

Méthodes de calcul et statistiques

Cours

Marie-Christine Angonin

2011-2012

Table des matières

1	Nombres complexes	5
1.1	Rappels sur le repérage de l'espace	5
1.1.1	<i>Orientation d'un plan et d'un trièdre.</i>	5
1.1.2	<i>Repères</i>	6
1.2	Définition	7
1.3	Représentation géométrique d'un nombre complexe	8
1.3.1	<i>Représentation géométrique d'un nombre complexe en coordonnées polaires : module et argument</i>	9
1.4	Exponentielle complexe, formule de Moivre	10
1.5	Propriétés de l'exponentielle complexe	11
1.6	Utilisation des nombres complexes	12
1.6.1	<i>Racines d'une équation du second degré</i>	12
1.6.2	<i>Racines n-ième de l'unité</i>	14
1.6.3	<i>Racines d'une équation du n-ième degré</i>	14
2	Les matrices et applications	15
2.1	Définitions	15
2.1.1	<i>Différentes matrices</i>	15
2.1.2	<i>Déterminant d'une matrice</i>	17
2.2	Calcul matriciel	21
2.2.1	<i>Addition de 2 matrices</i>	21
2.2.2	<i>Multiplication par un scalaire</i>	21
2.2.3	<i>Produit de matrices</i>	22
2.2.4	<i>Matrice inverse d'une matrice carrée</i>	24
2.3	Interprétation géométrique de matrices	25
2.3.1	<i>Représentation de vecteurs sous forme de matrices</i>	26
2.3.2	<i>Homothétie</i>	26
2.3.3	<i>Rotation et symétrie</i>	26
2.4	Application des matrices à la résolution de systèmes d'équations linéaires	28
2.4.1	<i>Définition</i>	28
2.4.2	<i>Résolution par échelonnement</i>	29
2.4.3	<i>Résolution d'un système régulier à l'aide de la matrice inverse</i>	31
3	Primitives et Intégrales	35
3.1	Définitions et exemples	35
3.2	Intégrale définie	36
3.3	Méthodes d'intégration	39
3.3.1	<i>Les tables et les ordinateurs.</i>	39
3.3.2	<i>La simplification du problème.</i>	39

3.3.3	<i>L'intégration par partie.</i>	40
3.3.4	<i>Le changement de variable.</i>	40
4	Les équations différentielles	43
4.1	Equations différentielles du premier ordre, linéaires, à coefficients constants	43
4.1.1	<i>Résolution de l'équation sans second membre</i>	43
4.1.2	<i>Résolution de l'équation avec second membre</i>	44
4.2	Equations différentielles du deuxième ordre, linéaires, à coefficients constants	45
4.2.1	<i>Résolution de l'équation sans second membre</i>	45
4.2.2	<i>Résolution de l'équation avec second membre</i>	46
5	Fonctions de plusieurs variables	49
5.1	Dérivée	49
5.1.1	<i>Dérivée partielles du premier ordre.</i>	49
5.1.2	<i>Dérivées partielles d'ordre supérieurs à 1.</i>	50
5.1.3	<i>Généralisation et dimensions physiques.</i>	51
5.2	Différentielle	52
5.3	Expressions remarquables	53
5.4	Calcul d'incertitudes	54
5.4.1	<i>Rappels</i>	54
5.4.2	<i>Calcul d'incertitudes</i>	55
5.5	Différentielles et dérivées de vecteurs	56
5.5.1	<i>Définitions.</i>	56
5.5.2	<i>Propriétés.</i>	57
5.6	Champs de vecteurs	57
5.6.1	<i>Circulation d'un vecteur.</i>	58
5.6.2	<i>Gradient d'une fonction.</i>	59
6	Probabilités et statistiques	63
6.1	Notion d'expérience, d'événement et conséquences	63
6.1.1	<i>Qu'est-ce qu'une expérience ?</i>	63
6.1.2	<i>Opérations sur les événements</i>	64
6.2	Probabilités	64
6.2.1	<i>Définition</i>	64
6.2.2	<i>Probabilités conditionnelles et combinées</i>	66
6.2.3	<i>Rappels d'analyse combinatoire</i>	68
6.3	Variables aléatoires discrètes	69
6.3.1	<i>Définitions</i>	69
6.3.2	<i>Paramètres caractéristiques</i>	70
6.3.3	<i>Principales lois des variables discrètes</i>	72
6.4	Variables aléatoires continues	73
6.4.1	<i>Définitions</i>	73
6.4.2	<i>Paramètres caractéristiques</i>	74
6.4.3	<i>Principales lois de probabilité</i>	75
6.5	Statistique, échantillonnage et estimation	78
6.5.1	<i>Définitions</i>	78
6.5.2	<i>Caractéristiques d'un échantillon statistique</i>	79
6.5.3	<i>Estimateur</i>	80
6.5.4	<i>Intervalle de confiance</i>	80

Préface

Le but du cours est d'introduire et d'approfondir certaines notions mathématiques utilisées dans les cours de physique, biologie ou chimie. Les connaissances de mathématiques correspondantes seront supposées assimilées pour l'examen.

Dans le présent fascicule, il y aura parfois quelques rappels, mais il est donc fortement conseillé de se reporter aux cours de mathématiques du lycée pour reprendre quelques notions de base si nécessaire.

Le programme spécifique à la Licence 1^{ère} année concerne principalement :

- **Les fonctions de plusieurs variables**
- **Les différentielles**
- **Les intégrales** et les méthodes d'intégration
- Les différentielles et **les dérivées de vecteurs**, les champs de vecteurs
- **Les matrices** et leurs applications.
- **Les équations différentielles**
- Les **lois de probabilité** et les notions **d'échantillonnage et d'estimation**

Symboles de comparaison

- Nous utiliserons parfois le symbole " $:=$ " au lieu de " $=$ ".

Par exemple, définissant l'accélération nous écrirons " $\gamma_x := d^2x/dt^2$ ". Aucune loi n'est exprimée par cette égalité, elle ne représente pas une équation ni un résultat. C'est seulement **une définition**.

Par contre pour présenter la loi fondamentale de la dynamique (la seconde loi de Newton) nous écrirons " $\vec{F} = m\vec{\gamma}$ ".

Ainsi $x^3 - 7 := 0$ doit être considéré comme la relation de définition de x , tandis que $x^3 - 7 = 0$ est une équation à résoudre dont la solution devra être notée x_0 par exemple ("Soit x_0 la solution de l'équation $x^3 - 7 = 0$ " s'écrit " $x_0^3 - 7 := 0$ ").

Lors de démonstrations, nous ferons également usage de " $:=$ " pour souligner que telle égalité est bien établie et qu'elle ne fait pas l'objet de la démonstration en cours (par opposition aux égalités notées " $=$ ").

Nous utiliserons le symbole " $:=$ " dans un souci de concision ou de clarification ; nous n'en ferons pas un usage systématique.

- Nous utiliserons le symbole " \propto " pour signifier "**équivalent à...**" ou "**varie comme...**" ou encore "**proportionnel à...**".

Considérons par exemple la fonction $F = x^n(1 + \varepsilon(x))$ où $\varepsilon(x)$ décroît vers 0 lorsque x tend vers ∞ . Dans ces conditions nous poserons $F \propto x^n$ (on dit " F équivalent à x^n "). Cela ne signifie pas que $F - x^n \rightarrow 0$. Pour $n = 2$ et $\varepsilon(x) = 1/x$ il vient $F = x^2 + x$ et par conséquent $F - x^2 = x \rightarrow \infty$ lorsque $x \rightarrow \infty$, cependant, dans ces conditions $F \propto x^2$.

Une grandeur physique P , s'exprime en fonction des variables ρ, T , etc. On écrit $P = F(\rho, T, \dots)$. Supposons que l'on ait $P = \rho^\gamma f(T, \dots)$ où f est indépendant de ρ . On écrira $P \propto \rho^\gamma$ (on dit " P varie comme ρ^γ "). Bien que le symbole " \propto " puisse prendre des sens différents, aucune ambiguïté n'est à redouter dans un contexte donné.

- Très souvent, nous ne sommes pas intéressés par la valeur précise d'une quantité physique mais par son "**ordre de grandeur**".
Ainsi le rayon terrestre, R_{\oplus} est de l'ordre de 10 000 km (plus précisément 6 400 km). Nous écrirons $R_{\oplus} \sim 10\,000$ km. Remarquer que le diamètre terrestre est du même ordre. Une goutte d'eau dont le diamètre est $d \sim 1$ mm est 10 ordres de grandeur plus petite que la Terre $((1\text{ mm})/(10\,000\text{ km}) = 10^{-10})$. Nous distinguerons les relations " \simeq " (à peu près égale à) et " \sim " (de l'ordre de) : par exemple $1254 \simeq 1,3 \cdot 10^3 \sim 10^3$.
- Les symboles " \gtrsim " et " \lesssim " seront employés pour signifier respectivement "**supérieur à un terme de l'ordre de...**" et "**inférieur à un terme de l'ordre de...**" tandis que " \gg " et " \ll " signifient "**beaucoup plus grand que...**" et "**beaucoup plus petit que...**".

Seules des grandeurs de même nature peuvent être comparées entre elles : une longueur avec une autre longueur, une durée avec une autre durée et **jamais** une durée avec une longueur. Les grandeurs de même nature ont mêmes unités.

Chapitre 1

NOMBRES COMPLEXES

Ce chapitre porte sur les nombres complexes : définitions, propriétés. Différentes utilisations de ces nombres sont présentées à la suite ou dans des chapitres ultérieurs. Auparavant, quelques rappels sur les repérages dans l'espace sont présentés.

1.1 Rappels sur le repérage de l'espace

1.1.1 Orientation d'un plan et d'un trièdre.

Etant donné un plan, P , on choisit un vecteur, \vec{u} , orthogonal à ce plan. Deux choix sont possibles. Le vecteur \vec{u} étant choisi, on dispose alors **la main droite** comme il est indiqué sur la figure 1.1. L'orientation positive des angles du plan s'en déduit.

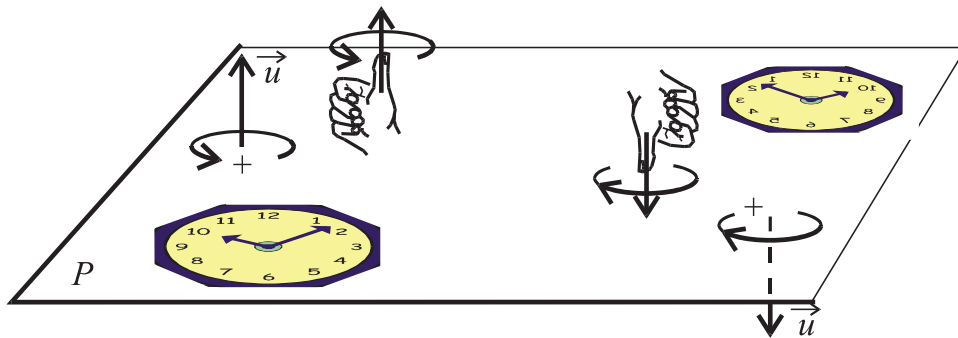


fig. 1.1 : Les deux orientations possibles d'un plan.

Considérons maintenant trois vecteurs non coplanaires, \vec{v}_1 , \vec{v}_2 et \vec{v}_3 de même origine. Le plan formé par \vec{v}_1 et \vec{v}_2 est orienté de telle sorte que le sinus de l'angle (\vec{v}_1, \vec{v}_2) soit positif. On en déduit le vecteur unitaire, \vec{u} , qui oriente le plan.

Si le vecteur \vec{v}_3 est dans le même demi espace que \vec{u} , le trièdre $(\vec{v}_1, \vec{v}_2, \vec{v}_3)$ est un **"trièdre direct"** dans le cas contraire le trièdre est **"rétrograde"**.

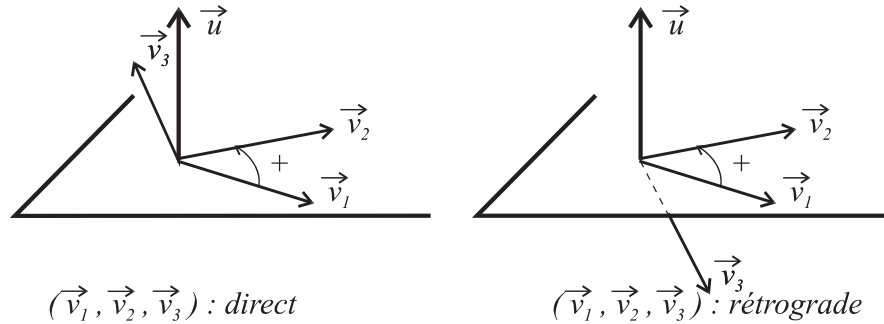


fig. 1.2 : Trièdres direct et rétrograde.

Attention! L'ordre des vecteurs est important. Si le trièdre $(\vec{v}_1, \vec{v}_2, \vec{v}_3)$ est direct il en est de même de $(\vec{v}_3, \vec{v}_1, \vec{v}_2)$ et $(\vec{v}_2, \vec{v}_3, \vec{v}_1)$ mais $(\vec{v}_1, \vec{v}_3, \vec{v}_2)$ est alors rétrograde ainsi que $(-\vec{v}_1, \vec{v}_2, \vec{v}_3)$, par exemple.

1.1.2 Repères

Pour repérer la position des points de l'espace, on se donne un trièdre orthonormé direct, $(\vec{u}_x, \vec{u}_y, \vec{u}_z)$ et un point O , choisi comme origine. Le point M est repéré par les composantes (x, y, z) du vecteur \vec{OM} : $\vec{OM} = x\vec{u}_x + y\vec{u}_y + z\vec{u}_z$. Les trois valeurs x , y et z sont les "**coordonnées cartésiennes**" du point M associées au repère donné (l'abscisse est x , l'ordonnée est y et la cote est z).

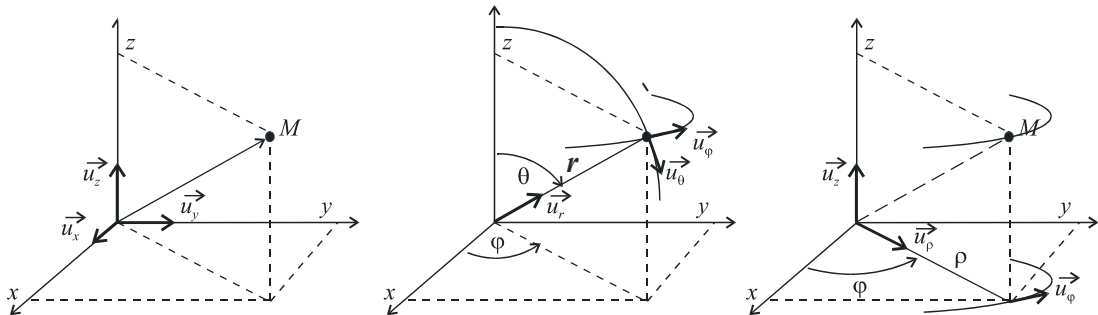


fig. 1.3 : Coordonnées cartésiennes, sphériques et cylindriques.

Les "**coordonnées sphériques**" sont r, θ et φ .

$$\begin{aligned} x &= r \sin \theta \cos \varphi, & y &= r \sin \theta \sin \varphi, & z &= r \cos \theta \text{ avec} \\ r &\geq 0, & 0 &\leq \theta \leq \pi, & 0 &\leq \varphi < 2\pi \end{aligned}$$

M étant repéré par ses coordonnées sphériques, on définit le repère "local" $(\vec{u}_r, \vec{u}_\theta, \vec{u}_\varphi)$ qui est un repère orthonormé direct (cf. fig. 1.3)

On utilise aussi les "**coordonnées cylindriques**" ρ, φ, z où le trièdre local $(\vec{u}_\rho, \vec{u}_\varphi, \vec{u}_z)$ est orthonormé direct :

$$x = \rho \cos \varphi, \quad y = \rho \sin \varphi$$

Remarquons que les vecteurs $(\vec{u}_r, \vec{u}_\theta, \vec{u}_\varphi)$ de la représentation sphérique et les vecteurs $(\vec{u}_\rho, \vec{u}_\varphi)$ de la représentation cylindrique dépendent du point M considéré* tandis que les trois vecteurs de base $(\vec{u}_x, \vec{u}_y, \vec{u}_z)$ de la représentation cartésienne ne dépendent pas du point, M .

1.2 Définition

Les nombres complexes ont été introduits en mathématiques pour résoudre certaines équations qui n'ont pas de solution dans l'ensemble \mathbb{R} des réels, comme par exemple :

$$x^2 = -1$$

Les mathématiciens ont ainsi été amenés à définir le "**nombre imaginaire**" i comme une des solutions de cette équation :

$$i^2 := -1$$

Cette définition serait restée anecdotique si elle n'avait pas abouti à une propriété tout à fait intéressante : l'utilisation des nombres complexes permet de résoudre toute équation polynomiale, quel que soit son degré. On aurait pu imaginer qu'il faille des nombres de plus en plus compliqués pour déterminer les racines d'équations polynomiales en fonction de la valeur du degré. Mais non, les nombres complexes suffisent pour ce travail !

La première preuve de ce résultat a été proposée par le célèbre mathématicien Gauss dans sa thèse de doctorat en 1799. Nous nous contenterons donc d'admettre cette propriété...

Un "**nombre complexe**" z peut s'écrire sous la forme $z = a + ib$ avec $a, b \in \mathbb{R}$, $z \in \mathbb{C}$.

a est la "**partie réelle**" du nombre complexe, on écrit $a = \operatorname{Re}(z)$.

b est la "**partie imaginaire**" du nombre complexe, on écrit $b = \operatorname{Im}(z)$.

Un nombre complexe dont la partie réelle est nulle est dit "**imaginaire pur**".

Un nombre complexe dont la partie imaginaire est nulle est un "**réel**".

On définit le "**nombre complexe conjugué**" \bar{z} , comme étant le nombre complexe ayant la même partie réelle que z et l'opposé de sa partie imaginaire

$$\bar{z} := a - ib$$

Remarque : le conjugué d'un nombre réel est égal à ce nombre. Le conjugué d'un nombre imaginaire pur est l'opposé de ce nombre.

Soient deux nombres complexes $z = a + ib$ et $z' = x + iy$

– Leur *addition* est définie par :

$$z + z' = (a + ib) + (x + iy) = (a + x) + i(b + y)$$

*Pour cette raison les repères introduits en représentation sphérique et cylindrique ont été qualifiés de "locaux".

– Leur *multiplication* est définie par :

$$z \times z' = (a + ib) \times (x + iy) = ax + iay + ibx + i^2by = (ax - by) + i(ay + bx)$$

En particulier, en utilisant cette définition :

$$i^2 = (0 + 1i)(0 + 1i) = (0 \times 0 - 1 \times 1) + i(0 \times 1 + 1 \times 0) = -1$$

1.3 Représentation géométrique d'un nombre complexe

Un nombre complexe $z = a + ib$ est donc défini par deux nombres réels a et b . Ainsi, si l'on fait abstraction du i , on peut dire qu'il y a une correspondance entre un nombre complexe et un couple de nombres :

$$a + ib \rightarrow (a, b)$$

Le nombre imaginaire i devient, par cette correspondance, le couple $(0,1)$.

Cette identification entre un nombre complexe et le couple de nombres réels correspondant permet une représentation géométrique des nombres complexes dans un plan (\mathbb{R}^2), appelé le "**plan complexe**". Le couple de nombres réels devient alors les coordonnées d'un vecteur représentant le nombre complexe dans le plan orthonormé considéré. L'axe des abscisses est appelé "**axe des réels**" et l'axe des ordonnées, "**axe des imaginaires purs**". On peut aussi associer à un nombre complexe le point de coordonnées (a, b) , appelé "**affiche**" de ce nombre complexe.

Remarque : la représentation géométrique du conjugué d'un nombre complexe est le symétrique par rapport à l'axe des abscisses du vecteur (ou du point) représentant le nombre complexe.

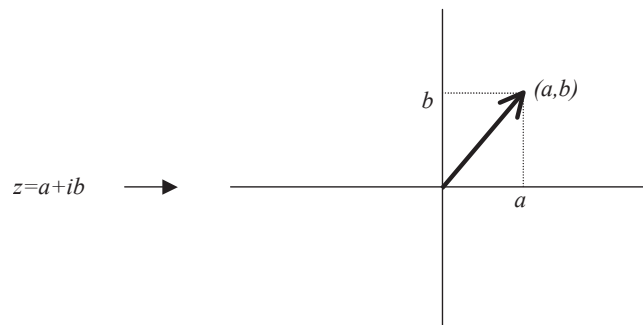


fig. 1.4 : Le plan complexe

La représentation par un vecteur a toutefois l'avantage de permettre d'avoir une analogie simple pour l'addition. En effet, la définition de l'addition de deux nombres complexes correspond exactement à l'addition des deux vecteurs dont les coordonnées coïncident avec les deux couples de nombres correspondants.

$$(a, b) + (x, y) \rightarrow (a + x, b + y)$$

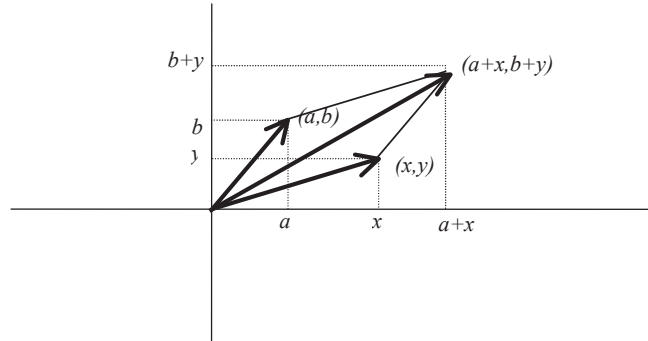


fig. 1.5 : Addition de deux nombres complexes

La multiplication est plus difficile à décrire dans ce système de coordonnées. Il faut alors se mettre dans des coordonnées polaires.

1.3.1 Représentation géométrique d'un nombre complexe en coordonnées polaires : module et argument

- A toutes coordonnées (a, b) , il est possible de faire correspondre un rayon r ($\in \mathbb{R}^+$) et un angle α ($\in [0, 2\pi[$) tels que :

$$(a, b) = (r \cos \alpha, r \sin \alpha)$$

$r = \sqrt{a^2 + b^2}$ est alors le module du vecteur associé à (a, b) et α l'angle que forme le vecteur avec l'axe des abscisses.

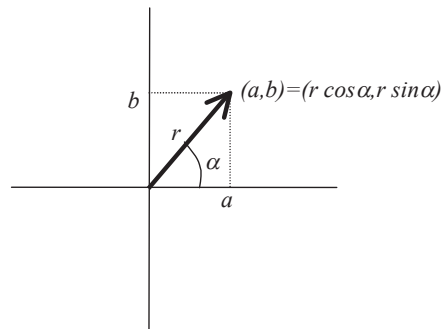


fig. 1.6 : Coordonnées polaires

- Le nombre complexe associé peut alors s'écrire de deux façons différentes

$$z = a + ib = r \cos \alpha + i r \sin \alpha$$

On définit r comme étant le "**module**" de z , il sera noté $|z| = r$, et α comme étant l'"**argument**" de z .

Le module d'un nombre complexe est unique et appartient à \mathbb{R}^+ . Par contre, la valeur de l'argument peut dépendre de l'intervalle de définition que l'on choisit. Ainsi, nous avons choisi $\alpha \in [0, 2\pi[$, mais il est tout à fait possible de choisir un autre intervalle de définition, par exemple : $\alpha \in]-\pi, +\pi]$, du moment que ce dernier permette de décrire un cercle trigonométrique complet.

Nous avons, par ailleurs, la relation

$$z \times \bar{z} = (a + ib) \times (a - ib) = a^2 + b^2 + iab - iab = a^2 + b^2 = |z|^2$$

– Revenons à la multiplication de deux nombres complexes.

On définit de la même façon les coordonnées polaires s et β du vecteur complexe (x, y)

$$(x, y) = (s \cos \beta, s \sin \beta)$$

Si l'on reprend la définition de la multiplication de deux nombres complexes, on obtient, d'après les formules trigonométriques vues en début de chapitre

$$(a, b) \times (x, y) = (rs (\cos \alpha \cos \beta - \sin \alpha \sin \beta), rs (\sin \alpha \cos \beta + \cos \alpha \sin \beta)) = (rs \cos(\alpha + \beta), rs \sin(\alpha + \beta))$$

La multiplication de deux nombres complexes donne un vecteur complexe dont le module est la multiplication des modules des deux vecteurs complexes et l'angle d'orientation est la somme des deux angles. Cette propriété nous amène à la définition des exponentielles complexes.

1.4 Exponentielle complexe, formule de Moivre

A partir des calculs précédents, nous pouvons nous arrêter sur le cas remarquable du produit de deux nombres complexes de module unité (égal à 1)

$$(\cos \alpha + i \sin \alpha) \times (\cos \beta + i \sin \beta) = \cos(\alpha + \beta) + i \sin(\alpha + \beta)$$

La fonction complexe $\alpha \rightarrow f(\alpha) = \cos \alpha + i \sin \alpha$ a donc la propriété suivante

$$f(\alpha) f(\beta) = f(\alpha + \beta)$$

Si cette fonction avait comme variable uniquement un réel, nous aurions dit que cette propriété était caractéristique d'une exponentielle. De fait, cette fonction complexe possède un certain nombre de propriétés communes avec l'exponentielle réelle. Ces propriétés nous montrent qu'il est possible de définir une "**fonction exponentielle complexe**"

$$e^{i\theta} := \cos \theta + i \sin \theta \quad \text{avec } \theta \in \mathbb{R}$$

Cette définition constitue ce que l'on appelle la formule de Moivre.

Remarque (hors programme) : Cette définition est en fait une égalité démontrable. Lorsque nous verrons les développements limités dans le chapitre sur les fonctions, nous saurons qu'il est possible d'écrire pour l'exponentielle réelle :

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots + \frac{x^n}{n!} + \dots$$

Il se trouve, de plus, que cette expression n'est pas seulement vraie pour x proche de 0, mais pour tout x de \mathbb{R} . Si, par analogie, on remplace x par ix , on obtient (en tenant compte du fait que $i^2 = -1$, $i^3 = -i$, $i^4 = 1$, etc.) :

$$e^{ix} = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \dots + i \left(x - \frac{x^3}{3!} + \frac{x^5}{5!} + \dots \right)$$

où l'on retrouve l'expression des développements limités des fonctions cosinus et sinus soit

$$e^{ix} = \cos x + i \sin x$$

1.5 Propriétés de l'exponentielle complexe

– Le module d'une exponentielle complexe est égal à 1

Démonstration : soit un nombre complexe $z = e^{i\theta}$ avec $\theta \in \mathbb{R}$.

Le module de ce nombre est

$$|z| = |e^{i\theta}| = |\cos \theta + i \sin \theta| = \sqrt{\cos^2 \theta + \sin^2 \theta} = 1$$

– Tout nombre complexe peut s'écrire sous la forme :

$$z = r \cos \alpha + ir \sin \alpha = r e^{i\alpha}$$

où r est le module de z et α son argument.

– Si nous posons $\theta = 2\pi$, nous obtenons

$$e^{i2\pi} = \cos 2\pi + i \sin 2\pi = 1$$

Remarque : Cette relation est certainement l'une des plus spectaculaires des mathématiques : un nombre transcendant puissance un autre nombre transcendant multiplié par un nombre imaginaire donne l'unité!

