4 Differentiablility

Definition. Let I be an interval containing more than one point, $I \subseteq \mathbb{R}$. Let $f : I \to \mathbb{R}$. Then f is called **differentiable** at a point $x \in I$ if

$$\lim_{t \to x} \frac{f(t) - f(x)}{t - x}$$

exists, and is finite. We say that this limit is the **derivative** of f at x, denoted by $f'(x)$ or sometimes $\frac{d}{dx} f(x)$.

Facts.

1. For appropriate t, x, we say that the expression $\dfrac{f(t) - f(x)}{t - x}$ is called the differential quotient.
2. We write $f'^{-}(x) = \lim_{t \to x^-} \dfrac{f(t) - f(x)}{t - x}$, $f'^{+}(x) = \lim_{t \to x^+} \dfrac{f(t) - f(x)}{t - x}$, then of course a necessary condition (even equivalent) for $f'(x)$ to exist is that $f'^{+}(x) = f'^{-}(x)$, or x is a boundary point of I[1].
3. Equivalently, f is differentiable at $x \in I$ if and only if
$$\exists L \in \mathbb{R} : \forall (t_n)_{n \in \mathbb{N}^*} \subseteq (I \setminus \{x\})^{\mathbb{N}^*} : \left(t_n \overset{n \to \infty}{\longrightarrow} x \implies \left(\frac{f(t_n) - f(x)}{t_n - x} \right)_{n \in \mathbb{N}^*} \overset{n \to \infty}{\longrightarrow} L \right)$$
4. Differentiability is a local property, that is, f is differentiable at x if and only if for some $\delta > 0$ appropriately chosen to such that $(-\delta + x, \delta + x) \subseteq I$ for $f : I \to \mathbb{R}$, $f|_{(-\delta+x, \delta+x)}$ is differentiable at x

Example.

1. Let $n \in \mathbb{N}^*$, $f : \mathbb{R} \to \mathbb{R}$, $x \mapsto x^n$. Then $f'(x) = nx^{n-1}$ as a consequence of the binomial theorem. Let us prove it slightly differently than in the lecture by using the equivalent definition of a derivative,
$$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}$$
Consider that
$$\begin{aligned} \frac{f(x+h) - f(x)}{h} &= \frac{(x+h)^n - x^n}{h} \\ &= \frac{1}{h} \left(\sum_{k=0}^{n} \binom{n}{k} x^{n-k} h^k - x^n \right) \\ &= \frac{1}{h} \left(\sum_{k=1}^{n} \binom{n}{k} x^{n-k} h^k \right) \\ &= \frac{1}{h} \left(nx^{n-1} h + \left(\sum_{k=2}^{n} \binom{n}{k} x^{n-k} h^k \right) \right) \\ &= nx^{n-1} \left(\sum_{k=2}^{n} \binom{n}{k} x^{n-k} h^{k-1} \right) \\ &\overset{h \to 0}{\longrightarrow} nx^{n-1} \end{aligned}$$
Hence $\lim_{h \to 0} \dfrac{f(x+h) - f(x)}{h} = nx^{n-1} = \lim_{t \to x} \dfrac{f(t) - f(x)}{t - x}$.
2. Special case: let $c \in \mathbb{R}$, let $f : \mathbb{R} \to \mathbb{R}$, $f(x) := c$. Then obviously
$$\lim_{t \to x} \frac{f(t) - f(x)}{t - x} = \lim_{t \to x} \frac{\overbrace{c - c}^{0}}{t - x} = 0$$

---

[1] We might define a boundary point as follows: $x \in I$ is called a boundary point of I if $\forall \varepsilon > 0 : (\exists y \in B_\varepsilon(x) : y \in I) \wedge (\exists y \in B_\varepsilon(x) : y \notin I)$

Hence we can extend the previous example as follows: Let $n \in \mathbb{N}$, $f : \mathbb{R} \to \mathbb{R}$, $x \mapsto x^n$. Then $f'(x) = \begin{cases} 0 & \text{if n = 0} \\ nx^{n-1} & \text{otherwise} \end{cases}$. We draw the conclusion that polynomials are differentiable on $\mathbb{R}$.

3. Consider $f : \mathbb{R} \to \mathbb{R}$, $x \mapsto |x|$. Observe that

$$\lim_{t \to 0^-} \frac{f(t) - f(0)}{t} = \lim_{t \to 0^-} \frac{|t| - |0|}{t} = \lim_{t \to 0^-} \frac{-t}{t} = -1$$

Furthermore,

$$\lim_{x \to 0^+} \frac{f(t) - f(0)}{t} = \lim_{x \to 0^+} \frac{t}{t} = 1$$

Hence $f'(0)$ does not exist (f is not differentiable at 0).

