[1]

CALCULUS 3 NOTES

FRÉDÉRIC LATRÉMOLIÈRE

Contents

1. Using derivatives to construct local approximations

Theorem 1.1: *Characterization of first derivatives*

Let $f : I \to \mathbb{R}$ be a function defined on an open interval and $a \in I$. The function f is differentiable at a, with derivative number $f'(a)$, if and only if there exists a function $\rho : I \to \mathbb{R}$ such that:

(1) $\lim_{x\to a} \rho(x) = 0$,
(2) $\forall x \in I \quad f(x) = f(a) + (x-a)f'(a) + (x-a)\rho(x)$.

Proof. First, suppose that there exists $\rho : I \to \mathbb{R}$, with $\lim_a \rho = 0$ and:

$$\forall x \in I \quad f(x) = f(a) + (x-a)f'(a) + (x-a)\rho(x)$$

Date: September 6, 2018.

for some number $f'(a)$. Then:

$$\frac{f(x)-f(a)}{x-a} = \frac{(x-a)(f'(a)+\rho(x))}{x-a} = f'(a)+\rho(x) \xrightarrow{x\to a} f'(a),$$

so f is indeed differentiable at a, with derivative number $f'(a)$.

Now, assume conversely that f is differentiable at a, with derivative number $f'(a)$. For all $x \in I$, we set:

$$\rho(x) = \begin{cases} \frac{f(x)-f(a)}{x-a} - f'(a) \text{ if } x \neq a, \\ 0\text{otherwise.} \end{cases}$$

By definition of f being differentiable at a, we note that:

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

On the other hand, a direct computation shows that:

$$f(x) = f(a) + (x-a)f'(a) + (x-a)\rho(x),$$

as desired. □

Numerical Experiment 1.2: tan

We have $\tan(x) = \tan(0) + \tan'(0)x + x\rho(x)$ for $x \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ and ρ some function with limit 0 at 0. Since $\tan(0) = 0$ and $\tan'(0) = 1$, we get $\tan(x) \approx 0 + 1\cdot x = x$. We notice:

x	$\tan(x)$
0.1	0.1003...
0.01	0.100003...

We may apply the Taylor-Young to compute limits.

Example 1.3: $\lim_{x\to 0} \frac{\sqrt[3]{1+x}-1}{x}$

By the Taylor-Young Theorem, there exists a function ρ with $\lim_0 \rho = 0$ such that:

$$\begin{aligned}\sqrt[3]{1+x} &= \sqrt[3]{1+0} + \frac{d\sqrt[3]{1+x}}{dx}(0)\cdot x + x\rho(x) \\ &= 1 + \frac{1}{3}x + x\rho(x) = 1 + \frac{1}{3}x + x\rho(x).\end{aligned}$$

Therefore:

$$\begin{aligned}\frac{\sqrt[3]{1+x}-1}{x} &= \frac{\frac{1}{3}x + x\rho(x)}{x} \\ &= \frac{1}{3} + \rho(x) \xrightarrow{x\to 0} \frac{1}{3}.\end{aligned}$$

Example 1.4: $\lim_{x\to 0}\frac{\sqrt{4+x^2}-2}{x}$

By the Taylor-Young Theorem, we have $\sqrt{1+x}=1+\frac{1}{2}x+x\rho(x)$ for some function ρ with $\lim_0 \rho = 0$. We then compute:

$$\begin{aligned}\frac{\sqrt{4+x^2}-2}{x} &= \frac{2(\sqrt{1+\left(\frac{x}{2}\right)^2}-1)}{x}\\ &= 2\frac{\frac{1}{2}\left(\frac{x}{2}\right)^2+\left(\frac{x}{2}\right)^2\rho\left(\left(\frac{x}{2}\right)^2\right)}{x}\\ &= \frac{x}{8}+\frac{x}{4}\rho\left(\left(\frac{x}{2}\right)^2\right)\\ &\xrightarrow{x\to 0} 0.\end{aligned}$$

Example 1.5: $\lim_{x\to 0}\frac{\sqrt{9+x^2}-2}{x^2}$

$$\begin{aligned}\frac{\sqrt{9+x^2}-2}{x^2} &= \frac{3(\sqrt{1+\left(\frac{x}{3}\right)^2}-1)}{x}\\ &= 3\frac{\frac{1}{2}\left(\frac{x}{3}\right)^2+\left(\frac{x}{3}\right)^2\rho\left(\left(\frac{x}{3}\right)^2\right)}{x}\\ &= \frac{1}{18}+\frac{1}{9}\rho\left(\left(\frac{x}{3}\right)^2\right)\\ &\xrightarrow{x\to 0} \frac{1}{18}.\end{aligned}$$

Example 1.6: *Alternate computation of* $\lim_{x\to 0}\frac{\sqrt{4+x^2}-2}{x}$

Let $f(x)=\sqrt{4+x^2}$. Then $f'(x)=\frac{2x}{2\sqrt{4+x^2}}$ and thus there exists ρ with $\lim_0 \rho = 0$ such that $\sqrt{4+x^2}=\sqrt{4}+0\cdot x+x\rho(x)$. Hence:

$$\frac{f(x)-2}{x}=\frac{2+x\rho(x)-2}{x}=\rho(x)\xrightarrow{x\to 0} 0.$$

Example 1.7: $\lim_{x\to 0}\frac{\exp(2x^9)-1}{x^9}$

$$\begin{aligned}\frac{\exp(2x^9)-1}{x^9} &= \frac{1+2x^9+2x^9\rho(2x^9)-1}{x^9}\\ &= 2+\rho(x^9)\xrightarrow{x\to 0} 2.\end{aligned}$$

2. L'Hôpital's rule

Theorem 2.1: *Cauchy's mean value theorem*

If $f : [a,b] \to \mathbb{R}$ and $g : [a,b] \to \mathbb{R}$ are two continuous functions which are differentiable on (a,b), then there exists $c \in (a,b)$ such that:
$$f'(c)\,(g(b)-g(a)) = (f(b)-f(a))\,g'(c).$$

Proof. Let $\varphi : x \in [a,b] \mapsto f(x)(g(b)-g(a)) - g(x)(f(b)-f(a))$. By assumption, φ is continuous on $[a,b]$ and differentiable on (a,b). By the Mean Value Thoerem, there exists $c \in (a,b)$ such that:
$$0 = \varphi(b)-\varphi(a) = \varphi'(c)(b-a) = f'(c)(g(b)-g(a)) - g'(c)(f(b)-f(a)),$$
as desired. $\square$

Theorem 2.2: *L'Hôpital's rule*

Let $a < b \in \mathbb{R}$. Let $f : (a,b) \to \mathbb{R}$ and $g : (a,b) \to \mathbb{R}$ be two differentiable functions on (a,b). If $\lim_a f = \lim_a g = 0$, if $g'(x) \neq 0$ for all $x \in (a,b)$, and if $\frac{f'}{g'}$ has a limit on the right of a, then:
$$\lim_{x\to a^+} \frac{f(x)}{g(x)} = \lim_{x\to a^+} \frac{f'(x)}{g'(x)}.$$

Proof. Let:
$$F : x \in [a,b) \mapsto \begin{cases} f(x) \text{ if } x \in (a,b), \\ 0 \text{ if } x = a, \end{cases} \quad \text{and } G : x \in [a,b) \mapsto \begin{cases} g(x) \text{ if } x \in (a,b), \\ 0 \text{ if } x = a. \end{cases}$$
By assumption, since f and g are continuous since differentiable on (a,b) and both have limit 0 at a, we conclude that F and G are both continuous on $[a,b)$. Moreover, both are differentiable on (a,b) with $F' = f'$ and $G' = g'$.

Let $x \in [a,b)$. The function g is continuous on $[a,x]$ and differentiable on (a,x) so by the mean value theorem, there exists $\gamma \in (a,x)$ such that $g(x) = g(x)-g(a) = g'(\gamma)(x-a) \neq 0$. So g is not zero on (a,b), and thus G is not zero on (a,b).

For all $x \in (a,b)$, by the Cauchy's mean value theorem, there exists $\gamma_x \in (a,x)$ such that $G'(\gamma_x)(F(x)-F(a)) = F'(\gamma_x)(G(x)-G(a))$, which by assumption leads to $G'(\gamma_x)F(x) = F'(\gamma_x)G(x)$. Therefore:
$$\forall x \in (a,b) \quad \exists \gamma_x \quad 0 < \gamma_x - a \leqslant x - a \text{ and } \frac{F(x)}{G(x)} = \frac{F'(\gamma_x)}{G'(\gamma_x)}.$$

Let $l = \lim_{a^+} \frac{f'}{g'}$. First, by construction, we of course have $l = \lim_{a^+} \frac{F'}{G'}$. Let $\varepsilon > 0$. There exists $\delta > 0$ such that for all $x \in I$ such that $0 < x - a < \delta$, we have $\left|\frac{F'(x)}{G'(x)} - l\right| < \varepsilon$. Therefore, if $0 < x-a < \delta$, then $0 < \gamma_x - a < \delta$ so it follows that:
$$\left|\frac{f(x)}{g(x)} - l\right| = \left|\frac{F(x)}{G(x)} - l\right| = \left|\frac{F'(\gamma_x)}{G'(\gamma_x)} - l\right| < \varepsilon.$$
This concludes our proof. $\square$

Corollary 2.3

Let $b < a \in \mathbb{R}$. Let $f : (b,a) \to \mathbb{R}$ and $g : (b,a) \to \mathbb{R}$ be two differentiable functions on (b,a). If $\lim_a f = \lim_a g = 0$, if $g'(x) \neq 0$ for all $x \in (b,a)$, and if $\frac{f'}{g'}$ has a limit on the left of a, then:

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

Proof. We apply L'Hôpital's rule to the functions $h : x \in [a, 2a-b) \mapsto f(2a-x)$ and $j : x \in [a, 2a-b) \mapsto g(2a-x)$. □

Corollary 2.4

Let $d < a < b \in \mathbb{R}$. Let $J = (d,a) \cup (a,b)$. Let $f : J \to \mathbb{R}$ and $g : J \to \mathbb{R}$ be two functions which are differentiable on (d,a) and (a,b). If $\lim_a f = \lim_a g = 0$, if $g'(x) \neq 0$ for $x \in J$ and if $\frac{f'}{g'}$ has a limit at a, then:

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

Proof. This follows from the characterization of two-sided limits using one-sided limits, and the L'Hôpital's rule. □

Remark 2.5

We note that much effort was put in writing L'Hôpital rule with the assumption that neither the numerator nor the denominator function need to be differentiable at the point where the limit is sought. If we assume that we are given two functions $f : (b,d) \to \mathbb{R}$ and $g : (b,d) \to \mathbb{R}$ which are differentiable at some $a \in (b,d)$, if $g'(a) \neq 0$, if $g(a) = f(a) = 0$, and if $g(x) \neq 0$ for all $x \in (b,d)$, then:

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

This is a much simpler result, with an easy proof, which does not even assume that f and g have a derivative anywhere but at a.

Corollary 2.6

Let $f : [b,d] \to \mathbb{R}$ be continuous, and differentiable on (b,a) and (a,d) for some $a \in (b,d)$. If f' has a limit l at a then f is differentiable at a and $f'(a) = l$.

Proof. Let $J = (b,a) \cup (a,d)$. We note that $x \in J \mapsto f(x) - f(a)$ has limit 0 at a, as does $x \in J \mapsto (x-a)$. Moreover, the latter function has derivative 1 on J. By

L'Hôpital's rule, we note:

$$\lim_{x\to a}\frac{f(x)-f(a)}{x-a}=\lim_{x\to a}\frac{f'(x)}{1}=l$$

as desired. ☐

THEOREM 2.7: *L'Hôpital's rule for ∞ limits*

Let $a<b$. If $f:(a,b)\to\mathbb{R}$ and $g:(a,b)\to\mathbb{R}$ are two differentiable functions such that:

(1) $\lim_{a^+} f=\lim_{a^+} g=\infty$,
(2) $g'(x)\neq 0$ for all $x\in(a,b)$,
(3) $\frac{f'}{g'}$ has a limit on the right of a,

then:

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

Proof. We first note that by reducing b is necessary, we may assume that $g(x)\neq 0$. Write $l=\lim_{x\to a,x>a}\frac{f'(x)}{g'(x)}$.