– A partir des deux équations (en se servant des propriétés de parité des sinus et cosinus)

$$\begin{aligned} e^{ix} &= \cos x + i \sin x \\ e^{-ix} &= \cos x - i \sin x \end{aligned}$$

Il est possible de faire la demi-somme et la demi-différence. On obtient alors

$$\begin{aligned} \cos x &= \frac{e^{ix} + e^{-ix}}{2} \\ \sin x &= \frac{e^{ix} - e^{-ix}}{2i} \end{aligned}$$

– Lorsque nous multiplions un nombre complexe par $e^{i\theta}$, nous obtenons le nombre complexe dont la représentation dans le plan complexe est le vecteur déduit du vecteur initial par la rotation d'angle θ autour de l'origine avec l'orientation positive habituelle : dans le sens des aiguilles d'une montre.

Démonstration : soit le nombre complexe $z = r e^{i\alpha}$

Si nous le multiplions par $e^{i\theta}$, nous obtenons

$$e^{i\theta} \times z = e^{i\theta} \times r e^{i\alpha} = r e^{i(\alpha+\theta)}$$

Le nombre complexe $e^{i\theta} \times z$ possède le même module que z et son argument est $\alpha + \theta$. Sa représentation dans le plan complexe est donc

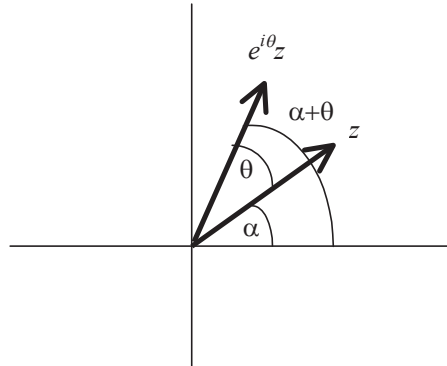


fig. 1.7 : Représentation dans le plan complexe

Nous voyons que le vecteur associé à $e^{i\theta} \times z$ est la rotation du vecteur associé à z d'un angle θ .

- Le nombre complexe $e^{i\omega t}$, où t désigne le temps, peut se représenter par un vecteur unitaire qui tourne dans le sens positif à la vitesse angulaire ω autour de l'origine.
- Nous pouvons écrire maintenant l'exponentielle d'un nombre complexe quelconque.

Soit un nombre complexe $z = a + ib$ avec $a \in \mathbb{R}$, $b \in \mathbb{R}$

$$e^z = e^{a+ib} = e^a \times e^{ib} = e^a \times (\cos b + i \sin b)$$

On montre que cette écriture est parfaitement légitime. De plus, comme $e^a \in \mathbb{R}^+$, nous pouvons déduire de l'expression précédente que e^a n'est autre que le module de ce nombre complexe.

1.6 Utilisation des nombres complexes

Comme nous l'avons vu au début de ce chapitre, les nombres complexes ont été découverts pour résoudre des équations (en particulier les équations du second degré). Nous allons voir maintenant comment les utiliser concrètement. Les nombres complexes ne représentent donc pas uniquement une belle construction mathématique, ils ont de nombreuses autres applications. Nous en aborderons d'autres ultérieurement, lorsque nous décrirons la technique de résolution des équations différentielles.

1.6.1 Racines d'une équation du second degré

- Soit une équation du second degré dont les coefficients a , b et c sont réels

$$ax^2 + bx + c = 0$$

Cette équation peut s'écrire sous la forme (en se servant des identités remarquables)

$$ax^2 + bx + c = a \left(x^2 + \frac{b}{a}x + \frac{c}{a} \right) = a \left[\left(x + \frac{b}{2a} \right)^2 - \frac{b^2}{4a^2} + \frac{c}{a} \right] = 0$$

L'équation de départ est donc équivalente à

$$\left(x + \frac{b}{2a}\right)^2 = \frac{b^2 - 4ac}{4a^2}$$

Dans le cas où tous les coefficients de l'équation sont réels et que l'on cherche des solutions où x est aussi réel, la démarche pour résoudre le problème est donc la démarche classique (qui est à connaître) :

- 1- Nous calculons le discriminant : $\Delta := b^2 - 4ac$
- 2- Nous déterminons le signe du discriminant.
- 3- Si le discriminant est positif, il y a deux solutions ou racines :

$$x_1 = \frac{-b + \sqrt{\Delta}}{2a} \text{ et } x_2 = \frac{-b - \sqrt{\Delta}}{2a}$$

Si le discriminant est nul, la solution est unique (racine double) : $x = \frac{-b}{2a}$
 Si le discriminant est négatif, il n'y a pas de racine.

- Nous pouvons généraliser cette méthode au cas où la solution de l'équation est complexe. Pour ne pas confondre avec le cas précédent, nous allons écrire l'équation avec une nouvelle inconnue

$$az^2 + bz + c = 0 \quad \text{avec } a \in \mathbb{R}, b \in \mathbb{R}, c \in \mathbb{R} \text{ et, maintenant, } z \in \mathbb{C}$$

De la même façon que pour l'équation précédente, nous pouvons dire que résoudre cette équation est équivalent à résoudre l'équation

$$\left(z + \frac{b}{2a}\right)^2 = \frac{b^2 - 4ac}{4a^2} \quad \text{avec } z \in \mathbb{C}$$

La méthode est, au départ, la même que précédemment, la nouveauté réside dans la troisième étape :

- 1- Nous calculons le discriminant : $\Delta := b^2 - 4ac$
- 2- Nous déterminons le signe du discriminant.
- 3- Si le discriminant est positif, il y a deux racines :

$$z_1 = \frac{-b + \sqrt{\Delta}}{2a} \text{ et } z_2 = \frac{-b - \sqrt{\Delta}}{2a}$$

Si le discriminant est nul, la solution est unique (racine double) : $z = \frac{-b}{2a}$
 Si le discriminant est négatif, il y a **deux racines** :

$$z_1 = \frac{-b + i\sqrt{-\Delta}}{2a} \text{ et } z_2 = \frac{-b - i\sqrt{-\Delta}}{2a}$$

Une équation du second degré à coefficients réels a ainsi **toujours** des racines complexes.

Remarque : Pour retrouver le résultat dans le cas où le discriminant est négatif, on peut utiliser le moyen mnémotechnique suivant : si Δ est négatif alors $-\Delta$ est positif. Or $\Delta = (-1) \times (-\Delta) = i^2 \times (\sqrt{-\Delta})^2$. Il suffit alors d'utiliser les formules correspondant au cas où le discriminant est positif en remplaçant $\sqrt{\Delta}$ par $i\sqrt{-\Delta}$.

1.6.2 Racines n-ième de l'unité

Chercher la racine n-ième de l'unité consiste à chercher les solutions complexes de l'équation suivante

$$z^n = 1$$

Supposons donc que $z \in \mathbb{C}$ est solution de cette équation. Il a un module $r \in \mathbb{R}^+$ et un argument α (ici, nous ne fixons pas d'intervalle de définition de l'argument volontairement car nous allons chercher toutes les solutions possibles).

$$z = r e^{i\alpha}$$

Cela implique la relation

$$z^n = r^n e^{in\alpha} = 1$$

Prenons le module des nombres de part et d'autre de la relation

$$|z^n| = |r^n e^{in\alpha}| = |r^n| |e^{in\alpha}| = |r^n| = r^n = 1$$

Si n est non nul, cela implique que $r = 1$. **Les racines n-ième de l'unité ont un module égal à 1.** Leur représentation dans le plan complexe est donc située sur le cercle trigonométrique. Nous avons donc finalement :

$$z^n = e^{in\alpha} = 1 \implies \cos(n\alpha) + i \sin(n\alpha) = 1 \implies \cos(n\alpha) = 1 \text{ et } \sin(n\alpha) = 0$$

L'argument α d'une racine n-ième de l'unité s'écrit

$$\alpha = \frac{2k\pi}{n} \quad \text{avec} \quad k \in \mathbb{Z}, n \neq 0$$

Au total, les racines n-ième de l'unité s'écrivent

$$z = \exp\left(\frac{2k\pi}{n}\right) \quad \text{avec} \quad k \in \mathbb{Z}, n \neq 0$$

Remarque : si z est une racine n-ième de l'unité alors son conjugué est aussi une racine n-ième de l'unité. Les représentations dans le plan complexe des racines n-ième de l'unité sont donc symétriques par rapport à l'axe des abscisses.

1.6.3 Racines d'une équation du n-ième degré

Les résultats obtenus pour une équation du second degré sont, en fait, très généraux. On peut en effet montrer -mais la démonstration ne sera pas faite ici- le théorème suivant (Cauchy, 1799) :

Soit l'équation du n-ième degré à coefficients complexes ($\forall i \in \mathbb{N}$, $a_i \in \mathbb{C}$; de plus, $a_n \neq 0$) :

$$\sum_{i=0}^n a_i z^i = a_n z^n + a_{n-1} z^{n-1} + \dots + a_1 z + a_0 = 0$$

Cette équation possède toujours n racines complexes : z_1, \dots, z_n . Ces racines ne sont pas toujours toutes distinctes : il est possible que certaines racines soient doubles, triples, ou plus. Par ailleurs, il est toujours possible d'écrire :

$$\sum_{i=0}^n a_i z^i = a_n z^n + a_{n-1} z^{n-1} + \dots + a_1 z + a_0 = a_n (z - z_1)(z - z_2) \dots (z - z_n)$$

Chapitre 2

LES MATRICES ET APPLICATIONS

Comme les nombres complexes, les matrices sont un outil mathématique utile pour de nombreuses applications : géométrie, systèmes d'équations à plusieurs inconnues, mécanique quantique, ... Dans ce chapitre, nous allons nous attacher à deux applications particulières : leur rôle dans les transformations géométriques que nous présenterons sous forme de quelques cas particuliers et la résolution des systèmes d'équations à deux ou trois inconnues. Les exemples de situations où ces systèmes apparaissent sont multiples : les études des cartes, des populations, des statistiques, etc. mettent en oeuvre plusieurs variables qui dépendent les unes des autres. Nous proposons ici une méthode générale de résolution de ces systèmes qui pourra s'appliquer en toutes circonstances. Mais auparavant, il convient de définir soigneusement les matrices et de décrire leurs propriétés essentielles.

2.1 Définitions

2.1.1 Différentes matrices

- Une "**matrice**" A peut être définie comme un tableau de nombres réels ou complexes (m et n sont des entiers naturels)

$$A := \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

Les a_{ij} (avec $i \in \mathbb{N}$, $j \in \mathbb{N}$) sont appelés les "**coefficients**" de la matrice. Suivant le cas, ils peuvent être réels ou complexes.

Nous pouvons remarquer que l'ordre des indices est très important : il désigne de façon non-ambiguë l'emplacement du coefficient dans le tableau de la matrice.

Lorsque la matrice est carrée ($m = n$), les coefficients a_{kk} dont l'indice est deux fois le même nombre forment la "**diagonale principale**" de la matrice.

- L'ensemble des coefficients $a_{11}, a_{12}, \dots, a_{1n}$ correspondent à une "**ligne**" de la matrice. Nous désignerons une ligne de la matrice de la façon suivante

$$(L_i) := (a_{i1} \quad a_{i2} \quad \dots \quad a_{in})$$

De même, l'ensemble des coefficients $a_{11}, a_{21}, \dots, a_{m1}$ correspondent à une "**colonne**" de la matrice. Nous désignerons une colonne de la matrice de la façon suivante

$$(C_j) := \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix}$$

La matrice A possède donc n colonnes et m lignes, on l'appelle une matrice $m \times n$.

Exemples :

$A = \begin{pmatrix} 1 & i \\ -i & 0 \end{pmatrix}$ est une matrice 2×2 .

$B = \begin{pmatrix} 2 & 2 & 2 \\ 2 & 2 & 2 \end{pmatrix}$ est une matrice 2×3 .

Remarque : Les (L_i) et les (C_j) sont elles-mêmes des cas particuliers de matrices :

- Les (L_i) font partie des matrices $1 \times n$, que l'on appelle matrices-lignes.
- Les (C_i) font partie des matrices $m \times 1$, que l'on appelle matrices-colonnes.

Les coefficients de la matrice A pour la ligne i et la colonne j seront notés $(A)_{ij}$ (dans le cas présent, $(A)_{ij} = a_{ij}$).

Remarque : Une matrice 1×1 n'a qu'un seul coefficient. Par usage, on ne parle alors plus de matrice, mais du nombre ou "**scalaire**" correspondant.

- Nous allons définir deux types de matrices associées à A : la matrice transposée et la matrice conjuguée.

On appelle "**matrice transposée**" de A , tA , la matrice obtenue en échangeant les lignes et les colonnes.

$${}^tA = \begin{pmatrix} {}^t a_{11} & {}^t a_{12} & \dots & {}^t a_{1m} \\ {}^t a_{21} & {}^t a_{22} & \dots & {}^t a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ {}^t a_{n1} & {}^t a_{n2} & \dots & {}^t a_{nm} \end{pmatrix} := \begin{pmatrix} a_{11} & a_{21} & \dots & a_{m1} \\ a_{12} & a_{22} & \dots & a_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \dots & a_{mn} \end{pmatrix}$$

soit : $({}^tA)_{ij} = {}^t a_{ij} = a_{ji} = (A)_{ji}$

Remarque : si la matrice A est une matrice $m \times n$, sa transposée tA est une matrice $n \times m$.

On appelle "**matrice conjuguée**" de A , A^* , la matrice obtenue en prenant le conjugué de chaque coefficient.

$$A^* = \begin{pmatrix} a_{11}^* & a_{12}^* & \dots & a_{1n}^* \\ a_{21}^* & a_{22}^* & \dots & a_{2n}^* \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}^* & a_{m2}^* & \dots & a_{mn}^* \end{pmatrix} := \begin{pmatrix} \bar{a}_{11} & \bar{a}_{12} & \dots & \bar{a}_{1n} \\ \bar{a}_{21} & \bar{a}_{22} & \dots & \bar{a}_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \bar{a}_{m1} & \bar{a}_{m2} & \dots & \bar{a}_{mn} \end{pmatrix}$$

soit : $(A^*)_{ij} = a_{ij}^* = \bar{a}_{ij}$

Remarque : si la matrice A est une matrice $m \times n$, sa conjuguée A^* est une matrice $m \times n$.

- Examinons quelques matrices particulières.

Matrice carrée : matrice dont le nombre n de lignes est égal au nombre de colonnes. On dit que la matrice est d'"ordre" n .

Exemple : $A = \begin{pmatrix} 2+i & -3 \\ 9 & 8 \end{pmatrix}$ est une matrice carrée d'ordre 2.

Matrice triangulaire supérieure (inférieure) : matrice carrée dont les éléments sous (sur) la diagonale principale sont nuls.

Exemple : $B = \begin{pmatrix} 2+i & -3 & 1 \\ 0 & 8 & 0 \\ 0 & 0 & 9 \end{pmatrix}$ une matrice triangulaire supérieure.

$C = \begin{pmatrix} 2+i & 0 \\ 9 & 8 \end{pmatrix}$ est une matrice triangulaire inférieure.

$D = \begin{pmatrix} 0 & -3 \\ 9 & 8 \end{pmatrix}$ n'est pas une matrice triangulaire.

Matrice diagonale : matrice carrée dont tous les coefficients qui ne sont pas sur la diagonale principale sont nuls.

Exemple : $E = \begin{pmatrix} 2+i & 0 & 0 \\ 0 & 8 & 0 \\ 0 & 0 & 9 \end{pmatrix}$ une matrice diagonale, mais $F = \begin{pmatrix} 0 & 0 & 3 \\ 0 & 8 & 0 \\ 5 & 0 & 0 \end{pmatrix}$

n'est pas une matrice diagonale.

Matrice unité ou identité d'ordre n : matrice diagonale d'ordre n dont tous les éléments diagonaux sont égaux à 1. A l'usage, cette matrice est parfois représentée par un I ou I_n comme identité, ou par un 1.

$$I = I_n = 1 = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}$$

Matrice nulle : matrice dont tous les coefficients sont égaux à 0. A l'usage, cette matrice est représentée par un simple 0.

Matrice échelonnée : matrice dont le nombre de 0 consécutifs à gauche augmente de ligne en ligne. Une matrice triangulaire supérieure est un cas particulier de matrice échelonnée.

2.1.2 Déterminant d'une matrice

Cas d'une matrice 2×2

Soit A une matrice carrée 2×2

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

Le "**déterminant**" de la matrice A est le nombre suivant (les différentes notations peuvent être utilisées)

$$\det A = |A| = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} := a_{11} \times a_{22} - a_{12} \times a_{21}$$

Exemple :

Calculons le déterminant de $A = \begin{pmatrix} 1 & i \\ -i & 0 \end{pmatrix}$

$$\det A = |A| = 1 \times 0 - i \times (-i) = -1$$

Cas d'une matrice 3×3

Soit B une matrice carrée 3×3

$$B = \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix}$$

Le "**déterminant**" de la matrice B est noté

$$\det B = |B| = \begin{vmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{vmatrix}$$

Son calcul est un peu plus compliqué que dans le cas précédent, aussi nous allons voir différentes méthodes pour l'obtenir. Chacun choisira la méthode qui lui paraît la plus pratique à l'usage.

– 1^{ère} méthode : **Développement à partir d'une ligne ou d'une colonne.**

Cette méthode est très générale, elle est valide quel que soit l'ordre n d'une matrice carrée.

Il faut tout d'abord se rapporter à une quantité que nous connaissons déjà : un déterminant de matrice 2×2 . C'est pourquoi nous avons besoin de la définition d'un mineur. On appelle "**mineur**" M_{ij} de l'élément b_{ij} de B le déterminant de la matrice issue de la matrice B à laquelle on a retiré la ligne i et la colonne j .

Exemple : Pour la matrice B ,
$$M_{23} = \begin{vmatrix} b_{11} & b_{12} \\ b_{31} & b_{32} \end{vmatrix}$$

Pour obtenir ce résultat, il faut prendre le déterminant de la matrice B à laquelle on a enlevé la deuxième ligne et la troisième colonne.

Le déterminant de la matrice B se calcule en le développant à partir d'une ligne ou d'une colonne.

Si l'on développe à partir d'une ligne i

$$\det B = |B| = b_{i1} (-1)^{i+1} M_{i1} + b_{i2} (-1)^{i+2} M_{i2} + \dots + b_{in} (-1)^{i+n} M_{in}$$

Si l'on développe à partir d'une colonne j

$$\det B = |B| = b_{1j} (-1)^{1+j} M_{1j} + b_{2j} (-1)^{2+j} M_{2j} + \dots + b_{nj} (-1)^{n+j} M_{nj}$$

Le terme $(-1)^{i+j} M_{ij}$ est appelé le "**cofacteur**" de l'élément b_{ij}

$$\text{cof } b_{ij} = (-1)^{i+j} M_{ij}$$

Dans le cas d'une matrice 3×3 , nous avons

Si l'on développe à partir d'une ligne i

$$\det B = |B| = b_{i1} (-1)^{i+1} M_{i1} + b_{i2} (-1)^{i+2} M_{i2} + b_{i3} (-1)^{i+3} M_{i3}$$

Si l'on développe à partir d'une colonne j

$$\det B = |B| = b_{1j} (-1)^{1+j} M_{1j} + b_{2j} (-1)^{2+j} M_{2j} + b_{3j} (-1)^{3+j} M_{3j}$$

Exemple : Calculons le déterminant de la matrice

$$C = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$$

Développons à partir de la première ligne

$$\begin{aligned} \det C &= |C| = 1 \times M_{11} - 2 \times M_{12} + 3 \times M_{13} \\ &= 1 \times \begin{vmatrix} 5 & 6 \\ 8 & 9 \end{vmatrix} - 2 \times \begin{vmatrix} 4 & 6 \\ 7 & 9 \end{vmatrix} + 3 \times \begin{vmatrix} 4 & 5 \\ 7 & 8 \end{vmatrix} \\ \det C = |C| &= 1 \times (5 \times 9 - 6 \times 8) - 2 \times (4 \times 9 - 6 \times 7) + 3 \times (4 \times 8 - 5 \times 7) \\ &= -3 + 12 - 9 = 0 \end{aligned}$$

Développons à partir de la première colonne

$$\begin{aligned} \det C &= |C| = 1 \times M_{11} - 4 \times M_{21} + 7 \times M_{31} \\ &= 1 \times \begin{vmatrix} 5 & 6 \\ 8 & 9 \end{vmatrix} - 4 \times \begin{vmatrix} 2 & 3 \\ 8 & 9 \end{vmatrix} + 7 \times \begin{vmatrix} 2 & 3 \\ 5 & 6 \end{vmatrix} \\ \det C = |C| &= 1 \times (5 \times 9 - 6 \times 8) - 4 \times (2 \times 9 - 3 \times 8) + 7 \times (2 \times 6 - 3 \times 5) \\ &= -3 + 24 - 21 = 0 \end{aligned}$$

- 2^{ème} méthode : **La méthode de Sarrus.**

Attention : cette méthode n'est valide que pour une matrice 3×3 .

Soit B une matrice carrée 3×3

$$B = \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix}$$

Cette méthode consiste à faire le tableau suivant où l'on reproduit les deux premières colonnes après avoir écrit une première fois les éléments de la matrice

$$\oplus \Rightarrow \begin{pmatrix} b_{11} & b_{12} & b_{13} & b_{11} & b_{12} \\ & \searrow & & \searrow & \\ b_{21} & b_{22} & b_{23} & b_{21} & b_{22} \\ & & \searrow & & \searrow \\ b_{31} & b_{32} & b_{33} & b_{31} & b_{32} \end{pmatrix}$$

Le déterminant de la matrice B est l'addition des produits des termes situés sur les diagonales principales (\searrow) à laquelle on soustrait les produits des termes situés sur les autres diagonales du tableau (\swarrow).

$$\ominus \Rightarrow \begin{pmatrix} b_{11} & b_{12} & b_{13} & b_{11} & b_{12} \\ & & \swarrow & \swarrow & \\ b_{21} & b_{22} & b_{23} & b_{21} & b_{22} \\ & \swarrow & & \swarrow & \\ b_{31} & b_{32} & b_{33} & b_{31} & b_{32} \end{pmatrix}$$

Au final, nous avons donc l'expression du déterminant de la matrice B

$$\begin{aligned}
 \det B = |B| &= \begin{vmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{vmatrix} \\
 &= \underbrace{b_{11} \times b_{22} \times b_{33} + b_{12} \times b_{23} \times b_{31} + b_{13} \times b_{21} \times b_{32}}_{\substack{\text{diagonales} \searrow \text{ du tableau } \oplus}} \\
 &\quad - \underbrace{b_{31} \times b_{22} \times b_{13} - b_{32} \times b_{23} \times b_{11} - b_{33} \times b_{21} \times b_{12}}_{\substack{\text{diagonales} \nearrow \text{ du tableau } \ominus}}
 \end{aligned}$$

Exemple : Calculons le déterminant de la matrice C par la méthode de Sarrus

$$C = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$$

Ecrivons les tableaux correspondants

$$\begin{aligned}
 \oplus \Rightarrow & \begin{pmatrix} 1 & 2 & 3 & 1 & 2 \\ & \searrow & \searrow & \searrow & \\ 4 & & 5 & 6 & 4 & 5 \\ & & \searrow & \searrow & \searrow & \\ 7 & 8 & 9 & 7 & 8 \end{pmatrix} \\
 \ominus \Rightarrow & \begin{pmatrix} 1 & 2 & 3 & 1 & 2 \\ & & \nearrow & \nearrow & \nearrow & \\ 4 & & 5 & 6 & 4 & 5 \\ & \nearrow & \nearrow & \nearrow & \nearrow & \\ 7 & 8 & 9 & 7 & 7 \end{pmatrix}
 \end{aligned}$$

L'expression du déterminant est donc

$$\begin{aligned}
 \det C = |C| &= 1 \times 5 \times 9 + 2 \times 6 \times 7 + 3 \times 4 \times 8 - 7 \times 5 \times 3 - 8 \times 6 \times 1 - 9 \times 4 \times 2 \\
 &= 45 + 84 + 96 - 105 - 48 - 72 = 225 - 225 = 0
 \end{aligned}$$

Nous obtenons le même résultat qu'en utilisant les développements à partir d'une ligne ou d'une colonne.

Propriétés des déterminants

Soit une matrice carrée A d'ordre n .

- $|{}^t A| = |A|$
- S'il y a une colonne ou une ligne constituée uniquement de 0, $|A| = 0$
- Si A a 2 (ou plus) lignes ou 2 (ou plus) colonnes identiques, $|A| = 0$
- Si l'on multiplie une ligne ou une colonne de A par un nombre quelconque k , le déterminant de la matrice obtenue vaut $k|A|$
- Si l'on ajoute à une ligne (ou colonne) une combinaison linéaire des autres lignes (ou colonnes), le déterminant de la matrice résultante est le même.

Remarque : De la dernière propriété associée à la deuxième, nous déduisons que si une ligne (ou une colonne) est égale à la combinaison linéaire d'autres lignes (ou colonnes), le déterminant de la matrice est nul.

2.2 Calcul matriciel

Nous nous replaçons, à présent, dans le cas général de matrices quelconques (pas seulement carrées). Il est possible de définir l'addition et la multiplication de matrices. Mais ces opérations n'ont pas tout à fait les mêmes propriétés que l'addition et la multiplication de nombres. Il convient donc d'apprendre soigneusement, en plus de la méthode de calcul, ce que l'on a le droit de faire et ce qu'il ne faut pas faire avec ces opérations. La pratique de quelques exercices peut être très utile pour assimiler ces règles.

2.2.1 Addition de 2 matrices

Remarque importante :

Nous ne pouvons additionner **que** des matrices de **même** taille.

Définition

Soient deux matrices A et B , toutes deux sont des matrices $m \times n$. Les coefficients de A sont notés a_{ij} et les coefficients de B sont notés b_{ij} .

On définit la matrice **addition** de A et B de la façon suivante

$$(A + B)_{ij} = a_{ij} + b_{ij}$$

Exemple :

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \text{ et } B = \begin{pmatrix} 0 & 1 \\ 2 & 3 \end{pmatrix} \quad \Longrightarrow \quad A + B = \begin{pmatrix} 1 & 3 \\ 5 & 7 \end{pmatrix}$$

Propriétés

- La matrice addition de deux matrices $m \times n$ donne une matrice $m \times n$.
- L'ordre ne compte pas dans l'addition de matrices :
 $A + B = B + A \quad \Rightarrow \quad$ L'addition est *commutative*.
- $(A + B) + C = A + (B + C) \quad \Rightarrow \quad$ L'addition est *associative*.
- $0 + A = A + 0 = A$ On dit que 0 est *élément neutre* pour l'addition.

2.2.2 Multiplication par un scalaire

Définition

Soit une matrice A , supposons qu'elle soit une matrice $m \times n$. Les coefficients de A sont notés a_{ij} .

On définit la multiplication de A par le scalaire (nombre quelconque) k de la façon suivante

$$(k \times A)_{ij} = (kA)_{ij} = ka_{ij}$$

Exemple :

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \text{ et } k = -2 + i$$

$$\Longrightarrow \quad kA = (-2 + i)A = \begin{pmatrix} -2 + i & -4 + 2i \\ -6 + 3i & -8 + 4i \end{pmatrix}$$

Propriétés

– Quand on multiplie une matrice $m \times n$ par un scalaire, on obtient une matrice $m \times n$.

– Cas particuliers

Si $k = 1$, $1A = A$

Si $k = 0$, $0A = 0$

Si $k = -1$, $(-1)A = -A$

On peut remarquer que $A + (-A) = 0$.

Cela signifie que $-A$ est l'opposé de A pour l'addition.

– $k \times (A + B) = kA + kB$

$(k + k') \times A = kA + k'A$

Ces deux dernières propriétés montrent la *distributivité* de la multiplication par un scalaire par rapport à l'addition de deux matrices et de deux scalaires.

2.2.3 Produit de matrices

Remarque importante : Le produit de matrices n'existe que dans des cas particuliers. De plus, **l'ordre de multiplication compte**. Il faut que **le nombre de colonnes de la première matrice soit égal au nombre de lignes de la deuxième matrice**.