4. Will prove $\exp' = \exp$ in the tutorial. This needs a little bit of effort since we need to use the property of exchanging limits, which we know very little about. We can use the fact that exp is uniformly continuous, which almost by definition justifies the interchange of limits. The upshot here is that eventually we can derive rules for exchanging the 'derivative operator'. We can already kind of see why the result $\exp' = \exp$ is expected, by assuming we can indeed exchange the derivative:

$$\begin{aligned} \frac{d}{dx} \exp(x) &= \frac{d}{dx} \sum_{n=0}^{\infty} \frac{x^n}{n!} \\ &= \frac{d}{dx} 1 + \frac{d}{dx} \sum_{n=1}^{\infty} \frac{x^n}{n!} \\ &= \sum_{n=1}^{\infty} \frac{d}{dx} \left(\frac{x^n}{n!} \right) \\ &= \sum_{n=1}^{\infty} \frac{nx^{n-1}}{n!} \\ &= \sum_{n=1}^{\infty} \frac{x^{n-1}}{(n-1)!} \\ &= \sum_{n=0}^{\infty} \frac{x^n}{n!} \\ &= \exp(x) \end{aligned}$$

5. Defined sin, cos on a problem sheet, and as we will show in the tutorial we have that $\sin' = \cos$, $\cos' = -\sin$.

Proposition. *Let $f : I \to \mathbb{R}$, I an interval of more than one point. Assume $f'(x)$ exists for $x \in I$. Then f is continuous at x (and the converse does not hold in general).*

Proof. We can see this by observing that

$$\lim_{t \to x} \frac{f(t) - f(x)}{t - x} (t - x) = 0 f'(x) = 0 = \lim_{t \to x} (f(t) - f(x)) \implies \lim_{t \to x} f(t) = f(x)$$

□

Proposition. *Let I be an interval of more than one point. Let $f, g : I \to \mathbb{R}$. Let $x \in I$ and assume that f and g are differentiable at x. Let $c \in \mathbb{R}$. Then*

- *$f + g$ is differentiable and $(f + g)'(x) = f'(x) + g'(x)$.*
- *$f \cdot g$ is differentiable and $(f \cdot g)'(x) = f'(x)g(x) + g'(x)f(x)$.*
- *$c \cdot f$ is differentiable and $(c \cdot f)'(x) = c \cdot f'(x)$ (hence '$\frac{d}{dx}$ is linear').*

- *If there exists some* $\delta > 0$ *for which it holds that* $(-\delta+x, \delta+x) \subseteq I$ *and* $\forall y \in (-\delta+x, \delta+x) : g(y) \neq 0$, *then we have that* $\dfrac{f}{g}$ *is differentiable, and*

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

Proof. Tedious exercise, done in lectures. □

Example. If we consider $f : \mathbb{R} \setminus \{0\} \to \mathbb{R}$, $x \mapsto x^n$ for $n \in \mathbb{Z}$ and $n < 0$, we have that f is differentiable everywhere (let $x \in \mathbb{R} \setminus \{0\}$), since $f(x) = x^n = \dfrac{1}{x^{-n}}$, then by the quotient rule we have that

$$f'(x) = \frac{-(-nx^{-n-1})}{(x^{-n})^2} = nx^{-n-1}x^{2n} = nx^{n-1}$$

So finally we can restate our example more generally again as follows: let $n \in \mathbb{Z}$, $f : \mathbb{R} \to \mathbb{R}$, $x \mapsto x^n$. Then $f'(x) = \begin{cases} 0 & \text{if n = 0} \\ nx^{n-1} & \text{otherwise} \end{cases}$. Later, we will define exponentiation for real numbers in the exponent, and we will see that this rule even works for any real number (instead of $n \in \mathbb{Z}$).

Proposition. *Let* I, J *be intervals of more than one point Let* $f : I \to \mathbb{R}$, $g : J \to \mathbb{R}$. *Assume that* $f(I) \subseteq J$. *Let* $x \in I$. *Assume that* f *is differentiable at* x *and* g *is differentiable at* $f(x) \in J$. *Then the function* $g \circ f$ *is differentiable at* x, *and*

$$(g \circ f)'(x) = g'(f(x))f'(x)$$

Example. Fake application of the chain rule, Let I, J be intervals of more than one point, let $f : I \to J$ and assume that f is bijective and continuous, such that we also have that f^{-1} is continuous. Let $x \in I$ and suppose f is differentiable at x, and that $f'(x) \neq 0$. Then f^{-1} is differentiable at $f(x) \in J$, and we have that

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

Proof (fake proof). Consider that $1 = \mathrm{id}'(x) = (f^{-1} \circ f)'(x) = (f^{-1})'(f(x))f'(x)$, hence the result is immediate.
The fallacy here is that f^{-1} might not be differentiable at $f(x)$, which is part of the claim, and we also need to prove that part before applying the chain rule, by the assumptions of the chain rule.

