UNIVERSIDAD DE SANTIAGO DE CHILE

CIRAS

Mathematical reminders

Ileyk EL MELLAH

August 2024

Introduction

Contents

Chapter 1

Notation and lexicon

My notations are largely personal and I do not claim that they are fully consistent or even the only way to properly describe the mathematical and physical notions, and the different steps of a reasoning. Instead, the role of this section is to provide you with a reference you can check in case of doubt on the meaning of what I wrote, and to invite you in designing your own compact, coherent and clean linguistic and conceptual shortcuts. Doing so will help you a lot in memorizing information and activating it at the right moment in the right place.

1.1 Greek alphabet

While taking notes during classes, you will be facing a practical though unnecessary hindrance: Greek alphabet. Most of you probably never studied Greek so it is somewhat a source of distraction when you stumble into one of these letters. In order to be able to take fast and seamless notes, you should practice identifying Greek letters and writing them in a clear and intelligible form. Fig. 1.1 provides a convenient basis to do so.

1.2 Mathematical jargon

What we call "jargon" is the unofficial though widespread expressions used in a field. It is both a shortcut which speeds up the exchange and facilitates the memorization, and a social construct which fosters the bounds within a group[1]. Let us start with a representative piece of mathematical jargon, a convenient formulation which saves space. The resp. notation draws a parallel between two statements. For instance, instead of saying:

"The electric field is a true vector and the magnetic field is a pseudo vector."

I will rather write:

"The electric (resp. magnetic) field is a true (resp. pseudo) vector."

Also, we commonly use the following notations:

- **e.g.** - short for "for example", used at the beginning of a parenthesis. For instance: "Members of the ruling class defend their own interests first and foremost (e.g. the Macayas, the Coxes, the Freis, the Lagos, the Alessandris, the Piñeras and the Aylwins)".
- **i.e.** - short for "that is to say", used at the beginning of a parenthesis to provide an equivalent statement. For instance, "Oppressive policies are prevalent in colonial societies (i.e. those where the power is hold by an allochthonous group who bases its legitimacy on a dismissive representation of the historical inhabitants)".

[1] To some extent, the language itself is a jargon.

Alpha Α α	Beta Β β	Gamma Γ γ	Delta Δ δ
Epsilon Ε ε	Zeta Ζ ζ	Eta Η η	Theta Θ θ
Iota Ι ι	Kappa Κ κ	Lambda Λ λ	Mu Μ μ
Nu Ν ν	Xi Ξ ξ	Omicron Ο ο	Pi Π π
Rho Ρ ρ	Sigma Σ σ	Tau Τ τ	Upsilon Υ υ
Phi Φ φ	Chi Χ χ	Psi Ψ ψ	Omega Ω ω

FIGURE 1.1: Greek letters and their names. Their cursive writing is shown in this video.

- **a.k.a.** - short for "also known as", used at the beginning of a parenthesis to provide an alternative denomination. For instance, "the first Gulf War (a.k.a. Operation Desert Storm) was an attempt from the US to secure their access to cheap oil, at the expense of regional stability for the decades to come".
- **N.B.** - short for "nota bene", which can be interpreted as "side note", that I use when I need to add a minor comment. For instance, "N.B.: Views and opinions expressed in this textbook are those of the author only and do not necessarily reflect those of their employer".

To re-inject an expression $A = B$ in an equation $C = A$ means replacing the term A with the term B in the latter expression.

In order to simplify the calculation, I will often, in the middle of a reasoning, make an assumption which does not affect the conclusions in the sense that the complementary hypothesis would have led to very similar steps: I will use expressions such as "without loss of generality" to indicate such a procedure. For instance, let us assume that we work on a 3D system perfectly symmetric with respect to the plane $z = 0$ in a Cartesian frame. Then, we can focus on the half-space $z > 0$ without loss of generality, in the sense that the computation would be totally equivalent if we were focusing on the half-space $z < 0$.

A back-of-the-envelope calculation is a calculation which is not fully rigorous and only provides a preliminary and approximate evaluation of the result. It can serve to later guide a more rigorous reasoning.

An equality sign separates two sides, the left hand side and the right hand side that we will regularly refer to.

Commas matter: "let's eat, kids" does not mean the same thing as "let's eat kids". More generally, commas serve to guide the reading and structure the information, like parenthesis in algebra and in logic (section 1.3).

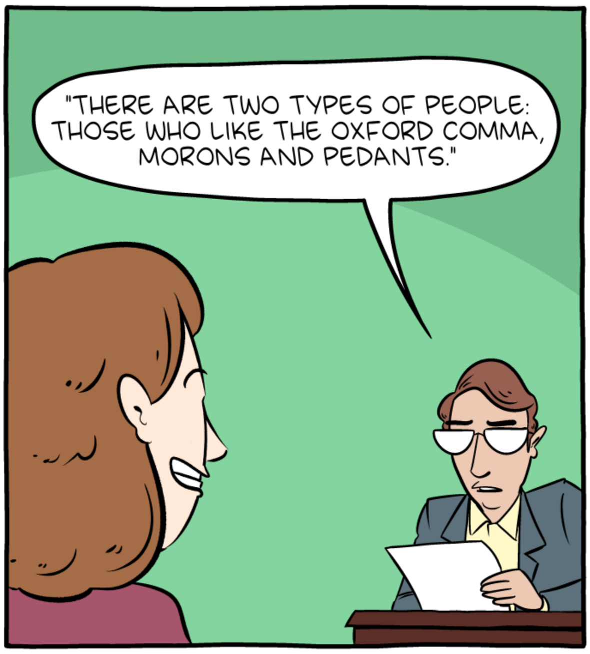

FIGURE 1.2: From Saturday Morning Breakfast Cereal.

> The Oxford comma is a comma placed immediately after the penultimate term and before the coordinating conjunction (and or or) in a series of three or more terms. For instance, the comma in bold red font in "Marraquetas, hallullas**,** and sopaipillas" is an Oxford comma. I don't use it but it is common in scientific literature.

Sometimes, I will use the notation $\bullet$ to mention a fiducial element whose nature does not rely matter. For instance, if I say "the divergence operator $\boldsymbol{\nabla} \cdot \bullet$", I want to emphasize on the operator itself while downplaying the importance of its argument (the bullet). It would have been equivalent, though more cumbersome and less prone to put the operator in the spotlight, to write "the divergence operator $\boldsymbol{\nabla} \cdot \mathbf{u}$, where $\mathbf{u}$ is a vector field".

1.3 Logic

To define terms or introduce a new notation, I will use the $\hat{=}$ symbol. For instance, the velocity vector $\mathbf{v}$ is defined as the time derivative of the position vector $\mathbf{r}$:

$$\mathbf{v} \hat{=} \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t} \tag{1.1}$$

where t is the time.

Sometimes, I will use the $\coloneqq$ symbol in order to emphasize the link of causality between the two sides of the equality sign. For instance, Newton's second law can be written as:

$$\mathbf{a} \coloneqq \frac{\mathbf{F}}{m} \tag{1.2}$$

where $\mathbf{a}$ is the acceleration vector, $\mathbf{F}$ is the force and m is the inertial mass. Here, I insist on the fact that the acceleration (i.e. what is on the left hand side) is produced by the force exerted on the mass m (the right hand side). The right hand side is the cause and the left hand side is the consequence. In computer science, this notation is particularly important since it specifies which variable is assigned which value. For instance, the code:

```
a=1
b=2
b=a
c=a+b
print(c)
```

provides the answer `2` while the code:

```
a=1
b=2
a=b
c=a+b
print(c)
```

provides the answer `4`. The only lines which differ are in bold font. Contrary to the mathematical equality $=$ operator, in computer science, the equality $=$ symbol is not comutative.

I will sometimes use the equivalence symbol $\Longleftrightarrow$ or the statement "if and only if" to define a notion (rather than to establish an equivalence between two statements). In this case, the logical operator somewhat looses its commutativity.

1.4 Algebra

Most of the time, when I write the product between two numbers a and b, I will omit the multiplication symbol $\cdot$ (a.k.a. $\times$) and simply write ab. It can produce confusion when combined with parenthesis. For instance, it can be ambiguous whether the equation $a(b+1)$ is the product of a by $b+1$, or whether a is a function of the variable $b+1$. The context generally enables you to make the difference but to avoid confusion, in some cases, I will write the multiplication symbol $\cdot$, for instance in equation (3.68).

The opposite of a number a is a number b such as $a+b=0$. It is written $-a$. Provided $a \neq 0$, the inverse of a number a is a number b such as $a \cdot b = 0$. It is written $1/a$ (or a^{-1}).

The exponent above a trigonometric function (or a logarithm) means that the function is raised to this exponent. For instance, $\cos^2 x = (\cos x) \cdot (\cos x)$.

The $\pm$ and $\mp$ symbols in an equation mean that we wrote two equations in one: one where the signs are the upper ones, one where the signs are the lower ones.

I write the minimum and maximum of a variable x as x_m and x_M respectively.

I will seldom use the symbols $\geqslant$ (for "greater than or equal to") and $\leqslant$ ("less than or equal to") because most of the time in Physics, a strict equality between two measurable quantities is pointless due to unavoidable perturbations. Rather, I will use $>$ ("greater than") and $<$ ("smaller than").

In Astrophysics, we pay a specific attention to the scaling of variables with respect to each other: how does the variation of X affect $Y(X)$? A simple way to represent this is through the proportional $\propto$ symbol. For instance, Y can be proportional to X through a power-law (see exercises below for other examples):

$$Y \propto X^{\alpha} \quad \text{with} \quad \alpha \text{ a constant exponent} \tag{1.3}$$

it means that Y is proportional to X^{α}. If X doubles for instance, Y will be multiplied by 2^{α}. It is less informative than the full equality relation since we do not know the value of the proportionality constant[2] K such as:

$$Y = KX^{\alpha} \tag{1.4}$$

[2] K does not have to be a constant, it can simply be an independent variable which does not depend on X and Y.

but K can be found if Y and X are known for a specific case. Indeed, let us assume that we have the point of reference (X_0, Y_0). Then, we have:

$$Y_0 = K X_0^\alpha \tag{1.5}$$

$$K = \frac{Y_0}{X_0^\alpha} \tag{1.6}$$

that we can re-inject in equation (1.4) to obtain:

$$Y = \left(\frac{Y_0}{X_0^\alpha}\right) X^\alpha \tag{1.7}$$

$$Y = Y_0 \left(\frac{X}{X_0}\right)^\alpha \tag{1.8}$$

> Graphically, a proportionality relation gives the shape of the curve $Y(X)$. The scale is provided by the constant K. For instance, the power-law (1.3) gives the slope of the straight line $Y(X)$ in a log-log diagram (see section 1.7). Then, a single point of reference (X_0, Y_0) is enough to entirely determine the curve.

1.4.1 Scaling relations

1.5 Differential calculus

1.6 Vector calculus

I write vectors in bold font, something you cannot do during the in-class exams (and probably not when you take notes). Please, use left-to-right arrows above vectors to make the difference with scalars.

The notation $\hat{\mathbf{u}}$ means that we are working with a vector of magnitude unity (a.k.a. a unit vector). We can construct a unit vector from any vector $\mathbf{u}$ via:

$$\hat{\mathbf{u}} = \frac{\mathbf{u}}{|\mathbf{u}|} \tag{1.9}$$

In particular, when we work with a basis $(\hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}})$ in Cartesian coordinates, $(\hat{\mathbf{r}}, \hat{\boldsymbol{\theta}}, \hat{\mathbf{z}})$ in cylindrical coordinates or $(\hat{\mathbf{r}}, \hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\phi}})$ in spherical coordinates, the hat indicates that this is the unit vector in the direction of the axis corresponding to the coordinates. For instance, $\hat{\mathbf{z}}$ is the unit vector carried by the axis along which only the z coordinate varies.

> An important point to keep in mind is that the $\hat{\mathbf{r}}$ and $\hat{\boldsymbol{\theta}}$ axis in cylindrical coordinates, and the three axis in spherical coordinates have a different behavior than the other axis: they are associated to the point we study and their orientation changes with the coordinates of the point. It is the reason why their time derivative is non-null (see section 4.4).

I use the same symbol $\cdot$ for the dot product between two vectors $(\mathbf{u} \cdot \mathbf{v})$ as for the product between two scalars $(a \cdot b)$, except that I generally omit the latter. When $\boldsymbol{\nabla}$ is immediately followed by a $\cdot$, it refers to the divergence operator. I use the $\wedge$ symbol for vector product (while others sometimes use $\times$).

The decomposition of a vector $\mathbf{u}$ in an orthonormal base $(\hat{\mathbf{a}}, \hat{\mathbf{b}}, \hat{\mathbf{c}})$ is given by:

$$\mathbf{u} = u_a\, \hat{\mathbf{a}} + u_b\, \hat{\mathbf{b}} + u_c\, \hat{\mathbf{c}} \tag{1.10}$$

where $u_i \hat{=} \mathbf{u} \cdot \hat{\mathbf{i}}$, with $i = a$, b or c.

The $\perp$ (resp. $/\!\!/$) symbol between two vectors means that they are orthogonal (resp. colinear). The adjective normal, orthogonal and perpendicular are somewhat synonym, although perpendicular should be reserved to straight lines while the two first ones can be used for vectors and straight lines.

1.7 Representations

1.8 Physics

In this section, I emphasize on the notations and lexicons which are more specific to physical aspects.

A scalar is a number, with or without dimension.

When we refer to the velocity, we generally refer to a vector. On the contrary, the speed is the magnitude of the vector and its dimension are length unit per time unit (e.g. cm$\cdot$s^{-1} in CGS).

1.8.1 Dimensional analysis

A necessary (but not sufficient) condition for a physical equation to be meaningful is to be dimensionally homogeneous, that is to say that both sides of the equality sign must have the same dimension (e.g. a length $[L]$, a mass $[M]$ or a time $[T]$): apples and oranges cannot be compared. The most basic strategy to verify the dimensional homogeneity of a function is to decompose each quantity in its dimensions. For instance, let us consider the escape speed v_{esc} of a test-mass located at a distance R from the center of a spherically-symmetric mass distribution of total mass M:

$$v_{\text{esc}} = \sqrt{\frac{2GM}{R}} \tag{1.11}$$

where $G \sim 6.67\text{cm}^3\cdot\text{g}^{-1}\cdot\text{s}^{-2}$ is the gravitational constant. Then, the dimensional analysis gives:

$$\begin{aligned}
\frac{[L]}{[T]} &= \sqrt{\frac{[L]^3[M]^{-1}[T]^{-2}[M]}{[L]}} && (1.12)\\
&= \sqrt{[L]^2[T]^{-2}} && (1.13)\\
&= \frac{[L]}{[T]} && (1.14)
\end{aligned}$$

so this equation is dimensionally homogeneous. However, this approach is lengthy, cumbersome and prone to mistakes. For instance, it requires that you know the dimension of the gravitational constant G, a difficulty which can be alleviated. A safer strategy is to recognize intermediate equations you are familiar with. For instance, you know that GM^2/R^2 is homogeneous to a force, and that Newton's second law implies that this force is homogeneous to the product of a mass m by an acceleration $\mathbf{a}$. Therefore, we can raise equation (1.11) to the square and multiply the result by a mass divided by a length to obtain something homogeneous to $m\,\mathbf{a}$ on the left hand side and try to recover the force on the right hand side:

$$v_{\mathrm{esc}}^2 = \frac{2GM}{R} \tag{1.15}$$

$$v_{\mathrm{esc}}^2 \frac{M}{R} = \frac{2GM^2}{R^2} \tag{1.16}$$

which shows that this equation is dimensionally homogeneous.

> The dimension of a variable, not to be confused with the dimension of a vector (2D, 3D...), is either a length, a mass, an energy or a time for instance. It is different from units, which are either centimeters, grams, ergs or seconds for instance, but could also be foot, pounds, joules or weeks.

Strictly speaking, the argument of a function (be it an exponential, a log, a trigonometric or a hyperbolic function) has to be dimensionless. Power-laws are the only one which can apply to a variable with a physical dimension. For instance, the argument of the square-root function in equation (1.11) has the dimension of a velocity squared. However, if we write $\sin(2\pi\omega t)$, necessarily, $2\pi\omega t$ is dimensionless.

> Angles (resp. solid angles) have no dimension but they are expressed in radians (resp. steradians). Therefore, an angular speed defined by:
>
> $$\omega = \frac{\mathrm{d}\theta}{\mathrm{d}t} \tag{1.17}$$
>
> is homogeneous to the inverse of a time (with θ an angle and t the time).

An apparent exception to the "dimensionless arguments" rule you will commonly encounter is when a logarithm is used. For instance:

$$\frac{\mathrm{d}x}{x} = \omega\,\mathrm{d}t \tag{1.18}$$

with ω a constant angular speed, t the time and x the position. The left hand side and right hand side are both dimensionless. If we integrate (section 3.2.4) this equation from time t_1 to time t_2, with the boundary conditions $x_1 = x(t = t_1)$ and $x_2 = x(t = t_2)$, we obtain:

$$\int_{x_1}^{x_2} \frac{\mathrm{d}x}{x} = \int_{t_1}^{t_2} \omega\,\mathrm{d}t \tag{1.19}$$

$$\int_{x_1}^{x_2} \mathrm{d}\left(\ln x\right) = \int_{t_1}^{t_2} \mathrm{d}\left(\omega t\right) \tag{1.20}$$

$$\ln x_2 - \ln x_1 = \omega\left(t_2 - t_1\right) \tag{1.21}$$

In the last equation, the arguments of the logarithms are homogeneous to a length, but it is tolerated to write such a thing because we can immediately use the logarithmic fundamental property to reformulate it into:

$$\ln\left(\frac{x_2}{x_1}\right) = \omega\left(t_2 - t_1\right) \tag{1.22}$$

where this time, the argument of the logarithm is dimensionless.

1.8.2 Systems of units

The laws of physics do not depend on the units we use to measure quantities: the fall of an apple is not modified if we measure its altitude in meters or in foot.

The reference system of units in Physics is the International System of Units (a.k.a. SI, or MKS for "Meter, Kilogram, Second"). It defines base units from which other units can be deduced. The main ones to be known are meter (m), kilogram (kg) and second (s). In MKS, energies are in Joule (J).

Yet, in Astrophysics, we commonly use another system of units called Gaussian (a.k.a. CGS, for "Centimeter, Gram, Second"). In CGS, energies are in erg (erg). Both in MKS and CGS, we measure temperature in Kelvin (K).

> Particle physicists often use a third system of units, the Heaviside-Lorentz one.

Lhe laws of mechanics (e.g. Newton's second law) look the same in both system of units, but not the laws of electromagnetism (e.g. Maxwell laws). The latter look somewhat simpler in CGS because we measure electric charges in statcoulomb (statC) instead of coulomb (C) in MKS.

Sometimes, when the unit of a quantity is not relevant, I will simply write [MKS] or [CGS] to indicate that the value given is in MKS or CGS. This is not good practice, please try to avoid doing so, except if it is really an intermediate quantity whose exact unit does not matter. In appendix XXX, I give the value of common constants and quantities in CGS.

Chapter 2
Algebra

Algebra, a branch of mathematics somewhat more abstract than analysis, is the manipulation of operations between elements of ensembles (e.g. the real numbers $\mathbb{R}$).

Linear algebra only.

2.1 Basic reminders

2.1.1 Manipulation of additions and multiplications

2.1.2 System of equations

2.2 Complex numbers

de Moivre's formula (find it on the campus)

$$sin(n\theta)$$

2.3 Polynomials

2.3.1 Definition and properties

2.3.2 Second-order polynomials

$$x^4 + x^2 + 1 = 0$$

2.3.3 Classic polynomials

Legendre polynomials

Chebyshev polynomials

2.4 Algebraic fractions

2.4.1 Definition

2.4.2 Partial fraction decomposition

2.5 Matrices

Chapter 3

Analysis

In this section, I give basic reminders of analysis regarding the definition of a function, the notations and the basic concepts, but I elude many important properties, theorems and pieces of information you need to know about but that you already saw in first year: for instance, L'Hôpital's rule, continuous and bounded functions, integration by parts, and the main analytic functions. Instead, here, I focus on the aspects directly related to differential calculus, the manipulation of infinitesimal quantities.

3.1 Functions

3.1.1 Definition

A function is a transformation defined on a certain set X and which assigns to each element x of X an element $f(x)$:

$$\begin{aligned} f : X &\to Y \\ x &\mapsto f(x) \end{aligned}$$

where Y is another set called the codomain (a.k.a. set of destination) of f.

> A function necessarily maps an element x on a single element $f(x)$ of Y. For instance, the curve $\mathcal{C}$ represented below in Fig 3.1 cannot be associated to a function (hence the interest of the notion of parametric equation).

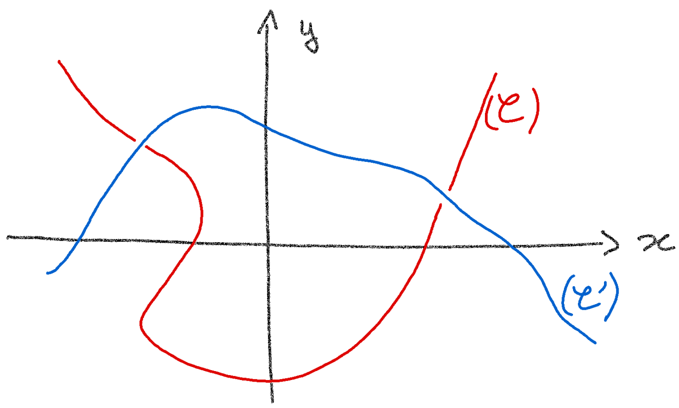

FIGURE 3.1: Two curves are represented but only one corresponds to a function.

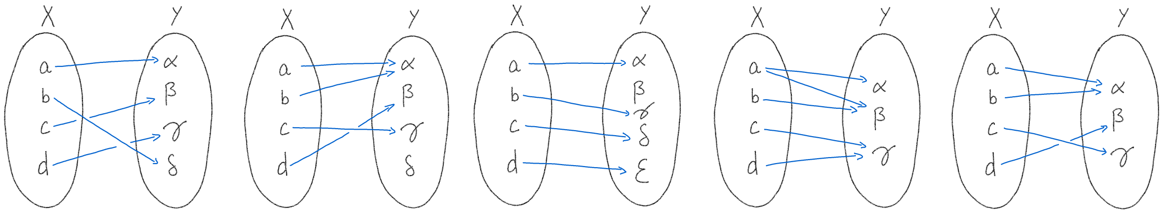

FIGURE 3.2: Mapping diagrams.

Mapping diagram

An abstract way to represent a function is through mapping diagrams which show how elements of a domain X map on the codomain Y. We can interpret X as the inputs of the function and Y as the outputs of the function. Among the mapping diagrams in Fig. 3.2, which of them cannot correspond to a function?

> This representation and the function composition operator $\circ$ defined in equation (3.3) are largely used in the functional programming paradigm.

Functions can be added or multiplied like real numbers, and when applied to functions, these operations share the same mathematical properties as when applied to real numbers. In particular, they are commutative. Indeed, let f and g be two functions[1], we have:

$$(f+g)(x) = f(x) + g(x) = g(x) + f(x) = (g+f)(x) \tag{3.1}$$
$$(fg)(x) = f(x)g(x) = g(x)f(x) = (gf)(x) \tag{3.2}$$

Functions can also be combined together through the $\circ$ operator:

$$(f \circ g)(x) = f(g(x)) \tag{3.3}$$

Is the $\circ$ operator commutative?

Usual functions

The mathematical properties, the shape and the limits of the following family of functions should be know by heart: the exponential functions a^x, the logarithmic functions $a \ln x$, the power laws x^a, the 3 trigonometric functions (cosine $\cos$, sine $\sin$ and tangent $\tan$) and their inverse (arc cosine $\arccos$, arc sine $\arcsin$ and arc tangent $\arctan$) and the 3 hyperbolic functions (hyperbolic cosine $\cosh$, hyperbolic sine $\sinh$ and hyperbolic tangent $\tanh$) and their inverse (area hyperbolic cosine arccosh, area hyperbolic sine arcsinh and area hyperbolic tangent arctanh).

A function $f\colon X \to Y$ is injective if and only if it maps different elements of X to different elements of Y. More rigorously:

$$f\colon X \to Y \text{ injective} \iff (\forall x_1, x_2 \in X \quad x_1 \neq x_2 \implies f(x_1) \neq f(x_2)) \tag{3.4}$$

[1] In this chapter, I bypass all the details relative to the domains of definition. For instance, here, these equalities only make sense on a domain X where both f and g are defined.

A function $f\colon X \to Y$ is surjective if and only all elements of Y are mapped by at least one element of X. More rigorously:

$$f\colon X \to Y \text{ surjective} \iff \forall y \in Y \quad \exists x \in X \quad y = f(x) \tag{3.5}$$

The function f is automatically surjective provided we define Y as the image of X:

$$Y = \{\forall x \in X, f(x)\} \tag{3.6}$$

Colloquially, we write $Y = f(X)$.