Définition

Soient deux matrices A et B : A est une matrice $m \times n$ et B est une matrice $n \times p$. Les coefficients de A sont notés a_{ij} et les coefficients de B sont notés b_{ij} .

On définit la matrice **produit** de A et B de la façon suivante

$$(A \times B)_{ij} = (AB)_{ij} = \sum_{k=1}^n a_{ik} \times b_{kj}$$

Concrètement, pour trouver l'élément ij du produit AB , il faut prendre la ligne i de A et la colonne j de B et multiplier le premier coefficient de la ligne par le premier coefficient de la colonne, puis lui additionner le deuxième coefficient de la ligne multiplié par le deuxième coefficient de la colonne, et ainsi de suite. On voit ici qu'il est impératif que le nombre de colonnes de A soit égal au nombre de lignes de B sinon ce calcul ne pourrait se faire correctement jusqu'au bout. Un moyen mnémotechnique peut consister à présenter le produit sous forme du tableau proposé dans l'exemple : on voit qu'il faut, pour chaque coefficient du produit, associer une ligne de A avec une colonne de B .

Exemple :

	$\times \nearrow$	$\begin{pmatrix} 0 & 1 \\ 2 & 3 \end{pmatrix} = B$
$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$		$\begin{pmatrix} 4 & 7 \\ 8 & 15 \end{pmatrix} = \begin{pmatrix} 1 \times 0 + 2 \times 2 & 1 \times 1 + 2 \times 3 \\ 3 \times 0 + 4 \times 2 & 3 \times 1 + 4 \times 3 \end{pmatrix} = A \times B = AB$

Propriétés

- La matrice produit d'une matrice $m \times n$ par une matrice $n \times p$ est une matrice $m \times p$.

Exemple : Le produit d'une matrice 3×1 par une matrice 1×3 donne une matrice 3×3 .

	$\times \nearrow$	$(+1 \quad +2 \quad +3) = B$
$A = \begin{pmatrix} -2 \\ 0 \\ +2 \end{pmatrix}$		$\begin{pmatrix} -2 & -4 & -6 \\ 0 & 0 & 0 \\ +2 & +4 & +6 \end{pmatrix} = \begin{pmatrix} -2 \times 1 & -2 \times 2 & -2 \times 3 \\ 0 \times 1 & 0 \times 2 & 0 \times 3 \\ 2 \times 1 & 2 \times 2 & 2 \times 3 \end{pmatrix} = A \times B = AB$

Le produit d'une matrice 1×3 par une matrice 3×1 donne une matrice 1×1 , c'est-à-dire un scalaire.

	$\times \nearrow$	$\begin{pmatrix} -2 \\ 0 \\ +2 \end{pmatrix} = A$
$B = (+1 \quad +2 \quad +3)$		$(4) = 4 = (-2 \times 1 + 0 \times 2 + 2 \times 3) = B \times A = BA$

- Le produit d'une matrice n'est *pas commutatif*. L'exemple précédent en est un exemple flagrant : la matrice AB n'a même pas la même taille que la matrice BA ! De plus, comme le nombre de colonnes de la première matrice doit être égal au nombre de lignes de la deuxième matrice dans le produit, dans le cas général de matrices quelconques, **si AB existe alors BA n'existe pas forcément**. Il faut être extrêmement vigilant sur ce point.

Exemple : $A = \begin{pmatrix} 1 & 0 \\ -1 & 2 \end{pmatrix}$ et $B = \begin{pmatrix} 3 \\ 4 \end{pmatrix}$

Il est possible de calculer AB

	$\times \nearrow$	$\begin{pmatrix} 3 \\ 4 \end{pmatrix} = B$
$A = \begin{pmatrix} 1 & 0 \\ -1 & 2 \end{pmatrix}$		$\begin{pmatrix} 3 \\ 5 \end{pmatrix} = \begin{pmatrix} 1 \times 3 + 0 \times 4 \\ -1 \times 3 + 2 \times 4 \end{pmatrix} = AB$

Par contre, il est impossible de calculer BA !

- $(A \times B) \times C = A \times (B \times C) \quad \Rightarrow \quad$ La multiplication est *associative*.
- $A \times (B + C) = A \times B + A \times C$
- $(A + B) \times C = A \times C + B \times C \quad \Rightarrow \quad$ La multiplication est *distributive* par rapport à l'addition.
- $k \times (A \times B) = kA \times B = A \times kB \quad \Rightarrow \quad$ La multiplication de deux matrices est *associative* par rapport à la multiplication par un scalaire.
- Si I est la matrice identité : $I \times A = A \times I = A \quad \Rightarrow \quad$ La matrice identité est *élément neutre* pour la multiplication.

2.2.4 Matrice inverse d'une matrice carrée

Définition

Soit A une matrice carrée d'ordre n .

On dit que A est "**inversible**" s'il est possible de trouver une matrice A^{-1} telle que : $AA^{-1} = A^{-1}A = I$, I étant la matrice identité.

La matrice A^{-1} est alors appelée "**matrice inverse**" de A .

On démontre -mais nous ne le ferons pas ici- qu'une matrice A est inversible, si et seulement si son déterminant est non nul : $|A| \neq 0$.

Remarque : La matrice inverse de AB est $(AB)^{-1} = B^{-1}A^{-1}$.

Démonstration : $ABB^{-1}A^{-1} = A(BB^{-1})A^{-1} = AA^{-1} = I$

De même, $B^{-1}A^{-1}AB = B^{-1}(A^{-1}A)B = B^{-1}B = I$

Calcul de la matrice inverse

- Cas général (définition d'une comatrice)

Si la matrice carrée A d'ordre n est inversible (c'est-à-dire, $|A| \neq 0$) alors sa matrice inverse A^{-1} s'écrit

$$A^{-1} = \frac{{}^t \text{com}A}{|A|}$$

${}^t \text{com}A$ est la transposée de la comatrice de A .

La "**comatrice**" de A , noté $\text{com}A$ ou A_C , est la matrice dont les coefficients sont les cofacteurs des coefficients correspondants de A :

$$\text{com}A = A_C = \begin{pmatrix} \text{cofa}_{11} & \dots & \text{cofa}_{1n} \\ \vdots & \ddots & \vdots \\ \text{cofa}_{n1} & \dots & \text{cofa}_{nn} \end{pmatrix}$$

- Cas d'une matrice 2×2

Nous allons effectuer le calcul de la matrice inverse d'une matrice 2×2 quelconque, puis nous vérifierons que la matrice inverse obtenue correspond bien à sa définition.

Soit la matrice $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$

Calculons les cofacteurs (il faut, dans chaque cas, enlever les coefficients de la ligne et de la colonne concernées) :

$$\begin{aligned} \text{cofa}_{11} &= (-1)^{1+1}a_{22} = a_{22} \\ \text{cofa}_{12} &= (-1)^{1+2}a_{21} = -a_{21} \\ \text{cofa}_{21} &= (-1)^{2+1}a_{12} = -a_{12} \\ \text{cofa}_{22} &= (-1)^{2+2}a_{11} = a_{11} \end{aligned}$$

La comatrice s'écrit donc

$$\text{com}A = A_C = \begin{pmatrix} a_{22} & -a_{21} \\ -a_{12} & a_{11} \end{pmatrix}$$

Sa transposée (symétrique par rapport à la diagonale principale) donne :

$${}^t_{com}A = {}^tA_C = \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}$$

Le déterminant de A s'écrit, par ailleurs

$$|A| = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11} \times a_{22} - a_{12} \times a_{21}$$

Au final, la matrice inverse de A s'écrit

$$A^{-1} = \frac{{}^t_{com}A}{|A|} = \frac{\begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}}{a_{11} \times a_{22} - a_{12} \times a_{21}}$$

Vérification :

Nous voulons vérifier que la matrice inverse satisfait bien à la relation

$$\begin{aligned} AA^{-1} &= A^{-1}A = I \\ AA^{-1} &= \frac{\begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}}{a_{11} \times a_{22} - a_{12} \times a_{21}} \times \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \\ &= \frac{\begin{pmatrix} a_{11} \times a_{22} - a_{12} \times a_{21} & a_{12} \times a_{22} - a_{12} \times a_{22} \\ a_{11} \times a_{21} - a_{11} \times a_{21} & a_{11} \times a_{22} - a_{12} \times a_{21} \end{pmatrix}}{a_{11} \times a_{22} - a_{12} \times a_{21}} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I \end{aligned}$$

De même

$$\begin{aligned} A^{-1}A &= \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \times \frac{\begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}}{a_{11} \times a_{22} - a_{12} \times a_{21}} \\ &= \frac{\begin{pmatrix} a_{11} \times a_{22} - a_{12} \times a_{21} & a_{12} \times a_{11} - a_{12} \times a_{11} \\ a_{22} \times a_{21} - a_{22} \times a_{21} & a_{11} \times a_{22} - a_{12} \times a_{21} \end{pmatrix}}{a_{11} \times a_{22} - a_{12} \times a_{21}} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I \end{aligned}$$

2.3 Interprétation géométrique de matrices

Les matrices sont un outil mathématique qui peut servir pour décrire des transformations géométriques. Il faut pour cela définir un repère dans un plan (représentation en 2 dimensions, que nous appellerons 2D) ou dans l'espace (représentation en 3 dimensions ou 3D). La matrice va y constituer ce que l'on appelle une forme bilinéaire, c'est-à-dire qu'elle va représenter une application linéaire des coordonnées d'un vecteur, soit encore un système d'équations dont les variables sont les coordonnées d'un vecteur. Pour être plus rigoureux, il faudrait développer un formalisme assez lourd dont nous n'avons pas l'utilité. Nous nous contenterons donc de quelques exemples très précis qui ont leur utilité pour décrire des mouvements dans un plan (par exemple : la tectonique des plaques) ou dans l'espace (par exemple : mouvement des bras et des jambes).

2.3.1 Représentation de vecteurs sous forme de matrices

Soit \vec{V} un vecteur

On associe à ce vecteur une matrice colonne dont les coefficients sont les coordonnées cartésiennes du vecteur dans le repère orthonormé choisi.

$$\begin{aligned} \text{Cas 2D : } \vec{V} &= \begin{pmatrix} x \\ y \end{pmatrix} \\ \text{Cas 3D : } \vec{V} &= \begin{pmatrix} x \\ y \\ z \end{pmatrix} \end{aligned}$$

Attention : Du fait de l'habitude de noter les coordonnées d'un vecteur ou d'un point par un couple (x,y) ou un triplet (x,y,z) -ce que l'on trouve parfois dans les livres-, la tendance naturelle est souvent d'associer un vecteur à une matrice ligne. Cela représente une erreur grave. Dans le cadre de ce polycopié, les matrices lignes n'ont aucune contrepartie géométrique et ne doivent pas apparaître dans les calculs.

Il ne faut donc pas confondre le couple (x, y) et la matrice ligne $(x \ y)$, ni le triplet (x,y,z) avec la matrice $(x \ y \ z)$

2.3.2 Homothétie

L'homothétie est une transformation géométrique qui affecte de la même façon chaque coordonnée d'un vecteur.

On définit donc l'homothétie de rapport k comme étant l'application qui, à un vecteur \vec{V} associe le vecteur $k\vec{V}$.

Sous la forme de matrice, cela revient à une multiplication de la matrice colonne représentant le vecteur par le scalaire k .

$$\begin{aligned} \text{Cas 2D : } \vec{V} &= \begin{pmatrix} x \\ y \end{pmatrix} & \xrightarrow{\text{Homothétie de rapport } k} & k\vec{V} &= k \begin{pmatrix} x \\ y \end{pmatrix} &= \begin{pmatrix} k \times x \\ k \times y \end{pmatrix} \\ \text{Cas 3D : } \vec{V} &= \begin{pmatrix} x \\ y \\ z \end{pmatrix} & \xrightarrow{\text{Homothétie de rapport } k} & k\vec{V} &= k \begin{pmatrix} x \\ y \\ z \end{pmatrix} &= \begin{pmatrix} k \times x \\ k \times y \\ k \times z \end{pmatrix} \end{aligned}$$

Une homothétie est l'équivalent d'un effet de "zoom" grossissant si le rapport $k > 1$.

Une homothétie est l'équivalent d'un effet de "zoom" inversé si le rapport $k < 1$.

La matrice représentative d'une homothétie de rapport k est donc k fois la matrice identité, c'est-à-dire une matrice diagonale dont tous les coefficients non-nuls sont égaux à k . En particulier, une homothétie de rapport $k = 1$ est l'identité.

2.3.3 Rotation et symétrie

Rotation

Soit \vec{V} un vecteur

Une rotation d'un angle θ de ce vecteur est représentée par la multiplication de ce vecteur par une matrice.

- La matrice représentant la rotation d'un angle θ dans le cas 2D s'écrit

$$\mathcal{R}_\theta = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$$

Exemple : Supposons que $\vec{V} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \vec{i}$. Il s'agit en fait du vecteur unitaire suivant l'axe des x .

La représentation matricielle de la rotation du vecteur \vec{V} d'un angle θ va consister à multiplier le vecteur par la matrice \mathcal{R}_θ . Effectuons cette multiplication

$$\vec{V}_\theta = \mathcal{R}_\theta \times \vec{V} = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

$\times \nearrow$	$\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \vec{V}$
$\mathcal{R}_\theta = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$	$\begin{pmatrix} \cos \theta \\ \sin \theta \end{pmatrix} = \vec{V}_\theta$

Nous obtenons $\vec{V}_\theta = \begin{pmatrix} \cos \theta \\ \sin \theta \end{pmatrix}$ dont les coordonnées sont bien celles d'un vecteur unitaire ayant un angle θ par rapport à l'axe des x et donc le résultat d'une rotation d'un angle θ du vecteur \vec{V} .

– Cas 3D

La matrice rotation d'un angle θ autour de l'axe des x s'écrit :

$$\mathcal{R}_\theta^x = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{pmatrix}$$

La matrice rotation d'un angle θ autour de l'axe des y s'écrit :

$$\mathcal{R}_\theta^y = \begin{pmatrix} \cos \theta & 0 & -\sin \theta \\ 0 & 1 & 0 \\ \sin \theta & 0 & \cos \theta \end{pmatrix}$$

La matrice rotation d'un angle θ autour de l'axe des z s'écrit :

$$\mathcal{R}_\theta^z = \begin{pmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Symétrie

Soit \vec{V} un vecteur 2D

La transformation donnant le symétrique de ce vecteur par rapport à un axe est représentée par la multiplication de ce vecteur par une matrice.

La matrice représentant une symétrie par rapport à l'axe des x est :

$$\mathcal{S}_x = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

La matrice représentant une symétrie par rapport à l'axe des y est :

$$\mathcal{S}_y = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$$

Combinaison de transformations

Si l'on veut appliquer sur un vecteur plusieurs transformations successives, il est possible d'obtenir le résultat en **multipliant dans le bon ordre** les matrices correspondant à chacune des transformations.

Exemple : La matrice d'une rotation d'un angle $\frac{\pi}{2}$ s'écrit

$$\mathcal{R}_{\frac{\pi}{2}} = \begin{pmatrix} \cos \frac{\pi}{2} & -\sin \frac{\pi}{2} \\ \sin \frac{\pi}{2} & \cos \frac{\pi}{2} \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

Si nous effectuons deux fois la rotation d'un angle $\frac{\pi}{2}$, nous devons en principe retrouver une rotation d'angle π . Dans la représentation matricielle, cela signifie que la multiplication de deux matrices $\mathcal{R}_{\frac{\pi}{2}} \times \mathcal{R}_{\frac{\pi}{2}}$ devrait donner une matrice \mathcal{R}_{π} . Vérifions-le

	$\times \nearrow$	$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \mathcal{R}_{\frac{\pi}{2}}$
$\mathcal{R}_{\frac{\pi}{2}} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$		$\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = \mathcal{R}_{\frac{\pi}{2}} \times \mathcal{R}_{\frac{\pi}{2}}$

Nous obtenons donc bien

$$\mathcal{R}_{\frac{\pi}{2}} \times \mathcal{R}_{\frac{\pi}{2}} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} \cos \pi & -\sin \pi \\ \sin \pi & \cos \pi \end{pmatrix} = \mathcal{R}_{\pi}$$

2.4 Application des matrices à la résolution de systèmes d'équations linéaires

2.4.1 Définition

Un système d'équations linéaires ou "**système linéaire**" (S) est une famille de n équations linéaires faisant intervenir p inconnues. Ces inconnues x_1, x_2, \dots, x_p sont habituellement regroupées dans le membre de gauche de chacune des équations. Le membre de droite est alors composé de constantes. Ces constantes et les coefficients de ces équations sont en général réels, mais les méthodes décrites ici sont valables aussi pour les complexes. Nous allons noter a_{ij} le coefficient de la variable x_j dans l'équation (i) et b_i la constante dans le membre de droite de l'équation (i).

$$(S) \begin{cases} a_{11} x_1 + a_{12} x_2 + \dots + a_{1p} x_p = b_1 & (1) \\ a_{21} x_1 + a_{22} x_2 + \dots + a_{2p} x_p = b_2 & (2) \\ \vdots & \\ a_{n1} x_1 + a_{n2} x_2 + \dots + a_{np} x_p = b_n & (n) \end{cases}$$

Il est utile d'associer à ce système une **représentation matricielle**. Nous pouvons ainsi regrouper les inconnues en une matrice colonne X et les éléments du second membre en une matrice colonne B . Les coefficients des équations forment de la même manière une matrice $n \times p$ A , appelée la matrice du système.

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{pmatrix} \quad B = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} \quad A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1p} \\ a_{21} & a_{22} & \dots & a_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{np} \end{pmatrix}$$

Le système (S) s'écrit alors sous la forme : $\boxed{A \times X = B}$.
Lorsque la matrice B est nulle, le système est dit "**homogène**".

Notre but est de chercher les solutions d'un tel système. Trois cas peuvent se présenter :

- 1- Il n'y a aucune solution.
- 2- Il y a une valeur unique pour chaque inconnue. La solution est unique.
- 3- Il y a une infinité de solutions. Ce qui ne veut pas dire que n'importe quelle solution soit valable. Mais nous verrons cela un peu plus loin.

Propriétés

Un système homogène a toujours au moins une solution : $X = 0$

Si $n < p$ (il y a moins d'équations que d'inconnues), il ne peut pas y avoir de solution unique : soit le système a une infinité de solution, soit il n'en a pas.

2.4.2 Résolution par échelonnement

On montre que remplacer une équation dans un système par une combinaison linéaire de cette équation avec les autres équations du système ne change pas les solutions. Cette méthode consiste à effectuer ce genre de remplacement pour simplifier la résolution du système. Il est alors important de veiller à garder le même nombre d'équations lors de ce travail. Le but est d'obtenir un système d'équations échelonné, c'est-à-dire dont la matrice représentative est échelonnée, plus simple à résoudre. Il est possible de travailler avec le système d'équations directement ou bien sur la matrice représentative (on dit alors que c'est la méthode du "pivot de Gauss") suivant ses préférences. C'est la méthode la plus simple et la plus universelle. Une illustration par un exemple permet d'expliquer plus précisément la méthode.

Exemple : Nous cherchons à résoudre le système suivant :

$$\begin{cases} x + 2y + 3z = 10 & (1) \\ -x + y + z = 0 & (2) \\ x + y + z = 6 & (3) \end{cases}$$

La représentation matricielle de ce système est :

$$\underbrace{\begin{pmatrix} 1 & 2 & 3 \\ -1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}}_A \underbrace{\begin{pmatrix} x \\ y \\ z \end{pmatrix}}_X = \underbrace{\begin{pmatrix} 10 \\ 0 \\ 6 \end{pmatrix}}_B$$

- *Résolution échelonnée directe du système* :

Il faut donc combiner les différentes équations pour obtenir un système "triangulaire", c'est-à-dire une équation où apparaît une inconnue, une équation où apparaissent deux inconnues et la troisième équation avec trois inconnues. Nous allons remplacer les équations du système par des combinaisons linéaires de ces équations avec les autres, mais il faut veiller à garder trois équations jusqu'à la fin.

Ainsi, nous pouvons remplacer l'équation (1) par la combinaison permettant d'isoler la variable z : $2 \times (1) - (2) - 3 \times (3)$ ce qui donne le système suivant.

$$\begin{cases} 2z = 2 & 2 \times (1) - (2) - 3 \times (3) = (1)' \\ -x + y + z = 0 & (2) \\ x + y + z = 6 & (3) \end{cases}$$

De même, nous pouvons remplacer l'équation (2) par la combinaison permettant d'isoler les variables y et z : (2) + (3) ce qui donne le système suivant.

$$\begin{cases} 2z = 2 & (1)' \\ 2y + 2z = 6 & (2) + (3) = (2)' \\ x + y + z = 6 & (3) \end{cases}$$

Ce système d'équation est un système échelonné. La solution découle alors directement.

L'équation (1)' nous indique que : $z = 1$

En remplaçant z par sa valeur dans (2)', on obtient une équation en y :

$$2y + 2 = 6 \implies y = 2$$

En remplaçant y et z par leur valeur dans (3), nous obtenons l'équation en x :

$$x + 2 + 1 = 6 \implies x = 3$$

- *Résolution échelonnée par pivot de Gauss :*

Cette méthode va suivre le même principe que la précédente. La différence réside sur le fait que nous allons manipuler (faire des combinaisons linéaires) de lignes de matrice au lieu de manipuler des équations. La notation est alors allégée. Il faut, en premier lieu, créer une matrice composée de la juxtaposition des matrices A et B , soit si l'on suit l'exemple

$$\left(\begin{array}{ccc|c} \overbrace{1 \quad 2 \quad 3}^A & \vdots & \overbrace{10}^B & \\ \hline -1 & 1 & 1 & 0 \\ \hline 1 & 1 & 1 & 6 \\ \hline & & & \vdots \end{array} \right) \begin{array}{l} (L_1) \\ (L_2) \\ (L_3) \end{array}$$

Nous allons remplacer progressivement les lignes de cette matrice par des combinaisons linéaires des autres lignes avec celle-ci afin de transformer la partie venant de A en matrice triangulaire. Toute combinaison linéaire est a priori acceptable, mais il existe une méthode qui fonctionne de façon efficace

1- Laisser la première ligne (L_1) en l'état.

2- Combiner (L_1) avec chacune des autres lignes, une à une, afin que ces dernières soient remplacées par des lignes dont, au moins, le premier coefficient est nul. Dans notre exemple, cela donne

$$\left(\begin{array}{cccc} 1 & 2 & 3 & 10 \\ 0 & 3 & 4 & 10 \\ 0 & -1 & -2 & -4 \end{array} \right) \begin{array}{l} (L_1) \\ (L_2) + (L_1) = (L_2)' \\ (L_3) - (L_1) = (L_3)' \end{array}$$

3- (L_2)' est maintenant fixée à son tour

4- Combiner (L_2)' avec chacune des autres lignes situées en dessous, une à une, afin que ces dernières soient remplacées par des lignes dont, au moins, les deux premiers coefficients sont nuls. Dans notre exemple, cela donne

$$\left(\begin{array}{cccc} 1 & 2 & 3 & 10 \\ 0 & 3 & 4 & 10 \\ 0 & 0 & -2 & -2 \end{array} \right) \begin{array}{l} (L_1) \\ (L_2)' \\ 3 \times (L_3)' + (L_2)' = (L_3)'' \end{array}$$

Et ainsi de suite jusqu'à ce qu'il ne reste au plus que deux coefficients non-nuls sur la dernière ligne.

Après ces transformations, nous pouvons réécrire le système d'équations correspondant à la matrice obtenue

$$\begin{cases} x + 2y + 3z = 10 & (1) \\ 3y + 4z = 10 & (2)' \\ -2z = -2 & (3)'' \end{cases}$$

Nous savons que ce système d'équations possède les mêmes solutions que le système que nous avons au départ. On en déduit

$$\begin{aligned} (3)'' &\iff z = 1 \\ (2)' &\implies 3y + 4 = 10 \iff y = 2 \\ (1) &\implies x + 4 + 3 = 10 \iff x = 3 \end{aligned}$$

Nous retrouvons bien les solutions précédemment obtenues par la résolution directe du système.

2.4.3 Résolution d'un système régulier à l'aide de la matrice inverse

Cette méthode n'est valide que pour les matrices carrées. Nous pouvons toutefois remarquer que toute matrice représentative d'un système d'équations peut être rendue artificiellement carrée en rajoutant des lignes et/ou des colonnes de 0 (c'est-à-dire en rajoutant des équations dont les coefficients sont nuls et/ou des variables "fantômes" dont les coefficients sont toujours nuls).

On dispose donc d'un système dont la représentation matricielle est : $A \times X = B$

Deux possibilités se présentent alors :

1/ Le déterminant de A est non-nul. A est inversible. La solution du système est unique et se calcule à partir des matrices A^{-1} et B . En effet, nous avons

$$\begin{aligned} A X &= B \\ \underbrace{A^{-1}A}_I X &= A^{-1}B \\ X &= A^{-1}B \end{aligned}$$

2/ Le déterminant de A est nul. A n'est pas inversible. Le système n'a pas de solution, ou, au contraire, en a une infinité. Une résolution par échelonnement permet de distinguer les deux cas.

Exemple 1 : Cherchons à résoudre le système suivant :

$$\begin{cases} x + 2y = 2 & (1) \\ 10x + 20y = 20 & (2) \end{cases}$$

La représentation matricielle du système est :

$$\underbrace{\begin{pmatrix} 1 & 2 \\ 10 & 20 \end{pmatrix}}_A \underbrace{\begin{pmatrix} x \\ y \end{pmatrix}}_X = \underbrace{\begin{pmatrix} 2 \\ 20 \end{pmatrix}}_B$$

Calculons le déterminant de A :

$$\det A = \begin{vmatrix} 1 & 2 \\ 10 & 20 \end{vmatrix} = 20 - 20 = 0$$

La matrice A n'est pas inversible. Il faut donc déterminer si le système a des solutions ou non.

Pour déterminer cela, il faut utiliser la même méthode que pour la résolution par échelonnement : effectuer des combinaisons linéaires des équations du système pour échelonner le système.

Ici, nous pouvons remplacer l'équation (2) par : $(2) - 10 \times (1)$ ce qui donne

$$\begin{cases} x + 2y = 2 & (1) \\ 0 = 20 - 20 = 0 & (2) - 10 \times (1) = (2)' \end{cases}$$

L'équation $(2)'$ est vraie pour toute valeur de x et y . Cela signifie que les deux équations (1) et $(2)'$ sont équivalentes et ont les mêmes solutions. Le système revient donc à une seule équation à deux inconnues : il y a une infinité de solutions. x peut prendre n'importe quelle valeur du moment que y satisfait à l'équation (1) ou $(2)'$.

Exemple 2 : Cherchons à résoudre le système suivant :

$$\begin{cases} x + 2y = 2 & (1) \\ 10x + 20y = 10 & (2) \end{cases}$$

La représentation matricielle du système est :

$$\underbrace{\begin{pmatrix} 1 & 2 \\ 10 & 20 \end{pmatrix}}_A \underbrace{\begin{pmatrix} x \\ y \end{pmatrix}}_X = \underbrace{\begin{pmatrix} 2 \\ 10 \end{pmatrix}}_B$$

Calculons le déterminant de A :

$$\det A = \begin{vmatrix} 1 & 2 \\ 10 & 20 \end{vmatrix} = 20 - 20 = 0$$

La matrice A n'est pas inversible. Il faut donc déterminer si le système a des solutions ou non.

Pour déterminer cela, il faut utiliser la même méthode que pour la résolution par échelonnement : effectuer des combinaisons linéaires des équations du système pour échelonner le système.

Ici, nous pouvons remplacer l'équation (2) par : $(2) - 10 \times (1)$ ce qui donne

$$\begin{cases} x + 2y = 2 & (1) \\ 0 = 10 - 20 = -10 & (2) - 10 \times (1) = (2)' \end{cases}$$

L'équation $(2)'$ est impossible. Il n'y a pas de solution.

