# BASIC CONCEPTS OF CALCULUS 1

FRÉDÉRIC LATRÉMOLIÈRE

ABSTRACT. ©2024 Frédéric Latrémolière, PhD.
This document is only for the current students in Calculus 1 MATH 1951 at DU. No redistribution or online distribution is allowed.

Contents

## 1. A problem

Let $f : x \in \mathbb{R} \mapsto x^2 - 2$. Given any $x_0 \neq x_1 \in \mathbb{R}$, we can compute an equation for the secant to the graph of $f$ passing through $(x_0, f(x_0))$, $(x_1, f(x_1))$:

$$y - f(x_0) = \frac{f(x_1) - f(x_0)}{x_1 - x_0}(x_1 - x_0).$$

For instance, if $x_0 = -1$ and $x_1 = 0$, then our secant has equation $y = x - 2$.

Heuristically, we observe that, fixing $x_0$ and moving $x_1$ closer to $x_0$, the secant seems to move toward a particular line. However, we can not compute the slope of this line by purely algebraic means, since we can not divide by 0. Yet, as long as $x_1 \neq x_0$:

$$\begin{aligned}\frac{f(x_1) - f(x_0)}{x_1 - x_0} &= \frac{x_1^2 - 2 - (x_0^2 - 2)}{x_1 - x_0} \\ &= \frac{(x_1 - x_0)(x_1 + x_0)}{x_1 - x_0} = x_1 + x_0.\end{aligned}$$

Fixing $x_0$, it stands to reason that the “limit” of the slope of the secant, as $x_1$ goes to $x_0$, is $2x_0$. Indeed, if $x_0 = 1$, we can see that the natural “limit" of the secant lines above seems to be a line of slope 2. This leads us to an equation for a special limit line, called the *tangent*, with $y = 2(x - 1) + f(1) = 2x - 3$.

*Limits*, the foundation of analysis/calculus the core contribution of topology, are defined as values a function could assume at a given input in such a way as the function is (locally) *continuous*. For instance, the tangent line seems to be the "natural" guess to complete the picture given by the evolution of the secants, as we move $x_1$ to $x_0$. Similarly, $2x_0$

*Date*: July 22, 2024.

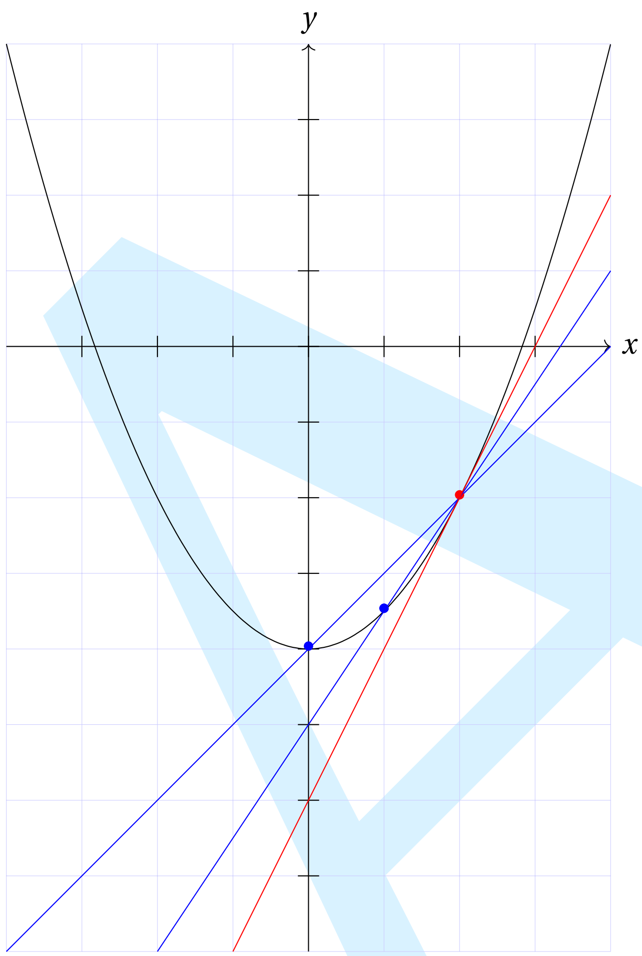


FIGURE 1. Two secants (blue) and the tangent (red) to $x \mapsto x^2 - 2$ at $(1, -1)$

is the natural guess as to which value we would insert at $x_0$ for the slope of the secant above, even though the algebraic formula for the slope is no longer defined at this point.

Tangent lines and their slopes have many applications. For instance, an object in free fall will see its height as a function of time evolved as $z : t \longmapsto gt^2 + v_0 t + z_0$, for some constant values $g$, $v_0$ and $z_0$. If we wish to compute how fast this object moves downward at a given time, we encounter the same problem as above. For two different times $t_0$ and $t_1$, the speed of the object between these two instants is:

$$\frac{z(t_1) - z(t_0)}{t_1 - t_0} = \frac{g(t_1^2 - t_0^2) + v_0(t_1 - t_0)}{t_1 - t_0}.$$

This formula does not make sense when $t_1 = t_0$. However, after trivial simplication,

$$\frac{z(t_1) - z(t_0)}{t_1 - t_0} = \frac{g(t_1^2 - t_0^2) + v_0(t_1 - t_0)}{t_1 - t_0}$$
$$= g\frac{(t_1 - t_0)(t_1 + t_0)}{t_1 - t_0} + v_0 = g(t_1 + t_0) + v_0.$$

Remarkably, this latter formula makes sense even when $t_1 = t_0$. *Be careful: the equalities above do not make any sense when $t_0 = t_1$; only the final formula does. So the speed of our object and the final formula defines different functions, although the only difference is that one makes no sense at $t_1 = t_0$ while the other does!* So, it is now natural to "extend by continuity" the speed formula and set that the velocity at $t_0$ is given by $2gt_0 + v_0$.

In general, if we have some function $f : D \subseteq \mathbb{R} \to \mathbb{R}$ and $t \in D$, where $D$ is some open interval in $\mathbb{R}$, which we use to model any phenomenon, then we may want to know how $f$ varies at $t$: for instance, if $f$ is a position on some axis and the variable is time, then this variation is some speed; if $f$ is a temperature, or a concentration of some compound in a solution during a reaction, or an electric charge in some conductor, or some energy produced by a system, then this variation is the variation in temperature, the speed of production/consumption of the compound, or the current in the conductor, or the power produced by this system. In all these cases and many more, we end up with asking whether a ratio of the form

$$\frac{f(t_1) - f(t_0)}{t_1 - t_0}$$

can be extended to a "natural" value, i.e. in a "continuous manner", when $t_1$ gets taken toward $t_0$ (while never equating it). We will later on call this value the *derivative* of $f$ at $t_0$; as seen above, the derivative is the concept behind many notions in practice (currents, powers, chemical rate of reactions, velocity, acceleration — the derivative of the velocity, ...). Mathematics is the study of these deep concepts which emerge in various areas, abstracted from their origin: everything true for derivatives will be true for any of these practical manifestations of derivatives.

It is not always as easy as using basic algebra to extend an ill-behaved ratio. For instance, let us look at the variation of sin near 0. The ratio becomes:

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

seen as a function of $x$. As $x$ gets toward 0, it may not be clear what the value of this ratio becomes. However, if we were to plot this ratio, we note

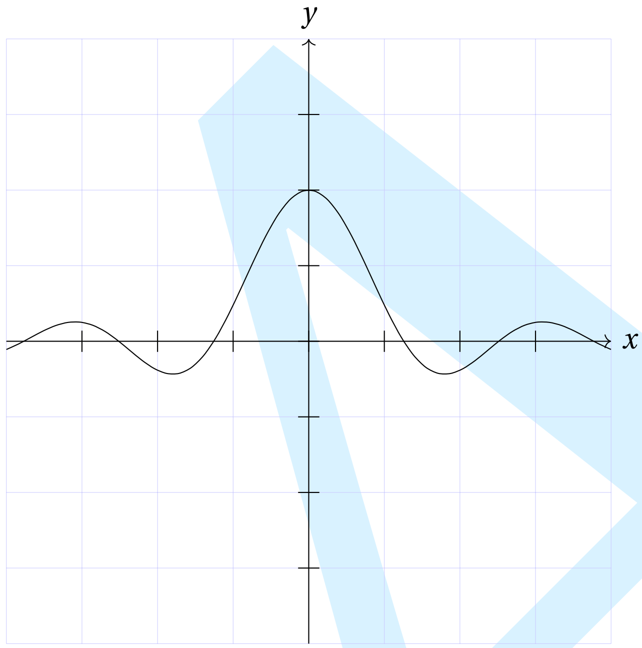


FIGURE 2. The function $x \in \mathbb{R} \setminus \{0\} \mapsto \frac{\sin(x)}{x}$, while not defined at 0, certainly seems extendable at 0.

that this ratio does not show any singular behavior at 0 — in fact, it seems to take the value 1 there!

The concept of limit is fundamentally connected with the notion of continuity. It has taken thousands of mathematical research to define an appropriately rigorous notion of continuity. The reason is that we are now dealing with infinity and infinitesimal, and these concepts require great care in order not to introduce all sort of nonsense. Calculus is the subject which handles this issue. Among other issues to keep in mind: as we saw, in general, algebra is vastly insufficient to handle the matter. Moreover, functions may have no limit at some points, and it turns out to be quite delicate to determine which function has a limit. Our theory needs to account for situations like $x \neq 0 \mapsto \sin\left(\frac{1}{x}\right)$, which does not seem to have a particular definite limit at 0.

## 2. LIMITS AND CONTINUITY

### 2.1. Continuity.

**Definition 2.1.** The *closure* $\mathrm{cl}(A)$ of a subset $A \subseteq \mathbb{R}$ is the set of points $x \in \mathbb{R}$ such that no open interval containing $x$ lies entirely in the complement of $A$.

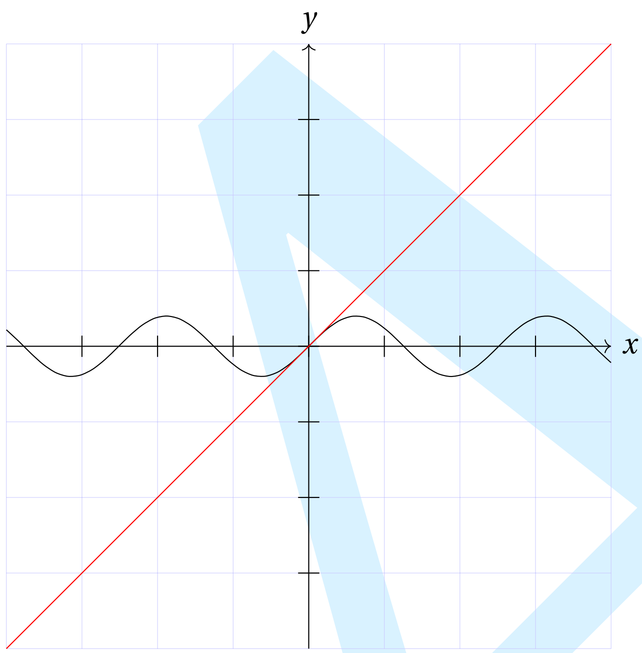


FIGURE 3. sin and its tangent at 0

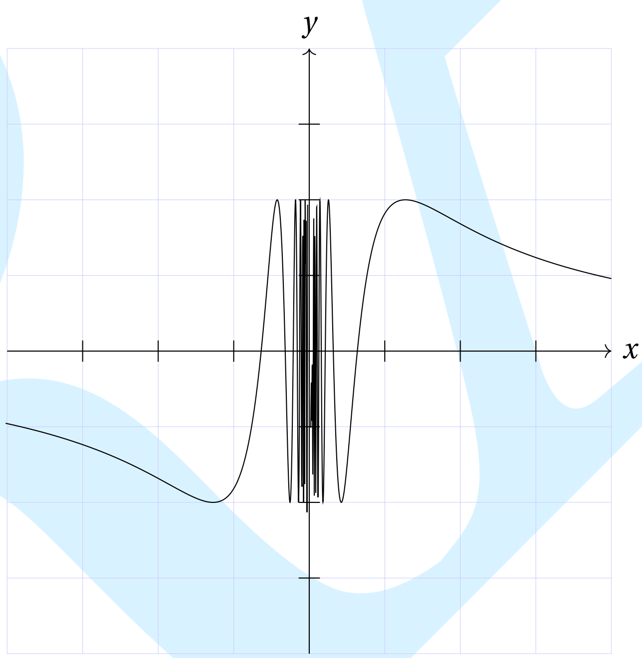


FIGURE 4. The function $x \mapsto \sin\left(\frac{1}{x}\right)$: no guess for the value at 0

**Theorem 2.2.** *A point $x \in \mathbb{R}$ is in the closure of a subset $A \subseteq \mathbb{R}$ of $\mathbb{R}$ if, and only if, for all $\varepsilon > 0$, there exists $a_\varepsilon \in A$ such that $|x - a_\varepsilon| < \varepsilon$.*

*Proof.* Assume that $x \in \mathrm{cl}(A)$. Let $\varepsilon > 0$. The open interval $(x-\varepsilon, x+\varepsilon)$ is not contained in the complement of $A$, i.e there exists $a_\varepsilon \in A$ also in $(x-\varepsilon, x+\varepsilon)$, i.e. $a_\varepsilon \in A$ and $|x - a_\varepsilon| < \varepsilon$.

