

Mathématiques pour l'ingénieur

Daniel Choi¹

LMNO

Groupe Mécanique Modélisation Mathématique et Numérique
Université de Caen, Bld Maréchal Juin, 14032 Caen Cedex, France

Version 2015

1. daniel.choi@unicaen.fr

Avertissement

Ce document est inspiré de divers cours enseignés par l'auteur à l'**Université de Caen** et de divers manuels classiques de Mathématiques tels que [Rud95], [Car85], [Sch67], [Que64], [Dix76]. Il était anciennement intitulé *Mathématiques pour la Licence de Mécanique*.

Initialement destiné aux étudiants en Mécanique, il est destiné également à tout étudiant en science de niveau L3 et Master ainsi qu'aux élèves des grandes écoles d'ingénieurs. Il peut également être utile aux étudiants en mathématiques qui trouveront des versions simplifiées ainsi que des exemples d'applications des éléments de base de leurs études.

Il s'agit principalement de définir les bases mathématiques nécessaires à la modélisation, l'analyse avant même l'éventuelle résolution (essentiellement numérique) d'un problème en ingénierie. Le document n'est naturellement pas exhaustif. Il manque notamment les probabilités et statistiques ainsi que les méthodes numériques.

Ainsi, nous avons identifiés les notions d'*algèbre linéaire*, de *calcul différentiel*, de *calcul intégral*, de la *résolution de systèmes différentiels ordinaires*, comme bases indispensables à maîtriser pour tout étudiant en mécanique ou en sciences de l'ingénieur.

Viennent ensuite diverses notions comme la *théorie des fonctions d'une variable complexe* qui vont essentiellement servir à la recherche de solutions analytiques au niveau L3 ou qui pourront avoir un rôle plus large et théorique au niveau Master. La théorie de l'*intégrale de Lebesgue*, base de l'*analyse fonctionnelle*, est fondamentale pour justifier sur un plan mathématique toute la modélisation en Mécanique des milieux continus et notamment la justification d'un problème bien posé. Elle permet également de bien définir la *transformation de Fourier* qui est un outil très important permettant la résolution analytique ou semi-analytique d'un grand nombre de problèmes différentiels ou aux dérivées partielles. Mais il n'est pas nécessaire pour un étudiant en Mécanique ou à un élève ingénieur de la maîtriser : on se contentera de présenter les résultats les plus importants pour les applications. Nous choisissons de présenter également la *théorie des distributions*, permettant de représenter notamment les cas de charges ponctuelles donnant un sens généralisé à la notion de dérivée. Nous terminons par une très courte introduction à l'analyse fonctionnelle en présentant les espaces de Hilbert et en particulier les espaces de Sobolev donnant le cadre théorique des problèmes bien posés et au delà de la théorie, les bases de la méthode des éléments-finis, omniprésents dans la résolution numérique des problèmes issus de la mécanique.

Ce document est encore et toujours **incomplet**, il manque un certain nombre

de résultats, ne contient pratiquement pas démonstrations et surtout ce document manque cruellement d'exemples.

Il est possible (probable) qu'il contienne des erreurs, même graves. Toutes contributions suggestions ou commentaires sont les bienvenus : envoyez moi un email à daniel.choi@unicaen.fr

Une version au format pdf est disponible ici : <http://www.meca.unicaen.fr/Enseignement/Document/CoursChoi/mathmeca.pdf>

Une version avec cadres est disponible ici : <http://www.meca.unicaen.fr/Enseignement/Document/CoursChoi/mathmeca-frame/index.html>

Table des matières

Conventions et notations	4
Notations	4
Convention de sommation suivant les indices répétés	4
Notation des dérivées partielles	5
Indices et exposant Grecs ou Latins	6
1 Algèbre linéaire : Espaces vectoriels et applications linéaires	7
1.1 Espaces vectoriels	7
1.1.1 Espace vectoriel : une présentation intuitive	7
1.1.2 Exemples d'espace vectoriel	8
1.1.3 Espace vectoriel : la définition	8
1.1.4 Sous-espace vectoriel	10
1.1.5 Indépendance linéaire : vecteurs libres, vecteurs liés	10
1.1.6 Produit d'espaces vectoriels	11
1.1.7 Cas des espaces vectoriels réels de dimension finie	11
1.2 Cas des espaces \mathbb{R}^n	12
1.2.1 Produit scalaire Euclidien et norme Euclidienne dans \mathbb{R}^n	12
1.3 Application linéaire	14
1.3.1 Opérateurs linéaires	15
1.3.2 Quelques exemples d'espaces vectoriels et d'applications linéaires	16
1.4 Matrice d'une application linéaire	17
1.4.1 Bijection entre $\mathcal{L}(X, \mathbb{R})$ et X	18
1.4.2 Norme d'une application linéaire	19
1.4.3 Produit de deux matrices et composition de deux applications linéaires	19
1.4.4 Changement de bases et Matrices de passage	20

1.5	Cas des opérateurs linéaires – Matrices carrés	22
1.5.1	Adjoint d'un opérateur linéaire	22
1.5.2	Partie symétrique et antisymétrique d'un opérateur linéaire de \mathbb{R}^n	22
1.6	Déterminant	23
1.6.1	Déterminant de n vecteurs de \mathbb{R}^n	23
1.6.2	Déterminant d'une matrice carré – Déterminant d'un opérateur linéaire	24
1.6.3	Règles de calcul d'un déterminant – Développement suivant une ligne ou une colonne	25
1.7	Éléments de théorie spectrale	26
1.7.1	Spectre d'un opérateur linéaire	26
1.7.2	Réduction d'un opérateur linéaire	27
1.8	Cas des opérateurs auto-adjoint ou matrices symétriques	28
1.9	Annexe : quelques preuves	29
1.9.1	Preuve du théorème 1.3.4	29
1.9.2	Preuve du théorème 1.3.8	30
1.9.3	Preuve de la proposition 1.4.4	31
2	Calcul différentiel	32
2.1	Éléments de topologie métrique	32
2.2	Fonctions continues de plusieurs variables	34
2.3	Fonction différentiable – Application dérivée	34
2.3.1	Cas des fonctions réelles définie sur un intervalle réel	34
2.3.2	Cas général	35
2.3.3	Matrice Jacobienne et dérivées partielles	38
2.4	Gradient, divergence, rotationnel	39
2.4.1	Gradient	39
2.4.2	Divergence	39
2.4.3	Rotationnel	40
2.4.4	Quelques remarques et propriétés des opérateurs différentiels	40
2.5	Dérivation le long d'une courbe – Dérivation partielle dans une direction	41
2.6	Fonction dérivée	42
2.7	Dérivée seconde	42
2.8	Principaux théorèmes sur les fonctions de plusieurs variables	43
2.9	Formule de Taylor	43
2.10	Coordonnées cylindriques et sphériques	44
2.10.1	Expressions en coordonnées Cylindriques	44
2.10.2	Expressions en coordonnées sphériques	47
2.11	Extrema d'une fonctions différentiable	47
2.11.1	Dérivation d'une fonctionnelle d'énergie	48

3	Intégrale de Riemann	49
3.1	Définition de l'intégrale de Riemann	49
3.2	Principales propriétés	50
3.3	Intégration et dérivation	51
3.4	Formule intégrale de Taylor	51
3.5	Intégrale sur un contour	51
3.6	Intégrale sur un pavé de \mathbb{R}^n .	52
3.7	Intégrale multiple – Formule de Jacobi	52
4	Théorème de Stokes et Formule de la divergence	54
4.1	Formes différentielles	54
4.2	Différentielle d'une forme différentielle	56
4.3	Théorème de Stokes – Formule de la divergence	56
4.4	Formule de la divergence : définition	57
4.4.1	Cas où Ω est un domaine de \mathbb{R}	57
4.4.2	Cas où Ω est un domaine de \mathbb{R}^2	57
4.4.3	Cas où Ω est un domaine de \mathbb{R}^3	58
4.4.4	Formules de Stokes	59
4.4.5	Formules de Green	59
5	Calcul différentiel sur un tenseur d'ordre 2	61
5.1	Tenseur d'ordre 2	61
5.2	Gradient d'un tenseur d'ordre 2	62
5.3	Divergence d'un tenseur d'ordre 2	62
5.4	Tenseur en coordonnées cylindriques	63
5.5	Tenseur en coordonnées sphériques	65
6	Systèmes différentiels ordinaires	66
7	Fonction d'une variable complexe	67
7.1	Nombres Complexes	67
7.2	Fonctions holomorphes	69
7.2.1	Fonctions holomorphes	69
7.2.2	Conditions de Cauchy-Riemann	70
7.3	Suite et Séries de nombres complexes	70
7.4	Divers	71
7.4.1	Fonctions exponentielle et trigonométrique	71
7.5	Intégrales complexes	72
7.5.1	Formule intégrale de Cauchy	72
7.5.2	Série de Laurent	74
7.5.3	Théorèmes des résidus	75
7.5.4	Classification des Pôles singuliers	75