Exemple 3 : Reprenons le système déjà résolu précédemment par échelonnement :

$$\begin{cases} x + 2y + 3z = 10 \\ -x + y + z = 0 \\ x + y + z = 6 \end{cases}$$

La représentation matricielle de ce système est :

$$\underbrace{\begin{pmatrix} 1 & 2 & 3 \\ -1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}}_A \underbrace{\begin{pmatrix} x \\ y \\ z \end{pmatrix}}_X = \underbrace{\begin{pmatrix} 10 \\ 0 \\ 6 \end{pmatrix}}_B$$

A est une matrice carrée. On peut donc calculer son déterminant et voir si elle est inversible.

$$\det A = \begin{vmatrix} 1 & 2 & 3 \\ -1 & 1 & 1 \\ 1 & 1 & 1 \end{vmatrix} = 1 \times (1 - 1) - (-1) \times (2 - 3) + 1 \times (2 - 3) = -2$$

Le déterminant n'est pas nul. La matrice est inversible. Cherchons sa matrice inverse :

$$\begin{aligned} \text{com } A &= \begin{pmatrix} 0 & 2 & -2 \\ 1 & -2 & 1 \\ -1 & -4 & 3 \end{pmatrix} \implies {}^t \text{com } A = \begin{pmatrix} 0 & 1 & -1 \\ 2 & -2 & -4 \\ -2 & 1 & 3 \end{pmatrix} \\ A^{-1} &= \frac{{}^t \text{com } A}{\det A} = \frac{1}{(-2)} \times \begin{pmatrix} 0 & 1 & -1 \\ 2 & -2 & -4 \\ -2 & 1 & 3 \end{pmatrix} \end{aligned}$$

La solution du système s'obtient en multipliant les matrices :

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \frac{1}{(-2)} \times \begin{pmatrix} 0 & 1 & -1 \\ 2 & -2 & -4 \\ -2 & 1 & 3 \end{pmatrix} \begin{matrix} X = A^{-1}B \\ \begin{pmatrix} 10 \\ 0 \\ 6 \end{pmatrix} \end{matrix} = \begin{pmatrix} \frac{-6}{-2} \\ \frac{20-24}{-2} \\ \frac{-20+18}{-2} \end{pmatrix} = \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}$$

La solution est, bien entendu, la même que celle obtenue à l'aide de la méthode par échelonnement.

Chapitre 3 PRIMITIVES ET INTÉGRALES

3.1 Définitions et exemples

Soient deux fonctions $F(x)$ et $f(x)$ de la variable x . On dit que F est une **primitive** de f si $f = \frac{dF}{dx}$.

Par exemple $\frac{d \sin x}{dx} = \cos x$. On dira que $\sin x$ est une primitive de $\cos x$

$$F = \text{primitive de } f \Leftrightarrow F'(x) := \frac{dF(x)}{dx} = f(x)$$

Les primitives, $F_1(x)$ et $F_2(x)$ d'une même fonction $f(x)$ diffèrent d'une constante

$$F_2(x) - F_1(x) = C$$

où C est une constante arbitraire.

En outre on vérifie aisément les relations

$$\left. \begin{array}{l} U(x) = \text{primitive de } u(x) \\ V(x) = \text{primitive de } v(x) \\ a = \text{constante} \end{array} \right\} \Rightarrow U + aV = \text{primitive de } (u + av)$$

Les tableaux ci-dessous donnent une primitives des fonctions usuelles

Fonction $f =$	a	$x^n \ (n \neq -1)$	$x^{1/2} = \sqrt{x}$	$x^{-1} = \frac{1}{x}$
Primitive de $f =$	ax	$\frac{x^{n+1}}{n+1}$	$\frac{2}{3} x \sqrt{x}$	$\ln x $

Fonction $f =$	e^{ax}	$\frac{u'}{u}$	$\ln x $
Primitive de $f =$	$\frac{1}{a} e^{ax}$	$\ln u $	$x \ln x - x$

Fonction $f =$	$\sin(ax)$	$\cos(ax)$	$\tan(ax) = \frac{\sin(ax)}{\cos(ax)}$
Primitive de $f =$	$-\frac{1}{a} \cos(ax)$	$\frac{1}{a} \sin(ax)$	$-\frac{1}{a} \ln \cos(ax) $

3.2 Intégrale définie

Considérons la fonction $f(x)$ et les deux valeurs $x = a$ et $x = b$. L'intervalle $[a, b]$ est divisé en n intervalles élémentaires dont les extrémités sont :
 $x_0 = a, x_1 = a + \delta x, \dots, x_k = a + k \delta x, \dots, x_n = a + n \delta x = b$ avec $\delta x = (b - a)/n$.

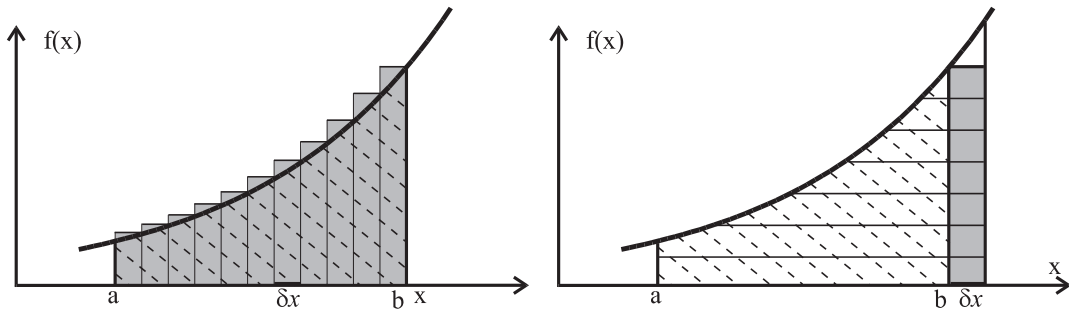


fig. 3.1

Nous considérons $I_n = f(x_1)\delta x + f(x_2)\delta x + \dots = \sum f(x_k)\delta x$. La somme I_n représente l'aire grise des rectangles de la première figure 3.1. Lorsque $n \rightarrow \infty$ (c'est à dire $\delta x \rightarrow 0$) l'aire grise se confond avec l'aire hachurée de la figure 3.1, comprise entre le graphe de $f(x)$ et l'axe des x . Cette limite est "**l'intégrale définie**" $I = \int_a^b f(x)dx$.

Remarquons que **le nom de la variable d'intégration est indifférent** :

$$I = \int_a^b f(x) dx = \int_a^b f(t) dt = \int_a^b f(u) du = \dots$$

- Supposons que f soit une combinaison linéaire des fonctions $g(x)$ et $h(x)$, c'est à dire $f(x) = g(x) + \lambda h(x)$ où λ est une constante. Il vient

$I_n = \sum f(x_k)\delta x = \sum g(x_k)\delta x + \lambda \sum h(x_k)\delta x$. A la limite $n \rightarrow \infty$ on trouve

$$\boxed{\int_a^b (g + \lambda h)(t) dt = \int_a^b g(t) dt + \lambda \int_a^b h(t) dt}$$

- Sur la figure 3.1 nous avons considéré le cas $b > a$. Lorsque $b < a$ la quantité δx change de signe alors que les valeurs de $f(x_k)$ restent inchangées, ce qui implique

$$\boxed{\int_a^b f(t) dt = -\int_b^a f(t) dt \quad \text{et donc} \quad \int_a^a f(t) dt = 0}$$

- En outre, si on introduit un troisième point d'abscisse c , on vérifie graphiquement la relation

$$\boxed{\int_a^b f(t) dt + \int_b^c f(t) dt = \int_a^c f(t) dt}$$

valable quelle que soit la position relative des points d'abscisse a, b et c .

Lorsque la fonction $f(x)$ ne conserve pas un signe constant, certains termes de la somme I_n sont positifs, d'autres sont négatifs. L'aire I apparaît donc comme une aire algébrique.

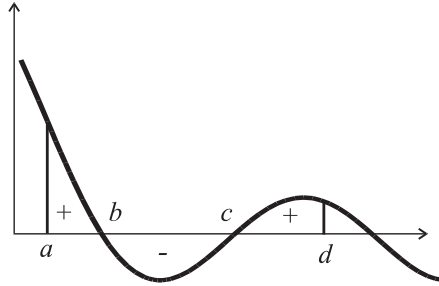


fig. 3.2

L'intégrale $\int_a^d f(x) dx$ où $f(x)$ est représentée sur la figure 3.2, se décompose en une somme de trois intégrales : $\int_a^d = \int_a^b + \int_b^c + \int_c^d$. Considérant la série I_n qui approxime chacune des intégrales on obtient les relations $\int_a^b f(x) dx > 0$, $\int_b^c f(x) dx < 0$ et $\int_c^d f(x) dx > 0$.

Sur la figure 3.3 nous représentons divers cas possibles pour le signe de $\int_a^b f(x) dx$.

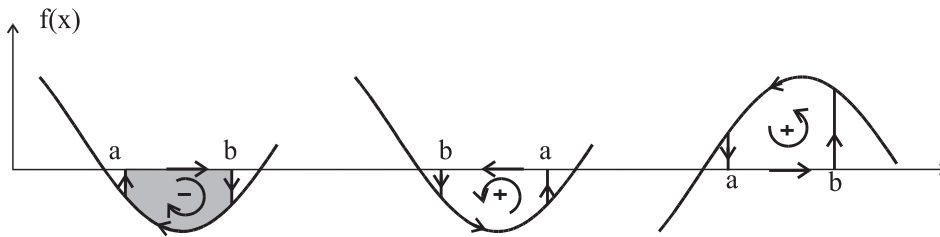


fig. 3.3

Dans le premier cas, $f(x) < 0$ et $b > a$ (l'axe des x est décrit dans le sens positif indiqué par la flèche sur l'axe). L'intégrale est négative. Ce résultat apparaît clairement en considérant le signe des termes des séries I_n . Cette intégrale représente l'aire algébrique de la surface grise. Le contour qui borde cette surface est orienté par continuité à partir de la flèche sur l'axe des x . L'orientation est rétrograde (sens des aiguilles d'une montre*). Cette orientation fixe le signe de l'intégrale. On vérifie aisément dans les deux autres cas que l'orientation du bord est celle du sens direct ; les intégrales correspondantes sont positives.

On peut considérer que b , borne supérieure de l'intégrale est une variable, x (fig.3.1). Dans ce cas l'intégrale est une fonction de x : $\Phi(x) := \int_a^x f(t) dt$. Il serait

*Un montre à aiguille "normale".

maladroit ici de noter la variable d'intégration du même nom que la borne supérieure; nous employons donc t pour la désigner.

N.B. La borne supérieure, x , peut être plus grande ou plus petite que la borne inférieure a . La fonction $\Phi(x)$ peut être positive ou négative.

La dérivée de $\Phi(x)$ en $x = b$ est la limite de $(\Phi(x + \delta x) - \Phi(x)) / \delta x$ lorsque $\delta x \rightarrow 0$. La différence entre les deux aires hachurées de la second figure 3.1 est précisément la quantité $\Phi(x + \delta x) - \Phi(x)$. C'est pratiquement l'aire $f(b) \delta x$ du rectangle gris (à la limite $\delta x \rightarrow 0$). Par conséquent $\Phi'(b) = f(b)$:

$$\boxed{\frac{d}{dx} \int_a^x f(t) dt = f(x)}$$

$\Phi(x) := \int_a^x f(t) dt$ est donc une primitive de $f(x)$. C'est la primitive telle que $\Phi(a) = 0$.

Considérons une primitive quelconque $F(x)$ de $f(x)$: $F(x) = \Phi(x) + C$ où C est une constante.

$$\text{Formons } F(b) - F(a) = (\Phi(b) + C) - (\Phi(a) + C) = \Phi(b) - \Phi(a).$$

$$\text{De la relation } \Phi(a) = 0 \text{ on déduit } F(b) - F(a) = \Phi(b) = \int_a^b f(t) dt :$$

$$\boxed{F(x) = \text{une primitive de } f(x) \Rightarrow \int_a^b f(t) dt = F(b) - F(a)}$$

$F(b) - F(a)$ est donc l'intégrale définie de f sur l'intervalle (a, b) . On utilise généralement la notation $[F(x)]_a^b := F(b) - F(a)$.

Premier exemple. Calculons l'aire comprise entre un arc de sinusöide et l'axe des x . Plus précisément calculons $A = \int_0^\pi \sin x dx$. Une primitive de $\sin x$ est $-\cos x$. On écrit $A = [-\cos x]_0^\pi := (-\cos \pi) - (-\cos 0) = 1 + 1 = 2$.

Second exemple. Nous voulons calculer le moment d'inertie d'une barre homogène, par rapport à son extrémité. La barre est homogène, sa masse est M et sa longueur L . Nous décomposons la barre en éléments de longueur $d\ell$.

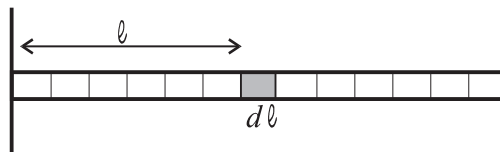


fig. 3.4

Pour calculer le moment d'inertie, I , il faut (par définition) sommer toutes les quantités $\ell^2 dm$ où dm est la masse de l'élément de longueur $d\ell$. La barre étant homogène il vient $dm = M \cdot d\ell / L$.

$$I = \int_0^L \ell^2 \frac{M}{L} d\ell = \frac{M}{L} \left[\frac{\ell^3}{3} \right]_0^L = \frac{1}{3} ML^2$$

3.3 Méthodes d'intégration

Une primitive quelconque de la fonction f est notée $\int f(x)dx$, c'est une " **intégrale indéfinie** ".

Par exemple $\int e^x dx = e^x + C$ (C désigne une constante appelée "**constante d'intégration**").

Il n'est pas toujours possible d'exprimer une intégrale indéfinie sous la forme d'une combinaison de fonctions connues. Ainsi $I = \int \frac{1}{x^2+1} dx$ définit une nouvelle fonction qui ne s'exprime pas au moyen des fonctions usuelles (polynômes, exponentielles, fonctions trigonométriques, etc.). C'est une fonction notée $\text{Arctan } x = y$ telle que $y \in (-\pi, +\pi)$ et satisfait la relation $\tan y = x$. La fonction Arctan est donc la fonction inverse de la fonction \tan .

Plusieurs méthodes permettent l'intégration des fonctions simples nous les résumons rapidement.

3.3.1 Les tables et les ordinateurs.

Il existe des livres qui donnent les primitives connues des fonctions simples; ce sont des "tables d'intégrales". Ces mêmes tables fournissent également la valeur numérique des intégrales définies remarquables comme $\int_0^\infty e^{-x^2} dx = \frac{1}{2}\sqrt{\pi}$.

Aujourd'hui nous disposons de logiciels puissants, d'utilisation commode, qui permettent de trouver (au moyen d'un ordinateur) la réponse à la plupart des problèmes d'intégration connus.

L'essentiel de l'effort doit donc porter sur la formulation des problèmes, sur la nature des réponses cherchées plus que sur les techniques de résolution elles-mêmes. Les méthodes élémentaires doivent néanmoins être connues.

3.3.2 La simplification du problème.

On utilise directement la relation

$$\int \left(\frac{df}{dx} \right) dx := \boxed{\boxed{\int df = f(x) + C}} \quad (3.1)$$

après avoir simplifié l'expression sous le signe \int .

Premier exemple. Soit à calculer $I = \int \tan^2 x dx$.

On sait que $\frac{d}{dx} \tan x = 1 + \tan^2 x$, on en déduit $I = \int \left(\frac{d}{dx} \tan x - 1 \right) dx$ soit $I = \int d(\tan x - x)$; d'où la solution $I = \tan x - x + C$.

Second exemple. $I = \int \frac{x^2}{x^2-1} dx$.

On écrit $\frac{x^2}{x^2-1} = \frac{x^2-1}{x^2-1} + \frac{1}{x^2-1} = 1 + \frac{1}{(x-1)(x+1)}$.

On peut démontrer que toute fraction rationnelle de la forme

$Q = \frac{1}{(x-a_1)(x-a_2)\dots(x-a_n)}$ s'écrit encore $Q = \frac{A_1}{(x-a_1)} + \frac{A_2}{(x-a_2)} + \dots + \frac{A_n}{(x-a_n)}$ où A_k est une constante. Dans le cas présent, il vient $\frac{1}{(x-1)(x+1)} = \frac{1}{2(x-1)} - \frac{1}{2(x+1)}$.

Par conséquent

$I = \int \left(1 + \frac{1}{2(x-1)} - \frac{1}{2(x+1)} \right) dx = x + \frac{1}{2} \ln|x-1| - \frac{1}{2} \ln|x+1| + C$ soit encore

$I = x + \ln \sqrt{\frac{|x-1|}{|x+1|}} + C$.

3.3.3 L'intégration par partie.

Démontrons la relation

$$\int_a^X u \, dv = [uv]_a^X - \int_a^X v \, du \quad (3.2)$$

où u et v sont des fonctions de la variable x . Par conséquent $du = \frac{du}{dx} dx$ et $dv = \frac{dv}{dx} dx$.

Nous utilisons la relation $\frac{d}{dx}(uv) = v \frac{du}{dx} + u \frac{dv}{dx}$, il vient

$[uv]_a^X = \int_a^X \frac{d}{dx}(uv) \, dx = \int_a^X v \frac{du}{dx} dx + \int_a^X u \frac{dv}{dx} dx = \int_a^X v \, du + \int_a^X u \, dv$ ce qui conduit à l'expression 3.2.

La relation 3.2 est généralement écrite sous la forme

$$\boxed{\int u \, dv = uv - \int v \, du} \quad (3.3)$$

Premier exemple. Calculons $\int t e^t dt$. Nous posons $u = t$ et $dv = e^t dt$. Il vient donc $du = dt$ et $v = e^t$.

La relation 3.3 s'écrit alors $\int t (e^t dt) = t e^t - \int e^t dt = t e^t - e^t + C$.

Second exemple. Calculons $\int t^2 e^t dt$. Nous posons $u = t^2$ et $e^t dt = dv$ soit $du = 2t dt$ et $v = e^t$. Il vient donc $\int t^2 e^t dt = t^2 e^t - \int 2te^t dt = t^2 e^t - 2(t e^t - e^t) + C$

3.3.4 Le changement de variable.

Le changement de variables est la méthode que l'on rencontre le plus souvent ; c'est donc la méthode la plus importante.

Démontrons la relation suivante

$$\boxed{F(x) := \int f(x) \cdot dx \Rightarrow F(g[u]) = \int f(g[u]) \cdot \left(\frac{dg}{du}\right)_u du} \quad (3.4)$$

La définition de F implique $\frac{dF}{dx} = f(x)$ ou encore $\frac{dF}{dg} = f(g)$, avec le changement de notation $x \rightarrow g$.

Si g est une fonction de u il vient $\left(\frac{dF}{du}\right)_u = \left(\frac{dF}{dg}\right)_{g(u)} \cdot \left(\frac{dg}{du}\right)_u$ d'où

$\left(\frac{dF}{du}\right)_u = f(g[u]) \cdot \left(\frac{dg}{du}\right)_u$ ce qui signifie que F , considérée comme une fonction de u (c'est à dire $F(g[u])$) est une primitive de $f(g[u]) \left(\frac{dg}{du}\right)_u$. C'est précisément le sens de la seconde égalité dans 3.4.

Premier exemple. Soit à calculer $I = \int \ln x \, dx := \int f(x) dx$.

Posons $x := g(u) := e^u$. Il vient $f(x) := \ln x = \ln(e^u) = u = f(g[u])$ et $dx = \left(\frac{dg}{du}\right) du = e^u du$.

On obtient $I = \int f(g[u]) \left(\frac{dg}{du}\right)_u du = \int u e^u du$. Cette intégrale a été calculée au paragraphe précédent : $I = u e^u - e^u + C$. La relation $x = e^u \Leftrightarrow u = \ln x$ conduit à la solution $I = \ln x \cdot x - x + C$.

Plutôt que la formule 3.4 **il est important de retenir la méthode mise en oeuvre.**

Pour intégrer $F(x) := \int f(x) dx$

1. nous avons posé $x = g(u)$ (où de façon équivalente $u = h(x)$ où h et g sont les fonctions inverses l'une de l'autre),

2. nous avons traité dx comme une différentielle et nous avons exprimé $f(x) dx$ en fonction de u et du , nous avons donc posé $f(x) dx = f(g[u]) \left(\frac{dg}{du}\right) du := \varphi(u) du$,

3. nous avons intégré $\int \varphi(u) du = \Phi(u)$,

4. nous avons exprimé u en fonction de x : $F(x) = \Phi[h(x)]$.

Second exemple. Soit à calculer $F(x) := \int x e^{-x^2} dx$.

1. Posons $x^2 = u$.

2. Dans ces conditions $du = 2x dx$ et $x e^{-x^2} dx = \frac{1}{2} e^{-u} du$ ($= \varphi(u) du$).

3. Il vient $F = \int \frac{1}{2} e^{-u} du = \frac{-1}{2} e^{-u} + C$.

4. Nous exprimons u en fonction de x . Nous obtenons $F = \frac{-1}{2} e^{-x^2} + C$.

Pour conclure, soulignons que les calculs d'intégration sont parfois fastidieux tandis que les outils informatiques donnent immédiatement la solution dans la plupart des cas usuels. Cependant, si l'informatique peut résoudre les problèmes liés aux techniques de calcul, les questions conceptuelles restent essentielles. L'informatique permet d'aller très vite et d'éviter les erreurs de calcul, mais ignorer ce qu'est une dérivée partielle, une intégrale, une asymptote ne permet même pas de comprendre la nature des questions qui se posent ni, *a fortiori*, d'imaginer ce que peuvent en être les réponses.

Avec ou sans outils informatiques, le problème à traiter doit être bien formulé et l'objectif du calcul clairement défini. En outre, il ne faut pas perdre de vue que le travail d'un ordinateur doit être contrôlé.

La connaissance des méthodes courantes, la capacité à résoudre des problèmes simples font partie d'une culture scientifique indispensable.

Chapitre 4

LES ÉQUATIONS DIFFÉRENTIELLES

Les équations différentielles sont des équations reliant une ou plusieurs fonctions avec leurs dérivées. Elles apparaissent dans la plupart des problèmes étudiant les variations d'une quantité, par exemple par rapport au temps (évolution de populations, mouvement d'un ressort ou d'un solide, décroissance radioactive d'un élément) ou par rapport à une longueur (absorption des rayonnements par la matière). Nous allons nous intéresser à quelques équations différentielles particulières.

On appelle "**ordre**" d'une équation différentielle le degré de dérivation de la fonction dans l'équation : une équation du premier ordre concerne une fonction et sa dérivée ; une équation du deuxième ordre concerne la fonction, sa dérivée et sa dérivée seconde. On qualifie ces équations de "**linéaires**" lorsqu'elles sont des combinaisons linéaires de la fonction et de ses dérivées. Nous allons ainsi voir les techniques usuelles pour résoudre deux types d'équations différentielles.

NB : Nous allons considérer des équations différentielles à coefficients réels. Dans certains cas, les méthodes employées sont aussi valables si les coefficients sont complexes.

4.1 Équations différentielles du premier ordre, linéaires, à coefficients constants

Les équations différentielles du premier ordre, linéaires, à coefficients constants peuvent s'écrire sous la forme :

$$\frac{df}{dx} + a f(x) = g(x)$$

Avec a constante réelle connue et g une fonction connue de x . Nous cherchons à déterminer la fonction $f(x)$.

La méthode de résolution de cette équation suit, en général, deux étapes.

4.1.1 Résolution de l'équation sans second membre

L'équation sans "**second membre**" correspond au cas où $g(x) = 0$.

$$\frac{df_0}{dx} + a f_0(x) = 0$$

Les solutions de ce type d'équation s'obtiennent de la façon suivante :

$$\frac{df_0}{dx} = (-a) f_0(x) \implies f_0(x) = C e^{-ax}$$

C est une constante réelle, elle peut prendre a priori n'importe quelle valeur. La solution $f_0(x)$ n'est donc pas unique. Si l'on donne une indication supplémentaire, par exemple la valeur de $f_0(x)$ pour une valeur donnée de x , alors il est possible de fixer la valeur de la constante C .

4.1.2 Résolution de l'équation avec second membre

Pour résoudre l'équation complète :

$$\frac{df}{dx} + a f(x) = g(x)$$

nous allons nous inspirer du résultat précédent. Plus précisément, on montre qu'il est possible de rechercher les solutions en les écrivant de la forme :

$$f(x) = C(x) e^{-ax}$$

Ici, $C(x)$ n'est plus une constante, mais une fonction de la variable x . Cette méthode est appelée méthode de la variation de la constante et est valide dans ce cas particulier. Il s'agit en fait d'un changement de fonction dans le but de nous simplifier les calculs : nous recherchons donc l'expression de $C(x)$ dont découlera l'expression de $f(x)$.

Nous avons donc

$$\frac{df}{dx} = \frac{dC}{dx} e^{-ax} - a C(x) e^{-ax}$$

En remplaçant l'expression de f et de $\frac{df}{dx}$ dans l'équation différentielle, nous obtenons

$$\frac{df}{dx} + a f(x) = \frac{dC}{dx} e^{-ax} - a C(x) e^{-ax} + a C(x) e^{-ax} = g(x)$$

Soit encore

$$\begin{aligned} \frac{dC}{dx} e^{-ax} = g(x) &\implies \frac{dC}{dx} = g(x) e^{ax} \\ C(x) &= \int g(x) e^{ax} dx + K \\ f(x) = C(x) e^{-ax} &= \left(\int g(x) e^{ax} dx + K \right) e^{-ax} \end{aligned}$$

A condition, bien sûr, que la primitive $\int g(x) e^{ax} dx$ existe, K est alors une constante réelle. K peut prendre n'importe quelle valeur à priori. Comme pour l'équation sans second membre, si l'on donne une indication supplémentaire, il est alors possible de déterminer la valeur de K .

Exemple : $\frac{df}{dx} - 2 f(x) = 5 e^x$

Dans un premier temps, il faut donc résoudre l'équation sans second membre, puis suivre la méthode de la variation de la constante.

Equation sans second membre :

$$\frac{df_0}{dx} - 2 f_0(x) = 0 \implies f_0(x) = C e^{2x}$$

f_0 nous indique quelle forme doivent prendre les solutions de l'équation différentielle avec second membre

$$f(x) = C(x) e^{2x}$$

On a alors

$$\frac{df}{dx} - 2 f(x) = \frac{dC}{dx} e^{2x} + 2 C(x) e^{2x} - 2C(x) e^{2x} = \frac{dC}{dx} e^{2x} = 5 e^x$$

Soit

$$\frac{dC}{dx} = 5 e^{-x} \implies C(x) = \int 5 e^{-x} dx + K = -5 e^{-x} + K$$

Nous avons ainsi les solutions de l'équation différentielle

$$f(x) = C(x) e^{2x} = (-5 e^{-x} + K) e^{2x} = -5 e^x + K e^{2x}$$

4.2 Equations différentielles du deuxième ordre, linéaires, à coefficients constants

Une équation différentielle du deuxième ordre, linéaire, à coefficients constants peut toujours s'écrire

$$a \frac{d^2 f}{dx^2} + b \frac{df}{dx} + c f(x) = g(x)$$

avec a réel non-nul, b et c réels fixés et $g(x)$ fonction connue de la variable x .