Let $\varepsilon>0$. There exists $\delta_1>0$ such that if $0<x-a<\delta_1$ then $\left|\frac{f'(x)}{g'(x)}-l\right|<\frac{\varepsilon}{3+\varepsilon}$. Set $c=a+\delta$. By the Cauchy mean value theorem, for all $x\in(a,c)$, there exists $\gamma_x\in(x,c)$ such that $f'(\gamma_x)(g(c)-g(x))=g'(\gamma_x)(f(c)-f(x))$. Since $\lim_{a^+} g=\infty$, we have $\lim_{a^+}\frac{g(c)}{g}=\lim_{a^+}\frac{f(c)}{g}=0$, so there exists $\delta_2>0$ such that if $0<x-a<\delta_2$, we have:

$$\left|\frac{f(c)}{g(x)}\right|<\frac{\varepsilon}{3}\text{ and }\left|\frac{g(c)}{g(x)}\right|<\min\left\{\frac{\varepsilon}{3},\frac{\varepsilon}{3(|l|+1)}\right\}.$$

Moreover, we note that for all $x\in(a,c)$ we have $g(c)-g(x)\neq 0$ by the mean value thoerem since g' is never 0 on (a,c). Now if $0<x-a<\min\{\delta_1,\delta_2\}$, we estimate:

$$\begin{aligned}\left|\frac{f(x)}{g(x)}-l\right| &\leqslant \left|\frac{f(x)-f(c)}{g(x)}-l\right|+\left|\frac{f(c)}{g(x)}\right|\\ &=\left|\frac{f(x)-f(c)}{g(x)-g(b)}-l\frac{g(x)}{g(x)-g(c)}\right|\left|\frac{g(x)-g(c)}{g(x)}\right|+\left|\frac{f(b)}{g(x)}\right|\\ &\leqslant\left|\frac{f(x)-f(c)}{g(x)-g(b)}-l\right|\left|\frac{g(x)-g(c)}{g(x)}\right|+\left|l\frac{g(c)}{g(x)}\right|+\left|\frac{f(b)}{g(x)}\right|\\ &=\left|\frac{f'(\gamma_x)}{g'(\gamma_x)}-l\right|\left|1-\frac{g(c)}{g(x)}\right|+\left|l\frac{g(c)}{g(x)}\right|+\left|\frac{f(c)}{g(x)}\right|\\ &\leqslant\varepsilon.\end{aligned}$$

This concludes our proof. ☐

Corollary 2.8

Let $b < a$. If $f : (b, a) \to \mathbb{R}$ and $g : (b, a)) \to \mathbb{R}$ are two differentiable functions such that:

(1) $\lim_{a^-} f = \lim_{a^-} g = \infty$,
(2) $g'(x) \neq 0$ for all $x \in (b, a)$,
(3) $\frac{f'}{g'}$ has a limit on the left of a,

then:

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

Proof. We replace f and g with $x \in (a, 2a - b) \mapsto f(2a - x)$ and $x \in (a, 2a - b) \mapsto g(2a - x)$ and apply our previous theorem. □

Corollary 2.9

Let $a > 0$. Let $f : (a, \infty) \to \mathbb{R}$ and $g : (a, \infty) \to \mathbb{R}$ be two differentiable functions with $g'(x) \neq 0$ for $x > a$. If $\frac{f'}{g'}$ has a limit at ∞ and if $\lim_\infty f = \lim_\infty g = 0$ (or if $\lim_\infty f = \lim_\infty g = \pm\infty$) then:

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

Proof. Let $l = \lim_\infty \frac{f}{g}$. Let $h : x \in (0, \frac{1}{a}) \mapsto h\left(\frac{1}{x}\right)$ and $j : x \in (0, \frac{1}{a}) \mapsto g\left(\frac{1}{x}\right)$. We note that h and j are differentiable with $h'(x) = \frac{-1}{x^2} f'\left(\frac{1}{x}\right)$ and $j'(x) = \frac{-1}{x^2} g'\left(\frac{1}{x}\right)$ for $x \in (0, a^{-1})$. Therefore:

$$\frac{h'(x)}{j'(x)} = \frac{f'(\frac{1}{x})}{g'(\frac{1}{x})} \xrightarrow{x \to 0^+} l.$$

Moreover by assumption, $\lim_{0^+} h = \lim_{0^+} j = 0$ (or both are infinity). Consequently, we may apply L'Hôpital's rule to deduce:

$$\lim_{x \to 0, x > 0} \frac{h(x)}{j(x)} = l.$$

Therefore $\lim_\infty \frac{f}{g} = l$ as desired. □

3. Taylor-Young Theorem

To motive our result, we begin with:

Theorem 3.1

If P is a polynomial of degree n then for any $a \in \mathbb{R}$:

$$\forall x \in \mathbb{R} \quad P(x) = \sum_{j=0}^{n} \frac{P^{(j)}(a)}{j!} (x - a)^j.$$

Proof. Let $a \in \mathbb{R}$. We make the induction hypothesis that if P is a polynomial of degree at most n then:

$$\forall x \in \mathbb{R} \quad P(x) = \sum_{j=0}^{n} \frac{P^{(j)}(a)}{j!}(x-a)^j.$$

The result holds for $P = 0$ trivially. If $n = 0$, i.e. P is a nonzero constant polynomial, then $P(x) = P(a)$ for any $x \in \mathbb{R}$. Assume our hypothesis holds for some $n \in \mathbb{N}$. Let P be a polynomial of degree at most $n+1$, written as $P(X) = \sum_{j=0}^{n} c_j X^j$ for $c_0, \ldots, c_{n+1} \in \mathbb{R}$. Then P' has degree at most n, so by our induction hypothesis:

$$\forall x \in \mathbb{R} \quad P'(x) = \sum_{j=0}^{n} \frac{P^{(j+1)}(a)}{j!}(x-a)^j.$$

On the other hand:

$$P'(x) = \sum_{j=1}^{n+1} j c_j (x-a)^{j-1} = \sum_{j=0}^{n} (j+1) c_{j+1} (x-a)^j.$$

for all $x \in \mathbb{R}$. By uniqueness of the coefficients of polynomial, we conclude that $c_{j+1} = \frac{P^{(j+1)}(a)}{(j+1)\cdot j!} = \frac{P^{(j+1)}(a)}{(j+1)!}$ for $j \in \{0, \ldots, n\}$. We also note again that $c_0 = P(a)$. We thus have proven our induction hypothesis for all $n \in \mathbb{N}$. $\square$

The main theorem of this section is:

Theorem 3.2: *Taylor-Young*

If $f : I \rightarrow \mathbb{R}$ is a function continuous on an open interval I and $a \in I$, and if f has n derivatives on (a, b), then there exists $\rho_a : I \rightarrow \mathbb{R}$ with $\lim_a \rho_a = \rho_a(a) = 0$ and:

$$f(x) = \sum_{j=0}^{n} \frac{f^{(j)}(a)}{j!}(x-a)^j + (x-a)^n \rho(x).$$

Remark 3.3

The conclusion of Theorem (3.2) can be restated as:

$$\lim_{x \to a} \frac{f(x) - \sum_{j=0}^{n} \frac{f^{(j)}(a)}{j!}(x-a)^j}{(x-a)^n} = 0. \tag{3.1}$$

This limit obviously follows from Theorem (3.2). On the other hand, if Equation (3.1) holds, then setting:

$$\rho : x \in I \longmapsto \begin{cases} \frac{f(x) - \sum_{j=0}^{n} \frac{f^{(j)}(a)}{j!}(x-a)^j}{x^n} & \text{if } x \neq a, \\ 0 \text{ if } x = a. \end{cases}$$

We will employ either formulation whenever convenient.

We propose two proofs of this important theorem.

First proof of Taylor-Young's theorem. We make the induction hypothesis that if $n \in \mathbb{N}$ and $f : I \to \mathbb{R}$ is a function with n derivatives at $a \in I$ for some open interval I then:

$$\lim_{x\to a} \frac{f(x) - \sum_{j=0}^{n} \frac{f^{(j)}(a)}{j!}(x-a)^j}{(x-a)^n} = 0.$$

For $n = 1$, we note that the induction hypothesis holds by the very definition of the derivative. Assume now that our hypothesis holds for some $n \in \mathbb{N}$. Let $f : I \to \mathbb{R}$ be some function $(n+1)$ times differentiable at some a in an open interval I. In particular, f' is well-defined on some open interval $J \subseteq I$ with $a \in J$, and has n derivatives at a. By induction hypothesis:

$$\lim_{x\to a} \frac{f'(x) - \sum_{j=0}^{n} \frac{f^{(j+1)}(a)}{j!}(x-a)^j}{(x-a)^n} = 0.$$

Now, the function $x \in J \mapsto (x-a)^{n+1}$ is differentiable on J with nonzero derivative except at a. By L'Hôpital rule, we then have:

$$\lim_{x\to a} \frac{f(x) - \sum_{j=0}^{n+1} \frac{f^{(j)}(a)}{j!}(x-a)^j}{(x-a)^{n+1}} = \lim_{x\to a} \frac{1}{n+1} \frac{f'(x) - \sum_{j=0}^{n} \frac{f^{(j+1)}(a)}{j!}(x-a)^j}{(x-a)^n} = 0.$$

This complete our proof by induction. $\square$

Our second proof requires a few lemmas.

Lemma 3.4

If $f : [a,b] \to \mathbb{R}$ and $g : [a,b] \to \mathbb{R}$ are continuous functions over $[a,b]$, differentiable on (a,b), and if for all $x \in (a,b)$:

$$f'(x) \leqslant g'(x)$$

then:

$$f(b) - f(a) \leqslant g(b) - g(a).$$

Proof. Let $h = f - g$. Note that h is continuous on $[a,b]$ and differentiable on (a,b), so by the mean value theorem, there exists $c \in (a,b)$ such that $h(b) - h(a) = h'(c)(b-a)$. Now, $h' = f' - g' \leqslant 0$, so:

$$\begin{aligned} f(b) - f(a) - (g(b) - g(a)) &= f(b) - g(b) - (f(a) - g(a)) \\ &= h(b) - h(a) \\ &\leqslant 0 \end{aligned}$$

as desired. $\square$

Lemma 3.5: *Comparative growth rate inequality*

If $f : [a,b] \to \mathbb{R}$ and $g : [a,b] \to \mathbb{R}$ are continuous functions over $[a,b]$, differentiable on (a,b), and if for all $x \in (a,b)$:

$$|f'(x)| \leqslant g'(x)$$

then:

$$|f(b) - f(a)| \leqslant g(b) - g(a).$$

Proof. By assumption, $f' \leqslant g'$ on (a,b), so $f(b) - f(a) \leqslant g(b) - g(a)$. Moreover, again by assumption, $-f' = (-f)' \leqslant g'$, so $(-f)(b) - (-f)(a) \leqslant g(b) - g(a)$, i.e. $f(a) - f(b) \leqslant g(b) - g(a)$. Hence $|f(b) - f(a)| \leqslant g(b) - g(a)$ as desired. □

Lemma 3.6

If $f : I \to \mathbb{R}$ with I an open interval and $a \in I$, and if, some some $n \in \mathbb{N}$, the function f has $n+1$ derivative at a, and if moreover there exists a polynomial P of degree n such that:

$$\lim_{x \to a} \frac{f'(x) - P(x)}{(x-a)^n} = 0$$

then, for $Q = x \in \mathbb{R} \mapsto \int_a^x P$, the following holds as well:

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

Proof. Let f be a function with $(n+1)$ derivatives at a. We set:

$$g : x \in I \mapsto f(x) - Q(x).$$

We note that g is differentiable on some closed interval $[a-\delta, a+\delta]$ for some $\delta > 0$ and, if $x \in [a-\delta, a+\delta]$ then:

$$\begin{aligned} g'(x) &= f'(x) - Q'(x) \\ &= f'(x) - P(x). \end{aligned}$$

By assumption, $\lim_{x\to a} \frac{f'(x)-P(x)}{(x-a)^n} = 0$. Let $\varepsilon > 0$. There exists $\eta \in (0,\delta)$ such that if $|x-a| < \eta$ then $|g'(x)| < |x-a|^n \varepsilon$. Consequently by Lemma (3.5), $|g(x) - g(a)| < |x-a|^{n+1}\varepsilon$. Hence:

$$\begin{aligned} \frac{1}{|x-a|^{n+1}} |f(x) - Q(x)| &= |g(x) - g(a)| \\ &\leqslant \frac{1}{|x-a|^{n+1}} |x-a|^{n+1}\varepsilon = \varepsilon. \end{aligned}$$

Hence we have proven our Lemma. □

Second proof of Taylor-Young's theorem. Characterization of the first derivative proves our present result for $n = 1$ and the result is trivial for $n = 0$ by continuity if f at a. We now assume that for some $n \in \mathbb{N}$, if h is any function with n derivatives at a, then:

$$\lim_{x \to a} \frac{h(x) - \sum_{j=0}^{n} \frac{h^{(j)}(a)}{j!}(x-a)^j}{(x-a)^n} = 0.$$

Let now f be a function with $(n+1)$ derivatives at a. Set:

$$Q : x \in \mathbb{R} \longmapsto \sum_{j=0}^{n+1} \frac{f^{(j)}(a)}{j!}(x-a)^j.$$

Then we note that:

$$Q' : x \in \mathbb{R} \longmapsto \sum_{j=1}^{n+1} \frac{f^{(j)}(a)}{j!} \cdot j(x-a)^{j-1}$$

$$= \sum_{j=0}^{n} \frac{f^{(j+1)}(a)}{j!}(x-a)^j.$$

Now, since f' has n derivatives at a, we have by our induction hypothesis that:

$$\lim_{x\to a} \frac{f'(x) - Q'(x)}{(x-a)^n} = 0.$$

Consequently, by Lemma (3.6), we conclude that:

$$\lim_{x\to a} \frac{f(x) - Q(x)}{(x-a)^{n+1}} = 0,$$

as desired. Thus our induction hypothesis is proved by induction.

Our theorem is thus proven. □

Definition 3.7: *Taylor Polynomials*

Let $f : I \to \mathbb{R}$ be a function with n derivatives at $a \in I$. The *Taylor polynomial* of f at a is:

$$x \in \mathbb{R} \longmapsto \sum_{j=0}^{n} \frac{f^{(j)}(a)}{j!}(x-a)^j.$$

Lemma 3.8

If P is a polynomial of degree at most $n \in \mathbb{N}$, if $a \in \mathbb{R}$ and if $x \in \mathbb{R} \setminus \{a\} \mapsto \frac{P(x-a)}{(x-a)^n}$ has a limit at a then $P = 0$.

