Essential Mathematics and Physics for a working understanding of Fluid Mechanics

Damien Violeau

École des Ponts ParisTech

Table des matières

1 Essential Mathematics 4

1.1 Matrix calculus 4

1.1.1 Introduction – Plane rotations 4
1.1.2 Endomorphisms of $\mathbb{R}^n$ 8
1.1.3 Eigenvectors and eigenvalues 17

1.2 Tensor Algebra 24

1.2.1 The concept of tensors 24
1.2.2 Tensor contraction 26

1.3 Integro-differential tools 30

1.3.1 Differential operators 30
1.3.2 Manipulating operators 36
1.3.3 Integrals 38
1.3.4 Ordinary differential equations 47
1.3.5 Definition and properties 50
1.3.6 Applications 56

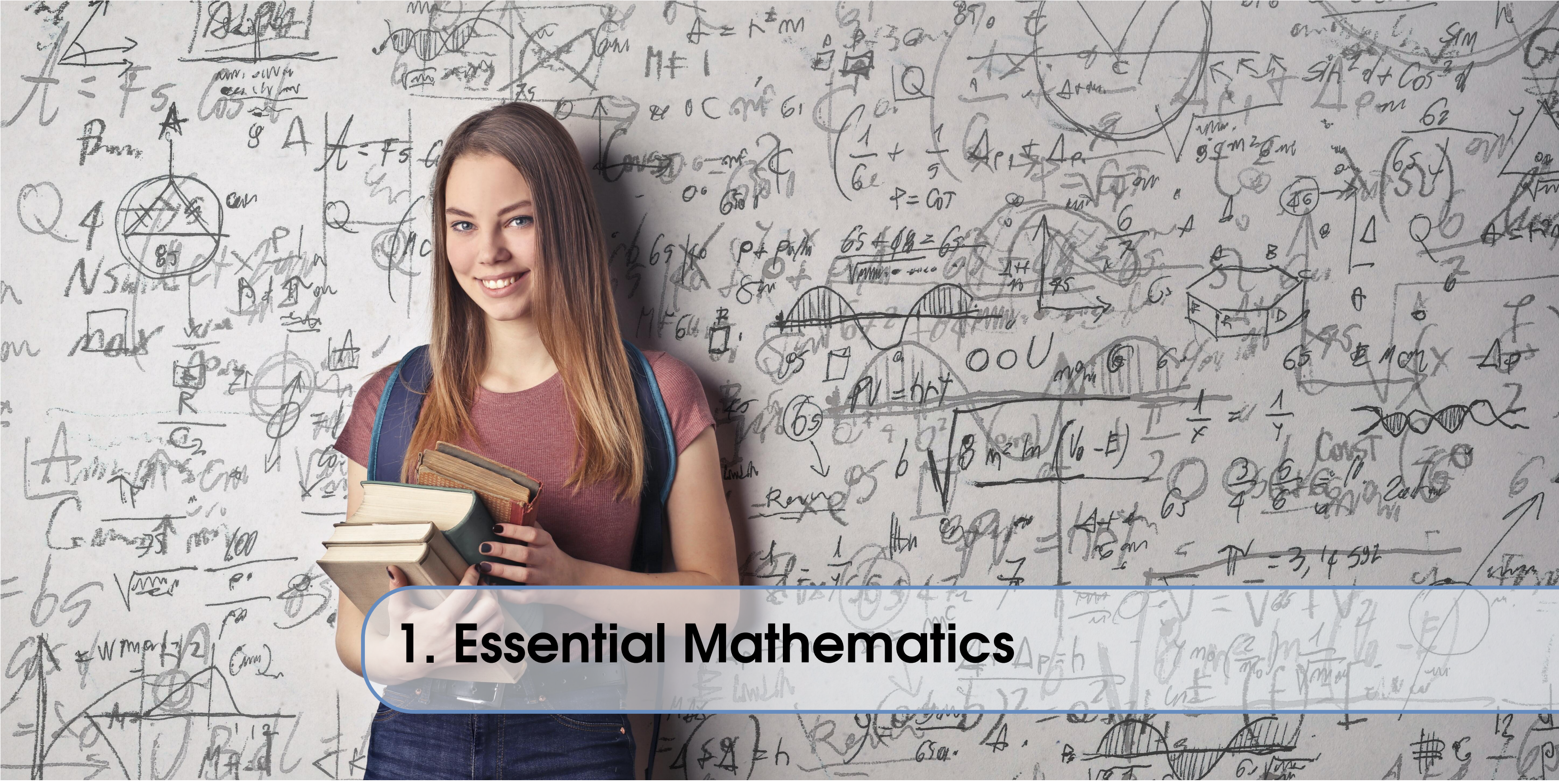
2 Essential physics 58

2.1 Mechanics of continuous media 58

2.1.1 The Cauchy stress tensor 58
2.1.2 Balance equations 62

2.2 Fluid mechanics 64

2.2.1 Navier-Stokes Equations 64
2.2.2 Two examples of analytical solutions 69
2.2.3 Propagation of surface waves 72



1. Essential Mathematics

1.1 Matrix calculus

1.1.1 Introduction – Plane rotations

Consider an orthonormal basis for the plane $\mathscr{B} \doteqdot (\underline{e}_x, \underline{e}_y)$ and a vector with the components $(x,y)^T$. To begin, we must bear in mind that $(x,y)^T$ is not the vector itself, but simply its representation in the basis $\mathscr{B}$. The vector is

$$\underline{X} = x\underline{e}_x + y\underline{e}_y. \tag{1.1}$$

> **R** In this document, the symbol $\doteqdot$ is used to designate a relationship of equality by definition, as opposed to the familiar symbol = which denotes a relationship of equality as a property which may result from pre-existing definitions. The notation $(x,y)^T$ is used to represent a plane vector; it will be explained (especially the meaning of the exponent T) in paragraph 1.1.2.

It is therefore technically inaccurate to call this a « vector $(x,y)^T$ »; we should really write « the vector $(x,y)^T_{\mathscr{B}}$ », but from an ontological perspective a vector is the entirety of its possible representations in all possible bases.

If we take another orthonormal basis $\mathscr{B}' \doteqdot (\underline{e}_{x'}, \underline{e}_{y'})$ whose vectors form an angle θ with the vectors of $\mathscr{B}$ (Figure 1.1), we can break down $\underline{e}_{x'}$ and $\underline{e}_{y'}$ into $\mathscr{B}$ just as we did for $\underline{X}$:

$$\begin{aligned} \underline{e}_{x'} &= (\cos\theta)\,\underline{e}_x + (\sin\theta)\,\underline{e}_y, \\ \underline{e}_{y'} &= -(\sin\theta)\,\underline{e}_x + (\cos\theta)\,\underline{e}_y. \end{aligned} \tag{1.2}$$

As for the vector $\underline{X}$, its representation in $\mathscr{B}'$ is written $(x',y')^T$. It can be expressed as a function of $(x,y)^T$ by substituting (1.2) in the breakdown of $\underline{X}$ over $\mathscr{B}'$:

$$\begin{aligned} \underline{X} &= x'\underline{e}_{x'} + y'\underline{e}_{y'} \\ &= x'\left((\cos\theta)\,\underline{e}_x + (\sin\theta)\,\underline{e}_y\right) + y'\left(-(\sin\theta)\,\underline{e}_x + (\cos\theta)\,\underline{e}_y\right) \\ &= \left(x'\cos\theta - y'\sin\theta\right)\underline{e}_x + \left(x'\sin\theta + y'\cos\theta\right)\underline{e}_y. \end{aligned} \tag{1.3}$$

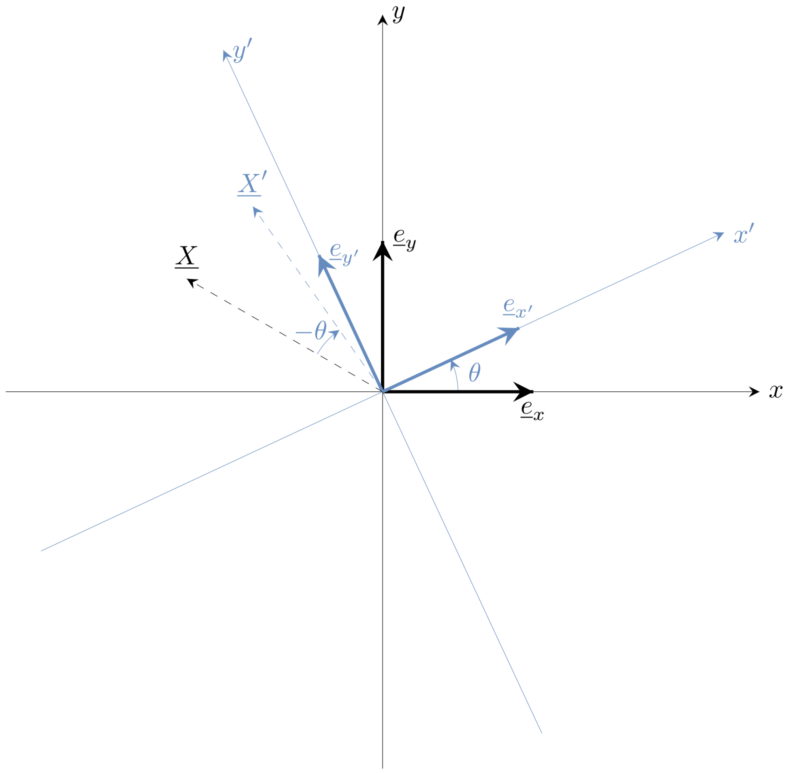


FIGURE 1.1 – Changing an orthonormal basis.

By identifying the factors which come before $\underline{e}_{x'}$ and $\underline{e}_{y'}$ in the first and last lines of this calculation, we find that

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x'\cos\theta - y'\sin\theta \\ x'\sin\theta + y'\cos\theta \end{pmatrix}, \tag{1.4}$$

which is to say that the coordinates of $\underline{X}$ in $\mathscr{B}$ and In $\mathscr{B}'$ are linked by a transformation *linéaire* ; we call this *endomorphism* because the transformed data belong to the same space as the original data (in this case, the plane $\mathbb{R}^2$). We can introduce the *matrix* for this endomorphism into the basis $\mathscr{B}$, by rewriting the above in the following symbolic form :

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix}, \tag{1.5}$$

This constitutes a table of 2×2 cells, which we call a matrix. Symbolically placing the representation $(x', y')^T$ to the right of this matrix suggests that a vector may multiply a matrix to its right as follows :

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \doteqdot \begin{pmatrix} a_{11}x + a_{12}y \\ a_{21}x + a_{22}y \end{pmatrix}. \tag{1.6}$$

For the time being, this formula only makes sense within a given basis. Later on, we shall see that it is in fact universal in nature. The endomorphism whose matrix is represented by (1.5) in $\mathscr{B}$ is a geometric transformation : a central rotation through an angle θ. We thus write

$$\underline{\underline{P}}_{\theta} \doteqdot \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}. \tag{1.7}$$

This is the matrix of the rotation in basis $\mathscr{B}$. Which gives us

$$\underline{X} = \underline{\underline{P}}_{\theta}\underline{X}'. \tag{1.8}$$

As $(x', y')^T$ represents the vector $\underline{X}$ in the basis $\mathscr{B}'$, we can determine the vector $\underline{X}' \doteqdot x'\underline{e}_x + y'\underline{e}_y$, which in $\mathscr{B}$ has the same coordinates as $\underline{X}$ in $\mathscr{B}'$ (Figure 1.1). It can be derived from $\underline{X}$ by means of inverse rotation, i.e. through the angle $-\theta$. In other words, we have

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}. \tag{1.9}$$

In formal terms, this suggests that we should define the *inverse matrix* of $\underline{\underline{P}}_{\theta}$, written $\underline{\underline{P}}_{\theta}^{-1}$ and defined in (1.9) by the reciprocal relation $\underline{X}' = \underline{\underline{P}}_{\theta}^{-1}\underline{X}$. It is clear from these definitions that $\underline{\underline{P}}_{\theta}^{-1} = \underline{\underline{P}}_{-\theta}$, which is consistent with our observations.

If we consider two rotations through angles θ_1 and θ_2 (signs not defined), their successive application gives us a rotation of angle $\theta_1 + \theta_2$ (N.B. this is generally not true in the 3rd dimension) If $\underline{X} = \underline{\underline{P}}_{\theta_1}\underline{X}'$ and $\underline{X}' = \underline{\underline{P}}_{\theta_2}\underline{X}''$, the successive application can be written as follows $\underline{X} = \underline{\underline{P}}_{\theta_1}\underline{\underline{P}}_{\theta_2}\underline{X}'' = \underline{\underline{P}}_{\theta_1+\theta_2}\underline{X}''$, which gives us $\underline{\underline{P}}_{\theta_1}\underline{\underline{P}}_{\theta_2} = \underline{\underline{P}}_{\theta_1+\theta_2}$.[1] Breaking it down into components gives us

$$\begin{pmatrix} \cos\theta_1 & -\sin\theta_1 \\ \sin\theta_1 & \cos\theta_1 \end{pmatrix} \begin{pmatrix} \cos\theta_2 & -\sin\theta_2 \\ \sin\theta_2 & \cos\theta_2 \end{pmatrix} \tag{1.10}$$
$$= \begin{pmatrix} \cos\theta_1\cos\theta_2 - \sin\theta_1\sin\theta_2 & -\sin\theta_1\cos\theta_2 - \cos\theta_1\sin\theta_2 \\ \sin\theta_1\cos\theta_2 + \cos\theta_1\sin\theta_2 & \cos\theta_1\cos\theta_2 - \sin\theta_1\sin\theta_2 \end{pmatrix}.$$

For plane endomorphisms in general, this suggests that we should define the products of their two matrices (in a given basis) in the following format :

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \doteqdot \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{pmatrix}. \tag{1.11}$$

It should also be noted that if we invert the transformation resulting from the product $\underline{\underline{P}}_{\theta_1}\underline{\underline{P}}_{\theta_2}$ then we need to invert both of the transformations, and we also need to invert the order. We can observe this by writing $\underline{\underline{P}}_{\theta_1+\theta_2}^{-1}\underline{X} = \underline{\underline{P}}_{\theta_2}^{-1}\underline{\underline{P}}_{\theta_1}^{-1}\underline{X} = \underline{X}''$, or :

$$\left(\underline{\underline{P}}_{\theta_1}\underline{\underline{P}}_{\theta_2}\right)^{-1} = \underline{\underline{P}}_{\theta_2}^{-1}\underline{\underline{P}}_{\theta_1}^{-1}. \tag{1.12}$$

1. From a terminological perspective, all of the $\underline{\underline{P}}_{\theta}$ collectively constitute an additive group.

We have not yet made use of the fact that our selected bases are orthonormal. If we apply (1.8) to $\underline{X} \doteq \underline{e}_{i'}$ we get

$$\underline{e}_{i'} = \sum_{k=1}^{n} (P_\theta)_{i'k}\, \underline{e}_k, \tag{1.13}$$

where $(P_\theta)_{ij}$ are the components of $\underline{\underline{P}}_\theta$ in $\mathscr{B}$. This means, as we saw earlier (equation (1.2)) that they are also components of vectors $\underline{e}_i$ in the basis $\mathscr{B}'$, thus $(P_\theta)_{i'k} = \underline{e}_{i'} \cdot \underline{e}_k$. This gives us

$$\begin{aligned} \underline{e}_{i'} \cdot \underline{e}_{j'} &= \sum_{k=1}^{n} (P_\theta)_{i'k}\, \underline{e}_k \cdot \sum_{\ell=1}^{n} (P_\theta)_{j'\ell}\, \underline{e}_\ell \\ &= \sum_{k=1}^{n} \sum_{\ell=1}^{n} (P_\theta)_{i'k} (P_\theta)_{j'\ell}\, \underline{e}_k \cdot \underline{e}_\ell. \end{aligned} \tag{1.14}$$

The fact that the bases are orthonormal can be written $\underline{e}_k \cdot \underline{e}_\ell = \delta_{k\ell}$ and $\underline{e}_{i'} \cdot \underline{e}_{j'} = \delta_{i'j'}$ where ù δ_{ij} is the Kronecker symbol ($\delta_{ij} = 1$ if $i = j$, or else 0). Calculation (1.14) thus gives us

$$\delta_{i'j'} = \sum_{k=1}^{n} (P_\theta)_{i'k} (P_\theta)_{j'k}. \tag{1.15}$$

At this stage, we need to introduce two more concepts/notations Firstly, the *identity matrix* 2×2 :

$$\underline{\underline{I}} \doteq \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \underline{\underline{P}}_0, \tag{1.16}$$

which is equivalent to a rotation around the origin of an angle zero, i.e. the *status quo* remains unchanged. We can then introduce the *transpose* of a 2×2 matrix :

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}^T \doteq \begin{pmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{pmatrix}, \tag{1.17}$$

which can be obtained by symmetrizing the components with respect to the main diagonal of the coefficients $\{a_{ii}\}_{i=1,2}$. The relationship (1.15) can thus be expressed

$$\underline{\underline{I}} = \underline{\underline{P}}_\theta \underline{\underline{P}}_\theta^T, \tag{1.18}$$

which gives us

$$\underline{\underline{P}}_\theta^{-1} = \underline{\underline{P}}_\theta^T. \tag{1.19}$$

We can see from (1.5) and (1.9) that this property is satisfied. It is important to note that this is only valid for rotation matrices. We call these kinds of matrices *orthogonal*. They are *transition matrices* which allow us to express the change of coordinates which occurs when changing orthonormal basis.

1.1.2 Endomorphisms of $\mathbb{R}^n$

In this section we will look at the broader application of the concepts described above. A linear transformation within a Euclidean space of arbitrary dimension n (endomorphism of $\mathbb{R}^n$) is associated with a matrix $\underline{\underline{A}}$ in a given basis :

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \doteqdot \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^{n} a_{1i}x_i \\ \vdots \\ \sum_{i=1}^{n} a_{ni}x_i \end{pmatrix}, \tag{1.20}$$

Which can also be written

$$\underline{X} = \underline{\underline{A}}\,\underline{X}'. \tag{1.21}$$

The matrix $\underline{\underline{A}}$ has an inverse matrix $\underline{\underline{A}}^{-1}$ if

$$\underline{\underline{A}}\,\underline{\underline{A}}^{-1} = \underline{\underline{A}}^{-1}\underline{\underline{A}} = \underline{\underline{I}}, \tag{1.22}$$

the identity matrix in dimension n is defined by

$$\underline{\underline{I}} \doteqdot \begin{pmatrix} 1 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 1 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 1 \end{pmatrix}. \tag{1.23}$$

Its coefficients are given by the Kronecker symbol δ_{ij}, as in dimension 2. The key difference with plane rotation matrices is that there is still no inverse matrix for a standard endomorphism. Inversion, where possible, is in fact an involution. $(\underline{\underline{A}}^{-1})^{-1} = \underline{\underline{A}}$.

■ **Exemple 1.1** Consider the following matrix in a given basis in the 3rd dimension :

$$\underline{\underline{A}} \doteqdot \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{1.24}$$

We can see that it transforms $(x,y,z)^T$ into $(x,y,0)^T$: this is a projection onto the plane (x,y). Obviously, there can be no inverse projection because all of the vectors have the same x and y components, and thus the same image irrespective of z. We will look at a more formal explanation for this later on. On a more general level, the term *projector* is used to describe an endomorphism whose matrix satisfies $\underline{\underline{A}}^2 = \underline{\underline{A}}$: they cannot be inverted. ■

Exercice 1.1 Check that the matrix (1.24) satisfies $\underline{\underline{A}}^2 = \underline{\underline{A}}$.

Solution 1.1

$$\underline{\underline{A}}^2 = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{1.25}$$

$$= \begin{pmatrix} 1\times1+0\times1+0\times0 & 1\times1+0\times1+0\times0 & 0\times1+0\times1+0\times0 \\ 1\times0+0\times0+0\times0 & 1\times0+0\times0+0\times0 & 0\times0+0\times0+0\times0 \\ 1\times0+0\times0+0\times0 & 1\times0+0\times0+0\times0 & 0\times0+0\times0+0\times0 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

■

If a matrix can be inverted, the formula (1.21) describes the inverted endomorphism in the basis in question :

$$\underline{X}' = \underline{\underline{A}}^{-1}\underline{X}. \tag{1.26}$$

The product of two matrices is given by the generic formula :

Définition 1.1 *Product of matrices.* The product of two matrices with the coefficients $\{a_{ij}\}_{i,i=1,\ldots,n}$ and $\{b_{ij}\}_{i,i=1,\ldots,n}$ in the same basis is the following matrix, also in the same basis :

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & \ddots & \vdots \\ b_{n1} & \cdots & b_{nn} \end{pmatrix} \doteq \begin{pmatrix} \sum_{i=1}^n a_{1i}b_{i1} & \cdots & \sum_{i=1}^n a_{1i}b_{in} \\ \vdots & \ddots & \vdots \\ \sum_{i=1}^n a_{ni}b_{i1} & \cdots & \sum_{i=1}^n a_{ni}b_{in} \end{pmatrix}. \tag{1.27}$$

In other words, the $ij-$th coefficient is given by $\sum_{k=1}^n a_{ik}b_{kj}$.

Let us take a closer look at some of its properties. To begin with, the product of matrix multiplication is associative, as is the composition of its functions (in this case linear functions) :

$$(\underline{\underline{A}}\,\underline{\underline{B}})\underline{\underline{C}} = \underline{\underline{A}}(\underline{\underline{B}}\,\underline{\underline{C}}) \doteq \underline{\underline{A}}\,\underline{\underline{B}}\,\underline{\underline{C}}. \tag{1.28}$$

However, it is generally not commutative :

$$\underline{\underline{A}}\,\underline{\underline{B}} \neq \underline{\underline{B}}\,\underline{\underline{A}}. \tag{1.29}$$

Exercice 1.2 Calculate the products of $\underline{\underline{A}}\,\underline{\underline{B}}$ and $\underline{\underline{B}}\,\underline{\underline{A}}$ when

$$\underline{\underline{A}} \doteqdot \begin{pmatrix} 0 & 1 & -4 \\ 1 & -1 & 2 \\ 3 & 0 & 6 \end{pmatrix} \quad ; \quad \underline{\underline{B}} \doteqdot \begin{pmatrix} -2 & -1 & 3 \\ 0 & 5 & 7 \\ -3 & 0 & 1 \end{pmatrix}. \tag{1.30}$$

Discuss these results.

Solution 1.2

$$\underline{\underline{A}}\,\underline{\underline{B}} = \begin{pmatrix} 12 & 5 & 3 \\ -8 & -6 & -2 \\ -24 & -3 & 15 \end{pmatrix}, \tag{1.31}$$

$$\underline{\underline{B}}\,\underline{\underline{A}} = \begin{pmatrix} 8 & -1 & 24 \\ 26 & -5 & 52 \\ 3 & -3 & 18 \end{pmatrix}. \tag{1.32}$$

These two matrices therefore do not commute. ■

We can thus observe that the identity matrix is the neutral element in this multiplication :

$$\forall \underline{\underline{A}}, \quad \underline{\underline{A}}\,\underline{\underline{I}} = \underline{\underline{I}}\,\underline{\underline{A}} = \underline{\underline{A}}. \tag{1.33}$$

As for the inverse of a product, it can be calculated in the same way as plane rotations, i.e. by inverting both the matrices and their order of application :

$$\left(\underline{\underline{A}}\,\underline{\underline{B}}\right)^{-1} = \underline{\underline{B}}^{-1}\underline{\underline{A}}^{-1}. \tag{1.34}$$

These last two results are very simple, as demonstrated in the exercises.

Définition 1.2 *Transposition.* Let us define the transpose of a matrix, like the 2-dimensional matrix shown above :

$$\begin{pmatrix} a_{11} & \cdots & a_{1i} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{i1} & \cdots & a_{ii} & \cdots & a_{in} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{ni} & \cdots & a_{nn} \end{pmatrix}^T \doteqdot \begin{pmatrix} a_{11} & \cdots & a_{i1} & \cdots & a_{n1} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{1i} & \cdots & a_{ii} & \cdots & a_{ni} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{1n} & \cdots & a_{in} & \cdots & a_{nn} \end{pmatrix}. \tag{1.35}$$

the $ij-$th coefficient is therefore a_{ji}.

It should thus be obvious that transposition is an exercise in involution : $\left(\underline{\underline{A}}^T\right)^T = \underline{\underline{A}}$. The transpose of a product can be written

$$\left(\underline{\underline{A}}\,\underline{\underline{B}}\right)^T = \underline{\underline{B}}^T\underline{\underline{A}}^T. \tag{1.36}$$

Exercice 1.3 Demonstrate this.

Solution 1.3 The ij−th coefficient of $\left(\underline{\underline{A}}\,\underline{\underline{B}}\right)^T$ is given by

$$\begin{aligned}\left(\underline{\underline{A}}\,\underline{\underline{B}}\right)^T_{ij} &= \left(\underline{\underline{A}}\,\underline{\underline{B}}\right)_{ji} \\ &= \sum_{k=1}^{n} a_{jk} b_{ki} \\ &= \sum_{k=1}^{n} \left(\underline{\underline{A}}\right)_{jk} \left(\underline{\underline{B}}\right)_{ki} \\ &= \sum_{k=1}^{n} \left(\underline{\underline{B}}^T\right)_{ik} \left(\underline{\underline{A}}^T\right)_{kj} \\ &= \left(\underline{\underline{B}}^T \underline{\underline{A}}^T\right)_{ij}.\end{aligned} \tag{1.37}$$

■

This definition of transposition also extends to non-square matrices (not covered here). Hence the notation used thus far to write the coordinates of vectors in lines, using (1.36) :

$$\begin{pmatrix} x_1 & \cdots & x_i & \cdots & x_n \end{pmatrix}^T \doteqdot \begin{pmatrix} x_1 \\ \vdots \\ x_i \\ \vdots \\ x_n \end{pmatrix}. \tag{1.38}$$

This enables us to write a matrix-vector product such as (1.20) while including the transpose of a vector to the left of the matrix :

$$\begin{aligned}\underline{X}^T \underline{\underline{A}} &= \left(\underline{\underline{A}}^T \underline{X}\right)^T \\ &= \begin{pmatrix} \sum_{i=1}^n x_i a_{i1} & \cdots & \sum_{i=1}^n x_i a_{in} \end{pmatrix}.\end{aligned} \tag{1.39}$$

If we change basis to an arbitrary dimension n, the transition matrix is no longer as simple as it is for planes, but the calculation (1.14)–(1.15) which gives us (1.19) remains applicable for orthonormal bases :

Théorème 1.1 For a transition matrix $\underline{\underline{P}}$ between orthonormal bases : $\underline{\underline{P}}^{-1} = \underline{\underline{P}}^T$.

However, it is also possible with non-orthonormal bases (in which case the formula shown above is not applicable). As we shall see later on, mechanics provides a framework within which we can generally work with orthonormal bases, which has major advantages.

Just as a vector has a different representation in each basis, an endomorphism has a different matrix in each basis. When we change basis, a matrix $\underline{\underline{A}}$ will change its representation and can instead be written $\underline{\underline{A}}'$. To see how the two are connected, we must assume that the endormorphism remains unchanged by the choice of basis (like a vector, its identity is absolute). For example, in the case of a plane rotation a change of basis does not affect the angle of rotation; the same applies to homothety, for example. Let us then consider the image $\underline{X}$ of $\underline{X}'$ by $\underline{\underline{A}}$ (equation (1.26)) : if we bring this vector back to its first coordinate, then transform it through endomorphism before moving onto the second coordinate, we get $\underline{\underline{P}}\,\underline{\underline{A}}\,\underline{\underline{P}}^{-1}\underline{X}'$. Expressed in the new coordinate, the vector can also be written $\underline{\underline{A}}'\underline{X}'$. This is true for any vector, and as such we can infer that

$$\underline{\underline{A}} = \underline{\underline{P}}^{-1}\underline{\underline{A}}'\underline{\underline{P}}. \tag{1.40}$$

It is here that the definition (1.20) of the matrix-vector product, as well as the definition (1.27) of the product of two matrices, reveal their full meaning, as these products are independent of basis (they are universal, as noted above). For example, we can write

$$\begin{aligned}\underline{\underline{A}}\,\underline{\underline{B}} &= \left(\underline{\underline{P}}^{-1}\underline{\underline{A}}'\underline{\underline{P}}\right)\left(\underline{\underline{P}}^{-1}\underline{\underline{B}}'\underline{\underline{P}}\right)\\ &= \underline{\underline{P}}^{-1}\underline{\underline{A}}'\underline{\underline{B}}'\underline{\underline{P}}\\ &= \underline{\underline{P}}^{-1}\left(\underline{\underline{A}}\,\underline{\underline{B}}\right)'\underline{\underline{P}}.\end{aligned} \tag{1.41}$$

As such, $\underline{\underline{A}}\,\underline{\underline{B}}$ transforms in the same manner as $\underline{\underline{A}}$ and $\underline{\underline{B}}$. We can thus talk in terms of the product (in the compositional sense) of endomorphisms, a more general and coherent concept than the product of matrices. The identity matrix also comes into its own here, as it remains unchanged when we change basis, since $\underline{\underline{P}}^{-1}\underline{\underline{I}}\,\underline{\underline{P}} = \underline{\underline{P}}^{-1}\underline{\underline{P}} = \underline{\underline{I}}$; we can thus regard it as an identity endomorphism.