4.2.1 Résolution de l'équation sans second membre

Ce que l'on appelle l'équation sans second membre correspond au cas où $g(x) = 0$.

$$a \frac{d^2 f_0}{dx^2} + b \frac{df_0}{dx} + c f_0(x) = 0$$

On montre que les solutions peuvent s'écrire de la forme

$$f_0(x) = C e^{rx}$$

où C est une constante réelle. Recherchons les valeurs de r .

$$\begin{aligned} \frac{df_0}{dx} &= r C e^{rx} \\ \frac{d^2 f_0}{dx^2} &= r^2 C e^{rx} \end{aligned}$$

Cela nous donne

$$\begin{aligned} a \frac{d^2 f_0}{dx^2} + b \frac{df_0}{dx} + c f_0(x) &= (a r^2 + b r + c) C e^{rx} = 0 \\ \implies &\boxed{a r^2 + b r + c = 0} \end{aligned}$$

Cette dernière équation du deuxième degré est appelée "**équation caractéristique**" de l'équation différentielle. Nous avons vu comment résoudre une telle équation. Trois cas peuvent se présenter : suivant le signe du discriminant : $\Delta := b^2 - 4ac$ l'équation caractéristique peut avoir deux racines réelles, une racine double réelle ou deux racines complexes. Examinons chaque cas.

- *Le discriminant est positif* : l'équation caractéristique a deux racines réelles que nous appelons r_1 et r_2 .

On peut montrer alors que la solution générale de l'équation différentielle sans second membre est la combinaison linéaire des solutions correspondant à chacune des racines obtenues

$$f_0(x) = C_1 e^{r_1 x} + C_2 e^{r_2 x}$$

C_1 et C_2 sont des constantes réelles arbitraires, c'est-à-dire, pouvant a priori prendre n'importe quelle valeur sauf si l'on fixe deux valeurs particulières de la fonction $f_0(x)$ ou de sa dérivée première. Suivant le signe de r_1 et r_2 , la solution de l'équation différentielle est donc une fonction exponentielle croissante ou décroissante ou encore une combinaison des deux.

- *Le discriminant est nul* : l'équation caractéristique a une racine double réelle que nous appelons r_0 .

On peut montrer alors que la solution générale de l'équation différentielle sans second membre est

$$f_0(x) = (C_0 x + C'_0) e^{r_0 x}$$

C_0 et C'_0 sont des constantes réelles arbitraires, c'est-à-dire, pouvant a priori prendre n'importe quelle valeur sauf si l'on fixe deux valeurs particulières de la fonction $f_0(x)$ ou de sa dérivée première.

- *Le discriminant est négatif* : l'équation caractéristique a deux racines complexes que nous appelons $z_1 = \alpha + i\beta$ et $z_2 = \alpha - i\beta$ (α et β sont des réels).

Comme dans le cas des racines réelles, on peut montrer alors que la solution générale de l'équation différentielle sans second membre est la combinaison linéaire des solutions correspondant à chaque des racines obtenues

$$f_0(x) = C_1 e^{z_1 x} + C_2 e^{z_2 x} = C_1 e^{(\alpha+i\beta)x} + C_2 e^{(\alpha-i\beta)x} = e^{\alpha x} (C_1 e^{i\beta x} + C_2 e^{-i\beta x})$$

En utilisant la formule de Moivre, on obtient

$$f_0(x) = e^{\alpha x} (K_1 \cos \beta x + K_2 \sin \beta x)$$

Ce qui peut aussi se mettre sous la forme

$$f_0(x) = \Gamma e^{\alpha x} \cos(\beta x + \varphi)$$

où $\Gamma = \sqrt{K_1^2 + K_2^2}$ et $\tan \varphi = -\frac{K_2}{K_1}$

Il faut choisir la forme de la solution que l'on préfère ; on définit alors, suivant le cas, un couple (C_1, C_2) ou (K_1, K_2) ou (Γ, φ) de constantes réelles arbitraires, c'est-à-dire, pouvant a priori prendre n'importe quelle valeur que l'on peut fixer avec deux valeurs particulières de la fonction $f_0(x)$ ou de sa dérivée première.

4.2.2 Résolution de l'équation avec second membre

Nous cherchons à résoudre l'équation complète

$$a \frac{d^2 f}{dx^2} + b \frac{df}{dx} + c f = g(x)$$

On peut montrer que cette équation admet une solution générale $f(x)$ qui est la somme de la solution générale de l'équation sans second membre $f_0(x)$ et d'une solution particulière de l'équation avec second membre $f_p(x)$.

$$f(x) = f_0(x) + f_p(x)$$

Dans la plupart des cas, cette solution particulière est de même type que le second membre de l'équation lui-même.

Au final, tous les solutions d'une équation différentielle du deuxième ordre comprennent deux constantes arbitraires, de même qu'il y a une constante arbitraire dans toute solution d'une équation différentielle du premier ordre.

Exemple : Nous cherchons les solutions de l'équation différentielle

$$\frac{d^2 f}{dt^2} - 5 \frac{df}{dt} + 6 f = 50 \cos 4t$$

Nous pouvons remarquer que le second membre de l'équation est un cosinus. Ce genre d'équation est souvent rencontré dans les problèmes liés aux oscillateurs et ondes. Dans ce type de problème, le coefficient devant f est une fréquence au carré, c'est la fréquence propre de l'oscillateur et le coefficient devant la dérivée de f est proportionnel à l'amortissement ou au frottement que subit le système.

Recherchons les solutions de l'équation sans second membre

$$\frac{d^2 f_0}{dt^2} - 5 \frac{df_0}{dt} + 6 f_0 = 0$$

L'équation caractéristique est

$$r^2 - 5r + 6 = 0$$

Le discriminant est égal à 1 ($\Delta = 25 - 4 \times 6 = 1$)

Il y a donc deux racines réelles : $r_1 = 2$ et $r_2 = 3$

Nous obtenons donc

$$f_0(x) = C_1 e^{2t} + C_2 e^{3t}$$

Ce sont toutes les solutions de l'équation sans second membre. Nous n'avons pas d'autre élément pour fixer les valeurs des constantes C_1 et C_2 qui restent donc arbitraires.

Remarque : Dans un problème de physique, les solutions exponentielles croissantes (comme celles que nous obtenons) ne sont généralement pas réalistes. Cet argument suffit souvent à en déduire que les constantes sont nulles. D'une autre côté, si nous avons obtenu des exponentielles décroissantes, nous aurions pu dire qu'au bout d'un certain temps, ces solutions deviennent négligeables par rapport à la solution particulière. Au final, des solutions d'une équation différentielle sans second membre dont les racines de l'équation caractéristique sont réelles sont rarement intéressantes du point de vue physique... Cependant nous les gardons pour résoudre notre exemple.

Cherchons maintenant une solution particulière de l'équation différentielle

$$\frac{d^2 f}{dt^2} - 5 \frac{df}{dt} + 6 f = 50 \cos 4t$$

Nous cherchons une solution du même type que le second membre

$$\begin{aligned} f_p(t) &= A \cos 4t + B \sin 4t \\ \frac{df_p}{dt} &= -4A \sin 4t + 4B \cos 4t \\ \frac{d^2 f_p}{dt^2} &= -16 \times (A \cos 4t + B \sin 4t) \end{aligned}$$

Remplaçons ces expressions dans l'équation différentielle

$$\begin{aligned} -16(A \cos 4t + B \sin 4t) - 5 \times (-4A \sin 4t + 4B \cos 4t) + 6 \times (A \cos 4t + B \sin 4t) \\ = 50 \cos 4t \\ \cos 4t (-16A - 20B + 6A - 50) = -\sin 4t (-16B + 20A + 6B) \end{aligned}$$

Cette relation doit être vraie pour toute valeur de t . Les deux parenthèses sont donc nulles.

$$\begin{cases} -16A - 20B + 6A - 50 = -10A - 20B - 50 = 0 \\ -16B + 20A + 6B = -10B + 20A = 0 \end{cases}$$

La deuxième équation indique que

$$B = 2A$$

En remplaçant dans la première équation, on obtient

$$\begin{aligned} -10A - 40A - 50 &= 0 \\ A = -1 \text{ et } B &= -2 \end{aligned}$$

La solution particulière s'écrit donc

$$f_p(t) = -\cos 4t - 2 \sin 4t$$

Au final, les solutions de l'équation différentielle que nous cherchons s'écrivent

$$f(t) = f_0(t) + f_p(t) = C_1 e^{2t} + C_2 e^{3t} - \cos 4t - 2 \sin 4t$$

avec C_1 et C_2 constantes réelles.

Chapitre 5

FONCTIONS DE PLUSIEURS VARIABLES

Dans de nombreux problèmes en physique, chimie ou biologie, il ne suffit pas de savoir utiliser une fonction à une variable pour pouvoir décrire correctement les phénomènes. L'espace a trois dimensions et nécessite donc trois variables (x, y, z) pour le décrire, parfois le temps peut intervenir. La thermodynamique utilise aussi un nombre de variables élevés (pression, température, volume, ...). Nous allons ainsi nous pencher dans ce chapitre sur les fonctions de plusieurs variables. Le cas est beaucoup plus complexe que dans le cas des fonctions d'une variable. En fait, il n'est pas possible de faire une étude complète d'une fonction de plusieurs variables de façon directe. Il faut aborder les propriétés de la fonction suivant différentes approches, par exemple en fixant la valeur de certaines variables ou en faisant des diagrammes 2D ou 3D. Nous allons donc voir quelques outils pour pouvoir utiliser et décrire ces fonctions.

5.1 Dérivée

5.1.1 Dérivée partielles du premier ordre.

Soit une fonction de deux variables $f(x, y)$. Dans un système d'axes orthonormés $\{Ox, Oy, Oz\}$, nous considérons le point de coordonnées $\{x, y, 0\}$ du plan xOy . En ce point, parallèlement à Oz , nous portons le point M de cote $z = f(x, y)$. Lorsque x et y varient, le point M décrit la surface S .

Les "**lignes de niveau**" sont les courbes dans le plan xOy telles que $f(x, y) = z_0$, z_0 étant une constante définissant la ligne de niveau. L'usage le plus courant des lignes de niveau est la représentation de l'altitude sur une carte géographique. Une ligne de niveau est aussi l'intersection de la surface S avec le plan $z = z_0$.

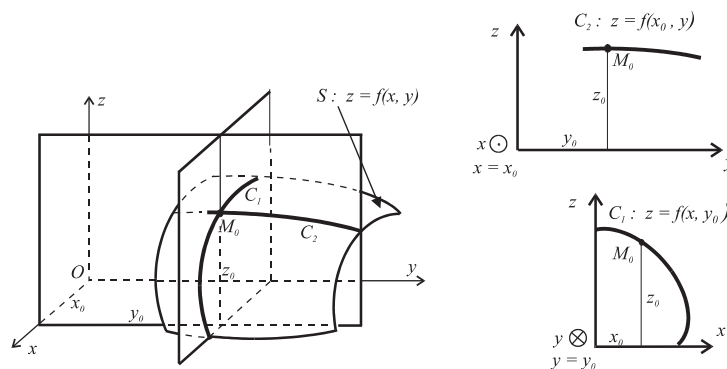


fig. 5.1

Considérons le point M_0 de coordonnées $x_0, y_0, z_0 := f(x_0, y_0)$. Coupons la surface S par le plan parallèle au plan xOz qui passe par M_0 . L'intersection est la courbe C_1 . Les points de C_1 ont tous la même ordonnée y_0 . Leur cote, z , dépend de leur abscisse $x : z = f(x, y_0)$. On peut dériver cette fonction de x ; on obtient la "**dérivée partielle**" de f par rapport à x . Cette dérivée partielle est notée $\left(\frac{\partial f}{\partial x}\right)_{(x, y_0)}$; c'est une fonction de x seulement car y_0 est fixé. On peut cependant, **après avoir calculé la dérivée partielle**, considérer que y_0 peut prendre une valeur arbitraire y . On obtient alors une fonction des deux variables x et $y : \left(\frac{\partial f}{\partial x}\right)_{(x, y)}$ que l'on note plus simplement $\left(\frac{\partial f}{\partial x}\right)$.

De même, on peut définir la fonction $\left(\frac{\partial f}{\partial y}\right)$.

Exemple 1 : L'aire d'un rectangle de côtés x et y s'écrit $z = x \cdot y$, c'est une fonction des deux variables x et y .

Considérons y comme une donnée. La fonction z est alors une fonction de x dont on peut calculer la dérivée. La dérivée partielle de z par rapport à x est $\frac{\partial z}{\partial x} = y$.

Et la dérivée partielle de z par rapport à $y : \frac{\partial z}{\partial y} = x$.

Exemple 2 : Soit $f(x, t) := x \cdot \cos(\omega t/2)$.

Pour calculer $\partial f/\partial x$, nous considérons t comme un paramètre constant et nous dérivons la fonction de $x : x \mapsto \Phi(x) := f(x, t)_{t=cte}$. Il vient $\frac{\partial f}{\partial x} = \cos\left(\frac{\omega t}{2}\right)$.

Pour calculer $\partial f/\partial t$, nous considérons x comme un paramètre constant et nous dérivons la fonction de $t : t \mapsto \Psi(t) := f(x, t)_{x=cte}$. On trouve $\frac{\partial f}{\partial t} = -\frac{\omega}{2} \cdot x \cdot \sin\left(\frac{\omega t}{2}\right)$.

Donnons nous les valeurs $x_0 = a$ et $t_0 = \frac{\pi}{2\omega}$. On obtient

$$\left(\frac{\partial f}{\partial x}\right)_{(x_0, t_0)} = \frac{\sqrt{2}}{2}, \quad \left(\frac{\partial f}{\partial t}\right)_{(x_0, t_0)} = -\frac{\omega}{2} a \frac{\sqrt{2}}{2} = -\frac{\sqrt{2}}{4} a \omega$$

5.1.2 Dérivées partielles d'ordre supérieurs à 1.

Etant donnée une fonction de deux variables, $f(x, y)$, les dérivées partielles du premier ordre $\left(\frac{\partial f}{\partial x}\right)$ et $\left(\frac{\partial f}{\partial y}\right)$ sont des fonctions des deux variables x et y . On peut les dériver à leur tour. On obtient les dérivées partielles du second ordre. Il y en a quatre. Nous en donnons la liste ci-dessous

$$\begin{aligned} \frac{\partial}{\partial x} \left(\frac{\partial f}{\partial x}\right) &: = \frac{\partial^2 f}{\partial x^2}, & \frac{\partial}{\partial y} \left(\frac{\partial f}{\partial y}\right) &:= \frac{\partial^2 f}{\partial y^2} \\ \frac{\partial}{\partial y} \left(\frac{\partial f}{\partial x}\right) &: = \frac{\partial^2 f}{\partial y \partial x}, & \frac{\partial}{\partial x} \left(\frac{\partial f}{\partial y}\right) &:= \frac{\partial^2 f}{\partial x \partial y} \end{aligned}$$

Exemple : Soit $f(x, y) := x^2 \sin y$.

$$\begin{aligned} \frac{\partial f}{\partial x} &= 2x \sin y, & \frac{\partial f}{\partial y} &= x^2 \cos y \\ \frac{\partial^2 f}{\partial x^2} &= 2 \sin y, & \frac{\partial^2 f}{\partial y^2} &= -x^2 \sin y \\ \frac{\partial^2 f}{\partial y \partial x} &= 2x \cos y, & \frac{\partial^2 f}{\partial x \partial y} &= 2x \cos y, \end{aligned}$$

On remarque que $\frac{\partial^2 f}{\partial y \partial x} = \frac{\partial^2 f}{\partial x \partial y}$. Cette propriété, connue comme le "lemme de Schwartz", est très généralement vérifiée pour la plupart des fonctions que l'on rencontre en physique*. Selon ce lemme, l'ordre dans lequel on effectue les dérivations est sans importance.

On peut continuer à dériver et calculer les dérivées du troisième ordre de l'exemple précédent :

$$\frac{\partial^3 f}{\partial x^3} = 0, \quad \frac{\partial^3 f}{\partial x^2 \partial y} = 2 \cos y, \quad \frac{\partial^3 f}{\partial x \partial y^2} = -2x \sin y, \quad \frac{\partial^3 f}{\partial y^3} = -x^2 \cos y, \text{ etc}$$

On remarquera les égalités $\frac{\partial^3 f}{\partial y \partial x^2} = \frac{\partial^3 f}{\partial x \partial y \partial x} = \frac{\partial^3 f}{\partial x^2 \partial y}$, ce qui est une conséquence du lemme de Schwartz appliqué à f et à ses dérivées.

Dans le calcul d'une dérivée partielle d'ordre quelconque d'une fonction suffisamment régulière, **le résultat ne dépend pas de l'ordre des dérivations**.

5.1.3 Généralisation et dimensions physiques.

Les résultats précédents se généralisent aux fonctions de plusieurs variables en nombre quelconque. Si ces fonctions sont suffisamment régulières, leurs dérivées partielles peuvent être calculées à un ordre arbitraire et ces dérivations peuvent être effectuées dans un ordre quelconque.

La dérivée partielle $\frac{\partial f}{\partial x}$ est la limite du rapport $\frac{\delta f}{\delta x}$ lorsque δx tend vers zéro, les autres variables restant constantes. Les dimensions physiques de $\frac{\partial f}{\partial x}$ et $\frac{\delta f}{\delta x}$ sont donc les mêmes.

Premier exemple. La grandeur f est la pression atmosphérique dont la valeur dépend de la position du point considéré dans l'atmosphère : $f = f(x, y, z)$. La pression s'exprime en Pa (soit en $\text{kg m}^{-1} \text{s}^{-2}$). Les variables x, y, z s'expriment en m. Les dérivées premières de f s'expriment donc en Pa m^{-1} (soit en $\text{kg m}^{-2} \text{s}^{-2}$). Les dérivées secondes sont les dérivées des dérivées premières. Elles s'expriment donc en $(\text{Pa m}^{-1}) \text{m}^{-1} = \text{Pa m}^{-2}$. De façon générale les dérivées d'ordre n s'expriment en Pa m^{-n} .

Dans l'expression $\frac{\partial^n f}{\partial x^{n_1} \partial y^{n_2} \dots} := D$ il faut considérer la notation $\partial^n f$ comme la variation d'une variation d'une variation d'une etc... d'une variation de f , c'est à dire comme une grandeur de mêmes dimensions que f tandis que la notation $\partial x^{n_1} \partial y^{n_2} \dots$ doit être considérée comme un produit de variations $(\delta x)^{n_1} (\delta y)^{n_2} \dots$.

On remarquera que la somme $n_1 + n_2 + \dots$ est égale à l'ordre de variation n . La dimension physique de D est donc la même que celle de $\frac{f}{x^{n_1} \cdot y^{n_2} \dots}$.

Second exemple. La température T d'une masse donnée de gaz s'exprime en fonction de sa pression P et de son volume V : $T = T(P, V)$. Quelle est la dimension physique de $\frac{\partial^3 T}{\partial V^2 \partial P}$? Ce sont les mêmes dimensions que celles de $\frac{T}{V^2 P}$. Les unités en sont $\text{K m}^{-6} \text{Pa}^{-1} = \text{K kg}^{-1} \text{m}^{-5} \text{s}^2$.

*Fonctions dérivables deux fois à dérivées continues.

5.2 Différentielle

Considérons $f(x, y, z, t)$, fonction de quatre variables. Soit M_0 l'ensemble des 4 valeurs données (x_0, y_0, z_0, t_0) . On définit la différentielle de f , en M_0 , de la façon suivante

$$df := \left(\frac{\partial f}{\partial x} \right)_0 dx + \left(\frac{\partial f}{\partial y} \right)_0 dy + \left(\frac{\partial f}{\partial z} \right)_0 dz + \left(\frac{\partial f}{\partial t} \right)_0 dt$$

La différentielle df est une fonction linéaire des 4 variables arbitraires dx, dy, dz et dt tandis que les dérivées partielles sont des constantes calculées en M_0 .

La définition se généralise à un nombre quelconque de variables. Dans le cas des fonctions d'une seule variable on retrouve la définition de la section précédente.

On remarquera que $\left(\frac{\partial f}{\partial x} \right)$ a mêmes dimensions que f/x tandis que dx a mêmes dimensions que x . Par conséquent $\left(\frac{\partial f}{\partial x} \right) dx$ a mêmes dimensions physiques que f . On en déduit que **la différentielle de f a mêmes dimensions physiques que f** .

Exemple. Soit à calculer la différentielle de $f(x, y) = x^2y$ en $x_0 = 1, y_0 = 2$.

$$\begin{aligned} \frac{\partial f}{\partial x} &= 2xy, \quad \left(\frac{\partial f}{\partial x} \right)_0 = 4; \quad \frac{\partial f}{\partial y} = x^2, \quad \left(\frac{\partial f}{\partial y} \right)_0 = 1 \Rightarrow \\ df &= 4dx + dy \text{ en } M_0(1, 2) \end{aligned}$$

La valeur numérique de df s'obtient alors en remplaçant dx et dy par leur valeur numérique.

Si on ne spécifie pas la valeur numérique de x_0 et y_0 on écrit (en abandonnant l'indice " 0 ")

$$df = 2xy dx + x^2 dy$$

et de façon plus générale $f = f(x, y, z, \dots) \Rightarrow$

$$df := \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy + \frac{\partial f}{\partial z} dz + \dots$$

En physique on est souvent conduit à étudier de petites quantités notées parfois δG ou dG . Ce ne sont pas nécessairement les différentielles d'une fonction. Mais lorsque dG est la différentielle d'une fonction, pour insister sur ce point qui s'avère souvent important, on spécifie que c'est une **"différentielle totale"**.

Etant donnée une forme différentielle

$$\delta G = A(x, y) dx + B(x, y) dy$$

le lemme de Schwartz fournit une condition nécessaire pour que ce soit une différentielle totale. En effet si δG est la différentielle d'une fonction f , alors $A(x, y) = \frac{\partial f}{\partial x}$ et

$B(x, y) = \frac{\partial f}{\partial y}$, ce qui implique $\frac{\partial B}{\partial x} = \frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x} = \frac{\partial A}{\partial y}$. Il en découle la condition

$$\frac{\partial}{\partial y} A(x, y) = \frac{\partial}{\partial x} B(x, y)$$

On peut démontrer que cette condition nécessaire est également suffisante.

Dans le cas de fonctions de plusieurs variables, les petites variations sont pratiquement égales aux différentielles. Plus précisément soit $f(x, y, z)$ une fonction de trois variables. Les variables subissent les variations $x = a \rightarrow a + \delta x$, $y = b \rightarrow b + \delta y$, $z = c \rightarrow c + \delta z$. La fonction f subit alors l'accroissement δf , correspondant :
 $f(a, b, c) \rightarrow f(a + \delta x, b + \delta y, c + \delta z) := f(a, b, c) + \delta f$.

Lorsque $\delta x = dx$, $\delta y = dy$, $\delta z = dz$, ... sont assez petits, il vient

$$\delta f \simeq df := \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy + \frac{\partial f}{\partial z} dz + \dots$$

Premier exemple. Nous voulons calculer $K = 1,1 \cdot \sqrt{4,2}$.

Nous posons $f(x, y) = x\sqrt{y}$. Ce qui donne $K = f(x, y)$ pour $x = 1,1$ et $y = 4,2$. Il est aisé de calculer $f(a, b)$ pour $a = 1$ et $b = 4$. Nous posons donc $a = 1$, $b = 4$, $dx = 0,1$ et $dy = 0,2$.

$$K = (a + dx)\sqrt{b + dy} = a\sqrt{b} + \delta f \text{ avec } \delta f \simeq df.$$

Le calcul donne $f(a, b) = 2$ et $df = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy$ avec $\frac{\partial f}{\partial x} = \sqrt{y}$ et $\frac{\partial f}{\partial y} = \frac{x}{2\sqrt{y}}$.

Ici il vient $df = \sqrt{b} dx + \frac{a}{2\sqrt{b}} dy = 2dx + 0,25 dy = 0,2 + 0,05 = 0,25$ d'où $K \simeq 2,25$ (le résultat exact est 2,254...).

Deuxième exemple. L'aire, F , d'un rectangle de côtés x et y est $F = xy$.

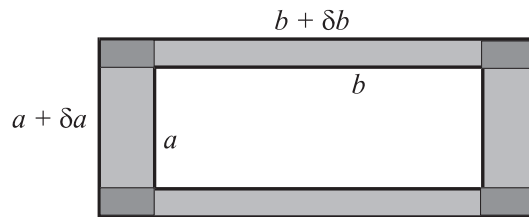


fig. 5.2

Le rectangle intérieur de la figure 5.2 a pour aire $F_0 = ab$. Le rectangle extérieur a pour aire $F_1 = (a + \delta a)(b + \delta b)$. La valeur exacte de l'accroissement est $\delta F = F_1 - F_0 = (a\delta b + b\delta a) + \delta a\delta b$. C'est l'aire de la bande grise (gris clair et gris foncé) entre les deux rectangles. δF est la somme de deux termes. Le terme $(a\delta b + b\delta a)$ ne contient que des termes du premier degré relativement aux accroissements δa et δb : c'est le terme du **premier ordre**. Le terme $\delta a\delta b$ est du second degré par rapport à δa et δb : c'est le terme du **second ordre**.

La différentielle de F en $(x = a, y = b)$ est $dF = a dy + b dx$, sa valeur pour $dx = \delta a$ et $dy = \delta b$ est $dF = a\delta b + b\delta a$. C'est l'aire des quatre rectangles gris clair. La différentielle de F fournit le terme du premier ordre de δF . En posant $\delta F \simeq dF$ on commet une erreur absolue du second ordre, $e = \delta a\delta b$. C'est l'aire des quatre rectangles gris foncé. Lorsque les variations sont petites ($\delta a \ll a$ et $\delta b \ll b$) cette erreur est négligeable comparée au terme du premier ordre que constitue la différentielle.

5.3 Expressions remarquables

Soient u et v deux fonctions d'une ou plusieurs variables et a une constante. En utilisant les définitions et les propriétés précédentes on obtient

$u = cte \iff du = 0$	$d(u + av) = du + adv$
$d(uv) = u dv + v du$	$d\left(\frac{u}{v}\right) = \frac{v du - u dv}{v^2}$
$d(u^n) = n u^{n-1} du$	$d \ln u = \frac{du}{u}$

Exemple. Le calcul de $d\left(\frac{1}{h}\right)$ peut se faire de diverses façons.

La relation $d(u^n) = n u^{n-1} du$ avec $n = -1$ et $u = h$ donne

$$d\left(\frac{1}{h}\right) = d(h^{-1}) = -h^{-2} dh = \frac{-1}{h^2} dh.$$

La relation $d\left(\frac{u}{v}\right) = \frac{v du - u dv}{v^2}$ avec $v = h$ et $u = 1$ (et donc $du = 0$) donne

$$d\left(\frac{1}{h}\right) = \frac{-dh}{h^2}.$$

Considérons la quantité $F = u^n \cdot v^m \cdot \dots$. Il vient $\ln |F| = n \ln |u| + m \ln |v| + \dots$

Nous calculons les différentielles de chaque membre (ce que l'on appelle **la différentielle logarithmique de F**).

En utilisant les relations $d \ln |F| = \frac{dF}{F}$ ainsi que $d \ln |u| = \frac{du}{u}$ et $d \ln |v| = \frac{dv}{v}$, nous obtenons

$$\boxed{F = u^n \cdot v^m \cdot \dots \implies \frac{dF}{F} = n \frac{du}{u} + m \frac{dv}{v} + \dots} \quad (5.1)$$

Exemple. $F = y \cdot \sqrt{x} / (x + 1)$. On écrit $F = y \cdot x^{1/2} \cdot (x + 1)^{-1}$.

La relation 5.1 s'écrit $dF/F = dy/y + \frac{1}{2} dx/x - d(x + 1)/(x + 1)$.

Avec $d(x + 1) = dx$, on obtient $\frac{dF}{F} = \frac{dy}{y} + \frac{1 - x}{2x(x + 1)} dx$.

