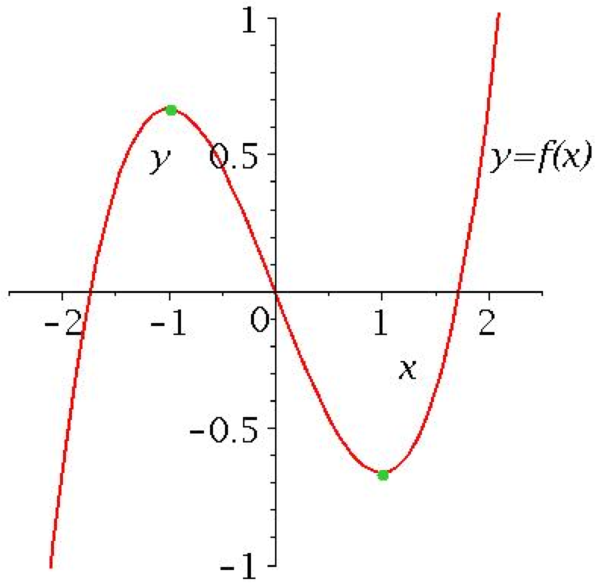
$$\frac{f(x_2) - f(x_1)}{x_2 - x_1} = f'(c) > 0 \ .$$

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 $f(x) = f(a)$. □

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 $c \in (a, b)$ such that $f(c) = s$.*

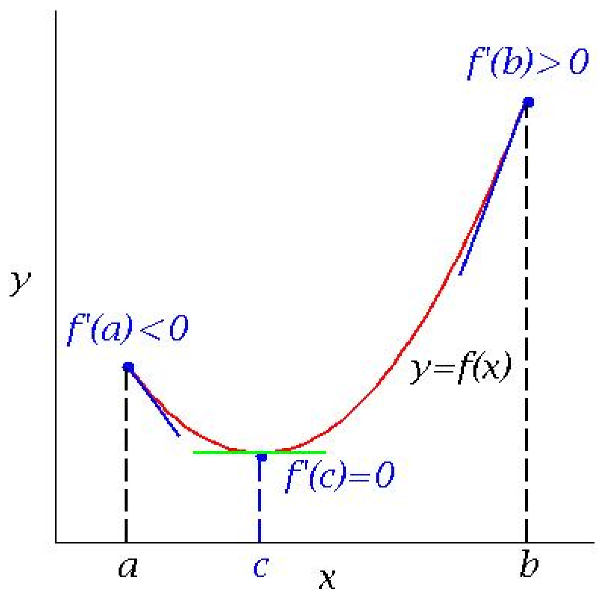
The following theorem looks very similar.

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

Lecture 6: 18/01/13

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; that is, we will show that if $f'(a) < 0 < f'(b)$, then there exists $c \in (a, b)$ such that $f'(c) = 0$:



Since f is differentiable on $[a, b]$, it is certainly continuous on $[a, b]$ (see Theorem 1.4), and therefore attains its minimum on $[a, b]$ (by a result in Convergence & Continuity).

Now $f'(a) < 0$, so there exists $a' > a$ such that

$$\frac{f(a') - f(a)}{a' - a} < 0\,.$$

(To see this, note that the definition $f'(a) = \lim_{x \to a} \frac{f(x)-f(a)}{x-a}$ implies there exists $\delta > 0$ such that if $a' \in [a, b]$ with $0 < |a' - a| < \delta$ then $\left| \frac{f(a')-f(a)}{a'-a} - f'(a) \right| < -f'(a)/2$.)
In particular, $f(a') < f(a)$.

Similarly, the fact that $f'(b) > 0$ means there exists $b' < b$ with $\frac{f(b)-f(b')}{b-b'} > 0$, and hence $f(b') < f(b)$.

So $f(a') < f(a)$ and $f(b') < f(b)$, therefore the minimum of the function $f : [a, b] \to \mathbb{R}$ cannot be attained at either of the endpoints a or b. Therefore the minimum of $f : [a, b] \to \mathbb{R}$ must be attained at some point $c \in (a, b)$. But f is differentiable at $c \in (a, b)$, so $f'(c) = 0$ by Theorem 2.1. This concludes the proof for the case $s = 0$. Now consider the general case of $s \neq 0$; that is, we assume that $f'(a) < s < f'(b)$, and will show there exists $c \in (a, b)$ such that $f'(c) = s$.

We can reduce this general case to the case $s = 0$ by considering the function $g : [a, b] \to \mathbb{R}$ defined by $g(x) = f(x) - sx$. Clearly g is differentiable on $[a, b]$, and $g'(x) = f'(x) - s$, so $g'(a) = f'(a) - s < 0$ and $g'(b) = f'(b) - s > 0$. Therefore, $g'(c) = 0$ for some $c \in (a, b)$, and hence $f'(c) = s$.

$\square$

Remark. In view of the Remark prior to the proof of Theorem 2.6, we may ask: what sort of function is differentiable everywhere yet does not have continuous derivative? One example is the function f from Example 3 in Chapter 1:

$$f : \mathbb{R} \to \mathbb{R}, x \mapsto \begin{cases} x^2 \sin \frac{1}{x} & x \neq 0 \\ 0 & x = 0 \end{cases}.$$

We saw that f is differentiable at 0, with $f'(0) = 0$, and for $x \neq 0$ we can calculate the derivative to be

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

So f is differentiable everywhere, but $\lim_{x \to 0} f'(x)$ does not exist (as suggested by the graph of f' on page 7), so we cannot say that $\lim_{x \to 0} f'(x) = f'(0)$, therefore f' is not continuous at the point 0.

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 an exponential function.*

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

In items (A) to (J) we shall collect properties of an exponential function (note that items (I) and (J) appear in Chapter 5).

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

Proof. Differentiate $h(x) = f(x)f(-x)$: Using the chain and product rules we have

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

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

(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 some $c \in \mathbb{R}$ such that $g = cf$.

Proof. Consider $h(x) = g(x)/f(x)$. By (B), the function h is defined on the whole of $\mathbb{R}$. The quotient rule implies that h is differentiable with

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

Therefore, by Theorem 2.5, h is constant, $h(x) = c$, and hence $g(x) = cf(x)$. □

(D) Definition 3.1 determines f uniquely.

Proof. Assume g satisfies Definition 3.1, i.e. that $g' = g$ and $g(0) = 1$. Then (C) implies that $g = cf$ for some $c \in \mathbb{R}$, and $g(0) = 1 = f(0)$ implies that $c = 1$, so $g = f$. □

Lecture 7: 21/01/13

Now that we have shown uniqueness (property (D)), 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 the function $g : \mathbb{R} \to \mathbb{R}$ defined by $g(x) = \exp(a+x)$ for all $x \in \mathbb{R}$. Then $g'(x) = \exp(a+x) = g(x)$, so $\exp(a+x) = c\exp(x)$ for some $c \in \mathbb{R}$ (by (C)). Letting $x = 0$, we find $\exp(a) = c$, so that $\exp(a+b) = c\exp(b) = \exp(a)\exp(b)$. $\square$

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

Proof. We use mathematical induction on 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

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

$\square$

(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 (from Convergence & Continuity) it follows that there exists $c \in \mathbb{R}$ such that $\exp(c) = 0$, contradicting property (B). $\square$

(F) exp is strictly increasing.

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

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

Proof. If $x < 0$ then by (E) we have $\exp(x) > 0 > x$, as required.