Exercice 1.4 Demonstrate that for any integer k

$$\underline{\underline{A}}^k = \underline{\underline{P}}^{-1}(\underline{\underline{A}}')^k\underline{\underline{P}}. \tag{1.42}$$

Solution 1.4 By aligning k times $\underline{\underline{P}}^{-1}\underline{\underline{A}}'\underline{\underline{P}}$, the products $\underline{\underline{P}}\,\underline{\underline{P}}^{-1}$ disappear, because they are equal to the neutral element $\underline{\underline{I}}$ by (1.22). ∎

Inverting a matrix is equivalent to explicitly expressing $\underline{X}'$ with reference to $\underline{X}$ using the formula (1.26), i.e. resolving the linear system (1.21) of n equations with n unknowns $\{x_i\}_{i=1,\dots,n}$. Consider this example $n = 2$:

$$a_{11}x + a_{12}y = x', \tag{1.43}$$
$$a_{21}x + a_{22}y = y', \tag{1.44}$$

Where the unknowns are x and y, and we know the other values. If a solution exists, it can be written :

$$x = \frac{a_{22}x' - a_{12}y'}{a_{11}a_{22} - a_{12}a_{21}}, \tag{1.45}$$
$$y = \frac{-a_{21}x' + a_{11}y'}{a_{11}a_{22} - a_{12}a_{21}}. \tag{1.46}$$

This is only possible if $a_{11}a_{22} \neq a_{12}a_{21}$.[2]

2. If, on the other hand, we have $a_{11}a_{22} = a_{12}a_{21}$, the left sides of both equations (1.43) and (1.44) are proportional, so the system can only be solved if the right sides x' and y' are also proportional. There is therefore a potentially infinite number of solutions for the right side of the equation $a_{11}x + a_{12}y = y'$.

Définition 1.3 *Determinant.* The determinant of a matrix, written $|\underline{\underline{A}}|$ or $\det\underline{\underline{A}}$, can be defined as follows. In 2 dimensions it is equal to

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \doteqdot \det\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \doteqdot a_{11}a_{22} - a_{12}a_{21}. \tag{1.47}$$

The same analysis for a system comprising three equations with three unknowns would yield a solution whose denominator is the determinant of the associated 3×3 matrix in the form

$$\begin{aligned}\det\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} &\doteqdot a_{11}\det\begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix} \\ &\quad - a_{12}\det\begin{pmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{pmatrix} + a_{13}\det\begin{pmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix}.\end{aligned} \tag{1.48}$$

For an arbitrary dimension n we can define the determinant by means of recurrence, using formulae analogous to (1.48).

The determinant of a 2×2 matrix can be regarded as the area of a parallelogram defined by the two vectors which constitute its columns, while the determinant of a 3×3 matrix is equivalent to the volume of a parallelepiped defined by the three vectors which constitute its columns (see Figure 1.2)[3]. Since the columns of a matrix are, by construction, endomorphic images of the basis vectors, this parallelogram (or parallelepiped) is the image of the identity square (or cube) derived from $\{\underline{e}_i\}_{i=1,\dots,n}$.

The preceding analysis demonstrates that a 2×2 or 3×3 matrix can only be inverted if its determinant is non-null. We can thus adopt the following theorem, which stipulates this to be true in all dimensions :

Théorème 1.2 A matrix $\underline{\underline{A}}$ can be inverted if and only if $\det\underline{\underline{A}} \neq 0$.

Exercice 1.5 Calculate $\det\underline{\underline{P}}_\theta$ (defined by the equation (1.7)). Discuss this result.

Solution 1.5

$$\begin{aligned}\det\underline{\underline{P}}_\theta &= \cos^2\theta + \sin^2\theta \\ &= 1.\end{aligned} \tag{1.49}$$

This confirms that the matrices for plane rotations are capable of inversion. ■

3. Strictly speaking, these surface or volume values should be preceded by a sign depending upon the mutual orientation of the vectors.

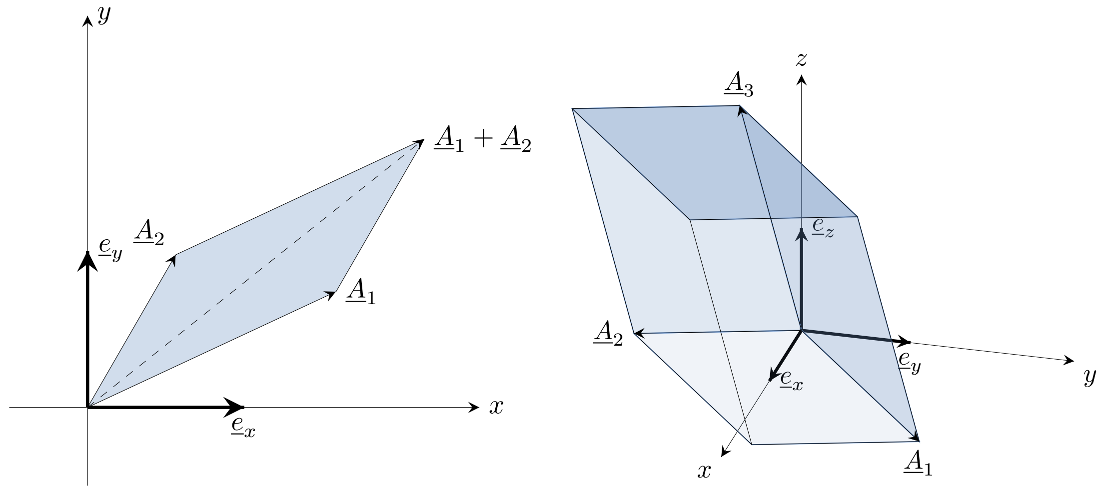


FIGURE 1.2 – Illustration of the notion of the determinant. On the left : $\det \underline{\underline{A}}$ is the oriented surface area of the parallelogram defined by the vectors $\underline{A}_1$ and $\underline{A}_2$ which form the columns of $\underline{\underline{A}}$. On the right : $\det \underline{\underline{A}}$ is the oriented volume of the parallelepiped defined by the vectors $\underline{A}_1$, $\underline{A}_2$ and $\underline{A}_3$ which form the columns of $\underline{\underline{A}}$.

Exercice 1.6 Calculate the determinant of the matrix given by (1.24). Discuss.

Solution 1.6

$$\det \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = 0. \tag{1.50}$$

This confirms that this matrix cannot be inverted. This is true of the matrices of all projectors. ■

We can thus also accept that :

Théorème 1.3

$$\det\left(\underline{\underline{A}}\,\underline{\underline{B}}\right) = \left(\det \underline{\underline{A}}\right)\left(\det \underline{\underline{B}}\right). \tag{1.51}$$

However, there is no formula which would allow us to simplify $\det\left(\underline{\underline{A}} + \underline{\underline{B}}\right)$ in all cases.

Exercice 1.7 Exercise. Demonstrate (1.51) for $n = 2$.

Solution 1.7

$$\det\left(\underline{\underline{A}}\,\underline{\underline{B}}\right) - \left(\det\underline{\underline{A}}\right)\left(\det\underline{\underline{B}}\right) = \left(a_{11}b_{11}+a_{12}b_{21}\right)\left(a_{21}b_{12}+a_{22}b_{22}\right) - \left(a_{11}b_{12}+a_{12}b_{22}\right)\left(a_{21}b_{11}+a_{22}b_{21}\right) - \left(a_{11}a_{22}-a_{12}a_{21}\right)\left(b_{11}b_{22}-b_{12}b_{21}\right) = 0. \quad (1.52)$$

■

Exercice 1.8 Show that

$$\det\left(\underline{\underline{A}}^{-1}\right) = \left(\det\underline{\underline{A}}\right)^{-1}. \quad (1.53)$$

Solution 1.8 The definition (1.23) demonstrates that $\det\underline{\underline{I}} = 1$. The result we are after is therefore a consequence of (1.33) and (1.51). ■

As with the product of two matrices, the determinant is independent of basis, so we can regard it as the determinant of an endomorphism (in much the same way that we would talk about the norm of a vector). Indeed :

$$\det\underline{\underline{A}} = \det\left(\underline{\underline{P}}^{-1}\underline{\underline{A}}'\underline{\underline{P}}\right) = \det\left(\underline{\underline{P}}^{-1}\right)\left(\det\underline{\underline{A}}'\right)\det\underline{\underline{P}} = \det\underline{\underline{A}}'. \quad (1.54)$$

The capacity for inversion is therefore not dependent upon the basis in which we are working. In order to determine the inverse of a matrix (where it exists), we still need one more element :

Définition 1.4 *Cofactors and comatrix.* For a matrix $\underline{\underline{A}}$ with the coefficients $\left\{a_{ij}\right\}_{i,i=1,\ldots,n}$, each coefficient has a cofactor defined by

$$\text{cofact}(a_{ij}) \doteqdot (-1)^{i+j}\det\begin{pmatrix} a_{1,1} & \cdots & a_{1,j-1} & a_{1,j+1} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,j-1} & a_{i-1,j+1} & \cdots & a_{i-1,n} \\ a_{i+1,1} & \cdots & a_{i+1,j-1} & a_{i+1,j+1} & \cdots & a_{i+1,n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & \cdots & a_{n,j-1} & a_{n,j+1} & \cdots & a_{nn} \end{pmatrix}. \quad (1.55)$$

We therefore have, sign for sign, the determinant of the sub-matrix obtained by removing the

$i-$th line and the $j-$th column of $\underline{\underline{A}}$. The comatrix of $\underline{\underline{A}}$ is thus defined by

$$\operatorname{comat} \underline{\underline{A}} \doteqdot \begin{pmatrix} \operatorname{cofact}(a_{11}) & \cdots & \operatorname{cofact}(a_{1n}) \\ \vdots & \ddots & \vdots \\ \operatorname{cofact}(a_{n1}) & \cdots & \operatorname{cofact}(a_{nn}) \end{pmatrix}. \tag{1.56}$$

Exercice 1.9 Calculate the comatrix of an arbitrary 2×2 matrix.

Solution 1.9

$$\operatorname{comat} \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{pmatrix} a_{22} & -a_{21} \\ -a_{12} & a_{11} \end{pmatrix}. \tag{1.57}$$

■

Théorème 1.4 For any matrix

$$\left(\operatorname{comat} \underline{\underline{A}}\right)^T \underline{\underline{A}} = \left(\det \underline{\underline{A}}\right) \underline{\underline{I}}. \tag{1.58}$$

Exercice 1.10 Verify this in the 2nd dimension.

Solution 1.10

$$\begin{aligned} \left(\operatorname{comat} \underline{\underline{A}}\right)^T \underline{\underline{A}} - \left(\det \underline{\underline{A}}\right) \underline{\underline{I}} &= \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \\ &- \begin{pmatrix} a_{11}a_{22} - a_{12}a_{21} & 0 \\ 0 & a_{11}a_{22} - a_{12}a_{21} \end{pmatrix} \\ &= \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}. \end{aligned} \tag{1.59}$$

■

From this we can deduce the following important result :

Théorème 1.5 *Matrice inverse.* For any invertible matrix $\underline{\underline{A}}$, the inverse matrix is given by

$$\underline{\underline{A}}^{-1} = \frac{1}{\det \underline{\underline{A}}} \left(\operatorname{comat} \underline{\underline{A}}\right)^T. \tag{1.60}$$

Exercice 1.11 Calculate the inverse of

$$\underline{\underline{A}} \doteqdot \begin{pmatrix} -1 & 0 & 1 \\ 0 & 3 & 2 \\ -2 & 4 & 1 \end{pmatrix}, \tag{1.61}$$

Then check your result by calculating the product of the matrices.

Solution 1.11 We start by calculating

$$\det \underline{\underline{A}} = 11, \tag{1.62}$$

Demonstrating that it is invertible, then

$$\text{comat}\, \underline{\underline{A}} = \begin{pmatrix} -5 & -4 & 6 \\ 4 & 1 & 4 \\ -3 & 2 & -3 \end{pmatrix}, \tag{1.63}$$

which gives us

$$\underline{\underline{A}}^{-1} = \frac{1}{11} \begin{pmatrix} -5 & 4 & -3 \\ -4 & 1 & 2 \\ 6 & 4 & -3 \end{pmatrix}. \tag{1.64}$$

Finally, we check that $\underline{\underline{A}}\,\underline{\underline{A}}^{-1} = \underline{\underline{I}}$. ■

1.1.3 Eigenvectors and eigenvalues

Let us consider the endomorphism of $\mathbb{R}^2$ whose representation in a basis $\mathscr{B}$ is the following matrix :

$$\underline{\underline{A}} \doteqdot \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}. \tag{1.65}$$

Figure 1.3 is a graphical representation in this basis, in the same format as Figure 1.2 : we can see that the vectors colinear to $(1,1)^T$ are dilated by a factor of 2 without undergoing rotation, and the vectors colinear to $(1,-1)^T$ remain unchanged. This property is obviously very interesting from a geometric perspective, as it allows us to characterize the dilation or contraction of a material in a given direction, for example. In fact, studying these properties has far broader mathematical and physical implications, the broad outlines of which we shall attempt to define hereunder.

Let's start with a few definitions.

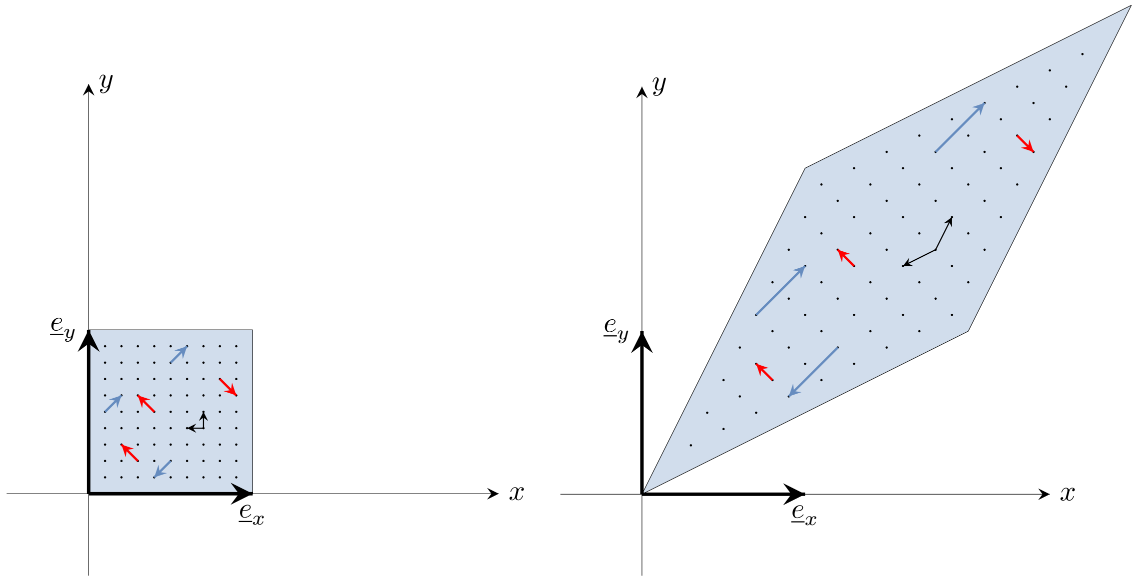


FIGURE 1.3 – Illustration of the concept of eigenvalues and eigenvectors : the matrix (1.65). The image on the right includes the images of several endomorphism vectors The vectors shown in blue are eigenvectors associated with the eigenvalue $\lambda_1 \doteqdot 2$ (the red correspond to $\lambda_2 \doteqdot 1$). The other vectors, for example those in black, change direction. Taken from Wikipedia

Définition 1.5 *Trace.* The *trace* of a matrix is the sum of its diagonal coefficients :

$$\operatorname{Tr} \underline{\underline{A}} \doteqdot \sum_{i=1}^{n} a_{ii}. \tag{1.66}$$

Exercice 1.12 Show that

$$\operatorname{Tr}(\underline{\underline{A}}\,\underline{\underline{B}}) = \operatorname{Tr}(\underline{\underline{B}}\,\underline{\underline{A}}). \tag{1.67}$$

Solution 1.12

$$\begin{aligned}
\operatorname{Tr}(\underline{\underline{A}}\,\underline{\underline{B}}) &= \sum_{i=1}^{n} (\underline{\underline{A}}\,\underline{\underline{B}})_{ii} \\
&= \sum_{i=1}^{n}\sum_{k=1}^{n} a_{ik} b_{ki} \\
&= \sum_{k=1}^{n}\sum_{i=1}^{n} b_{ki} a_{ik} \\
&= \operatorname{Tr}(\underline{\underline{B}}\,\underline{\underline{A}}).
\end{aligned} \tag{1.68}$$

■

Exercice 1.13 Show that the trace is not dependent on the basis.

Solution 1.13 We can call this the trace of the endomorphism, because the result of the preceding exercise shows us that

$$\begin{aligned}\operatorname{Tr}\underline{\underline{A}} &= \operatorname{Tr}(\underline{\underline{P}}^{-1}\underline{\underline{A}}'\underline{\underline{P}}) \\ &= \operatorname{Tr}(\underline{\underline{P}}^{-1}\underline{\underline{P}}\,\underline{\underline{A}}') \\ &= \operatorname{Tr}\underline{\underline{A}}'.\end{aligned} \tag{1.69}$$

■

Définition 1.6 *Diagonal matrix.* A matrix is said to be *diagonal* when it takes the following form :

$$\underline{\underline{D}} \doteqdot \begin{pmatrix} \lambda_1 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_i & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & \lambda_n \end{pmatrix}. \tag{1.70}$$

This is sometimes written

$$\underline{\underline{D}} \doteqdot \operatorname{diag}\{\lambda_i\}_{i=1,\ldots,n}. \tag{1.71}$$

We can thus easily agree upon the following properties :

$$\operatorname{Tr}\underline{\underline{D}} = \sum_{i=1}^{n}\lambda_i, \tag{1.72}$$

$$\det\underline{\underline{D}} = \prod_{i=1}^{n}\lambda_i, \tag{1.73}$$

$$\forall k,\ \ \underline{\underline{D}}^k = \operatorname{diag}\left\{\lambda_i^k\right\}_{i=1,\ldots,n}. \tag{1.74}$$

Définition 1.7 *Eigenvectors and eigenvalues.* A non-null vector $\underline{X}$ is said to be the *eigenvector* or characteristic vector of a matrix $\underline{\underline{A}}$ if it is colinear with its image $\underline{\underline{A}}$:

$$\exists \underline{X} \neq \underline{0},\quad \underline{\underline{A}}\,\underline{X} = \lambda\underline{X}. \tag{1.75}$$

The proportionality factor λ is the *eigenvalue* of $\underline{\underline{A}}$ associated with $\underline{X}$.

In the example at the top of this paragraph, $\lambda_1 \doteqdot 2$ and $\lambda_2 \doteqdot 1$ are eigenvalues of $\underline{\underline{A}}$ associated with the respective eigenvectors $\underline{X}_1 \doteqdot (1,1)^T$ and $\underline{X}_2 \doteqdot (1,-1)^T$.

Exercice 1.14 Demonstrate that the eigenvalues and eigenvectors are not dependent upon the basis in which you are working, in other words that there are eigenvalues and eigenvectors specific to endomorphisms.

Solution 1.14 On the one hand

$$\begin{aligned} \underline{\underline{A}}\,\underline{X} &= (\underline{\underline{P}}^{-1}\underline{\underline{A}}'\underline{\underline{P}})(\underline{\underline{P}}^{-1}\underline{X}') \\ &= \underline{\underline{P}}^{-1}\underline{\underline{A}}'\underline{X}', \end{aligned} \tag{1.76}$$

and on the other hand

$$\begin{aligned} \lambda\underline{X} &= \lambda(\underline{\underline{P}}^{-1}\underline{X}') \\ &= \underline{\underline{P}}^{-1}(\lambda\underline{X}'), \end{aligned} \tag{1.77}$$

Which shows that the relation which defines the eigenvector remains unchanged when we change basis, as is the associated eigenvalue. We still need to make clear that $\underline{X} \neq 0$ implies that $\underline{X}' \neq 0$ because transition matrices are invertible. ■

It should now be clear that, for each eigenvalue, all of the vectors which are colinear to an associated eigenvector are also eigenvectors associated with the same eigenvalue. At a more general level, any linear combination of eigenvectors associated with λ is an eigenvector associated with λ. We can thus talk of *eigenspaces or characteristic spaces* (in this case the lines : you may also encounter the term *eigendirections*). It is also easy to see why an eigenvector cannot be associated with two discrete eigenvalues, because by subtraction that would give us $(\lambda_1 - \lambda_2)\underline{X} = \underline{0}$, which would contradict the fact that $\underline{X} \neq \underline{0}$ by definition. As such, the intersection of two eigenspaces is reduced by the null vector. Eigenspaces thus form a direct sum, and the sum of their dimensions can never exceed the dimension of the workspace. In the previous example, the plane was the direct sum of two distinct, characteristic directions. We can thus conclude that a single endomorphism $\mathbb{R}^n$ cannot have more than n eigenvalues.

In order to move forward, let us make the following observation :

$$\exists \underline{X} \neq \underline{0}, \quad \underline{\underline{A}}\,\underline{X} = \underline{0} \Longleftrightarrow \det\underline{\underline{A}} = 0. \tag{1.78}$$

If $\underline{\underline{A}}$ were invertible we could deduce that $\underline{X} = \underline{\underline{A}}^{-1}\underline{0} = \underline{0}$, contradicting our hypothesis. This suggests that $\underline{\underline{A}}$ is not invertible, and so $\det\underline{\underline{A}} = 0$ according to the theorem 1.2. As such, by writing (1.75) in the form $\exists\underline{X} \neq \underline{0}, (\underline{\underline{A}} - \lambda\underline{\underline{I}})\underline{X} = 0$ we can deduce the following :

Théorème 1.6 An eigenvector-eigenvalue pairing $(\underline{X}, \lambda)$ of $\underline{\underline{A}}$ satisfies

$$\mathscr{P}[\underline{\underline{A}}](\lambda) = 0, \tag{1.79}$$

where $\mathscr{P}[\underline{\underline{A}}]$ is the *characteristic polynomial* of $\underline{\underline{A}}$, defined by

$$\mathscr{P}[\underline{\underline{A}}](\lambda) \doteqdot \det(\underline{\underline{A}} - \lambda\underline{\underline{I}}). \tag{1.80}$$

As the determinant is independent of basis, this result is consistent with the fact that eigenvalues and eigenvectors are attached to endomorphisms rather than to any particular matricial representation.

That $\mathscr{P}[\underline{\underline{A}}](\lambda)$ is a polynomial of λ (of a degree equal to the dimension n of the workspace) is a consequence of the definition 1.3 of the determinant.

Exercice 1.15 Determine the characteristic polynomial of a 2-dimensional matrix based on its determinant and its trace.

Solution 1.15

$$\begin{aligned}\mathscr{P}\left[\underline{\underline{A}}\right](\lambda) &= \begin{vmatrix} a_{11}-\lambda & a_{12} \\ a_{21} & a_{22}-\lambda \end{vmatrix} \\ &= \lambda^2 - (a_{11}+a_{22})\lambda + a_{11}a_{22} - a_{12}a_{21} \\ &= \lambda^2 - \left(\mathrm{Tr}\,\underline{\underline{A}}\right)\lambda + \det \underline{\underline{A}}\end{aligned} \tag{1.81}$$

This confirms, at least for 2 dimensions, that the characteristic polynomial is not dependent on the basis, much like $\det \underline{\underline{A}}$ and $\mathrm{Tr}\,\underline{\underline{A}}$. ■

Adding the 3rd dimension, we can check this result :

$$\mathscr{P}\left[\underline{\underline{A}}\right](\lambda) = -\lambda^3 + \left(\mathrm{Tr}\,\underline{\underline{A}}\right)\lambda^2 - \tfrac{1}{2}\left(\left(\mathrm{Tr}\,\underline{\underline{A}}\right)^2 - \mathrm{Tr}\left(\underline{\underline{A}}^2\right)\right)\lambda + \det \underline{\underline{A}}. \tag{1.82}$$

Théorème 1.7 *Cayley-Hamilton Theorem.* A matrix cancels out its characteristic polynomial :

$$\mathscr{P}[\underline{\underline{A}}](\underline{\underline{A}}) = \underline{\underline{0}}. \tag{1.83}$$

By this we mean that the matrix satisfies its own characteristic equation if we replace λ by $\underline{\underline{A}}$ in the polynomial, as long as $\underline{\underline{A}}^0 = \underline{\underline{I}}$.

Exercice 1.16 Verify this theorem in 2 dimensions.

Solution 1.16

$$\begin{aligned}P\left[\underline{\underline{A}}\right]\left(\underline{\underline{A}}\right) &= \underline{\underline{A}}^2 - \left(\mathrm{Tr}\,\underline{\underline{A}}\right)\underline{\underline{A}} + \left(\det \underline{\underline{A}}\right)\underline{\underline{I}} \\ &= \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}^2 - (a_{11}+a_{22})\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} + (a_{11}a_{22} - a_{12}a_{21})\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}\end{aligned} \tag{1.84}$$

Doing the same thing with a 3rd dimension is more time-consuming, but still possible to do manually. Consult more specialist works for demonstrations of this theorem in all dimensions, invoking notions which cannot be covered here. ■

Now let us get to the most important point in this paragraph.

Définition 1.8 *Diagonalizable matrix.* A matrix (or, to be more accurate, an endomorphism) is said to be *diagonalizable* if and only if there exists a basis (not necessarily orthogonal) in which its representation is diagonal :

$$\begin{aligned} \exists \underline{\underline{P}}, \quad \underline{\underline{A}} &= \underline{\underline{P}}^{-1} \underline{\underline{D}}\,\underline{\underline{P}} \\ \underline{\underline{D}} &\doteqdot \text{diag}\left\{\lambda_i\right\}_{i=1,\ldots,n} \\ \forall i, \quad \mathscr{P}\left[\underline{\underline{A}}\right](\lambda_i) &= 0 \end{aligned} \tag{1.85}$$

Exercice 1.17 Show proof of theorem 1.7 for all dimensions for a diagonalizable endomorphism.

Solution 1.17 Let us begin by noting that the formula (1.42) allows us to assert that $\mathscr{P}[\underline{\underline{A}}](\underline{\underline{P}}^{-1}\underline{\underline{D}}\,\underline{\underline{P}}) = \mathscr{P}[\underline{\underline{A}}](\underline{\underline{D}})$. Moreover, with (1.74) and (1.80) we have $\mathscr{P}[\underline{\underline{A}}](\underline{\underline{D}}) = \text{diag}\left\{\mathscr{P}[\underline{\underline{A}}](\lambda_i)\right\}_{i=1,\ldots,n} = \underline{\underline{0}}$. Now, because the characteristic polynomial is not dependent on the basis, by using the diagonalization basis we get

$$\begin{aligned} \mathscr{P}\left[\underline{\underline{A}}\right]\left(\underline{\underline{A}}\right) &= \mathscr{P}\left[\underline{\underline{A}}\right]\left(\underline{\underline{P}}^{-1}\underline{\underline{DP}}\right) \\ &= \mathscr{P}\left[\underline{\underline{A}}\right]\left(\underline{\underline{D}}\right) \\ &= \underline{\underline{0}}. \end{aligned} \tag{1.86}$$

■

N.B. For a diagonalizable endomorphism, eigenvalues are roots of the characteristic polynomial and the latter can thus be expressed as follows

$$\mathscr{P}[\underline{\underline{A}}](\lambda) = \prod_{i=1}^{n}(\lambda - \lambda_i), \tag{1.87}$$

With certain λ_i which may be identical if there are multiple roots. For dimensions 2 and 3, this is consistent with (1.81) and (1.82) with (1.72)–(1.73).

If we want to diagonalize a matrix, identifying its eigenvalues is not enough ; we also need to find a basis in which the endomorphism takes a diagonal form, which is to say that we need to identify a suitable transition matrix $\underline{\underline{P}}$. There may be several possibilities if some eigenvalues are associated with eigenspaces with a dimension greater than 1 ; they are obviously the roots of a characteristic polynomial with a degree of multiplicity greater than 1. We can choose a $\underline{\underline{P}}$ using the following theorem :

Théorème 1.8 If a transition matrix $\underline{\underline{P}}$ towards a diagonalization basis $\underline{\underline{A}}$ is given by (1.85), then its inverse $\underline{\underline{P}}^{-1}$ is composed of eigenvectors which form a family, arranged in a single column.

Preuve. By multiplying (1.85) on the right by $\underline{\underline{P}}^{-1}$ we find that

$$\underline{\underline{A}}\,\underline{\underline{P}}^{-1} = \underline{\underline{P}}^{-1}\underline{\underline{D}}. \tag{1.88}$$

If we assign the symbols $\{\underline{X}_i\}_{i=1,\ldots,n}$ to the vectors which make up the columns of $\underline{\underline{P}}^{-1}$, it soon becomes clear that (1.88) is made up of columns in the form $\underline{\underline{A}}\,\underline{X}_i = \lambda_i \underline{X}_i$. ■

To conclude, here are two theorems which should not need demonstration :

Théorème 1.9 A (real) and symmetrical matrix is diagonalizable in an orthonormal basis.

Théorème 1.10 Two diagonalizable matrices will commute (i.e. $\underline{\underline{A}}\,\underline{\underline{B}} = \underline{\underline{B}}\,\underline{\underline{A}}$) if and only if they share a common diagonalization basis.