5.4 Calcul d'incertitudes

5.4.1 Rappels

Le résultat d'une mesure est en général donné sous la forme

$$F = f \pm \Delta f$$

Par exemple pour un longueur : $F = 5,231 \text{ m} \pm 2 \cdot 10^{-3} \text{ m}$. L'"**incertitude absolue**" sur la mesure est Δf ; c'est une grandeur positive tandis que f est l'estimation de la grandeur mesurée. Une telle expression a même signification que les inégalités

$$f - \Delta f < F < f + \Delta f$$

On définit aussi l'"**incertitude relative**" $\Delta f / |F|$. La vraie valeur de F n'est pas connue, on estime donc l'incertitude relative au moyen de la grandeur $\Delta f / |f|$.

L'incertitude absolue s'exprime avec les mêmes unités que f ; par contre l'incertitude relative est un nombre pur, sans unités.

Il ne faut pas confondre incertitude et erreur. L'"**erreur**", e , est défini par la relation $e := f - F$; elle est inconnue. Si e était connue, $F = f - e$ serait connu sans incertitude car f est connu.

Pour déterminer les incertitudes, on peut estimer un majorant des erreurs. On peut aussi répéter plusieurs fois la même mesure et apprécier les petites variations des résultats obtenus. Cette méthode conduit à la notion d'incertitude standard.

5.4.2 Calcul d'incertitudes

Considérons une grandeur physique, F , fonction de deux variables, x et y , que l'on mesure. Le résultat des mesures fournit les estimations \tilde{x} et \tilde{y} des vraies grandeurs x et y . Nous posons $x = \tilde{x} + \delta x$ et $y = \tilde{y} + \delta y$. Nous ne connaissons ni δx ni δy . Nous savons seulement que ces quantités satisfont les relations $|\delta x| < \Delta x$ et $|\delta y| < \Delta y$. Les incertitudes Δx et Δy sont connues ou tout au moins estimées convenablement.

De tels résultats s'expriment sous la forme $x = \tilde{x} \pm \Delta x$ et $y = \tilde{y} \pm \Delta y$.

Nous cherchons à estimer F et son incertitude. La quantité $\tilde{F} := F(\tilde{x}, \tilde{y})$ est une estimation de F . Au premier ordre $\delta F := F - \tilde{F} \simeq dF$.

Il vient donc $F \simeq \tilde{F} + dF = \tilde{F} + \left(\frac{\partial F}{\partial x}\right)_{(\tilde{x}, \tilde{y})} \delta x + \left(\frac{\partial F}{\partial y}\right)_{(\tilde{x}, \tilde{y})} \delta y$.

La seule indication que nous ayons est $|\delta x| < \Delta x$; c'est à dire $-\Delta x < \delta x < \Delta x$,

ce qui implique $-\left|\left(\frac{\partial F}{\partial x}\right)_{(\tilde{x}, \tilde{y})}\right| \Delta x < \left(\frac{\partial F}{\partial x}\right)_{(\tilde{x}, \tilde{y})} \delta x < \left|\left(\frac{\partial F}{\partial x}\right)_{(\tilde{x}, \tilde{y})}\right| \Delta x$.

De même $-\left|\left(\frac{\partial F}{\partial y}\right)_{(\tilde{x}, \tilde{y})}\right| \Delta y < \left(\frac{\partial F}{\partial y}\right)_{(\tilde{x}, \tilde{y})} \delta y < \left|\left(\frac{\partial F}{\partial y}\right)_{(\tilde{x}, \tilde{y})}\right| \Delta y$.

En additionnant membres à membres les inégalités précédentes nous obtenons $-\Delta F < dF < \Delta F$ avec

$$\Delta F := \left|\left(\frac{\partial F}{\partial x}\right)_{(\tilde{x}, \tilde{y})}\right| \cdot \Delta x + \left|\left(\frac{\partial F}{\partial y}\right)_{(\tilde{x}, \tilde{y})}\right| \cdot \Delta y$$

Par conséquent il vient $F = \tilde{F} + \delta F$ avec $|\delta F| \simeq |dF| < \Delta F$. On écrit ce résultat sous la forme

$$F = \tilde{F} \pm \Delta F$$

La quantité ΔF est une estimation de l'incertitude sur F .

Exemple. $F = y \cdot \sqrt{x} / (x + 1)$ avec $y = 1 \pm 0,1$ et $x = 4 \pm 0,2$ soit $\tilde{F} = \frac{2}{5}$;

$$\left(\frac{\partial F}{\partial x}\right) = \frac{1-x}{2(1+x)^2 \sqrt{x}} y, \quad \left(\frac{\partial F}{\partial x}\right)_{(\tilde{x}, \tilde{y})} = -\frac{3}{100}; \quad \left(\frac{\partial F}{\partial y}\right) = \sqrt{x} / (x+1), \quad \left(\frac{\partial F}{\partial y}\right)_{(\tilde{x}, \tilde{y})} = \frac{2}{5}.$$

On en déduit $\Delta F = \frac{3}{100} \cdot 0,2 + \frac{2}{5} \cdot 0,1 = 0,046 \simeq 0,05$.

$$F \simeq \frac{2}{5} \pm 0,05 = 0,4 \pm 0,05$$

On peut aussi opérer de la façon suivante :

$$\delta F/F \simeq \delta y/y + \frac{1}{2}\delta x/x - \delta(x+1)/(x+1).$$

Au voisinage de $x = \tilde{x} = 4$ et $y = \tilde{y} = 1$ il vient

$$\delta F/F \simeq \delta y + \frac{1}{8}\delta x - \frac{1}{5}\delta x = \delta y - \frac{3}{40}\delta x.$$

Avec $F \simeq \tilde{F} = \frac{2}{5}$ on obtient la petite variation de F due aux variations de x et y : $\delta F = \tilde{F} \cdot \left(\delta y - \frac{3}{40}\delta x \right) = \frac{2}{5}\delta y - \frac{3}{100}\delta x$.

Les relations $|\delta x| < 0,2$ et $|\delta y| < 0,1$ permettent d'obtenir le minimum de δF (pour $\delta y = -0,1$ et $\delta x = 0,2$) et le maximum de δF (pour $\delta y = 0,1$ et $\delta x = -0,2$). On obtient alors $-0,046 < \delta F < 0,046$ ce qui est le résultat précédent.

On remarquera que les incertitudes sur \sqrt{x} et sur $1/(x+1)$ ne s'additionnent pas. Ce sont les petites variations possibles qui s'additionnent.

5.5 Différentielles et dérivées de vecteurs

L'un des domaines d'application des vecteurs en physique est l'étude des mouvements.

Afin d'étudier le mouvement d'un mobile, il faut décider ce qu'est un **référentiel fixe** : nous devons choisir une origine, O , et un trièdre orthonormé $(\vec{u}_x, \vec{u}_y, \vec{u}_z)$ que nous déclarons être fixes. Les mouvements étudiés seront alors les **mouvements par rapport à ce référentiel**.

Comment choisir concrètement le référentiel fixe ? Les critères sont soit de nature mathématique, dans ce cas c'est une question de définition et de vocabulaire, soit de nature physique et alors la recherche d'un référentiel fixe s'identifie à la recherche d'un référentiel privilégié dans lequel les lois de la physique seront "aussi simple que possible".

Longtemps on a considéré que le référentiel lié à la Terre était le "bon" référentiel. C'était acceptable pour décrire la physique terrestre mais pas pour comprendre les mouvements des planètes : le repère héliocentrique s'imposa. De toutes façons, avec l'amélioration des performances expérimentales, il aurait fallu abandonner le repère terrestre, en l'absence même de toutes considérations astronomiques.

5.5.1 Définitions.

Le référentiel fixe étant choisi, considérons un mobile ponctuel, M . Le vecteur \vec{OM} varie avec le temps (ses composantes (X, Y, Z) sont des fonctions du temps).

On peut, sans considération de temps, imaginer également des vecteurs dont les coordonnées sont fonction d'un ou plusieurs paramètres. Par exemple, le vecteur \vec{u}_φ introduit dans le premier chapitre est fonction de la variable φ , tandis que les vecteurs \vec{u}_r et \vec{u}_θ dépendent des deux variables θ et φ .

Considérons le vecteur $\vec{V} = X\vec{u}_x + Y\vec{u}_y + Z\vec{u}_z$ dont les composantes sont des fonctions de la variable t . La différentielle de \vec{V} est le vecteur $d\vec{V}$

$$d\vec{V} := dX\vec{u}_x + dY\vec{u}_y + dZ\vec{u}_z$$

$d\vec{V}$ représente la variation du vecteur \vec{V} lorsque ses composantes varient respectivement de dX, dY, dZ .

La dérivée de \vec{V} par rapport à t est, par définition,

$$\boxed{\frac{d\vec{V}}{dt} := \frac{dX}{dt} \vec{u}_x + \frac{dY}{dt} \vec{u}_y + \frac{dZ}{dt} \vec{u}_z}$$

5.5.2 Propriétés.

Les propriétés des différentielles et des dérivées sont résumées ci-dessous.

$$\boxed{\begin{aligned} d(\vec{V} + f\vec{W}) &= d\vec{V} + df\vec{W} + f d\vec{W} \implies \\ \frac{d}{dt}(\vec{V} + f\vec{W}) &= \frac{d\vec{V}}{dt} + \frac{df}{dt}\vec{W} + f \frac{d\vec{W}}{dt} \end{aligned}} \quad (5.2)$$

où $f = f(t)$ est une fonction de la variable t .

$$\boxed{\begin{aligned} d(\vec{V} \cdot \vec{W}) &= d\vec{V} \cdot \vec{W} + \vec{V} \cdot d\vec{W} \implies \\ \frac{d}{dt}(\vec{V} \cdot \vec{W}) &= \frac{d\vec{V}}{dt} \cdot \vec{W} + \vec{V} \cdot \frac{d\vec{W}}{dt} \end{aligned}} \quad (5.3)$$

$$\boxed{\begin{aligned} d(\vec{V} \wedge \vec{W}) &= d\vec{V} \wedge \vec{W} + \vec{V} \wedge d\vec{W} \implies \\ \frac{d}{dt}(\vec{V} \wedge \vec{W}) &= \frac{d\vec{V}}{dt} \wedge \vec{W} + \vec{V} \wedge \frac{d\vec{W}}{dt} \end{aligned}}$$

Exemple. Le vecteur $\vec{u}_\rho = \cos \varphi \vec{u}_x + \sin \varphi \vec{u}_y$ est un vecteur de norme constante dont la dérivée par rapport à φ est le vecteur

$$\frac{d\vec{u}_\rho}{d\varphi} = -\sin \varphi \vec{u}_x + \cos \varphi \vec{u}_y = \vec{u}_\varphi.$$

On peut vérifier que \vec{u}_φ est orthogonal à \vec{u}_ρ . C'est généralement le cas lorsqu'on dérive un vecteur de norme constante.

En effet, soit $\vec{u} = \vec{u}(t)$ tel que $\vec{u} \cdot \vec{u} = C$ et $\vec{n} := \frac{d\vec{u}}{dt}$. La relation $\vec{u} \cdot \vec{u} = C$ implique $\frac{d}{dt}(\vec{u} \cdot \vec{u}) = 0$. En utilisant la relation 5.3 il vient $0 = \frac{d}{dt} \vec{u} \cdot \vec{u} + \vec{u} \cdot \frac{d}{dt} \vec{u} = 2 \vec{u} \cdot \vec{n}$. Les vecteurs \vec{n} et \vec{u} sont donc orthogonaux..

5.6 Champs de vecteurs

Certaines grandeurs physiques sont représentées par des vecteurs, c'est le cas de la vitesse d'un mobile ponctuelle, de son accélération ou de la force qui agit sur un solide. Parfois, la connaissance du vecteur associé à une grandeur physique ne suffit pas à sa description.

Dans le cas d'une force, si nous connaissons le vecteur associé \vec{F} , c'est à dire sa direction, son sens et son intensité, nous ne pouvons pas prédire complètement son effet tant que nous ne connaissons pas son point d'application, A . On définit "**la ligne d'action de la force**" comme la droite, D , parallèle à \vec{F} qui passe par A . L'expérience montre que la force \vec{F} appliquée en un autre point B de sa ligne d'action produit le même effet. Une force est donc définie par l'ensemble $\{\vec{F}, D\}$ on dit que c'est un "**vecteur**

glissant" car on peut faire glisser la force le long de sa ligne d'action sans en modifier l'effet.

Considérons maintenant le cas d'une répartition donnée de charges électriques. Nous disposons d'une charge ponctuelle q que nous déplaçons à volonté. En chaque point, la charge q est soumise à une force \vec{F} , résultante des attractions et des répulsions de la répartition de charges donnée. La force \vec{F} dépend du point M où nous plaçons la charge q ; c'est une fonction des coordonnées, x, y, z de M . On dit que l'on est en présence d'un **"champ de force"**. Le vecteur qui représente la force n'a pas de signification physique indépendamment du point M où se situe la charge q sur laquelle s'exerce cette force. C'est donc l'ensemble $\{\vec{F}, M\}$ qui a un sens physique. La force est représenté ici par un **"vecteur lié"**. Un champ de force est donc décrit par un ensemble de vecteurs liés.

5.6.1 Circulation d'un vecteur.

Considérons une courbe C le long de laquelle est défini un vecteur \vec{F} en chaque point. Soit s l'abscisse curviligne d'un point courant, M et \vec{u} le vecteur tangent unitaire à la courbe en ce point. Le vecteur \vec{F} dépend du point M considéré : \vec{F} est donc une fonction de s . Il en est de même de \vec{u} .

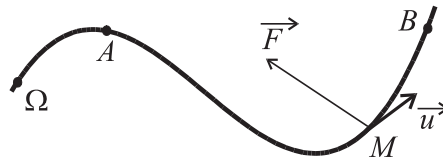


fig. 5.3

On pose $dW = \vec{F} \cdot \vec{u} ds$. **La circulation**, $W_{AB/C}$, du vecteur \vec{F} le long de la courbe C de A à B est l'intégrale de dW

$$W_{AB/C} = \int_{s_A}^{s_B} \vec{F} \cdot \vec{u} ds$$

Lorsque \vec{F} est une force dont le point d'application est M , la circulation est appelée **"travail"** de la force \vec{F} .

Exemple. Considérons le champ de vecteurs $\vec{F}(x, y, z) = ay \vec{u}_x$ où a est une constante positive. Nous considérons les deux courbes, (1) et (2), de la figure 5.4 qui relie l'origine O au point Q de coordonnées (ℓ, ℓ)

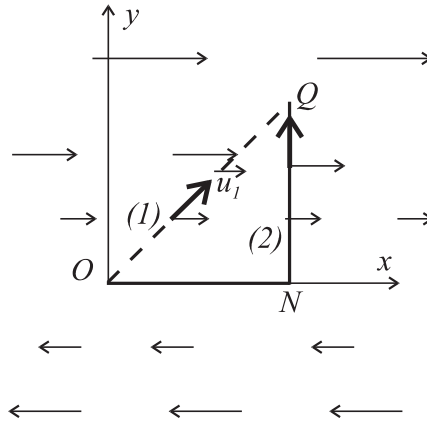


fig. 5.4

Sur la figure 5.4 nous avons représenté quelques uns des vecteurs du champ considéré. L'origine de chaque vecteur représenté est le point de coordonnées (x, y, z) où $\vec{F} = \vec{F}(x, y, z)$ a été calculé.

Le long de ON les vecteurs sont nuls (car $y = 0$). Le long de NQ le vecteur \vec{F} est orthogonal au vecteur \vec{u} en tout point. Par conséquent, $dW = \vec{F} \cdot \vec{u} ds = 0$ en tout point de la courbe ONQ et $W_{ONQ} = 0$.

Le long de OQ , le vecteur tangent unitaire est $\vec{u}_1 = \frac{\sqrt{2}}{2} (\vec{u}_x + \vec{u}_y)$ en chaque point.

Nous choisissons l'origine des abscisses curvilignes en O . Il vient $y = s \frac{\sqrt{2}}{2}$ en chaque point d'ordonnée y de la courbe OQ . On en déduit $ds = \sqrt{2} dy$.

En ce même point $\vec{F} \cdot \vec{u}_1 = ay \frac{\sqrt{2}}{2}$. On en déduit $dW = \vec{F} \cdot \vec{u}_1 ds = ay dy$. Ce qui implique $W_{OQ} = \int_{y_O}^{y_Q} ay dy$ avec $y_O = 0$ et $y_Q = \ell$. En intégrant on trouve $W_{OQ} = [\frac{1}{2} ay^2]_0^\ell = \frac{1}{2} a \ell^2 \neq W_{ONQ}$

On en déduit que **la circulation d'un champ de vecteurs entre deux points dépend le plus souvent du chemin emprunté.**

Remarquons que pour calculer la circulation d'un vecteur le long de OQ , il n'est pas nécessaire que ce vecteur appartienne à un champ de vecteurs mais seulement qu'il soit défini en tout point de OQ .

Nous allons maintenant étudier un cas important où la circulation d'un champ de vecteurs ne dépend que des points A et B et non du chemin entre A et B .

5.6.2 Gradient d'une fonction.

Considérons une fonction des coordonnées $V(x, y, z)$. Nous définissons le champ de vecteurs "gradient de V " :

$$\boxed{\vec{\text{grad}} [V] := \frac{\partial V}{\partial x} \vec{u}_x + \frac{\partial V}{\partial y} \vec{u}_y + \frac{\partial V}{\partial z} \vec{u}_z}$$

Exemple. $V(x, y, z) = 1/r$ où $r = \sqrt{x^2 + y^2 + z^2}$ (coordonnée radiale de la repré-

sentation sphérique). Le calcul donne $\frac{\partial r}{\partial x} = \frac{x}{r}$, $\frac{\partial r}{\partial y} = \frac{y}{r}$, $\frac{\partial r}{\partial z} = \frac{z}{r}$. On en déduit

$$\begin{aligned}\overrightarrow{\text{grad}} [r] &= \frac{x}{r} \vec{u}_x + \frac{y}{r} \vec{u}_y + \frac{z}{r} \vec{u}_z := \vec{u}_r \\ \overrightarrow{\text{grad}} \left[\frac{1}{r} \right] &= -\frac{1}{r^2} \left(\frac{\partial r}{\partial x} \vec{u}_x + \frac{\partial r}{\partial y} \vec{u}_y + \frac{\partial r}{\partial z} \vec{u}_z \right) = -\frac{1}{r^2} \vec{u}_r\end{aligned}$$

où \vec{u}_r est le vecteur radial de la représentation sphérique.

Cherchons la différentielle de V . On vérifie directement la relation

$$\boxed{\boxed{dV := \frac{\partial V}{\partial x} dx + \frac{\partial V}{\partial y} dy + \frac{\partial V}{\partial z} dz = \overrightarrow{\text{grad}} [V] \cdot d\overrightarrow{OM}}} \quad (5.4)$$

où $d\overrightarrow{OM} = dx \vec{u}_x + dy \vec{u}_y + dz \vec{u}_z$.

La circulation, W , du vecteur $\overrightarrow{\text{grad}} [V]$ le long d'une courbe, entre les points A et B , s'exprime sous la forme $W := \int_{s_A}^{s_B} \overrightarrow{\text{grad}} [V] \cdot \vec{u} ds$

La fonction V et son gradient dépendent de la position, sur la courbe, du point M où ils sont calculés : ce sont des fonctions de l'abscisse curviligne s du point M . Le vecteur \vec{u} est le vecteur unitaire tangent à la courbe : $\vec{u} = \frac{d\overrightarrow{OM}}{ds}$. C'est également une fonction de s .

Le vecteur $\vec{u} ds$ est la différentielle de \overrightarrow{OM} . En effet $\overrightarrow{OM} = \overrightarrow{OM}(s) \Rightarrow d\overrightarrow{OM} = \frac{d\overrightarrow{OM}}{ds} ds := \vec{u} ds$.

En utilisant la relation 5.4 il vient $\overrightarrow{\text{grad}} [V] \cdot \vec{u} ds = \overrightarrow{\text{grad}} [V] \cdot d\overrightarrow{OM} = dV$. La relation 3.1 implique $W = \int_{s_A}^{s_B} dV = V(s_B) - V(s_A)$.

$V(s)$ est la valeur de V au point d'abscisse curviligne s par conséquent $V(s_B) - V(s_A) = V(B) - V(A)$ ne dépend que de A et B et non du chemin intermédiaire :

$$\boxed{\boxed{\int_A^B \overrightarrow{\text{grad}} [V] \cdot \vec{u} ds = V(B) - V(A)}}$$

Exemple. Soit le champ de vecteurs $\vec{F} = ay \vec{u}_x + ax \vec{u}_y := F_x \vec{u}_x + F_y \vec{u}_y$.

Si \vec{F} est le gradient d'une fonction V , le lemme de Schwartz implique :

$F_x = \partial V / \partial x$ et $F_y = \partial V / \partial y \Rightarrow \partial F_x / \partial y = \frac{\partial^2 V}{\partial y \partial x} = \frac{\partial^2 V}{\partial x \partial y} = \partial F_y / \partial x$. Cette condition est nécessaire ; il s'avère qu'elle est également suffisante.

Ici $\partial F_x / \partial y = a = \partial F_y / \partial x \Rightarrow$ le champ de vecteur est le gradient d'une fonction, $V(x, y, z)$. On dit que \vec{F} **dérive du potentiel** V .

Pour calculer, par exemple, la circulation entre les points $O(0, 0)$ et $Q(\ell, \ell)$ on peut choisir un chemin particulier et calculer la circulation W le long de ce chemin car le résultat ne dépend pas du chemin choisi.

On peut aussi chercher le potentiel V .

Dans ce but on considère que y est une constante (arbitraire mais donnée). Le potentiel V est donc une fonction de x : $V(x, y)_{y=cte} = \Phi(x)$. Par conséquent $\frac{d\Phi}{dx} = \frac{\partial V}{\partial x} = ay$. En intégrant il vient $V = \Phi = ayx + C$ (car, ici, y est une constante). La constante d'intégration C est constante dans les conditions considérées, c'est à dire dans

des conditions telles que seul x varie. C est donc éventuellement une fonction de y (mais non de x). On en déduit $V(x, y) = ayx + C(y)$. On considère maintenant que y seul varie et que x est maintenu constant. La relation $\frac{\partial V}{\partial y} = ax$ s'écrit $ax + \frac{dC}{dy} = ax$ soit $\frac{dC}{dy} = 0$. La quantité C , susceptible d'être une fonction de y seulement, est en fait une constante car sa dérivée (par rapport à y est nulle). On trouve donc $V = axy + cte$.

La circulation de \vec{F} entre O et Q vaut $V(Q) - V(O) = ax_Q y_Q - ax_O y_O = al^2$.

Remarque importante. Si une fonction V est constant, cela signifie que V n'est pas une fonction de x , c'est à dire $\partial V/\partial x = 0$ et de même $\partial V/\partial y = 0 = \partial V/\partial z$ d'où

$$\boxed{\boxed{\overrightarrow{\text{grad}} [V] = \vec{0} \iff V = cte}}$$

N.B. On trouve souvent l'écriture $\vec{\nabla} f$ pour $\overrightarrow{\text{grad}} [f]$ que nous avons utilisé ici.

Chapitre 6

PROBABILITÉS ET STATISTIQUES

6.1 Notion d'expérience, d'événement et conséquences

6.1.1 Qu'est-ce qu'une expérience ?

Toute expérience scientifique doit être menée de façon rigoureuse en transmettant toutes les informations nécessaires pour que toute personne s'intéressant au sujet de l'expérience puisse en vérifier les résultats.

Cette transmission des informations est nommée "**protocole**" de l'expérience, c'est-à-dire la description de l'ensemble des matériels et des méthodes employés pour mener à bien l'expérience.

Lorsqu'elle ne dépend pas du temps ou n'est pas perturbée par la mesure, une expérience est généralement reproductible (par exemple, une mesure de distances, de tailles, de poids, ...). Mais il existe aussi des expériences non reproductibles (par exemple, des mesures dépendantes du temps, des caractéristiques météorologiques, transmission d'un gène par hérédité, ...).

Théoriquement, il est possible de définir deux sortes d'expériences :

- celles qui sont dites déterministes, qui suivent une loi bien précise (loi de physique par exemple) et dont les résultats sont souvent prévisibles.
- les expériences aléatoires, dont les résultats sont par définition imprévisibles. La plus connue est le résultat d'un jet de dé ou d'un jeu de "pile ou face", mais cela peut être aussi la réaction d'un individu au traitement de sa maladie ou le nombre de fruits produits par un arbre, etc.

Cette dichotomie est purement théorique. Une expérience déterministe n'existe qu'en pensée : pour mesurer une distance, l'instrument n'est jamais parfait, il peut donner des résultats légèrement différents suivant la température, le taux d'humidité ; la manière dont l'expérimentateur l'utilise peut induire aussi des résultats différents. L'ensemble des résultats se répartissent "aléatoirement" autour de la loi que l'expérience illustre. Interviennent alors les notions d'incertitudes et d'erreurs que nous avons déjà décrites lors du chapitre sur les fonctions. Mais, ici, nous allons un peu plus loin en élaborant des moyens de décrire comment les mesures se répartissent au sein de ces incertitudes et comment estimer à partir des mesures obtenues la valeur recherchée par l'expérience.

Inversement, nous avons des moyens mathématiques de décrire les résultats d'une expérience aléatoire. Cela est encore plus vrai lorsque l'expérience est reproduite un grand nombre de fois.

En résumé, toute mesure issue d'une expérience comporte une part plus ou moins grande d'aléatoire. L'objet de ce chapitre est d'apprendre des outils mathématiques qui permettent d'utiliser ces mesures pour en déduire les résultats recherchés.

Mais avant tout, nous établissons quelques définitions.

- On appelle une "**épreuve**", le protocole d'une expérience dont le résultat est aléatoire.
- Chaque résultat d'une épreuve définit un "**événement élémentaire**".
- L'ensemble des événements élémentaires constitue l'"**univers**" que nous écrivons $[E]$.

L'univers est ainsi l'ensemble de toutes les mesures possibles.

Exemple : pour un jet de dé : $[E] = \{1,2,3,4,5,6\}$

- On appelle "**événement**" un sous-ensemble de $[E]$. Il est composé d'un ou plusieurs événements élémentaires. On note $[\Omega]$ l'ensemble des événements.

Exemple : On considère l'événement A pour un jet de dé consistant à obtenir un chiffre pair.

$A = \{2,4,6\}$ A est composé de trois événements élémentaires.

L'univers des événements devient, dans ce cadre :

$[\Omega] = \{\text{chiffres pairs, chiffres impairs}\}$

6.1.2 Opérations sur les événements

Ce paragraphe nécessite, dans un premier temps, de développer une liste de définitions amenant un certain nombre de symboles de logique. Il convient de s'habituer à cette écriture qui permet d'alléger les notations.

Soient deux événements A et B correspondant à une même épreuve. Ils peuvent être élémentaires ou non.

- L'événement **{A ou B}** est réalisé si ou A, ou B ou les deux simultanément sont réalisés. On le note $A \cup B$.
- L'événement **{A et B}** est réalisé si A et B sont réalisés à la fois. On le note $A \cap B$. On remarquera que l'événement $A \cap B$ est inclus dans l'événement $A \cup B$ et donc que si $A \cap B$ est réalisé alors $A \cup B$ est réalisé (l'inverse n'étant pas systématiquement vrai).
- L'événement "**impossible**" \emptyset est celui qui ne peut pas être réalisé à la suite de l'épreuve.
- Un événement "**certain**" est celui qui est réalisé à chaque épreuve. Il contient tous les événements élémentaires et est donc identique à l'univers $[E]$.
- A et B sont des événements "**incompatibles**" lorsque, en effectuant l'épreuve, il est impossible de réaliser à la fois A et B. Il n'y a pas d'élément élémentaire commun à A et à B. A et B sont disjoints : $A \cap B = \emptyset$
- L'événement "**contraire**" de A est l'événement réalisé lorsque A ne l'est pas. La liste d'événements élémentaires qui le caractérise est composé des événement de $[E]$ n'appartenant pas à la liste de A. On l'appelle aussi parfois non-A, noté $\neg A$ ou \bar{A}

6.2 Probabilités

6.2.1 Définition

La notion de probabilité est liée à la notion de fréquence. Le plus simple pour les introduire est sans doute de le faire à partir d'un exemple de répétition d'une expérience aléatoire.

