# Lecture notes, Analysis 2

Yossi Shamai

March 17, 2010

## 1 Introduction

These lecture notes are aimed for first year's students (**maths/non maths**) that complete the module MA131 Analysis II, in the university of Warwick. In this module, the basis theory of real-valued functions in one variable is studied, in which the underlying concept is the concept of limits, either of sequences or functions. The study of classes of functions, such as continuous or differentiable functions, which play an important role in analysis, relies in a better, more thorough understanding of limits. These lecture notes are complimented by a collection of home assignments, each of which gathers standard problems for every subject. Students are required to submit these assignments, and study the lecture material. For more information, please refer to my homepage.

### 1.1 Some notations

We will commonly use throughout this module the following notations:

| Notation | Meaning |
|---|---|
| $\mathbb{N}$ | Naturals $\{1, 2, \dots\}$ |
| $\mathbb{Z}$ | Integers $\{0, \pm 1, \pm 2, \dots\}$ |
| $\mathbb{Q}$ | Rationals $\left\{\frac{n}{m} : n, m \in \mathbb{Z}, m \neq 0\right\}$ |
| $\mathbb{R}$ | The real line |

In particular, we assume that these sets of numbers are defined a priori. We do not, for example, state the axioms of the ordered field, nor the axiom of a complete field, that are necessary for a rigorous mathematical definition of the set of real numbers. Subsets of $\mathbb{R}$ are often denoted by $A, B$, and points in $\mathbb{R}$ are denoted by $x, y, z$. When we want to emphasise that a point/number is fixed, we will add the subscript zero to it, e.g, $x_0$ or $\varepsilon_0$. Intervals in $\mathbb{R}$ are denoted by

| | |
|---|---|
| closed | $[a, b] = \{x \in \mathbb{R} : a \leq x \leq b\}$ |
| open | $(a, b) = \{x \in \mathbb{R} : a < x < b\}$ |
| partially open | $(a, b] = \{x \in \mathbb{R} : a < x \leq b\}$ |
| open and unbounded | $(a, \infty) = \{x \in \mathbb{R} : a < x < \infty\}$ |
| closed and unbounded | $(-\infty, b] = \{x \in \mathbb{R} : -\infty < x \leq b\}$ |
| $\vdots$ | |

and so on. Functions are often denoted by the letters $f, g$ or $\phi, \psi$, and their domains by $E, G \subseteq \mathbb{R}$. If not stated otherwise, the domain of a function is always assumed to be an interval. Sequences of numbers are denoted by $(a_n)_{n=1}^{\infty}$ or $(x_n)_{n=17}^{\infty}$. Whenever the indices are missing, you may assume that the index starts from one; e.g, $(a_n) = (a_n)_{n=1}^{\infty}$. In particular, the characters $n, m$ and $k, l$ always denote natural numbers, unless stated otherwise.

# 2 Continuity of functions of a real variable

The notion of a continuous curve is very intuitive. A curve is continuous if one can sketch its figure without ever 'disconnecting' his pen from the paper. Likewise, a real-valued function is continuous if its graph is a continuous curve (see Figure 1). Unfortunately, 'never disconnect your pen' is not a mathematical statement. To give a proper mathematical definition for continuity, it is often convenient to first define it for a given, fixed point on the real line, $x_0$, and then extend the definition for arbitrary subsets of the real line (such as intervals). We begin with a definition for a real-valued function,

***Definition 2.1*** *Let $E$ be a given subset of $\mathbb{R}$. A* ***real-valued function*** *$f : E \to \mathbb{R}$ is a 'rule', that assigns each element $x \in E$ a (unique) value $y \in \mathbb{R}$, denoted by $y = f(x)$, and called '**the image of** $x$'. The set $E$ is called the* ***Domain*** *of the function $f$.*

Note that the word 'rule' is used with commas. The reason is that the above definition is not completely rigorous. To give a proper definition for

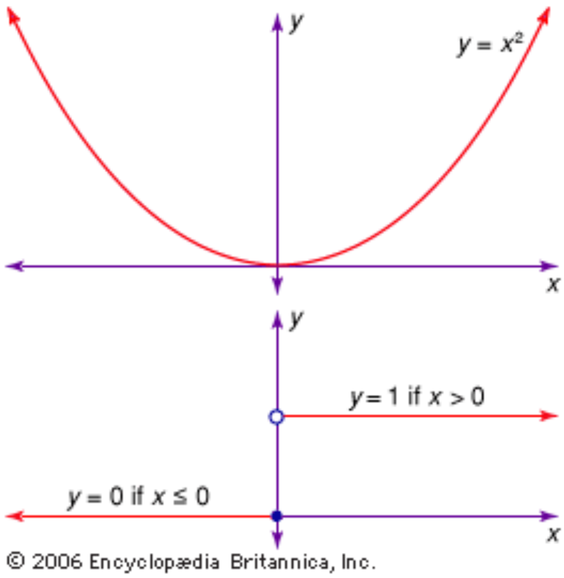



Figure 1: Graphs of continuous and discontinuous functions

a 'function' one should first specify carefully the axioms of set theory, and that is out of the scope of this module. Instead, think of a function as a 'black box' which receives an 'input' $x \in E$ and produces a single 'output' $y = f(x)$. Note also that neither the 'input' $x$ nor the 'output' $y$ do not have to be real numbers, but we assume it as it is the only relevant case for this talk.

**Definition 2.2** *The **image** of a real-valued function $f : E \to \mathbb{R}$ is defined as the set*

$$\mathrm{Im}(f) = \{f(x) : x \in E\} \subseteq \mathbb{R}.$$

**Definition 2.3** *The **Graph** of a real-valued function $f : E \to \mathbb{R}$ is defined as the set*

$$\mathrm{gr}(f) = \{(x, f(x)) : x \in E\} \subseteq \mathbb{R}^2.$$

In this module, the domain $E$ is always a subset of $\mathbb{R}$, in which case the graph of $f$ is a subset of the 2-dimensional plain $\mathbb{R}^2$ (see Figure 6). The domain will often be an interval; either open $E = (a, b)$ , closed $E = [a, b]$, partially closed (e.g $E = [a, b))$ or unbounded (e.g, $E = [a, \infty))$. It may also be a (finite) union of such intervals. When the domain of $f$ is not specified, we always assume that $E$ is the maximal subset of $\mathbb{R}$ for which $f$ may be defined. For example, if not specified otherwise, then the domain of the function $f(x) = \frac{1}{x}$ is the set $E = \mathbb{R} \setminus \{x_0\} = (-\infty, 0) \cup (0, \infty)$.

***Definition 2.4*** *A real-valued function $f : E \to \mathbb{R}$ is said to be* ***continuous*** *at a point $x_0 \in E$, if for every $\varepsilon > 0$ there exists a $\delta > 0$, such that for every $x \in E$,*

$$|x - x_0| < \delta \qquad \textit{implies} \qquad |f(x) - f(x_0)| < \varepsilon. \tag{2.1}$$

It is often convenient to write the definition of continuity with the use of the quantifiers $\forall$ (**for all**) and $\exists$ (**exists**),

$$\forall \varepsilon > 0 \; \exists \delta > 0 : \forall x \in E, \; |x - x_0| < \delta \Rightarrow |f(x) - f(x_0)| < \varepsilon.$$

We start with some basic and important observations:

(a) If $\varepsilon_1 < \varepsilon_2$ then $|f(x) - f(x_0)| < \varepsilon_1$ implies $|f(x) - f(x_0)| < \varepsilon_2$. Since (2.1) is required to hold for every $\varepsilon > 0$ (with $\delta$ depending on $\varepsilon$), then 'for every' means 'as small as you want'.

(b) The parameter $\delta$ depends in general on $\varepsilon$, but also on the point $x_0$ and on the function $f$.

(c) If $\delta_1 < \delta_2$ then $|x - x_0| < \delta_1$ implies $|x - x_0| < \delta_2$. Hence, the choice of $\delta$ is not unique (we can always replace $\delta$ by a smaller positive value), and can be also considered as a 'small parameter'. Be careful of writing $\delta = \delta(\varepsilon)$ as $\delta$ is not uniquely defined by $\varepsilon$ (and therefore is not a function).

As $|x - y|$ is simply the 'distance' from $x$ to $y$, the above observations suggest to read the definition of continuity as follows: for every $\varepsilon > 0$ (as small as you want) there exists a (sufficiently small) $\delta > 0$, such that if $x$ is sufficiently close to $x_0$ (i.e, $|x - x_0| < \delta$), then $f(x)$ is sufficiently close to $f(x_0)$ (i.e, $|f(x) - f(x_0)| < \varepsilon$).
The definition of continuity is extended to arbitrary subsets of $E$ as follows,

***Definition 2.5*** *A function $f : E \to \mathbb{R}$ is* ***continuous*** *in a subset $A \subseteq E$, if it is continuous (see Definition 2.4) at every $x_0 \in A$.*

Note that if $x_0 \notin E$, then the statement '$f$ is continuous at $x_0$' is meaningless; for example, the function $f : [0,1] \to \mathbb{R}$ defined by $f(x) = x^2$ cannot be continuous at any point $x_0 \notin [0,1]$ (it is not defined there!). Therefore, we can only give meaning for a continuity property in subsets of $E$.

***Definition 2.6*** *A function $f : E \to \mathbb{R}$ is said to be* ***continuous****, if it is continuous (see Definition 2.4) at every point $x_0 \in E$*[1]

We proceed with some examples below,

***Example 2.1*** *The function $f(x) = 2x$ is continuous at every point $x_0 \in \mathbb{R}$.*

*Proof*: Let $\varepsilon > 0$ and set $\delta = \varepsilon/2$. Let $x$ be a point for which $|x - x_0| < \delta$. Then

$$|f(x) - f(x_0)| = |2x - 2x_0| = 2|x - x_0| < 2\delta = \varepsilon.$$

Since the choice of $\varepsilon$ was arbitrary, the above holds for every $\varepsilon > 0$, as required. ■

Remember that one may replace $\delta = \varepsilon/2$ by any smaller value of $\delta$; indeed, the above proof remains valid, if for example $\delta = \varepsilon/10$.

***Example 2.2*** *The function $f(x) = x^2$ is continuous at every point $x_0 \in \mathbb{R}$.*

*Proof*: Given $\varepsilon > 0$ we set $\delta = \min\left(1, \frac{\varepsilon}{1+2|x_0|}\right)$. Let $x$ be a point for which $|x - x_0| < \delta$. Then

$$\begin{aligned} |f(x) - f(x_0)| &= |x^2 - x_0^2| = |x - x_0||x + x_0| \\ &\le |x - x_0|\left(|x| + |x_0|\right) < \delta \cdot \left(|x| + |x_0|\right), \end{aligned} \tag{2.2}$$

where we have used the Triangle inequality. Note also that as $\delta \le 1$ then by the Triangle inequality,

$$|x| = |x - x_0 + x_0| \le |x - x_0| + |x_0| < 1 + |x_0|.$$

Substituting the above inequality into the right-hand side of (2.2) yields

$$|f(x) - f(x_0)| < \delta \cdot (1 + 2|x_0|) \le \varepsilon.$$

[1] The function $f(x) = 1/x$ is continuous as it is continuous at its domain, $E = \mathbb{R} \setminus \{0\}$. Is $f$ continuous at $x_0 = 0$? of course not. It isn't defined there! nevertheless we use the terminology that $f$ is continuous to refer to the domain of $f$ and not to $\mathbb{R}$.

∎

Note that, as in the above example, the choice for a correct $\delta$ cannot in general be 'guessed' a priori; it is derived from the algebraic manipulations of upper bounding $|f(x) - f(x_0)|$.

Next, we give an example of a function which is discontinuous. To make it more convenient, we invert Definition 2.4,

*A function $f : E \to \mathbb{R}$ is* ***discontinuous*** *at $x_0 \in E$, if there exists an $\varepsilon_0 > 0$ such that for every $\delta > 0$, there exists a point $x \in E$ such that $|x - x_0| < \delta$ but*

$$|f(x) - f(x_0)| \geq \varepsilon_0.$$

Rewriting this definition with quantifiers reads

$$\exists \varepsilon_0 > 0 \ : \ \forall \delta > 0 \ \exists x \in \mathbb{E} \ : \ |x - x_0| < \delta \text{ but } |f(x) - f(x_0)| \geq \varepsilon_0.$$

Note that the word 'but' is only used for convenience, and formally should be replaced by the word 'and'.

***Example 2.3*** *Let $f : \mathbb{R} \to \mathbb{R}$ be defined by*

$$f(x) = \begin{cases} \frac{x}{|x|} & x \neq 0 \\ 0 & x = 0 \end{cases}$$

*The 'gap' at $x_0 = 0$ (see Figure (2)) suggests that $f$ is discontinuous at $x_0 = 0$. To show that, we set $\varepsilon_0 = 1$. Let $\delta > 0$, and set $x = \delta/2$. Then $|x - 0| = \delta/2 < \delta$ and*

$$|f(x) - f(0)| = +1 - 0 \geq \varepsilon_0.$$

*Thus, $f$ is discontinuous at $x_0 = 0$. It is easy to verify that $f$ is continuous at any point $x \neq 0$ (but lets prove it anyway in the next proposition)*

***Proposition 2.1*** *Let $f : \mathbb{R} \to \mathbb{R}$ be a real-valued function. Suppose that there exists an $r > 0$, such that $f$ is constant on the interval $(x_0 - r, x_0 + r)$. Then $f$ is continuous at $x_0$.*

*Proof*: Given $\varepsilon > 0$, set $\delta = r$. Then for every $x \in \mathbb{R}$ such that

$$|x - x_0| < \delta = r.$$

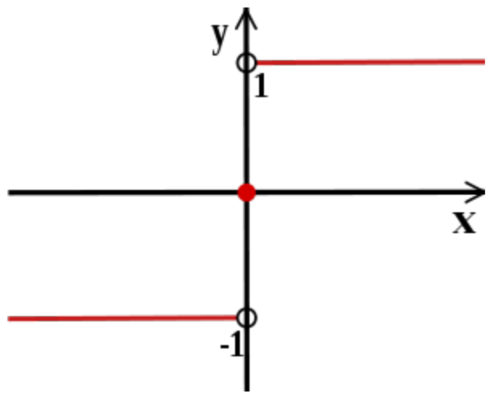



Figure 2: The function $f(x)$ is often referred to as the 'sign' function; it is equal to $+1$ for any positive number, $-1$ for any negative number, and arbitrarily set to 0 for $x = 0$.

Then $x \in (x_0 - r, x_0 + r)$, hence

$$|f(x) - f(x_0)| = 0 < \varepsilon.$$

■

We proceed with the fundamental example of a **nowhere continuous** function,

***Example 2.4*** *(the* Dirichlet *function) Let $D : \mathbb{R} \to \mathbb{R}$ be defined by*

$$D(x) = \begin{cases} 1 & x \in \mathbb{Q} \\ 0 & x \notin \mathbb{Q} \end{cases} \tag{2.3}$$

***Proposition 2.2*** *The Dirichlet function is discontinuous at every point $x_0 \in \mathbb{R}$.*

*Proof*: Let $x_0 \in \mathbb{R}$ and assume first that $x_0 \notin \mathbb{Q}$. Set $\varepsilon_0 = 1$ and let $\delta > 0$. Recall that the set of rational numbers is dense in $\mathbb{R}$, hence there exists a rational $x \in \mathbb{Q}$ such that $|x - x_0| < \delta$, and by the definition of $D(x)$ and $D(x_0)$,

$$|D(x) - D(x_0)| = |1 - 0| \geq \varepsilon_0.$$

Thus, we have shown that $D(x)$ is discontinuous at any irrational point $x_0 \notin \mathbb{Q}$. Now assume that $x_0 \in \mathbb{Q}$. Set again $\varepsilon_0 = 1$ and let $\delta > 0$. Recall that the set of irrational numbers is also dense in $\mathbb{R}$, hence there exists an irrational $x \notin \mathbb{Q}$ such that $|x - x_0| < \delta$, and by the definition of $D(x)$ and $D(x_0)$,

$$|D(x) - D(x_0)| = |0 - 1| \geq \varepsilon_0.$$

■

## 2.1 Sequential continuity

An alternative approach to define continuity of a function at $x_0$ is via the use of *sequences* (remember sequences?).

***Definition 2.7*** *A function $f : E \to \mathbb{R}$ is said to be **sequentially continuous** at $x_0 \in E$, if for any sequence $(x_n)_{n=1}^{\infty}$ in $E$,*

$$\lim_{n\to\infty} x_n = x_0 \qquad \textit{implies} \qquad \lim_{n\to\infty} f(x_n) = f(x_0).$$

Lets recall the definition for a converging sequence,

***Definition 2.8*** *A sequence $(a_n)_{n=1}^{\infty}$ in $\mathbb{R}$ **converges** to a limit $L \in \mathbb{R}$, iff*

$$\forall \varepsilon > 0 \; \exists N \in \mathbb{N} \; : \; n \in \mathbb{N}, n > N \Rightarrow |a_n - L| < \varepsilon. \tag{2.4}$$

The 'secret' behind the definition of sequential continuity it uses a notion of limits which is already defined, hence avoids the formal formulation using the $\varepsilon - \delta$ language. The main assertion on sequential continuity is the following:

***Theorem 2.1*** *A real-valued function $f : E \to \mathbb{R}$ is continuous at $x_0$ if and only if it is sequentially continuous at $x_0$.*

In other words, continuity and sequential continuity are exactly the same. Then why do they have different names? well, the right way of showing two things are the same is first giving them different names like $a, b$ and then proving that $a = b$.

*Proof*: We first assume that $f$ is continuous at $x_0$ and prove that it is sequentially continuous. According to Definition 2.1,

$$\forall \varepsilon > 0 \; \exists \delta > 0 \; : \; |x_n - x_0| < \delta \Rightarrow |f(x_n) - f(x_0)| < \varepsilon.$$

Let $(x_n)$ be a sequence in $E$ which converges to $x_0$. According to Definition 2.4 with $\delta > 0$ replacing the role of $\varepsilon$,

$$\exists N \in \mathbb{N} \ : n \in \mathbb{N}, n > N \Rightarrow |x_n - x_0| < \delta.$$

Putting these two conditions together, then

$$\forall \varepsilon > 0 \ \exists N \in N \ : \ n > N \Rightarrow |f(x_n) - f(x_0)| < \varepsilon,$$

which reads $\lim_{n\to\infty} f(x_n) = f(x_0)$.
Now to prove the opposite direction; assume that $f$ is sequentially continuous at $x_0$, and suppose, by contradiction, that $f$ is discontinuous at $x_0$. To arrive at the contradiction, we show that there exists a sequence $(x_n)_{n=1}^{\infty}$ such that $\lim_{n\to\infty} x_n = x_0$ but $f(x_n)$ does not converge to $f(x_0)$: as $f$ is discontinuous at $x_0$, then

$$\exists \varepsilon_0 > 0 \ \forall \delta > 0 \ \exists x \in E \ : \ |x - x_0| < \delta \text{ but } |f(x) - f(x_0)| \geq \varepsilon_0. \tag{2.5}$$

Consider $\delta_n = \frac{1}{n}$. Then, according to the above, there exists a point $x_n \in E$ such that

$$|x_n - x_0| < \frac{1}{n} \text{ but } |f(x_n) - f(x_0)| \geq \varepsilon_0.$$

Consider the sequence $(x_n)_{n=1}^{\infty}$. As $|x_n - x_0| < \frac{1}{n}$ for every $n$, then $\lim_{n\to\infty} x_n = x_0$. As $|f(x_n) - f(x_0)| \geq \varepsilon_0$ for every $n$, then the sequence $(f(x_n))$ cannot converge to $f(x_0)$ (and may not even converge at all). ■
Henceforth we will use the two equivalent definitions for continuity (2.4) and (2.7) interchangeably.
Sequential continuity may be extremely useful for proving that a given function is discontinuous at $x_0$; indeed, it is sufficient to find two sequences $(x_n), (y_n)$ which converge to $x_0$, such that the following limits exist and are not equal,

$$\lim_{n\to\infty} f(x_n) \neq \lim_{n\to\infty} f(y_n). \tag{2.6}$$

***Example 2.5*** *The Dirichlet function is discontinuous at every point $x_0 \in \mathbb{R}$.*

Proof*: Let $x_0$ be given. As the sets of rational and irrational numbers are both dense in $\mathbb{R}$, there exist sequences $(x_n)$ in $\mathbb{Q}$ and $(y_n)$ in $\mathbb{R} \setminus \mathbb{Q}$ which converge to $x_0$. However,*

$$\lim_{n\to\infty} D(x_n) = 1 \neq 0 = \lim_{n\to\infty} D(y_n).$$

*Thus $D(x)$ is discontinuous at $x_0$.* ■

***Example 2.6*** *Let $a \in \mathbb{R}$ and set*

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

*Then $f$ is discontinuous at $x_0 = 0$, for every value of $a$.*

*Proof*: As explained above, it is sufficient to find two sequences $(x_n)$ and $(y_n)$ which converge to zero, such that (2.6) holds; set

$$x_n = \frac{1}{2\pi n + \frac{\pi}{2}} \qquad \text{and} \qquad y_n = \frac{1}{2\pi n + \frac{3\pi}{2}},$$

and recall that $\sin\left(2\pi n + \frac{\pi}{2}\right) = \sin\left(\frac{\pi}{2}\right) = 1$ and that $\sin\left(2\pi n + \frac{3\pi}{2}\right) = \sin\left(\frac{3\pi}{2}\right) = -1$, hence

$$f(x_n) = 1 \to 1 \qquad \text{and} \qquad f(y_n) = -1 \to -1,$$

as required. ∎

***Example 2.7*** *Let $f : \mathbb{R} \to \mathbb{R}$ be the real-valued function given by*

$$f(x) = \begin{cases} x & x \in \mathbb{Q} \\ -x & x \notin \mathbb{Q} \end{cases}$$

*Then $f$ is continuous at $x_0 = 0$ and discontinuous at any other point $x_0 \neq 0$.*

*Proof*: We start by proving that $f$ is continuous at $x_0 = 0$; let $\varepsilon > 0$, and set $\delta = \varepsilon$. As $|f(x)| = |x|$ for every $x$, then for every $|x| < \delta$, we have

$$|f(x) - f(0)| = |x| < \delta = \varepsilon,$$

i.e, $f$ is continuous at $x_0 = 0$.
Now, to show that $f$ is discontinuous for every $x_0 \neq 0$, let $(x_n)$ and $(y_n)$ be two sequences of rational and irrational numbers which converge to $x_0$. The existence of $(x_n)$ and $(y_n)$ is guaranteed by the fact that both $\mathbb{Q}$ and $\mathbb{R} \setminus \mathbb{Q}$ are dense subsets in $\mathbb{R}$. As $x_0 \neq 0$,

$$\lim_{n\to\infty} f(x_n) = x_n \to x_0 \neq -x_0 = \lim_{n\to\infty} f(y_n),$$

which implies that $f$ is discontinuous at $x_0$. ∎
Another useful and easy-to-prove application of the alternative definition for continuity using sequences, is the algebraic properties for continuous functions, which is treated in the next section.

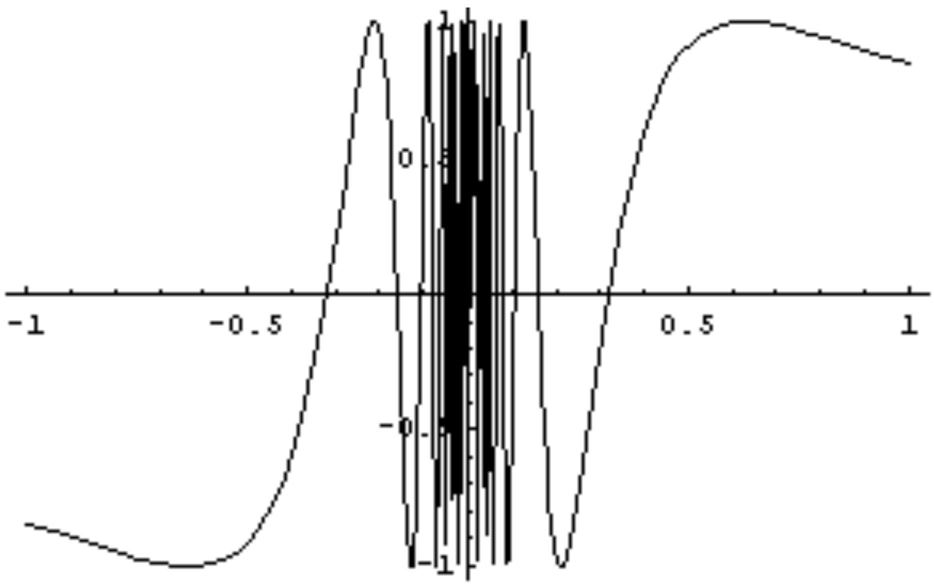



Figure 3: The function $f(x) = \sin \frac{1}{x}$. Discontinuous at zero, no matter how $f(0)$ is defined.

## 2.2 Algebra of continuous functions

In this section we show that basic algebraic operations (such as sums and products) of continuous functions are also continuous. In 'fancy words', we show that the class of continuous functions is 'closed' under basic algebraic operations. We start by recalling similar algebraic properties of sequences,

***Theorem 2.2*** *(Algebra of sequences) Let $(a_n)_{n=1}^{\infty}$ and $(b_n)_{n=1}^{\infty}$ be two converging sequences, and denote by*

$$a = \lim_{n\to\infty} a_n, \qquad \text{and} \qquad b = \lim_{n\to\infty} b_n.$$

*Then,*

1. *The sequence $A_n = a_n + b_n$ converges and $\lim_{n\to\infty} A_n = a + b$.*
2. *The sequence $B_n = a_n \cdot b_n$ converges and $\lim_{n\to\infty} B_n = a \cdot b$.*
3. *If $b_n \neq 0$ for every $n$, and $b \neq 0$, then the sequence $C_n = a_n/b_n$ converges and $\lim_{n\to\infty} C_n = a/b$.*

*Proof*: See Term I. ∎

Applying the above theorem, we obtain a similar argument for continuous functions,

***Theorem 2.3*** *Let $f : E \to \mathbb{R}$ and $g : E \to \mathbb{R}$* [2] *be two real-valued functions that are continuous at $x_0 \in E$. Then the functions (i) $f + g$ and (ii) $f \cdot g$ are continuous at $x_0$. If, in addition, $g(x) \neq 0$ for every $x \in E$, then the function (iii) $f/g$ is continuous at $x_0$.*

*Proof*: Given a sequence $(x_n)$ in $E$ which converges to $x_0$, then by the continuity of $f$ and $g$ we have

$$\lim_{n\to\infty} f(x_n) = f(x_0) \qquad \text{and} \qquad \lim_{n\to\infty} g(x_n) = g(x_0),$$

and consequently, by Theorem 2.2 we have

(i)

$$\begin{aligned}\lim_{n\to\infty}(f+g)(x_n) &= \lim_{n\to\infty}(f(x_n)+g(x_n)) = \lim_{n\to\infty} f(x_n) + \lim_{n\to\infty} g(x_n)\\ &= f(x_0)+g(x_0) = (f+g)(x_0).\end{aligned}$$

(ii)

$$\begin{aligned}\lim_{n\to\infty}(f\cdot g)(x_n) &= \lim_{n\to\infty}(f(x_n)\cdot g(x_n)) = \lim_{n\to\infty} f(x_n) \cdot \lim_{n\to\infty} g(x_n)\\ &= f(x_0)\cdot g(x_0) = (f\cdot g)(x_0),\end{aligned}$$

and

(iii)

$$\begin{aligned}\lim_{n\to\infty}(f/g)(x_n) &= \lim_{n\to\infty}(f(x_n)/g(x_n)) = \lim_{n\to\infty} f(x_n) / \lim_{n\to\infty} g(x_n)\\ &= f(x_0)/g(x_0) = (f/g)(x_0),\end{aligned}$$

∎

***Example 2.8*** *The function $h(x) = x^2$ is continuous, as it is simply the multiplication of the identity function $f(x) = x$ by itself* [3]

[2] Note that $f$ and $g$ must have the same domain

[3] the identity function is clearly continuous.