Proof. Let $(t_n)_{n \in \mathbb{N}^*}$ be a sequence in $J \setminus \{f(x)\}$ such that $t_n \overset{n\to\infty}{\longrightarrow} f(x)$. Since f^{-1} is continuous, we have that $f^{-1}(t_n) \overset{n\to\infty}{\longrightarrow} f^{-1}(f(x)) = x$. Observe that

$$\begin{aligned} \lim_{t \to f(x)} \frac{f^{-1}(t) - f^{-1}(f(x))}{t - f(x)} &= \lim_{n\to\infty} \frac{f^{-1}(t_n) - f^{-1}(f(x))}{t_n - f(x)} \\ &= \lim_{n\to\infty} \left(\frac{t_n - f(x)}{f^{-1}(t_n) - x}\right)^{-1} \\ &= \left(\lim_{n\to\infty} \frac{f(f^{-1}(t_n)) - f(x)}{f^{-1}(t_n) - x}\right)^{-1} \\ &= (f'(f^{-1}(f(x)))^{-1} = \frac{1}{f'(x)} \end{aligned}$$

Hence $(f^{-1})'(f(x) = \dfrac{1}{f'(x)}$. □

Definition. For $f : I \to \mathbb{R}$ differentiable on I, we can define a function f' given by

$$f' : I \to \mathbb{R}$$

$$x \mapsto f'(x)$$

called the **first derivative** of f. Then recursively define the following for $n \in \mathbb{N}^*$:

$$f^{(n)} = \begin{cases} f & \text{if } n = 0 \\ \left(f^{(n-1)}\right)' & \text{otherwise} \end{cases}$$

if it exists (that is, well-defined, it might be the case that $f^{(n-1)}$ at some points in I is not differentiable). If it does, then we say that f is n-times differentiable.

Remark. In order to successfully define higher order derivatives, we not only need differentiability in one point, but in a neighbourhood around one, otherwise we could not take a limit.

Definition. Let $f : I \to \mathbb{R}$, I interval more than one point. we say that f is **continuously differentiable** if f' exists and is continuous on I. Furthermore, we define the set (could also do this for general metric spaces)

$$C^1(I;\mathbb{R}) := \{(f : I \to \mathbb{R}) \in \mathcal{P}(I \times \mathbb{R}) \mid f \text{ is continuously differentiable}\}$$

Furthermore, we define the set (for $n \in \mathbb{N}^*$)

$$C^n(I;\mathbb{R}) := \{(f : I \to \mathbb{R}) \in \mathcal{P}(I \times \mathbb{R}) \mid f^{(n-1)} \text{ exists and is continuously differentiable}\}$$

Lastly, we define:

$$P(f) :\iff f \in C^1(I;\mathbb{R}) \wedge (\forall n \in \mathbb{N}^* : f \in C^n(I;\mathbb{R}) \implies f \in C^{n+1}(I;\mathbb{R}))$$

$$C^\infty(I;\mathbb{R}) := \{(f : I \to \mathbb{R}) \in \mathcal{P}(I \times \mathbb{R}) \mid P(f)\}$$

Equivalently, we can say that

$$C^\infty(I;\mathbb{R}) = \bigcap \{C^n(I,\mathbb{R}) \in \mathcal{P}(\mathcal{P}(I \times \mathbb{R})) \mid n \in \mathbb{N}^*\}$$

4.2 Mean value theorem

Stuff

4.3 Elementary functions

Recall: $\exp : \mathbb{C} \to \mathbb{C}$, $x \mapsto \displaystyle\sum_{n=0}^{\infty} \frac{x^n}{n!}$ is uniformly continuous on compact sets, converges for every $x \in \mathbb{R}$.

Proposition. *1)* $\exp(0) = 1$*, which is clear by definition.*

2) $\forall x \geq 0 : \exp(x) > 0$ *is clear, for* $x \in \mathbb{R}$ *we have proven this in the tutorial by observing that* $\exp(x) = \dfrac{1}{\exp(-x)}$ *for* $x < 0$*, and then* $\exp(-x) > 0$ *hence* $\exp(x) > 0$.

3) $\exp$ *is differentiable, and* $\exp' = \exp$.

4) $\exp \in \mathrm{C}^\infty(\mathbb{R},\mathbb{R})$ *by induction. Clearly* $\exp^{(1)} = \exp$*, then suppose* $\exp^{(n)} = \exp$ *for some* $n \in \mathbb{N}^*$*, we have that* $\exp^{(n+1)} = (\exp^{(n)})' = \exp' = \exp$ *and* $\exp$ *is continuous.*

5) $t \mapsto \exp(it)$ *is differentiable in* $(\mathbb{C}, d_2)$

6) $\forall x, y \in \mathbb{C} : \exp(x)\exp(y) = \exp(x + y)$ *as shown in the tutorial*

7) $\forall x \in \mathbb{C} : \exp(-x) = \dfrac{1}{\exp(x)}$ *(since* $\exp(x)\exp(-x) = 1$ *as shown in the tutorial, keep in mind that this is not a consequence of the previous item, since that is actually a consequence of this identity)*

8) $\exp(\mathbb{R}) = (0, \infty)$*,* $\lim_{x\to\infty} \exp(x) = \infty$*,* $\lim_{x\to-\infty} \exp(x) = 0$ *by the intermediate value theorem.*

9) Since '$\exp' = \exp > 0$'*, we know that* $\exp$ *is strictly increasing, hence* $\forall x, y \in \mathbb{R} : x < y \implies \exp(x) < \exp(y)$

10) If we force $\exp$ *to be surjective, we also know that it is injective by the increasing property as shown on the exam. Hence* $\exp : \mathbb{R} \to (0, \infty)$ *is invertible, the inverse is continuous and increasing, we define* $\ln := \exp^{-1}$.