7.5.5	Représentation conforme	76
8	Théorie restreinte de la mesure	
	La mesure de Lebesgue	77
8.1	Ensemble élémentaire	78
8.1.1	Pavés et mesure d'un pavé	78
8.1.2	Ensemble élémentaire et mesure d'un ensemble élémentaire	78
8.2	Ensemble mesurable	79
8.2.1	Mesure extérieure	79
8.2.2	Ensemble mesurable	79
8.2.3	Mesure d'un ensemble mesurable	80
8.2.4	Ensemble de mesure nulle	80
8.3	Fonctions mesurables	80
8.3.1	Limites de fonctions continues	80
8.3.2	Propriétés des fonctions mesurables	81
8.3.3	Classe des fonctions nulles presque partout	81
8.4	Quelques remarques	81
8.4.1	Sur les fonctions non-mesurables	81
8.4.2	Sur la théorie abstraite de la mesure	81
9	Intégrale de Lebesgue	
	Théorème de convergence dominée	82
9.1	Fonctions étagées	82
9.1.1	Fonctions élémentaires	82
9.1.2	Fonctions étagées	83
9.1.3	Intégrale d'une fonction étagée	84
9.1.4	Intégrale de Lebesgue d'une fonction mesurable	84
9.2	Théorèmes de Lebesgue	85
9.2.1	Théorème de convergence monotone	85
9.2.2	Lemme de Fatou	86
9.2.3	Théorème de convergence dominée	87
9.2.4	Théorème de dérivation sous le signe intégral	88
9.2.5	Théorème de Fubini	88
9.3	Espaces \mathbb{L}^1 et \mathbb{L}^2	89
10	Transformation de Fourier des fonctions	90
10.1	Preliminaires	90
10.1.1	Quelques intégrales utiles	90
10.1.2	Espace $\mathcal{S}(\mathbb{R}^n)$	90
10.1.3	Produit de convolution	91
10.1.4	Suites régularisantes	91

10.2	Transformation de Fourier	92
10.3	Rapports avec la convolution	93
10.4	Exemples	94
11	Introduction à la théorie des distributions	95
11.1	Introduction	95
11.2	Espace des fonctions infiniment dérivables à support compact	96
11.3	Théorie des distributions	96
11.3.1	Exemples de distributions	97
11.3.2	Dérivation des distributions	97
11.3.3	Opération sur les distributions	98
11.3.4	Convergence au sens des distributions	98
11.3.5	Support d'une distribution	98
11.4	Variantes des distributions	98
11.4.1	Distribution d'ordre finie	98
11.4.2	Distribution à support compact	99
11.4.3	Distribution tempérée	99
11.5	Produit de convolution de deux distributions	99
12	Petite introduction à l'analyse fonctionnelle	100
12.1	Espace de Banach	101
12.1.1	Quelques caractérisations abstraites d'un espace de Banach	101
12.1.2	Quelques exemples d'espaces de Banach	102
12.1.3	Quelques exemples d'espaces qui ne sont pas des espaces de Banach	102
12.2	Espace de Hilbert	102
12.2.1	Quelques exemples d'espaces de Hilbert	103
12.2.2	Principaux résultats sur les Hilberts	103

Conventions et notations

Dans ce document, nous serons parfois amené à utiliser quelques conventions de notations propres à la mécanique. En particulier nous utiliserons la convention de sommation par rapport aux indices répétés (on lit également *convention d'Einstein*) et nous noterons souvent les dérivées partielles à l'aide d'un indice précédé d'une virgule.

Nous commençons toutefois par quelques notations utilisés dans ce document qui sont du reste tout à fait usuel.

Notations

Nous n'avons pas cherché à faire dans l'originalité, ainsi dans tout le document \mathbb{R} désigne l'espace des nombres réels tandis que \mathbb{C} sera l'espace des nombres complexes.

Les quantités scalaire seront systématiquement noté en italique, tandis que les objets vectoriels seront noté soit avec une flèche soit en caractère soulignés :

On désigne généralement par f ou g une *fonction scalaire*, c'est à dire à valeur dans \mathbb{R} (ou éventuellement \mathbb{C}).

Les **objets vectoriels** seront souvent désigné par \underline{u} dont les composantes notées u_i sont des quantités scalaires.

Le symbole Ω représentera généralement un domaine à bord régulier de \mathbb{R}^d , où la dimension d , en mécanique, désigne souvent les nombres 2 ou 3.

Convention de sommation suivant les indices répétés

La convention d'Einstein sur la sommation sur les indices ou exposants répétés est une convention destiné à alléger les écritures dans les formules mathématiques sans pour autant les rendre ambigu.

La convention implique une sommation sur des termes produits dès lors qu'ils présentent des indices répétés :

Ainsi, par exemple, pour $\underline{x} = [x_1, x_2, \dots, x_n]^\top$ et $\underline{y} = [y_1, y_2, \dots, y_n]^\top$, deux vecteurs de \mathbb{R}^n , le produit scalaire :

$$\underline{x} \cdot \underline{y} = \sum_{i=1}^n x_i y_i,$$

où l'on remarque l'indice i qui apparaît répété, sera noté plus simplement :

$$\underline{x} \cdot \underline{y} = x_i y_i.$$

De même pour un produit de matrices $C = BA$:

$$c_{kj} = \sum_{i=1}^m b_{ki} a_{ij} \quad \longrightarrow \quad c_{kj} = b_{ki} a_{ij}.$$

Si un vecteur \underline{x} a pour composantes (x_1, x_2, \dots, x_n) dans la base $\underline{e}_1, \underline{e}_2, \dots, \underline{e}_n$, on écrit

$$\underline{x} = \sum_{i=1}^n x_i \underline{e}_i \quad \longrightarrow \quad \underline{x} = x_i \underline{e}_i$$

Si on note, dans \mathbb{R}^3 le produit mixte des vecteurs de la base canonique :

$$\varepsilon_{ijk} = (\underline{e}_i, \underline{e}_j, \underline{e}_k) = \underline{e}_i \cdot (\underline{e}_j \wedge \underline{e}_k)$$

On peut écrire le produit mixte de trois vecteurs \underline{a} , \underline{b} et \underline{c} par

$$(\underline{a}, \underline{b}, \underline{c}) = \sum_{i=1}^3 \sum_{j=1}^3 \sum_{k=1}^3 \varepsilon_{ijk} a_i b_j c_k \quad \longrightarrow \quad (\underline{a}, \underline{b}, \underline{c}) = \varepsilon_{ijk} a_i b_j c_k,$$

où on a appliqué la convention sur les trois indices répétés i , j et k .

En réalité, ce manuscrit a d'abord été rédigé sans utiliser cette convention. Ainsi, on ne la trouve pas dans la plupart des chapitres. Nous avons cependant ajouté en bleu, dans la mesure du possible.

Notation des dérivées partielles

En mécanique et en mathématiques en général, nous avons souvent affaire à des systèmes d'équations aux dérivées partielles. Ainsi pour alléger les notations on préférera utiliser la notation en indice précédé d'une virgule :

$$\begin{aligned} \frac{\partial f}{\partial x} &= f_{,x} \\ \frac{\partial u_i}{\partial x_j} &= u_{i,j} \\ \frac{\partial^2 u_i}{\partial x_j \partial x_k} &= u_{i,jk}. \end{aligned}$$

Si on considère une fonction scalaire f défini sur $\Omega \subset \mathbb{R}^n$ (à valeur dans \mathbb{R}) alors pour toute direction $\underline{h} = [h_1, h_2, \dots, h_n]^\top$, on écrit la dérivée de f dans la direction \underline{h} :

$$\nabla f(x)(\underline{h}) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(x) h_i \quad \longrightarrow \quad \nabla f(x)(\underline{h}) = f_{,i}(x) h_i.$$

où on a également utilisé la convention de sommation suivant les indices répétés.

Indices et exposant Grecs ou Latins

Bien que la plupart des théories mathématique soit présenté dans un espace abstrait de dimension n , en Mécanique les problèmes sont généralement posés dans les variables d'espace. C'est à dire qu'on travaille en dimension 3 en général et en dimension 2 pour des problèmes plans.

Alors, il est une façon bien commode si un problème est en dimension 2 ou 3 :

- l'utilisation réservée des indices et exposants Grecs en dimension 2,
- l'utilisation réservée des indices et exposants Latins en dimension 3.

Ainsi le système

$$\sigma_{ij,j} + f_j = 0 \quad i = \{1, 2, 3\}$$

identifie immédiatement un problème à trois dimension (avec une sommation sur j), tandis que dans

$$\sigma_{\alpha\beta,\beta} + f_\alpha = 0 \quad \alpha = \{1, 2\}$$

on reconnait un problème en dimension 2 (avec une sommation sur β).

CHAPITRE 1

Algèbre linéaire : Espaces vectoriels et applications linéaires

Une bonne connaissance de l'algèbre linéaire est essentielle en Mathématiques appliquées. c'est une théorie sur laquelle repose le calcul différentiel ou encore la résolution de systèmes linéaires, indispensable pour une éventuelle approximation numérique des problèmes de la mécanique. Nous présentons ici les définitions et les résultats principaux. Nous renvoyons par exemple à [Que64] ou n'importe quel cours d'algèbre de premier cycle universitaire ou de classes préparatoires tel que [?].

1.1 Espaces vectoriels

Nous commençons par la notion d'espace vectoriel sans introduire des notions usuelles d'algèbre telles que les *groupes*, les *anneaux* ou encore les *corps*. Elles seront tout de même rappelées, nous préférons une présentation intuitive basée sur l'ensemble des nombres réels \mathbb{R} ou des nombres complexes \mathbb{C} , qui sont des exemples particuliers de *corps*.

1.1.1 Espace vectoriel : une présentation intuitive

Un espace vectoriel est un ensemble d'éléments $\underline{x}, \underline{y}, \dots$, appelés *vecteurs*, sur lequel sont définis les opérations linéaires : une *addition* et une *multiplication* par un scalaire

(un nombre réel ou complexe), avec les règles usuelles liées à la notion d'addition et de multiplication : associativité, commutativité, élément neutre, transitivité,...