1. Reformulate the definition of an injective function using the contraposition of the last part of the definition.

2. Provide the mathematically formal definition of a non-injective function and a non-surjective function.

A function $f\colon X \to Y$ is bijective if and only all elements of Y are mapped by a single element of X. More rigorously:

$$f\colon X \to Y \text{ bijective} \iff \forall y \in Y \quad \exists! x \in X \quad y = f(x) \tag{3.7}$$

Bijective = injective + surjective

Show that a function is bijective if and only if it is both injective and surjective.

1. Are the functions below injective? surjective? bijective?

$$\begin{aligned} f\colon \mathbb{R} &\to \mathbb{R} \\ x &\mapsto f(x) = x^2 \end{aligned} \tag{3.8}$$

$$\begin{aligned} f\colon \mathbb{R} &\to \mathbb{R}^+ \\ x &\mapsto f(x) = x^2 \end{aligned} \tag{3.9}$$

$$\begin{aligned} f\colon \mathbb{R}^+ &\to \mathbb{R}^+ \\ x &\mapsto f(x) = x^2 \end{aligned} \tag{3.10}$$

$$\begin{aligned} f\colon \mathbb{R}^+ &\to \mathbb{R} \\ x &\mapsto f(x) = \sqrt{x} \end{aligned} \tag{3.11}$$

$$\begin{aligned} f\colon \mathbb{R}^+ &\to \mathbb{R}^+ \\ x &\mapsto f(x) = \sqrt{x} \end{aligned} \tag{3.12}$$

$$\begin{aligned} f\colon [0;42] &\to \mathbb{R} \\ x &\mapsto f(x) = e^x \end{aligned} \tag{3.13}$$

$$\begin{aligned} f\colon \mathbb{R}^{+,*} &\to \mathbb{R} \\ x &\mapsto f(x) = \ln x \end{aligned} \tag{3.14}$$

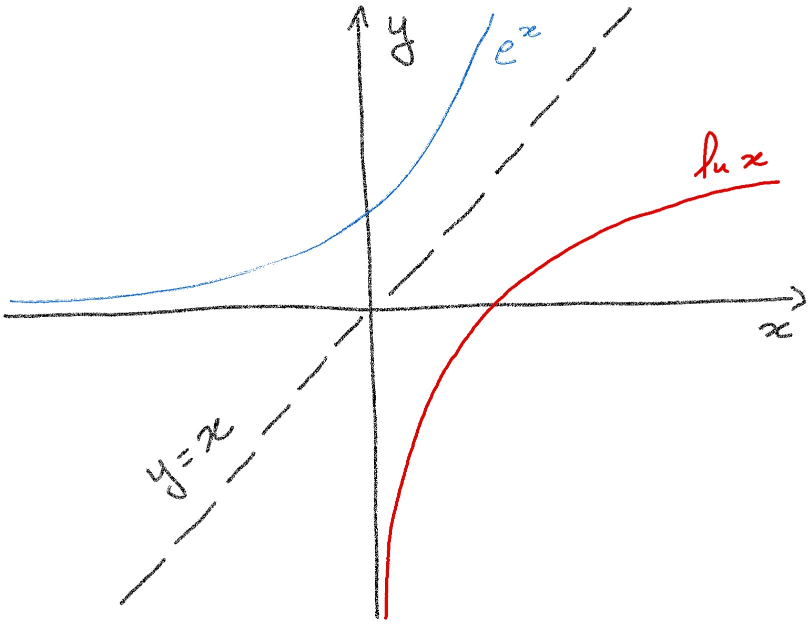

FIGURE 3.3: The exponential and logarithmic functions are inverse: their curves are symmetric with respect to the axis $y = x$.

$$\begin{aligned} f\colon \mathbb{R}^* &\to \mathbb{R} \\ x &\mapsto f(x) = 1/x \end{aligned} \tag{3.15}$$

$$\begin{aligned} f\colon \mathbb{R}^* &\to \mathbb{R}^* \\ x &\mapsto f(x) = 1/x \end{aligned} \tag{3.16}$$

$$\begin{aligned} f\colon \mathbb{R} &\to [-2; +\infty] \\ x &\mapsto f(x) = x^2 - 4x + 2 \end{aligned} \tag{3.17}$$

2. In the mapping diagrams in Fig. 3.2, which are the functions injective? surjective? bijective?

Let $f\colon X \to Y$ be an injective function. Then, we can define the inverse f^{-1} of the function f as the function:

$$\begin{aligned} f^{-1}\colon f(X) &\to X \\ x &\mapsto f^{-1}(x) \end{aligned}$$

such as:

$$\forall x \in X \quad (f^{-1} \circ f)(x) = x \tag{3.18}$$

It means that if we write $y = f(x)$ and that f is injective, we can either see y as a function of x or, equivalently, x as a function of y. When working on stellar interiors, it is what enabled us to work either with radial or mass-coordinates.

> Graphically, the inverse of a function is obtained by performing a planar symmetry with respect to the straight line $y = x$.

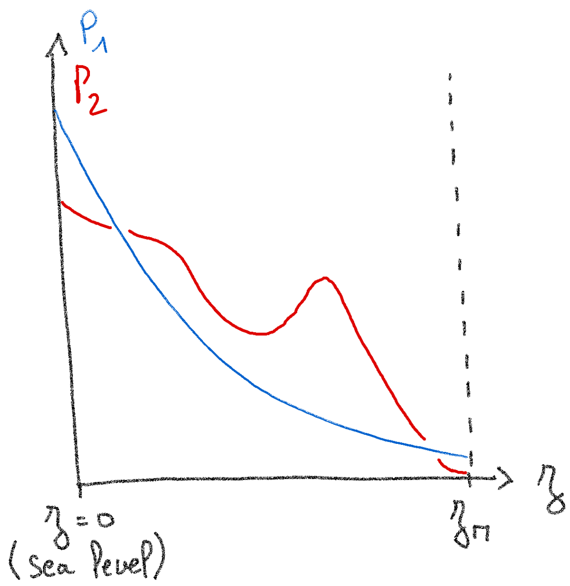

FIGURE 3.4: Pressure profiles as a function of the altitude z.

1. Convince yourself that if a function is not injective, we cannot define its inverse.
2. Fig. 3.4 represent two different pressure profiles $P_1(z)$ and $P_2(z)$ in an atmosphere, between the altitudes $z = 0$ and $z = z_M$. Can we substitute to the z-coordinate the pressure P as a coordinate in the case of P_1? in case P_2?

An involution is a function $f\colon X \to X$ that is its own inverse, that is to say when f is applied twice, it brings one back to the starting point:

$$f\colon X \to X \quad \text{involution} \iff (f \circ f)(x) \hat{=} f(f(x)) = x \tag{3.19}$$

where $\circ$ is the composition symbol.

Show that the function f of x defined by:

$$f(x) = \frac{a - x}{1 + bx} \tag{3.20}$$

with a and b two constants, is an involution. Specify the condition on the product ab for this result to be valid.

Wave optics

The Fourier transform is an involution: the Fourier transform of the Fourier transform of a function is the original function itself (apart from a dimensionless factor). You can visualize it in Wave Optics, where Huygens–Fresnel principle leads to the conclusion that light diffraction is the outcome of the superposition of multiple waves. Let an aperture be the source of diffraction. We can introduce the concept of transparency function f which models the shape of the aperture and expresses how easily light passes through it. For instance, for a hole of diameter R in an opaque cardboard, it

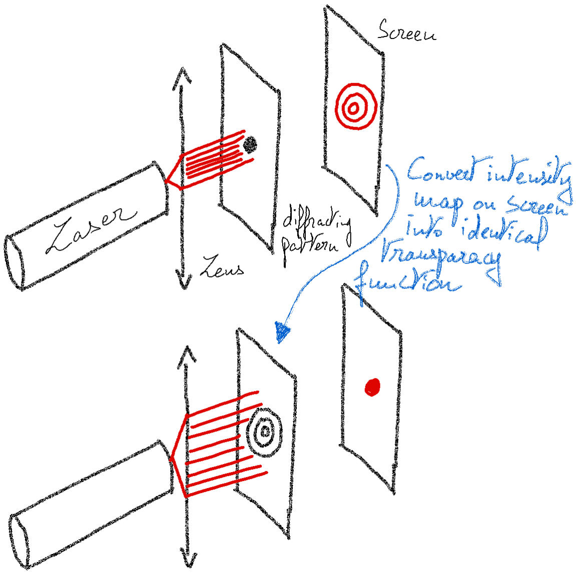

FIGURE 3.5: Sketch of a diffraction experiment through a circular hole (upper panel) and through the transparency function corresponding to the diffraction pattern obtained on the screen in the upper panel (bottom panel).

could be modeled as a function of r only, the distance to the hole's center:

$$f: r \mapsto f(r) = H_R(r) \tag{3.21}$$

where $H_R(r)$ is the Heaviside step function defined by:

$$H_R: r \mapsto H_R(r) = \begin{cases} 0 \text{ if } r > R \\ 1 \text{ if } r < R \end{cases} \tag{3.22}$$

It means that the obstacle fully lets the light through (resp. fully block the light) within the hole (resp. outside the hole).

> Strictly speaking, the transparency function is also a function of the wavelength. Indeed, the glass of a window is transparent in optical, but not in UV nor in IR (see Fig.3 top in the appendix Spectroscopy of the Stellar Astrophysics textbook for the electromagnetic spectrum and the Earth atmosphere transparency function).

Huygens-Fresnel principle applied to an aperture says that the diffraction pattern is the spatial Fourier transform of the transparency function f (in the Fraunhofer far-field approximation). Since the Fourier transform is an involution, it means that if you use the diffraction pattern itself as a transparency function, and perform diffraction through this transparency function, you retrieve the initial transparency function f (Fig. 3.5).

> A convenient way to encrypt information...

A function $f: x \mapsto f(x)$ is even (resp. odd) if and only if $\forall x \in X \quad f(-x) = f(x)$ (resp. $f(-x) = -f(x)$). The graph of an even (resp. odd) function presents a plane symmetry with respect to the ordinate $\hat{\mathbf{y}}$-axis (resp. a point symmetry with respect to the origin

In most functions you have been working with, the domain X was the domain $\mathbb{R}$, the domain of all real numbers, or a subdomain of $\mathbb{R}$ (e.g. $\mathbb{R}^+$). However, X is not necessarily one-dimensional. Instead, it could be $\mathbb{R}^n$ with n a positive integer. In this case, the variable x is actually a vector $\mathbf{x}$ of size n and the function f is a function of n variables. For instance, the 2D scalar field in Fig.**??** is a function $\mathbb{R}^2 \to \mathbb{R}$. In this case, we can also subdivide the vector $\mathbf{x}$ in its components $(x_1, x_2, ..., x_N)$ and write:

$$f\colon \mathbb{R}^n \to \mathbb{R}$$
$$\mathbf{x} = (x_1, x_2, ..., x_N) \mapsto f(x_1, x_2, ..., x_N)$$

Typically, N will be either 2 or 3 when we work with vectors such as the location $\mathbf{r}$, the velocity $\mathbf{v}$, the acceleration $\mathbf{a}$, the angular momentum $\mathbf{L}$, etc. However, when you work in phase space (e.g. in non-linear physics or in statistical mechanics), the dimensionality of the inputs is typically higher than 3. Furthermore, the output domain Y can also be of dimension $n > 1$, which is the case of vector fields for instance (section 4.6.2). Hereafter, unless specified otherwise, we work with $f\colon \mathbb{R} \to \mathbb{R}$.

3.1.2 Limit

The limit of a function $f\colon X \to Y$ in x_0 describes the behavior of that function near a point x_0 which might be finite or not. It is written[2]:

$$\lim_{x \longrightarrow x_0} f(x) = L \tag{3.23}$$

where L can be finite or not. An alternative and somewhat lighter notation is:

$$f(x) \underset{x \longrightarrow x_0}{\longrightarrow} L \tag{3.24}$$

We also say that "the function f tends towards L when x tends towards x_0. The formal definition of the notion of limit are the following ones:

- If x_0 is $+\infty$ but L is finite:

$$\lim_{x \longrightarrow +\infty} f = L \iff \forall \epsilon > 0 \quad \exists b > 0 \quad \forall x \in X \; (x > b \implies |f(x) - L| < \epsilon) \tag{3.25}$$

  It means that however close from the limit L you want f to be, there is always a b such as beyond this value, f is indeed as close from L as you wanted (Fig. **??**).

  > The formal definition in $-\infty$ is analogous. Write it.

- If x_0 is finite but L is $+\infty$:

$$\lim_{x \longrightarrow x_0} f = +\infty \iff \forall a > 0 \quad \exists \delta > 0 \quad \forall x \in X \; (0 < |x - x_0| < \delta \implies f(x) > a) \tag{3.26}$$

  > The formal definition towards $-\infty$ is analogous. Write it.

- If both x_0 and L are $+\infty$:

$$\lim_{x \longrightarrow +\infty} f = +\infty \iff \forall a > 0 \quad \exists b > 0 \quad \forall x \in X \; (x > b \implies f(x) > a) \tag{3.27}$$

[2] Hereafter, we omit the (x) in $f(x)$.

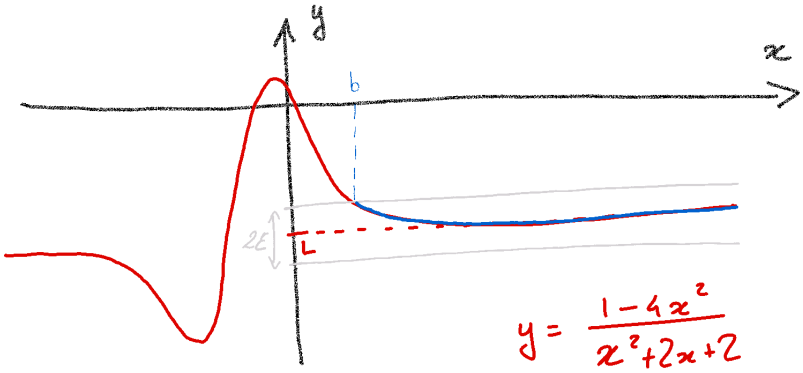

FIGURE 3.6: Illustration of the definition of a limit in the case (3.25).

- If x_0 and L are both finite:

$$\lim_{x \longrightarrow x_0} f = L \Longleftrightarrow \forall \epsilon > 0 \quad \exists \delta > 0 \quad \forall x \in X \; (0 < |x - x_0| < \delta \implies |f(x) - L| < \epsilon) \tag{3.28}$$

> In the definition (3.28), we see that the point x_0 where we look for the limit does not need to belong to the domain X where the function f is defined. Indeed, we never evaluate the function f in x_0 directly.

Cardinal sinus

The cardinal sinus sinc is an omnipresent function in Wave Physics, defined by:

$$\begin{aligned} \mathrm{sinc}\colon \mathbb{R} &\to \mathbb{R} \\ x &\mapsto \mathrm{sinc}(x) = \begin{cases} \dfrac{\sin x}{x} & \text{if } x \neq 0 \\ 1 \quad \text{if} = 0 \end{cases} \end{aligned}$$

Is this function continuous (i.e. $\mathscr{C}^0$) in $x = 0$?

3.1.3 Asymptotic behavior

Two functions f and g are said to be equivalent in x_0, finite or not, if and only if their ratio tends towards 1 in x_0:

$$f \underset{x \to x_0}{\sim} g \Longleftrightarrow \lim_{x \longrightarrow x_0} \frac{f}{g} = 1 \tag{3.29}$$

We say that f and g are asymptotically equal in x_0.

Hyperbolic functions

Give an equivalent of the functions cosh, sinh and tanh in $+\infty$.

The function f is little-o of the function g in x_0, finite or not, if and only if the ratio of f by g tends towards 0 in x_0:

$$f \underset{x \to x_0}{=} o(g) \Longleftrightarrow \lim_{x \longrightarrow x_0} \frac{f}{g} = 0 \tag{3.30}$$

We say that f is dominated by g asymptotically in x_0.

1. Show that $\sin x = x + o(x)$ in $x = 0$.
2. Show that $f \underset{x \to x_0}{\sim} g \iff f - g \underset{x \to x_0}{=} o(g)$.

The function f is big-O of the function g in x_0, finite or not, if and only if the ratio of f by g tends towards a non-zero finite value in x_0. We say that f is is asymptotically bounded by g in x_0 and we write:

$$f \underset{x \to x_0}{=} O(g) \tag{3.31}$$

Polynomial

Let $f(x) = a_0 + a_1 x + a_2 x^2 + ... + a_n x^n$ be a polynomial of degree n, with a_i the constant coefficients of the polynomial.

1. Show that $f \underset{x \to \pm\infty}{\sim} a_n x^n$.
2. Show that $f \underset{x \to \pm\infty}{=} O(x^n)$.
3. Show that $\forall k > n \quad f \underset{x \to \pm\infty}{=} o(x^k)$.

> The equivalence relation $\sim$ and the big-O O are both reflexive. For instance: $f \underset{x \to x_0}{\sim} g \iff g \underset{x \to x_0}{\sim} f$.

Usual functions can be compared between each other through these tools. One can show that:

$$\forall \alpha, \beta \in \mathbb{R} \quad \ln^\beta x \underset{x \to +\infty}{=} o(x^\alpha) \tag{3.32}$$

$$\forall \alpha \in \mathbb{R} \quad \forall a > 1 \quad x^\alpha \underset{x \to +\infty}{=} o(a^x) \tag{3.33}$$

$$\forall \alpha \in \mathbb{R} \quad \alpha^n \underset{n \to +\infty}{=} o(n!) \tag{3.34}$$

where the last relation is only valid if we work with $n \in \mathbb{Z}$ since $n!$ is not defined for $n \in \mathbb{R}$.

> Notice that the function a^x is of exponential type since:
>
> $$a^x = \exp(x \ln a) \tag{3.35}$$
>
> while the function x^α is polynomial.

It is important to memorize the comparison relations between these functions in $+\infty$:

$$\boxed{\text{log} \ll \text{polynomial} \ll \text{exp} \ll \text{factorial}} \tag{3.36}$$

We commonly say that "exponential functions grow faster than polynomial functions which grow faster than logarithmic functions", which serves to determine the limit of a ratio between functions. For instance:

$$\lim_{x \longrightarrow +\infty} \frac{x^2}{e^x} = 0 \quad \text{while} \quad \lim_{x \longrightarrow +\infty} \frac{e^x}{x^2} = +\infty \tag{3.37}$$

Give an equivalent of the functions below:

1. $f(x) = x + 1 + \ln x$ in 0 and in $+\infty$.
2. $f(x) = \cos(\sin x)$ in 0.
3. $f(x) = \cosh(\sqrt{x})$ in $+\infty$.
4. $f(x) = \frac{(\sin x)\ln(1+x^2)}{x\tan x}$ in 0.
5. $f(x) = \ln(1 + \sin x)$ in 0.
6. $f(x) = \ln(\cos x)$ in 0.

> These notions are instrumental in determining the order of convergence of a numerical scheme. They are also tightly bound to the concept of computational complexity in algorithmic sciences (e.g. when we talk about a $N \log N$ algorithms). See the amazing article "Visualizing algorithms" for insightful perks on the algorithmic world.

3.2 Differential calculus

Differential (a.k.a. infinitesimal) calculus is a sub-branch of analysis. It addresses the manipulation of infinitesimal quantities to construct derivatives and compute integrals (section **??**).

3.2.1 Derivative

Let $f \colon \mathbb{R} \to \mathbb{R}$ be the function plotted in Fig. 3.7. The derivative of f with respect to x is the function which associates to each x the rate of change of the function f, in x specifically. An intuitive way to understand the notion of rate of change is to adopt a global approach. If we consider two points x_1 and $x_2 > x_1$, we can define the rate of change of the function f between x_1 and x_2 as:

$$\frac{f(x_2) - f(x_1)}{x_2 - x_1} \tag{3.38}$$

where, by convention, we work with "final value minus initial value". This quantity is the slope of the thin black straight line in the upper panel in Fig. 3.7. Indeed, a straight line is necessarily described by an affine function g:

$$g(x) = ax + b \tag{3.39}$$

with b the value of the function in $x = 0$ and a the slope of the straight line.

1. Why is a straight line parametrized by two degrees-of-freedom (here, the constants a and b)?
2. How many constrains do we need to determine their value?
3. Geometrically, how much information do you need to determine which is the specific straight line we are talking about?

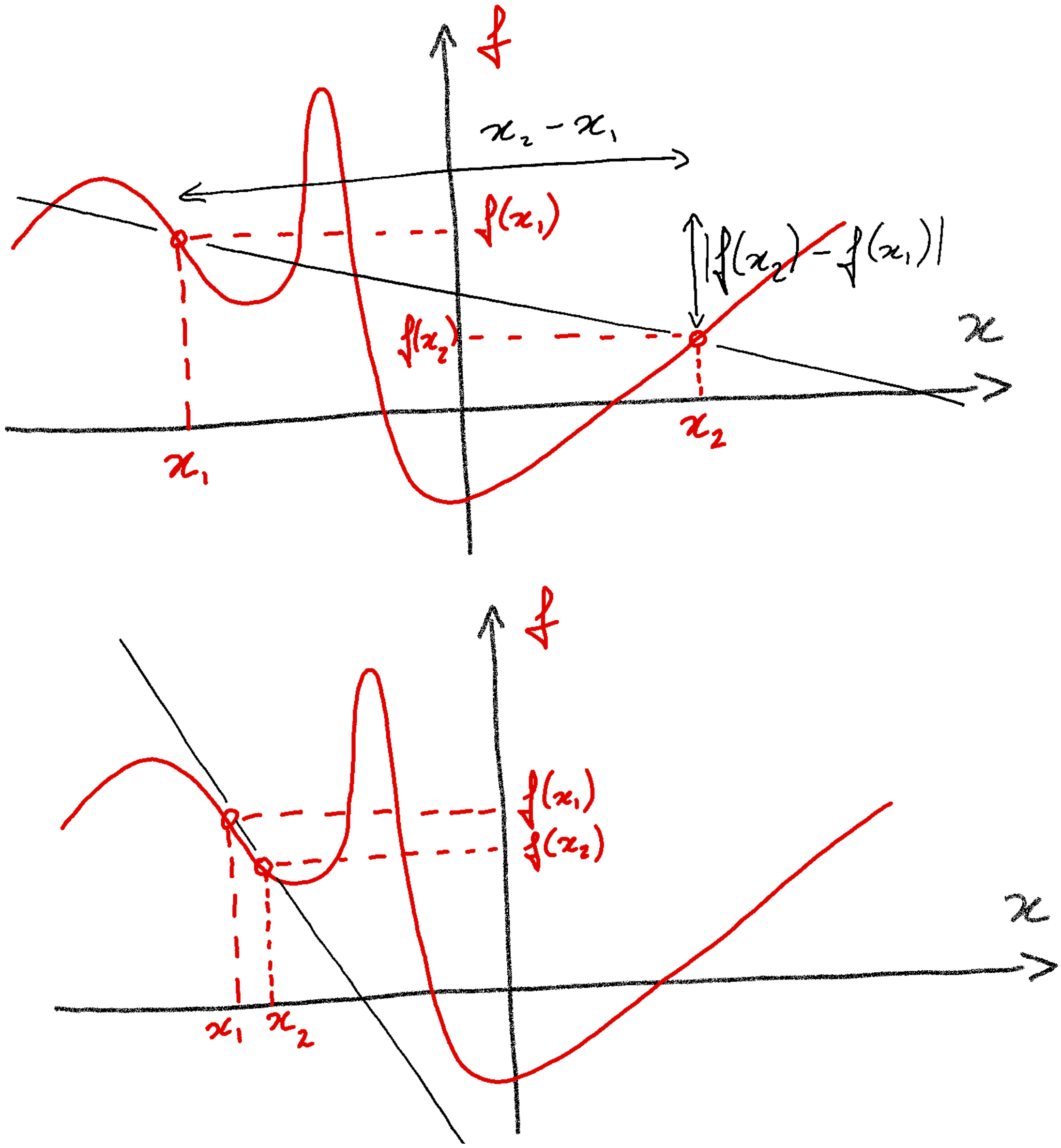

FIGURE 3.7: The slope of the straight line between two points can be obtained from their coordinates (upper panel). As the two points get closer, the straight line looks more and more like the tangent to the curve (lower panel).