***Example 2.9*** *The functions* $f(x) = \frac{x}{1+x^2}$ *and* $g(x) = 1+x+x^2$ *are continuous. This follows directly from the above theorem by noting* $f, g$ *are derived by elementary algebraic operations of the constant function* 1 *and the identity, which are both continuous.*

***Example 2.10*** *If* $f$ *is continuous at* $x_0$ *then* $-f = (-1) \cdot f$ *is continuous at* $x_0$*, as any constant function is continuous.*

Note that the assumption that $g(x)$ does not vanish for every $x \in E$ is necessary for the quotient function $f(x)/g(x)$ to be defined in $E$. However, it is sufficient to only require that $g$ does not vanish at $x_0$, as continuous functions have the property that if the do not vanish at some point, they have to remain different than zero on an open interval containing that point. To show that, we first state our assertion more accurately,

***Lemma 2.1*** *Let* $f : E \to \mathbb{R}$ *be a real-valued functions that is continuous at a point* $x_0 \in E$*. Suppose that* $f(x_0) > 0$*. Then, there exists a* $\delta > 0$ *for which*

$$f(x) > 0, \qquad \forall x \in (x_0 - \delta, x_0 + \delta).$$

*Proof*: Let $\varepsilon = f(x_0)/2 > 0$. As $f$ is continuous at $x_0$, there exists a $\delta > 0$ such that

$$|f(x) - f(x_0)| < f(x_0)/2 \qquad , \forall x \in (x_0 - \delta, x_0 + \delta).$$

i.e,

$$-f(x_0)/2 < f(x) - f(x_0) < f(x_0)/2.$$

Adding $f(x_0)$ to both sides of the left-hand side of the above inequality gives

$$f(x) > f(x_0) - f(x_0)/2 = f(x_0)/2 > 0,$$

that is, $f(x) > 0$ for every $x \in (x_0 - \delta, x_0 + \delta)$.

■

An immediate result is the opposite assertion,

***Corollary 2.1*** *Let* $f : E \to \mathbb{R}$ *be a real-valued functions that is continuous at* $x_0 \in E$*. Suppose that* $f(x_0) < 0$*. Then, there exists a* $\delta > 0$ *for which*

$$f(x) < 0 \qquad , \forall x \in (x_0 - \delta, x_0 + \delta).$$

*Proof*: Apply Lemma 2.1 to the function $g(x) = -f(x)$, which is continuous and satisfies $g(x_0) > 0$. ■

Now, suppose that $f, g : E \to \mathbb{R}$ are continuous functions. If $g(x_0) \neq 0$, then, by Lemma 2.1 and Corollary 2.1, there exists a $\delta > 0$ such that $g(x) \neq 0$ for every $x \in (x_0 - \delta, x_0 + \delta)$. Consequently, $f/g$ is defined on this interval. Applying Theorem 2.3 to $f, g$ and this interval, we obtain that $f/g$ is continuous at $x_0$.

***Example 2.11*** *Every polynomial $p(x) = \sum_{n=0}^{N} x^n$ is continuous in $\mathbb{R}$; it is equal to a finite number of sums and multiplications of the identity function $f(x) = x$ and the constant function $g(x) = 1$, which are trivially continuous (take $\delta = \varepsilon$) in Definition (2.4)). Similarly, every rational function $Q(x) = \frac{p(x)}{q(x)}$ where $p(x), q(x)$ are polynomials is continuous at every point $x_0 \in \mathbb{R}$, provided that $q(x_0) \neq 0$.*

Another algebraic operation which preserves the property of continuity is composition,

***Theorem 2.4*** *Let $f : E \to \mathbb{R}$ be a continuous function at $x_0 \in E$, and let $g : G \to \mathbb{R}$. Assume that $\operatorname{Im} f \subseteq G$ and that $g$ is continuous at $f(x_0) \in G$. Then the function $g \circ f$, defined by*

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

*is continuous at $x_0$.*

*Proof*: Let $(x_n)_{n=1}^{\infty}$ be a sequence in $E$ which converges to $x_0$. As $f$ is continuous at $x_0$ then

$$\lim_{n\to\infty} f(x_n) = f(x_0).$$

Now, $(f(x_n))_{n=1}^{\infty}$ is a sequence contained in $G$ which converges to $f(x_0)$. As $g$ is continuous at $f(x_0)$, then

$$\lim_{n\to\infty} g\left(f(x_n)\right) = g\left(f(x_0)\right),$$

i.e, $g \circ f(x_n)$ converges to $g \circ f(x_0)$. Thus, $g \circ f$ is continuous at $x_0$. ■

***Example 2.12*** *Lets 'believe' for now that the function $x \mapsto \sin x$ is continuous. Then the function $f(x) = \sin(1 + x^2)$ is also continuous, as it is the composition of two continuous functions. Imagine trying to prove this directly from the definition!*

## 2.3 The Intermediate Value Theorem

The intermediate value theorem has various versions, all of which are equivalent. In principle, the theorem states that that a continuous function attains all intermediate values that are in between any two of its values (see Figure 4).

**Theorem 2.5** *(The intermediate value theorem version I) Let $f : [a, b] \to \mathbb{R}$ be a continuous function. Suppose that $f(a) \le 0 \le f(b)$. Then there exists a point $c \in [a, b]$ such that $f(c) = 0$.*

*Proof*: Consider the set

$$S = \{x \in [a, b] : f(x) \le 0\}.$$

Note that $S$ is non-empty as $a \in S$. As $S \subset [a, b]$ it is bounded from above by $b$. It follows that it has an upper bound $c = \sup S$ and $a \le c \le b$. We now show that $f(c) = 0$. Suppose, by contradiction, that $f(c) \ne 0$.
If $f(c) < 0$, then by the assumption that $f(b) \ge 0$ we have $a \le c < b$. Applying Corollary 2.1, there exists a $\delta > 0$ for which $f(x) < 0$ for every $x \in [a, b]$ such that $c - \delta < x < c + \delta$. In particular, there exists a point $c < x < b$ such that $f(x) < 0$. This contradicts the property that $c$ is an upper bound for $S$.
if $f(c) > 0$, then by the assumption that $f(a) \le 0$ we have $a < c \le b$. applying Lemma 2.1, there exists a $\delta > 0$ for which $f(x) > 0$ for every $x \in (c - \delta, c + \delta)$. In particular, for every $x \in S$ we have $x < c - \delta$. This contradicts the property that $c$ a minimal upper bound. ■

**Corollary 2.2** *(The intermediate value theorem version II) Let $f : [a, b] \to \mathbb{R}$ be a continuous function. Suppose that $f(a) < f(b)$. Then for every $f(a) \le v \le f(b)$ there exists a point $a \le c \le b$ for which $f(c) = v$.*

*Proof*: Apply Theorem 2.5 for $g(x) = f(x) - v$. Note that $g$ is continuous and satisfies $g(a) \le 0 \le g(b)$, hence there exists a point $c \in [a, b]$ for which

$$g(c) = 0 \Rightarrow f(c) = v.$$

■

We proceed with some applications of the intermediate value theorem,

**Theorem 2.6** *(the fixed-point theorem) Let $f : [a, b] \to [a, b]$ be a continuous function [4] Then, there exists a point $c \in [a, b]$ such that $f(c) = c$.*

[4] i.e, $f : [a, b] \to \mathbb{R}$ and $\operatorname{Im} f \subseteq [a, b]$,

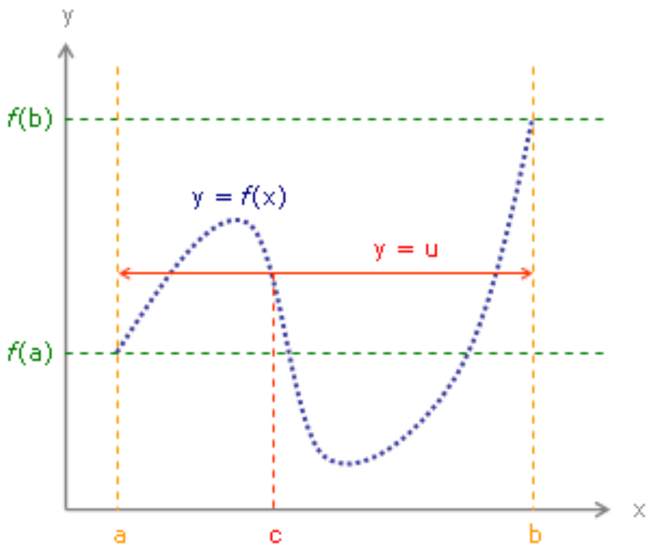



Figure 4: The IVT

*Proof*: Consider the function $g(x) = f(x) - x$. By Theorem 2.3, $g$ is continuous in $[a, b]$. Note that $g(a) = f(a) - a \geq 0$ and $g(b) = f(b) - b \leq 0$. By Theorem 2.5, there exists a point $c \in [a, b]$ for which $g(c) = 0$, i.e, $f(c) = c$. ■

In simple words, the fixed-point theorem states that if you sketch a continuous curve in a square, starting from one side and ending on the other, you must intersect the diagonal at least once (see Figure 5).

The intermediate-value theorem is also useful for showing that some algebraic equations have solutions.

***Example 2.13*** *There exists a solution for the equation* $\sin x + x = 1$.

*Proof*: Consider the function $f(x) = \sin x + x - 1$. Note that $f(0) = -1$ and that $f(\pi) = \sin \pi + \pi - 1 = \pi - 1 > 0$. As $f$ is continuous, then by Theorem 2.2 there exists a point $x \in [0, \pi]$ (and more precisely, $x \in (0, \pi)$) for which $f(x) = 0$, i.e, $\sin x + x = 1$. ■

Remark: note that the solution might not be unique. Indeed, the IVT (in all versions) does not guarantee that $c \in [a, b]$ is unique. It may be that $f(c) = v$ is satisfied for more than one point (see Figure 4)).

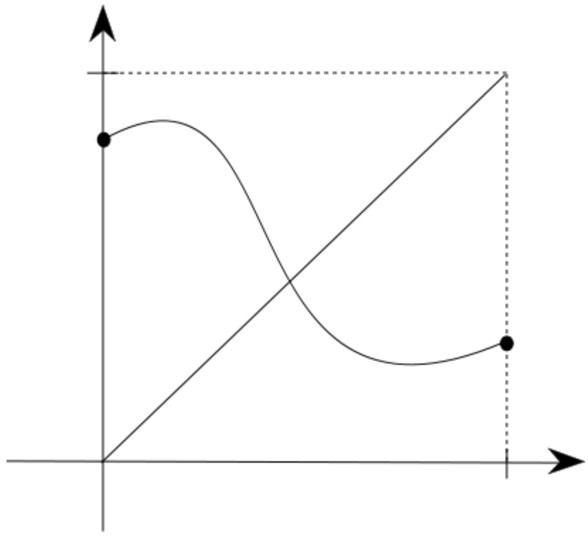


Figure 5: The fixed-point theorem. The graph of $f$ intersects the line $y = x$ at some intermediate point $c$

***Example 2.14*** *The equation* $e^x - x + \frac{1}{x} = 0$ *has a solution.*

*Proof*: Set $f(x) = e^x - x + \frac{1}{x}$. The function $f$ is continuous, by the algebra of continuous functions (and assuming, at this point, that $e^x$ is continuous). Also, $f(-\frac{1}{10}) = e^{-1/10} + \frac{1}{10} - 10 < 1.1 - 10 < 0$ and $f(-10) = e^{-10} + 10 - \frac{1}{10} > 9.9 > 0$. Thus, there exists a point $c \in [-10, -\frac{1}{10}]$ for which $f(c) = 0$, i.e, $e^c - c + \frac{1}{c} = 0$. ■

We remark that the proof will not remain valid if, for example, we consider the interval $[-\frac{1}{10}, 1]$ (in which $f$ changes signature), as $f$ is not continuous on this interval (it is undefined at zero!).

## 2.4 The Extreme Value Theorem (or the Max-Min Theorem)

The Bolzano-Weierstrass theorem (remember term I?) states that any sequence $(x_n)_{n=1}^{\infty}$ that is bounded has a converging subsequence. This fundamental theorem, together with the notion of sequential continuity, can lead to various results on continuous functions which are defined in closed intervals. Here the assumption that the interval is closed is essential. But why closed intervals and not, for example, partially closed or open? we will try to answer this question by the end of this section.

***Definition 2.9***

1. *A real-valued function $f : E \to \mathbb{R}$ is said to be* ***bounded from above****, if there exists a constant $M \in \mathbb{R}$ such that $f(x) \le M$ for every $x \in E$.*

2. *A real-valued function $f : E \to \mathbb{R}$ is said to be* ***bounded from below****, if there exists a constant $m \in \mathbb{R}$ such that $f(x) \ge m$ for every $x \in E$.*

3. *A real-valued function $f : E \to \mathbb{R}$ is said to be* ***bounded****, if there exists a constant $K \ge 0$ such that $|f(x)| \le K$.*

4. *A point $x_1 \in E$ is said to be* ***minimal****, if $f$ attains a minimum*[5] *at $x_1$, namely,*
$$f(x_1) \le f(x), \qquad \forall x \in E.$$

5. *A point $x_2 \in E$ is said to be* ***maximal****, if $f$ attains a maximum at $x_2$, namely,*
$$f(x) \le f(x_2), \qquad \forall x \in E.$$

Note that a bounded function is bounded from above and from below, as

$$|f(x)| \le K \Rightarrow -K \le f(x) \le K, \qquad \forall x \in E.$$

Vice versa, if $f$ is bounded from above and from below then $f$ is bounded, as

$$m \le f(x) \le M \Rightarrow |f(x)| \le K,$$

where $K = \max(|m|, |M|)$.

***Theorem 2.7*** *(The Max-Min Theorem) Let $f : [a, b] \to \mathbb{R}$ be a continuous real-valued function. Then*

1. *$f$ is bounded.*

[5] sometimes is also called a 'global minimum'.

2. *There exist points* $x_1, x_2 \in [a,b]$ *such that*

$$f(x_1) \le f(x) \le f(x_2), \qquad \forall x \in [a,b],$$

*that is, there exists a point* $x_1 \in [a,b]$ *for which* $f$ *attains its minimum, and a point* $x_2 \in [a,b]$ *for which* $f$ *attains its maximum.*

*Proof*:

1. Suppose, by contradiction, that $f$ is not bounded. Then, for every $n \in \mathbb{N}$ there exists a point $x_n \in [a,b]$ for which $|f(x_n)| > n$. Consider the sequence $(x_n)_{n=1}^{\infty}$. By the Bolzano-Weierstrass Theorem, it has a converging sub-sequence $(x_{n_k})_{k=1}^{\infty}$ and denote by $x_0 = \lim_{k\to\infty} x_{n_k}$. Note that $x_0 \in [a,b]$. As $|f(x_{n_k})| > n_k \to \infty$ then the sequence $(f(x_{n_k}))$ is not bounded. On the other hand, as $f$ is continuous at $x_0$, we have $\lim_{k\to\infty} f(x_{n_k}) = f(x_0)$. This contradicts the fact that a converging sequence has to be bounded.

2. Consider the following set

$$S = \{f(x) : a \le x \le b\} = \mathrm{Im}(f).$$

We showed in 1 that $S$ is a bounded set, and clearly $S \ne \phi$. Then let

$$M = \sup S \qquad \text{and} \qquad m = \inf S.$$

There exists a sequence $(x_n)_{n=1}^{\infty}$ in $[a,b]$ such that $f(x_n) \to M$. Now, by the B-W theorem, $(x_n)$ has a converging subsequence $(x_{n_k})_{k=1}^{\infty}$ and let $x_2 = \lim_{k\to\infty} x_{n_k}$. Recall that a subsequence of a converging sequence converges to the same limit, i.e, we have

$$\lim_{k\to\infty} f(x_{n_k}) = M,$$

Finally, as $f$ is continuous at $x_2$ we have $M = f(x_2)$. Repeating the same argument for $m = \inf S$, there exists a point $x_1 \in [a,b]$ such that $f(x_1) = m$.

■

***Example 2.15*** *The function* $f(x) = \frac{1}{x}$ *is continuous on the interval* $(0,1)$ *but is unbounded from above. The interval is not closed, as is required in the proof of Theorem 2.7 part* 1.

***Example 2.16*** *The function $f : [0,1) \to \mathbb{R}$ defined by $f(x) = x$ does not have a maximum on $[0,1)$. Likewise, the function $g : \mathbb{R} \to \mathbb{R}$ defined by $g(x) = x$ does not have a maximum (nor a minimum) in $\mathbb{R}$. Thus, the Min-Max theorem does not apply in general for open or unbounded intervals.*

Looking back at the proof of Theorem 2.7, note where we have used the fact that the domain of $f$ is a closed interval, rather than open, for example; closed intervals have the property that for any converging sequence $x_n \to x_0$ in $[a,b]$, the limit $x_0$ is contained in $[a,b]$ (and, in particular, belongs to the domain of $f$). This property does not apply for domain which are open intervals, where which a sequence from within the interval can converge to a limit which is outside of the domain. The property of continuity is also crucial, as is demonstrated in the following example

***Example 2.17*** *The function $f : [-1,1] \to \mathbb{R}$ defined by*

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

*does not have a maximum nor minimum in $[-1,1]$. Indeed, it is not continuous at $x_0 = 0$.*

## 2.5 The inverse-function Theorem

When schematically thinking of a function al a collection of arrows from one set to another (i.e, from the domain of $f$ to its image), then the inverse of $f$ is the function that is obtained by reversing the arrows. However, not every function has an inverse function; the arrows have to be one-to-one. In this section we prove that an inverse function of a continuous function is continuous. We begin with some definitions,

***Definition 2.10*** *A real-valued function $f : E \to \mathbb{R}$ is said to be* ***injective****, or* ***one-to-one****, if for every $x, y \in E$,*

$$x \neq y \Rightarrow f(x) \neq f(y).$$

Given a point $y$ in the image of an injective function, then there exists a unique source $x \in E$ satisfying $f(x) = y$, in which case it is possible to consider $x$ as a function of $y$.

***Definition 2.11*** *Let $f : E \to \mathbb{R}$ be an injective, real-valued function. The inverse of $f$, denoted by $f^{-1} : \operatorname{Im} f \to \mathbb{R}$ is defined by*

$$f^{-1}(y) = x \qquad \textit{if and only if} \qquad f(x) = y.$$

Observe that as $y \in \operatorname{Im} f$ then there exists a point $x \in E$ such that $f(x) = y$. As $f$ is injective, this $x$ must be unique, and therefore $f^{-1}(y)$ is a well-defined function.

***Theorem 2.8*** *Let $f : [a, b] \to \mathbb{R}$ be a continuous, injective real-valued function. Then the inverse function $f^{-1} : \operatorname{Im} f \to \mathbb{R}$ is continuous.*

*Proof*: Let $y_0 \in \operatorname{Im} f$ and let $(y_n)_{n=1}^{\infty}$ be a sequence in $\operatorname{Im} f$ which converges to $y_0$. We show that $\lim_{n\to\infty} f^{-1}(y_n) = f^{-1}(y_0)$. Write $y_0 = f(x_0)$ and $y_n = f(x_n)$. Then, by the definition of the inverse function, $x_0 = f^{-1}(y_0)$ and $x_n = f^{-1}(y_n)$. Hence, we should show that $x_n \to x_0$. Suppose, by contradiction, that $(x_n)$ does not converge to $x_0$. Then there exists a subsequence $x_{n_k}$ which converge to a limit $x^*$ satisfying $x^* \neq x_0$. As $f$ is continuous at $x^*$ we have

$$\lim_{k\to\infty} f(x_{n_k}) = f(x^*).$$

Note that the sequence $y_{n_k} = f(x_{n_k})$ is a sub-sequence of $y_n$, and therefore converges to the same limit $y_0 = f(x_0)$. We thus have

$$\lim_{k\to\infty} f(x_{n_k}) = f(x_0),$$

and by the uniqueness property of limits,

$$f(x^*) = f(x_0).$$

Finally, as $f$ is injective we have $x_0 = x^*$, which is a contradiction. ∎

***Example 2.18*** *The function $g(x) : [0, \infty) \to \mathbb{R}$ defined by $g(x) = \sqrt{x}$ is continuous; Indeed, $g(x)$ is the inverse function of $f : [0, \infty) \to \mathbb{R}$ defined by $f(x) = x^2$, which is continuous. Similarly, the $n$-th root $g(x) = x^{1/n}$ is continuous for all $x \geq 0$.*

## 2.6 The Trigonometric functions

In this section we introduce the set of trigonometric functions $\sin x$ and $\cos x$ are prove that they are continuous. The **unit circle** is the circle of radius 1 in the 2-dimensional plain centred at $(0, 0)$. The most natural measure for angles is by radians,

***Definition 2.12*** *A **radian** is the angle that corresponds to an arc on the unit circle of length* 1*. Consequently,* $x$ *radians is the angle that corresponds to an arc on the unit circle of length* $x$*.*

As the length of the circumference of the unit circle is $2\pi$, then $x$ must satisfy $0 \le x \le 2\pi$.

***Definition 2.13*** *Let* $0 \le x \le 2\pi$*. Let* $(a, b)$ *denote the point on the unit circle that is set by an angle of* $x$ *radians from the 'x'-axis. We define*

$$\cos x = a \qquad \textit{and} \qquad \sin x = b.$$

We extend the definition of $\sin x$ and $\cos x$ to arbitrary $x \in \mathbb{R}$ by periodicity, i.e such that $\sin x = \sin(x + 2\pi)$ and $\cos x = \cos(x + 2\pi)$; for example, for $-2\pi \le x \le 0$ we define $\sin x = \sin(x + 2\pi)$ and note that the right-hand side is already defined as $0 \le x + 2\pi \le 2\pi$.

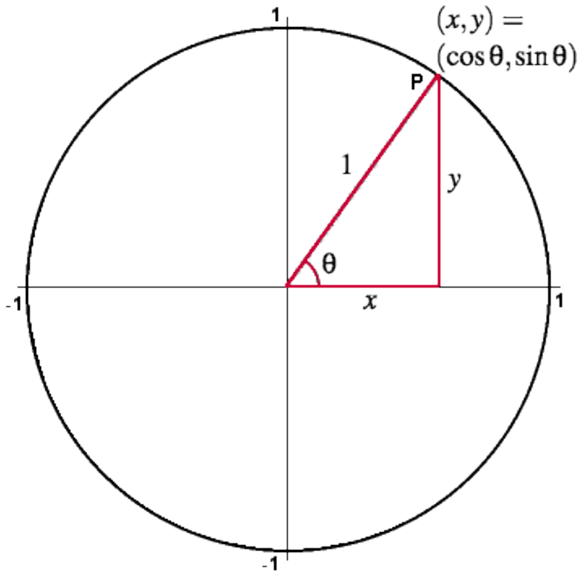



Figure 6: The unit circle

From the definition of $\sin x, \cos x$ and the Pythagoras theorem, some immediate trigonometric identities follow. We list the most common ones below. You are welcome to verify each one of them:

- $\sin 0 = 0$, $\sin \frac{\pi}{2} = 1$, $\cos 0 = 1$, $\cos \frac{\pi}{2} = 0$.
- $|\sin x|, |\cos x| \leq 1$.
- $\cos x = \sin(x + \frac{\pi}{2})$.
- $\sin^2 x + \cos^2 x = 1$.
- $|\sin x| \leq |x|$ for every $x \in \mathbb{R}$.
- $\sin x$ is one-to-one on the interval $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$.
- $\cos x$ is one-to-one on the interval $[0, \pi]$.

It can also be shown, using geometric arguments that for every $x, y \in \mathbb{R}$

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

Using the above observation, we can now prove

***Proposition 2.3*** *The function $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = \sin x$ is continuous.*

*Proof*: Let $x_0 \in \mathbb{R}$ and let $\varepsilon > 0$ be given. Set $\delta = \varepsilon/2$. Let $x \in \mathbb{R}$ such that $|x - x_0| < \delta$ and set $y = x - x_0$. Then

$$\begin{aligned} |\sin x - \sin x_0| = |\sin(y + x_0) - \sin x_0| &= |\sin y \cos x_0 + \sin x_0 \cos y - \sin x_0| \\ &\leq |\sin y| + |\cos y - 1| \end{aligned}$$

where we have used the triangle inequality and the property that $|\sin x_0|, |\cos x_0| \leq 1$. As $\cos y - 1 = -2\sin^2 \frac{y}{2}$ and $|\sin \frac{y}{2}| \leq 1$ we have $|\cos y - 1| \leq 2|\sin \frac{y}{2}|$. Thus,

$$|\sin x - \sin x_0| \leq |\sin y| + 2|\sin \frac{y}{2}| \leq 2|y| < 2\delta = \varepsilon.$$

■

Consequently, the function $\cos x = \sin(x + \pi/2)$ is continuous, by algebra of continuous functions and similarly $\tan x = \frac{\sin x}{\cos x}$ is continuous for every $x$ for which $\cos x \neq 0$.
By the inverse function theorem, the inverse functions $\arcsin x : [-1, 1] \to \mathbb{R}$ and $\arccos x : [-1, 1] \to \mathbb{R}$ are well defined.

