# Mathematics

**Prof. Christophe BOURLIER**
Directeur de Recherche CNRS au laboratoire IETR, Site de Nantes

Courses
*MASTER WET (Wireless Embeded Technology)*

July 3, 2018

# Contents

# List of Figures

# 1 Function of a single variable and complex number

## 1.1 Limit

**Definition 1.1** *Suppose that $f$ a real function defined on $\mathbb{R}$ and $(x_0, L) \in \mathbb{R}^2$. It is said the limit of $f$, as $x$ approaches $x_0$, is $L$ and written*

$$\lim_{x \to x_0} f(x) = L,$$

*if the following property holds:*

*For every real $\epsilon > 0$, there exists a real $\delta > 0$ such that for all real $x$, $0 < |x - x_0| < \delta$ implies $|f(x) - L| < \epsilon$.*

**Definition 1.2** *Alternatively $x$ may approach $x_0$ from above (right) or below (left), in which case the limits may be written as*

$$\lim_{x \to x_0^+} f(x) = L,$$

*or*

$$\lim_{x \to x_0^-} f(x) = L,$$

*respectively. If these limits exist at $x_0$ and are equal there, then this can be referred to as the limit of $f(x)$ at $x_0$. If the one-sided limits exist at $x_0$, but are unequal, there is no limit at $x_0$ (the limit at $x_0$ does not exist). If either one-sided limit does not exist at $x_0$, the limit at $x_0$ does not exist.*

**Exercice 1.1** *Let $f$ be the function defined by $f(x) = x + \sqrt{x^2}/x$. Determine the domain of existence of $f$, named $D_f$, derive the limits of $f$ at $\{0^+, 0^-\}$ and simplify $f$ for $x \in D_f$.*

**Definition 1.3** *If a function $f$ is real-valued, then the limit of $f$ at $x_0$ is $L$ if and only if both the right-handed limit and left-handed limit of $f$ at $x_0$ exist and are equal to $L$.*

**Definition 1.4** *The function $f$ is continuous at $x_0$ if and only if the limit of $f(x)$ as $x$ approaches $x_0$ exists and is equal to $f(x_0)$.*

**Theorem 1.1** *If the limit of $f(x)$ as $x$ approaches $x_0$ is $L$ and the limit of $g(x)$ as $x$ approaches $x_0$ is $P$, then the limit of $f(x) + g(x)$ as $x$ approaches $x_0$ is $L + P$. If $a$ is a scalar, then the limit of $af(x)$ as $x$ approaches $x_0$ is $aL$.*

**Theorem 1.2** *If $f$ is a real-valued (or complex-valued) function, then taking the limit is compatible with the algebraic operations, provided the limits on the right sides of the equations below exist (the last identity only holds if the denominator is non-zero). This fact is often called the algebraic limit theorem.*

$$\lim_{x\to x_0} [f(x) \pm g(x)] = \lim_{x\to x_0} f(x) \pm \lim_{x\to x_0} g(x),$$
$$\lim_{x\to x_0} [f(x) \times g(x)] = \lim_{x\to x_0} f(x) \times \lim_{x\to x_0} g(x),$$
$$\lim_{x\to x_0} \frac{f(x)}{g(x)} = \frac{\lim_{x\to x_0} f(x)}{\lim_{x\to x_0} g(x)}.$$

**Theorem 1.3** *Limit of a composition of functions. If $(a, L, L') \in \mathbb{R}^3$, $\lim_{x\to a} f = L$ and $\lim_{x\to L} g = L'$, then $\lim_{x\to a} g(f(x)) = L'$.*

**Theorem 1.4** *Limit of a rational function. If and $n$ and $p$ are positive integers, then*

$$\lim_{x\to\pm\infty} \frac{a_1x + a_2x^2 + \ldots + a_nx^n}{b_1x + b_2x^2 + \ldots + b_px^p} = \lim_{x\to\pm\infty} \frac{a_nx^n}{b_px^p}.$$

**Exercice 1.2** *Show the above theorem. Derive the limits for the cases $n = p$, $n > p$ and $n < p$.*

The limits of indeterminate forms (sometimes difficult to derive) are

$$\boxed{\frac{0}{0}, \frac{\infty}{\infty}, 0 \times \infty, 1^{\infty}, +\infty - \infty}, \tag{1.1}$$

independently of the signs. The indetermination may be solved as:

1. a variable transformation on the function and the use of theorem 1.3,
2. a variable transformation on the value $x_0$, for which the limit is derived,
3. for a function with square roots, by multiplying the function by its conjugate expression,
4. using a Taylor series expansion of the function near $x_0$, for which the limit is derived,
5. using the Hospital'rule (particular case of a Taylor series expansion): If the expression $\frac{f(x)}{g(x)}$ has the form $\frac{\infty}{\infty}$ or $\frac{0}{0}$ for $x = a$ and if $g'(a) \neq 0$ (derivative of $g$ on $a$), then

$$\boxed{\lim_{x\to a} \frac{f(x)}{g(x)} = \lim_{x\to a} \frac{f'(a)}{g'(a)}}. \tag{1.2}$$

**Exercice 1.3** *Derive the following limits:*

$$\begin{cases} \lim_{x\to 2} \dfrac{\sqrt{2x} - 2}{\sqrt{x+1} - \sqrt{2x-1}} & \lim_{x\to 0} \dfrac{\sin(x)}{x} \\ \lim_{x\to 0^+} x \ln x & \lim_{x\to +\infty} \left(\dfrac{x+a}{x+b}\right)^x \end{cases}.$$

## 1.2 Derivative

| Function | Derivative |
|---|---|
| $\cos(x)$ | |
| $\sin(x)$ | |
| $\tan(x)$ | |
| $f(x) + g(x)$ | |
| $f(x)g(x)$ | |
| $f(x)/g(x)$ | |
| $\ln[f(x)]$ | |
| $\exp[f(x)]$ | |
| $[f(x)]^{g(x)}$ | |
| $\sqrt{f(x)}$ | |
| $[f(x)]^p$ | |

Table 1.1: Usual derivatives.

Differentiation is the action of computing a derivative. The derivative of a function $y = f(x)$ of a variable $x$ is a measure of the rate at which the value $y$ of the function changes with respect to the change of the variable $x$. It is called the derivative of $f$ with respect to $x$. If $x$ and $y$ are real numbers, and if the graph of $f$ is plotted versus $x$, the derivative is the slope of this graph at each point.

**Definition 1.5** *Derivative at the point $x_0$. A function $f$ has a derivative at $x_0$ if and only if the rate growth of $f$ at the point $x$ has a finite value when $x$ tends toward $x_0$. Then*

$$\boxed{f'(x_0) = \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} = \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h}}. \tag{1.3}$$

**Exercice 1.4** *Give a geometrical interpretation of the derivative and give the equation of the tangent at the point $a$ of the representative curve of $f$, named $C_f$.*

**Exercice 1.5** *From the definition of the derivative calculate the derivative of the function $f$ defined as $f(x) = \sqrt{x}$.*

**Definition 1.6** *Alternatively $x$ may approach $x_0$ from above (right) or below (left), in which case the limits may be written as*

$$f'(x_0) = \lim_{h\to 0^+} \frac{f(x_0+h) - f(x_0)}{h},$$

*or*

$$f'(a) = \lim_{h\to 0^-} \frac{f(x_0+h) - f(x_0)}{h},$$

*respectively. If these limits exist at 0 and are equal there, then this can be referred to the derivative of $f$ at $x_0$. If the one-sided limits exist at 0, but are unequal, the function is not derivable at $x_0$. If either one-sided limit does not exist at 0, the function is not derivable at $x_0$.*

**Theorem 1.5** *Increasing and decreasing of a function. If $f$ is a differentiable function on the range $[a;b]$, then $f$ is an* ***increasing*** *function on $[a;b]$, if and only if $\forall x \in [a;b]$, $f'(x) \geq 0$ and $f$ is an* ***decreasing*** *function on $[a;b]$, if and only if $\forall x \in [a;b]$, $f'(x) \leq 0$.*

If the inequality becomes strict, then $f$ is **strictly** an increasing and decreasing function.

If at the point $x_0$, the derivative exits and change of sign, then $f$ has an extremum (minimum or maximum) at $x_0$.

An inflexion point is a point, for which, the curve crosses its tangent.

**Theorem 1.6** *Inflexion point. If $f$ is a function two times differentiable on $[a;b]$, then $f$ is* ***convex*** *on $[a;b]$ if and only if $f''(x) \geq 0$. It is* ***concave*** *on $[a;b]$ if and only if $f''(x) \leq 0$. If $f''(x)$ vanishes and changes of sign at $x_0$, then $f$ has at $x_0$ an inflexion point.*

**Exercice 1.6** *Give the inflexion point of the function $f(x) = x^3$.*

**Exercice 1.7** *Fill the following table of the usual derivatives.* ***To know.***

Table 1.2 presents de different notations met in the books.

| Origin | Notation |
|---|---|
| Leibniz | $\frac{dy}{dx}$, $\frac{df}{dx}(x)$, $\frac{d}{dx}f(x)$ |
| Lagrange | $f'(x)$ |
| Newton | $\dot{y}$ |
| Euler | $D_x y$, $D_x f(x)$ |

Table 1.2: Different notations of the derivatives of the function $y = f(x)$.

## 1.3 Taylor series expansion

**Theorem 1.7** *The Taylor series expansion of a real function $f$ that is infinitely differentiable at a real $x_0$ is the power series*

$$\boxed{f(x) = f(x_0) + (x - x_0)f'(x_0) + \ldots + \frac{(x - x_0)^n}{n!} f^{(n)}(x_0) + o\left[(x - x_0)^n\right]}. \tag{1.4}$$

Notation of Landau : We note $y(x) = \mathrm{o}\left[(x - x_0)^n\right]$ if $\lim_{x \to x_0} \frac{y(x)}{(x - x_0)^n}$ vanishes.

The formula of Mac-Laurin is obtained for $x_0 = 0$. The table below gives the aylor series expansions of usual functions.

| Functions | aylor series expansion |
|---|---|
| $\cos(x)$ | $1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \ldots + (-1)^n \frac{x^{2n}}{(2n)!} + \mathrm{o}(x^{2n})$ |
| $\sin(x)$ | $x - \frac{x^3}{3!} + \frac{x^5}{5!} - \ldots + (-1)^n \frac{x^{2n+1}}{(2n+1)!} + \mathrm{o}(x^{2n+1})$ |
| $\exp(x)$ | $1 + x + \frac{x^2}{2!} + \ldots + \frac{x^n}{n!} + \mathrm{o}(x^n)$ |
| $\frac{1}{1+x}$ | $1 - x + x^2 - x^3 + \ldots + (-1)^n x^n + \mathrm{o}(x^n)$ |
| $\sqrt{1+x}$ | $1 + \frac{x}{2} - \frac{x^2}{8} + \frac{x^3}{16} + \mathrm{o}(x^3)$ |

Table 1.3: Taylor series expansion near zero of usual functions.

**Exercice 1.8** *Show the Taylor series expansions listed in above table 1.3.*

**Exercice 1.9** *Calculate the Taylor series expansion of the real function $f(x) = \sqrt{1 + \sqrt{1 + x}}$ at the order 2 and near $x = x_0 = 0$.*

**Exercice 1.10** *Calculate the Taylor series expansion of the real function $f(x) = \tan(x)$ at the order 3 and near $x = x_0 = 0$.*

## 1.4 Integral

Integrals appear in many practical situations. If a swimming pool is rectangular with a flat bottom, then from its length, width, and depth we can easily determine the volume of water it can contain (to fill it), the area of its surface (to cover it), and the length of its edge (to rope it). But if it is oval with a rounded bottom, all of these quantities call for integrals. Practical approximations may suffice for such trivial examples, but precision engineering (of any discipline) requires exact and rigorous values for these elements.

### 1.4.1 Geometrical interpretation

To start off, consider the curve $y = f(x)$ between $x = 0$ and $x = 1$ with $f(x) = \sqrt{x}$ (see figure 1.1). We ask: What is the area under the function $f$ and the equation $y = 0$, in the interval from 0 to 1? And call this (yet unknown) area the (definite) integral of $f$. The notation for this integral will be

$$\int_0^1 \sqrt{x}dx.$$

The exact value of this integral is $2/3 \approx 0.66667$.

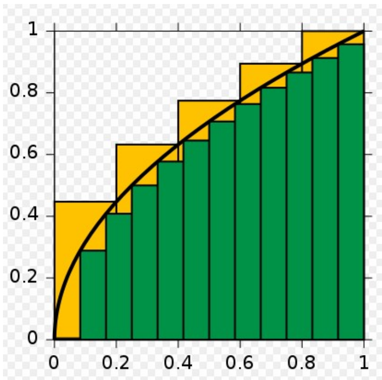



Figure 1.1: Approximations to integral of $\sqrt{x}$ from 0 to 1, with 5 rectangles (yellow) and 12 rectangles (green).

As a first approximation, look at the unit square given by the sides $x = 0$ to $x = 1$ and $y = f(0) = 0$ and $y = f(1) = 1$. Its area is exactly 1. As it is, the true value of the integral must be somewhat less than 1. Decreasing the width of the approximation rectangles and increasing the number of rectangles shall give a better result; so cross the interval in five steps, using the approximation points 0, 1/5, 2/5, and so on to 1. Fit a box for each step using the right end height of each curve piece, thus $\sqrt{1/5}$, $\sqrt{2/5}$, and so on to $\sqrt{1} = 1$. Summing the areas of these rectangles, we get a better approximation for the sought integral, namely

$$\sqrt{\frac{1}{5}}\left(\frac{1}{5} - 0\right) + \sqrt{\frac{2}{5}}\left(\frac{2}{5} - \frac{1}{5}\right) + \cdots + \sqrt{\frac{5}{5}}\left(\frac{5}{5} - \frac{4}{5}\right) \approx 0.7497.$$

We are taking a sum of finitely many function values of $f$, multiplied with the differences of two subsequent approximation points. We can easily see that the approximation is still too large. Using more steps produces a closer approximation, but will never be exact: replacing the **five** subintervals by **twelve** in the same way, but with the left end height of each piece, we will get an approximate value for the area of 0.6203, which is too small. The key idea is the transition from adding finitely many differences of approximation points multiplied by their respective function values to using infinitely many fine, or infinitesimal steps.

## 1.4.2 Riemann integral

**Definition 1.7** *The Riemann integral is defined in terms of Riemann sums of functions with respect to tagged partitions of an interval. Let* $[a, b]$ *be a closed interval of the real line; then a tagged partition of* $[a_1; a_2]$ *is a finite sequence (see figure 1.2)*

$$a_1 = x_0 \leq \xi_1 \leq x_1 \leq \xi_2 \leq x_2 \leq \cdots \leq x_{n-1} \leq \xi_n \leq x_n = a_2.$$

*This division into sub-intervals* $[x_{i-1}; x_i]$ *indexed by* $i$*, each of which is "tagged" with a distinguished point* $\xi_i \in [x_{i-1}; x_i]$*.* ***A Riemann sum*** *of a function* $f$ *with respect to such a tagged partition is defined as*

$$\sum_{i=1}^{n} f(\xi_i)\Delta_i = S_n$$

*Thus each term of the sum is the area of a rectangle with height equal to the function value at the distinguished point of the given sub-interval, and width the same as the sub-interval* $i$ *defined as* $\Delta_i = x_i - x_{i-1}$*.*

*Then, the limit of* $S_n$*, when* $n$ *tends toward the infinity, that is* $\Delta_i$ *tends to zero, is named* ***definite integral*** *of the function* $f$ *assumed to be continuous on the integration range* $[a_1; a_2]$*; it writes*

$$\boxed{I = \int_{a_1}^{a_2} f(x)dx = \lim_{n \to +\infty} \sum_{i=1}^{i=n} y_i \Delta x_i}, \tag{1.5}$$

*where* $dx$ *is the differential of the variable* $x$

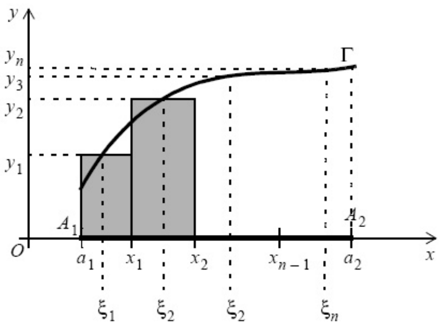



Figure 1.2: Definition of a defined integral.

## 1.4.3 Properties

If $f$ is a function defined and continued on $[a; b]$ $(a < b)$, then

$$\int_a^b f(x)dx = -\int_b^a f(x)dx.$$

If $c \in [a; b]$ and if $f$ is integrable on $[a; c]$ and $[c; b]$, then

$$\int_a^b f(x)dx = \int_a^c f(x)dx + \int_c^b f(x)dx.$$

If $\lambda \in \mathbb{R}$ (constant), then

$$\int_a^b \lambda f(x)dx = \lambda \int_a^b f(x)dx.$$

If $f$ is an even and odd fonction, respectively, then

$$\int_{-a}^a f(x)dx = 2\int_0^a f(x)dx \qquad \int_{-a}^a f(x)dx = 0.$$

If two functions $f$ and $g$ are integrable on $[a; b]$ then

$$\int_a^b [f(x) \pm g(x)]\, dx = \int_a^b f(x)dx \pm \int_a^b g(x)dx.$$

### 1.4.4 Primitive or antiderivative

Roughly speaking, the operation of integration is the reverse of differentiation. For this reason, the term integral may also refer to the related notion of the "antiderivative", a function $F$ whose derivative is the given function $f$. Here, the term "primitive" will be used.

**Definition 1.8** *Link between integral and primitive. We call primitive of the function $f$, assumed to be definite and continuous, any function $F$ which satisfies*

$$\boxed{F'(x) = \frac{dF}{dx} = f(x)}. \tag{1.6}$$

**Theorem 1.8** *If the function $F$ is a primitive of the function $f$, then*

$$\boxed{\int_a^b f(x)dx = F(b) - F(a)}. \tag{1.7}$$

**Exercice 1.11** *Fill the following table 1.4 of the usual primitives.* ***To know.***

### 1.4.5 Usual methods of integration

If the calculation of the indefinite (no limits) integral $\int f(x)dx$ is impossible, a variable transformation $x = g(t)$, where $dx = g'(t)dt$, can be used. Then

$$\boxed{\int f(x)dx = \int f\,[g(t)]\, g'(t)dt = \int \phi(t)dt}, \tag{1.8}$$

| Function | Primitive |
|---|---|
| $\cos(x)$ | |
| $\sin(x)$ | |
| $\frac{1}{x}$ | |
| $e^x$ | |
| $x^n$, $n > 0$ integer | |
| $\frac{1}{1+x^2}$ | |
| $\frac{2ax+b}{ax^2+bx+c}$, $(a \neq 0, b, c) \in \mathbb{R}^3$ | |
| $f'(x)f(x)^n$, $n > 0$ integer | |
| $\frac{f'(x)}{f(x)}$ | |

Table 1.4: Usual Primitives.

where the primitive of the function $\phi(t)$ is known. For a definite integral, we have

$$\boxed{\int_a^b f(x)dx = \int_\alpha^\beta f\left[g(t)\right] g'(t)dt = \int_\alpha^\beta \phi(t)dt \quad \text{with} \quad \left\{ \begin{array}{l} a = g(\alpha) \\ b = g(\beta) \end{array} \right.}. \tag{1.9}$$

**Exercice 1.12** *Calculate the integral* $I = \int_{-1}^{1} \sqrt{1-x^2}dx$. *We can set* $x = cos(t)$.

**Theorem 1.9** *Integration by parts. Let $u$ and $v$ be two derivative functions of $x$ such as $f(x) = u(x)v'(x)$, then*

$$\boxed{\int_a^b f(x)dx = \int_a^b u(x)v'(x)dx = \left[u(x)v(x)\right]_a^b - \int_a^b u'(x)v(x)dx}. \tag{1.10}$$

**Exercice 1.13** *Calculate the primitive of* $\ln x$ *for* $x \in ]0; +\infty[$.

### 1.4.6 Integration of a rational function

**Definition 1.9** *Rational function. A rational function is defined as the ratio of two polynomial functions:*

$$\frac{a_0 + a_1x + a_2x^2 + \ldots + a_nx^n}{b_0 + b_1x + b_2x^2 + \ldots + b_dx^d} = \frac{N(x)}{D(x)},$$

*The integer $n > 0$ stands for the degree of the numerator whereas the integer $d > 0$ stands for the degree of the denominator. In addition, the coefficients $a_{i\in[1;n]}$ and $a_{j\in[1;d]}$ are real.*

If $n \leq d$, then it is necessary to perform the Euclidean division of $N$ by $D$, using polynomial long division, giving $N(x) = E(x)D(x) + R(x)$ with $r < n$, in which $r$ is the degree of the polynomial function $R$. Dividing by $Q(x)$ this gives

$$\frac{N(x)}{D(x)} = E(x) + \frac{R(x)}{D(x)},$$

where $E$ is a polynomial function. A primitive of $E$ is then obtained from the idendity

$$\boxed{\int x^n dx = C + \frac{x^{n+1}}{n+1} \quad \text{with} \quad n+1 \neq 0 \quad \text{and} \quad C \in \mathbb{R}}. \tag{1.11}$$

The shape of the partial fraction decomposition of $R/D$ depends on the number and of the nature (complex or real) of the roots $x_i$ of $D$ ($D(x_i) = 0$, where $x_i$ can be a complex number).