11) By some earlier theorem, we know that $\ln'(\exp(t)) = \frac{1}{\exp'(t)} = \frac{1}{\exp(t)}$. *But since* $\exp$ *is bijective, we can set* $x =: \exp(t)$ *for a given* $x \in (0, \infty)$*, hence we can observe that*

$$\ln' : (0, \infty) \to (0, \infty)$$

$$x \mapsto x^{-1}$$

12) By some facts in the tutorial we know that if we define $a^x := \exp(\ln(a)x)$ *for some* $x \in \mathbb{R}$, $a \in (0, \infty)$*, and then as a special case consider for* $x \geq 0$

$$0^x := \lim_{a \to 0^+} \exp(\ln(a)x) = \begin{cases} 1 & \textit{if } x = 0, \\ 0 & \textit{if } x > 0. \end{cases}$$

The case where $x < 0$ *leads to inconsistencies since the limit is not a real number and if it was, then we would violate some other properties we have shown about* $\exp$*. This general definition of* a^x *is consistent with the definition of* $a^{\frac{p}{q}}$ *for* $p \in \mathbb{Z}$, $q \in \mathbb{N}^*$, $a \in \mathbb{R}$ *as we knew it. Then we also have that by definition* $e^x = \exp(\ln(e)x) = \exp(x)$*, hence* $e^x = \exp(x)$*. This is a consequence of* $\exp(1) = e$ *and hence* $1 = \ln(e)$.

Definition. Let $a > 0$. Let $x \in \mathbb{C}$. Then define the symbol

$$a^x := \exp(\ln(a)x)$$

Induce the function:

$$a^{\cdot} : \mathbb{C} \to \mathbb{C} \quad , \quad x \mapsto a^x$$

Proposition. *Let* $a > 0$.

1) For $m \in \mathbb{Z}$, $n \in \mathbb{N}^*$*, we have that* $a^{\frac{m}{n}} = \sqrt[n]{m}$

2) $\forall x, y > 0 : \ln(xy) = \ln(x) + \ln(y)$*. Also,* $\ln(\frac{1}{x}) = -\ln(x)$.

3) $\lim_{x \to 0^+} \ln(x) = -\infty$, $\lim_{x \to \infty} \ln(x) = \infty$.

4) $\ln(1) = 0$, $\ln(e) = 1$.

Definition. Define $\sin, \cos : \mathbb{R} \to \mathbb{R}$, $\sin(x) := \frac{\exp(ix) - \exp(-ix)}{2i}$, $\cos(x) := \frac{\exp(ix) + \exp(-ix)}{2}$.

Well-definition of mapping to $\mathbb{R}$ has been shown in the tutorials. We can also then observe that $\sin(x) = \mathrm{Im}(\exp(ix))$, $\cos(x) = \mathrm{Re}(\exp(ix))$, hence Euler's formula $\exp(ix) = \cos(x) + i\sin(x)$. Furthermore, we have the series expansions of sin and cos given by

$$\sin(x) = \sum_{n=0}^{\infty} \frac{x^{2n+1}}{(2n+1)!}(-1)^n$$

$$\cos(x) = \sum_{n=0}^{\infty} \frac{x^{2n}}{(2n)!}(-1)^n$$

Proposition. *1)* $\sin(0) = 0$, $\cos(0) = 1$ *(clear from series expansion).*

2) $\exp(ix) = \cos(x) + i\sin(x)$

3) $\exp(inx) = (\cos(x) + i\sin(x))^n = \cos(nx) + i\sin(nx)$

4)

$$\cos(x+y) = \cos(x)\cos(y) - \sin(x)\sin(y)$$

$$\sin(x+y) = \sin(x)\cos(y) + \cos(x)\sin(y)$$

(tutorial)

5) $\sin(-x) = -\sin(x)$ *(*$\sin$ *is an odd function),* $\cos(-x) = \cos(x)$ *(*$\cos$ *is an even function) (clear from series expansions, since* $(-x)^2n = x^2n$ *and* $(-x)^{2n+1} = -x \cdot x^2n = -x^{2n+1}$*).*

6) For $t \in [0,2]$*,* $n \in \mathbb{N}^*$*, we have*

$$\frac{t^n}{n!} \geq \frac{t^{n+2}}{(n+2)!}$$

7) $\exists x_0 \in (0,\infty) : \cos(x_0) = 0$*, define*

$$\pi := 2\min\{x_0 \in (0,\infty) \mid \cos(x_0) = 0\}$$

8) Recall (next lecture now) that we defined π *such that in particular* $\cos(\frac{\pi}{2}) = 0$*. Furthermore, in the tutorial we showed that* $\forall t \in \mathbb{R} : \sin(t + \frac{\pi}{2}) = \cos(t)$*. Similarly,* $\cos(t + 2\pi) = \cos(t)$ *and the same for* $\sin$*. We can then also show that the smallest period of these functions is* 2π*.*