If $x = 0$ then $\exp(x) = 1 > 0 = x$, as required.

If $x > 0$ then by the Mean Value Theorem (Theorem 2.3) applied to $[0, x]$, there exists $c \in (0, x)$ such that

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

Moreover, by (F) we know $\exp(c) > \exp(0) = 1$ by (F), therefore

$$\exp(x) - 1 = x \exp(c) > x \ ,$$

and thus $\exp(x) > x + 1 > x$, as required. □

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

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

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

Case 1: $c = 1$.

This case follows since $\exp(0) = 1$.

Case 2: $c > 1$.

We have $\exp(0) = 1 < c < \exp(c)$, by Theorem 3.3. By the Intermediate Value Theorem applied to $[0, c]$, there exists $x \in (0, c)$ such that $\exp(x) = c$.

Case 3: $0 < c < 1$.

Now $1/c > 1$ and as in Case 2 we can deduce that there exists an $x \in (0, 1/c)$ such that $\exp(x) = 1/c$. By (A) we know that $\exp(x)\exp(-x) = 1$, therefore $\exp(-x) = c$. □

Before we prove the next property of the exponential function we need to recall some things from Convergence and Continuity. Lecture 8: 24/01/13

Definition 3.4 (Convergence of a sequence)**.** *Let $(a_n)_{n=1}^{\infty}$ be a sequence of real numbers and let $a \in \mathbb{R}$. We say that $\underline{(a_n)_{n=1}^{\infty} \text{ converges to } a}$ if*

$$\forall \varepsilon > 0,\ \exists N \in \mathbb{N} \text{ such that if } n \geq N \text{ then } |a_n - a| < \varepsilon.$$

In this case we write $\lim\limits_{n\to\infty} a_n = a$.

Also, recall the following two results from Convergence and Continuity.

Theorem 3.5. *Let $a, c, d \in \mathbb{R}$ and suppose $(a_n)_{n=1}^{\infty}$ is a sequence of real numbers such that $\lim\limits_{n\to\infty} a_n = a$.*

- *If $a_n \geq c$ for all $n \in \mathbb{N}$ then $a \geq c$.*
- *If $a_n \leq d$ for all $n \in \mathbb{N}$ then $a \leq d$.*

Theorem 3.6. *If $(a_n)_{n=1}^{\infty}$ is an increasing sequence which is bounded above then $(a_n)_{n=1}^{\infty}$ converges to some real number. Similarly, if $(a_n)_{n=1}^{\infty}$ is a decreasing sequence which is bounded below then $(a_n)_{n=1}^{\infty}$ converges to some real number.*

We are now ready to prove the following crucial property of the exponential function.

(H) $\exp(1) = e$, where $e := \lim\limits_{n\to\infty} \left(1 + \dfrac{1}{n}\right)^n$.

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 $\lim\limits_{n\to\infty} \left(1 + \dfrac{1}{n}\right)^n$ exists:

   (a) $a_n = \left(1 + \dfrac{1}{n}\right)^n$ is an increasing sequence:
   A calculation gives

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

from which 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, as well as the calculation $(1-\frac{1}{n})(1+\frac{1}{n-1}) = 1$.

(b) $b_n = \left(1+\frac{1}{n}\right)^{n+1}$ is a decreasing sequence:
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) Note that $4 = b_1 \geq b_n \geq a_n$ for all $n \in \mathbb{N}$ by (b). Thus, Theorem 3.6 implies that $\lim\limits_{n\to\infty} a_n$ exists. Similarly, $2 = a_1 \leq a_n \leq b_n$ for all $n \in \mathbb{N}$ by (a). So Theorem 3.6 implies that $\lim\limits_{n\to\infty} b_n$ exists.

Moreover,

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

2) Show that, for all $n \in \mathbb{N}$,

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

Consider any $n \in \mathbb{N}$. The Mean Value Theorem for exp on $[0, 1/n]$ implies that there exists $c \in (0, 1/n)$ such that

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

so that $\exp(1/n) = 1 + \exp(c)/n$. As $1 \leq \exp(c) \leq \exp(1/n)$ (since exp is strictly increasing by property (F)), 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

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

by the Corollary to Theorem 3.2 (setting $a = 1/n$).

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

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

again using the Corollary to Theorem 3.2.

Having now established the inequality

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

we may apply Theorem 3.5 to get

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

So it follows that

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

as required. □

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

Lecture 9: 25/01/13

Proof. If $n \in \mathbb{N}$ then $\exp(n) = (\exp(1))^n = e^n$, using property (H) and the Corollary to Theorem 3.2.
If $n = 0$ then $\exp(0) = 1 = e^0$, using Definition 3.1.
If $-n \in \mathbb{N}$ then combining the above with property (A) gives

$$\exp(n) = 1/\exp(-n) = 1/e^{-n} = e^n\,.$$

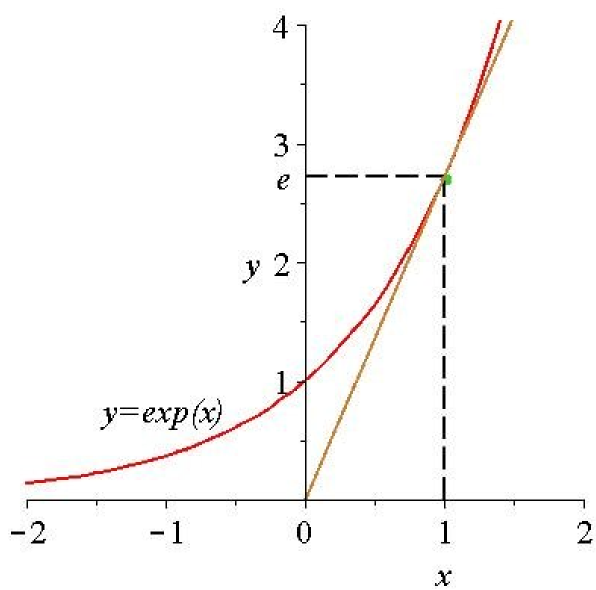
□

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 3.7. *(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 $f : \mathcal{D} \to \mathbb{R}$, and let $\mathcal{E} = f(\mathcal{D})$ be the image of f. Then f is* *invertible* *if there exists $g : \mathcal{E} \to \mathbb{R}$ such that*

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

The function g is called an *inverse* *of f.*

Properties of the inverse:

1) The inverse is uniquely defined.

Proof. Let $\mathcal{E} = f(\mathcal{D})$ and $g_1, g_2 : \mathcal{E} \to \mathbb{R}$ 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 $g_1 = g_2$. □

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

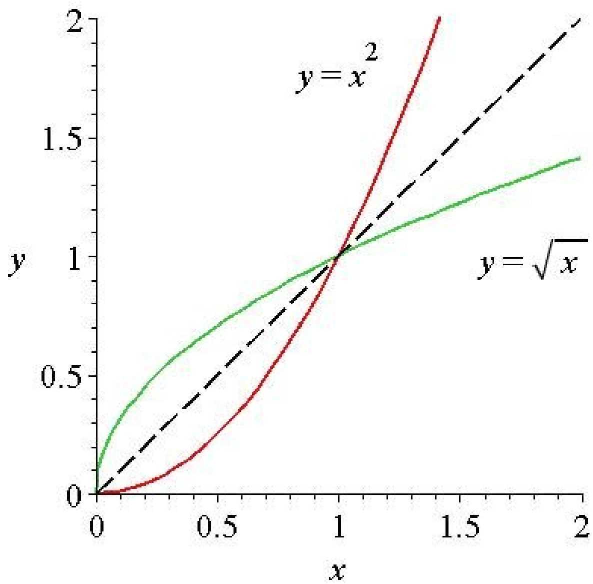
2) If f is invertible, then f^{-1} is invertible as well, and $(f^{-1})^{-1} = f$.