#### 1.4.6.1 Case for which $D$ has $N$ single and real roots

If $D$ has $N$ **single** and **real** roots, then

$$\boxed{\frac{R(x)}{D(x)} = \sum_{n=1}^{n=N} \frac{A_n}{x - x_n}}. \tag{1.12}$$

The constant $A_n$ can be derived either by identifying the powers equal $x$ or by calculating the following limit

$$\lim_{x \to x_n} \frac{R(x)}{D(x)}(x - x_n) = A_n. \tag{1.13}$$

The primitive of $\frac{R(x)}{D(x)}$ is then

$$\boxed{\int \frac{R(x)}{D(x)} dx = C + \sum_{n=1}^{n=N} A_n \ln|x - x_n|}. \tag{1.14}$$

**Exercice 1.14** *Show that*

$$f(x) = \frac{2x^4 - 6x^3 + 7x^2 - 8x + 6}{x^2 - 3x + 2} = 2x^2 + 3 + \frac{x}{x^2 - 3x + 2}.$$

*Applying a partial fraction decomposition of $x/(x^2 - 3x + 2)$, calculate a primitive of $f$ over the domain of existence $D_f$.*

#### 1.4.6.2 Case for which $D$ has one multiple $M$ and real root

If $D$ has **one multiple** $M$ and **real** root, then

$$\boxed{\frac{R(x)}{D(x)} = \sum_{m=1}^{m=M} \frac{B_m}{(x-x_1)^m}}. \tag{1.15}$$

The constant $B_n$ can be derived either by identifying the powers equal $x$ or by calculating the following limit

$$\begin{aligned} B_M &= \lim_{x \to x_1} \left[ \frac{R(x)}{D(x)} (x-x_1)^M \right], \\ B_{M-1} &= \lim_{x \to x_1} \frac{d}{dx} \left[ \frac{R(x)}{D(x)} (x-x_1)^M \right], \\ B_{M-m} &= \frac{1}{m!} \lim_{x \to x_1} \frac{d^m}{dx^m} \left[ \frac{R(x)}{D(x)} (x-x_1)^M \right]. \end{aligned}$$

The primitive of $\frac{R(x)}{D(x)}$ is then

$$\boxed{\int \frac{1}{(x-x_1)^m} dx = C + \frac{(x-x_1)^{1-m}}{1-m} \quad \text{with} \quad m \neq 1 \quad \text{and } C \in \mathbb{R}}. \tag{1.17}$$

**Exercice 1.15** *Make a partial fraction decomposition of the function $f(x) = \frac{x}{(x-2)^2}$ and derive a primitive of $f$.*

#### 1.4.6.3 Case for which $D$ has $N$ single and complex roots

If $D$ has $N$ **single** and **complex** roots, then the roots are complex conjugates. For $N$ roots, $\{x_n = \alpha_n + j\beta_n, \bar{x}_n = \alpha_n - j\beta_n\}$, the partial fraction decomposition has the shape

$$\boxed{\frac{R(x)}{D(x)} = \sum_{n=1}^{n=N} \frac{xA_n + B_n}{(x-x_n)(x-\bar{x}_n)} = \sum_{n=1}^{n=N} \frac{xA_n + B_n}{x^2 + \alpha_n^2 + \beta_n^2 - 2x\alpha_n}}. \tag{1.18}$$

Moreover

$$x^2 + \alpha_n^2 + \beta_n^2 - 2x\alpha_n = (x-\alpha_n)^2 + \beta_n^2 = \beta_n^2 \left[ 1 + \left( \frac{x-\alpha_n}{\beta_n} \right)^2 \right].$$

The variable transformation $t = \frac{x-\alpha_n}{\beta_n}$ where $dt = \frac{dx}{\beta_n}$ leads to

$$\int \frac{R(x)}{D(x)} dx = \frac{1}{\beta_n} \int \frac{A_n(t\beta_n + \alpha_n) + B_n}{1+t^2} dt = A_n \int \frac{tdt}{1+t^2} + \frac{A_n\alpha_n + B_n}{\beta_n} \int \frac{dt}{1+t^2}.$$

Then

$$\boxed{\int \frac{R(x)}{D(x)} dx = C + \frac{A_n}{2} \ln \left| 1+t^2 \right| + \frac{A_n\alpha_n + B_n}{\beta_n} \arctan(t) \quad \text{avec} \quad t = \frac{x-\alpha_n}{\beta_n}}. \tag{1.19}$$

**Exercice 1.16** *Derive a primitive of the function $f(x) = \dfrac{x}{x^2 - 4x + 5}$.*

#### 1.4.6.4 Much more complicated cases

For more complicated cases, the resulting partial fraction decomposition is a linear combinaison of the previous cases. For example, if $D(x) = (x-1)(x-2)(x-3)^2(x^2+1)(x^2+4)^2$, the single real roots are $\{1, 2\}$, the one multiple real root is $x = 3$ (its multiplicity is 2), the one single complex roots are $x = \pm j$ and the one multiple complex roots are $x = \pm 2j$ (its multiplicity is 2), leading to the following partial fraction decomposition

$$\frac{R(x)}{D(x)} = \frac{A_1}{x-1} + \frac{A_2}{x-2} + \frac{B_1}{x-3} + \frac{B_2}{(x-3)^2}\frac{\alpha_1 + \beta_1 x}{x^2+1} + \frac{\alpha_2 + \beta_2 x}{x^2+4} + \frac{\alpha_3 + \beta_3 x}{(x^2+4)^2}.$$

The constants $\{A_1, A_2\}$ are calculated from Eq. (1.13), the constants $\{B_1, B_2\}$ are calculated from Eq. (1.16) and the constants $\{\alpha_i, \beta_i\}$ $(i = \{1, 2, 3\})$ by identifying the powers of $x$ and/or by taking limits as particular points.

### 1.4.7 Integration of a rational function of trigonometric functions

**Definition 1.10** *You want to derive the following indefinite integral*

$$\int R(\sin(x), \cos(x))dx = \int f(x)dx,$$

*where $R$ is a rational function.*

The use of rules's Bioche allows us to convert this integral into a conventional rational function. Then

1. If $f(-x)d(-x) = f(x)dx$, then the variable transformation $t = \cos(x)$ is applied.
2. If $f(\pi - x)d(\pi - x) = f(x)dx$, then the variable transformation $t = \sin(x)$ is applied.
3. If $f(\pi + x)d(\pi + x) = f(x)dx$, then the variable transformation $t = \tan(x)$ is applied.

In addition $d(\pi \pm x) = d\pi + d(\pm x) = d\pi \pm dx = \pm dx$ since $d\pi = 0$ ($\pi$ is a constant). In general, $df = f'(x)dx$ since $f'(x) = df/dx$.

It is always possible to apply the variable transformation $t = \tan\left(\frac{x}{2}\right)$ because

$$\boxed{\begin{cases} \sin(x) = \dfrac{2t}{1+t^2} \qquad \cos(x) = \dfrac{1-t^2}{1+t^2} \\ \\ dt = \dfrac{dx}{2}\left[1 + \tan^2\left(\frac{x}{2}\right)\right] \Rightarrow dx = \dfrac{2dt}{1+t^2} \end{cases}}. \tag{1.20}$$

Then, the integral is converted into the integral of a rational function over $t$.

If the rational function as only $\tan(x)$ terms, then the variable transformation $t = \tan(x)$ is applied and $dt = dx(1 + t^2)$ since $(\tan(x))' = 1 + \tan^2(x) = 1 + t^2$.

**Exercice 1.17** *Calculate the primitive of the function* $R(x) = \frac{1}{\sin(x)}$ *by using two methods:*

- *The Bioche rule.*
- *By setting* $t = \tan\left(\frac{x}{2}\right)$.

# 1.5 Ordinary differential equation

**Definition 1.11** *Ordinary differential equation (ODE) of* $n$*-th order. An Ordinary differential equation (ODE) of* $n$*-th order satisfies by the function* $y(x)$ *has the shape*

$$\boxed{F\left(x, y(x), \frac{dy}{dx}, \frac{d^2y}{dx^2}, \ldots, \frac{d^{n-1}y}{dx^{n-1}}\right) = 0}, \tag{1.21}$$

*where* $x$ *is the independent variable. The general solution is a function* $y = f(x, C_1, C_2, \ldots, C_n)$, *which depends on* $n$ ***arbitrary constants****. These constants are determined from particular conditions, which are the values taken by* $y$, $\frac{dy}{dx}$, $\frac{d^2y}{dx^2}$, $\ldots$, $\frac{d^{n-1}y}{dx^{n-1}}$ *for given values of* $x$*. If* $x = 0$, *the conditions are named* ***initial conditions****.*

*An* ***ordinary*** *differential equation (ODE) is an equation containing a function of one independent variable and its derivatives. The term "ordinary" is used in contrast with the term partial differential equation which may be with respect to more than one independent variable.*

An ODE is **linear** if the unknown function and its derivatives have degree 1 (products of the unknown function and its derivatives are not allowed) and nonlinear otherwise.

An ODE is **homogeneous** if the right-hand side of the equality is zero.

ODEs are described by their **order**, determined by the term with the highest derivatives. An equation containing only first derivatives is a first-order ODE, an equation containing the second derivative is a second-order ODE, and so on.

**Exercice 1.18** *Give a "name" (linear or not, order, homogeneous or inhomogeneous) to the following ODEs :*

1. $xy'(x) + y(x) = 0$.
2. $y'(x) + y(x) = 2x$.
3. $y''(x) + y^2(x) = 0$.
4. $y'(x) + \sin(y(x)) = \exp(x)$.

## 1.5.1 First-order linear ODE

**Definition 1.12** *First-order linear ODE. A first-order linear ODE is defined as*

$$\boxed{y'(x) + B(x)y(x) = \Phi(x)}. \tag{1.22}$$

The function $\Phi(x)$ stands for the right-hand side of the equality. The ODE $y'(x)+B(x)y(x) = 0$ is the homogeneous ODE, that is $\Phi = 0$.

The general solution can be written as

$$\boxed{y(x) = y_H(x) + y_P(x)}, \tag{1.23}$$

where $y_H$ is the **homogeneous** solution of the ODE and $y_P$ is the particular solution of the **inhomogeneous** ODE.

#### 1.5.1.1 Derivation of the homogeneous solution

By definition, $y_H$ satisfies

$$y_H'(x) + B(x)y_H(x) = 0,$$

which leads for $y_H \neq 0$ to

$$\frac{dy_H(x)}{y_H(x)} = -B(x)dx \Rightarrow \ln|y_H(x)| = -\int B(x)dx + C_1 \quad \text{with} \quad C_1 \in \mathbb{R}.$$

Then

$$\boxed{y_H(x) = C \times u(x) \quad \text{with} \quad u(x) = \exp\left(-\int B(x)dx\right) \quad \text{and} \quad C = \exp(C_1) \in \mathbb{R}}. \tag{1.24}$$

We can note that $y_H \neq 0$ because the exponential function differs from zero.

#### 1.5.1.2 Derivation of the particular solution

To find $y_P$, the method of the variation of the constant of Lagrange is applied. This consists in considering that the constant $C$ is an **unknown** solution of the variable variable $x$. Then

$$y_P(x) = C \times u(x) \to y_P(x) = C(x) \times u(x) \quad \text{and} \quad \begin{aligned} y_P'(x) &= [C(x) \times u(x)]' \\ &= C(x)u'(x) + u(x)C'(x). \end{aligned}$$

Reporting this equation in the **inhomogeneous** ODE, $y'(x)+B(x)y(x) = \Phi(x)$ with $y(x) = y_P(x)$, it leads to

$$C(x)\left[u'(x) + B(x)u(x)\right] + u(x)C'(x) = \Phi(x).$$

In addition, since the function $Cu(x)$ satisfied the **homogeneous** ODE, we have $u'(x) + B(x)u(x) = 0$, which implies that

$$C'(x) = \frac{\Phi(x)}{u(x)} \Rightarrow C(x) = \int \frac{\Phi(x)}{u(x)}dx + K \quad \text{with} \quad K \in \mathbb{R}.$$

Then

$$\boxed{y_P(x) = C(x)u(x) = \left[\int \frac{\Phi(x)}{u(x)}dx + K\right] u(x)}. \tag{1.25}$$

| Condition | Solution of the characteristic equation | $y_H(x)$ |
|---|---|---|
| $\Delta > 0$ | 2 real roots: $r_{1,2} = \dfrac{-b \pm \sqrt{\Delta}}{2a} = -\lambda \pm \Omega$ with $\lambda = \dfrac{b}{2a}$ et $\Omega = \dfrac{\Delta}{2a}$ | $C_1 e^{r_1 x} + C_2 e^{r_2 x} = e^{-\lambda x}(C_1 e^{\Omega x} + C_2 e^{-\Omega x})$ |
| $\Delta = 0$ | 1 double real root: $r_1 = r_2 = -\dfrac{b}{2a} = -\lambda$ | $e^{-\lambda x}(C_1 + C_2 x)$ |
| $\Delta < 0$ | 2 complex conjugate roots: $r_{1,2} = \dfrac{-b \pm j\sqrt{-\Delta}}{2a} = -\lambda \pm j\omega$ with $\lambda = \dfrac{b}{2a}$ et $\omega = \dfrac{\sqrt{-\Delta}}{2a}$ | $e^{-\lambda x}\left[C_1 \cos(\omega x) + C_2 \sin(\omega x)\right]$ |

Table 1.5: Solution of $y_H$ for a second-oder ODE with constant coefficients.

#### 1.5.1.3 General solution

The General solution is then

$$\begin{aligned} y(x) &= y_H(x) + y_P(x) = Cu(x) + \left[\int \frac{\Phi(x)}{u(x)}dx + K\right] u(x) \\ &= \left[\int \frac{\Phi(x)}{u(x)}dx + K_1\right] u(x) \quad \text{with} \quad K_1 = (C+K) \in \mathbb{R} \end{aligned} \tag{1.26}$$

**Exercice 1.19** *Solve the following ODE* $(1+x^2)y'(x) + xy(x) = x$.

### 1.5.2 Second-order linear ODE

**Definition 1.13** *Second-order linear ODE. A second-order linear ODE is defined as*

$$\boxed{a(x)y''(x) + b(x)y'(x) + c(x)y(x) = \Phi(x)}, \tag{1.27}$$

*where* $a$, $b$, $c$ *and* $\Phi$ *are four known functions. The homogeneous ODE is defined as* $a(x)y''(x) + b(x)y'(x) + c(x)y(x) = 0$, *that is* $\Phi(x) = 0$.

As previously, the general solution can be written as

$$\boxed{y(x) = y_H(x) + y_P(x)}, \tag{1.28}$$

where $y_H$ is the **homogeneous** solution of the ODE and $y_P$ is the particular solution of the **inhomogeneous** ODE.

To illustrate the method of the variation of the constant of Lagrange and to simplify the problem, we assume that the functions $(a, b, c)$ are constants (independent of the variable $x$).

| Expression of $\Phi(x)$ | Particular solution |
|---|---|
| $\Phi(x) = P(x)$ where $P$ is a polynomial function of degree $n$. | $y_P$ is a polynomial function of degree<br>- $n$, if $c \neq 0$.<br>- $n+1$, if $c = 0$ and $b \neq 0$.<br>- $n+2$, if $c = 0$ and $b = 0$. |
| $\Phi(x) = \Phi_0 e^{\beta x}$ where $\Phi_0 \in \mathbb{R}$. | If $\beta$ is not a root of the haracteristic equation, then $y_P(x) = Ce^{\beta x}$.<br>If $\beta$ is a single root, then $y_P(x) = Cxe^{\beta x}$.<br>If $\beta$ is a double root, then $y_P(x) = Cx^2 e^{\beta x}$. |
| $\Phi(x) = \Phi_1 \cos(\beta x) + \Phi_2 \sin(\beta x)$. | $y_P(x) = C_1 \cos(\beta x) + C_2 \sin(\beta x)$. |
| $\Phi(x) = P(x)e^{\beta x}$, where $P$ is a polynomial function of degree $n$. | $y_P(x) = x^k Q(x) e^{\beta x}$, where $Q$ is a polynomial function of degree $n$ and<br>- $k = 0$, if $\beta$ is not a root of the characteristic equation.<br>- $k = 1$, if $\beta$ is a single root.<br>- $k = 2$, if $\beta$ is a double root. |

Table 1.6: Particular solution for a second-order ODE with constant coefficients.

#### 1.5.2.1 Derivation of the homogeneous solution

By definition $y_H$ is defined as

$$ay''_H(x) + by'_H(x) + c = 0.$$

The solution has the form $y_H(x) = e^{rx}$ where $r \in \mathbb{R}$. The homogeneous ODE then becomes $e^{rx}(ar^2 + br + c) = 0$. As $e^{rx}$ differs from zero, the previous relation is satisfied for any $x$ if $r$ is a root of the following second-degree equation

$$ar^2 + br + c = 0,$$

named **characteristic equation**. Three cases can be distinguished versus the sign of the discriminant $\Delta^2 = b^2 - 4ac$ (table 1.5).

In general, the homogeneous solution can be written as $y_H(x) = C_1 y_1(x) + C_2 y_2(x)$, where $C_1$ and $C_2$ are two constants and $y_1(x)$, $y_2(x)$ are two solutions **linearly dependent** of the homogeneous solution. The particular solution $y_P(x)$ is found from the method of the variation of the constant of Lagrange, which states that the constant $C_1$ and $C_2$ are functions that depend on the variable $x$. Then, we can show that the functions $C_1(x)$ and $C_2(x)$ satisfy the ODE system of two unknowns defined as

$$\begin{cases} y_1(x)C'_1(x) + y_2(x)C'_2(x) = 0 \\ a\left[y'_1(x)C'_1(x) + y'_2(x)C'_2(x)\right] = \Phi(x) + c \end{cases}.$$

With respect the expression of $\Phi(x)$, table 1.6 is obtained.

This method can be extended if the constants $(a, b, c)$ become functions of the variable $x$.

## 1.6 Complex number

**Definition 1.14** *A complex number is a number that can be expressed in the form* $a + bi = a + ib = a + jb$*, where* $a$ *and* $b$ *are real numbers and* $(i, j)$ *is the imaginary unit, satisfying the equation* $i^2 = j^2 = -1$*. In this expression,* $a$ *is the* ***real*** *part and* $b$ *is the* ***imaginary*** *part of the complex number. If* $z = a + bi$*, then we write* $Re(z) = a$ *and* $Im(z) = b$*.*

*Complex numbers extend the concept of the one-dimensional number line to the two-dimensional complex plane by using the horizontal axis for the real part and the vertical axis for the imaginary part. The complex number* $a + bi$ *can be identified with the point* $(a, b)$ *in the complex plane (see figure 1.3). A complex number whose real part is zero is said to be* ***purely imaginary****, whereas a complex number whose imaginary part is zero is a real number.*

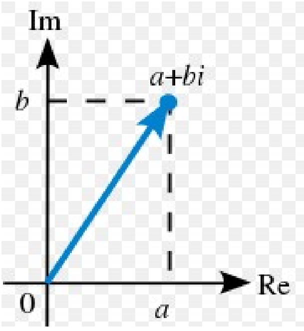



Figure 1.3: A complex number $z = a + bi$ can be visually represented as a pair of numbers $(a, b)$ forming a vector in the Cartesian system $(x, y)$.

Complex numbers allow solutions to certain equations that have no solutions in real numbers. For example, the equation $(x + 1)^2 = -9$ has no real solution, since the square of a real number cannot be negative. Complex numbers provide a solution to this problem. The idea is to extend the real numbers with the imaginary unit $i$ where $i^2 = -1$, so that solutions to equations like the preceding one can be found. In this case the solutions are $-1 + 3i$ and $-1 - 3i$.

**Definition 1.15** *Polar representation. A complex number* $z = x + iy$ *may also be defined in terms of its magnitude* $r = |z|$ *(distance) and direction relative to the origin* $\phi = \arg(z)$*. These are emphasized in a complex number's* ***polar form*** *(see figure 1.4). The angle* $\phi$ *is obtained by calculating the argument, named* arg*, of* $z$ *and the amplitude* $r$ *by calculating the modulus, named* $||$*, of* $z$*. Then,*

$$z = |z|e^{j\arg(z)} = re^{j\phi} = r\left[\cos(\phi) + j\sin(\phi)\right],$$

*where*

$$r = \sqrt{x^2 + y^2} \quad \cos\phi = \frac{x}{r} \quad \sin\phi = \frac{y}{r}.$$

**Definition 1.16** *Complex conjugate. The complex conjugate of the complex number* $z = x + yi$ *is defined to be* $x - yi$*. It is denoted by either* $\bar{z}$ *or* $z^*$*. Formally, for any complex number z:*

$$\bar{z} = Re(z) - Im(z)i.$$

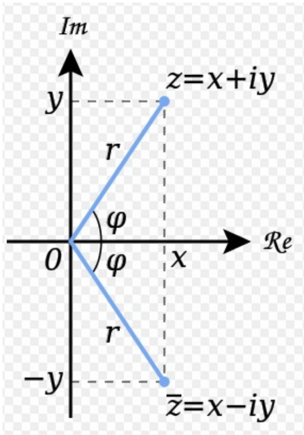



Figure 1.4: A complex number $z = x + yi = re^{j\varphi}$ in polar coordinates.

Geometrically $\bar{z}$ is the "reflection" (see Fig. 1.4) of $z$ about the real axis. Conjugating twice gives the original complex number: $\bar{\bar{z}} = z$.

The real and imaginary parts of a complex number $z$ can be extracted using the conjugate:

$$\text{Re}(z) = \frac{1}{2}(z + \bar{z}) \quad \text{Im}(z) = \frac{1}{2i}(z - \bar{z})$$

Moreover, a complex number is real if and only if it equals its conjugate.

Some properties. **To know**.

$$\overline{z_1 \pm z_2} = \bar{z}_1 \pm \bar{z}_2.$$
$$\overline{z_1 z_2} = \bar{z}_1 \bar{z}_2.$$
$$\overline{\frac{z_1}{z_2}} = \frac{\bar{z}_1}{\bar{z}_2}.$$

$$z_1 \pm z_2 = (a_1 + b_1 i) \pm (a_2 + b_2 i) = (a_1 \pm a_2) + (b_1 \pm b_2)i.$$

$$\begin{aligned} z_1 z_2 &= (a_1 + b_1 i)(a_2 + b_2 i) \\ &= a_1 a_2 + a_1 b_2 i + b_1 a_2 i - b_1 b_2 \\ &= a_1 a_2 - b_1 b_1 + (a_1 b_2 + b_1 a_2)i, \end{aligned}$$

then $\text{Re}(z_1 z_2) \neq \text{Re}(z_1)\text{Re}(z_2)$ and $\text{Im}(z_1 z_2) \neq \text{Im}(z_1)\text{Im}(z_2)$.

$$\begin{aligned} \frac{z_1}{z_2} &= \frac{a_1 + b_1 i}{a_2 + b_2 i} \\ &= \frac{(a_1 + b_1 i)(a_2 - b_2 i)}{(a_2 + b_2 i)(a_2 - b_2 i)} \\ &= \frac{a_1 a_2 - a_1 b_2 i + b_1 a_2 i + b_1 b_2}{a_2^2 + b_2^2} \\ &= \frac{a_1 a_2 + b_1 b_2}{a_2^2 + b_2^2} + \frac{b_1 a_2 - a_1 b_2}{a_2^2 + b_2^2} i \end{aligned}$$

In polar representation, $z_1 = r_1 e^{j\phi_1}$ and $z_2 = r_2 e^{j\phi_2}$, we have

$$z_1 z_2 = r_1 r_2 e^{j(\phi_1 + \phi_2)} \tag{1.29}$$

$$\frac{z_1}{z_2} = \frac{r_1}{r_2} e^{j(\phi_1 - \phi_2)} \tag{1.30}$$

**Exercice 1.20**

1. *Calculate the real and imaginary parts of* $\frac{1+2i}{1-3i}$.
2. *Calculate the modulus and the argument of* $1+i$, $-1+i$, $(1+i)(-1+i)$ *and* $(1+i)/(-1+i)$.
3. *Give the real and imaginary part of* $1/(a+z)$ *where* $z = x + yi$ *and* $(a, x, y) \in \mathbb{R}^3$.
4. *Show that* $z_1 z_2 = r_1 r_2[\cos(\phi_1 + \phi_2) + \sin(\phi_1 + \phi_2)]$. *Representing then* $z_1 z_2$ *in the complex plane from* $z_1$ *and* $z_2$.
5. *Calculate* $|z|$ *where* $z = a + be^{j\phi}$ *and* $(a, b) \in \mathbb{R}^2$.
6. *Show that* $[\cos(\phi) + \sin(\phi)i]^n = \cos(n\phi) + \sin(n\phi)i$.
7. *We set* $\sqrt{x + yi} = \pm(a + bi)$. *Calculate* $a$ *and* $b$ *from* $x$ *and* $y$, *where* $(x, y, a, b) \in \mathbb{R}^4$.