Dans le jargon mathématicien l'addition désigne une opération interne entre deux vecteurs donnant un vecteur, tandis que la multiplication est une opération externe entre un scalaire et un vecteur donnant encore un vecteur.

On parle d'*espace vectoriel réel* si les scalaires sont les nombres réels. De même, on parle d'*espace vectoriel complexe* si les scalaires sont les nombres complexes. Dans toute la suite, sauf mention contraire, les espaces vectoriels seront réels.

1.1.2 Exemples d'espace vectoriel

1. Le premier exemple d'espace vectoriel réel est l'ensemble des nombres réels \mathbb{R} lui-même, avec évidemment les addition et multiplication usuelles. Dans ce cas particulier, les nombres réels sont à la fois considérés comme des scalaires *et* comme des vecteurs en tant qu'éléments de l'espace vectoriel.

Il en va de même pour l'ensemble des nombres complexes \mathbb{C} . Cependant, dans ce cas, on peut distinguer deux possibilités : \mathbb{C} en tant qu'espace vectoriel complexe, ou bien \mathbb{C} en tant qu'espace vectoriel réel (on parle de plan complexe pour cette dernière).

2. Autre exemple, l'ensemble des fonctions définies sur un intervalle de $I \subset \mathbb{R}$ à valeur réelle : on définit naturellement l'addition de deux fonctions f et g par :

$$h = f + g \quad \Leftrightarrow \quad h(x) = f(x) + g(x) \quad \forall x \in I.$$

La multiplication d'une fonction f par un scalaire λ est définie par

$$h = \lambda f \quad \Leftrightarrow \quad h(x) = \lambda f(x) \quad \forall x \in I.$$

En fait dans cet exemple, un point crucial est que l'addition est une opération interne : autrement dit, l'addition de deux fonctions doit encore être une fonction définie sur I à valeur réelle. Il en va de même pour la multiplication. C'est ce qu'on appelle la *stabilité linéaire*.

Une conséquence est qu'un espace vectoriel contient nécessairement un élément neutre pour l'addition, à savoir le vecteur nul.

Ainsi, le sous-ensemble des fonctions valant 1 en 0, n'est pas un espace vectoriel pour l'addition et la multiplication usuelle, puisque la fonction nulle n'en fait pas partie.

1.1.3 Espace vectoriel : la définition

Pour information et par souci de complétude, donnons la définition exacte et précise et abstraite¹ de la notion d'espace vectorielle :

1. Comme on pourra le remarquer, en Mathématique plus on est abstrait et plus on est précis ;-)... C'est comme ça ! Et encore : ici, nous nous sommes restreint aux cas réel ou complexe...

Définition 1.1.1. Soit E , un ensemble abstrait, muni d'une opération interne notée '+' et d'une opération externe notée '.' avec un scalaire réel (resp. complexe).

On dit que $(E, '+', '.')$ possède une structure d'espace vectoriel sur \mathbb{R} (resp. sur \mathbb{C}) si les conditions suivantes sont satisfaites :

1. $\forall \underline{x}, \underline{y} \in E, \underline{x} + \underline{y} \in E$
2. $\forall \underline{x} \in E, \forall \lambda \in \mathbb{R}$ (resp. $\in \mathbb{C}$), $\lambda.\underline{x} \in E$
3. $\forall \underline{x}, \underline{y}, \underline{z} \in E, (\underline{x} + \underline{y}) + \underline{z} = \underline{x} + (\underline{y} + \underline{z})$
4. $\forall \underline{x}, \underline{y} \in E, \underline{x} + \underline{y} = \underline{y} + \underline{x}$
5. $\forall \underline{x} \in E, \forall \lambda, \mu \in \mathbb{R}$ (resp. $\in \mathbb{C}$), $(\lambda\mu).\underline{x} = \lambda.(\mu.\underline{x})$
6. $\forall \underline{x}, \underline{y} \in E, \forall \lambda \in \mathbb{R}$ (resp. $\in \mathbb{C}$), $\lambda.(\underline{x} + \underline{y}) = \lambda.\underline{x} + \lambda.\underline{y}$
7. $\forall \underline{x} \in E, \forall \lambda, \mu \in \mathbb{R}$ (resp. $\in \mathbb{C}$), $(\lambda + \mu).\underline{x} = \lambda.\underline{x} + \mu.\underline{x}$
8. Il existe $0_E \in E$ tel que $\forall \underline{x} \in E, \underline{x} + 0_E = 0_E + \underline{x} = \underline{x}$
9. $\forall \underline{x} \in E, \exists \underline{y} \in E$ tel que $\underline{x} + \underline{y} = 0_E$

A la première lecture, cela fait beaucoup de choses à apprendre et à connaître, mais il est facile de ne retenir que l'essentiel : l'ensemble E muni de son opération interne (addition) est un groupe commutatif : l'addition doit être commutative (2), associative (3), posséder un élément neutre (8) et pour chaque élément de E il existe un symétrique pour l'addition (9) ; on parle également d'élément opposé. On y ajoute l'opération externe (produit par un scalaire) possédant les propriétés de distributivité à gauche (6), d'associativité mixte (5) et de distributivité à droite (7).

Exercice. Montrer que dans tout espace vectoriel E , réel ou complexe :

$$\forall \underline{x} \in E, \quad 0.\underline{x} = \underline{0}_E$$

Le point fondamental dans une structure d'espace vectoriel et qui le différencie d'une structure de groupe est l'opération externe, produit par un scalaire. Autrement dit un espace vectoriel, sauf s'il est réduit à un élément neutre, est un ensemble infini.

Il faut noter que l'addition entre deux scalaires (7) est notée de la même façon que l'opérateur interne avec le symbole +, cependant, il ne s'agit pas en général de la même opération. Cet abus d'écriture peut désorienter certains, mais dans la pratique, il n'y a jamais d'ambiguïté.

Un élément d'un espace vectoriel est appelé vecteur, et hormis les cas triviaux d'espaces vectoriels tels que \mathbb{R} , nous choisissons de noter les vecteurs de façon soulignée.

Remarquons que la notion espace vectoriel est une notion de structure d'un ensemble muni d'opérateurs. On désigne usuellement de manière abusive (mais sans ambiguïté) l'espace vectoriel $(E, '+', '.')$ tout simplement par E , car les opérations interne (addition) et externe (multiplication par un scalaire) sont généralement implicites et ne nécessitent pas d'être rappelées.

C'est bien ce que nous indiquions dans la sous-section précédente : **un espace vectoriel est en gros un ensemble pour lequel est défini une addition interne et une multiplication par un nombre scalaire avec les propriétés de stabilité linéaire (1) et (2).**

Exercice. Montrer que les exemples 1.1.2 sont bien des espaces vectoriels.

1.1.4 Sous-espace vectoriel

Définition 1.1.2. Soit E un espace vectoriel réel et soit un sous-ensemble non vide $X \subset E$. X est un [sous-]espace vectoriel réel s'il satisfait aux conditions de stabilité linéaire, c'est à dire :

$$\begin{aligned}\underline{x} + \underline{y} &\in X \\ \alpha \underline{x} &\in X\end{aligned}$$

pour tout $\underline{x}, \underline{y} \in X$ et pour tout scalaire $\alpha \in \mathbb{R}$.

Remarque 1. On a naturellement une définition analogue pour les sous-espaces d'espaces vectoriels complexe.

Par exemple, toute droite de \mathbb{R}^2 passant par 0 est un sous-espace vectoriel de \mathbb{R}^2 ou encore toute droite ou tout plan de \mathbb{R}^3 passant par 0 est un sous-espace vectoriel de \mathbb{R}^3 .

1.1.5 Indépendance linéaire : vecteurs libres, vecteurs liés

Soit E un espace vectoriel (réel ou complexe) et soient p vecteurs $\underline{x}_1, \underline{x}_2, \dots, \underline{x}_p$. Si $\alpha_1, \dots, \alpha_p$ sont des scalaires, alors on dit que le vecteur

$$\alpha_1 \underline{x}_1 + \dots + \alpha_p \underline{x}_p$$

est une *combinaison linéaire* des vecteurs $\underline{x}_1, \dots, \underline{x}_p$.

Définition 1.1.3. On dit que les vecteurs $\underline{x}_1, \underline{x}_2, \dots, \underline{x}_p$ forment une famille libres, ou bien qu'ils sont linéairement indépendants si toute combinaison linéaire non nulle de ces vecteurs est nécessairement non nulle :

$$\text{Si } \alpha_1 \underline{x}_1 + \dots + \alpha_p \underline{x}_p = 0 \text{ alors, } \alpha_1 = \dots = \alpha_p = 0.$$

Définition 1.1.4. On dit que p vecteurs $\underline{x}_1, \underline{x}_2, \dots, \underline{x}_p$ sont liés si ils ne sont pas libres.

Proposition 1.1.5. Soit $\underline{x}_1, \underline{x}_2, \dots, \underline{x}_p$ une famille de p vecteurs. Si ces vecteurs sont libres (ou linéairement indépendants), alors la décomposition de

$$\underline{x} = \alpha_1 \underline{x}_1 + \dots + \alpha_p \underline{x}_p$$

est unique.