Let us consider an example of diagonalization. Take the following matrix :

$$\underline{\underline{A}} \doteqdot \begin{pmatrix} -8 & 0 & 5 \\ 10 & 2 & -5 \\ -10 & 0 & 7 \end{pmatrix}. \tag{1.89}$$

To begin with, the characteristic polynomial can be written, according to (1.82), as follows :

$$\begin{aligned} \mathscr{P}\left[\underline{\underline{A}}\right](\lambda) &= -\lambda^3 + \lambda^2 + 8\lambda - 12 \\ &= -(\lambda - 2)^2 (\lambda + 3). \end{aligned}$$

The eigenvalues are thus $\lambda_1 \doteqdot 2$ and $\lambda_2 \doteqdot -3$. The matrix will be diagonalizable if we can identify a characteristic plane associated with λ_1. To do that, we need to find the eigenvectors. $\underline{\underline{A}}\,\underline{X} = 2\underline{X}$ s'écrit

$$-8x + 5z = 2x, \tag{1.90}$$
$$10x + 2y - 5z = 2y, \tag{1.91}$$
$$-10x + 7z = 2z. \tag{1.92}$$

By grouping all of the terms in the left side, we can see that these equations are in fact identical, so $2x = z$. The system (1.90) thus has two degrees of freedom, for example x and y, after which z is determined by this equation. A characteristic plane basis associated with λ_1 is thus given by the equation $\underline{X}_1 \doteqdot (1,0,2)^T, \underline{X}_2 \doteqdot (1,1,2)^T$. As for the characteristic line associated with λ_2, it is guided by a vector which satisfies the equation $\underline{\underline{A}}\,\underline{X} = -3\underline{X}$:

$$-8x + 5z = -3x, \tag{1.93}$$
$$10x + 2y - 5z = -3y, \tag{1.94}$$
$$-10x + 7z = -3z. \tag{1.95}$$

This time, only two of the equations are identical, and we are left with

$$-x + z = 0 \tag{1.96}$$
$$2x + y - z = 0. \tag{1.97}$$

We thus have just a single degree of freedom, for example x, after which y and z are determined by these two equations. We can thus select the eigenvector $\underline{X}_3 \doteqdot (1,-1,-1)^T$. The theorem 1.8 tells us that a transition matrix towards a diagonalization basis is given by

$$\underline{\underline{P}}^{-1} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & -1 \\ 2 & 2 & 1 \end{pmatrix}. \tag{1.98}$$

Its determinant is equivalent to $\det \underline{\underline{P}}^{-1} = -1$, and its comatrix can be easily calculated to find its inverse $\underline{\underline{P}}$:

$$\underline{\underline{P}} = \begin{pmatrix} -3 & -1 & 2 \\ 2 & 1 & -1 \\ 2 & 0 & -1 \end{pmatrix}. \tag{1.99}$$

As the diagonal form of the matrix is $\underline{\underline{D}} \doteqdot \operatorname{diag}\{2,2,-3\}$, we check that $\underline{\underline{P}}^{-1}\underline{\underline{D}}\underline{\underline{P}} = \underline{\underline{A}}$. We can also check that $\operatorname{Tr} \underline{\underline{A}} = \operatorname{Tr} \underline{\underline{D}} = 1$ and that $\det \underline{\underline{A}} = \det \underline{\underline{D}} = -12$, i.e. that the Cayley–Hamilton theorem is satisfied.

1.2 Tensor Algebra

1.2.1 The concept of tensors

A tensor is a mathematical object used to group and generalize notions involving numbers, vectors and matrices. A number (otherwise known as a « scalar ») is a tensor of order 0, a vector is a tensor of order 1, and a matrix is a tensor of order 2. We should also make it clear that all of the above (with the exception of scalars, which are naturally independent of the working basis) must be reformulated in order to give these objects an absolute character. A tensor of order p can thus be defined as a table with p inputs in a space of dimension n, whose components vary depending on the working basis, generalizing from the rules defined in Section 1.

A tensor of order p is underlined p times, and its components are indexed p times : scalar A, vector $\underline{A}$ with components A_i ($i = 1, \ldots n$), matrix (sometimes simply called a tensor, although this is technically inaccurate) $\underline{\underline{A}}$ with components A_{ij}, etc.

> **Définition 1.9** *Writing vectors in a basis.* We use $\underline{e}_i$ to denote vectors within an orthonormal basis. A vector can thus be written $\underline{A} = \sum_{i=1}^{n} A_i \underline{e}_i = A_1 \underline{e}_1 + \ldots + A_n \underline{e}_n = A_x \underline{e}_x + A_y \underline{e}_y + A_z \underline{e}_z$ pour $n = 3$.

The tensor product of two vectors in this basis can thus be written $\underline{e}_i \otimes \underline{e}_j$. Simply put, this represents the intersection of the i−iþligne and the j−iþcolumn of an empty matrix :

$$\begin{pmatrix} \underline{e}_1 \otimes \underline{e}_1 & & \cdots & & \underline{e}_1 \otimes \underline{e}_n \\ & \ddots & & & \\ \vdots & & \underline{e}_i \otimes \underline{e}_j & & \vdots \\ & & & \ddots & \\ \underline{e}_n \otimes \underline{e}_1 & & \cdots & & \underline{e}_n \otimes \underline{e}_n \end{pmatrix}. \tag{1.100}$$

The matrix can thus be written $\underline{\underline{A}} = \sum_{i,j=1}^{n} A_{ij} \underline{e}_i \otimes \underline{e}_j$.

■ **Exemple 1.2** Consider the following second order tensor :

$$3\underline{e}_1 \otimes \underline{e}_1 - \underline{e}_1 \otimes \underline{e}_2 + 8\underline{e}_2 \otimes \underline{e}_2 = \begin{pmatrix} 3 & -1 \\ 0 & 8 \end{pmatrix}. \tag{1.101}$$

■

Strictly speaking, the preceding example only shows the representation of a tensor in a specific basis ; a full definition would require more information regarding said basis. For a more general example, taking into consideration the Kronecker symbol δ_{ij}, the identity matrix in dimension n, as defined in (1.23), can be written thus

$$\underline{\underline{I}} = \sum_{i,j=1}^{n} \delta_{ij} \underline{e}_i \otimes \underline{e}_j, \tag{1.102}$$

irrespective of basis.

Définition 1.10 *Einstein notation (for repeated indices).* When an index variable is repeated (i.e. appears twice) within the same term, it implies summation of that term over all the values of the index. This helps to avoid excessive use of the summation symbol $\sum$.

For example, the trace of a matrix $\mathrm{Tr}\underline{\underline{A}} = \sum_{i,j=1}^{n} A_{ii}$ can simply be written A_{ii}. By the same token, the tensor $\underline{\underline{I}} = \sum_{i,j=1}^{n} \delta_{ij} \underline{e}_i \otimes \underline{e}_j$ can be simplified to $\underline{\underline{I}} = \delta_{ij} \underline{e}_i \otimes \underline{e}_j = \underline{e}_i \otimes \underline{e}_i$. In the latter example, we have two pairs of indices ; we can make use of this summarizing technique as many times as necessary within the same term.

The tensor can thus be generalized to any order. A tensor of order p is written

$$\underline{A}_{(p)} = \underbrace{A_{ij.....k}}_{p \text{ indices}} \underbrace{\underline{e}_i \otimes \cdots \otimes \underline{e}_s}_{p \text{ vecteurs}}, \tag{1.103}$$

with summation of the p indices. As noted above, this is still not a comprehensive definition of a tensor. The data for the coefficients is not yet sufficient : they need to be transformed in line with clearly-defined rules when changing basis. The generalization of these rules (1.21) and (1.40) will be covered a little later on.

R If we were working in a non-orthonormal basis, we would have to make a clear distinction between two types of components : covariants and contravariants, with indices positioned at the bottom and at the top respectively. This is particularly important when working with a Riemannian manifold to analyze curved spaces. In fluid mechanics this distinction is unimportant (except in certain digital methods) ; we thus ignore it in the present document.

Définition 1.11 *Tensor products.* If we take two tensors p and q of any order, their tensor product (or exterior product) can be determined by placing them side to side, separated by the $\otimes$ symbol. This gives us a tensor of order $p+q$:

$$\underline{A}_{(p)} \otimes \underline{B}_{(q)} \doteqdot \underbrace{A_{ij.....k}}_{p \text{ indices}} \underbrace{B_{\ell.....rs}}_{q \text{ indices}} \underbrace{\underline{e}_i \otimes \underline{e}_j \otimes \cdots \otimes \underline{e}_r \otimes \underline{e}_s}_{p+q \text{ vecteurs}}. \tag{1.104}$$

■ **Exemple 1.3** The exterior product of two vectors is a matrix :

$$\begin{aligned} \underline{A} \otimes \underline{B} &= (A_i \underline{e}_i) \otimes \left(B_j \underline{e}_j\right) = A_i B_j \underline{e}_i \otimes \underline{e}_j \\ &= \begin{pmatrix} A_1 B_1 & \cdots & A_1 B_n \\ \vdots & \ddots & \vdots \\ A_n B_1 & \cdots & A_n B_n \end{pmatrix}. \end{aligned} \tag{1.105}$$

The exterior product of a matrix and a vector is a third-order tensor :

$$\underline{\underline{A}} \otimes \underline{B} = A_{ij} B_k \underline{e}_i \otimes \underline{e}_j \otimes \underline{e}_k. \tag{1.106}$$

■

1.2.2 Tensor contraction

Définition 1.12 *First order contraction.* To contract two tensors of any order in a single operation, we align them as we would do to calculate the tensor product, but this time with no symbol separating them. We then use the same letter for the last index of the first tensor and the first index of the second tensor. Finally, we use Einstein notation for this index :

$$\underline{A}_{(p)} \cdot \underline{B}_{(q)} \doteqdot \underbrace{A_{ij.....k}}_{p \text{ indices}} \underbrace{B_{k.....\ell m}}_{q \text{ indices}} \underbrace{\underline{e}_i \otimes \underline{e}_j \otimes \cdots \otimes \underline{e}_\ell \otimes \underline{e}_m}_{p+q-2 \text{ vectors}}. \tag{1.107}$$

The result is a tensor of order $p+q-2$.

■ **Exemple 1.4** Consider these three common examples :
$\underline{A} \cdot \underline{B} = A_i B_i$ is the *scalar* product of two vectors $(1+1-2=0)$.
$\underline{\underline{A}} \cdot \underline{B} = A_{ij} B_j \underline{e}_i$ is the matrix-vector product, a vector $(2+1-2=1)$.
$\underline{\underline{A}} \cdot \underline{\underline{B}} = A_{ik} B_{kj} \underline{e}_i \otimes \underline{e}_j$ is the matrix product, a matrix $(2+2-2=2)$. ■

We can now look at the rule by which all components of a tensor must abide when changing basis.

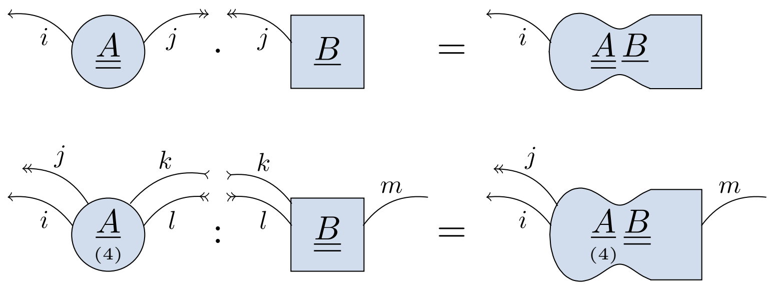


FIGURE 1.4 – Illustrating the contracted product principle with two examples of the « amoeba » technique in action : each tensor is represented by an amoeba of a particular species, with the number of flagella corresponding to the order of the tensor. The flagella disappear through contraction. We thus have the simple product of a matrix and a vector (top, giving us a vector) and the double product of a 4th-order tensor and a matrix (bottom, giving us a 3rd-order tensor).

Théorème 1.11 During a change of orthonormal basis using an orthogonal transition matrix $\underline{\underline{P}}$, a tensor of order p written $\underline{A}_{(p)}$ must follow this rule :

$$\underline{A}_{(p)} = \underbrace{A'_{ij....k}}_{p \text{ indices}} \underbrace{\underline{\underline{P}}^T \underline{e}_i \otimes \cdots \otimes \underline{\underline{P}}^T \underline{e}_k}_{p \text{ produits matrice-vecteur}} . \tag{1.108}$$

Exercice 1.18 Check that this confirms the 2nd-order case (1.67).

Solution 1.18 The $k-$th component of the vector $\underline{e}_i$ is δ_{ik}, so

$$\begin{aligned} \underline{\underline{P}}^T \underline{e}_i &= P_{k\ell}\delta_{ik}\underline{e}_\ell \\ &= P_{i\ell}\underline{e}_\ell. \end{aligned} \tag{1.109}$$

So for $p = 2$ (1.108) we write

$$\begin{aligned} \underline{\underline{A}} &= A'_{ij}\underline{\underline{P}}^T \underline{e}_i \otimes \underline{\underline{P}}^T \underline{e}_j \\ &= A'_{ij}P_{i\ell}\underline{e}_\ell \otimes P_{jm}\underline{e}_m \\ &= P_{i\ell}A'_{ij}P_{jm}\underline{e}_\ell \otimes \underline{e}_m \\ &= \underline{\underline{P}}^T \underline{\underline{A}}' \underline{\underline{P}}. \end{aligned} \tag{1.110}$$

■

Définition 1.13 *Contraction in an arbitrary order k ; double contraction.* The process demonstrated above for first-order contraction can also be applied to the last k indices of the first tensor and the first k indices of the second, giving us a tensor of order $p+q-2k$. For the purposes of this seminar, we will focus on the double contracted product ($k = 2$) :

$$\underline{A}_{(p)} : \underline{B}_{(q)} \doteqdot \underbrace{A_{ij.....k\ell}}_{p \text{ indices}} \underbrace{B_{\ell k.....rs}}_{q \text{ indices}} \underbrace{\underline{e}_i \otimes \underline{e}_j \otimes \cdots \otimes \underline{e}_r \otimes \underline{e}_s}_{p+q-4 \text{ vectors}}, \tag{1.111}$$

The result is a tensor of order $p + q - 4$.

We will also restrict this exercise to the double product of two matrices, which is a scalar $(2 + 2 - 4 = 0)$:

$$\underline{\underline{A}} : \underline{\underline{B}} = A_{ij}B_{ji}. \tag{1.112}$$

It is a sort of scalar product of two matrices. It is also equal to the trace of their product :

$$\mathrm{Tr}(\underline{\underline{A}}\,\underline{\underline{B}}) = A_{ij}B_{ji} = \underline{\underline{A}} : \underline{\underline{B}}. \tag{1.113}$$

It is important to note that the products of tensors are tensors themselves, in that they also transform in accordance with the rule (1.108). If you're feeling brave, you can attempt to demonstrate this result with an exercise.

Usually, to understand the principle of the contracted product we can refer to the conceptual graph shown in Figure 1.4. The following theorem is very intuitive, and may be of use for exercise purposes :

Théorème 1.12 Tensor and contracted products (of whatever order) are linear.

Exercice 1.19 Calculate the following :

$$(3\underline{e}_1 \otimes \underline{e}_1 - \underline{e}_1 \otimes \underline{e}_2 + 8\underline{e}_2 \otimes \underline{e}_2) : (\underline{e}_1 \otimes \underline{e}_1 + 2\underline{e}_1 \otimes \underline{e}_2 + 5\underline{e}_2 \otimes \underline{e}_1 - 4\underline{e}_2 \otimes \underline{e}_2). \tag{1.114}$$

Solution 1.19

$$\begin{aligned} \ldots &= \begin{pmatrix} 3 & -1 \\ 0 & 8 \end{pmatrix} : \begin{pmatrix} 1 & 2 \\ 5 & -4 \end{pmatrix} \\ &= 3 \times 1 + (-1) \times 5 + 0 \times 2 + 8 \times (-4) \\ &= -34. \end{aligned} \tag{1.115}$$

■

Exercice 1.20 Taking the three vectors $\underline{A}$, $\underline{B}$ and $\underline{C}$, simplify the formulation $(\underline{A} \otimes \underline{B}) \cdot \underline{C}$. Consider the case $\underline{C} = \underline{B}$.

Solution 1.20

$$(\underline{A} \otimes \underline{B}) \cdot \underline{C} = (A_i B_j \underline{e}_i \otimes \underline{e}_j) \cdot (C_k \underline{e}_k) = A_i B_j C_j \underline{e}_i = (\underline{B} \cdot \underline{C})\underline{A}. \tag{1.116}$$

If $\underline{C} = \underline{B}$, this gives us :

$$(\underline{A} \otimes \underline{B}) \cdot \underline{B} = |\underline{B}|^2 \underline{A}. \tag{1.117}$$

■

Exercice 1.21 Show that the product $\underline{\underline{A}} : \underline{\underline{A}}^T$ is the square of a norm (i.e. a positive value, which only disappears if the tensor is null).

Solution 1.21

$$\underline{\underline{A}} : \underline{\underline{A}}^T = A_{ij} A^T_{ji} = A_{ij} A_{ij} = \sum_{i,j=1}^{n} A^2_{ij}. \tag{1.118}$$

■

More generally, we write $\left(A_{ij\ldots rs}\underline{e}_i \otimes \underline{e}_j \otimes \cdots \otimes \underline{e}_r \otimes \underline{e}_s\right)^T \doteqdot A_{sr\ldots ji}\underline{e}_i \otimes \underline{e}_j \otimes \cdots \otimes \underline{e}_r \otimes \underline{e}_s$ for the transpose of a tensor. This operator only works for order two and above, as it inverts the order of all of the indices. We can easily demonstrate that it also inverts the order of the contracted product :

$$\begin{aligned} \underline{A}^T_{(p)} \cdot \underline{B}^T_{(q)} &= A^T_{ij.....k} B^T_{k.....\ell m} \underline{e}_i \otimes \underline{e}_j \otimes \cdots \otimes \underline{e}_\ell \otimes \underline{e}_m \\ &= A_{k.....ji} B_{\ell m.....k} \underline{e}_i \otimes \underline{e}_j \otimes \cdots \otimes \underline{e}_\ell \otimes \underline{e}_m \\ &= B_{m\ell.....k} A_{k.....ij} \underline{e}_i \otimes \underline{e}_j \otimes \cdots \otimes \underline{e}_\ell \otimes \underline{e}_m \\ &= \left(\underline{B}_{(q)} \cdot \underline{A}_{(p)}\right)^T. \end{aligned} \tag{1.119}$$

This is also true of the double product.

Exercice 1.22 Verify this with

1. $\underline{\underline{A}}^T \cdot \underline{\underline{B}}^T$,
2. $\underline{\underline{A}}^T \cdot \underline{B}$,
3. $\underline{\underline{A}} : \underline{\underline{B}}^T$.

Solution 1.22 We can write out these expressions and the expressions we are looking for, and demonstrate their equality, by manipulating the indices a little :

$$\begin{aligned} \underline{\underline{A}}^T \cdot \underline{\underline{B}}^T &= A^T_{ik} B^T_{kj} \underline{e}_i \otimes \underline{e}_j = A_{ki} B_{jk} \underline{e}_i \otimes \underline{e}_j, \\ \left(\underline{\underline{B}} \cdot \underline{\underline{A}}\right)^T &= \left(B_{ik} A_{kj} \underline{e}_i \otimes \underline{e}_j\right)^T = B_{jk} A_{ki} \underline{e}_i \otimes \underline{e}_j. \end{aligned} \tag{1.120}$$

Then :

$$\begin{aligned} \underline{\underline{A}}^T \cdot \underline{B} &= A^T_{ij} B_j \underline{e}_i = A_{ji} B_j \underline{e}_i \\ \underline{B}^T \cdot \underline{\underline{A}} &= B_i A_{ij} \underline{e}_j = A_{ij} B_i \underline{e}_j, \end{aligned} \tag{1.121}$$

Which are identical once the mute indices have been inverted (in this case there is no need for transposition, since the result is a vector). Finally :

$$\begin{aligned} \underline{\underline{A}} : \underline{\underline{B}}^T &= A_{ij} B^T_{ji} = A_{ij} B_{ij} \\ \underline{\underline{B}} : \underline{\underline{A}}^T &= B_{ij} A^T_{ji} = A_{ij} B_{ij}. \end{aligned} \tag{1.122}$$

■

In this latter case, we might also observe that $\text{Tr}(\underline{\underline{A}}\,\underline{\underline{B}}) = \text{Tr}(\underline{\underline{B}}\,\underline{\underline{A}})$ (equation (1.67)). At a more general level, the trace is not sensitive to the transposition of a tensor in any order. In these calculations, and all calculations making use of tensors, we can adjust the values with indices as if they were ordinary numbers - because they are. This is the advantage of indicial notation.

R In matrix algebra we define the transpose of a vector as a vector whose components are written in a line, not in columns (equation (1.38)). This is not the purpose of the transpose operation introduced here. Nevertheless, it does coincide with the usual transposition of matrices.

So far, we have not given any specifications regarding the potential dependence of these tensors on space and/or time. All of the above remains valid in such cases. This leads us onto the subject of tensor fields :

Définition 1.14 *Tensor fields.* A field is a function of coordinates representing space $\underline{x} = (x, y, z)$ and time t. We talk about tensor fields when this function assigns a point in space-time $(\underline{x}, t)$ to a tensor. For example : the temperature within a material is a scalar field $T(\underline{x}, t)$, the velocity of a fluid is a field of the vectors $\underline{u}(\underline{x}, t)$. Time t is thus considered separately, as the components of the tensor only concern the spatial dimensions (in other words : the indices go from 1 to n, the dimension of the space).

Exercice 1.23 Calculate

$$\begin{pmatrix} 4x \\ xy + y^2 \end{pmatrix} \otimes \begin{pmatrix} y \\ 1 - 8x \end{pmatrix}. \tag{1.123}$$

Solution 1.23

$$\begin{pmatrix} 4xy & 4x(1-8x) \\ (x+y)y^2 & (x+y)(1-8x)y \end{pmatrix}. \tag{1.124}$$

■

1.3 Integro-differential tools

There is no room in this brief introduction to explore the strict fundamentals of differentiation and integration theory. We will thus have to skip over some of these subjects, but you are encouraged to seek out more specialist works to understand the conditions which underpin the existence of derivatives and integrals. To simplify slightly, the functions described here are regular, which is to say that they would represent smooth, unbroken lines on a graph, with the exception of a certain number (finite or infinite) or isolated points.

1.3.1 Differential operators

Remember the definition of the derivative of a single-variable function $A(x)$, which represents its local slope, i.e. the angle of the tangent in relation to its representative curve (positive when the function is increasing, negative when it is decreasing) :

$$\frac{\mathrm{d}A}{\mathrm{d}x}(x) \doteqdot \lim_{\varepsilon\to 0}\frac{A(x+\varepsilon)-A(x)}{\varepsilon}. \tag{1.125}$$

We can also define the partial derivatives of a function with n variables $A(x,y,\ldots)$, bearing in mind that when we have a derivative from one variable the others remain fixed. In the case $n=2$, or $A(x,y)$, which we will now use throughout this section as an illustration, this gives us

$$\begin{aligned}\frac{\partial A}{\partial x}(x,y) &\doteqdot \lim_{\varepsilon\to 0}\frac{A(x+\varepsilon,y)-A(x,y)}{\varepsilon}\\ \frac{\partial A}{\partial y}(x,y) &\doteqdot \lim_{\varepsilon\to 0}\frac{A(x,y+\varepsilon)-A(x,y)}{\varepsilon}.\end{aligned} \tag{1.126}$$

There are 3 second derivatives (for a function with two variables) :

$$\begin{aligned}\frac{\partial^2 A}{\partial x^2}(x,y) &\doteqdot \frac{\partial}{\partial x}\left(\frac{\partial A}{\partial x}\right)(x,y)\\ \frac{\partial^2 A}{\partial y^2}(x,y) &\doteqdot \frac{\partial}{\partial y}\left(\frac{\partial A}{\partial y}\right)(x,y),\\ \frac{\partial^2 A}{\partial x\partial y}(x,y) &\doteqdot \frac{\partial}{\partial x}\left(\frac{\partial A}{\partial y}\right)(x,y)=\frac{\partial}{\partial y}\left(\frac{\partial A}{\partial x}\right)(x,y).\end{aligned} \tag{1.127}$$

As for the definition in the third line of (1.127), it is easy to see that these partial derivatives commute, which is to say that we can inverse the order of the derivations. This is the crux of the following theorem :

Théorème 1.13 *Schwarz Theorem.* Partial derivatives commute, i.e. they share the following characteristic :

$$\frac{\partial^2 A}{\partial x\partial y}(x,y)=\frac{\partial^2 A}{\partial y\partial x}(x,y).$$

Exercice 1.24 Verify this with

$$A(x,y) \doteqdot x^2\cos(x-4y)+3\sqrt{xy^2-1} \tag{1.128}$$

Solution 1.24 Begin by calculating the first partial derivatives :

$$\begin{aligned}\frac{\partial A}{\partial x}(x,y) &= 2x\cos(x-4y)-x^2\sin(x-4y)+\frac{3y^2}{2\sqrt{xy^2-1}},\\ \frac{\partial A}{\partial y}(x,y) &= 4x^2\sin(x-4y)+\frac{3xy}{\sqrt{xy^2-1}}.\end{aligned} \tag{1.129}$$

Next, calculate the second cross derivative in two ways :

$$\begin{aligned}\frac{\partial^2 A}{\partial x \partial y}(x,y) &= \frac{\partial}{\partial x}\frac{\partial A}{\partial y}(x,y) \\ &= 8x\sin(x-4y) + 4x^2\cos(x-4y) + \frac{3y}{\sqrt{xy^2-1}} - \frac{3xy^3}{2(xy^2-1)^{3/2}}, \\ \frac{\partial^2 A}{\partial y \partial x}(x,y) &= \frac{\partial}{\partial y}\frac{\partial A}{\partial x}(x,y) \\ &= 8x\sin(x-4y) + 4x^2\cos(x-4y) + \frac{3y}{\sqrt{xy^2-1}} - \frac{3xy^3}{2(xy^2-1)^{3/2}}.\end{aligned} \tag{1.130}$$

■

Using this, we can define the gradient of a scalar field as a field of vectors whose components are partial derivatives of the scalar field, which in two dimensions gives us :

$$(\underline{\mathrm{grad}}A)(x,y) \doteqdot \begin{pmatrix} \frac{\partial A}{\partial x}(x,y) \\ \frac{\partial A}{\partial y}(x,y) \end{pmatrix}. \tag{1.131}$$

We can also generalize this for an arbitrary dimension :

Définition 1.15 *Gradient of a scalar field* The gradient of a scalar field is a field of vectors defined by

$$\underline{\mathrm{grad}}A \doteqdot \frac{\partial A}{\partial x_i}\underline{e}_i. \tag{1.132}$$

R From this point on we will generally omit the explicit dependencies between fields and coordinates, in order to simplify the notation. For example, we will write A instead of $A(x,y)$.

■ **Exemple 1.5** Consider that $A(x,y) = 3x^2 + 2y^2$. Calculating the gradient gives us

$$\underline{\mathrm{grad}}A = \begin{pmatrix} \frac{\partial}{\partial x}(3x^2+2y^2) \\ \frac{\partial}{\partial y}(3x^2+2y^2) \end{pmatrix} = \begin{pmatrix} 6x \\ 4y \end{pmatrix}. \tag{1.133}$$

■

Théorème 1.14 The gradient of a scalar field is locally orthogonal with the contour lines, or isolines, of that field.

Preuve. The variation of A is given by its differential :

$$\begin{aligned}\mathrm{d}A &= \frac{\partial A}{\partial x}\mathrm{d}x + \frac{\partial A}{\partial x}\mathrm{d}y \\ &= \underline{\mathrm{grad}}A \cdot \mathrm{d}\underline{x}.\end{aligned} \tag{1.134}$$

But with the isolines of A, by definition $\mathrm{d}A = 0$ when we move away from $\mathrm{d}\underline{x}$. ■

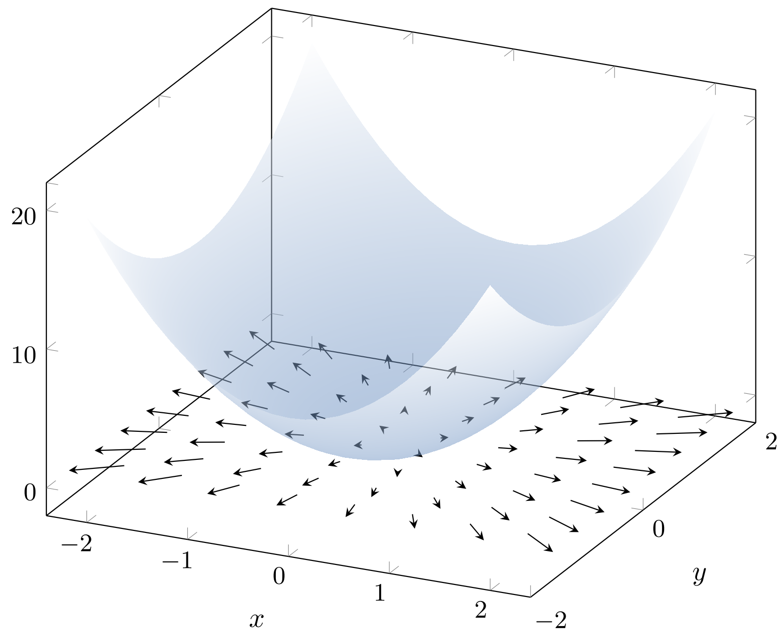


FIGURE 1.5 – graph of the function $A(x,y) = 3x^2 + 2y^2$ and its gradient field (1.133).

Figure 1.5 shows the gradient field for example (1.133). We can see that the vectors move away from the local minimum of the function, a general property consistent with what we have seen above :

Théorème 1.15 The vectors which make up the gradient field of a scalar field tend towards their local maxima, and away from their local minima.