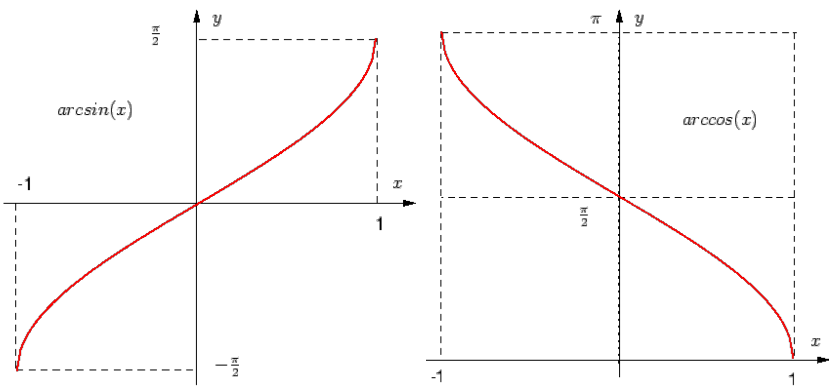



Figure 7: The graphs of arcsine and arccosine

# 3 Limits of functions

The intuition behind the notion of continuity that was introduced in the previous section was that $f(x)$ ‘approaches’ $f(x_0)$ as $x$ ‘approaches’ $x_0$. This intuition was translated to formal mathematical content in two equivalent ways; via the use of sequence and via the use of an $\varepsilon - \delta$ formulation. But what if the value of $f(x)$ approaches some value $L$ (which may or may not be $f(x_0)$)? we can, as we are already used for sequences, define a 'limit' of a function $\lim_{x\to x_0} f(x) = L$, and then translate continuity at $x_0$ as the requirement that the limit $\lim_{x\to x_0} f(x) = L$ exists and $L = f(x_0)$. After finding a satisfying definition for the limit of a function at $x_0$, then in particular, $f$ would be discontinuous at $x_0$, if the limit $\lim_{x\to x_0} f(x)$ does not exists or that it exists but $\lim_{x\to x_0} f(x) \neq f(x_0)$. To define a limit of a function at a point $x_0$, it is not necessary to require that $x_0$ belongs to the domain of $f$ (i.e, $f(x_0)$ does not have to be defined). However, $x_0$ must be an accumulation point,

***Definition 3.1*** *Let $E$ be a subset of $\mathbb{R}$. A point $x_0 \in \mathbb{R}$ is said to be an **accumulation point** of $E$, if for any $\delta > 0$ there exists a point $x \in E$ such that $0 < |x - x_0| < \delta$.*

In other words, an accumulation point of a set $E$ is a point for which you can approach by points from the set $E$. We proceed with some simple examples,

- Let $E = (a, b]$. Then every point $x_0 \in E$ is an accumulation point.
- Let $E = \left\{\frac{1}{n} : n \in \mathbb{N}\right\}$. Then $x_0 = 0$ is an accumulation point (and the only one).
- Let $E = \mathbb{N}$. There are no accumulation points.
- Let $E = \mathbb{Q}$. Then every point $x_0 \in \mathbb{R}$ is an accumulation point (recall that $\mathbb{Q}$ is a dense set in $\mathbb{R}$).

Note that the accumulation point $x_0 \in \mathbb{R}$ does not have to be a point in the set $E$. Henceforth we will assume that $x_0$ is an accumulation point of a set $E$, unless stated otherwise. We are now ready to give the definition for a limit of a function,

***Definition 3.2*** *A real-valued function $f : E \to \mathbb{R}$ has a **limit** $L$ at $x_0$ if for every $\varepsilon > 0$ there exists a $\delta > 0$ such that*

$$0 < |x - x_0| < \delta \qquad \textit{implies} \qquad |f(x) - L| < \varepsilon.$$

*In which case we denote by $L = \lim_{x\to x_0} f(x)$.*

Pay attention for the differences of Definition 3.2 from the definition of continuity 2.4; The function $f$ may not be defined in $x_0$ (which may lay outside of the domain $E$), hence is the additional requirement that $0 < |x - x_0|$. If $f$ is continuous at $x_0 \in E$ then clearly $f(x_0)$ is a limit of $f$ at $x_0$. The converse also holds; if $L$ is a limit of $f$ at $x_0$, $f(x_0)$ is defined and $f(x_0) = L$, then clearly $f$ is continuous at $x_0$. However, it may be that the limit $\lim_{x\to x_0} f(x) = L$ exists, $f(x_0)$ is defined but is different than $L$ (see Figure 8). Recall that the

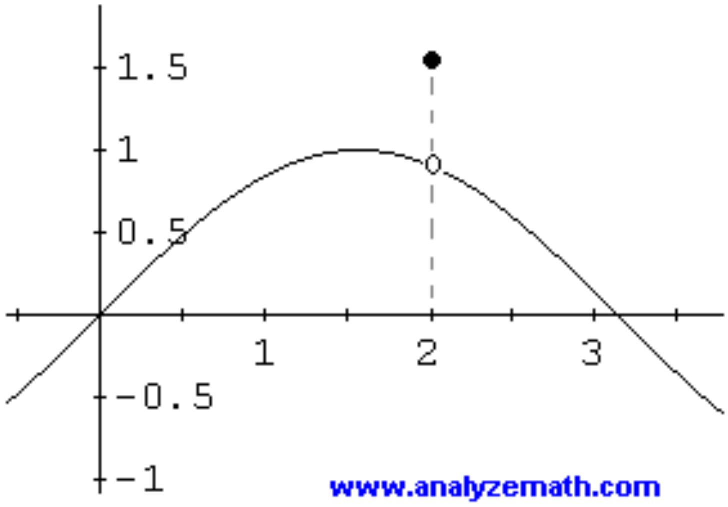



Figure 8: A discontinuous function: the limit $\lim_{x\to 2} f(x) = 1 \neq f(2)$. In which case the discontinuity of $f$ at $x_0$ is said to be 'removable'.

notation $L = \lim_{x\to x_0} f(x)$ is senseless without a uniqueness argument for a limit, which we prove in the following statement,

***Proposition 3.1*** *(uniqueness) Let $f : E \to \mathbb{R}$. Suppose that $L_1$ and $L_2$ are limits of $f$ at $x_0$. Then $L_1 = L_2$.*

*Proof*: Let $\varepsilon > 0$. By definition 3.2 there exist $\delta_1, \delta_2 > 0$ [6] such that for every $x \in E$ such that $0 < |x - x_0| < \delta_i$ we have

$$|f(x) - L_i| < \varepsilon/2, \qquad i = 1, 2.$$

[6] Note that $\delta_1$ and $\delta_2$ may be different

Setting $\delta = \min(\delta_1, \delta_2)$ then for every $x \in E$ such that $0 < |x - x_0| < \delta$ both inequalities hold. By the triangle inequality,

$$|L_1 - L_2| \leq |f(x) - L_1| + |f(x) - L_2| < \varepsilon/2 + \varepsilon/2 = \varepsilon.$$

Finally, as the above holds for every $\varepsilon > 0$, then $L_1 = L_2$. . [7] ■

Similarly to the way continuity is defined via sequences in the previous section, we can give a definition for a sequential limit,

***Definition 3.3*** *Let $f : E \to \mathbb{R}$. Then $L$ is said to be the* ***sequential limit*** *of $f$ at $x_0$ if for every sequence $(x_n)$ in $E$ such that $x_n \to x_0$ and $x_n \neq x_0$ for every $n$, we have $\lim_{n\to\infty} f(x_n) = L$.*

Note that $f$ may not be defined in $x_0$, thus we add the additional requirement that $x_n \neq x_0$ for every $n$. The proof for the following theorem is identical to the proof of Theorem 2.1, except that $f(x_0)$ is replaced by $L$,

***Theorem 3.1*** *Let $f : E \to \mathbb{R}$. Then $L$ is a sequential limit of $f$ at $x_0$ if and only if $\lim_{x\to x_0} f(x) = L$.*

The following theorem is straightforward from the definition of a limit, in any of the equivalent formulations (either the $\varepsilon - \delta$ or the sequences formulation). We state it as a theorem just for convenience, but it is actually more of a remark,

***Theorem 3.2*** *Let $f : E \to \mathbb{R}$, and set*

$$\bar{f}(x) = \begin{cases} f(x) & x \neq x_0 \\ L & x = x_0 \end{cases}$$

*Then, $\bar{f}$ is continuous at $x_0$ if and only if $\lim_{x\to x_0} f(x) = L$.*

One of the main advantages of the definition of a limit of a function, is that limits can be considered also when $f$ is not continuous, or even undefined at the point $x_0$.

***Example 3.1***

$$\lim_{x\to 0} x \sin \frac{1}{x} = 0.$$

[7] Whats wrong with this proof? $L_1 = \lim_{x\to x_0} f(x)$, $L_2 = \lim_{x\to x_0} f(x)$. Therefore $L_1 = L_2$.

*Proof*: Let $\varepsilon > 0$ be given, and set $\delta = \varepsilon$. For every $0 < |x| < \delta$, we have

$$|x \sin \frac{1}{x} - 0| = |x \sin \frac{1}{x}| \leq |x| < \delta = \varepsilon,$$

i.e, $\lim_{x \to 0} x \sin \frac{1}{x}$ exists and is equal to 0. ■

The calculation of limits as the above limit can follow more easily when using the so called Squeeze rule (or Sandwich rule). The proof uses the Squeeze rule for sequence, which reads

***Theorem 3.3*** *(the Squeeze rule for sequences) Let $a_n, b_n, c_n$ be three sequences such that*

$$b_n \leq a_n \leq c_n.$$

*Suppose that the limits $\lim_{n\to\infty} b_n$ and $\lim_{n\to\infty} c_n$ exist and are equal to $L$. Then the sequence $a_n$ is convergent, and $\lim_{n\to\infty} a_n = L$.*

We can now state,

***Theorem 3.4*** *(the Squeeze rule) Let $f, g, h : E \to \mathbb{R}$ be three real-valued function such that*

$$g(x) \leq f(x) \leq h(x), \qquad \forall x \in E.$$

*Suppose that the limits $\lim_{x\to x_0} g(x)$ and $\lim_{x\to x_0} h(x)$ exist and are equal to $L$. Then the limit of $f$ at $x_0$ exists, and $\lim_{x\to x_0} f(x) = L$.*

Note that the main assertion of the above theorem is for the existence of the limit; Indeed, if we assume that the limit exists and denote $L' = \lim_{x\to x_0} f(x)$ then trivially $L \leq L' \leq L \Rightarrow L = L'$.

*Proof*: Let $x_n \to x_0$ be a sequence converging to $x_0$ in $E \setminus \{x_0\}$. Then for every $n$ we have

$$g(x_n) \leq f(x_n) \leq g(x_n).$$

By Theorem 3.1 we have

$$\lim_{n\to\infty} g(x_n) = \lim_{n\to\infty} h(x_n) = L,$$

and by the Squeeze rule for sequences (remember Term I?), the limit $\lim_{n\to\infty} f(x_n)$ exists and equals to $L$. As this applies to any such sequence, then by Theorem 3.1, the limit $\lim_{x\to x_0} f(x)$ exists and equals to $L$. ■

***Example 3.2***

$$\lim_{x\to 0} x \sin\frac{1}{x} = 0$$

*Proof*: We use the Squeeze rule twice; The function $f(x) = x\sin\frac{1}{x}$ satisfies $0 \le |f(x)| \le |x|$ for every $x \neq 0$. Thus, by the squeeze rule we have $\lim_{x\to 0}|f(x)| = 0$. As $-|f(x)| \le f(x) \le |f(x)|$ for every $x \neq 0$, then applying the squeeze rule again we obtain that $\lim_{x\to 0} f(x) = 0$. ■

Note that the squeeze rule implies that if $\lim_{x\to x_0}|f(x)| = 0$ then $\lim_{x\to x_0} f(x) = 0$.

Why is the squeeze rule so useful? when ever proving that $L$ is a limit of a function $f$ at $x_0$, we show that $|f(x) - L| < \varepsilon$ as long as $x$ satisfies $0 < |x - x_0| < \delta$. But what we actually do is upper bound the term $|f(x) - L|$ be a term which we already know that converges to zero, such as $|x - x_0|$ or $|x - x_0|^2$. Once doing that we are ready to set $\delta$ as $\varepsilon$ or $\sqrt{\varepsilon}$ or whatever will fit. Therefore, the Squeeze rule just saves as the setting of $\delta$ (and the time of writing 'given $\varepsilon > 0$').

Another property that holds for limits and follows directly from the definition of sequential limits is the algebra of limits; for example, if $\lim_{x\to x_0} f(x) = L_1$ and $\lim_{x\to x_0} g(x) = L_2$ then $\lim_{x\to x_0}(f(x)+g(x))$ exists and equals to $L_1+L_2$. We can now calculate various limits,

***Example 3.3*** *Calculate the following limit (if exists)*

$$\lim_{x\to 0} \sin x \cos x.$$

**Solution** The function $f(x) = \sin x \cos x$ is continuous. Thus, $\lim_{x\to 0} f(x) = f(0) = 0$.

***Example 3.4***

$$\lim_{x\to 0} \sin x \cos\frac{1}{x} = 0.$$

*Proof*: For every $x \neq 0$, we have

$$0 \le |\sin x \cos(1/x)| \le |\sin x| \le |x|.$$

Thus, by the squeeze rule, $\lim_{x\to 0} \sin x \cos\frac{1}{x}$ exists and equals to 0. ■

***Example 3.5*** *Calculate the following limit (if exists)*

$$\lim_{x\to 1}\frac{x^2+x-2}{x-1}.$$

*Proof*: Note that $x^2+x-2=(x-1)(x+2)$ thus

$$\lim_{x\to 1}\frac{x^2+x-2}{x-1}=\lim_{x\to 1}(x+2)=3.$$

■

## 3.1 Left and right limits

The idea behind left and right limits is to consider limits that are restricted only to points $x\in E$ that are either left or right from $x_0$. The same idea applies for right and left continuity.

***Definition 3.4*** *Let $f:E\to\mathbb{R}$. A number $L\in\mathbb{R}$ is said to be a* ***right-limit (left-limit)*** *of $f$ at $x_0$, if for every $\varepsilon>0$ there exists a $\delta>0$ such that for every $x\in E$ such that $x_0<x$ $(x<x_0)$,*

$$0<|x-x_0|<\delta\Rightarrow|f(x)-L|<\varepsilon.$$

*In which case we denote $\lim_{x\to x_0+}f(x)=L$ ( $\lim_{x\to x_0-}f(x)=L$). Consequently, $f$ is called* ***right-continuous*** *if $\lim_{x\to x_0+}f(x)=f(x_0)$ and* ***left-continuous*** *if $\lim_{x\to x_0-}f(x)=f(x_0)$.*

It can be shown by repeating the proof of Proposition 3.1, with the additional restriction that $x_0<x$ $(x<x_0)$ that if a right (left) limit exists, then it is unique. Thus, the notations $\lim_{x\to x_0+}f(x)$ and $\lim_{x\to x_0-}f(x)$ are justified. We give a simple example (though not shown in class but is worthwhile)

***Example 3.6*** *The function $f(x)=[x]$ is right-continuous.*

*Proof*: Recall that you have shown in your exercise that $f$ is continuous at every point $x\notin\mathbb{N}$. It is therefore sufficient to show that $f$ is right-continuous at every $n\in\mathbb{N}$. However,

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

■

Note that $f$ is not continuous from the left at $x = n$, as

$$\lim_{x \to n-} f(x) = \lim_{x \to n-} (n-1) = n - 1 \neq f(n).$$

Note that by Definition 3.4, if $\lim_{x \to x_0} f(x) = L$ then clearly $\lim_{x \to x_0+} f(x) = L$ and $\lim_{x \to x_0-} f(x) = L$. The converse also holds,

***Proposition 3.2*** *Let $f : E \to \mathbb{R}$ be a real-valued function. Then $\lim_{x \to x_0} f(x) = L$ if and only if $\lim_{x \to x_0+} f(x) = L$ and $\lim_{x \to x_0-} f(x) = L$.*

*Proof*: The first direction is straightforward. Suppose now that $\lim_{x \to x_0+} f(x) = L$ and $\lim_{x \to x_0-} f(x) = L$ and let $\varepsilon > 0$. Then there exist $\delta_1, \delta_2 > 0$ such that

$$0 < |x - x_0| < \delta_1, x < x_0 \Rightarrow |f(x) - L| < \varepsilon.$$
$$0 < |x - x_0| < \delta_2, x > x_0 \Rightarrow |f(x) - L| < \varepsilon.$$

Setting $\delta = \min(\delta_1, \delta_2)$ then for every $0 < |x - x_0| < \delta$ we either have $x < x_0$ or $x > x_0$, either way $|f(x) - L| < \varepsilon$. ■

Consequently, $f$ is continuous if and only if it is right and left continuous.

***Example 3.7*** *Find all values of $a \in \mathbb{R}$ for which the function*

$$f(x) = \begin{cases} ax + 1 & x > 1 \\ 2x - a & x \leq 1 \end{cases}$$

*is continuous.*

*Proof*: By the algebra of continuous functions, clearly $f$ is continuous at every point $x_0 \neq 1$. For $x_0 = 1$ we have

$$\lim_{x \to 1+} f(x) = \lim_{x \to 1+} (ax + 1) = a + 1$$

and

$$\lim_{x \to 1-} f(x) = \lim_{x \to 1+} (2x - a) = 2 - a$$

Thus, $f$ is continuous if and only if $a + 1 = 2 - a$, i.e, $a = \frac{1}{2}$. ■

## 3.2 Finite limits at infinity

Similarly to the definition of a limit of a function at a point $x_0$, we can define a limit at $\infty$ or $-\infty$,

***Definition 3.5*** *Let $f : E \to \mathbb{R}$ be a real-valued function. Suppose that the domain $E$ is not bounded from above. We say that $f$ has a **limit** $L$ as $x$ approaches $\infty$, if for every $\varepsilon > 0$ there exists $M \in \mathbb{R}$ such that for every $x \in E$ such that $x > M$ we have*

$$|f(x) - L| < \varepsilon.$$

*In which case we define* $\lim_{x\to\infty} f(x) = L$.

In this module, when considering $\lim_{x\to\infty} f(x) = L$ then often $f$ will be defined on some interval of the form $[a, \infty)$ for some $a \in \mathbb{R}$. Similarly (copy-pasting the above definition and changing just a bit)

***Definition 3.6*** *Let $f : E \to \mathbb{R}$ be a real-valued function. Suppose that the domain $E$ is not bounded from **below**. We say that $f$ has a **limit** $L$ as $x$ approaches $-\infty$, if for every $\varepsilon > 0$ there exists $M \in \mathbb{R}$ such that for every $x \in E$ such that $x < M$ we have*

$$|f(x) - L| < \varepsilon.$$

*In which case we define* $\lim_{x\to-\infty} f(x) = L$.

Like in previous sections, here too sequential limits at infinity are defined and are shown to be equivalent to the usual $\varepsilon - M$ definitions. Also, the usual properties of uniqueness, the algebra of limits and the squeeze rule applies hold; The proofs follow directly from similar properties for sequences. You are welcome to verify you can recover these proofs for yourself.

***Example 3.8*** *Prove that* $\lim_{x\to\infty} \frac{1}{x} = 0$.

*Proof*: Let $\varepsilon > 0$ be given and set $M = \frac{1}{\varepsilon}$. Then for every $x > M$ we have

$$|f(x)| = \frac{1}{x} < \frac{1}{M} = \varepsilon.$$

■

*Example 3.9* *Calculate the limit (if exists)* $\lim_{x\to\infty} \frac{-x^2+x+1}{2x^2+\sqrt{x}+2}$.

*Proof*: We have

$$\frac{-x^2+x+1}{2x^2+\sqrt{x}+2} = \frac{-1+\frac{1}{x}+\frac{1}{x^2}}{2+\frac{1}{x^{3/2}}+\frac{2}{x^2}}.$$

By the algebra of limits, and as $\frac{1}{x} \to 0$ as $x \to \infty$, we have

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

■

## 3.3 Infinite limits

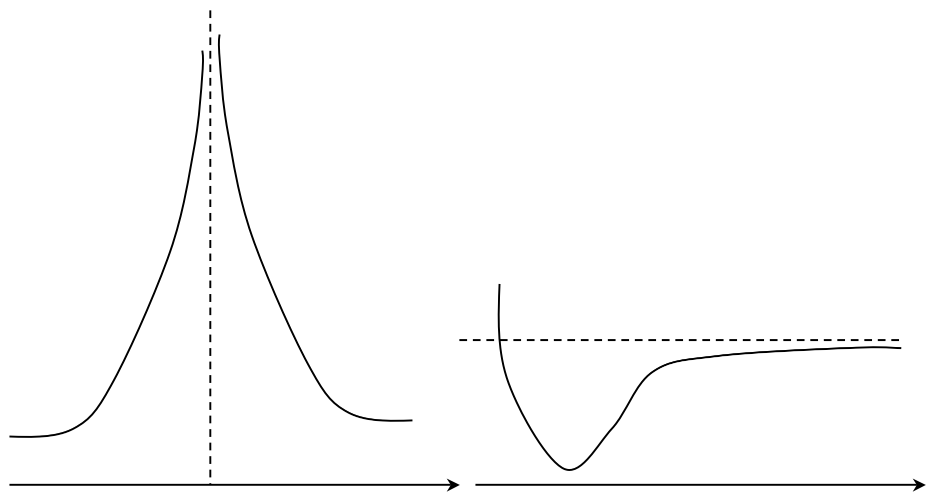


Figure 9: An infinite limit and a limit at infinity

*Definition 3.7* *Let* $f : E \to \mathbb{R}$ *and let* $x_0$ *be an accumulation point of* $E$. *We say that* $\lim_{x\to x_0} f(x) = \infty$, *if* $\forall M > 0$ $\exists \delta > 0$ *such that* $\forall x \in E$ *with* $0 < |x - x_0| < \delta$ *we have*

$$f(x) > M.$$

The equivalent sequential definition is:

***Definition 3.8*** *Let $f : E \to \mathbb{R}$ and let $x_0$ be an accumulation point of $E$. We say that $\lim_{x\to x_0} f(x) = \infty$, if for every sequence $(x_n)$ in $E \setminus x_0$ such that $x_n \to x_0$ we have $\lim_{n\to\infty} f(x_n) = \infty$.*

Likewise, (although I did not define this in class)

***Definition 3.9*** *Let $f : E \to \mathbb{R}$ and suppose that the domain $E$ is not bounded from above. We say that $\lim_{x\to\infty} f(x) = \infty$, if $\forall M > 0$ $\exists N > 0$ such that for every $x \in E$ such that $x > N$ we have*

$$f(x) > M.$$

and the equivalent sequential form is

***Definition 3.10*** *Let $f : E \to \mathbb{R}$ and suppose that the domain $E$ is not bounded from above. We say that $\lim_{x\to\infty} f(x) = \infty$,if for every sequence $(x_n)$ in $E$ such that $x_n \to \infty$ we have $\lim_{n\to\infty} f(x_n) = \infty$.*

Verify you understand the intuition behind the above definitions. Recover the equivalent sequential definitions. Similarly, you should be able to recover definitions for limits like

- $\lim_{x\to-\infty} f(x) = -\infty$.
- $\lim_{x\to\infty} f(x) = -\infty$
- $\lim_{x\to x_0} f(x) = -\infty$.
- $\lim_{x\to x_0+} f(x) = \infty$.

***Example 3.10***

$$\lim_{x\to 1+} \frac{1}{x-1} = \infty.$$

*Proof*: Let $M \in \mathbb{R}$ be given. We may assume without loss of generality that $M > 0$ (if we are greater than 0 then we are greater than any negative number). Set $\delta = \frac{1}{M} > 0$. Then $\forall x$ such that $1 < x < 1 + \delta$ we have

$$\frac{1}{x-1} > \frac{1}{\delta} = M.$$

■

# 4 Differentiation

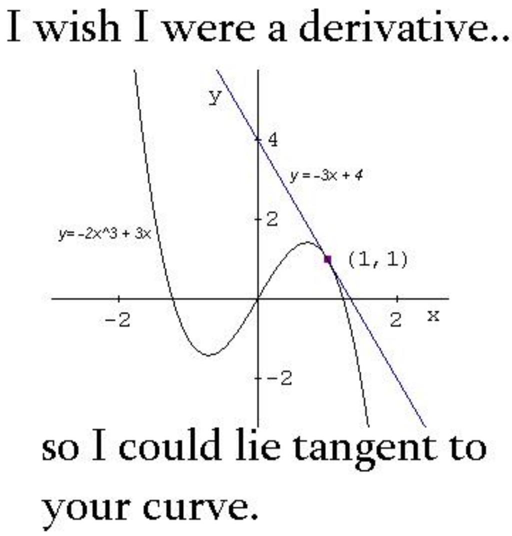



Figure 10: The graph of a differentiable function

The notion of differentiability refers to a function having a 'smooth' graph. The smoothness of the graph of $f$ at point $x_0$ can be interpreted as the existence of a straight line that is tangent to the graph of $f$ at $(x_0, f(x_0))$ (see Figure 10). If there does not exist a tangent line, or if the tangent line is not unique, then $f$ is said to be not differentiable at $x_0$. If the function is continuous but not differentiable, then the graph of $f$ has the shape of a spearhead at $x_0$ (see Figure 12).

***Definition 4.1*** *A real-valued function $f : E \to \mathbb{R}$ is said to be* ***differentiable*** *at a point $x_0 \in E$, if the limit*

$$\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} \tag{4.1}$$

*exists. In which case, we denote this limit by $f'(x_0)$.*

Given any set of points $x_1, x_2$ and $y_1, y_2$, the fraction $a = \frac{y_1 - y_2}{x_1 - x_2}$ is the slope of the straight line that crosses the points $(x_1, y_1)$ and $(x_2, y_2)$ in the plain.

Therefore, the limit (4.1) is interpreted as the slope of the tangent line to the graph of $f$ at the point $(x_0, f(x_0))$ (see Figure 11).

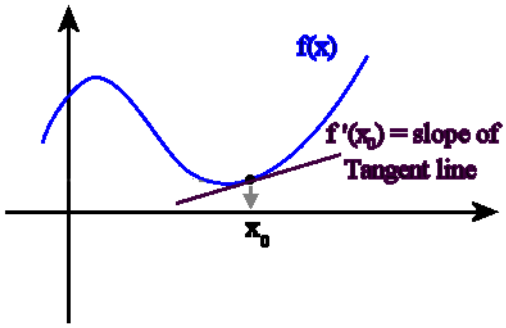



Figure 11: The graph of a differentiable function

*Remarks*

- Denoting by $h = x - x_0$, then $x \to x_0$ if and only if $h \to 0$. Thus, the limit (4.1) can be re-written as

$$\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} = \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h}.$$

- If the limit (4.1) exists, then the numerator has to converge to zero (because the denomirator does!). Thus, if $f$ is differentiable at $x_0$ then $\lim_{x \to x_0}(f(x) - f(x_0)) = 0$, i.e, $f$ is continuous at $x_0$ (in other words, a discontinuous function cannot be differentiable).