Proof. We make the following induction hypothesis: if $n \in \mathbb{N}$, $a \in \mathbb{R}$ and P is a polynomial of degree at most n then $x \in \mathbb{R} \setminus \{a\} \mapsto \frac{P(x-a)}{(x-a)^n}$ has a limit at a then $P = 0$. The result is trivial for $P = 0$. If $n = 0$ then P is contant, and the function $x \neq a \mapsto \frac{P(0)}{x-a}$ does not have a limit at a unless $P = 0$. Now, assume our hypothesis is valid for some $n \in \mathbb{N}$ and let P be a polynomial of degree at most $n+1$. First we note that if $x \in \mathbb{R} \setminus \{a\} \mapsto \frac{P(x-a)}{(x-a)^{n+1}}$ has a limit at a then we must have:

$$P(0) = \lim_{x\to a} P(x-a) = \lim_{x\to a} \frac{P(x-a)}{(x-a)^{n+1}}(x-a)^{n+1} = 0.$$

Hence there exists a polynomial Q such that $P(x) = xQ(x)$ for all $x \in \mathbb{R}$. Note that Q has degree at most n and moreover:

$$\frac{Q(x-a)}{(x-a)^n} = \frac{P(x-a)}{(x-a)^{n+1}} \xrightarrow{x\to a} 0.$$

By induction hypothesis, $Q = 0$ and therefore, $P = 0$. □

Theorem 3.9: *Uniqueness of Taylor Polynomials*

Let $f : I \to \mathbb{R}$ for some open interval I, and $a \in I$. If f has n derivatives at a, if $p : \mathbb{R} \to \mathbb{R}$ is a polynomial, and if ρ is some function on I with limit 0 at a, such that:

$$\forall x \in I \quad f(x) = p(x-a) + (x-a)^n \rho(x)$$

then $x \mapsto p(x-a)$ is the Taylor polynomial of order n of f.

Proof. By Taylor-Young theorem, there exists $\eta : I \to \mathbb{R}$ with $\lim_a \eta = 0$ such that for all $x \in I$ we have:

$$f(x) = \sum_{j=0}^{n} \frac{f^{(j)}(a)}{j!}(x-a)^j + (x-a)^n \eta(x).$$

Hence for all $x \in I$, we have:

$$p(x-a) + (x-a)^n \rho(x) = \sum_{j=0}^{n} \frac{f^{(j)}(a)}{j!}(x-a)^j + (x-a)^n \eta(x)$$

so for $x \in I$, $x \neq a$:

$$\frac{p(x-a) - \sum_{j=0}^{n} \frac{f^{(j)}(a)}{j!}(x-a)^j}{(x-a)^n} = \eta(x) - \rho(x) \xrightarrow{x \to a} 0.$$

Now, $x \mapsto p(x-a) - \sum_{j=0}^{n} \frac{f^{(j)}(a)}{j!}(x-a)^j$ is a polynomial of degree at most n by assumption. Thus by Lemma (3.8), it must be null, as desired. □

4. Computation of some Taylor Polynomials

Convention 4.1

We reserve the letter ρ for *any* function defined near 0, with $\rho(0) = 0 = \lim_0 \rho$.

Theorem 4.2

For all $n \in \mathbb{N}$ there exists $\rho : \mathbb{R} \to \mathbb{R}$ with $\rho(0) = \lim_0 \rho$ such that:

$$\forall x \in \mathbb{R} \quad \exp(x) = \sum_{j=0}^{n} \frac{x^j}{j!} + x^n \rho(x) = 1 + x + \frac{x^2}{2} + \ldots + \frac{x^n}{n!} + x^n \rho(x).$$

Proof. This follows from a direct application the function exp, which has infinitely many derivatives on $\mathbb{R}$ given by $\exp^{(n)} = \exp$ for all $n \in \mathbb{N}$ by a trivial induction ($\exp^{(0)} = \exp$ and if $\exp^{(n)} = \exp$ then $\exp^{(n+1)} = \exp' = \exp$). We conclude noting that $\exp(0) = 1$. □

Corollary 4.3

For all $n \in \mathbb{N}$ there exists $\rho : \mathbb{R} \to \mathbb{R}$ with $\rho(0) = \lim_0 \rho$ such that:

$$\forall x \in \mathbb{R} \quad \cosh(x) = \sum_{j=0}^{n} \frac{x^{2j}}{(2j)!} + x^{2n}\rho(x) = 1 + \frac{x^2}{2} + \ldots + \frac{x^{2n}}{(2n)!} + x^{2n}\rho(x).$$

Proof. Let $n \in \mathbb{N}$ and let ρ be given by Theorem (4.2) so that:

$$\exp(x) = \sum_{j=0}^{2n} \frac{x^j}{j!} + x^{2n}\rho(x).$$

Therefore:

$$\begin{aligned}
\forall x \in \mathbb{R} \quad \cosh(x) &= \frac{\exp(x) + \exp(-x)}{2} \\
&= \frac{1}{2}\left(\sum_{j=0}^{n} \frac{x^{2j}}{(2j)!} + \sum_{j=0}^{n} \frac{(-x)^{2j}}{(2j)!}\right) + x^{2n}\rho(x) \\
&= \frac{1}{2}\sum_{j=0}^{n} \frac{(1 + (-1)^j)x^{2j}}{(2j)!} + x^{2n}\rho(x) \\
&= \sum_{j=0}^{n} \frac{x^{2j}}{(2j)!} + x^{2n}\rho(x)
\end{aligned}$$

since:

$$1 + (-1)^j = \begin{cases} 0 \text{ if } j = 1, \\ 2 \text{ otherwise.} \end{cases}$$

By Theorem (3.9), we have proven our theorem. □

Corollary 4.4

For all $n \in \mathbb{N}$ there exists $\rho : \mathbb{R} \to \mathbb{R}$ with $\rho(0) = \lim_0 \rho$ such that:

$$\begin{aligned}
\forall x \in \mathbb{R} \quad \sinh(x) &= \sum_{j=0}^{n} \frac{x^{2j+1)}}{(2j+1)!} + x^{2n+1}\rho(x) \\
&= x + \frac{x^3}{3!} + \ldots + \frac{x^{2n+1}}{(2n+1)!} + x^{2n+1}\rho(x).
\end{aligned}$$

Proof. The proof is similar as the one for cosh. □

Theorem 4.5

If $n \in \mathbb{N}$, there exists $\rho : \mathbb{R} \to \mathbb{R}$ such that $\lim_0 \rho = \rho(0) = 0$ and:

$$\begin{aligned}\forall x \in \mathbb{R} \quad \sin(x) &= \sum_{j=0}^{n} \frac{(-1)^j x^{2j+1}}{(2j+1)!} + x^{2n+1}\rho(x) \\ &= x - \frac{x^3}{3!} + \frac{x^5}{5!} + \ldots + \frac{(-1)^n x^{2n+1}}{(2n+1)!} + x^{2n+1}\rho(x).\end{aligned}$$

If $n \in \mathbb{N}$, there exists $\rho : \mathbb{R} \to \mathbb{R}$ such that $\lim_0 \rho = \rho(0) = 0$ and:

$$\begin{aligned}\forall x \in \mathbb{R} \quad \cos(x) &= \sum_{j=0}^{n} \frac{(-1)^j x^{2j}}{(2j)!} + x^{2n}\rho(x) \\ &= 1 - \frac{x^2}{2!} + \frac{x^4}{4!} + \ldots + \frac{(-1)^n x^{2n}}{(2n)!} + x^{2n}\rho(x).\end{aligned}$$

Proof. We set our induction hypothesis: if $n \in \mathbb{N}$ then $\sin^{(2n)}(x) = (-1)^n \sin(x)$ and $\sin^{(2n+1)}(x) = (-1)^n \cos(x)$ for all $x \in \mathbb{R}$. A direct computation shows this hypothesis holds for $n = 0$. Assume that this hypothesis holds for some $n \in \mathbb{N}$. Then:

$$\sin^{2(n+1)} = \sin^{2n+2} = \left(\sin^{(2n)}\right)' = (-1)^n \cos' = (-1)^{n+1} \sin .$$

Similarly:

$$\sin^{2(n+1)+1} = \sin^{2n+3} = \left(\sin^{(2n+2)}\right)' = (-1)^{n+1} \sin' = (-1)^{n+1} \cos .$$

Thus our hypothesis holds for all $n \in \mathbb{N}$ by induction.

In particular, $\sin^{(2n)}(0) = 0$ and $\sin^{(2n+1)}(0) = (-1)^n$ for all $n \in \mathbb{N}$. From this, our theorem follows by applyting Taylor-Young theorem.

The proof goes along similar lines for cos. □

Definition 4.6: *Binomial coefficients*

For any $\alpha \in \mathbb{R}$ and $n \in \mathbb{N}$, $n > 0$, we set:

$$\binom{\alpha}{n} = \prod_{j=1}^{n} \frac{\alpha - (j-1)}{j} = \frac{\alpha(\alpha-1)\cdots(\alpha-n+1)}{n!}.$$

We also set $\binom{\alpha}{0} = 1$ by convention.

Remark 4.7

If $n, k \in \mathbb{N}$ and $k < n$ then $\binom{k}{n} = 0$.

Theorem 4.8: *Binomial Taylor expansion*

Let $\alpha \in \mathbb{R} \setminus \{0\}$. If $n \in \mathbb{N}$, then there exists $\rho : (-1, \infty) \to \mathbb{R}$ such that $\rho(0) = 0 = \lim_0 \rho$ and:

$$\forall x > -1 \quad (1+x)^\alpha = \sum_{j=0}^{n} \binom{\alpha}{j} x^j + x^n \rho(x).$$

Proof. We make the induction hypothesis: if $n \in \mathbb{N}$, $n > 0$ then for all $x > -1$, we have $\frac{d^n (1+x)^\alpha}{dx^n} = \prod_{j=1}^{n} (\alpha - (j-1))(1+x)^{\alpha - n}$. This hypothesis holds for $n = 1$ by direct computation. If this hypothesis holds for some $n \in \mathbb{N}$, $n > 0$, then for any $x > -1$:

$$\begin{aligned}
\frac{d^{n+1}(1+x)^\alpha}{dx^{n+1}} &= \frac{d \prod_{j=1}^{n} (\alpha - (j-1))(1+x)^{\alpha - n}}{dx} \\
&= \prod_{j=1}^{n} (\alpha - (j-1)) \frac{d(1+x)^{\alpha-n}}{dx} \\
&= \prod_{j=1}^{n} (\alpha - (j-1)) \cdot (\alpha - n) \frac{d(1+x)}{dx} \cdot (1+x)^{\alpha - n - 1} \\
&= \prod_{j=1}^{n+1} (\alpha - (j-1))(1+x)^{\alpha - (n+1)}.
\end{aligned}$$

Thus our hypothesis holds for all $n \in \mathbb{N}, n > 0$. Now, for $n \in \mathbb{N}, n > 0$, by Taylor-Young theorem, there exists $\rho : (-1, \infty) \to \mathbb{R}$ such that $\lim_0 \rho = 0 = \rho(0)$ and:

$$\begin{aligned}
\forall x > -1 \quad (1+x)^\alpha &= 1 + \sum_{j=1}^{n} \frac{\prod_{j=1}^{n} (\alpha - (j-1))}{j!} x^j + x^n \rho(x) \\
&= \sum_{j=0}^{n} \binom{\alpha}{j} x^j + x^n \rho(x)
\end{aligned}$$

as desired. □

The binomial expansion gives a few interesting corollaries.

Corollary 4.9: *Newton's Binomial formula*

If $x, y \in \mathbb{R}$ and $n \in \mathbb{N}$ then:

$$(x+y)^n = \sum_{j=0}^{n} \binom{n}{j} x^j y^{n-j}.$$

Calculus based proof. We note that there is a very general algebraic proof of this formula valid in any commutative ring, but here we present a proof based on Taylor's expansion.

Let $x, y \in \mathbb{R}$ and $n \in \mathbb{N}$, $n > 0$. Note first that if $y = 0$ then the result is trivial (the only nonzero term in the right hand side sum is x^n, i.e. the only term without a positive power of y). Assume now $y \neq 0$.

We then have $(x+y)^n = y^n \left(1 + \frac{x}{y}\right)^n$. Now, $x \mapsto \left(1 + \frac{x}{y}\right)^n$ is a polynomial of degree n, so it equals its Taylor polynomial at order n. Now by Theorem (4.8):

$$\forall x > -1 \quad \left(1 + \frac{x}{y}\right)^n = \sum_{j=0}^{n} \binom{n}{j} \frac{x^j}{y^j}.$$

As two polynomials which agree on an infinite subset of $\mathbb{R}$ must agree on all of $\mathbb{R}$, we have shown that:

$$\forall x \in \mathbb{R} \quad \left(1 + \frac{x}{y}\right)^n = \sum_{j=0}^{n} \binom{n}{j} \frac{x^j}{y^j}.$$

Therefore:

$$\forall x \in \mathbb{R} \quad (x+y)^n = \sum_{j=0}^{n} \binom{n}{j} x^j y^{n-j},$$

as desired. □

Corollary 4.10: *Geometric sums*

If $n \in \mathbb{N}$, then there exists $\rho : (-1, \infty) \to \mathbb{R}$ with $\lim_0 \rho = \rho(0) = 0$ and:

$$\forall x > -1 \quad \frac{1}{1+x} = \sum_{j=0}^{n} (-1)^j x^j + x^n \rho(x).$$

First proof: Binomial formula. By Theorem (4.8), there exists $\rho : (-1, 1) \to \mathbb{R}$ with $\lim_0 \rho = rho(0) = 0$ and:

$$\forall x \in (-1, 1) \quad \frac{1}{1+x} = \sum_{j=0}^{n} \binom{-1}{j} x^j + x^n \rho(x) = \sum_{j=0}^{n} (-1)^j x^j + x^n \rho(x).$$

□

Second proof: algebraic manipulation. For $x \in \mathbb{R}$:

$$\begin{aligned}(1-x)\sum_{j=0}^{n} x^j &= \sum_{j=0}^{n} x^j - \sum_{j=0}^{n} x^{j+1} \\ &= 1 - x^{n+1}\end{aligned}$$

so for all $x \neq 1$ we have $\frac{1-x^{n+1}}{1-x} = \sum_{j=0}^{n} x^j$.. Hence, for all $x \neq -1$ in particular:

$$\frac{1}{1-x} = \sum_{j=0}^{n} x^j + x^n \left(\frac{x}{1-x}\right)$$

and since $\lim_{x \to 0} \frac{x}{1-x} = 0$, so by uniqueness of the Taylor polynomial, the Taylor polynomial of $x \mapsto \frac{1}{1-x}$ at 0 of order n is $x \mapsto \sum_{j=0}^{n} x^j$. Again by uniqueness of the Taylor polynomial, we conclude our theorem by substituting $-x$ for x. □