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

Proof. Graph$(f) = \{(x, f(x)) : x \in \mathcal{D}\}$ and 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. □

Example.

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



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. Suppose $x_1, x_2 \in \mathcal{D}$ are such that $f(x_1) = f(x_2)$. Then $x_1 = f^{-1} \circ f(x_1) = f^{-1} \circ f(x_2) = x_2$. Therefore f is injective.

"⇐" 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 $y = f(x)$; let $g(y) = x$. This defines a function $g : \mathcal{E} \to \mathbb{R}$. Then

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

So g is the inverse of f. □

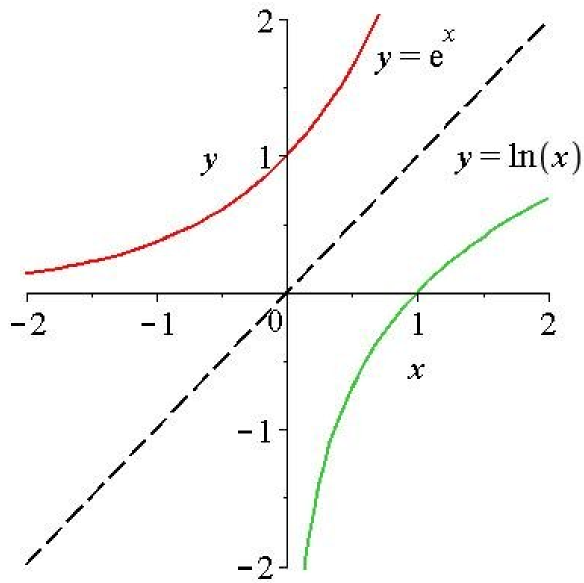
Lecture 10: 28/01/13

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

Proof. Suppose f is strictly increasing. If $x_1 \neq x_2$ then either $x_1 < x_2$, in which case $f(x_1) < f(x_2)$, or $x_2 < x_1$, in which case $f(x_2) < f(x_1)$; in either case we have $f(x_1) \neq f(x_2)$, so f is injective.

The proof for f strictly decreasing is very similar: if $x_1 \neq x_2$ then either $x_1 < x_2$, in which case $f(x_1) > f(x_2)$, or $x_2 < x_1$, in which case $f(x_2) > f(x_1)$; in either case we have $f(x_1) \neq f(x_2)$, so f is injective. □

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



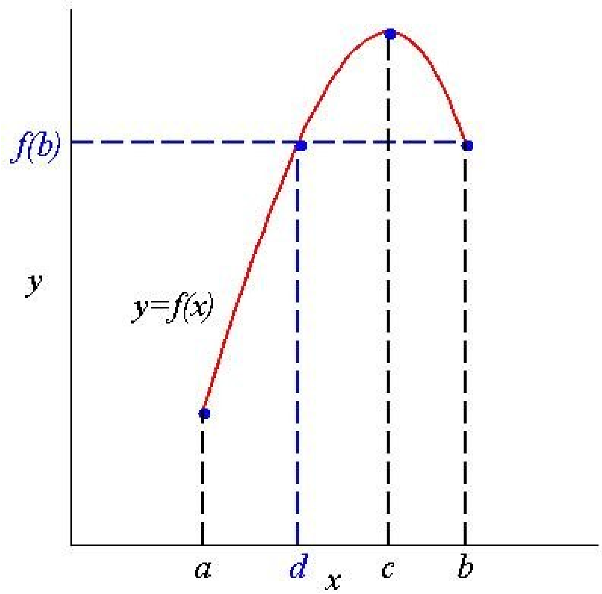
$$\exp(\mathbb{R}) = \mathbb{R}^+ \qquad\qquad \exp^{-1} = \log : \mathbb{R}^+ \to \mathbb{R}\,.$$

Let I be an interval (i.e. it has the property that if $a, b \in I$ then $a \le c \le b \Rightarrow c \in I$). Note that if $f : I \to \mathbb{R}$ is continuous then its image $f(I)$ is an interval, by the Intermediate Value Theorem.

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

Proof. Injectivity of f means that $f(a) \neq f(b)$. Without loss of generality, suppose $f(a) < f(b)$. In this case we aim to show that f attains its minimum at a and its maximum at b.

Now f is continuous, so we know f attains its maximum at some $c \in [a, b]$ (by a result from Convergence & Continuity), and clearly c cannot equal a (because $f(a) < f(b)$).



We wish to show that $c = b$. If this is not the case, i.e. if $c < b$, then $f(a) < f(b) < f(c)$ and by the Intermediate Value Theorem there exists some $d \in (a, c)$ such that $f(d) = f(b)$. But $d < c < b$ implies $d \neq b$, contradicting the injectivity of f. Thus $c = b$, and f attains its maximum at b, as required. An analogous argument shows that f attains its minimum at a. $\square$

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

Proof. (1) First consider the case of a closed bounded interval $I = [a, b]$ and assume without loss of generality that $f(a) < f(b)$. We then wish to show that f is strictly increasing. Let $x, y \in I$ with $x < y$. Then, by Theorem 4.3, f attains its maximum at b, and therefore $f(x) \leq f(b)$. Restricted to the interval $[x, b]$, the minimum of

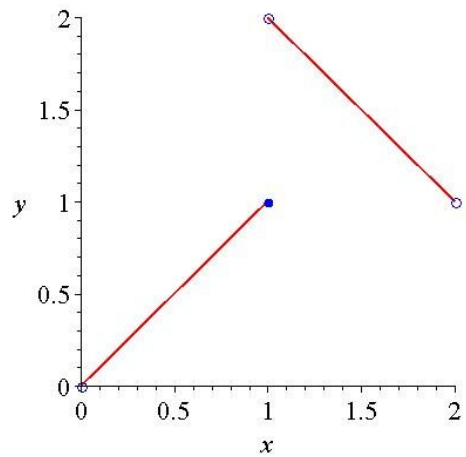
f is attained at x, and thus $f(x) \leq f(y)$. But equality $f(x) = f(y)$ is impossible, because f is injective, so in fact $f(x) < f(y)$; in other words, f is strictly increasing.

(2) Consider now an arbitrary interval I.

Fix any $u, v \in I$ with $u < v$, and assume without loss of generality that $f(u) < f(v)$. We then wish to prove that f is strictly increasing. To show this, consider any $x, y \in I$ with $x < y$. Now choose a closed interval $[a, b] \subseteq I$ which contains each of x, y, u, v. We know that f is strictly increasing or strictly decreasing on $[a, b]$ by (1). However, it cannot be strictly decreasing, because $f(u) < f(v)$ and $u < v$, therefore it must be strictly increasing on $[a, b]$. Therefore $f(x) < f(y)$. It follows that $f : I \to \mathbb{R}$ is strictly increasing. $\square$

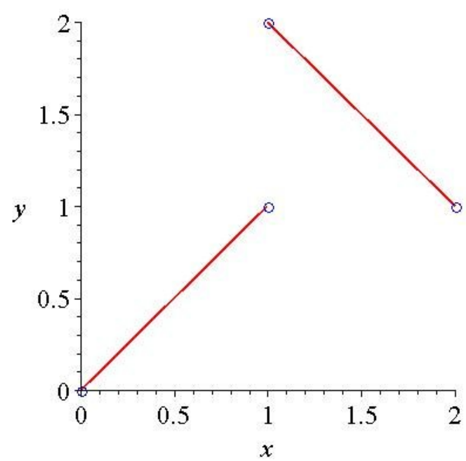
Examples.