# 1.7 Homework

Make the exercises on a copy with a **clean** presentation and **underline** the final results. Do not forgot to write your name and surname on all sheets. The copy will be read by the Professor in the next session (course).

## 1.7.1 Limit

Calculate the following limits:

1. $\lim_{x \to 0} \frac{1 - \cos x}{x(2 - x)\tan(2x)}$.
2. $\lim_{x \to 0} \frac{(1 - e^x)\sin x}{x^2 + x^3}$.
3. $\lim_{x \to \frac{1}{2}} (2x^2 - 3x + 1)\tan(\pi x)$.
4. $\lim_{x \to 0} \frac{\ln[\cos(ax)]}{\ln[\cos(bx)]}$ avec $a$ et $b$ reals.
5. $\lim_{x \to 0} \frac{a^x - b^x}{x}$ avec $(a, b) \in \mathbb{R}^{+,*} \times \mathbb{R}^{+,*}$.

## 1.7.2 Taylor series expansion

Calculate the following Taylor series expansion:

1. $f(x) = \dfrac{x}{e^x - 1}$ near 0 and at the order 2.

2. $f(x) = e^{\cos x}$ near 0 and at the order 3.

3. $f(x) = (1 + x)^{\frac{1}{x}}$ near 0 and at the order 2.

4. $f(x) = \dfrac{x}{\sin x}$ near 0 and at the order 3.

### 1.7.3 Indefinite integral: Primitive

Let $D_f$ be the domain of existence of $f$ defined as $F(x) = \int f(x)dx$. Calculate $D_f$ and $F$ of:

1. $\displaystyle\int x \ln x dx.$

2. $\displaystyle\int \ln\left(x + \sqrt{a^2 + x^2}\right) dx$ with $a \neq 0$.

3. $\displaystyle\int \frac{dx}{a^2 - x^2}.$

4. $\displaystyle\int \frac{dx}{(a - x)(b - x)^2} dx.$

5. $\displaystyle\int \frac{\sin^2 \theta d\theta}{\cos \theta}.$

6. $\displaystyle\int \tan^2 \theta d\theta.$

7. $\displaystyle\int \sin^3 \theta \cos^2 \theta d\theta.$

### 1.7.4 Definite integral

Calculate the following definite integrals:

1. $I_1 = \displaystyle\int_{-1}^{0} e^x \sqrt{1 - e^x} dx$. We can set $t = e^x$.

2. $I_2 = \displaystyle\int_{-1}^{+1} \left(1 + x^2\right) \sqrt{1 - x^2} dx$. We can set $x = \sin t$.

3. $I_3 = \displaystyle\int_{0}^{\pi/4} \frac{\tan x}{1 + \cos x} dx.$

4. $I_4 = \int_0^1 \frac{x^3 + 2x^2 + 1}{x^2 + x + 1} dx.$

### 1.7.5 ODE

1. Solve $y'(x) - y(x) = x^2$. Give the solution for $y(0) = 0$.
2. Solve $xy'(x) - y(x) = \ln(x)$ $(x > 0)$.
3. Solve $y''(x) - \omega_0^2 y(x) = 0$ with $\omega_0$ real. Give the solution for $y(0) = a \in \mathbb{R}$ and $y'(0) = 0$.
4. Solve $y''(x) - 4y'(x) + 3y(x) = xe^x$.

# 2 Function of several variables and vectorial calculus

In this chapter, the concepts defined in chapter 1 are extended to several variables.

## 2.1 Partial derivative

In mathematics, a partial derivative of a function of several variables is its derivative with respect to one of those variables, with the others held **constant** (as opposed to the total derivative, in which all variables are allowed to vary).

The partial derivative of a function $f(x, y, \ldots)$ with respect to the variable $x$ is mainly denoted by

$$f'_x \quad \partial_x f \quad \frac{\partial f}{\partial x}.$$

The partial-derivative symbol is $\partial$.

Suppose that $f$ is a function of more than one variable. For instance,

$$z = f(x, y) = x^2 + xy + y^2.$$

The graph (see figure 2.1(a)) of this function defines a surface in Euclidean space. To every point on this surface, there is an infinite number of tangent lines. Partial differentiation is the act of choosing one of these lines and finding its slope. Usually, the lines of most interest are those that are parallel to the $xz$-plane, and those that are parallel to the $yz$-plane (which result from holding either $y$ or $x$ constant, respectively.)

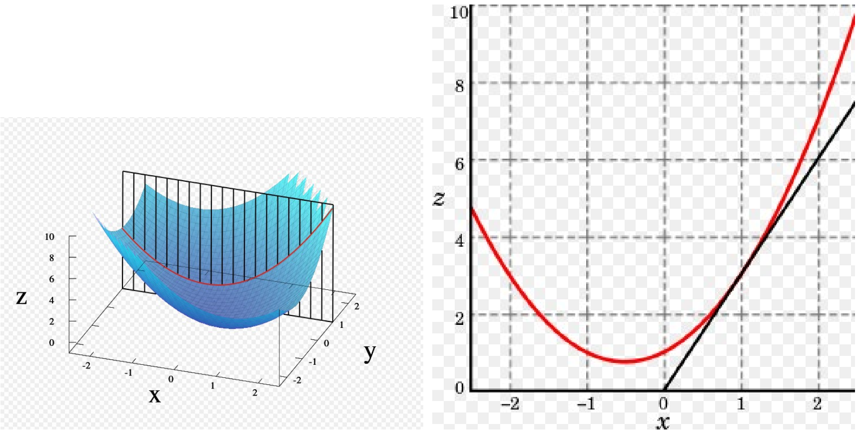
To find the slope of the line tangent to the function at $P(1, 1)$ that is parallel to the $xz$-plane, the $y$ variable is treated as **constant**. On the graph below it, we see the way the function looks on the plane $y = 1$. By finding the derivative of the equation while assuming that $y$ is a constant, the slope of $f$ at the point $(x, y)$ is found to be:

$$\frac{\partial f}{\partial x} = 2x + y.$$

So at $(1, 1)$, by substitution, the slope is 3. Therefore

$$\frac{\partial f}{\partial x} = 3,$$

at the point $(1, 1)$. That is, the partial derivative of $z$ with respect to $x$ at $(1, 1)$ is 3, as shown in the graph (see figure 2.1(b)).




(a) $z = f(x, y) = x^2 + xy + y^2$. (b) $z = f(x, y) = x^2 + xy + y^2$ for $y = 1$.

Figure 2.1: On the left, $z(x, y)$ versus $(x, y)$. On the left, $z(x, y)$ versus $x$ for $y = 1$.

**Definition 2.1** *Partial derivative. In general, the partial derivative of an n-variables function $f(x_1, ..., x_n)$ in the direction $x_i$ at the point $(a_1, ..., a_n)$ is defined to be:*

$$\begin{aligned} \left.\frac{\partial f}{\partial x_i}\right|_{x_i=a_i} &= \frac{\partial f}{\partial x_i}(a_i) = g(a_i) \\ &= \lim_{x_i \to a_i} \frac{f(x_i, a_1, \ldots, a_{n\neq i}) - f(a_i, a_1, \ldots, a_{n\neq i})}{x_i - a_i} \\ &= \lim_{h \to 0} \frac{f(a_i + h, a_1, \ldots, a_{n\neq i}) - f(a_1, a_2, \ldots, a_n)}{h}. \end{aligned} \tag{2.1}$$

*In the above difference quotient, all the variables except $x_i$ are held fixed. That choice of fixed values determines a function of one variable $g(x_i) = f(x_i, a_1, a_2, \ldots, a_{n\neq i})$, where the $n-1$ variables $\{a_{n\neq i}\}$ are **constants**.*

*Even if all partial derivatives $\frac{\partial f}{\partial x_i}(\overrightarrow{a})$ exist at a given point $\overrightarrow{a} = (a_1, a_2, \ldots, a_n)$, the function need not be **continuous** there. However, if all partial derivatives exist in a neighborhood of $\overrightarrow{a}$ and are continuous there, then $f$ is totally differentiable in that neighborhood and the total derivative is continuous. In this case, it is said that $f$ is a $\mathcal{C}^1$ function.*

The partial derivative $\frac{\partial f}{\partial x_i}$ can be seen as another function and can again be partially differentiated. If all mixed second order partial derivatives are continuous at a point (or on a set), $f$ is termed a $\mathcal{C}^2$ function at that point (or on that set); in this case, the partial derivatives can be exchanged

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}.$$

**Exercice 2.1** *Let $f$ be the function defined by $f(x, y) = x^3 - 3x^2y - 2y^3$. Determine de domain of existence of $f$. Calculate the partial derivatives of orders one and two.*

## 2.2 Primitive or antiderivative

There is a concept for partial derivatives that is analogous to primitives or antiderivatives for regular derivatives. Given a partial derivative, it allows for the partial recovery of the original function.

Consider the example of $\frac{\partial z}{\partial x} = 2x + y$. The "partial" integral can be taken with respect to $x$ (treating $y$ as constant, in a similar manner to partial differentiation):

$$z = \int \frac{\partial z}{\partial x} dx = x^2 + xy + g(y).$$

Here, the "constant" of integration is no longer a **constant**, but instead a **function** of all the variables of the original function except $x$. The reason for this is that all the other variables are treated as constant when taking the partial derivative, so any function which does not involve $x$ will disappear when taking the partial derivative, and we have to account for this when we take the antiderivative. The most general way to represent this is to have the "constant" represent an unknown function of all the other variables.

## 2.3 Taylor series expansion

Let $f$ be a function defined on $D_f \subset \mathbb{R}^n$ and of class $\mathcal{C}^p$, we note

$$\left[\sum_{i=1}^{i=n} h_i \frac{\partial f}{\partial x_i}(x_{10}, \ldots, x_{n0})\right]^{[p]} = \sum_{i_1+\ldots+i_n=p} \frac{p!}{i_1! \ldots i_n!} h_1^{i_1} \ldots h_n^{i_n} \frac{\partial^p f}{\partial x_1^{i_1} \ldots x_n^{i_n}}(x_{10}, \ldots, x_{n0}).$$

**Theorem 2.1** *Taylor-Lagrange series expansion. Let $f$ be a function defined on $D_f \subset \mathbb{R}^n$ (opened) of class $\mathcal{C}^p$. Then*

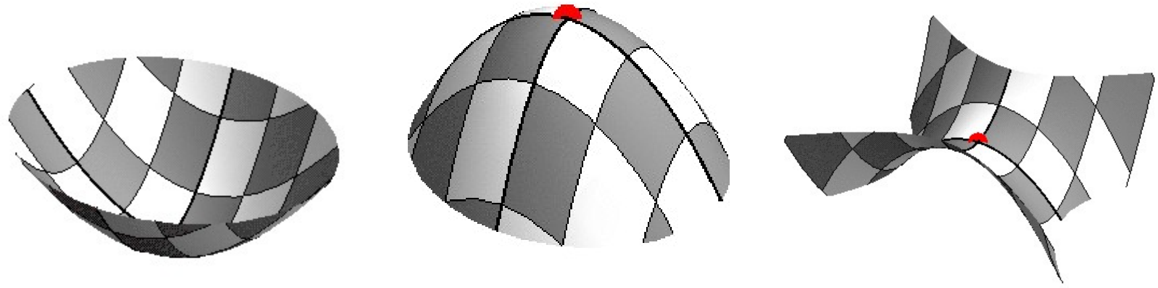
$$\begin{aligned} f(x_1, \ldots, x_n) &= f(a_1, \ldots, a_n) + \left[\sum_{i=1}^{i=n} h_i \frac{\partial f}{\partial x_i}(a_1, \ldots, a_n)\right] \\ &+ \frac{1}{2!}\left[\sum_{i=1}^{i=n} h_i \frac{\partial f}{\partial x_i}(a_1, \ldots, a_n)\right]^{[2]} + \ldots \\ &+ \frac{1}{k!}\left[\sum_{i=1}^{i=n} h_i \frac{\partial f}{\partial x_i}(a_1, \ldots, a_n)\right]^{[k]} \\ &+ \mathcal{O}\left(\left[h_1^2 + \ldots + h_n^2\right]^{k/2}\right), \end{aligned} \tag{2.2}$$

*where $h_i = x_i - a_i$ is close to zero.*

**Exercice 2.2** *Write a Taylor series expansion up the order two of a function $f(x, y)$ of two variables near the point $(0, 0)$.*

**Exercice 2.3** *Write a Taylor series expansion up to the order two of a function $f(x, y) = xe^{xy}$ near the point $(0, 0)$.*

## 2.4 Extremum



(a) Minimum. (b) Maximum. (c) Saddle point.

Figure 2.2: Three possible cases at a given critical point.

**Definition 2.2** *Relative extremum. If $f : D_f \subset \mathbb{R}^n \to \mathbb{R}$ has a **relative extremum** at the point $\overrightarrow{a}$ of $D_f$ and if $f$ is differentiable at $\overrightarrow{a}$, then*

$$\frac{\partial f}{\partial x_i}(\overrightarrow{a}) = 0 \text{ for all } i.$$

**Theorem 2.2** *Let $f(x, y) : D_f \subset \mathbb{R}^2 \to \mathbb{R}$ be the function of class $\mathcal{C}^2$, for which $\frac{\partial f}{\partial x}(\overrightarrow{a}) = \frac{\partial f}{\partial y}(\overrightarrow{a}) = 0$. Setting*

$$\boxed{r = \frac{\partial^2 f}{\partial x^2}(a) \quad s = \frac{\partial^2 f}{\partial x \partial y}(a) \quad t = \frac{\partial^2 f}{\partial y^2}(a)}. \tag{2.3}$$

*Then*

- *If $rt - s^2 > 0$ and $r > 0$, then $\overrightarrow{a} = (x_0, y_0)$ is a local minimum of $f$ (case (a) of figure 2.2).*
- *If $rt - s^2 > 0$ and $r < 0$, then $\overrightarrow{a} = (x_0, y_0)$ is a local maximum of $f$ (case (b) of figure 2.2).*
- *If $rt - s^2 < 0$, then $\overrightarrow{a} = (x_0, y_0)$ is a saddle point of $f$ (case (c) of figure 2.2).*
- *If $rt - s^2 = 0$, then the second derivative test is inconclusive, and the point $\overrightarrow{a} = (x_0, y_0)$ could be any of a minimum, maximum or saddle point.*

This theorem can be generalized by introducing the Hessian matrix, for which the elements $(i, j)$ (row, column) of the matrix is $\dfrac{\partial^2}{\partial x_i \partial x_j}$ and by calculating the eigenvalues of the Hessian matrix at the critical points.

**Exercice 2.4** *Study the relative extrema of the function $f$ defined by $f(x, y) = x^3 + y^3 - 3xy + 1$.*

## 2.5 Multiple integral

The multiple integral is a definite integral of a function of more than one real variable, for example, $f(x, y)$ or $f(x, y, z)$. Integrals of a function of two variables over a region in $\mathbb{R}^2$ are called double integrals, and integrals of a function of three variables over a region of $\mathbb{R}^3$ are called triple integrals.

We focus on the case of a function of two variables.

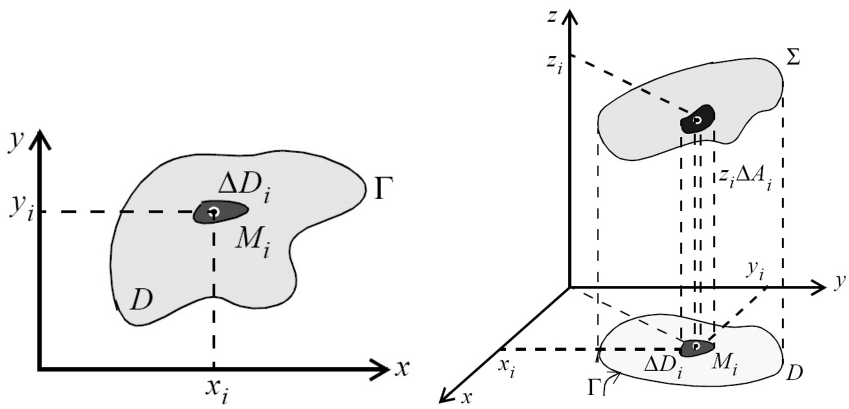
### 2.5.1 Definition

Let $(0x, Oy)$ be a Cartesian space, for which a domain $D$ limited by a closed line $\Gamma$ (figure 2.3(a)). $D$ is decomposed into $n$ non-overlapping elementary domains $\Delta D_i$ of area $\Delta A_i$ ($i = \{1, 2, \dots, n\}$) and let $M_i(x_i, y_i)$ be a point belonging $\Delta D_i$. Let $f(x, y)$ be a function **defined** and **continued** on $D$. At the point $M_i \in \Delta D_i$ it takes the value $f(x_i, y_i)$. Let be the sum

$$S_n = \sum_{i=1}^{i=n} f(x_i, y_i)\Delta A_i.$$

**Definition 2.3** *Double integral. If the limit of $S_n$ exists when $n \to \infty$ and all the $\Delta D_i$ tend toward zero, then the sum is called ordinary double integral of $f$ on the domain $D$; It writes*

$$\boxed{I = \iint_D f(x, y)dxdy = \lim_{n\to\infty} \sum_{i=1}^{i=n} f(x_i, y_i)\Delta A_i}. \tag{2.4}$$




(a) Definition. (b) geometrical interpretation.

Figure 2.3: On the left, definition of a double integral. On the right, geometrical interpretation of a double integral.

Let $\Sigma$ be the representative surface of the function $z = f(x, y)$ (see figure 2.3(b)). The sum

$$S_n = \sum_{i=1}^{i=n} z_i\Delta A_i,$$

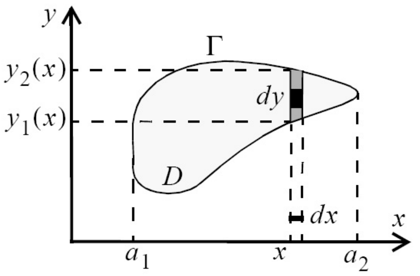
stands for a approximated value of the volume $V$ ranged from the plane $(Ox, Oy)$, the surface $\Sigma$ and the lines parallel to the axis $(Oz)$ lying on the contour $\Gamma$ of $D$. If $n \to \infty$, the limit **exactly** corresponds to the volume $V$; Thus

$$\boxed{I = \iint_D z(x,y)dA = \lim_{n\to\infty} \sum_{i=1}^{i=n} z(x_i, y_i)\Delta A_i = V}. \tag{2.5}$$

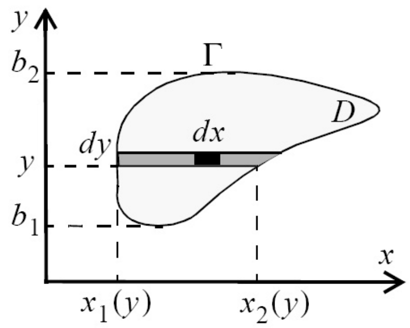
### 2.5.2 Calculation in Cartesian coordinates

This method is applicable to any domain $D$ for which (see figure 2.4):

- The projection of $D$ onto either the $x$-axis or the $y$-axis is bounded by the two values, $(a_1, a_2)$ and $(b_1, b_2)$, respectively.
- Any line perpendicular to this axis that passes between these two values intersects the domain in an interval whose endpoints are given by the graphs of two functions, $(y_1(x), y_2(x))$ and $(x_1(y), x_2(y))$, respectively.




(a) $x$ constant and $y$ moves.




(b) $y$ constant and $x$ moves.

Figure 2.4: Illustration of the Fubini theorem.

Such a domain will be here called a normal domain.

**$x$-axis** If the domain $D$ is normal with respect to the $x$-axis, and $f : D \to \mathbb{R}$ is a continuous function; then $y_1(x)$ and $y_2(x)$ (both of which are defined on the interval $[a_1; a_2]$) are the two functions that determine $D$. Then, by Fubini's theorem

$$\boxed{I = \iint_D f(x,y)dxdy = \int_{a_1}^{a_2} dx \left[\int_{y_1(x)}^{y_2(x)} f(x,y)dy\right]}. \tag{2.6}$$

**$y$-axis** If the domain $D$ is normal with respect to the $y$-axis, and $f : D \to \mathbb{R}$ is a continuous function; then $x_1(y)$ and $x_2(y)$ (both of which are defined on the interval $[b_1; b_2]$) are the two functions that determine $D$. Then, by Fubini's theorem

$$\boxed{I = \iint_D f(x,y)dxdy = \int_{b_1}^{b_2} dy \left[ \int_{x_1(y)}^{x_2(y)} f(x,y)dx \right]}. \tag{2.7}$$

This procedure can be easily generalize to a function of $n$ variables.

**Example 2.1** *Consider the domain $D$ defined by $D = \{(x,y) \in \mathbb{R}^2, x \geq 0, y \leq 1, y \geq x^2\}$ (see figure 2.5). Calculate the definite integral*

$$I = \iint_D (x+y)dxdy.$$

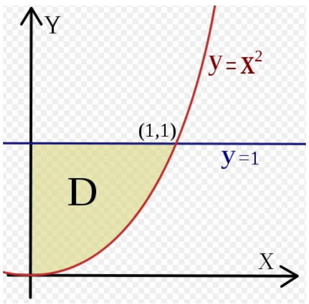



Figure 2.5: Domain $D = \{(x,y) \in \mathbb{R}^2, x \geq 0, y \leq 1, y \geq x^2\}$.