Corollary 4.11

If $n \in \mathbb{N}$ then there exists $\rho : (-1, \infty) \to \mathbb{R}$ with $\lim_0 \rho = \rho(0) = 0$ and:

$$\forall x > -1 \quad \ln(1+x) = \sum_{j=1}^{n} \frac{(-1)^{j-1}x^j}{j} + x^n \rho(x).$$

Proof. Let $n \in \mathbb{N}$, $n > 0$. Since:

$$\lim_{x \to 0} \frac{\frac{1}{1+x} - \sum_{j=0}^{n-1}(-1)^j x^j}{x^{n-1}} = 0,$$

and since $\frac{d \ln(1+x)}{dx} = \frac{1}{1+x}$ for all $x > -1$, we conclude by Lemma (3.6) that:

$$\lim_{x \to 0} \frac{\ln(1+x) - \sum_{j=1}^{n} \frac{(-1)^{j-1}x^j}{j}}{x^n} = 0,$$

and thus our corollary follows from the uniqueness of the Taylor polynomial. $\square$

Corollary 4.12

If $n \in \mathbb{N}$ then there exists $\rho : (-1, \infty) \to \mathbb{R}$ with $\lim_0 \rho = \rho(0) = 0$ and:

$$\forall x > -1 \quad \arctan(x) = \sum_{j=1}^{n} \frac{(-1)^j x^{2j+1}}{2j+1} + x^{2n+1} \rho(x).$$

Proof. By Corollary (4.10), there exists ρ with the desired property and:

$$\forall x > -1 \quad \frac{1}{1+x} = \sum_{j=0}^{n} (-1)^j x^j + x^n \rho(x)$$

so:

$$\forall x > -1 \quad \frac{1}{1+x^2} = \sum_{j=0}^{n} (-1)^j x^{2j} + x^{2n} \rho(x)$$

Since $\arctan'(x) = \frac{1}{1+x^2}$ for all $x \in \mathbb{R}$ and since:

$$\lim_{x \to 0} \frac{\frac{1}{1+x^2} - \sum_{j=0}^{n}(-1)^j x^{2j}}{x^{2n}} = 0$$

we conclude by Lemma (3.6) that:

$$\lim_{x \to 0} \frac{\arctan(x) - \sum_{j=0}^{n} \frac{(-1)^j x^{2j+1}}{2j+1}}{x^{2n+1}} = 0$$

and therefore, by uniqueness of the Taylor polynomial, we conclude our proof. $\square$

Corollary 4.13

If $n \in \mathbb{N}$, then there exists $\rho : (-1,1) \to \mathbb{R}$ such that $\lim_0 \rho = \rho(0) = 0$ and:

$$\forall x \in (-1,1) \quad \arcsin(x) = \sum_{j=0}^{n} \frac{\binom{-\frac{1}{2}}{j} x^{2j+1}}{2j+1} + x^{2n+1}\rho(x),$$

and thus:

$$\forall x \in (-1,1) \quad \arccos(x) = \frac{\pi}{2} - \sum_{j=0}^{n} \frac{\binom{-\frac{1}{2}}{j} x^{2j+1}}{2j+1} + x^{2n+1}\rho(x).$$

Proof. By Theorem (4.8), we have:

$$\lim_{x\to 0} \frac{\frac{1}{\sqrt{1-x^2}} - \sum_{j=0}^{n} \binom{-\frac{1}{2}}{j} x^{2j}}{x^{2n}} = 0$$

thus by Lemma (3.6), we thus have:

$$\lim_{x\to 0} \frac{\arcsin(x) - \sum_{j=0}^{n} \frac{\binom{-\frac{1}{2}}{j}}{2j+1} x^{2j+1}}{x^{2n+1}} = 0,$$

which concludes our corollary by uniqueness of the Taylor polynomial. ☐

We conclude with a more complicated example. As a first step, we note:

Lemma 4.14

If $f : I \to \mathbb{R}$ with I a nonempty open interval symmetric with respect to 0, if f is odd, and if f has $2n+1$ derivatives on I then:

$$\lim_{x\to 0} \frac{f(x) - \sum_{j=0}^{n} \frac{f^{(2j+1)}(0)}{(2j+1)!} x^{2j+1}}{x^{2n+1}} = 0$$

for some $a_0, \ldots, a_n \in \mathbb{R}$. If instead f is even then:

$$\lim_{x\to 0} \frac{f(x) - \sum_{j=0}^{n} \frac{f^{(2j)}(0)}{(2j)!} x^{2j}}{x^{2n}} = 0.$$

Proof. If f is an odd function and is differentiable on I then, since $f(-x) = -f(x)$ for $x \in I$, we conclude $-f'(-x) == f'(x)$ for $x \in I$ so f' is even. Similarly if f is even then f' is odd. Thus by an easy induction, if f is odd then $f^{(2n+1)}$ is even and $f^{(2n)}$ is odd if f has $2n+1$ derivatives on I. Thus in particular, $f^{(2j)}(0) = 0$ for all $j \in \{0, \ldots, n\}$. A similar proof applies to f even. ☐

Corollary 4.15

Let:
$$\begin{cases} a_0 = 1, \\ a_n = \frac{1}{2n+1}\left(\sum_{j=0}^{n-1} a_j a_{n-j-1}\right) \text{ if } n > 0. \end{cases}$$
Note that $(a_n)_{n\in\mathbb{N}} = \left(1, \frac{1}{3}, \frac{2}{15}, \frac{17}{315}, \ldots\right)$. If $n \in \mathbb{N}$ then:
$$\lim_{x\to 0} \frac{\tan(x) - \sum_{j=0}^{n} a_j x^{2j+1}}{x^{2n+1}} = 0.$$

Proof. Let $I = \left(-\frac{1}{2}, \frac{1}{2}\right)$. We shall use the relation: $\tan' = 1 + \tan^2$. As tan has infinitely many derivatives on I, Theorem (3.2) holds for all $n \in \mathbb{N}$. Let $P_n(x) = \sum_{j=0}^{2n} a_j x^{2j+1}$ be the Taylor polynomial of tan of order $2n+1$, centered at 0 — note that tan is odd. An easy computation shows that $P_1(x) = x$ so $a_0 = 1$. Let now $n \geqslant 1$. As $\tan'$ has $2n$ derivatives at 0 and since $P = \sum_{j=0}^{2n+1} \frac{\tan^{(2j+1)}(0)}{(2j+1)!} X^{2j+1}$, we note that $P' = \sum_{j=0}^{2n} \frac{(\tan')^{(2j)}(0)}{(2j)!} X^{2j}$ is the Taylor polynomial of $\tan'$ at 0 of order $2n$, and thus by Theorem (3.2):
$$\lim_{x\to 0} \frac{\tan'(x) - P_n'(x)}{x^{2n}} = 0$$
so
$$\lim_{x\to 0} \frac{1 + \tan^2(x) - P_n'(x)}{x^{2n}} = 0.$$
Since $\lim_{x\to 0} \frac{\tan(x) - P_n(x)}{x^{2n+1}} = 0$, we have $\lim_{x\to 0} \frac{\tan(x) - P_n(x)}{x^n} = 0$ so:
$$\frac{\tan^2(x) - P_n^2(x)}{x^{2n}} = \frac{\tan(x) - P_n(x)}{x^n} \frac{\tan(x) + P_n(x)}{x^n} \xrightarrow{n\to\infty} 0.$$
Therefore:
$$\lim_{x\to 0} \frac{1 + P_n^2(x) - P_n'(x)}{x^{2n}} = 0.$$
In particular, the coefficient of x^{2n} is given by $-(2n+1)a_n + \sum_{j=0}^{n-1} a_j a_{n-j-1}$ and must be zero by Lemma (3.8). ☐

5. Sequences

Definition 5.1: *Sequence of real numbers*

A *sequence* of real numbers is a function from $\mathbb{N}$ to $\mathbb{R}$.

Notation 5.2

A sequence is some function $f : \mathbb{N} \to \mathbb{R}$. We will adopt the family notation when working with sequences, where instead of writing f, we will explicitly list its values as indexed by $n \in \mathbb{N}$, i.e. we write f as $(x_n)_{n\in\mathbb{N}}$, where $x_n = f(n)$ for all $n \in \mathbb{N}$. This notation is here to help our intuitions that sequences are a sort of infinite (countable) tuple, which we can enumerate for as many entries as we want.

Definition 5.3: *Truncated sequence of real numbers*

A *truncated sequence* $(x_n)_{n \geqslant N}$ of real numbers is a real-valued function from the subset $\{n \in \mathbb{N} : n \geqslant N\}$ of natural numbers greater of equal to some $N \in \mathbb{N}$.

Convention 5.4

If $(x_n)_{n \geqslant N}$ is some truncated sequence, then setting $x_j = 0$ for all $j \in \{0, \ldots, N-1\}$ extend $(x_n)_{n \in \mathbb{N}}$ to a sequence. As we will only be interested in the problem of limits of sequences at infinity, we will adopt this convention when working with extended sequences all throughout.

Definition 5.5

If $(x_n)_{n \in \mathbb{N}}$ and $(y_n)_{n \in \mathbb{N}}$ are two sequences, then their sum is $(x_n + y_n)_{n \in \mathbb{N}}$ and their product is $(x_n y_n)_{n \in \mathbb{N}}$.

6. Subsequences

A subsequence is simply a sequence built from another one by "removing" certain entries (altogether, possibly infinitely many). We will write such sequences by specifying the function which selects which entries we keep. Note that entries of a subsequence are kept in the same order as in the original sequence and can not appear more often than they already do in the original sequence.

Lemma 6.1

If $\varphi : \mathbb{N} \to \mathbb{N}$ is a strictly increasing function, then for all $n \in \mathbb{N}$ we have $\varphi(n) \geqslant n$.

Proof. By definition, $\varphi(0) \in \mathbb{N}$ so $\varphi(0) \geqslant 0$. Assume $\varphi(n) \geqslant n$ for some $n \in \mathbb{N}$. Then $\varphi(n+1) > \varphi(n) \geqslant n$ so $\varphi(n+1) \geqslant n+1$ (as $n+1$ is the smallest element of $\mathbb{N}$ strictly greater than n). □

Definition 6.2

A subsequence of a sequence $(x_n)_{n \in \mathbb{N}}$ is a sequence of the form $(x_{\varphi(n)})_{n \in \mathbb{N}}$ for some strictly increasing function $\varphi : \mathbb{N} \to \mathbb{N}$.

7. Convergence of Sequence

The intuition for the convergence of sequence is as follows. Suppose we plot the sequence $\left(\frac{1}{n}\right)_{n>1}$ on a real line with a pencil. After putting the first few points, all other entries basically become indistinguishable from 0. This is because while none of the entries of this sequence are null, there is some error in measurement when placing our points — for instance because our pencil has some thickness. We could distinguish more entries if we "zoom" on our graph by changing our scale and using more precise instruments, but the phenomenon occurs again, just later on. In contrast, a sequence like $((-1)^n)_{n\in\mathbb{N}}$ or $(2, 3, 5, 7, 11, 13, \ldots)$ do not possess this property. In the first case, while all entries are indistinguishable from *two* points, we keep going from one of these values to the other has we plot the sequence, that is: to the precision at which we work, the sequence never seem to become constant up to minute errors. In the second case, no entry is particularly close to the other ones, so with an appropriate scale, we can keep drawing our sequence — though we will discuss that the fact we need ever more paper to do so is a hint of a interesting behavior at infinity as well.

Let us formalize our intuition saying that a sequence $(x_n)_{n\in\mathbb{N}}$ converges to some real number l when, no matter how small an open interval we draw around l — our way to model our "measurement error near l" — only finitely many entries of $(x_n)_{n\in\mathbb{N}}$ ever appear *outside* of this small interval. Another way to say this is that, after a while, all entries of $(x_n)_{n\in\mathbb{N}}$ lie within the given interval.

More formally, if I is some open interval containing l then there exists $N \in \mathbb{N}$ such that if $n \geqslant N$ then $x_n \in I$. Very importantly, the choice of N depends on I, and if N is a good choice, so is K for any $K \geqslant N$.

One last observation is helpful. An open interval I containing l always contains an open interval of the form $(l - \varepsilon, l + \varepsilon)$ for some $\varepsilon > 0$. Hence it is easy to check that we may as well work with symmetric intervals around l in our definition. Putting all this together, we get:

Definition 7.1: *Convergence of a sequence*

A sequence $(x_n)_{n\in\mathbb{N}}$ in $\mathbb{R}$ converges to $l \in \mathbb{R}$ when:
$$\forall \varepsilon > 0 \quad \exists N \in \mathbb{N} \quad \forall n \in \mathbb{N} \quad (n \geqslant N) \implies |x_n - l| < \varepsilon.$$

When a sequence does not converge, it is said to *diverge*.

Theorem 7.2: *Uniqueness of the limit of a sequence*

If a sequence $(x_n)_{n\in\mathbb{N}}$ in $\mathbb{R}$ converges to $l \in \mathbb{R}$ and $l' \in \mathbb{R}$ then $l = l'$.