Exemple. La mesure d'une longueur est répétée de nombreuses fois. On a obtenu la liste des 18 résultats fiables (appelés $\{g_k\}$) suivante $\{10, 4 \text{ m}, 10, 7 \text{ m}, 10, 9 \text{ m}, 10, 9 \text{ m}, 11 \text{ m}, 11 \text{ m}, 11 \text{ m}, 11, 1 \text{ m}, 11, 1 \text{ m}, 11, 1 \text{ m}, 11, 1 \text{ m}, 11, 2 \text{ m}, 11, 2 \text{ m}, 11, 3 \text{ m}, 11, 3 \text{ m}, 11, 5 \text{ m}, 11, 7 \text{ m}, 12 \text{ m}\}$. Les résultats sont représentés ci-dessous sur l'axe des G où un point représente un résultat.

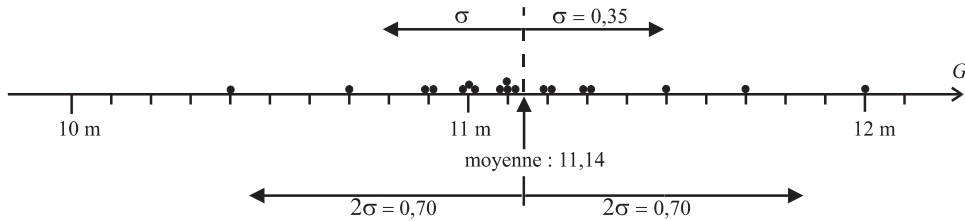


fig. 6.1 : Résultats de la mesure d'une longueur répétée 18 fois.

La valeur $g = 11\text{m}$ y apparaît trois fois, on dit que sa **"fréquence"** est $3/18 \approx 0,17$ (on l'appelle parfois **"fréquence relative"** par opposition à la **"fréquence absolue"** qui est le nombre d'observation soit ici 3)

Nous devons distinguer clairement la liste des $\{g_k\}$ où apparaissent des répétitions et la liste des valeurs différentes prises par les g_k . Le tableau ci-dessous donne les valeurs explicites des dix-huit g_k

10,4 m	10,7 m	10,9 m	11 m	11,1 m	11,2 m	11,3 m	11,5 m	11,7 m	12 m
g_1	g_2	g_3 et g_4	g_5 à g_7	g_8 à g_{11}	g_{12} et g_{13}	g_{14} et g_{15}	g_{16}	g_{17}	g_{18}

Maintenant augmentons le nombre de mesures. Nous obtenons le tableau suivant

Nombre n de mesures	Nombre k de fois que $g = 11\text{m}$ a été obtenu	Fréquence $f = \frac{k}{n}$
18	3	$3/18 \approx 0,17$
100	17	$17/100 = 0,17$
1000	165	$165/1000 = 0,165$
10000	1662	$1662/10000 = 0,1662$

Nous pouvons remarquer que lorsque le nombre de mesures augmente, la fréquence se stabilise et tend vers la valeur limite de $1/6$. On appelle **"probabilité objective"**, issue de l'expérience, cette valeur limite de la fréquence. C'est le résultat de la loi des grands nombres pour les expériences répétées de façon identique.

$$p = \lim_{n \rightarrow \infty} \left(f = \frac{k}{n} \right)$$

Il existe une définition mathématique axiomatique (c'est-à-dire non démontrable) de la probabilité :

- 1- La **"probabilité"** $P(A)$ de tout événement A associé à une épreuve est un nombre compris entre 0 et 1.
 $0 \leq P(A) \leq 1$
- 2- Si deux événements A et B sont incompatibles, la probabilité de l'événement $(A \cup B)$ est égale à la somme des probabilités de A et de B .
 $A \cap B = \emptyset$ alors $P(A \cup B) = P(A) + P(B)$
- 3- La probabilité d'un événement certain est égale à 1.
 $P(\Omega) = 1$

On en déduit que la probabilité d'un événement impossible est égale à 0.

$$P(\emptyset) = 0$$

Démonstration : $P(\emptyset \cup \Omega) = P(\Omega) = P(\Omega) + P(\emptyset) \implies P(\emptyset) = 0$

Par ailleurs, l'axiome 2 peut se généraliser au cas de plusieurs événements incompatibles :

Soient A_1, A_2, \dots, A_n événements incompatibles 2 à 2

$$P(A_1 \cup A_2 \cup \dots \cup A_n) = P(A_1) + P(A_2) + \dots + P(A_n)$$

Dans le cas où une épreuve mène à un nombre fini n d'événements élémentaires pouvant être considérés comme symétriques sur le plan de leur réalisation, la probabilité de chacun d'eux est la même et vaut $\frac{1}{n}$. On dit alors que les événements sont "**équiprobables**".

Exemple : Les lancers de dé non-pipé sont équiprobables et la probabilité d'obtenir une valeur particulière est $\frac{1}{6}$.

Remarque : **Tout calcul conduisant à des valeurs de probabilités négatives ou supérieures à 1 est faux.**

6.2.2 Probabilités conditionnelles et combinées

– La probabilité du contraire d'un événement est égale à 1 moins la probabilité de cet événement

$$P(\bar{A}) = 1 - P(A)$$

Démonstration : Un événement et son contraire sont incompatibles : $A \cap \bar{A} = \emptyset$

$$P(A) + P(\bar{A}) = P(A \cup \bar{A}) = P(\Omega) = 1$$

– Si un événement A est inclus dans un événement B (ce qui se note $A \subset B$), alors la probabilité de A est inférieure à celle de B .

Démonstration : Si A est inclus dans B , B peut se décomposer en deux événements incompatibles : A et $(\bar{A} \cap B)$. On a donc

$$\begin{aligned} P(B) &= P(\bar{A} \cap B) + P(A) \\ P(\bar{A} \cap B) &\geq 0 \implies P(B) \geq P(A) \end{aligned}$$

– La probabilité de l'événement $A \cup B$ est égale à la somme des probabilités de A et de B , moins celle de $(A \cap B)$.

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Démonstration : A peut se décomposer en deux événements incompatibles : $(A \cap B)$ et $(A \cap \bar{B})$

$$P(A) = P(A \cap B) + P(A \cap \bar{B})$$

B peut se décomposer en deux événements incompatibles : $(A \cap B)$ et $(\bar{A} \cap B)$

$$\begin{aligned} P(B) &= P(A \cap B) + P(\bar{A} \cap B) \\ \implies P(A) + P(B) &= 2 P(A \cap B) + P(\bar{A} \cap B) + P(A \cap \bar{B}) \end{aligned}$$

Or, $A \cup B$ peut se décomposer en trois événements incompatibles : $(A \cap \bar{B})$, $(A \cap B)$ et $(\bar{A} \cap B)$

$$\begin{aligned} P(A \cup B) &= P(A \cap B) + P(\bar{A} \cap B) + P(A \cap \bar{B}) \\ \implies P(A \cup B) &= P(A) + P(B) - P(A \cap B) \end{aligned}$$

Exemple : Pour le tirage d'une carte dans un jeu de 32 cartes, nous recherchons la probabilité d'obtenir un trèfle ou un roi.

On suppose l'équiprobabilité des cartes.

$$P(\text{trèfle}) = \frac{8}{32} = \frac{1}{4} \quad P(\text{roi}) = \frac{4}{32} = \frac{1}{8} \quad P(\text{roi} \cap \text{trèfle}) = P(\text{roi de trèfle}) = \frac{1}{32}$$

Les événements ne sont donc pas incompatibles. Nous avons donc

$$P(\text{roi ou trèfle}) = P(\text{roi} \cup \text{trèfle}) = P(\text{roi}) + P(\text{trèfle}) - P(\text{roi} \cap \text{trèfle})$$

Il faut soustraire $P(\text{roi} \cap \text{trèfle})$ car sinon le roi de trèfle est compté à la fois dans la probabilité de tirer un roi et dans celle de tirer un trèfle.

$$P(\text{roi ou trèfle}) = \frac{1}{4} + \frac{1}{8} - \frac{1}{32} = \frac{8 + 4 - 1}{32} = \frac{11}{32}$$

– Considérons deux événements A et B, au cours d'une expérience aléatoire. La probabilité de réaliser A sachant que B est réalisé s'appelle "**probabilité conditionnelle**" de A sachant B et s'écrit $P(A/B)$. Elle est parfois appelée probabilité "**a posteriori**".

On remarquera que : $P(\bar{A}/B) = 1 - P(A/B)$

– Il peut arriver que l'information apportée par la réalisation ou la non-réalisation de B ne modifie pas la probabilité de réalisation de A soit $P(A)$

$$P(A/B) = P(A/\bar{B}) = P(A)$$

On dit que A et B sont "**indépendants**", alors

$$P(A \cap B) = P(A) P(B)$$

Attention : Il ne faut pas confondre indépendance et incompatibilité. Si $P(A \cap B) \neq 0$, les deux événements sont compatibles, mais ils peuvent être dépendants ou indépendants.

Exemple : Soient les deux événements :

A = aimer lire

B = aimer le chocolat

A et B sont compatibles (on peut à la fois aimer lire et le chocolat), mais A et B sont indépendants (le fait d'aimer lire n'apporte aucune information sur le fait d'aimer le chocolat et réciproquement).

– De façon plus générale, si l'on considère deux événements A et B quelconques, il est possible de calculer la probabilité d'avoir A et B

$$P(A \cap B) = P(B \cap A) = P(A) P(B/A) = P(B) P(A/B)$$

Exemple : Sur une population de 100 enfants, une enquête aboutit aux résultats suivants :

- 1- 87 enfants aiment le chocolat.
- 2- 63 enfants aiment lire.
- 3- Parmi les 63 enfants qui aiment lire, 55 aiment le chocolat.

Prenons un enfant au hasard dans cette population. Nous voulons déterminer les probabilités pour que cet enfant corresponde aux trois événements suivants :

- A : l'enfant aime le chocolat
- B : l'enfant aime lire
- $A \cap B$: l'enfant aime lire et manger du chocolat

D'après les résultats de l'enquête, nous déduisons :

$$P(A) = 87/100 \quad P(B) = 63/100 \quad P(A \cap B) = 55/100$$

Si l'on sait au départ que l'enfant aime lire, alors l'ensemble de référence est restreint à 63 enfants. La probabilité que l'enfant aime le chocolat sachant qu'il aime lire est donc

$$P(A/B) = 55/63$$

Or, en utilisant le théorème précédent, nous avons

$$P(A/B) = P(A \cap B) / P(B) = (55/100) / (63/100) = 55/63$$

Nous retrouvons bien le résultat que nous avons obtenu directement.

De la même façon, la probabilité que l'enfant aime lire s'il aime le chocolat est

$$P(B/A) = 55/87 = (55/100) / (87/100) = P(A \cap B) / P(A)$$

6.2.3 Rappels d'analyse combinatoire

Les notions présentées ici sont des révisions du lycée. Elles sont donc présentées rapidement comme aide-mémoire pour faciliter la résolution des exercices. Cette partie permet aussi d'évaluer la nécessité (ou non) de revoir ces notions dans des documents issus du secondaire.

- *Notion de factorielle* : Si $n \in \mathbb{N}$

$$n! = n (n-1) (n-2) \dots 1$$

Par convention : $0! = 1$

- *Permutations de n objets discernables* : toute disposition ordonnée des n ($\in \mathbb{N}$) objets.

P_n nombre de permutations d'un ensemble de n éléments discernables : $P_n = n!$

- *Arrangements de p éléments parmi n ($p \leq n$)* : toute disposition ordonnée sans répétition (sans remise, exhaustif) de p ($\in \mathbb{N}$) objets parmi n ($\in \mathbb{N}$) objets discernables.

A_n^p nombre d'arrangements de p éléments pris parmi n : $A_n^p = \frac{n!}{(n-p)!}$

- *Combinaisons de p éléments parmi n ($p \leq n$)* : toute disposition non-ordonnée de p ($\in \mathbb{N}$) objets pris parmi n ($\in \mathbb{N}$) objets discernables, chaque objet ne pouvant intervenir qu'une fois.

C_n^p nombre de combinaisons de p éléments pris parmi n : $C_n^p = \frac{n!}{p!(n-p)!}$

Propriétés des C_n^p :

- Si $n \geq 1$ $C_n^0 = 1$ $C_n^n = 1$ $C_n^1 = n$

- $C_n^p = C_n^{n-p}$

- Si $1 \leq p \leq n$ $C_n^p = C_{n-1}^p + C_{n-1}^{p-1}$

- Si $a \in \mathbb{R}, b \in \mathbb{R}$ $(a+b)^n = \sum_{p=0}^n C_n^p a^{n-p} b^p$
 $= C_n^0 a^n + C_n^1 a^{n-1} b + \dots + C_n^{n-1} a b^{n-1} + C_n^n b^n$

6.3 Variables aléatoires discrètes

6.3.1 Définitions

Une variable "**aléatoire**" est une épreuve menant à des événements élémentaires qui sont des nombres.

Par convention, nous allons désigner la variable aléatoire par une majuscule et le résultat de la mesure de cette variable aléatoire par une minuscule.

Ainsi, une variable aléatoire X est définie dans son univers Ω , en associant à chaque résultat de l'expérience aléatoire un nombre réel x caractéristique de ce résultat. L'événement $X = x$ prend en compte tous les résultats pour lesquels la variable X (en majuscule) prend la valeur x (en minuscule).

Une variable aléatoire X est dite "**discrète**", si l'ensemble des réalisations possibles x_1, x_2, \dots, x_n est fini ou dénombrable.

Dans le cas d'une variable discrète, la probabilité, associée à la réalisation de l'événement $X = x$, s'écrit $P(X = x)$

Remarque : Soit g une fonction mathématique quelconque. Si l'on pose $Y = g(X)$, on définit une nouvelle variable aléatoire Y fonction de X .

La "**loi de probabilité**" d'une variable aléatoire X discrète est définie en donnant l'ensemble des valeurs $p_1, p_2, \dots, p_n, \dots$ correspondant aux probabilités des différentes éventualités $x_1, x_2, \dots, x_n, \dots$

On a donc $p_i = P(X = x_i)$ avec $0 \leq p_i \leq 1$ et $\sum_i p_i = 1$

Exemple : Dans le cas d'un jet de dé, on a : $p_1 = p_2 = p_3 = p_4 = p_5 = p_6 = \frac{1}{6}$

La "**fonction de répartition**" associée à la variable aléatoire X est la fonction, notée F ou F_X , définie dans \mathbb{R} dans l'intervalle $[0,1]$ par

$$F(x) = F_X(x) = P(X \leq x) = \sum_{x_i \leq x} P(X=x_i)$$

On peut montrer aisément que :

$$0 \leq F(x) \leq 1$$

$F(x)$ est croissante : si $a \leq b$ alors $F(a) \leq F(b)$

$$\lim_{x \rightarrow -\infty} F(x) = 0 \quad \text{et} \quad \lim_{x \rightarrow \infty} F(x) = 1$$

6.3.2 Paramètres caractéristiques

Soient une variable aléatoire X et sa loi de probabilité $P(X = x)$

- La "**moyenne**" \bar{X} de la variable aléatoire X est la moyenne des résultats qu'on obtiendrait en répétant indéfiniment l'épreuve associée à X .

$$\bar{X} = x_1 \times p_1 + x_2 \times p_2 + \dots + x_n \times p_n = \sum_{i=1}^n x_i \times p_i$$

Exemple : La moyenne des tirs d'un dé est :

$$1 \times \frac{1}{6} + 2 \times \frac{1}{6} + 3 \times \frac{1}{6} + 4 \times \frac{1}{6} + 5 \times \frac{1}{6} + 6 \times \frac{1}{6} = 3,5$$

- L'"**espérance mathématique**" d'une variable aléatoire X est notée $E(X)$. C'est une question de vocabulaire. Cette expression est utilisée en calcul des statistiques et probabilités comme un synonyme de moyenne. Pour être plus précis, l'espérance mathématique désigne la moyenne issue de la loi de probabilité alors que, quand on parle de la moyenne, on peut tout aussi bien désigner la moyenne "théorique" issue de la loi de probabilité que la moyenne "expérimentale" des résultats obtenus pendant l'expérience.

$$\bar{X} = E(X)$$

L'espérance mathématique, et donc la moyenne, présente quelques propriétés.

Soient X et Y deux variables aléatoires

$$E(aX+b) = a E(X) + b \quad \text{avec } a \in \mathbb{R} \text{ et } b \in \mathbb{R}$$

$$E(X+Y) = E(X) + E(Y)$$

Si, de plus, X et Y sont des variables aléatoires indépendantes

$$E(X \times Y) = E(X) \times E(Y)$$

- La "**variance**" $\sigma^2(X)$ de la variable aléatoire X est la moyenne des carrés des écarts entre X et sa valeur moyenne $E(X)$.

$$\sigma^2 = E[(X-E(X))^2] = \sum_i p_i (x_i - \bar{X})^2 = E(X^2) - E(X)^2 = E(X^2) - \bar{X}^2$$

L'"**écart-type**" $\sigma(X)$ est la racine carrée de la variance.

Attention : La moyenne de X^2 n'est pas le carré de la moyenne de X .

Démonstration : $E(X^2) = \sum_i p_i x_i^2 \neq \left(\sum_i p_i x_i \right)^2 = E(X)^2$

Propriétés de la variance :

$$\sigma^2(X) \geq 0$$

$$\sigma^2(aX) = a^2 \sigma^2(X) \quad \text{avec } a \in \mathbb{R}$$

$$\sigma^2(X + a) = \sigma^2(X) \quad \text{avec } a \in \mathbb{R}$$

Si, de plus, X et Y sont des variables aléatoires indépendantes

$$\sigma^2(X + Y) = \sigma^2(X - Y) = \sigma^2(X) + \sigma^2(Y)$$

- L'espérance d'une variable aléatoire est un indicateur de position sur la distribution de probabilité de X . Les résultats de X tombent autour de cette espérance. L'écart-type permet de caractériser comment les valeurs de X sont dispersées autour de la valeur moyenne. Si cet indicateur est nul, cela veut dire que le résultat de X est certain.

- La "**médiane**" est la valeur de la variable aléatoire X qui, pour une série de résultats étudiée, permet de séparer en deux groupes égaux en nombre les résultats. 50 % des résultats ont une valeur inférieure ou égale à la médiane et 50 % des résultats ont une valeur supérieure ou égale à la médiane.
- Le "**mode**" ou la "**valeur modale**" de la variable aléatoire X , noté Mo , est la (les) valeur(s) de la réalisation de X qui a (ont) la plus grande probabilité de réalisation

$$P(X=Mo) \geq P(X=x_i) \quad \forall x_i$$

Exemple : Reprenons l'exemple des mesures de longueur de la figure 6.1. Supposons que la répartition des mesures observées corresponde exactement à la loi de probabilité de ces mesures. La fréquence de chaque mesure correspond donc à sa probabilité. Reprenons le tableau des mesures en indiquant dessous chaque mesure sa probabilité.

10,4 m	10,7 m	10,9 m	11 m	11,1 m	11,2 m	11,3 m	11,5 m	11,7 m	12 m
$\frac{1}{18}$	$\frac{1}{18}$	$\frac{2}{18}$	$\frac{3}{18}$	$\frac{3}{18}$	$\frac{2}{18}$	$\frac{2}{18}$	$\frac{1}{18}$	$\frac{1}{18}$	$\frac{1}{18}$

Pour cette loi de probabilité, l'espérance est

$$E(X) = \frac{10,4}{18} + \frac{10,7}{18} + \frac{10,9}{9} + \frac{11}{6} + \frac{11,1}{6} + \frac{11,2}{9} + \frac{11,3}{9} + \frac{11,5}{18} + \frac{11,7}{18} + \frac{12}{18} = 11,14$$

L'écart-type vaut

$$\sigma(X) = \frac{10,4^2}{18} + \frac{10,7^2}{18} + \frac{10,9^2}{9} + \frac{11^2}{6} + \frac{11,1^2}{6} + \frac{11,2^2}{9} + \frac{11,3^2}{9} + \frac{11,5^2}{18} + \frac{11,7^2}{18} + \frac{12^2}{18} - 11,14^2 \approx 0,35$$

La médiane de ces résultats est : 11,1, car la moitié des résultats est inférieure ou égale à cette valeur et l'autre moitié supérieure ou égale.

Les modes de ces résultats sont : 11 et 11,1 car, pour ces deux valeurs, il y a trois occurrences ce qui est le plus grand nombre d'occurrences de ces résultats.

Remarque : Suivant la forme de la loi de probabilité, la moyenne, la médiane et le mode peuvent avoir parfois des valeurs très différentes.

- Une variable aléatoire "**centrée**" est une variable aléatoire dont l'espérance est nulle.

Soit X une variable aléatoire, la variable $X - E(X)$ est centrée.

Une variable aléatoire "**réduite**" est une variable aléatoire dont l'écart-type est égal à 1

Soit X une variable aléatoire, la variable aléatoire sans unité $X/\sigma(X)$ est réduite.

Soit X une variable aléatoire, la variable aléatoire $\frac{X - E(X)}{\sigma(X)}$ est centrée et réduite.

6.3.3 Principales lois des variables discrètes

Loi de Bernoulli

Une variable aléatoire B ayant deux valeurs possibles 0 et 1 qu'elle prend avec les probabilités respectives q et p (avec $q = 1 - p$) suit la loi de Bernoulli. On parle alors de "**variable de Bernoulli**" de paramètre p .

La moyenne de la loi de Bernoulli est p .

La variance de la loi de Bernoulli est $p(1 - p) = pq$

L'écart-type de la loi de Bernoulli est \sqrt{pq}

On parle parfois de loi de Bernoulli pour des cas où la variable aléatoire prend d'autres valeurs que 0 et 1 (la moyenne et la variance peuvent alors changer de valeur). La loi de Bernoulli sert dans tous les cas où l'on traite d'une simple alternative dont on sait suivant quel pourcentage elle se produit.

Loi Binomiale

Une variable aléatoire Bin suit une "**loi binomiale**" de paramètres p et n lorsqu'elle prend les valeurs 0, 1, ..., n avec les probabilités p_1, p_2, \dots, p_n suivant la formule

$$p_i = C_n^i p^i q^{n-i}$$

avec $q = 1 - p$

La moyenne de la loi binomiale est np .

La variance de la loi binomiale est npq .

L'écart-type de la loi binomiale est \sqrt{npq}

Le mode de la loi binomiale est : $np + p - 1 \leq Mo \leq np + p$

La démonstration de ces résultats est lourde quoique faisable, nous les admettrons donc.

La somme de n variables de Bernoulli indépendantes de même paramètre p est une variable aléatoire qui suit la loi de Bernoulli. De façon générale, la loi binomiale est utilisée dès qu'il s'agit d'un décompte de succès parmi n épreuves répétées, identiques et indépendantes à deux issues : échec ou succès.

Si X et Y sont deux variables aléatoires suivant des lois binomiales de paramètres respectivement (n_X, p) et (n_Y, p) , la somme de ces deux variables aléatoires suit la loi binomiale de paramètres $(n_X + n_Y, p)$

Loi de poisson

Lorsqu'une variable aléatoire X suit la "**loi de Poisson**", ses valeurs possibles sont 0, 1, 2, ..., k , ... La probabilité p_k d'obtenir la valeur k est donnée par la formule

$$P(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}$$

On appelle λ le paramètre de la loi de Poisson.

La moyenne de la loi de Poisson est λ .

La variance de la loi de Poisson est λ .

L'écart-type de la loi de Poisson est $\sqrt{\lambda}$

Le mode de la loi de Poisson est : $\lambda - 1 \leq Mo \leq \lambda$

La démonstration de ces résultats ainsi que la vérification du fait qu'il s'agit bien d'une loi de probabilité nécessite le développement en série entière de la fonction exponentielle et est donc hors-programme. Elle est présente toutefois dans la plupart des livres sur les probabilités.

La loi de Poisson a de très nombreuses applications, puisqu'elle permet de représenter la survenue d'événements rares dans le temps et/ou dans l'espace. On peut la faire intervenir, par exemple, dans les titrages des solutions bactériennes, pour comptabiliser des accidents ou pour la pharmacovigilance.

Si X et Y sont deux variables aléatoires indépendantes suivant des lois de Poisson de paramètre respectif λ et μ , la somme de ces variables est une variable aléatoire suivant une loi de Poisson de paramètre $\lambda + \mu$.

Remarque : Si l'on a une valeur élevée pour n , en gardant p petit et np fini, alors la loi binomiale de paramètres (n, p) peut être approximée par la loi de Poisson de paramètre $\lambda = np$.

Suivant la précision recherchée, les limites pour ces approximations restent à discuter et dépendent du contexte. On donne souvent des valeurs de l'ordre de 3 ou 5 pour la valeur supérieure de np et des valeurs de plusieurs dizaines (typiquement 30) pour la valeur inférieure de n .

6.4 Variables aléatoires continues

6.4.1 Définitions

Une variable aléatoire est dite "**continue**" si son domaine de variation est l'ensemble des réels \mathbb{R} ou un intervalle de cet ensemble.

Une variable aléatoire continue X peut être définie par sa "**fonction de répartition**" F ou F_X , définie comme pour les variables aléatoires discrètes

$$F(x) = F_X(x) = P(X \leq x)$$

La fonction de répartition F d'une variable aléatoire continue a les propriétés suivantes

$F(x)$ est continue

$F(x)$ est croissante : $x_1 \leq x_2 \implies F(x_1) \leq F(x_2)$

$0 \leq F(x) \leq 1$

$\lim_{x \rightarrow -\infty} F(x) = 0$ et $\lim_{x \rightarrow \infty} F(x) = 1$

La fonction f telle que $f(x) = \frac{dF(x)}{dx}$ où F est la fonction de répartition d'une variable aléatoire continue X est appelée fonction "**densité de probabilité**" de la variable aléatoire X . On parle parfois aussi de "**fonction de distribution**".

On définit aussi la densité de probabilité de la manière suivante

$$f(x) dx = P(x \leq X \leq x+dx)$$

Ces deux définitions sont équivalentes.

La densité de probabilité d'une variable aléatoire continue présente toujours les propriétés suivantes

$f(x) \geq 0$

f est une fonction intégrable et $\int_{-\infty}^{+\infty} f(x) dx = 1$

6.4.2 Paramètres caractéristiques

Les paramètres caractéristiques sont les mêmes que dans le cas des variables discrètes. Cependant, il est utile de préciser leur définition dans le cas continu.

- L'"**espérance**" d'une variable aléatoire continue X de densité de probabilité f est donnée par

$$E(X) = \int_{-\infty}^{+\infty} x f(x) dx$$

si cette intégrale existe.

L'espérance mathématique, et donc la moyenne, présente quelques propriétés.

Soient X et Y deux variables aléatoires

$$E(aX+b) = a E(X) + b \quad \text{avec } a \in \mathbb{R} \text{ et } b \in \mathbb{R}$$

$$E(X+Y) = E(X) + E(Y)$$

Si, de plus, X et Y sont des variables aléatoires indépendantes

$$E(X \times Y) = E(X) \times E(Y)$$

Une variable aléatoire continue "**centrée**" est une variable aléatoire dont l'espérance est nulle.

Soit X une variable aléatoire, la variable $X - E(X)$ est centrée.