9)

$$\{x \in \mathbb{R} \mid \cos(x) = 0\} = \left\{\frac{\pi}{2} + k\pi \,\middle|\, k \in \mathbb{Z}\right\}$$

$$\{x \in \mathbb{R} \mid \sin(x) = 0\} = \{k\pi \mid k \in \mathbb{Z}\}$$

10) Since $\cos(0) = 1$ *and* $\cos(\frac{\pi}{2}) = 0$ *and cosine is continuous, and* $\cos' = -\sin$*, we have that* $\cos$ *is strictly decreasing on* $[0, \frac{\pi}{2}]$ *and that* $\sin$ *is positive on the same interval. Making more of these observations and using periodicity, we can sketch cosine and sine.*

11) $\lim_{x\to\pm\infty}\cos(x)$ *does not exist (pick for example a sequence* $a_n := \frac{\pi}{2} + n\pi$*, and a sequence* $b_n := 2n\pi$*, then clearly as* $n \to \infty$*,* $a_n \to \infty$ *and* $b_n \to \infty$ *but* $\cos(a_n) = 0$ *and* $\cos(b_n) = 1$*, then we have found two sequences converging to* ∞ *that do not agree on the limit).*

12) $\cos$ *has local maxima at every* $2k \cdot \pi$ *and local minima at* $(2k+1) \cdot \pi$

4.4 The theorem by Taylor

Motivation: find sufficient conditions for local extrema, and how to approximate sufficiently differentiable functions by polynomials.

Definition. Let I be an interval of more than one point. Let $y \in I$. Let $f : I \to \mathbb{R}$ n-times differentiable at y, for some $n \in \mathbb{N}$ (including zero). Then the polynomial $T_{n,y}(f)$ given by

$$T_{n,y}(f) : \mathbb{R} \to \mathbb{R}$$

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

is called the **Taylor polynomial** of f of nth order at y.

Theorem. *Let* $n \in \mathbb{N}$*. Let* I *be an interval of more than one point. Let* $y \in I$*. Let* $f : I \to \mathbb{R}$*,* $f \in C^n(I,\mathbb{R})$*, and* $f^{(n)}$ *is assumed to be differentiable on non-boundary points of* I*. Then for every* $x, y \in I$ *and* $x \neq y$ *there exists some* $\xi \in (\min\{x,y\}, \max\{x,y\})$ *such that*

$$(R_{n,y}(f))(x) := f(x) - (T_{n,y}(f))(x) = \frac{(x-y)^{n+1}}{(n+1)!} f^{(n+1)}(\xi)$$

($R_{n,y}(f)$ *is called the 'rest term' or 'remainder' of the Taylor polynomial* $T^{n,y}(f)$*)*

Proof. Define

$$F, G : I \to \mathbb{R}$$

$$F(t) := f(x) - \sum_{k=0}^{n} \frac{(x-t)^k}{k!} f^{(k)}(t)$$

$$G(t) := (x-t)^{n+1}$$

Since $f \in C^n$, f is in particular continuous in $[\min\{x,y\}, \max\{x,y\}]$ and differentiable in $(\min\{x,y\}, \max\{x,y\})$, and the same holds for G trivially. Observe that

$$G'(t) = -(n+1)(x-t)^n \neq 0$$

for $t \neq x$, hence $t \in (\min\{x,y\}, \max\{x,y\})$ is sufficient. Then we apply the general mean value theorem, and we first make the observations that $F(x) = 0$, $F(y) = (R_{n,y}(f))(x)$, $G(x) = 0$. Furthermore, $F'(t) = \ldots = -\frac{(x-t)^n}{n!} f^{(n+1)}(t)$. We have that there exists some $\xi \in (\min\{x,y\}, \max\{x,y\})$ such that

$$\frac{f(x) - (T_{n,y}(f))(x)}{(x-y)^{n+1}} = \frac{F(y) - F(x)}{G(y) - G(x)} = \frac{F'(\xi)}{G'(\xi)} = \frac{\frac{(x-\xi)^n}{n!} f^{(n+1)}(\xi)}{(n+1)(x-\xi)^n} = \frac{f^{(n+1)}(\xi)}{(n+1)!}$$

Hence we are done. ☐

Remark. (i) By choosing G differently, one can find alternative formulations/representations of the remainder term.

(ii) If we assume $\forall z \in I : f^{(n+1)}(z) = 0$, then $f = T_{n,y}$. This proves that polynomials of degree at most n are the only solutions to $f^{(n+1)}(x) = 0$ for all $x \in I$ (as a differential equation).

(iii) Observe that if we apply Taylors theorem around 0 with exp we find that

$$(T_{n,0}(\exp))(x) = \sum_{k=0}^{n} \frac{x^k}{k!} \underbrace{\exp^{(k)}(0)}_{\exp(0)=1} = \sum_{k=0}^{n} \frac{x^k}{k!}$$

Hence

$$(R_{n,0}(\exp))(\xi) = \sum_{k=n+1}^{\infty} \frac{x^k}{k!}$$

for appropriate ξ.