Proof. Assume $l \neq l'$. Let $\varepsilon = \frac{1}{2}|l - l'|$. Then there exists $N \in \mathbb{N}$ such that for all $n \geqslant N$, we have $|x_n - l| < \varepsilon$ and there exists $N' \in \mathbb{N}$ such that for all $n \geqslant N'$ we have $|x_n - l'| < \varepsilon$. Then, for $n = \max\{N, N'\}$ we have:
$$|l - l'| \leqslant |l - x_n| + |x_n - l'| < 2\varepsilon = |l - l'|.$$
We have reached a contradiction, so $l = l'$. $\square$

Notation 7.3: *Limit notation*

If $(x_n)_{n\in\mathbb{N}}$ is a sequence in $\mathbb{R}$ which converges to l, we denote l by $\lim_{n\to\infty} x_n$. We will also say that $(x_n)_{n\in\mathbb{N}}$ has a limit, or has limit l (at infinity).

Theorem 7.4

Let $r \in \mathbb{N}, r > 0$ and $M \in \mathbb{R}$. The sequence $\left(\frac{M}{N^r}\right)_{n\in\mathbb{N}}$ converges to 0.

Proof. Let $\varepsilon > 0$. Let $\eta = \sup\{x \in \mathbb{R} : x^r \leqslant \varepsilon\}$, which exists as the supremum of a nonempty bounded subset of $\mathbb{R}$. Since $\mathbb{R}$ is Archimedean, there exists $N \in \mathbb{N}$ such that $N \geqslant \frac{|M|}{\eta}$. Thus, for all $n \geqslant N$ we have:

$$\left|\frac{M}{n^r}\right| \leqslant \frac{|M|}{N^r} < \eta^r \leqslant \varepsilon.$$

This concludes our proof. ☐

Theorem 7.5: *Finite changes do not impact convergence*

Let $(x_n)_{n\in\mathbb{N}}$ and $(y_n)_{n\in\mathbb{N}}$ be two sequences. If there exists $N \in \mathbb{N}$ such that for all $n \geqslant N$ we have $x_n = y_n$, then $(x_n)_{n\in\mathbb{N}}$ converges to $l \in \mathbb{R}$ if and only if $(y_n)_{n\in\mathbb{N}}$ converges to l.

Proof. Assume that $(x_n)_{n\in\mathbb{N}}$ converges to l. Let $\varepsilon > 0$. There exists $N' \in \mathbb{N}$ such that for all $n \geqslant N'$ we have $|x_n - l| < \varepsilon$. Let $n \geqslant \max\{N, N'\}$. Then $y_n = x_n$ and thus $|y_n - l| < \varepsilon$. Thus $(y_n)_{n\in\mathbb{N}}$ converges to l. The result is symmetric in the role of $(x_n)_{n\in\mathbb{N}}$ and $(y_n)_{n\in\mathbb{N}}$. ☐

Remark 7.6

We can thus use Theorem (7.5) to work with truncated sequences, i.e. functions from $\{n \in \mathbb{N} : n \geqslant N\}$ to $\mathbb{R}$ for an arbitrary $N \in \mathbb{N}$: such truncated sequences can be regarded as sequences by completing them with 0 for the first N entries, without changing the behavior at infinity.

Theorem 7.7: *Characterization of convergence*

A sequence $(x_n)_{n\in\mathbb{N}}$ converges to $l \in \mathbb{R}$ if and only if $(|x_n - l|)_{n\in\mathbb{N}}$ converges to 0.

Proof. This is by definition since $||x|| = |x|$ for all $x \in \mathbb{R}$. ☐

Theorem 7.8: *Convergence of subsequences*

Let $(x_n)_{n\in\mathbb{N}}$ be a sequence in $\mathbb{R}$. If $(x_n)_{n\in\mathbb{N}}$ converges, then every subsequence of $(x_n)_{n\in\mathbb{N}}$ converges to the same limit.

Proof. Let $l = \lim_{n\to\infty} x_n$. Let $\varphi : \mathbb{N} \to \mathbb{N}$ be strictly increasing. Let $\varepsilon > 0$. There exists $N \in \mathbb{N}$ such that for all $n \geqslant N$ we have $|x_n - l| < \varepsilon$. Since $\varphi(n) \geqslant n$ for all $n \in \mathbb{N}$ by Lemma (6.1), we conclude that for all $n \geqslant N$ we have $|x_{\varphi(n)} - l| < \varepsilon$. ☐

Theorem 7.9: *Another characterization of convergence*

A sequence $(x_n)_{n\in\mathbb{N}}$ converges to $l \in \mathbb{R}$ if and only if every subsequence of $(x_n)_{n\in\mathbb{N}}$ has a subsequence converging to l.

Proof. Theorem (7.8) shows that the condition is necessary: if $(x_n)_{n\in\mathbb{N}}$ converges, so does every subsequence of $(x_n)_{n\in\mathbb{N}}$.

Suppose now that $(x_n)_{n\in\mathbb{N}}$ does not converge to $l \in \mathbb{R}$. Thus, there exists $\varepsilon > 0$ such that for all $N \in \mathbb{N}$, there exists $n \geqslant N$ such that $|x_n - l| \geqslant \varepsilon$. Thus, for any $N \in \mathbb{N}$, the set $\{n \in \mathbb{N} : n \geqslant N \text{ and } |x_n - l| \geqslant \varepsilon\}$ is not empty.

Let $\varphi(0) = \min\{n \in \mathbb{N} : |x_n - l| \geqslant \varepsilon\}$. Assume that for some $n \in \mathbb{N}$ we have constructed $\varphi(0) < \varphi(1) < \ldots < \varphi(n)$ such that $|x_{\varphi(n)} - l| \geqslant \varepsilon$. Let:

$$\varphi(n+1) = \min\left\{k \in \mathbb{N} : k \geqslant \varphi(n) + 1 \text{ and } |x_k - l| \geqslant \varepsilon\right\}.$$

We thus have constructed $\varphi : \mathbb{N} \to \mathbb{N}$ strictly increasing such that for all $n \in \mathbb{N}$, we have $|x_{\varphi(n)} - l| \geqslant \varepsilon$. The subsequence $(x_{\varphi(n)})_{n\in\mathbb{N}}$ has no subsequence converging to l, thus proving our theorem. □

8. Order and Limit

Theorem 8.1: *Convergence implies bounded*

If a sequence $(x_n)_{n\in\mathbb{N}}$ in $\mathbb{R}$ converges, then it is bounded.

Proof. Let $l = \lim_{n\to\infty} x_n$. There exists $N \in \mathbb{N}$ such that for all $n \geqslant N$, we have $|x_n - l| < 1$ so $|x_n| < |l| + 1$. Let $M = \max\{|l| + 1, |x_n| : n \in \mathbb{N}, n \leqslant N\}$. Then for all $n \in \mathbb{N}$ we have $|x_n| \leqslant M$. □

Theorem 8.2: *Comparison of limits*

Let $(x_n)_{n\in\mathbb{N}}$ and $(y_n)_{n\in\mathbb{N}}$ be two convergent sequences in $\mathbb{R}$. If there exists $N \in \mathbb{N}$ such that for all $n \geqslant N$ we have $x_n \leqslant y_n$, then $\lim_{n\to\infty} x_n \leqslant \lim_{n\to\infty} y_n$.

Proof. Let $l_x = \lim_{n\to\infty} x_n$ and $l_y = \lim_{n\to\infty} y_n$. If $l_x > l_y$ then let $\varepsilon = \frac{l_x - l_y}{2} > 0$. There exists $N_x \in \mathbb{N}$ such that for all $n \in \mathbb{N}$ with $n \geqslant N_x$, we have $l_x - \varepsilon < x_n < l_x + \varepsilon$. In particular, $x_n > \frac{l_x + l_y}{2}$ for all $n \geqslant N_x$.

There exists $N_y \in \mathbb{N}$ such that for all $n \geqslant N_y$ we have $l_y - \varepsilon < y_n < l_y + \varepsilon$, so in particular $\frac{l_x + l_y}{2} > y_n$.

For $n \geqslant \max\{N_x, N_y\}$, we have $x_n > y_n$, which proves our theorem by contraposition. □

Theorem 8.3: *Lower bound*

If $(x_n)_{n\in\mathbb{N}}$ converges to $l > 0$ then there exists $N \in \mathbb{N}$ such that for all $n \geqslant N$ we have $x_n > \frac{l}{2}$.

Proof. Let $\varepsilon = \frac{1}{2}l$. There exists $N \in \mathbb{N}$ such that for all $n \geqslant N$ we have $x_n > l - \varepsilon = \frac{l}{2}$. □

Theorem 8.4: *Squeeze Theorem*

Let $(x_n)_{n\in\mathbb{N}}$, $(y_n)_{n\in\mathbb{N}}$ and $(z_n)_{n\in\mathbb{N}}$ be three sequences in $\mathbb{R}$ such that:

(1) there exists $N \in \mathbb{N}$ such that for all $n \geqslant N$ we have $x_n \leqslant y_n \leqslant z_n$,
(2) the sequences $(x_n)_{n\in\mathbb{N}}$ and $(z_n)_{n\in\mathbb{N}}$ converge and $\lim_{n\to\infty} x_n = \lim_{n\to\infty} z_n$.

Then $(y_n)_{n\in\mathbb{N}}$ converges and $\lim_{n\to\infty} y_n = \lim_{n\to\infty} x_n$.

Proof. Let $l = \lim_{n\to\infty} x_n = \lim_{n\to\infty} z_n$. Let $\varepsilon > 0$. There exists $N_x \in \mathbb{N}$ such that for all $n \geqslant N_x$ we have $|x_n - l| < \varepsilon$ so $l - \varepsilon < x_n$. There exists $N_z \in \mathbb{N}$ such that for all $n \geqslant N_z$ we have $|z_n - l| < \varepsilon$ so $z_n < l + \varepsilon$. Let $n \geqslant \max\{N, N_x, N_z\}$. Then:

$$l - \varepsilon < x_n \leqslant y_n \leqslant z_n < l + \varepsilon.$$

This concludes our proof. □

Corollary 8.5

If $(x_n)_{n\in\mathbb{N}}$ converges to $l \in \mathbb{R}$ then $(|x_n|)_{n\in\mathbb{N}}$ converges to $|l|$.

Proof. For all $n \in \mathbb{N}$ we have $0 \leqslant ||x_n| - |l|| \leqslant |x_n - l|$. Conclude by applying Theorem (8.4). □

9. Limits and Algebra

Theorem 9.1: *Linearity of limits*

If $(x_n)_{n\in\mathbb{N}}$ and $(y_n)_{n\in\mathbb{N}}$ are two convergent sequences and $t \in \mathbb{R}$ then $(x_n + ty_n)_{n\in\mathbb{N}}$ converges to $\lim_{n\to\infty} x_n + t \lim_{n\to\infty} y_n$.

Proof. Let $l_x = \lim_{n\to\infty} x_n$ and $l_y = \lim_{n\to\infty} y_n$. Let $\varepsilon > 0$. There exists $N_x \in \mathbb{N}$ such that for all $n \geqslant N_x$ we have $|x_n - l_x| < \frac{\varepsilon}{2}$. There exists $N_y \in \mathbb{N}$ such that for all $n \geqslant N_y$ we have $|y_n - l_y| \leqslant \frac{\varepsilon}{2(|t|+1)}$.

Let $n \geqslant \max\{N_x, N_y\}$. Then:

$$|(x_n + ty_n) - (l_x + tl_y)| \leqslant |x_n - l_x| + |t||y_n - l_y| < \varepsilon.$$

□

Theorem 9.2: *Limits and products*

If $(x_n)_{n\in\mathbb{N}}$ and $(y_n)_{n\in\mathbb{N}}$ are two convergent sequences and $t \in \mathbb{R}$ then $(x_n y_n)_{n\in\mathbb{N}}$ converges to $\lim_{n\to\infty} x_n \cdot \lim_{n\to\infty} y_n$.

Proof. Let $l_x = \lim_{n\to\infty} x_n$ and $l_y = \lim_{n\to\infty} y_n$. Since $(y_n)_{n\in\mathbb{N}}$ converges, it is bounded: there exists $M \in \mathbb{R}$ such that for all $n \in \mathbb{N}$ we have $|y_n| \leqslant M$. For all $n \in \mathbb{N}$ we have:

$$\begin{aligned} 0 \leqslant |x_n y_n - l_x l_y| &= |x_n y_n - l_x y_n + l_x y_n - l_x l_y| \\ &\leqslant |x_n - l_x||y_n| + |y_n - l_y||l_x| \\ &\leqslant |x_n - l_x|M + |y_n - l_y||l_x|. \end{aligned}$$

By assumption, $(|x_n - l_x|)_{n\in\mathbb{N}}$ and $(|y_n - l_y|)_{n\in\mathbb{N}}$ converge to 0. By Theorem (9.1), $(|x_n - l_x|M + |y_n - l_y||l_x|)_{n\in\mathbb{N}}$ converges to 0 as well. By Theorem (8.4), we conclude $(|x_n y_n - l_x l_y|)_{n\in\mathbb{N}}$ converges to 0. Hence $(x_n y_n)_{n\in\mathbb{N}}$ converges to $l_x l_y$. ☐

Theorem 9.3: *Reciprocal and limits*

If $(x_n)_{n\in\mathbb{N}}$ is a sequence in $\mathbb{R}$ converging to some $l \in \mathbb{R}\setminus\{0\}$, then there exists $N \in \mathbb{N}$ such that for all $n \in \mathbb{N}$ we have $x_n \neq 0$, and moreover $\left(\frac{1}{x_n}\right)_{n\geqslant N}$ converges to $\frac{1}{l}$.