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



Here f is injective, but is neither strictly increasing nor strictly decreasing (it is not continuous).

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

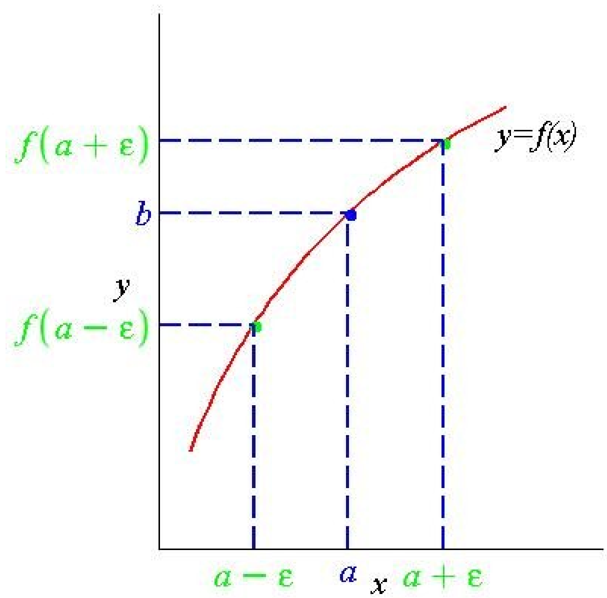


Here f is injective and continuous, but neither strictly increasing nor strictly decreasing ($(0,1)\cup(1,2)$ is not an interval).

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.* Lecture 11: 31/01/13

Proof. Theorem 4.4 inplies that f is strictly increasing or decreasing. Consider the case of strictly increasing f (the proof in the case of strictly decreasing f is similar). We need to show that f^{-1} is continuous at all $b \in f(I)$. Let us write $b = f(a)$ for some $a \in I$.

Fix $\varepsilon > 0$. We wish to show there exists $\delta > 0$ such that if $|y - b| < \delta$ then $|f^{-1}(y) - f^{-1}(b)| < \varepsilon$. If $y = f(x) \in f(I)$ satisfies $f(a - \varepsilon) < y < f(a + \varepsilon)$ then $a - \varepsilon < x < a + \varepsilon$, because f is strictly increasing.



Choose now

$$\delta := \min\{f(a + \varepsilon) - b, b - f(a - \varepsilon)\} > 0 .$$

Then $|y - b| < \delta$ implies $|x - a| < \varepsilon$, i.e. $|f^{-1}(y) - f^{-1}(b)| < \varepsilon$. So f^{-1} is continuous at b, as required. □

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 write $b = f(a)$.*

(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 using the chain rule to differentiate the equation $x = f^{-1}(f(x))$ 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.

The function f^{-1} is continuous on $f(I)$, by Theorem 4.5, 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 this function is continuous, so

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

so $(f^{-1})'(b) = \lim_{y \to b} A(y) = 1/(\lim_{y \to b} 1/A(y))$ exists and equals $1/f'(a)$.

□

Examples.

1) Consider $f : \mathbb{R} \to \mathbb{R}$, $x \mapsto x^3$. The function f is differentiable, and $f'(x) = 3x^2$. Moreover, $f(\mathbb{R}) = \mathbb{R}$ (and f is continuous by Theorem 1.4).

   Note that $f' > 0$ 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 $f^{-1} : \mathbb{R} \to \mathbb{R}$ 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 $x = 0$, but is differentiable at 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 $f : \mathbb{R} \to \mathbb{R}$, $x \mapsto \exp(x)$. Note that $f(\mathbb{R}) = \mathbb{R}^+$, f is differentiable, and $f'(x) = \exp(x) > 0$ for all $x \in \mathbb{R}$.

   Therefore Theorem 4.6 implies $f^{-1} : \mathbb{R}^+ \to \mathbb{R}$, $x \mapsto \log(x)$ is differentiable, and

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

General powers, exponentials, and logarithms

Lecture 12: 01/02/13

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

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

In particular, setting $b = e = \lim_{n\to\infty}(1+\frac{1}{n})^n$, we have

$$e^x = \exp(x\log(e)) = \exp(x\log(\exp(1))) = \exp(x1) = \exp(x)$$

for all $x \in \mathbb{R}$ (compare to Theorem 3.7 (c), where we showed this for $x \in \mathbb{Q}$).

We have $x^a = \exp(a\log(x))$ for $a \in \mathbb{R}$ and $x \in \mathbb{R}^+$, and differentiating using the chain rule (and Example 2 above) gives

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

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 $\log_b$, *logarithm to base* b, by

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

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 $x \in \mathbb{R}$

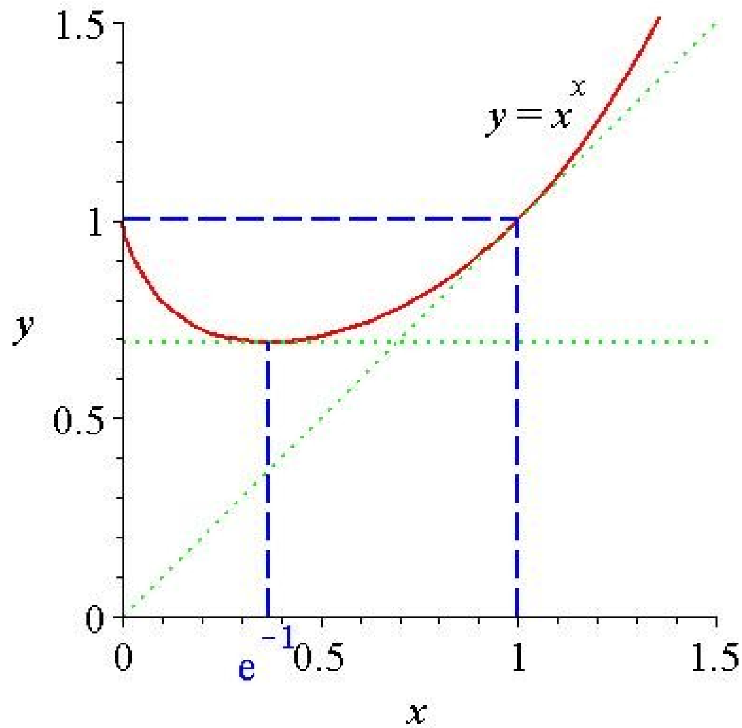
$$\log_b(b^x) = \frac{1}{\log(b)}\log(\exp(\log(b)x)) = \frac{1}{\log(b)}\log(b)x = x \ ,$$

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 applying the chain and product rules we have

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



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 $c \in (a, b)$ 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 $c \in (a, b)$ such that

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

and inserting the definition of h, we find

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

□

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* <u>*n-th derivative of f*</u>*. It is given by*

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

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.

Lecture 13: 04/02/13

Example. For $n \in \mathbb{N}$, let $f : \mathbb{R} \to \mathbb{R}$, $x \mapsto |x|x^n$. We claim that f is precisely n times differentiable (i.e. it is n times differentiable, but not $n+1$ times differentiable).