**$x$-axis** From figure 2.5, $(a_1, a_2) = (0,1)$ and $(y_1(x), y_2(x)) = (x^2, 1)$. Then

$$I = \int_0^1 \left[ \int_{x^2}^1 (x+y)dy \right] dx = \int_0^1 \left[ yx + \frac{y^2}{2} \right]_{x^2}^1 dx.$$

At first the second integral is calculated considering $x$ as a constant. The remaining operations consist of applying the basic techniques of integration:

$$I = \int_0^1 \left( x + \frac{1}{2} - x^3 - \frac{x^4}{2} \right) dx = \ldots = \frac{13}{20}.$$

**$y$-axis** From figure 2.5, if we choose normality with respect to the $y$-axis we could calculate

$$I = \int_0^1 \left[ \int_0^{\sqrt{y}} (x+y)dx \right] dy.$$

### 2.5.3 Theorem of the variable transformation

**Theorem 2.3** *Theorem of the variable transformation. From the variable transformation $x = x(u,v)$ et $y = y(u,v)$, the domain $D$ of the plane $(x,y)$ becomes the domain $D_{uv}$ of the plane*

*$(u, v)$; $J$ being the **Jacobian** of the transformation, we have*

$$\boxed{I = \iint_D f(x, y)dxdy = \iint_{D_{uv}} |J| \times f\left(x(u, v), y(u, v)\right) dudv}, \tag{2.8}$$

*where*

$$\boxed{J = \det\left(\underbrace{\begin{bmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{bmatrix}}_{\text{Jacobian matrix}}\right) = \frac{\partial x}{\partial u}\frac{\partial y}{\partial v} - \frac{\partial x}{\partial v}\frac{\partial y}{\partial u}}. \tag{2.9}$$

where the operator det stands for the determinant (il will be defined in the next chapter).

**Example 2.2** *Consider the domain $D$ defined by $D = \{(x, y) \in \mathbb{R}^2, (a \geq 0, b > a), a^2 \leq x^2 + y^2 \leq b^2, x \geq 0\}$ (see figure 2.6). Calculate the definite integral*

$$I = \iint_D e^{-x^2-y^2} dxdy.$$

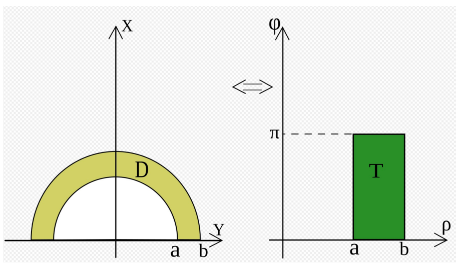



Figure 2.6: Domain $D = \{(x, y) \in \mathbb{R}^2, (a \geq 0, b > a), a^2 \leq x^2 + y^2 \leq b^2, x \geq 0\}$.

The function $x^2 + y^2 = R^2$ describes a circle of radius $R$ and of center $(0, 0)$. As shown in figure 2.6, the domain $D$ describes then a half (because $x \geq 0$) circular crown of radius $a$ and $b > a$. Using the polar coordinates system, $\{x = \rho\cos\phi, y = \rho\cos\phi\}$, the new domain $T$ depicted on the right of figure 2.6 is a rectangle of sides $(b - a)$ and $\pi$ (because $x \geq 0$). In addition, the Jacobian is

$$J = \det\begin{bmatrix} \frac{\partial x}{\partial \rho} & \frac{\partial x}{\partial \phi} \\ \frac{\partial y}{\partial \rho} & \frac{\partial y}{\partial \phi} \end{bmatrix} = \begin{bmatrix} \cos\phi & -\rho\sin\phi \\ \sin\phi & \rho\cos\phi \end{bmatrix} = \rho\left(\cos^2\phi + \sin^2\phi\right) = \rho. \tag{2.10}$$

Then, since $x^2 + y^2 = \rho^2$, the integral $I$ is

$$I = \int_a^b \int_0^\pi |r|e^{-r^2} d\phi dr = \pi \int_a^b re^{-r^2} dr = -\frac{1}{2}\left[e^{-r^2}\right]_a^b = \frac{e^{-a^2} - e^{-b^2}}{2}.$$

## 2.6 Vectorial calculus

### 2.6.1 Dot and cross products

A vector is a geometrical object that possesses both a magnitude and a direction. A vector can be pictured as an arrow. Its magnitude is its length, and its direction is the direction that the arrow points. The magnitude of a vector $\overrightarrow{u}$ is denoted by $\|\overrightarrow{u}\|$

Let $\overrightarrow{u}$ be a vector of $n$ components $(u_1, u_2, \ldots, u_n)$ and let $\overrightarrow{v}$ be a vector of $n$ components $(v_1, v_2, \ldots, v_n)$.

**Definition 2.4** *Dot product. The* ***dot*** *product of the two vectors* $\overrightarrow{u}$ *and* $\overrightarrow{v}$ *is a* ***scalar*** *defined as*

$$s = \overrightarrow{u} \cdot \overrightarrow{v} = \sum_{i=1}^{i=n} u_i v_i.$$

**Definition 2.5** *For example, in an unitary orthogonal Cartesian coordinates system* $(\overrightarrow{i}, \overrightarrow{j}, \overrightarrow{k})$*, if* $\overrightarrow{u} = (x_1, y_1, z_1)$ *and* $\overrightarrow{v} = (x_2, y_2, z_2)$*, then*

$$s = x_1 y_1 + x_2 y_2 + x_3 y_3,$$

*where* $\overrightarrow{u} \cdot \overrightarrow{u} = 1$ *for* $\overrightarrow{u} = \{\overrightarrow{i}, \overrightarrow{j}, \overrightarrow{k}\}$ *and* $\overrightarrow{i} \cdot \overrightarrow{j} = \overrightarrow{j} \cdot \overrightarrow{k} = \overrightarrow{k} \cdot \overrightarrow{i} = 0$.

*The dot scalar product can also be defined as*

$$s = \overrightarrow{u} \cdot \overrightarrow{v} = \|\overrightarrow{u}\| \, \|\overrightarrow{v}\| \cos(\theta),$$

*where* $\theta$ *is the angle between* $\overrightarrow{u}$ *and* $\overrightarrow{v}$*.*

*In particular, if* $\overrightarrow{u}$ *and* $\overrightarrow{v}$ *are orthogonal, then the angle between them is 90° and* $\overrightarrow{u} \cdot \overrightarrow{v} = 0$. *At the other extreme, if they are codirectional (or collinear), then the angle between them is 0° and* $\overrightarrow{u} \cdot \overrightarrow{v} = \|\overrightarrow{u}\| \, \|\overrightarrow{v}\|$*. In addition,* $\overrightarrow{u} \cdot \overrightarrow{u} = \|\overrightarrow{u}\|^2 = u^2$ *and* $u = \sqrt{\overrightarrow{u} \cdot \overrightarrow{u}}$*.*

***Exercice 2.5*** *Generalization of the Pythagore theorem. From figure 2.7, show that* $c^2 = a^2 + b^2 - 2ab\cos(\theta)$*.*

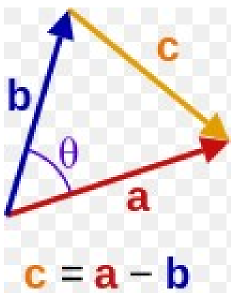



Figure 2.7: Generalization of the Pythagore theorem.

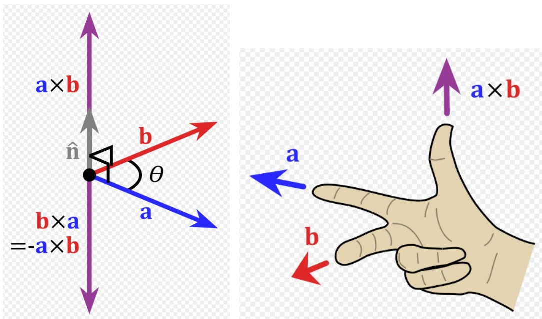
**Definition 2.6** *Cross product. In Cartesian coordinates, the* ***cross*** *product (it is defined only in three-dimensional space) is a* ***vector*** $\overrightarrow{w}$ *defined as*

$$\begin{aligned}\overrightarrow{w} &= \overrightarrow{u} \wedge \overrightarrow{v} = \left(x_1 \overrightarrow{i} + y_1 \overrightarrow{j} + z_1 \overrightarrow{k}\right) \wedge \left(x_2 \overrightarrow{i} + y_2 \overrightarrow{j} + z_2 \overrightarrow{k}\right) \\ &= (y_1 z_2 - y_2 z_1)\, \overrightarrow{i} + (z_1 x_2 - x_1 z_2)\, \overrightarrow{j} + (x_1 y_2 - x_2 y_1)\, \overrightarrow{k}.\end{aligned}$$

*The cross product* $\overrightarrow{u} \wedge \overrightarrow{v}$ *is defined as a vector* $\overrightarrow{w}$ *that is perpendicular to both* $\overrightarrow{u}$ *and* $\overrightarrow{v}$ *(see figure 2.8), with a direction given by the right-hand rule and a magnitude equal to the area of the parallelogram that the vectors span :*

$$\overrightarrow{u} \wedge \overrightarrow{v} = \|\overrightarrow{u}\|\, \|\overrightarrow{v}\| \sin(\theta)\, \overrightarrow{n},$$

*where* $\theta$ *is the angle between* $\overrightarrow{u}$ *and* $\overrightarrow{v}$ *in the plane containing them (hence, it is between 0° and 180°),* $\|\overrightarrow{u}\|$ *and* $\|\overrightarrow{v}\|$ *are the magnitudes of vectors* $\overrightarrow{u}$ *and* $\overrightarrow{v}$*, and* $\overrightarrow{n}$ *is an unit vector perpendicular to the plane containing* $\overrightarrow{u}$ *and* $\overrightarrow{v}$ *in the direction given by the right-hand rule (see figure 2.8).*




(a) Definition. (b) Right-hand rule.

Figure 2.8: On the left, definition of the cross product. On the right, illustration of the right-hand rule.

In an unitary orthogonal Cartesian coordinates system, $\overrightarrow{u} \wedge \overrightarrow{u} = \overrightarrow{0}$ for $\overrightarrow{u} = \{\overrightarrow{i}, \overrightarrow{j}, \overrightarrow{k}\}$ and $\overrightarrow{i} \wedge \overrightarrow{j} = \overrightarrow{k}$, $\overrightarrow{k} \wedge \overrightarrow{i} = \overrightarrow{j}$ and $\overrightarrow{j} \wedge \overrightarrow{k} = \overrightarrow{i}$. The cross product of two collinear vectors vanishes.

**Property 2.1**

- *The sum the two vectors* $\overrightarrow{u}$ *and* $\overrightarrow{v}$, $\overrightarrow{u} + \overrightarrow{v}$ *is also a vector of* $n$ *components equal* $(u_1 + v_1, u_2 + v_2, \ldots, u_n + v_n)$.
- *The dot and cross products are distributive.*
- $\overrightarrow{u} \cdot \overrightarrow{u} = u^2$ *and* $\overrightarrow{u} \wedge \overrightarrow{u} = \overrightarrow{0}$.
- $\overrightarrow{u} \cdot \overrightarrow{v} = \overrightarrow{v} \cdot \overrightarrow{u}$ *and* $\overrightarrow{u} \wedge \overrightarrow{v} = -\overrightarrow{v} \wedge \overrightarrow{u}$.
- $\overrightarrow{u} \wedge (\overrightarrow{v} \wedge \overrightarrow{w}) = \overrightarrow{v}(\overrightarrow{u} \cdot \overrightarrow{w}) - \overrightarrow{w}(\overrightarrow{u} \cdot \overrightarrow{v})$.

### 2.6.2 Scalar and vector fields

**Definition 2.7** ***Scalar*** *field* $f = f(x, y, z)$. *A scalar field associates a scalar value to every point in a space. The scalar may either be a mathematical number or a physical quantity. Examples of scalar fields in applications include the temperature distribution throughout space or the pressure distribution in a fluid. These fields are the subject of scalar field theory.*

**Definition 2.8** ***Vectorial*** *field* $\overrightarrow{F}(x, y, z)$. *A vector field is an assignment of a vector to each point in a subset of space. A vector field in the plane, for instance, can be visualized as a collection of arrows with a given magnitude and direction each attached to a point in the plane. Vector fields are often used to model, for example, the speed and direction of a moving fluid throughout space, or the strength and direction of some force, such as the magnetic and electric fields, as it changes from point to point.*

### 2.6.3 Vectorial operators

To express the vectorial operator, it is very convenient to introduce the nabla operator defined in Cartesian coordinates as

$$\overrightarrow{\nabla} = \frac{\partial f}{\partial x}\overrightarrow{i} + \frac{\partial f}{\partial y}\overrightarrow{j} + \frac{\partial f}{\partial z}\overrightarrow{k}.$$

#### 2.6.3.1 Gradient

**Definition 2.9** ***Gradient vectorial*** *operator. In mathematics, the gradient is a multi-variable generalization of the derivative. While a derivative can be defined on functions of a single variable, for functions of several variables, the gradient takes its place. The gradient is a vector-valued function, as opposed to a derivative, which is a scalar-valued. If* $f(x_1, \ldots, x_n)$ *is a differentiable, real-valued function of several variables, its gradient is the vector whose components are the* $n$ *partial derivatives of* $f$. *Then*

$$\overrightarrow{F} = \overrightarrow{grad}(f) = \sum_{i=1}^{i=n} \frac{\partial f}{\partial x_i}\overrightarrow{e}_i.$$

*The vector* $\{\overrightarrow{e}_i\}$ *are the unit vectors which define the Euclidian space.*

In physics, the scalar and vectorial fields depend in general of the spatial coordinates $(x, y, z)$. Then

$$\boxed{\overrightarrow{\mathrm{grad}}(f) = \frac{\partial f}{\partial x}\overrightarrow{i} + \frac{\partial f}{\partial y}\overrightarrow{j} + \frac{\partial f}{\partial z}\overrightarrow{k} = \overrightarrow{\nabla} f}. \tag{2.11}$$

The gradient operator convert a **scalar** field $f$ into a **vectorial** field $\overrightarrow{F} = \overrightarrow{\mathrm{grad}} f$.

**Example 2.3** *if* $f(x, y, z) = 2x + e^{ay} + \sin(z)$, *then* $\overrightarrow{grad} f = 2\overrightarrow{i} + ae^{ay}\overrightarrow{j} + \cos(z)\overrightarrow{k}$.

#### 2.6.3.2 Divergence

**Definition 2.10** ***Divergence scalar*** *operator. In vector calculus, divergence is a vector operator that produces a signed scalar field giving the quantity of a vector field's source at each point. In mathematics, if* $\overrightarrow{F} = \sum F_i \overrightarrow{e}_i$ *is a continuously differentiable vector field of several variables, its divergence is the scalar defined as*

$$\boxed{div\overrightarrow{F} = \sum_{i=1}^{i=n} \frac{\partial F_i}{\partial x_i}}. \tag{2.12}$$

In physics, the scalar and vectorial fields depend in general of the spatial coordinates $(x, y, z)$. If $\overrightarrow{F} = F_x \overrightarrow{i} + F_y \overrightarrow{j} + F_z \overrightarrow{k}$, then

$$\boxed{\text{div}\overrightarrow{F} = \frac{\partial F_x}{\partial x} + \frac{\partial F_y}{\partial y} + \frac{\partial F_z}{\partial z} = \overrightarrow{\nabla} \cdot \overrightarrow{F}}. \tag{2.13}$$

The divergence operator convert a **vectorial** field $\overrightarrow{F}$ into a **scalar** field $f = \text{div}\overrightarrow{F}$.

**Example 2.4** *if* $\overrightarrow{F}(x, y, z) = 2x\overrightarrow{i} + e^{ay}\overrightarrow{j} + \sin(z)\overrightarrow{k}$, *then* $div\overrightarrow{F} = 2 + ae^{ay} + \cos(z)$.

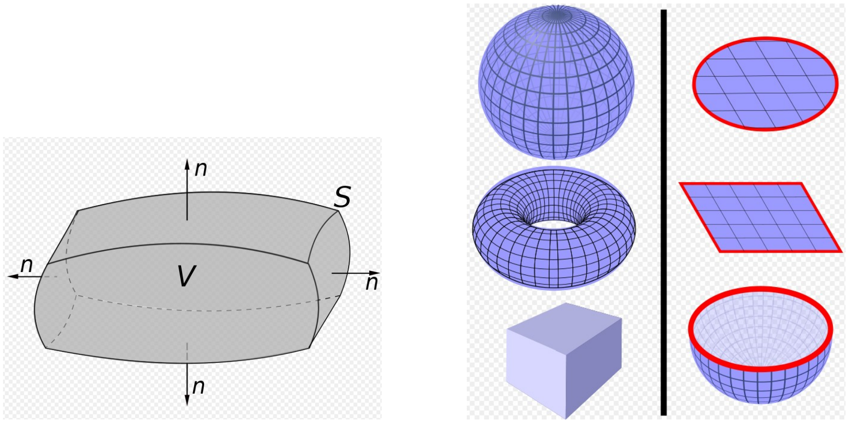
**Theorem 2.4** *Divergence theorem (see figure 2.9(a)). The divergence theorem, also known as Gauss's theorem or Ostrogradsky's theorem, is a result that relates the flow (that is, flux) of a vector field through a surface to the behavior of the vector field inside the surface. More precisely, the divergence theorem states that the outward flux of a vector field through a closed surface is equal to the volume integral of the divergence over the region inside the surface. Intuitively, it states that the sum of all sources (with sinks regarded as negative sources) gives the net flux out of a region. If* $\overrightarrow{F}$ *is a continuously differentiable vector field defined on a neighborhood of* $V$*, then*

$$\boxed{\iiint div\overrightarrow{F}\, dV = \oint\!\!\!\oint_S \overrightarrow{F} \cdot \overrightarrow{dS}}.$$

The left side is a volume integral over the volume $V$, the right side is the surface integral over the boundary of the volume $V$. The closed surface $S$ is the boundary of $V$ oriented by outward-pointing normals, and $\overrightarrow{n}$ is the outward pointing unit normal field of the boundary $S$ ($\overrightarrow{dS} = dS\overrightarrow{n}$). The symbol within the two integrals stresses once more that $S$ is a **closed** surface. In terms of the intuitive description above, the left-hand side of the equation represents the total of the sources in the volume $V$, and the right-hand side represents the total flow across the boundary $S$.

#### 2.6.3.3 Curl

**Definition 2.11** ***Curl vectorial*** *operator. In vector calculus, the curl is a vector operator that describes the infinitesimal rotation of a 3-dimensional vector field. At every point in the field, the curl of that point is represented by a vector. The attributes of this vector (length and direction)*




(a) Divergence theorem. (b) Condition of application.

Figure 2.9: Illustration of the divergence theorem. The divergence theorem can be used to calculate a flux through a closed surface that fully encloses a volume, like any of the surfaces on the left of (b). It can not directly be used to calculate the flux through surfaces with boundaries, like those on the right of (b) (Surfaces are blue, boundaries are red).

*characterize the rotation at that point. In mathematics, if $\overrightarrow{F}(x,y,x) = F_x\overrightarrow{i} + F_y\overrightarrow{j} + F_z\overrightarrow{k}$ is a differentiable, real-valued vectorial function, its curl is defined as*

$$\boxed{\overrightarrow{curl}\overrightarrow{F} = \left(\frac{\partial F_z}{\partial y} - \frac{\partial F_y}{\partial z}\right)\overrightarrow{i} + \left(\frac{\partial F_x}{\partial z} - \frac{\partial F_z}{\partial x}\right)\overrightarrow{j} + \left(\frac{\partial F_y}{\partial x} - \frac{\partial F_x}{\partial y}\right)\overrightarrow{k} = \overrightarrow{\nabla} \wedge \overrightarrow{F}}. \tag{2.14}$$

The curl operator convert a **vectorial** field $\overrightarrow{F}$ into a **vectorial** field $\overrightarrow{F}_1 = \overrightarrow{\text{curl}}\overrightarrow{F}$.

**Exercice 2.6** *If $\overrightarrow{F}(x,y,z) = 2x\overrightarrow{i} + e^{ay}\overrightarrow{j} + \sin(z)\overrightarrow{k}$, then calculate $\overrightarrow{curl}\overrightarrow{F}$.*

**Exercice 2.7** *Show from two methods that $\overrightarrow{curl}[\overrightarrow{grad}f] = \overrightarrow{0}$, where $f$ is a scalar function of class $\mathcal{C}^2$.*

**Exercice 2.8** *Show from two methods that $div(\overrightarrow{curl}\overrightarrow{F}) = 0$, where $\overrightarrow{F}$ is a vectorial function of class $\mathcal{C}^2$.*

If $\text{div}\overrightarrow{F} = 0$, then $\overrightarrow{F} = \overrightarrow{\text{curl}}\overrightarrow{A}$; We said that $\overrightarrow{F}$ derive from a potential vector $\overrightarrow{A}$. In fact, the vector $\overrightarrow{A}$ is defined to a gradient since if $\overrightarrow{A}' = \overrightarrow{A} + \overrightarrow{\text{grad}}f$, then

$$\overrightarrow{\text{curl}}\overrightarrow{A}' = \overrightarrow{\text{curl}}\left(\overrightarrow{A} + \overrightarrow{\text{grad}}f\right) = \overrightarrow{\text{curl}}\overrightarrow{A} + \underbrace{\overrightarrow{\text{curl}}\left(\overrightarrow{\text{grad}}f\right)}_{\overrightarrow{0}\ \forall f} = \overrightarrow{\text{curl}}\overrightarrow{A}.$$

If $\overrightarrow{\text{curl}}\overrightarrow{F} = \overrightarrow{0}$, then $\overrightarrow{F} = -\overrightarrow{\text{grad}}f$; We said that $\overrightarrow{F}$ derive from a potential scalar $f$.

**Exercice 2.9** *Let be the vectorial field $\overrightarrow{F}$ defined on $\mathbb{R}^3$ by $\overrightarrow{F}(x,y,z) = 2xz\overrightarrow{e}_x + yz\overrightarrow{e}_y + (x^2 + y^2/2)\overrightarrow{e}_z$. Show that $\overrightarrow{curl}\overrightarrow{F} = \overrightarrow{0}$ and deduce the associated scalar function $f$.*

**Theorem 2.5** *Stokes's theorem (see figure 2.10). The Stokes theorem relates the surface integral of the curl of a vector field $\overrightarrow{F}$ over a surface $S$ in Euclidean three-space to the line integral of the vector field over its boundary $S$:*

$$\boxed{\oint_{\Gamma} \overrightarrow{F} \cdot \overrightarrow{dl} = \iint_{S} \overrightarrow{curl}\overrightarrow{F} \cdot \overrightarrow{dS}.}. \tag{2.15}$$

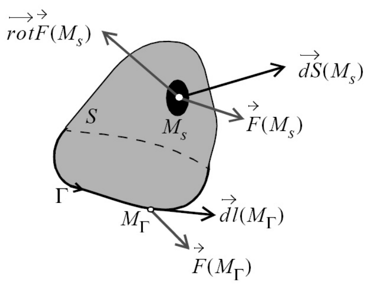



Figure 2.10: Illustration of the Stokes theorem. $M_S$ is a point on the opened surface and $M_\Gamma$ is a point on the closed contour.