Proof. If $l < 0$, then replace $(x_n)_{n\in\mathbb{N}}$ by $(-x_n)_{n\in\mathbb{N}}$, so that we may assume $l > 0$. By Theorem (8.3), there exists $N \in \mathbb{N}$ such that for all $n \geqslant N$ we have $x_n \geqslant \frac{l}{2} > 0$. Now let $\varepsilon > 0$. There exists $P \in \mathbb{N}$ such that for all $n \geqslant P$ we have $|x_n - l| < \frac{2\varepsilon}{l^2}$. Thus, for all $n \geqslant \max\{P, N\}$ we have:

$$\begin{aligned} \left|\frac{1}{x_n} - \frac{1}{l}\right| &= \frac{|l - x_n|}{l|x_n|} \\ &\leqslant \frac{|x_n - l|}{\frac{1}{2}l^2} \\ &< \varepsilon. \end{aligned}$$

This completes our proof. ☐

10. Monotone Sequences

Definition 10.1: *Decreasing Sequences*

A sequence $(x_n)_{n\in\mathbb{N}}$ is increasing when for all $n \leqslant m \in \mathbb{N}$ we have $x_n \leqslant x_m$.

Definition 10.2: *Increasing Sequences*

A sequence $(x_n)_{n\in\mathbb{N}}$ is decreasing when $(-x_n)_{n\in\mathbb{N}}$ is increasing.

Definition 10.3: *Monotone Sequences*

A sequence is monotone if it is either decreasing or increasing.

Definition 10.4: *Strict monotonicity*

A sequence $(x_n)_{n\in\mathbb{N}}$ is strictly increasing (resp. strictly decreasing) when for all $n \leqslant m \in \mathbb{N}$ we have $x_n < x_m$ (resp. $x_m < x_n$).

Theorem 10.5: *Monotone Subsequence Theorem*

If $(x_n)_{n\in\mathbb{N}}$ is a sequence in $\mathbb{R}$ then there exists a subsequence $(x_{\varphi(n)})_{n\in\mathbb{N}}$ which is monotone.

Proof. Let:

$$\mathscr{P} = \{n \in \mathbb{N} : \forall k \in \mathbb{N} \quad k \geqslant n \implies x_k \leqslant x_n\}.$$

- If $\mathscr{P}$ is infinite, then let $\varphi(0) = \min \mathscr{P}$. Assume that for some $n \in \mathbb{N}$ we have constructed $\varphi(0) < \ldots < \varphi(n) \in \mathscr{P}$. The set $\mathscr{P} \setminus \{\varphi(j) : 0 \leqslant j \leqslant n\}$ is not empty since $\mathscr{P}$ is infinite. Let $\varphi(n+1)$ be its smallest element. Now $\varphi : \mathbb{N} \to \mathbb{N}$ is strictly increasing, and by construction, if $n \leqslant m$ then $x_{\varphi(m)} \leqslant x_{\varphi(n)}$. Thus $(x_{\varphi(n)})_{n\in\mathbb{N}}$ is decreasing.
- If $\mathscr{P}$ is finite, then $\mathscr{P}$ is bounded by M. Let $\varphi(0) = M+1$. Now assume we have constructed $\varphi(0) < \ldots < \varphi(n)$ for some $n \in \mathbb{N}$ with $x_{\varphi(j)} \leqslant x_{\varphi(j+1)}$ for all $0 \leqslant j \leqslant n-1$. Since $\varphi(n) \notin \mathscr{P}$ there exists $\varphi(n+1) > \varphi(n)$ such that $x_{\varphi(n+1)} \geqslant x_{\varphi(n)}$. The sequence $(x_{\varphi(n)})_{n\in\mathbb{N}}$ thus constructed by induction is increasing.

Our theorem is thus proven. □

Theorem 10.6: *Monotone Convergence Theorem*

A monotone sequence $(x_n)_{n\in\mathbb{N}}$ is convergent if and only if it is bounded.

Proof. Any convergent sequence is bounded by Theorem (8.1). Assume now that $(x_n)_{n\in\mathbb{N}}$ is a bounded increasing sequence. The set $X = \{x_n : n \in \mathbb{N}\}$ is thus nonempty and bounded, so it has a supremum l. Let $\varepsilon > 0$. Since $l - \varepsilon$ is not an upper bound for X, there exists $N \in \mathbb{N}$ such that $x_N \geqslant l - \varepsilon$. Since $(x_n)_{n\in\mathbb{N}}$ is increasing and l is an upper bound of the sequence, we get for all $n \geqslant N$ that $l \geqslant x_n \geqslant x_N \geqslant l - \varepsilon$. Hence for all $n \in \mathbb{N}$ with $n \geqslant N$ we have $|x_n - l| < \varepsilon$, so $(x_n)_{n\in\mathbb{N}}$ converges to l.

If $(x_n)_{n\in\mathbb{N}}$ is bounded decreasing, then $(-x_n)_{n\in\mathbb{N}}$ is bounded increasing, and we conclude with the method above and Theorem (9.1). □

Theorem 10.7: *Geometric Sequences*

Let $a \in \mathbb{R}$. The sequence $(a^n)_{n\in\mathbb{N}}$ (with the convention that $a^0 = 1$ for this theorem, even if $a = 0$) converges if and only if $a \in (-1, 1]$; moreover if $|a| < 1$ then $\lim_{n\to\infty} a^n = 0$.

Proof. For all $n \in \mathbb{N}$ we note that $x_{n+1} = ax_n$. We note that if $(a^n)_{n\in\mathbb{N}}$ converges to l, then:

$$\begin{aligned} l - al &= \lim_{n\to} a^n - a \lim_{n\to\infty} a^n \\ &= \lim_{n\to} a^n - \lim_{n\to\infty} a^{n+1} \\ &= 0 \text{ since } (a^{n+1})_{n\in\mathbb{N}} \text{ is a subsequence of } (a^n)_{n\in\mathbb{N}}. \end{aligned} \tag{10.1}$$

Thus $l = al$. If $a \neq 1$ then this equation is solved if and only if $l = 0$.

The sequence $(1)_{n\in\mathbb{N}}$ converges to 1; we shall henceforth assume $a \neq 1$.

Assume first $a \geqslant 0$. We have $x_0 = 1 \geqslant 0$ and if, for some $n \in \mathbb{N}$ we have $x_n \geqslant 0$ then $x_{n+1} = ax_n \geqslant 0$. By induction $x_n \geqslant 0$ for all $n \in \mathbb{N}$. Moreover $x_{n+1} - x_n = (1-a)x_n$ for all $n \in \mathbb{N}$.

Therefore, if $a \in [0,1)$ then $(x_n)_{n\in\mathbb{N}}$ is a decreasing sequence, bounded below by 0, so it converges by Theorem (10.5), and its limit must be 0 by Equation (10.1).

Now, if $a > 1$ then $(x_n)_{n\in\mathbb{N}}$ is increasing. If it did converge, then it would have limit 0; yet $x_n \geqslant x_0 > 1$ for all $n \in \mathbb{N}$, so by Theorem (8.2), $(x_n)_{n\in\mathbb{N}}$ may not converge to 0; hence it does not converge.

Last, assume $a \leqslant 0$. If $a \in (-1, 0]$, then $(|a|^n)_{n\in\mathbb{N}} = (|a^n|)_{n\in\mathbb{N}}$ converges to 0, so $(a^n)_{n\in\mathbb{N}}$ does as well. If $a < -1$ then $a^2 \in (1,\infty)$ so $(a^{2n})_{n\in\mathbb{N}}$ does not converge, and thus $(a_n)_{n\in\mathbb{N}}$ does not converge by Theorem (7.8). Last, $((-1)^n)_{n\in\mathbb{N}}$ admits a subsequence convergent to 1 and a subsequence convergent to -1, so it can not converge, by Theorem (7.8). □

REMARK 10.8

For $a > 1$ we have shown that $(a^n)_{n\in\mathbb{N}}$ is not bounded.

11. Cauchy Sequences

DEFINITION 11.1: *Cauchy Sequences*

A sequence $(x_n)_{n\in\mathbb{N}}$ in $\mathbb{R}$ is Cauchy when:

$$\forall \varepsilon > 0 \quad \exists N \in \mathbb{N} \quad \forall p, q \in \mathbb{N} \quad (p, q \geqslant N) \implies |x_p - x_q| < \varepsilon.$$

THEOREM 11.2: *Convergence implies Cauchy*

If a sequence $(x_n)_{n\in\mathbb{N}}$ converges in $\mathbb{R}$, then it is Cauchy.

Proof. Let $l = \lim_{n\to\infty} x_n$. Let $\varepsilon > 0$. There exists $N \in \mathbb{N}$ such that for all $n \geqslant N$, we have $|x_n - l| < \frac{1}{2}\varepsilon$. Thus for all $p, q \geqslant N$, we have:

$$|x_p - x_q| \leqslant |x_p - l| + |l - x_q| < \varepsilon.$$

□

THEOREM 11.3: *Cauchy implies bounded*

If a sequence $(x_n)_{n\in\mathbb{N}}$ is Cauchy, then it is bounded.

Proof. There exists $N \in \mathbb{N}$ such that for all $p, q \geqslant N$ we have $|x_p - x_q| < 1$. Thus for all $p \geqslant N$ we have $|x_p| \leqslant |x_N| + 1$. If $M = \max\{|x_N| + 1, |x_j| : j \leqslant N\}$ then for all $p \in \mathbb{N}$ we have $|x_p| \leqslant M$. □

Theorem 11.4: *Convergence criterion for Cauchy sequences*

If $(x_n)_{n\in\mathbb{N}}$ is a Cauchy sequence which admits a subsequence converging to $l \in \mathbb{R}$ then $(x_n)_{n\in\mathbb{N}}$ converges to l.

Proof. Let $(x_{\varphi(n)})_{n\in\mathbb{N}}$ be a subsequence of $(x_n)_{n\in\mathbb{N}}$ which converges to some $l \in \mathbb{R}$. Let $\varepsilon > 0$. Since $(x_{\varphi(n)})_{n\in\mathbb{N}}$ converges to l, there exists $N \in \mathbb{N}$ such that for all $n \geqslant N$ we have $|x_{\varphi(n)} - l| < \frac{1}{2}\varepsilon$. Since $(x_n)_{n\in\mathbb{N}}$ is Cauchy, there exists $P \in \mathbb{N}$ such that for all $p, q \geqslant P$ we have $|x_p - x_q| < \frac{1}{2}\varepsilon$. Let $n \geqslant \max\{N, P\}$ and note that, by Lemma (6.1), we have $\varphi(n) \geqslant n \geqslant \max\{N, P\}$. Then:

$$|x_n - l| \leqslant |x_n - x_{\varphi(n)}| + |x_{\varphi(n)} - l| < \frac{1}{2}\varepsilon + \frac{1}{2}\varepsilon = \varepsilon.$$

This completes our proof. □

Theorem 11.5: *Completeness of* $\mathbb{R}$

A sequence $(x_n)_{n\in\mathbb{N}}$ is Cauchy in $\mathbb{R}$ if and only if $(x_n)_{n\in\mathbb{N}}$ converges in $\mathbb{R}$.

Proof. By Theorem (11.2), every convergent sequence is Cauchy. Let now $(x_n)_{n\in\mathbb{N}}$ be a Cauchy sequence. By Theorem (11.3), the sequence $(x_n)_{n\in\mathbb{N}}$ is bounded. By Theorem (10.5), there exists a monotone subsequence $(x_{\varphi(n)})_{n\in\mathbb{N}}$ of $(x_n)_{n\in\mathbb{N}}$. Since $(x_{\varphi(n)})_{n\in\mathbb{N}}$ is bounded monotone, it converges by Theorem (10.6). Since $(x_n)_{n\in\mathbb{N}}$ is a Cauchy sequence with a convergent subsequence, it must converge as well by Theorem (11.4). □

12. Series

Definition 12.1: *Series*

Let $(x_n)_{n\in\mathbb{N}}$ be a sequence. The *series* $\left(\sum x_n\right)_{n\in\mathbb{N}}$ *of general term* $(x_n)_{n\in\mathbb{N}}$ is the sequence $\left(\sum_{j=0}^n x_j\right)_{n\in\mathbb{N}}$ where, for each $n \in \mathbb{N}$, the real number:

$$\sum_{j=0}^{n} x_j = x_0 + x_1 + \ldots + x_n$$

is called the n^{th}*-partial sum of the series* $\left(\sum x_n\right)_{n\in\mathbb{N}}$.

Notation 12.2: *I*

a series $\left(\sum x_n\right)_{n\in\mathbb{N}}$ converges, then its limit is denoted by $\sum_{n=0}^{\infty} x_n$, rather than the equivalent $\lim_{n\to\infty} \sum_{j=0}^{n} x_j$. Do note that the infinite sum is a *limit*, hence must be manipulated with care. This limit is called the *sum* of the series.

Warning 12.3

The notation $\sum_{n=0}^{\infty} x_n$ should not be used unless the series converge — or in the special case of a series of positive numbers (see below).

Theorem 12.4

If $(\sum x_n)_{n\in\mathbb{N}}$ and $(\sum y_n)_{n\in\mathbb{N}}$ are two convergent series then for any real number λ:

$$\sum_{n=0}^{\infty}(\lambda x_n + y_n) = \lambda\sum_{n=0}^{\infty} x_n + \sum_{n=0}^{\infty} y_n.$$

Proof. This follows from the linearity of limits and the observation that for all $n \in \mathbb{N}$:

$$\sum_{j=0}^{n}(\lambda x_j + y_j) = \lambda\sum_{j=0}^{n} x_j + \sum_{j=0}^{n} y_j.$$

□

Warning 12.5: *Finite changes*

Let $(\sum x_n)_{n\in\mathbb{N}}$ and $(\sum y_n)_{n\in\mathbb{N}}$ be two series. If there exists $N \in \mathbb{N}$ such that for all $n \geqslant N$, we have $x_n = y_n$, then $(\sum x_n)_{n\in\mathbb{N}}$ converges if and only if $(\sum y_n)_{n\in\mathbb{N}}$ does, though not in general to the same limit — their sums would differ by $\sum_{j=0}^{N}(x_j - y_j)$. It is therefore fine to ignore finitely many terms of a series when studying convergence, but not when estimating the sum.

There is a simple a sufficient condition for divergence, or better stated, a necessary condition for convergence.