1.1.6 Produit d'espaces vectoriels

Soient E_1 et E_2 deux espaces vectoriels réels (resp. complexes). On définit l'espace produit par l'ensemble

$$E_1 \times E_2 = \{(\underline{x}_1, \underline{x}_2) \text{ tels que } \underline{x}_1 \in E_1 \text{ et } \underline{x}_2 \in E_2\}.$$

On munit cet espace produit d'une addition et d'une multiplication induite par celles de E_1 et de E_2 :

$$\begin{aligned}(\underline{x}_1, \underline{x}_2) + (\underline{y}_1, \underline{y}_2) &= (\underline{x}_1 + \underline{y}_1, \underline{x}_2 + \underline{y}_2) \\ \lambda(\underline{x}_1, \underline{x}_2) &= (\lambda\underline{x}_1, \lambda\underline{x}_2)\end{aligned}$$

Munis de ces opérations linéaires, l'espace produit $E_1 \times E_2$ est encore un espace vectoriel réel (resp. complexe).

Ainsi, par exemple on note $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$ et par récurrence :

$$\mathbb{R}^n = \mathbb{R} \times \mathbb{R}^{n-1} \quad \forall n = 2, 3, \dots$$

1.1.7 Cas des espaces vectoriels réels de dimension finie

Définition 1.1.6. On dit qu'un espace vectoriel E est de dimension n s'il possède une famille libre de n vecteurs et que toute famille de $n + 1$ vecteurs est liée.

Définition 1.1.7. On dit qu'une famille B de vecteurs de E est génératrice ou encore que cette famille engendre E si tout vecteur de E est une combinaison linéaire des vecteurs de B .

Définition 1.1.8. Une base d'un espace vectoriel E est une famille libre et génératrice, c'est-à-dire qu'une base est constituée d'une famille de vecteurs linéairement indépendants et qui engendre E .

Ainsi, si $B = (\underline{b}_1, \dots, \underline{b}_n)$ est une base de E alors tout vecteur \underline{x} de E se décompose de manière *unique* en une combinaison linéaire des vecteurs de la base B :

$$\underline{x} = \sum_{i=1}^{i=n} x_i \underline{b}_i.$$

Les coefficients x_1, \dots, x_n sont les *coordonnées* de \underline{x} dans la base B .

Théorème 1.1.9. Soit E un espace vectoriel de dimension n .

1. Une famille de n vecteurs de E engendre E si et seulement si c'est une famille libre.
2. E possède au moins une base et toute base est constitué de n vecteurs.
3. Si $Y = (\underline{y}_1, \dots, \underline{y}_m)$ ($1 \leq m \leq n$) est un système libre de E , alors il existe une base de E contenant ce système. Autrement dit, on peut toujours compléter une famille libre en une base.

1.2 Cas des espaces \mathbb{R}^n

Généralement, les vecteurs de \mathbb{R}^n sont notés par une représentation dite matricielle :

$$\underline{x} \longleftrightarrow \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = [x_1, \dots, x_n]^\top,$$

où le symbole $[\dots]^\top$ indique une matrice transposée (i.e les lignes sont remplacées par des colonnes).

La base canonique de \mathbb{R}^n est constituée des vecteurs $\underline{e}_i = [0, \dots, 1, \dots, 0]^\top$, dont toutes les coordonnées sont nulles sauf la $i^{\text{ème}}$ qui vaut 1.

Si $\underline{x} \in \mathbb{R}^n$ avec $\underline{x} = [x_1, \dots, x_n]^\top$, alors

$$\underline{x} = \sum_{i=1}^{i=n} x_i \underline{e}_i$$

La base $(\underline{e}_1, \dots, \underline{e}_n)$ est appelée *base canonique* de \mathbb{R}^n .

On a alors de cette façon :

Théorème 1.2.1. *L'espace \mathbb{R} est de dimension n . On note $\dim \mathbb{R}^n = n$.*

Théorème 1.2.2. *Tout espace vectoriel de dimension n est isomorphe à \mathbb{R}^n .*

1.2.1 Produit scalaire Euclidien et norme Euclidienne dans \mathbb{R}^n

Soit $E = \mathbb{R}^n$ et soit Pour tout \underline{x} et \underline{y} appartenant à E et dont les composantes dans la base canonique $\{\underline{e}_i\}$ de E sont notées respectivement x_i et y_i :

$$\underline{x} = \sum_{i=1} x_i \underline{e}_i, \quad \underline{y} = \sum_{i=1} y_i \underline{e}_i.$$

On définit le produit scalaire (Euclidien) de \underline{x} par \underline{y} dans \mathbb{R}^n

$$(\underline{x}, \underline{y})_{\mathbb{R}^n} = \sum_{i=1}^n x_i y_i. \quad (1.1)$$

Il s'agit donc d'une application qui associe à deux éléments \underline{x} et \underline{y} de \mathbb{R}^n , un nombre réel (un scalaire) qu'on note dans ce document $(\underline{x}, \underline{y})_{\mathbb{R}^n}$.

Il y a beaucoup de notations possibles pour le produit scalaire de deux vecteurs. C'est une source importante de confusion chez certains étudiants :

$$\begin{aligned}(\underline{x}, \underline{y})_{\mathbb{R}^n} &= (\underline{x}, \underline{y}) \\ &= \langle \underline{x}, \underline{y} \rangle \\ &= \langle \underline{x}, \underline{y} \rangle \\ &= \underline{x} \cdot \underline{y}.\end{aligned}$$

Les notations les plus courantes sont $\langle \underline{x}, \underline{y} \rangle$ et $(\underline{x}, \underline{y})$, adoptée dans ce document, où la notion de produit scalaire Euclidien dans \mathbb{R}^n est implicite (encore un abus de notation).

On vérifie aisément les propriétés de bilinéarité et de symétrie :

Proposition 1.2.3. Soient $\underline{x}, \underline{y}, \underline{z} \in \mathbb{R}^n$ et soit un réel λ . On a

$$\begin{aligned}(\underline{x}, \underline{y}) &= (\underline{y}, \underline{x}) \\ (\underline{x} + \underline{z}, \underline{y}) &= (\underline{x}, \underline{y}) + (\underline{z}, \underline{y}) \\ (\lambda \underline{x}, \underline{y}) &= \lambda(\underline{x}, \underline{y})\end{aligned}$$

Le produit scalaire induit une forme quadratique et nous avons trivialement la propriété de positivité :

Proposition 1.2.4. Soit $\underline{x} \in \mathbb{R}^n$, alors $(\underline{x}, \underline{x}) \geq 0$.

Théorème 1.2.5. Inégalité de Cauchy-Schwarz – $\forall \underline{x}, \underline{y} \in \mathbb{R}^n$,

$$|(\underline{x}, \underline{y})| \leq (\underline{x}, \underline{x})^{\frac{1}{2}} (\underline{y}, \underline{y})^{\frac{1}{2}} \quad (1.2)$$

Démonstration. La preuve de cette inégalité repose sur une astuce classique utilisée dans un grand nombre de démonstration mathématiques utilisant un paramètre scalaire $\lambda \in \mathbb{R}$: Considérons la quantité, nécessairement positive, quel que soit λ

$$(\underline{x} + \lambda \underline{y}, \underline{x} + \lambda \underline{y}) \geq 0$$

En développant, on obtient

$$(\underline{x}, \underline{x}) + 2\lambda(\underline{x}, \underline{y}) + \lambda^2(\underline{y}, \underline{y}) \geq 0.$$

On reconnaît donc un polynôme en λ du second degré qui est toujours positif. Cela signifie que le discriminant de ce polynôme est nécessairement négatif, c'est à dire

$$4(\underline{x}, \underline{y})^2 - 4(\underline{x}, \underline{x})(\underline{y}, \underline{y}) \leq 0$$

d'où

$$(\underline{x}, \underline{y})^2 \leq (\underline{x}, \underline{x})(\underline{y}, \underline{y}).$$

□

Définition 1.2.6. *Le produit scalaire dans \mathbb{R}^n induit la norme Euclidienne :*

$$\forall \underline{x} \in \mathbb{R}^n, \quad \|\underline{x}\|_{\mathbb{R}^n} = \left[\sum_{i=1}^n x_i^2 \right]^{\frac{1}{2}}.$$

On notera, sans ambiguïté :

$$\|\underline{x}\| = \|\underline{x}\|_{\mathbb{R}^n}.$$

Grâce à l'inégalité de Cauchy-Schwarz, on montre facilement que la norme Euclidienne est bien une norme, c'est à dire qu'elle satisfait aux conditions, $\forall \underline{x}, \underline{y} \in \mathbb{R}^n$ et $\forall \lambda \in \mathbb{R}$:

$$\begin{aligned} \|\underline{x}\| &= 0 \Leftrightarrow \underline{x} = \underline{0} \\ \|\lambda \underline{x}\| &\leq |\lambda| \|\underline{x}\| \\ \|\underline{x} + \underline{y}\| &\leq \|\underline{x}\| + \|\underline{y}\|. \end{aligned}$$

1.3 Application linéaire

Dans toute la suite, sauf indication contraire, un espace vectoriel désignera un espace vectoriel réel.

Définition 1.3.1. *Une application A d'un espace vectoriel E vers un espace vectoriel F est une application linéaire si*

1. $A(\underline{x}_1 + \underline{x}_2) = A(\underline{x}_1) + A(\underline{x}_2)$ pour tout $\underline{x}_1, \underline{x}_2 \in E$.
2. $A(\alpha \underline{x}) = \alpha A(\underline{x})$ pour tout $\underline{x} \in E$ et tout scalaire α .

Pour désigner une application linéaire, on lit parfois *homomorphisme*, dans un langage plus savant².

On note souvent $A\underline{x}$, l'image d'un vecteur \underline{x} par une application linéaire A , au lieu de $A(\underline{x})$. C'est essentiellement parce qu'en représentation matricielle (voir plus loin) l'application ou la composition s'assimile à un produit de matrice.