***Example 4.1*** *The function $f(x) = x$ and $g(x) = x^2$ are differentiable for every $x_0 \in \mathbb{R}$. Moreover, $f'(x_0) = 1$ and $g'(x_0) = 2x_0$.*

*Proof*: By Definition 4.1 we have

$$f'(x_0) = \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} = \lim_{x \to x_0} \frac{x - x_0}{x - x_0} = 1.$$

and

$$g'(x_0) = \lim_{x \to x_0} \frac{g(x) - g(x_0)}{x - x_0} = \lim_{x \to x_0} \frac{(x - x_0)(x + x_0)}{x - x_0} = \lim_{x \to x_0} (x + x_0) = 2x_0.$$

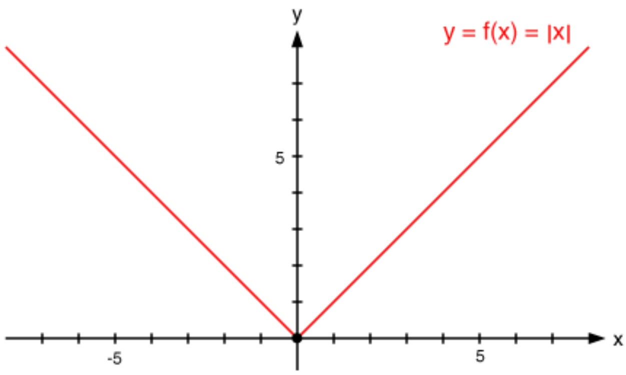



Figure 12: The function $|x|$ is not differentiable at $x_0 = 0$.

■

Next is an example of a continuous function that is not differentiable.

***Example 4.2*** *The function* $f(x) = |x|$ *is not differentiable at* $x_0 = 0$.

*Proof*: Consider the right and left limits,

$$\lim_{x \to 0+} \frac{f(x) - f(0)}{x - 0} = 1 \qquad \text{and} \qquad \lim_{x \to 0-} \frac{f(x) - f(0)}{x - 0} = -1.$$

Thus, by Proposition 3.2 the limit $\lim_{x \to 0} \frac{f(x)-f(0)}{x-0}$ does not exist, which implies, according to Definition 4.1, that the function is not differentiable at $x_0 = 0$. ■

***Example 4.3*** *The function* $f(x) = x^2 D(x)$*, where* $D(x)$ *is the Dirichelet function given by* (2.3) *is differentiable only at* $x_0 = 0$.

*Proof*: For any $x_0 \neq 0$ the function $f(x)$ is discontinuous, and therefore not differentiable. To show that it is indeed discontinuous, then suppose by contradiction it were continuous for some $x_0 \neq 0$. Then the function $D(x) = \frac{f(x)}{x^2}$ is continuous by the algebra of continuous function. However we have seen that the Dirichlet function is nowhere continuous (see Example 2.4). To show that $f$ is continuous at $x_0 = 0$ we use the fact that $|D(x)| \leq 1$, hence

$$0 \leq \frac{|f(h) - f(0)|}{|h|} = \frac{h^2 D(|h|)}{|h|} \leq |h|.$$

Thus, by the Squeeze rule we have $\lim_{h\to 0} \frac{|f(h)-f(0)|}{|h|} = 0$, hence

$$f'(0) = \lim_{h\to 0} \frac{f(h) - f(0)}{h} = 0.$$

■

## 4.1 Algebra of differentiable functions

Much like in Section 2.3, sums, multiplications and compositions of differentiable functions is respectively differentiable. Now we should additionally provide a formula for the derivative of each algebraic operation of two functions, in terms of the derivatives of the functions.

***Theorem 4.1*** *Let $f, g : E \to \mathbb{R}$ be two real-valued functions that are differentiable at a point $x_0 \in E$. Then*

1. *The function $h_1 = f + g$ is differentiable at $x_0$, and*

$$h_1'(x_0) = f'(x_0) + g'(x_0).$$

2. *(**Leibnitz rule**) The function $h_2 = f \cdot g$ is differentiable at $x_0$, and*

$$h_2'(x_0) = f'(x_0)g(x_0) + g'(x_0)f(x_0).$$

3. *If $g(x_0) \neq 0$ then the function $h_3 = \frac{f}{g}$ is differentiable at $x_0$, and*

$$h_3'(x_0) = \frac{f'(x_0)g(x_0) - f(x_0)g'(x_0)}{g^2(x_0)}.$$

*Proof*:

1. We have for every $x \neq x_0$,

$$\frac{h_1(x) - h_1(x_0)}{x - x_0} = \frac{f(x) - f(x_0)}{x - x_0} + \frac{g(x) - g(x_0)}{x - x_0},$$

   and therefore by the algebra of limits,

$$\begin{aligned} h_1'(x_0) &= \lim_{x\to x_0} \frac{h_1(x) - h_1(x_0)}{x - x_0} = \lim_{x\to x_0} \frac{f(x) - f(x_0)}{x - x_0} + \lim_{x\to x_0} \frac{g(x) - g(x_0)}{x - x_0} \\ &= f'(x_0) + g'(x_0). \end{aligned}$$

2. We have for every $x \neq x_0$,

$$\frac{h_2(x) - h_2(x_0)}{x - x_0} = \frac{f(x)g(x) - f(x_0)g(x_0)}{x - x_0}$$
$$= \frac{f(x) - f(x_0)}{x - x_0} g(x) + \frac{g(x) - g(x_0)}{x - x_0} f(x_0).$$

Taking the limit $x \to x_0$ and using the fact that $g$ is differentiable, hence continuous at $x_0$, and the algebra of limits, we obtain

$$h_2'(x_0) = \lim_{x \to x_0} \frac{h_2(x) - h_2(x_0)}{x - x_0}$$
$$= \lim_{x \to x_0} \left[\frac{f(x) - f(x_0)}{x - x_0} g(x)\right] + \lim_{x \to x_0} \left[\frac{g(x) - g(x_0)}{x - x_0} f(x_0)\right]$$
$$= f'(x_0)g(x_0) + g'(x_0)f(x_0).$$

3. Note that we only need to prove this part for $h_3(x) = \frac{1}{g(x)}$, and then apply (2) for $f(x) \cdot \frac{1}{g(x)}$. Note also that as $g(x_0) \neq 0$ then the continuity of $g$ at $x_0$ implies that $g(x) \neq 0$ in some open interval centred at $x_0$, thus $h_3(x)$ is defined on this interval. We have for every $x \neq x_0$ in this interval,

$$\frac{h_3(x) - h_3(x_0)}{x - x_0} = \frac{\frac{1}{g(x)} - \frac{1}{g(x_0)}}{x - x_0}$$
$$= \frac{g(x_0) - g(x)}{(x - x_0)g(x)g(x_0)}$$

Thus, by the algebra of limits and the continuity of $g$ at $x_0$, we obtain

$$h_3'(x_0) = \lim_{x \to x_0} \frac{h_3(x) - h_3(x_0)}{x - x_0} = -\frac{g'(x_0)}{g^2(x_0)}.$$

■

With his theorem in hand, we can already differentiate a large number of functions, [8]

***Example 4.4*** *The derivative of a linear combination,*

$$(af + bg)' = af' + bg', \qquad a, b \in \mathbb{R}$$

[8] For as long as we remember to never drink and derive ;-)

*Proof*: We only need to verify that if $f$ is differentiable at $x_0$ then so is $af$, and $(af)' = af'$. This follows from the derivative of the product, and from $a' = 0$ (constant function). ■

***Example 4.5*** *Let* $n \in \mathbb{N}$ *and consider the function* $f_n(x) = x^n$. *Then* $f_n'(x) = nx^{n-1}$.

*Proof*: We show this inductively. We already know that his is true for $n = 0$ and $n = 1$. Suppose this were true for $n$, then $f_{n+1}(x) = x^{n+1} = x \cdot f_n(x)$, and by the differentiation rule for products,

$$f_{n+1}'(x) = 1 \cdot f_n(x) + x \cdot f_n'(x) = x^n + x \cdot nx^{n-1} = (n+1)x^n.$$

■

***Example 4.6***

$$\left(2 + x - 2x^{17}\right)' = 2' + x' - 2\left(x^{17}\right)' = 1 - 34x^{16}.$$

***Example 4.7*** *We can differentiate a product of more than two functions, by applying the Leibniz rule more than once, for example*

$$(f \cdot g \cdot h)' = ((f \cdot g) \cdot h)' = (f \cdot g)'h + (f \cdot g)h' = (f'g + fg')h + fgh' = f'gh + fg'h + fgh'.$$

Suppose we take for granted that

$$\sin' = \cos \qquad \text{and} \qquad \cos' = -\sin.$$

Then we can calculate the derivative of, say, $\sin^3 x$, by applying the rule of a product of three function. But what about the derivative of $\sin x^3$? Here we need a rule for how to differentiate compositions.

***Theorem 4.2*** *(**The Chain rule**) Let* $f : E \to G$ *(that is,* $\operatorname{Im} f \subseteq G$*), and* $g : G \to \mathbb{R}$ *be two real-valued functions, Suppose that* $f$ *is differentiable at* $x_0$*, and that* $g$ *is differentiable at* $y_0 = f(x_0)$*. Then the composition* $g \circ f : E \to \mathbb{R}$ *is differentiable at* $x_0$*, and*

$$(g \circ f)'(x_0) = g'(f(x_0)) \cdot f'(x_0).$$

Before giving the proof for the above theorem, lets try to calculate the derivative of $g \circ f$ at $x_0$ naively from the definition, like we did for other algebraic operations. We then look at the function

$$A(h) = \frac{g(f(x_0 + h)) - g(f(x_0))}{h},$$

and calculate its limit as $h \to 0$. For small values of $h$, we expect the argument $f(x_0 + h)$ and $f(x_0)$ of $g$ to be very close. This suggests the following treatment,

$$A(h) = \frac{g(f(x_0 + h)) - g(f(x_0))}{f(x_0 + h) - f(x_0)} \cdot \frac{f(x_0 + h) - f(x_0)}{h}.$$

Since $f(x_0 + h) - f(a) \to 0$ as $h \to 0$, it seems that this product tends to $g'(f(x_0)) \cdot f'(x_0)$ (i.e, we are happy). The problem with this argument is that while the limit $h \to 0$ means that the case $h = 0$ is not to be considered, there is nothing to prevent the denominator $f(x_0+h)-f(x_0)$ from vanishing, rendering this expression meaningless (i.e, we were too happy). Yet, the result is still correct, and it only takes a little more subtlety to prove it. Actually, there is more than one correct way to do so, and perhaps the simplest way uses the following Lemma,

***Lemma 4.1*** *Let $f : E \to \mathbb{R}$ be a real-valued function. Then $f$ is differentiable at $x_0 \in E$ if and only if there exists a function $\Phi : E \to \mathbb{R}$ such that*

$$f(x) = f(x_0) + \Phi(x)(x - x_0) \tag{4.2}$$

*and $\Phi$ is continuous at $x_0$. In which case, $\Phi(x_0) = f'(x_0)$.*

*Proof*: Assume first that $f$ is differentiable at $x_0$ and set

$$\Phi(x) = \begin{cases} \frac{f(x)-f(x_0)}{x-x_0} & x \neq x_0 \\ f'(x_0) & x = x_0 \end{cases}.$$

Clearly, (4.2) holds for every $x \in E$ (when $x = x_0$ both sides vanish). Also, by Definition 4.1 we have $\lim_{x\to x_0} \Phi(x) = f'(x_0) = \Phi(x_0)$, i.e, $\Phi$ is continuous. Now assume the existence of such function $\Phi$. Then by (4.2),

$$\lim_{x\to x_0} \frac{f(x) - f(x_0)}{x - x_0} = \lim_{x\to x_0} \Phi(x) = \Phi(x_0)$$

where we have used the assumption that $\Phi$ is continuous at $x_0$. Thus, $f$ is differentiable at $x_0$ and $f'(x_0) = \Phi(x_0)$. ■

We are now ready to prove the Chain rule,

*Proof*: Setting $y_0 = f(x_0)$, then applying (4.2) for $f$ and $g$ about $x_0$ and $y_0$ respectively gives

$$\begin{aligned} f(x) &= f(x_0) + \Phi(x)(x - x_0), \\ g(y) &= g(y_0) + \Psi(y)(y - y_0). \end{aligned} \tag{4.3}$$

Susbtituting $y = f(x)$ and $y_0 = f(x_0)$ in (4.3) yields

$$\begin{aligned} h(x) = g(f(x)) &= g(f(x_0)) + \Psi(f(x))(f(x) - f(x_0)) \\ &= h(x_0) + [\Psi(f(x)) \cdot \Phi(x)] (x - x_0). \end{aligned}$$

The function $H(x) = \Psi(f(x)) \cdot \Phi(x)$ is continuous at $x_0$ by the algebra of continuous functions (and that $f$ is differentiable at $x_0$ implies that $f$ is continuous at $x_0$). Finally, by Lemma 4.2 we have that $h$ is differentiable at $x_0$ and

$$h'(x_0) = H(x_0) = \Psi(f(x_0)) \cdot \Phi(x_0) = g'(f(x_0)) \cdot f'(x_0).$$

■

***Example 4.8*** *The function $f(x) = \sin x^3$ is differentiable as it is the composition of two differentiable functions (again, we assume for now that $\sin' x = \cos x$). Using the Chain rule,*

$$f'(x) = \sin' x^3 \cdot (x^3)' = \cos x^3 \cdot 3x^2.$$

The last operation which would be of interest to differentiate is the inverse operation, allowing us to derive inverse function such as $\arcsin x$ and $\arctan x$. The perhaps naive way to differentiate inverse functions is by applying the Chain rule; let $f : E \to \mathbb{R}$ be an injective function, and let $f^{-1} : \operatorname{Im} f \to \mathbb{R}$ be its inverse function. By definition,

$$\left(f^{-1} \circ f\right)(x) = x.$$

Differentiating both sides of the equations and applying the Chain rule to the left-hand side yields

$$\left(f^{-1}\right)'(f(x)) \cdot f'(x) = 1,$$

i.e,

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

and, assuming that the right-hand side is defined (i.e, that $f'(x) \neq 0$), we obtain a formula for the derivative of the inverse function. Unfortunately, we cannot use the Chain rule without already knowing that the composed functions are differentiable; we did not yet prove that the inverse of a differentiable function is differentiable. To prove that, we need to go back in some sense to the definition of differentiability, and be slightly more careful about our assumptions on $f$.

***Theorem 4.3*** *(The inverse rule) Let $f : [a, b] \to \mathbb{R}$ be an injective function, and suppose that $f$ is differentiable in $[a, b]$[9] and that $f'(x_0) \neq 0$ for some $x_0 \in [a, b]$. Then the inverse function $f^{-1} : \operatorname{Im} f \to \mathbb{R}$ is differentiable at $y_0 = f(x_0)$. Moreover,*

$$\left(f^{-1}\right)'(y_0) = \frac{1}{f'(x_0)}. \tag{4.4}$$

*Proof*: Denoting by $y = f(x) \iff x = f^{-1}(y)$ for every $x \in [a, b]$, we have

$$\begin{aligned}\lim_{y \to y_0} \frac{f^{-1}(y) - f^{-1}(y_0)}{y - y_0} &= \lim_{y \to y_0} \frac{x - x_0}{f(x) - f(x_0)} \\ &= \lim_{y \to y_0} \frac{1}{\left[\frac{f(x) - f(x_0)}{x - x_0}\right]}.\end{aligned}$$

It remains to explain why we can replace the limit $y \to y_0$ by $x \to x_0$ on the right-hand side. To justify this replacement, we need to prove that $y \to y_0 \iff x \to x_0$. In other words, that $x \to x_0$ if and only if $y \to y_0$. As $f$ is differentiable at $x_0$, then $f$ is continuous at $x_0$. Thus,

$$\lim_{x \to x_0} y = \lim_{x \to x_0} f(x) = f(x_0) = y_0.$$

To show the opposite direction, then by the Inverse function Theorem 2.8, $f^{-1}$ is continuous at $y_0$. Thus

$$\lim_{y \to y_0} x = \lim_{y \to y_0} f^{-1}(y) = f^{-1}(y_0) = x_0.$$

[9] This assumption can be weakened. We only use the assumption that $f$ is differentiable at every point in $[a, b]$ to apply the Inverse function theorem.

■

We proceed with a simple application of the inverse rule.

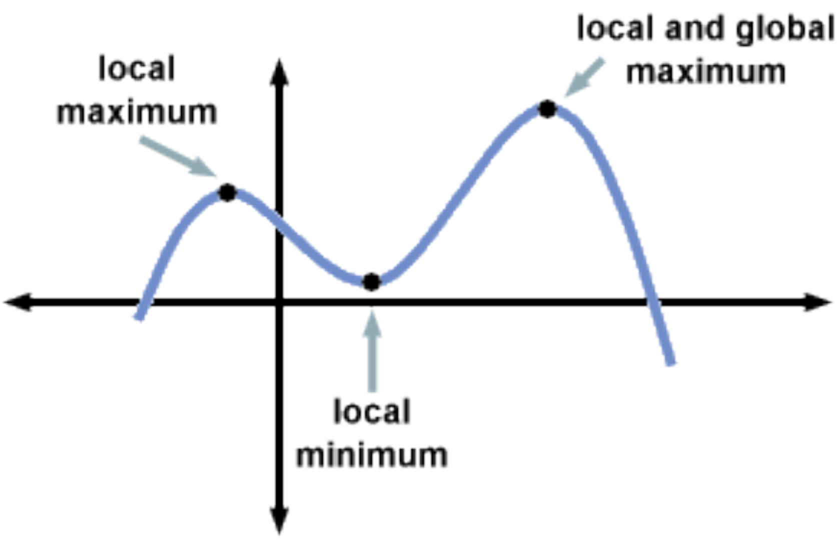
**Example 4.9** *Let $f : [0, \infty) \to \mathbb{R}$ be the function given by $f(x) = x^2$. Then $f$ is injective and $f^{-1} : [0, \infty) \to \mathbb{R}$ is given by $f^{-1}(y) = \sqrt{y}$. Following the notations $y = f(x)$ and $x = f^{-1}(y)$, the inverse rule (4.4) reads*

$$(\sqrt{y})' = \frac{1}{f'(x)} = \frac{1}{2x} = \frac{1}{2\sqrt{y}}, \qquad y \neq 0.$$

## 4.2 Properties of differentiable functions

**Definition 4.2** *Let $f : E \to \mathbb{R}$ be a real-valued function. A point $x_0 \in (a, b)$ is said to be a **local maximum** (a **local minimum**) if*

1. *There exists an open interval $I = (a, b) \subset E$ such that $x_0 \in I$.*
2. $$f(x) \leq f(x_0) \qquad (f(x) \geq f(x_0)), \qquad \forall x \in I.$$




**Theorem 4.4** *(Fermat's Theorem) Let $f : E \to \mathbb{R}$. If $x_0$ is a local maximum (minimum) of $f$, and $f$ is differentiable at $x_0$, then $f'(x_0) = 0$.*

*Proof*: Set $L = f'(x_0)$. We want to show that $L = 0$. Consider the function defined for every $x \in E$ such that $x \neq x_0$ by

$$A(x) = \frac{f(x) - f(x_0)}{x - x_0}.$$

As $x_0$ is a local maximum, then there exists an open interval $I = (a, b) \subseteq E$ such that

$$f(x) - f(x_0) \leq 0, \qquad \forall x \in I.$$

Thus, for every $x > x_0$ such that $x \in I$ we have

$$A(x) \leq 0.$$

Hence,

$$L = \lim_{x \to x_0+} A(x) \leq 0.$$

Similarly, for every $x < x_0$ such that $x \in I$ we have

$$A(x) \geq 0.$$

Thus,

$$L = \lim_{x \to x_0-} A(x) \geq 0.$$

Putting the two inequalities together, we obtain that $L = 0$, as required. If $x_0$ is a local minimum, then the same hold with the opposite inequalities. ■

***Example 4.10*** *The function $f : [0, 1] \to \mathbb{R}$ defined by $f(x) = x$ has a maximum at $x_0 = 1$. However, this maximum is not a local maximum, and $f'(1) = 1$. Indeed, Fermat's theorem requires that the maximal point has some open interval around it for which $f$ is defined, hence we can consider both the right and left limits of $A(x)$ as $x \to x_0$.*

***Theorem 4.5*** *(Rolle's Theorem) Let $f : [a, b] \to \mathbb{R}$ be a real-valued function, satisfying the following requirements:*

1. *$f$ is differentiable on $(a, b)$.*
2. *$f$ is continuous on $[a, b]$.*
3. *$f(a) = f(b)$.*

*Then, there exists a point $c \in (a, b)$ such that $f'(c) = 0$ (see Figure 14)*

*Proof*: It follows from the continuity of $f$ that it has a maximum and a minimum point in $[a, b]$ (the Max-Min theorem). If the maximum or the minimum occurs at some point $c \in (a, b)$, then by Fermat theorem we have $f'(c) = 0$, and we are done. The only remaining alternative is that $a$ and $b$ are both minima and maxima. In which case, the assumption that $f(a) = f(b)$ implies that the maximal and the minimal values of $f$ in $[a, b]$ are the same, hence $f$ is constant and its derivative satisfies $f'(c) = 0$ for every $c \in [a, b]$. ■

Figure 13: Michel Rolle

Clearly, the proof remains valid under the stronger assumption that $f$ be differentiable on $[a, b]$, and this would be the case in most of our applications.

***Example 4.11*** *The requirement that $f$ be differentiable at $(a, b)$ is imperative; Consider the absolute value function $f(x) = |x|$. It is continuous at $[-1, 1]$, satisfies $f(-1) = f(1) = 1$, but there is no point $c$ for which $f'(c) = 0$.*

***Example 4.12*** *Rolle's theorem implies that differentiable functions have the following property: between every two distinct zeros of $f$ there exists a zero of $f'$.*

Rolle's theorem is a useful tool for proving uniqueness of solutions for given algebraic equations.

***Example 4.13*** *The equation $\sin x + 17x = 1$ has a unique solution in $\mathbb{R}$.*

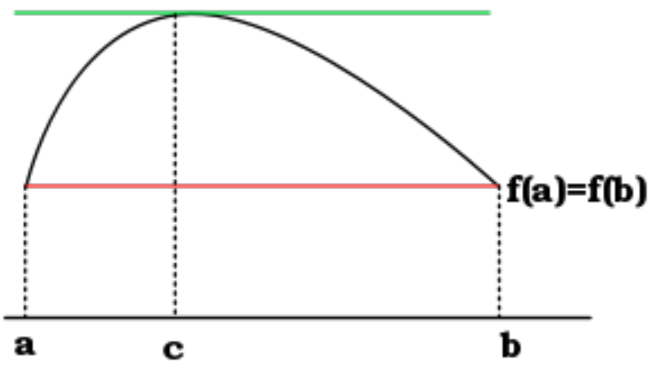



Figure 14: Rolle's theorem

*Proof*: Let $f(x) = \sin x + 17x$. Then $f$ is continuous, and clearly $f(-1) < 1$ and $f(1) > 1$. Thus by the Intermediate-Value theorem, the equation $f(x) = 1$ has a solution. To show that it is unique, then suppose by contradiction there exist two solutions $a < b$ in $\mathbb{R}$. As $f$ is differentiable, then by Rolle's theorem, there exists a point $c \in (a, b)$ such that $f'(c) = 0$. However,

$$f'(x) = \cos x + 17 \geq -1 + 17 > 0, \qquad \forall x \in \mathbb{R}.$$

■

Figure 15: Joseph-Louis Lagrange

***Theorem 4.6*** *(Lagrange's Mean-Value Theorem) Let* $f : [a, b] \to \mathbb{R}$ *be a real-valued function, satisfying the following requirements:*

1. $f$ *is differentiable on* $(a, b)$.

2. $f$ *is continuous on* $[a, b]$.

*Then, there exists a point* $c \in (a, b)$ *such that*

$$\frac{f(b) - f(a)}{b - a} = f'(c). \tag{4.5}$$

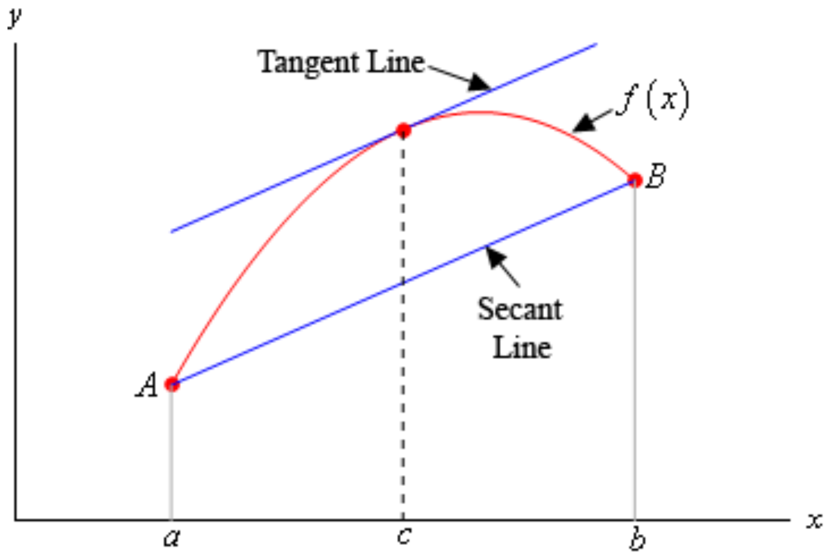



Figure 16: The Mean-Value theorem

The Mean-Value theorem has a simple geometric interpretation; recall that the fraction $\frac{f(b)-f(a)}{b-a}$ is the slope of the secant line that connects the points $(a, f(a))$ and $(b, f(b))$ on the graph of $f$. The Mean-Value theorem then states that this slope is equal to the slope of the tangent line at some interior point $c \in (a, b)$ (see Figure 16). In light of that, the proof for the Mean-Value theorem can be described in an intuitive way; simply rotate your head so that the secant line would become horizontal, then apply Rolle's theorem. Indeed, in the proof we give below we define a function $F$ which is simply the corresponding appropriate rotation of $f$, *Proof*: Consider the function $F : [a, b] \to \mathbb{R}$ defined by

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

$F$ is continuous on $[a, b]$ and differentiable in $(a, b)$. Moreover, $F(a) = F(b) = f(a)$. Thus, by Rolle's theorem there exists a point $c \in (a, b)$ such that

$$0 = F'(c) = f'(c) - \frac{f(b) - f(a)}{b - a}.$$

■

We proceed with some observations,

- The Mean-Value theorem is a generalisation of Rolle's theorem, which follows immediately from (4.5) with the additional assumption that $f(a) = f(b)$. As Rolle's theorem is also applied in the proof of the Mean-Value theorem, they are in fact equivalent theorems. This is not surprising in light of their geometric interpretations.