Then, if this line passes by the points $(x_1, f(x_1))$ and $(x_2, f(x_2))$, we have:

$$\begin{cases} g(x_1) = f(x_1) = ax_1 + b \\ g(x_2) = f(x_2) = ax_2 + b \end{cases} \tag{3.40}$$

so once we compute the difference between these two equations, we obtain:

$$a = \frac{f(x_2) - f(x_1)}{x_2 - x_1} \tag{3.41}$$

This graphic representation of the rate of change of the function f between x_1 and x_2 is non-local: it is the average of the rate of change of the function f between x_1 and x_2 (something you will show formally in section 3.2.4, once we define the average). On the contrary, the notion of derivative as defined above is local because it is defined at a given x. Therefore, we need the points x_1 and x_2 to be infinitely close from each other for the slope (3.41) to correspond to the derivative, which leads us to the definition of the derivative in x_1:

$$\frac{\mathrm{d}f}{\mathrm{d}x}(x_1) = \lim_{x_2 \longrightarrow x_1} \frac{f(x_2) - f(x_1)}{x_2 - x_1} \tag{3.42}$$

where we generally omit the "(x_1)" on the left hand side, except if the location where the derivative is evaluated is ambiguous. Alternatively, we can define the derivative of the function f with respect to x evaluated in x as:

$$\boxed{\frac{\mathrm{d}f}{\mathrm{d}x} = \lim_{\mathrm{d}x \longrightarrow 0} \left[\frac{f(x + \mathrm{d}x) - f(x)}{\mathrm{d}x} \right]} \tag{3.43}$$

where we introduced the physics notation of an infinitesimal, informally defined as:

$$\text{"}\ \mathrm{d}\bullet = \lim_{\bullet_2 \longrightarrow \bullet_1} (\bullet_2 - \bullet_1)\ \text{"} \tag{3.44}$$

The definition (3.43) emphasizes that a derivative is strictly the same thing as a ratio between two infinitesimals.

> These results are only true if the function is derivable (i.e. $\mathscr{C}^1$), which we assumed to be the case.

Usual derivatives to know by heart

XXX

The concept of derivative enables us to identify whether in a given x, the function f is increasing ($f'(x) > 0$) or decreasing ($f'(x) < 0$). Furthermore, we have:

$$x \text{ is a local extremum} \iff f'(x) = 0 \tag{3.45}$$

To determine whether this extremum is a maximum or a minimum, we need to look at the sign of the second order derivative, defined exactly as in equation (3.43) but with f' instead of f:

$$\frac{\mathrm{d}^2 f}{\mathrm{d}x^2} = \frac{\mathrm{d}f'}{\mathrm{d}x} = \lim_{\mathrm{d}x \longrightarrow 0} \frac{f'(x+\mathrm{d}x) - f'(x)}{\mathrm{d}x} \tag{3.46}$$

$$= \lim_{\mathrm{d}x \longrightarrow 0} \left[\frac{\lim_{\mathrm{d}x_0 \longrightarrow 0} \frac{f(x+\mathrm{d}x+\mathrm{d}x_0) - f(x+\mathrm{d}x)}{\mathrm{d}x_0} - \lim_{\mathrm{d}x_0 \longrightarrow 0} \frac{f(x+\mathrm{d}x_0) - f(x)}{\mathrm{d}x_0}}{\mathrm{d}x} \right] \tag{3.47}$$

where we write $\mathrm{d}x$ and $\mathrm{d}x_0$ to differentiate the two limits: $\mathrm{d}x_0$ tends to 0 at constant $\mathrm{d}x$ and then, $\mathrm{d}x$ tends to zero.

Here, it is interesting to notice the asymmetry between the notations in the numerator and in the denominator. The numerator $\mathrm{d}^2 f$ indicates that we work with a second order infinitesimal (i.e. a limit of limit) while in the denominator, we write $\mathrm{d}x^2$ to emphasize that there is a product of $\mathrm{d}x$ at the denominator. Therefore, from a physical dimensionality point-of-view, the unit of f'' is the same as the unit of f divided by the unit of x^2. From an infinitesimal calculus point of view, it is much easier to work with first order derivatives than with higher order derivatives. For instance, for f and g two functions of x, we can write:

$$\frac{\mathrm{d}f}{\mathrm{d}x} = g(x) \quad \Longleftrightarrow \quad \mathrm{d}f = g(x)\mathrm{d}x \tag{3.48}$$

and then, we can integrate from x_1 to x_2 (section **??**). However, we cannot do the same with the $\mathrm{d}x^2$ element at the denominator:

$$\frac{\mathrm{d}^2 f}{\mathrm{d}x^2} = g(x) \quad \not\Longleftrightarrow \quad \mathrm{d}^2 f = g(x)\mathrm{d}x^2 \tag{3.49}$$

Instead, one would need to work with the first-order derivative and write:

$$\frac{\mathrm{d}f'}{\mathrm{d}x} = g(x) \quad \Longleftrightarrow \quad \mathrm{d}f' = g(x)\mathrm{d}x \tag{3.50}$$

and then integrate from a given x_0 to a fiducial x to obtain:

$$f'(x) - f'(x_0) = \int_{x_0}^{x} g(x)\mathrm{d}x \quad \Longleftrightarrow \quad \frac{\mathrm{d}f}{\mathrm{d}x}(x) = \frac{\mathrm{d}f}{\mathrm{d}x}(x_0) + \int_{x_0}^{x} g(x)\mathrm{d}x \tag{3.51}$$

$$\Longleftrightarrow \mathrm{d}f = \left[\frac{\mathrm{d}f}{\mathrm{d}x}(x_0)\right]\mathrm{d}x + \left[\int_{x_0}^{x} g(x)\mathrm{d}x\right]\mathrm{d}x \tag{3.52}$$

and we can now integrate a second time, this time between x_1 and x_2 to obtain:

$$f(x_2) - f(x_1) = f'(x_0)(x_2 - x_1) + \int_{x_1}^{x_2} \left[\int_{x_0}^{x} g(x)\mathrm{d}x\right]\mathrm{d}x \tag{3.53}$$

since $f'(x_0)$ does not depend on x (since it was already evaluated in x_0) so it can be extracted from the integral.

Then, we have the definition of a local maximum and minimum[3]:

$$\boxed{x \text{ local maximum (resp. minimum)} \Longleftrightarrow \begin{cases} f'(x) = 0 \\ f''(x) < 0 \text{ (resp. } f''(x) > 0) \end{cases}} \tag{3.54}$$

If the second order derivative cancels out in x, then x is neither a local maximum nor minimum (e.g. the the function $f\colon x \mapsto x^3$ in $x = 0$). A local maximum (resp. minimum) which

[3]One can remember this result by considering the minimum of the function $f\colon x \mapsto x^2$ in $x = 0$.

is higher (resp. lower) in algebraic value[4] than all the other local maxima (resp. minima) over the domain X over which the function is defined is the absolute maximum (resp. minimum).

> In Physics, you used equation (3.54) to determine whether an equilibrium was stable (local minimum), unstable (local maximum) or meta-stable (second derivative null). We will see in section 3.2.2 that it is a consequence of the way the system reacts to a perturbation, which can be written as a Taylor expansion.

Let $f\colon X \to Y$ be a function, we further define:

$$f \text{ strictly increasing (resp. decreasing)} \iff \forall x \in X \quad f'(x) > 0 \text{ (resp. } f'(x) < 0) \tag{3.55}$$

A function which is either strictly increasing or strictly decreaing is said to be monotonic.

1. Show that a function $f\colon X \to f(X)$ is monotonic if and only if it is injective.
2. In a stellar interior where the mass density profile ρ is non-monotonic, can we work both with radial and mass coordinates?
3. Would have it been legit if we had been working with an electrically charged sphere?

> We left aside the marginal case where the function f is constant to avoid confusion. Furthermore, the result above on the equivalence between monotonicity and injectivity is only valid if the function is continuous.

One can always choose the variable with respect to which a function is derived. Let us consider the function:

$$f\colon x \mapsto f(x) = e^{-x^2} \tag{3.56}$$

Then, we can compute the derivative of f with respect to x using the chain rule like explained below, but we can also consider the derivative of f with respect to another variable y. For instance:

- With the change of variable $y = e^{-x^2}$, the function $f(y)$ is:

$$f\colon y \mapsto f(y) = y \tag{3.57}$$

  and its derivative is:

$$\frac{\mathrm{d}f}{\mathrm{d}y} = \frac{\mathrm{d}f}{\mathrm{d}\left(e^{-x^2}\right)} = 1 \tag{3.58}$$

- With the change of variable $y = -x^2$, the function $f(y)$ is:

$$f\colon y \mapsto f(y) = e^{y} \tag{3.59}$$

  and its derivative is:

$$\frac{\mathrm{d}f}{\mathrm{d}y} = \frac{\mathrm{d}f}{\mathrm{d}\left(-x^2\right)} = e^{y} = e^{-x^2} \tag{3.60}$$

- With the change of variable $y = x^2$, the function $f(y)$ is:

$$f\colon y \mapsto f(y) = e^{-y} \tag{3.61}$$

[4] That is to say accounting for the sign.

and its derivative is[5]:
$$\frac{\mathrm{d}f}{\mathrm{d}y} = \frac{\mathrm{d}f}{\mathrm{d}\left(x^2\right)} = -e^{-y} = -e^{-x^2} \tag{3.62}$$

These results are perfectly coherent together once we reformulate the denominators since:

- $\frac{\mathrm{d}\left(e^{-x^2}\right)}{\mathrm{d}x} = -2xe^{-x^2}$ so $\mathrm{d}\left(e^{-x^2}\right) = -2xe^{-x^2}\,\mathrm{d}x$ and equation (3.58) gives:
$$\mathrm{d}f = -2xe^{-x^2}\,\mathrm{d}x \tag{3.63}$$
- $\frac{\mathrm{d}\left(-x^2\right)}{\mathrm{d}x} = -2x$ so $\mathrm{d}\left(-x^2\right) = -2x\,\mathrm{d}x$ and equation (3.60) gives:
$$\mathrm{d}f = -2xe^{-x^2}\,\mathrm{d}x \tag{3.64}$$
- $\frac{\mathrm{d}\left(x^2\right)}{\mathrm{d}x} = 2x$ so $\mathrm{d}\left(x^2\right) = 2x\,\mathrm{d}x$ and equation (3.62) gives:
$$\mathrm{d}f = -2xe^{-x^2}\,\mathrm{d}x \tag{3.65}$$

In the three cases, we retrieve the same expression for $\mathrm{d}f$. As a conclusion, we can always choose the "block" with respect to which we derive a function. Let us use this information to determine the derivative of the composition between two function. Let f and g be two functions and $h = f \circ g$ their composition. The function h can either be seen as a function of x directly:
$$h\colon x \mapsto h(x) \tag{3.66}$$
or as a function f of a function g:
$$h\colon x \mapsto f(g(x)) \tag{3.67}$$
In the latter case, we can use the chain rule to decompose the derivative of h with respect to x into two intermediate parts:
$$\frac{\mathrm{d}h}{\mathrm{d}x}(x) = \frac{\mathrm{d}f}{\mathrm{d}g}(g(x)) \cdot \frac{\mathrm{d}g}{\mathrm{d}x}(x) \tag{3.68}$$
where we specified the variable at which each derivative is evaluated since in this case, it is not obvious. It means that to determine the derivative of a composed function $f \circ g$, the strategy is the following:

1. First, determine the derivative of the inner function, $g(x)$. It gives the second factor in the right hand side of the equation (3.68).
2. Then, consider $g(x)$ as a block. Said otherwise, replace $g(x)$ with a variable y and compute the derivative of f with respect to y.
3. In the derivative of f with respect to y you just computed, replace back y with $g(x)$, and you obtain the first factor in the right hand side of the equation (3.68).

[5]Beware, here, $\mathrm{d}\left(x^2\right) = 2x\,\mathrm{d}x$ should not be confused with the $\mathrm{d}x^2$ element which intervenes in the notation of the second order derivative.

Equation (3.68) is strictly the same as the more famous (and compact) one you saw in mathematics:

$$(f \circ g)'(x) = (f' \circ g)(x) \cdot g'(x) = f'(g(x))g'(x) \tag{3.69}$$

but it is more explicit regarding the variable with respect to which we derive (either x or g), and the method to follow when we want to derive a composed function.

The chain rule can be applied as many times as needed so you can also work with successive compositions $f \circ g \circ h \circ \ldots$.

1. Remind the domain X over which the main analytic functions are defined, and their output domain Y: trigonometric (3 functions and their inverse), hyperbolic (3 functions and their inverse), exponential a^x (with $a > 1$), logarithmic $a \ln x$, power-law x^a...
2. Remind their derivatives over these domains.
3. From these 15 functions, combine them together 2-by-2 to obtain ~ 120 combinations, for instance:

$$f(x) = \arctan(\ln x) \tag{3.70}$$

   or

$$f(x) = \left(1 + e^{x^2}\right)^{-1/2} \tag{3.71}$$

   You will pay attention to the domain over which these combined functions are defined.
4. For each of these combination, determine the derivative using the chain rule.
5. Let $f\colon x \mapsto \sin\left[\exp\left(ax^2 + bx + c\right)\right]$ be a function, with a, b and c three constants.
   (a) Write f as the composition between three main analytic functions f_1, f_2 and f_3.
   (b) Deduce the derivative $f'(x)$ from this decomposition.

Baywatch

Fig. 3.8 represents a flat sand beach (top part) and the sea (bottom part) as seen from above. A lifeguard is located at point $A(x_A, y_A)$ on the beach, where we introduced the $\hat{\mathbf{x}}$-axis colinear to the sea side and pointing towards the right, and the $\hat{\mathbf{y}}$-axis normal to the sea side and pointing towards the beach. A tourist located at point $B(x_B, y_B)$ in the sea is about to drown. Through the problem, the tourist remains fixed. The lifeguard, who wants to reach the tourist as fast as possible, can run with speed v_1 and swim with speed $v_2 < v_1$. We write $I(x_I, y_I)$ the point where the lifeguard enters the sea.

1. What is the shortest path from A to B? We write $C(x_C, y_C)$ the intersection of this path with the sea side. Is $x_C >, <$ or $= x_I$?
2. Why is this problem effectively one-dimensional?
3. Introduce a convenient origin to simplify the problem.

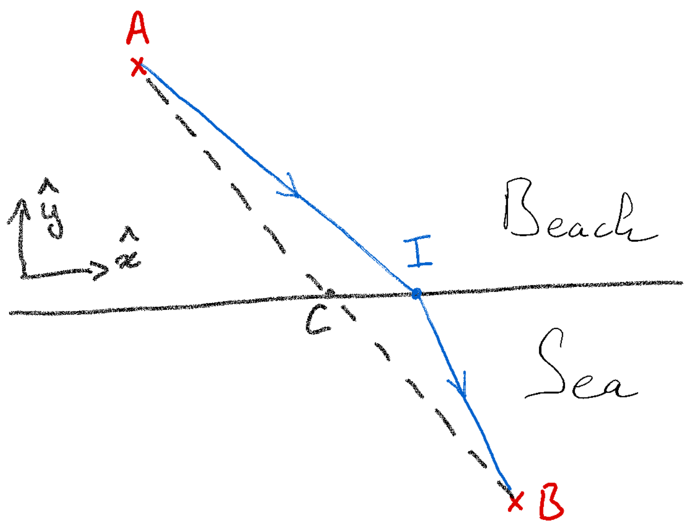

FIGURE 3.8: Trajectory (in blue) of the lifeguard A to the tourist B.

4. For a fiducial point I, what is the duration Δt taken by the lifeguard to reach the tourist?
5. For which specific position of the point I is this duration Δt minimal? Is it coherent with the limit case when $v_2 = v_1$?
6. Re-express this condition based only on the angles i_1 and i_2, and the velocities v_1 and v_2. Which law do you recognize?

> In geometric optics, the principle of least time states that between points A and B, a light ray follows the path that can be traveled in the least time. It is a special case of the more general principle of least action that you saw in analytic mechanics. The laws of refraction can be deduced from the principle of least time.

The functions we worked with until now were functions of one variable x only. Therefore, the changes of f are necessarily computed with to the unique independent variable (be it x or any other variable obtained from x by a change of variable): in Fig. 3.7, there is only one direction, one axis with respect to which we can move and study the variation of f. In this case, we can use the d symbol in the definition of the derivative of f with respect to x (or any other variable obtained from x by a change of variable). Now, let f be a function of two variables:

$$\begin{aligned} f \colon \mathbb{R}^2 &\to \mathbb{R} \\ (x, y) &\mapsto f(x, y) \end{aligned}$$

represented in Fig.XXX. In this case, the derivative is ambiguous since we can choose to derive the function f either:

- with respect to x, assuming that y is constant: $\left.\dfrac{\partial f}{\partial x}\right|_y$
- with respect to y, assuming that x is constant: $\left.\dfrac{\partial f}{\partial y}\right|_x$

Each of these derivatives is called partial and we use the ∂ symbol. Contrary to the exact differential d (e.g. in $\mathrm{d}f$), we cannot manipulate and separate the numerator and denominator in the partial derivatives. Each partial derivative is equivalent to perform twice the computations we computed for a function of one variable only: first along the $\hat{\mathbf{x}}$-axis (i.e. at constant y) and then along the $\hat{\mathbf{y}}$-axis (i.e. at constant x). We will come back on these aspects in section 5.1 devoted to the gradient operator.

Compute the partial derivatives of the following functions:

1.

3.2.2 Taylor expansion

Now that we introduced the concepts of asymptotic behavior and derivative, we can explain the notion of Taylor expansion which plays a key-role in performing approximate evaluations in Physics in general and in Hydrodynamics in particular. For the sake of simplicity, we introduce the notation $f^{(n)}$ to write the n^{th}-order derivative, with n an integer higher or equal to zero: $f^{(0)} = f$, $f^{(1)} = f'$, $f^{(2)} = f''$...

The idea of a Taylor expansion is to approximate, in a given point $x = a$, a function f by a polynomial f_n of n^{th}-order (with n an integer $\leqslant 0$). For the sake of simplicity, let us first work in $x = 0$. The general expression of the polynomial f_n is:

$$f_n\colon x \mapsto f_n(x) = a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n \tag{3.72}$$

$$= \sum_{i=0}^{n} a_i x^i \tag{3.73}$$

Let us design a procedure to find polynomials which approximates a function $f\colon X \to Y$ in $x = 0$. We will proceed step by step and extrapolate the result.

The first, crudest and simple most way to approximate the function f in $x = 0$ would be with a function f_0 defined by:

$$f_0\colon x \mapsto f_0(x) = f(0) \tag{3.74}$$

This function is constant and has the value of the function f at the point $x = 0$ where we perform the approximation (Fig. 3.9). As soon as x is a bit different from 0, this approximation is no longer accurate at all (except in the very specific case where the function f we are trying to approximate is itself constant). The function f_0 is called the 0^{th}-order Taylor expansion of the function f in 0, because it is a 0^{th}-order polynomial in x. If we compare it to equation (3.72), we get the first coefficient of the polynomial f_n:

$$a_0 = f(0) \tag{3.75}$$

The next step is to approximate the function f in 0 taking into account the local slope through the first-order derivative $f^{(1)}$. Doing so, we will rely on the tangent to the function f in 0 defined by:

$$f_1\colon x \mapsto f_1(x) = f(0) + x f^{(1)}(0) \tag{3.76}$$

The function f_1 is called the 1^{st}-order Taylor expansion of the function f in 0, because it is a 1^{st}-order polynomial in x. It is a better approximation of the function f near $x = 0$ than the

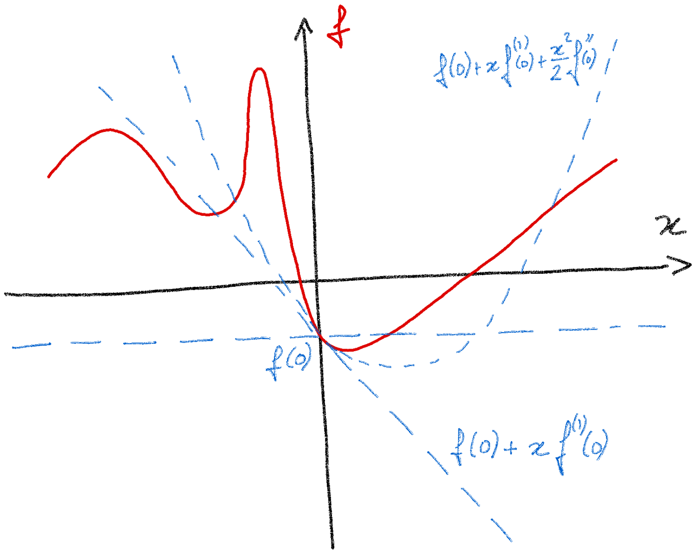

FIGURE 3.9: Successive Taylor expansion of the function f in $x = 0$.

function f_0. Indeed, we have:

$$\frac{f(x) - f_1(x)}{f(x) - f_0(x)} = \frac{f(x) - f(0) - xf^{(1)}(0)}{f(x) - f(0)} \tag{3.77}$$

$$= 1 - \frac{xf^{(1)}(0)}{f(x) - f(0)} \tag{3.78}$$

The last term seems to have an undetermined limit when $x \to 0$ because the numerator and the denominator both tend towards 0. Yet, the definition (3.43) of the derivative tells us that:

$$\frac{xf^{(1)}(0)}{f(x) - f(0)} = \frac{x}{f(x) - f(0)} \cdot \lim_{x \longrightarrow 0} \left[\frac{f(x) - f(0)}{x} \right] \tag{3.79}$$

so the limit of this term when $x \to 0$ is 1. Therefore, the ratio in equation (3.77) tends to 0, which tells us that f_1 tends towards f faster than f_0 when $x \to 0$. If we compare the function f_1 to equation (3.72), we get the second coefficient of the polynomial f_n. To get rid of a_0 and focus on a_1, we derive the polynomial f_n once and evaluate it in $x = 0$:

$$f_n^{(1)}(0) = a_1 \tag{3.80}$$

and we know from equation (3.76) that:

$$f_1^{(1)}(0) = f^{(1)}(0) \tag{3.81}$$

so:

$$a_1 = f^{(1)}(0) \tag{3.82}$$

Let us try to generalize the expression (3.76) to higher orders by following the same procedure to determine the a_i. For instance, for a_2, we identify the second-order derivative of the polynomial f_n to the second-order derivative of the function f:

$$f^{(2)}(0) = 2a_2 \tag{3.83}$$

so:

$$a_2 = \frac{f^{(2)}(0)}{2} \tag{3.84}$$

Show by recurrence that:

$$a_i = \frac{f^{(i)}(0)}{i!} \tag{3.85}$$

We re-inject the expression of the a_i coefficients in the polynomial f_n to obtain the n^{st}-order Taylor expansion of the function f in 0:

$$f_n(x) = \sum_{i=0}^{n} x^i \frac{f^{(i)}(0)}{i!} \tag{3.86}$$

In general[6], the higher the order, the better the approximation up to a larger distance from $x = 0$ (Fig. 3.9).

> Generally, in Physics, we simply use the $\sim$ symbol to approximate the function by its n^{st}-order Taylor expansion and we say that "$f \sim f_n$ in 0".

Ideally, if we wanted the approximation to be perfect in any point $x \in X$, we would need to go up to an infinite order. As a matter of fact, it can be shown that a function f is equal to the Taylor series:

$$\boxed{f(x) = \sum_{i=0}^{+\infty} x^i \frac{f^{(i)}(0)}{i!}} \tag{3.87}$$

We can generalize this result to obtain the n^{th}-order Taylor expansion of the function f in $x = x_0 \neq 0$ (with $x_0 \in X$ and therefore, necessarily finite):

$$\boxed{f(x) = \sum_{i=0}^{+\infty} (x - x_0)^i \frac{f^{(i)}(x_0)}{i!}} \tag{3.88}$$

Demonstrate equation (3.88) by following the same procedure as before but in $x = x_0 \neq 0$.

Usual Taylor expansions to know by heart

XXX

> Generally, we will use these expansions in $x = 0$.

[6]These results are only true if the function is $\mathscr{C}^\infty$ over X, which we assumed to be the case.

Taylor expansions and limits

1. Show that $(1+x)^{1/x} \underset{x\to 0}{\sim} e$
2. Show that $e^x - 1 \underset{x\to 0}{\sim} x$
3. Show that $\lim_{x\to 0^+} x^x = 1$
4. Show that $\lim_{x\to 0} \frac{\ln(1+ax)}{bx} = \frac{a}{b}$, with a and b two constants.

> As illustrated by this exercise, Taylor expansions represent a convenient way to determine (or retrieve) the limit of a function.

> Taylor expansion is a type of series expansion: we write a function as an infinite sum. Fourier expansion is another type of series you are familiar with, particularly relevant when we work with periodic functions. On a sphere, necessarily, a function of the azimuthal ϕ coordinate is at least 2π periodic. The analogous to Fourier series on a sphere are spherical harmonics like the ones you used in Quantum mechanics to describe the atomic orbitals of the atom of hydrogen, and that you will use to solve Laplace equation in equation (**??**) of this course. In Astrophysics, spherical harmonics lie at the core of asteroseismology which describe how stars pulsate and how we can use these pulsations to determine the internal structure of a star.