(a) We first claim that

$$f'(x) = (n+1)|x|x^{n-1}\,.$$

To prove this, consider three cases:

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

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

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

(b) Secondly, we claim that if $0 \le k \le n$ then

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

To prove this we use mathematical induction in k:

First we check the statement is true for $k = 0$:

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

Next we show that if the statement is true for $k(< n)$, then it is also true for $k+1$:

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

where we used part (a) to deduce the second equality.

So $f^{(k)}$ exists for all $0 \le k \le n$; in other words, f is n times differentiable. Now $f^{(n)}(x) = \left(\prod_{i=0}^{n-1}(n+1-i)\right)|x|$, and $x \mapsto |x|$ is not differentiable, so $f^{(n)}$ is not differentiable; therefore f is *not* $n+1$ times differentiable.

Theorem 5.3 (Taylor's Theorem)**.** *Let $n \geq 0$ be an integer. 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 $c \in (a, x)$ 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)).

Remark. When $n = 0$, Taylor's Theorem becomes precisely the Mean Value Theorem.

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 the product rule for differentiation gives:

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

Now define $g : [a, x] \to \mathbb{R}$ by $g(t) = (x - t)^{n+1}$. Applying the Second Mean Value Theorem (Theorem 5.1) to F and g on $[a, x]$ shows that there exists $c \in (a, x)$ 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

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

which becomes

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

so that

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

But from the definition of F we see that

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

so

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

as required. □

Remark. We call

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

the <u>n-th degree Taylor polynomial of f at a</u> and

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

the <u>Lagrange form of the remainder term</u>. The equation

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

is also called <u>Taylor's formula</u>, and

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

is called the <u>Taylor series of f at a</u> (whenever $f^{(k)}(a)$ exists for all $k \geq 0$).

Examples.

Lecture 14: 7/02/13

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

For $f(x) = \exp(x)$, we have $f^{(k)}(x) = \exp(x)$ for all $k \geq 0$, and thus for any $n \geq 0$,

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

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 $c \in (0,1)$ such that

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

Recall that in Chapter 3 we showed that $\exp(1) \leq (1+1/m)^{m+1}$ for all $m \in \mathbb{N}$. So $\exp(c) < \exp(1) \leq (1+1/1)^2 = 4$, and thus

$$\sum_{k=0}^{n} \frac{1}{k!} < e < \sum_{k=0}^{n} \frac{1}{k!} + \frac{4}{(n+1)!} .$$

Evaluating this chain of inequalities for $n = 11$ gives the bounds

$$2.718281826 < e < 2.718281901 .$$

Moreover, as

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

for all n, we see that

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

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 $c \in \mathbb{R}$ with $|c| < |x|$ such that

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

Now $\lim_{n\to\infty} \frac{x^n}{n!} = 0$ (see Convergence & Continuity), so $R_n \to 0$ as $n \to \infty$. Thus, $\exp(x) = \sum_{k=0}^{\infty} \frac{x^k}{k!}$.

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

For $f(x) = \log(x)$, we have $f'(x) = 1/x$, $f''(x) = -1/x^2$, $f'''(x) = 2/x^3$, $f^{(4)}(x) = -6/x^4$, ... From this we conjecture that for $k \geq 1$

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

holds and prove this via mathematical induction (this is a standard argument which we omit here). Now choose $a = 1$ and use Taylor's Theorem to 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 $c \in (1, x) \subseteq (1, 2)$ such that

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

Now $0 < x - 1 \leq 1$ and $1 < c < x \leq 2$, so that $\left|\frac{x-1}{c}\right| < 1$. Therefore $R_n \to 0$ as $n \to \infty$. In other words, $\log(x) = \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k}(x-1)^k$ for $1 < x \leq 2$, as required.

(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) $\exp(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!}$.

Proof. From Example 2) above. □

(J) $\lim_{x\to\infty} x^n \exp(-x) = 0$ for all $n \in \mathbb{N}_0$. Lecture 15: 08/02/13

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 as $x \to \infty$,

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

□

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 for all $k \geq 0$,

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

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 $c_n \in (0, x)$ 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 (of Theorem 5.4). We use mathematical induction in k. The case $k = 0$ is clearly true, since we can 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 use the product rule to 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 $\lim\limits_{x\nearrow 0} \dfrac{f^{(k)}(x) - f^{(k)}(0)}{x - 0} = 0$ and find

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

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

Lecture 16: 11/02/13

Theorem 5.5 (L'Hospital's Rule). *For $a \in \mathbb{R}$ and $\varepsilon > 0$, let $f, g : \mathcal{D} \to \mathbb{R}$ be differentiable on $(a-\varepsilon, a+\varepsilon)$, and suppose $g'(x) \neq 0$ for $0 < |x-a| < \varepsilon$. If $\lim_{x\to a} f(x) = \lim_{x\to a} g(x) = 0$ and if $\lim_{x\to a} \frac{f'(x)}{g'(x)}$ exists, then $\lim_{x\to a} \frac{f(x)}{g(x)}$ 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 $g(a) = \lim_{x\to a} g(x) = 0$. If $g(b) = 0$ for some b with $0 < |b-a| < \varepsilon$, then we can apply Rolle's Theorem to g and find that there exists some c between a and b such that $g'(c) = 0$, but this contradicts the assumption that $g'(x) \neq 0$ for $0 < |x-a| < \varepsilon$. So in fact such a b does not exist.

Next, by the Second Mean Value Theorem applied to f and g, there exists some 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

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

□

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

6 Definition of the Riemann Integral

Lecture 17: 14/02/13

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. Note that $P_1 \subseteq P_2$ 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 $P_1 = \left\{0, \frac{1}{n}, \frac{2}{n}, \ldots, \frac{n}{n}\right\}$; both P_1 and P_2 are partitions of $[0, 1]$. Here $\sigma(P_2) = \frac{1}{2n} < \sigma(P_1) = \frac{1}{n}$. Note that $P_3 = \left\{0, \frac{1}{n+1}, \frac{2}{n+1}, \ldots, \frac{n+1}{n+1}\right\}$ is also a partition of $[0, 1]$, but is not a refinement of P_1.

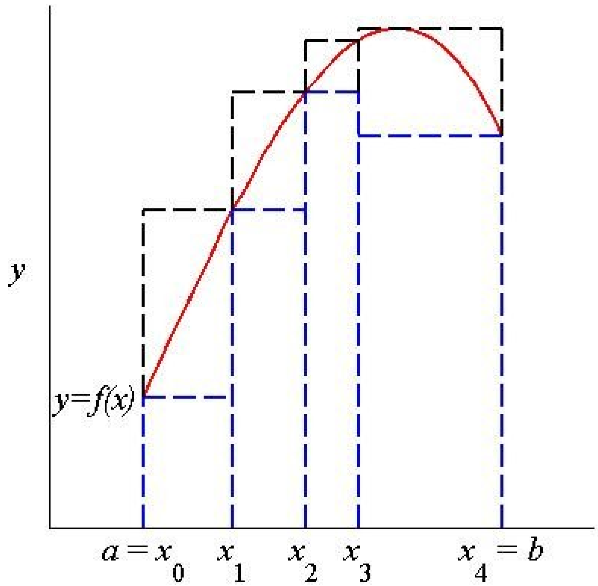
Definition 6.1. *Let $f : [a, b] \to \mathbb{R}$ be bounded and $P = \{x_0, x_1, \ldots, x_n\}$ be a partition of $[a, b]$. We define the upper sum of f with respect to P*