Corollary. *Let $m \in \mathbb{N}^*$, $m > 1$. Let $a, b \in \mathbb{R}$, $a < b$, $f : (a,b) \to \mathbb{R}$ m-times differentiable. Assume $f^{(m)}$ is continuous at $x \in (a,b)$. If $f'(x) = f''(x) = \ldots = f^{(m-1)}(x) = 0$ and $f^{(m)}(x) \neq 0$, then:*

$$\begin{cases} m \text{ even} \implies & \begin{cases} \text{local maximum} & \text{if } f^{(m)}(x) < 0 \\ \text{local minimum} & \text{if } f^{(m)}(x) > 0 \end{cases} \\ m \text{ odd} \implies & f \text{ has no local extremum at } x \end{cases}$$

Proof. Use Taylor's theorem for $m = n+1$, $y = x$. ☐

5 Riemann & Darboux Integral

Motivation: we would like to define/formalize the concept of an ‘area underneath function’, and ‘revert’ differentiation in some way. The main idea is to draw a fancy picture of small rectangles that approximate a function, and then make sure our definition of an integral lies between these for any sufficiently small choice of rectangles.

Definition. Let $a, b \in \mathbb{R}$ and $a < b$. Let $S \subseteq [a, b]$. S is called a **partition** of $[a, b]$ if:

1. S is finite, and
2. $a, b \in S$

The convention is to write the elements of a partition P using the total ordering of the real numbers, and give them an index accordingly. For example, suppose P is a partition of $[a, b]$ and $|P| = N + 1$. Then we have $t_0, \dots, t_N \in P$ such that $t_0 = a$, $t_N = b$ and for all $n \in \{0, \dots, N-1\}$ we have that $t_n < t_{n+1}$.

Definition. Let $a, b \in \mathbb{R}$, $a < b$, $f : [a, b] \to \mathbb{R}$ such that f is bounded. Let P be a partition of $[a, b]$ and write $P = \{t_0, \dots, t_N\}$. We define the following numbers (called the upper and lower sum of f with respect to partition P, respectively)

$$U_P(f) := \sum_{i=1}^{N} (t_i - t_{i-1}) \sup_{t \in (t_{i-1}, t_i)} f(t)$$

$$L_P(f) := \sum_{i=1}^{N} (t_i - t_{i-1}) \inf_{t \in (t_{i-1}, t_i)} f(t)$$

Furthermore, we define the following (if they exist):

$$\overline{\int_{[a,b]}} f := \inf\{U_P(f) \in \mathbb{R} \mid P \text{ partition of } [a, b]\}$$

$$\underline{\int_{[a,b]}} f := \sup\{U_P(f) \in \mathbb{R} \mid P \text{ partition of } [a, b]\}$$

Definition. Let $a, b \in \mathbb{R}$, $a < b$, $f : [a, b] \to \mathbb{R}$ such that f is bounded. Then f is called **integrable** if there exists some $L \in \mathbb{R}$ such that

$$L = \overline{\int_{[a,b]}} f = \underline{\int_{[a,b]}} f$$

If this L exists, we introduce the notation:

$$\int_a^b f(x)\, dx := L$$

Facts.

1. We have

$$\begin{aligned} L_P(f) &= \sum_{i=1}^{N} (t_i - t_{i-1}) \inf_{t \in (t_{i-1}, t_i)} f(t) \\ &\geq \sum_{i=1}^{N} (t_i - t_{i-1}) \inf_{t \in [a,b]} f(t) \\ &= \inf_{t \in [a,b]} f(t) \sum_{i=1}^{N} (t_i - t_{i-1}) \\ &= \inf_{t \in [a,b]} f(t)(t_N - t_0) \\ &= \inf_{t \in [a,b]} f(t)(b - a) \end{aligned}$$

   Similar procedure for $U_P(f)$.

2. If $P = t_0, \dots, t_N$ is a partition of $[a,b]$ and $f : [a,b] \to \mathbb{R}$ such that for all $i \in \{1, \dots, N\}$ (where $N := |P|$) we have that there exists some $C_i \in \mathbb{R}$ such that for all $t \in (t_{i-1}, t_i)$ we have $f(t) = C_i$, then f is called a **step function**, and f is integrable. This is pretty clear since the upper and lower sums are bounded by the sums over this particular partition, and this can only get smaller by choosing 'finer' partitions, but then the value of the upper and lower sums are still the same.
3. $f : [a,b] \to \mathbb{R}$ is integrable if and only if for all $\varepsilon > 0$ there exists some partition P of $[a,b]$ such that
$$U_P(f) - L_P(f) < \varepsilon$$
This follows quite quickly from the definitions of inf and sup.
4. $f : [a,b] \to \mathbb{R}$ integrable if and only if there exists sequences of step functions g_n and h_n such that for all $n \in \mathbb{N}^*$ and all $x \in [a,b]$ we have $g_n(x) \leq f(x) \leq h_n(x)$ and $\lim_{n\to\infty} \int_a^b g_n(x)\,dx = \lim_{n\to\infty} \int_a^b h_n(x)\,dx =: L$ and then $\int_a^b f(x)\,dx = L$.