> Numerical evaluation of typical functions (e.g. exponential) is based on high order series expansions. It is how Human computers used to do before actual computers became the norm.

3.2.3 Primitive

We saw in section 3.2.1 that a derivative was the same thing as the ratio between two infinitesimals. In equations (3.63) to (3.65), we saw that we could separate the infinitesimals in the numerator (e.g. $\mathrm{d}f$) and in the denominator (e.g. $\mathrm{d}x$) to obtain quantities such as:

$$g(x)\,\mathrm{d}x \tag{3.89}$$

where g was a function of x. We can generalize the computation we made by saying that any infinitesimal quantity $\mathrm{d}\left(F(x)\right)$ (where F is a function of x) can be "developed" such as we extract things out of the differential d, provided we first determine the derivative f of the function F. Then, we have:

$$\boxed{\mathrm{d}\left(F(x)\right) = f(x)\,\mathrm{d}x} \tag{3.90}$$

The $\mathrm{d}x$ at the end is vital (i) because it indicates the variable with respect to which we derived the function F (after all, we could have performed a change of variable and derive with respect to another variable), and (ii) because without it, we would have an infinitesimal on the left hand side and a finite quantity on the right hand side, which is impossible (it is a kind of "homogeneity" argument: two objects of different nature cannot be equal).

Differentials

Develop the following differentials:

1. $\mathrm{d}\left(x^2/2\right)$
2. $\mathrm{d}\left(e^{-x^2}\right)$
3. $\mathrm{d}\left(\arctan x\right)$
4. $\mathrm{d}\left(\arctan\left(e^{-x^2}\right)\right)$

The inverse of a derivative is a primitive in the sense that the derivative of a primitive (and vice versa) is the function itself. From the left hand side of equation (3.90) to the right hand side, we derived the function F. On the reverse, if we want to go from the right hand side to the left hand side, we need to find the primitive F of the function f with respect to x, that is to say the function F which, ones derived with respect to x, yields f.

Write the following expressions as the differential of a function:

1. $x\,\mathrm{d}x$
2. $-2xe^{-x^2}\,\mathrm{d}x$
3. $\frac{1}{1+x^2}\,\mathrm{d}x$
4. $\frac{-2x}{1+e^{x^2}}\,\mathrm{d}x$

Usual primitives to know by heart

XXX

3.2.4 Integration

On purpose, I separated the question of integration (in this section) from the question of primitives (in the previous sections). The goal is to make it clear that we can manipulate infinitesimals without caring about integrating quantities. Infinitesimals are insightful quantities in their own right, as made clear by the central role played by differential equations in Physics. Let us first understand the concept of integral through a concrete problem: the need for a mathematical way to compute a surface.

Imagine you are tasked with doubling the size of a rectangular-shaped pucará. The first question you should ask is to precise this ambiguous instruction since it could either mean...:

- ... that you have to double the length of each side of the rectangle and in this case, the surface of the new pucará will be four times larger.
- ... that you have to double the surface of the pucará.

In the latter case, you will need a versatile way to measure the surface of the initial and final pucarás. A common practice in ancient times was to use thin wood sticks and to align them in a way similar to the left panel in Fig. 3.10: the length of the wood sticks correspond to the length of the shortest side of the rectangle, and the width of the wood sticks being very

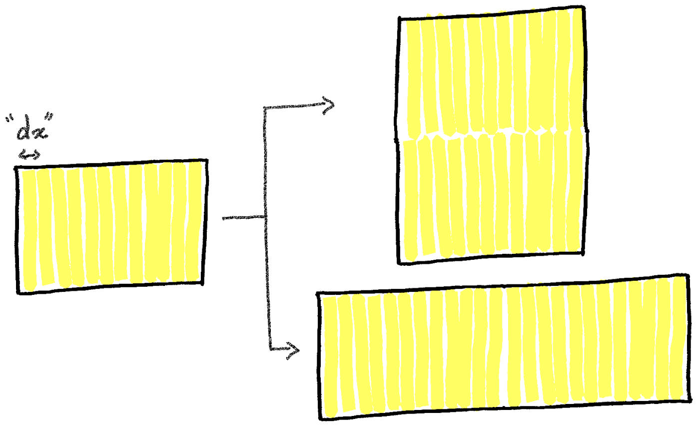

FIGURE 3.10: Measure of the surface of a pucará with thin wood sticks (in yellow) and methods to double it.

small, we can write it[7] $\mathrm{d}x$. The surface is defined by the number of wood sticks needed, and the simplemost way to double it is to build a surface which encompasses twice more of these wood sticks.

In 1D, to define a length, we can adopt a similar approach: a length can be defined as the sum of tiny segments of length $\mathrm{d}x$, which yields the fundamental relation between integrals and differentials:

$$\boxed{\int_{x_1}^{x_2} \mathrm{d}x = x_2 - x_1} \tag{3.91}$$

with $x_2 > x_1$. This equality means that the continuous sum of infinitesimal elements of length $\mathrm{d}x$ from position x_1 to position x_2 gives the length $x_2 - x_1$.

> As a matter of fact, the integral symbol $\int$ is a stylized version of the letter "s" which stands for "sum", exactly like the Σ used in discrete summation (since Σ is the Greek equivalent of the Latin letter s). The integral is the continuous counterpart of discrete summation.

We now have a general method to compute the integral, with respect to a variable x, of any function f:

$$\int_{x_1}^{x_2} f(x)\,\mathrm{d}x \tag{3.92}$$

The idea is to move the function f inside the differential d. To do so, we look for the primitive F of f with respect to x to replace $f(x)\,\mathrm{d}x$ with $\mathrm{d}\,(F(x))$. We are left with:

$$\int_{x_1}^{x_2} f(x)\,\mathrm{d}x = \int_{F(x_1)}^{F(x_2)} \mathrm{d}\,(F(x)) \tag{3.93}$$

where we carefully adapted the bounds of the integral to the content of the differential element d. Indeed, the bounds of the integral should always be two given values of the function inside the differential element d, which was simply x in equation (3.91) but which is $F(x)$ in the left hand side of equation (3.93).

[7] Strictly speaking, it is not legit because however thin the wood sticks are, they still have a finite width.

> On a side-note, it is always more comfortable and less prone to errors to work with a dimensionless expression $f(x)\,\mathrm{d}x$, by extracting the scales (necessarily constant) out of the integral.

Then, we can use equation (3.91) to deduce the value of the integral:

$$\int_{F(x_1)}^{F(x_2)} \mathrm{d}\left(F(x)\right) = F(x)|_{F(x_1)}^{F(x_2)} = F(x_2) - F(x_1) \tag{3.94}$$

An alternative method, though risky and somewhat less elegant, is to skip the bounds of the integral and add a constant of integration C that you will determine later on, for instance with a boundary or an initial condition:

$$\int_{x_1}^{x_2} f(x)\,\mathrm{d}x = \int \mathrm{d}\left(F(x)\right) = F(x) + C \tag{3.95}$$

Integrals obey a bunch of properties similar to derivatives such as linearity. Also, the computation above implies that:

$$\int_{x_1}^{x_2} f(x)\,\mathrm{d}x = -\int_{x_2}^{x_1} f(x)\,\mathrm{d}x \tag{3.96}$$

The integral of a function f with respect to x, from x_1 to x_2, can be interpreted graphically as the surface if $x_2 > x_1$ (or the opposite of the surface is $x_2 < x_1$) below the graph of f and between the points x_1 and x_2 (Fig.XXX). Consequently, the parity of a function enables us to simplify an integral:

$$f \text{ is even} \iff \forall x \in X \int_{-x}^{x} f(x)\,\mathrm{d}x = 2\int_{0}^{x} f(x)\,\mathrm{d}x \tag{3.97}$$

and more importantly:

$$f \text{ is odd} \iff \forall x \in X \int_{-x}^{x} f(x)\,\mathrm{d}x = 0 \tag{3.98}$$

Evaluate the integrals below:

1. $\int_{-x}^{x} \arctan\left[e^{x^2}\sin x\right]\mathrm{d}x$ for any $x \in \mathbb{R}$.
2. $\int_{\pi/2-x}^{\pi/2+x} \cos(3x)\sin^2(4x)\,\mathrm{d}x$ for any $x \in \mathbb{R}$.
3. $\int_{-x}^{x} |x|\,\mathrm{d}x$ for any $x \in \mathbb{R}$.

> Beware, the later function is not $\mathcal{C}^1$ over the whole $\mathbb{R}$ domain because of the absolute value. Here, it is not a problem but if you want to perform other computation, like a Taylor expansion, it is always safer to treat the function over individual sub-domains where it is $\mathcal{C}^\infty$ (here, $\mathbb{R}^+$ and $\mathbb{R}^-$. See also the "Mind the gap" insert below.

Also, we have the following useful property that we will use to derive the continuity equation in section **??**:

$$\forall a, b \in X \int_{a}^{b} f(x)\,\mathrm{d}x = 0 \iff f = \tilde{0} \tag{3.99}$$

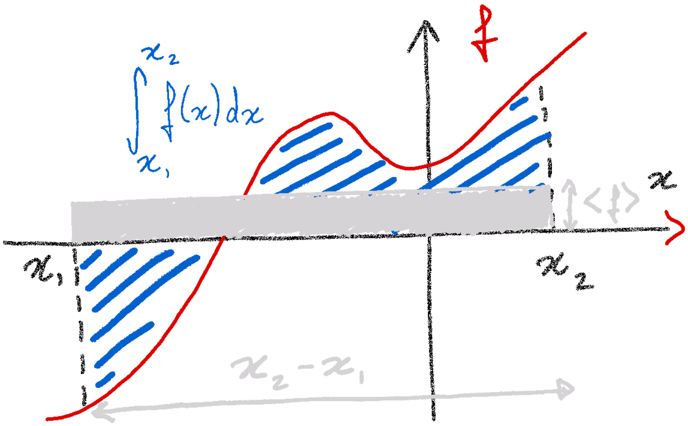

FIGURE 3.11: The integral of a function between x_1 and $x_2 > x_1$ is equal to the surface under the curve (in blue), counted positively (resp. negatively) where $f(x) > 0$ (resp. $f(x) < 0$). The average $\langle f \rangle$ of f between x_1 and x_2 is such that the rectangle of sides $x_2 - x_1$ and $\langle f \rangle$ has the same surface (in grey) as the surface under the curve.

where $\tilde{0}$ is the null function defined over X by $x \mapsto \tilde{0}(x) = 0$.

Mind the gap

Primitives and derivatives are local operations, contrary to integrals. Therefore, we should always make sure beforehand that over the interval $[x_1, x_2]$ we consider, the function f we integrate is at least continuous, at best $\mathcal{C}^\infty$. As you know, there are tricky integrals where the function is not defined in x_1, x_2 or any point in-between, and they do not necessarily converge. Online tools such as Wolfram integrator will not always warn you about these subtilities that you need to check beforehand.

The integral enables us to define the average (a.k.a. mean) $\langle f \rangle_{[x_1,x_2]}$ of the function f over a given interval $[x_1, x_2]$ (with $x_2 > x_1$) as:

$$\boxed{\langle f \rangle_{[x_1,x_2]} \hat{=} \frac{\int_{x_1}^{x_2} f(x)\,\mathrm{d}x}{\int_{x_1}^{x_2} \mathrm{d}x} = \frac{\int_{x_1}^{x_2} f(x)\,\mathrm{d}x}{x_2 - x_1}} \tag{3.100}$$

Graphically, it turns out that $\langle f \rangle_{[x_1,x_2]}$ is the value such as the rectangle of sides $\langle f \rangle_{[x_1,x_2]}$ and $x_2 - x_1$ have the same (algebraic) surface as the surface below the graph of f and between the points x_1 and x_2 (Fig. 3.11). It can be seen immediately from equation (3.100) by writing:

$$\langle f \rangle_{[x_1,x_2]} \cdot (x_2 - x_1) = \int_{x_1}^{x_2} f(x)\,\mathrm{d}x \tag{3.101}$$

Show that equation (3.38) corresponds to the average of the derivative $\frac{\mathrm{d}f}{\mathrm{d}x}$ between x_1 and x_2.

3.3 Differential equations

XXX all linear and with exact derivatives

> Non-linear as an entry point to chaos theory.

3.3.1 First order

Homogeneous

with coefficients constant or not

Heterogeneous

with coefficients constant

3.3.2 Second order

with coefficients constant

Homogeneous

characteristic equation

Heterogeneous

sinusoidal excitation and complex method resonance

Chapter 4

Geometry

In this section, we focus on vector calculus, of direct interest for this class, and largely elude other aspects of geometry, either simpler (e.g. Cartesian, cylindrical and spherical basis) or less central in the study of Hydrodynamics (e.g. conicals).

4.1 Vectors

Vector calculus can be understood as a sub-branch of either algebra (through the notion of vector space) or geometry (through their graphic interpretation). On purpose, we provide here a simplified (and inaccurate) definition of vectors which bypasses the fundamental nature of a vector as an object which obeys specific transformation. We do so for we believe that the rigorous mathematical definition is not decisive for most applications met in undergraduate studies[1] (except when we compute vector fields from distributions where it plays a minor role, section 4.6.3).

4.1.1 Definition

A vector $\mathbf{u}$ is a mathematical object which contains three information: its magnitude (a.k.a. norm) $|u|$, its direction, which is a straight line, and its orientation, which indicates in which of the two directions it points along the straight line. It can be written as the product of its magnitude by a unit vector:

$$\mathbf{u} = |u| \cdot \frac{\mathbf{u}}{|u|} \tag{4.1}$$

A vector does not have unit, but its magnitude can have one. A unit vector is a vector of norm unity (so its magnitude does not have any dimension). By construction, the vector $\frac{\mathbf{u}}{|u|}$ on the right hand side of equation (4.1) has a norm of 1.

> Notice that among the information contained in a vector, there is nothing about its location: a vector does not exist in a specific place. In Fig. 4.1, vectors $\mathbf{u}$ and $\mathbf{v}$ are strictly the same because they have the same magnitude, the same direction and the same orientation: $\mathbf{u} = \mathbf{v}$. Therefore, you can move a vector as much as you want, as long as you do not change its magnitude, direction and orientation. It is the lack of location which will drive us into introducing the notion of vector field.

From a linear algebra point-of-view, a vector can be interpreted as a uni-dimensional array containing as many coefficients as the dimension of the vector space we work in. For instance, we we are working in the classic 3D space, it contains three coefficients, but if we work in the phase space $(\mathbf{r}, \mathbf{v})$, where $\mathbf{r}$ is the location and $\mathbf{v}$ is the velocity, it can contain up to 6 coefficients. If we stick to a simple case where we work in the classic 2D space, we can use equation (4.1) to elaborate a graphic interpretation of what a vector is: a vector indicates how high is a quantity (through its norm) and in which direction it is pointing (through the

[1] It will be central though when you study general relativity.

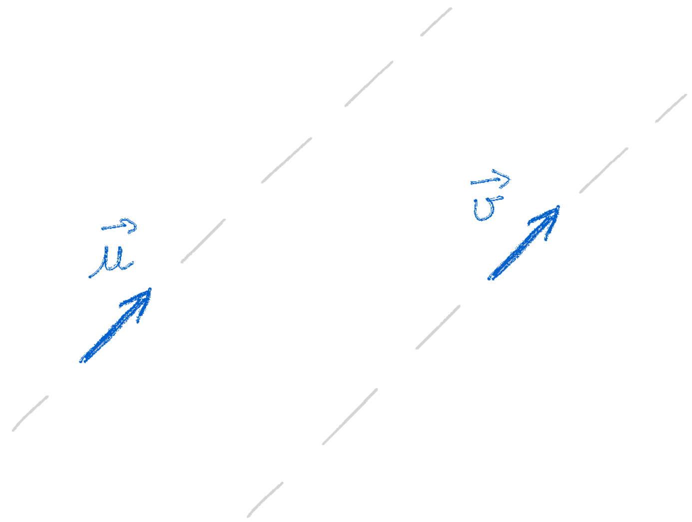

FIGURE 4.1: The vectors $\mathbf{u}$ and $\mathbf{v}$ are the same, even if they do not have the same origin, because they have the same magnitude, direction and orientation.

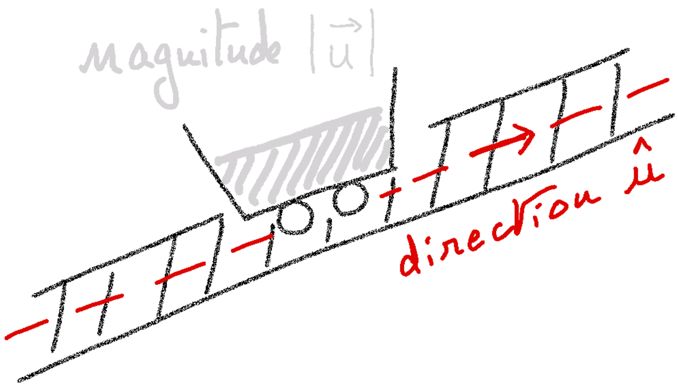

FIGURE 4.2: The vector can be seen as the combination between an information on the direction taken by a mine cart, and the amount of material it contains.

unit vector). A mining metaphor would be that the unit vector stands for the rails while the magnitude represents the amount contained by the mine cart (Fig. 4.2). More generally, when we face the quantity:

$$\alpha\,\mathbf{u} \quad \text{with } \alpha \in \mathbb{R} \tag{4.2}$$

we say that "the number α is carried by the vector $\mathbf{u}$". Beware, if $\alpha < 0$, the vector $\alpha\,\mathbf{u}$ is in the opposite direction of $\mathbf{u}$. If you aren't into extractivism, you can also see equation (4.2) as a mathematical description of a bead on a wire: the wire sets the direction while the size (or the mass) of the bead stands for the magnitude. In section 4.6.2, we will see yet another graphic interpretation through the concept of vector field.

4.1.2 Operations

We can define several operations involving vectors. First, vectors can be added together, which gives a vector. Second, we can multiply a scalar and a vector, like in equation (4.2). Then, we can perform dot and vector products.

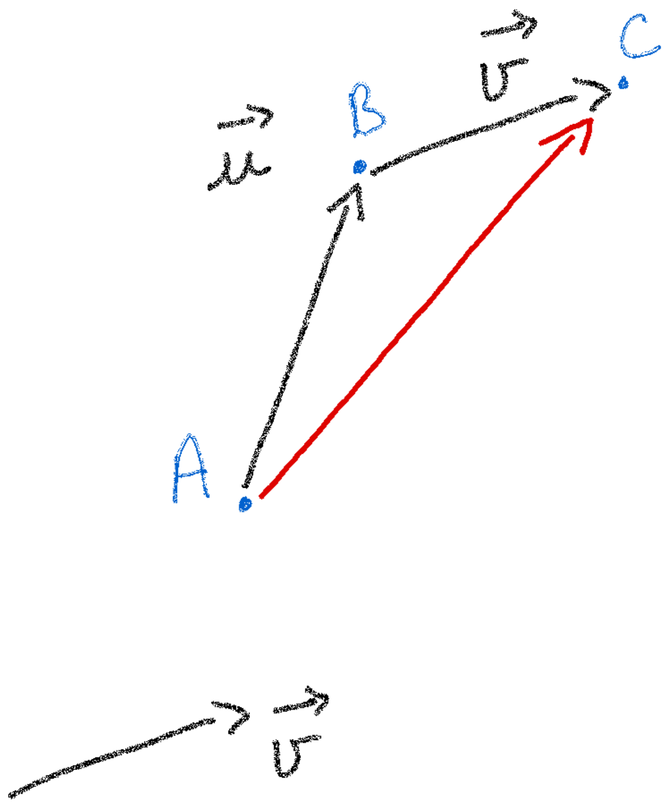

FIGURE 4.3: Two vectors $\mathbf{u}$ and $\mathbf{v}$ can be summed, and the result is obtained after shifting one vector's origin (here, the origin of $\mathbf{v}$) such as the ending point of one is the starting point of the other (here, point B). Then, the result of the sum is obtained by joining the starting point of the whole path (here, A) to the ending point of the whole path (here, C).

Addition

Two vectors $\mathbf{u}$ and $\mathbf{v}$ can be added, which yields a vector $\mathbf{w}$ that we determine from Chasles' relation (Fig. 4.3). Since vectors do not have fixed position in space, we can always move $\mathbf{u}$ and $\mathbf{v}$ such as the ending point of $\mathbf{u}$ is the starting point of $\mathbf{v}$, or the reverse. Without loss of generality, let us consider the former case. We write, after having "moved" the vectors:

- A the starting point of $\mathbf{u}$.
- B the ending point of $\mathbf{u}$ and starting point of $\mathbf{v}$.
- C the ending point of $\mathbf{v}$.

We obtain the triangle ABC in Fig. 4.3 where $\mathbf{u} = \mathbf{AB}$ and $\mathbf{v} = \mathbf{BC}$. Then, Chasles' relation states that the sum of the two vectors is obtained by connecting the starting point of $\mathbf{u}$ to the ending point of $\mathbf{v}$:

$$\mathbf{AB} + \mathbf{BC} = \mathbf{AC} \tag{4.3}$$

which gives $\mathbf{w} = \mathbf{AC}$.

> The use of points A, B and C is just a mere convenience here, keep in mind that the vectors do not have a fixed starting and ending points, they can be located anywhere as long as they keep the same magnitude, direction and orientation.

The opposite of a vector $\mathbf{u}$ is the vector $-\mathbf{u}$ of same magnitude and direction but opposite orientation such as $\mathbf{u} + (-\mathbf{u}) \hat{=} \mathbf{u} - \mathbf{u} = \mathbf{0}$.

The addition between two vectors is associative:

$$\mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w} \tag{4.4}$$

and commutative:

$$\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u} \tag{4.5}$$

> A vector and a scalar cannot be added together.

Multiplication by a scalar

Let $\mathbf{u}$ be a vector and $\alpha \in \mathbb{R}$ a scalar. Then, we can multiply them together as we informally did in section 4.1.1. It yields a vector $\alpha\,\mathbf{u}$ with a magnitude $|\alpha| \cdot |\,\mathbf{u}|$, the same direction as $\mathbf{u}$ (i.e. $\alpha\,\mathbf{u} /\!/ \mathbf{u}$) and the same orientation (resp. the opposite orientation) is $\alpha > 0$ (resp. if $\alpha < 0$). This operation can be understood as loading (resp. unloading) a mine cart if $|\alpha| > 1$ (resp. $|\alpha| < 1$) through the metaphor presented in section 4.1.1 since this operation modifies the magnitude of the vector (and its orientation if $\alpha < 0$).

> It is correct though inelegant to write the vector first and the scalar after, $\mathbf{u}\alpha$.

Dot product

The dot product $\cdot$ between two vectors $\mathbf{u}$ and $\mathbf{v}$ gives a scalar and is defined by:

$$\boxed{\mathbf{u} \cdot \mathbf{v} = |\,\mathbf{u}| \cdot |\,\mathbf{v}| \cdot \cos\theta} \tag{4.6}$$

where θ is the angle between the two vectors, with $\theta < \pi$ (Fig. **??**. Also, notice that the dot $\cdot$ in the left hand side stands for the dot product between two vectors while the dot $\cdot$ in the right hand side simply represents the product between scalars. Since cosine is an even function, the direction of the angle (i.e. θ or $-\theta$) does not matter.

There is an important geometrical interpretation of the dot product which will play a key-role in the definition of the component of a vector with respect to another in section 4.1.3 on basis, and on the definition of the cosine in section 4.2 on trigonometry. To find the orthogonal projection of a vector $\mathbf{u}$ on a vector $\mathbf{v}$, we place the two vectors such as their starting points coincide in a point A (Fig. 4.5). Let us note $\mathbf{u} = \mathbf{AB}$ and $\mathbf{v} = \mathbf{AC}$, and $\mathcal{D}$ the straight line defined by the direction of $\mathbf{v}$. Then, we trace the straight line $\mathcal{D}'$ perpendicular to $\mathcal{D}$ and passing by B. The point $H = \mathcal{D} \cap \mathcal{D}'$ is called the orthogonal projection of B onto $(\mathcal{D})$. The dot product between $\mathbf{u}$ and $\mathbf{v}$ is the length AH, counted positively (resp. negatively) if $|\theta| < \pi/2$ (resp. if $|\theta| > \pi/2$).