Remarquons que si A est une application linéaire alors on a nécessairement

$$A(\underline{0}) = \underline{0}.$$

Proposition 1.3.2. *Une application linéaire est entièrement déterminée par les images des vecteurs d'une base.*

2. la notion d'application linéaire peut être considérée comme le prolongement d'un homomorphisme de groupe à un espace vectoriel

En effet, si $(\underline{x}_1, \dots, \underline{x}_n)$ est une base de E alors tout vecteur \underline{x} se décompose de manière unique

$$\underline{x} = \sum_{i=1}^{i=n} \alpha_i \underline{x}_i.$$

Alors la linéarité de A permet de calculer $A(\underline{x})$ à partir des vecteurs $A(\underline{x}_i)$:

$$A(\underline{x}) = \sum_{i=1}^{i=n} \alpha_i A(\underline{x}_i).$$

1.3.1 Opérateurs linéaires

Une application linéaire d'un espace vectoriel E à image dans E est appelée *opérateur linéaire* sur E (on lit aussi *endomorphisme* sur E).

Définition 1.3.3. Si A est un opérateur linéaire sur E qui établit une bijection de E vers E , on dit que A est inversible. On peut alors définir l'application réciproque A^{-1} de E vers E par la relation

$$A^{-1}(A\underline{x}) = \underline{x}.$$

On a également $A(A^{-1}\underline{x}) = \underline{x}$ et A^{-1} est linéaire.

Théorème 1.3.4. Soit A un opérateur linéaire sur un espace vectoriel de dimension finie E , alors les conditions suivantes sont équivalentes

- A est injectif
- A est surjectif
- A est bijectif

Preuve en annexe 1.9.1.

Signalons tout de même que ce théorème n'est plus vrai si E est de dimension infinie : pensez à l'espace vectoriel des suites numériques et l'application qui enlève le premier élément de la suite³ qui est une application surjective de l'espace des suites vers lui même mais qui n'est pas injective puisque le noyau n'est pas réduit à 0, c'est à dire la suite infinie de 0..

Encore quelques définitions :

Définition 1.3.5. On note $\mathcal{L}(E, F)$ l'ensemble des applications linéaires d'un espace vectoriel E vers un espace vectoriel F .

Dans le cas des opérateurs linéaires, au lieu de $\mathcal{L}(E, E)$ on écrit plus simplement $\mathcal{L}(E)$.

3. c'est l'opérateur *shift* en anglais

Si A_1 et A_2 sont dans $\mathcal{L}(E, F)$ et si α_1 et α_2 sont deux scalaires, on définit l'application $\alpha_1 A_1 + \alpha_2 A_2$ par

$$(\alpha_1 A_1 + \alpha_2 A_2)\underline{x} = \alpha_1 A_1 \underline{x} + \alpha_2 A_2 \underline{x}, \quad \underline{x} \in E.$$

Naturellement on a $(\alpha_1 A_1 + \alpha_2 A_2) \in \mathcal{L}(E, F)$.

Définition 1.3.6. Si E, F et Z sont trois espaces vectoriels et si $A \in \mathcal{L}(E, F)$, $B \in \mathcal{L}(F, Z)$, on définit leur produit BA comme étant la composée de A et B :

$$(BA)\underline{x} = B(A\underline{x}) \quad \underline{x} \in E.$$

On a alors $BA \in \mathcal{L}(E, Z)$.

Soulignons que même si $E = F = Z$, le produit d'opérateurs linéaires ne commute pas, i.e.

$$AB \neq BA \quad \text{en général.}$$

Définition 1.3.7. Pour $A \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$, on peut définir une norme de A par le sup de tous les vecteurs $\|A\underline{x}\|$ quand \underline{x} parcourt la boule unité de \mathbb{R}^n centrée en 0 :

$$\|A\| = \sup_{\|\underline{x}\|=1} \|A\underline{x}\|.$$

Par linéarité, on a alors toujours l'inégalité :

$$\|A\underline{x}\| \leq \|A\| \|\underline{x}\|.$$

Théorème 1.3.8. — Si $A \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ alors $\|A\| < +\infty$ et A est une application uniformément continue sur \mathbb{R}^n .

— Si $A, B \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ et si α est un scalaire alors

$$\|A + B\| \leq \|A\| + \|B\| \quad \|\alpha A\| = |\alpha| \|A\|.$$

— Si $A \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ et $B \in \mathcal{L}(\mathbb{R}^m, \mathbb{R}^p)$ alors

$$\|BA\| \leq \|A\| \|B\|.$$

Preuve en annexe 1.9.2.

1.3.2 Quelques exemples d'espaces vectoriels et d'applications linéaires

1. On note $\mathbb{R}_n[X]$ l'ensemble des polynômes réels de degré $\leq n$.

2. L'espace des fonctions définies sur un intervalle de \mathbb{R} (c'est un e.v. de dimension infinie)
3. L'ensemble des applications linéaires de \mathbb{R}^n dans \mathbb{R}^m .
4. L'espace des matrices $n \times m$ (de dimension mn).
5. L'ensemble des solutions d'une équation différentielle ordinaire homogène linéaire à coefficients constants ou non (dimension finie).
6. L'ensemble des solutions d'une équation aux dérivées partielles linéaires homogène.

Exercice : Pour chacun de ces exemples donner un sous-espace vectoriel.

1.4 Matrice d'une application linéaire

Soient X et Y deux espaces vectoriels de dimension respective n et m .

Soient $(\underline{x}_1, \dots, \underline{x}_n)$ et $(\underline{y}_1, \dots, \underline{y}_m)$ deux bases respectives de X et Y . Toute application linéaire $A \in \mathcal{L}(X, Y)$ détermine un ensemble de coefficients notés a_{ij} tels que

$$A(\underline{x}_j) = \sum_{i=1}^m a_{ij} \underline{y}_i \quad (1 \leq j \leq n). \quad (1.3)$$

Il est usuel et commode de représenter ces coefficients dans un tableau rectangulaire de m lignes et de n colonnes appelé *matrice* $m \times n$ noté $M(A)$ ⁴, on parle alors de *représentation matricielle* de l'application linéaire :

$$M(A) = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}.$$

Remarquons que les coordonnées a_{ij} du vecteur $A(\underline{x}_j)$ apparaissent dans la j^{me} colonne de la matrice $M(A)$. Les vecteurs $A(\underline{x}_j)$ sont donc parfois appelés *vecteurs colonnes* de la matrice $M(A)$.

On appelle *rang* de A , la dimension de l'image de X par A . Il est donc égal à la dimension de l'espace engendré par les vecteurs colonnes de $M(A)$.

A l'aide des coefficients de la matrice on peut déterminer l'image de tout vecteur \underline{x} de X par A . En effet, si $\underline{x} = \sum \alpha_i \underline{x}_i$, on déduit par linéarité de A que

$$A(\underline{x}) = \sum_{i=1}^m \left(\sum_{j=1}^n a_{ij} \alpha_j \right) \underline{y}_i. \quad (1.4)$$

4. On le note également d'autres façon telles que (A) ou encore M ou bien $[A]$, pour être tout à fait rigoureux il faudrait indiquer $\text{Mat}(A, \{\underline{x}_j\}, \{\underline{y}_i\})$; c'est-à-dire indiquer la matrice de l'application linéaire A de la base de départ $\{\underline{x}_j\}$ dans la base d'arrivée $\{\underline{y}_i\}$.

Réciproquement, donnons nous maintenant une matrice $m \times n$ de coefficients a_{ij} , notée $M(A)$. Si on définit l'application A par la formule (1.4), on remarque que $A \in \mathcal{L}(X, Y)$, où nous rappelons que X et Y sont des espaces vectoriels de dimension n et m respectivement. Ainsi :

Théorème 1.4.1. *Il y a une bijection entre $\mathcal{L}(X, Y)$ et l'ensemble des matrices $m \times n$, m lignes, dimension de l'espace d'arrivée, n colonnes, dimension de l'espace de départ.*

Exemple : Plaçons nous dans \mathbb{R}^2 et \mathbb{R}^3 , si relativement à des bases $(\underline{e}_1, \underline{e}_2)$ et $(\underline{f}_1, \underline{f}_2, \underline{f}_3)$, une application linéaire est définie comme

$$\begin{aligned} A(\underline{e}_1) &= \underline{f}_1 + \underline{f}_2 + \underline{f}_3 \\ A(\underline{e}_2) &= \underline{f}_1 - \underline{f}_2. \end{aligned}$$

Alors, la matrice de l'application linéaire A dans les bases considérées s'écrit :

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

Ainsi, si par exemple

$$\underline{x} = \alpha_1 \underline{e}_1 + \alpha_2 \underline{e}_2,$$

alors

$$A(\underline{x}) = (\alpha_1 + \alpha_2) \underline{f}_1 + (\alpha_1 - \alpha_2) \underline{f}_2 + \alpha_1 \underline{f}_3.$$

Ce qu'on note également noter sous la forme :

$$A \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} = \begin{pmatrix} \alpha_1 + \alpha_2 \\ \alpha_1 - \alpha_2 \\ \alpha_1 \end{pmatrix}.$$

1.4.1 Bijection entre $\mathcal{L}(X, \mathbb{R})$ et X

Enchaînons par un résultat fondamental :

Théorème 1.4.2. *Soit X un espace vectoriel réel de dimension finie. Étant donné une base de X , on peut définir le produit scalaire : $\forall \underline{x} = x_i \underline{x}_i \quad \forall \underline{y} = y_i \underline{x}_i$*

$$\underline{x} \cdot \underline{y} = x_i y_i.$$

Ainsi pour tout $\underline{x} \in X$, on définit une forme linéaire sur X . Réciproquement, toute forme linéaire sur X , via sa représentation matricielle, peut être présentée comme un produit scalaire par un vecteur de X .