- For any $a \neq b$ we have

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

  Hence we can re-state the Mean-Value theorem as follows: for any $x \neq y$ there exists a point $c$ between $x$ and $y$ such that

  $$\frac{f(x) - f(y)}{x - y} = f'(c).$$

- The point $c$ may not be unique.

The Mean-Value theorem is a very useful tool in analysis; it provides a relation between the values of $f'$ (at perhaps an unknown point $c$) and the values of $f$.

***Example 4.14*** *Let $f : [a, b] \to \mathbb{R}$ be a function for which $f'(c) = 0$ for every $c \in [a, b]$. Then $f$ is constant.*

*Proof*: For every $x \neq y$ in $[a, b]$ we can apply the Mean-Value theorem for the interval $[x, y]$ (or $[y, x]$), thus there exists a point $c$ between $x$ and $y$ such that

$$\frac{f(x) - f(y)}{x - y} = f'(c) = 0 \Rightarrow f(x) = f(y).$$

■

***Example 4.15*** *The function* $\sin x$ *satisfies the inequality*

$$|\sin x - \sin y| \leq |x - y| \qquad \forall x, y \in \mathbb{R}.$$

*Proof*: If $x = y$ then the inequality is trivial. Otherwise, by the Mean-Value theorem there exists a point $c$ between $x$ and $y$ such that

$$\frac{|\sin x - \sin y|}{|x - y|} = |\sin' c| = |\cos c| \leq 1.$$

Multiplying both sides of the inequality by $|x - y|$ gives the result. ■

***Example 4.16*** *For every* $x > 1$ *we have*

$$1 - \frac{1}{x} < \ln x < x - 1.$$

*Proof*: Recall that $\ln x : (0, \infty) \to \mathbb{R}$ is the inverse function of $e^x$, and by the inverse rule (assuming we know that $(e^x)' = e^x$),

$$\ln' x = \frac{1}{e^{lnx}} = \frac{1}{x}.$$

Applying the Mean-Value theorem for the interval $[1, x]$, then there exists a point $1 < c < x$ such that

$$\frac{\ln x - \ln 1}{x - 1} = \ln' c = \frac{1}{c},$$

and as $\ln 1 = 0$ and $1 < c < x$, we obtain

$$1 - \frac{1}{x} = \frac{x - 1}{x} < \ln x = \frac{x - 1}{c} < x - 1.$$

■

Note that in particular, we have every $x > 1$

$$\frac{\ln x}{x} < 1 - \frac{1}{x} < 1,$$

i.e, $\ln x < x$.

***Definition 4.3*** *A function $f : E \to \mathbb{R}$ is said to be* ***monotonically increasing (monotonically decreasing)*** *if for every $x, y \in E$ such that $x < y$ we have $f(x) \leq f(y)$ ($f(x) \geq f(y)$). If the latter inequality is satisfied strongly (that is, $f(x) < f(y)$) for every $x < y$, then $f$ is said to be* ***strongly monotonically increasing****.*

***Example 4.17*** *If $f'(x) \geq 0$ on $[a, b]$ then $f$ is monotonically increasing.*

*Proof*: By the Mean-Value theorem, for every $x < y$ in $[a, b]$ there exists a point $x < c < y$ such that

$$\frac{f(y) - f(x)}{y - x} = f'(c) \geq 0.$$

Multiplying both sides by the denominator $y - x > 0$ we obtain that $f(x) \geq f(y)$. ∎

By the same argument as above, if $f'(x) > 0$ on $[a, b]$ then $f$ is strongly monotonically increasing. The assumption that $f'(x) \geq 0$ for every $x \in [a, b]$ is imperial; we can construct a function $f$ such that $f'(x_0) > 0$ but $f$ is not monotonically increasing at any interval containing $x_0$.

What about the converse?

***Proposition 4.1*** *Let $f : [a, b] \to \mathbb{R}$ be a differentiable function, that is monotonically increasing. Let $x_0 \in [a, b]$. Then $f'(x_0) \geq 0$.*

*Proof*: It is easy to verify that if $f$ is monotonically increasing, then the function $A(x) = \frac{f(x)-f(x_0)}{x-x_0}$ is nonnegative. Thus, $f'(x_0) = \lim_{x \to x_0} A(x) \geq 0$. ∎

The last theorem in this sequence of statements on differentiable functions defined in closed intervals is a generalization of Lagrange's Mean-Value theorem,

***Theorem 4.7*** *(Cauchy's Mean-Value Theorem) Let $f, g : [a, b] \to \mathbb{R}$ be two real-valued functions, satisfying the following requirements:*

1. *$f, g$ are differentiable on $(a, b)$.*

2. *$f, g$ are continuous on $[a, b]$.*

*Then, there exists a point $c \in (a, b)$ such that*

$$g'(c)\,[f(b) - f(a)] = f'(c)\,[g(b) - g(a)]\,. \tag{4.6}$$

Figure 17: Augustin Louis Cauchy

Note that if $g'(x) \neq 0$ for every $x \in [a,b]$ and if $g(b) - g(a) \neq 0$, then (4.6) reads

$$\frac{f(b) - f(a)}{g(b) - g(a)} = \frac{f'(c)}{g'(c)}. \tag{4.7}$$

Taking $g(x) = x$ then (4.7) reduces to (4.5). In other words, Cauchy's Mean-Value theorem is a generalization of Lagrange's Mean-Value theorem.

*Proof*: Consider the function $h : [a,b] \to \mathbb{R}$ defined by

$$h(x) = g(x)\left[f(b) - f(a)\right] - f(x)\left[g(b) - g(a)\right].$$

By the algebra of continuous and differentiable functions, $h$ is continuous on $[a,b]$ and differentiable on $(a,b)$. Moreover, $h(a) = h(b) = g(a)f(b) - f(a)g(b)$. Thus, by Rolle's theorem, there exists a point $c \in (a,b)$ such that

$$0 = h'(c) = g'(c)\left[f(b) - f(a)\right] - f'(c)\left[g(b) - g(a)\right].$$

■

If you consider the curve in $\mathbb{R}^2$ defined by the values of $(f(x), g(x))$ as $x$ changes from $a$ to $b$, then Cauchy's Mean-Value theorem has a geometric interpretation with respect to this curve (see Figure 18).

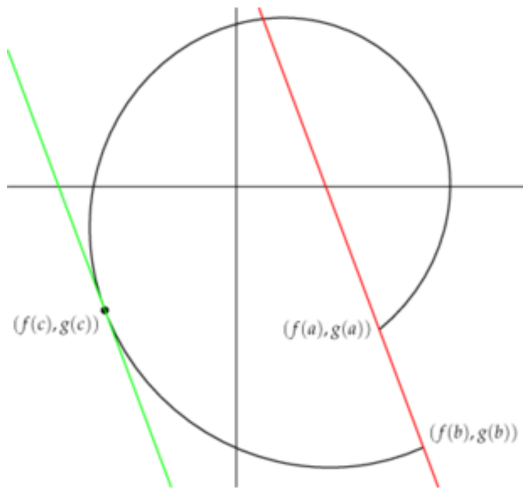



Figure 18: Cauchy's theorem.

## 4.3 One-sided derivatives

Just like we defined right and left continuity by using the one-sided definitions for limits, we define the right and left derivatives of a function,

*Definition 4.4* *A function $f : E \to \mathbb{R}$ is said to be* ***right differentiable*** *at a point $x_0 \in E$, if the limit*

$$\lim_{x \to x_0+} \frac{f(x) - f(x_0)}{x - x_0}$$

*exists. In which case we denote this limit by $f'_+(x_0)$. Similarly, $f$ is said to be* ***left differentiable*** *at $x_0$, if the limit*

$$\lim_{x \to x_0-} \frac{f(x) - f(x_0)}{x - x_0}$$

*exists. In which case we denote this limit by $f'_-(x_0)$.*

By Proposition 3.2, then $f$ is differentiable at $x_0$ if and only if $f'_-(x_0)$ and $f'_+(x_0)$ exist and $f'(x_0) = f'_-(x_0) = f'_+(x_0)$ (assuming that $x_0$ is not an edge point). [10]

[10] Suppose that $f : [a, b] \to \mathbb{R}$. Then the limit of the derivative at $a$, $\lim_{x \to a} \frac{f(x) - f(a)}{x - a}$ coincides with the right limit at $a$, $\lim_{x \to a+} \frac{f(x) - f(a)}{x - a}$. In which case the left limit at $a$ is not defined as $f(x)$ is not defined for $x < a$.

***Example 4.18*** *The function $f(x) = |x|$ is not differentiable at $x_0 = 0$, as $f'_+(0) = 1$ and $f'_-(0) = -1$.*

***Example 4.19*** *Let*

$$f(x) = \begin{cases} 2x^2 & x \geq 0 \\ ax + b & x < 0 \end{cases}.$$

*For which values of $a, b \in \mathbb{R}$ is $f$ differentiable at $x_0 = 0$?*

**Solution** Well, first $f$ has to be continuous, hence

$$b = \lim_{x \to 0-} f(x) = \lim_{x \to 0+} f(x) = 0.$$

Secondly, the right and left derivatives of $f$ at zero have to be equal, hence

$$0 = f'_+(0) = f'_-(0) = a.$$

## 4.4 Higher order derivatives

Fermat's theorem states that if $x_0 \in E$ is a local maximum/minimum and if $f$ is differentiable at $x_0$, then $f'(x_0) = 0$. But how can we tell if $x_0$ is a local maximum or a local minimum? there are several ways to do that. A very convenient one is to consider the values of $f''(x_0)$ (assuming it exists).
Suppose that $f : E \to \mathbb{R}$ is differentiable. Then the derivative $f'(x)$ is defined for any choice of $x \in E$. One can consider the derivative as a function of $x$, i.e, $f' : E \to \mathbb{R}$, and ask questions addressing the function $f'(x)$ (e.g, is it continuous? monotonic? differentiable?) if the derivative $f'(x)$ is differentiable at $x_0 \in E$, then we say that $f$ is twice differentiable at $x_0$. In which case, the second derivative, $f''(x_0)$ is defined at $x_0$. The second derivative may be used to sort points of $f$ as either local minima or local maxima of $f$.

***Definition 4.5*** *Let $f : E \to \mathbb{R}$. A point $x_0 \in E$ is called* ***critical*** *if $f'(x_0) = 0$, or if $f'(x_0)$ does not exist (i.e, $f$ is not differentiable at $x_0$).*

***Theorem 4.8*** *Suppose that $f : (a, b) \to \mathbb{R}$ is differentiable and let $x_0 \in (a, b)$ be a critical point. Suppose that $f$ is twice differentiable at $x_0$. Then*

1. *If $f''(x_0) > 0$ then $x_0$ is a local minimum of $f$.*
2. *If $f''(x_0) < 0$ then $x_0$ is a local maximum of $f$.*

The above theorem does not give any information on critical points for which $f''(x_0) = 0$. Such points can be either local maxima, local minima or 'saddle' points, as we shall see in examples below.

*Proof*:

1. Suppose that $f''(x_0) > 0$. Then, by definition, the function $A(x) = \frac{f'(x)-f'(x_0)}{x-x_0}$ defined for every $x \in (a, b)$ such that $x \neq x_0$ satisfies $\lim_{x\to x_0} A(x) = f''(x_0) > 0$. Thus, there exists a $\delta > 0$ such that for every $x \in (x_0 - \delta, x_0 + \delta)$ we have $A(x) > 0$. Note that as $x_0$ is critical then $A(x) = \frac{f'(x)}{x-x_0}$.

   Suppose first that $x_0 < x < x_0+\delta$. Then as $x-x_0 > 0$ and $A(x) > 0$ we have that $f'(x) > 0$. This applies for every such $x$. By the Mean-Value theorem (4.5), there exists a point $x_0 < c < x$ for which

   $$f(x) - f(x_0) = f'(c)(x - x_0).$$

   As $c$ satisfies $x_0 < c < x_0 + \delta$ then we have $f'(c) > 0$. Thus, $f(x) - f(x_0) > 0$.

   Now, suppose that $x_0-\delta < x < x_0$. Then as $x-x_0 < 0$ and $A(x) > 0$ we have that $f'(x) < 0$. This applies for every such $x$. By the Mean-Value theorem (4.5), there exists a point $x < c < x_0$ for which

   $$f(x) - f(x_0) = f'(c)(x - x_0).$$

   As $c$ satisfies $x_0 - \delta < c < x_0$ then we have $f'(c) < 0$, and as $x - x_0 < 0$ we have $f(x) - f(x_0) > 0$.

   We have shown that for every $x \in (x_0 - \delta, x_0 + \delta)$ such that $x \neq x_0$ we have $f(x) - f(x_0) > 0$. Thus, $x_0$ is a local minimum.

2. Suppose now that $f''(x_0) < 0$, and consider the function $g(x) = -f(x)$. Then $g$ is differentiable in $(a, b)$ and twice differentiable at $x_0$, and $g''(x_0) = -f''(x_0) > 0$. Thus, $x_0$ is a local minimum of $g$, hence it is a local maximum of $f = -g$.

■

***Example 4.20*** *The function $f(x) = x^4$ has a minimum at $x_0 = 0$, yet $f'(0) = f''(0) = 0$. The function $f(x) = -x^4$ has a maximum at $x_0 = 0$, yet*

*$f'(0) = f''(0) = 0$. Thus, we cannot determine whether a critical point $x_0$ is a maximum or a minimum if $f''(x_0) = 0$, just by looking at $f''(x_0)$. In fact, higher derivatives of $f$ have to be considered (such as $f'''$).*

***Example 4.21*** *The function $f(x) = x^3$ satisfies $f'(0) = f''(0) = 0$ but zero is neither a minimum nor a maximum. In which case $x_0$ is called a saddle point (see Figure **??**).*

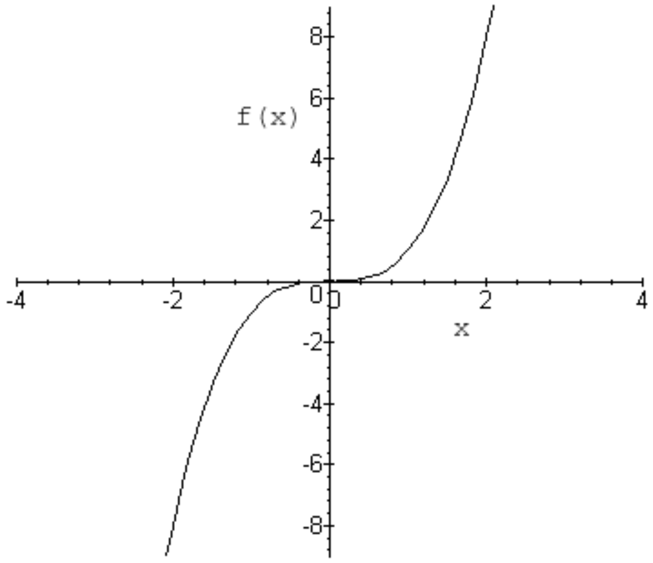



Figure 19: The graph of $f(x) = x^3$. The point $x_0 = 0$ is a critical point which is neither a maximum nor a minimum (saddle point)

***Definition 4.6*** *We define the sets*

$$C^0[a,b] = \{f : [a,b] \to \mathbb{R} : f \text{ is cont. in } [a,b]\}$$

*and*

$$\begin{aligned} C^1[a,b] &= \{f : [a,b] \to \mathbb{R} : f \text{ is diff. in } [a,b], \text{ and } f'(x) \text{ is continuous}\} \\ &= \left\{f : [a,b] \to \mathbb{R} : f' \in C^0[a,b]\right\}. \end{aligned}$$

*We define recursively the set*

$$C^{n+1}[a,b] = \{f : [a,b] \to \mathbb{R} : f' \in C^n[a,b]\}.$$

In other words, $C^n[a,b]$ is the set of functions for which the $n$-th order derivative is defined and continuous on $[a,b]$. Note that since every differentiable function is continuous, then $f \in C^{n+1}[a,b]$ implies that $f \in C^n[a,b]$, thus $C^{n+1}[a,b] \subseteq C^n[a,b]$.

***Definition 4.7*** *We define the set*

$$C^\infty[a,b] = \{f : [a,b] \to \mathbb{R} : f \in C^n[a,b],\ \forall n \in \mathbb{N}\}.$$

A similar definition may be given for $C^n(E)$ and $C^\infty(E)$ for any subset $E \subseteq \mathbb{R}$.

***Example 4.22*** *The function $f(x) = x^{17} + 2x + 1$ is in $C^\infty(\mathbb{R})$. Indeed, you can differentiate it as many times as you like at any point $x \in \mathbb{R}$.*

***Example 4.23*** *Consider the following function*

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

1. *Is $f$ continuous?*
2. *Is $f$ differentiable?*
3. *Is $f \in C^1(\mathbb{R})$?*

**Solution**

1. Yes. Clearly, $f$ is continuous at every $x \neq 0$ by the algebra of continuous functions. To show that it is continuous at $x = 0$, recall that the Squeeze rule implies that the product of a function converging to zero by a bounded function is a function which converges to zero; As $\lim_{x\to 0} x^2 = 0$ and $|\sin \frac{1}{x}| \leq 1$ we have by the Squeeze rule

   $$\lim_{x\to 0} f(x) = 0 = f(0),$$

   hence $f$ is continuous as $x = 0$.

2. Yes. Clearly, $f$ is differentiable at every $x \neq 0$ by the algebra of differentiable functions. To check whether $f$ is differentiable at $x = 0$ we consider the limit of the derivative
$$\lim_{x\to 0} \frac{f(x) - f(0)}{x - 0} = \lim_{x\to 0} x \sin \frac{1}{x} = 0,$$
where we have used the squeeze rule. Thus, $f$ is differentiable at zero and $f'(0) = 0$.

3. To determine whether $f'(x)$ is continuous, we calculate $f'$ using the rules of differentiations for $x \neq 0$, and obtain
$$f'(x) = \begin{cases} 2x \sin \frac{1}{x} - \cos \frac{1}{x} & x \neq 0 \\ 0 & x = 0 \end{cases}.$$
While the limit of the first term $\lim_{x\to 0} 2x \sin \frac{1}{x}$ exists and is equal to zero, the second term $\cos \frac{1}{x}$ does not converge as $x \to 0$ (see example 2.6). Thus, the limit $\lim_{x\to 0} f'(x)$ does not exist, hence $f'$ is not continuous at $x = 0$, which implies that $f \notin C^1(\mathbb{R})$.

# 5 Power series

Power series are a special class of series of function, which, in their most general representation take the form
$$\sum_{n=0}^{\infty} f_n(x), \tag{5.1}$$
where $x \in \mathbb{R}$ is some number for which $f_n(x)$ is defined for all $n \in \mathbb{N}$ (assuming such $x$ exists). Fixing such point $x$, the sequence $(f_n(x))_{n=1}^{\infty}$ is a sequence of numbers. If the corresponding series of numbers converges, it is convenient to denote its limit as a function of $x$, namely,
$$f(x) = \sum_{n=0}^{\infty} f_n(x) = \lim_{n\to\infty} \sum_{k=0}^{n} f_k(x).$$
The limit function $f$ is defined for every $x$ for which the above series converges. Unfortunately, obtaining an analytic expression for the function $f$ , given analytic expressions for the functions $f_n$, is often impossible, even when the functions $f_n$ are relatively very simple[11].

[11]for example, there is (so far) no analytic expression for the value of $\sum_{n=1}^{\infty} \frac{1}{n^3}$

A simpler problem that one may address, is of finding the domain of $f$, namely, the maximal set of points $E \subseteq \mathbb{R}$ for which the series (5.1) converges. Alternatively, one may ask for which values of $x$ does the series (5.1) converge? Unfortunately, determining the set $E$ may also be also be a hard problem. [12]

***Definition 5.1*** *A **power series** is a series of the form*

$$\sum_{n=0}^{\infty} a_n (x - x_0)^n, \tag{5.2}$$

*for some $x_0 \in \mathbb{R}$ and sequence $(a_n)_{n=0}^{\infty}$ in $\mathbb{R}$.*

*Remarks*

- Power series are a special class series of functions (5.1), where $f_n(x) = a_n(x - x_0)^n$.
- It is conventional to write the first term as $a_0 = a_0(x - x_0)^0$, thus the series (5.2) is interpreted as

$$\sum_{n=0}^{\infty} a_n (x - x_0)^n = a_0 + a_1(x - x_0) + a_2(x - x_0)^2 + \dots$$

- Denoting by

$$E = \left\{ x \in \mathbb{R} : \sum_{n=0}^{\infty} a_n (x - x_0)^n \text{ converges} \right\},$$

then $E \subseteq \mathbb{R}$ and $x_0 \in \mathbb{R}$. In fact, setting $f(x) = \sum_{n=0}^{\infty} a_n(x - x_0)^n$, then $E$ is the domain of the function $f$, the 'base point' $x_0$ always belongs to the domain and $f(x_0) = a_0$.

***Example 5.1*** *Consider the power series $\sum_{n=0}^{\infty} n!(x - x_0)^n$, where we use the convention that $0! = 1$. If $x = x_0$ then the series obviously converges to the value of 1. Otherwise, if $x \neq x_0$ then the $n$-th term of the series $n!(x - x_0)^n$ does not converge to zero, hence the series diverges. Thus, $E = \{0\}$.*

[12] Note that determining whether a series converges is equivalent to determining whether the corresponding sequence of partial sums converges. Similarly, every sequence $(S_n)_{n=0}^{\infty}$ can be viewed as a sequence of partial sums of a series defined by $a_n = S_{n+1} - S_n$ for $n \geq 1$ and $a_0 = S_0$. Therefore, a series is just way to view a sequence.

The next examples are applications of the so-called *ratio test*, which we state below

***Theorem 5.1*** *(The ratio test) Let $\sum_{n=0}^{\infty} a_n$ be a series in $\mathbb{R}$, such that $a_n \neq 0$ for every $n$. Suppose that the following limit exists,*

$$L = \lim_{n\to\infty} \frac{|a_{n+1}|}{|a_n|}.$$

*Then*

- *If $L < 1$ then the series converges.*
- *If $L > 1$ then the series diverges.*

***Example 5.2*** *Consider the geometric series $\sum_{n=0}^{\infty} x^n$. We have*

$$\frac{|x^{n+1}|}{|x^n|} = |x|$$

*hence, by the ratio test, the series converges for every $|x| < 1$ and diverges for every $|x| > 1$. If $|x| = 1$ it is easy to verify that the series diverges, hence*

$$E = \{x : -1 < x < 1\} = (-1, 1).$$

***Example 5.3*** *Consider the series $\sum_{n=0}^{\infty} \frac{x^n}{n!}$. Setting $x$, and denoting the $n$th term by $c_n = \frac{x^n}{n!}$, then*

$$\frac{|c_{n+1}|}{|c_n|} = \frac{|x|}{n+1},$$

*thus, $L = \lim_{n\to\infty} \frac{|c_{n+1}|}{|c_n|} = 0$, and by the ratio test, the series converges for every $x \in \mathbb{R}$, hence $E = \mathbb{R}$. It is often convenient to define the exponential function as*

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}. \tag{5.3}$$

*Taking (5.3) as the definition for $e^x$, then by the above $e^x$ is defined for every $x$. We will later show that $e^x = \lim_{n\to\infty} \left(1 + \frac{x}{n}\right)^n$.*

***Example 5.4*** *Consider the following series* $\sum_{n=0}^{\infty} n\sin(nx)$*. Note that this is not a power series. Clearly, the series converges for any* $x = k\pi$ *where* $k$ *is an integer. It can be verified (but take it as forgranted) that for any choice of* $x$ *not in the form of* $x = k\pi$*, the* $n$*th term does not converge to zero, and in fact, the limit* $\lim_{n\to\infty} n\sin(nx)$ *does not exist at all. Hence*

$$E = \{\pi k : k \in \mathbb{Z}\}\,. \tag{5.4}$$

We will show in the next few sections that given any power series, the set of points for which it converges is always an interval. It may be either closed, open, partially closed or unbounded, but may never take the form of (5.4)

## 5.1 Superior and Inferior limits

In this section we quickly review the definitions of the superior and the inferior limits of a sequence $(a_n)_{n=1}^{\infty}$ in $\mathbb{R}$. The set of partial limits of $(a_n)$ is defined as

$$S = \{s \in \mathbb{R} \cup \{\pm\infty\} : \text{ there exists a subsequence of } a_n \text{which converges to } s\}\,. \tag{5.5}$$

Note that we include $\infty$ and $-\infty$ as legitimate partial limits. The set of partial limits is never empty; when the sequence is bounded, then the B-W theorem states that there exists a converging subsequence of $(a_n)$ (to a finite partial limit). If the sequence is unbounded, then there exists a subsequence converging to either infinity or minus infinity. Thus, every sequence has a partial limit in the extended sense. We now define

***Definition 5.2***

$$\limsup_{n\to\infty} a_n = \sup S \leq \infty \qquad \textit{and} \qquad \liminf_{n\to\infty} a_n = \inf S \geq -\infty.$$

An alternative, commonly used notation is also

$$\overline{\lim_{n\to\infty}}\, a_n = \limsup_{n\to\infty} a_n \qquad \text{and} \qquad \underline{\lim}_{n\to\infty} a_n = \liminf_{n\to\infty} a_n$$

Note that by definition, we always have $\liminf_{n\to\infty} a_n \leq \limsup_{n\to\infty} a_n$. Recall that a sequence $(a_n)$ converges if and only if it has a unique partial limit, namely, that the set $S$ of partial limits has only one element. Thus, a sequence $(a_n)$ converges if and only if $\liminf_{n\to\infty} a_n = \limsup_{n\to\infty} a_n$. The perhaps surprising result about superior and inferior limits is the following theorem, from which we omit the proof,

***Theorem 5.2*** *Let $(a_n)_{n=1}^{\infty}$ be a sequence in $\mathbb{R}$, and let $S$ be the corresponding set of partial limits, given by* (5.5). *Then, there exist subsequences $(a_{n_k})_{k=1}^{\infty}$ and $(a_{m_l})_{l=1}^{\infty}$ of $(a_n)$, such that $\lim_{k\to\infty} a_{n_k} = \sup S$ and $\lim_{l\to\infty} a_{m_l} = \inf S$.*

Repeating the above statement in simpler words, then the set of partial limits $S$ has a both maximum and a minimum. Thus, we may define $\limsup_{n\to\infty} a_n$ as the maximal partial limit, and $\liminf_{n\to\infty} a_n$ as the minimal partial limit of $(a_n)$.