In agreement with this geometric interpretation, the dot product enables us to determine whether two vectors are orthogonal (a.k.a. normal), that is to say whether the lines carrying these vectors are perpendicular:

$$\mathbf{u} \perp \mathbf{v} \iff \mathbf{u} \cdot \mathbf{v} = 0 \tag{4.7}$$

The dot product between two vectors is commutative:

$$\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u} \tag{4.8}$$

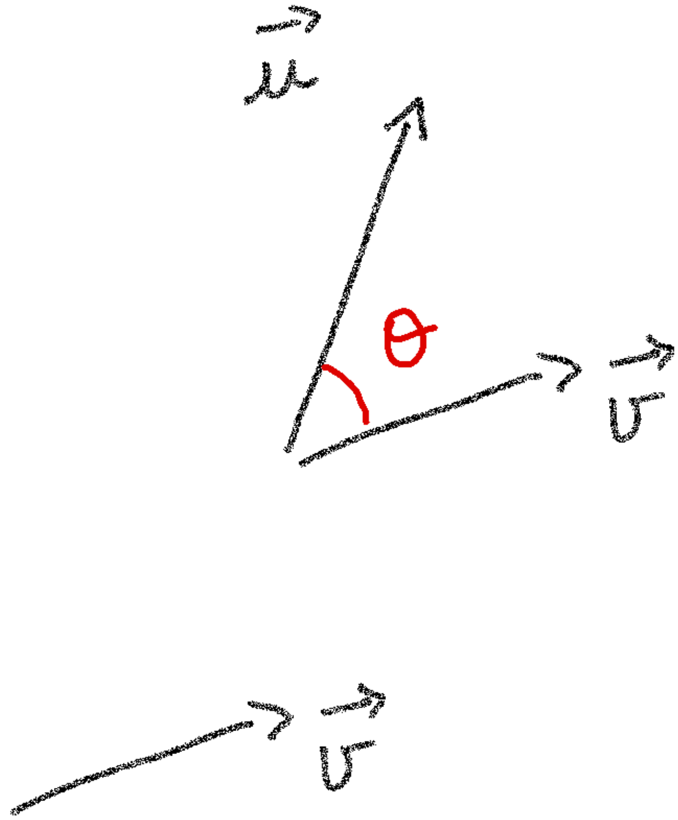

FIGURE 4.4: The angle θ between two vectors, which intervenes in the dot product, can be found by shifting one vector's origin (here, the origin of $\mathbf{v}$) such as their origins coincide.

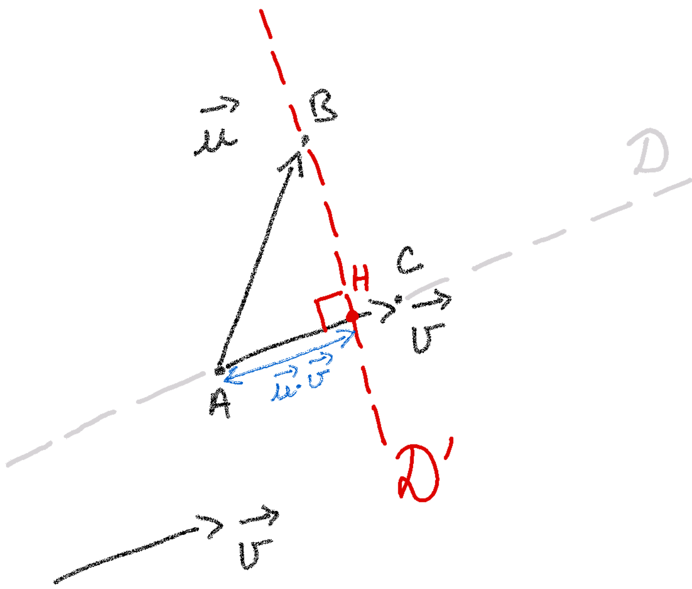

FIGURE 4.5: The orthogonal projection of a point B on an axis $\mathcal{D}$ associated to a vector $\mathbf{v}$ is the point of intersection between $\mathcal{D}$ and the line $\mathcal{D}'$, orthogonal to $\mathcal{D}$ and passing by B.

Prove that the geometric interpretation above gives the same result if we consider the projection of $\mathbf{v}$ on $\mathbf{u}$ as if we consider the projection of $\mathbf{u}$ on $\mathbf{v}$.

The dot product between two vectors is distributive:

$$\mathbf{u} \cdot (\mathbf{v} + \mathbf{w}) = \mathbf{u} \cdot \mathbf{v} + \mathbf{u} \cdot \mathbf{w} \tag{4.9}$$

The dot product provides an important way to determine the magnitude of a vector when we have an expression of $\mathbf{u}$ as a function of other vectors, from $|u| = \sqrt{\mathbf{u} \cdot \mathbf{u}}$ (as illustrated by the exercises below), or the angle between two vectors, from equation (4.6) (as illustrated by the exercises in section 4.1.3).

Triangles

1. Let ABC be a right triangle of hypotenuse BC. Demonstrate the Pythagorean theorem from the chain rule (4.3) and a dot product.
2. Let ABC be a triangle. Demonstrate the Al-Kāshī theorem (a.k.a. law of cosines):
$$|BC|^2 = |AB|^2 + |AC|^2 - 2|AB| \cdot |AC| \cdot \cos\theta \tag{4.10}$$
with θ the angle between $\mathbf{AB}$ and $\mathbf{AC}$.

Diffusion

Assume that the incoherent motion of a drunk person can be described as successive increments of a vector $\mathbf{r_i}$, where i is the index of the movement. Each step is isotropic (i.e. all the directions are equally likely), has the same magnitude (i.e. $\forall i |\mathbf{r_i}| = r$ constant) and is independent of the previous one.

1. After N iterations, what is the vector position $\mathbf{R}$ of the person?
2. After N iterations, with $N \gg 1$, what is the distance $R = |\mathbf{R}|$ traveled by the drunk person from their starting point (as a function of r and N)?
3. Verify these results numerically.

> Isotropic particle diffusion is described through the same model.

Find the angle between the face diagonals of a cube (Fig. 4.6).

Vector product

The vector (a.k.a. cross) product $\wedge$ between two vectors $\mathbf{u}$ and $\mathbf{v}$ gives a vector and is defined by:

$$\boxed{\mathbf{u} \wedge \mathbf{v} = |\mathbf{u}| \cdot |\mathbf{v}| \cdot |\sin\theta| \cdot \mathbf{w}} \tag{4.11}$$

where θ is still the angle between the two vectors $\mathbf{u}$ and $\mathbf{v}$, and $\mathbf{w}$ is the unit vector obtained by applying the right-hand rule: your thumb stands for $\mathbf{u}$, your index for $\mathbf{v}$ and your middle finger gives you $\mathbf{w}$ (Fig. 4.7). Also, the vector product enables us to determine whether two

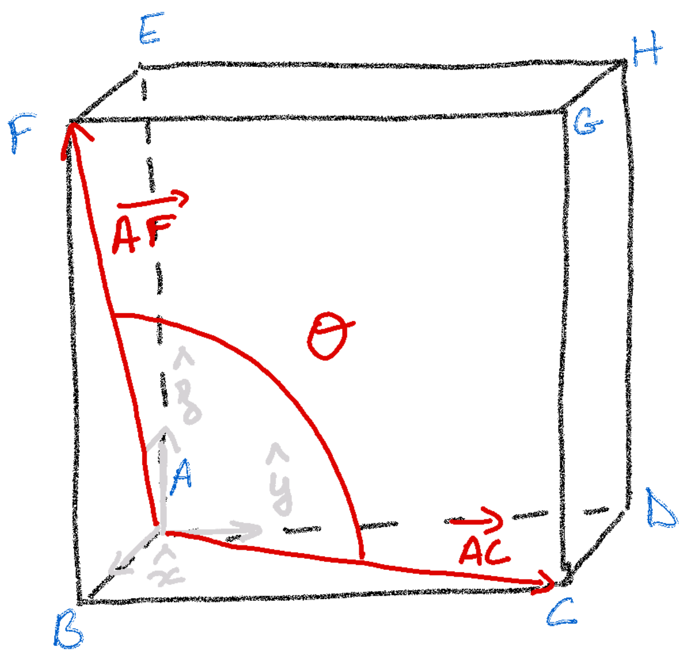

FIGURE 4.6: What is the angle θ between the diagonals of two faces of a cube?

vectors are co-linear, that is to say whether the lines carrying these vectors are parallel:

$$\mathbf{u} /\!/ \mathbf{v} \iff \mathbf{u} \wedge \mathbf{v} = \mathbf{0} \tag{4.12}$$

Is the vector product associative?

The vector product is not commutative:

$$\mathbf{u} \wedge \mathbf{v} = -\,\mathbf{v} \wedge \mathbf{u} \tag{4.13}$$

since the right-hand rule gives the opposite direction when we permute the index and the thumb[2]. We sometimes say that the vector product is anti-commutative.

The vector product is distributive:

$$\mathbf{u} \wedge (\mathbf{v} + \mathbf{w}) = \mathbf{u} \wedge \mathbf{v} + \mathbf{u} \wedge \mathbf{w} \tag{4.14}$$

> There is also another type of product between two vectors which yields a matrix, the outer product, which is of prime importance when we want to write the conservation equations in a universal form called the conservative form, a key-prerequisite to do numerical hydrodynamics.

Properties

When we compute the scalar triple product $\mathbf{u} \cdot (\mathbf{v} \wedge \mathbf{w})$ (which gives a scalar) and the vector triple product $\mathbf{u} \wedge (\mathbf{v} \wedge \mathbf{w})$ (which gives a vector), there are several properties and permutation rules to know (circular permutation, swapping, Lagrange's formula...).

[2]Please, do not try this *literally*.

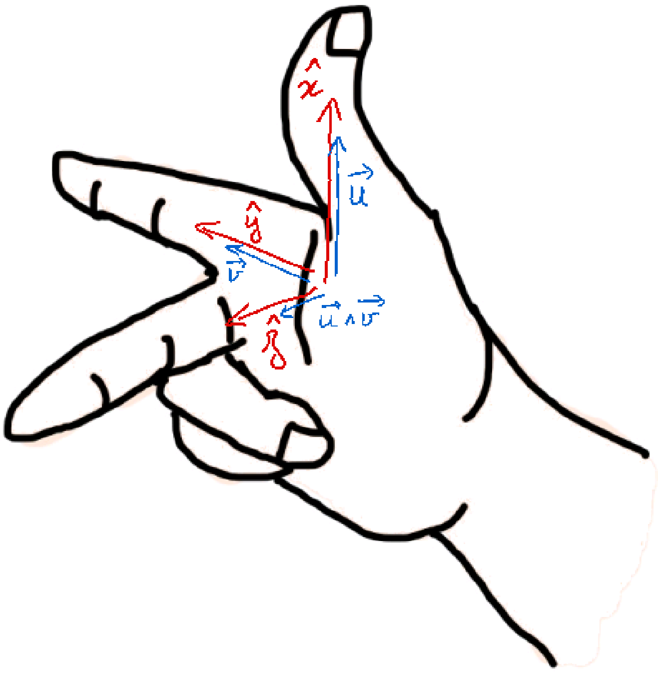

FIGURE 4.7: Right-hand convention, where the basis $(\hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}})$ is direct if and only if its order is the same as the order of the thumb, index and middle finger of the right hand. The direction and orientation of the vector product between $\mathbf{u}$ and $\mathbf{v}$ is found by using the thumb for $\mathbf{u}$ and the index for $\mathbf{v}$.

Show the following equalities:

$$\mathbf{u} \cdot (\mathbf{v} \wedge \mathbf{w}) = \mathbf{v} \cdot (\mathbf{w} \wedge \mathbf{u}) = \mathbf{w} \cdot (\mathbf{u} \wedge \mathbf{b}) \tag{4.15}$$

$$\mathbf{u} \wedge (\mathbf{v} \wedge \mathbf{w}) = (\mathbf{u} \cdot \mathbf{w})\, \mathbf{v} - (\mathbf{u} \cdot \mathbf{v})\, \mathbf{w} \quad \text{(Lagrange's formula)} \tag{4.16}$$

> By symmetry, $\mathbf{u} \wedge (\mathbf{v} \wedge \mathbf{w})$ could not be equal to $(\mathbf{v} \cdot \mathbf{w})\, \mathbf{u} - (\mathbf{u} \cdot \mathbf{v})\, \mathbf{w}$ or $(\mathbf{u} \cdot \mathbf{w})\, \mathbf{v} - (\mathbf{w} \cdot \mathbf{v})\, \mathbf{u}$ for instance, since $\mathbf{v}$ and $\mathbf{w}$ have the same status, different from $\mathbf{u}$.

$$(\mathbf{a} \wedge \mathbf{b}) \cdot (\mathbf{c} \wedge \mathbf{d}) = (\mathbf{a} \cdot \mathbf{c})\, (\mathbf{b} \cdot \mathbf{d}) - (\mathbf{b} \cdot \mathbf{c})\, (\mathbf{a} \cdot \mathbf{d}) \tag{4.17}$$

$$(\mathbf{a} \wedge \mathbf{b}) \wedge (\mathbf{c} \wedge \mathbf{d}) = [\mathbf{a} \cdot (\mathbf{b} \wedge \mathbf{d})]\, \mathbf{c} - [\mathbf{a} \cdot (\mathbf{b} \wedge \mathbf{c})]\, \mathbf{d} \tag{4.18}$$

A vector can be derived with respect to time (but we cannot derive a vector or a scalar with respect to a vector), but only if we specify the basis in which we perform the derivation. It is what enables you, in point mechanics, to determine the fictitious forces and accelerations which appear in a non-inertial frame of reference (see section **??**).

4.1.3 Basis

Hereafter, we define properties of orthogonal and orthonormal basis in general which are relative to the operations between vectors. For the classic Cartesian, cylindrical and spherical basis in particular, see section 4.3.

Operations between basis vectors

Let $\mathcal{B} = (\hat{\mathbf{a}}, \hat{\mathbf{b}}, \hat{\mathbf{c}})$ be a basis of three vectors. Then, we introduce the following definition:

$$\mathcal{B} \text{ is orthogonal} \iff \text{each basis vector is normal to the other two} \iff \begin{cases} \mathbf{a} \cdot \mathbf{b} = 0 \\ \mathbf{a} \cdot \mathbf{c} = 0 \\ \mathbf{b} \cdot \mathbf{c} = 0 \end{cases} \quad (4.19)$$

The vector product between basis vectors obeys a cyclic permutation rule:

$$\begin{cases} \mathbf{a} \wedge \mathbf{b} = \mathbf{c} \\ \mathbf{b} \wedge \mathbf{c} = \mathbf{a} \\ \mathbf{c} \wedge \mathbf{a} = \mathbf{b} \end{cases} \quad (4.20)$$

and since the vector product is anti-commutative (4.13):

$$\begin{cases} \mathbf{b} \wedge \mathbf{a} = -\mathbf{c} \\ \mathbf{c} \wedge \mathbf{b} = -\mathbf{a} \\ \mathbf{a} \wedge \mathbf{c} = -\mathbf{b} \end{cases} \quad (4.21)$$

Also:

$$\boxed{\mathcal{B} \text{ is orthonormal} \iff \begin{cases} \mathcal{B} \text{ is orthogonal} \\ |\,\mathbf{a}| = |\,\mathbf{b}| = |\,\mathbf{c}| = 1 \end{cases}} \quad (4.22)$$

> Unless stated otherwise, Cartesian, cylindrical and spherical basis are always orthonormal.

Cubes

1. How can you use a Cartesian basis $\mathcal{B} = (\hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}})$ to write, in a cube of side length unity:
   (a) the 6 small diagonals of the 6 faces?
   (b) the 3 big diagonals (Fig. 4.8)?
2. Deduce the length of:
   (a) the 6 small diagonals of the 6 faces.
   (b) the 3 big diagonals.

Decomposition of a vector on a basis

Until now, all the formulas that we gave where totally independent of any coordinate system (Cartesian, cylindrical or spherical). There was no orthonormal vector basis involved. Actually, it is the main strength of vectors: to some extent, they can be manipulated without specifying a basis. However, there are many operations, such as integration, which cannot be performed with vectors. It motivates us to decompose them, which does require the introduction of a basis. To properly decompose a vector, we need to project it on each vector

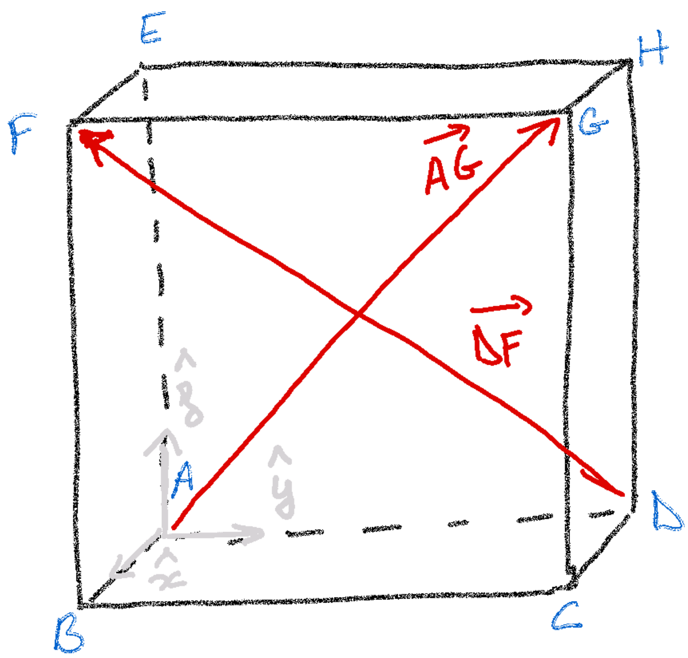

FIGURE 4.8: What is the length of the 3 big diagonals of a cube of side unity? What are the angles between them?

of the orthonormal basis we work with. In this section, we will assume that we work in 3 dimensions of space and we will call the vector of the orthonormal basis $\mathcal{B} = (\hat{\mathbf{a}}, \hat{\mathbf{b}}, \hat{\mathbf{c}})$. It could either be a Cartesian basis $(\hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}})$, a cylindrical basis $(\hat{\mathbf{r}}, \hat{\boldsymbol{\theta}}, \hat{\mathbf{z}})$, a spherical basis $(\hat{\mathbf{r}}, \hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\phi}})$ or any other orthonormal basis. The concept of projection is tightly bound to the dot product defined in equation (4.6). Indeed, the projection of a vector $\mathbf{u}$ on a vector $\mathbf{v}$ is defined as $\mathbf{u} \cdot \mathbf{v}$ and we commonly write $u_v = \mathbf{u} \cdot \mathbf{v}$.

> Since the dot product is commutative, the projection of a vector $\mathbf{u}$ on a vector $\mathbf{v}$ is also the projection of a vector $\mathbf{v}$ on a vector $\mathbf{u}$.

The decomposition of the vector $\mathbf{u}$ in the orthonormal basis $\mathcal{B}$ is:

$$\mathbf{u} = (\mathbf{u} \cdot \hat{\mathbf{a}})\, \hat{\mathbf{a}} + \left(\mathbf{u} \cdot \hat{\mathbf{b}}\right) \hat{\mathbf{b}} + (\mathbf{u} \cdot \hat{\mathbf{c}})\, \hat{\mathbf{c}} \tag{4.23}$$

$$= u_a\, \hat{\mathbf{a}} + u_b\, \hat{\mathbf{b}} + u_c\, \hat{\mathbf{c}} \tag{4.24}$$

where u_i is called the component (a.k.a. the coordinate) of $\mathbf{u}$ along the $\hat{\mathbf{i}}$-axis (where $i = a, b, c$), which can be positive or negative.

> This expression is a universal way to decompose a vector into its components on a basis. It should always be kept in mind when working with vectors, to be used in case of confusion in the middle of a vector calculus.

We can use the convenient notation:

$$\mathbf{u} = \begin{pmatrix} u_a \\ u_b \\ u_c \end{pmatrix}_{\mathcal{B}} \tag{4.25}$$

that we will use, among others, in chapter **??**. The subscript $\mathcal{B}$ in equation (4.25) is sometimes omitted when the basis we are working in is clear and unambiguous.

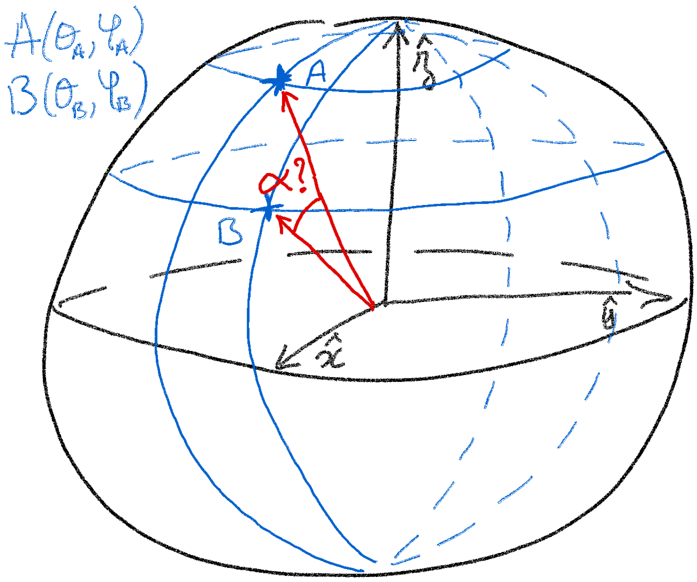

FIGURE 4.9: What is the angle between two points A and B of coordinates (θ_A, ϕ_A) and (θ_B, ϕ_B) respectively on a sphere?

1. Express the angle between the points $A(\theta_A, \phi_A)$ and $B(\theta_B, \phi_B)$ at the surface of a sphere, as a function of their angular coordinates (Fig. 4.9). Deduce the distance between points A and B.

2. Find the angle between the body diagonals of a cube.

3. Use the vector product to find the components of the unit vector $\mathbf{n}$ perpendicular to the plane shown in Fig. 4.10.

A single projection is a loss of information since a scalar contains only one number while a vector contains three, but working with the three components u_a, u_b and u_c of $\mathbf{u}$ in the basis $\mathcal{B}$ is equivalent, information-wise, to working with the vector $\mathbf{u}$ directly. However, the advantage is that we are now allowed to use operations such as integration since u_a, u_b and u_c are scalars.

> By extension, we say that we can project a vector equation $\mathbf{u} = \mathbf{v}$ (e.g. Newton's second law) on the vectors of a basis to obtain three scalar equations, easier to manipulate (but somewhat less elegant owing to their fragmented nature).

Dot product

The results on dot and vector products between basis vectors given in equations (4.19)-(4.22) enable us to define the scalar and vector products based on the components of the vectors

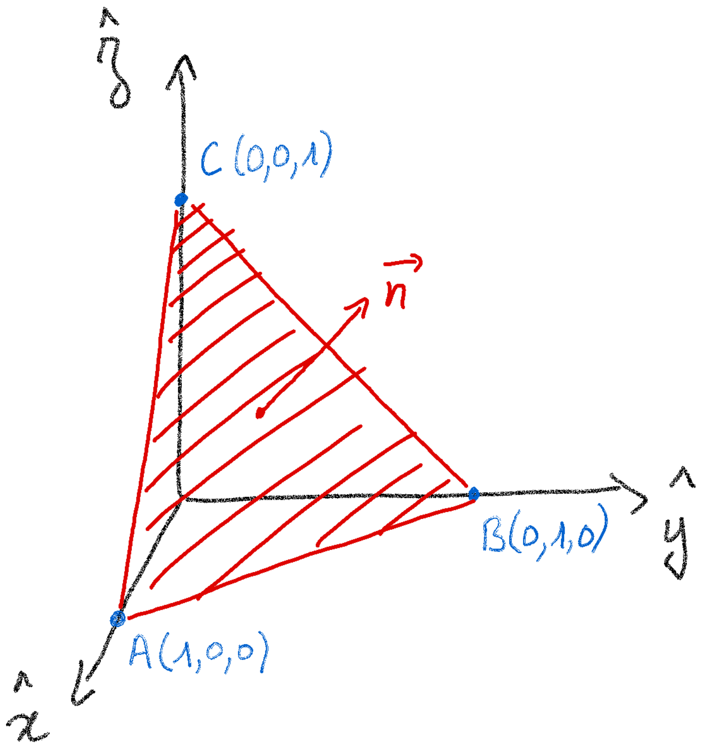

FIGURE 4.10: What are the components of the vector $\mathbf{n}$ normal to the surface passing by the points $(1, 0, 0)$, $(0, 1, 0)$ and $(0, 0, 1)$?

involved in the basis we are working in. For the scalar product, we have:

$$\begin{aligned}
\mathbf{u}\cdot\mathbf{v} &= \left(u_a\,\hat{\mathbf{a}} + u_b\,\hat{\mathbf{b}} + u_c\,\hat{\mathbf{c}}\right)\cdot\left(v_a\,\hat{\mathbf{a}} + v_b\,\hat{\mathbf{b}} + v_c\,\hat{\mathbf{c}}\right) \\
&= u_a v_a\,\hat{\mathbf{a}}\cdot\hat{\mathbf{a}} + u_a v_b\,\hat{\mathbf{a}}\cdot\hat{\mathbf{b}} + u_a v_c\,\hat{\mathbf{a}}\cdot\hat{\mathbf{c}}\,... \\
&\;...+ u_b v_a\,\hat{\mathbf{b}}\cdot\hat{\mathbf{a}} + u_b v_b\,\hat{\mathbf{b}}\cdot\hat{\mathbf{b}} + u_b v_c\,\hat{\mathbf{b}}\cdot\hat{\mathbf{c}}\,... \\
&\;...+ u_c v_a\,\hat{\mathbf{c}}\cdot\hat{\mathbf{a}} + u_c v_b\,\hat{\mathbf{c}}\cdot\hat{\mathbf{b}} + u_c v_c\,\hat{\mathbf{c}}\cdot\hat{\mathbf{c}} \\
&= u_a v_a + u_b v_b + u_c v_c \qquad (4.26)\\
&= \sum_{i=a,b,c} u_i v_i \qquad (4.27)
\end{aligned}$$

We can also use the notation (4.25) to simplify the computation:

$$\begin{array}{cc}
 & \begin{pmatrix} v_a \\ v_b \\ v_c \end{pmatrix} \\
\begin{pmatrix} u_a & u_b & u_c \end{pmatrix} & u_a v_a + u_b v_b + u_c v_c
\end{array}$$

> This is a special case of matrix product where the matrices are of dimension 1.