Conversely, let $x \in \mathbb{R}$ such that, for all $\varepsilon > 0$, there exists $a_\varepsilon \in A$ such that $|x - a_\varepsilon| < \varepsilon$. Let $(a, b)$ be an open interval with $x \in (a, b)$. Let $\varepsilon = \min\{x - a, b - x\}$. Then $(x-\varepsilon, x+\varepsilon) \subseteq (a, b)$. By assumption, there exists $a_\varepsilon \in A \cap (x-\varepsilon, x+\varepsilon)$, so $(a, b)$ is not contained in the complement of $A$. □

**Example 2.3.** If $x \in A$, then any open interval containing $x$ obviously contains a point of $A$ (namely, $x$), so it is never contained in the complement of $A$. So in particular, any point in $A$ is in the closure of $A$.

**Example 2.4.** The point 0 is not in the open interval $(0, 1)$, but it is in its closure: indeed, any open interval containing 0 is of the form $(a, b)$ with $b > 0$ and $a < 0$; in particular it contains, say, $\min\{\frac{1}{2}, b\}$, which is also in $(a, b)$.

**Example 2.5.** Any real number is in the closure of the set of rationals.

**Example 2.6.** 0 is in the closure of $\left\{\frac{1}{n} : n \in \mathbb{N} \setminus \{0\}\right\}$.

A subset $F \subseteq \mathbb{R}$ of $\mathbb{R}$ is closed when it equals its closure. The complement of a closed set is open. Some sets are neither close nor open (like $(0, 1]$) and some are both (like $\mathbb{R}$).

**Definition 2.7.** A function $f : D \to \mathbb{R}$ is *continous at $x$* when, for any subset $A \subseteq D$ of $D$ such that $x$ is in the closure of $A$, the value $f(x)$ is in the closure of $f(A)$ — the range of the restriction of $f$ to $A$.

**Definition 2.8.** A function $f : D \to \mathbb{R}$ is *continuous over $D$* when it is continuous at all $x \in D$.

**Theorem 2.9.** *A function $f : D \to \mathbb{R}$ is* continous over $D$ *if, and only if, for any subset $A \subseteq D$ of $D$, the restriction of $f$ to* $\mathrm{cl}(A) \cap D$ *is contained in the closure of $f(A)$.*

*Proof.* Assume that $f$ is continuous over $D$. Let $A \subseteq D$, not empty (if $A$ is empty, the result is obvious). Let $x \in \mathrm{cl}(A) \cap D$. Then $f(x)$ is in the closure of $f(A)$. Since $x$ is arbtirary in $A$, we have shown that $f(\mathrm{cl}(A) \cap D) \subseteq \mathrm{cl}\big(f(A)\big)$.

Assume now that for all $A \subseteq D$, we have $f(\mathrm{cl}(A) \cap D) \subseteq \mathrm{cl}\big(f(A)\big)$. Let $x \in D$ Let $A \subseteq D$ with $x \in \mathrm{cl}(D)$. Then $f(x) \in \mathrm{cl}\big(f(A)\big)$ by assumption. Hence, $f$ is continuous at $x \in D$. □

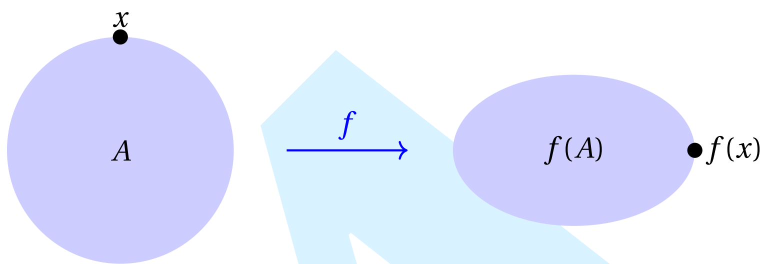


FIGURE 5. A continuous function $f$ maps $x \in \mathrm{cl}(A)$ to $f(x) \in \mathrm{cl}\big(f(A)\big)$.

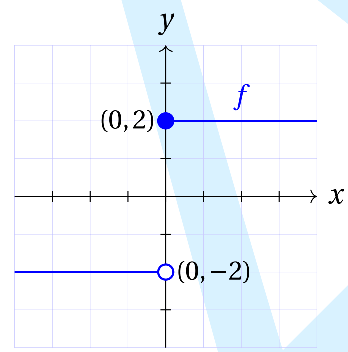


FIGURE 6. Graph of the discontinuous function

$$f : x \in \mathbb{R} \mapsto \begin{cases} -2 \text{ if } x < 0, \\ 2 \text{ if } x \geqslant 0. \end{cases}$$

Note that $0 \in \mathrm{cl}((-\infty, 0))$, and $f(-\infty, 0) = \{-2\}$, yet $f(0) = 2 \notin \mathrm{cl}(\{-2\}) = \{-2\}$.

**Theorem 2.10.** *A function $f : D \to \mathbb{R}$ is continuous over at $x \in D$ if, and only if:*

$$\forall \varepsilon > 0 \quad \exists \delta > 0 \quad \forall t \in D \quad |t - x| < \delta \implies |f(t) - f(x)| < \varepsilon. \tag{2.1}$$

*Proof.* Assume that Expression (**??**) holds. Let $A \subseteq D$ with $x \in \mathrm{cl}(A)$. Let $\varepsilon > 0$. There exists $\delta > 0$ such that, for all $t \in D$, if $|t - x| < \delta$, then $|f(t) - f(x)| < \varepsilon$. Since $x \in \mathrm{cl}(A) \cap D$, there exists indeed $t \in A$ with $|t - x| < \delta$, and thus there exists an element in $f(A)$ — namely $f(t)$ — such that $|f(t) - f(x)| < \varepsilon$. By Theorem (**??**), we conclude that $f(x) \in \mathrm{cl}\big(f(A)\big)$.

Assume that Expression (**??**) does not hold. Thus, there exists $\varepsilon > 0$ such that, for all $\delta > 0$, there exists $t \in D$ with $|t - x| < \delta$ and $|f(t) - f(x)| \geqslant \varepsilon$. In other words, $A := \{t \in D : |f(t) - f(x)| < \varepsilon\}$ is not empty, and in fact, for all $\delta > 0$, there exists $t_\delta \in A$ such that $|t_\delta - x| < \delta$. Therefore, $x \in \mathrm{cl}(A)$. Yet, by construction, $(f(x) - \varepsilon, f(x) + \varepsilon)$ lies in the complement of $f(A)$, so $f(x)$ is not in the closure of $f(A)$. Thus, $f$ is not continuous over $D$. $\square$

**Corollary 2.11.** *A function $f : D \to \mathbb{R}$ is continuous over $D$ when*

$$\forall \varepsilon > 0 \quad \in D \quad \exists \delta > 0 \quad \forall t \in D \quad |t - x| < \delta \implies |f(x) - f(t)| < \varepsilon.$$

*Proof.* Obvious by definition. □

**Definition 2.12.** A function $f : D \to \mathbb{R}$ is $k$*-Lipschitz*, for $k > 0$, when for all $x, y \in D$,

$$|f(x) - f(y)| \leqslant k|x - y|.$$

More generally, a function $f : D \to \mathbb{R}$ is $(k, \alpha)$*-Hölder*, for $k, \alpha > 0$ when for all $x, y \in D$,

$$|f(x) - f(y)| \leqslant k|x - y|^{\alpha}.$$

**Theorem 2.13.** *If $f : D \to \mathbb{R}$ is $(k, \alpha)$-Hölder, then it is continuous over $D$.*

*Proof.* Let $\varepsilon > 0$ and $x \in D$. If we set $\delta := \left(\frac{\varepsilon}{k}\right)^{\frac{1}{\alpha}} > 0$, then for all $y \in D$ with $|x - y| < \delta$, we compute:

$$\begin{aligned} |f(x) - f(y)| &\leqslant k|x - y|^{\alpha} \\ &< k\delta^{\alpha} = k\left(\left(\frac{\varepsilon}{k}\right)^{\frac{1}{\alpha}}\right)^{\alpha} \\ &= \varepsilon. \end{aligned}$$

Our result follows from Theorem (**??**). □

2.2. **Limits.** Limits are, intuitively, values one would choose to extend a function by continuity, but only in a "local" manner: we take advantage of Theorem (**??**), and the quantification over the input variable in the domain, to make our notion local by removing this quantification and fixing a value not necessarily in the domain, but in its closure (close enough).

**Definition 2.14.** A function $f : D \to \mathbb{R}$ has a limit $l$ at $a \in \mathrm{cl}(D)$ when

$$\forall \varepsilon > 0 \quad \exists \delta > 0 \quad \forall x \in D \quad 0 < |x - a| < \delta \implies |f(x) - l| < \varepsilon.$$

Limits, if they exist, are unique.

**Theorem 2.15.** *If $f : D \to \mathbb{R}$ has limit $l$ and $l'$ at $a \in \mathrm{cl}(D)$, then $l = l'$.*

*Proof.* Assume that $l \neq l'$, and that $f$ has both limit $l$ and $l'$ at $a$. Set $\varepsilon = |l - l'| > 0$. By Theorem (**??**), there exists $\delta > 0$ such that, if $y \in D$ with $|y - a| < \delta$, then $|f(y) - l| < \frac{\varepsilon}{2}$. There also exists $\delta' > 0$ such that, if $y \in D$ with $|y - a| < \delta'$, then $|f(y) - l'| < \frac{\varepsilon}{2}$. Since $a \in \mathrm{cl}(D)$, there exists $t \in D$ with $|t - a| < \min\{\delta, \delta'\}$; we then get:

$$|l - l'| \leqslant |l - f(t)| + |f(t) - l'| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon = |l - l'|.$$

We have reached a contradiction. □

*Notation* 2.16. Let $f : D \to \mathbb{R}$. Let $a \in \mathbb{R}$. If $a$ is in the closure of $D \setminus \{a\}$, and if the restriction of $f$ to $D \setminus \{a\}$ has a limit $l$ at $a$, then we denote this limit by $\lim_{x \to a} f(x)$ (or $\lim_a f$). We say that $f$ has limit $\lim_a f$ at $a$.

If $a$ is in the closure of $D \cap (-\infty, a)$, and if the restriction of $f$ to $D \cap (-\infty, a)$ has a limit $l$ at $a$, then we denote this limit by $\lim_{\substack{x \to a \\ x < a}} f(x)$, or $\lim_{x \to a^-} f(x)$ (or $\lim_{a^-} f$). This limit is called the *left limit* of $f$ at $a$.

If $a$ is in the closure of $D \cap (a, \infty)$, and if the restriction of $f$ to $D \cap (a, \infty)$ has a limit $l$ at $a$, then we denote this limit by $\lim_{\substack{x \to a \\ x > a}} f(x)$, or $\lim_{x \to a^+} f(x)$ (or $\lim_{a^+} f$). This limit is called the *right limit* of $f$ at $a$.

**Theorem 2.17.** *A function $f : D \to \mathbb{R}$ has limit $l$ at $a$ if, and only if, $f$ has left and right limit $l$ at $a$.*

*Proof.* Assume $\lim_{a^-} f = \lim_{a^+} f = l$. Let $\varepsilon > 0$. By Theorem (**??**), there exists $\delta > 0$ such that, if $x \in D \cap (-\infty, a)$ and $|x - a| < \delta$, then $|f(x) - l| < \varepsilon$, and there exists $\delta' > 0$ such that, if $x \in D \cap (a, \infty)$ and $|x - a| < \delta$, then $|f(x) - l| < \varepsilon$. So if $x \in D \setminus \{a\}$ with $|x - a| < \min\{\delta, \delta'\}$, we have $|f(x) - l| < \varepsilon$, so by Theorem (**??**), $f$ has limit $l$ at $a$.

The converse is similar (and easier). □

A key property of limits is that they are a *local concept*, which means the following.