In other words: The circulation of the vectorial field $\overrightarrow{F}$ along of any **closed** contour $\Gamma$ equals the flux of the vectorial field $\overrightarrow{\mathrm{curl}}\overrightarrow{F}$ within the **opened** surface lying on $\Gamma$.

#### 2.6.3.4 Scalar Laplacian

**Definition 2.12** ***Laplacian scalar*** *operator. The Laplace operator is a second order differential operator in the $n$-dimensional Euclidean space. Thus if $f$ is a twice-differentiable real-valued function, then the scalar Laplacian of $f$ is defined by*

$$\boxed{\nabla^2 f = div\left(\overrightarrow{grad}f\right) = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2}}. \tag{2.16}$$

The notation $\nabla^2$ stands for the scalar Laplacian and is obtained from $\overrightarrow{\nabla}$ as

$$\mathrm{div}\left(\overrightarrow{\mathrm{grad}}f\right) = \overrightarrow{\nabla} \cdot \left(\overrightarrow{\nabla} f\right) = \overrightarrow{\nabla} \cdot \overrightarrow{\nabla} f = \nabla^2 f.$$

#### 2.6.3.5 Vectorial Laplacian

**Definition 2.13** ***Laplacian vectorial*** *operator. The Laplace operator is a second order differential operator in the $n$-dimensional Euclidean space. Thus if $\overrightarrow{F}$ is a twice-differentiable real-valued vectorial function, then the vectorial Laplacian of $\overrightarrow{F} = F_x\overrightarrow{i} + F_y\overrightarrow{j} + F_z\overrightarrow{k}$ is defined by*

$$\boxed{\nabla^2\overrightarrow{F} = \overrightarrow{grad}\left(div\overrightarrow{F}\right) - \overrightarrow{curl}\left(\overrightarrow{curl}\overrightarrow{F}\right) = \nabla^2 F_x\overrightarrow{i} + \nabla^2 F_y\overrightarrow{j} + \nabla^2 F_z\overrightarrow{k}}. \tag{2.17}$$

**Exercice 2.10** *From two methods, show that*

$$\begin{aligned}\vec{\nabla}^2\vec{F} &= \left(\frac{\partial^2 F_x}{\partial x^2}+\frac{\partial^2 F_x}{\partial y^2}+\frac{\partial^2 F_x}{\partial z^2}\right)\vec{i}\\ &+ \left(\frac{\partial^2 F_y}{\partial x^2}+\frac{\partial^2 F_y}{\partial y^2}+\frac{\partial^2 F_y}{\partial z^2}\right)\vec{j}\\ &+ \left(\frac{\partial^2 F_z}{\partial x^2}+\frac{\partial^2 F_z}{\partial y^2}+\frac{\partial^2 F_z}{\partial z^2}\right)\vec{k}\,.\end{aligned}$$

### 2.6.4 Operators in cylindrical and spherical coordinates

When we solve a problem in physics, the choice of the coordinates system is very important. It is therefore necessary to know the expression of the vector operators for such systems.

Let $(x, y, z)$ be the Cartesian (rectangular) coordinates of a point, which depend on $(q_1, q_2, q_3)$ and defined by

$$\begin{cases} x = x(q_1, q_2, q_3) \\ y = y(q_1, q_2, q_3) \\ z = z(q_1, q_2, q_3) \end{cases}.$$

The functions $x$, $y$, $z$, $q_1$, $q_2$ and $q_3$ are assumed to be bijective and have continuous partial derivatives. Then, the correspondance between $(x, y, z)$ and $(q_1, q_2, q_3)$ is unique. The functions $(q_1, q_2, q_3)$ are named the **curviligne coordinates** of the point $M(x, y, z)$.

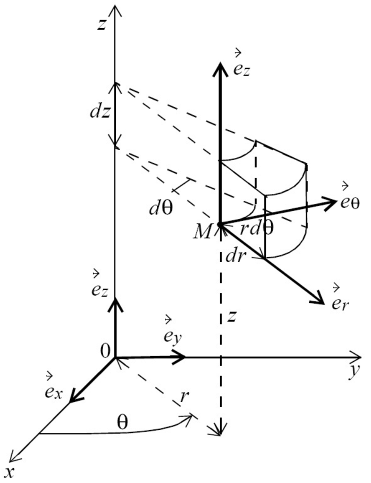
The metric coefficients $\{h_i\}$ $(i = \{1, 2, 3\})$, which depend on the coordinates system, are defined by

$$\boxed{h_i = \sqrt{\left(\frac{\partial x}{\partial q_i}\right)^2 + \left(\frac{\partial y}{\partial q_i}\right)^2 + \left(\frac{\partial z}{\partial q_i}\right)^2}}. \tag{2.18}$$

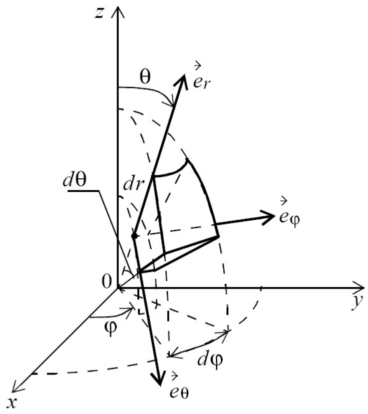
Then, for the cylindrical and spherical coordinates, table 2.1 is obtained.

| | Cylindrical Coordinates | Spherical Coordinates |
|---|---|---|
| Definition | $\begin{cases} x = r\cos\theta \\ y = r\sin\theta \\ z = z \end{cases}$ et $\begin{cases} q_1 = r \\ q_2 = \theta \\ q_3 = z \end{cases}$ | $\begin{cases} x = r\sin\theta\cos\phi \\ y = r\sin\theta\sin\phi \\ z = r\cos\theta \end{cases}$ et $\begin{cases} q_1 = r \\ q_2 = \theta \\ q_3 = \phi \end{cases}$ |
| Local basis | $(\vec{e}_r, \vec{e}_\theta, \vec{e}_z)$ | $(\vec{e}_r, \vec{e}_\theta, \vec{e}_\phi)$ |
| Metric Coefficients | $\begin{cases} h_1 = h_r = 1 \\ h_2 = h_\theta = r \\ h_3 = h_z = 1 \end{cases}$ | $\begin{cases} h_1 = h_r = 1 \\ h_2 = h_\theta = r \\ h_3 = h_\phi = r\sin\theta \end{cases}$ |
| Elementary displacement | $\begin{cases} dx = dr \\ dy = rd\theta \\ dz = dz \end{cases}$ | $\begin{cases} dx = dr \\ dy = rd\theta \\ dz = r\sin\theta d\phi \end{cases}$ |
| Geometrical representation | Figure 2.11(a) | Figure 2.11(b) |

Table 2.1: Metric coefficients in cylindrical and spherical coordinates.




(a) Cylindrical.




(b) Spherical.

Figure 2.11: Cylindrical and spherical coordinates.

**Exercice 2.11** *Calculate the metric coefficients* $\{h_i\}$ *in cylindrical and spherical coordinates.*

In a curvilinear orthogonal coordinates system ($\vec{e}_1 \cdot \vec{e}_2 = \vec{e}_1 \cdot \vec{e}_3 = \vec{e}_2 \cdot \vec{e}_3 = 0$ and $\vec{e}_1 \wedge \vec{e}_2 = \vec{e}_3$), the gradient operator is expressed as

$$\boxed{\overrightarrow{\text{grad}} f = \sum_{i=1}^{i=3} \frac{1}{h_i} \frac{\partial f}{\partial q_i} \vec{e}_i}. \tag{2.19}$$

The divergence operator is expressed as

$$\boxed{\text{div} \vec{F} = \frac{1}{h_1 h_2 h_3} \sum_{i=1}^{i=3} \frac{\partial}{\partial q_i} \left( h_1 h_2 h_3 \frac{F_i}{h_i} \right)}. \tag{2.20}$$

The curl operator is expressed as

$$\boxed{\overrightarrow{\text{curl}} \vec{F} = \frac{1}{h_1 h_2 h_3} \left\| \begin{bmatrix} h_1 \vec{e}_1 & h_2 \vec{e}_2 & h_3 \vec{e}_3 \\ \frac{\partial}{\partial q_1} & \frac{\partial}{\partial q_2} & \frac{\partial}{\partial q_3} \\ h_1 F_1 & h_2 F_2 & h_3 F_3 \end{bmatrix} \right\| = \left| \begin{array}{l} \frac{1}{h_2 h_3} \left[ \frac{\partial}{\partial q_2}(h_3 F_3) - \frac{\partial}{\partial q_3}(h_2 F_2) \right] \\ \\ \frac{1}{h_3 h_1} \left[ \frac{\partial}{\partial q_3}(h_1 F_1) - \frac{\partial}{\partial q_1}(h_3 F_3) \right] \\ \\ \frac{1}{h_1 h_2} \left[ \frac{\partial}{\partial q_1}(h_2 F_2) - \frac{\partial}{\partial q_2}(h_1 F_1) \right] \end{array} \right.}. \tag{2.21}$$

Note that the equation for each component, $\overrightarrow{\text{curl}} \vec{F} \cdot \vec{e}_i$ can be obtained by exchanging each occurrence of a subscript 1, 2, 3 in cyclic permutation: $1 \to 2$, $2 \to 3$, and $3 \to 1$.

**Exercice 2.12** *In a curvilinear orthogonal coordinates system, show that the scalar Laplacian is expressed as*

$$\nabla^2 = \frac{1}{h_1 h_2 h_3}\left[\frac{\partial}{\partial q_1}\left(\frac{h_2 h_3}{h_1}\frac{\partial f}{\partial q_1}\right) + \frac{\partial}{\partial q_2}\left(\frac{h_3 h_1}{h_2}\frac{\partial f}{\partial q_2}\right) + \frac{\partial}{\partial q_3}\left(\frac{h_1 h_2}{h_3}\frac{\partial f}{\partial q_3}\right)\right].$$

**Exercice 2.13** *From table 2.1, show that the gradient operator is expressed in cylindrical coordinates as*

$$\overrightarrow{grad} f = \frac{\partial f}{\partial r}\overrightarrow{e}_r + \frac{1}{r}\frac{\partial f}{\partial \theta}\overrightarrow{e}_\theta + \frac{\partial f}{\partial z}\overrightarrow{e}_z.$$

**Exercice 2.14** *From table 2.1, show that the divergence operator is expressed in spherical coordinates as*

$$div\overrightarrow{F} = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2 F_r\right) + \frac{1}{r\sin\theta}\left[\frac{\partial}{\partial \theta}\left(F_\theta \sin\theta\right) + \frac{\partial F_\phi}{\partial \phi}\right].$$

# 2.7 Homework

Make the exercises on a copy with a **clean** presentation and **underline** the final results. Do not forgot to write your name and surname on all sheets. The copy will be read by the Professor in the next session (course).

## 2.7.1 Partial derivative

1. Let $f$ be the function defined by $f(x,y) = x^3 y + e^{xy^2}$. Calculate $\frac{\partial f}{\partial x} = f'_x$, $\frac{\partial f}{\partial y} = f'_y$, $\frac{\partial^2 f}{\partial x^2} = f''_x$, $\frac{\partial^2 f}{\partial y^2} = f''_y$, $\frac{\partial^2 f}{\partial x \partial y} = f''_{xy}$ and $\frac{\partial^2 f}{\partial y \partial x} = f''_{yx}$.

2. Show that $U(x,y,x) = (x^2 + y^2 + z^2)^{-\frac{1}{2}}$, with $x$, $y$ and $z$ different of zero, satisfies the Laplace equation $\frac{\partial^2 U}{\partial x^2} + \frac{\partial^2 U}{\partial y^2} + \frac{\partial^2 U}{\partial z^2} = 0$.

3. If $f(x,y) = x^2 \arctan\left(\frac{x}{y}\right)$, calculate then $\frac{\partial^2 f}{\partial x \partial y}$ at the point $(1,1)$.

## 2.7.2 Extremum

1. Write the Taylor series expansion up to the order 2 at the point $(0, \pi/2)$ of the function $f$ defined by $f(x,y) = e^x \sin(x+y)$.

2. Find the maxima and the minima of the function $f$ defined by $f(x,y) = x^3 + y^3 - 3x - 12y + 20$.

3. Consider a rectangular box of volume $a^3$ ($a > 0$). What are the dimensions of the sides to have a minimum total surface?

### 2.7.3 Double integral

1. Consider the domain $D = \{0 \leq y \leq x^2, 0 \leq x \leq 1\}$. Plot the domain and calculate the following integral
$$I = \iint_D (x^2 + y^2) dxdy.$$

2. Consider the domain $D = \{0 \leq y \leq -2x + 2, 0 \leq x \leq 1\}$. Plot the domain and calculate the following integral
$$I = \iint_D (2x + y^2) dxdy,$$

3. Consider the domain $D = \{|x| + |y| \leq 1\}$. Plot the domain and calculate the following integral
$$I = \iint_D e^{x+y} dxdy,$$

4. Consider the domain $D = \{x^2 + y^2 \geq 4, x^2 + y^2 \leq 9\}$. Plot the domain and calculate the following integral
$$I = \iint_D \sqrt{x^2 + y^2} dxdy,$$

5. Consider the domain $D = \mathbb{R}^2$. Calculate the following integral
$$I = \iint_D e^{-(\alpha x + a)^2 - (\beta y + b)^2} dxdy,$$
where $\alpha \neq 0$, $\beta \neq 0$ and $(a, b) \in \mathbb{R}^2$. The variable transformation $\alpha x + a = r \cos \phi$ and $\beta y + b = r \sin \phi$ can be used.

### 2.7.4 Vectorial calculus

1. Show that $\overrightarrow{u} \wedge (\overrightarrow{v} \wedge \overrightarrow{w}) = \overrightarrow{v}(\overrightarrow{u} \cdot \overrightarrow{w}) - \overrightarrow{w}(\overrightarrow{u} \cdot \overrightarrow{v})$.

2. From the Lentz law
$$e = -\frac{\partial \Phi}{\partial t} = \oint_\Gamma \overrightarrow{E} \cdot \overrightarrow{dl},$$
where $\Phi$ is the flux of the vector magnetic field $\overrightarrow{B}$, show that $\overrightarrow{\text{curl}} \overrightarrow{E} = -\dfrac{\partial \overrightarrow{B}}{\partial t}$.

3. From the Gauss theorem
$$\oiint_S \overrightarrow{E} \cdot \overrightarrow{dS} = \frac{Q}{\epsilon_0},$$

where $\overrightarrow{E}$ is the vector electric field and $Q$ the total charge inside the volume and the divergence theorem, show that $\text{div}\overrightarrow{E} = \rho/\epsilon_0$ where $\rho$ is the charge volume density and $\epsilon_0$ is the permittivity of the vacuum.

4. From table 2.1, show that the curl operator in spherical coordinates is expressed as

$$\begin{aligned}\overrightarrow{\text{curl}}\overrightarrow{F} &= \frac{1}{r\sin\theta}\left[\frac{\partial}{\partial\theta}(\sin\theta F_\theta) - \frac{\partial F_\phi}{\partial\phi}\right]\overrightarrow{e}_r \\ &+ \frac{1}{r}\left[\frac{1}{\sin\theta}\frac{\partial F_r}{\partial\phi} - \frac{\partial}{\partial r}(rF_\phi)\right]\overrightarrow{e}_\theta \\ &+ \frac{1}{r}\left[\frac{\partial}{\partial r}(rF_\theta) - \frac{\partial F_r}{\partial\theta}\right]\overrightarrow{e}_\phi.\end{aligned}$$

5. For a linear, homogeneous and isotropic medium without charges, the Maxwell equations are given by

$$\overrightarrow{\nabla} \wedge \overrightarrow{H} = \frac{\partial\overrightarrow{D}}{\partial t}$$
$$\overrightarrow{\nabla} \wedge \overrightarrow{E} = -\frac{\partial\overrightarrow{B}}{\partial t}$$
$$\overrightarrow{\nabla} \cdot \overrightarrow{B} = 0$$
$$\overrightarrow{\nabla} \cdot \overrightarrow{D} = 0$$

where

$$\overrightarrow{D} = \epsilon_0\overrightarrow{E}$$
$$\overrightarrow{B} = \mu_0\overrightarrow{H}$$

and $\epsilon_0$ and $\mu_0$ are constant.

(a) Show that the wave propagation is expressed as :

$$\overrightarrow{\nabla}^2\overrightarrow{E} - \epsilon_0\mu_0\frac{\partial^2\overrightarrow{E}}{\partial t^2} = \overrightarrow{0}$$

You can use the identity $\overrightarrow{\nabla} \wedge (\overrightarrow{\nabla} \wedge \overrightarrow{A}) = -\overrightarrow{\nabla}^2\overrightarrow{A} + \overrightarrow{\nabla}(\overrightarrow{\nabla} \cdot \overrightarrow{A})$ for any vector $\overrightarrow{A}$.

(b) We assume that $\overrightarrow{E}(\overrightarrow{r}, t) = \overrightarrow{E}_0(\overrightarrow{r})e^{-j\omega t}$, $\overrightarrow{r} = (x, y, z)$. Show then

$$\left(\overrightarrow{\nabla}^2 + k_0^2\right)\overrightarrow{E}_0(\overrightarrow{r}) = \overrightarrow{0}$$

where $k_0 = \omega/c$, in which $c = 1/\sqrt{\epsilon_0\mu_0}$ is the wave speed in vacuum.

# 3 Matrix

## 3.1 Rectangular matrices

### 3.1.1 Definition

**Definition 3.1** *Definition of a matrix. In mathematics, a matrix (plural matrices) is a rectangular array of numbers, symbols, or expressions, arranged in rows and columns. It is expressed as*

$$[A] = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix},$$

where $\{a_{ij}\}$ are real numbers (it can be complex) and are called elements or entries.

In this course, a matrix will be named within brackets.

The **horizontal** elements

$$(a_{11}, a_{12}, \dots, a_{1n}), (a_{21}, a_{22}, \dots, a_{2n}), \dots, (a_{m1}, a_{m2}, \dots, a_{mn}),$$

are the **rows** of the matrix and the **vertical** elements

$$\begin{pmatrix} a_{11} \\ a_{21} \\ \dots \\ a_{m1} \end{pmatrix}, \begin{pmatrix} a_{12} \\ a_{22} \\ \dots \\ a_{m2} \end{pmatrix}, \dots, \begin{pmatrix} a_{1n} \\ a_{2n} \\ \dots \\ a_{mn} \end{pmatrix},$$

are the **columns**.

The element $a_{ij}$ corresponds to the intersection of the $i$-th row with the $j$-th column. A matrix being $m$ rows and $n$ columns is called a matrix of size $m \times n$, where the integer numbers $(m, n)$ give the **dimensions** or size of the marix.

Two matrices $[A]$ and $[B]$ of **same sizes** are equal if $\forall i \in [1; m],\ \forall j \in [1; n],\ a_{ij} = b_{ij}$.

**Exercice 3.1** *Let be the matrix defined as*

$$[A] = \begin{bmatrix} 1 & -3 & 4 \\ 0 & 5 & -2 \end{bmatrix}.$$

*Give the rows and the columns of* $[A]$.

**Exercice 3.2** *We have :*

$$\begin{bmatrix} x+y & 2z+w \\ x-y & z-w \end{bmatrix} = \begin{bmatrix} 3 & 5 \\ 1 & 4 \end{bmatrix}.$$

*Derive the values of* $x$, $y$, $w$ *and* $z$.

A matrix having one row can be considered as a *row vector* and a matrix having one column can be considered as a *column vector*.

### 3.1.2 Addition and multiplication by a scalar

Let be $[A]$ et $[B]$ two matrices of **same** size $m \times n$

$$[A] = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \quad \text{et} \quad [B] = \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \dots & b_{mn} \end{bmatrix}.$$

**Definition 3.2** *Addition of two matrices. The* ***addition*** *of* $[A]$ *and* $[B]$, $[A]+[B]$, *is the matrix obtained by adding the elements of the two matrices. Then*

$$\boxed{[A]+[B] = \begin{bmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \dots & a_{1n}+b_{1n} \\ a_{21}+b_{21} & a_{22}+b_{22} & \dots & a_{2n}+b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}+b_{m1} & a_{m2}+b_{m2} & \dots & a_{mn}+b_{mn} \end{bmatrix}}. \tag{3.1}$$

**Definition 3.3** *Matrix product by a scalar. The matrix product* $[A]$ *by a real scalar* $k$, $k[A]$, *is the matrix obtained by multiplying each element by* $k$. *Then*

$$\boxed{k[A] = \begin{bmatrix} ka_{11} & ka_{12} & \dots & ka_{1n} \\ ka_{21} & ka_{22} & \dots & ka_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ ka_{m1} & ka_{m2} & \dots & ka_{mn} \end{bmatrix}}. \tag{3.2}$$

We can note that $[A]+[B]$ and $k[A]$ are matrices $m \times n$. We can also define

$$-[A] = (-1) \times [A] \quad \text{and} \quad [A]-[B] = [A] + (-[B]).$$

The addition of two matrices of different sizes is not **defined**.

**Exercice 3.3** *We have*

$$[A] = \begin{bmatrix} 1 & -2 & 3 \\ 4 & 5 & -6 \end{bmatrix} \quad and \quad [B] = \begin{bmatrix} 3 & 0 & 2 \\ -7 & 1 & 8 \end{bmatrix}.$$

*Calculate* $[A] + [B]$*,* $3[A]$ *and* $2[A] - 3[B]$*.*

The null matrix is defined by $\forall i \in [1; m],\ \forall j \in [1; n],\ a_{ij} = 0$. It is named $[0]$. Then

$$[A] + [0] = [0] + [A] = [A].$$

### 3.1.3 Matrix product

The matrix product $[A]$ by $[B]$, $[A][B]$, is more complicated an then, we start by studying a particular case.

The matrix product $[A][B]$ of a **row** vector $[A]$ by a **column** vector $[B]$, having the same number of elements, is defined as

$$\boxed{[a_1, a_2, \ldots, a_n] \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} = a_1 b_1 + a_2 b_2 + \ldots + a_n b_n = \sum_{i=1}^{i=n} a_i b_i}. \tag{3.3}$$

We can note that the product $[A][B]$ is a real. The product $[A][B]$ is not defined if $[A]$ and $[B]$ have a number of elements that **differs**.