The divergence of a vector field is defined, in two dimensions, by

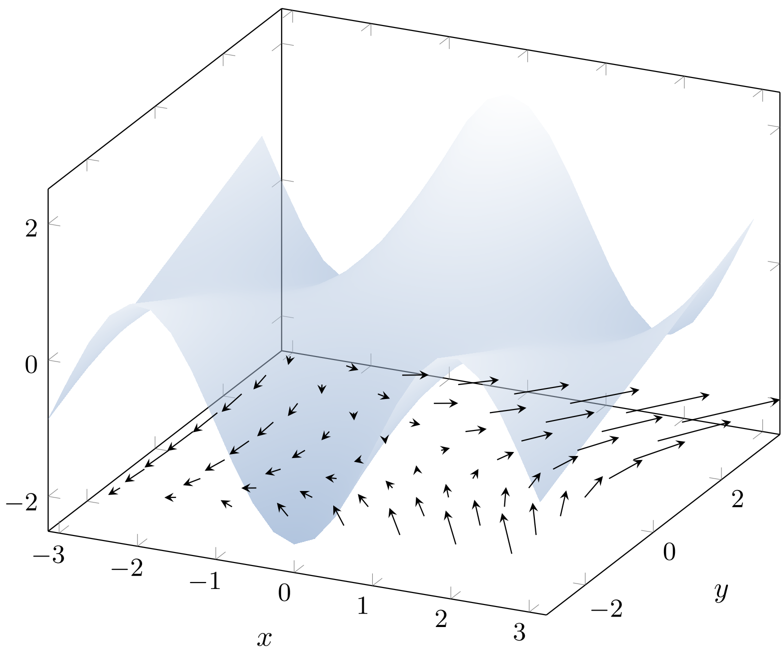
$$\mathrm{div}\,\underline{A} \doteqdot \frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y}, \tag{1.135}$$

which can be generalized for an arbitrary dimension :

Définition 1.16 *Divergence in a vector field.* Divergence in a vector field is in fact a scalar field :

$$\mathrm{div}\,\underline{A} \doteqdot \frac{\partial A_i}{\partial x_i}. \tag{1.136}$$

As the name suggests, the divergence of a vector field that diverges from a point M is positive at point M, and vice versa.



Graph of the vector field (1.137) and its divergence (1.138).

Exercice 1.25 Calculate the divergence of the following vector field :

$$\underline{A}(x,y) = \begin{pmatrix} \frac{1}{6}xy + \frac{1}{2}y + y\sin y \\ x + \frac{1}{4}y^2\cos x \end{pmatrix}. \tag{1.137}$$

Solution 1.25

$$\begin{aligned} \operatorname{div}\underline{A}(x,y) &= \frac{\partial}{\partial x}\left(\tfrac{1}{6}xy + \tfrac{1}{2}y + y\sin y\right) + \frac{\partial}{\partial y}\left(x + \tfrac{1}{4}y^2\cos x\right) \\ &= \tfrac{1}{6}y + \tfrac{1}{2}y\cos x. \end{aligned} \tag{1.138}$$

■

Figure 1.3.1 shows the gradient field for example (1.137). These definitions can be extended to higher-order tensors. For the purpose of this course on fluid mechanics, we will stick to the following two operators :

Définition 1.17 *Gradient operators and divergence in higher orders.* The gradient of a vector field and the divergence of a matrix field are defined, respectively, by

$$\underline{\underline{\text{grad}}}\,\underline{A} \doteq \begin{pmatrix} \frac{\partial A_x}{\partial x} & \frac{\partial A_x}{\partial y} \\ \frac{\partial A_y}{\partial x} & \frac{\partial A_y}{\partial y} \end{pmatrix} = \frac{\partial A_i}{\partial x_j}\underline{e}_i \otimes \underline{e}_j, \tag{1.139}$$

et

$$\underline{\operatorname{div}}\,\underline{\underline{A}} \doteq \begin{pmatrix} \frac{\partial A_{xx}}{\partial x} + \frac{\partial A_{xy}}{\partial y} \\ \frac{\partial A_{yx}}{\partial x} + \frac{\partial A_{yy}}{\partial y} \end{pmatrix} = \frac{\partial A_{ij}}{\partial x_j}\underline{e}_i. \quad (1.140)$$

Broadly speaking, gradient increases by one unit (and divergence reduces by one unit) the order of a tensor field.

Exercice 1.26 Calculate the gradient of a vector field (1.137). Calculate the divergence of the following matrix field :

$$\underline{\underline{B}}(x,y) = \begin{pmatrix} x-y & xy+8y^3 \\ 4x^2y-7 & 9(x+y) \end{pmatrix}. \quad (1.141)$$

Solution 1.26

$$\underline{\underline{\operatorname{grad}}}\,\underline{A}(x,y) = \begin{pmatrix} -5\sin y & -5x\cos y \\ -2\sin 2x + 4y^2 & 8xy \end{pmatrix}, \quad (1.142)$$

$$\underline{\operatorname{div}}\,\underline{\underline{B}}(x,y) = \begin{pmatrix} 1+5x\sin y \\ -4\sin 2x + 8x \end{pmatrix}. \quad (1.143)$$

■

Exercice 1.27 Give a more simple expression of the trace of a vector gradient.

Solution 1.27

$$\operatorname{Tr}\left(\underline{\underline{\operatorname{grad}}}\,\underline{A}\right) = \frac{\partial A_i}{\partial x_i} = \operatorname{div}\underline{A}. \quad (1.144)$$

■

Théorème 1.16 *The chain rule of derivatives* When we change the coordinates $(x,y)^T \to (x',y')^T$, the partial derivatives are transformed by the following rules :

$$\begin{aligned} \frac{\partial A}{\partial x'} &= \frac{\partial A}{\partial x}\frac{\partial x}{\partial x'} + \frac{\partial A}{\partial y}\frac{\partial y}{\partial x'}, \\ \frac{\partial A}{\partial y'} &= \frac{\partial A}{\partial x}\frac{\partial x}{\partial y'} + \frac{\partial A}{\partial y}\frac{\partial y}{\partial y'}. \end{aligned} \quad (1.145)$$

When changing orthonormal basis, i.e. from $\mathscr{B}$ to $\mathscr{B}'$, since the transformation $\underline{X} = \underline{\underline{P}}\,\underline{X}'$ is linear it follows that (1.145) will give us

$$(\underline{\operatorname{grad}}\,A)_{\mathscr{B}'} = \underline{\underline{P}}(\underline{\operatorname{grad}}\,A)_{\mathscr{B}}, \quad (1.146)$$

Which shows that the field $\underline{\text{grad}} A$ transforms in the way we would expect of a vector.

> R The preceding remarks can be applied on a more general scale : the operators introduced here, along with their arbitrary generalizations, transform tensors into tensors, in so far as the gradient and divergence of tensors obey the rule (1.108). For brevity's sake, we cannot get into a demonstration of this important result here.

1.3.2 Manipulating operators

Tensors can be used to perform many operations, particularly for composing various differential operators. For example

$$\text{div}\,\underline{\text{grad}} A = \text{div}\left(\frac{\partial A}{\partial x_i}\underline{e}_i\right) = \frac{\partial}{\partial x_i}\left(\frac{\partial A}{\partial x_i}\right) = \frac{\partial^2 A}{\partial x_i \partial x_i}, \tag{1.147}$$

$$\underline{\text{div}}\,\underline{\underline{\text{grad}}}\,\underline{A} = \underline{\text{div}}\left(\frac{\partial A_i}{\partial x_j}\underline{e}_i \otimes \underline{e}_j\right) = \frac{\partial}{\partial x_j}\left(\frac{\partial A_i}{\partial x_j}\right)\underline{e}_i = \frac{\partial^2 A_i}{\partial x_j \partial x_j}\underline{e}_i.$$

The operator thus obtained is known as a Laplace operator :

Définition 1.18 *The Laplace operator*. The Laplace operator is given the symbol Δ and is defined by

$$\Delta \underline{A}_{(p)} \doteq \frac{\partial^2 \underline{A}_{(p)}}{\partial x_j \partial x_j} \tag{1.148}$$

It is applicable to tensors of all order, and respects order.

> R In the formulae (1.147) we wrote $\frac{\partial^2}{\partial x_i \partial x_i}$ instead of $\frac{\partial^2}{\partial x_i^2}$, making use of Einstein notation. Some documents use the latter form, for simplicity's sake.

Exercice 1.28 Calculate the Laplace operator of $A(x,y) = 8\sin^2 x - 4x^2\cos y$.

Solution 1.28

$$\begin{aligned}
\Delta A &= \frac{\partial^2 A}{\partial x^2} + \frac{\partial^2 A}{\partial y^2} \\
&= \frac{\partial}{\partial x}(8\sin 2x - 8x\cos y) + \frac{\partial}{\partial y}\left(4x^2 \sin y\right) \\
&= (16\cos 2x - 8\cos y) + \left(4x^2\cos y\right) \\
&= 16\cos 2x + \left(4x^2 - 8\right)\cos y.
\end{aligned} \tag{1.149}$$

■

Définition 1.19 *The Del or nabla operator.* This operator is written ∇ and defined by

$$\nabla \underline{A}_{(p)} = \frac{\partial \underline{A}_{(p)}}{\partial x_i} \otimes \underline{e}_i. \tag{1.150}$$

This gives us $\underline{\text{grad}}\,A = \nabla A$, $\text{div}\,\underline{A} = \nabla \cdot \underline{A}$, $\Delta A = \nabla \cdot (\nabla A) = \nabla^2 A$.

We can perform operations involving tensor fields, using the calculation rules set out above.

Exemple 1.6

$$\begin{aligned}\text{div}\,(A\underline{B}) &= \text{div}\,(AB_i\underline{e}_i)\\ &= \frac{\partial}{\partial x_i}(AB_i)\\ &= \frac{\partial A}{\partial x_i}B_i + A\frac{\partial B_i}{\partial x_i}\\ &= (\underline{\text{grad}}\,A)\cdot\underline{B} + A\,\text{div}\,\underline{B}.\end{aligned} \tag{1.151}$$

■

Exercice 1.29 Expand upon the following :

1. $\underline{\text{div}}\,(A\underline{\underline{I}})$,
2. $\underline{\text{grad}}\,(\underline{A}\cdot\underline{B})$,
3. $\text{div}\,(\underline{\underline{A}}\cdot\underline{B})$.

Solution 1.29

$$\begin{aligned}\underline{\text{div}}\,(A\underline{\underline{I}}) &= \underline{\text{div}}\,(A\delta_{ij}\underline{e}_i\otimes\underline{e}_j)\\ &= \frac{\partial(A\delta_{ij})}{\partial x_j}\underline{e}_i\\ &= \frac{\partial A}{\partial x_j}\delta_{ij}\underline{e}_i + A\frac{\partial\delta_{ij}}{\partial x_j}\underline{e}_i\\ &= \frac{\partial A}{\partial x_i}\underline{e}_i\\ &= \underline{\text{grad}}\,A,\end{aligned} \tag{1.152}$$

$$\begin{aligned}\underline{\text{grad}}\,(\underline{A}\cdot\underline{B}) &= \underline{\text{grad}}\,(A_iB_i)\\ &= \frac{\partial(A_iB_i)}{\partial x_j}\underline{e}_j\\ &= \frac{\partial A_i}{\partial x_j}B_i\underline{e}_j + A_i\frac{\partial B_i}{\partial x_j}\underline{e}_j\\ &= \underline{B}^T\cdot\underline{\underline{\text{grad}}}\,\underline{A} + \underline{A}^T\cdot\underline{\underline{\text{grad}}}\,\underline{B},\end{aligned} \tag{1.153}$$

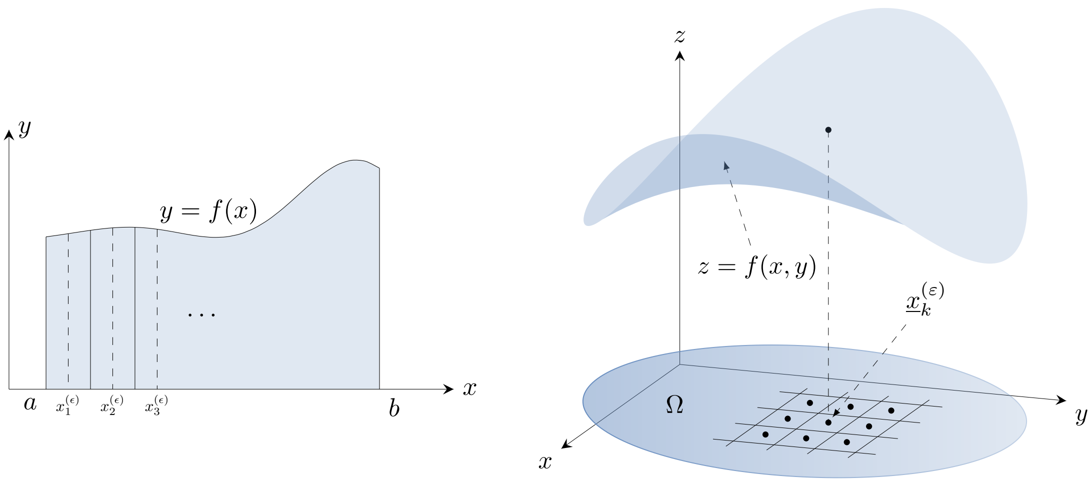


FIGURE 1.6 – Illustration of the definition of integrals : $n = 1$ left, $n = 2$ right.

$$\begin{aligned} \operatorname{div}\left(\underline{\underline{A}} \cdot \underline{B}\right) &= \operatorname{div}\left(A_{ij} B_j \underline{e}_i\right) \\ &= \frac{\partial}{\partial x_i}\left(A_{ij} B_j\right) \\ &= \frac{\partial A_{ij}}{\partial x_i} B_j + A_{ij} \frac{\partial B_j}{\partial x_i} \\ &= \left(\operatorname{div} \underline{\underline{A}}\right) \cdot \underline{B} + \left(\underline{\underline{\operatorname{grad}}}\, \underline{B}\right) : \underline{\underline{A}}^T . \end{aligned} \tag{1.154}$$

■

1.3.3 Integrals

Before getting onto the topic of multidimensional integrals, let us first remind ourselves of the definition, and three well-known properties, of one-dimensional integrals.

Définition 1.20 *Integral of a single variable function*. For a real function $A(x)$ to be integrated in a segment $[a, b]$, this is defined by

$$\begin{aligned} \int_a^b A(x)\, dx &\doteqdot \lim_{\varepsilon \to 0} \left(\sum_{k=0}^{N^{(\varepsilon)}-1} A\left(x_k^{(\varepsilon)}\right) \varepsilon \right), \\ N^{(\varepsilon)} &= \left\lfloor \frac{b-a}{\varepsilon} \right\rfloor, \\ x_k^{(\varepsilon)} &= a + k\varepsilon . \end{aligned} \tag{1.155}$$

In this simplified definition, x_k run through the segment $[a, b]$ in regular fashion, starting from a, as seen in Figure 1.6.

Théorème 1.17

$$\int_a^b \frac{\mathrm{d}A}{\mathrm{d}x}\mathrm{d}x = [A]_a^b \doteqdot A(b) - A(a)\,. \tag{1.156}$$

Preuve. We start by writing this trivial formula :

$$\sum_{k=0}^{N^{(\varepsilon)}-1}\left[A\left(x_{k+1}^{(\varepsilon)}\right) - A\left(x_k^{(\varepsilon)}\right)\right] = A\left(x_{N^{(\varepsilon)}}^{(\varepsilon)}\right) - A\left(x_0^{(\varepsilon)}\right), \tag{1.157}$$

Which we then multiply and divide by ε before making it tend to zero :

$$\lim_{\varepsilon\to 0}\left(\sum_{k=0}^{N^{(\varepsilon)}-1}\frac{A\left(x_k^{(\varepsilon)}+\varepsilon\right) - A\left(x_k^{(\varepsilon)}\right)}{\varepsilon}\varepsilon\right) = A\left(\lim_{\varepsilon\to 0} x_{N^{(\varepsilon)}}^{(\varepsilon)}\right) - A(a)\,. \tag{1.158}$$

On the left side we now have the derivative of A ; to the right, the last point x_k gets closer to b when ε decreases. So :

$$\lim_{\varepsilon\to 0}\left(\sum_{k=0}^{N^{(\varepsilon)}-1}\frac{dA}{dx}\left(x_k^{(\varepsilon)}\right)\varepsilon\right) = A(b) - A(a)\,, \tag{1.159}$$

And the left side is indeed the definition of the integral of $\frac{\mathrm{d}A}{\mathrm{d}x}$. ■

Théorème 1.18 *Integration by parts.*

$$\int_a^b \frac{\mathrm{d}A}{\mathrm{d}x}B(x)\mathrm{d}x = [AB]_a^b - \int_a^b A(x)\frac{\mathrm{d}B}{\mathrm{d}x}\mathrm{d}x. \tag{1.160}$$

Preuve. This useful formula, which can be used to calculate many integrals, is derived from the previous theorem and the formula which gives us the derivative of a product :

$$\frac{\mathrm{d}AB}{\mathrm{d}x} = \frac{\mathrm{d}A}{\mathrm{d}x}B + A\frac{\mathrm{d}B}{\mathrm{d}x}. \tag{1.161}$$

By applying (1.156) to (1.161) we get (1.160). ■

Exercice 1.30 Calculate the following integral :

$$I \doteqdot \int_0^2 \ln x\,\mathrm{d}x. \tag{1.162}$$

Solution 1.30 A tip which often works : make use of the fact that $A(x) = 1 \times A(x) = \frac{\mathrm{d}x}{\mathrm{d}x}A(x)$:

$$\begin{aligned} I &= \int_0^2 \tfrac{\mathrm{d}x}{\mathrm{d}x} \ln x \,\mathrm{d}x \qquad (1.163)\\ &= [x \ln x]_0^2 - \int_0^2 x \frac{\mathrm{d}\ln x}{\mathrm{d}x}\mathrm{d}x \\ &= 2\ln 2 - \int_0^2 x \frac{1}{x}\mathrm{d}x \\ &= 2\ln 2 - 2. \end{aligned}$$

■

Théorème 1.19 *Changing variables.* Consider a new variable X defined by $X \doteqdot \phi^{-1}(x)$, where ϕ is a continuous function of the reciprocal ϕ^{-1}. For any function $A(x)$ we thus have

$$\int_a^b A(x)\mathrm{d}x = \int_{\phi^{-1}(a)}^{\phi^{-1}(b)} A(\phi(X))\phi'(X)\mathrm{d}X, \qquad (1.164)$$

où $\phi'(X) = \frac{\mathrm{d}\phi}{\mathrm{d}X}$ est la dérivée de $\phi(X)$.

Preuve. This formula can be obtained by writing $\mathrm{d}x = \frac{\mathrm{d}x}{\mathrm{d}X}\mathrm{d}X = \phi'(X)\mathrm{d}X$. ■

Exercice 1.31 Calculate the following integral :

$$I \doteqdot \int_0^{\pi/2} \ln\cos x \,\mathrm{d}x. \qquad (1.165)$$

Solution 1.31 Change the variable $X \doteqdot \frac{\pi}{2} - x$:

$$\begin{aligned} I &= \int_{\pi/2}^0 \ln\cos\left(\tfrac{\pi}{2} - X\right)(-1)\,\mathrm{d}X \qquad (1.166)\\ &= \int_0^{\pi/2} \ln\sin X \,\mathrm{d}X, \end{aligned}$$

Then add the two integrals :

$$\begin{aligned} 2I &= \int_0^{\pi/2} \ln\cos x\,\mathrm{d}x + \int_0^{\pi/2} \ln\sin X\,\mathrm{d}X \qquad (1.167)\\ &= \int_0^{\pi/2} \ln(\sin x\cos x)\,\mathrm{d}x \\ &= \int_0^{\pi/2} (\ln\sin 2x - \ln 2)\,\mathrm{d}x. \end{aligned}$$

Now change the variable $X \doteqdot 2x$:

$$\begin{aligned} 2I &= \int_0^{\pi} (\ln\sin X)\,\tfrac{1}{2}\mathrm{d}X - \tfrac{\pi}{2}\ln 2 \qquad (1.168)\\ &= \tfrac{1}{2}\int_0^{\pi/2} \ln\sin X\,\mathrm{d}X + \tfrac{1}{2}\int_{\pi/2}^{\pi} \ln\sin X\,\mathrm{d}X - \tfrac{\pi}{2}\ln 2. \end{aligned}$$

The first integral is equal to I, as per (1.166). The second integral, thanks to the last change of variable $Y \doteqdot X - \frac{\pi}{2}$, is also equal to I. We can deduce that $I = -\frac{\pi}{2}\ln 2$. ■

This can be generalized to a field of arbitrary dimension. For example, the (double) integral of a function with $n = 2$ variables $A(x,y)$ in a domain Ω of the plane (x,y), is given by

$$\int_\Omega A(x,y)\,\mathrm{d}x\mathrm{d}y \doteqdot \lim_{\varepsilon\to 0}\left(\sum_j\sum_k A\left(x_j^{(\varepsilon)}, y_k^{(\varepsilon)}\right)\varepsilon^2\right), \tag{1.169}$$

where points (x_j, y_k) run through the integration domain Ω as in Figure 1.6. It should be noted that we are not obliged to used Cartesian coordinates : using polar coordinates (r denoting the distance from the origin and θ the angle between the axis x and the radius vector running from the origin to the point of interest), surface element dxdy becomes $r\mathrm{d}r\mathrm{d}\theta$:

$$\int_\Omega A(x,y)\,\mathrm{d}x\mathrm{d}y = \int_\Omega A(r,\theta)\,r\mathrm{d}r\mathrm{d}\theta. \tag{1.170}$$

Exercice 1.32 Write out the following double integral in polar coordinates, then calculate it :

$$I \doteqdot \int_\Omega \mathrm{e}^{-(x^2+y^2)/2}\mathrm{d}x\mathrm{d}y, \tag{1.171}$$

where Ω is the positive quadrant of the plane ($x \geq 0, y \geq 0$).

Solution 1.32

$$\begin{aligned} I &= \int_0^{\pi/2}\int_0^{+\infty} \mathrm{e}^{-r^2/2} r\mathrm{d}r\mathrm{d}\theta \\ &= \int_0^{\pi/2}\left[-\mathrm{e}^{-r^2/2}\right]_0^{+\infty}\mathrm{d}\theta \\ &= \int_0^{\pi/2} 1\mathrm{d}\theta \\ &= \frac{\pi}{2}. \end{aligned} \tag{1.172}$$

■

It can also be written by separating the Cartesian variables :

$$\begin{aligned} I &= \int_0^{+\infty}\int_0^{+\infty} \mathrm{e}^{-(x^2+y^2)/2}\mathrm{d}x\mathrm{d}y \\ &= \int_0^{+\infty}\left(\int_0^{+\infty}\mathrm{e}^{-x^2/2}\mathrm{d}x\right)\mathrm{e}^{-y^2/2}\mathrm{d}y \\ &= \left(\int_0^{+\infty}\mathrm{e}^{-x^2/2}\mathrm{d}x\right)^2. \end{aligned} \tag{1.173}$$

Identifying (1.172) and (1.173) we find that

$$\int_0^{+\infty}\mathrm{e}^{-x^2/2}\mathrm{d}x = \sqrt{\frac{\pi}{2}}. \tag{1.174}$$

More generally, for a function with n variables, we can proceed in the same manner :

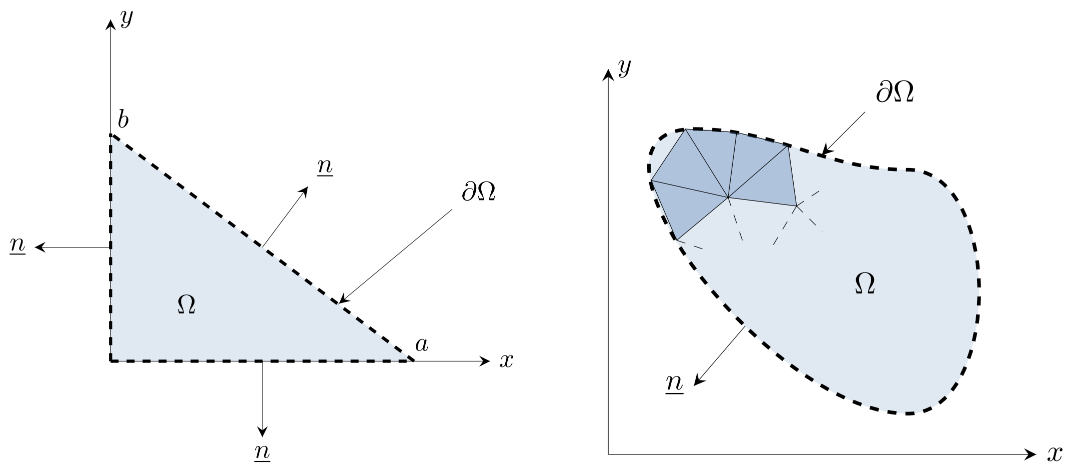


FIGURE 1.7 – Illustration of the proof cited here for the divergence theorem for $n = 2$. In this case $\partial\Omega$ is a curve, for $n = 3$ it is a surface.

Définition 1.21 *Volume integral of a tensor field.* The volume integral of a function with n dimensions is written

$$\int_{\Omega} A\left(\underline{x}, t\right) \mathrm{d}\Omega, \tag{1.175}$$

Where $\mathrm{d}\Omega$ is the infinitesimal limit of ε^n. Integrals of this kind can be defined for tensors of all orders.

■ **Exemple 1.7** An integral represents the sum of an infinite number of contributions in order to estimate an extensive (i.e. additive) macroscopic value whose density varies in space. For example, if ρ denotes the field of volumetric mass density (or specific mass) :

$$\int_{\Omega} \rho\left(\underline{x}, t\right) \mathrm{d}\Omega \quad \text{is the mass contained in } \Omega. \tag{1.176}$$

■

Exercice 1.33 Express the momentum contained in Ω in the form of an integral.

Solution 1.33 Using $\underline{u}$ to denote speed, and omitting spatio-temporal dependencies to simplify the notation :

$$\int_{\Omega} \rho \underline{u} \,\mathrm{d}\Omega \quad \text{is the momentum contained in} \Omega. \tag{1.177}$$

■

In physics, the volume integral is every bit as important as the border integral :

Définition 1.22 *Border integral for a tensor field.* For a volume Ω whose border (or boundary) $\partial\Omega$ is sufficiently regular and admitting of direction, for an arbitrary field A we can define its border integral using $\partial\Omega$:

$$\int_{\partial\Omega} A\left(\underline{x},t\right)\mathrm{d}S = \lim_{\sigma\to 0}\left(\sum_k A\left(\underline{x}_k^{(\sigma)}\right)\sigma\right), \tag{1.178}$$

Where the points $\underline{x}_k$ run across the integration surface, and $\mathrm{d}S$ is the infinitesimal limit of $\sigma^n \sim \varepsilon^{n-1}$, since $\partial\Omega$ is of a lower spatial dimension than Ω.

The following result confirms the general applicability of the theorem 1.17 :

Théorème 1.20 *This is known as the divergence theorem, or Gauss' theorem, or sometimes the Green-Ostrogradski theorem.* For a volume Ω (which may be mobile), $\underline{n}(\underline{x},t)$ denotes the vector which is unitary, orthogonal and exterior to $\partial\Omega$ at every point. So for any vector field $\underline{A}(\underline{x},t)$,

$$\int_{\Omega(t)} \mathrm{div}\,\underline{A}\left(\underline{x},t\right)\mathrm{d}\Omega = \int_{\partial\Omega(t)} \underline{A}\left(\underline{x},t\right)\cdot\underline{n}\left(\underline{x},t\right)\mathrm{d}S. \tag{1.179}$$

Furthermore, this result can be applied to tensors of all orders, as long as we use the appropriate divergence operator. This theorem represents a generalization of (1.156) to arbitrary dimensions.

Preuve. For our purposes we will limit ourselves to $n=2$ and a single vector field. We begin by demonstrating this result for a right-angled triangle $\Omega = \left\{(x,y), x\geqslant 0, y\geqslant 0, \frac{x}{a}+\frac{y}{b}\leqslant 1\right\}$, as in Figure 1.7. To do this, we go round the triangle in two different directions for each of the terms of $\mathrm{div}\,\underline{A}$, and we use 1.17 :

$$\begin{aligned}\int_{\Omega} \mathrm{div}\,\underline{A}\left(\underline{x},t\right)\mathrm{d}\Omega &= \int_0^b\left(\int_0^{a(1-y/b)}\frac{\partial A_x}{\partial x}\mathrm{d}x\right)\mathrm{d}y + \int_0^a\left(\int_0^{b(1-x/a)}\frac{\partial A_y}{\partial y}\mathrm{d}y\right)\mathrm{d}x \\ &= \int_0^b A_x\left(a\left(1-y/b\right),y\right)\mathrm{d}y + \int_0^a A_y\left(x,b\left(1-x/a\right)\right)\mathrm{d}x \\ &\quad -\int_0^a A_y\left(x,0\right)\mathrm{d}x - \int_0^b A_x\left(0,y\right)\mathrm{d}y.\end{aligned} \tag{1.180}$$

In Figure 1.7 it is immediately apparent that

$$\begin{aligned}-\int_0^a A_y\left(x,0\right)dx &= \int_{\partial\Omega_x}\underline{A}\cdot\underline{n}\mathrm{d}S, \\ -\int_0^b A_x\left(0,y\right)dy &= \int_{\partial\Omega_y}\underline{A}\cdot\underline{n}\mathrm{d}S.\end{aligned} \tag{1.181}$$

The last segment $\partial\Omega_{xy}$ of the border has the equation $\frac{x}{a}+\frac{y}{b}=1$, which gives us $b\mathrm{d}x + a\mathrm{d}y = 0$ (compare this with (1.134)). The integration vector and unit normal vector (which is constant here) are thus equal to

$$\begin{aligned}\mathrm{d}S &= \sqrt{\mathrm{d}x^2+\mathrm{d}y^2} = \sqrt{1+\left(\frac{b}{a}\right)^2}\mathrm{d}x = \sqrt{\left(\frac{a}{b}\right)^2+1}\,\mathrm{d}y, \\ \underline{n} &= \frac{1}{\sqrt{a^2+b^2}}\begin{pmatrix} b \\ a\end{pmatrix}.\end{aligned} \tag{1.182}$$

Putting all of this together, we get $\underline{A}\cdot\underline{n}\mathrm{d}S = \frac{bA_x+aA_y}{\sqrt{a^2+b^2}}\mathrm{d}S = A_x\mathrm{d}y + A_y\mathrm{d}x$, then

$$\int_{\partial\Omega_{xy}} \underline{A}\cdot\underline{n}\mathrm{d}S = \int_0^b A_x\left(a\left(1-y/b\right),y\right)\mathrm{d}y + \int_0^a A_y\left(x,b\left(1-x/a\right)\right)\mathrm{d}x, \tag{1.183}$$

Which wraps things up for our right-angled triangle. As every triangle can be split into two right-angled triangles, this result is applicable to all triangles through the additivity of integrals for contiguous and disjointed domains (in this case the vectors $\underline{n}$ are opposites on the shared border) :

$$\int_{\Omega\cup\Omega'} \mathrm{div}\,\underline{A}\mathrm{d}\Omega = \int_{\Omega} \mathrm{div}\,\underline{A}\mathrm{d}\Omega + \int_{\Omega'} \mathrm{div}\,\underline{A}\mathrm{d}\Omega, \tag{1.184}$$
$$\int_{\partial\Omega\cup\partial\Omega'} \underline{A}\cdot\underline{n}\mathrm{d}S = \int_{\partial\Omega} \underline{A}\cdot\underline{n}\mathrm{d}S + \int_{\partial\Omega'} \underline{A}\cdot\underline{n}\mathrm{d}S.$$

In order to deduce that the theorem is valid for any domain Ω, we need to divide the domain into a mosaic of small triangles as in Figure 1.7, where error decreases with size ε (in fact we are not obliged to reduce the size of the triangles in the middle of the domain, just those toward the edges). ■

Exercice 1.34 Ω is a side 1 in the bottom left corner, located at the origin and

$$\underline{A}(x,y) = \begin{pmatrix} xy \\ x^2+y^2 \end{pmatrix}. \tag{1.185}$$

Verify that (1.179) is indeed equal by explicitly calculating both sides.