Il existe donc une bijection entre $\mathcal{L}(X, \mathbb{R})$ et X .

Corollaire 1.4.3. *Dans le cas où $X = \mathbb{R}^n$ le produit scalaire rapporté à la base canonique définit la bijection naturelle entre $\mathcal{L}(\mathbb{R}^n, \mathbb{R})$ et \mathbb{R}^n .*

Cette bijection est essentielle pour bien comprendre le calcul différentiel : Ainsi un réel a peut être identifié à une forme linéaire réelle définie par l'application qui associe à tout réel x , la valeur ax .

De même un vecteur $\underline{a} \in \mathbb{R}^2$ peut être identifié à une forme linéaire réelle définie par l'application qui associe à tout vecteur \underline{x} , la valeur réelle $\underline{a} \cdot \underline{x}$.

1.4.2 Norme d'une application linéaire

Terminons cette section par une proposition permettant d'obtenir une estimation de la norme d'une application linéaire grâce à sa représentation matricielle

Proposition 1.4.4. *Soit a_{ij} les coefficients de la matrice de $A \in \mathcal{L}(X, Y)$ de la base $(\underline{x}_1, \dots, \underline{x}_n)$ de X dans $(\underline{y}_1, \dots, \underline{y}_n)$ de Y . On a la majoration :*

$$\|A\| \leq \left(\sum_{i,j} a_{ij}^2 \right)^{1/2}.$$

Preuve en annexe 1.9.3

1.4.3 Produit de deux matrices et composition de deux applications linéaires

Considérons à présent un troisième espace vectoriel Z de dimension p avec une base $(\underline{z}_1, \dots, \underline{z}_p)$. Si A est définie par (1.3), définissons de la même manière une application linéaire $B \in \mathcal{L}(Y, Z)$ à l'aide de coefficients b_{ki} :

$$B(\underline{y}_i) = \sum_{k=1}^p b_{ki} \underline{z}_k \quad (1 \leq i \leq m). \quad (1.5)$$

On peut alors définir l'application composée $BA \in \mathcal{L}(X, Z)$ par

$$(BA)(\underline{x}_j) = \sum_{k=1}^p c_{kj} \underline{z}_k.$$

Mais comme

$$\begin{aligned}
 (BA)(\underline{x}_j) &= B(A\underline{x}_j) \\
 &= B\left(\sum_{i=1}^m a_{ij}\underline{y}_i\right) \\
 &= \sum_{i=1}^m a_{ij}B(\underline{y}_i) \\
 &= \sum_{i=1}^m a_{ij} \sum_{k=1}^p b_{ki}z_k,
 \end{aligned}$$

on a finalement

$$(BA)(\underline{x}_j) = \sum_{k=1}^p \sum_{i=1}^m b_{ki}a_{ij}z_k,$$

c'est à dire :

$$c_{kj} = \sum_{i=1}^m b_{ki}a_{ij}. \quad (1.6)$$

On dit aussi que la matrice $p \times n$, notée C , de coefficients c_{kj} est le *produit de la matrice A par la matrice B*.

La formule (1.6) donne la règle usuelle du produit de deux matrices.

Terminons par la règle : le produit d'une matrice $n \times m$ par une matrice $m \times p$, donne une matrice $n \times p$.

$$(n \times m)(m \times p) = (n \times p).$$

Si les dimensions des espaces ne rentrent pas dans le cadre de cette règle, le produit n'est pas défini, les dimensions d'espaces étant incompatibles.

1.4.4 Changement de bases et Matrices de passage

Dans la pratique, et notamment en Mécanique, il apparaît souvent judicieux de passer d'un système de coordonnées à un autre suivant les besoins. Il est donc nécessaire de maîtriser les outils permettant ces passages.

Soit X un espace vectoriel de dimension n et soit un vecteur $\underline{x} \in X$ de coordonnées α_i dans la base $(\underline{x}_1, \dots, \underline{x}_n)$. Soit $(\underline{x}'_1, \dots, \underline{x}'_n)$ une autre base de X . Supposons connues les coordonnées des vecteurs \underline{x}'_i dans la base $(\underline{x}_1, \dots, \underline{x}_n)$:

$$\underline{x}'_j = \sum_{i=1}^n p_{ij}\underline{x}_i.$$

On a donc défini une matrice P représentant l'application identité de X , muni de la base $(\underline{x}'_1, \dots, \underline{x}'_n)$, dans X , muni de la base $(\underline{x}_1, \dots, \underline{x}_n)$:

$$Id_X(\underline{x}'_j) = \underline{x}'_j = \sum_{i=1}^n p_{ij} \underline{x}_i.$$

Théorème 1.4.5. *La matrice P ayant pour j^{me} colonne les coordonnées de \underline{x}'_j dans la base $(\underline{x}_1, \dots, \underline{x}_n)$ est inversible. On l'appelle la matrice de passage de la base $(\underline{x}_1, \dots, \underline{x}_n)$ vers la base $(\underline{x}'_1, \dots, \underline{x}'_n)$.*

De plus, si α_i sont les “anciennes” coordonnées de \underline{x} dans $(\underline{x}_1, \dots, \underline{x}_n)$, formant la matrice colonne $(n \times 1)$ $[\underline{x}]_x$, on peut les exprimer en fonction des “nouvelles” coordonnées α'_j de \underline{x} dans $(\underline{x}'_1, \dots, \underline{x}'_n)$, notées également par une matrice colonne $[\underline{x}]_{x'}$ avec la formule :

$$[\underline{x}]_x = P[\underline{x}]_{x'} \iff P^{-1}[\underline{x}]_x = [\underline{x}]_{x'}.$$

En résumé, si

$$\underline{x} = \sum_{j=1}^n \alpha_j \underline{x}_j = \sum_{j=1}^n \alpha'_j \underline{x}'_j, \quad \text{et} \quad \underline{x}'_j = \sum_{i=1}^n p_{ij} \underline{x}_i,$$

alors

$$\alpha_i = \sum_{j=1}^n p_{ij} \alpha'_j.$$

Autrement dit, pour obtenir les “nouvelles” coordonnées en fonction des anciennes, il faut passer par la matrice de passage P^{-1} de la base $(\underline{x}'_1, \dots, \underline{x}'_n)$ vers la base $(\underline{x}_1, \dots, \underline{x}_n)$.

La notation P^{-1} est justifiée par le fait que P^{-1} est nécessairement la matrice inverse de P . En effet, on doit avoir

$$PP^{-1} = P^{-1}P = I_n,$$

où I_n est la matrice identité $n \times n$.

Considérons maintenant une application linéaire $A \in \mathcal{L}(X, Y)$ où X et Y sont toujours des espaces vectoriels de dimension n et m respectivement.

On muni X des bases $(\underline{x}_1, \dots, \underline{x}_n)$ et $(\underline{x}'_1, \dots, \underline{x}'_n)$, avec P la matrice de passage de $(\underline{x}_1, \dots, \underline{x}_n)$ vers $(\underline{x}'_1, \dots, \underline{x}'_n)$.

On muni de même Y des bases $(\underline{y}_1, \dots, \underline{y}_m)$ et $(\underline{y}'_1, \dots, \underline{y}'_m)$, avec Q la matrice de passage de $(\underline{y}_1, \dots, \underline{y}_m)$ vers $(\underline{y}'_1, \dots, \underline{y}'_m)$. Alors, on a la proposition :

Proposition 1.4.6. *Soit M la matrice de A de la base $(\underline{x}_1, \dots, \underline{x}_n)$ de X vers la base $(\underline{y}_1, \dots, \underline{y}_m)$ de Y . Alors la matrice M' de A de la base $(\underline{x}'_1, \dots, \underline{x}'_n)$ de X vers la base $(\underline{y}'_1, \dots, \underline{y}'_m)$ de Y , se décompose :*

$$M' = Q^{-1}MP.$$

Démonstration. En effet, on a

$$[A(\underline{x})]_{y'} = M' [\underline{x}]_{x'}$$

et

$$[A(\underline{x})]_{y'} = Q^{-1} [A(\underline{x})]_y = Q^{-1} M [\underline{x}]_x = Q^{-1} M P [\underline{x}]_{x'}.$$

□

Dans le cas particulier où A est un opérateur linéaire (on dit aussi endomorphisme) sur X , si P est la matrice de passage de la base $(\underline{x}_1, \dots, \underline{x}_n)$ vers la base $(\underline{x}'_1, \dots, \underline{x}'_n)$, alors si M et M' sont les matrices de A dans $(\underline{x}_1, \dots, \underline{x}_n)$ et $(\underline{x}'_1, \dots, \underline{x}'_n)$ respectivement, alors on a

$$M' = P^{-1} M P.$$

1.5 Cas des opérateurs linéaires – Matrices carrés

Dans le cas d'un opérateur linéaire (*i.e.* un endomorphisme) de \mathbb{R}^n , la représentation matricielle est une matrice $n \times n$. Une telle matrice, ayant le même nombre de ligne et de colonne, est dite carrée.