**Exercice 3.4** *Calculate the product*

$$[8, -4, 5] \begin{bmatrix} 3 \\ 2 \\ -1 \end{bmatrix}.$$

From the previous definition, it is then easy to calculate the matrix product for a general case.

**Definition 3.4** *Suppose that* $[A]$ *and* $[B]$ *are matrices, whose the number of columns of* $[A]$ ***equals*** *the number of rows of* $[B]$*, that is* $[A]$ *is matrix* $m \times p$ *and* $[B]$ *a matrix* $p \times n$*. Then, the matrix product* $[A][B]$ *is a matrix* $m \times n$*, where the* $ij-th$ *element is obtained multiplying the* $i-th$ *row of* $[A]$ *by the* $j-th$ *column of* $[B]$*. Then (see also figure 3.1)*

$$[A][B] = \begin{bmatrix} a_{11} & \dots & a_{1p} \\ \vdots & \ddots & \vdots \\ a_{i1} & \dots & a_{ip} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mp} \end{bmatrix} \begin{bmatrix} b_{11} & \dots & b_{1j} & \dots & b_{1n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ b_{p1} & \dots & b_{pj} & \dots & b_{pn} \end{bmatrix} = \begin{bmatrix} c_{11} & \dots & c_{1n} \\ \vdots & \ddots & \vdots \\ \vdots & c_{ij} & \vdots \\ \vdots & \ddots & \vdots \\ c_{m1} & \dots & c_{mn} \end{bmatrix} = [C], \tag{3.4}$$

where

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \ldots + a_{ip}b_{pj} = \sum_{k=1}^{k=p} a_{ik}b_{kj}. \tag{3.5}$$

It is important no note that the product $[A][B]$ is not defined if $[A]$ is a matrix $n \times p$ and $[B]$ a matrix $q \times n$, where $p \neq q$.

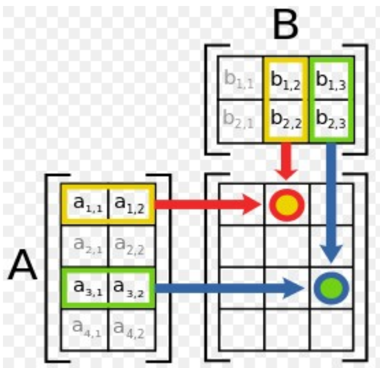



Figure 3.1: Illustration of the matrix product.

**Exercice 3.5** *Calculate the product*

$$\begin{bmatrix} r & s \\ t & u \end{bmatrix} \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{bmatrix}.$$

**Exercice 3.6** *Calculate the products and conclude.*

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix} \quad \textit{and} \quad \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}.$$

### 3.1.4 Transpose of a matrix

**Definition 3.5** *Transpose of a matrix. The transpose of an $m \times n$ matrix $[A]$ is the $n \times m$ $[A]^T$ formed by turning rows into columns and vice versa:*

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}^{\mathrm{T}} = \begin{bmatrix} a_{11} & a_{21} & \dots & a_{m1} \\ a_{12} & a_{22} & \dots & a_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \dots & a_{mn} \end{bmatrix}. \tag{3.6}$$

In other words, $b_{ij} = a_{ji}$ for all $i$ and $j$.

The transpose of a row vector is a column vector and vice versa.

### 3.1.5 Matrix and linear equation system

Consider the linear system of $m$ equations with $n$ unknowns

$$\begin{cases} a_{11}x_1 & +a_{12}x_2 & +\dots & +a_{1n}x_n & = b_1 \\ a_{21}x_1 & +a_{22}x_2 & +\dots & +a_{2n}x_n & = b_2 \\ \vdots & +\vdots & +\ddots & +\vdots & =\vdots \\ a_{m1}x_1 & +a_{m2}x_2 & +\dots & +a_{mn}x_n & = b_m \end{cases}.$$

This system can be casted into

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} \quad \text{or} \quad [A][X] = [B],$$

where the matrix $[A]$ is called the **coefficient** matrix, $[X]$ the column matrix (column vector) containing the **unknowns** and $[B]$ the column matrix (column vector) containing the **constants**.

**Exercice 3.7** *Cast the following linear system into matrix form:*

$$\begin{cases} 2x + 3y - 4z = 7 \\ x - 2y - 5z = 3 \end{cases}.$$

### 3.1.6 Summary of the properties

Consider $[A]$, $[B]$ and $[C]$ three matrices of same size and $k_1$, $k_2$ two reals, then

$$\begin{cases} ([A]+[B])+[C]=[A]+([B]+[C]) \\ [A]+[0]=[A] \\ [A]+(-[A])=[0] \\ [A]+[B]=[B]+[A] \\ k_1([A]+[B])=k_1[A]+k_1[B] \\ (k_1+k_2)[A]=k_1[A]+k_2[A] \\ 1\times[A]=[A] \\ 0\times[A]=[0] \end{cases}, \tag{3.7}$$

$$\begin{cases} ([A][B])[C]=[A]([B][C]) \\ [A]([B]+[C])=[A][B]+[A][C] \\ ([B]+[C])[A]=[B][A]+[C][A] \\ k([A][B])=(k[A])[B]=[A](k[B]) \\ [A][B]\neq[B][A] \\ [0][A]=[A][0]=[0] \end{cases}. \tag{3.8}$$

$$\begin{cases} ([A]+[B])^{\mathrm{T}}=[A]^{\mathrm{T}}+[B]^{\mathrm{T}} \\ \left([A]^{\mathrm{T}}\right)^{\mathrm{T}}=[A] \\ (k[A])^{\mathrm{T}}=k[A]^{\mathrm{T}} \\ ([A][B])^{\mathrm{T}}=[B]^{\mathrm{T}}[A]^{\mathrm{T}} \end{cases}. \tag{3.9}$$

## 3.2 Square matrices

**Definition 3.6** *Square matrices. A square matrix is a matrix with the same number of rows and columns. A square matrix $n\times n$ is called a matrix of dimension $n$.*

The addition, multiplication and transposition operations of a rectangular matrix remain valid for a square matrix. In addition, the resulting operation is also a square matrix of same dimension.

**Exercice 3.8** *We have*

$$[A]=\begin{bmatrix} 1 & 2 & 3 \\ -4 & -4 & -4 \\ 5 & 6 & 7 \end{bmatrix} \quad \text{and} \quad [B]=\begin{bmatrix} 2 & -5 & 1 \\ 0 & 3 & -2 \\ 1 & 2 & -4 \end{bmatrix}.$$

*Calculate $[A]+[B]$, $2[A]$, $[A]^T$ and $[A][B]$.*

**Definition 3.7** *Commuting matrix. Two matrices* $[A]$ *and* $[B]$ *are said to commute if* $[A][B] = [B][A]$.

**Exercice 3.9** *Consider*

$$[A] = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \quad et \quad [B] = \begin{bmatrix} 5 & 4 \\ 6 & 11 \end{bmatrix}.$$

*Calculate* $[A][B]$ *and* $[B][A]$. *Conclude.*

### 3.2.1 Linear transformation

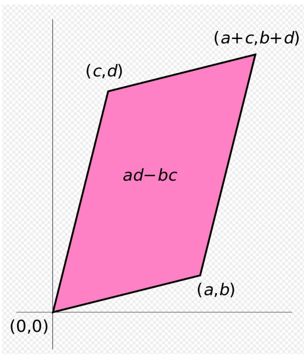



Figure 3.2: The vectors represented by a $2 \times 2$ matrix correspond to the sides of a unit square transformed into a parallelogram.

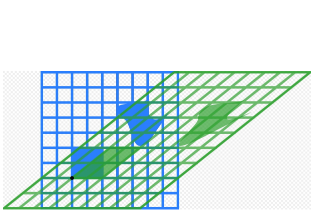
Matrices and matrix multiplication reveal their essential features when related to linear transformations, also known as linear maps.

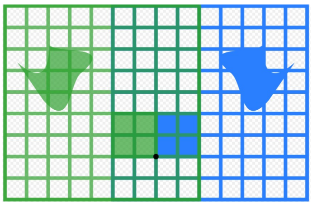
For instance, the $2 \times 2$ matrix

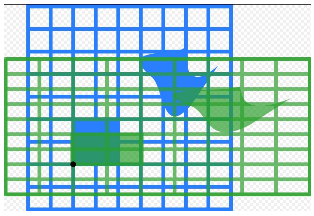
$$[A] = \begin{bmatrix} a & c \\ b & d \end{bmatrix}, \quad \text{then} \quad [A] \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} ax + cy \\ bx + dy \end{bmatrix},$$

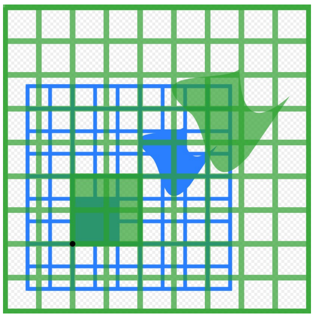
can be viewed as the transform of the unit square into a parallelogram with vertices at $(0, 0)$, $(a, b)$, $(a+c, b+d)$, and $(c, d)$. The parallelogram pictured in figure 3.2 is obtained by multiplying $[A]$ with each of the column vectors $[0\ 0]^{\mathrm{T}}$, $[1\ 0]^{\mathrm{T}}$, $[1\ 1]^{\mathrm{T}}$ and $[0\ 1]^{\mathrm{T}}$ in turn. These vectors define the vertices of the unit square.

Figure 3.3 shows a number of $2 \times 2$ matrices with the associated linear maps in a Cartesian system $(x, y)$. The blue original is mapped to the green grid and shapes. The origin $(0, 0)$ is marked with a black point.


(a) Horizontal shear with $m = 1.25$.


(b) Reflection through the vertical axis.


(c) Squeeze mapping with $r = 3/2$.


(d) Scaling by a factor of $3/2$.

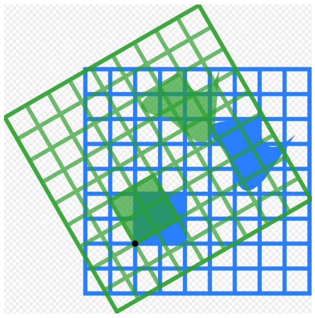

(e) Rotation by $\pi/6$.

Figure 3.3: Some linear transformations. The matrix of transformations are expressed below.

$$[A] = \begin{bmatrix} 1 & m \\ 0 & 1 \end{bmatrix} \quad \text{Horizontal shear with } m = 1.25,$$

$$[A] = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{Reflection through the vertical axis,}$$

$$[A] = \begin{bmatrix} r & 0 \\ 0 & 1/r \end{bmatrix} \quad \text{Squeeze mapping with } r = 3/2,$$

$$[A] = \begin{bmatrix} s & 0 \\ 0 & s \end{bmatrix} \quad \text{Scaling by a factor } s = 3/2,$$

$$[A] = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \quad \text{Rotation by } \theta = \pi/6.$$

### 3.2.2 Diagonal, trace of a matrix and identity matrix

**Definition 3.8** *Diagonal and trace of a matrix. The diagonal of a square matrix* $[A]$ *is the vector of components* $\{a_{11}, a_{22}, \ldots, a_{nn}\}$. *Its trace, trace*$[A]$*, is the sum of its diagonal elements. Then*

$$\boxed{\text{trace}[A] = a_{11} + a_{22} + \ldots + a_{nn} = \sum_{i=1}^{i=n} a_{ii}}. \tag{3.10}$$

**Definition 3.9** *Identity matrix. The identity matrix, $[I]$, is the $n \times n$ matrix in which all the elements on the main diagonal are equal to 1 and all other elements are equal to 0.*

In addition

$$\boxed{[A][I] = [I][A] = [A]}. \tag{3.11}$$

**Example 3.1** *The Kronecker symbol $\delta_{ij}$ is defined as*

$$\delta_{ij} = \begin{cases} 0 \ si\ i \neq j \\ 1 \ si\ i = j \end{cases} .$$

*Then, the elements of the identity matrix can be written as $[I]_{ij} = \delta_{ij}$ .*

The scalar matrices defined as $[A] = k[I]$ of dimensions, 2, 3 and 4 corresponding to the scalar $k = 5$ are respectively

$$\begin{bmatrix} 5 & 0 \\ 0 & 5 \end{bmatrix} \quad \begin{bmatrix} 5 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 5 \end{bmatrix} \quad \begin{bmatrix} 5 & & & \\ & 5 & & \\ & & 5 & \\ & & & 5 \end{bmatrix}.$$

### 3.2.3 Power and polynomial of a matrix

**Definition 3.10** *Power of a matrix. If $[A]$ is a square matrix of dimension $n$, then the power of $[A]$ is defined as*

$$\boxed{[A]^2 = [A][A] \quad [A]^3 = [A]^2[A] \quad [A]^{n+1} = [A]^n[A] \qquad \text{and} \quad [A]^0 = [I]}. \tag{3.12}$$

**Exercice 3.10** *Consider*

$$[A] = \begin{bmatrix} 1 & 2 \\ 3 & -4 \end{bmatrix}.$$

*Calculate $[A]^2$ and $[A]^3$.*

**Definition 3.11** *polynomial of a matrix. For any polynomial function defined as*

$$f(x) = a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n,$$

*where $\{a_i\}$ are real numbers, the polynomial of the square matrix $[A]$ is written as*

$$\boxed{f([A]) = a_0[I] + a_1[A] + a_2[A]^2 + \ldots + a_n[A]^n}. \tag{3.13}$$

We can note that $f([A])$ is obtained from $f(x)$ changing the variable $x$ by the matrix $[A]$ and the real $a_n$ by the scalar matrix $a_n[I] = a_n\delta_{ij}$.

For the case where $f([A])$ is a null matrix, the matrix $[A]$ is said a **zero** or a **root** of $f(x)$.

**Exercice 3.11** *Consider*

$$[A] = \begin{bmatrix} 1 & 2 \\ 3 & -4 \end{bmatrix}.$$

1. *Calculate* $[A]^2$ *and* $[A]^3$.
2. *Show then*

$$f([A]) = \begin{bmatrix} 16 & -18 \\ -27 & 61 \end{bmatrix},$$

*where* $f(x) = 2x^2 - 3x + 5$.

**Theorem 3.1** *If* $f(x)$ *and* $g(x)$ *are two polynomial functions and if* $[A]$ *is a square matrix of dimension* $n$*, then*

- $(f+g)([A]) = f([A]) + g([A])$.
- $(fg)([A]) = f([A])\, g([A])$.
- $f([A])\, g([A]) = g([A])\, f([A])$.

### 3.2.4 Diagonal, triangular, symmetric and orthogonal matrices

**Definition 3.12** *Diagonal matrix. A square matrix* $[D]$ *is said* ***diagonal*** *if all the off-diagonal elements equal zero. It can be noted as* $[D] = Diag(d_{11}, d_{22}, \ldots, d_{nn})$*, where* $\{d_{ii}\}$ *are real numbers.*

**Example 3.2**

$$\begin{bmatrix} 3 & 0 & 0 \\ 0 & -7 & 0 \\ 0 & 0 & 2 \end{bmatrix} \quad \begin{bmatrix} 4 & 0 \\ 0 & -5 \end{bmatrix} \quad \begin{bmatrix} 6 & & & \\ & 0 & & \\ & & -9 & \\ & & & 1 \end{bmatrix}.$$

The identity matrix is a particular case of a diagonal matrix.

**Definition 3.13** *Upper and lower triangular matrices. If all elements of the square matrix* $[A]$ ***below*** *the main diagonal are zero,* $[A]$ *is called an* ***upper*** *triangular matrix. Similarly, if all the elements of* $[A]$ ***above*** *the main diagonal are zero,* $[A]$ *is called a* ***lower*** *triangular matrix.*

The upper triangular matrices of dimensions 2, 3 and 4 are respectively

$$\begin{bmatrix} a_{11} & a_{12} \\ & a_{22} \end{bmatrix} \quad \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ & a_{22} & a_{23} \\ & & a_{33} \end{bmatrix} \quad \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ & a_{22} & a_{23} & a_{24} \\ & & a_{33} & a_{34} \\ & & & a_{44} \end{bmatrix}.$$

The blocks of zeros are removed.

**Definition 3.14** *Orthogonal matrix. An orthogonal matrix is a square matrix with real elements whose columns and rows are orthogonal unit vectors (that is, orthonormal vectors). Equivalently, a matrix $[A]$ is orthogonal if its transpose is equal to its inverse:*

$$[A]^{\mathrm{T}} = [A]^{-1} \quad \text{then} \quad [A][A]^{\mathrm{T}} = [I].$$

**Exercice 3.12** *Consider*

$$[A] = \frac{1}{9}\begin{bmatrix} 1 & 8 & -4 \\ 4 & -4 & -7 \\ 8 & 1 & 4 \end{bmatrix}.$$

*Show that $[A]$ is an orthogonal matrix.*

Let be a matrix of dimension 3

$$[A] = \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix}.$$

If $[A]$ is orthogonal, then

$$[A][A]^{\mathrm{T}} = \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix}\begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

This leads to

$$\begin{cases} a_1^2 + a_2^2 + a_3^2 = 1 & a_1b_1 + a_2b_2 + a_3b_3 = 0 & a_1c_1 + a_2c_2 + a_3c_3 = 0 \\ b_1a_1 + b_2a_2 + b_3a_3 = 0 & b_1^2 + b_2^2 + b_3^2 = 1 & b_1c_1 + b_2c_2 + b_3c_3 = 0 \\ c_1a_1 + c_2a_2 + c_3a_3 = 0 & c_1b_1 + c_2b_2 + c_3b_3 = 0 & c_1^2 + c_2^2 + c_3^2 = 1 \end{cases}.$$

Then

$$\begin{cases} \vec{u}_1 \cdot \vec{u}_1 = 1 & \vec{u}_1 \cdot \vec{u}_2 = 0 & \vec{u}_1 \cdot \vec{u}_3 = 0 \\ \vec{u}_2 \cdot \vec{u}_1 = 0 & \vec{u}_2 \cdot \vec{u}_2 = 1 & \vec{u}_2 \cdot \vec{u}_3 = 0 \\ \vec{u}_3 \cdot \vec{u}_1 = 0 & \vec{u}_3 \cdot \vec{u}_2 = 0 & \vec{u}_3 \cdot \vec{u}_3 = 1 \end{cases},$$

where $\vec{u}_1 = [a_1\ a_2\ a_3]$, $\vec{u}_2 = [b_1\ b_2\ b_3]$, $\vec{u}_3 = [c_1\ c_2\ c_3]$ are the rows of the matrices $[A]$ (row vectors). The row vectors $\vec{u}_1$, $\vec{u}_2$ and $\vec{u}_3$ are mutually orthogonal and of unit norms. In other words, the vectors $\vec{u}_1$, $\vec{u}_2$ and $\vec{u}_3$ form an orthogonal unit basis. The condition $[A][A]^{\mathrm{T}}$ also shows that the column vectors form an orthogonal unit basis.

In compact form, we have

$$\boxed{\vec{u}_i \cdot \vec{u}_j = \delta_{ij}}. \tag{3.14}$$

### 3.2.5 Determinant of a matrix

In linear algebra, the determinant is a useful value that can be computed from the elements of a square matrix. In this course, the determinant of a matrix [A] is denoted det[A]. It can be viewed as the scaling factor of the transformation described by the matrix.

For a given square matrix [A] of dimension $n$, we can calculate a real number called *determinant* of [A]. Then

$$\det \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}.$$

Le determinant is not a matrix but a real number.

The use of determinants in calculus includes the Jacobian determinant in the change of variables rule for integrals of functions of several variables. Determinants are also used to define the characteristic polynomial of a matrix, which is essential for eigenvalue problems in linear algebra.

#### 3.2.5.1 Determinants of orders 1 and 2

The determinants of orders 1 and 2 are defined as follows

$$\boxed{\det[A] = a_{11} \quad \det[A] = \det \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = a_{11}a_{22} - a_{12}a_{21}}. \tag{3.15}$$

As shown in figure 3.2, the absolute value of $ad - bc$ is the area of the parallelogram, and thus represents the scale factor by which areas are transformed by [A], where

$$[A] = \begin{bmatrix} a & c \\ b & d \end{bmatrix}.$$

**Exercice 3.13** *Show that*

$$\det \begin{bmatrix} 5 & 4 \\ 2 & 3 \end{bmatrix} = 7 \quad \textit{and} \quad \det \begin{bmatrix} 2 & 1 \\ -4 & 6 \end{bmatrix} = 16.$$

A direct application of the determinant is the resolution of linear system of two unknowns. If

$$\begin{cases} a_1x + b_1y = c_1 \\ a_2x + b_2y = c_2 \end{cases} \Leftrightarrow \underbrace{\begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \end{bmatrix}}_{[A]} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}.$$

Then, the solution can be expressed as

$$\left\{\begin{array}{l} x = \dfrac{D_x}{\det[A]} = \dfrac{\det\begin{bmatrix} c_1 & b_1 \\ c_2 & b_2 \end{bmatrix}}{\det[A]} = \dfrac{c_1 b_2 - c_2 b_1}{a_1 b_2 - a_2 b_1} \\ y = \dfrac{D_y}{\det[A]} = \dfrac{\det\begin{bmatrix} a_1 & c_1 \\ a_2 & c_2 \end{bmatrix}}{\det[A]} = \dfrac{a_1 c_2 - a_2 c_1}{a_1 b_2 - a_2 b_1} \end{array}\right. . \tag{3.16}$$

The numerators $D_x$ et $D_y$ giving $x$ et $y$, respectively, can be obtained changing the column of the coefficients of the unknowns to determine, in the matrix of the coefficients, by the column of the constant terms.

**Exercice 3.14** *Show that the solution of the following linear system:*

$$\left\{\begin{array}{l} 2x - 3y = 7 \\ 3x + 5y = 1 \end{array}\right. ,$$

*is $x = 2$ and $y = -1$.*

#### 3.2.5.2 Determinant of order 3

The determinant of order 3 can be expressed from determinants of order 2 by linear combinations. then

$$\begin{aligned} \det[A] &= \det\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \\ &= +a_{11}\det\begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix} - a_{12}\det\begin{bmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{bmatrix} + a_{13}\det\begin{bmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} \\ &= a_{11}\left(a_{22}a_{33} - a_{23}a_{32}\right) - a_{12}\left(a_{21}a_{33} - a_{23}a_{31}\right) + a_{13}\left(a_{21}a_{32} - a_{22}a_{31}\right). \end{aligned} \tag{3.17}$$

The rule of Sarrus is a mnemonic for the $3\times3$ matrix determinant: the sum of the products of three diagonal north-west to south-east lines of matrix elements, minus the sum of the products of three diagonal south-west to north-east lines of elements, when the copies of the first two columns of the matrix are written beside it as in figure 3.4.