Solution 1.34 We begin by calculating :

$$\mathrm{div}\,\underline{A} = \frac{\partial\left(xy\right)}{\partial x} + \frac{\partial\left(x^2+y^2\right)}{\partial y} = y + 2y = 3y, \tag{1.186}$$

Which allows us to calculate the volume integral, using (1.156) :

$$\int_{\Omega} \mathrm{div}\,\underline{A}\mathrm{d}\Omega = \int_0^1 \left(\int_0^1 3y\mathrm{d}x\right)\mathrm{d}y = \int_0^1 \left[3yx\right]_0^1 \mathrm{d}y \tag{1.187}$$
$$= \int_0^1 3y\mathrm{d}y = \left[\tfrac{3}{2}y^2\right]_{y=0}^{y=1} = \tfrac{3}{2}.$$

We need to calculate the border integral in segments. For the $\partial\Omega_{x-}$ side, on the axis of x :

$$\int_{\partial\Omega_{x-}} \underline{A}\cdot\underline{n}\mathrm{d}S = \int_0^1 \begin{pmatrix} xy \\ x^2+y^2 \end{pmatrix}_{y=0} \cdot \begin{pmatrix} 0 \\ -1 \end{pmatrix}\mathrm{d}x = -\int_0^1 x^2\mathrm{d}x \tag{1.188}$$
$$= -\left[\tfrac{1}{3}x^3\right]_{x=0}^{x=1} = -\tfrac{1}{3}.$$

On $\partial\Omega_{x+}$, which runs parallel :

$$\int_{\partial\Omega_{x+}} \underline{A}\cdot\underline{n}\mathrm{d}S = \int_0^1 \begin{pmatrix} xy \\ x^2+y^2 \end{pmatrix}_{y=1} \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix}\mathrm{d}x = \int_0^1 \left(x^2+1\right)\mathrm{d}x \tag{1.189}$$
$$= \left[\tfrac{1}{3}x^3 + x\right]_{x=0}^{x=1} = \tfrac{4}{3}.$$

For the $\partial\Omega_{y-}$ side, on the axis of y :

$$\int_{\partial\Omega_{y-}} \underline{A}\cdot\underline{n}\mathrm{d}S = \int_0^1 \begin{pmatrix} xy \\ x^2+y^2 \end{pmatrix}_{x=0} \cdot \begin{pmatrix} -1 \\ 0 \end{pmatrix} \mathrm{d}y = -\int_0^1 0\mathrm{d}y = 0. \tag{1.190}$$

Finally, on $\partial\Omega_{y+}$, which runs parallel :

$$\begin{aligned}\int_{\partial\Omega_{y+}} \underline{A}\cdot\underline{n}\mathrm{d}S &= \int_0^1 \begin{pmatrix} xy \\ x^2+y^2 \end{pmatrix}_{x=1} \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} \mathrm{d}y = \int_0^1 y\mathrm{d}y \\ &= \left[\tfrac{1}{2}x^2\right]_{x=0}^{x=1} = \tfrac{1}{2}.\end{aligned} \tag{1.191}$$

By adding up these values we get the border integral, which is in fact identical to the volume integral :

$$\int_{\partial\Omega} \underline{A}\cdot\underline{n}\mathrm{d}S = -\tfrac{1}{3}+\tfrac{4}{3}+0+\tfrac{1}{2} = \tfrac{3}{2}. \tag{1.192}$$

■

Exercice 1.35 Repeat the same task as in the previous exercise, but this time assume that Ω is a semicircle with a radius of 1 with its centre at the origin and its border diameter on the x axis, and

$$\underline{A}(x,y) = \begin{pmatrix} x^3y+y \\ xy^4-x \end{pmatrix}. \tag{1.193}$$

Solution 1.35 This time we find that

$$\mathrm{div}\underline{A} = \frac{\partial\left(x^3y+y\right)}{\partial x} + \frac{\partial\left(xy^4-x\right)}{\partial y} = 3x^2y+4xy^3, \tag{1.194}$$

therefore

$$\begin{aligned}\int_{\Omega} \mathrm{div}\underline{A}\mathrm{d}\Omega &= \int_{-1}^1 \left(\int_0^{\sqrt{1-x^2}} \left(3x^2y+4xy^3\right) dy\right) \mathrm{d}x = \int_{-1}^1 \left[\tfrac{3}{2}x^2y^2+xy^4\right]_{y=0}^{y=\sqrt{1-x^2}} \mathrm{d}x \\ &= \int_{-1}^1 \left(\tfrac{3}{2}x^2\left(1-x^2\right)+x\left(1-x^2\right)^2\right) \mathrm{d}x \\ &= \left[\tfrac{1}{2}x^2+\tfrac{1}{2}x^3-\tfrac{1}{2}x^4-\tfrac{3}{10}x^5+\tfrac{1}{6}x^6\right]_{x=-1}^{x=1} = \tfrac{2}{5}.\end{aligned} \tag{1.195}$$

The integral for the section of the border on the x axis is

$$\begin{aligned}\int_{\partial\Omega_-} \underline{A}\cdot\underline{n}\mathrm{d}S &= \int_{-1}^1 \begin{pmatrix} x^3y+y \\ xy^4-x \end{pmatrix}_{y=0} \cdot \begin{pmatrix} 0 \\ -1 \end{pmatrix} \mathrm{d}x = \int_{-1}^1 x\mathrm{d}x \\ &= \left[\tfrac{1}{2}x^2\right]_{x=-1}^{x=1} = 0.\end{aligned} \tag{1.196}$$

On the circular section of the border, we can calculate $\underline{n}(x,y)$ as the (normalized) gradient of the function $B(x,y) \doteqdot x^2+y^2-1$. The isolines of B confirm that $dB = xdx + ydy = 0$, so

$$\mathrm{d}S = \frac{1}{\sqrt{1-x^2}}\mathrm{d}x, \tag{1.197}$$
$$\underline{n} = \frac{1}{\sqrt{x^2+y^2}}\begin{pmatrix} x \\ y \end{pmatrix}.$$

Ainsi :

$$\begin{aligned}\int_{\partial\Omega_+} \underline{A}\cdot\underline{n}\mathrm{d}S &= \int_{-1}^{1} \left(\left(\begin{pmatrix} x^3y+y \\ xy^4-x \end{pmatrix}\cdot\frac{1}{\sqrt{x^2+y^2}}\begin{pmatrix} x \\ y \end{pmatrix}\right)\right)_{y=\sqrt{1-x^2}} \frac{1}{\sqrt{1-x^2}}\mathrm{d}x \\ &= \int_{-1}^{1}\left(x^5+x^4-2x^3+x\right)\mathrm{d}x \\ &= \left[\tfrac{1}{6}x^6+\tfrac{1}{5}x^5-\tfrac{1}{2}x^4+\tfrac{1}{2}x^2\right]_{x=-1}^{x=1} = \tfrac{2}{5},\end{aligned}$$

Which allows us to conclude ■

Exercice 1.36 Show that

$$\int_{\partial\Omega} \underline{n}\,\mathrm{d}S = 0. \tag{1.198}$$

Solution 1.36 We must postulate that $\underline{A} = \underline{cst}$ (constant vector) in (1.179). ■

Exercice 1.37 Show that

$$\int_{\Omega} \underline{\mathrm{grad}} A \mathrm{d}\Omega = \int_{\partial\Omega} A\underline{n}\,\mathrm{d}S. \tag{1.199}$$

Solution 1.37 We must use the formula (1.152). The divergence theorem then gives us the result we are looking for, since $\underline{\underline{I}}\,\underline{n} = \underline{n}$. This formula naturally leads us to Archimedes' principle. ■

Integration by parts works in multiple dimensions as well as in one dimension. There are numerous variants of multidimensional integration by parts.

■ **Exemple 1.8** Using (1.151) and (1.179) gives us

$$\begin{aligned}\int_{\Omega} A\,\mathrm{div}\,\underline{B}\,\mathrm{d}\Omega &= \int_{\Omega}\mathrm{div}\,(A\underline{B})\mathrm{d}\Omega - \int_{\Omega}(\underline{\mathrm{grad}}\,A)\cdot\underline{B}\,\mathrm{d}\Omega \\ &= \int_{\partial\Omega} A\underline{B}\cdot\underline{n}\,\mathrm{d}S - \int_{\Omega}(\underline{\mathrm{grad}}\,A)\cdot\underline{B}\,\mathrm{d}\Omega.\end{aligned} \tag{1.200}$$

■

1.3.4 Ordinary differential equations

In physics, we often find that our reasoning does not allow us to explicitly obtain the function we are looking for (atmospheric temperature, speed in a river etc.), but instead a relationship involving derivatives of that function. For example, if $y(x)$ is a value we wish to ascertain, which depends solely upon the variable x, suppose that after a few calculations we end up with the following equation :

$$y' = a, \tag{1.201}$$

where a is a known constant, connected with the parameters of the problem, the prime denoting the derivative (in this example there is just one variable, x). In this case we can simply integrate in x to obtain $y(x) = ax + b$, and the constant b can be determined using a boundary condition, i.e. the result of $y(x = x_0)$ for a suitable x_0. If the variable is time t (instead of x), we call this the initial condition.

But what happens if our calculations lead us to an equation such as this one ?

$$y' = ay. \tag{1.202}$$

In this case we can still solve the problem simply by dividing by y (which we assume to be non-null), leaving us with $(\ln y)' = a$, so $\ln y = ax + b$, or else :

$$y(x) = C\mathrm{e}^{ax}, \tag{1.203}$$

avec $C \doteqdot \mathrm{e}^b$. The equation (1.202) is a first order (because it only contains the first derivative of y) ordinary differential equation (ODE).

Définition 1.23 An ordinary differential equation (ODE) is an equation whose unknown is a function of a single variable, represented by its derivatives. The derivative of the highest order defines the order of the ODE, by default.

■ **Exemple 1.9** The movement of a pendulum (a mass hanging from a string of length ℓ, the other end of which is fixed to a point) obeys the following ODE :

$$\ell\theta'' = -g\sin\theta, \tag{1.204}$$

where $\theta(t)$ is the angle of the string to the vertical axis, a function of time t, and g is the acceleration of gravity. This is a second order ODE, where $\theta''(t)$ plays the role of $y''(x)$. ■

As we have seen, (1.202) is compatible with a family of solutions. The example above was only easy because the ODE in question is linear, which is to say that any linear combination of solutions is also a solution. Moreover, the single coefficient a is a constant, which also makes things easier. Let us summarize what we have found :

Théorème 1.21 *Homogeneous, linear, first-order ODE.* The homogeneous, linear, first-order ODE with the constant coefficient $y' = ay$ allows of a family of solutions $y(x) = C\mathrm{e}^{ax}$ forming a vector line (i.e. they are all proportional. The coefficient C must be solved on a case-by-case basis, using a boundary condition.

The equation (1.202) is regarded as *homogeneous*, because it does not contain any additive terms which are not dependent on the unknown function $y(x)$. Here is an example of a non-homogeneous ODE :

$$y' = ay + c(x), \tag{1.205}$$

where $c(x)$ is an arbitrary, but known, function. Once again we can solve this equation, noting that without this function we would find ourselves with (1.202). It is therefore tempting to look for solutions to (1.205) in the form

$$y(x) = C(x)\mathrm{e}^{ax}. \tag{1.206}$$

Exercice 1.38 Determine the value of $C(x)$ so that (1.206) satisfies (1.205). Deduce from this the solutions to (1.205).

Solution 1.38 By injecting (1.206) into (1.205) we find

$$C(x) = \int_{x_0}^{x} c(\tilde{x})\mathrm{e}^{-a\tilde{x}}\mathrm{d}\tilde{x}, \tag{1.207}$$

where x_0 is a constant (changing it would be equivalent to adding a constant to $C(x)$). As such, the overall solution to (1.205) is given by

$$y(x) = \mathrm{e}^{ax}\int_{x_0}^{x} c(\tilde{x})\mathrm{e}^{-a\tilde{x}}\mathrm{d}\tilde{x}. \tag{1.208}$$

■

Exercice 1.39 Solve $y' = -y + x$ with the boundary condition $y(0) = 1$.

Solution 1.39 This ODE takes the form (1.205) (first order linear, non-homogeneous), where $a = -1$ and $c(x) \doteqdot x$. The solution (1.208) can be written

$$\begin{aligned} y(x) &= \mathrm{e}^{-x}\int_{x_0}^{x} \tilde{x}\mathrm{e}^{\tilde{x}}\mathrm{d}\tilde{x} \\ &= \mathrm{e}^{-x}\left((x-1)\mathrm{e}^{x} + D\right) \\ &= x - 1 + D\mathrm{e}^{-x}. \end{aligned} \tag{1.209}$$

We can directly verify that this general solution is consistent with the original ODE. We set the constant D by determining that $y(0) = 1$, or $D = 2$:

$$y(x) = x - 1 + 2\mathrm{e}^{-x}. \tag{1.210}$$

■

First order ODEs are used fairly frequently in physics, but we also encounter second order ODEs. For a linear case with constant coefficients, a second order ODE would look like this :

$$y'' = ay' + by. \tag{1.211}$$

The example (1.204) given above is non-linear on account of the sine function, but we can linearize it by focusing only on the small values of angle θ (small oscillations) : $\ell\theta'' = -g\theta$, which falls into the category (1.211) where $a = 0, b = -g/\ell$. Building on our experience with first order ODEs, we can try our luck at finding solutions in the form $y(x) = \mathrm{e}^{\lambda x}$, with λ to be determined. Injecting this ansatz into (1.211) gives us a second degree polynomial equation in λ :

$$\lambda^2 - a\lambda - b = 0, \tag{1.212}$$

with the discriminant $\Delta = a^2 + 4b$. There are three possible options here :

— If $\Delta > 0$, (1.212) has two real roots λ_1 and λ_2, and we have two real solutions $y_1(x) \doteqdot \mathrm{e}^{\lambda_1 x}$, $y_2(x) \doteqdot \mathrm{e}^{\lambda_2 x}$.

— If $\Delta = 0$, (1.212) has a real double root $\lambda = \frac{a}{2}$, and we can obtain a first real solution, $y_1(x) \doteqdot \mathrm{e}^{\lambda x}$. But we then realize that $y_2(x) \doteqdot x\mathrm{e}^{\lambda x}$ is also the solution to the ODE.

— If $\Delta < 0$, (1.212) has two complex conjugated roots $\lambda_1 \pm \mathrm{i}\lambda_2$, and the two solutions can be recombined (by linearity) to give two real solutions $y_1(x) \doteqdot \mathrm{e}^{\lambda_1 x}\cos(\lambda_2 x)$, $y_2(x) \doteqdot \mathrm{e}^{\lambda_1 x}\sin(\lambda_2 x)$.

There are thus two real solutions to this problem, and even a “double infinity” of solutions (a 2nd-dimension vector space) by means of linear combination :

$$y(x) = C_1 y_1(x) + C_2 y_2(x). \tag{1.213}$$

But, for the time being, we do not know if there are other types of solutions to (1.211). To answer that question, let us consider a solution $y(x)$, and let us define the *Wronskian* of the ODE for the pair (y_1, y) :

$$W(x) \doteqdot \begin{vmatrix} y_1(x) & y(x) \\ y_1'(x) & y'(x) \end{vmatrix} = y_1(x)y'(x) - y(x)y_1'(x). \tag{1.214}$$

Exercice 1.40 Demonstrate that W follows a first order linear ODE. Deduce from this a formula which explicitly gives the value of $W(x)$ (to one undetermined multiplication constant).

Solution 1.40 Let us now calculate the derivative of the Wronskian :

$$\begin{aligned} W'(x) &= y_1(x)y''(x) - y(x)y_1''(x) \\ &= y_1(x)(ay'(x) + by(x)) - y(x)(ay_1'(x) + by_1(x)) \\ &= aW(x). \end{aligned} \tag{1.215}$$

W thus complies with the equation (1.202). From what we already know, we can deduce that :

$$W(x) = C\mathrm{e}^{ax}. \tag{1.216}$$

■

According to the definition (1.214) of W, this gives us

$$y'(x) - \lambda_1 y(x) = C\exp((a - \lambda_1)x), \tag{1.217}$$

Which is a first-order, non-homogeneous ODE. As such, the overall solution to (1.208) is

$$\begin{aligned} y(x) &= \exp(\lambda_1 x)\int_{x_0}^{x} \exp((a - 2\lambda_1)\tilde{x})\mathrm{d}\tilde{x} \\ &= C_1\exp(\lambda_1 x) + C_2\exp((a - \lambda_1)x). \end{aligned} \tag{1.218}$$

We can also see that the sum of the roots of the polynomial (1.212) is a, which means that $a - \lambda_1 = \lambda_2$, and thus $y(x)$ does indeed take the form (1.213). So :

Théorème 1.22 *Homogeneous, linear, second-order ODE.* Let λ_1, λ_2 be the roots (potentially overlapping) of the trinomial $\lambda^2 - a\lambda - b$. The homogeneous, linear, second-order ODE with constant coefficients $y'' = ay' + by$ allows for a family of solutions $y(x) = C_1 e^{\lambda_1 x} + C_2 e^{\lambda_2 x}$ if $a^2 + 4b \neq 0$, or $y(x) = (C_1 + C_2 x)e^{\lambda_1 x}$ otherwise. These solutions thus always constitute a vector plane. Coefficients C_1, C_2 must be solved on a case-by-case basis, using two boundary conditions.

R The trinomial $\lambda^2 - a\lambda - b$ is known as the *characteristic polynomial* of the second order ODE. As for the boundary conditions, in most cases they involve y and/or y'.

Exercice 1.41 Solve the ODE $y'' + y = 0$ where $y(0) = 1$ and $y'(0) = 2$.

Solution 1.41 In this case the characteristic trinomial can be written $\lambda^2 + 1$ $(a = 0, b = -1)$. Its roots are $\lambda_{1,2} = \pm \mathrm{i}$. The general solution is thus

$$\begin{aligned} y(x) &= C_1 e^{\mathrm{i}x} + C_2 e^{-\mathrm{i}x} \\ &= D_1 \cos x + D_2 \sin x. \end{aligned} \tag{1.219}$$

The constants which satisfy the boundary conditions are $D_1 = 1$ and $D_2 = 2$. This type of equation typically applies to non-dissipative and unforced linear oscillators, as in the example of a pendulum performing small oscillations for which we can overlook the effect of friction. ■

R There are many other types of ODE : higher order linear, non-homogeneous, non-constant coefficients, non-linear etc. When working with fields with multiple variables, a system of equations connecting the partial derivatives of the fields in question is known as a partial differential equation (PDE).

the Fourier Transform

Once again, our aim here is not to exhaustively define the context of this tool, but rather to list its properties without too much mathematical detail.

1.3.5 Definition and properties

Définition 1.24 *Fourier transform.* Taking a function $A(x)$ with a real variable x, its Fourier transform with respect to x (if it has been determined) is the following function of k :

$$\mathscr{F}_x[A(x)](k) \doteq \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} A(x) e^{-\mathrm{i}kx} \,\mathrm{d}x. \tag{1.220}$$

■ **Exemple 1.10** So the function $P(x) = 1$ if $|x| \leq \frac{1}{2}$, or 0 if not (known as the *rectangular or gate* *function*, Figure 1.8). The Fourier transform is fairly easy to calculate :

$$\begin{aligned}
\mathscr{F}_x[P(x)](k) &= \tfrac{1}{\sqrt{2\pi}} \int_{-1/2}^{1/2} \mathrm{e}^{-\mathrm{i}kx}\,\mathrm{d}x \\
&= \tfrac{1}{\sqrt{2\pi}} \left[-\frac{\mathrm{e}^{-\mathrm{i}kx}}{\mathrm{i}k} \right]_{-1/2}^{1/2} \\
&= \tfrac{1}{\sqrt{2\pi}} \frac{-\mathrm{e}^{-\mathrm{i}k/2} + \mathrm{e}^{\mathrm{i}k/2}}{\mathrm{i}k} \\
&= \tfrac{1}{\sqrt{2\pi}} \frac{\sin(k/2)}{k/2}.
\end{aligned} \tag{1.221}$$

■

For the next step, we will make use of a tool which is very important in mathematics for physicists :

Définition 1.25 *The Dirac delta function.* It is written $\delta(x)$ and defined as follows :

$$\delta(x) \doteqdot \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} P\left(\frac{x}{\varepsilon}\right), \tag{1.222}$$

where P is the rectangular function. This definition is strange, because we divide by ε before making it tend to zero. As we shall see, δ is not an ordinary function. Consider the following integral :

$$I \doteqdot \int_{-\infty}^{+\infty} \phi(x)\delta(x)\,\mathrm{d}x, \tag{1.223}$$

where ϕ is an arbitrary but continuous function. With a little manipulation (particularly changing the variable $x \doteqdot \varepsilon X$), we get

$$\begin{aligned}
I &= \int_{-\infty}^{+\infty} \phi(x) \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} P\left(\frac{x}{\varepsilon}\right) \mathrm{d}x \qquad (1.224) \\
&= \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \int_{-\infty}^{+\infty} \phi(x) P\left(\frac{x}{\varepsilon}\right) \mathrm{d}x \\
&= \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \int_{-\varepsilon/2}^{+\varepsilon/2} \phi(x)\,\mathrm{d}x. \\
&= \lim_{\varepsilon \to 0} \int_{-1/2}^{+1/2} \phi(\varepsilon X)\,\mathrm{d}X. \\
&= \int_{-1/2}^{+1/2} \phi(0)\,\mathrm{d}X. \\
&= \phi(0). \qquad (1.225)
\end{aligned}$$

This is the true definition of the Dirac delta function : a « function » which, when we multiply it by another function then integrate the real line, gives the value of the other function at zero. We can thus see that it is not a standard function, since formally speaking it is null everywhere and infinite at zero (see Figure 1.8).

Exercice 1.42 Calculate the Fourier transform of $\delta(x)$.

Solution 1.42

$$\begin{aligned}\mathscr{F}_x[\delta(x)](k) &= \tfrac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}\delta(x)\mathrm{e}^{-\mathrm{i}kx}\,\mathrm{d}x \\ &= \tfrac{1}{\sqrt{2\pi}},\end{aligned} \tag{1.226}$$

since $\mathrm{e}^0 = 1$. ■

We can generalize the above by changing the variable $X \doteqdot x - a$:

$$\begin{aligned}\mathscr{F}_x[\delta(x-a)](k) &= \tfrac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}\delta(x-a)\mathrm{e}^{-\mathrm{i}kx}\,\mathrm{d}x \\ &= \tfrac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}\delta(X)\mathrm{e}^{-\mathrm{i}k(X+a)}\,\mathrm{d}X \\ &= \tfrac{1}{\sqrt{2\pi}}\mathrm{e}^{-\mathrm{i}ka},\end{aligned} \tag{1.227}$$

Which gives us (1.226) for $a = 0$.

Let us now look at some of the simple properties of the Fourier transform (demonstrated in the exercises) :

Théorème 1.23 *Essential properties of the Fourier transform.*

— The Fourier transform is linear.
— A real, even function has a real, even Fourier transform.
— A real, odd function has an odd but purely imaginary Fourier transform.
— Translation :

$$\mathscr{F}_x[A(x-a)](k) = \mathrm{e}^{\mathrm{i}ka}\mathscr{F}_x[A(x)](k). \tag{1.228}$$

— Dilatation :

$$\mathscr{F}_x[A(ax)](k) = \frac{1}{|a|}\mathscr{F}_x[A(x)]\left(\frac{k}{a}\right). \tag{1.229}$$

N.B. (1.228) replicates (1.227) from (1.226). Now we come to a very important, and more subtle, property of the Fourier transform.

Théorème 1.24 *Derivative of the transform, and inversely.* For a function $A(x)$ which tends to zero with two infinites :

$$\frac{\mathrm{d}}{\mathrm{d}k}\mathscr{F}_x[A(x)](k) = -\mathrm{i}\mathscr{F}_x[xA(x)](k), \tag{1.230}$$

$$\mathscr{F}_x\left[\frac{\mathrm{d}}{\mathrm{d}x}A(x)\right](k) = \mathrm{i}k\mathscr{F}_x[A(x)](k). \tag{1.231}$$

Preuve. Derivation beneath the integral (which refers to another variable) :

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}k}\mathscr{F}_x[A(x)](k) &= \frac{1}{\sqrt{2\pi}}\frac{\mathrm{d}}{\mathrm{d}k}\int_{-\infty}^{+\infty} A(x)\mathrm{e}^{-\mathrm{i}kx}\,\mathrm{d}x \\ &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} A(x)\frac{\mathrm{d}}{\mathrm{d}k}\mathrm{e}^{-\mathrm{i}kx}\,\mathrm{d}x \\ &= -\frac{1}{\sqrt{2\pi}}\mathrm{i}\int_{-\infty}^{+\infty} xA(x)\mathrm{e}^{-\mathrm{i}kx}\,\mathrm{d}x \\ &= -\mathrm{i}\mathscr{F}_x[xA(x)](k).\end{aligned} \tag{1.232}$$

Then, integrating by parts (with $\lim_{x\to\pm\infty}A(x)=0$) :

$$\begin{aligned}\mathscr{F}_x\left[\frac{\mathrm{d}}{\mathrm{d}x}A(x)\right](k) &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}\frac{\mathrm{d}}{\mathrm{d}x}A(x)\mathrm{e}^{-\mathrm{i}kx}\,\mathrm{d}x \\ &= -\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} A(x)\frac{\mathrm{d}}{\mathrm{d}x}\mathrm{e}^{-\mathrm{i}kx}\,\mathrm{d}x \\ &= \frac{1}{\sqrt{2\pi}}\mathrm{i}k\int_{-\infty}^{+\infty} A(x)\mathrm{e}^{-\mathrm{i}kx}\,\mathrm{d}x \\ &= \mathrm{i}k\mathscr{F}_x[A(x)](k)\end{aligned} \tag{1.233}$$

■

Exercice 1.43 We define the *Heaviside distribution* $H(x)=1$ if $x\geq 0$, or 0 otherwise (see Figure 1.8). We can easily verify that its derivative exhibits Dirac distribution. Calculate the Fourier transform of $H(x-a)$.

Solution 1.43 We use (1.227) and (1.231) :

$$\begin{aligned}\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\mathrm{i}ka} &= \mathscr{F}_x[\delta(x-a)](k) \\ &= \mathscr{F}_x\left[\frac{\mathrm{d}}{\mathrm{d}x}H(x-a)\right](k) \\ &= \mathrm{i}k\mathscr{F}_x[H(x-a)](k),\end{aligned} \tag{1.234}$$

hence

$$\mathscr{F}_x[H(x-a)](k) = \frac{1}{\sqrt{2\pi}}\frac{\mathrm{e}^{-\mathrm{i}ka}}{\mathrm{i}k}. \tag{1.235}$$

■

All of which brings us to the most important and most remarkable property of the Fourier transform. You can find full demonstrations of this property online - they are too long to be included here :

Définition 1.26 *The inverse Fourier transform.* For a given function $A(x)$ we have :

$$A(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}\mathscr{F}_x[A(x)](k)\mathrm{e}^{\mathrm{i}kx}\,\mathrm{d}k. \tag{1.236}$$

In other words, the Fourier transform can be inverted and its inverse is written

$$\mathscr{F}_k^{-1}[B(k)](x) \doteqdot \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} B(k)\mathrm{e}^{\mathrm{i}kx}\,\mathrm{d}k. \tag{1.237}$$

Note ths + sign in the exponential, contrary to the original transform.