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

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-valued then the area A between the x-axis and the graph of f from a to b should satisfy

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

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

$$U(f, P) = 6 \cdot 1 + 2 \cdot 2 = 10 \ ,$$

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

Lecture 18: 15/02/13

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) $U(f, P_2) \le U(f, P_1)$, and

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

Proof. Let $P_1 = \{x_0, x_1, \ldots, x_n\}$. First consider the special case where the refinement P_2 is obtained from P_1 by adding a single new point y; i.e. $P_2 = P_1 \cup \{y\}$ for some $y \notin P_1$. Let i be such that $x_{i-1} < y < x_i$. Then

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

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

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

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

$$U(f, P_1) = U(f, Q_0) \ge U(f, Q_1) \ge \ldots \ge U(f, Q_r) = U(f, P_2) \ .$$

A very similar argument leads to $L(f, P_2) \ge L(f, P_1)$. □

Corollary. *Let $f : [a, b] \to \mathbb{R}$ be bounded and let P_1, P_2 be partitions of $[a, b]$. Then*

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

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

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

$\square$

Corollary. *Let $f : [a, b] \to \mathbb{R}$ be bounded. $\{U(f, P) : P \in \mathcal{P}\}$ 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. We have that,

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

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 19: 25/02/13

Theorem 6.5 (Riemann's Condition). *A bounded function $f : [a,b] \to \mathbb{R}$ 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

$$0 \leq \int_a^{*b} f(x)\,dx - \int_{*a}^{b} f(x)\,dx \leq U(f,P) - L(f,P) < \varepsilon \, .$$

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

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

so f is Riemann integrable.

□

Examples.

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

For any partition $P = \{x_0, x_1, \ldots, x_n\}$ 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 $f : [a,b] \to \mathbb{R}$, $x \mapsto \begin{cases} 1 & x \in \mathbb{Q} \ , \\ 0 & x \notin \mathbb{Q} \ . \end{cases}$

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

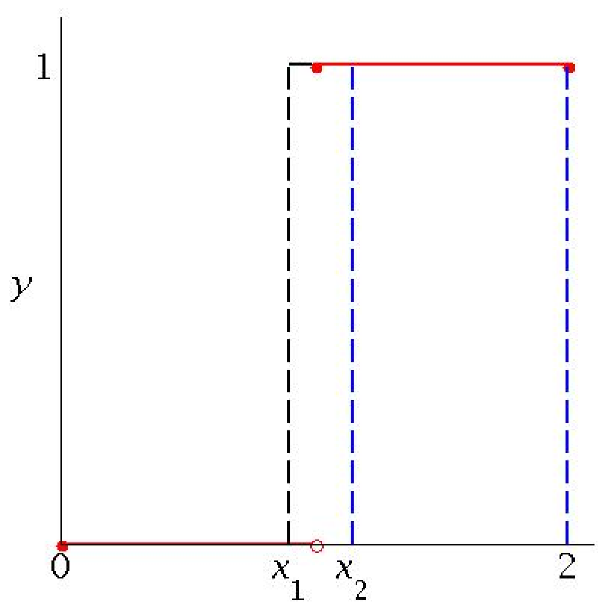
$$U(f,P) = \sum_{i=1}^{n} M_i \Delta x_i = \sum_{i=1}^{n} \Delta x_i = (b-a)$$

and

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

Therefore f is not Riemann integrable.

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



Let $\varepsilon > 0$. 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

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

and

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

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.* Lecture 20: 28/02/13

Proof. Assume that f is increasing (the argument is very similar if f is decreasing). Then $f(a) \le f(x) \le f(b)$ for $x \in [a, b]$, so f is bounded.

Let $\varepsilon > 0$. Choose a partition $P = \{x_0, x_1, \ldots, x_n\}$ of $[a, b]$. with a mesh

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

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

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

By Riemann's Condition (Theorem 6.5), f is Riemann integrable. □

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. Directly from the definitions, 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 (Bolzano-Weierstraß)**.** *Every bounded sequence has a convergent subsequence.*

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

Proof. Suppose f is continuous on $\mathcal{D} = [a, b]$ 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) $\lim_{r\to\infty} c_{n_r} = d$ for some $d \in [a, b]$,

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

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

((c) follows from (a) and (b) since f is continuous.) 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 21: 01/03/13

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 of $[a, b]$ 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

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

By Riemann's Condition (Theorem 6.5), f is Riemann integrable. □

Examples.

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

   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

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

Further,

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

Therefore

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

Since

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

we have

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

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

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

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

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

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

Therefore

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

Similarly

$$U(f,P_n) = \sum_{i=1}^{n}\frac{1}{q^{i-1}}(q-1)q^{i-1} = \sum_{i=1}^{n} q-1 = n(q-1)\ .$$

Thus,

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

Since

$$\int_{*1}^{a} f(x)\,dx \leq \int_{1}^{*a} f(x)\,dx,$$

we have

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

7 Properties of the Riemann Integral

Lecture 22: 04/03/13

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 by Riemann's Condition (Theorem 6.5), there exists a partition P of $[a,b]$ such that $U(f,P) - L(f,P) < \varepsilon$. If we define P' by

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

then P' is a refinement of P, so

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

Now let

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

Note that P'' is a partition of $[c,d]$, with

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

Thus f is Riemann integrable on $[c,d]$, by Riemann's Condition (Theorem 6.5). □

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 $[a,b]$, by Riemann's Condition (Theorem 6.5). Moreover, as

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

we have

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

Clearly we also have

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

and taking differences leads to

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

or, equivalently,

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

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) $U(f + g, P) \leq U(f, P) + U(g, P)$, *and*

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

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

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

Thus

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

(b) Similarly,

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

$\square$

Lecture 23: 07/03/13

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*

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

Proof. (a) Let $\varepsilon > 0$. By Riemann's Condition 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

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

By Theorem 7.3 it follows that

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

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

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

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

we have

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

Clearly we also have

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

and taking differences leads to

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

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

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

so that

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

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

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

□

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

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

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 by Theorem 7.4 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 $a < x_1 < c < x_2 < b$ 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, by Riemann's Condition, χ_c is Riemann integrable. As $L(\chi_c, P) = 0$ for any partition P, we have

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

□

Lecture 24: 08/03/13

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

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

Proof. As $g(x) - f(x) \ge 0$, we find that for any partition P of $[a,b]$,

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

□

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

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

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

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

we can show (exercise problem) that

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

Therefore

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

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

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

implies by Theorem 7.6 that

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

□

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 $K \in \mathbb{R}$ such that $|f(x)| \leq K$ 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

$$M_i(f^2) - m_i(f^2) = (M_i(|f|) + m_i(|f|))(M_i(|f|) - m_i(|f|)) \leq 2K(M_i(|f|) - m_i(|f|))\;.$$

Thus

$$U(f^2, P) - L(f^2, P) \leq 2K\,(U(|f|, P) - L(|f|, P))\;,$$

and hence f^2 is Riemann integrable. □

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 $fg = \frac{1}{4}\left((f + g)^2 - (f - g)^2\right)$ is Riemann integrable. □

8 The Fundamental Theorem of Calculus

Lecture 25: 11/03/13

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

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

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

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

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

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

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

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

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

Lecture 26: 14/03/13

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

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. □

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 df for $d \in \mathbb{R}$. Moreover,*