As shown in figure 3.5, the volume of the parallelepiped is the absolute value of the determinant of the matrix formed by the rows constructed from the vectors $\overrightarrow{r}_1$, $\overrightarrow{r}_2$ and $\overrightarrow{r}_3$.

**Exercice 3.15** *Show that the determinant of the following matrix:*

$$[A] = \begin{bmatrix} 2 & 1 & -1 \\ 1 & 1 & 1 \\ 1 & -2 & -3 \end{bmatrix},$$

*is 5.*

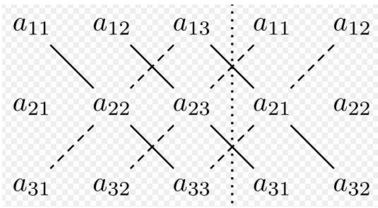



Figure 3.4: Sarrus'rule: The determinant of the three columns on the left is the sum of the products along the solid diagonals minus the sum of the products along the dashed diagonals.

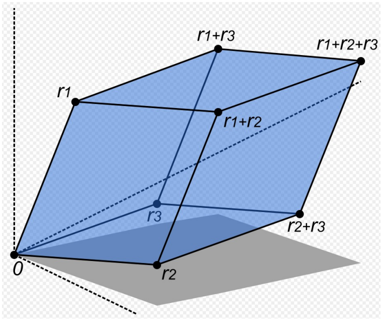



Figure 3.5: Volume of the parallelepiped.

#### 3.2.5.3 Determinant of order $n$

**Definition 3.15** *Cofactor. Consider a square matrix $[A]$ of dimension $n$. We define the square matrix $[M]$ of dimension $n-1$ that results from $[A]$ by removing the $i$-th row and the $j$-th column. The determinant of $[M]$ is called minor of the element $a_{ij}$ of the matrix $[A]$. The cofactor of $a_{ij}$, denoted as $c_{ij}$, is the minor $m_{ij}$ affected of its signature defined by*

$$\boxed{c_{ij} = (-1)^{i+j} m_{ij} = (-1)^{i+j} \det[M]}. \tag{3.18}$$

We can note that the signs $(-1)^{i+j}$ of the minors form a chessboard arrangement, with the sign + on the main diagonal. Then

$$\begin{bmatrix} + & - & + & - & \dots \\ - & + & - & + & \dots \\ + & - & + & - & \dots \\ \ddots & \ddots & \ddots & \ddots & \ddots \end{bmatrix}.$$

**Exercice 3.16** *Show that $m_{23} = -6$ and $c_{23} = 6$ of the following square matrix:*

$$[A] = \begin{bmatrix} 2 & 3 & 4 \\ 5 & 6 & 7 \\ 8 & 9 & 1 \end{bmatrix}.$$

**Theorem 3.2** *Determinant of order n. The determinant of the square matrix [A] is equal to the sum of the products obtained multiplying the elements of an any row (or column, respectively) by their respective cofactors. Then*

$$\boxed{\det[A] = \begin{cases} a_{i1}c_{i1} + a_{i2}c_{i2} + \ldots + a_{in}c_{in} = \displaystyle\sum_{j=1}^{j=n} a_{ij}c_{ij} \\ a_{1j}c_{1j} + a_{2j}c_{2j} + \ldots + a_{nj}c_{nj} = \displaystyle\sum_{j=1}^{j=n} a_{ij}c_{ij} \end{cases}}. \tag{3.19}$$

**Exercice 3.17** *Show that the determinant of the following matrix:*

$$[A] = \begin{bmatrix} 1 & 2 & 3 & -2 \\ -1 & 3 & 1 & -1 \\ 0 & 1 & -1 & 3 \\ -2 & 0 & 1 & 1 \end{bmatrix}.$$

is equal to $-63$.

A direct application of the determinant is the resolution of a linear system of $n$ unknowns. Consider

$$\begin{cases} a_{11}x_1 & +a_{12}x_2 & +\ldots & +a_{1n}x_n & = b_1 \\ a_{21}x_1 & +a_{22}x_2 & +\ldots & +a_{2n}x_n & = b_2 \\ \vdots & +\vdots & +\ddots & +\vdots & = \vdots \\ a_{n1}x_1 & +a_{n2}x_2 & +\ldots & +a_{nn}x_n & = b_n \end{cases}.$$

This linear system can be casted into matrix form as $[A][X] = [B]$, where $[A]$ is a square matrix of dimension $n$ whose elements are $a_{ij}$, $[X]$ the column vector of components the unknown $x_i$ and $[B]$ the column vector of components the constants $b_i$.

Let $[M]$ be the matrix obtained from the matrix $[A]$ changing the $i-$th column by the column vector $[B]$:

$$[M] = \begin{bmatrix} a_{11} & a_{12} & \ldots & b_1 & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & b_2 & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \ldots & b_n & \ldots & a_{nn} \end{bmatrix}.$$

**Theorem 3.3** *The previous linear system has an unique solution if, and only if $D = \det[M] \neq 0$ and the solution is given by*

$$\boxed{x_1 = \frac{D_1}{D} \quad x_2 = \frac{D_2}{D} \quad \ldots x_n = \frac{D_n}{D}}. \tag{3.20}$$

In addition:

**Theorem 3.4** *The system $[A][X] = [0]$ has a non-null solution if, and only if $D = \det[A] = 0$.*

**Exercice 3.18** *Show that the solution of following linear system:*

$$\begin{cases} 2x + y - z = 3 \\ x + y + z = 1 \\ x - 2y - 3z = 4 \end{cases},$$

*is* $x = 2$, $y = -1$ *and* $z = 0$.

### 3.2.6 Inverse of a matrix

**Definition 3.16** *Adjoint matrix. Consider the square matrice* $[A]$ *of dimension* $n$. *The adjoint matrix of* $[A]$, *called Adj*$[A]$, *is the transpose of the cofactors matrix* $[C]$ *of elements* $a_{ij}$ *of* $[A]$. *Then*

$$\boxed{\mathrm{Adj}[A] = \begin{bmatrix} c_{11} & c_{12} & \dots & c_{1n} \\ c_{21} & c_{22} & \dots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{n1} & c_{n2} & \dots & c_{nn} \end{bmatrix}^{\mathrm{T}} = \begin{bmatrix} c_{11} & c_{21} & \dots & c_{n1} \\ c_{12} & c_{22} & \dots & c_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ c_{1n} & c_{2n} & \dots & c_{nn} \end{bmatrix}}. \tag{3.21}$$

**Exercice 3.19** *We define*

$$[A] = \begin{bmatrix} 2 & 3 & -4 \\ 0 & -4 & 2 \\ 1 & -1 & 5 \end{bmatrix}.$$

*Then show*

$$Adj\,[A] = \begin{bmatrix} -18 & -11 & -10 \\ 2 & 14 & -4 \\ 4 & 5 & -8 \end{bmatrix}.$$

**Definition 3.17** *Inverse of a matrix. A square matrix* $[A]$ *is called invertible or non-singular if there exists a matrix* $[B]$ *such that* $[A][B] = [I]$. *Then* $[B] = [A]^{-1}$ *and*

$$\boxed{[A]^{-1} = \frac{1}{\det[A]}\mathrm{Adj}[A] = \frac{1}{\det[A]}[C]^{\mathrm{T}}}, \tag{3.22}$$

if $\det[A] \neq 0$.

If $[A]$ is an orthogonal matrix, then $[A][A]^{\mathrm{T}} = [I]$ and $[A]^{-1} = [A]^{\mathrm{T}}$.

**Exercice 3.20** *Consider*

$$[A] = \begin{bmatrix} 2 & 3 & -4 \\ 0 & -4 & 2 \\ 1 & -1 & 5 \end{bmatrix}.$$

*Then show*

$$[A]^{-1} = \begin{bmatrix} \frac{9}{23} & \frac{11}{46} & \frac{5}{23} \\ -\frac{1}{23} & -\frac{7}{23} & \frac{2}{23} \\ -\frac{2}{23} & -\frac{5}{46} & \frac{4}{23} \end{bmatrix}.$$

## 3.2.7 Properties

Consider $[A]$ and $[B]$ two square matrices of same dimension and $k$ a real number. Then

$$\begin{cases} ([A][B])^{-1} = [B]^{-1}[A]^{-1} \\ \det\left([A]^{\mathrm{T}}\right) = \det[A] \\ \det([A][B]) = \det[A]\det[B] \\ \text{If } [A] \text{ has 1 row (or 1 column) of zeros, then } \det[A] = 0 \\ \text{If } [A] \text{ has 2 identical rows (or 2 columns), then } \det[A] = 0 \\ \text{If } [A] \text{ is a triangular matrix, then } \det[A] = \prod_{i=1}^{i=n} a_{ii} \\ \text{If 2 rows (or 2 columns) of } [A] \text{ is swapped giving } [B] \text{, then } \det[B] = -\det[A] \\ \text{If 1 row (or 1 column) of } [A] \text{ is multiplied by } k \text{, then } \det[B] = k\det[A] \end{cases} . \tag{3.23}$$

and

$$\begin{cases} \mathrm{trace}([A]+[B]) = \mathrm{trace}([A]) + \mathrm{trace}([B]) \\ \mathrm{trace}(k[A]) = k\mathrm{trace}([A]) \\ \mathrm{trace}\left([A]^{\mathrm{T}}\right) = \mathrm{trace}([A]) \\ \mathrm{trace}([A][B]) = \mathrm{trace}([B][A]) \end{cases} . \tag{3.24}$$

## 3.2.8 Diagonalisation : Eigen values and eigen vectors

### 3.2.8.1 Eigen values and eigen vectors

**Definition 3.18** *Eigen values and eigen vectors. Consider a square matrix $[A]$. A scalar $\lambda$ is called* ***eigen value*** *of $[A]$ if is exist a non nul column vector that satisfies*

$$\boxed{[A][v] = \lambda[v]}. \tag{3.25}$$

*The vector $[v]$ satisfying this equation is called* ***eigen vector*** *associated to the eigen value $\lambda$.*

From this definition, it is easy to show that the vector $k[v]$ is also an eigen vector of eigen value $\lambda$.

As shown in figure 3.6, in a Cartesian basis $(x, y)$, an eigen vector is parallel or collinear to $[x]$.

**Example 3.3** *We consider the square matrix*

$$[A] = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$

*We will show that the eigen values of $[A]$ are $\lambda_1 = 1$ and $\lambda_2 = 3$ and the associated eigen vectors are $[v_1] = [1\ \ -1]^T$ and $[v_1] = [1\ 1]^T$. Figure 3.7 then shows the linear transformation $[A][v]$ and the links with the eigen values and eigen vectors.*

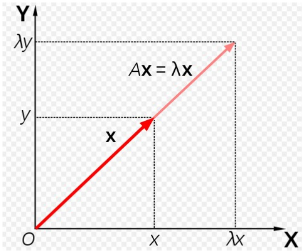
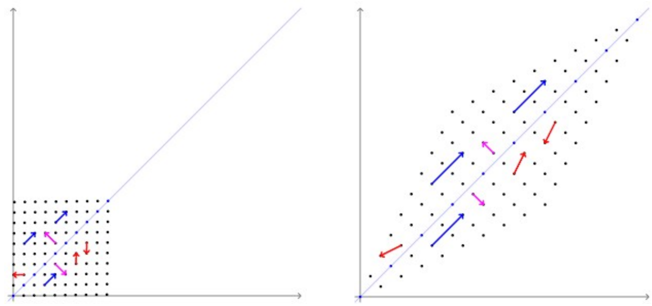



Figure 3.6: Matrix $[A]$ acts by stretching the vector $[x]$, not changing its direction, so $[x]$ is an eigenvector of $[A]$.



(a) Map of the input vector.

(b) Map of the output vector.

Figure 3.7: On the left, map of the vector $[v] = [x\ y]^{\mathrm{T}}$. On the right, map of the transformation $[A][v]$; the transformation matrix $[A]$ preserves the direction of vectors parallel to $[v_1] = [1\ -1]^{\mathrm{T}}$ (in purple, where $\lambda_1 = 1$) and $[v_1] = [1\ 1]^{\mathrm{T}}$ (in blue, where $\lambda_2 = 3$). The vectors in red are not parallel to either eigenvector, so, their directions are changed by the transformation.

**Theorem 3.5** *Diagonalization of a matrix. A square matrix of dimension $n$ is diagonalizable, if and only if it hase $n$ eigen vectors* ***linearly independent****. The resulting diagonal matrix $[D]$ of $[A]$ have for elements the eigen values and the matrice $[P]$ (of transformation) such as $[D] = [P]^{-1}[A][P]$ has for columns the corresponding eigen vectors.*

If a square matrix $[A]$ is diagonalisable, that is $[P]^{-1}[A][P] = [D]$, in which the matrix $[D]$ is diagonal, the following equation is very useful to write $[A]$ from a diagonal matrix. It is called **diagonal decomposition** of $[A]$. Then

$$\boxed{[A] = [P][D][P]^{-1}}. \tag{3.26}$$

For instance, this formula is very useful to calculate the power of a matrix since

$$\boxed{[A]^m = \left([P][D][P]^{-1}\right)^m = [P][D]^m[P]^{-1} = [P]\mathrm{Diag}\left(k_1^m, k_2^m, \ldots, k_n^m\right)[P]^{-1}}, \tag{3.27}$$

where $[D] = \mathrm{Diag}\left(k_1, k_2, \ldots, k_n\right)$.

**Exercice 3.21** *We consider*

$$[A] = \begin{bmatrix} 3 & 1 \\ 2 & 2 \end{bmatrix} \quad [v_1] = \begin{bmatrix} 1 \\ -2 \end{bmatrix} \quad [v_2] = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

1. *Show that the vectors* $[v_1]$ *and* $[v_2]$ *are eigen vectors of the matrix* $[A]$ *and calculate the respective eigen values,* $\lambda_1$ *and* $\lambda_2$.

2. *Deduce the associated transformation matrix* $[P]$ *and show that*

$$[P]^{-1} = \begin{bmatrix} \frac{1}{3} & -\frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} \end{bmatrix}.$$

3. *Calculate* $[P]^{-1}[A][P]$. *Conclude.*

4. *Show that* $[P][D][P]^{-1} = [A]$.

5. *Show that*

$$[A]^4 = \begin{bmatrix} 171 & 85 \\ 170 & 86 \end{bmatrix}.$$

6. *Show that*

$$[A]^{\frac{1}{2}} = \begin{bmatrix} \frac{5}{3} & \frac{1}{3} \\ \frac{2}{3} & \frac{4}{3} \end{bmatrix}.$$

#### 3.2.8.2 Characteristic polynomial

**Definition 3.19** *Characteristic polynomial. Let* $[A]$ *be a square matrix of dimension* $n$ *and let* $[M]$ *be the matrix defined as* $[M] = [A] - \lambda[I]$, *where* $[I]$ *is the identity matrix and* $\lambda$ *an unknown number. Then, the polynomial of degree* $n$ *defined as*

$$\boxed{D(\lambda) = \det(\lambda[I] - [A]) = (-1)^n \det([A] - \lambda[I])}, \tag{3.28}$$

*is called characteristic polynomial of the matrix* $[A]$.

**Theorem 3.6** *Cayley-Hamilton theorem. Any mmatirx* $[A]$ *is root of its characteristic polynomial. Then* $D([A]) = [0]$.

**Exercice 3.22** *Consider*

$$[A] = \begin{bmatrix} 1 & 3 \\ 4 & 5 \end{bmatrix}.$$

1. *Show that* $D(\lambda) = \lambda^2 - 6\lambda - 7$.

2. *Show that*

$$[A]^2 = \begin{bmatrix} 13 & 18 \\ 24 & 37 \end{bmatrix}.$$

3. *Show that* $D([A]) = [0]$.

**Exercice 3.23** *If* $[A]$ *is a square matrix of dimension 2, then show*

$$\boxed{D(\lambda) = \lambda^2 - \lambda trace[A] + \det[A]}. \tag{3.29}$$

**Exercice 3.24** *If $[A]$ is a square matrix of dimension 3, then show*

$$\boxed{D(\lambda) = \lambda^3 - \lambda^2 trace[A] + \lambda\left(c_{11} + c_{22} + c_{33}\right) - \det[A]}, \tag{3.30}$$

*where the reals $\{c_{ii}\}$ are the cofactors of the elements $\{a_{ii}\}$ of $[A]$.*

#### 3.2.8.3 Calculation of the eigen values and vectors

This section provides an algorithm to calculate the eigen values and eigen vectors of a square matrix and establishes l'existence of a regular (or invertible) matrix $[P]$ such as the matrix $[P]^{-1}[A][P]$ is diagonal.

1. Find the characteristic polynomial $D(\lambda)$ of $[A]$.
2. Determine the roots of $D(\lambda)$, which are the eigen values $\{\lambda_i\}$ of $[A]$.
3. For each of the eigen value $\lambda_i$ of $[A]$, repeat the following two items:
   a) Construct the matrix $[M_i] = [A] - \lambda_i[I]$.
   b) Determine the solution of the homogeneous linear system $[M_i][X] = [0]$: this vector is an eigen vector of $[A]$ linearly independent and of eigen value $\lambda_i$.
4. Study the system $S = \{[v_1], [v_2], \ldots, [v_m]\}$ obtained at the third step:
   a) If $m \neq n$, $[A]$ is not diagonalisable.
   b) If $m = n$, $[A]$ is diagonalisable. Construct then $[P]$, whose the columns are the column vectors $[v_1], [v_2], \ldots, [v_n]$. Then

   $$D(\lambda) = [P]^{-1}[A][P] = \text{Diag}\left(\lambda_1, \lambda_2, \ldots, \lambda_n\right), \tag{3.31}$$

   where $\lambda_i$ is the eigen value of the eigen vector $[v_i]$.

**Exercice 3.25** *We consider*

$$[A] = \begin{bmatrix} 4 & 2 \\ 3 & -1 \end{bmatrix}.$$

1. *Show that $D(\lambda) = (\lambda - 5)(\lambda + 2)$.*
2. *Calculate the eigen values $\lambda_1$ and $\lambda_2 < \lambda_1$ of the matrix $[A]$.*
3. *Show that the associated eigen vectors are*

   $$[v_1] = \begin{bmatrix} 2 \\ 1 \end{bmatrix} \quad [v_2] = \begin{bmatrix} -1 \\ 3 \end{bmatrix}.$$

4. *Deduce the matrix $[P]$ and calculate $[P]^{-1}$.*
5. *Then show that*

   $$[D] = [P]^{-1}[A][P] = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}.$$

#### 3.2.8.4 Diagonalisation of real symmetric matrices

**Theorem 3.7** *If $[A]$ is a (real) symmetric matrix, then the roots of the polynomial characteristic are real.*

**Theorem 3.8** *If $[u]$ and $[v]$ are two eigen vectors of a (real) symmetric matrix $[A]$, then $[u]$ and $[v]$ are orthogonal.*

These two theorems lead to the following fundamental theorem:

**Theorem 3.9** *If $[A]$ is a (real) symmetric matrix, then it exists an **orthogonal** matrix $[P]$ such as the matrix $[D] = [P]^{-1}[A][P]$ is diagonal with $[P]^{-1} = [P]^T$.*

As shown latter, the orthogonal matrix $[P]$ is obtained by **normalizing** the eigen vectors.

**Exercice 3.26** *We consider the following symmetric matrix*

$$[A] = \begin{bmatrix} 2 & -2 \\ -2 & 5 \end{bmatrix}.$$

1. *Show that the eigen values of the matrix $[A]$ are $\lambda_1 = 6$ and $\lambda_2 = 1$.*
2. *Show that the associated eigen vectors are*

$$[v_1] = \begin{bmatrix} 1 \\ -2 \end{bmatrix} \quad [v_2] = \begin{bmatrix} 2 \\ 1 \end{bmatrix}.$$

3. *Show that the vectors $[v_1]$ and $[v_2]$ are orthogonal.*
4. *Calculate the associated normalized vector $[\hat{v}_1]$ and $[\hat{v}_2]$.*
5. *Deduce de matrix $[P]$.*
6. *Verify that $[P]$ is an orthogonal matrix.*
7. *Show that*

$$[D] = [P]^{-1}[A][P] = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}.$$

#### 3.2.8.5 Property of the eigen values

Let $[A]$ be an arbitrary $n$ square complex matrix of eigen values $\{\lambda_{i\in[1;n]}\}$.
The trace of $[A]$, defined as the sum of its diagonal elements, is also the sum of all eigenvalues:

$$\text{trace}[A] = \sum_{i=1}^{n} a_{ii} = \sum_{i=1}^{n} \lambda_i.$$

The determinant of $[A]$ is the product of all its eigenvalues:

$$\det[A] = \lambda_1 \lambda_2 \dots \lambda_n.$$

### 3.2.9 Hermitian matrix

**Definition 3.20** *Hermitian matrix. A Hermitian matrix (or self-adjoint matrix) is a complex square matrix that is equal to its own conjugate transpose, that is, the element in the $i$-th row and $j$-th column is equal to the complex conjugate of the element in the $j$-th row and $i$-th column, for all indexes $i$ and $j$:*

$$a_{ij} = \bar{a}_{ji}, \text{ or } [A] = \overline{[A]^T} = [A]^H \text{ in matrix form.}$$

**Example 3.4**

$$[A]^H = \begin{bmatrix} 2 & 1+i & -2+2i \\ 1-i & 7 & i \\ -2-2i & -i & 6 \end{bmatrix}.$$

The diagonal elements must be real, as they must be their own complex conjugate.

**Property 3.1**

- *The entries on the main diagonal (top left to bottom right) of any Hermitian matrix are necessarily real, because they have to be equal to their complex conjugate.*
- *A matrix that has only real elements is Hermitian if and only if it is a symmetric matrix, that is, if it is symmetric with respect to the main diagonal. A real and symmetric matrix is simply a special case of a Hermitian matrix.*
- *All eigenvalues of a Hermitian matrix $[A]$ with dimension $n$ are **real**, and that $[A]$ has $n$ linearly independent eigenvectors. Moreover, Hermitian matrix has orthogonal eigenvectors for distinct eigenvalues.*
- *For an arbitrary complex valued vector $[v]$ the product $[v]^H[A][v]$ is real because of $[v]^H[A][v] = ([v]^H[A][v])^H$.*
- *If $n$ orthonormal eigenvectors $[v_1], [v_2], \ldots, [v_n]$ of a Hermitian matrix are chosen and written as the columns of the matrix $[P]$, then one eigen decomposition of $[A]$ is $[A] = [P][D][P]^H$ where $[P][P]^H = [P]^H[P] = [I]$ and therefore*

  $$[A] = \sum_{i=1}^{n} \lambda_i [v_i][v_i]^H,$$

  *where $\{\lambda_i\}$ are the eigenvalues on the diagonal of the diagonal matrix $[D]$.*

The last property it is a generalization of equation (3.26) valid for a real symmetric matrix.