■ **Exemple 1.11** We can verify this with the function $\frac{1}{\sqrt{2\pi}}\frac{\sin(k/2)}{k/2}$ which appears in (1.221) as a transform of the rectangular function. The first observation is that (1.235) gives us

$$\mathscr{F}_k^{-1}\left[\frac{\mathrm{e}^{-\mathrm{i}ka}}{\mathrm{i}k}\right] = \sqrt{2\pi}H(x-a). \tag{1.238}$$

So :

$$\begin{aligned} \mathscr{F}_k^{-1}\left[\frac{1}{\sqrt{2\pi}}\frac{\sin(k/2)}{k/2}\right](x) &= \frac{1}{\sqrt{2\pi}}\mathscr{F}_k^{-1}\left[\frac{1}{\mathrm{i}k}\left(\mathrm{e}^{\mathrm{i}k/2}-\mathrm{e}^{-\mathrm{i}k/2}\right)\right](x) \\ &= H(x+\tfrac{1}{2})-H(x-\tfrac{1}{2}), \end{aligned} \tag{1.239}$$

which coincides with $P(x)$, as becomes clear in graph form. ■

Exercice 1.44 Using (1.227), calculate the Fourier transform of $\mathrm{e}^{\mathrm{i}k_0x}$.

Solution 1.44 (1.227) gives us

$$\begin{aligned} \delta(x-a) &= \frac{1}{\sqrt{2\pi}}\mathscr{F}_k^{-1}\left[\mathrm{e}^{-\mathrm{i}ka}\right](x) \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty}\mathrm{e}^{\mathrm{i}k(x-a)}\,\mathrm{d}k. \end{aligned} \tag{1.240}$$

Now let us repeat that formula with some changes $(x,k,a) \leftarrow (k_0,x,k)$:

$$\begin{aligned} \sqrt{2\pi}\delta(k_0-k) &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\mathrm{e}^{\mathrm{i}k_0x}\mathrm{e}^{-\mathrm{i}kx}\,\mathrm{d}x \\ &= \mathscr{F}_x\left[\mathrm{e}^{\mathrm{i}k_0x}\right](k). \end{aligned} \tag{1.241}$$

■

This result is extremely important, and leads us to the following theorem :

Théorème 1.25 *Spectral decomposition.* The Fourier transform of a function allows us to ascertain its oscillation behaviour.

Preuve. For a harmonic function with wavenumber k_0, i.e. $\mathrm{e}^{\mathrm{i}k_0x}$, (1.241) shows that the transform exhibits Dirac distribution, with an infinite peak at $k=k_0$, and null value elsewhere. By means of linearity, the Fourier transform for a discrete sum of oscillations has as many Dirac peaks as it does oscillation modes, with each indicating a wavenumber. For an arbitrary function, the Fourier transform is continuous and may be interpreted as a spectrum of the initial function, a bit like the spectrum of a light or sound signal. Figure 1.8 shows a few examples. ■

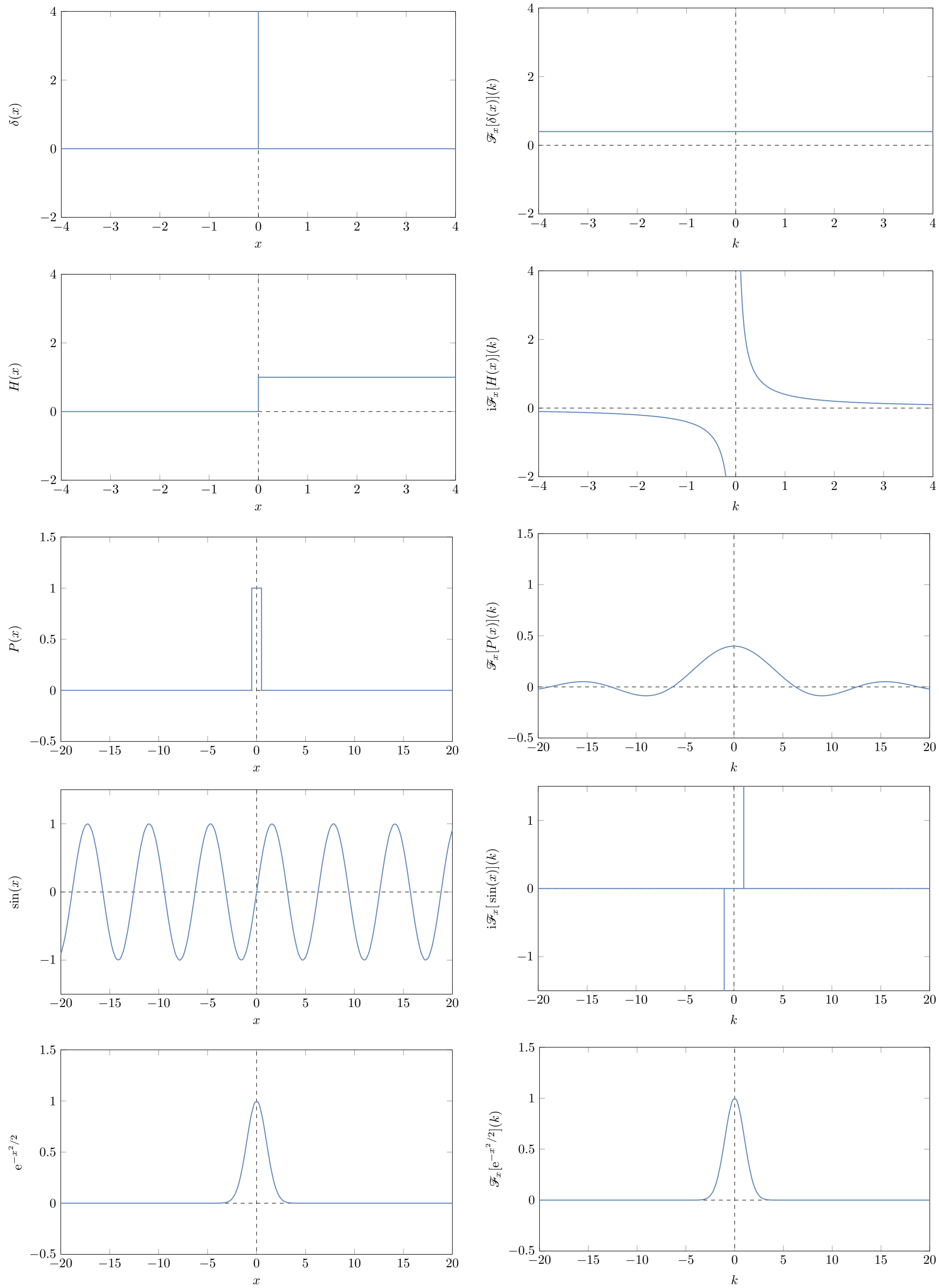


FIGURE 1.8 – The graphs on the right hand side represent the Fourier transforms (or their product with i) of the functions on the left. Top to bottom : Dirac distributions, Heaviside function, rectangular function, sine, Gaussian. N.B. The scale may change from one line to the next.

The Fourier transform has many other properties, which cannot all be listed here. You will find various tables online listing the transforms of standard functions. Care is needed here because there are various notation conventions in use (at least three of them are considered standard; for our purposes I have used the one I consider to be most useful). Remember also that the Fourier transform can be applied in multiple spatial dimensions, or with reference to time.

1.3.6 Applications

The Fourier transform is closely related to linear differential equations. Now let us take a look at some linear ODEs, returning to the problem we saw in Exercise (1.41) : $y'' + y = 0$ where $y(0) = 1$ and $y'(0) = 2$. Let us begin by applying the Fourier transform to the ODE :

$$\mathscr{F}_x\left[\frac{\mathrm{d}^2 y}{\mathrm{d}x^2}(x)\right](k) + \mathscr{F}_x[y(x)](k) = 0. \tag{1.242}$$

Writing $Y(k) \doteqdot \mathscr{F}_x[y(x)](k)$ and using (1.231) :

$$(\mathrm{i}k)^2 Y(k) + Y(k) = 0, \tag{1.243}$$

or

$$(1-k^2)Y(k) = 0. \tag{1.244}$$

As such, $Y(k)$ is non-null if and only if $k = \pm 1$; it is thus a linear combination of $\delta(x-1)$ and $\delta(x+1)$:

$$Y(k) = C^- \delta(k-1) + C^+ \delta(k+1). \tag{1.245}$$

Taking the inverse transform of this result, we can use (1.241) to find a linear combination of $\mathrm{e}^{\mathrm{i}x}$ and $\mathrm{e}^{-\mathrm{i}x}$, i.e. a linear combination of $\cos x$ and $\sin x$ as above. The boundary conditions do the rest.

We can thus use the Fourier transform to solve ODEs, because the Fourier transforms derivatives into multiplications. N.B. it is not always this easy. For example, resolving the problem (1.39) soon proves to be difficult because the solution does not have a well-defined Fourier transform. Moreover, this method is not always easier (or quicker) than the direct method detailed in the preceding paragraph, especially if we are seeking to solve non-homogeneous equations. We shall see later on that the Fourier transform is particularly well-suited to solving certain equations involving partial derivatives.

Exercice 1.45 Use two methods to solve the general ODE for a linear, damped, free pendulum : $y'' + 2\gamma y' + \omega^2 y = 0$ (in physical terms, γ is a friction coefficient and ω a characteristic frequency; we can presume that $\omega \neq 0$).

Solution 1.45 Using the direct method : the characteristic polynomial is $\lambda^2 - 2\gamma\lambda - \omega^2 = 0$, with the roots

$$\lambda_\pm = \gamma \pm \sqrt{\gamma^2 + \omega^2}. \tag{1.246}$$

They are distinct because $\omega \neq 0$, and the theorem 1.22 gives us the general solution :

$$y(x) = C_+ \mathrm{e}^{\lambda_+ x} + C_- \mathrm{e}^{\lambda_- x}. \tag{1.247}$$

Using the Fourier transform : establishing $Y(k) \doteqdot \mathscr{F}_x[y(x)](k)$, the transform of the ODE can be written

$$(-k^2 + 2\mathrm{i}\gamma k + \omega^2)Y(k) = 0. \tag{1.248}$$

So $Y(k)$ is non-null if and only if $k^2 - 2\mathrm{i}\gamma k - \omega^2 = 0$, so :

$$k = \mathrm{i}\left(2\gamma \pm \sqrt{\gamma^2 + \omega^2}\right) \doteqdot k_{1,2}. \tag{1.249}$$

In other words :

$$Y(k) = C_1\delta(k - k_1) + C_2(k - k_2), \tag{1.250}$$

Which gives us the expected solution if we apply the inverse Fourier transform, by means of a calculation analogous to (1.227). ■

Let us conclude this paragraph by noting that ODEs may render a similar service to Fourier transforms : the best way to calculate certain transforms is to write an appropriate ODE. The most famous example is that of the Gaussian function :

Théorème 1.26 *The transformation of a Gaussian function is another example.*

$$\mathscr{F}_x\left[\mathrm{e}^{-x^2/2}\right](k) = \mathrm{e}^{-k^2/2}. \tag{1.251}$$

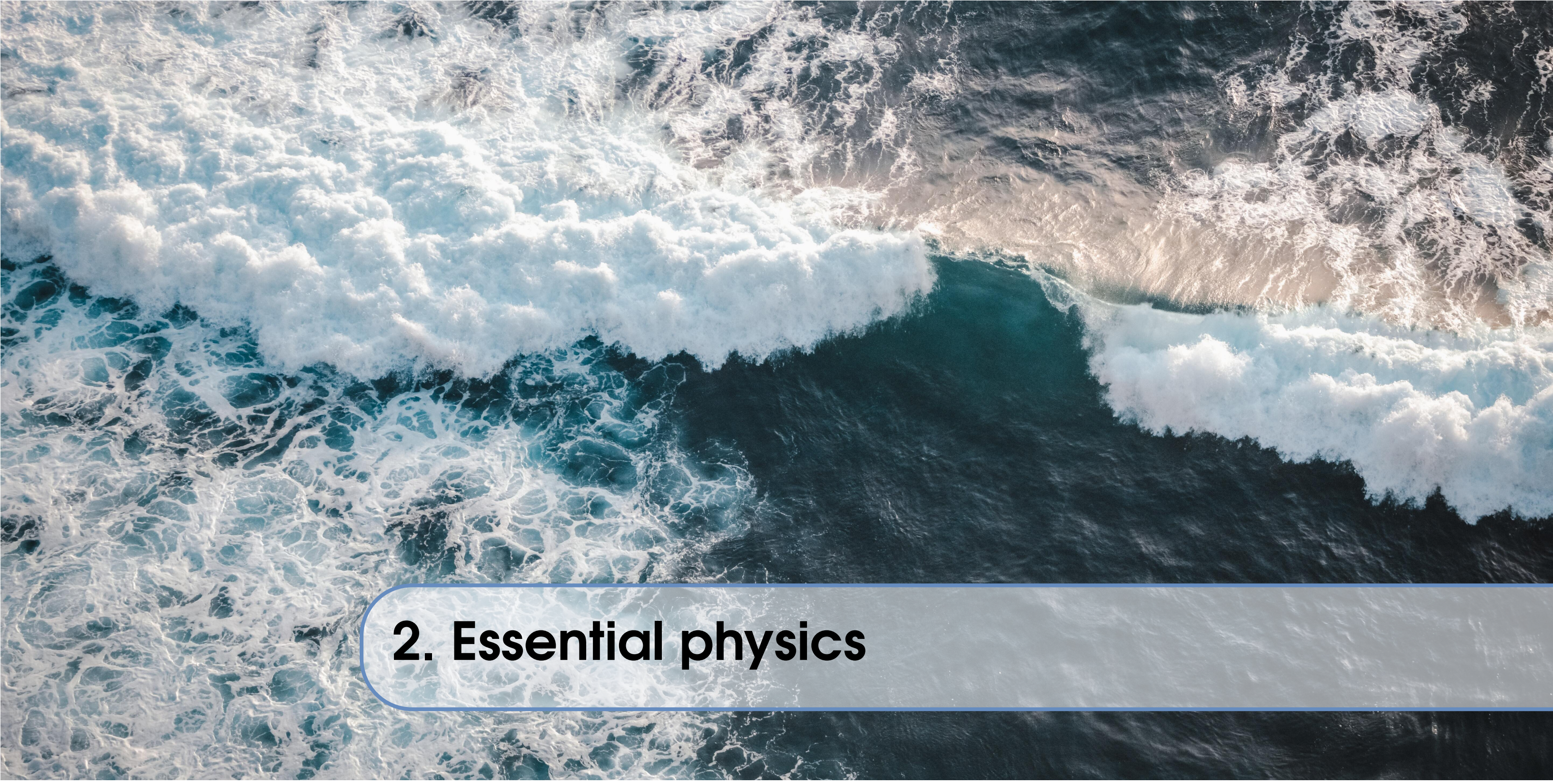
Preuve. Let us write $\phi(x) \doteqdot \mathrm{e}^{-x^2/2}$. We can begin by noting that this confirms the ODE $\phi' + x\phi = 0$ (check it for yourself !). With the help of this ODE, we can derive its transform using (1.230) and (1.231) :

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}k}\mathscr{F}_x[\phi(x)](k) &= -\mathrm{i}\mathscr{F}_x[x\phi(x)](k) \\ &= \mathrm{i}\mathscr{F}_x\left[\frac{\mathrm{d}\phi}{\mathrm{d}x}(x)\right](k) \\ &= -k\mathscr{F}_x[\phi(x)](k).\end{aligned}$$

As such, noting that $\Phi(k) \doteqdot \mathscr{F}_x[\phi(x)](k)$, this function satisfies the ODE $\Phi' + k\Phi = 0$, where the prime denotes the derivative in relation to k. This is the same which satisfies ϕ. Since this is a homogeneous first order ODE, we know from theorem 1.21 that the solutions are all proportional). To deduce the result we are looking for, $\Phi = \phi$, we still need to demonstrate that these two functions coincide at a given point, e.g. That $\Phi(0) = \phi(0) = 1$. This is true, because

$$\begin{aligned}\Phi(0) = \mathscr{F}_x\left[\mathrm{e}^{-x^2/2}\right](0) &= \tfrac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} \mathrm{e}^{-x^2/2}\,\mathrm{d}x \\ &= 1,\end{aligned} \tag{1.252}$$

thanks to (1.174). ■



2. Essential physics

2.1 Mechanics of continuous media

2.1.1 The Cauchy stress tensor

Mechanics is the study of the movement of a system with reference to the causes that determine this movement, i.e. the forces at work within the system. If we take a section of a continuous medium in its entirety, and apply the fundamental law of mechanics (mass $\times$ acceleration = the sum of external forces) we need to establish the sum total of the external forces acting upon the medium. These come in two varieties : external fields and forces arising from contact with contiguous media.

— External fields are created by the action of force fields (usually gravitational or electromagnetic) governed by distant systems, so that it is not unreasonable to talk of action at a distance, even though this is not strictly the case. For the purpose of this course on fluid mechanics, we will focus exclusively on the earth's gravitational field, represented by its acceleration $\underline{g}$. As for the gravity specific to the medium, i.e. the gravitational forces that the different sections of the medium exert on one another, these are negligible in most cases and will be treated as such in this course.

— Contact forces with the contiguous medium are heavily dependent on the nature of the material, particularly its molecular structure (elastic solid, liquid, gas, plasma etc.). To generalize, we can represent these forces using a vector field known as stress vectors $\underline{T}$. If we consider a point $\underline{r}$ in a medium at a given moment in time t, along with an imaginary plane containing this point and oriented by the unit normal vector $\underline{n}$ (Figure 2.1), locally and virtually separating the medium into two sides, then the force exerted by one of these two sides on the other, by unit of mass, is written $\underline{T}(\underline{x},t,\underline{n})$.

By convention, $\underline{T}(\underline{x},t,\underline{n})$ is exerted *by* the side towards which $\underline{n}$ is pointing on the other side, so the action-reaction principle gives us $\underline{T}(\underline{x},t,-\underline{n})$ for the force exerted *on* the side towards which $\underline{n}$ is pointed, by the other side.

Now consider a material right-angled tetrahedron, with three orthogonal axes selected as axes of the frame (Figure 2.1), and with sides ε. The fundamental law of dynamics, applied to this material

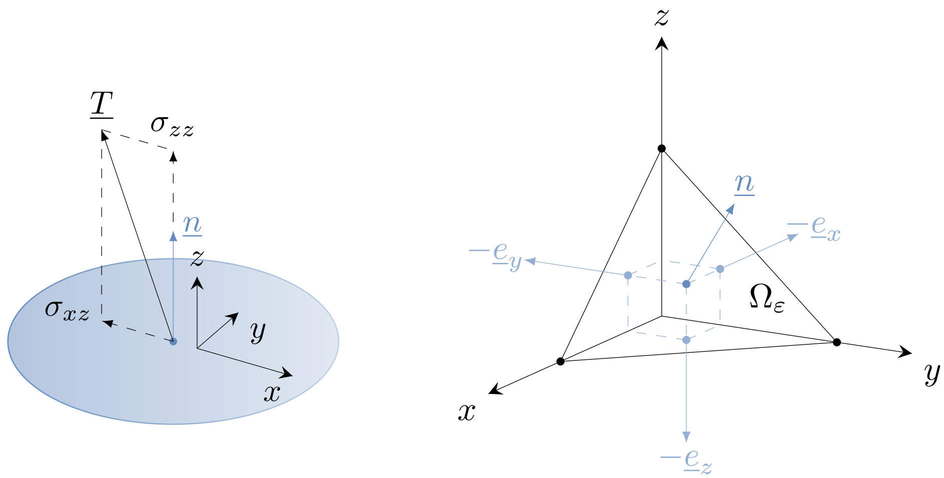


FIGURE 2.1 – Illustration of the definition of the vector $\underline{T}$ (left) and the Cauchy *tetrahedron lemma* (right).

element Ω_ε, can thus be written

$$\text{mass}(\varepsilon) \times \text{acceleration} = \int_\Omega \rho \underline{g} \mathrm{d}\Omega + \int_{\partial\Omega_\varepsilon} \underline{T}(\underline{x}, t, \underline{n}) \mathrm{d}S. \tag{2.1}$$

In keeping with the above, we note that the external forces are the sum of two contributing factors : external fields (volume integral) and contact forces (border integral). Our goal now is to make ε tend to zero ; in doing so, the mass of the tetrahedron and the integral for the gravitational force decrease with ε^3, as they are proportional to the material volume. However, the final integral decreases like ε^2 because it varies with the surface of the material element. We can thus deduce that the border integral must be null [1] :

$$\int_{\partial\Omega_\varepsilon} \underline{T}(\underline{n}) \mathrm{d}S = 0. \tag{2.2}$$

N.B. We have omitted the explicit dependence in relation to $\underline{x}$ and t for simplicity's sake. As for $\underline{x}$, we must now assume that it no longer varies, because the diminutive size of the tetrahedron allows us to assimilate all values of $\underline{x}$ with the position of the tetrahedron's centre of mass. If we now break down the integral into four contributions, one for each face, then we can assume the stress vector to be constant for each of them, giving us

$$S\underline{T}(\underline{n}) + S_i \underline{T}(-\underline{e}_i) = 0, \tag{2.3}$$

using Einstein notation. This essentially means that the system in question is in mechanical equilibrium. The three faces borne by the planes (x,y), (y,z) and (z,x) have as their exterior normal vectors the opposites of the basis vectors $\underline{e}_i$, and their areas can be noted S_i, whereas the face whose

1. To be more precise, it must be an infinitely small value of orderε^3, which implies that the principal contribution, in terms of ε^2, must be null.

three summits are shared among the axes has an exterior normal vector $\underline{n}$ and an area S (Figure 2.1). Let us begin by noting that, by virtue of the law of action and reaction, we can see that

$$\underline{T}(\underline{n}) = \frac{S_i}{S}\underline{T}(\underline{e}_i), \tag{2.4}$$

While the equation (1.198) gives us

$$\underline{n} = \frac{S_i}{S}\underline{e}_i. \tag{2.5}$$

We can now introduce the following tensor :

$$\underline{\underline{\sigma}} \doteq \underline{T}(\underline{e}_i)\otimes\underline{e}_i, \tag{2.6}$$

Once again, using Einstein notation.

Exercice 2.1 Write (2.6) in expanded form.

Solution 2.1 Bearing in mind that the $k-$th coordinate of $\underline{e}_i$ is equal to δ_{ik}, the components of $\underline{\underline{\sigma}}$ can be written :

$$\begin{aligned}\underline{\underline{\sigma}} &= T_j(\underline{e}_i)\delta_{ik}\underline{e}_j\otimes\underline{e}_k \\ &= T_j(\underline{e}_k)\underline{e}_j\otimes\underline{e}_k,\end{aligned} \tag{2.7}$$

Or, in expanded form :

$$\underline{\underline{\sigma}} = \begin{pmatrix} T_x(\underline{e}_x) & T_x(\underline{e}_y) & T_x(\underline{e}_z) \\ T_y(\underline{e}_x) & T_y(\underline{e}_y) & T_y(\underline{e}_z) \\ T_z(\underline{e}_x) & T_z(\underline{e}_y) & T_z(\underline{e}_z) \end{pmatrix}, \tag{2.8}$$

as if we had concatenated the three vectors $\underline{T}(\underline{e}_i)$ written in columns. ■

R We are now in dimension $n = 3$, or $n = 2$ if the flow in question exhibits translation invariance in a given direction, and follows the perpendicular plane. We often use the indices x, y, z instead of $i = 1, 2, 3$.

Now let us multiply (2.6) by $\underline{n}$, successively making use of (2.5), (1.116) and (2.4) :

$$\begin{aligned}\underline{\underline{\sigma}}\cdot\underline{n} &= (\underline{n}\cdot\underline{e}_i)\underline{T}(\underline{e}_i) \\ &= \frac{S_j}{S}(\underline{e}_j\cdot\underline{e}_i)\underline{T}(\underline{e}_i) \\ &= \frac{S_j}{S}\delta_{ij}\underline{T}(\underline{e}_i) \\ &= \frac{S_i}{S}\underline{T}(\underline{e}_i) \\ &= \underline{T}(\underline{n}).\end{aligned} \tag{2.9}$$

This result is valid at all points and all times, and thus deserves to be elevated to the rank of Theorem :

Théorème 2.1 *Cauchy Theorem.* At all points $\underline{x}$ within a continuous medium, and at all times t, the stress vector associated with an arbitrary normal vector $\underline{n}$ is a linear function of the latter, which is to say that there is a tensor $\underline{\underline{\sigma}}(\underline{x},t)$ as there is for any $\underline{n}$:

$$\underline{T}(\underline{n}) = \underline{\underline{\sigma}}(\underline{x},t)\cdot\underline{n}. \tag{2.10}$$

The field $\underline{\underline{\sigma}}(\underline{x},t)$ is known as the Cauchy stress tensor field.

According to the established definition, we can see that each component σ_{ij} represents the force (by surface unit) along axis i exerted on a small imaginary plane oriented by axis j. As such, on the left-hand graph in Figure 2.1, σ_{zz} represents compression or traction, while σ_{xz} is friction as per x (we have positioned the local frame so that the friction from σ_{yz} is null). This gives us the result :

Définition 2.1 *Compression/traction and shearing.* At a given point and moment in time, the diagonal values of $\underline{\underline{\sigma}}(\underline{x},t)$ represent forces of compression or traction, depending on their sign (by convention we assign $+$ to traction and $-$ to pressure [a]). As for the extradiagonal components, they represent friction forces known as *shearing*.

a. The opposite is true in soil mechanics.

Note that this demonstration of the existence of $\underline{\underline{\sigma}}$ is similar to our demonstration of the divergence theorem (using a tetrahedron, instead of a triangle, to work in three dimensions). This is because the divergence theorem is a generalized form of the law of action and reaction, or rather the complementarity of flows. To understand this, we can draw a comparison with a simpler case : a heat balance equation. If we posit that at thermal equilibrium the sum total of the quantities of heat Q_i passing through each face of the tetrahedron is null, then :

$$Q(\underline{n}) = \frac{S_i}{S}Q(\underline{e}_i), \tag{2.11}$$

Which, with (2.5), leads us to

$$Q(\underline{n}) = -\underline{q}\cdot\underline{n}, \tag{2.12}$$

where $\underline{q} \doteqdot -Q(\underline{e}_i)\underline{e}_i$ is the heat flow vector. These formulae (2.10) and (2.12) lead us, when integrated into a macroscopic domain, to the following equalities :

$$\int_{\partial\Omega} \underline{T}\,\mathrm{d}S = \int_{\Omega} \underline{\mathrm{div}\,\underline{\underline{\sigma}}}\,\mathrm{d}\Omega, \tag{2.13}$$

$$\int_{\partial\Omega} Q\,\mathrm{d}S = \int_{\Omega} \mathrm{div}\,\underline{q}\,\mathrm{d}\Omega. \tag{2.14}$$

The similarities between these two formulae are obvious ; this is due to the fact that, in both cases, internal contributions (forces, or the amount of heat transferred) cancel themselves out through asymmetry (the law of action and reaction for forces), so the volume integrals are akin to border integrals. This is equivalent to saying that a divergent term only acts by means of a flow through the boundary of the domain in question. This important observation will be explored further in the next paragraph.

2.1.2 Balance equations

A balance equation encapsulates the evolution over time of a value integrated into a domain, for example the temporal variation of the total concentration of oxygen in a lake. The principal balance equations used in fluid mechanics concern mass, momentum and energy. For mass, for example, we can write that within a fixed domain Ω (i.e. of unchanging volume), the mass of an incompressible fluid remains constant. The result is that anything which enters this volume must be matched by that which comes out. But the flow of material passing through a small surface element $\mathrm{d}S$ within a given unit of time, dictated by the output vector $\underline{n}$ is $-\rho\underline{u}\cdot\underline{n}\mathrm{d}S$, ρ and $\underline{u}(\underline{x},t)$ represents the density and velocity of the fluid at the point in question. The chosen sign will result in a positive flow if the speed is opposed to $\underline{n}$, i.e. if the fluid enters the domain Ω locally, so the integral of this flow gives us the temporal variation of the mass contained within the domain :

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{\Omega}\rho\,\mathrm{d}\Omega = -\int_{\partial\Omega}\rho\underline{u}\cdot\underline{n}\mathrm{d}S. \tag{2.15}$$

But since the mass contained in Ω is constant, its derivative is null. Using the divergence theorem to transform the right side of the equation, we get :

$$\int_{\Omega}\mathrm{div}\,(\rho\underline{u})\mathrm{d}S = 0. \tag{2.16}$$

Since this relation holds true for any imaginary volume Ω, the integrand must be null everywhere : $\mathrm{div}\,(\rho\underline{u}) = 0$. Furthermore, if the fluid is incompressible its density is constant. Using (1.151) avec $A=\rho$ and $\underline{B}=\underline{u}(\underline{x},t)$, we can obtain the following result :

Théorème 2.2 *Continuity equation for incompressible flow.* For a fluid flow (or movement in a continuous medium) without modification of the density (i.e. incompressible and with no variable concentration of the substance), velocity must obey, at all points and all moments, the following continuity equation :

$$\mathrm{div}\,\underline{u} = 0. \tag{2.17}$$

As we can see, a bit of simple reasoning - with the help of some of the mathematical formulae we covered in Chapter 1 - can help us to establish laws for the values we wish to determine, which ultimately serve as equations governing the temporal evolution of those values. Let us briefly consider the case of momentum. Writing (2.18) for an arbitrary domain Ω and using (2.13), we find that

$$\int_{\Omega}\rho\frac{\mathrm{d}\underline{u}}{\mathrm{d}t}\mathrm{d}\Omega = \int_{\Omega}\left(\underline{\mathrm{div}}\,\underline{\underline{\sigma}}+\rho\underline{g}\right)\mathrm{d}\Omega. \tag{2.18}$$

Here again, as for any Ω , we can take away the integrals and we are left with a relation which is valid at all points and all moments in time :

Théorème 2.3 *The Cauchy momentum equation.* For any continuous medium, the variation in velocity, at all points and all moments, is given by

$$\frac{\mathrm{d}\underline{u}}{\mathrm{d}t} = \frac{1}{\rho}\underline{\mathrm{div}\,\underline{\underline{\sigma}}} + \underline{g}. \tag{2.19}$$

At this point it is important to note that, in the preceding demonstration, $\underline{\underline{\sigma}}$ appears beneath a divergence, like the flow of heat in (2.14). It is thus a flow of momentum. It should be noted that conventional notation means that it is effectively $-\underline{\underline{\sigma}}$ which represents the flow. Let us remember that :

Définition 2.2 *The notion of flow.* Within a tensor field, when we write the equation representing its temporal derivative, the value shown beneath a divergence represents the flow of the field, and is a tensor of the order just above : 1) heat flow (associated with temperature, a scalar field) is the vector field $\underline{q}$; 2) the flow of momentum (vector field) is the matrix field $-\underline{\underline{\sigma}}$. Flow in a field $\underline{A}_{(p)}$ simply spatially redistributes this field over time, without effect on the integral evolution of $\underline{A}_{(p)}$ in an insulated domain. If the domain is open, the evolution of this integral can be attributed exclusively to the integral of $\underline{A}_{(p)} \cdot \underline{n}$ on its boundary.