- La "**variance**" d'une variable aléatoire continue X est définie par

$$\sigma^2(X) = \int_{-\infty}^{+\infty} [x - E(X)]^2 f(x) dx = E(X^2) - E(X)^2$$

L'"**écart-type**" $\sigma(X)$ est la racine carrée de la variance.

Propriétés de la variance :

$$\sigma^2(X) \geq 0$$

$$\sigma^2(aX) = a^2 \sigma^2(X) \quad \text{avec } a \in \mathbb{R}$$

$$\sigma^2(X + a) = \sigma^2(X) \quad \text{avec } a \in \mathbb{R}$$

Si, de plus, X et Y sont des variables aléatoires indépendantes

$$\sigma^2(X + Y) = \sigma^2(X - Y) = \sigma^2(X) + \sigma^2(Y)$$

Une variable aléatoire "**réduite**" est une variable aléatoire

dont l'écart-type est égal à 1

Soit X une variable aléatoire, la variable aléatoire sans unité $X/\sigma(X)$ est réduite.

Soit X une variable aléatoire continue, la variable aléatoire $\frac{X - E(X)}{\sigma(X)}$ est centrée et réduite

- La "**médiane**" d'une variable aléatoire continue est le nombre réel Me tel que

$$F(Me) = 0,5$$

La médiane correspond à la valeur de X qui divise la distribution en deux parts équiprobables

$$P(X \leq Me) = P(X \geq Me) = 0,5$$

- Le "**mode**" Mo d'une variable aléatoire continue est défini comme la (ou les) valeur(s) qui correspond(ent) à un maximum local de la fonction densité de probabilité.

Une fonction de densité de probabilité ne présentant qu'un seul maximum est dite "**unimodale**".

6.4.3 Principales lois de probabilité

Loi uniforme

La distribution uniforme est la plus simple loi de variable aléatoire continue. Elle correspond à une densité de probabilité constante pour tout l'ensemble de valeurs de la variable aléatoire. Elle est notamment rencontrée lors d'un tirage d'un nombre aléatoire compris dans un intervalle $[a, b]$.

La fonction densité de probabilité d'une "**loi uniforme**" sur l'intervalle $[a, b]$, $a < b$, est définie par

$$f(x) = \begin{cases} 0 & \text{si } x < a \\ \frac{1}{b-a} & \text{si } a \leq x \leq b \\ 0 & \text{si } x > b \end{cases}$$

Nous pouvons calculer la fonction de répartition F associée à une distribution uniforme

$$F(x) = P(X \leq x) = \int_{-\infty}^x f(y) dy$$

$$F(x) = \begin{cases} 0 & \text{si } x < a \\ \frac{x-a}{b-a} & \text{si } a \leq x \leq b \\ 1 & \text{si } x > b \end{cases}$$

$F(x)$ est donc composée de trois tronçons de droites : la droite horizontale à 0 jusqu'au point a , la droite horizontale à 1 à partir du point b et la droite qui relie les deux précédentes.

Nous pouvons de même calculer les paramètres caractéristiques :

$$E(X) = \int_{-\infty}^{+\infty} x f(x) dx = \int_a^b \frac{x-a}{b-a} dx = \frac{1}{b-a} \left[\frac{x^2}{2} \right]_a^b = \frac{b^2 - a^2}{2(b-a)} = \frac{a+b}{2}$$

c'est-à-dire le milieu du segment $[a, b]$ ce qui est assez intuitif.

$$\sigma^2(X) = E(X^2) - E(X)^2 = \int_{-\infty}^{+\infty} x^2 f(x) dx - E(X)^2 = \frac{(b-a)^2}{12}$$

Loi normale

Remarque : Sont synonymes :

"loi normale", "loi de Gauss", "loi de Laplace", "loi de Laplace-Gauss"

On dit aussi d'une variable aléatoire suivant une loi normale qu'elle est gaussienne, et la courbe de la distribution est souvent qualifiée de courbe en cloche (the bell curve).

Pourquoi tant de noms? Tout simplement parce que la loi normale est la plus utilisée de toutes les fonctions de distribution. Dans la plupart des cas, lorsque l'on nous donne un intervalle d'incertitude sur une mesure d'une quantité, on considère que la distribution des erreurs associées suit une loi normale.

Une variable aléatoire X suit une "**loi normale**" si sa densité de probabilité s'écrit

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\bar{X})^2}{2\sigma^2}}$$

Elle est définie pour $-\infty < x < +\infty$.

Les deux paramètres \bar{X} et σ de la densité de probabilité sont respectivement la *moyenne* et l'*écart-type*.

La fonction de répartition F de la loi normale est donnée par

$$F(x) = P(X \leq x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{(y - \bar{X})^2}{2\sigma^2}} dy$$

Les figures 6.2 et 6.3 donnent l'allure de ces deux fonctions pour des variables centrées et réduites; dans ce cas précis, on compare parfois la fonction de répartition $F(x)$ à une fonction mathématique classique (et tabulée dans de nombreux ouvrages), la fonction erf

$$\operatorname{erf}(x) := \frac{2}{\sqrt{\pi}} \int_0^x e^{-u^2} du = 2F(x\sqrt{2}) - 1$$

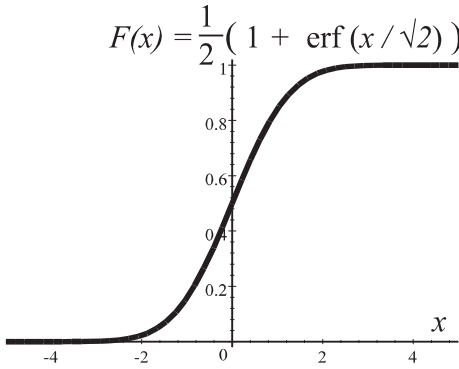


fig. 6.2 : $F(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-t^2/2} dt$

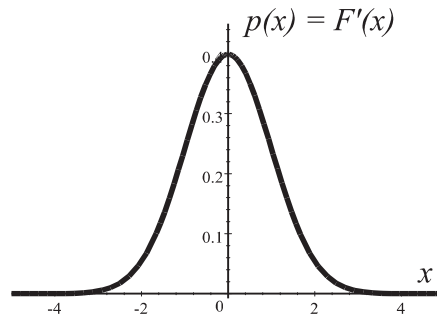


fig. 6.3 : $p(x) = \frac{dF}{dx} = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$

Si X et Y sont deux variables aléatoires indépendantes distribuées suivant la loi normale, toute combinaison linéaire $aX + bY + c$ de ces deux variables est une variable aléatoire qui suit la loi normale.

La moyenne est alors $a\bar{X} + b\bar{Y} + c$.

Et la variance devient $a^2 \sigma_X^2 + b^2 \sigma_Y^2$.

Cette dernière propriété nous permet de rapporter toute étude de variable aléatoire distribuée normalement de paramètres quelconques en l'étude d'une variable aléatoire suivant la loi normale centrée réduite. En effet, soit X une variable aléatoire continue suivant la loi normale de moyenne \bar{X} et d'écart-type σ , $Z = \frac{X - \bar{X}}{\sigma}$ est une variable aléatoire centrée réduite ainsi que nous l'avons déjà vu. Si on applique alors le changement de variable $X \rightarrow Z$ (en effectuant le changement de borne correspondant $x \rightarrow z = \frac{x - \bar{X}}{\sigma}$ dans l'intégrale), on a

$$F_X(x) = P(X \leq x) = P\left(\frac{X - \bar{X}}{\sigma} \leq \frac{x - \bar{X}}{\sigma}\right) = P(Z \leq z) = F_Z(z)$$

$F_Z(z)$ étant une fonction de répartition centrée réduite.

Pour résoudre des exercices comportant des variables suivant une ou des lois normales, il est donc précieux d'avoir sous la main une table comme celle de la figure 6.4 et de savoir l'utiliser.

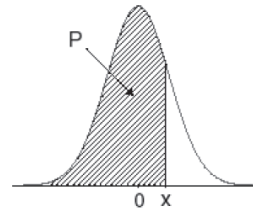
Table de la loi normale centrée-réduite

La table indique, pour $x \geq 0$, la valeur $F(x)$ de la fonction de répartition de la loi normale centrée-réduite définie par

$$F(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy.$$

Pour $x < 0$, $F(x) = 1 - F(-x)$.

La table retourne la valeur de $F(x)$ pour la valeur de x lue comme la somme des valeurs figurant en tête de la ligne et de la colonne correspondantes. Ex. : $x = 0,83 = 0,8 + 0,03 \rightarrow F(x) = 0,7967$.



x	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9 ³ 03	0.9 ³ 06	0.9 ³ 10	0.9 ³ 13	0.9 ³ 16	0.9 ³ 18	0.9 ³ 21	0.9 ³ 24	0.9 ³ 26	0.9 ³ 29
3.2	0.9 ³ 31	0.9 ³ 34	0.9 ³ 36	0.9 ³ 38	0.9 ³ 40	0.9 ³ 42	0.9 ³ 44	0.9 ³ 46	0.9 ³ 48	0.9 ³ 50
3.3	0.9 ³ 52	0.9 ³ 53	0.9 ³ 55	0.9 ³ 57	0.9 ³ 58	0.9 ³ 60	0.9 ³ 61	0.9 ³ 62	0.9 ³ 64	0.9 ³ 65
3.4	0.9 ³ 66	0.9 ³ 68	0.9 ³ 69	0.9 ³ 70	0.9 ³ 71	0.9 ³ 72	0.9 ³ 73	0.9 ³ 74	0.9 ³ 75	0.9 ³ 76
3.5	0.9 ³ 77	0.9 ³ 78	0.9 ³ 78	0.9 ³ 79	0.9 ³ 80	0.9 ³ 81	0.9 ³ 81	0.9 ³ 82	0.9 ³ 83	0.9 ³ 83
3.6	0.9 ³ 84	0.9 ³ 85	0.9 ³ 85	0.9 ³ 86	0.9 ³ 86	0.9 ³ 87	0.9 ³ 87	0.9 ³ 88	0.9 ³ 88	0.9 ³ 89
3.7	0.9 ³ 89	0.9 ³ 90	0.9 ⁴ 00	0.9 ⁴ 04	0.9 ⁴ 08	0.9 ⁴ 12	0.9 ⁴ 15	0.9 ⁴ 18	0.9 ⁴ 22	0.9 ⁴ 25
3.8	0.9 ⁴ 28	0.9 ⁴ 31	0.9 ⁴ 33	0.9 ⁴ 36	0.9 ⁴ 38	0.9 ⁴ 41	0.9 ⁴ 43	0.9 ⁴ 46	0.9 ⁴ 48	0.9 ⁴ 50
3.9	0.9 ⁴ 52	0.9 ⁴ 54	0.9 ⁴ 56	0.9 ⁴ 58	0.9 ⁴ 59	0.9 ⁴ 61	0.9 ⁴ 63	0.9 ⁴ 64	0.9 ⁴ 66	0.9 ⁴ 67

N.B. : La notation 0,9³03, par exemple, équivaut à 0,99903.

fig. 6.4 : Table de la loi normale centrée réduite

Exemple : Cherchons $P(Z \leq 1,17)$.

Sur la table de la figure 6.4, nous devons chercher dans la colonne de gauche le chiffre des unités et le premier chiffre après la virgule de Z . Il s'agit ici de 1,1. Il faut suivre maintenant la ligne correspondante jusqu'à la colonne correspondant au deuxième chiffre après la virgule de Z . Ici, il s'agit de 0,07. Au croisement de la ligne 1,1 et de la colonne 0,07, nous lisons la probabilité recherchée : $P(Z \leq 1,17) = 0,8790$

Théorème central limite

Ce théorème est essentiel pour comprendre l'importance de la loi normale et son omniprésence dans tous les domaines.

Soient X_1, X_2, \dots, X_n , n variables aléatoires indépendantes suivant des lois de probabilité quelconques, chacune d'espérance $E(X_i)$ et de variance finie σ_i^2 . La loi suivie par la variable aléatoire : $X = \sum_{i=1}^n X_i$ peut être approximée pour n grand (et sous certaines conditions) par une loi normale de paramètres

$$\bar{X} = \sum_{i=1}^n E(X_i) \quad \text{et} \quad \sigma = \sqrt{\sum_{i=1}^n \sigma_i^2}$$

Cela signifie concrètement que dès que l'on a un grand nombre de mesures d'une même quantité, ces mesures vont se distribuer suivant une loi normale quelles que soient les origines des variations dans les mesures. L'expérience montre que le théorème central limite s'applique à la quasi totalité des mesures effectuées dans le domaine industriel, économique ou scientifique, sciences naturelles ou sciences humaines.

6.5 Statistique, échantillonnage et estimation

6.5.1 Définitions

Les "**statistiques**" repose sur des ensembles de données, d'observations : recensement, cadastre, sondage, ... (ce ne sont donc que des chiffres).

Les probabilités, ainsi que nous l'avons vu, forment une branche des mathématiques et représentent donc une description rigoureuse et exacte; pour cela, elles travaillent sur des objets mathématiques parfaitement définis et abstraits (bien que d'origine concrète).

La statistique est donc la science qui utilise les méthodes mathématiques (dont beaucoup viennent des probabilités) pour étudier et analyser des statistiques en vue :

- d'en accroître les connaissances scientifiques
- de planifier des stratégies
- d'aider à des prises de décisions, etc.

Cette approche est valable quel que soit le domaine abordé : biologie, médecine, physique, mécanique, chimie, ...

La "**population**" est l'ensemble étudié par la statistique.

L'"**individu**" est un élément de la population.

Le "**caractère**", le "**facteur**" ou la "**variable**" qualifient toute caractéristique prise par les individus de la population.

L'"**échantillon**" est un sous-ensemble de la population.

On parle de "**séries statistiques**" lorsque l'on fait la correspondance entre les individus de la population et les valeurs des facteurs que l'on étudie sur cette population.

La notion d'échantillon apparaît souvent dans le langage courant : on le retrouve quotidiennement dans les journaux à propos des innombrables sondages qui visent à donner une image de l'opinion de la population globale à partir des opinions recueillies sur un nombre limité de personnes. On le retrouve aussi, par exemple, lorsque l'on veut acheter du papier peint pour tapisser une pièce, on cherche à déterminer le résultat final à partir d'un petit morceau de papier appelé également "*échantillon*".

L'échantillon statistique est constitué d'un ensemble de sujets tirés au sort dans la population à laquelle on s'intéresse. Comme dans l'exemple du papier peint, l'échantillon a pour objet de donner des informations les plus fiables possibles sur l'ensemble de la population. Un des problèmes essentiels consiste à déterminer la bonne taille de l'échantillon pour répondre à une question donnée. Reprenons l'exemple du papier peint : s'il comporte un motif périodique, l'échantillon minimal doit avoir la taille de ce motif ; si, par contre, le papier peint est uni, un échantillon plus réduit peut suffire pour tester la couleur. Une fois, l'échantillon déterminé, il faut en déduire une estimation des paramètres de la population et donner la précision que l'on peut obtenir sur ces paramètres ; c'est l'objet de ce cours.

6.5.2 Caractéristiques d'un échantillon statistique

Un "**échantillon statistique**" de taille n d'une variable aléatoire X suivant une loi de paramètres \bar{X} et σ est obtenu en répétant n fois de façons indépendantes l'épreuve qui mène à X . Cet échantillon est noté (X_1, X_2, \dots, X_n) . Il s'agit donc d'une variable aléatoire à n dimensions. La variable X_i de l'échantillon suit la même loi que X par définition. On a donc $E(X_i) = \bar{X}$ et $\sigma_{X_i} = \sigma$. On peut noter parfois (x_1, x_2, \dots, x_n) les valeurs des variables aléatoires une fois les mesures faites. On l'appelle alors de l'échantillon réalisé.

La loi de la variable aléatoire X qu'on étudie grâce à un échantillon est dite "**loi parente**".

La moyenne M_n d'un échantillon de n variables aléatoires (X_1, X_2, \dots, X_n) est la moyenne arithmétique de ses valeurs

$$M_n = \frac{X_1 + X_2 + \dots + X_n}{n}$$

M_n est une variable aléatoire qui a pour espérance $E(M_n) = \bar{X}$ et pour écart-type $\sigma_{M_n} = \sigma/n$.

Démonstration : D'après les propriétés de l'espérance et de la variance,

$$E(M_n) = \frac{E(X_1) + E(X_2) + \dots + E(X_n)}{n} = \frac{n E(X)}{n} = E(X) = \bar{X}$$

$$\sigma_{M_n}^2 = \frac{1}{n^2} \sigma^2 (X_1 + X_2 + \dots + X_n) = \frac{n \sigma^2}{n^2} = \frac{\sigma^2}{n}$$

Si l'on ajoute à cela que la variable X suit une loi normale, M_n suit aussi une loi normale de moyenne \bar{X} et de variance $\frac{\sigma^2}{n}$.

D'après le théorème central limite vu à la section précédente, plus la taille de l'échantillon est grande, plus la distribution de la moyenne M_n des valeurs de l'échantillon s'approche d'une loi normale quelle que soit la loi de la variable X . M_n suit donc une loi normale de moyenne \bar{X} et de variance $\frac{\sigma^2}{n}$ pour des valeurs de n supérieures typiquement à 30.

6.5.3 Estimateur

Un "**estimateur**" est une valeur calculée à partir des différentes valeurs x_1, x_2, \dots, x_n de l'échantillon. Il est destiné à donner une valeur la plus proche possible du paramètre inconnu de la population.

On appelle "**biais**" d'un estimateur la différence entre la valeur moyenne qu'il prendrait sur une infinité d'échantillons et la valeur recherchée du paramètre à estimer. Un estimateur est dit sans biais si sa valeur est en moyenne égale à la vraie valeur du paramètre de la population, pour une taille d'échantillon fixée. Dans le cas contraire, l'estimateur est dit biaisé.

Un estimateur est "**convergent**" si la valeur moyenne des carrés des différences entre les estimations et le paramètre recherché devient de plus en plus petite au fur et à mesure que la taille de l'échantillon augmente.

On démontre que :

- le meilleur estimateur de la moyenne de la population est :

$$m = M_n \text{ moyenne de l'échantillon.}$$

- le meilleur estimateur de la variance de la population est :

$$s^2 = \frac{n}{n-1} \sigma_{M_n}^2$$

- le meilleur estimateur de la probabilité dans la population est :

$$f, \text{ fréquence observée}$$

Ces trois estimateurs sont sans biais.

Par ailleurs, si la variable X suit une loi normale, la variable $\frac{M_n - \bar{X}}{s/\sqrt{n}}$ est une variable aléatoire qui suit une loi connue appelée loi de Student à $(n-1)$ degrés de liberté ($\nu = n-1$), dont la table est présentée dans la figure 6.5.

6.5.4 Intervalle de confiance

Supposons que l'on réalise un sondage électoral. On dispose d'un échantillon de 850 personnes dont 422 affirment vouloir voter pour A. On fournit une estimation "**ponctuelle**" des intentions de vote dans la population si on fournit un seul chiffre : ce serait ici 49,6 %. Cette présentation est souvent utilisée dans les journaux, mais elle n'est pas satisfaisante car elle ne dit rien sur la pertinence de l'échantillon choisi et donc sur la fiabilité de ce chiffre.

Il arrive que, par prudence, les journaux préfèrent donner une "**fourchette**", c'est-à-dire, un intervalle censé encadrer le pourcentage réel de la population. Par exemple, de 47% à 52%. Mais cet intervalle ne doit pas être choisi n'importe comment, il faut trouver une taille raisonnable qui ne doit pas être trop grande, sinon l'information ne devient plus intéressante (comme dans le cas : de 0% à 100%), ni trop petit car alors on prend un grand risque de se tromper. Ainsi doit-on se contenter de fournir un intervalle de confiance dans lequel on peut avoir une certaine confiance que la vraie valeur se trouve bien qu'il y ait un certain risque estimable de se tromper.

On appelle "**intervalle de confiance**" à $1-\alpha$ (on dit aussi intervalle de confiance à risque α) d'une variable aléatoire X (ou d'un paramètre) un intervalle $[a, b]$ dans lequel cette variable (ou ce paramètre) a une probabilité α de se trouver.

Remarques :

- Si la loi de répartition de la variable est symétrique, la valeur centrale c est souvent prise comme milieu de $[a, b]$. L'intervalle prend alors la forme symétrique $[c-l_\alpha, c+l_\alpha]$, l_α dépendant de la loi de la variable, de α et de la taille de l'échantillon. Dans ce

cas, on définit un *coefficient de précision* : $\frac{l_\alpha}{c}$

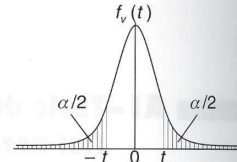
- Si la loi n'est pas symétrique, on centrera l'intervalle sur la probabilité.

- On peut chercher des intervalles de confiance de la forme $[-\infty, b]$ ou $[a, +\infty]$.

A2-Table de distribution de T (loi de Student)

$\alpha = P(|T| \geq t)$

$v =$ nombre de degrés de liberté



$v \backslash \alpha$	0,90	0,80	0,70	0,60	0,50	0,40	0,30	0,20	0,10	0,05	0,02	0,01	0,001
1	0,158	0,325	0,510	0,727	1,000	1,376	1,963	3,078	6,314	12,706	31,821	63,657	636,619
2	0,142	0,289	0,445	0,617	0,816	1,061	1,386	1,886	2,920	4,303	6,965	9,925	31,598
3	0,137	0,277	0,424	0,584	0,765	0,978	1,250	1,638	2,353	3,182	4,541	5,841	12,929
4	0,134	0,271	0,414	0,569	0,741	0,941	1,190	1,533	2,132	2,776	3,747	4,604	8,610
5	0,132	0,267	0,408	0,559	0,727	0,920	1,156	1,476	2,015	2,571	3,365	4,032	6,869
6	0,131	0,265	0,404	0,553	0,718	0,906	1,134	1,440	1,943	2,447	3,143	3,707	5,959
7	0,130	0,263	0,402	0,549	0,711	0,896	1,119	1,415	1,895	2,365	2,998	3,499	5,408
8	0,130	0,262	0,399	0,546	0,706	0,889	1,108	1,397	1,860	2,306	2,896	3,355	5,041
9	0,129	0,261	0,398	0,543	0,703	0,883	1,100	1,383	1,833	2,262	2,821	3,250	4,781
10	0,129	0,260	0,397	0,542	0,700	0,879	1,093	1,372	1,812	2,228	2,764	3,169	4,587
11	0,129	0,260	0,396	0,540	0,697	0,876	1,088	1,363	1,796	2,201	2,718	3,106	4,437
12	0,128	0,259	0,395	0,539	0,695	0,873	1,083	1,356	1,782	2,179	2,681	3,055	4,318
13	0,128	0,259	0,394	0,538	0,694	0,870	1,079	1,350	1,771	2,160	2,650	3,012	4,221
14	0,128	0,258	0,393	0,537	0,692	0,868	1,076	1,345	1,761	2,145	2,624	2,977	4,140
15	0,128	0,258	0,393	0,536	0,691	0,866	1,074	1,341	1,753	2,131	2,602	2,947	4,073
16	0,128	0,258	0,392	0,535	0,690	0,865	1,071	1,337	1,746	2,120	2,583	2,921	4,015
17	0,128	0,257	0,392	0,534	0,689	0,863	1,069	1,333	1,740	2,110	2,567	2,898	3,965
18	0,127	0,257	0,392	0,534	0,688	0,862	1,067	1,330	1,734	2,101	2,552	2,878	3,922
19	0,127	0,257	0,391	0,533	0,688	0,861	1,066	1,328	1,729	2,093	2,539	2,861	3,883
20	0,127	0,257	0,391	0,533	0,687	0,860	1,064	1,325	1,725	2,086	2,528	2,845	3,850
21	0,127	0,257	0,391	0,532	0,686	0,859	1,063	1,323	1,721	2,080	2,518	2,831	3,819
22	0,127	0,256	0,390	0,532	0,686	0,858	1,061	1,321	1,717	2,074	2,508	2,819	3,792
23	0,127	0,256	0,390	0,532	0,685	0,858	1,060	1,319	1,714	2,069	2,500	2,807	3,767
24	0,127	0,256	0,390	0,531	0,685	0,857	1,059	1,318	1,711	2,064	2,492	2,797	3,745
25	0,127	0,256	0,390	0,531	0,684	0,856	1,058	1,316	1,708	2,060	2,485	2,787	3,725
26	0,127	0,256	0,390	0,531	0,684	0,856	1,058	1,315	1,706	2,056	2,479	2,779	3,707
27	0,127	0,256	0,389	0,531	0,684	0,855	1,057	1,314	1,703	2,052	2,473	2,771	3,690
28	0,127	0,256	0,389	0,530	0,683	0,855	1,056	1,313	1,701	2,048	2,467	2,763	3,674
29	0,127	0,256	0,389	0,530	0,683	0,854	1,055	1,311	1,699	2,045	2,462	2,756	3,659
30	0,127	0,256	0,389	0,530	0,683	0,854	1,055	1,310	1,697	2,042	2,457	2,750	3,646
40	0,126	0,255	0,388	0,529	0,681	0,851	1,050	1,303	1,684	2,021	2,423	2,704	3,551
80	0,126	0,254	0,387	0,527	0,679	0,848	1,046	1,296	1,671	2,000	2,390	2,660	3,460
120	0,126	0,254	0,386	0,526	0,677	0,845	1,041	1,289	1,658	1,980	2,358	2,617	3,373
∞	0,126	0,253	0,385	0,524	0,674	0,842	1,036	1,282	1,645	1,960	2,326	2,576	3,291

fig. 6.5 : Loi de Student

On suppose, à présent et pour toute la suite, que la variable aléatoire X suit une loi normale de moyenne \bar{X} et de variance σ^2 . Les résultats donnés sont toutefois aussi vrais pour n'importe quelle loi si n est très grand (au moins supérieur à 30).

- *Intervalle de confiance de la variable X à $1-\alpha$*

Trois cas doivent être envisagés :

- \bar{X} et σ^2 sont connus :

$$[\bar{X} - z_\alpha \sigma ; \bar{X} + z_\alpha \sigma]$$

- \bar{X} est inconnu mais estimé par M_n et σ^2 est connu :

$$[M_n - z_\alpha \sigma ; M_n + z_\alpha \sigma]$$

Dans les deux cas, z_α est la valeur de z trouvée dans la table de la figure 6.4 pour une probabilité α .

- \bar{X} et σ^2 sont inconnus mais estimé par M_n et s^2 et n suffisamment grand :

$$[M_n - t_{\alpha, n-1} s ; M_n + t_{\alpha, n-1} s]$$

$t_{\alpha, n-1}$ étant issu de la loi de Student à $n-1$ degré de liberté dont nous avons déjà parlé (fig. 6.5).

- *Intervalle de confiance de la variable M_n à $1-\alpha$*

$$[\bar{X} - z_\alpha \frac{\sigma}{\sqrt{n}} ; \bar{X} + z_\alpha \frac{\sigma}{\sqrt{n}}]$$

z_α est la valeur de z trouvée dans la table de la figure 6.4 pour une probabilité α .

- *Intervalle de confiance de la variable \bar{X} à $1-\alpha$*

Deux cas doivent être envisagés :

- \bar{X} est inconnu mais estimé par M_n et σ^2 est connu :

$$[M_n - z_\alpha \frac{\sigma}{\sqrt{n}} ; M_n + z_\alpha \frac{\sigma}{\sqrt{n}}]$$

z_α est la valeur de z trouvée dans la table de la figure 6.4 pour une probabilité α .

- \bar{X} et σ^2 sont inconnus mais estimé par M_n et s^2 et n suffisamment grand :

$$[M_n - t_{\alpha, n-1} \frac{s}{\sqrt{n}} ; M_n + t_{\alpha, n-1} \frac{s}{\sqrt{n}}]$$

$t_{\alpha, n-1}$ étant issu de la loi de Student à $n-1$ degré de liberté dont nous avons déjà parlé (fig. 6.5).