## 3.3 Homework

Make the exercises on a copy with a **clean** presentation and **underline** the final results. Do not forgot to write your name and surname on all sheets. The copy will be read by the Professor in the next session (course).

### 3.3.1 Matrix product

1. We consider

$$[A] = \begin{bmatrix} 2 & 3 & -1 \\ 4 & -2 & 5 \end{bmatrix} \quad \text{and} \quad [B] = \begin{bmatrix} 2 & -1 & 0 & 6 \\ 1 & 3 & -5 & 1 \\ 4 & 1 & -2 & 2 \end{bmatrix}.$$

   Calculate $[A][B]$.

2. Find a non-diagonal $2 \times 2$ matrix $[A]$ whose elements are non null and such as $[A]^2$ is diagonal.

3. Find an upper triangular real matrix $[A]$ such as

$$[A]^3 = \begin{bmatrix} 8 & -57 \\ 0 & 27 \end{bmatrix}.$$

### 3.3.2 Determinant and linear system

1. Calculate the determinant of the following matrices:

$$[A] = \begin{bmatrix} 2 & 3 & 4 \\ 5 & 4 & 3 \\ 1 & 2 & 1 \end{bmatrix} \quad [B] = \begin{bmatrix} 1 & -2 & 3 \\ 2 & 4 & -1 \\ 1 & 5 & -2 \end{bmatrix} \quad [C] = \begin{bmatrix} 1 & 3 & -5 \\ 3 & -1 & 2 \\ 1 & -2 & 1 \end{bmatrix}.$$

2. From the determinant method, solve the following system

$$\begin{cases} 3y - 2x = z + 1 \\ 3x + 2z = 8 - 5y \\ 3z - 1 = x - 2y \end{cases}.$$

### 3.3.3 Eigen values and eigen vectors

1. Calculate the characteristic polynomial $D$ and its roots of the following matrices:

$$[A] = \begin{bmatrix} 2 & 5 \\ 4 & 1 \end{bmatrix} \quad [B] = \begin{bmatrix} 3 & -2 \\ 9 & -3 \end{bmatrix}.$$

2. Calculate the characteristic polynomial $D$ and its roots of the following matrix:

$$[A] = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 0 & 4 \\ 6 & 4 & 5 \end{bmatrix}.$$

   We can note that $\lambda = -2$ is a root of $D$

3. We consider

$$[A] = \begin{bmatrix} 3 & -4 \\ 2 & -6 \end{bmatrix}.$$

   (a) Calculate the eigen values and eigen vectors of $[A]$.

   (b) Find a regular matrix (invertible) $[P]$ and a diagonal matrix $[D]$ such as $[D] = [P]^{-1}[A][P]$.

   (c) Deduce $[A]^4$.

   (d) Calculate $f([A])$ where $f(x) = 2 + 3x + x^4$.

# A Practical work 1: Initiation to the software MatLab

The software MatLab is a relevant tool to make numerical computation. The syntax programming is very easy and intuitive because the names of the commands are the direct translation in English of the truncated term. For instance, if you want to compute an integral, use the command int.

The command **help** is very useful to obtain further information on the command (for example, help int).

## A.1 Vector (of 1 dimension)

The power of MatLab is to make operation on a matrix and then on a vector, which is a particular of a matrix.

### A.1.1 Vectors and their manipulation

1. Write $a = 0 : 0.2 : 1$ and $b = 3 * a + 2$, $a(1)$, $a(2)$. Comment the obtained results.
2. Compute $size(a)$, $length(a)$.
3. Compute $a + b$, $a - b$.
4. Write $C = [1, 2, 3]$, $D = [C, 4, 5, 6]$, $length(C)$, $length(D)$, $E = D(1 : 2)$ and $length(E)$. Explain the interest of these operations.
5. Compute $sum(a)$, $prod(a)$, $mean(a)$. Check the results by "hand".
6. Compare: $a' * b$, $a * b'$, $a. * b$. Well understand the difference between these operations (careful between $*$ and $.*$).

### A.1.2 MatLab file and plotting a function

1. Click on "file", "New" and "Script" and write the following lines:

```
clear all ; close all
```

```
x= -pi :0.1 :pi;
y=sin(x);
plot(x,y);
grid
xlabel('x')
ylabel('y')
title('y = f(x)')
```

2. Save the program in a chosen directory.

### A.1.3 Break figure into sub-figures

Write a program which does:

1. Break a figure window in two parts (help *subplot*).
2. Plot the curve $y = sin(2 * pi * 5 * t)$, for $0 \leq t \leq 1$ in the upper part.
3. Plot the curve $y = f(t)$ with $t = 1 : 1000$ and $y = randn(size(t))$ (a vector of a Gaussian random variable of size $t$) in the lower part.

### A.1.4 Statistics analysis of a random variable

Write a program which does:

1. Define a vector, $y$, composed of 50000 samples of a Gaussian random variable (help *randn*).
2. Compute the mean value and the standard deviation of $y$ (help *mean* and *std*) .
3. Now $z = 2 * y + 3$. Compute the mean value and the standard deviation of $z$ and compare with those of $y$. Conclude.
4. Plot the histogram of $y$ (upper part) and $z$ (lower part) (help *hist*) with a number of bars (rectangle) equals 50.

### A.1.5 Complex numbers

Write a program which does:

1. Define $z_1 = 1 + 2i$ and $z_2 = 1 - 3i$.
2. Calculate $a = \mathrm{Re}(z_1)$, $b = \mathrm{Im}(z_1))$, $|z_1|$, $\arg(z_1)$, $\bar{z}_1$, $z_1^2$, $z_1/z_2$ and $z_1 z_2$ (help *real*, *imag*, *abs* and *angle*). Check the results by "hand".

### A.1.6 Polynomial

Matlab represents a polynomial with the help of an array, whose elements are the coefficients of the polynomial in descending powers. The polynomial $P(x) = x^2 - 6x + 4$ is represented, for example, from the array (vector): $P = [1 \ -6 \ 4]$.

1. Compute the value of the polynomial $P$ for $x = 1$ and its roots (help *polyval* and *roots*).

2. Plot the polynomial for $-5 \leq x \leq 5$.

3. Solve the equation $z^4 - 3(1 + 2i)z^2 - 8 + 6i = 0$. Compute the modulus and the phase of the solution.

## A.2 Matrices

### A.2.1 Basic manipulations

1. Write $A = [1, 2; 3, 4]$.

2. Do $A(1, :)$, $A(:, 1)$, $A(:)$. Understand these syntaxes.

3. Extract the first column and the first row of $A$.

4. Do $D = [A; 5, 6]$. What do you observe?

5. Extract the first two columns and the first two rows of $D$.

### A.2.2 Basic calculations

The power of MatLab is to make mathematical computations directly on vectors or matrices.

1. Write $A = [1, 2; 3, 4]$ and $B = [3, 4; 5, 6]$.

2. Compute $A + B$, $A - B$, $A * B$ and $A. * B$ and compare the results with those obtained by "hand" to check. What is the difference between $.*$ and $*$.

3. Inverse the matrix $A$ (help *inv*) and compare the result with that obtained by "hand" to check.

4. Compute the determinant, the trace, the eigen values and the eigen vectors of $A$ and compare the results with those obtained obtained by "hand" to check.

5. Define $P$, the matrix containing the column eigen vectors and $S$ the diagonal matrix of the eigen values. Verify that $[A] = [P][S][P]^{-1}$.

### A.2.3 Linear system

1. We consider the following linear system:

$$\begin{cases} 2x_1 + 3x_2 = 8 \\ x_1 + 2x_2 = -3 \end{cases} .$$

2. From MatLab, represent this system by $[A][x] = [b]$ and solve it from $[x] = [A]^{-1}[b]$.
3. Compare the solution with that obtained by "hand" to check (from the determinant method).

# B Practical work 2: Solve $f(x) = 0$ from the dichotomy method

The bisection method in mathematics is a root-finding method that repeatedly bisects an interval and then selects a subinterval in which a root must lie for further processing. It is a very simple and robust method, but it is also relatively slow. This method is also called the interval halving method, the binary search method, or the dichotomy method.

## B.1 Principle

The method is applicable for numerically solving the equation $f(x) = 0$ for the real variable $x$, where $f$ is a continuous function defined on an interval $[a; b]$ and where $f(a)$ and $f(b)$ have opposite signs. In this case $a$ and $b$ are said to bracket a root since, by the intermediate value theorem, the continuous function $f$ must have at least one root in the interval $[a; b]$ (see figure B.1).

At each step the method divides the interval in two by computing the midpoint $c = (a + b)/2$ of the interval and the value of the function $f(c)$ at that point. Unless $c$ is itself a root (which is very unlikely, but possible) there are now only two possibilities: either $f(a)$ and $f(c)$ have opposite signs and bracket a root, or $f(c)$ and $f(b)$ have opposite signs and bracket a root. The method selects the subinterval that is guaranteed to be a bracket as the new interval to be used in the next step. In this way an interval that contains a zero of $f$ is reduced in width by 50% at each step. The process is continued until the interval is sufficiently small.

Explicitly, if $f(a)$ and $f(c)$ have opposite signs, then the method sets $c$ as the new value for $b$, and if $f(b)$ and $f(c)$ have opposite signs then the method sets $c$ as the new $a$. If $f(c) = 0$ then $c$ may be taken as the solution and the process stops. In both cases, the new $f(a)$ and $f(b)$ have opposite signs, so the method is applicable to this smaller interval.

## B.2 Iteration tasks

The input for the method is a continuous function $f$, an interval $[a; b]$ and the expected precision $\epsilon$. The function values are of opposite sign (there is at least one zero crossing within the interval). Each iteration performs these steps:

1. Calculate $c$, the midpoint of the interval, $c = (a + b)/2$.

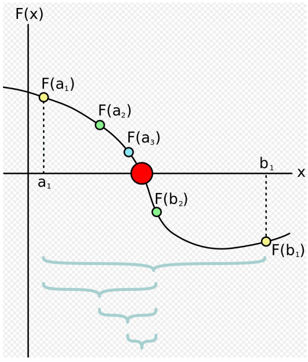



Figure B.1: A few steps of the bisection method applied over the starting range $[a_1; b_1]$ ($a = a_1$, $b = b_1$ and $f = F$). The bigger red dot is the root of the function.

2. Calculate the function value at the midpoint, $f(c)$.
3. If convergence is satisfactory (that is, $|f(c)| < \epsilon$ is sufficiently small), return $c$ and stop iterating.
4. Examine the sign of $f(c)$ and replace either $(a, f(a))$ or $(b, f(b))$ with $(c, f(c))$ so that there is a zero crossing within the new interval.

## B.3 Exercice

To illustrate the method, the function $f$ defined as $f(x) = \sin(x)e^x + x$ is considered in the interval $[1; 3.5]$.

1. Write a function on MatLab named "Function_f.m" allowing us to calculate $f$.
2. Write a function on MatLab named "Function_Dichotomy.m" allowing us to calculate $x_0$ such as $f(x_0) = 0$. Clearly defined the inputs and the outputs are: $x_0$, $f(x_0)$ and $n$ the number of iterations to reach the convergence.
3. In a main program named "Main1_Dichotomy.m", plot the function $f$ on the interval $[1; 3.5]$ and deduce an approximated value of $x_0$. In the same program, compute from the function "Function_Dichotomy.m" an approximated value of $x_0$ for $\epsilon = 10^{-4}$.
4. In a main program named "Main2_Dichotomy.m", plot $n$ versus $\epsilon$ (top), plot $x_0$ versus $\epsilon$ (middle) and plot $f(x_0)$ (bottom) versus $\epsilon$. The figure is broken into three parts and the horizontal axis is in logarithmic scale (help *semilogx*). The values of $\epsilon$ are $\{10^{-2}, 5 \times 10^{-3}, 10^{-3}, 5 \times 10^{-4}, 10^{-4}, 5 \times 10^{-5}, 10^{-5}\}$.

# C Practical work 3: Numerical integration

The trapezoidal rule is one of a family of formulas for numerical integration called Newton-Cotes formulas, of which the midpoint rule is similar to the trapezoid rule. Simpson's rule is another member of the same family, and in general has faster convergence than the trapezoidal rule for functions which are twice continuously differentiable, though not in all specific cases. However for various classes of rougher functions (ones with weaker smoothness conditions), the trapezoidal rule has faster convergence in general than Simpson's rule.

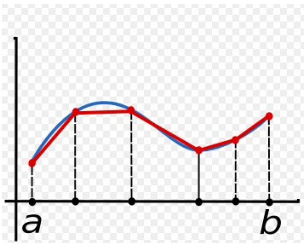



Figure C.1: Illustration of trapezoidal rule used on a sequence of samples (in this case, a non-uniform grid).

## C.1 Principle

For a domain discretized into $N$ spaced panels, or $N+1$ grid points $a = x_1 < x_2 < \ldots < x_{N+1} = b$, the integral of $f$ on the interval $[a;b]$ can be approximated as

$$I = \int_a^b f(x)dx \approx \frac{1}{2} \sum_{n=1}^{n=N} (x_{n+1} - x_n) \left[ f(x_n) + f(x_{n+1}) \right].$$

As shown in figure C.1, the number $(x_{n+1} - x_n) \left[ f(x_{n+1}) + f(x_n) \right] /2$ is the trapezium area assumed to be equal to the area under the curve, above the line $x = 0$ and on the interval $[x_n; x_{n+1}]$.

## C.2 Exercice

To illustrate the method, the function $f$ defined as $f(x) = xe^x$ is considered in the interval $[0; 3]$.

1. Calculate $I = I_{ana}$ analytically. It is the exact value or the reference value.
2. Write a function on MatLab named "Function_f.m" allowing us to calculate $f$.
3. Write a function on MatLab named "Function_Trapezoidal.m" allowing us to calculate $I$ for an equally spacing, that is $h = x_{n+1} - x_n =$ constant. Clearly defined the inputs and the outputs.
4. In a main program named "Main1_Trapezoidal.m", plot the function $f$ on the interval $[0; 3]$. In the same program, compute from the function "Function_Trapezoidal.m" an approximated value of $I$ for $h = 0.05$.
5. In a main program named "Main2_Trapezoidal.m", plot $I$ versus $h$ (top) and plot $|I - I_{ana}|/I_{ana}$ (relative error) versus $h$ (bottom). The figure is broken into two parts and the horizontal axis is in logarithmic scale (help *semilogx*). The values of $h$ are $\{10^{-2}, 5 \times 10^{-3}, 10^{-3}, 5 \times 10^{-4}, 10^{-4}, 5 \times 10^{-5}, 10^{-5}\}$. Conclude.

# D Practical work 4: Ordinary differential equation

Consider the problem of calculating the shape of an unknown curve which starts at a given point and satisfies a given ordinary differential equation (ODE). Here, an ODE can be thought of as a formula by which the slope of the tangent line to the curve can be computed at any point on the curve, once the position of that point has been calculated.

The idea is that while the curve is initially unknown, its starting point, which we denote by $A_0$, is known (see figure D.1). Then, from the ODE, the slope to the curve at $A_0$ can be computed, and so, the tangent line.

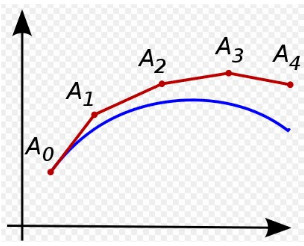



Figure D.1: Illustration of the Euler method. The unknown curve is in blue, and its polygonal approximation is in red.

## D.1 Principle

Take a small step along that tangent line up to a point $A_1$ (see figure D.1). Along this small step, the slope does not change too much, so $A_1$ will be close to the curve. If we pretend that $A_1$ is still on the curve, the same reasoning as for the point $A_0$ above can be used. After several steps, a polygonal curve $A_0$, $A_1$, $A_2$, $A_3$, ... is computed. In general, this curve does not diverge too far from the original unknown curve, and the error between the two curves can be made small if the step size is small enough and the interval of computation is finite:

$$y' = f(x, y), \;\; y(x_0) = y_0.$$

Choose a value $h$ for the size of every step and set $x_n = x_0 + nh$. Now, one step of the more simple method, named Euler, from $x_n$ to $x_{n+1} = x_n + h$ is:

$$y_{n+1} \approx y_n + hy'_n = y_n + hf(x_n, y_n).$$

The value of $y_n$ is an approximation of the solution to the ODE at $x_n$: $y_n \approx y(x_n)$.

## D.2 Exercice

To illustrate different methods, the ODE $y' - y = x$ is considered with $y(0) = 1$ and $x \in [0; 2]$.

1. From the Lagrange method, derive analytically the function $y = y_{ana}$.
2. We want to test the three following methods:

$$\text{Euler: } y_{n+1} = y_n + hf(x_n, y_n).$$

$$\text{Improved Euler: } y_{n+1} = y_n + hf\left(x_n + \frac{h}{2}, y_n + \frac{h}{2}f(x_n, y_n)\right).$$

Runge-Kutta at the order 4: Used the function *ode*45 on MatLab.

3. For the Euler and Improved Euler methods, write 2 functions on MatLab named "Function_Euler.m" and "Function_IEuler.m". Clearly defined the inputs and the outputs.
4. In a main program named "Main1_ODE.m" for $h = 0.05$, plot the function $y$ on the interval $[0; 2]$. In the same program, compute $y$ from the three methods and plot the solutions in the same figure as $y_{ana}$ to compare. Conclude. Compare also their computing time (help *tic* and *toc*).

# E Exam of Mathematics, 20 December 2017, duration 2H00

**The only authorized document is a double-sided A4 paper.**

**The mathematical demonstrations must be rigorous, the copy must be clean (not a draft) and the final result must be underlined.**

## E.1 Integral (5 points)

The function $f$ is defined as $f(x) = \frac{1}{a+b\tan(x)}$, where $(a,b) \in \mathbb{R} \times \mathbb{R}^*$. We want to derive the integral $F(x) = \int f(x)dx$.

1. Give the domain of definition of $f$.
2. Give the domain of definition of $F$ by justifying your response.
3. Calculate $F$. You can set $t = \tan(x)$.

## E.2 Taylor series expansion (3 points)

The function $f$ is defined as $f(x) = \frac{x}{\sin(x)}$. Calculate the series Taylor expansion up to the order 4 near 0.

## E.3 ODE (4 points)

We consider the ordinary differential equation (ODE) defined as $y'(x)x^2 + y(x) = x^2 e^{\frac{1}{x}}$.

1. Give a "name" of the ODE.
2. Calculate $y(x)$ from the constant variation method.

## E.4 Double integral (4 points)

We consider the following integral:

$$I = \iint_{D_{xy}} y^2 e^{-(x^2+y^2)^2} dxdy.$$

where the domain $D_{xy}$ is defined as $D_{xy} = \{x \geq 0, 0 \leq y \leq x\}$.

1. By justifying your response, plot $D_{xy}$ in the $(x, y)$ Cartesian plane.
2. By justifying your response, plot the new domain $D_{r\theta}$ associated to $D_{xy}$ in the $(r, \theta)$ plane, where $(r, \theta)$ are the polar coordinates.
3. Calculate $I$.

## E.5 Eigen values and vectors (4 points)

We consider the matrix $[A]$ defined as

$$[A] = \begin{bmatrix} 1-a & 2a \\ 2a & 1-a \end{bmatrix},$$

where $a > 0$.

1. Calculate the eigen values $(\lambda_1, \lambda_2)$ of $[A]$, where $\lambda_1 < \lambda_2$ (2 points).
2. Calculate the associated eigen vectors $([v_1], [v_2])$.

## E.6 Formula

$$\sin(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} + \mathrm{o}(x^5).$$

$$\frac{1}{1-x} = 1 + x + x^2 + x^3 + x^4 + x^5 + \mathrm{o}(x^5).$$

# Bibliography


[1] Xavier Gourdon, Les maths en tête, Ellipses, 1994.

[2] Louis Gacôgne, *Algébre et analyse cours de mathématiques tome 1*, Eyrolles, 1990.

[3] André Baumy et Michel Bonnaud, *Mathématique pour le physicien tome 1*, McGraw-Hill (Paris), 1989.

[4] Gabriel Soum, Raymond Jagut et Pierre Dubouix, *Techniques mathématiques pour la physique - I*, Hachette, 1995.

[5] Gabriel Soum, Raymond Jagut et Pierre Dubouix, *Techniques mathématiques pour la physique - II*, Hachette, 1995.

[6] Murray R. Spiegel, *Analyse vectorielle cours et problème*, McGraw-Hill (New-York), douzième édition, 1973.

[7] Murray R. Spiegel, *Analyse cours et problème*, McGraw-Hill (New-York), dix-septième édition, 1973.

[8] Frank Ayres, *Equations différentielles cours et problème*, McGraw-Hill (New-York), quinzième édition, 1972.

[9] G. Hirsch et G. Eguether, *Fonctions de plusieurs variables*, Masson, 1994.

[10] K. Arbenz et A. Wohlhauser, *Complèments d'analyse*, Presses polytechniques et universitaires romandes, 1993.

[11] M. Chossat, *Aide Mémoire de Mathématiques de l'ingénieur*, Dunod, 1996.

[12] Marie-Pascale Avignon et Jacques Rogniaux, *Analyse 369 exercices corrigés*, Ellipses, 1991.

[13] Seymour Lipschutz et Marc Lipson, *Algèbre linéaire - 3ème édition*, Dunod, 2003.

[14] Seymour Lipschutz, *Algèbre linéaire, Cours et problèmes - 2ème édition*, McGraw-Hill (New-York), 1994.

[15]
- https://en.wikipedia.org/wiki/Limit_of_a_function
- https://en.wikipedia.org/wiki/Derivative
- https://en.wikipedia.org/wiki/Integral
- https://en.wikipedia.org/wiki/Differential_equation
- https://en.wikipedia.org/wiki/Complex_number
- https://en.wikipedia.org/wiki/Partial_derivative

- https://en.wikipedia.org/wiki/Second_partial_derivative_test
- https://en.wikipedia.org/wiki/Multiple_integral
- https://en.wikipedia.org/wiki/Vector_calculus
- https://en.wikipedia.org/wiki/Gradient
- https://en.wikipedia.org/wiki/Divergence_theorem
- https://en.wikipedia.org/wiki/Laplace_operator
- https://en.wikipedia.org/wiki/Matrix_(mathematics)
- https://en.wikipedia.org/wiki/Determinant
- https://en.wikipedia.org/wiki/Eigenvalues_and_eigenvectors
- https://en.wikipedia.org/wiki/Bisection_method
- https://en.wikipedia.org/wiki/Trapezoidal_rule
- https://en.wikipedia.org/wiki/Euler_method
- https://en.wikipedia.org/wiki/Hermitian_matrix