Vector product

For the vector product, we have:

$$
\begin{aligned}
\mathbf{u} \wedge \mathbf{v} &= \left(u_a\,\hat{\mathbf{a}} + u_b\,\hat{\mathbf{b}} + u_c\,\hat{\mathbf{c}}\right) \wedge \left(v_a\,\hat{\mathbf{a}} + v_b\,\hat{\mathbf{b}} + v_c\,\hat{\mathbf{c}}\right) \\
&= u_a v_a\,\hat{\mathbf{a}} \wedge \hat{\mathbf{a}} + u_a v_b\,\hat{\mathbf{a}} \wedge \hat{\mathbf{b}} + u_a v_c\,\hat{\mathbf{a}} \wedge \hat{\mathbf{c}} \ldots \\
&\ldots + u_b v_a\,\hat{\mathbf{b}} \wedge \hat{\mathbf{a}} + u_b v_b\,\hat{\mathbf{b}} \wedge \hat{\mathbf{b}} + u_b v_c\,\hat{\mathbf{b}} \wedge \hat{\mathbf{c}} \ldots \\
&\ldots + u_c v_a\,\hat{\mathbf{c}} \wedge \hat{\mathbf{a}} + u_c v_b\,\hat{\mathbf{c}} \wedge \hat{\mathbf{b}} + u_c v_c\,\hat{\mathbf{c}} \wedge \hat{\mathbf{c}} \\
&= u_a v_b\,\hat{\mathbf{c}} - u_a v_c\,\hat{\mathbf{b}} \ldots \\
&\ldots - u_b v_a\,\hat{\mathbf{c}} + u_b v_c\,\hat{\mathbf{a}} \ldots \\
&\ldots + u_c v_a\,\hat{\mathbf{b}} - u_c v_b\,\hat{\mathbf{a}} \qquad (4.28) \\
&= \sum_{i=a,b,c} \sum_{j=a,b,c} \sum_{k=a,b,c} \epsilon_{i,j,k} u_j v_k\,\hat{\mathbf{i}} \qquad (4.29)
\end{aligned}
$$

where we used the Levi-Civita symbol in the last equation, defined as:

$$
\epsilon_{i,j,k} = \begin{cases} 0 & \text{if } i = j \text{ or } i = k \text{ or } j = k \\ +1 & \text{if } (i,j,k) \text{ is } (a,b,c) \text{ or a circular permutation of } (a,b,c) \\ -1 & \text{if } (i,j,k) \text{ is } (a,c,b) \text{ or a circular permutation of } (a,c,b) \end{cases} \qquad (4.30)
$$

Similarly to the scalar product, we can use the notation (4.25) to simplify the computation:

$$
\begin{aligned}
\mathbf{u} \wedge \mathbf{v} &= \begin{pmatrix} u_a \\ u_b \\ u_c \end{pmatrix} \wedge \begin{pmatrix} v_a \\ v_b \\ v_c \end{pmatrix} \\
&= \begin{pmatrix} u_b v_c - u_c v_b \\ u_c v_a - u_a v_c \\ u_b v_c - u_c v_b \end{pmatrix} \qquad (4.31)
\end{aligned}
$$

4.2 Trigonometry

Trigonometry stems from a basic question: how much longer would it take to follow a circular arc rather than going straight from point A to point B? Using ropes, ancient builders realized that the ratio between the circular and half of the straight paths does not depend on the distance between the two points and decided to call it π. We provide a brief reminder of the trigonometric functions (sine, cosine and tangent) through a geometric point-of-view (see chapter 3 for the analytic reminder).

4.2.1 Angles

Let $\mathcal{A}$ be a circular arc (i.e. a portion of a circle) of length L, and R the length of its two radii (Fig. 4.11). Then, the angle θ is defined as:

$$
\boxed{\theta \hat{=} \frac{L}{R}} \qquad (4.32)
$$

and it is counted in radians. A full turn is 2π radians, such as we retrieve the perimeter of a circle $L = 2\pi R$. Therefore, any arc length L can be measured through equation (4.32). We

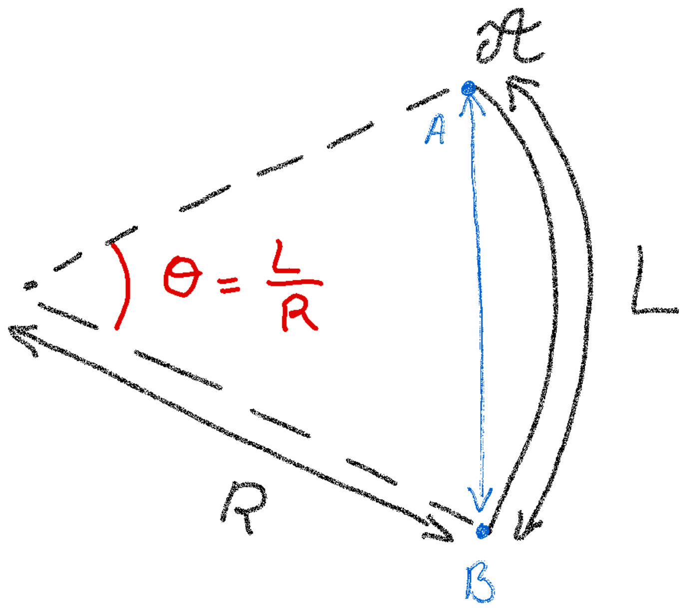

FIGURE 4.11: Definition of an angle θ from the length L of the arc $\mathcal{A}$ between two points A and B, subtended by radii of length R. The chord AB is the segment which connects A and B, not to be confused with the arc.

also retrieve the definition of π given in introduction of section 4.2: the ratio of the circular path between A and B to the $AB/2$ (i.e. the radius) is equal to π. By convention, angles counted in the anti-clock wise (resp. clock wise) direction are positive (resp. negative). The anti-clock wise direction is also known as the trigonometric direction.

Arc $\neq$ chord

The arc between two points A and B is a fraction of a circle passing by both A and B (Fig. 4.11). Its length is different from the chord between A and B, which is the straight segment joining A and B.

1. Show that there are exactly two circles passing by both A and B.
2. What is the relation between the length L of the arc and the length AB of the chord between A and B?

> On the circle of radius unity (a.k.a. unit circle, or trigonometric circle), we can locate a point by the angle θ with respect to a reference axis: we say that θ is the argument of the point on the unit circle.

4.2.2 Cosine and sine

Projections in the unit circle

In section 4.1.2, we defined the dot and vector products in equations (4.6) and (4.31) respectively. Conversely, the former can be used to define the cosine of an angle, based on the unit circle in which we drew two orthogonal axis, $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$ (Fig. 4.12). By convention, the argument θ of a point $M(\theta)$ is taken with respect to $\hat{\mathbf{x}}$. Then, we take $\mathbf{u} = \hat{\mathbf{x}}$ and $\mathbf{v}$ is the vector from the center of the circle to the point M of argument θ. Since the circle is of radius

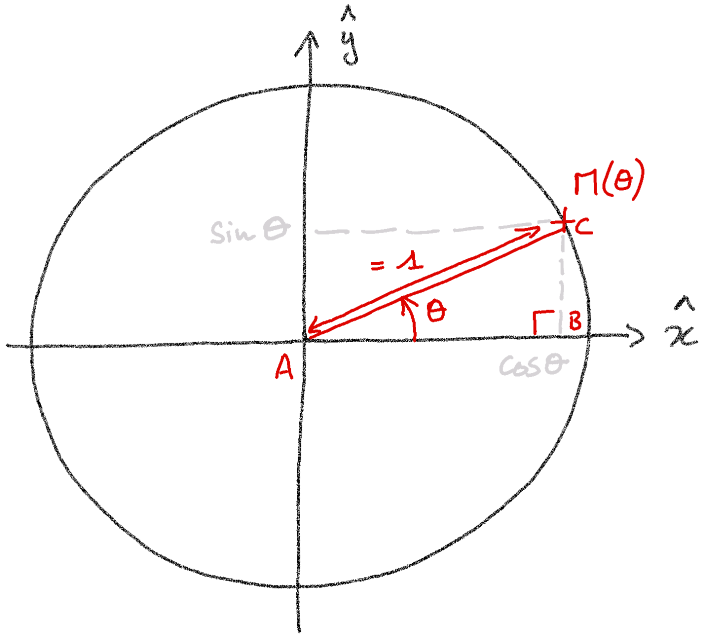

FIGURE 4.12: Trigonometric circle with a point M of argument θ. The cosine and sine of θ correspond to the abscissa and ordinate of the point M respectively.

unity, $|\mathbf{u}| = |\mathbf{v}| = 1$ and we have:

$$\cos\theta = \mathbf{u}\cdot\mathbf{v} \quad \text{with } \theta < \pi \tag{4.33}$$

$$= v_x \tag{4.34}$$

such as $\cos\theta$ is the orthogonal projection of the point M located by the vector $\mathbf{v}$ on the $\hat{\mathbf{x}}$-axis. In the same way, $\sin\theta$ can be defined as the projection of $\mathbf{v}$ on the $\hat{\mathbf{y}}$-axis. Sine and cosine are thus in the $[-1; 1]$ range and from $\theta = 0$ to $\theta = \pi/2$, cos (resp. sin) is a decreasing (resp. increasing) function.

> In a circle, all the results are unchanged by full turns (i.e. adding or subtracting an integer of times 2π): for instance, $-\pi/2 \equiv 3\pi/2[2\pi]$ and $-\pi \equiv \pi[2\pi]$, where we use the notation $\bullet \equiv \bullet[\bullet]$ for congruences (see chapter 7).

Assuming that we repeat this operation in a circle of radius R, we will obtain a projection $R\cos\theta$ (resp. $R\sin\theta$) on the $\hat{\mathbf{x}}$-axis (resp. the $\hat{\mathbf{y}}$-axis), which explains the classic relations in a right triangle:

$$\cos\theta = \frac{\text{adjacent}}{\text{hypotenuse}} \tag{4.35}$$

$$\sin\theta = \frac{\text{opposite}}{\text{hypotenuse}} \tag{4.36}$$

Fundamental relation

The fundamental relation between sine and cosine is a direct consequence of the Pythagorean theorem in the triangle ABC in Fig. 4.12, of sides 1 (the radius of the circle), $\cos\theta$ (the horizontal side of the triangle) and $\sin\theta$ (the vertical side of the triangle):

$$\boxed{\cos^2\theta + \sin^2\theta = 1} \tag{4.37}$$

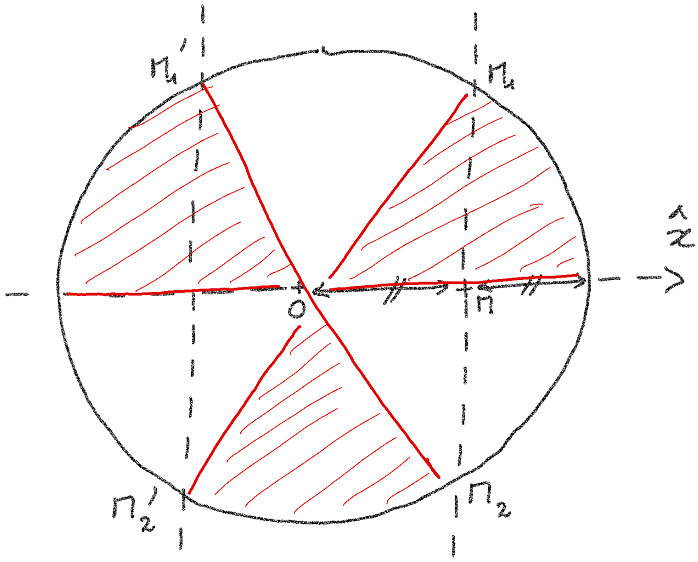

FIGURE 4.13: A classic method to cut a pie in 6 provides a recipe to remember that $\cos(\pi/3) = 1/2$.

Of pi(es) and men

We can immediately read the values of the cos and sin functions for $\theta = 0$, $\pm\pi/2$ and π as projections onto the $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$-axis, summarized in Table 4.1 (green cells). To memorize the values colored in light orange in Table 4.1, one can memorize the following method to share a pie[3] into 6 equal portions (Fig. 4.13):

1. Locate the center O of the pie and a point M half-way between the center and the edge.
2. With your knife, draw the straight line (M_1M_2) passing by M and perpendicular to the ray (OM) of the pie.
3. Repeat the operation with M', diametrically opposed to M, to obtain $(M'_1M'_2)$.
4. Draw a line perpendicular to both lines you just drew, the one passing by the center O (the $\hat{\mathbf{x}}$-axis of the circle).
5. Cut the pie along the diagonal straight lines $(M_1M'_2)$ and (M'_1M_2), and the horizontal line. You obtain 6 equal portions.
6. Enjoy your equally-divided portion of the pie.

The first step of this operation will lead to 6 portions because $1/2 = \cos(\pi/3)$. The process would be totally equivalent if you were to work with horizontal lines in the second and thirds steps, and a vertical line in the fourth step (i.e. if the pie were rotated by $\pi/2$), hence the cells in light orange in Table 4.1.

Finally, the value for $\theta = \pi/4$ (yellow cells) can be obtained from the fundamental relation (4.37) and the fact that, by symmetry, $\cos(\pi/4) = \sin(\pi/4)$.

[3] Make sure you read this part on March 14[th] (i.e. 03/14).

θ	0	$\pi/6$	$\pi/4$	$\pi/3$	$\pi/2$	π	$3\pi/2$
$\cos\theta$	1	$\sqrt{3}/2$	$\sqrt{2}/2$	1/2	0	-1	0
$\sin\theta$	0	1/2	$\sqrt{2}/2$	$\sqrt{3}/2$	1	0	-1
$\tan\theta$	0	$1/\sqrt{3}$	1	$\sqrt{3}$	$\pm\infty$	0	$\pm\infty$

TABLE 4.1: Main values of the trigonometric functions, in order of increasing θ. The second row can be deduced from the first one with the fundamental relation (**??**) (and vice versa), while the third row is deduced from the two first ones and the definition of tangent as $\tan\theta = \sin\theta/\cos\theta$. See text for color-code.

Reflections and rotations

Hereafter, we use a fiducial $\theta < \pi/4$, without loss of generality. We "see" the following properties directly from rotations and planar symmetries[4] in Fig. 4.14:

$$\cos(-\theta) = \cos\theta \quad \text{(i.e. cosine is an even function)} \tag{4.38}$$
$$\sin(-\theta) = -\sin\theta \quad \text{(i.e. sine is an odd function)} \tag{4.39}$$
$$\cos(\theta + \pi/2) = -\sin\theta \tag{4.40}$$
$$\sin(\theta + \pi/2) = -\cos\theta \tag{4.41}$$
$$\cos(\pi - \theta) = -\cos\theta \tag{4.42}$$
$$\sin(\pi - \theta +) = \sin\theta \tag{4.43}$$
$$\cos(\pi + \theta) = -\cos\theta \tag{4.44}$$
$$\sin(\pi + \theta +) = -\sin\theta \tag{4.45}$$
$$\sin(\theta - \pi/2) = -\sin(\pi/2 - \theta) = -\cos\theta \tag{4.46}$$
$$\cos(\theta - \pi/2) = \cos(\pi/2 - \theta) = \sin\theta \tag{4.47}$$

where the two last lines can be guessed from the equality between the two orange cells and the two white cells in the two first rows of Table 4.1. In Fig. 4.15, we plot the sine and cosine functions. We see that near their extrema, these functions are slowly varying. It means that starting from $\theta = 0$, we need to go beyond half-way to $\theta = \pi/2$ (i.e. above $\theta = \pi/4$) for the projection on the $\hat{\mathbf{x}}$-axis to be smaller than half of the radius (i.e. for $\cos\theta < 1/2$). Similarly, we need to go beyond half-way to $\theta = \pi/2$ (i.e. above $\theta = \pi/4$) for the projection on the $\hat{\mathbf{y}}$-axis to be larger than half of the radius (i.e. for $\sin\theta > 1/2$). It is the difference between following an arc length and a straight line (which would yield a proportional relation): trigonometric functions are not linear since $\sin(2\theta) \neq 2\sin\theta$ for instance. More specifically, we will memorize:

$$\cos(\pi/4) = \sqrt{2}/2 > 1/2 \tag{4.48}$$
$$\sin(\pi/4) = \sqrt{2}/2 > 1/2 \tag{4.49}$$

where the second can be deduced from the first one with the fundamental relation (**??**) (and vice versa). It is an alternative strategy to memorize the cells in yellow in Table 4.1.

[4]These identities must in no case be memorized but rather retrieved from mental visualization of the unit circle.

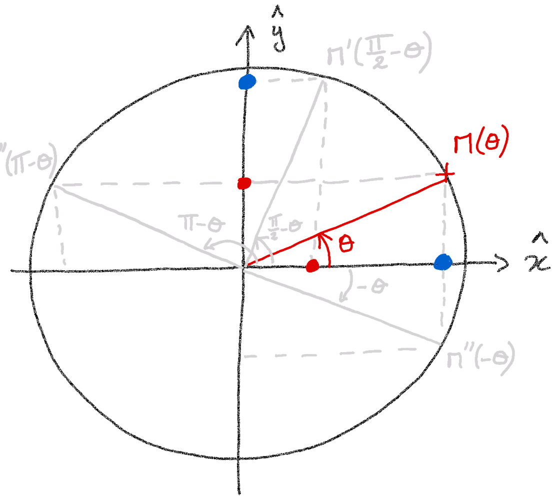

FIGURE 4.14: The symmetries in the trigonometric circle provide the first trigonometric identities given in equation (4.38) to (4.47).

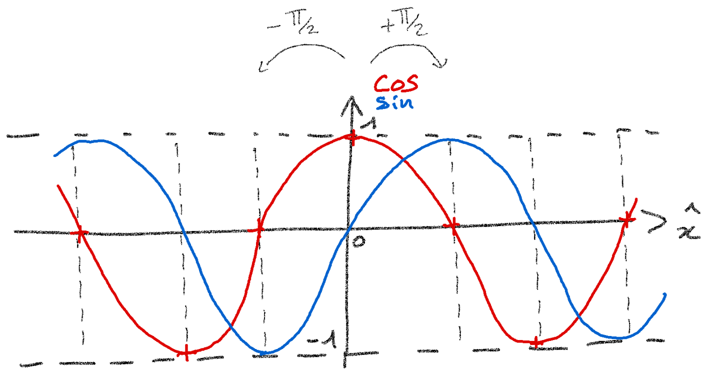

FIGURE 4.15: Curves of the cosine and sine functions.

Sum of angles

The following relations, less obvious to retrieve, must be memorized:

$$\boxed{\begin{aligned} \cos(\alpha+\beta) &= \cos\alpha\cos\beta - \sin\alpha\sin\beta \\ \sin(\alpha+\beta) &= \sin\alpha\cos\beta + \cos\alpha\sin\beta \end{aligned}} \tag{4.50}$$

from which we deduce, since cos is even and sin is odd:

$$\cos(\alpha-\beta) = \cos\alpha\cos\beta + \sin\alpha\sin\beta \tag{4.51}$$
$$\sin(\alpha-\beta) = \sin\alpha\cos\beta - \cos\alpha\sin\beta \tag{4.52}$$

From equations (4.50), we also deduce:

$$\cos(2\theta) = 2\cos^2\theta - 1 \tag{4.53}$$
$$\sin(2\theta) = 2\sin\theta\cos\theta \tag{4.54}$$

Determine the sine and cosine of $\pi/8$ and $\pi/12$.

By combining the identities for sum and differences of angles (4.50)-(4.52), we have the following sum-to-product identities:

$$\sin\alpha \pm \sin\beta = 2\sin\left(\frac{\alpha\pm\beta}{2}\right)\cos\left(\frac{\alpha\mp\beta}{2}\right) \tag{4.55}$$
$$\cos\alpha + \cos\beta = 2\cos\left(\frac{\alpha+\beta}{2}\right)\cos\left(\frac{\alpha-\beta}{2}\right) \tag{4.56}$$
$$\cos\alpha - \cos\beta = -2\sin\left(\frac{\alpha+\beta}{2}\right)\sin\left(\frac{\alpha-\beta}{2}\right) \tag{4.57}$$

4.2.3 Tangent

The tangent is the ratio of the sine by the cosine:

$$\tan\theta \hat{=} \frac{\sin\theta}{\cos\theta} \tag{4.58}$$

Graphically, $\tan\theta$ corresponds to the y-coordinate of the point of intersection between the vertical line passing by $(1, 0)$, that is to say the tangent to the circle in $\theta = 0$, and the line from the circle's center O to the point M of argument θ on the circle. Therefore, the values of tan for $\theta = 0$, $\pm\pi/2$ and π can be guessed. Furthermore, from equations (4.35), we retrieve the classic relation in a right triangle:

$$\tan\theta = \frac{\text{opposite}}{\text{adjacent}} \tag{4.59}$$

and the identities above can be combined to obtain identities for the tangent.

Show that:

$$\tan(2\theta) = \frac{2\tan\theta}{1-\tan^2\theta} \tag{4.60}$$

4.3 Coordinates

4.3.1 Basis

> polar = (r, θ) from cylindrical

Position vector $\mathbf{r}$ from the origin of the reference frame to the point we study.

4.3.2 Elementary displacement vector

How to find it

The expression of the elementary displacement vector in Cartesian, cylindrical and spherical coordinates is respectively:

$$\boxed{\mathrm{d}\boldsymbol{l} = \begin{pmatrix} \mathrm{d}x \\ \mathrm{d}y \\ \mathrm{d}z \end{pmatrix}_{\text{Cart.}} = \begin{pmatrix} \mathrm{d}r \\ r\,\mathrm{d}\theta \\ \mathrm{d}z \end{pmatrix}_{\text{cyl.}} = \begin{pmatrix} \mathrm{d}r \\ r\,\mathrm{d}\theta \\ r\sin\theta\,\mathrm{d}\phi \end{pmatrix}_{\text{sph.}}} \tag{4.61}$$

where coordinates have their classic meaning.

4.3.3 Elementary surfaces and volume

range of variation

4.4 Time derivative of a vector

Non-inertial frames

4.4.1 Velocity and acceleration vectors

v and a in different coordinate systems

4.5 Solid angles

4.6 Fields

4.6.1 Scalar fields

A scalar field is a function which associates to each point of space a scalar:

$$f : \mathbb{R}^n \to \mathbb{R} \tag{4.62}$$

where $n = 1$, 2 or 3 depending on the dimensionality of the space we are looking at. Graphically, it can be represented:

- **if $n = 2$**, as a color map (Fig. 4.16), via iso-contours (see exercise below) and/or with numbers at specific locations (Fig. 4.17).
- **if $n = 3$**, through volume rendering techniques (Fig. 4.18).