A flow generally constitutes an additional unknown ; it is therefore necessary to find a formula to express the flow of a value $\underline{A}_{(p)}$ with respect to $\underline{A}_{(p)}$, in order to close the system of equations, i.e. to have the same number of unknowns and equations. Such formulae are known as *constitutive equations*, and can only be adequately constructed with proper reference to empirical and/or heuristic considerations. In the case of $\underline{\underline{\sigma}}$, the constitutive equation defines the behaviour of a material in relation to other materials, or else a material in a given state (e.g. fluid in turbulent conditions).

As such, $\underline{\underline{\sigma}}$ cannot be calculated without serious reflection, as we shall see in class. Nevertheless, on a very general level, a kinetic movement balance would allow us to demonstrate that almost all materials, [2] have symmetrical stress tensors, which is to say that they are equal to their respective transposes :

$$\underline{\underline{\sigma}} = \underline{\underline{\sigma}}^T, \quad \text{soit} \quad \forall i, j, \; \sigma_{ij} = \sigma_{ji}. \tag{2.20}$$

In spite of this, the system of equations (2.17) and (2.19) still contains more unknowns than equations, hence the need for a constitutive equation.

R As it is symmetrical and real, the stress tensor (for a given point in space and time) can be diagonalized in an orthonormal basis (cf. Chapter 1). The eigenvalues are also known as *principal constraints*.

Exercice 2.2 Consider the system (2.17) and (2.19). Compare the number of (scalar) equations and the number of unknown fields.

Solution 2.2 We have 4 equations in 3 dimensions, since the equation (2.19) is a vector equation. At constant density (as per the continuity equation), we have 3 unknowns for speed and 6 unknown constraints, since $\underline{\underline{\sigma}}$ is symmetrical. This gives us 9 unknowns and just 4 equations. ■

2. with the exception of piezoelectric materials.

2.2 Fluid mechanics

2.2.1 Navier-Stokes Equations

Fluids, like other continuous media, abide by the Cauchy momentum equation (2.19). As we will see in class, the constitutive equation for the stress tensor, for a viscous fluid in an incompressible state, can be written :

$$\underline{\underline{\sigma}} = -p\underline{\underline{I}} + \mu\left(\underline{\underline{\text{grad}}}\,\underline{u} + (\underline{\underline{\text{grad}}}\,\underline{u})^T\right), \tag{2.21}$$

where μ is a constant specific to the fluid at a given temperature, known as dynamic molecular viscosity. The term $-p\underline{\underline{I}}$ is isotropic, which is to say rotation-invariant ; it is a pressure force (the name we will give to the new field $p(\underline{x},t)$), on account of the negative sign. The proportional term μ is a matrix for shearing, i.e. friction.

Exercice 2.3 Write the stress tensor (2.21) in expanded form.

Solution 2.3

$$\underline{\underline{\sigma}} = \begin{pmatrix} -p+2\mu\dfrac{\partial u_x}{\partial x} & \mu\dfrac{\partial u_x}{\partial y}+\mu\dfrac{\partial u_y}{\partial x} & \mu\dfrac{\partial u_x}{\partial z}+\mu\dfrac{\partial u_z}{\partial x} \\ \mu\dfrac{\partial u_y}{\partial x}+\mu\dfrac{\partial u_x}{\partial y} & -p+2\mu\dfrac{\partial u_y}{\partial y} & \mu\dfrac{\partial u_y}{\partial z}+\mu\dfrac{\partial u_z}{\partial y} \\ \mu\dfrac{\partial u_z}{\partial x}+\mu\dfrac{\partial u_x}{\partial z} & \mu\dfrac{\partial u_z}{\partial y}+\mu\dfrac{\partial u_y}{\partial z} & -p+2\mu\dfrac{\partial u_z}{\partial z} \end{pmatrix}, \tag{2.22}$$

and we can see that it is indeed symmetrical. ∎

Exercice 2.4 Calculate the divergence of (2.21). Use (1.147) and (1.152).

Solution 2.4 Since μ is constant, it can be taken out of the divergence operator. By linearity :

$$\underline{\text{div}}\,\underline{\underline{\sigma}} = -\underline{\text{div}}\left(p\underline{\underline{I}}\right) + \mu\underline{\text{div}}\left(\underline{\underline{\text{grad}}}\,\underline{u}\right) + \mu\underline{\text{div}}\left((\underline{\underline{\text{grad}}}\ \underline{u})^T\right). \tag{2.23}$$

The first of these three terms is equal to $-\underline{\text{grad}}\,p$, on account of the formula (1.152). The second term can be obtained by means of the Laplacian (1.147) :

$$\underline{\text{div}}\left(\underline{\underline{\text{grad}}}\,\underline{u}\right) = \Delta\underline{u}. \tag{2.24}$$

To get the third and final term we need to return to the definitions of our operators :

$$\begin{aligned}\underline{\text{div}}\left((\underline{\underline{\text{grad}}}\ \underline{u})^T\right) &= \underline{\text{div}}\left(\frac{\partial u_j}{\partial x_i}\underline{e}_i\otimes\underline{e}_j\right) = \frac{\partial}{\partial x_j}\left(\frac{\partial u_j}{\partial x_i}\right)\underline{e}_i = \frac{\partial}{\partial x_i}\left(\frac{\partial u_j}{\partial x_j}\right)\underline{e}_i \\ &= \frac{\partial}{\partial x_i}\left(\text{div}\,\underline{u}\right).\end{aligned} \tag{2.25}$$

Ultimately, this means that we have inverted the two partial derivatives, and since $\text{div}\,\underline{u} = 0$ (for incompressible flow) this term is actually null ; it remains :

$$\underline{\text{div}}\,\underline{\underline{\sigma}} = -\underline{\text{grad}}\,p + \mu\Delta\underline{u}. \tag{2.26}$$

∎

By injecting this into the Cauchy equation (2.19), we obtain an equation for the momentum of an incompressible, viscous fluid. It is important to note that the temporal derivative on the left hand side of (2.19) is a material derivative, which is to say it follows the movement of the material ; as we shall see in class, it can be written

$$\frac{\mathrm{d}\underline{u}}{\mathrm{d}t} = \frac{\partial \underline{u}}{\partial t} + (\underline{\underline{\mathrm{grad}}}\,\underline{u})\underline{u}, \tag{2.27}$$

The non-linear term $(\underline{\underline{\mathrm{grad}}}\,\underline{u})\underline{u}$ is called *inertia*. It represents the fact that a unit of fluid is in movement, carrying with it information regarding its own speed. Add to this the continuity equation (2.17), and we obtain the Navier-Stokes equations :

Théorème 2.4 *Navier-Stokes Equations*. A viscous fluid, in incompressible state, will abide by the following equations :

$$\mathrm{div}\,\underline{u} = 0, \tag{2.28}$$

$$\frac{\partial \underline{u}}{\partial t} + (\underline{\underline{\mathrm{grad}}}\,\underline{u})\underline{u} = -\tfrac{1}{\rho}\underline{\mathrm{grad}}\,p + \nu\Delta\underline{u} + \underline{g}, \tag{2.29}$$

Where $\nu \doteqdot \frac{\mu}{\rho}$ is the kinematic molecular viscosity of the fluid at a given temperature. In expanded form, using $(u,v,w) = (u_x,u_y,u_z)$ and aligning the z axis with the ascending vertical, these equations can be written :

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} = 0, \tag{2.30}$$

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} + v\frac{\partial u}{\partial y} + w\frac{\partial u}{\partial z} = -\tfrac{1}{\rho}\frac{\partial p}{\partial x} + \nu\left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2}\right), \tag{2.31}$$

$$\frac{\partial v}{\partial t} + u\frac{\partial v}{\partial x} + v\frac{\partial v}{\partial y} + w\frac{\partial v}{\partial z} = -\tfrac{1}{\rho}\frac{\partial p}{\partial y} + \nu\left(\frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2} + \frac{\partial^2 v}{\partial z^2}\right), \tag{2.32}$$

$$\frac{\partial w}{\partial t} + u\frac{\partial w}{\partial x} + v\frac{\partial w}{\partial y} + w\frac{\partial w}{\partial z} = -\tfrac{1}{\rho}\frac{\partial p}{\partial z} + \nu\left(\frac{\partial^2 w}{\partial x^2} + \frac{\partial^2 w}{\partial y^2} + \frac{\partial^2 w}{\partial z^2}\right) - g. \tag{2.33}$$

Collectively, these equations form a system of PDEs with initial conditions which depend on the nature of the case, and the boundary conditions. To put it simply (we will look at this in greater detail in class) :

- the pressure is null on a free surface. In reality it its (generally) constant, but pressure is defined to the nearest constant because it comes below the gradient operator,
- speed on one side is equal to the speed of the side (due to impermeability and the adherence of viscous fluid). This means it is null along immobile solid surfaces (as per this definition).

Exercice 2.5 Flow between two infinite, parallel plates, separated by a distance of $2d$. Ignore gravity. Assume that velocity is constant over time, and parallel to the plates. We can also assume that it remains within a plane which we will write (x,y), as the y axis is perpendicular to the plates (the component $w = u_z$ is therefore null) and nothing depends on z ; we can thus work within the plane (x,y). Write the expanded Navier-Stokes equations in the most simple form possible, taking

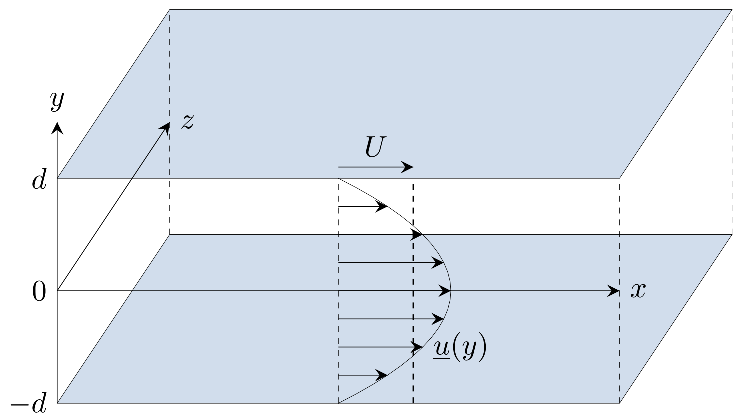


FIGURE 2.2 – Laminar plane Poiseuille flow.

these hypotheses into account. Show that the pressure varies in linear fashion with reference to x, and calculate the spatial distribution of velocity.

Solution 2.5 In the equation (2.30) the last two terms are null because $v = w = 0$, hypothetically. This also shows us that u is not dependent on x, and thus must be solely dependent on y, as per our hypotheses. Based on this result, our three other equations give us :

$$0 = -\tfrac{1}{\rho}\frac{\partial p}{\partial x} + \nu \frac{\partial^2 u}{\partial y^2}, \tag{2.34}$$

$$0 = -\tfrac{1}{\rho}\frac{\partial p}{\partial y}. \tag{2.35}$$

The second of these equations tells us that p does not depend on y, so must be solely dependent on x, according to our hypotheses. The first equation can thus be rewritten as follows, and the partial derivatives become ordinary derivatives because $u(y)$ and $p(x)$ are both single-variable functions :

$$\frac{\mathrm{d}p}{\mathrm{d}x} = \mu \frac{\mathrm{d}^2 u}{\mathrm{d}y^2}. \tag{2.36}$$

Each side of this equation is a function of a different variable from the other side, and thus the two must be constant : we can demonstrate this by deriving (2.36) in relation to x, for example. From this we can deduce that $\frac{\mathrm{d}p}{\mathrm{d}x} = -\alpha$, so pressure varies in a linear fashion according to x : $p(x) = p_0 - \alpha x$. We can then deduce that $\mu \frac{\mathrm{d}^2 u}{\mathrm{d}y^2} = -\alpha$, so speed follows a parabolic trajectory on y, becoming null at both sides. But, to the nearest multiplicative constant, the only parabola which reaches zero at $y = \pm d$ (we locate the origin of y in the middle of the plates) is $d^2 - y^2$. This leaves us with :

$$u(y) = \tfrac{\alpha}{2\mu}(d^2 - y^2). \tag{2.37}$$

This is the velocity profile for *Poiseuille laminar flow*. ■

Returning to the constitutive equation (2.21), the equilibrium (2.36) can also be written :

$$\frac{\mathrm{d}p}{\mathrm{d}x} = \frac{\partial \sigma_{xy}}{\partial y}. \tag{2.38}$$

This is an equilibrium of forces, which is normal because we are in a steady state. The pressure gradient as per x is thus in equilibrium with the shear stress. The physical interpretation of this equilibrium is straightforward : the pressure gradient is the term which drives the flow, the shear represents a brake, compensating the driving force to maintain a steady state. The fact that the flow is caused by the pressure gradient is clear from the final formula (2.37), and the flow moves towards the point of lowest pressure. Recall that the gradient of a vector field directed towards the local maxima of p, here the component x, $\frac{\mathrm{d}p}{\mathrm{d}x} = -\alpha$, is positive if pressure decreases with x. In other words, if we push from the left then the fluid moves to the right...

Exercice 2.6 For the example of Poiseuille laminar flow, calculate the stress tensor and its divergence.

Solution 2.6 Starting with (2.22) and adding in the pressure and velocity distributions given above, the only non-null component of the velocity gradient is $\frac{\mathrm{d}u}{\mathrm{d}y} = -\frac{\alpha}{\mu}y$ (still working in the plane (x,y)) :

$$\underline{\underline{\sigma}} = \begin{pmatrix} -p_0 + \alpha x & -\alpha y \\ -\alpha y & -p_0 + \alpha x \end{pmatrix}. \tag{2.39}$$

We thus find that :

$$\underline{\operatorname{div}}\, \underline{\underline{\sigma}} = \begin{pmatrix} \frac{\partial \sigma_{xx}}{\partial x} + \frac{\partial \sigma_{xy}}{\partial y} \\ \frac{\partial \sigma_{yx}}{\partial x} + \frac{\partial \sigma_{yy}}{\partial y} \end{pmatrix} = \underline{0}. \tag{2.40}$$

As per the Cauchy equation, this confirms that the system is in equilibrium. This relies on several hypotheses, particularly that the flow is permanent and the velocity field runs parallel to the plates (laminar flow), which allows us to dispense completely with the left hand side of the momentum equation. ■

R Such situations are very rare, as we shall see in class when we take a closer look at the validity of the laminar hypothesis. The disappearance of the inertia term makes the equations linear, which is what makes them so easy to solve analytically.

The stress tensor allows us to calculate the stress on the lower solid surface, for example. As this is a vector, we need only consider the shear stress as given by (2.10), with $\sigma_{xy}n_y$, $n_y = -1$ the y component of the exterior normal vector. This stress is usually written τ_p :

$$\tau_p = (\sigma_{xy} n_y)_{y=-d} = \alpha d. \tag{2.41}$$

We can check this using the constitutive equation and the velocity profile :

$$\tau_p = \left(\mu \frac{\mathrm{d}u}{\mathrm{d}y}\right)_{y=-d} = \alpha d. \tag{2.42}$$

R On the upper solid surface we find a stress with the opposite sign, because in this case $n_y = +1$. The signs attributed to this stress are dictated by convention, as described above. For friction, it is the absolute value which matters.

Exercice 2.7 For the same case, sketch the velocity profile and calculate the mean U for a transverse profile connecting the two plates, otherwise known as the *flow velocity*.

Solution 2.7 We know that

$$\begin{aligned} U &= \tfrac{1}{2d}\int_{-d}^{d} u(y)\mathrm{d}y \\ &= \tfrac{\alpha}{4\mu d}\int_{-d}^{d}(d^2-y^2)\mathrm{d}y \\ &= \tfrac{\alpha d^2}{4\mu}\int_{-1}^{1}(1-\xi^2)\mathrm{d}\xi \\ &= \tfrac{\alpha d^2}{4\mu}\left[\xi-\tfrac{1}{3}\xi^3\right]_{\xi=-1}^{\xi=1} \\ &= \tfrac{\alpha d^2}{3\mu}. \end{aligned} \tag{2.43}$$

The Poiseuille velocity profile (2.37) can thus be rewritten

$$u(y) = \tfrac{3}{2}U\left(1-\left(\frac{y}{d}\right)^2\right). \tag{2.44}$$

We can see that the maximum velocity, occurring at $y=0$, is equal to $\frac{3}{2}U$ (see Figure 2.2). ■

Exercice 2.8 Calculate the Poiseuille profile using the Cauchy momentum equation in tensor form and the constitutive equation for a viscous fluid (with the help of the continuity equation). We can assume that $\underline{\underline{\tau}} \doteqdot \mu\,\underline{\underline{\mathrm{grad}}}\,\underline{u} + \mu(\underline{\underline{\mathrm{grad}}}\,\underline{u})^T$ (shear tensor). Show that

$$\tau_{xy} = \alpha y, \tag{2.45}$$

And calculate the friction stress on the solid surfaces.

Solution 2.8 Based on the hypotheses of Poiseuille laminar flow, the Cauchy equation becomes, as we have seen, $\underline{\mathrm{div}}\,\underline{\underline{\sigma}}$, which with the constitutive equation gives us :

$$\underline{0} = -\underline{\mathrm{grad}}\,p + \underline{\mathrm{div}}\,\underline{\underline{\tau}}, \tag{2.46}$$

Or, in expanded form :

$$0 = -\frac{\partial p}{\partial x} + \frac{\partial \tau_{xx}}{\partial x} + \frac{\partial \tau_{xy}}{\partial y}, \tag{2.47}$$

$$0 = -\frac{\partial p}{\partial y} + \frac{\partial \tau_{yx}}{\partial x} + \frac{\partial \tau_{yy}}{\partial y}. \tag{2.48}$$

We can then proceed as above : the continuity equation tells us that velocity depends only on y, and so too do its derivatives, so $\underline{\underline{\tau}}$ also. Moreover, as the only non-null component of the velocity gradient is $\frac{\mathrm{d}u}{\mathrm{d}y}$, we have $\tau_{yy}=0$. This leaves us with :

$$0 = -\frac{\partial p}{\partial x} + \frac{\partial \tau_{xy}}{\partial y}, \tag{2.49}$$

$$0 = -\frac{\partial p}{\partial y}. \tag{2.50}$$

We also know the second equation, which tells us that $p = p(x)$, while the first allows us to deduce that the two sides are constant, so we can integrate with reference to y :

$$\tau_{xy} = \alpha y + \tau_0, \tag{2.51}$$

where τ_0 is a constant. If $y = \pm d$, then $\tau_{xy} = \tau_0 \pm \alpha d$. But the friction stress τ_p exerted by a fluid on a solid surface is $\tau_{xy} n_y$, where $n_y = \pm 1$ is the vertical component of the normal vector for that surface. From this we can deduce that $\tau_p = \pm\tau_0 + \alpha d$, which implies that τ_0 et $\alpha = \frac{\tau_p}{d}$, so the shear stress is distributed in linear fashion between the two plates (Figure 2.2) :

$$\tau_{xy}(y) = \tau_p \frac{y}{d}. \tag{2.52}$$

Taking the above and incorporating the definition of τ_{xy}, we can obtain the parabolic velocity profile.
■

2.2.2 Two examples of analytical solutions

In the following exercises we propose two solutions to Navier-Stokes equations, using some of the notions explained in the paragraph of Chapter 1 devoted to differential equations, and highlighting the role of viscous forces. The flow here is planar as with the Poiseuille model ; the figures can thus be drawn in two dimensions.

Exercice 2.9 Work in the plane (x,y) and disregard gravity. Consider the following vector field :

$$\underline{u}(x,y,t) = U(t) \begin{pmatrix} -\cos(Kx)\sin(Ky) \\ \sin(Kx)\cos(Ky), \end{pmatrix} \tag{2.53}$$

where K is an arbitrary wavenumber and U is an unknown function. We also need to adopt a pressure field :

$$p(x,y,t) = P(t)(\cos(2Kx) + \cos(2Ky)). \tag{2.54}$$

Find $U(t)$ and $P(t)$ so that these fields comply with the Navier-Stokes equations for an incompressible flow (Taylor-Green vortices, Figure 2.3).

Solution 2.9 We begin by calculating the divergence of the velocity field :

$$\operatorname{div}\underline{u} = \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} = KU(t)\left(\sin(Kx)\sin(Ky) - \sin(Kx)\sin(Ky)\right) = 0. \tag{2.55}$$

So the continuity equation is respected. The $\underline{u}$ time partial derivative is

$$\frac{\partial \underline{u}}{\partial t} = U'(t) \begin{pmatrix} -\cos(Kx)\sin(Ky) \\ \sin(Kx)\cos(Ky), \end{pmatrix} \tag{2.56}$$

where U' is the derivative of U. Now we can calculate $\underline{\underline{\operatorname{grad}}}\,\underline{u}$:

$$\underline{\underline{\operatorname{grad}}}\,\underline{u} = \begin{pmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{pmatrix} = KU(t) \begin{pmatrix} \sin(Kx)\sin(Ky) & -\cos(Kx)\cos(Ky) \\ \cos(Kx)\cos(Ky) & -\sin(Kx)\sin(Ky) \end{pmatrix}, \tag{2.57}$$

and the inertia term can thus be written

$$(\underline{\underline{\mathrm{grad}}}\,\underline{u})\underline{u} = -\tfrac{1}{2}KU^2(t)\begin{pmatrix}\sin(2Kx)\\ \sin(2Ky)\end{pmatrix}. \tag{2.58}$$

As for the pressure gradient, it is equal to

$$\underline{\mathrm{grad}}\,p = P(t)\begin{pmatrix}\dfrac{\partial p}{\partial x}\\ \dfrac{\partial p}{\partial y}\end{pmatrix} = -2KP(t)\begin{pmatrix}\sin(2Kx)\\ \sin(2Ky)\end{pmatrix}. \tag{2.59}$$

Finally, we can calculate $\Delta\underline{u}$:

$$\Delta\underline{u} = U(t)\begin{pmatrix}\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}\\ \frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}\end{pmatrix} = -2K^2U(t)\begin{pmatrix}-\cos(Kx)\sin(Ky)\\ \sin(Kx)\cos(Ky)\end{pmatrix}. \tag{2.60}$$

We can thus observe that $\Delta\underline{u} = -2K^2\underline{u}$. Plugging all of this into the momentum equation, we get

$$\left(U'(t)+2\nu K^2U(t)\right)\begin{pmatrix}-\cos(Kx)\sin(Ky)\\ \sin(Kx)\cos(Ky)\end{pmatrix} = \left(\tfrac{2K}{\rho}P(t)+\tfrac{1}{2}KU^2(t)\right)\begin{pmatrix}\sin(2Kx)\\ \sin(2Ky)\end{pmatrix}. \tag{2.61}$$

In the x component of this vector equation, the left hand side depends on y but the right hand side does not ; as such they are constant, which is only possible if each of the scalar fields in front of the vectors is null :

$$U'(t)+2\nu K^2U(t) = 0, \tag{2.62}$$
$$\tfrac{2K}{\rho}P(t)+\tfrac{1}{2}KU^2(t) = 0. \tag{2.63}$$

The equation (2.62) is a differential equation whose unknown is the function U. Its solutions are $U(t) = U_0\exp(-2\nu K^2t)$, with the constant U_0 determined by the initial conditions. The equation (2.63) thus gives us $P(t) = \frac{1}{4}\rho U_0^2\exp(-4\nu K^2t)$. ■

R The preceding exercise demonstrates that viscosity ν has the effect of slowing down flow ; we already saw as much with Poiseuille planar flow. It is also interesting to note that pressure is dimensionally homogeneous with the product of density and velocity squared.

Exercice 2.10 Let us continue working in the plane (x,y) with no gravity. The half-plane $y \geq 0$ is filled with a viscous fluid. All of a sudden, we set in motion at a constant velocity of $U\underline{e}_x$ the surface which is coextensive with the x axis. Calculate the velocity field, assuming that the pressure field is null and working on the hypothesis that the velocity field runs parallel to the solid surface. It will be useful, at a certain point, to change the variable $\xi \doteqdot \frac{x}{\sqrt{\nu t}}$. Calculate the stress on the surface τ_p over time.

Solution 2.10 With these hypotheses, only the x component of the momentum equation is useful for our purposes. As in the Poiseuille flow case, the continuity equation shows us that u is solely dependent on y, so the inertia is null and we are left with

$$\frac{\partial u}{\partial t} = \nu\frac{\partial^2 u}{\partial y^2}. \tag{2.64}$$

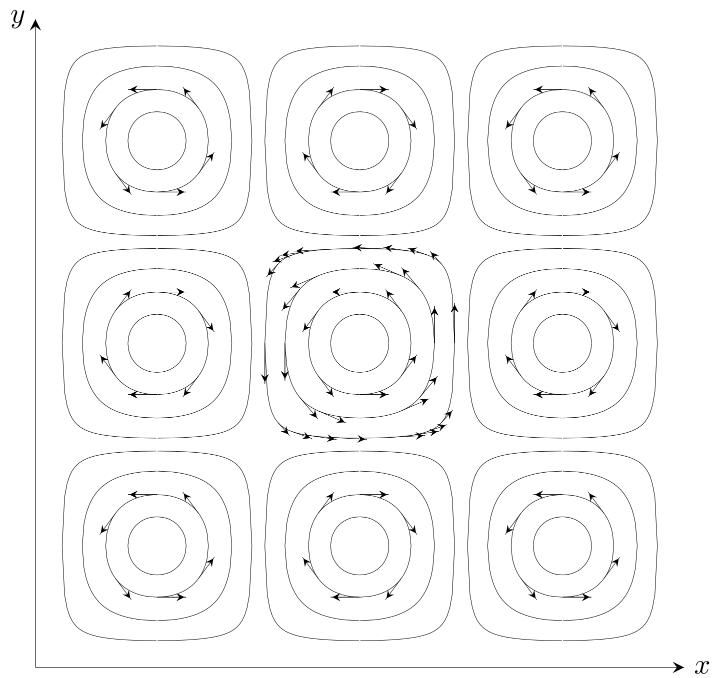


FIGURE 2.3 – Taylor-Green vortices–(Exercise 2.9).

Performing the recommended change of variable, we find that

$$2\frac{\partial^2 u}{\partial \xi^2} + \xi \frac{\partial u}{\partial \xi} = 0. \tag{2.65}$$

This is a first order linear ODE with a non-constant coefficient. It can be integrated by dividing by $\frac{\partial u}{\partial \xi}$:

$$\frac{\partial u}{\partial \xi} = C \exp\left(-\tfrac{1}{4}\xi^2\right), \tag{2.66}$$

where C is a constant. A quick online search tells us that the primitive of the function which for ξ has $\frac{2}{\sqrt{\pi}}\exp(-\xi^2)$ can be obtained using the error function « erf », which gives us

$$u(\xi) = a + b\,\mathrm{erf}\,\frac{\xi}{2}, \tag{2.67}$$

We postulate that $u(0) = U$ and $\lim_{y\to+\infty} u(y) = 0$ (away from the surface its effect is not felt), which, bearing in mind that $\mathrm{erf}(0) = 0$ et $\lim_{y\to+\infty} \mathrm{erf}\,(y) = 1$, gives us $a = -b = U$:

$$u(y,t) = U\left(1 - \mathrm{erf}\,\frac{y}{2\sqrt{\nu t}}\right). \tag{2.68}$$

The stress on the solid surface is calculated using the velocity derivative, which we already know :

$$\begin{aligned}\tau_p(t) &= \mu\left(\frac{\partial u}{\partial y}\right)_{y=0} \\ &= -\frac{\mu U}{\sqrt{\pi \nu t}} \exp\left(-\frac{y^2}{4\nu t}\right).\end{aligned} \tag{2.69}$$

■

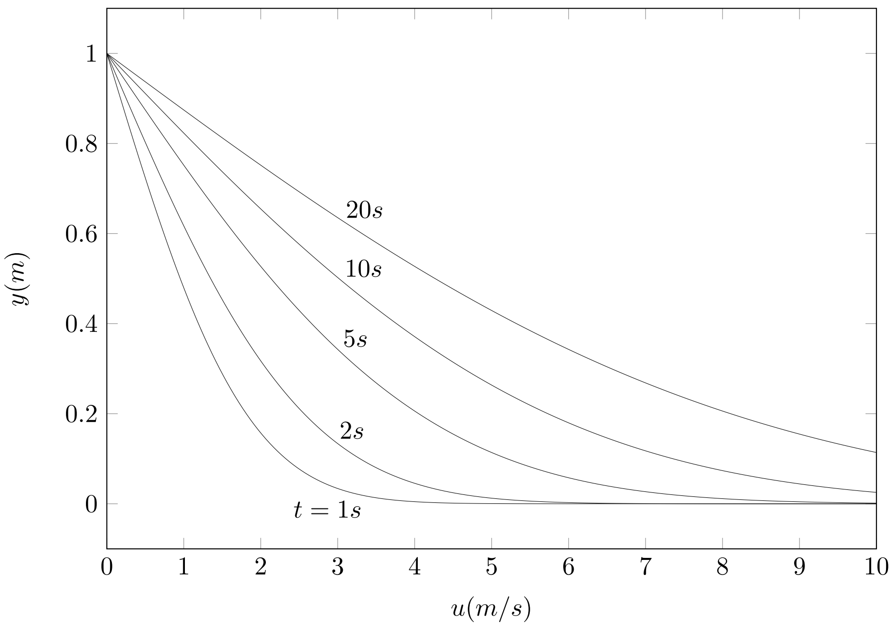


FIGURE 2.4 – Longitudinal velocity profile (2.68) with reference to y at several points in time$t$, where $U = 1\,m/s$ and $\nu = 1\,\mathrm{m}^2/s$. The $y = 0$ axis takes on the role of the lower solid surface, moving to the right and taking the fluid with it by means of viscous friction.