**Theorem 2.18.** *Let $f : D \to \mathbb{R}$, $a \in \mathrm{cl}(D)$ and $\delta > 0$. The function $f$ has limit $l$ at $a$ if, and only if, the restriction of $f$ to $D \cap (a - \delta, a + \delta)$ has limit $l$ at $a$.*

*Proof.* Assume that the restriction of $f$ to $(a - \delta, a + \delta)$ has limit $l$ at $a$, for some $\delta > 0$. Let $\varepsilon > 0$. There exists $\alpha > 0$ such that, if $x \in D \cap (a - \delta, a + \delta)$ and $|x - a| < \alpha$, then $|f(x) - l| < \varepsilon$. So for all $x \in D$ with $|x - a| < \min\{\delta, \alpha\}$, we have $|f(x) - l| < \varepsilon$, so $f$ has limit $l$ at $a$.

The converse is analogous. □

In the sequel, all our proofs work whenever we restrict the domain of the function, so in particular, they apply to left limits, right limits, and are valid if we restrict the functions to some open intervals centered on the point at which we take our limit. This is useful to keep in mind.

Local continuity is simply the statement that the limit of a function is its value.

**Theorem 2.19.** *A function $f : D \to \mathbb{R}$ is continuous at $x \in D$ if, and only if,* $\lim_{t \to x} f(t) = f(x)$.

*A function $f : D \to \mathbb{R}$ is continuous over $D$ if, and only if,* $\lim_{t \to x} f(t) =$ *$f(x)$ for every $x \in D$.*

*Proof.* This is Theorem (**??**). ☐

Limits live in the closure of ranges, which is a useful observation.

**Theorem 2.20.** *If $f : D \to \mathbb{R}$ has limit $l$ at $a \in \mathrm{cl}(D)$, then $a \in \mathrm{cl}\big(f(D)\big)$.*

*Proof.* Let $\varepsilon > 0$. By definition, there exists $\delta > 0$ such that, if $x \in D$ and $|x-a| < \delta$, then $|f(x) - l| < \varepsilon$. Since $a \in \mathrm{cl}(D)$, we conclude that indeed, there exists $x \in D$ with $|x-a| < \delta$, and therefore, such that $|f(x) - l| < \varepsilon$. So by Theorem (**??**), we conclude $l \in \mathrm{cl}\big(f(D)\big)$. ☐

Since determining exact ranges of functions is challenging, the above result is often used in the following weaker but more applicable form.

**Corollary 2.21.** *If $f : D \to \mathbb{R}$ has limit $l$ at $a \in \mathrm{cl}(D)$, and if the range of $f$ is contained in $A$, then $l \in \mathrm{cl}(A)$.*

We then observe the following useful result.

**Corollary 2.22.** *If $f : D \to [k,\infty)$ has limit $l$ at $a \in \mathrm{cl}(D)$, then $l \geqslant k$. If $f : D \to (-\infty, k]$ has limit $l$ at $a \in \mathrm{cl}(D)$, then $l \leqslant k$.*

2.3. **Computation of Limits.** We now establish results about how to compute limits. First, we note the following:

**Corollary 2.23.** *If $f : D \to \mathbb{R}$ is a Hölder function over $D$ and $a \in D$, then* $\lim_{x\to a} f(x) = f(a)$.

*Proof.* Since $f$ is Hölder, it is continuous over $D$, and thus continuous at $a \in D$, hence the result. ☐

**Corollary 2.24.** *If $f : x \in \mathbb{R} \mapsto \alpha x + \beta$, for some constant $\alpha, \beta$, then for all $a \in \mathbb{R}$, we have* $\lim_{x\to a} f(x) = f(a) = \alpha a + \beta$.

*Proof.* The affine function $f$ is Lipschitz on $\mathbb{R}$ — $|f(x) - f(y)| = |\alpha||x-y|$ for all $x, y \in \mathbb{R}$, hence our result. ☐

*Remark* 2.25. Constant functions are continuous, as a special case of affine functions, or even, Lipschitz functions.

We now turn to the basic manipulations of limits. First, since limits extend the idea of evaluation, rules about evaluations can be extended, under the assumption that limits indeed exist.

**Theorem 2.26** (Composition)**.** *If $f : D \to \mathbb{R}$ is a function continuous at $x \in D$, and if $g : E \to \mathbb{R}$ is a function with limit $x$ at $a \in \mathrm{cl}(E)$, and if $g(E) \subseteq D$, then* $\lim_{t\to x} f(g(t)) = f(x)$.

*Proof.* Let $\varepsilon > 0$. There exists $\delta_f > 0$ such that, for all $t \in D$, if $|t - x| < \delta_f$, then $|f(t) - f(x)| < \varepsilon$. There exists $\delta_g > 0$ such that, for all $y \in E$, if $|y - a| < \delta_g$, then $|g(y) - x| < \delta_f$. So for all $y \in E$, if $|y - a| < \delta_g$, then $|f(g(y)) - f(x)| < \varepsilon$, as claimed. □

**Theorem 2.27** (Linearity of Limits)**.** *If $f : D \to \mathbb{R}$ and $g : D \to \mathbb{R}$ are two functions with limits at some $a \in \mathrm{cl}(D)$, then for all $\lambda \in \mathbb{R}$, the function $\lambda f + g$ has a limit at a, and*

$$\lim_{x \to a} (\lambda f(x) + g(x)) = \lambda \lim_{x \to a} f(x) + \lim_{x \to a} g(x).$$

*Proof.* Let $l := \lim_a f$ and $l' := \lim_a g$. Fix $\lambda \in \mathbb{R}$. Let $\varepsilon > 0$. By definition, there exists $\delta > 0$ such that, if $x \in D$ with $|x - a| < \delta$ then $|f(x) - l| < \frac{\varepsilon}{2|\lambda|+1}$, and there exists $\delta' > 0$ such that, if $x \in D$ with $|x - a| < \delta'$ then $|g(x) - l| < \frac{\varepsilon}{2}$. Thus, if $x \in D$ with $|x - a| < \min\{\delta, \delta'\}$, then

$$\begin{aligned}|(\lambda f + g)(x) - (\lambda l + l')| &= |\lambda(f(x) - l) + (g(x) - l')| \\ &\leqslant |\lambda||f(x) - l| + |g(x) - l'| < |\lambda| \frac{\varepsilon}{2|\lambda| + 1} + \frac{\varepsilon}{2} < \varepsilon.\end{aligned}$$

Our theorem is proven by definition. □

**Lemma 2.28.** *The function $x \in \mathbb{R} \mapsto x^2$ is continuous over $\mathbb{R}$.*

*Proof.* Let $\varepsilon > 0$. Let $\delta := \min\left\{1, \frac{\varepsilon}{1+2|a|}\right\} > 0$. If $x \in \mathbb{R}$ and $|x - a| < \delta$, then $|x + a| \leqslant |x - a| + 2|a| \leqslant 1 + 2|a|$, and thus

$$|x^2 - a^2| = |x - a||x + a| \leqslant \delta \frac{1}{1 + 2|a|} \leqslant \varepsilon.$$

This completes our proof. □

**Corollary 2.29.** *If $f : D \to \mathbb{R}$ and $g : D \to \mathbb{R}$ are two functions with limits at $a \in \mathrm{cl}(D)$, then $fg$ has a limit at a, and*

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

*Proof.* Let $l := \lim_a f$ and $l' := \lim_a g$.

By linearity, $f + g$ and $f - g$ have limits, respectively, $l + l'$ an $l - l'$ at $a$. By composition, $(f + g)^2$ and $(f - g)^2$ have limits $(l + l')^2$ and $(l - l')^2$, respectively. Using linearity of limits again, since:

$$fg = \frac{1}{4}\left((f + g)^2 - (f - g)^2\right),$$

we conclude that $fg$ has limit $\frac{1}{4}((l + l')^2 - (l - l')^2) = ll'$. □

**Corollary 2.30.** *All polynomials are continuous over $\mathbb{R}$.*

*Proof.* As they are affine functions (hence Lipschitz), $x \in \mathbb{R} \mapsto 1$ and $x \in \mathbb{R} \mapsto x$ are both continuous on $\mathbb{R}$. Assume that, for some $n \in \mathbb{N}$, the function $x \in \mathbb{R} \mapsto x^n$ is continuous. By Corollary (**??**), the function $x \in \mathbb{R} \mapsto x^{n+1} = x^n x$ is also continuous. By induction, for all $n \in \mathbb{N}$, the function $x \in \mathbb{R} \mapsto x^n$ is continuous over $\mathbb{R}$.

Polynomials are linear combinations of the functions $x \in \mathbb{R} \mapsto x^n$ ($n \in \mathbb{N}$), so our result follows from Theorem (**??**). □

**Lemma 2.31.** *The function $x \in (-\infty,0) \cup (0,\infty) \mapsto \frac{1}{x}$ is continuous over $(-\infty,0) \cup (0,\infty)$.*

*Proof.* Let $a \in (-\infty,0) \cup (0,\infty)$ and let $\varepsilon > 0$. Set $\delta := \min\left\{\frac{|a|^2\varepsilon}{2}, \frac{|a|}{2}\right\} > 0$. If $x \in (-\infty,0) \cup (0,\infty)$ with $|x-a| < \delta$, then we observe the following. First, $|x| > \frac{|a|}{2} > 0$. Second,

$$\begin{aligned} \left|\frac{1}{x} - \frac{1}{a}\right| &= \frac{|a-x|}{|xa|} \\ &< \frac{2\delta}{|a|^2} \leqslant \varepsilon. \end{aligned}$$

This concludes our proof. □

**Corollary 2.32.** *If $f : D \to \mathbb{R}$ has limit $l$ at $a \in \mathrm{cl}(D)$, then $\frac{1}{f}$ is well defined on a domain of the form $D \cap (a-\delta, a+\delta)$, for some $\delta > 0$, and moreover $\frac{1}{f}$ has limit $\frac{1}{l}$ at $a$ on this domain.*

*Proof.* For $\varepsilon = \frac{|l|}{2} > 0$, there exists $\delta > 0$ such that, if $x \in D$ and $|x-a| < \delta$, then $|f(x) - l| < \varepsilon = \frac{|l|}{2}$, and thus $|f(x)| > \frac{|l|}{2}$. Now, by composition, we obtain the desired limit. □

**Corollary 2.33.** *If $f : D \to \mathbb{R}$ has limit $l$ at $a \in \mathrm{cl}(D)$, and if $g : D \to \mathbb{R}$ has limit $l'$ at $a \in \mathrm{cl}(D)$, then $\frac{g}{f}$ is well defined on a domain of the form $D \cap (a-\delta, a+\delta)$, for some $\delta > 0$, and moreover $\frac{g}{f}$ has limit $\frac{l'}{l}$ at $a$ on this domain.*

*Proof.* This follows from Corollary (**??**) and Corollary (**??**). □

**Corollary 2.34.** *Every rational function is continuous on its domain.*

*Proof.* This follows from Corollary (**??**) and the continuity of polynomials. □

Since roots are Hölder functions, we conclude, by composition:

**Corollary 2.35.** *Every algebraic function is continuous over its domain.*

Algebraic manipulations have limitations, and a powerful tool can be used to handle many more situations.

**Theorem 2.36** (Squeeze Theorem)**.** *If $f, g, h : D \to \mathbb{R}$ are three functions, such that:*

*(1) $f(x) \leqslant g(x) \leqslant h(x)$ for all $x \in D$,*
*(2) $f$ and $h$ have limit $l \in \mathbb{R}$ at $a \in \mathrm{cl}(D)$,*

*then $g$ has limit $l$ at $a$.*

*Proof.* Let $\varepsilon > 0$. By Definition (**??**), there exists $\delta > 0$ such that, for all $x \in D$ with $|x-a| < \delta$, we have $|f(x) - l| < \varepsilon$; in particular, $f(x) \geqslant l - \varepsilon$. By Definition (**??**), there exists $\delta' > 0$ such that, for all $x \in D$ with $|x-a| < \delta$, we have $|h(x) - l| < \varepsilon$; in particular, $h(x) \leqslant l + \varepsilon$.

If $x \in D$ with $|x - a| < \min\{\delta, \delta'\}$, we have

$$l - \varepsilon < f(x) \leqslant g(x) \leqslant h(x) < l + \varepsilon,$$

and thus $|g(x) - l| < \varepsilon$, concluding our proof. □

**Corollary 2.37.** *If $f : D \to \mathbb{R}$, $l \in \mathbb{R}$, and $g : D \to [0, \infty)$ are such that:*

*(1) $|f(x) - l| \leqslant g(x)$ for all $x \in D$,*
*(2) $g$ has limit $0$ at $a \in \mathrm{cl}(D)$,*

*then $f$ has limit $l$ at $a$.*

*Proof.* By Theorem (**??**), the function $x \in D \mapsto |f(x) - l|$ has limit 0 at $a$, which is equivalent to $\lim_a f = l$. □

**Theorem 2.38.** *The functions* $\sin$, $\cos$, $\tan$, $\cot$, $\sec$ *and* $\csc$ *are continuous on their domains.*

*Proof.* Let $x \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$. Since a straight line is the shortest path between two points, $2|\sin(x)| \leqslant 2|x|$, i.e. $|\sin(x)| \leqslant |x|$. By Theorem (**??**), we conclude that $\lim_0 \sin = 0 = \sin(0)$.