***Example 5.5*** *Let $(a_n)_{n=1}^{\infty}$ be given by $a_n = (-1)^n$. Then it can be verified that $S = \{-1, 1\}$. Thus, $\limsup_{n\to\infty} a_n = 1$ and $\liminf_{n\to\infty} a_n = 1$.*

***Example 5.6*** *Let $(a_n)_{n=1}^{\infty}$ be given by*

$$a_n : 0, 1, -1, 2, 0, -2, 3, 0, -3, \ldots$$

*Then it can be verified that $S = \{-\infty, 0, +\infty\}$, hence $\limsup_{n\to\infty} a_n = \infty$ and $\liminf_{n\to\infty} a_n = -\infty$.*

The next Lemma is given without proof,

***Lemma 5.1*** *Let $(a_n)_{n=1}^{\infty}$ be a sequence, and denote by $L = \limsup_{n\to\infty} a_n$. Then $\forall \varepsilon > 0$, $\exists N \in \mathbb{N}$ such that $\forall n > N$ we have*

$$a_n < L + \varepsilon.$$

In other words, there is only a finite number of elements of $(a_n)$ which satisfy $a_n \geq L + \varepsilon$.

***Theorem 5.3*** *(The root test) Let $\sum_{n=0}^{\infty} a_n$ be a series of numbers and let $L = \limsup_{n\to\infty} \sqrt[n]{|a_n|}$. Then*

1. *If $L < 1$ the series converges.*

2. *If $L > 1$ the series diverges.*

*Proof*:

1. Let $\varepsilon > 0$ be sufficiently small such that $L + \varepsilon < 1$. By Lemma (5.1), there exists a natural $N \in \mathbb{N}$ such that for every $n > N$ we have $\sqrt[n]{|a_n|} < L + \varepsilon$, hence
$$|a_n| < (L + \varepsilon)^n.$$
Denoting by $q = L+\varepsilon$, then the above reads $|a_n| < q^n$. As the geometric series $\sum_{n=N}^{\infty} q^n$ converges, then, by the comparison theorem, the series $\sum_{n=N}^{\infty} |a_n|$ converges, hence the series $\sum_{n=0}^{\infty} a_n$ converges absolutely, which implies that the series converges.

2. Let $\varepsilon > 0$ be sufficiently small such that $L - \varepsilon > 1$. Let $(a_{n_k})_{k=1}^{\infty}$ be a subsequence of $(a_n)$ for which $\sqrt[n_k]{|a_{n_k}|}$ converges to $L$. Then, there exists a $N \in \mathbb{N}$ such that for every $k > N$, we have
$$\sqrt[n_k]{|a_{n_k}|} > L - \varepsilon > 1 \Rightarrow |a_{n_k}| > 1,$$
which implies that the sequence $a_n$ does not converge to zero, hence the series $\sum a_n$ diverges.

■

## 5.2 Radius of convergence

Back to power series, then the main theorem, also known as Hadarmard's Theorem, states the following,

**Theorem 5.4** *(Hadarmard's theorem) Let $\sum_{n=0}^{\infty} a_n(x - x_0)^n$ be a power series about $x_0$. Then there exists an $R \in [0, \infty) \cup \{\infty\}$ such that the series converges for every $x \in \mathbb{R}$ such that $|x - x_0| < R$ and diverges for every $x \in \mathbb{R}$ such that $|x - x_0| > R$. Moreover, $R$ is given by the following formula, (known as Hadarmard's Formula)*

$$R = \frac{1}{\limsup_{n\to\infty} \sqrt[n]{|a_n|}}. \tag{5.6}$$

The above theorem needs to be stated with more details and clarifications. Instead, we clarify below:

- As the superior limit of a sequence always exists, then $R$ is always defined; If $\limsup_{n\to\infty} \sqrt[n]{|a_n|} = 0$ then we define (or use the convention that) $R = \infty$. If $\limsup_{n\to\infty} \sqrt[n]{|a_n|} = \infty$ then we define (or use the convention that) $R = 0$.

- If $R = \infty$ then the theorem states that the series converges for every $x \in \mathbb{R}$. If $R = 0$ then the theorem states that the series converges for only $x = x_0$.

- If the limit $\lim_{n\to\infty} \sqrt[n]{|a_n|}$ exists then $\limsup_{n\to\infty} \sqrt[n]{|a_n|} = \lim_{n\to\infty} \sqrt[n]{|a_n|}$, and the superior limit in (5.6) can be replaced by the limit.

- $R$ is referred to as the **radius of converges** of the power series. Note that it does not depend on $x_0$ but only on the sequences $(a_n)_{n=0}^{\infty}$.

- Note that the theorem states that the power series converges on the open interval $(x_0 - R, x_0 + R)$ and diverges outside of the closed interval $[x_0 - R, x_0 + R]$. It fails to give any information about the boundary points $x_0 - R$ and $x_0 + R$ of the interval. These two points have to be treated separately.

We are now ready to prove Theorem 5.4.

*Proof*: Let $L = \limsup_{n\to\infty} \sqrt[n]{|a_n|} = \frac{1}{R}$ (where, as before, we use the convention that if $R = 0$ implies $L = \infty$ and vise versa.

1. Suppose that $|x - x_0| < R$. Then,

$$\limsup_{n\to\infty} \sqrt[n]{|a_n(x-x_0)^n|} = \limsup_{n\to\infty} \sqrt[n]{|a_n|} \cdot |x - x_0| = L|x - x_0| < 1.$$

   Thus, by the root test, the series $\sum_{n=0}^{\infty} a_n(x - x_0)^n$ converges. Note that if $R = 0$ then the above statement is empty.

2. Suppose that $|x - x_0| > R$. Then,

$$\limsup_{n\to\infty} \sqrt[n]{|a_n(x-x_0)^n|} = \limsup_{n\to\infty} \sqrt[n]{|a_n|} \cdot |x - x_0| = L|x - x_0| > 1.$$

   Thus, by the root test, the series $\sum_{n=0}^{\infty} a_n(x - x_0)^n$ diverges. Not that if $R = \infty$ then the above statement is empty.

■

We proceed with some examples,

***Example 5.7*** *Consider again the geometric series* $\sum_{n=0}^{\infty} x^n$. *We have* $a_n = 1$, *and*

$$\limsup_{n\to\infty} \sqrt[n]{|a_n|} = \lim_{n\to\infty} \sqrt[n]{|a_n|} = 1$$

*hence, the radius of converges is* $R = 1$. *Thus the series converges for* $|x| < 1$ *and diverges for* $|x| > 1$. *It should be again verified that the series diverges on the boundary points* $\pm 1$.

***Example 5.8*** *Consider the geometric series* $\sum_{n=0}^{\infty} \frac{x^n}{2^n}$. *We have* $a_n = \frac{1}{2^n}$, *and*

$$L = \limsup_{n\to\infty} \sqrt[n]{|a_n|} = \lim_{n\to\infty} \sqrt[n]{\frac{1}{2^n}} = \frac{1}{2},$$

*hence, the radius of converges is* $R = \frac{1}{L} = 2$, *and the series converges for* $|x| < 2$ *and diverges for* $|x| > 2$. *We now check the boundary points; if* $x = 2$ *then the series reduces to* $\sum \frac{2^n}{2^n} = \sum 1$, *hence diverges. if* $x = -2$ *then the series reduces to* $\sum \frac{(-1)^n 2^n}{2^n} = \sum (-1)^n$, *hence diverges. Thus, the domain of convergence is* $E = (-2, 2)$.

***Example 5.9*** *Consider the power series* $\sum_{n=1}^{\infty} \frac{x^n}{n}$. *We have* $a_n = \frac{1}{n}$, *hence*

$$\frac{1}{R} = \limsup_{n\to\infty} \sqrt[n]{1/n} = \lim_{n\to\infty} \frac{1}{\sqrt[n]{n}} = 1,$$

*where we have used the fact that* $\lim_{n\to\infty} \sqrt[n]{n} = 1$. *Thus* $R = 1$. *Checking the boundary points, then for* $x = 1$ *the series reduces to the harmonic sum* $\sum \frac{1}{n}$, *hence diverges, while for* $x = -1$ *the series reduces to the alternating series* $\sum \frac{(-1)^n}{n}$, *hence converges. Thus,* $E = [-1, 1)$.

***Example 5.10*** *Consider the power series* $\sum_{n=0}^{\infty} \frac{x^n}{n!}$. *We have* $a_n = \frac{1}{n!}$. *Although estimating the value of* $\limsup \sqrt[n]{|a_n|}$ *is hard (in fact, it requires the use of the so-called Stirling formula), note that for every* $x \in \mathbb{R}$,

$$\lim_{n\to\infty} \frac{|a_{n+1}x^{n+1}|}{|a_n x^n|} = \lim_{n\to\infty} \frac{|x|}{n+1} = 0.$$

*Thus, by the ratio test, the series converges for every $x \in \mathbb{R}$ (i.e, $R = \infty$). It is convenient to define the exponential map as*

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}. \tag{5.7}$$

Let us recall the ratio test for series in its simplest form,

**Theorem 5.5** *(The ratio test) Let $\sum_{n=0}^{\infty} a_n$ be a series in $\mathbb{R}$. Suppose that the following limit exists (or equals to $\infty$),*

$$L = \lim_{n\to\infty} \frac{|a_{n+1}|}{|a_n|}.$$

*Then*

1. *If $L < 1$ the series converges.*
2. *If $L > 1$ the series diverges.*

We remark that the above limit for the ratio does not always exist. It is worth mentioning that the ratio test can be reformulated using limsup and liminf instead of limits, assuring that the value of $L$ is always defined.

**Example 5.11** *Consider the following power series: $\sum_{n=0}^{\infty} a_n x^n$, where $a_n$ is given by*

$$a_n = \begin{cases} k^n & n = 2k \\ 0 & \text{otherwise} \end{cases}$$

*Then $\limsup_{n\to\infty} \sqrt[n]{|a_n|} = \lim_{n\to\infty} \sqrt[n]{k^n} = \lim_{n\to\infty} k = \infty$, and consequently $R = 0$. Thus, the series converges if and only if $x = 0$.*

**Example 5.12** *Consider the following power series: $\sum_{n=0}^{\infty} 2^k x^{2^k}$. Then $\limsup_{n\to\infty} \sqrt[n]{|a_n|} = \lim_{n\to\infty} \sqrt[2^k]{2^k} = 1$, and consequently $R = 1$. It is easy to verify that the series diverges on the boundary points $x = \pm 1$, hence $E = (-1, 1)$. Note that in this case the limit $\lim_{n\to\infty} \sqrt[n]{|a_n|}$ does not exists, as, for example, the subsequence of odd indices $(a_{2k+1})_{k=0}^{\infty}$ is equal to zero.*

## 5.3 differentiation of power series

Let $f(x) = \sum_{n=0}^{\infty} a_n(x-x_0)^n$ be a power series, and denote by $R$ its converges radius. A natural question to ask is whether the function $f$, which is defined on $(x_0 - R, x_0 + R)$ (and possibly on the boundary) is continuous/differentiable, and if so, what is $f'(x)$? Fortunately, the answer to this question is always yes. In fact, every function which has a power series expansion is in $C^{\infty}(x_0 - R, x_0 + R)$. Moreover, the derivative of $f$ is given by a simple rule of term-by-term differentiation. To prove that, we first state a lemma,

***Lemma 5.2*** *Let $\sum_{n=0}^{\infty} a_n(x-x_0)^n$ be a power series with a radius of convergence $R$. Then the series of derivatives, defined by $\sum_{n=0}^{\infty} a_n n(x-x_0)^{n-1}$ has a radius of convergence of $\tilde{R} = R$.*

- Note that the $n$th term of the series of derivatives, $a_n n(x-x_0)^{n-1}$, is equal to the derivative of the $n$th term of the series, $a_n(x-x_0)^n$.
- Note that the series of derivatives can be re-written as

$$\sum_{n=0}^{\infty} a_n n(x-x_0)^{n-1} = \sum_{n=1}^{\infty} a_n n(x-x_0)^{n-1} = \sum_{n=0}^{\infty} a_{n+1}(n+1)(x-x_0)^n.$$

To proof Lemma 5.2, we use the following statement about sequences, from which we omit the proof,

***Lemma 5.3*** *Let $(A_n)_{n=1}^{\infty}$ and $(B_n)_{n=1}^{\infty}$ be two sequences in $\mathbb{R}$. Suppose that $B_n$ converges and $\lim_{n\to\infty} B_n = B > 0$. Then*

$$\limsup_{n\to\infty} (A_n \cdot B_n) = \limsup_{n\to\infty} A_n \cdot B.$$

We now prove Lemma 5.2,

*Proof*: Applying Hadarmard's formula 5.4 to the series of derivatives, with the (notational) conventions that $\frac{1}{0} = \infty$ and $\frac{1}{\infty} = 0$, we obtain

$$\begin{aligned}\frac{1}{\tilde{R}} &= \limsup_{n\to\infty} \sqrt[n]{|a_{n+1}(n+1)|} = \limsup_{n\to\infty} \sqrt[n-1]{|a_n n|} \\ &= \limsup_{n\to\infty} \sqrt[n]{|a_n n|} = \limsup_{n\to\infty} \sqrt[n]{|a_n|} \cdot \lim_{n\to\infty} \sqrt[n]{n} = \frac{1}{R} \cdot 1\end{aligned}$$

where we applied Lemma 5.3 and used the fact that $\sqrt[n]{n} \to 1$. Thus, $R = \tilde{R}$.[13] ■

An straightforward application is the following,

***Corollary 5.1*** *Let $\sum_{n=0}^{\infty} a_n (x - x_0)^n$ be a power series with a radius of convergence $R$. Then the series of double derivatives, defined by $\sum_{n=2}^{\infty} a_n n(n-1)(x - x_0)^{n-2}$ has a radius of convergence of $\tilde{R} = R$.*

*Proof*: Apply Lemma 5.2 twice. ■

We are now ready to prove the main assertion for this section,

***Theorem 5.6*** *(term-by-term differentiation) Let $f(x) = \sum_{n=0}^{\infty} a_n (x - x_0)^n$ be a power series, and denote by $R$ its converges radius. Let $x \in (x_0 - R, x_0 + R)$ [14]. Then $f$ is differentiable at $x$, and*

$$f'(x) = \sum_{n=1}^{\infty} a_n n (x - x_0)^{n-1}. \tag{5.8}$$

Note that by Lemma 5.2, the series of derivatives $\sum_{n=1}^{\infty} a_n n (x - x_0)^{n-1}$ has a radius $R$, hence converges at any $x \in (x_0 - R, x_0 + R)$.

*Proof*: We assume, for notational simplicity, that $x_0 = 0$ (otherwise, replace $x$ by $x - x_0$ and repeat the proof). Let $\epsilon > 0$. We show that there exists a $\delta > 0$, depending on $\varepsilon$ (and $x$) such that for any $y$ such that $|y - x| < \delta$ we have

$$\left| \frac{f(y) - f(x)}{y - x} - \sum_{n=1}^{\infty} a_n n x^{n-1} \right| < \varepsilon.^{15}$$

[13] The assertion that $\limsup_{n\to\infty} \sqrt[n-1]{|a_n n|} = \limsup_{n\to\infty} \sqrt[n]{|a_n n|}$ has to be justified thoroughly, but we omit the details. This follows from the fact that the function $x^y$, where ever defined, is a continuous function of $x$ and $y$ (where ever $(x, y) \neq (0, 0)$), or from equivalent algebraic properties of sequences.

[14] If $R = 0$ then this statement is empty. If you are bothered by that, you may add the assumption that $R > 0$. Note that in which case $f$ is defined on $x$, which lays within the radius of convergence

[15] Note that this is simply the definition for $\lim_{y\to x} \frac{f(y)-f(x)}{y-x} = \sum_{n=1}^{\infty} a_n n x^{n-1}$

Substituting the expansion for $f$ as a power series, we obtain

$$\left|\frac{f(y)-f(x)}{y-x}-\sum_{n=1}^{\infty}a_nnx^{n-1}\right| = \left|\frac{\sum_{n=0}^{\infty}a_ny^n-\sum_{n=0}^{\infty}a_nx^n}{y-x}-\sum_{n=1}^{\infty}a_nnx^{n-1}\right|$$
$$= \left|\frac{\sum_{n=0}^{\infty}a_n(y^n-x^n)}{y-x}-\sum_{n=1}^{\infty}a_nnx^{n-1}\right| = \left|\sum_{n=0}^{\infty}a_n\frac{y^n-x^n}{y-x}-\sum_{n=1}^{\infty}a_nnx^{n-1}\right|$$
$$= \left|\sum_{n=1}^{\infty}a_n\frac{y^n-x^n}{y-x}-\sum_{n=1}^{\infty}a_nnx^{n-1}\right| = \left|\sum_{n=1}^{\infty}a_n\left[\frac{y^n-x^n}{y-x}-nx^{n-1}\right]\right|.$$

By the Mean-Value Theorem, for every $n$ there exists a point $c_n$ between $x$ and $y$ such that

$$\frac{y^n-x^n}{y-x}=nc_n^{n-1}.$$

Thus,

$$\left|\sum_{n=1}^{\infty}a_n\left[\frac{y^n-x^n}{y-x}-nx^{n-1}\right]\right| = \left|\sum_{n=1}^{\infty}a_nn\left(c_n^{n-1}-x^{n-1}\right)\right|.$$

Applying again the Mean-Value Theorem for the points $x$ and $c_n$, then there exists a point $d_n$ between $x$ and $c_n$ such that

$$\frac{c_n^{n-1}-x^{n-1}}{c_n-x}=(n-1)d_n^{n-2}.$$

Thus,

$$\left|\sum_{n=1}^{\infty}a_nn\left(c_n^{n-1}-x^{n-1}\right)\right| = \left|\sum_{n=1}^{\infty}a_nn(n-1)d_n^{n-2}(c_n-x)\right|,$$

and by the Triangle inequality we have

$$\left|\sum_{n=1}^{\infty}a_nn(n-1)d_n^{n-2}(c_n-x)\right| \le \sum_{n=1}^{\infty}|a_n|n(n-1)|d_n|^{n-2}\cdot|c_n-x|.$$

Note that $|c_n-x|<\delta$ and that $|d_n|<r$, thus,

$$\sum_{n=1}^{\infty}|a_n|n(n-1)|d_n|^{n-2}\cdot|c_n-x|\le\delta\cdot\alpha,$$

where

$$\alpha = \sum_{n=1}^{\infty} |a_n| n(n-1) r^{n-2}$$

is defined as the right-hand side converges; Indeed, the series $\sum_{n=0}^{\infty} |a_n| x^n$ has a radius of convergence which is equal to $R$. Consequently, Corollary 5.1 implies that the series $\sum_{n=1}^{\infty} |a_n| n(n-1) r^{n-2}$ converges for $r \in (-R, R)$. ■

If $f(x) = \sum_{n=0}^{\infty} a_n (x-x_0)^n$ is a power series with a radius $R$, then $f(x_0) = a_0$. If $R > 0$, then we can use the above theorem, and by plugging $x = x_0$ in (5.8), we obtain that $f'(x_0) = a_1$. Similarly, by the above theorem we have that

$$f''(x) = \sum_{n=2}^{\infty} a_n n(n-1)(x-x_0)^{n-2}.$$

Plugging $x = x_0$ into the the above equation for $f''(x)$ yields $f''(x_0) = a_2 \cdot 2$, or $a_2 = \frac{f''(x_0)}{2}$. Similarly, by induction we obtain that

$$a_n = \frac{f^{(n)}(x_0)}{n!}. \tag{5.9}$$

In particular, we arrive to the following conclusion:

***Corollary 5.2*** *(uniqueness of power series expansion) Let $f(x) = \sum_{n=0}^{\infty} a_n (x-x_0)^n$ be a power series, with a radius $R > 0$. Then, $f \in C^{\infty}(x_0 - R, x_0 + R)$, and $a_n = \frac{f^{(n)}(x_0)}{n!}$. In particular, if $\sum_{n=0}^{\infty} a_n (x-x_0)^n = \sum_{n=0}^{\infty} b_n (x-x_0)^n$ (in some open interval centred at $x_0$), then $a_n = b_n$ for every $n$.*

***Example 5.13*** *Find the radius of convergence of the following series $\sum_{n=1}^{\infty} \frac{x^n}{n}$. Find an explicit formula for $f(x) = \sum_{n=1}^{\infty} \frac{x^n}{n}$ for every $x \in (-R, R)$.*

**Solution** As we have already seen in example 5.10, the radius is $R = 1$. For every $x \in (-1, 1)$ we have by Theorem 5.6 that $f$ is differentiable at $x$ and

$$f'(x) = \sum_{n=1}^{\infty} n \frac{x^{n-1}}{n} = \sum_{n=1}^{\infty} x^{n-1} = \sum_{n=0}^{\infty} x^n.$$

The above geometric series converges for every $|x| < 1$ to the function $\frac{1}{1-x}$, thus,

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

which implies that $f(x) = -\ln(1-x)+C$ for some constant $C$. Substituting $x = 0$ into the series, we obtain that $C = 0$, i.e, $f(x) = -\ln(1-x)$ for any $x \in (-1, 1)$.
We end this section with a useful observation,

***Proposition 5.1*** *Let $\sum_{n=0}^{\infty} a_n(x-x_0)^n$ be a power series with radius $R$, and suppose that the following limit exists*

$$L = \lim_{n\to\infty} \frac{|a_{n+1}|}{|a_n|}.$$

*Then $L = \frac{1}{R}$, with the agreement that $L = 0$ implies $R = \infty$ and vice versa.*

*Proof*: Denote by $\tilde{R} = \frac{1}{L}$, and we show that $\tilde{R} = R$. We apply the ratio test 5.5 for the series $\sum_{n=0}^{\infty} a_n(x-x_0)^n$ at a given point $x \in \mathbb{R}$,

$$\lim_{n\to\infty} \frac{|a_{n+1}(x-x_0)^{n+1}|}{|a_n(x-x_0)^n|} = L \cdot |x-x_0| \frac{|x-x_0|}{\tilde{R}}.$$

Thus, by the ratio test, the series converges for $|x-x_0| < \tilde{R}$ and diverges for $|x-x_0| > \tilde{R}$. Thus, $\tilde{R}$ must be the radius of converges, i.e, $\tilde{R} = R$. ■

## 5.4 The exponential function

We defined in (20) the exponential function as

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}, \qquad x \in \mathbb{R}.$$

Note that, as we have seen, the exponential function is defined for every $x \in \mathbb{R}$. We give some basic properties of the exponential function in the following statement

***Proposition 5.2*** *The exponential function satisfies the following properties:*

1. *$e^x$ is differentiable and $(e^x)' = e^x$.*
2. *$e^0 = 1$.*
3. *$e^{x+y} = e^x \cdot e^y$ for every $x, y \in \mathbb{R}$.*

4. $e^{-x} = \frac{1}{e^x}$.

5. $e^x > 0$ *for every* $x \in \mathbb{R}$.

6. $\lim_{x\to\infty} e^x = \infty$ *and* $\lim_{x\to-\infty} e^x = 0$.

7. *The function* $e^x$ *is strictly monotonically increasing in* $\mathbb{R}$.

*Proof*:

1. By Theorem 5.6, $e^x$ is differentiable and

$$
\begin{aligned}
(e^x)' &= \sum_{n=1}^{\infty} n\frac{x^{n-1}}{n!} = \sum_{n=1}^{\infty} \frac{x^{n-1}}{(n-1)!} \\
&= \sum_{n=0}^{\infty} \frac{x^n}{n!} = e^x.
\end{aligned}
$$

2. Immediate.

3. Fix $y \in \mathbb{R}$ and set $f(x) = e^{x+y}$, then $f'(x) = f(x)$. Thus, by the rules of derivation, the function $e^{-x} \cdot f(x)$ is differentiable and

$$\left(e^{-x} f(x)\right)' = -e^{-x} \cdot f(x) + e^{-x} f'(x) = -e^{-x} \cdot f(x) + e^{-x} f(x) = 0,$$

which implies that $e^{-x} f(x)$ is constant, hence

$$e^{-x} f(x) = e^0 f(0) = e^y,$$

i.e,

$$f(x) = e^x e^y.$$

4. Applying the above formula for $y = -x$ we obtain

$$1 = e^0 = e^{x-x} = e^x \cdot e^{-x},$$

i.e,

$$e^{-x} = \frac{1}{e^x}. \tag{5.10}$$

5. If $x \geq 0$ then all the terms in the series are positive, hence $e^x > 0$. If $x < 0$ then the equation $e^x \cdot e^{-x} = 1$ implies that $e^x$ and $e^{-x}$ must share the same signature.

6. Truncating the series at $n = 1$, we get for every $x \geq 0$ that

$$e^x \geq 1 + x, \qquad x \geq 0,$$

which yields that $\lim_{x\to\infty} e^x \geq \lim_{x\to\infty}(1+x) = \infty$. Using the formula (5.10) we obtain that

$$\lim_{x\to-\infty} e^x = \lim_{x\to\infty} e^{-x} = \lim_{x\to\infty} \frac{1}{e^x} = 0.$$

7. Follows directly from $(e^x)' = e^x > 0$ for every $x \in \mathbb{R}$, and from Example 4.17.

■

As $e^x$ is strictly increasing, then it is injective. As $e^x$ is continuous and as $\lim_{x\to-\infty} e^x = 0$, and $\lim_{x\to\infty} e^x = \infty$, then we have $\operatorname{Im} e^x = (0, \infty)$. The inverse function, $\ln x : (0, \infty) \to \mathbb{R}$ (which is sometimes denoted by $\log x$) is well defined. We have seen that $(\ln x)' = \frac{1}{x}$ for every $x > 0$ (see Example 4.16), hence $\ln x$ is strictly monotonically increasing in $(0, \infty)$. The fact that $\lim_{x\to\infty} \ln x = \infty$ and $\lim_{x\to 0+} \ln x = -\infty$ follows directly from the corresponding properties $\lim_{x\to\infty} e^x = \infty$ and $\lim_{x\to-\infty} e^x = 0$ (see Figure 20).