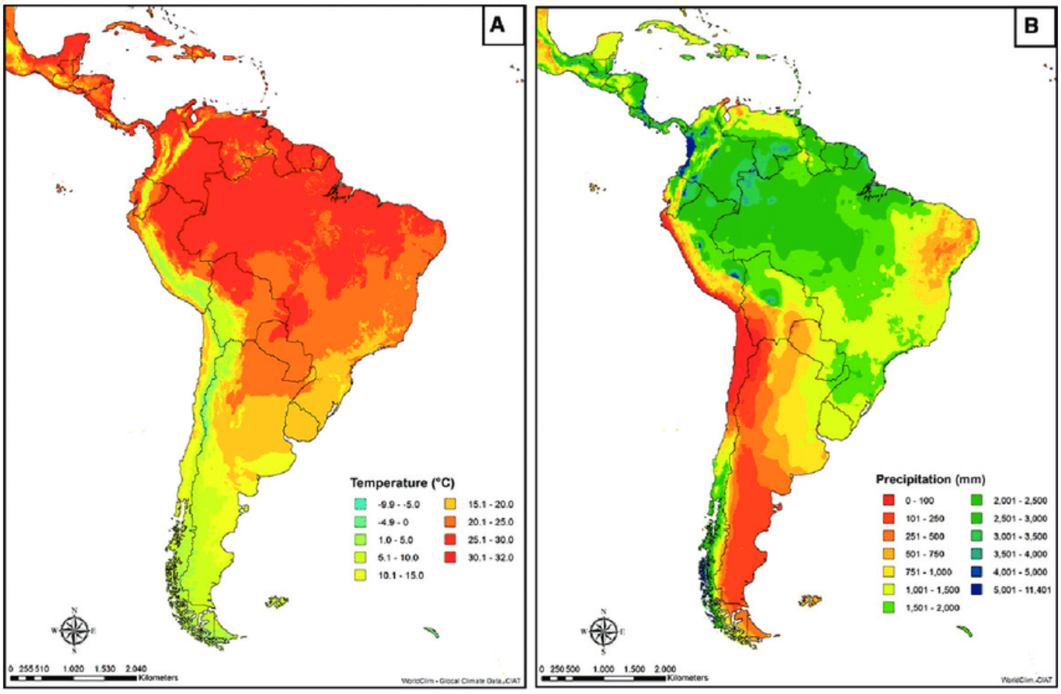

FIGURE 4.16: Color maps of the average temperatures (left panel) and precipitation (right panel) in Abya Yala.

47	57	58	55	51	45	39	34	27
53	61	69	59	53	46	41	36	29
55	65	63	57	51	43	38	37	39
57	59	55	53	48	40	39	39	35

FIGURE 4.17: Number representation of a 2D scalar field.

Iso-contours

An iso-contour in a 2D scalar field is a line along which the scalar field does not vary.

1. In the 2D scalar field in Fig. 4.17, draw approximate iso-contours from 30 to 60, by increment of 10.

2. Is it possible for two different iso-contours to cross?

> Similarly, we can define iso-surfaces in a 3D scalar field.

Topological maps

A good example of a scalar field is given by topological maps. The surface of the Earth's crust is bi-dimensional. Although its global shape is spherical, its local structure can be very irregular as illustrated by the Cordillera. Locally, a topological map works in Cartesian

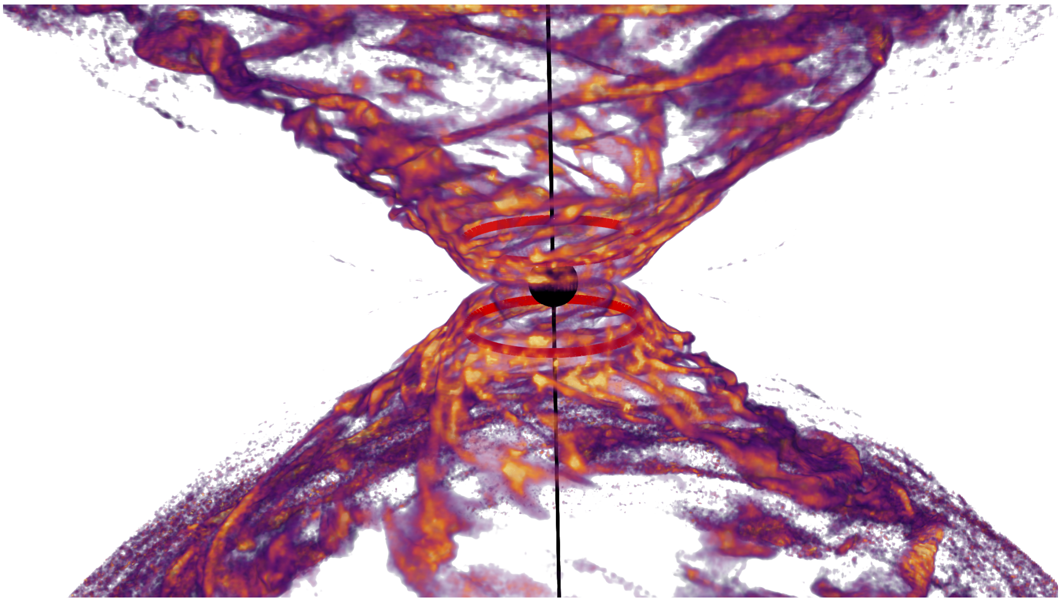

FIGURE 4.18: 3D volume rendering of the plasma density, which can be interpreted as a scalar field $\rho(x, y, z)$, around a spinning black hole whose event horizon is represented by a black sphere in the center.

coordinates and represents the altitude z of the rocky ground with respect to a reference point (e.g. the sea level). The map itself is in the $z = 0$ plane, where the coordinates are given for a $(\hat{\mathbf{x}}, \hat{\mathbf{y}})$ Cartesian basis, and the altitude is shown with iso-contours (i.e. lines along which the altitude remains constant). Doing so, we have constructed a 2D scalar field $z(x, y)$ defined as:

$$\begin{aligned} z\colon \mathbb{R}^2 &\to \mathbb{R} \\ (x, y) &\mapsto z(x, y) \end{aligned}$$

Practice on this website to identify the main structures: peaks, holes, saddle points (i.e. a mountain pass), valleys, cliffs, etc.

> On a topological map, the slope is steep (resp. gentle) when the iso-contours are close together (resp. far apart).

Topological maps

In Fig. 4.19, associate the topological maps (on the left) to the vertical cross-sections (on the right).

4.6.2 Vector fields

A vector field is a function which associates to each point of space a vector:

$$f\colon \mathbb{R}^n \to \mathbb{R}^m \tag{4.63}$$

where $n = 1$, 2 or 3, and $m = 2$ or 3. We can define the field lines of a vector field $\mathbf{u}$ as the curves tangent to the vectors in any point:

$$\mathrm{d}\boldsymbol{l} \wedge \mathbf{u} = \mathbf{0} \tag{4.64}$$

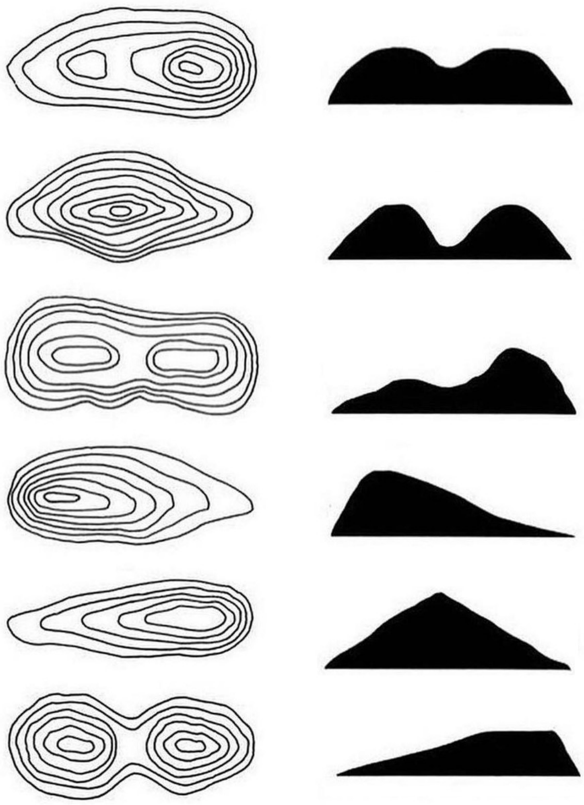

FIGURE 4.19: Topological maps (left) and vertical cross-sections (right). Which corresponds to which? For more info on how to read topological maps, see this website.

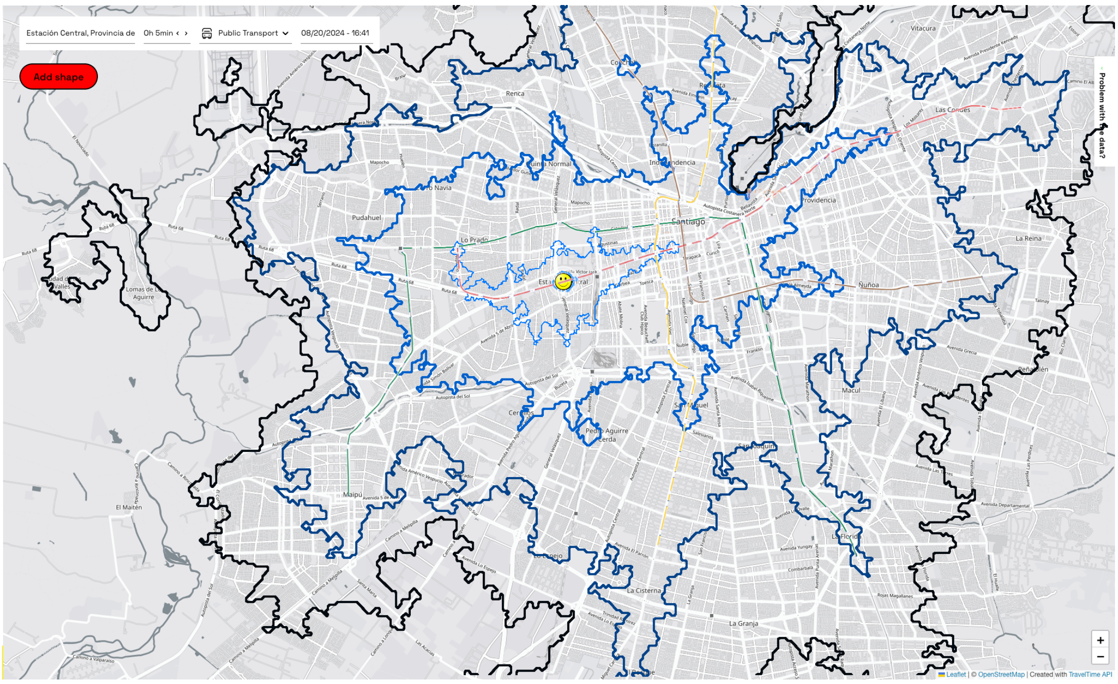

FIGURE 4.20: Isochrones are iso-contours for the time travel to a given point, which can be represented as a scalar field $\mathbf{r} \mapsto t(\mathbf{r})$. If we take Santiago's Estación central as a reference point (smiley), this figure represents isochrones from 15-minutes (light blue) to 1-hour distance (black) by public transports, by increment of 15 minutes. The main metro lines are represented with their usual color. From TravelTime's API.

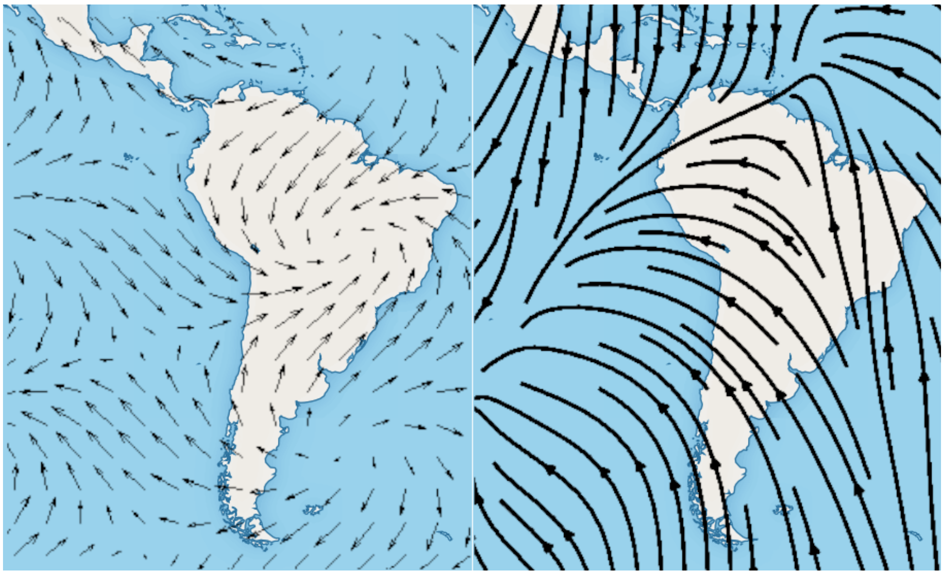

FIGURE 4.21: The horizontal winds in Abya Yala can be represented as a vector field with a quiver plot (left), where the vectors are represented at the nodes of a regularly spaced grid, or as a stream plot (right). N.B.: the vector fields represented are different in both panels.

with $\mathrm{d}\boldsymbol{l}$ the elementary displacement vector (section 4.3.2). As we move along the curve defined by the elementary displacement vector in equation (4.64), we remain co-linear to the local vectors of the vector field.

In 2D, a vector field can be represented as a collection of arrows with given magnitudes and directions, each attached to a point in space. For instance, it would be the privileged way to represent a map of the horizontal winds at the surface of the Earth (Fig. 4.21 and 4.22). Field lines are also a convenient way to visualize a vector field, in particular in 3D where a collection of arrows is cumbersome.

Dipolar magnetic field lines

The electric and magnetic fields are vector fields. In Electromagnetism, you have seen the dipolar magnetic field produced by a loop of electric current (Fig.XXX). You shown that its expression in a point located by the vector $\mathbf{r}$ with respect to the center of the loop is, in MKSA:

$$\mathbf{B}(\mathbf{r}) = \frac{\mu_0}{4\pi}\left[\frac{3\,(\mathbf{m}\cdot\mathbf{r})\,\mathbf{r}}{r^5} - \frac{\mathbf{m}}{r^3}\right] \tag{4.65}$$

where $r = |\mathbf{r}|$, μ_0 is the vacuum permeability constant and $\mathbf{m}$ is the magnetic momentum of the dipole, linked to the underlying currents which produce it. It is invariant by rotation around the axis defined by the magnetic momentum of the dipole (see section 4.6.3), which enables us to define a spherical coordinate system.

1. Give the three components of the dipolar magnetic field in the spherical basis.

2. Determine the parametric equation $r(\theta)$ of the magnetic field lines and plot them.

3. Define a cylindrical basis and give the three components of the dipolar magnetic field in this basis.

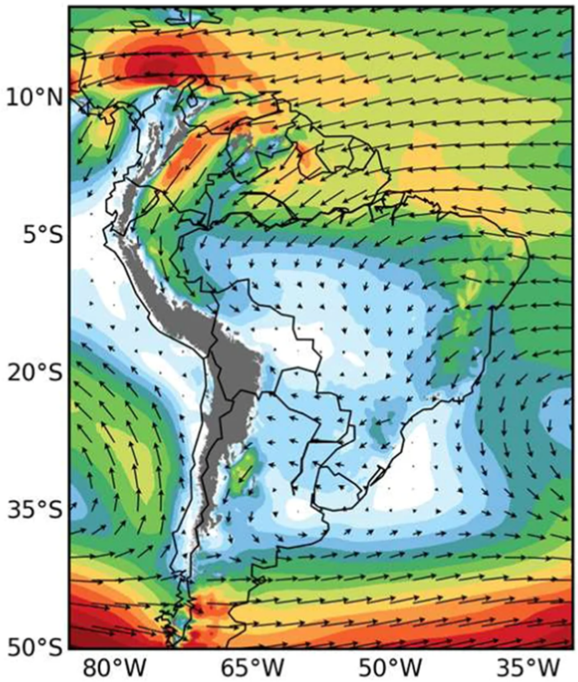

FIGURE 4.22: Map of the horizontal winds in Abya Yala, which combines both a quiver plot, to represent the vector field of the velocity, and a color map, to represent the scalar field the local wind speed (i.e. the magnitude of the velocity).

> Many celestial bodies have an essentially dipolar magnetic field: some planets of the Solar system (e.g. the Earth), neutron stars, white dwarfs...

> In sections 4.6.1 and 4.6.2, we used the notion of field in a slightly ambivalent way. For instance, mathematically speaking, the velocity of a point-mass and a velocity field are two very different objects: the former is a vector, the latter is a field, that is to say a function which associates to each point of space a vector. However, we commonly extend the notions which apply to vectors (e.g. the operations) to vector fields. For instance, we will talk about the magnitude of the velocity field but what we actually mean is the magnitude of the vectors mapped onto by the velocity field. Similarly, the temperature of a thermodynamic system is not of the same nature as a temperature field which describes the temperature in any point of space simultaneously. The distinction between both approaches, which is fundamental, will be made clear in section **??**.

4.6.3 Invariances and symmetries

Invariances

Be it scalar or vector, a field is said to be invariant with respect to a transformation if and only if it remains unchanged by this transformation. The invariances set the geometry of the problem. Indeed, in order to simplify the computation, we always determine the invariances of the field we work on first and then, we define a basis where the invariances can be expressed in a simple way.

For instance, let $T(\mathbf{r})$ be a scalar field in 3D space. Assume that it is invariant by translation along an axis $\hat{\mathbf{a}}$. Then, we construct a basis $\mathcal{B} = (\hat{\mathbf{a}}, \hat{\mathbf{b}}, \hat{\mathbf{c}})$ such as the scalar field T depends only on the coordinates along the $\hat{\mathbf{b}}$ and $\hat{\mathbf{c}}$-axis: $T(b, c)$. The situation is analogous for a vector field.

> The coordinates along the $\hat{\theta}$-axis in cylindrical and the $\hat{\theta}$ and $\hat{\phi}$-axis in spherical are angles, not positions. Therefore, along these axis, the counterpart of a translation is a rotation.

Let $\rho(\mathbf{r})$ be a mass density field. If we assume that it is spherically-symmetric, how can we simplify its dependences?

For the dipolar magnetic field, express the invariance of the field with respect to the axis defined by the magnetic momentum of the dipole...

1. ... using the spherical basis.
2. ... using a Cartesian basis you will define.

Symmetries

Symmetries are relevant for a vector field such as the gravitational, electric and magnetic fields. They enable to determine which component of the field can be omitted. According to the Curie principle, the consequences inherit the symmetries of the causes. It means that a gravitational field $\mathcal{G}(\mathbf{r})$, an electric field $\mathbf{E}(\mathbf{r})$ and a magnetic field $\mathbf{B}(\mathbf{r})$ will have (at least) the symmetries of the distribution of mass $\rho(\mathbf{r})$, of electric charges $\rho_e(\mathbf{r})$ and of electric currents $\mathbf{j}(\mathbf{r})$ which produced them respectively. True vectors (resp. pseudo vectors) like the gravitational or the electric field (resp. the magnetic field) are symmetric with respect to a plane of symmetry (resp. a plane of anti-symmetry) and anti-symmetric with respect to a plane of anti-symmetry (resp. a plane of symmetry). Therefore:

- In the left panel (resp. the right panel) in Fig. 4.23, the true vector $\mathbf{E}$ (resp. the pseudo vector $\mathbf{B}$) transforms into $\mathbf{E}'$ (resp. $\mathbf{B}'$).
- In the left panel (resp. the right panel) in Fig. 4.23, the pseudo vector $\mathbf{B}$ (resp. the true vector $\mathbf{E}$) transforms into $\mathbf{B}''$ (resp. $\mathbf{E}''$).

These properties have direct consequences on the direction of true and pseudo vectors on the plane of symmetry (or anti-symmetry) itself. Indeed, in Fig. 4.23, imagine bringing the points where we want to determine $\mathbf{E}$ and $\mathbf{B}$ closer and closer from the planes. The only manner to avoid any discontinuity of the components of $\mathbf{E}$ and $\mathbf{B}$ is to have the following:

- The normal component of $\mathbf{E}$ canceling on the plane of symmetry (left panel).
- The non-normal components of $\mathbf{B}$ canceling on the plane of symmetry (left panel).
- The non-normal components of $\mathbf{E}$ canceling on the plane of anti-symmetry (right panel).
- The normal component of $\mathbf{B}$ canceling on the plane of anti-symmetry (right panel).

Therefore, once a distribution of sources (e.g. ρ, ρ_e or $\mathbf{j}$) is given, the first thing to do is to look for the planes of symmetry and of anti-symmetry (Fig. 4.24 to 4.27).

> Vectors defined through a vector product (e.g. the magnetic field and the angular momentum) are pseudo vectors.

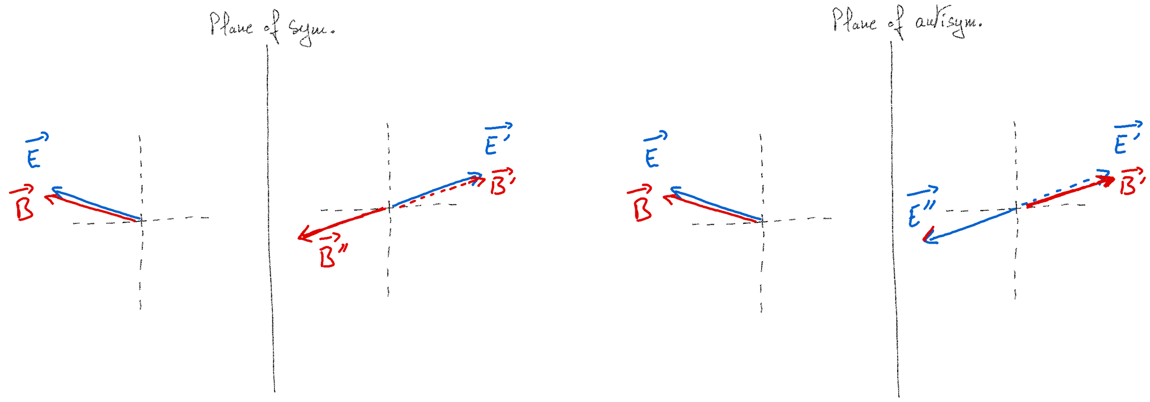

FIGURE 4.23: Transformation of a true vector **E** and of a pseudo-vector **B** across a plane of symmetry (left panel) and anti-symmetry (right panel).

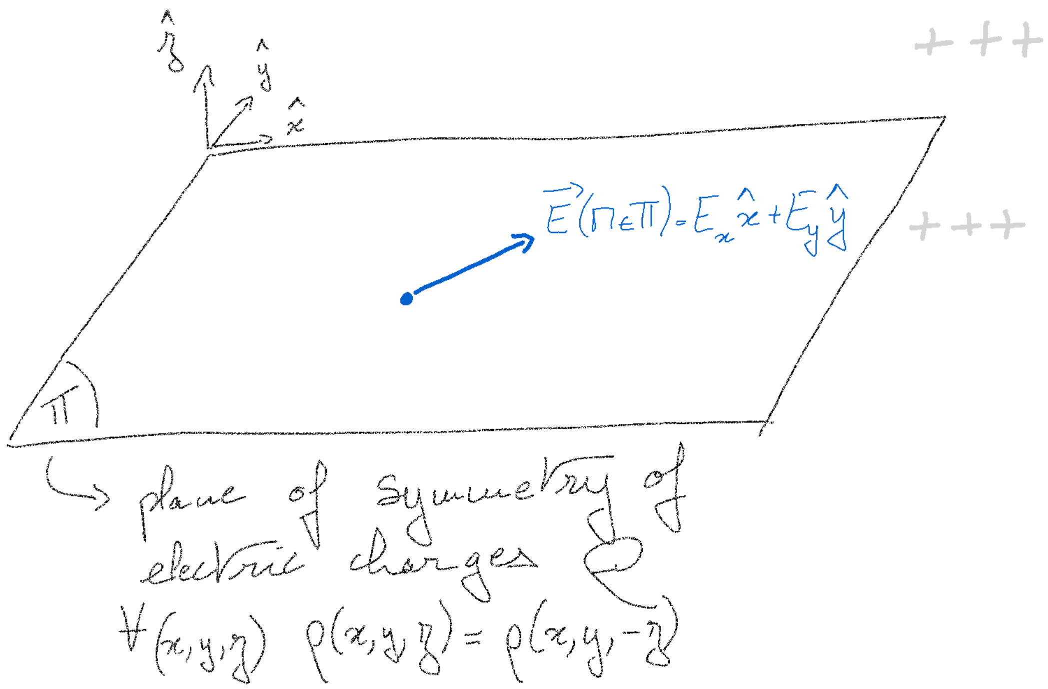

FIGURE 4.24: On a plane of symmetry Π, the normal component of a true vector **E** is null.