Now, for $x \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$. Since $\cos(x) = \sqrt{1 - \sin^2(x)}$, by the continuity of the square root (as a Hölder function), by composition, linearity and multiplicativity, we conclude that

$$\lim_{x \to 0} \sqrt{1 - \sin^2(x)} = \sqrt{1 - 0} = 1.$$

Let now $x \in \mathbb{R}$. Since, for all $y \in \mathbb{R}$, writing $t = y - x$,

$$\begin{aligned} |\sin(x) - \sin(y)| &= |\sin(x) - \sin(x+t)| \\ &= |\sin(x) - \sin(x)\cos(t) - \cos(x)\sin(t)| \\ &\leqslant |\sin(x)(1 - \cos(t))| + |\sin(t)| \\ &\xrightarrow{t \to 0} 0. \end{aligned}$$

So $\lim_x \sin(y) = \sin(x)$, so sin is continuous over $\mathbb{R}$.

A similar computation shows that cos is continuous over $\mathbb{R}$. Application of Corollary (**??**) completes our proof. □

By composition, any algebraic function of trigonometric functions, including trigonometric polynomials, are continuous on their domains.

## 3. Derivation

### 3.1. The derivative.

**Definition 3.1.** The interior $\operatorname{int}(D)$ of a set $D$ is the set of all points $a \in D$ such that there exists $\delta > 0$ (depening on $a$) such that $(a-\delta, a+\delta)$ is a subset of $D$.

*Remark* 3.2. The complement of the interior of $D$ is a closed set, i.e. a set equal to its closure. In fact, the interior is exactly the complement of the closure of the complement of $D$. Sets which are equal to their interior are called open sets, they are exactly the complements of the closed sets.

**Definition 3.3.** Let $f : D \to \mathbb{R}$, and let $a \in \operatorname{int}(D)$. The *derivative of $f$ at $a$*, if it exists, is the number $f'(a)$ defined by:

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

*Remark* 3.4. A trivial computation — or the use of Theorem (**??**) — shows that if $f'(a)$ exists, then

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

**Theorem 3.5.** *If $f : D \to \mathbb{R}$ is differentiable at $a \in \operatorname{int}(D)$, then $f$ is continuous at $a$.*

*Proof.* Since $a \in \operatorname{int}(D)$, there exists $\delta > 0$ such that $(a-\delta, a+\delta) \subseteq D$. For all $x \in (a-\delta, a+\delta)$, we then have

$$\begin{aligned} f(x) &= f(x) - f(a) + f(a) \\ &= \frac{f(x) - f(a)}{x - a}(x - a) + f(a) \\ &\xrightarrow{x \to a} f'(a) \cdot 0 + f(a) = f(a). \end{aligned}$$

This completes our proof. □

**Definition 3.6.** The *domain of differentiability* of $f : D \to \mathbb{R}$ is the subset $U \subseteq \operatorname{int}(D)$ of $\operatorname{int}(D)$ consisting of all points in $\operatorname{int}(D)$ where $f$ is differentiable. The *derivative* of $f$ is the function $f : x \in U \mapsto f'(x)$.

*Remark* 3.7. A function is continuous on its domain of differentiability.

*Notation* 3.8. There are various notations used to designate the derivative of a function $f : D \to \mathbb{R}$. The most common is to write $f'$ for the derivative. Alternatively, following Newton, one could also write $\dot{f}$ for $f'$, which is still used at times when the variable is assumed to represent time. Leibniz' notation is also common. In that case, the value $f'(x)$ is denoted by $\frac{df(x)}{dx}$. This notation can be confusing, since the derivative is $x \mapsto \frac{df(x)}{dx}$ in that case. A common abuse of language is to write $\frac{df}{dx}$ for $f'$ but it is flawed (as it invokes the variable $x$, which has no intrinsic meaning), so we will refrain from this latter one, but $\frac{df(x)}{dx}$ for $f'(x)$ is at times quite helpful, when $x$ is clearly quantified.

**Lemma 3.9.** *The derivative of $x \mapsto \alpha x + \beta$ is the constant function $x \mapsto \alpha$.*

*Proof.* For all $a \in \mathbb{R}, x \in \mathbb{R} \setminus \{a\}$:

$$\frac{\alpha x + \beta - (\alpha a + \beta)}{x - a} = \alpha \frac{x-a}{x-a} = \alpha \xrightarrow{x \to a} \alpha.$$

□

### 3.2. Computation of the derivative.

**Theorem 3.10** (Linearity)**.** *Let $f : D \to \mathbb{R}$ and $g : D \to \mathbb{R}$. If $f$ and $g$ are differentiable at $x \in D$, and if $\lambda \in \mathbb{R}$, then $\lambda f + g$ is differentiable at $x$, and:*

$$(\lambda f + g)'(x) = \lambda f'(x) + g'(x).$$

*Proof.* Fix $x \in \operatorname{int}(D)$ a point where $f$ is differentiable. Let $t \in D, t \neq x$. We compute:

$$\begin{aligned}\frac{(\lambda f + g)(t) - (\lambda f + g)(x)}{t - x} &= \lambda \frac{f(t) - f(x)}{t - x} + \frac{g(t) - g(x)}{t - x} \\ &\xrightarrow{t \to x} \lambda f'(x) + g'(x).\end{aligned}$$

□

**Theorem 3.11** (Chain rule)**.** *If $f : D \to \mathbb{R}$ and $g : E \to \mathbb{R}$ are functions such that $f(D) \subseteq E$, $f$ is differentiable at $a \in D$, and $g$ is differentiable at $f(a) \in E$, then $g \circ f$ is differentiable at $a \in D$, and moreover:*

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

*Proof.* We define the following function $\varphi$ on $E$:

$$\varphi : y \in E \mapsto \begin{cases} \frac{g(y) - g(f(a))}{y - f(a)} & \text{if } y \neq f(a), \\ g'(f(a)) & \text{otherwise.} \end{cases}$$

Let $x \in D$. We compute:

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

By construction, $\lim_{y\to f(a)} \varphi(y) = g'(f(a))$, so $\varphi$ is continuous at $f(a)$. By Theorem (**??**), we conclude that $\lim_a \varphi \circ f = g'(f(a))$.

Hence, for all $x \neq a$:

$$\begin{aligned} \frac{g(f(x))-g(f(a))}{x-a} &= \frac{(f(x)-f(a))\varphi(f(x))}{x-a} \\ &= \frac{f(x)-f(a)}{x-a}\varphi(f(x)) \\ &\xrightarrow{x\to a} f'(a)g'(f(a)), \end{aligned}$$

and our result is proven.

□

**Lemma 3.12.** *The derivative of* $x \in \mathbb{R} \mapsto x^2$ *is* $x \in \mathbb{R} \mapsto 2x$.

*Proof.* Fix $x \in \mathbb{R}$. For all $y \in \mathbb{R} \setminus \{x\}$, we compute:

$$\begin{aligned} \frac{y^2-x^2}{y-x} &= \frac{(y-x)(y+x)}{y-x} \\ &= y+x \xrightarrow{y\to x} 2x. \end{aligned}$$

□

**Corollary 3.13** (Leibniz' rule)**.** *If* $f, g : D \to \mathbb{R}$ *are two differentiable functions at* $x \in D$*, then*

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

*Proof.* Fix $x \in \operatorname{int}(D)$ be a point where both $f$ and $g$ are differentiable at $x$. By linearity, $f+g$ and $f-g$ are both differentiable at $x$, and together with the chain rule, we get:

$$\left[(f+g)^2\right]'(x) = 2(f(x)+g(x))(f+g)'(x) = 2(f(x)+g(x))(f'(x)+g'(x)).$$

and

$$\left[(f-g)^2\right]'(x) = 2(f(x)-g(x))(f'(x)-g'(x)).$$

Hence, using linearity again:

$$\begin{aligned} (fg)'(x) &= \left[\frac{1}{4}\left((f+g)^2-(f-g)^2\right)\right]'(x) \\ &= \frac{1}{4}\left(\left[(f+g)^2\right]'(x) - \left[(f-g)^2\right]'(x)\right) \\ &= \frac{1}{4}\left(2(f(x)+g(x))(f'(x)+g'(x)) - 2(f(x)-g(x))(f'(x)-g'(x))\right) \end{aligned}$$

$$= f(x)g'(x) + f'(x)g(x).$$

□

**Corollary 3.14.** *For all $x \in \mathbb{R}$,*

$$\frac{dx^n}{dx} = nx^{n-1},$$

*so for any $f : D \to \mathbb{R}$, if $f$ is differentiable at $x \in D$, and for all $n \in \mathbb{N}$:*

$$\left(f^n\right)'(x) = nf'(x)f^{n-1}(x).$$

*Proof.* We could obtain the derivative of $x \mapsto x^n$ using algebra:

$$\frac{x^n - a^n}{x-a} = x^{n-1} + ax^{n-2} + \cdots + a^{n-1}x + a^n \xrightarrow{x\to a} na^{n-1}.$$

Alternatively, $\frac{dx}{dx} = 1$ and $\frac{dx^2}{dx} = 2x$. Assume that $\frac{dx^n}{dx} = nx^{n-1}$ for some $n$. Then $\frac{dx^{n+1}}{dx} = x\frac{dx^n}{dx} + x^n\frac{dx}{dx} = nx^n + x^n = (n+1)x^n$, so our resut follows by induction.

Applying the chain rule, we then get $\left(f^n\right)'(x) = nf^{n-1}(x)f'(x)$. □

**Lemma 3.15.** *For all $x \neq 0$, we have:*

$$\frac{d\frac{1}{x}}{dx} = -\frac{1}{x^2}.$$

*Proof.* Let $a \neq 0$ and $x \in \mathbb{R} \setminus \{a, 0\}$. We compute:

$$\begin{aligned}\frac{\frac{1}{x} - \frac{1}{a}}{x-a} &= \frac{\frac{a-x}{ax}}{x-a}\\ &= -\frac{1}{xa} \xrightarrow{x\to a} -\frac{1}{a^2}.\end{aligned}$$

□

**Corollary 3.16.** *If $f : D \to \mathbb{R}$ is differentiable at $x \in D$ and such that $f(x) \neq 0$, then:*

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

*Proof.* Let $x \in D$ so that $f(x) \neq 0$ and $f$ is differentiable at $x$. We compute, via the chain rule:

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

□

*Remark* 3.17. For all $x \in D$ with $f(x) \neq 0$, we compute: $(f^{-1})'(x) = (-1)f'(x)f^{-1-1}(x)$, so we can extend the power rule to negative powers.

**Corollary 3.18.** *If $f, g : D \to \mathbb{R}$ are two functions which are differentiable at $x \in D$ and such that $g(x) \neq 0$, then $\frac{f}{g}$ is well-defined in some open interval around a and differentiable at a, with:*

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

*Proof.* Since $g$ is differentiable at $a$, it is continuous, and thus $g(a) \neq 0$ implies $\lim_a g \neq 0$, so there exists $\delta > 0$ such that $g \neq 0$ on $(a-\delta, a+\delta)$. Then, using the Leibniz rule and the quotient rule at $x$:

$$\begin{aligned}\left(\frac{f}{g}\right)'(x) &= \left(f\frac{1}{g}\right)'(x)\\ &= f(x)\left(\frac{1}{g}\right)'(x) + f'(x)\frac{1}{g(x)}\\ &= \frac{f'(x)}{g(x)} + f(x)\frac{-g'(x)}{g^2(x)}\\ &= \frac{f'(x)g(x) - f(x)g'(x)}{g^2(x)}.\end{aligned}$$

□

**Theorem 3.19.** *If $f : D \to E$ is a bijection, and if $x \in E$ such that $f$ is differentiable at $f^{-1}(x)$ and $f'(f^{-1}(x)) \neq 0$, then $f^{-1}$ is differentiable at $x$, and:*

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

*In particular, $f^{-1}$ is continuous at $x$ as well.*

*Proof.* Since $f$ is injective, we can compute:

$$\begin{aligned}\frac{f^{-1}(x) - f^{-1}(a)}{x-a} &= \frac{f^{-1}(x) - f^{-1}(a)}{f(f^{-1}(x)) - f(f^{-1}(a))}\\ &\xrightarrow{x \to a} \frac{1}{f'(f^{-1}(a))}.\end{aligned}$$

□

**Theorem 3.20.** *For all $x > 0$ and $n \in \mathbb{N}$, we have $\frac{d\sqrt[n]{x}}{dx} = \frac{1}{n\sqrt[n]{x}^{n-1}}$.*

*Proof.* Set $f : x \in \mathbb{R} \mapsto x^n$, so $f'(x) = nx^{n-1}$, and since $\sqrt[n]{\cdot} = f^{-1}$:

$$\begin{aligned}\frac{d\sqrt[n]{x}}{dx} &= \frac{1}{f'(\sqrt[n]{x})}\\ &= \frac{1}{n\sqrt[n]{x}^{n-1}}.\end{aligned}$$

□

*Remark* 3.21. We thus have shown that $\frac{dx^{\frac{1}{n}}}{dx} = \frac{1}{n}x^{\frac{1}{n}-1}$. Now, if $p \in \mathbb{N}$, then

$$\frac{dx^{\frac{p}{n}}}{dx} = p\frac{dx^{\frac{1}{n}}}{dx}x^{\frac{p-1}{n}} = \frac{p}{n}x^{\frac{p}{n}-1}.$$

All in all, we thus have proven that for all $q \in \mathbb{Q}$ and $q > 0$, we have $\frac{dx^q}{dx} = qx^{q-1}$.