# 6 Taylor series

As soon after we give a proper definition for a Taylor series, you will note that a Taylor series is power series, and conversely, every power series is a Taylor series. Consequently, a Taylor series is just a different terminology for a power series. There are two different ways to address a power series; one is via the elements of the series, i.e, $a_n \cdot (x - x_0)^n$, like we have done in the previous section. The second alternative is to approach its limit, defined by a function $f(x)$. Consequently, we may assign a given function $f$ the following question: does there exist a power series $\sum_{n=0}^{\infty} a_n(x-x_0)^n$ such that $f(x) = \sum_{n=0}^{\infty} a_n(x-x_0)^n$, where $x$ lies in some domain of convergence? sadly, or gladly (depending on how you look at life) the answer to this question is no. When $f$ has a power series expansion in some domain, we say that $f$ is **analytic** (in this domain). Obviously, if $f$ is not a $C^\infty$ function in this domain, then it cannot be analytic. Moreover, there exists a function $f$ that

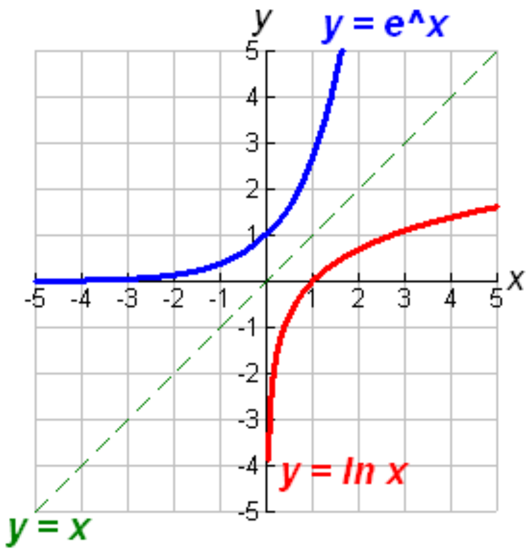



Figure 20: The graph of $e^x$ versus the graph of $\ln x$.

is $C^\infty(\mathbb{R})$ but not analytic. The set of elementary functions, however, such as $e^x$ and $\sin x$ are analytic. We make all of these assertions more precise below,

***Definition 6.1*** *Let $f : E \to \mathbb{R}$ be a real-valued function, and let $x_0 \in E$. Let $n \in \mathbb{N} \cup 0$. Suppose that $f$ is differentiable $n$ times at $x_0$. The* ***Taylor polynomial*** *of $f$, of order $n$ about $x_0$ is defined as*

$$T_n(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2!}(x - x_0)^2 + \cdots + \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n.$$

*Using the conventions that $0! = 1$ and $f^{(0)}(x) = f(x)$, we may rewrite this polynomial as*

$$T_n(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k. \tag{6.1}$$

***Definition 6.2*** *Let $f : E \to \mathbb{R}$ be a real-valued function, and let $x_0 \in E$. Let $n \in \mathbb{N} \cup 0$. Suppose that $f$ is differentiable $n$ times at $x_0$. The* ***Taylor***


Figure 21: Searching for Brook Taylor's photo on google, I have found two possible candidates ;-)

***remainder*** *of $f$, of order $n$ about $x_0$ is defined as*

$$R_n(x) = f(x) - T_n(x). \tag{6.2}$$

*Remarks*

- $T_n(x)$ is a polynomial of degree $\leq n$ (note that the last coefficient $\frac{f^{(n)}(x_0)}{n!}$ may be zero).
- Note that by the uniqueness argument (5.9), we have $T^{(k)}(x_0) = f^{(k)}(x_0)$ for every $0 \leq k \leq n$. Therefore, collecting the data of $f$ at $x_0$, $T_n(x)$ can be thought as reasonable estimate for $f(x)$ at nearby points $x$, by a polynomial of degree $\leq n$ (see Figure 22), and $R_n(x)$ as the error of this estimate.

***Definition 6.3*** *Let $f : E \to \mathbb{R}$ be a real-valued function, and let $x_0 \in E$. Suppose that $f$ is differentiable $n$ times at $x_0$ for every $n$ (i.e, $f$ is*

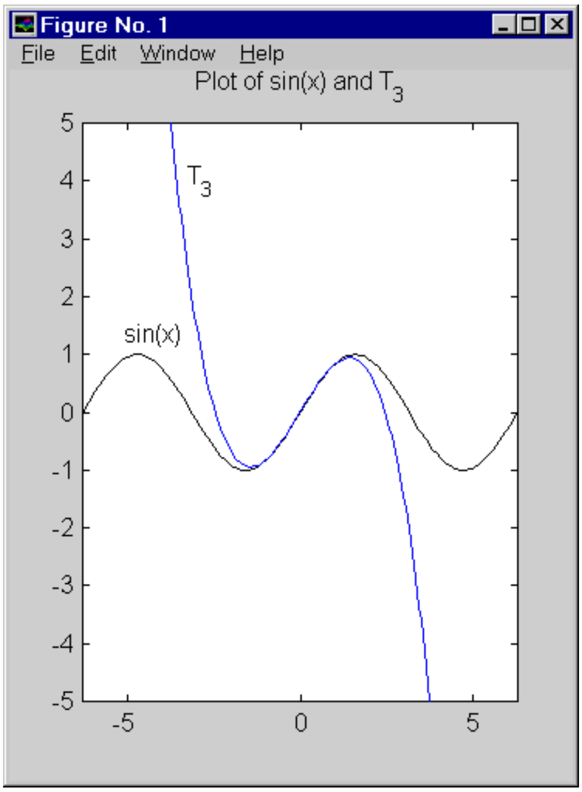



Figure 22: The Taylor polynomial $T_n(x)$ is an estimate for $f(x)$. As $x$ tends to $x_0$ the precision increases. Thank you Matlab ;-)

*differentiable $\infty$ number of times at $x_0$). The **Taylor series** of $f$ about $x_0$ is defined as*

$$T_f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n. \tag{6.3}$$

*Remarks*

- The Taylor series may only be defined for functions which are differentiable $\infty$ number of times at $x_0$.
- The Taylor series is a power series.
- Note that $T_f(x) = f(x)$ if and only if $\lim_{n\to\infty} T_n(x) = f(x)$, or $\lim_{n\to\infty} R_n(x) = 0$.
- If $f$ is a function defined by a power series, $f(x) = \sum_{n=0}^{\infty} a_n(x - x_0)^n$, within some radius of convergence $R > 0$, then by the uniqueness argument (5.9) we have that $f$ is differentiable $\infty$ number of times at $x_0$ and $T_f(x) = f(x)$. Therefore, every power series is a taylor series of the limit function $f(x)$ defined within the radius of convergence.

***Definition 6.4*** *A function $f$ is said to be **real analytic** (or just analytic) in some interval $I = (x_0 - R, x_0 + R)$ where $R > 0$, if $f \in C^\infty(x_0 - R, x_0 + R)$ and $f(x) = T_f(x)$ for all $x \in I$. If the latter holds for every $x \in \mathbb{R}$, we say that $f$ is analytic in $\mathbb{R}$.*

As mentioned at the beginning of the section, not every function is analytic. Clearly, if $f$ is not $C^\infty$ then it cannot be analytic (within the corresponding domain). But what if $f \in C^\infty(x_0 - R, x_0 + R)$? lets postpone this question to the end of this section. Instead, lets start with some simple examples of analytic functions and their Taylor series/polynomials,

***Example 6.1*** *Take $f(x) = e^x$ and $x_0 = 0$. Clearly, $f^{(n)}(0) = 1$ for all $n$ and thus,*

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

*and*

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

*Note that $T(x)$ is exactly the definition for $e^x$, hence $f(x) = T(x)$ for every $x \in \mathbb{R}$ by definition, hence the exponential function is analytic in $\mathbb{R}$.*

**Example 6.2** *Take $f(x) = \frac{1}{1-x}$ and $x_0 = 0$. Recall that*

$$f(x) = \sum_{n=0}^{\infty} x^n,$$

*where the equality holds if and only if $|x| < 1$. Thus, by the uniqueness of Taylor expansions, we have $f(x) = T_f(x) = \sum_{n=0}^{\infty} x^n$ for every $|x| < 1$. Thus $f$ is analytic in $(-1, 1)$. But what if we wanted to take a different $x_0$? well, for $x_0 = 1$ we have a problem; $f$ is not $C^\infty$ at any interval containing $x_0$ (it is not even defined there). If we take $x_0 = -1$ instead, then we can apply the following trick,*

$$f(x) = \frac{1}{1-x} = \frac{1}{2-y},$$

*where $y = x + 1$. We further develop,*

$$\frac{1}{2-y} = \frac{1}{2(1-z)} = \frac{1}{2}\sum_{n=0}^{\infty} z^n, \tag{6.4}$$

*where $z = y/2$. Thus,*

$$f(x) = \frac{1}{2}\sum_{n=0}^{\infty} z^n = \frac{1}{2}\sum_{n=0}^{\infty} \frac{y^n}{2^n} = \sum_{n=0}^{\infty} \frac{(x+1)^n}{2^{n+1}},$$

*and by the uniqueness argument, $T(x) = \sum_{n=0}^{\infty} \frac{(x+1)^n}{2^{n+1}}$. For which values of $x$ do we have $T(x) = f(x)$? well, equality (6.4) holds if and only if $|z| < 1$, and since $(x+1)/2 = z$, this remains valid if and only if $|x+1| < 2$, i.e, for $x \in (-3, 1)$. Thus, we have shown that $f$ is analytic in the interval $(-3, 1)$.*

**Example 6.3** *Take $f(x) = \frac{1}{1+x^2}$ and $x_0 = 0$. Denoting by $z = -x^2$ we have*

$$f(x) = \frac{1}{1-z} = \sum_{n=0}^{\infty} z^n = \sum_{n=0}^{\infty} (-1)^n x^{2n},$$

*where ever the series converges, i.e, $|z| < 1 \iff |x| < 1$.*

***Example 6.4*** *Take* $f(x) = \sin x$ *and* $x_0 = 0$*. It is simply a matter of induction to show that*

$$f^{(n)}(x) = \begin{cases} \sin x & n = 4k \\ \cos x & n = 4k+1 \\ -\sin x & n = 4k+2 \\ -\cos x & n = 4k+3 \end{cases}$$

*Thus,*

$$f^{(n)}(0) = \begin{cases} 0 & n = 2k \\ (-1)^k & n = 2k+1 \end{cases}$$

*and*

$$T_f(x) = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!} x^{2k+1} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \dots$$

*Taking* $f(x) = \cos x$ *and* $x_0 = 0$ *we get (by finding a corresponding formula for* $f^{(n)}(0)$*)*

$$T_f(x) = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!} x^{2k} = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \dots$$

But are $\sin x$ and $\cos x$ analytic functions? we need to show that

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \dots$$
$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \dots,$$

or, alternatively, that the Taylor remainders converge to zero for every $x$. In order to show that, we need a formula for the Taylor remainder. There are many different ways of writing a formula for $R_n(x)$, the most common form is named after Lagrange,

***Theorem 6.1*** *(Taylor's theorem) Let* $x_0 < x$ *be two different points and let* $f$ *be a function defined on the interval* $[x_0, x]$ *such that*

1. $f \in C^n[x_0, x]$
2. $f \in C^{n+1}(x_0, x)$.

*Then, there exists a point $c \in (x_0, x)$ such that*

$$R_n(x) = \frac{f^{n+1}(c)}{(n+1)!} \cdot (x - x_0)^{n+1} \tag{6.5}$$

- Formula (6.5) is known as **Lagrange's formula** for the Taylor remainder of $f$.

- Taylor's theorem can be similarly stated if $x < x_0$, and the same formula holds.

- Of course, if $f$ happens to be in $C^\infty[x_0, x]$ , or even in $C^{n+1}[x_0, x]$, then $f$ satisfies the requirements 1 and 2 for Lagrange's formula to hold for $R_n(x)$.

- Lets see what does Lagrange's formula imply for $n = 0$; the Taylor polynomial, which is of degree zero, satisfies $T_0(x) = f(x_0)$. The Taylor remainder takes the form

  $$R_0(x) = f'(c) \cdot (x - x_0)$$

  for some $c$ between $x$ and $x_0$. Recalling the definition, that $R_0(x) = f(x) - T_0(x)$, we obtain the Mean-Value theorem, (4.5). Note also that the assumptions on $f$ are also the same as assumptions stated by the Mean-Value theorem. Thus, Taylor's theorem is a generalisation (for higher derivatives) of the Mean-Value theorem. In particular, where ever the Mean-Value theorem may be applied, Taylor's theorem may be applied instead. So why did we state and prove the Mean-Value theorem, rather than stating Taylor's theorem instead? well, not surprisingly, the Mean-Value theorem (actually, in Cauchy's form 4.7) is applied in the proof of Taylor's theorem, (and here is the proof)

*Proof*: Fix $x > x_0$, and consider the function

$$\Phi(z) = f(x) - f(z) - f'(z)(x - z) - \frac{f''(z)}{2!}(x - z)^2 - \cdots - \frac{f^{(n)}(z)}{n!}(x - z)^n,$$

defined for every $x_0 \leq z \leq x$. Note that $\phi(x_0) = R_n(x)$, and that $\phi(x) = 0$. Moreover, $\Phi$ is differentiable for every $x_0 < z < x$, and

$$\begin{aligned}\Phi'(z) &= -f'(z) - [f''(z)(x-z) - f'(z)] - \left[\frac{f'''(z)}{2!} - f''(z)(x-z)\right] \\ &\quad - \cdots - \left[\frac{f^{(n+1)}(z)}{n!}(x-z)^n - \frac{f^{(n)}(z)}{(n-1)!}(x-z)^{n-1}\right] \\ &= -\frac{f^{(n+1)}(z)}{n!}(x-z)^n.\end{aligned} \tag{6.6}$$

Now, let $\psi$ be any function defined on $[x_0, x]$, differentiable on $(x_0, x)$ with a non-vanishing derivative. By Cauchy's mean-value theorem 4.7, there exists a point $x_0 < c < x$ such that

$$\frac{\Phi(x) - \Phi(x_0)}{\Psi(x) - \Psi(x_0)} = \frac{\Phi'(c)}{\Psi'(c)}.$$

Substituting (6.6), we get

$$R_n(x) = \frac{\Psi(x) - \Psi(x_0)}{\Psi'(c)} \cdot \frac{f^{(n+1)}(c)}{n!}(x-c)^n. \tag{6.7}$$

Now take $\Psi(z) = (x-z)^{n+1}$, then the above reduces to

$$R_n(x) = \frac{-(x-x_0)^{n+1}}{-(n+1)(x-c)^n} \frac{f^{(n+1)}(c)}{n!}(x-c)^n = \frac{f^{n+1}(c)}{(n+1)!} \cdot (x-x_0)^{n+1}.$$

■

Formula (6.7) is known as the general formula for the Taylor remainder. If, for example, instead of taking $\Psi(z) = (x-z)^{n+1}$ as in the above proof, we take $\Psi(z) = (x-z)$ and plug it into (6.7), we get the so called **Cauchy's formula** for the Taylor remainder,

$$R_n(x) = \frac{f^{n+1}(c)}{n!} \cdot (x-c)^n(x-x_0).$$

We continue with some applications of Taylor's theorem,

***Example 6.5*** *Estimate* $e^2$ *by a precision of* $10^{-2}$.

**Solution** Set $f(x) = e^x$ and $x_0 = 0$. We find an $n \geq 0$ such that the Taylor remainder of $f$ about $x_0 = 0$ satisfies $|R_n(2)| < 10^{-2}$, implying that $e^2 \approx T_n(2)$, with an error smaller than $10^{-2}$. By (6.5), there exists a point $0 < c < 2$ such that

$$R_n(2) = \frac{e^c}{(n+1)!} 2^{n+1},$$

hence, as $e < 3$ [16] and $c < 2$ we have

$$|R_n(2)| \leq \frac{9 \cdot 2^{n+1}}{(n+1)!}.$$

Note that the right-hand side converges to zero, as it is the $(n+1)$ term of the following converging series

$$\sum_{n=0}^{\infty} \frac{2^n}{n!},$$

which converges by the way to $e^2$. Thus, there exists an $n$ such that $9 \cdot 2^{n+1}(n+1)! < 10^{-2}$. Taking, for example, $n = 9$, then $10! = 3,608,800 > 3,000,000$, and

$$|R_n(2)| \leq \frac{9 \cdot 2^{n+1}}{(n+1)!} < \frac{10^4}{3 \cdot 10^6} = \frac{1}{3} 10^{-2} < 10^{-2}.$$

Thus,

$$e^2 \approx T_9(2) = \sum_{k=0}^{9} \frac{2^k}{k!} = 1 + 2 + \frac{2^2}{2!} + \cdots + \frac{2^9}{9!}.$$

A natural question is which value of $x_0$ should one take to estimate a given function $f$ at a point $x$. On one hand, one should prefer a point $x_0$ which is as close as possible to $x$, resulting in smaller remainder, and hence in a better precision. On the other hand, one should prefer to pick a point $x_0$ for which all of the information (i.e, the derivatives) is known analytically (or precisely). For example, taking $x_0 = 1$ and $f(x) = e^x$ results in the Taylor series

$$T(x) = \sum_{n=0}^{\infty} \frac{e \cdot (x-1)^n}{n!},$$

[16] this can be seen from the definition, as $e = \sum_{n=0}^{\infty} \frac{1}{n!} < 1 + \sum_{n=0}^{\infty} \frac{1}{2^n} = 3$.

and we have that $e^2 = T(2)$. Of course, one cannot estimate $T(2)$ without an a priori estimate for the number $e$. The advantage of picking $x_0 = 0$ in the previous example is that the calculation of the Taylor polynomial does not require any a priori estimate of $e$.

***Example 6.6*** *Estimate* $\ln(2)$ *by a precision of* $10^{-4}$.

**Solution** Set $f(x) = \ln(1+x)$. We wish to estimate $f(1)$. Set $x_0 = 0$. [17] To use the formula for the Taylor remainder, we need to obtain an expression for $f^{(n)}(x)$. Note that

$$f'(x) = \frac{1}{1+x} = (1+x)^{-1}, \qquad f''(x) = -(1+x)^{-2}$$
$$f'''(x) = (-1)\cdot(-2)\cdot(1+x)^{-3}, \dots$$

A simple induction reveals that

$$f^{(n)}(x) = (-1)\cdot(-2)\cdot\ldots(-(n-1))\cdot(1+x)^{-n} = (-1)^{n-1}(n-1)!(1+x)^{-n}.$$

Thus, by formula (6.5), for every $x > 0$ there exists a point $0 < c < x$ such that

$$|R_n(1)| = \left|\frac{f^{(n+1)}(c)}{(n+1)!}\right| \le \frac{1}{n+1},$$

and in particular, $R_n(1) \to 0$ as $n \to \infty$. For the remainder to be smaller than $10^{-4}$, we require that $n > 10^4$. Therefore, to obtain the first 5 digits of $\ln 2$, we should calculate $T_{10{,}000}(1)$. Note that $\frac{f^{(n)}(0)}{n!} = \frac{(-1)^{n-1}}{n}$, thus

$$T_n(x) = \sum_{k=1}^{n} \frac{(-1)^{k-1}}{k} x^k,$$

and therefore

$$\ln 2 \approx T_{10{,}000}(1) = \sum_{k=1}^{10{,}000} \frac{(-1)^{k-1}}{k}.$$

In particular, note that as $R_n(1) \to 0$ we have

$$\ln 2 = T(1) = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n}.$$

[17] Of course, this is just for convenience. We may alternatively take $f(x) = \ln x$ and $x_0 = 1$.

[18] One may think that calculating $T_{10,000}$ in a pocket calculator may take a while. That may have been the case if we wanted to calculate the first $10^6$ digits of $\ln 2$. Well, then we propose the following alternative for the calculation of $\ln 2$; recall that $\ln(xy) = \ln x + \ln y$ for any $x, y > 0$, and that $\ln 1 = 0$, hence $\ln 2 = -\ln \frac{1}{2}$. Setting $x = \frac{1}{2}$, $x_0 = 0$ and $f(x) = \ln(1+x)$ we obtain, by Taylor's theorem, that there exists a point $\frac{1}{2} < c < 1$ such that

$$|R_n(1)| = \left|\frac{f^{(n+1)}(c)}{(n+1)!}\right| \cdot \frac{1}{2^{n+1}} \leq \frac{1}{(n+1)2^{n+1}}.$$

For the remainder to be smaller than $10^{-4}$ it is now sufficient to take $n = 10$, and

$$\ln 2 = -\ln\frac{1}{2} \approx -T_{10}\left(\frac{1}{2}\right) = -\sum_{k=1}^{10} \frac{(-1)^{k-1}}{k}\frac{1}{2^k} = \sum_{k=1}^{10} \frac{(-1)^k}{k}\frac{1}{2^k}.$$

***Proposition 6.1*** *Let $f \in C^\infty[a,b]$, and suppose that there exists a constant $M \geq 0$ (which may depend on $[a,b]$) such that $|f^{(n)}(x)| \leq M$ for every $x \in [a,b]$. Then, $\lim_{n\to\infty} R_n(x) = 0$ for every $x \in [a,b]$.* [19]

*Proof*: Setting $x \in [a,b]$ and $x_0 = a$, then by Taylor's theorem, there exists a point $c$ between $a$ and $x$ such that

$$|R_n(x)| = \left|\frac{f^{(n+1)}(c)}{(n+1)!} \cdot (x-a)^{n+1}\right| \leq M\frac{|x-a|^{n+1}}{(n+1)!}$$

To show that $\lim_{n\to\infty} R_n(x) = 0$ it is sufficient to show (by the squeeze rule for sequences) that $\lim_{n\to\infty} \frac{|x-a|^n}{n!} = 0$. But this sequence is the $n$th terms of the series $\sum_{n=0}^\infty \frac{|x-a|^n}{n!}$ which converges (to the value of $e^{|x-a|}$). Finally, the $n$th term of a converging series must converge to zero. ■

[18] Recall that by (5.13) we already know that

$$\ln(1+x) = \sum_{n=1}^\infty \frac{(-1)^{n-1}x^n}{n},$$

whenever $x \in (-1, 1]$.

[19] or, in other words, $f$ is analytic in $[a,b]$.

*Corollary 6.1*

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \dots$$
$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \dots$$

***Example 6.7*** *Compute the following limit (if exists)*

$$\lim_{x\to 0} \frac{e^x - 1 - x}{x^2}.$$

**Solution** Note that both the denominator and the numerator converge to zero as $x \to 0$. Set $f(x) = e^x$ and $x_0 = 0$. The Taylor polynomial of $f$ of order 2 about $x_0$ is $T_3(x) = 1 + x + \frac{x^2}{2}$. Hence,

$$\lim_{x\to 0} \frac{e^x - 1 - x}{x^2} = \lim_{x\to 0} \frac{\frac{x^2}{2} + R_2(x)}{x^2} = \frac{1}{2} + \lim_{x\to 0} \frac{R_2(x)}{x^2}$$

By Lagrange's formula, there exists a point $c$ between $x$ and zero such that

$$|R_2(x)| = \left| \frac{e^c}{3!} x^3 \right| < \frac{e^x}{3!} |x|^3,$$

and, in particular,

$$0 \le \frac{|R_2(x)|}{x^2} \le \frac{e^x}{3!} x.$$

By the Squeeze rule, we have $\lim_{x\to 0} \frac{|R_2(x)|}{x^2} = 0$, hence $\lim_{x\to 0} \frac{R_2(x)}{x^2} = 0$, and

$$\lim_{x\to 0} \frac{e^x - 1 - x}{x^2} = \frac{1}{2}.$$

The above trick is not only specific to one example; By Lagrange's formula, it is always true that if $f^{(n+1)}(x)$ is bounded on $[x_0 - \delta, x_0 + \delta]$, then

$$\lim_{x\to x_0} \frac{R_n(x)}{(x - x_0)^n} = \lim_{x\to x_0} \left[ \frac{f^{(n+1)}(c)}{(n+1)!} (x - x_0) \right] = 0.$$

It only remains to decide which order to take for the Taylor polynomial; the order should be sufficiently large so that $R_n(x)/(x - x_0)^n$ converges to zero.

***Example 6.8*** *Compute the following limit (if exists)*

$$\lim_{x\to 0}\frac{\sin x - x}{x^3}.$$

**Solution** Set $f(x)=\sin x$, $x_0=0$ and $n=3$. Then $\sin x = x-\frac{x^3}{6}+R_3(x)$, thus

$$\lim_{x\to 0}\frac{\sin x - x}{x^3} = -\frac{1}{6}+\lim_{x\to 0}\frac{R_3(x)}{x^3} = -\frac{1}{6}.$$

***Example 6.9*** *Compute the following limit (if exists)*

$$\lim_{n\to\infty} n\ln\left(1+\frac{1}{n}\right).$$

**Solution** Set $f(x)=\ln(1+x)$ and $x_0=0$. By Example 5.13, then $\ln(1+x)=x-\frac{x^2}{2}+\cdots = x+R_1(x)$ and the equality is true for any $x\in(-1,1]$. Denoting by $y=\frac{1}{x}$ we obtain

$$\lim_{x\to\infty} x\ln\left(1+\frac{1}{x}\right) = \lim_{y\to 0+}\frac{1}{y}\ln(1+y) = \lim_{y\to 0+}\left[1+\frac{R_1(y)}{y}\right] = 1.$$

Finally, setting $x_n = n$ and using the sequential definition for limits, we obtain

$$\lim_{n\to\infty} n\ln\left(1+\frac{1}{n}\right) = \lim_{x\to\infty} x\ln\left(1+\frac{1}{x}\right) = 1.$$

We finish this section with a fundamental example of a non-analytic function,

***Example 6.10*** *Consider the following function,*

$$f(x)=\begin{cases}\exp\left(-\frac{1}{x}\right) & x>0\\ 0 & x\le 0\end{cases}$$

*It can be verified that $f$ is a $C^\infty(\mathbb{R})$ function, and that $f^{(n)}(0)=0$ for every $n$. In particular, $T_f(x)$, the Taylor series of $f$ about $x_0=0$, satisfies $T_f(x)=0$ for every $x$. However, $f(x)>0$ for every $x>0$, which implies that $f$ cannot be analytic in any open interval containing $x_0=0$.*

The nature of the exponential causes the function to look 'flatter' at zero then any polynomial $x^n$, yet while remaining positive for any $x > 0$ (see Figure 23). Another known example of a non-analytic function is

$$g(x) = \begin{cases} \exp\left(-\frac{1}{x^2}\right) & x \neq 0 \\ 0 & x = 0 \end{cases} \tag{6.8}$$

which also satisfies the same properties, that is, $g$ is $C^\infty$ while $T_g(x) = 0$ for every $x$, hence $g$ is non analytic. You will prove that in your home assignment.