1.5.1 Adjoint d'un opérateur linéaire

Définition 1.5.1. Soit A un opérateur linéaire sur \mathbb{R}^n , l'opérateur adjoint A^* est défini par :

$$\forall \underline{x}, \underline{y} \in \mathbb{R}^n, \quad \langle A^* \underline{x}, \underline{y} \rangle = \langle \underline{x}, A \underline{y} \rangle.$$

Proposition 1.5.2. Soit M la représentation matricielle d'un opérateur linéaire A de \mathbb{R}^n dans une certaine base orthonormée. Alors la représentation matricielle M^* de l'adjoint A^* dans cette même base est égale à la transposée de M .

$$M^* = M^\perp \quad \text{ou encore} \quad M_{ij}^* = M_{ji}.$$

1.5.2 Partie symétrique et antisymétrique d'un opérateur linéaire de \mathbb{R}^n

Définition 1.5.3. La partie symétrique d'un opérateur linéaire A de \mathbb{R}^n est égale à la somme de A et de A^* divisé par deux :

$$A_{sym} = \frac{1}{2} (A + A^*).$$

Définition 1.5.4. La partie antisymétrique d'un opérateur linéaire A de \mathbb{R}^n est égale à la somme de A et de $-A^*$ divisé par deux :

$$A_{antisym} = \frac{1}{2} (A - A^*).$$

Proposition 1.5.5.

$$\begin{aligned} A &= A_{sym} + A_{antisym}, \\ A_{sym}^* &= A_{sym}, \\ A_{antisym}^* &= -A_{antisym}. \end{aligned}$$

Proposition 1.5.6. La partie antisymétrique d'un opérateur linéaire A de \mathbb{R}^n définit de manière unique un vecteur \underline{v} de \mathbb{R}^n tel que

$$\forall \underline{x} \in \mathbb{R}^n, \quad A_{antisym} \underline{x} = \underline{v} \wedge \underline{x}.$$

Par exemple dans \mathbb{R}^3 , si $\underline{v} = (v_1, v_2, v_3)$, on a

$$\underline{v} \wedge \underline{x} = \begin{pmatrix} v_2 x_3 - v_3 x_2 \\ v_3 x_1 - v_1 x_3 \\ v_1 x_2 - v_2 x_1 \end{pmatrix} = \begin{pmatrix} 0 & -v_3 & v_2 \\ v_3 & 0 & -v_1 \\ -v_2 & v_1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

1.6 Déterminant

1.6.1 Déterminant de n vecteurs de \mathbb{R}^n

Le théorème suivant est également une définition :

Théorème 1.6.1. Soit $E = \mathbb{R}^n$, et soit $(\underline{x}_1, \dots, \underline{x}_n)$ une base de E . Il existe une unique n -forme alternée définie sur E^n appelée déterminant telle que

$$\det_{(\underline{x}_i)}(\underline{x}_1, \dots, \underline{x}_n) = 1.$$

E^n désigne le produit d'espace $E \times E \times \dots \times E$ (n fois). Le déterminant étant alterné par définition, le déterminant de n vecteurs de E change de signe si on permute deux vecteurs :

$$\det_{(\underline{x}_i)}(\underline{v}_1, \dots, \underline{v}_j, \dots, \underline{v}_k, \dots, \underline{v}_n) = -\det_{(\underline{x}_i)}(\underline{v}_1, \dots, \underline{v}_k, \dots, \underline{v}_j, \dots, \underline{v}_n).$$

Par conséquent, le déterminant de n vecteurs est nul si deux vecteurs sont égaux :

$$\det_{(\underline{x}_i)}(\underline{v}_1, \dots, \underline{v}_j, \dots, \underline{v}_j, \dots, \underline{v}_n) = 0.$$

Plus généralement, si une famille de n vecteurs est liée, c'est-à-dire que au moins un des vecteurs est une combinaison linéaire des autres, par linéarité le déterminant est nécessairement nul.

Réciproquement, si une famille est libre, son déterminant est une combinaison linéaire non nulle du déterminant de la base. On énonce ainsi un critère pour déterminer si une famille de n vecteurs de E est une base ou non :

Théorème 1.6.2. Une famille $(\underline{v}_1, \dots, \underline{v}_n)$ de n vecteurs de $E = \mathbb{R}^n$, est une base de E si et seulement si

$$\det_{(\underline{x}_i)}(\underline{v}_1, \dots, \underline{v}_n) \neq 0.$$

Dans toute la suite, on définira toujours le déterminant rapporté à la base canonique de E plutôt qu'une base quelconque. Ainsi le déterminant ne sera désigné que par le symbole \det .

1.6.2 Déterminant d'une matrice carré – Déterminant d'un opérateur linéaire

Soit A une matrice $n \times n$, elle est donc constituée de n colonnes de matrices $n \times 1$, chacune d'elles représentant un vecteur de \mathbb{R}^n (ou n'importe quel espace vectoriel de dimension n). Ainsi on peut définir le déterminant d'une matrice par l'intermédiaire de ses vecteurs colonnes :

Définition 1.6.3. Le déterminant d'une matrice carré est défini par le déterminant de ses vecteurs colonnes.

Proposition 1.6.4. Soit I la matrice identité $n \times n$, par définition $\det I = 1$.

Proposition 1.6.5. Soit A une matrice carré $n \times n$,

$$\det A = \det A^t$$

Proposition 1.6.6. Soient A et B deux matrices $n \times n$. Bien que A et B ne permutent pas en général, on a

$$\det(AB) = \det(BA) = \det A \det B.$$

Cette dernière propriété permet de définir d'un opérateur linéaire : Soit f un opérateur linéaire sur E et soit A une représentation matricielle de f dans une base (\underline{x}_i) et B la représentation matricielle de f dans une autre base (\underline{y}_i) . Par conséquent il existe une matrice de passage P de la base (\underline{x}_i) vers la base (\underline{y}_i) . Si bien que

$$\det A = \det(P^{-1}BP) = \det(PP^{-1}B) = \det B.$$

Autrement dit, si on définit le déterminant d'un opérateur linéaire par le déterminant de sa représentation matricielle dans une base, ce déterminant est indépendant du choix de la base.

Définition 1.6.7. Soit f un opérateur linéaire sur E et soit A une représentation matricielle de f dans une base (\underline{x}_i)

$$\det f = \det A.$$

Théorème 1.6.8. Un opérateur linéaire est bijectif ou inversible si et seulement si son déterminant est non nul.

1.6.3 Règles de calcul d'un déterminant – Développement suivant une ligne ou une colonne

Soit A une matrice carré $n \times n$ dont les coefficients sont a_{ij} . On définit des sous-matrices extraites de A utiles pour le calcul d'un déterminant :

Définition 1.6.9. On appelle mineur de a_{ij} dans A , le déterminant de la sous matrice extraite de A ôté de la i -ème ligne et la j -ème colonne. On le note Δ_{ij} .

Définition 1.6.10. On note $\text{Com}A$, la matrice dont les composantes notées A_{ij} sont les cofacteurs de A :

$$A_{ij} = (-1)^{i+j} \Delta_{ij}.$$

On dit que $\text{Com}A$ est la matrice des cofacteurs de A .

Théorème 1.6.11. Soit A une matrice carré $n \times n$ dont les coefficients sont a_{ij} et soient A_{ij} ses cofacteurs.

On a alors le développement du déterminant suivant la i -ème ligne :

$$\det A = \sum_{j=1}^n a_{ij} A_{ij},$$

et le développement suivant la j -ème colonne :

$$\det A = \sum_{i=1}^n a_{ij} A_{ij}.$$

Une conséquence immédiate est le calcul très simple du déterminant d'une matrice triangulaire supérieure (resp. inférieure), *i.e.* une matrice dont les coefficients au dessous (resp. au dessus) de la diagonale sont tous nuls :

Proposition 1.6.12. Soit A une matrice carré triangulaire, son déterminant est le produit de ses coefficients diagonaux.

Proposition 1.6.13. Soit A une matrice 2×2 de coefficients $a_{\alpha\beta}$,

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

Preuve : On vérifie aisément que la forme linéaire ainsi définie est alternée et satisfait à $\det I = 1$ où I désigne la matrice identité. Par définition, c'est le déterminant.

Ainsi pour calculer le déterminant d'une matrice 3×3 on peut développer suivant la 3-ème colonne (par exemple) :

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = (-1)^{1+3} a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} + (-1)^{2+3} a_{23} \begin{vmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{vmatrix} + (-1)^{3+3} a_{33} \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}.$$

On termine par une formule donnant l'inverse d'une matrice :

Théorème 1.6.14. *Soit A une matrice carré $n \times n$ inversible, alors si A^{-1} désigne son inverse on a*

$$A^{-1} = \frac{1}{\det A} \text{Com } A^t.$$

1.7 Éléments de théorie spectrale

1.7.1 Spectre d'un opérateur linéaire

Définition 1.7.1. *Soit A un opérateur linéaire sur \mathbb{R}^n , un scalaire $\lambda \in \mathbb{R}^n$ est une valeur propre de A si il existe un vecteur non nul \underline{x} tel que :*

$$A\underline{x} = \lambda\underline{x}.$$

On dit alors que \underline{x} est un vecteur propre associé à λ .

L'ensemble des valeurs réelles λ pour lesquels l'application $[A - \lambda I]$ est une bijection, est appelé *ensemble résolvant*. Le complémentaire sur \mathbb{R} de l'ensemble résolvant est le *spectre* de A .

Remarque 2. *Il est clair que l'ensemble des valeurs propres est contenu dans le spectre. Cette inclusion est une égalité dans le cas où A opère sur un espace de dimension finie, mais est stricte dans le cas général : Pensez à l'opérateur shift dans l'espace vectoriel des suites numériques de carré sommable : 0 est dans le spectre mais n'est pas une valeur propre !*

Définition 1.7.2. *Soit A un opérateur linéaire sur \mathbb{R}^n , le le déterminant*

$$P_A(X) = \det(A - XI_n),$$

c'est un polynôme de degré n . On l'appelle polynôme caractéristique de A .