If $q < 0$, then

$$\frac{dx^q}{dx} = \frac{d\frac{1}{x^{-q}}}{dx} = (-1)\frac{dx^{-q}}{dx}\frac{-1}{x^{-2q}} = -(-q)x^{-q-1}\frac{1}{x^{-2q}} = qx^{q-1}.$$

In conclusion, for all $q \in \mathbb{Q}$, we have shown that the power rule holds:

$$\frac{dx^q}{dx} = qx^{q-1}.$$

*Remark* 3.22. It may be helpful to see a direct computation of the derivative of the square root. For all $x, a > 0$, $x \neq a$:

$$\begin{aligned}
\frac{\sqrt{x}-\sqrt{a}}{x-a} &= \frac{\sqrt{x}-\sqrt{a}}{x-a}\frac{\sqrt{x}+\sqrt{a}}{\sqrt{x}+\sqrt{a}} \\
&= \frac{x-a}{(x-a)(\sqrt{x}+\sqrt{a})} \\
&= \frac{1}{\sqrt{x}+\sqrt{a}} \\
&\xrightarrow{x\to a} \frac{1}{2\sqrt{a}} = \frac{1}{2}a^{-\frac{1}{2}}.
\end{aligned}$$

## 4. Global properties of continuity and derivation

4.1. **Two topological theorems.** Continuity is the core property in the study of topology, which then establishes various persistence results where continuous functions preserve, in some manner or another, certain properties of sets. We will give two of these results without proof.

The most important theorem of calculus is the following result. It is at the core of the proofs of the mean value theorem and the fundamental theorem of calculus.

**Theorem 4.1** (Extreme Value Theorem)**.** *If $f : [a,b] \to \mathbb{R}$ is a continuous function, then there exists $m, M \in [a,b]$ such that the range of $f$ is $[f(m), f(M)]$.*

The computation of ranges of function is quite difficult in general, but the following result is helpful (and technically is used in the proof of the extreme value theorem above).

**Theorem 4.2** (Intermediate Value Theorem)**.** *The continuous image of an interval is an interval: if $f : I \to \mathbb{R}$ is continuous and if $I$ is an interval, then $f(I)$ is an interval.*

### 4.2. **The Mean Value Theorem.**

**Definition 4.3.** A point $a \in \operatorname{int}(D)$ is a *local minimum* of $f : D \to \mathbb{R}$ if there exists $\delta > 0$ such that $f(x) \geqslant f(a)$ for all $x \in D \cap (a-\delta, a+\delta)$. A point $a \in \operatorname{int}(D)$ is a *local maximum* of $f : D \to \mathbb{R}$ when it is a local minimum of $-f$. A point $a \in \operatorname{int}(D)$ is a *local extremum* of $f : D \to \mathbb{R}$ when it is a local maximum or a local minimum of $f$.

**Lemma 4.4** (Fermat's Lemma)**.** *If $f : D \to \mathbb{R}$ is differentiable at $a \in \operatorname{int}(D)$, and if $a$ is a local extremum, then $f'(a) = 0$.*

*Proof.* Assume that $a$ is local minimum of $f$, and let $\delta > 0$ such that $f(x) \geqslant f(a)$ for all $x \in D \cap (a-\delta, a+\delta)$. If $x \in D \cap (a-\delta, a+\delta)$ and $x > a$ then

$$f(x) - f(a) \leqslant 0 \text{ so } \frac{f(x)-f(a)}{x-a} \leqslant 0, \text{ so } f'(a) \leqslant 0;$$

if $x \in D \cap (a-\delta, a+\delta)$ and $x < a$ then

$$f(x) - f(a) \leqslant 0 \text{ so } \frac{f(x)-f(a)}{x-a} \geqslant 0, \text{ so } f'(a) \geqslant 0.$$

So $f'(a) = 0$.

The result hold for local maximum by applying the above to $-f$. □

**Theorem 4.5** (Rolle's Theorem)**.** *If $f : [a,b] \to \mathbb{R}$ is continuous on $[a,b]$ and differentiable on $(a,b)$, and if $f(a) = f(b)$, then there exists $c \in (a,b)$ such that $f'(c) = 0$.*

*Proof.* By Theorem (**??**), since $f$ is continuous over $[a,b]$, there exists $m, M \in [a,b]$ such that for all $x \in [a,b]$, we have $f(m) \leqslant f(x) \leqslant f(M)$. If $\{m, M\} \subseteq \{a,b\}$, since $f(a) = f(b)$, we conclude that $f(x) = f(m)$ for all $x \in [a,b]$; in that case $f' = 0$ on $(a,b)$ and we can pick $c = \frac{a+b}{2}$ (or any other $c \in (a,b)$). Otherwise, $m$ or $M$ lies in $(a,b)$. By Lemma (**??**), if $m \in (a,b)$ then $f'(m) = 0$; otherwise if $M \in (a,b)$ then $f'(M) = 0$. □

The following result is the most important result about the application of derivatives.

**Theorem 4.6** (Mean Value Theorem)**.** *If $f : [a,b] \to \mathbb{R}$ is continuous on $[a,b]$, differentiable on $(a,b)$, then there exists $c \in (a,b)$ such that*

$$f(b) - f(a) = f'(c)(b-a).$$

*Proof.* Let $g : x \in [a,b] \mapsto f(x) - \frac{f(b)-f(a)}{b-a}(x-a)$. The function $g$ is continuous on $[a,b]$ and differentiable on $(a,b)$. Moreover, $g(a) = f(a) -$

$\frac{f(b)-f(a)}{b-a}(a-a) = f(a)$ and $g(b) = f(b) - \frac{f(b)-f(a)}{b-a}(b-a) = f(a)$, so $g(a) = g(b)$. Therefore, by Rolle's Theorem (**??**), there exists $c \in (a,b)$ such that $g'(c) = 0$. Now, $g'(x) = f'(x) - \frac{f(b)-f(a)}{b-a}$ for all $x \in (a,b)$. So $0 = f'(c) - \frac{f(b)-f(a)}{b-a}$, so $f'(c)(b-a) = f(b) - f(a)$ as claimed. □

**Corollary 4.7.** *Let $f : [a,b] \to \mathbb{R}$ be a continuous function over $[a,b]$, differentiable on $(a,b)$.*

1. *If $f' \geqslant 0$ on $(a,b)$, then $f$ is increasing on $[a,b]$.*
2. *If $f'(x) > 0$ for all $x \in (a,b)$, then $f$ is strictly increasing on $[a,b]$.*
3. *If $f' \leqslant 0$ on $(a,b)$, then $f$ is decreasing on $[a,b]$.*
4. *If $f'(x) < 0$ for all $x \in (a,b)$, then $f$ is strictly decreasing on $[a,b]$.*
5. *If $f' = 0$ on $(a,b)$, then $f$ is constant on $[a,b]$.*

*Proof.* Assume that $f'(x) \geqslant 0$ for all $x \in (a,b)$. Let $x, y \in [a,b]$ with $x < y$. By the Mean Value Theorem (**??**), there exists $c \in (a,b)$ such that $f(y) - f(x) = f'(c)(y-x) \geqslant 0$. So $f$ is increasing on $[a,b]$. If, moreover, $f'(x) > 0$ for all $x \in (a,b)$, then for all $x, y \in [a,b]$ with $x < y$, then, again by Theorem (**??**), there exists $c \in (a,b)$ such that $f(y) - f(x) = f'(c)(y-x) > 0$ so $f$ is strictly increasing on $[a,b]$.

The cases for negative derivatives follows similarly, or by applying the above to $-f$. If $f' = 0$, then $f$ is both increaisng and decreasing, so it is constant. Alternatively, one may apply Theorem (**??**) again: for all $x \in (a,b]$, there exists $c \in (a,x)$ such that $f(x) - f(a) = f'(c)(x-a) = 0$ so $f(x) = f(a)$ for all $x \in [a,b]$. □

### 4.3. Exponential and Logarithm Functions.

**Theorem-Definition 4.8.** *There exists a unique function* $\exp : \mathbb{R} \to \mathbb{R}$ *such that* $\exp' = \exp$ *and* $\exp(0) = 1$.

**Corollary 4.9.** *For all $x, y \in \mathbb{R}$, we have*

$$\exp(x+y) = \exp(x)\exp(y);$$

*in particular* $\exp(x) > 0$ *and* $\exp(-x) = \frac{1}{\exp(x)}$ *for all* $x \in \mathbb{R}$.

*Proof.* Fix $x \in \mathbb{R}$. Let $e_x : t \in \mathbb{R} \mapsto \exp(x+t)\frac{1}{\exp(x)}$. We have $e_x'(t) = \frac{1}{\exp(x)}\exp(x+t) \cdot 1 = e_x(t)$. So $e_x' = e_x$ and $e_x(0) = 1$, so $e_x = \exp$. Hence, $\exp(t) = \frac{\exp(x+t)}{\exp(x)}$ for all $t, x \in \mathbb{R}$, as claimed.

Then, for all $x \in \mathbb{R}$, we have $\exp(x)\exp(-x) = \exp(x-x) = \exp(0) = 1$, so $\exp(-x) = \frac{1}{\exp(x)}$ — in particular, $\exp(x) \neq 0$.

On the other hand, for all $x \in \mathbb{R}$, $\exp(x) = \exp(\frac{x}{2} + \frac{x}{2}) = \exp(\frac{x}{2})^2 \geqslant 0$. Together, we have seen that $\exp(x) > 0$. □

**Corollary 4.10.** *The function* exp *is strictly increasing on* $\mathbb{R}$*. Moreover, the range of* exp *is* $(0,\infty)$*. In particular,* exp *is a continuous bijection from* $\mathbb{R}$ *onto* $(0,\infty)$*.*

*Proof.* By Theorem (**??**), since $\exp' = \exp > 0$, we conclude that exp is strictly increasing.

Since $\mathbb{R}$ is an interval, and since exp is continuous, Theorem (**??**) implies that the range of exp is an interval. Since $\exp > 0$, the range is a subset of $(0,\infty)$.

Now, let $n \in \mathbb{N}$. Note that $\exp(n) = \exp(1)^n$. Since exp is increasing and $\exp(0) = 1$, we conclude that $\exp(1) > 0$. So $\exp(n)$ is unbounded, i.e., for any $x > 0$ there exists $n \in \mathbb{N}$ such that $\exp(n) > x$. Similarly, $\exp(-n) = \frac{1}{\exp(n)}$, so if $x > 0$ there exists $n \in \mathbb{N}$ such that $\exp(-n) < x$. So we conclude the range of exp is $(0,\infty)$. ☐

**Definition 4.11.** The inverse function of exp is the *natural logarithm* $\ln : (0,\infty) \mapsto \mathbb{R}$.

**Theorem 4.12.** *The function* ln *is differentiable on* $(0,\infty)$*, hence continuous on* $(0,\infty)$*, and* $\ln'(x) = \frac{1}{x}$ *for all* $x > 0$*; moreover* $\ln(1) = 0$ *and*

$$\ln(xy) = \ln(x) + \ln(y)$$

*for all* $x, y > 0$*.*

*Proof.* As the inverse function of a strictly increasing function, ln is also strictly increasing. Moreover, $\exp(0) = 1$ so $\ln(1) = 0$. Since $\exp' = \exp > 0$, we also conclude ln is differentiable over its domain $(0,\infty)$, and for $x > 0$:

$$\ln'(x) = \frac{1}{\exp'(\exp(x))} = \frac{1}{x}.$$

Last, let $x, y \in (0,\infty)$. Since $(0,\infty)$ is the range of exp, there exists $a, b \in \mathbb{R}$ such that $x = \exp(a)$ and $y = \exp(b)$. Then

$$\ln(xy) = \ln(\exp(a)\exp(b)) = \ln(\exp(a+b)) = a + b = \ln(x) + \ln(y).$$

☐

The exponential and logarithm functions are useful to define arbitrary powers. Let $n \in \mathbb{N}$ and $x \in (0,\infty)$. Then

$$x^n = \exp(\ln(x))^n = \underbrace{\exp(\ln(x))\cdots\exp(\ln(x))}_{n \text{ times}} = \exp(n\ln(x)).$$

Of course, $\frac{1}{x} = \frac{1}{\exp(\ln(x))} = \exp(-\ln(x))$. Last, let $r \in \mathbb{N}$. We then note that $\exp(\frac{1}{r}\ln(x))^r = \exp(\ln(x)) = x$. So $\exp(\frac{1}{r}\ln x) = x^{\frac{1}{r}}$. All in all, we can thus check that if $q \in \mathbb{Q}$ then $x^q = \exp(q\ln(x))$. This justifies the following.