Theorem 12.6: *Divergence test*

If $(\sum x_n)_{n\in\mathbb{N}}$ is a convergent series then $\lim_{n\to\infty} x_n = 0$.

Proof. By hypothesis, $\left(\sum_{j=0}^{n} x_j\right)_{n\in\mathbb{N}}$ converges and we note for $n > 0$:

$$x_n = \sum_{j=0}^{n} x_j - \sum_{j=0}^{n-1} x_j \xrightarrow{n\to\infty} \sum_{n=0}^{\infty} x_n - \sum_{n=0}^{\infty} x_n = 0,$$

as needed. □

Example 12.7

The series $(\sum(-1)^n)_{n\in\mathbb{N}}$ diverges.

Example 12.8

The series $\left(\sum n\sin\left(\frac{1}{n}\right)\right)_{n\in\mathbb{N}}$ diverges. Indeed, there exists ρ with $\lim_o \rho = 0$ such that $\sin(x) = x + x\rho(x)$ for all $x \in \mathbb{R}$, and thus:

$$\begin{aligned} n\sin\left(\frac{1}{n}\right) &= n\left(\frac{1}{n} + \frac{1}{n}\rho(\frac{1}{n})\right) \\ &= 1 + \rho(\frac{1}{n}) \xrightarrow{n\to\infty} 1. \end{aligned}$$

Convergence, and even the sum, of certain series can be computed. Series with only finitely many nonzero entries always converge. More importantly for us, our first non-trivial example is the following:

Theorem 12.9: *Geometric Series*

The series $(\sum a^n)_{n\in\mathbb{N}}$ converges if and only if $|a| < 1$, and moreover if $|a| < 1$ then:

$$\sum_{n=0}^{\infty} a^n = \frac{1}{1-a}.$$

Proof. We know that $(a^n)_{n\in\mathbb{N}}$ converges to 0 if and only if $|a| < 1$, so the condition for convergence is clearly necessary by the divergence test.

On the other hand, for any $a \in \mathbb{R}$, we note that:

$$\begin{aligned} (1-a)\sum_{j=0}^{n} a^j &= \sum_{j=0}^{n} a^j - \sum_{j=1}^{n+1} a^j \\ &= 1 - a^{n+1}. \end{aligned}$$

Hence if $|a| < 1$ then for all $n \in \mathbb{N}$:

$$\sum_{j=0}^{n} a^j = \frac{1-a^{n+1}}{1-a} \xrightarrow{n\to\infty} \frac{1}{1-a},$$

as desired. □

Example 12.10

The series $(\sum(-1)^n 2^n)_{n\in\mathbb{N}}$ diverges while $\left(\sum \frac{1}{2^n}\right)_{n\in\mathbb{N}}$ converges to $\frac{1}{1-\frac{1}{2}} = 2$.

Series of positive terms are very important, and convergence can be tested in various ways thanks to the following key observation.

Lemma 12.11

If $(\sum x_n)_{n\in\mathbb{N}}$ is a series of positive general term, then it is increasing.

Proof. For any $n \in \mathbb{N}$ we have:

$$\sum_{j=0}^{n+1} x_j - \sum_{j=0}^{n} x_j = x_{n+1} \geqslant 0.$$

□

Consequently, a series of positive general term will either converge or sum to ∞.

13. The p-series

We now prove a deep criterion for convergence of series of a very particular form.

Theorem 13.1: *Comparison series / integral*

Let $f : [a, \infty) \to [0, \infty)$ be a positive *decreasing* function. The series $(f(n))_{n \in \mathbb{N}, n > a}$ converges if and only if the function:

$$x > 0 \mapsto \int_a^x f(t)\, dt$$

has a finite limit at ∞.

Proof. Let $n \in \mathbb{N}$, $n \geqslant a$. Since f is decreasing, we have for all $x \in [n, n+1]$:

$$f(n) \leqslant f(x) \leqslant f(n+1),$$

so

$$f(n) = \int_n^{n+1} f(n)\, dt \leqslant \int_n^{n+1} f(t)\, dt \leqslant \int_n^{n+1} f(n+1)\, dt = f(n+1)$$

and therefore:

$$\sum_{j=0}^{n} f(j) \leqslant \int_0^{n+1} f(t)\, dt \leqslant \sum_{j=0}^{n+1} f(j).$$

Let us assume first that $x > 0 \mapsto \int_0^x f$ has a finite limit. Note that this function is increasing since $f \geqslant 0$, so $\int_0^x f \leqslant \int_0^\infty f$ for all $x > 0$. Hence under this assumption, for all $n \in \mathbb{N}$:

$$\sum_{j=0}^{n} f(j) \leqslant \int_0^\infty f.$$

Now, $(\sum f(n))_{n \in \mathbb{N}}$ is increasing since it has a positive general term, and it is bounded, henceforth it converges.

On the other hand, let us assume that $x > 0 \mapsto \int_0^x f$ has no limit at infinity. This is equivalent to the sequence $\left(\int_0^n f\right)_{n \in \mathbb{N}}$ is unbounded: if it was bounded, then it would converge as it is increasing, and then $x > 0 \mapsto \int_0^x f$ would converge as well as it too is increasing.

Therefore, since $\int_0^{n+1} f(t)\, dt \leqslant \sum_{j=0}^{n+1} f(j)$ for all $n \in \mathbb{N}$, the series $(\sum f(n))_{n \in \mathbb{N}}$ is unbounded and thus it diverges. □

Corollary 13.2: *p-test*

The series $\left(\sum \frac{1}{n^p}\right)_{n > 0}$ converges if and only if $p > 1$.

Proof. We first proceed with $p \neq 1$. The function $f : x > 0 \mapsto \frac{1}{x^p}$ is decreasing as seen by taking its derivative. We then have for all $x > 0$:

$$\int_1^x \frac{1}{t^p}\,dt = \frac{1-p}{x^{p-1}} - (1-p)$$

which converges if and only if $p > 1$. Thus $\left(\sum \frac{1}{n^p}\right)_{n>0}$ converges if $p > 1$ and diverges if $p < 1$ by the integral-series comparison test.

Now, for $p = 1$, note that $t > 0 \mapsto \frac{1}{t}$ is decreasing and:

$$\int_1^x \frac{dt}{t} = \ln(x) \xrightarrow{x\to\infty} \infty$$

hence $\left(\sum \frac{1}{n}\right)_{n\in\mathbb{N}}$ diverges. □

14. Comparison and Limit Comparison

The following comparison theorem is at the core of many convergence tests.

Theorem 14.1: *Comparison Theorem*

Let $(\sum x_n)_{n\in\mathbb{N}}$ and $(\sum y_n)_{n\in\mathbb{N}}$ be two series of *positive general term.* If there exists $N \in \mathbb{N}$ such that for all $n \geqslant N$ we have:

$$x_n \leqslant y_n$$

then:

- if $(\sum y_n)_{n\in\mathbb{N}}$ converges then $(\sum x_n)_{n\in\mathbb{N}}$ converges.
- if $(\sum x_n)_{n\in\mathbb{N}}$ diverges then $(\sum y_n)_{n\in\mathbb{N}}$ diverges.

Proof. Suppose $(\sum y_n)_{n\in\mathbb{N}}$ converges. It is then bounded, so there exists $M \in \mathbb{N}$ such that for all $n \in \mathbb{N}$ we have $\sum_{j=0}^n x_j \leqslant M$. By assumption:

$$0 \leqslant \sum_{j=0}^n x_n \leqslant \sum_{j=0}^n y_n \leqslant M$$

and thus $(\sum x_n)_{n\in\mathbb{N}}$ is bounded as well; as it is increasing, it must converge as well.

The second assertion is the contrapositive of the first. So our theorem is proven. □

In many applications, the following corollary of the comparison theorem is helpful.

Theorem 14.2: *Limit comparison Theorem*

Let $(\sum x_n)_{n\in\mathbb{N}}$ and $(\sum y_n)_{n\in\mathbb{N}}$ be two series of *positive general term.* If there exists $N \in \mathbb{N}$ such that for all $n \geqslant N$ we have $y_n \neq 0$ and if for some $l > 0$:

$$\lim_{n\to\infty} \frac{x_n}{y_n} = l$$

then $(\sum x_n)_{n\in\mathbb{N}}$ converges if and only if $(\sum y_n)_{n\in\mathbb{N}}$ converges.

Proof. Let $\varepsilon = \frac{|l|}{2} > 0$. There exists $N \in \mathbb{N}$ such that for all $n \geqslant N$ we have:

$$\left|\frac{x_n}{y_n} - l\right| \leqslant \varepsilon = \frac{|l|}{2},$$

hence:

$$\frac{|l|}{2} \leqslant \frac{x_n}{y_n} \leqslant \frac{3|l|}{2}.$$

We now can simply apply the comparison theorem and linearity. □

Example 14.3: $\left(\sum \sin\left(\frac{1}{n}\right)\right)_{n>0}$

Does $\left(\sum \sin\left(\frac{1}{n}\right)\right)_{n>0}$ converge? Using Taylor-Young Theorem, there exists ρ with limit 0 at 0 such that $\sin(x) = x + x\rho(x)$. Therefore:

$$\begin{aligned}\frac{\sin\left(\frac{1}{n}\right)}{\frac{1}{n}} &= \frac{\frac{1}{n} + \frac{1}{n}\rho\left(\frac{1}{n}\right)}{\frac{1}{n}} \\ &= 1 + \rho\left(\frac{1}{n}\right) \xrightarrow{n\to\infty} 1.\end{aligned}$$

Since $\left(\sum \frac{1}{n}\right)_{n>1}$ diverges and all series involve have positive general terms, the series $\left(\sum \sin\left(\frac{1}{n}\right)\right)_{n>0}$ converges by the limit comparison theorem.

Example 14.4: $\left(\sum \left(1 - \cos\left(\frac{1}{n}\right)\right)\right)_{n>0}$

Does $\left(\sum \left(1 - \cos\left(\frac{1}{n}\right)\right)\right)_{n>0}$ converge? Using Taylor-Young Theorem, there exists ρ with limit 0 at 0 such that $\cos(x) = 1 - \frac{x^2}{2} + x^2\rho(x)$. Therefore:

$$\begin{aligned}\frac{1 - \cos\left(\frac{1}{n}\right)}{\frac{1}{n^2}} &= \frac{\frac{1}{2n^2} + \frac{1}{n^2}\rho\left(\frac{1}{n}\right)}{\frac{1}{n^2}} \\ &= \frac{1}{2} + \rho\left(\frac{1}{n}\right) \xrightarrow{n\to\infty} \frac{1}{2}.\end{aligned}$$

Since $\left(\sum \frac{1}{n^2}\right)_{n>1}$ converges and all series involve have positive general terms, the series $\left(\sum 1 - \cos\left(\frac{1}{n}\right)\right)_{n>0}$ converges by the limit comparison theorem.

15. Absolute Convergence of Series

Definition 15.1: *Absolute convergence*

A series $(\sum x_n)_{n\in\mathbb{N}}$ *converges absolutely* when $(\sum |x_n|)_{n\in\mathbb{N}}$ converges.

Theorem 15.2: *Absolute Convergence*

If $(\sum x_n)_{n\in\mathbb{N}}$ converges absolutely, then it converges.

Proof. For all $n \in \mathbb{N}$, we have $0 \leqslant x_n + |x_n| \leqslant 2|x_n|$. By the comparison test theorem, we thus conclude that if $(\sum |x_n|)_{n\in\mathbb{N}}$ converges, then $(\sum 2|x_n|)_{n\in\mathbb{N}}$ converges by linearity, and thus $(x_n + |x_n|)_{n\in\mathbb{N}}$ converges. By linearity, we then conclude that $(\sum x_n)_{n\in\mathbb{N}}$ converges, again since $(\sum |x_n|)_{n\in\mathbb{N}}$ does. □

Example 15.3

The series $\left(\sum \frac{\cos(n)}{n^2}\right)_{n>0}$ converges, since:

- for all $n \in \mathbb{N}$, $n > 0$ we have $\frac{|\cos(n)|}{n^2} \leqslant \frac{1}{n^2}$,
- the series $\left(\sum \frac{1}{n^2}\right)_{n>0}$ converges by the p-test,
- by comparison for positive term series, $\left(\sum \frac{|\cos(n)|}{n^2}\right)_{n>0}$ converges,
- hence $\left(\sum \frac{\cos(n)}{n^2}\right)_{n>0}$ converges.

16. Ratio and Root Tests

Bringing together geometric series, absolute convergence and comparison test, we get two new tests.

Theorem 16.1: *Ratio test*

If $(x_n)_{n\in\mathbb{N}}$ is a sequence of nonzero numbers, and if:

$$\lim_{n\to\infty}\left|\frac{x_{n+1}}{x_n}\right| = l$$

then:

(1) if $l < 1$ then $(\sum |x_n|)_{n\in\mathbb{N}}$, and therefore, $(\sum x_n)_{n\in\mathbb{N}}$ converge.
(2) if $l > 1$ then $(\sum x_n)_{n\in\mathbb{N}}$ diverges.

Proof. Assume first that $l < 1$. Let $\varepsilon > 0$ such that $l + \varepsilon < 1$. There exists $N \in \mathbb{N}$ such that for all $n \geqslant N$ we have:

$$\left|\frac{x_{n+1}}{x_n} - l\right| \leqslant \varepsilon$$

so for all $n \geqslant N$, we have:

$$\frac{|x_{n+1}|}{|x_n|} \leqslant l + \varepsilon.$$

Consequently, if $n \geqslant N$:

$$|x_{n+1}| \leqslant (l+\varepsilon)|x_n|$$

and thus by induction:

$$|x_{n+1}| \leqslant (l+\varepsilon)^{n-N+1}|x_N|.$$

Since $0 \leqslant l + \varepsilon < 1$, the geometric series $(\sum (l+\varepsilon)^n)_{n\in\mathbb{N}}$ converges, and thus by comparison theorem, so does $(\sum |x_n|)_{n\in\mathbb{N}}$. Henceforth $(\sum x_n)_{n\in\mathbb{N}}$ converges as well.