R The preceding exercise demonstrates that viscosity ν has a homogenizing effect : it tends to smooth out the velocity field. It is also interesting to note that viscosity is homogeneous with the relation between squares for distance and time, i.e. the product of distance and time.

2.2.3 Propagation of surface waves

To wrap up, we will look at a general problem illustrated by a classic example, in which we will have to accept a certain number of approximate estimates or simple hypotheses. To begin with, imagine that we have a flow with a velocity field $\underline{U}$, and then that a disturbance emerges in this initial flow, modifying the velocity field to $\underline{U} + \underline{u}$. We have already seen that the inertia term is non-linear, and can be broken down into four terms :

$$(\underline{\underline{\mathrm{grad}}}\,\underline{u})\underline{u} = (\underline{\underline{\mathrm{grad}}}\,\underline{U})\underline{U} + (\underline{\underline{\mathrm{grad}}}\,\underline{U})\underline{u} + (\underline{\underline{\mathrm{grad}}}\,\underline{u})\underline{U} + (\underline{\underline{\mathrm{grad}}}\,\underline{u})\underline{u}. \tag{2.70}$$

We often find that none of these terms is negligible in relation to the others, for example when studying turbulent flows. Nevertheless, in some cases the disturbance field $\underline{u}$ is so small that we can neglect the quadratic term, i.e. the last term of (2.70). We call this *linearization*, and the advantage is that it leaves us with linear equations which can be solved using simple techniques. For example if the initial flow $\underline{U}$ is constant, oriented by x and only varying in line with y, we are left with the following linear expression (u, v, w) :

$$\left(U\frac{\partial u}{\partial x} + v\frac{\partial U}{\partial y}\right)\underline{e}_x + U\frac{\partial v}{\partial x}\underline{e}_y + U\frac{\partial w}{\partial x}\underline{e}_z. \tag{2.71}$$

If the fluid is initially at rest ($\underline{U} = \underline{0}$), the inertia term simply disappears. Imagine such a case, with flow restricted exclusively to the plane (x,z) for a simple demonstration. Let us also suppose that the viscous phenomena are negligible for our purposes (we will discuss the conditions required for these approximations to be valid in class). The Navier-Stokes equations for the disturbed flow $\underline{u}$ thus become

$$\frac{\partial u}{\partial x} + \frac{\partial w}{\partial z} = 0, \tag{2.72}$$

$$\frac{\partial u}{\partial t} = -\frac{1}{\rho}\frac{\partial p}{\partial x}, \tag{2.73}$$

$$\frac{\partial w}{\partial t} = -\frac{1}{\rho}\frac{\partial p}{\partial z} - g. \tag{2.74}$$

R If we ignore the viscosity (we call this a perfect fluid model) and imagine that the fluid no longer sticks to the surfaces, the boundary condition at the surfaces is dictated by impermeability alone : $\underline{u} \cdot \underline{n} = 0$.

For example, we might find ourselves with a fluid which is initially at rest, with a free horizontal surface on which small waves begin to develop, as in Figure 2.5. In this case, vertical speed $w(x,z,t)$ is so low that we can ignore its temporal derivative before gravity, and (2.74) can be further simplified to give us

$$\frac{\partial p}{\partial z} = -\rho g, \tag{2.75}$$

We can integrate this, taking into account the absence of pressure on the free surface, to write the pressure distribution :

$$p(z) = \rho g(\eta - z), \tag{2.76}$$

Where η is the free surface value. This is what we call *hydrostatic pressure* : it grows in a linear fashion the further we descend beneath the surface. But the action of the waves modifies η , which in reality varies in space and time, and we should really write

$$p(x,z,t) = \rho g\left(\eta(x,y,t) - z\right). \tag{2.77}$$

R As we have seen, this approximation is only valid for surface waves whose form allows us to discount w, i.e. for long waves (we will look at a more detailed definition of this term in class).

Deriving (2.77), we find that

$$-\frac{1}{\rho}\frac{\partial p}{\partial x} = -g\frac{\partial \eta}{\partial x}. \tag{2.78}$$

Now let us attempt to connect the free surface value with the velocity field. It is clear that for low-amplitude waves like these (a necessary condition for linearization), the vertical velocity at the surface of the fluid is simply the temporal derivative of η :

$$\frac{\partial \eta}{\partial t} = w(x,h,t), \tag{2.79}$$

Where the vertical velocity is considered as $z = h$ and not $z = \eta(x,t)$, by linearization. We assume that the seabed is flat, with h the height of the water level at rest, and the origin of z lying on the seabed. The (2.79) equation is the *kinematic boundary condition* at the surface. In order to complete the construction of our simplified model of wave propagation, we can integrate (2.72) between the seabed and the surface, supposing that the velocity u depends only on (x,t) and making use of the fact that $w = 0$ on the bed (the impermeability condition). After linearization, this gives us :

$$w(x,h,t) = -h\frac{\partial u}{\partial x}, \tag{2.80}$$

Putting together (2.78), (2.79) and (2.80) and feeding them into the temporal derivative of (2.73), we find that

$$\frac{\partial^2 u}{\partial t^2} - c^2\frac{\partial^2 u}{\partial x^2} = 0, \tag{2.81}$$

Where $c \doteqdot \sqrt{gh}$. In order to understand the meaning of the PDE (2.81), which governs the spatio-temporal evolution of the velocity field associated with long waves of small amplitude, it can be useful to temporarily change the following variables :

$$\begin{aligned} X(x,t) &\doteqdot x - ct, \\ Y(x,t) &\doteqdot x + ct. \end{aligned} \tag{2.82}$$

The partial derivatives will be modified, as per the chain derivation rule (1.145) :

$$\begin{aligned} \frac{\partial}{\partial t} &= \frac{\partial X}{\partial t}\frac{\partial}{\partial X} + \frac{\partial Y}{\partial t}\frac{\partial}{\partial Y} = c\left(-\frac{\partial}{\partial X} + \frac{\partial}{\partial Y}\right), \\ \frac{\partial}{\partial x} &= \frac{\partial X}{\partial x}\frac{\partial}{\partial X} + \frac{\partial Y}{\partial x}\frac{\partial}{\partial Y} = \frac{\partial}{\partial X} + \frac{\partial}{\partial Y}. \end{aligned} \tag{2.83}$$

We can thus deduce the secondary derivatives :

$$\begin{aligned} \frac{\partial^2}{\partial t^2} &= c^2\left(-\frac{\partial}{\partial X} + \frac{\partial}{\partial Y}\right)^2 = c^2\left(\frac{\partial^2}{\partial X^2} - 2\frac{\partial^2}{\partial X \partial Y} + \frac{\partial^2}{\partial Y^2}\right), \\ \frac{\partial^2}{\partial x^2} &= \left(\frac{\partial}{\partial X} + \frac{\partial}{\partial Y}\right)^2 \qquad = \frac{\partial^2}{\partial X^2} + 2\frac{\partial^2}{\partial X \partial Y} + \frac{\partial^2}{\partial Y^2}. \end{aligned} \tag{2.84}$$

Using this notation system, (2.81) takes the following simplified form :

$$\frac{\partial^2 u}{\partial X \partial Y} = 0. \tag{2.85}$$

We start by integrating in relation to Y :

$$\frac{\partial u}{\partial X} = \Phi(X), \tag{2.86}$$

The integration « constant » Φ is constant in relation to Y, but *a priori* is variable in X. Now let us integrate in X, using ϕ to denote a primitive of Φ :

$$u(X,Y) = \phi(X) + \psi(Y), \tag{2.87}$$

This time, the integration « constant » ψ depends on the other variable, Y. Returning to the primitive variables x and t, this gives us

$$u(x,t) = \phi(x-ct) + \psi(x+ct). \tag{2.88}$$

This is the most general solution to (2.81). To understand what this means, imagine a person walking along the bank of a straight river (or canal) flowing at a constant velocity of c. If we let x represent the x-axis running along the canal, we can see that $x = x_0 + ct$, so $x - ct$ is constant for our observer, so if he was able to measure $\phi(x-ct)$ he would always find the same value. Now imagine another observer walking along the bank of the same canal at the same speed, but in the opposite direction. She would verify $x + ct = cst$, and would therefore always measure the same value $\psi(x+ct)$. In practice, when we resolve the equation (2.81) the functions ϕ and ψ are determined by the initial and/or boundary conditions (see the exercise below).

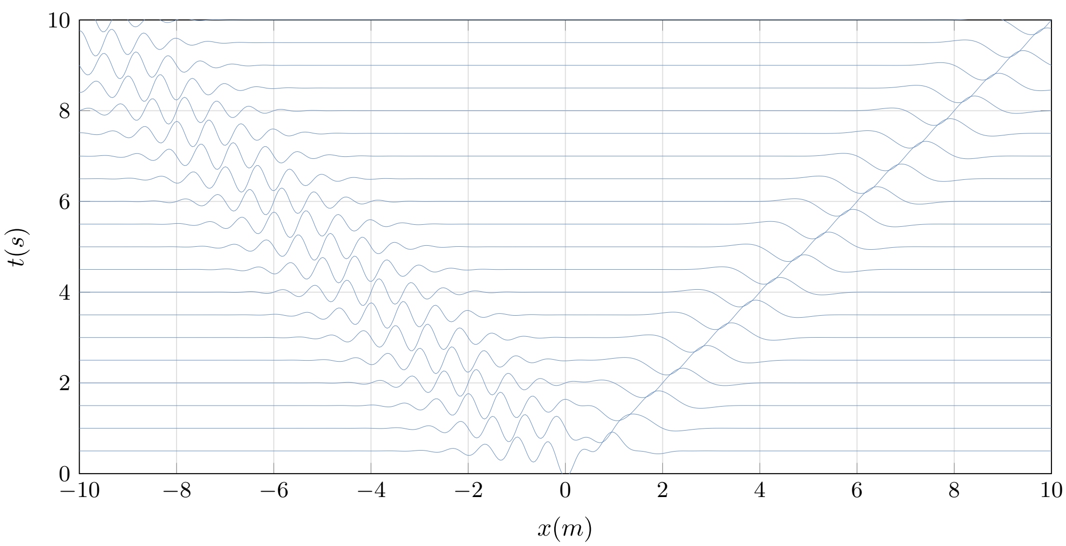


FIGURE 2.5 – Example of a solution to (2.81), consistent with (2.88), where $\phi(\xi) = 4\sin(\pi\xi)\exp(-\xi^2)$ and $\psi(\xi) = 3\sin(3\pi\xi)\exp(-\frac{1}{2}\xi^2)$, and $c = 1\,\text{m/s}$. In this space-time plane (x,t), we can clearly see the two components of the wave propagating in different directions from the point of origin, maintaining their respective forms. N.B. We have inflated the amplitude to make these low-amplitude waves visible ; as a result, the free surface appears much more highly ridged than it really is, which is to say that the wavelengths appear disproportionately small in relation to the amplitude.

From this reasoning we can deduce the following :

Théorème 2.5 *Wave propagation.* The solution (2.88) to the equation (2.81) is the superposition of two waves propagating at a speed of $c \doteqdot \sqrt{gh}$ in two opposing directions along the canal. As for the « trajectories » of these waves in space-time, they are the right sides of the respective $x \pm ct = cst$ equations ; we call them the *characteristic curves* for this problem. The equation (2.81) is a wave equation also known as *the D'Alembert equation*. An example of a solution is shown in Figure 2.5.

We can also deduce that :

Définition 2.3 *Propagation velocity of surface waves of great length and small amplitude.* For waves of this kind propagating over the surface of a deep body of water at rest h, this velocity is $c \doteqdot \sqrt{gh}$.

Exercice 2.11 Calculate the propagation velocity of a tsunami in the Pacific Ocean, whose depth is estimated to be $h = 4000m$.

Solution 2.11 We find that $c = \sqrt{9.81 \times 4000}$ m/s, i.e. 198 m/s or 713 km/h. ∎

R We must bear in mind that this does not represent the speed of the water, but rather the velocity of the wave (think of sports fans doing a Mexican wave : none of them actually gets up and runs around the stadium). We should also remember that the D'Alembert equation is not the only equation available for characterizing wave propagation (we will look at others in class). Some of these alternatives are non-linear and some are dispersive, which means that the speed of the waves depends on their length. This is not the case here, which explains why the waves propagate at a constant speed without changing form. One constant of wave equations for free surfaces is that the velocity c is a growing function of the water level, which has significant consequences that we will discuss in class.

Exercice 2.12 Show that the surface value $\eta(x,t)$ follows the same equation as $u(x,t)$, subject to the same hypotheses :

$$\frac{\partial^2 \eta}{\partial t^2} - c^2 \frac{\partial^2 \eta}{\partial x^2} = 0, \tag{2.89}$$

Solution 2.12 By deriving (2.79) with relation to t and invoking (2.80), (2.73) and (2.78), we end up with the equation we are looking for. ∎

It is interesting to consider the solution to the equation (2.81) for a monochromatic plane wave, i.e.$u(x,t) = \exp(\mathrm{i}Kx - \mathrm{i}\omega t)$ (its real part, in fact), with K and ω representing an arbitrary wavenumber and angular frequency. The temporal second derivative is thus replaced by $-\omega^2$ and the second derivative for x by $-K^2$, which gives us :

$$\omega^2 - c^2 K^2 = 0, \tag{2.90}$$

or

$$\omega = \pm cK. \tag{2.91}$$

Définition 2.4 *Dispersion relations.* The formula (2.91) is what we call a dispersion relation, which is to say that it is an equation which defines angular frequency on the basis of wavenumber, i.e. the period $T = \frac{2\pi}{\omega}$ with reference to wavelength $\ell = \frac{2\pi}{K}$. With the D'Alembert equation, this relation simply tells us that the speed of the waves is equal to $\frac{\ell}{T} = \frac{\omega}{K} = \pm c$, in accordance with the reasoning set out above.

Of course we could have obtained the same result by substituting identical ansätze into the original equations, without constructing a wave equation. We could also have performed a double Fourier transform in space and time, which would yield the same result.

Exercice 2.13 Do this for (2.72), (2.73) and (2.79), using the following ansätze :

$$u(x,t) = U\exp(\mathrm{i}Kx - \mathrm{i}\omega t), \tag{2.92}$$
$$w(x,t) = WKz\exp(\mathrm{i}Kx - \mathrm{i}\omega t),$$
$$p(x,t) = -\rho g\left(N\exp(\mathrm{i}Kx - \mathrm{i}\omega t) - z\right).$$

where U, W and N are constants. These formulae are justified by the hypotheses set out above (horizontal speed independent of z, hydrostatic pressure, continuity equation).

Solution 2.13 Now put the proposed forms into (2.72), (2.73) and (2.79) :

$$\mathrm{i}KU = -KW, \tag{2.93}$$
$$-\mathrm{i}\omega U = -\mathrm{i}KgN, \tag{2.94}$$
$$-\mathrm{i}\omega N = WKh. \tag{2.95}$$

Eliminating N, the formulae (2.94) and (2.95) give us $-\mathrm{i}\omega^2 U = WK^2gh$. With (2.93), this gives us $(Kc)^2 = \omega^2$ because $c \doteqdot \sqrt{gh}$. ■

Exercice 2.14 Using the hypotheses found in this paragraph, find a connection between u and η, then solve (2.81) with the following initial conditions : $\eta(x,0) = \eta_0(x) \doteqdot H\exp\left(-\frac{1}{2}K^2x^2\right)$ (H is a wave amplitude) and $u(x,0) = 0$.

Solution 2.14 Combining (2.73) and (2.78) gives us

$$\frac{\partial u}{\partial t} = -g\frac{\partial \eta}{\partial x}. \tag{2.96}$$

We know from the exercise 2.12 that η complies with the D'Alembert equation **??** if we maintain the same hypotheses, specifically long waves of low amplitude, which gives us low values for K and H (preceding h, as we shall see in class). We thus have $\eta(x,t) = \phi(x-ct) + \psi(x+ct)$ with two unknown functions ϕ and ψ. For $t = 0$, the initial condition for η leaves us with :

$$\phi(x) + \psi(x) = \eta_0(x). \tag{2.97}$$

On the other hand, the relation between u and η gives us $\frac{\partial u}{\partial t} = -g\phi'(x-ct) - g\psi'(x+ct)$, from which we can deduce :

$$u(x,t) = \frac{g}{c}\left(\phi(x-ct) - \psi(x+ct)\right). \tag{2.98}$$

The initial condition for u thus gives us

$$\phi(x) - \psi(x) = 0. \tag{2.99}$$

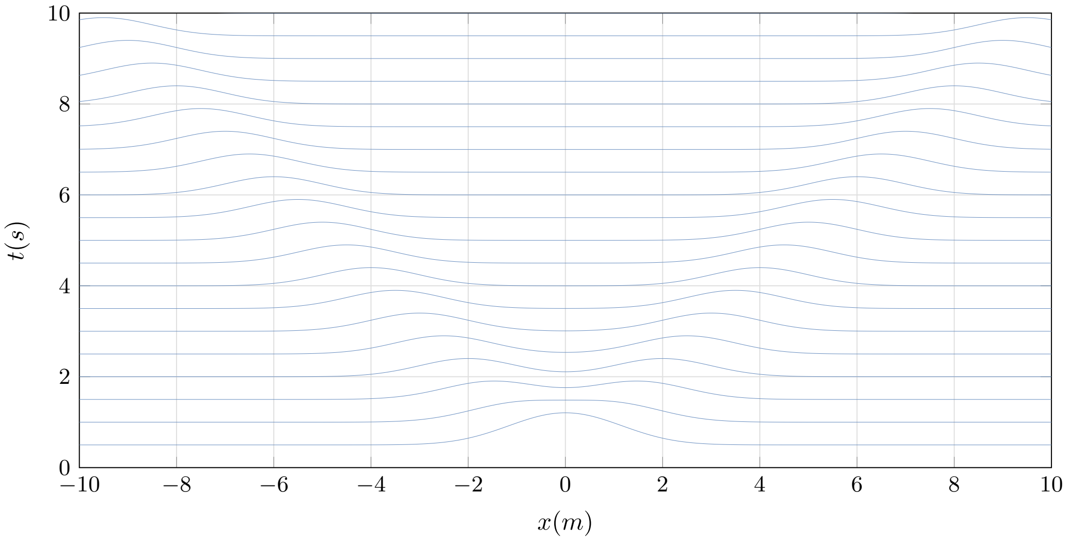


FIGURE 2.6 – Solution (2.102) for the plane (x,t), where $K = 1\,\mathrm{m}^{-1}$ and $c = 1\,\mathrm{m/s}$. N.B. We have inflated the amplitude to make these low-amplitude waves visible ; as a result, the free surface appears much more highly ridged than it really is, which is to say that the wavelengths appear disproportionately small in relation to the amplitude.

By combining (2.97) and (2.99) we get $\phi(x) = \psi(x) = \frac{1}{2}\eta_0(x)$, then

$$\eta(x,t) = \tfrac{1}{2}\left(\eta_0(x-ct) + \eta_0(x+ct)\right), \tag{2.100}$$

$$u(x,t) = \frac{g}{2c}\left(\eta_0(x-ct) - \eta_0(x+ct)\right). \tag{2.101}$$

This result is valid for all types of initial free surfaces $\eta_0(x)$ if the initial velocity is null. With the condition stipulated in the question :

$$\eta(x,t) = H\exp\left(-\tfrac{1}{2}K^2(x^2+c^2t^2)\right)\cosh(K^2xct), \tag{2.102}$$

$$u(x,t) = \frac{gH}{c}\exp\left(-\tfrac{1}{2}K^2(x^2+c^2t^2)\right)\sinh(K^2xct). \tag{2.103}$$

The free surface is represented in Figure 2.6. ■

Exercice 2.15 What happens to the wave equations (2.81) and (2.89) if the waves propagate in an initial state where velocity is not null but constant (horizontal), i.e. $U = cst \neq 0$? A clue : only the equations (2.73) and (2.79) are modified.

Solution 2.15 The equation (2.73) is modified by the addition of an inertia term, as in (2.71), while a convection terms appears in (2.79) because the vertical velocity at the surface gives us the temporal derivative of η following the mean current :

$$\frac{\partial u}{\partial t} + U\frac{\partial u}{\partial x} = -\frac{1}{\rho}\frac{\partial p}{\partial x}, \tag{2.104}$$

$$\frac{\partial \eta}{\partial t} + U\frac{\partial \eta}{\partial x} = w(x,h,t). \tag{2.105}$$

With (2.80) and (2.78), we get

$$\frac{\partial u}{\partial t} + U\frac{\partial u}{\partial x} = -g\frac{\partial \eta}{\partial x}, \tag{2.106}$$

$$\frac{\partial \eta}{\partial t} + U\frac{\partial \eta}{\partial x} = -h\frac{\partial u}{\partial x}. \tag{2.107}$$

We can introduce the operator for the *material derivative* $\frac{\mathrm{d}}{\mathrm{d}t} \doteqdot \frac{\partial}{\partial t} + U\frac{\partial}{\partial x}$:

$$\frac{\mathrm{d}u}{\mathrm{d}t} = -g\frac{\partial \eta}{\partial x}, \tag{2.108}$$

$$\frac{\mathrm{d}\eta}{\mathrm{d}t} = -h\frac{\partial u}{\partial x}. \tag{2.109}$$

By applying the operator $\frac{\mathrm{d}}{\mathrm{d}t}$ to (2.108) and using (2.109), we get the modified D'Alembert equation for an environment with a current :

$$\frac{\mathrm{d}^2 u}{\mathrm{d}t^2} - c^2\frac{\partial^2 u}{\partial x^2} = 0. \tag{2.110}$$

Adopting the same approach as we would without a current, we can factor the equation as follows :

$$\left(\frac{\partial}{\partial t} + (U+c)\frac{\partial}{\partial x}\right)\left(\frac{\partial}{\partial t} + (U-c)\frac{\partial}{\partial x}\right)u = 0, \tag{2.111}$$

With the same equation for η. This tells us that the waves always propagate at the same speed c in both directions, but *in relation to the current U*. The general solution can thus be written $\eta(x,t) = \phi(x-(U+c)t) + \psi(x-(U-c)t)$. Figure 2.7 demonstrates the consequences for the solution (2.102). ■

Finally, let us consider how the Fourier Transform can help us to solve the PDE (2.81). To do this we will perform a spatial transform, as we saw earlier on :

$$\begin{aligned}\widehat{u}(k,t) &\doteqdot \mathscr{F}_x[u(x,t)](k,t) \\ &= \tfrac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} u(x,t)\mathrm{e}^{-\mathrm{i}kx}\,\mathrm{d}x.\end{aligned} \tag{2.112}$$

The temporal dependency is left as it is. Avec la propriété (1.231), l'équation (2.81) dans l'espace de Fourier devient

$$\frac{\partial^2 \widehat{u}}{\partial t^2} + k^2c^2\widehat{u} = 0, \tag{2.113}$$

Which becomes an ODE : the derivative now only applies to time. It is formally identical to (1.22), and thus has the same solutions (written here in the form of complex exponentials) :

$$\widehat{u}(k,t) = C_+(k)\mathrm{e}^{\mathrm{i}kct} + C_-(k)\mathrm{e}^{-\mathrm{i}kct}. \tag{2.114}$$

As we solved the ODE with reference to t, the integration « constants » C_1 and C_2 are actually functions of the other variable, k. To solve the equation we need to make use of boundary conditions,

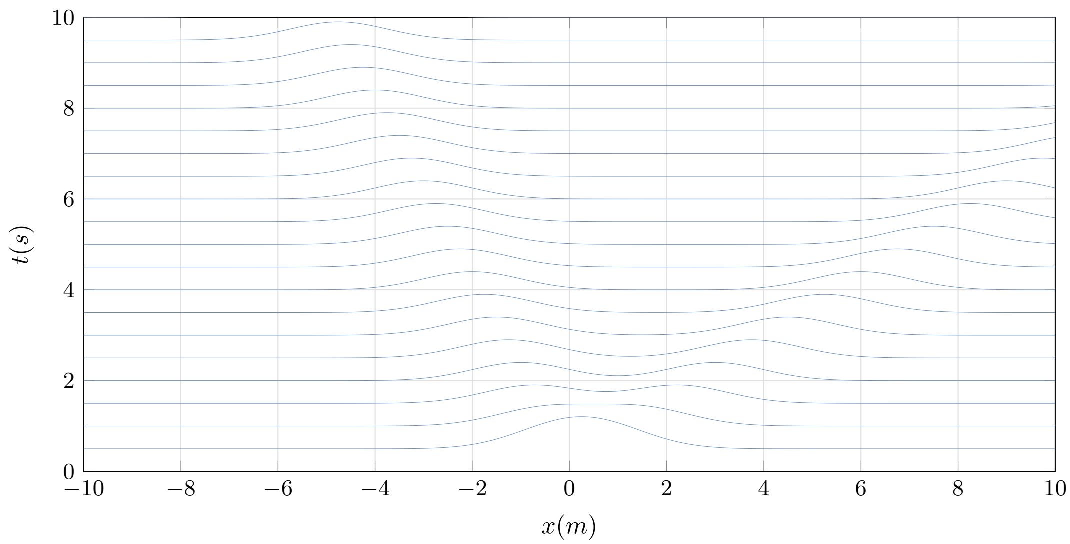


FIGURE 2.7 – Solution (2.102) for the plane (x,t), where $K = 1\,\text{m}^{-1}$ and $c = 1\,\text{m/s}$, modified by the introduction of a velocity value $U = 0,5\,\text{m/s}$ in keeping with the Exercise 2.15.

for example those from exercise 2.14. Condition $u(x,0) = 0$ in the Fourier space gives us $\widehat{u}(k,0) = 0$. Introduced into (2.114), this condition gives us $C_-(k) = -C_+(k)$, so

$$\widehat{u}(k,t) = 2\mathrm{i}C_+(k)\sin(kct). \tag{2.115}$$

Now let us write out the Fourier transform of (2.96) :

$$\frac{\partial \widehat{u}}{\partial t} = -\mathrm{i}kg\widehat{\eta}. \tag{2.116}$$

We can now bring in (2.115) and deduce that

$$\widehat{\eta}(k,t) = -\frac{2c}{g}C_+(k)\cos(kct). \tag{2.117}$$

The boundary condition $\eta(x,0) = \eta_0(x) \doteqdot H\exp\left(-\frac{1}{2}K^2x^2\right)$ gives us

$$\begin{aligned}\widehat{\eta}(k,0) &= \mathscr{F}_x[\eta_0(x)](k) \\ &= H\mathscr{F}_x[\mathrm{e}^{-K^2x^2/2}](k) \\ &= H\mathrm{e}^{-k^2/(2K^2)},\end{aligned} \tag{2.118}$$

Thanks to (1.251) and (1.229). Plugging this into (2.120) gives us

$$C_+(k) = -\frac{gH}{2c}\mathrm{e}^{-k^2/(2K^2)}, \tag{2.119}$$

hence

$$\widehat{\eta}(k,t) = H\mathrm{e}^{-k^2/(2K^2)}\cos(kct). \tag{2.120}$$

The solution can finally be obtained using an inverse Fourier transform

Exercice 2.16 Calculate it.

Solution 2.16

$$\begin{aligned}
\eta(x,t) &= H\mathscr{F}_k^{-1}\left[\mathrm{e}^{-k^2/(2K^2)}\cos(kct)\right](x,t) \\
&= \tfrac{1}{\sqrt{2\pi}}H\int_{-\infty}^{+\infty}\mathrm{e}^{-k^2/(2K^2)}\cos(kct)\mathrm{e}^{\mathrm{i}kx}\,\mathrm{d}k \\
&= \tfrac{1}{2\sqrt{2\pi}}H\int_{-\infty}^{+\infty}\mathrm{e}^{-k^2/(2K^2)}\left(\mathrm{e}^{\mathrm{i}k(x+ct)}+\mathrm{e}^{\mathrm{i}k(x-ct)}\right)\mathrm{d}k \\
&= \tfrac{1}{2}H\left(\mathscr{F}_k^{-1}\left[\mathrm{e}^{-k^2/(2K^2)}\right](x+ct)+\mathscr{F}_k^{-1}\left[\mathrm{e}^{-k^2/(2K^2)}\right](x-ct)\right) \\
&= \tfrac{1}{2}H\left(\mathrm{e}^{-K^2(x+ct)^2)/2}+\mathrm{e}^{-K^2(x-ct)^2)/2}\right),
\end{aligned} \tag{2.121}$$

Which finally corresponds to (2.102) after a bit of algebra. ■