**Definition 4.13.** If $a > 0$ and $x \in \mathbb{R}$, then $a^x = \exp(x\ln(a))$ .

**Theorem 4.14.** *Let $a > 0$. We have $a^0 = 1$, the function $x \in \mathbb{R} \mapsto a^x$ is strictly increasing, has range $(0,\infty)$, and for all $x, y \in \mathbb{R}$, we have $a^{x+y} = a^x a^y$; moreover this function is differentiable over $\mathbb{R}$ and*

$$\forall x \in \mathbb{R} \quad \frac{da^x}{dx} = \ln(a)a^x.$$

*The inverse function of $x \in \mathbb{R} \mapsto a^x$ is the* logarithm in base $a$, $\log_a : (0,\infty) \mapsto \mathbb{R}$*, which satisfies:*

*(1)* $\log_a(x) = \frac{\ln(x)}{\ln(a)}$ *for all $x \in \mathbb{R}$,*
*(2)* $\log_a(1) = 0$,
*(3) for all $x, y > 0$, we have* $\log_a(xy) = \log_a(x) + \log_a(y)$,
*(4)* $\frac{d\log_a(x)}{dx} = \frac{1}{x\ln(a)}$ *for all $x > 0$.*

*Proof.* All of these follow from an easy computation; for instance $a^x a^y = \exp(x\ln(a))\exp(y\ln(b)) = \exp((x+y)\ln(a))$ for all $x, y \in \mathbb{R}$, and

$$\frac{da^x}{dx} = \frac{d\exp(x\ln(a))}{dx} = \ln(a)\exp(x\ln(a)) = \ln(a)a^x.$$

A quick computation shows that $a^{\frac{\ln(x)}{\ln(a)}} = \exp(\ln(a)\frac{\ln(x)}{\ln(a)}) = \exp(\ln(x)) = x$ for all $x \in \mathbb{R}$, and similarly, $\log_a(a^x) = x$ for all $x \in \mathbb{R}$. From this, the rest of the computations are easy. □

*Notation* 4.15. We write $e := \exp(1)$, note that $e$ is a transcendantal number, with $e \approx 2.71828182846\ldots$. Thus $\exp(x) = e^x$ with the above notation. Importantly, $e^x$ only makes sense *after* we define the exponential function, for $x \in \mathbb{R} \setminus \mathbb{Q}$.

## 5. Convexity

5.1. **Convex Functions.** Fix $x < y$ in $\mathbb{R}$. If $a \in [x, y]$, then setting $t = \frac{a-x}{y-x}$, a basic computation shows that $t \in [0,1]$, and $a = tx + (1-t)y$. Conversely, if $t \in [0,1]$, then

$$x = tx + (1-t)x \leqslant tx + (1-t)y \leqslant ty + (1-t)y = y.$$

So $[x, y] = \{tx + (1-t)y : t \in [0,1]\}$. With this in mind:

**Definition 5.1.** A function $f : C \to \mathbb{R}$ is *convex* when $C$ is an interval and, for all $x, y \in C$ and $t \in [0,1]$:

$$f(tx + (1-t)y) \leqslant tf(x) + (1-t)f(y).$$

Note that to discuss convexity, we work on an interval, i.e. on a convex subset of real numbers. Geometrically, $f$ is convex when any secant to the graph of $f$ lies above the graph of $f$, per Definition (**??**).

Concavity is the opposite (not the complement!) of convexity; some functions are neither convex nor concave. We will discuss convexity in this section, and all the results about concave function can be trivially deduced from this simple definition.

**Definition 5.2.** A function $f : C \to \mathbb{R}$ is *concave* when $-f$ is convex.

A subset of the plane is *convex* when a line between any two points in the subset is entirely contained in that subset.

**Theorem 5.3.** *A function $f : C \to \mathbb{R}$ is convex if, and only if, the set*

$$\mathrm{epigraph}(f) := \{(x, y) : x \in D, y \geqslant f(x)\}$$

*is convex.*

*Proof.* Assume first that $f$ is convex. Let $(x, y)$ and $(w, z)$ in the epigraph of $f$. Let $t \in [0, 1]$. We then have by Definition (**??**):

$$t y + (1 - t)z \geqslant t f(x) + (1 - t) f(w) \geqslant f(t x + (1 - t)w)$$

so $(t x + (1 - t)w, t y + (1 - t)z)$ is also in the epigraph of $f$, so $f$ is convex.

Conversely, assume that the epigraph of $f$ is convex. Let $x, z \in D$ and $t \in [0, 1]$. Then, since $(x, f(x)$ and $(z, f(z))$ both are in the epigraph, by convexity, the segement from $(x, f(x))$ to $(z, f(z))$ is also in the epigragh. Therefore, by definition of the epigraph, $f(t x + (1 - t)z) \leqslant t f(x) + (1 - t) f(z)$, i.e. $f$ is convex by Definition (**??**). □

**Theorem 5.4.** *Let $f : D \to \mathbb{R}$, where $D$ is an interval. The following assertions are equivalent.*

- *the function $f : D \to \mathbb{R}$ is convex,*
- *for all $x, y, z \in D$, if $x < y < z$, then*
$$\frac{f(x) - f(y)}{x - y} \leqslant \frac{f(x) - f(z)}{x - z},$$
- *for all $x \in D$, and for all $y, z \in D \setminus \{x\}$, with $y < z$, we have*

$$\frac{f(x) - f(y)}{x - y} \leqslant \frac{f(x) - f(z)}{x - z}. \tag{5.1}$$

*Proof.* Assume that $f$ is convex. We first assume $x < y < z$. Let $t = \frac{z-y}{z-x}$. Note that $1 - t = \frac{y-x}{z-x}$ and $y = t x + (1 - t)z$. By convexity,

$$\begin{aligned} f(y) &= f(t x + (1 - t)z) \\ &\leqslant t f(x) + (1 - t) f(z) \\ &= \frac{z - y}{z - x} f(x) + \frac{y - x}{z - x} f(z). \end{aligned}$$

So

$$(z - x) f(y) \leqslant (z - y) f(x) + (y - x) f(z),$$

and thus

$$(z-x)(f(y)-f(x)) \leqslant (x-y)f(x)+(y-x)f(z)=(y-x)(f(z)-f(x)).$$

We therefore conclude:

$$\frac{f(y)-f(x)}{y-x} \leqslant \frac{f(z)-f(x)}{z-x}.$$

Assume now that $f$ satisfies our second assertion. Assume, morevoer, that $y<x<z$. Then by our assumption, we get:

$$\frac{f(x)-f(y)}{x-y} \leqslant \frac{f(z)-f(y)}{z-y},$$

so $(z-y)(f(x)-f(y)) \leqslant (x-y)(f(z)-f(y))$; therefore:

$$\begin{aligned}(z-x)(f(x)-f(y)) &\leqslant (x-y)(f(z)-f(y))+(y-x)(f(x)-f(y))\\ &=(x-y)(f(z)-f(y)+f(y)-f(x))\\ &=(x-y)(f(z)-f(x)),\end{aligned}$$

and thus we again obtain:

$$\frac{f(y)-f(x)}{y-x} \leqslant \frac{f(z)-f(x)}{z-x}.$$

Last, assume $y<z<x$. We can apply our assertion again to obtain:

$$\frac{f(z)-f(y)}{z-y} \leqslant \frac{f(y)-f(x)}{y-x}.$$

Then $(y-x)(f(z)-f(y)) \leqslant (z-y)(f(y)-f(x))$, so

$$(y-x)(f(z)-f(y)-(f(x)-f(y)) \leqslant (z-x)(f(y)-f(x))$$

so, noting that $y-x<0$,

$$\frac{f(z)-f(x)}{z-x} \geqslant \frac{f(y)-f(x)}{y-x},$$

and thus, we have proven that our second condition implies our third.

We note that it is trivial that our third assertion implies our second, since the latter is a special case of the former. So, unsurpisingly, we will only need our second assertion to prove convexity.

Now, assume that Expression (**??**) holds (we only it to hold for $x<y<z$ but we may as well assume it holds as stated). Let $x,z \in C$ and $t \in [0,1]$. Set $y=tx+(1-t)z$. We have

$$\begin{aligned}f(tx+(1-t)z) &= f(y)\\ &= f(y)-f(x)+f(x)\\ &= (y-x)\frac{f(y)-f(x)}{y-x}+f(x)\end{aligned}$$

$$\leqslant (y-x)\frac{f(z)-f(x)}{z-x}+f(x)$$
$$\leqslant (1-t)(z-x)\frac{f(z)-f(x)}{z-x}+f(x)$$
$$\leqslant (1-t)f(z)+(t-1)f(x)+f(x)=tf(x)+(1-t)f(z),$$

and therefore, $f$ is convex. □

Thus, if $s:(x,y)\mapsto\frac{f(y)-f(x)}{y-x}$, and noting that $s(x,y)=s(y,x)$ for all $x,y\in D$, Theorem (**??**) states that the convexity of $f$ is equivalent to $y\in D\mapsto s(x,y)$ is increasing for any fixed $x\in D$ (and thus by symmetry, so is $x\in D\mapsto s(x,y)$ for a fixed $y$).

**Corollary 5.5.** *A function $f:I\to\mathbb{R}$ is convex if, and only if, for all $x,y,z,t\in I$, with $y\leqslant z$, and $x\leqslant t$, $x\neq y$ and $z\neq t$:*

$$\frac{f(y)-f(x)}{y-x}\leqslant\frac{f(z)-f(t)}{z-t}. \tag{5.2}$$

*Proof.* Assume that $f$ is convex. By Theorem (**??**), we have:

$$\frac{f(y)-f(x)}{y-x}\leqslant\frac{f(z)-f(x)}{z-x}=\frac{f(x)-f(z)}{x-z}$$
$$\leqslant\frac{f(t)-f(z)}{t-z}.$$

Assume now that Expression (**??**) holds. Let $x\in D$, and let $y<z$ in $D$. Set $t=x$. Then

$$\frac{f(y)-f(x)}{y-x}\leqslant\frac{f(z)-f(x)}{z-x},$$

and thus $f$ is convex by Theorem (**??**). □

5.2. **Convexity and Extrema.**

**Theorem 5.6.** *If $f:C\to\mathbb{R}$ is convex, and if $x\in C$ is a local minimum of $f$, then $c$ is a global minimum of $f$.*

*Proof.* Let $\delta>0$ such that, if $y\in C$ and $|x-y|<\delta$, then $f(y)\geqslant f(x)$. Let $z\in C$. If $|z-x|<\delta$, then $f(z)\geqslant f(x)$. Assume now that instead, $|z-x|\geqslant\delta$. Define $w(t):=tx+(1-t)z$. We note that for $w(t)-x=(1-t)(z-x)$, so if $|z-x|\geqslant\delta$, then setting $t=1-\frac{\delta}{2|z-x|}$, we get $|w(t)-x|=\frac{\delta}{2}<\delta$. So $f(w(t))\geqslant f(x)$. Now, since $f$ is convex, by Theorem (**??**), if $z>x$ then $w(t)\leqslant z$ and :

$$\frac{f(z)-f(x)}{z-x}\geqslant\frac{f(w(t))-f(x))}{w(t)-x}$$

and the ratio on the right hand side is positive. Therefore $f(z)-f(x)\geqslant 0$ since $z>x$, and thus $f(z)\geqslant f(x)$.

If, instead, $z < x$, so $w(t) \geqslant z$, and then again by Theorem (**??**) since $f$ is convex, we have

$$\frac{f(z) - f(x)}{z - x} \leqslant \frac{f(w(t)) - f(x)}{w(t) - x}.$$

Thus time, $w(t) - x < 0$ and $f(w(t)) - f(x) \geqslant 0$ (as $x$ is a local minimum and $|w(t) - x| < \delta$), so $\frac{f(z)-f(x)}{z-x} \leqslant 0$. Since $z - x < 0$, we conclude that $f(z) \geqslant f(x)$.