Now if $l > 1$ then let $\varepsilon > 0$ such that $l - \varepsilon > 1$. We note that as above, for some $N \in \mathbb{N}$, if $n \geqslant N$ then:

$$|x_n| \geqslant (l+\varepsilon)^n |x_N|$$

and thus $(x_n)_{n\in\mathbb{N}}$ is unbounded; therefore it does not converge to 0 and so by the divergence test, $(\sum x_n)_{n\in\mathbb{N}}$ diverges. □

Theorem 16.2: *Root test*

If $(x_n)_{n\in\mathbb{N}}$ is a sequence of nonzero numbers, and if:

$$\lim_{n\to\infty} \sqrt[n]{|x_n|} = l$$

then:

(1) if $l < 1$ then $(\sum |x_n|)_{n\in\mathbb{N}}$, and therefore, $(\sum x_n)_{n\in\mathbb{N}}$ converge.
(2) if $l > 1$ then $(\sum x_n)_{n\in\mathbb{N}}$ diverges.

Proof. Assume $l < 1$. Let $\varepsilon > 0$ such that $l + \varepsilon < 1$. Then there exists $N \in \mathbb{N}$ such that if $n \geqslant N$ then $\sqrt[n]{|x_n|} \leqslant l + \varepsilon$, so $|x_n| \leqslant (l+\varepsilon)^n$. Since $(\sum (l+\varepsilon)^n)_{n\in\mathbb{N}}$ converges, we conclude that $(\sum x_n)_{n\in\mathbb{N}}$ converges absolutely and hence converges by comparison.

If $l > 1$ then $(\sum x_n)$ diverges by the divergence test. □

17. Alternating Sign Test

Theorem 17.1: *Alternating sign series test*

If $(x_n)_{n\in\mathbb{N}}$ is a decreasing sequence of nonnegative numbers such that $\lim_{n\to\infty} x_n = 0$, then $(\sum (-1)^n x_n)_{n\in\mathbb{N}}$ converges.

Proof. For all $n \in \mathbb{N}$, set $S_n = \sum_{j=0}^{n} (-1)^n x_n$. Note that for all $n \in \mathbb{N}$:

$$S_{2n+2} - S_{2n} = x_{2n+2} - x_{2n+1} \leqslant 0$$

since $(x_n)_{n\in\mathbb{N}}$ is decreasing. Moreover for all $n \in \mathbb{N}$:

$$S_{2n+3} - S_{2n+1} = x_{2n+2} - x_{2n+3} \geqslant 0$$

so $(S_{2n})_{n\in\mathbb{N}}$ is decreasing and $(S_{2n+1})_{n\in\mathbb{N}}$ is increasing. Moreover:

$$S_{2n+1} - S_{2n} = -x_{2n+1} \leqslant 0$$

for all $n \in \mathbb{N}$. Therefore, for all $n \in \mathbb{N}$:

$$S_1 \leqslant S_3 \leqslant \cdots \leqslant S_{2n-1} \leqslant S_{2n+1} \leqslant S_{2n} \leqslant S_{2n-2} \leqslant \ldots \leqslant S_2 \leqslant S_0$$

In particular, $(S_{2n+1})_{n\in\mathbb{N}}$ is increasing, and bounded above by S_0, so it converges to some $l \in \mathbb{R}$. Now $S_{2n} = S_{2n+1} + x_{2n+1} \xrightarrow{n\to\infty} l$. Thus $(S_n)_{n\in\mathbb{N}}$ converges to l as desired. □

18. Power Series

Definition 18.1: *Power Series*

Let $(a_n)_{n\in\mathbb{N}}$ be a sequence. The *power series* of coefficients $(a_n)_{n\in\mathbb{N}}$ centered as $c \in \mathbb{R}$ is the function defined by:

$$x \longmapsto \sum_{n=0}^{\infty} a_n(x-c)^n$$

for all input for which the series converges.

We adopt the convention that $[-\infty, \infty] = \mathbb{R}$ in the sections on power series.

Theorem 18.2: *Domain of Power Series*

If $f : x \longmapsto \sum_{n=0}^{\infty} a_n(x-c)^n$ for some $c \in \mathbb{R}$ and some sequence $(a_n)_{n\in\mathbb{N}}$ then there exists $R \in [0,\infty]$ such that the domain of f contains $(c-R, c+R)$ and is contained in $[-R,R]$. In particular, $\left(\sum a_n(x-c)^n\right)_{n\in\mathbb{N}}$ converges absolutely to $f(x)$ when $|x-c| < R$.

Proof. The set $A = \{r > 0 | \left(|a_n|r^n\right)_{n\in\mathbb{N}} \text{ is bounded}\}$ is not empty as it contains c, so it has a possibly infinite supremum which we denote by R.

Let $x \in \mathbb{R}$ with $|x| < R$. Choose $r \in A$ such that $x < r < R$. Since $r \in A$, there exists $M > 0$ such that for all $n \in \mathbb{N}$ we have $|a_n|r^n \leqslant M$. We note that for all $n \in \mathbb{N}$ we have:

$$0 \leqslant |a_n x^n| = |a_n| r^n \frac{|x|^n}{r^n} = \left(|a_n|r^n\right)\left(\frac{|x|}{r}\right)^n \leqslant M\left(\frac{|x|}{r}\right)^n.$$

Since $0 \leqslant \frac{|x|}{r} \leqslant 1$, the series $\left(\sum M\frac{|x|^n}{r^n}\right)_{n\in\mathbb{N}}$ converges, and thus by the comparison theorem, $\left(\sum a_n x^n\right)_{n\in\mathbb{N}}$ converges absolutely and hence converges, as desired.

Now, let $x \in \mathbb{R}$ with $|x| > R$. Then $\left(|a_n x^n|\right)_{n\in\mathbb{N}}$ is not bounded and thus cannot converge to 0; consequently by the divergence test, $\left(\sum a_n x^n\right)_{n\in\mathbb{N}}$ diverges. $\square$

The number $R > 0$ given by our previous theorem is necessarily unique.

Definition 18.3: *Radius of convergence*

The *radius of convergence* of a power series f is the unique nonnegative real number R such that the domain of f contains $(-R,R)$ and is contained in $[-R,R]$.

Theorem 18.4: *Continuity of power series*

A power series is continuous on its domain.

Theorem 18.5: *Integration term by term*

If $f : x \longmapsto \sum_{n=0}^{\infty} a_n(x-c)^n$ for some $c \in \mathbb{R}$ and some sequence $(a_n)_{n\in\mathbb{N}}$, and if R is the radius of convergence of f, and if:

$$g : x \longmapsto \sum_{n=0}^{\infty} \frac{a_n}{n+1} x^{n+1}$$

then g has radius of convergence R and $g' = f$ on $(-R, R)$.

Theorem 18.6: *Derivation term by term*

If $f : x \longmapsto \sum_{n=0}^{\infty} a_n(x-c)^n$ for some $c \in \mathbb{R}$ and some sequence $(a_n)_{n\in\mathbb{N}}$, and if R is the radius of convergence of f, and if:

$$g : x \longmapsto \sum_{n=0}^{\infty} (n+1)a_{n+1}x^n$$

then g has radius of convergence R and $f' = g$ on $(-R, R)$.

Definition 18.7: *Taylor Series*

If f is a function of class C^∞ on some interval $[a, b]$ and if $c \in [a, b]$ then the *Taylor series* of f centered at c is:

$$x \longmapsto \sum_{n=0}^{\infty} \frac{f^{(n)}(c)}{n!}(x-c)^n.$$

Theorem 18.8

If f is a the sum of a power series, then it is equal to the sum of its Taylor series.

19. Controlling the remainder: Taylor-Lagrange Theorem

Definition 19.1

A function $f : [a, b] \to \mathbb{R}$ is of class C^n for some natural number n if it has n derivatives on (a, b), each of them having a (necessarily unique) continuous extension to $[a, b]$.
A function $f : [a, b] \to \mathbb{R}$ is of class C^∞ when it has infinitely many derivatives on $[a, b]$.

Theorem 19.2: *Taylor-Lagrange theorem*

If $f : [a,b] \to \mathbb{R}$ be a function of class C^{n+1}, then:

$$f(b) = \sum_{j=0}^{n} \frac{f^{(j)}(a)}{j!}(x-a)^j + \int_a^b \frac{f^{(n+1)}(t)}{n!}(b-t)^n \, dt.$$

Proof. We proceed by induction on the order n. For $n = 1$, the fundamental theorem of calculus proves that if $f : [a,b] \to \mathbb{R}$ is C^1 then:

$$f(b) - f(a) = \int_a^n f'(t) \, dt$$

which is our result for $n = 1$.

Assume now that our result holds for some $n \in \mathbb{N}$. Let $f : [a,b] \to \mathbb{R}$ be of class C^{n+1}. In particular, f is of class C^n so by our induction hypothesis, we obtain:

$$f(b) = \sum_{j=0}^{n} \frac{f^{(j)}(a)}{j!}(x-a)^j + \int_a^b \frac{f^{(n+1)}(t)}{n!}(b-t)^n \, dt. \tag{19.1}$$

Now, $f^{(n)}$ has a continuous derivative on $[a,b]$ then by integration by part:

$$\begin{aligned}
&\int_a^b \frac{f^{(n+1)}(t)}{n!}(b-t)^n \, dt \\
&\qquad = \left[\frac{f^{(n+1)}(t)}{n!} \cdot \frac{-(b-t)^{n+1}}{n+1} \right]_{t=a}^{t=b} - \int_a^b \frac{f^{(n+2)}(t)}{n!} \cdot \frac{-(b-t)^{n+1}}{n+1} \, dt \\
&\qquad = \frac{f^{(n+1)}(a)}{(n+1)!}(b-a)^{n+1} + \int_a^b \frac{f^{(n+2)}(t)}{(n+1)!}(b-t)^{n+1} \, dt.
\end{aligned}$$

Thus replacing the remainder in Equation (19.1) by the above expression, we have proven our theorem for $n+1$ and thus for all $n \in \mathbb{N}$ by induction. $\square$

Remark 19.3

Taylor-Lagrange Theorem provides a formula for the remainder of the difference between a function and its Taylor polynomial on an entire interval — it is a global result, unlike Taylor-Young Theorem. However it does so at the cost of a stronger assumption: the existence of one more continuous derivative than the order of the Taylor polynomial approximation.

20. Application to the convergence of certain power series

Theorem 20.1

Let f be of class C^{n+1} on some interval $[a,b]$. Since $\left|f^{(n+1)}\right|$ is continuous on a compact interval $[a,b]$, it has some maximum value M. We then have for all $x \in [a,b]$:

$$\left|f(x) - \sum_{j=0}^{n} \frac{f^{(j)}(a)}{j!}(x-a)^j\right| \leqslant \frac{M|b-a|^n}{n!}.$$

Proof. By Taylor-Lagrange Theorem, we have for all $x \in [a,b]$:

$$f(x) - \sum_{j=0}^{n} \frac{f^{(j)}(a)}{j!}(x-a)^j = \int_a^b \frac{f^{(n+1)}(t)}{n!}(b-t)^n\,dt$$

so

$$\left|f(x) - \sum_{j=0}^{n} \frac{f^{(j)}(a)}{j!}(x-a)^j\right| \leqslant \int_a^b \frac{\left|f^{(n+1)}(t)\right|}{n!}(b-t)^n\,dt \leqslant \frac{M|b-a|^n}{n!}$$

as claimed. □

Example 20.2: *Controlling the numerical error*

We want to compute $\sin(0.1)$ to a precision of 10^{-5}. Note that all derivatives of sin take values between -1 and 1, so by the Taylor-Lagrange theorem, we have for all $x \in \mathbb{R}$ and $n \in \mathbb{N}$:

$$\left|\sin(x) - \sum_{j=0}^{n} \frac{\sin^{(j)}(0)}{j!}x^j\right| \leqslant \frac{|x|^n}{n!}.$$

We choose $n \in \mathbb{N}$ so that $\frac{|0.1|^n}{n!} \leqslant 10^{-5}$. The smallest such n is 4. The Taylor polynomial of order 4 at 0 for sin is $x \mapsto x - \frac{x^3}{6}$. In particular:

$$10^{-5} \geqslant \left|\sin(0.1) - \left(0.1 - \frac{0.0001}{6}\right)\right| = |\sin(0.1) - 0.9983|$$

so our estimate for $\sin(0.1)$ is 0.9983 ± 10^{-5}.

As a motivation for things to come, we wish to make sense of the following statement.

Corollary 20.3: *L*

t $f : [a,b] \to \infty$ be of class C^∞. If there exists $M > 0$ such that for all $n \in \mathbb{N}$, the maximum value of $\left|f^{(n)}\right|$ is no more than M, then for all $x \in [a,b]$:

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

Proof. There exists $N \in \mathbb{N}$ such that $|b-a| \leqslant N$. Therefore if $n > N$:

$$\begin{aligned} 0 \leqslant \frac{M|b-a|^n}{n!} &= M \prod_{j=0}^{n} \frac{|b-a|}{j} \\ &\leqslant M \prod_{j=0}^{N} \frac{|b-a|}{j} \cdot \frac{|b-a|}{n}, \end{aligned}$$

where we stress that $M \prod_{j=0}^{N} \frac{|b-a|}{j}$ is some constant independent of n. By the squeeze theorem, since $\lim_{n\to\infty} \frac{|b-a|}{n} = 0$, we conclude our corollary. □

References

Email address: frederic@math.du.edu
URL: http://www.math.du.edu/~frederic

Department of Mathematics, University of Denver, Denver CO 80208