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

Proof. Let F and G be antiderivatives of f and g respectively. $F' = f$ and $G' = g$ imply $(F+G)' = F' + G' = f + g$. Therefore

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

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

$$\int df(x)\, dx = dF(x) + c = d \int f(x)\, dx \ .$$

□

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

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

□

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.*

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

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

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

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

By the Fundamental Theorem of Calculus,

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

□

9 Sequences and Series of Functions

Lecture 27: 15/03/13

Let $\mathcal{D} \subseteq \mathbb{R}$. 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

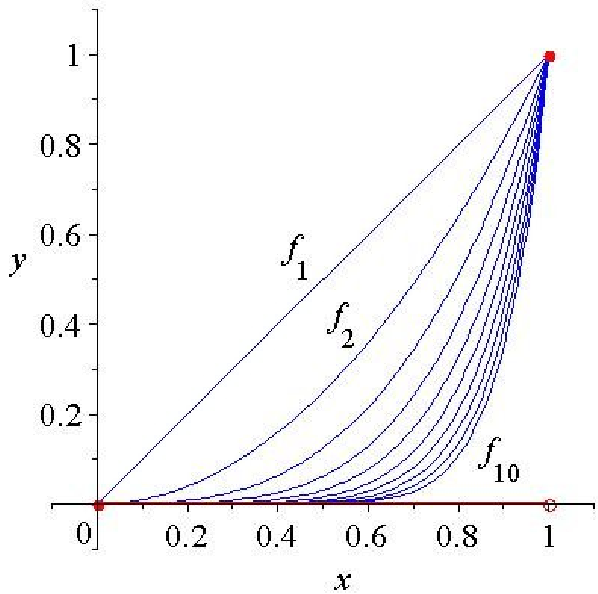
$$f = \lim_{n \to \infty} f_n \, .$$

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

By definition, 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 (for fixed x)

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

Thus f_n converges pointwise to the discontinuous function

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

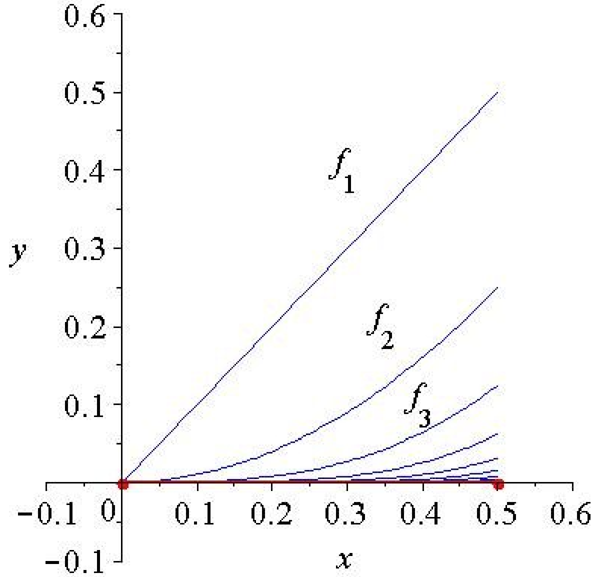
This convergence is <u>not</u> uniform: we need to show

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

Take $\varepsilon = 1/2$ and, for any n, consider the points $x = 2^{-1/n}$. Then:

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

(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 pointwise to

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

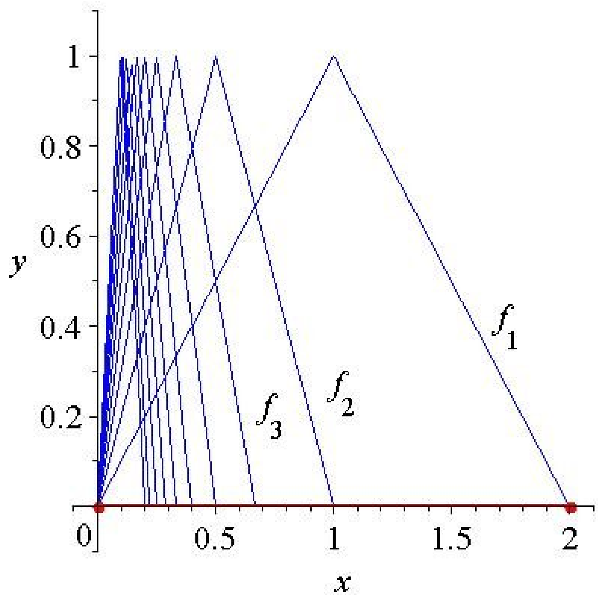
This convergence is uniform:

The difference between $f_n(x)$ and $f(x)$ is largest at $x = 1/2$. Therefore, given any $\varepsilon > 0$, 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

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

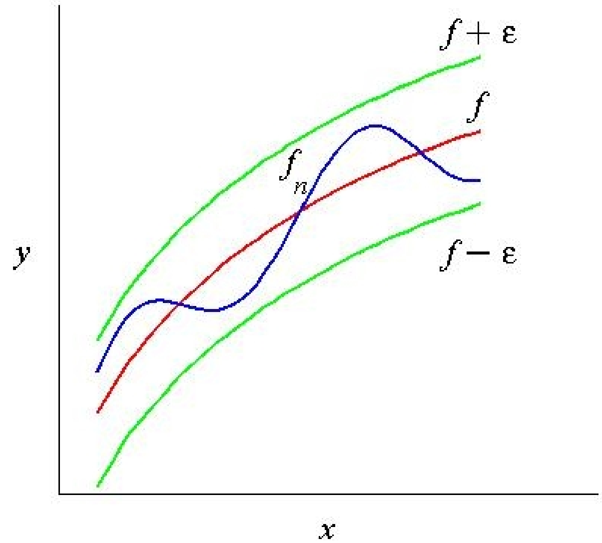
Thus f_n converges pointwise to

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

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

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

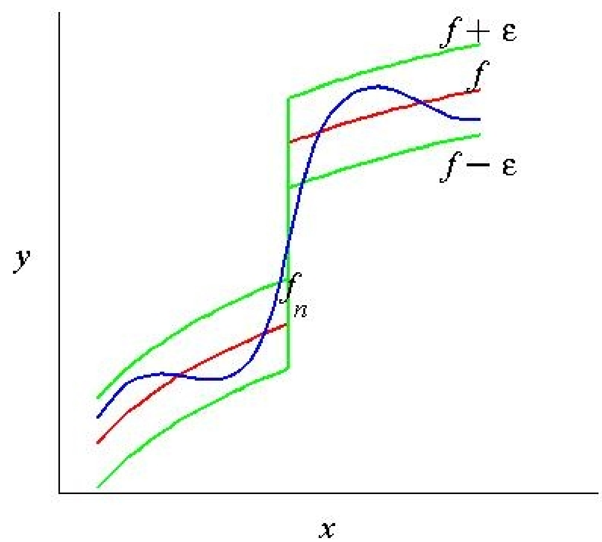
Remark. The following figures indicate the idea of an "ε-tube" around the limiting function f. Lecture 28: 18/03/13



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 this discontinuous 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 know that

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

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. Theorem 9.3 says that if the (pointwise) limiting function of a sequence of continuous functions is discontinuous, then the convergence cannot be uniform.

Examples (continued).

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

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

(3) Here the f_n are continuous, and 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:

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

Proof of Theorem 9.4

Lecture 29: 21/03/13

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

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

Therefore

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

Thus

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

Therefore f is Riemann integrable.