Proposition. *Let $a, b \in \mathbb{R}$, $a < b$ and $f, g : [a,b] \to \mathbb{R}$ integrable. Then*

- *For all $\lambda \in \mathbb{R}$ we have that $f + \lambda g$ is integrable and*
$$\int_a^b (f(x) + \lambda g(x))\,dx = \int_a^b f(x)\,dx + \lambda \int_a^b f(x)\,dx$$
(that is, the integral is linear).
- *If for all $x \in [a,b]$ we have $f(x) \leq g(x)$, then we have*
$$\int_a^b f(x)\,dx \leq \int_a^b g(x)\,dx$$
that is, the integral is monotonic.
- *Define $f_+(x) := \max\{f(x), 0\}$ and $f_-(x) := \min\{f(x), 0\}$ for all $x \in [a,b]$. Then we obviously have that $|f| = f_+ + f_-$ which is integrable because f is integrable and the zero function is as well, hence the absolute value function is also integrable.*
- *If we assume that $f \cdot g$ is integrable, and we have that there exists some $c > 0$ such that for all $x \in [a,b]$ we have $g(x) > c$, then also $\frac{f}{g}$ is integrable. We also have that*
$$\left| \int_a^b f(x)\,dx \right| \leq \int_a^b |f(x)|\,dx$$

Proof. Super trivial for step functions, since if we have partitions for those functions, we can take the union of those partitions to still have a partition of $[a,b]$ that still make f, g step functions, and then all the results follow immediately. If f and g are not step functions, then we know that we can at least bound f and g by sequences of step functions like in the facts. This is left as an excercise. ☐

Example. We take the Dirichlet function, which is defined as follows:

$$f : [0,1] \to \mathbb{R}$$
$$x \mapsto \begin{cases} 1 & \text{if } x \in \mathbb{Q} \cap [0,1], \\ 0 & \text{if } x \in (\mathbb{R} \setminus \mathbb{Q}) \cap [0,1]. \end{cases}$$

We claim that f is not integrable. We do this by showing that the upper integral is not equal to the lower integral.

Take any partition $P = \{t_0, \dots, t_N\}$ of $[0,1]$. Consider that both rationals and irrationals are dense in $[0,1]$, in other words, if we take $i \in 1, \dots, N$, we can always find some $x_0 \in (t_{i-1}, t_i)$ such that $f(x_0) = 0$ and x_1 similarly such that $f(x_1) = 1$. Hence $\sup f(t) = 1$ and $\inf f(t) = 0$. Then we have

$$U_P(f) = \sum_{i=1}^{N} (t_i - t_{i-1}) = t_N - t_0 = 1 - 0 = 1$$

and clearly, $L_P(f) = 0$, which holds for all partitions, so in particular we have that the upper and lower integrals are not equal, so f is not integrable.

Example. If we define a slightly different function, we still have an interesting result:

$$f : [0,1] \to \mathbb{R}$$

$$x \mapsto \begin{cases} \frac{1}{\text{denominator of } x} & \text{if } x \in \mathbb{Q} \cap [0,1], \\ 0 & \text{if } x \in (\mathbb{R} \setminus \mathbb{Q}) \cap [0,1]. \end{cases}$$

This does not seem well-defined, but: if $x \in \mathbb{Q}$, then there exists $p, q \in \mathbb{Z}$ such that $x = [(p,q)]_{\approx}$ as we have seen in the fundamentals of mathematics. It is also not particularly difficult to show that there exist $p', q' \in \mathbb{Z}$ such that $\gcd(p', q') = 1$ and $pq' = p'q$. Then we define the 'denominator of x' as just the number q', of course embedded into $\mathbb{R}$.
This function is called Thomae's function, which is interestingly integrable, even though it jumps 'infinitely often' and is not continuous.

Definition. Let $f : [a,b] \to \mathbb{R}$ be bounded. We say that f is **regulated** on $[a,b]$ if there exists a sequence $f_n : [a,b] \to \mathbb{R}$ such that $f_n \to f$ uniformly, and for all $n \in \mathbb{N}^*$ we have that f_n is a step function. We define the set

$$\operatorname{Reg}([a,b]) := \{(f : [a,b] \to \mathbb{R}) \in \mathcal{P}([a,b] \times \mathbb{R}) \mid f \text{ is regulated}\}$$

Theorem. *Let $f \in \operatorname{Reg}([a,b])$. Then f is integrable. Furthermore, if $f : [a,b] \to \mathbb{R}$ is continuous, then f is regulated.*

Proof. Let P_n be a partition for step function f_n. Define the following functions:

$$g_n, h_n : [a,b] \to \mathbb{R}$$

$$g_n(x) := \begin{cases} \inf_{t_{i-1}, t_i} f(t) & \text{if } x \in (t_{i-1}, t_i) \\ f(t_i) & \text{if } x = t_i \end{cases}$$

and analogous for h_n and the sup. Clearly we have the inequality '$g_n \leq f \leq h_n$'. Since $f_n \to f$, we can choose $n \in \mathbb{N}^*$ such that

$$\|f_n - f\|_\infty < \frac{\varepsilon}{\int_a^b 1 \, dx}$$

Then pretty quickly we can see that

$$h_n - g_n \leq \frac{2\varepsilon}{\int_a^b 1 \, dx}$$

Hence $\int_a^b h_n(x) g_n(x) \, dx < \varepsilon$, so those are both integrable, hence f is integrable.
Second part: we show that continuous functions are regulated. If we suppose $f : [a,b] \to \mathbb{R}$ is continuous, then it is uniformly continuous because it is defined over a compact interval. Then independently we can find $\delta > 0$ such that f is ε-close to points $x, y \in [a,b]$ such that $|x - y| < \delta$. Hence we partition as follows, let $t_0 := a$, $t_N := b$

$$t_i := t_{i-1} + \delta$$

The well-definedness is a short exercise. Using this partition, we get that on each interval, two points are δ-close, so then we can construct a step function for the appropriate ε such that $\|f_\varepsilon - f\|_\infty < \varepsilon$, hence we can construct a sequence, for example $a_n = \frac{1}{n}$, and substitute that for ε, then $f_n \to f$ and also uniformly, hence f is regulated. $\square$