We thus have shown that, for all $z \in C$, we have $f(z) \geqslant f(x)$, i.e. $x$ is a global minimum. □

5.3. **Convexity and Differentiability.**

**Theorem 5.7.** *If $f : (a, b) \to \mathbb{R}$ is differentiable over its domain, then $f$ is convex if, and only if, $f'$ is increasing on $(a, b)$.*

*Proof.* Fix $x < y$ in the domain of $f$. Let $\delta = \frac{y-z}{2} > 0$, and let $z \in (x-\delta, x+\delta)$ and $t \in (y - \delta, y + \delta)$. Note that by construction, $t > z$. Therefore, if $f$ is convex, then by Theorem (**??**):

$$\frac{f(z) - f(x)}{z - x} \leqslant \frac{f(t) - f(y)}{t - y}.$$

Thus, taking the limit as $z$ approaches $x$, and since $f$ is differentiable at $x$, we get:

$$f'(x) \leqslant \frac{f(t) - f(y)}{t - y},$$

for all $t, y > x$, $t \neq y$. Taking now the limit as $t$ goes to $y$, and since $f$ is differentiable at $y$, we conclude:

$$f'(x) \leqslant f'(y).$$

In other words, $f'$ is indeed increasing.

Assume now that $f'$ is increasing over $(a, b)$. Let $x < y < z$ in $D$. By the Mean Value Theorem **??**, there exists $c \in (x, y)$ and $d \in (y, z)$ such that $\frac{f(x)-f(y)}{x-y} = f'(c)$ and $\frac{f(y)-f(z)}{y-z} = f'(d)$. Since $f'$ is increasing and $c < d$, we conclude:

$$\frac{f(y) - f(x)}{y - x} \leqslant \frac{f(y) - f(z)}{z - y} = \frac{f(z) - f(y)}{z - y}.$$

We now proceed as in Theorem (**??**):

$$(f(y) - f(x))(z - y) = (f(y) - f(x))(z - x + x - y) \leqslant (f(z) - f(y))(y - x)$$

so

$$\begin{aligned}(f(y) - f(x))(z - x) &\leqslant (f(z) - f(y))(y - x) + (f(y) - f(x))(y - x) \\ &= (f(z) - f(x))(x - y)\end{aligned}$$

and therefore, $\frac{f(y)-f(x)}{y-x} \leqslant \frac{f(z)-f(x)}{z-x}$. So $f$ is convex by Theorem (**??**), as claimed. ☐

**Corollary 5.8.** *If $f : C \to \mathbb{R}$ is convex and differentiable on $(a,b) \subseteq C$, then $c \in (a,b)$ is a global minimum of $f$ if, and only if, $f'(c) = 0$.*

*Proof.* Let $c \in (a,b)$ with $f'(c) = 0$. By Theorem (**??**), it is sufficient to prove that $c$ is a local minimum. Now, since $f$ is convex, $f'$ is increasing on $(a,b)$ by Theorem (**??**). We conclude that $f' \leqslant 0$ on $(a,c)$, so by the Mean Value Theorem **??**, $f$ is decreasing on $(a,c)$, and so $f(x) \geqslant f(c)$ for all $x \in (a,c)$. On the other hand, $f' \geqslant 0$ on $(c,b)$ — again since $f'$ is increasing — so $f$ is increasing on $(c,b)$, so if $x > c$ then $f(x) \geqslant f(c)$. In summary, for all $x \in (a,b)$, we have $f(x) \geqslant f(c)$, i.e. $c$ is a local minimum of $f$.

Of course, by Lemma (**??**), if $c$ is a local minimum of $f$ then $f'(c) = 0$. ☐

**Corollary 5.9.** *If $f : (a,b) \to \mathbb{R}$ is convex and differentiable on $(a,b)$, then for all $x, y \in (a,b)$, we have:*

$$f(x) \geqslant f'(y)(x-y) + f(y),$$

*i.e. the graph of $f$ lies above its tangents.*

*Proof.* Assume first $x < y$. By the Mean Value Theorem (**??**), there exists $c \in (x,y)$ such that

$f(y)-f(x) = f'(c)(y-x) \leqslant f'(y)(y-x)$ since $f'$ is increasing by Theorem (**??**).

Hence $f(x) \geqslant f'(y)(x-y) + f(y)$ as claimed. The same reasoning applies for $x > y$. ☐

**Corollary 5.10.** *If $f : (a,b) \to \mathbb{R}$ is twice differentiable, then $f$ is convex if, and only if, $f'' \geqslant 0$.*

*Proof.* If $f$ is twice differentiable, then $f$ is in particular differentiable, and thus it is convex if, and only if, $f'$ is increasing, which is equivalent to its second derivative being positive. ☐

## 6. Taylor Theorem

If $f, g : I \to \mathbb{R}$ are both differentiable at $a$, if $f(a) = g(a) = 0$, and if $\frac{f'}{g'}$ has a limit at $a$, then

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

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

## 7. Limits of Sequences

**Definition 7.1.** The *sequence* $(x_n)_{n\in\mathbb{N}}$ is the function $n \in \mathbb{N} \mapsto x_n \in \mathbb{R}$.

**Definition 7.2.** A *subsequence* of a sequence $(x_n)_{n\in\mathbb{N}}$ is any sequence of the form $(x_{f(n)})_{n\in\mathbb{N}}$ for some *strictly increasing* function $f : \mathbb{N} \to \mathbb{N}$.

*Remark* 7.3. Think of a subsequence as taking the original sequence and *removing* entries, without reordering — just reindexing the entries — in such a way that we are still left with infinitely many entries.

**Definition 7.4.** A sequence $(x_n)_{n\in\mathbb{N}}$ *converges* to $l$ when:

$$\forall \varepsilon > 0 \quad \exists N \in \mathbb{N} \quad \forall n \in \mathbb{N} \quad n \geqslant N \implies |x_n - l| < \varepsilon.$$

*Remark* 7.5. Since $|a| = ||a| - 0|$ for any $a \in \mathbb{R}$, it is immediate that $(x_n)_{n\in\mathbb{N}}$ converges to $l$ if, and only if, $(|x_n - l|)_{n\in\mathbb{N}}$ converges to 0.

*Remark* 7.6. For any $l \in \mathbb{R}$, the constant sequence $(l)_{n\in\mathbb{N}}$ converges to $l$.

**Theorem 7.7.** *If $(x_{f(n)})_{n\in\mathbb{N}}$ is a subsequence of $(x_n)_{n\in\mathbb{N}}$, and if $(x_n)_{n\in\mathbb{N}}$ converges to $l$, then so does $(x_{f(n)})_{n\in\mathbb{N}}$.*

*Proof.* We first note that $f(0) \geqslant 0$ and if $f(n) \geqslant n$, then $f(n+1) > f(n) \geqslant n$; since $f(n+1) \in \mathbb{N}$, we conclude $f(n+1) \geqslant n+1$, so by induction, $f(n) \geqslant n$ for all $n \in \mathbb{N}$.

Let $\varepsilon > 0$. Since $(x_n)_{n\in\mathbb{N}}$ converges to $l$, there exists $N \in \mathbb{N}$ such that, if $n \geqslant N$, then $|x_n - l| < \varepsilon$. If $n \geqslant N$, then $f(n) \geqslant n \geqslant N$, and thus $|x_{f(n)} - l| < \varepsilon$. Our proof is complete. □

**Theorem 7.8** (Archimedean property)**.** *For any $\alpha > 0$, the sequence $\left(\frac{1}{(n+1)^\alpha}\right)_{n\in\mathbb{N}}$ converges to 0.*

*Proof.* Let $\varepsilon > 0$. There exists $N \in \mathbb{N}$ such that $\varepsilon^{\frac{1}{\alpha}} < N$. Therefore, $\frac{1}{N^\alpha} < \varepsilon$. So if $n \geqslant N$, we have $\frac{1}{(n+1)^\alpha} \leqslant \frac{1}{N^\alpha} < \varepsilon$. □

**Theorem 7.9.** *Boundedness If $(x_n)_{n\in\mathbb{N}}$ converges, then it is bounded.*

*Proof.* Assume $(x_n)_{n\in\mathbb{N}}$ converges to $l$. There exists $N \in \mathbb{N}$ such that, if $n \geqslant N$, then $|x_n - l| \leqslant 1$, so if $n \geqslant N$, then $|x_n| \leqslant |l| + 1$. Let $M := \max\{|l| + 1, |x_0|, \ldots, |x_N|\} \in \mathbb{N}$. By construction, $|x_n| \leqslant M$ for all $n \in \mathbb{N}$. □

**Theorem 7.10** (Squeeze)**.** *If $(y_n)_{n\in\mathbb{N}}$ converges to 0, and if there exists $N \in \mathbb{N}$ and $l \in \mathbb{R}$ such that*

$$\forall n \geqslant N \quad |x_n - l| \leqslant y_n$$

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

*Proof.* Observe that, if $n \geqslant N$, then $y_n \geqslant 0$.

Let $\varepsilon > 0$. Since $(y_n)_{n\in\mathbb{N}}$ converges to 0, there exists $N' \in \mathbb{N}$ such that, if $n \geqslant N'$, then $|y_n| < \varepsilon$. Let now $n \geqslant \max\{N, N'\}$. By assumption:

$$|x_n - l| \leqslant y_n < \varepsilon.$$

This concludes our proof. □

**Theorem 7.11** (Linearity)**.** *If $(x_n)_{n\in\mathbb{N}}$ converges to $l$ and $(y_n)_{n\in\mathbb{N}}$ converges to $l'$, then, for all $t \in \mathbb{R}$, the sequence $(tx_n + y_n)_{n\in\mathbb{N}}$ converges to $tl + l'$.*

*Proof.* The result is obvious if $t = 0$. We henceforth assume $t \neq 0$.

Let $\varepsilon > 0$. Since $(x_n)_{n\in\mathbb{N}}$ converges to $l$, there exists $N_1 \in \mathbb{N}$ such that, if $n \geqslant N_1$, then $|x_n - l| < \frac{\varepsilon}{2|t|}$. Since $(y_n)_{n\in\mathbb{N}}$ converges to $l'$, there exists $N_2 \in \mathbb{N}$ such that, if $n \geqslant N_2$, then $|y_n - l'| < \frac{\varepsilon}{2}$. Consequently, if $n \geqslant \max\{N_1, N_2\}$, then

$$|tx_n + y_n - (tl + l')| \leqslant |t||x_n - l| + |y_n - l'| < |t|\frac{\varepsilon}{2|t|} + \frac{\varepsilon}{2} = \varepsilon,$$

and our proof is complete. □

**Corollary 7.12** (Uniqueness of limits)**.** *If $(x_n)_{n\in\mathbb{N}}$ converges to both $l$ and $l'$, then $l = l'$.*

*Proof.* By linearity, $(0)_{n\in\mathbb{N}} = ((-1)x_n + x_n)_{n\in\mathbb{N}}$ converges to $l' - l$. If $l' - l \neq 0$, then $|0 - (l' - l)| = |l' - l| > 0$, which contradicts convergence [there exists $\varepsilon := |l' - l| > 0$ such that, for all $N \in \mathbb{N}$, there exists $n := N$ such that $|0 - (l' - l)| \not< \varepsilon$]. □

**Corollary 7.13.** *If $f : D \to \mathbb{R}$ is Lipschitz, and if $(x_n)_{n\in\mathbb{N}}$ is a sequence in $D$ which converges to $l \in D$, then $(f(x_n))_{n\in\mathbb{N}}$ converges to $f(l)$.*

*Proof.* Let $k > 0$ such that, for all $x, y \in D$, we have

$$|f(x) - f(y)| \leqslant k|x - y|.$$

Thus $|f(x_n) - f(l)| \leqslant k|x_n - l|$ for all $n \in \mathbb{N}$. Since $(|x_n - l|)_{n\in\mathbb{N}}$ converges to 0, linearity implies that $(k|x_n - l|)_{n\in\mathbb{N}}$ converges to 0 as well, and our conclusion follows from the Squeeze Theorem. □

**Corollary 7.14** (Multiplicativity)**.** *If $(x_n)_{n\in\mathbb{N}}$ converges to $l$ and if $(y_n)_{n\in\mathbb{N}}$ converges to $l'$, then $(x_n y_n)_{n\in\mathbb{N}}$ converges to $ll'$.*

*Proof.* Since $(x_n)_{n\in\mathbb{N}}$ converges, it is bounded. Let $M > 0$ such that, for all $n \in \mathbb{N}$, we have $|x_n| \leqslant M$.