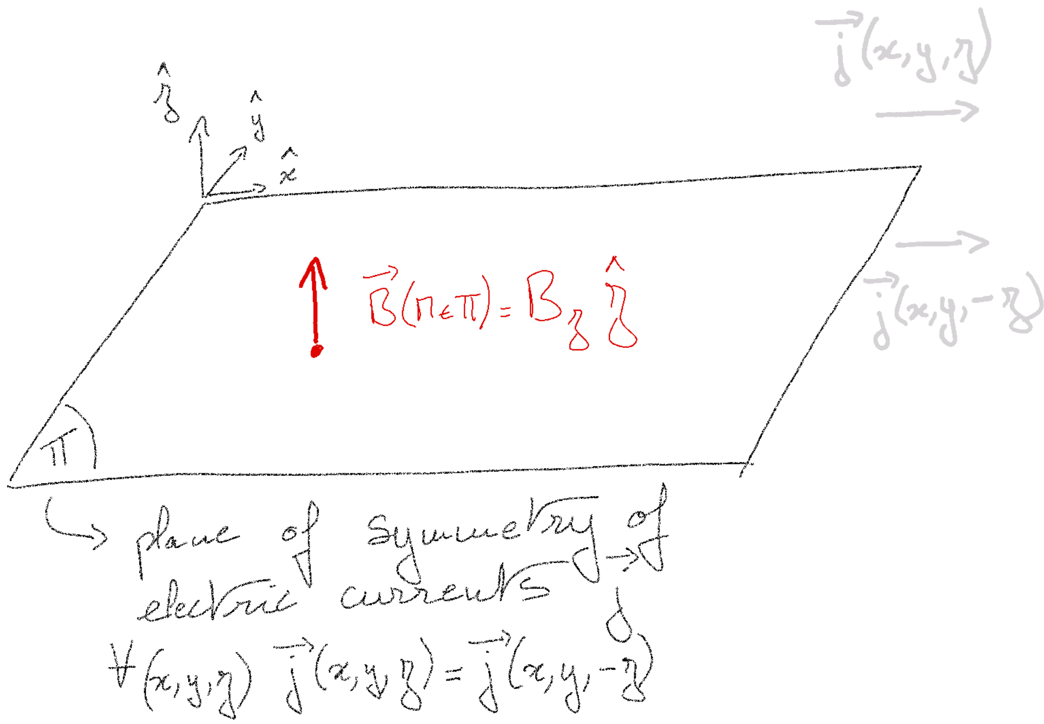

FIGURE 4.25: On a plane of symmetry Π, the only non-null component of a pseudo vector **B** is the normal one.

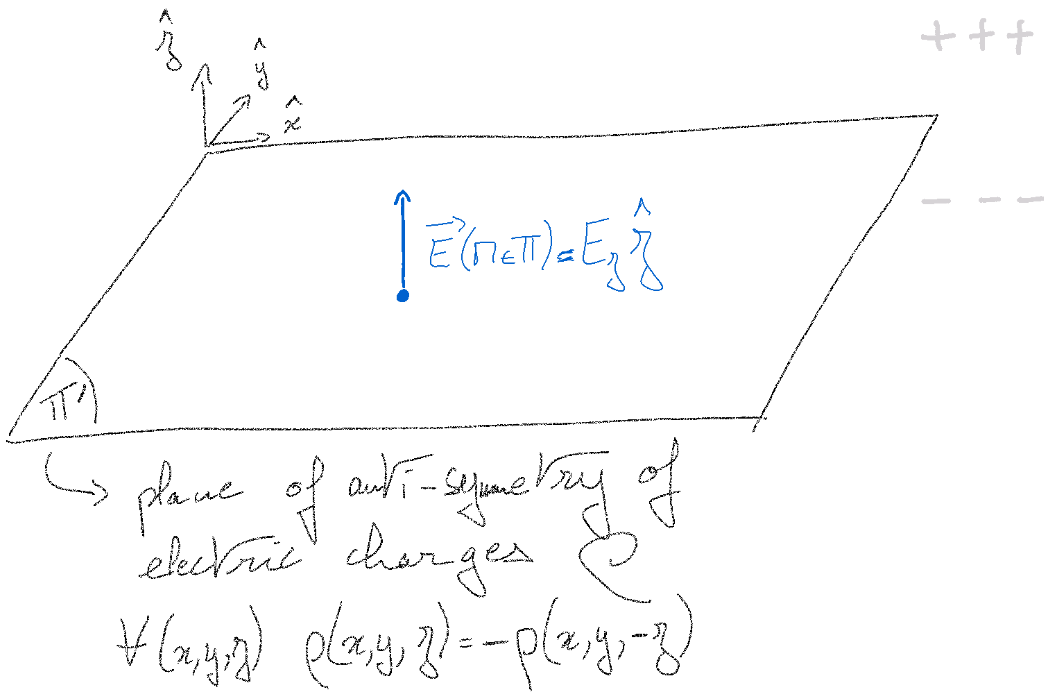

FIGURE 4.26: On a plane of anti-symmetry Π', the only non-null component of a true vector **E** is the normal one.

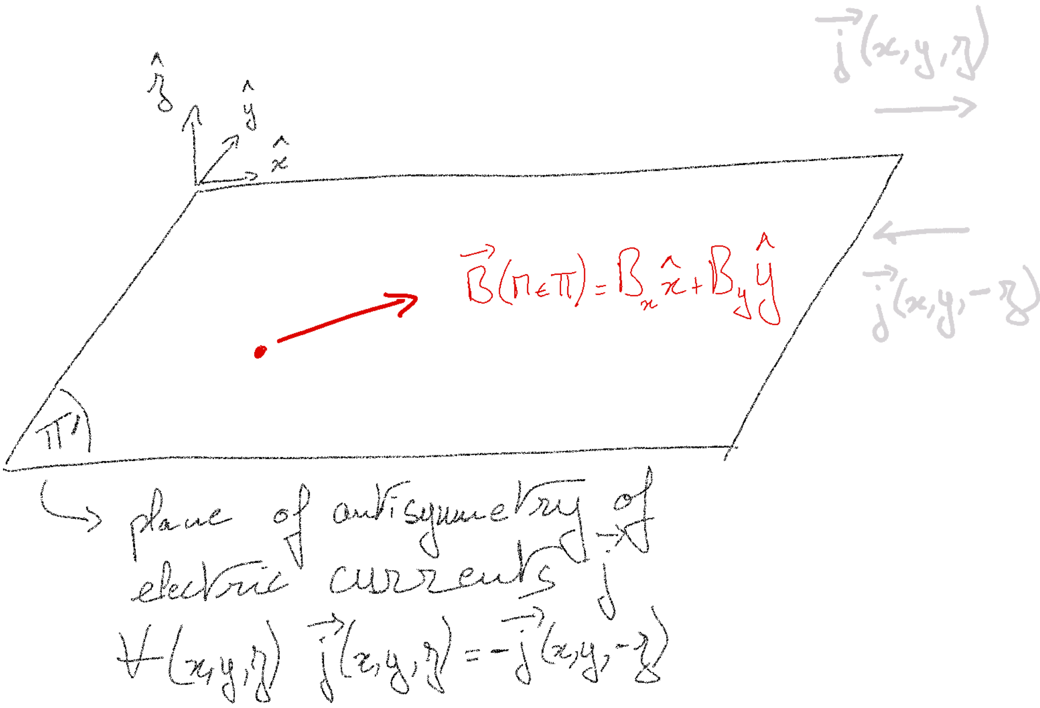

FIGURE 4.27: On a plane of anti-symmetry Π', the normal component of a pseudo vector $\mathbf{B}$ is null.

Chapter 5

Differential operators

Now that we have both reminders of differential analysis and vector calculus, we can combine them to introduce the notion of differential operators, a central tool in the study of any type of field (electric, magnetic, gravitational, velocity...). Differential operators apply to fields, but fields are function of several (space) variables. Hence, differential operators make use of the partial derivatives defined at the end of section 3.2.1.

5.1 Gradient

5.1.1 Definition

As mentioned at the end of section 3.2.1, the derivative of a function of several variables is ambiguous: with respect to which variable should we derive? The gradient $\boldsymbol{\nabla} f$ of a scalar field f is a way to estimate its derivative. The gradient is a vector field where each vector component quantifies the partial derivative of the scalar field in one of the directions given by the basis vectors. For instance, in a Cartesian basis $(\hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}})$, it is defined as:

$$\boxed{\boldsymbol{\nabla} f = \left.\frac{\partial f}{\partial x}\right|_{y,z} \hat{\mathbf{x}} + f(x) \left.\frac{\partial f}{\partial y}\right|_{x,z} \hat{\mathbf{y}} + \left.\frac{\partial f}{\partial z}\right|_{x,y} \hat{\mathbf{z}}} \tag{5.1}$$

In each component, we estimate how the scalar field f changes when one variable changes and the two others remain fix. Then, we "load" this quantity on the associated vector of the basis, and we sum the three vector together[1]. If we consider the example of the topological map $z(x, y)$ mentioned in 4.6.1, a marble left at rest on this surface will, in the uniform downwards gravitational field at the Earth surface, move in the direction of the local gradient[2] $-\boldsymbol{\nabla} f$ (if we neglect any solid friction). The division by a vector is not defined, so strictly speaking, it does not make sense, but the gradient of a scalar field f can be seen as:

$$\text{"}\ \boldsymbol{\nabla} f = \frac{\mathrm{d}f}{\mathrm{d}\boldsymbol{l}}\ \text{"} \tag{5.2}$$

where $\mathrm{d}\boldsymbol{l}$ is the infinitesimal displacement vector. Therefore, the dimension of the magnitude of a gradient is the dimension of f divided by a length. Furthermore, this formula provides a convenient way to retrieve the expression of the gradient in different coordinates: each component has, at the denominator, the component of the infinitesimal displacement vector in the corresponding direction. For instance, in spherical, the $\hat{\boldsymbol{\phi}}$ component of the gradient is given by:

$$(\boldsymbol{\nabla} f) \cdot \hat{\boldsymbol{\phi}} = \frac{1}{r \sin\theta} \frac{\partial f}{\partial \phi} \tag{5.3}$$

[1] In an analogy with the rails-and-mine-carts metaphor.

[2] In Mathematics, a derivative is positive when the function grows, so $\boldsymbol{\nabla} f$ points in the direction of fastest increase of f.

where the denominator is reminiscent of $\mathrm{d}\boldsymbol{l} \cdot \hat{\boldsymbol{\phi}} = r \sin\theta \, \mathrm{d}\phi$.

5.1.2 Properties

Although not rigorous, equation (5.2) suggests that we have:

$$\boxed{\boldsymbol{\nabla} f \cdot \mathrm{d}\boldsymbol{l} = \mathrm{d}f} \tag{5.4}$$

Combined with equation (3.91), equation (5.4) tells us that to determine the variation Δf of a 2D scalar field f between two points (x_1, y_1) and (x_2, y_2), we need to perform the scalar product of its gradient $\boldsymbol{\nabla} f$ by the infinitesimal displacement vector $\mathrm{d}\boldsymbol{l}$ whose expression depends on the coordinate systems, as reminded in equation (4.61), and then integrate:

$$\Delta f = f(x_2, y_2) - f(x_1, y_1) = \int_{f(x_1,y_1)}^{f(x_2,y_2)} \mathrm{d}f = \int_{(x_1,y_1)}^{(x_2,y_2)} \boldsymbol{\nabla} f \cdot \mathrm{d}\boldsymbol{l} \tag{5.5}$$

> Actually, equation (5.4) is the mathematical (and indirect) definition of the gradient.

The gradient enables us to extend the notion of Taylor expansion we saw in section 3.2.2 to functions of multiple variables. For instance, for a function $f\colon (x, y) \mapsto f(x, y)$, provided f is continuous in a point $M(x_0, y_0)$, we can use equation (5.5) to perform the Taylor expansion of f to 1$^{\text{st}}$-order:

$$f(x, y) = f(x_0, y_0) + \int_{(x_0,y_0)}^{(x,y)} \boldsymbol{\nabla} f \cdot \mathrm{d}\boldsymbol{l} \tag{5.6}$$

$$\sim f(x_0, y_0) + (x - x_0) \left.\frac{\partial f}{\partial x}\right|_{x_0,y_0} + (y - y_0) \left.\frac{\partial f}{\partial y}\right|_{x_0,y_0} \tag{5.7}$$

It gives the approximate expression of f near the point $M(x_0, y_0)$.

> The general expression up to the N$^{\text{th}}$-order is challenging since crossed derivatives intervene.

XXX

Compute the length of a path along a given $y(x)$ line.

XXX

Topological map of altitude. Guesstimate gradient in different points. Expression of z(x,y) is XXX. What is the expression of the gradient? We start at point X and we want to climb up to point Y. Intuitively, what is the path you want to follow? XXX On the reverse, assume that we followed the path parametrized by y(x) from X1 to X2. What is the associated variation of altitude Δz?

> Beware, in this specific exercise, the infinitesimal displacement vector is 2D: z is not a coordinate, it is merely the scalar field.

5.2 Divergence

5.2.1 Definition

The divergence $\boldsymbol{\nabla} \cdot \mathbf{u}$ of a vector field $\mathbf{u}$ is a scalar field which represents the volume density of the outward flux of $\mathbf{u}$ from an infinitesimal volume around a given point (Fig.XXX, inwards, outwards, rotation). In Cartesian coordinates, it is written:

$$\boxed{\boldsymbol{\nabla} \cdot \mathbf{u} = \begin{pmatrix} \partial_x \\ \partial_y \\ \partial_z \end{pmatrix} \cdot \begin{pmatrix} u_x \\ u_y \\ u_z \end{pmatrix}} \tag{5.8}$$

$$\boldsymbol{\nabla} \cdot \mathbf{u} = \left.\frac{\partial u_x}{\partial x}\right|_{y,z} + \left.\frac{\partial u_y}{\partial y}\right|_{x,z} + \left.\frac{\partial u_z}{\partial z}\right|_{x,y} \tag{5.9}$$

with u_x, u_y and u_z the components of the vector field $\mathbf{u}$ in the Cartesian basis $(\hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}})$. The unit of $\boldsymbol{\nabla} \cdot \mathbf{u}$ is given by the dimension of f divided by a length.

> The notation $\partial_x \,\hat{\mathbf{x}} + \partial_y \,\hat{\mathbf{y}} + \partial_z \,\hat{\mathbf{z}}$ does not mean that this is a vector. It is an operator which applies to a vector field.

We will see the physical interpretation of the divergence of the velocity field as the relative variation of fluid particles' volume in section **??**.

5.2.2 The divergence theorem

Also known as the Gauss theorem or the Green-Ostrogradsky theorem, the divergence theorem states that the integral of the divergence of a vector field $\mathbf{u}$ over a volume (V) is given by the integral of the vector field itself over the closed surface (S) surrounding the volume:

$$\boxed{\iiint_{(V)} (\boldsymbol{\nabla} \cdot \mathbf{u}) \,\mathrm{d}V = \oint\!\!\!\oint_{(S)} \mathbf{u} \cdot \mathrm{d}\,\mathbf{S}} \tag{5.10}$$

where $\mathrm{d}\,\mathbf{S}$ is the infinitesimal surface element vector pointing outwards and whose expression in Cartesian, cylindrical and spherical coordinates is given by combining two of the three components of the infinitesimal displacement vector $\mathrm{d}\boldsymbol{l}$ in equation (4.61).

5.3 Curl

5.3.1 Definition

The curl (a.k.a. rotational) $\boldsymbol{\nabla} \wedge \mathbf{u}$ of a vector field $\mathbf{u}$ is a vector field which represents its infinitesimal circulation. In Cartesian coordinates, it is written:

$$\boxed{\boldsymbol{\nabla} \wedge \mathbf{u} = \begin{pmatrix} \partial_x \\ \partial_y \\ \partial_z \end{pmatrix} \wedge \begin{pmatrix} u_x \\ u_y \\ u_z \end{pmatrix}} \tag{5.11}$$

$$\boldsymbol{\nabla} \wedge \mathbf{u} = \begin{pmatrix} \partial_y u_z - \partial_z u_y \\ \partial_z u_x - \partial_x u_z \\ \partial_x u_y - \partial_y u_x \end{pmatrix} \tag{5.12}$$

where we used the definition of the vector product (4.31), and with u_x, u_y and u_z the components of the vector field $\mathbf{u}$ in the Cartesian basis $(\hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}})$. The unit of $|\,\boldsymbol{\nabla} \wedge \mathbf{u}|$ is given by the dimension of f divided by a length. We will see the physical interpretation of the curl of the velocity field as the local rotation of fluid particles in section **??**.

> The specific expressions in cylindrical and spherical coordinates of the operators divergence $\boldsymbol{\nabla} \cdot \bullet$ and curl $\boldsymbol{\nabla} \wedge \bullet$ applied to a vector field, and of the operator Laplacian $\Delta\bullet$ are not to be known by heart. They can be found on the Wikipedia page Del in cylindrical and spherical coordinates.

5.3.2 The curl theorem

Also known as the Kelvin–Stokes theorem, the curl theorem states that the integral of the curl of a vector field $\mathbf{u}$ over an open surface (S) is given by the integral of the vector field itself over the closed contour $(\mathcal{C})$ surrounding the open surface:

$$\boxed{\iint_{(S)} (\boldsymbol{\nabla} \wedge \mathbf{u}) \cdot \mathrm{d}\,\mathbf{S} = \oint_{(\mathcal{C})} \mathbf{u} \cdot \mathrm{d}\boldsymbol{l}} \tag{5.13}$$

where $\mathrm{d}\,\mathbf{S}$ and $\mathrm{d}\boldsymbol{l}$ are consistently oriented following Ampère's right-hand grip rule.

5.4 Laplacian

The Laplacian is an operator which evaluate second-order partial derivatives. It can be applied...:

- ... either to a scalar field f, and is then defined, in Cartesian coordinates, as:

$$\Delta f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2} \tag{5.14}$$

- ... either to a vector field $\mathbf{u}$, and is then defined, in Cartesian coordinates, as:

$$\boldsymbol{\Delta}\,\mathbf{u} = \begin{pmatrix} \Delta u_x \\ \Delta u_y \\ \Delta u_z \end{pmatrix} \tag{5.15}$$

In Physics, it appears in the diffusion and in the wave equations for instance. The unit of Δf is given by the dimension of f divided by a length squared.

5.5 Vector calculus identities

The differential operators can be combined together. The following vector calculus identities are to be known by heart. The colors of the frames refer to the arrows in Fig. 5.1 which helps memorizing these identities.

Operator name	Notation	Input field	Output field
Gradient	$\boldsymbol{\nabla}\bullet$	Scalar	Vector
Divergence	$\boldsymbol{\nabla}\cdot\bullet$	Vector	Scalar
Curl	$\boldsymbol{\nabla}\wedge\bullet$	Vector	Vector
Laplacian (scalar)	$\Delta\bullet$	Scalar	Scalar
Laplacian (vector)	$\boldsymbol{\Delta}\bullet$	Vector	Vector

TABLE 5.1: Summary of the differential operators, and nature of their arguments and outputs.

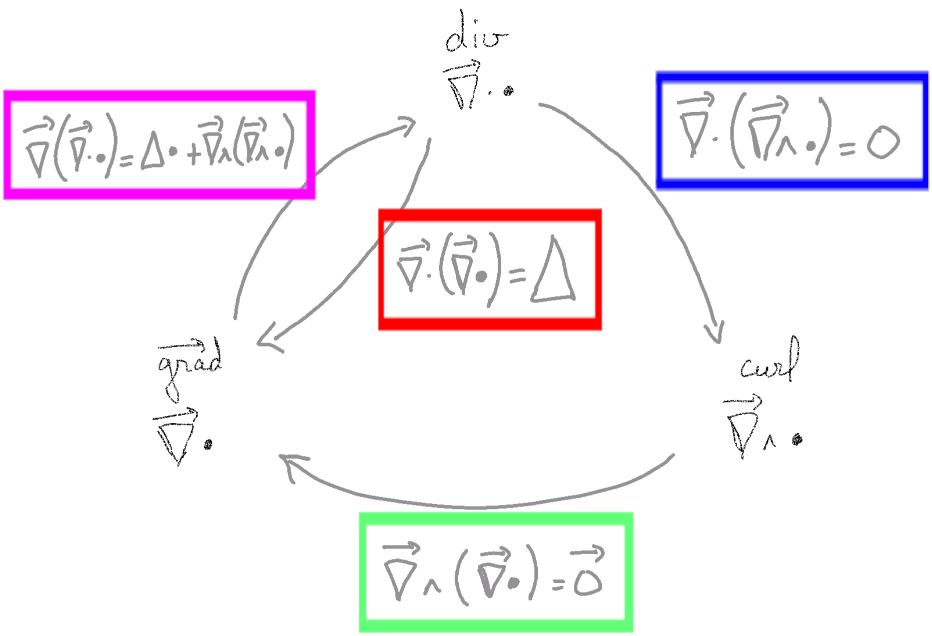

FIGURE 5.1: Mnemonic chart of all combinations between differential operators gradient, divergence and curl, given in equations (5.16) to (5.5).

- The divergence of the curl of any vector field $\mathbf{u}$ is equal to zero:

$$\boxed{\boldsymbol{\nabla} \cdot (\boldsymbol{\nabla} \wedge \mathbf{u}) = 0} \tag{5.16}$$

- The curl of the gradient of any scalar field f is equal to the null vector:

$$\boxed{\boldsymbol{\nabla} \wedge (\boldsymbol{\nabla} f) = \mathbf{0}} \tag{5.17}$$

- The divergence of the gradient of any scalar field f is the Laplacian of this scalar field:

$$\boxed{\boldsymbol{\nabla} \cdot (\boldsymbol{\nabla} f) = \Delta f} \tag{5.18}$$

- The gradient of the divergence of any vector field $\mathbf{u}$ is given by:

$$\boxed{\boldsymbol{\nabla} (\boldsymbol{\nabla} \cdot \mathbf{u}) = \boldsymbol{\Delta}\, \mathbf{u} + \boldsymbol{\nabla} \wedge (\boldsymbol{\nabla} \wedge \mathbf{u})} \tag{5.19}$$

Identities (5.16) and (5.5) have important consequences in Physics, respectively:

$$\boldsymbol{\nabla} \cdot \mathbf{u} = 0 \quad \Longleftrightarrow \quad \exists\, \mathbf{A} \text{ such as } \mathbf{u} = \boldsymbol{\nabla} \wedge \mathbf{A} \tag{5.20}$$

where $\mathbf{A}$ is called a vector potential, and

$$\boldsymbol{\nabla} \wedge \mathbf{u} = \mathbf{0} \quad \Longleftrightarrow \quad \exists V \text{ such as } \mathbf{u} = \boldsymbol{\nabla} V \tag{5.21}$$

where V is called a scalar potential. It can be easier to work with potentials than with the initial vector field itself, in particular in the latter case where a vector field is replaced by a scalar field.

> Any vector field $\mathbf{u}$ which verifies $\boldsymbol{\nabla} \cdot \mathbf{u} = 0$ is called divergence-free and presents important topological properties that we will see in chapter **??**.

Chapter 6

Logic

Mathematical logic is the formal study of reasoning (see chapter 6).

Let A and B be two logic statements. The two fundamental logic connections are the following:

- $\boldsymbol{A \implies B}$ - We usually pronounce "A implies B" or "if A, then B". It means that B is a necessary condition for A in the sense that, in the presence of A, we will necessarily have B. Equivalently, it means that A is a sufficient condition for B in the sense that A is enough to automatically trigger B. This relation, called an implication, is not commutative because we can perfectly have B without A. For example: "It rains (A)" and "I have my umbrella (B)".

- $\boldsymbol{A \iff B}$ - We usually pronounce "A if and only if B". It means that A is a necessary and sufficient condition for B. This relation, called an equivalence, is commutative and it serves to highlight the equivalence between two statements.

We define the Boolean operators which connects two logic statements A and B:

- **Conjunction `AND`** (a.k.a. $\cdot$) - A `AND` B is true if and only if A is true and B is true.

- **Disjunction `OR`** (a.k.a. $+$) - A `OR` B is true if and only if at least one of the two is true.

- **Non-disjunction `NOR`** - A `NOR` B is true if and only if exactly one of the two is true.

> In ensemble theory, the `AND` and `OR` Boolean operators are equivalent to the intersection $\cap$ and the reunion $\cup$ (Fig.XXX).

> These operators can be materialized and combined to produce sophisticated functions which lie at the core of computer science.

Formally, the negation (a.k.a. contrary) of a statement A, written $\bar{A}$, is the statement which is true when A is false, and false when A is true. It should not be confused with the contraposition of an implication: for a given implication $A \implies B$, the contraposition is defined as $\bar{B} \implies \bar{A}$. For instance, "if I don't have my umbrella, then it does not rain" is the contraposition of "if it rains, then I have my umbrella".

These definitions are somewhat intuitive, but their combination can be challenging to prove. We formalize them by introducing the concept of truth tables which enable us to identify on one hand, the true/false values of a statement made of two sub-statements A and B connected by a logical operator, and on the other hand, all the possible combinations of true/false values of the sub-statements A and B (Table XXX).

> I will sometimes use parenthesis to order and prioritize statements, like in algebra $(a \cdot (b + c))$ or like in linguistic with comas. Statements which are the more deeply buried into parenthesis are more tightly bound together. For instance, $A \implies (B \iff C)$ does not mean the same thing as $(A \implies B) \iff C)$.

· reasonings (contraposition, reductio ad absurdum, recursive, infinite regress)
Logic grid puzzles (The return of the Obra Dinn)

Chapter 7

Arithmetic