Moreover

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

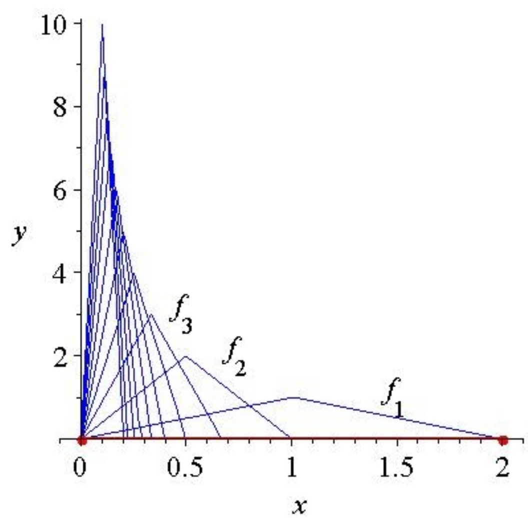
so

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

Example.

(4) Consider

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



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:

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

Proof. Consider $g_n = f_n'$. By assumption, g_n converges uniformly to g on $[a,b]$. Hence, since each g_n is continuous, Theorem 9.3 implies that 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 $f_n(x) = f_n(a) + (f_n(x) - f_n(a)) = f_n(a) + \int_a^x g_n(t)\,dt$ 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 we see that f is differentiable and $f' = g$.

□

Remarks.

(1) In Theorem 9.5, actually it suffices for f_n to be differentiable, i.e. f_n' need not be continous (proof omitted).

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

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

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

converges pointwise as $k \to \infty$.

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

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

converges uniformly as $k \to \infty$.

Remark.

In both cases we may write $\sum_{n=1}^{\infty} f_n(x) = \lim_{k\to\infty} s_k(x)$.

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

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

As $\frac{1}{2+x^2} \le \frac{1}{2}$ for all $x \in \mathbb{R}$, $\frac{1}{(2+x^2)^k} \to 0$ 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

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

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

Lecture 30: 22/03/13

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

Proof. For a fixed $x \in \mathcal{D}$, $|f_n(x)| \leq a_n$. So by the Comparison test (from Convergence and Continuity) $\sum_{n=1}^{\infty} |f_n(x)|$ converges. This implies (from a result in Convergence and Continuity) $\sum_{n=1}^{\infty} f_n(x)$ converges. So $\sum_{n=1}^{\infty} f_n(x)$ converges pointwise (i.e. $\sum_{n=1}^{\infty} f_n(x) = f(x)$ for some function f). We estimate

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

As $\sum_{n=1}^{\infty} a_n$ converges, the bound $\sum_{n=k+1}^{\infty} a_n \to 0$ as $k \to \infty$ independently of $x \in \mathcal{D}$. That is, given any $\varepsilon > 0$, there exists a $k_0 \in \mathbb{N}$ such that if $k \geq k_0$ then

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

So indeed $\sum_{n=1}^{\infty} f_n(x)$ converges uniformly to $f(x)$. □

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

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

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

(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*

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

(Note that a finite r *does 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 an $n_0 \in \mathbb{N}$ such that

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

Since $|\frac{x}{c}| < 1$, $\sum_{n=n_0}^{\infty} \left|\frac{x}{c}\right|^n$ converges. So by the Comparison test (from Convergence and Continuity) $\sum_{n=n_0}^{\infty} |a_n x^n|$ converges. Thus, $\sum_{n=0}^{\infty} |a_n x^n|$ converges.

(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.

Lecture 31: 25/03/13

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 $\mathcal{D} = \{x \in \mathbb{R} : |x| \leq \rho\}$.*

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

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, by Theorem 10.3, $\sum_{n=0}^{\infty} a_n x^n$ converges uniformly on $\mathcal{D} = \{x \in \mathbb{R} : |x| \leq \rho\}$. 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 $\mathcal{D} = \{x \in \mathbb{R} : |x| \leq \rho\}$. To apply Theorem 9.8(b), we need to show that $\sum_{n=1}^{\infty} n a_n x^{n-1}$ also converges uniformly on $\mathcal{D}$. Once this is established, it follows that f is differentiable on $\mathcal{D}$ and that $f'(x) = \sum_{n=1}^{\infty} n a_n x^{n-1}$.

Now pick ρ' such that $\rho < \rho' < r$. Then $\sum_{n=1}^{\infty} a_n \rho'^{n-1}$ converges absolutely, and

$$|n a_n x^{n-1}| \leq |n a_n \rho^{n-1}| = |a_n \rho'^{n-1}| \underbrace{\left| n \left(\frac{\rho}{\rho'}\right)^{n-1} \right|}_{\leq 1 \text{ for } n \geq n_0} \leq |a_n \rho'^{n-1}| .$$

The Weierstraß M-Test then implies the uniform convergence of $\sum_{n=1}^{\infty} n a_n x^{n-1}$ 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

$$\log(1+x) = \int_0^x \frac{1}{1+t}\, dt = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \dots = \sum_{n=0}^{\infty} (-1)^n \frac{x^{n+1}}{n+1}$$

for $|x| < 1$.

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.

(2) $\exp(-x^2) = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n}}{n!}$ for all $x \in \mathbb{R}$, so that

$$\int_0^x \exp(-t^2)\, dt = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{n!(2n+1)} \text{ for all } x \in \mathbb{R}.$$

We shall now connect power series to Taylor series. We note that

Lecture 32: 28/03/13

$$f(x) = \sum_{n=0}^{\infty} a_n (x-a)^n$$

converges for $|x-a| < r$, where $r > 0$ is the radius of convergence of $\sum_{n=0}^{\infty} a_n x^n$. 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.

Recall from Chapter 5 that for any $n \geq 0$ we define

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

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, and that Taylor's Theorem (Theorem 5.3) gives us:

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

We now give an alternative form of Taylor's Theorem:

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*

$$f(x) = T_{n,a}(x) + \int_a^x \frac{f^{(n+1)}(t)}{n!}(x-t)^n \, dt .$$

Remark. The term

$$I_n = \int_a^x \frac{f^{(n+1)}(t)}{n!}(x-t)^n \, dt$$

is called the integral form of the remainder term.

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

$$f(x) = T_{n,a}(x) + \int_a^x \frac{f^{(n+1)}(t)}{n!}(x-t)^n\,dt \ .$$

□

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 $0 \le t \le x < 1$, so that

$$0 \le \frac{x-t}{1+t} = x - t\frac{1+x}{1+t} \le x \; .$$

Similarly, if $x < 0$ we have $0 \ge t \ge x > -1$, so that

$$0 \ge \frac{x-t}{1+t} = x - t\frac{1+x}{1+t} \ge x \; .$$

Taken together, we conclude that inside the integral we can estimate

$$\left|\frac{x-t}{1+t}\right| \le |x| \; .$$

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 integration gives

$$\arcsin(x) = \int_0^x \frac{dt}{\sqrt{1-t^2}} = \sum_{k=0}^{\infty}\binom{-1/2}{k}\frac{(-1)^k}{2k+1}x^{2k+1}$$

$$= x + \frac{1}{2}\frac{x^3}{3} + \frac{1\cdot 3}{2\cdot 4}\frac{x^5}{5} + \frac{1\cdot 3\cdot 5}{2\cdot 4\cdot 6}\frac{x^7}{7} + \dots \; .$$