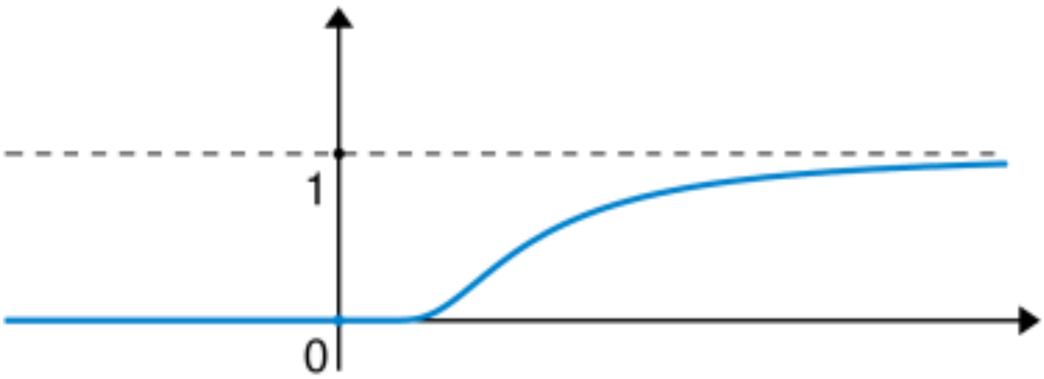



Figure 23: The graph of $f(x)$.

# 7 L'Hôpital's rule

The rule known after l'Hôpital's is a very convenient tool for obtaining a limit of fraction, when both numerator and denominator vanish in the limit. Roughly speaking, it states that in certain situations one may compute the quotient of the derivatives, instead of the quotient of the functions. However, l'Hôpital was not the discoverer of this rule. In 1694 he forged a deal with Johann Bernoulli. The deal was that l'Hôpital paid Bernoulli 300 Francs a year to tell him of his discoveries, which l'Hôpital described in his book. The widespread story that l'Hôpital tried to get credit for inventing this rule is false: he published his book anonymously, acknowledged Bernoulli's help in the introduction, and never claimed to be responsible for the rule.

Figure 24: Guillaume de l'Hôpital VS Johann Bernoulli.

***Theorem 7.1*** *(L'Hôpital's rule (the "$\frac{0}{0}$" variant)) Let $f, g$ be two functions that are differentiable in some interval centred at $x_0$, $(x_0 - r, x_0 + r)$, and except maybe at $x_0$. Suppose, in addition, that*

1. $g'(x) \neq 0$ *for every $x$ such that* $0 < |x - x_0| < r$.

2. $$\lim_{x \to x_0} f(x) = \lim_{x \to x_0} g(x) = 0,$$

3. $$L = \lim_{x \to x_0} \frac{f'(x)}{g'(x)} \qquad \textit{exists.}$$

*Then, the limit $\lim_{x \to x_0} \frac{f(x)}{g(x)}$ exists, and*

$$\lim_{x \to x_0} \frac{f(x)}{g(x)} = \lim_{x \to x_0} \frac{f'(x)}{g'(x)} = L.$$

*Proof*: Note that $f$ and $g$ are not necessarily defined at $x_0$; since they have a limit there, there will be no harm assuming that they are continuous at $x_0$, i.e, $f(x_0) = g(x_0) = 0$. First, we claim that $g$ does not vanish for every $x$ such that $0 < |x - x_0| < r$, for by Rolle's theorem, it would imply that $g'$ vanishes somewhere in this interval. More formally, if

$$g'(x) \neq 0 \qquad \text{whenever} \qquad 0 < |x - x_0| < r,$$

then

$$g(x) \neq 0 \qquad \text{whenever} \qquad 0 < |x - x_0| < r,$$

for if $g(x) = 0$, then as $g(x_0) = 0$, there exists a point $c$ between $x_0$ and $x$ such that $g'(c) = 0$. Now, let $x$ be a point such that $0 < |x - x_0| < r$, and suppose, for example, that $x_0 < x < x_0 + r$. Both $f$ and $g$ are continuous in $[x_0, x]$ and differentiable in $(x_0, x)$, and $g'$ does not vanish in this interval, hence by Cauchy's mean-value theorem (4.7), there exists a point $x_0 < c_x < x$ such that

$$\frac{f(x)}{g(x)} = \frac{f(x) - f(x_0)}{g(x) - g(x_0)} = \frac{f'(c_x)}{g'(c_x)}. \tag{7.1}$$

Repeating the same argument if $x < x_0$, we obtain (7.1) again. It only remains to let $x \to x_0$. Let's do it formally: since $\lim_{x \to x_0} \frac{f'(x)}{g'(x)} = L$,

$$\forall \varepsilon > 0 \; \exists \delta > 0 : 0 < |x - x_0| < \delta \text{ implies } \left| \frac{f'(x)}{g'(x)} - L \right| < \varepsilon, \tag{7.2}$$

and since $c_x$ satisfies $0 < |c_x - x_0| < \delta$ then

$$\forall \varepsilon > 0 \; \exists \delta > 0 : 0 < |x - x_0| < \delta \text{ implies } \left| \frac{f'(c_x)}{g'(c_x)} - L \right| < \varepsilon.$$

By (7.1),

$$\forall \varepsilon > 0 \; \exists \delta > 0 : 0 < |x - x_0| < \delta \text{ implies } \left| \frac{f(x)}{g(x)} - L \right| < \varepsilon,$$

which means that the limit

$$\lim_{x \to x_0} \frac{f(x)}{g(x)}$$

exists and is equal to $L$. ■

*Example 7.1* 1. *The following limit satisfies*

$$\lim_{x\to 0}\frac{\sin x}{x}=1.$$

2. *The following limit satisfies*

$$\lim_{x\to 0}\frac{\cos x-1}{x^2}=-\frac{1}{2}.$$

Recall that the Taylor polynomial is a tool for computing these limits. Nevertheless, we apply l'Hôpital's rule,

**Solution**

1. By l'Hôpital's rule, we have

$$\lim_{x\to 0}\frac{\sin x}{x}=\lim_{x\to 0}\frac{\cos x}{1}=1.$$

   Note that the existence of the limit is implicitly derived from the existence of the the right-hand limit.

2. By l'Hôpital's rule, we have

$$\lim_{x\to 0}\frac{\cos x-1}{x^2}=\lim_{x\to 0}\frac{-\sin x}{2x}=-\frac{1}{2}\lim_{x\to 0}\frac{\sin x}{x}=-\frac{1}{2}.$$

*Example 7.2* 1. *The following limit satisfies*

$$\lim_{x\to 0}\frac{e^x-1}{x}=1.$$

2. *The following limit satisfies*

$$\lim_{x\to 0}\frac{e^x-x-1}{x^2}=\frac{1}{2}.$$

**Solution**

1. Note that this limit is just the definition of the derivative of $e^x$ at $x_0=0$, hence is equal to $e^0=1$. Seeing this by l'Hôpital's,

$$\lim_{x\to 0}\frac{e^x-1}{x}=\lim_{x\to 0}\frac{e^x}{1}=1.$$

2. We have

$$\lim_{x\to 0}\frac{e^x-x-1}{x^2}=\lim_{x\to 0}\frac{e^x-1}{2x}=\frac{1}{2}\lim_{x\to 0}\frac{e^x}{1}=\frac{1}{2},$$

where we have used l'Hôpital's rule twice. Note, again, that the existence of the left-hand limit is guaranteed by the existence of the right-hand limit. Formally, the above equation should be read as follows: the limit $\lim_{x\to 0}\frac{e^x}{2}$ exists and equals to $\frac{1}{2}$. Hence, by l'Hôpital's rule, the limit $\lim_{x\to 0}\frac{e^x-1}{2x}$ exists and equals to $\frac{1}{2}$. Applying the rule again, the limit $\lim_{x\to 0}\frac{e^x-x-1}{x^2}$ exists and equals to $\frac{1}{2}$.

***Example 7.3*** *Compute the following limits, if exist*

*1.*

$$\lim_{x\to 0}\frac{\arctan x}{x}$$

*2.*

$$\lim_{x\to 0}\frac{\ln(1+x)}{x}$$

*3.*

$$\lim_{x\to\infty} x\ln\left(1+\frac{1}{x}\right)$$

**Solution**

1. As the derivative of $\arctan x$ is

$$(\arctan x)'=\frac{1}{1+x^2},$$

we have by l'Hôpital's rule

$$\lim_{x\to 0}\frac{\arctan x}{x}=\lim_{x\to 0}\frac{1/(1+x^2)}{1}=\lim_{x\to 0}\frac{1}{1+x^2}=1.$$

2. By l'Hôpital's rule,

$$\lim_{x\to 0}\frac{\ln(1+x)}{x}=\lim_{x\to 0}\frac{1}{1+x}=1.$$

3. Writing $y = \frac{1}{x}$, then $x \to \infty$ if and only if $y \to 0+$, hence

$$\lim_{x\to\infty} x \ln\left(1 + \frac{1}{x}\right) = \lim_{y\to 0+} \frac{1}{y} \ln(1+y) = 1.$$

As the above examples indicated, l'Hôpital's rule is a very useful tool for computing limits. Although all of the limits above can be computed via the Taylor's polynomial (as in the end of the previous section), l'Hôpital's rule saves us the computation of the Taylor polynomial (and more importantly, it saves as the need for carefully choosing the base point $x_0$ and the order of the polynomial). However, in cases where the limit is computed by applying l'Hôpital's rule more than once, Taylor's polynomial usually solves it in a single step. The following example is even more important, as it warns us when l'Hôpital's rule cannot be applied,

***Example 7.4*** *Compute the limit*

$$\lim_{x\to 0} \frac{\cos x}{x+1}.$$

Note that the above limit is trivial; both the numerator and the denominator do not converge to zero, hence the function $f(x) = \frac{\cos x}{x+1}$ is continuous at $x_0 = 0$, and $\lim_{x\to 0} f(x) = f(0) = 1$. Applying L'Hpital's rule, gives the limit of derivatives

$$\lim_{x\to 0} \frac{-\sin x}{1} = -1 \neq 1.$$

The reason l'Hôpital's rule cannot be applied is that the corresponding limit is not of the form "$\frac{0}{0}$", hence the assumptions of Theorem 7.1 are not fulfilled. However, when ever the limit of a fraction $\lim_{x\to x_0} \frac{f(x)}{g(x)}$ is nontrivial, then both $f$ and $g$ must tend to zero (assuming they are continuous); indeed, if $g$ does not tend to zero then the limit equals to $\frac{f(x_0)}{g(x_0)}$. Otherwise, if $f(x_0) \neq 0$ then the limit does not exist (or is not finite).
We proceed with an example where l'Hôpital's rule can be applied but is inefficient,

***Example 7.5*** *Compute the limit (if exists)*

$$\lim_{x\to 0} \frac{x^{17} \sin x}{\sin^{18} x + x \sin^{17} x}.$$

**Solution** The simplest way to solve this limit is perhaps to derive both the numerator and denominator by the factor of $x^18$, yielding

$$\lim_{x\to 0}\frac{x^{17}\sin x}{\sin^{18}x + x\sin^{17}x} = \lim_{x\to 0}\frac{\frac{\sin x}{x}}{\left(\frac{\sin x}{x}\right)^{18} + \left(\frac{\sin x}{x}\right)^{17}}.$$

As the limit $\lim_{x\to 0}\frac{\sin x}{x} = 1$, then by the algebra of limits we have

$$\lim_{x\to 0}\frac{x^{17}\sin x}{\sin^{18}x + x\sin^{17}x} = \frac{1}{1+1} = \frac{1}{2}.$$

If we have tried to apply l'Hôpital's rule to this limit, then we would have had to derive the denominator at least 18 times before we avoid the situation where the denominator converges to zero. Obviously, deriving a product 18 times is something we prefer not to do.
Theorem 7.1 has more than one variant. Just by looking carefully into the proof, we find the following observations,

- The proof remains valid if $x \to x_0$ is replaced by $x \to x_0+$ or $x \to x_0-$. In which case, the requirement that $x_0 < x$ should be added everywhere in the proof.
- The proof remains valid if $L = \infty$. In which case, (7.2) should be replaced by

$$\forall M > 0\ \exists \delta > 0 : 0 < |x - x_0| < \delta \text{ implies } \frac{f'(x)}{g'(x)} > M.$$

- The proof remains valid if $x_0 = \infty$ or $x_0 = -\infty$. To show that, we re-state the theorem,

***Theorem 7.2*** *(L'Hôpital's rule (the "$\frac{0}{0}$" variant with $x_0 = \infty$) Let $f, g$ be two functions that are differentiable in some interval of the form $(a, \infty)$ where $a > 0$. Suppose, in addition, that*

1. *$g'(x) \neq 0$ for every $x > a$.*
2. $$\lim_{x\to\infty} f(x) = \lim_{x\to\infty} g(x) = 0,$$

*3.*
$$L = \lim_{x\to\infty} \frac{f'(x)}{g'(x)} \qquad \textit{exists.}$$

*Then, the limit* $\lim_{x\to\infty} \frac{f(x)}{g(x)}$ *exists, and*

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

*Proof*: The proof in an application of theorem 7.1. Consider the following functions,
$$F(y) = f\left(\frac{1}{y}\right) \qquad \text{and} \qquad G(y) = g\left(\frac{1}{y}\right),$$
defined on the interval $\left(0, \frac{1}{a}\right)$. Note that by the Chain rule, $F$ and $G$ are differentiable, and

$$F'(y) = f'(y) \cdot \left(-\frac{1}{y^2}\right) \qquad \text{and} \qquad G'(y) = g'(y) \cdot \left(-\frac{1}{y^2}\right)$$

and, in particular, we have $G'(y) \neq 0$ for every $y \in \left(0, \frac{1}{a}\right)$. Also, we have

$$\lim_{y\to 0+} F(y) = \lim_{y\to 0+} f\left(\frac{1}{y}\right) = \lim_{x\to\infty} f(x) = 0,$$

where we have set $x = \frac{1}{y}$, and, similarly, we have $\lim_{y\to 0+} G(y) = 0$. Similarly,

$$\lim_{y\to 0+} \frac{F'(y)}{G'(y)} = \lim_{y\to 0+} \frac{f'\left(\frac{1}{y}\right)\cdot\left(-\frac{1}{y^2}\right)}{g'\left(\frac{1}{y}\right)\cdot\left(-\frac{1}{y^2}\right)} = \lim_{y\to 0+} \frac{f'\left(\frac{1}{y}\right)}{g'\left(\frac{1}{y}\right)} = \lim_{x\to\infty} \frac{f'(x)}{g'(x)} = L.$$

By theorem 7.1, then the limit $\lim_{y\to 0+} \frac{F(y)}{G(y)}$ exists and equals to $L$. Finally,

$$\lim_{x\to\infty} \frac{f(x)}{g(x)} = \lim_{y\to 0+} \frac{F(y)}{G(y)} = L.$$

■

Another important variant of l'Hôpital's rule is the following theorem,

***Theorem 7.3*** *(L'Hôpital's rule (the "$\infty$" variant) Let $f, g$ be two functions that are differentiable in some interval centred at $x_0$, $(x_0 - r, x_0 + r)$, and except maybe at $x_0$. Suppose, in addition, that*

1. *$g'(x) \neq 0$ for every $x$ such that $0 < |x - x_0| < r$.*

2. $$\lim_{x \to x_0} g(x) = \infty,$$

3. $$L = \lim_{x \to x_0} \frac{f'(x)}{g'(x)} \qquad \textit{exists.}$$

*Then, the limit $\lim_{x \to x_0} \frac{f(x)}{g(x)}$ exists, and*

$$\lim_{x \to x_0} \frac{f(x)}{g(x)} = \lim_{x \to x_0} \frac{f'(x)}{g'(x)} = L.$$

When we compare the above version of l'Hôpital's rule to the one in theorem 7.1 we find that the condition that $\lim_{x \to x_0} f(x) = \lim_{x \to x_0} g(x) = 0$ is replaced by the condition that $\lim_{x \to x_0} g(x) = \infty$. If $f$ is bounded, then clearly the condition that $g \to \infty$ implies that $\lim_{x \to x_0} \frac{f(x)}{g(x)} = 0$. Thus, if the limit $\lim_{x \to x_0} f(x)$ exists in the extended sense, the only interesting case is where $\lim_{x \to x_0} f(x) = \infty$. In which case we have a competition between $f$ and $g$, and the limit $\lim_{x \to x_0} \frac{f(x)}{g(x)}$ becomes non-trivial. Unfortunately, the proof for theorem 7.3 does not rely on the proof of 7.1,

*Proof*: Let $\varepsilon > 0$. If we show that

$$\limsup_{x \to x_0} \left| \frac{f(x)}{g(x)} - L \right| \leq \varepsilon, \tag{7.3}$$

then we would automatically get

$$0 \leq \liminf_{x \to x_0} \left| \frac{f(x)}{g(x)} - L \right| \leq \limsup_{x \to x_0} \left| \frac{f(x)}{g(x)} - L \right| \leq \varepsilon.$$

As the above holds for any $\varepsilon > 0$ we must have $\liminf_{x\to x_0} \left|\frac{f(x)}{g(x)} - L\right| = \limsup_{x\to x_0} \left|\frac{f(x)}{g(x)} - L\right| = 0$, hence the limit

$$\lim_{x\to x_0} \left|\frac{f(x)}{g(x)} - L\right|$$

exists and equals to zero, which is what we wanted to show. To show that (7.3) holds, let again $\varepsilon > 0$. As $L = \lim_{x\to x_0} \frac{f'(x)}{g'(x)}$, then there exists a $\delta > 0$ such that

$$0 < |x - x_0| < \delta \Rightarrow \left|\frac{f'(x)}{g'(x)} - L\right| < \varepsilon.$$

Now, suppose for example that $x_0 < x < x_0 + \delta$, and fix a point $y$ such that $x_0 < x < y < x_0 + \delta$. Now, both $f$ and $g$ are differentiable in the closed interval $[x, y]$ and $g'$ does not vanish on that interval, hence by Cauchy's mean-value theorem (4.7), there exists a point $x < c < y$ such that

$$\frac{f(x) - f(y)}{g(x) - g(y)} = \frac{f'(c)}{g'(c)}. \tag{7.4}$$

As the point $c$ satisfies $0 < |c - x_0| < \delta$, then we have

$$0 < |x - x_0| < \delta \Rightarrow \left|\frac{f'(c)}{g'(c)} - L\right| < \varepsilon,$$

and substituting (7.4) we get

$$0 < |x - x_0| < \delta \Rightarrow \left|\frac{f(x) - f(y)}{g(x) - g(y)} - L\right| < \varepsilon. \tag{7.5}$$

The above similarly holds if $x_0 - \delta < y < x < x_0$. Up to this point, the proof looks the same as the proof of theorem (7.1), only with $y$ replacing $x_0$. [20]

[20] Here is a suggested ending for this proof which has a fault: fix $y$ and let $x \to x_0$, then (7.5) looks like:

$$\lim_{x\to x_0} \frac{f(x) - f(y)}{g(x) - g(y)} = L.$$

Finally, we let $x \to x_0$. As $y$ is fixed and $\lim_{x\to x_0} g(x) = \infty$ we obtain

$$\lim_{x\to x_0} \frac{g(y)}{g(x)} = \lim_{x\to x_0} \frac{f(y)}{g(x)} = 0,$$

We thus write

$$\begin{aligned}\frac{f(x)}{g(x)} &= \frac{f(x)-f(y)}{g(x)-g(y)} \cdot \frac{g(x)-g(y)}{g(x)} + \frac{f(y)}{g(x)} \\ &= \frac{f(x)-f(y)}{g(x)-g(y)} \cdot \left[1 - \frac{g(y)}{g(x)}\right] + \frac{f(y)}{g(x)}\end{aligned} \tag{7.6}$$

Now, subtracting $L$, taking absolute values and using the Triangle inequality a few times yields

$$\begin{aligned}\left|\frac{f(x)}{g(x)} - L\right| &\le \left|\frac{f(x)-f(y)}{g(x)-g(y)}\left[1-\frac{g(y)}{g(x)}\right] - L\right| + \left|\frac{f(y)}{g(x)}\right| \\ &\le \left|\frac{f(x)-f(y)}{g(x)-g(y)} - L\right| + \left|\frac{f(x)-f(y)}{g(x)-g(y)} \cdot \frac{g(y)}{g(x)}\right| + \left|\frac{f(y)}{g(x)}\right| \\ &\le \left|\frac{f(x)-f(y)}{g(x)-g(y)} - L\right| + \left|\left(\frac{f(x)-f(y)}{g(x)-g(y)} - L\right) \cdot \frac{g(y)}{g(x)}\right| + L\left|\frac{g(y)}{g(x)}\right| + \left|\frac{f(y)}{g(x)}\right|\end{aligned}$$

Now, applying (7.5), we obtain

$$\left|\frac{f(x)}{g(x)} - L\right| \le \varepsilon + \varepsilon\left|\frac{g(y)}{g(x)}\right| + L\left|\frac{g(y)}{g(x)}\right| + \left|\frac{f(y)}{g(x)}\right|.$$

The above inequality holds for every $x, y$ such that $x_0 < x < y < x_0 + \delta$ or $x_0 - \delta < y < x < x_0$. Taking $x \to x_0$ (either from the left or from the right) we get that (note that now $y$ is fixed and depends on $\varepsilon$)

$$\limsup_{x\to x_0}\left|\frac{f(x)}{g(x)} - L\right| \le \varepsilon + \varepsilon \lim_{x\to x_0}\left|\frac{g(y)}{g(x)}\right| + L \cdot \lim_{x\to x_0}\left|\frac{g(y)}{g(x)}\right| + \lim_{x\to x_0}\left|\frac{f(y)}{g(x)}\right|$$

As $\lim_{x\to x_0} g(x) = \infty$ then the right-hand side limits exist, and

$$\lim_{x\to x_0}\left|\frac{g(y)}{g(x)}\right| = \lim_{x\to x_0}\left|\frac{f(y)}{g(x)}\right| = 0$$

---

and together with (7.6) we get, by algebra of limits, that the limit $\lim_{x\to x_0} \frac{f(x)}{g(x)}$ exists and

$$\lim_{x\to x_0} \frac{f(x)}{g(x)} = L \cdot 1 + 0 = L.$$

However, the above is statement is senseless as $y$ depends on $\delta$, and hence on $\varepsilon$ (so therefore $y$ is not really fixed). We can fix $y$ only after choosing $\varepsilon$, not before!

hence,

$$\limsup_{x\to x_0}\left|\frac{f(x)}{g(x)}-L\right|\le\varepsilon,$$

as required. ∎

It is easy to verify that Theorem 7.3 remains valid for $x_0=\infty$, by applying the trick of changing variables $y=\frac{1}{x}$.

***Example 7.6*** *Compute the limits (if exist)*

$$\lim_{x\to\infty}\frac{\ln x}{x}\qquad \textit{and}\qquad \lim x\to\infty x(\arctan x-\frac{\pi}{2}).$$

**Solution** Applying Theorem 7.3 for the first limit, we get

$$\lim_{x\to\infty}\frac{\ln x}{x}=\lim_{x\to\infty}\frac{1}{x}=0.$$

Applying Theorem (7.2) for the second limit, we get

$$\lim_{x\to\infty}x(\arctan x-\frac{\pi}{2})=\lim_{x\to\infty}\frac{\arctan x-\frac{\pi}{2}}{1/x}=\lim_{x\to\infty}\frac{\frac{1}{1+x^2}}{-1/x^2}=-\lim_{x\to\infty}\frac{x^2}{1+x^2}$$
$$=-\lim_{x\to\infty}\frac{1}{\frac{1}{x^2}+1}=-1.$$

***Example 7.7*** *Let $P(x)$ be a polynomial. Then*

$$\lim_{x\to\infty}\frac{P(x)}{e^x}=0.$$

*Proof*: Thus,

$$\lim_{x\to\infty}\frac{P(x)}{e^x}=\lim_{x\to\infty}\frac{P'(x)}{e^x}.$$

Repeating the step above $n+1$ times, where $n=\deg P$, we find that the denominator vanishes. Thus,

$$\lim_{x\to\infty}\frac{P(x)}{e^x}=0. \tag{7.7}$$

∎

***Example 7.8*** *Consider the function given in* (6.8),

$$g(x) = \begin{cases} \exp\left(-\frac{1}{x^2}\right) & x \neq 0 \\ 0 & x = 0 \end{cases}$$

*Then $g$ is a $C^\infty(\mathbb{R})$ function, and $g^{(n)}(0) = 0$ for every $n$.*

*Proof*: For every $x_0 \neq 0$, $g$ is differentiable $\infty$ number of times by the algebra of differentiable functions. It only remains to show that $g$ is differentiable at zero $\infty$ number of times, and $g^{(n)}(0) = 0$. We prove this statement by induction on $n$; If $n = 0$ then we show that $g$ is continuous at zero; indeed, as $\lim_{x\to 0} -\frac{1}{x^2} = -\infty$ we have

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

Now, it is easy to verify by induction that for every $n$ and $x \neq 0$ we have

$$g^{(n)}(x) = \frac{P(x)e^{-\frac{1}{x^2}}}{Q(x)},$$

where $P(x)$ and $Q(x)$ are polynomials. Suppose that $g$ is differentiable $n$ times at zero and that $g^{(n)}(0) = 0$. Then, by definition,

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

Now, there exist polynomials $P(x)$ and $Q(x)$ such that

$$\frac{g^{(n)}(x)}{x} = \frac{P(x)e^{-\frac{1}{x^2}}}{Q(x)}.$$

It only remains to show that given any two such polynomials, we have

$$\lim_{x\to 0} \left[\frac{P(x)e^{-\frac{1}{x^2}}}{Q(x)}\right] = 0.$$

Denoting by $y = \frac{1}{x}$ we have

$$\frac{P(x)e^{-\frac{1}{x^2}}}{Q(x)} = \frac{P(y^{-1})}{Q(y^{-1})} \exp\left(-y^2\right).$$

It is easy to verify that there exist polynomials $p(y)$ and $q(y)$ such that

$$\frac{P(y^{-1})}{Q(y^{-1})}\exp\left(-y^2\right) = \frac{p(y)}{q(y)\exp\left(y^2\right)}$$

Taking absolute values, for sufficiently small values of $x$ we have $|q(y)| > 1$, and $|p(y)| \leq p(y^2)$, hence

$$\left|\frac{P(x)e^{-\frac{1}{x^2}}}{Q(x)}\right| = \left|\frac{p(y)}{q(y)e^{y^2}}\right| \leq \left|\frac{p(y^2)}{e^{y^2}}\right|.$$

Setting $z = y^2$, note that $x \to 0$ if and only if $z = \frac{1}{x^2} \to \infty$, and, by (7.7),

$$\lim_{z\to\infty}\left|\frac{p(z)}{e^z}\right| = 0.$$

Thus, by the squeeze rule we have

$$\lim_{x\to 0}\left|\frac{P(x)e^{-\frac{1}{x^2}}}{Q(x)}\right| = 0.$$

■

**Acknowledgements** I wish to thank Dr Xue-Mei Li and Dr Oleg Zaboronski for giving me their lecture notes from previous years, on which I based my talks. Also, many thanks to all of the students who have given me their comments and remarks throughout the term (and currently do so as we speak).