Proposition. *Let $a < b$, $a, b \in \mathbb{R}$.*

$$f \in \operatorname{Reg}([a,b]) \iff \forall x \in [a,b] : \left(\lim_{t \to x^-} f(t) \in \mathbb{R} \wedge \lim_{t \to x^+} f(t) \in \mathbb{R} \right)$$

Example. Any piecewise continuous function (a function that is continuous except at finitely many points) is integrable.

Example. Using the proposition we can show that Thomae's function is integrable, apparently.

Theorem. ***(Fundamental Theorem of Calculus)***. *Let $a, b \in \mathbb{R}$, $a < b$, $f : [a,b] \to \mathbb{R}$ integrable. Then the function F defined by*

$$F : [a,b] \to \mathbb{R}$$
$$F(x) := \int_a^x f(y)\, dy$$

is continuous. If additionally we have that f is continuous at $x_0 \in (a,b)$, then F is differentiable at x_0, and $F'(x_0) = f(x_0)$.

Proof. Let $x \in [a,b]$, we show that F is continuous at x. Let $h \neq 0$ such that

$$\underbrace{[\min\{x, x+h\}, \max\{x, x+h\}]}_{=:I_h} \subseteq [a,b]$$

We clearly have $|I_h| = |h|$. We compute

$$\begin{aligned}
|F(x+h) - F(x)| &= \left| \int_x^{x+h} f(y)\, dy \right| \\
&\leq \int_{I_h} |f(y)|\, dy \\
&\leq \int_{I_h} \sup_{y \in I_h} |f(y)|\, dy \\
&= |h| \|f\|_\infty \xrightarrow{h \to 0} 0 \\
\implies \lim_{h \to 0} F(x+h) &= F(x) \\
\implies \lim_{t \to x} F(t) &= F(x)
\end{aligned}$$

We now suppose that f is continuous at some $x_0 \in (a,b)$. We compute

$$\begin{aligned}
\left| \frac{F(x+h) - F(x)}{h} - f(x) \right| &= \left| \frac{1}{h} \int_x^{x+h} f(y)\, dy - f(x) \right| \\
&= \left| \frac{1}{h} \left(\int_x^{x+h} f(y)\, dy - hf(x) \right) \right| \\
&= \left| \frac{1}{h} \int_x^{x+h} (f(y) - f(x))\, dy \right| \\
&\leq \frac{1}{|h|} \int_{I_h} |f(y) - f(x)|\, dy \\
&\leq \frac{1}{|h|} \int_{I_h} \max_{y \in I_h} |f(y) - f(x)|\, dy \\
&= \max_{y \in I_h} |f(y) - f(x)| \xrightarrow{h \to 0} 0 \\
\implies \lim_{h \to 0} \frac{F(x+h) - F(x)}{h} &= f(x)
\end{aligned}$$

□

Corollary. *Under the same assumptions as the FTC, we have that if f is continuous then it has an antiderivative, namely F. Furthermore, for any antiderivative G of f, we have that*

$$\int_a^b f(y)\,dx = G(b) - G(a)$$

Remark. Because integration involves finding an antiderivative, and as we have seen, that could be very difficult, people usually say that

> 'Differentiation is work, integration is art'

Differentiation usually has very clear rules to follow and apply, which basically compute the derivative of most nice functions, and otherwise some more involved tricks can be applied, but it can be done. For integration, it is sometimes not even possible to find an explicit antiderivative, and if it is, it is really easy to check your answer by differentiating, but really hard to come up with, as there is no general procedure that works.

Corollary.

1. *If $H(x) := \int_x^c f(y)\,dy$, then $H'(x) = -f(x)$.*
2. *If $f, g : [a,b] \to \mathbb{R}$ are continuously differentiable, then we have that*

$$\int_a^b f(x)g'(x)\,dx = f \cdot g\Big|_a^b - \int_a^b f'(x)g(x)\,dx$$

   *This is called **integration by parts**.*

3. *If $h : [\alpha, \beta] \to [a,b]$ differentiable, then*

$$\int_\alpha^\beta f(h(y))h'(y)\,dy = \int_{h(\alpha)}^{h(\beta)} f(y)\,dy$$

   *This is called the **substitution rule**.*

Remark. To apply those rules, you are allowed to use certain notations that you might have learnt in school, as long as you justify why the usage is allowed. Furthermore, it is advised to write it down maybe a bit more formally whenever possible.

Remark. The notation $f\Big|_a^b$ means $f(b) - f(a)$, and furthermore the notation $f(x)\Big|_{x=a}^{x=b}$ means the same.