For all $n \in \mathbb{N}$,

$$|x_n y_n - ll'| \leqslant |x_n||y_n - l'| + |x_n - l||l'| \leqslant M|y_n - l'| + |l'||x_n - l|.$$

By assumption, $(|x_n - l|)_{n\in\mathbb{N}}$ and $(|y_n - l'|)_{n\in\mathbb{N}}$ both converge to 0, so by linearity, $(M|y_n - l'| + |l'||x_n - l|)_{n\in\mathbb{N}}$ converges to $M0 + |l'|0 = 0$ as well. Our conclusion follows from the Squeeze Theorem. □

*Remark* 7.15. Since $x \mapsto x^2$ is Lipschitz on any bounded subset of $\mathbb{R}$ and since convergence sequences are bounded, we conclude that if $(x_n)_{n\in\mathbb{N}}$ converges to $l$, then $(x_n^2)_{n\in\mathbb{N}}$ converges to $l^2$. Therefore by linearity, if $(y_n)_{n\in\mathbb{N}}$ converges to $l'$, then $(x_n y_n)_{n\in\mathbb{N}} = \left(\frac{1}{4}\left((x_n+y_n)^2 - (x_n-y_n)^2\right)\right)_{n\in\mathbb{N}}$ converges to $ll'$.

**Corollary 7.16.** *If $f : D \to \mathbb{R}$ has a continuous derivative, $D$ is a bounded closed interval, then for any sequence $(x_n)_{n\in\mathbb{N}}$ converges to $l$, then $(f(x_n))_{n\in\mathbb{N}}$ converges to $f(l)$.*

**Corollary 7.17.** *If $(x_n)_{n\in\mathbb{N}}$ converges to $l \neq 0$, then $\left(\frac{1}{n}\right)_{n\in\mathbb{N}}$ and $(\ln(|x_n|))_{n\in\mathbb{N}}$ converge to $\frac{1}{l}$ and $\ln|l|$).*

**Theorem 7.18.** *Monotone Convergence Theorem If $(x_n)_{n\in\mathbb{N}}$ is monotone, then it converges if, and only if, it is bounded.*

## 8. Vector Spaces

**Definition 8.1.** A set $V$ endowed with two maps $+ : V \times V \to \mathbb{R}$ and $\cdot : \mathbb{R} \times V \to V$ is a *vector space* over $\mathbb{R}$ when all the following properties hold.

(1) $(u+v)+w = u+(v+w)$ for all $u, v, w \in V$,
(2) $u+v = v+u$ for all $u, v \in V$,
(3) there exists $0 \in V$ such that, for all $u \in V$, we have $0+u = u+0 = u$,
(4) for all $u \in V$, there exists $-v \in V$ such that $v+(-v) = 0$,
(5) $\lambda \cdot (u+v) = \lambda \cdot u + \lambda \cdot v$ for all $\lambda \in \mathbb{R}$ and for all $u, v \in V$,
(6) $(\lambda+\mu)\cdot v = \lambda \cdot v + \mu \cdot v$ for all $\lambda, \mu \in \mathbb{R}$ and $v \in V$,
(7) $(\lambda\mu)\cdot v = \lambda \cdot (\mu \cdot v)$,
(8) $1 \cdot v = v$ for all $v \in V$.

**Example 8.2.** If $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$, $\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in \mathbb{R}^n$ and $\lambda \in \mathbb{R}$, then we set:

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} := \begin{pmatrix} x_1+y_1 \\ \vdots \\ x_n+y_n \end{pmatrix} \text{ and } \lambda \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \lambda x_1 \\ \vdots \\ \lambda x_n \end{pmatrix}.$$

With these two operations, $\mathbb{R}^n$ is a vector space.

**Example 8.3** (Fundamental Example)**.** Let $S$ be a nonempty set, let $V$ be a vector space, and let $V^S$ be the set of all functions from $S$ to $V$. If $f, g : S \to V$ and $\lambda \in V$, then we define $f+g$ as the function:

$$x \in S \mapsto f(x)+g(x),$$

and the function $\lambda \cdot f$ as the function:

$$x \in S \mapsto \lambda f(x).$$

With these two operations, $V^S$ is a vector space.

**Theorem 8.4.** *Let $V$ be a vector space. The following algebraic rules hold.*

*(1) for all $u, v, w \in V$, if $u+v = u+w$, then $v = w$,*
*(2) $0 \cdot u = 0$ for all $u \in V$,*
*(3) if $u + v = u$ for some $u, v \in V$, then $v = 0$,*
*(4) if $u + v = 0$ for some $u, v \in V$, then $u = -v$,*
*(5) $(-1)v = -v$ for all $v \in V$,*
*(6) $n \cdot v = \underbrace{v + \cdots + v}_{n \text{ times}}$ and $-n \cdot v = \underbrace{(-v) + \cdots + (-v)}_{n \text{ times}}$ for all $n \in \mathbb{N}$.*

*Proof.* Assume $u + v = u + w$ for some $u, v, w \in V$. Since $V$ is a vector space, by Condition (4) in Definition (**??**), there exists $-u \in V$ such that $u + (-u) = 0$. Therefore:

$$\begin{aligned} v &= 0 + v \text{ (Cond. 2)} \\ &= (-u + u) + v \text{ (Cond 4)} \\ &= -u + (u + v) \text{ (Cond 1)} \\ &= -u + (u + w) \text{ (assumption on } u, v, w\text{)}, \\ &= (-u + u) + w \text{ (Cond 1)} \\ &= 0 + w \text{ (Cond 4)} \\ &= w \text{ (Cond. 2)} . \end{aligned}$$

This regularity property now implies all the other properties. If $v \in \mathbb{R}$ then:

$$\begin{aligned} 0 \cdot v + 0 &= 0 \cdot v \\ &= (0 + 0) \cdot v \\ &= 0 \cdot v + 0 \cdot v. \end{aligned}$$

By regularity, $0 = 0 \cdot v$.

Since $1 \cdot v = v$, we can now proceed by indction: if $n \cdot v = \underbrace{v + \cdot + v}_{n \text{ times}}$, then

$$(n+1) \cdot v = n \cdot v + 1 \cdot v = \underbrace{v + \cdot + v}_{n \text{ times}} + v = \underbrace{v + \cdot + v}_{n+1 \text{ times}}.$$

So our result holds by induction.

Now, if $u + v = u$ for some $u \in V$, then $u + v = u + 0$, and thus $v = 0$. Similarly, if $u + v = 0$ then $u + v = u + (-u)$ and thus $v = (-u)$.

If $v \in V$, then

$$\begin{aligned} 0 &= 0 \cdot v \\ &= (1 + (-1)) \cdot v \\ &= 1 \cdot v + (-1) \cdot v \end{aligned}$$

$$= v + (-1) \cdot v.$$

Therefore, $(-1) \cdot v = -v$. As a consequence, again by induction, $n \cdot v = \underbrace{(-v) + \cdots + (-v)}_{n \text{ times}}$. $\square$

## 9. Linear Maps

**Definition 9.1.** A *linear map* $\varphi : D \to C$ between two vector spaces $D$ and $C$ is a function form $D$ to $C$ such that, for any $u, v \in D$ and for any scalar $\lambda \in \mathbb{R}$,

$$\varphi(\lambda u + v) = \lambda \varphi(u) + \varphi(v).$$

**Exercise 9.2.** Which of the following functions is linear?

(1) $\varphi : \begin{pmatrix} x \\ y \end{pmatrix} \mapsto 3x + 4y.$

(2) $\varphi : \begin{pmatrix} x \\ y \end{pmatrix} \mapsto xy.$

(3) $\varphi : \begin{pmatrix} x \\ y \end{pmatrix} \mapsto x + 3y + 2.$

(4) $\varphi : \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} 2x + y \\ \pi x - 3y \end{pmatrix}.$

(5) $\varphi : \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} 3x + 4y \\ 3 \end{pmatrix}.$

(6) $\varphi : \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} p(x, y) \\ q(x, y) \end{pmatrix}$, where $p : \mathbb{R}^2 \to \mathbb{R}$ and $q : \mathbb{R}^2 \to \mathbb{R}$ are linear.

**Lemma 9.3.** *If $\varphi$ is linear, then $\varphi(0) = 0$.*

*Proof.* Note that

$$\varphi(0) + 0 = \varphi(0) = \varphi(0 + 0) = \varphi(1 \cdot 0 + 0) = 1\varphi(0) + \varphi(0).$$

Hence by regularity, $0 = \varphi(0)$. $\square$

**Lemma 9.4.** *If $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ and $v_1, \ldots, v_n \in V$ for some $n \in \mathbb{N} \setminus \{0\}$, then*

$$\varphi\left(\sum_{j=1}^{n} \lambda_j v_j\right) = \sum_{j=1}^{n} \lambda_j \varphi(v_j).$$

*Proof.* For $n = 1$, we have $u \in V$ and $\lambda \in \mathbb{R}$. Now

$$\varphi(\lambda u) = \varphi(\lambda u + 0) = \lambda \varphi(u) + \varphi(0) = \lambda \varphi(u) + 0 = \lambda \varphi(u).$$

Assume that $\varphi\left(\sum_{j=1}^{n} \lambda_j v_j\right) = \sum_{j=1}^{n} \lambda_j \varphi(v_j)$ for some $n \in \mathbb{N}$. Then

$$\varphi\left(\sum_{j=1}^{n+1} \lambda_j v_j\right) = \varphi\left(\lambda_1 v_1 + \sum_{j=2}^{n+1} \lambda_j v_j\right)$$

$$
\begin{aligned}
&= \lambda_1\varphi(v_1) + \varphi\left(\sum_{j=2}^{n+1}\lambda_j v_j\right)\\
&= \lambda_1\varphi(v_1) + \sum_{j=2}^{n+1}\lambda_j\varphi\left(v_j\right)\\
&= \sum_{j=1}^{n+1}\lambda_j\varphi(v_j).
\end{aligned}
$$

Our results holds by induction. □

**Corollary 9.5.** *If $\varphi:\mathbb{R}^n\to\mathbb{R}^m$ is linear, then*

$$\varphi\left(\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}\right) = \sum_{j=1}^{n} x_j\varphi(e_j)$$

*where* $e_j = \begin{pmatrix}0\\ \vdots\\ 1\\ \vdots\\ 0\end{pmatrix}$ *for each* $j\in\{1,\ldots,n\}$.

**Exercise 9.6.** Suppose that $\varphi:\mathbb{R}^3\mapsto\mathbb{R}^3$ is linear, and that

$$\varphi\left(\begin{pmatrix}1\\0\\0\end{pmatrix}\right) = \begin{pmatrix}2\\0\\0\end{pmatrix}, \varphi\left(\begin{pmatrix}0\\1\\0\end{pmatrix}\right) = \begin{pmatrix}2\\-3\\1\end{pmatrix} \text{ and } \varphi\left(\begin{pmatrix}0\\0\\1\end{pmatrix}\right) = \begin{pmatrix}1\\3\\-1\end{pmatrix}.$$

Compute:

(1) $\varphi\left(\begin{pmatrix}1\\2\\3\end{pmatrix}\right)$,
(2) $\varphi\left(\begin{pmatrix}\frac{3}{2}\\2\\2\end{pmatrix}\right)$,
(3) $\varphi\left(\begin{pmatrix}-1\\2\\3\end{pmatrix}\right)$,

**Exercise 9.7.**

**Theorem 9.8.** *If $\varphi,\psi: D\to C$ are two linear maps between vector spaces $D$ and $C$, and if $\lambda\in\mathbb{R}$, then:*

$$\lambda\varphi+\psi: u\in D\longmapsto \lambda\varphi(u)+\psi(u)$$

*is linear as well.*

*Proof.* Write $\theta := \lambda\varphi + \psi$. Let $u, v \in D$ and $\mu \in \mathbb{R}$.

$$\begin{aligned}\theta(\mu u + v) &= (\lambda\varphi + \psi)(\mu u + v)\\ &= \lambda(\varphi(\mu u + v)) + \psi(\mu u + v)\\ &= \lambda(\mu\varphi(u) + \varphi(v)) + \mu\psi(u) + \psi(v)\\ &= \mu(\lambda\varphi(u) + \psi(u)) + \lambda\varphi(v) + \psi(v)\\ &= \mu\theta(u) + \theta(v).\end{aligned}$$

This concludes our proof. □

**Theorem 9.9.** *If $\varphi : D \to C$ and $\psi : C \to V$ are two linear functions between vector spaces $D$, $C$ and $V$, then*

$$\psi \circ \varphi : x \in D \to \psi(\varphi(x))$$

*is linear from $D$ to $V$.*

*Proof.* Let $u, v \in D$ and $\lambda \in \mathbb{R}$.

$$\begin{aligned}\psi \circ \varphi(\lambda u + v) &= \psi(\varphi(\lambda u + v))\\ &= \psi(\lambda \underbrace{\varphi(u)}_{:=a} + \underbrace{\varphi(v)}_{:=b})\\ &= \psi(\lambda a + b)\\ &= \lambda\psi(a) + \psi(b)\\ &= \lambda\psi(\varphi(u)) + \psi(\varphi(v))\\ &= \lambda\psi \circ \varphi(u) + \psi \circ \varphi(v).\end{aligned}$$

Hence $\psi \circ \varphi$ is linear, as claimed. □

## 10. MATRICES

## 11. SUBSPACES

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

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF DENVER, DENVER CO 80208