Le polynôme caractéristique est utile dans la détermination des valeurs propres d'un opérateur linéaire :

Proposition 1.7.3. *Les racines du polynôme caractéristique d'un opérateur linéaire A sont les valeurs propres de A .*

Remarquons que les valeurs propres étant des racines d'un polynôme, il est tout à fait possible qu'elles n'existent pas toutes dans \mathbb{R} (c'est-à-dire que certaines sont complexes).

1.7.2 Réduction d'un opérateur linéaire

Définition 1.7.4. *On dit qu'un opérateur linéaire sur \mathbb{R}^n est diagonalisable si il existe une base de vecteurs propres. Dans une telle base la représentation matricielle de A est une matrice diagonale composée des valeurs propres de A .*

Proposition 1.7.5. *Soit A un opérateur linéaire sur \mathbb{R}^n , et soient $\underline{v}_1, \underline{v}_2, \dots, \underline{v}_n$, une base de vecteurs propres associées aux valeurs propres λ_i . Soit D la matrice, diagonale, dont tous les coefficients sur la diagonale est constitués des valeurs propres $\lambda_1, \lambda_2, \dots, \lambda_n$ et dont tous les autres coefficients sont nuls. Alors D est la représentation matricielle de A dans la base de vecteurs propres associée.*

$$D = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}$$

Si de plus P est la matrice de passage d'une base $\underline{p}_1, \underline{p}_2, \dots, \underline{p}_n$ vers la base de vecteurs propres $\underline{v}_1, \underline{v}_2, \dots, \underline{v}_n$, alors la représentation matricielle M de A dans la base $\underline{p}_1, \underline{p}_2, \dots, \underline{p}_n$ est donné par :

$$M = PDP^{-1}.$$

En général, un opérateur linéaire (resp. une matrice carrée) n'est pas diagonalisable, il suffit de penser aux cas de valeurs propres complexes. Il est cependant possible d'isoler quelques cas où il est possible de diagonaliser :

Proposition 1.7.6. *Soit A un opérateur linéaire sur E un espace vectoriel réel de dimension n . Si les valeurs propres sont toutes distinctes et réelles, alors A est diagonalisable.*

Naturellement la proposition précédente n'est pas du tout une condition nécessaire, est une application peut posséder des valeur propres multiples et être diagonalisable : pensez à l'opérateur identité.

Terminons par deux résultats très importants :

Théorème 1.7.7. *Il est toujours possible de trigonaliser un opérateur linéaire dans \mathbb{C} (resp. une matrice carrée).*

Il s'en suit le théorème

Théorème 1.7.8. *Cayley-Hamilton – Soit A un opérateur linéaire sur \mathbb{R}^n et P_A son polynôme caractéristique, alors $P_A(A) = 0$.*

En fait, la réduction d'un endomorphisme (un opérateur linéaire), s'applique également lorsqu'on a pas diagonalisation : c'est la décomposition spectrale.

Théorème 1.7.9. *Soit A un opérateur linéaire sur \mathbb{R}^n . On suppose que le polynôme minimal de A est scindé (pas de racines complexes). Alors on peut décomposer $A = D + N$ où D est diagonale et N est nilpotente. De plus N et D commutent. Cette décomposition est unique.*

1.8 Cas des opérateurs auto-adjoint ou matrices symétriques

Le cas des matrices symétriques est très utile en mécanique et en mathématiques pour l'ingénieur en général, puisque la plupart des objets rencontrés comme les formes d'énergie de déformation élastique (tenseur des contraintes, des déformations) pourront être représentées (dans leur version discrétisée) par de telles matrices, voir la méthode des éléments-finis par exemple.

Commençons par un résultat fondamental concernant les matrices symétriques :

Théorème 1.8.1. *Soit A un opérateur linéaire autoadjoint (i.e, symétrique) de \mathbb{R}^n alors toutes ses valeurs propres sont réelles et A est diagonalisable. De plus il existe une base orthonormée de vecteurs propres.*

Démonstration. Soit λ une valeur propre de A et soit \underline{v} un vecteur propre associé. On a alors que

$$(A\underline{v}, A\underline{v}) = \lambda^2(\underline{v}, \underline{v})$$

Autrement dit : $\lambda^2 \geq 0$, ce qui signifie que λ est nécessairement réel. On déduit que toutes les valeurs propres sont réelles et donc que A est diagonalisable.

Soit deux vecteurs propres \underline{v}_1 et \underline{v}_2 associés respectivement aux valeurs propres $\lambda_1 \neq \lambda_2$. Alors, on a

$$\begin{aligned} (\underline{v}_1, \underline{v}_2) &= \frac{1}{\lambda_1} (A\underline{v}_1, \underline{v}_2) \\ &= \frac{1}{\lambda_1} (\underline{v}_1, A^T \underline{v}_2) \\ &= \frac{1}{\lambda_1} (\underline{v}_1, A\underline{v}_2) \\ &= \frac{\lambda_2}{\lambda_1} (\underline{v}_1, \underline{v}_2) \end{aligned}$$

Si bien que puisque $\lambda_1 \neq \lambda_2$, alors nécessairement $(\underline{v}_1, \underline{v}_2) = 0$, c'est à dire qu'ils sont orthogonaux. \square

Définition 1.8.2. Soit A un opérateur linéaire sur \mathbb{R}^n .

On dit que A est positive si $\forall \underline{x} \in \mathbb{R}^n, A(\underline{x}) \cdot \underline{x} \geq 0$.

On dit que A est définie positive si il existe $c > 0$ telle que $\forall \underline{x} \in \mathbb{R}^n, A(\underline{x}) \cdot \underline{x} \geq c \|\underline{x}\|$.

Ainsi, soit $\underline{v}_1, \dots, \underline{v}_n$ une base orthonormée de vecteurs propres associés aux valeurs propres λ_i d'une matrice symétrique A . Alors pour tout $\underline{v} \in \mathbb{R}^n$, il existe une décomposition unique

$$\underline{v} = \alpha_i \underline{v}_i$$

de sorte que

$$(A\underline{v}, \underline{v}) = \lambda_i \alpha_i^2, \quad \text{et} \quad (\underline{v}, \underline{v}) = \alpha_i \alpha_i = \|\underline{v}\|^2$$

si bien que, si λ_{min} et λ_{max} désignent respectivement la plus petite et la plus grande des valeurs propres de A ,

$$\lambda_{min} \|\underline{v}\|^2 \leq (A\underline{v}, \underline{v}) \leq \lambda_{max} \|\underline{v}\|^2 \quad (1.7)$$

d'où

Théorème 1.8.3. Soit A un opérateur linéaire sur \mathbb{R}^n , A est définie positive si et seulement si toutes ses valeurs propres sont strictement positives.

Des algorithmes de résolution de systèmes linéaires pour de telles matrices sont particulièrement efficaces à l'exemple de la méthode du gradient conjugué basé sur une méthode de descente.

Terminons par un résultat d'optimisation quadratique fondamental :

Théorème 1.8.4. Soient $A \in \mathbb{R}^{n \times n}$ une matrice symétrique, $b \in \mathbb{R}^n$ et soit la fonctionnelle

$$J(\underline{v}) = \frac{1}{2}(A\underline{v}, \underline{v}) - (b, \underline{v}).$$

Alors le problème de minimisation J sur $\mathbb{R}^{n \times n}$ est équivalent au problème linéaire

$$A\underline{v} = b.$$

Si de plus A est définie positive, alors il existe une unique solution à ces problèmes.

1.9 Annexe : quelques preuves

1.9.1 Preuve du théorème 1.3.4

Supposons que X soit de dimension égale à n et soit $(\underline{x}_1, \dots, \underline{x}_n)$ une base de X . D'après la linéarité de A , l'image de A est engendré par le système de n vecteurs $S = (A\underline{x}_1, \dots, A\underline{x}_n)$, on en déduit que S engendre X (autrement dit que A est surjectif) si S est un système libre.

Montrons que si A est injectif alors A est surjectif.

Si A est injectif, cela signifie que $A\underline{x} = 0 \implies \underline{x} = 0$. Or si

$$\sum_{i=1}^{i=n} \alpha_i \underline{x}_i = 0$$

on a par linéarité de A

$$A \left(\sum_{i=1}^{i=n} \alpha_i \underline{x}_i \right) = 0 \implies \sum_{i=1}^{i=n} \alpha_i A(\underline{x}_i) = 0$$

mais comme les \underline{x}_i sont une famille libre, on obtient que les coefficients α_i sont nécessairement tous nuls. Autrement dit que le système S est libre.

Réciproquement, montrons que si A est surjectif alors A est injectif.

Si A est surjectif cela signifie que le système S est libre. Alors si $\underline{x} = \sum \alpha_i \underline{x}_i$ annule l'opérateur $A : A\underline{x} = 0$, alors par linéarité on a

$$0 = A(\underline{x}) = A \left(\sum_{i=1}^{i=n} \alpha_i \underline{x}_i \right) = \sum_{i=1}^{i=n} \alpha_i A(\underline{x}_i)$$

mais puisque S est une famille libre, cela signifie que les coefficients α_i sont tous nuls, autrement dit que \underline{x} est nécessairement nul, ce qui prouve que A est injectif.

1.9.2 Preuve du théorème 1.3.8

— Soit \underline{x} un vecteur tel que $\|\underline{x}\| = 1$. \underline{x} se décompose dans la base canonique sous la forme

$$\underline{x} = \sum_{i=1}^{i=n} \alpha_i \underline{e}_i.$$

Puisque $\|\underline{x}\| = \left(\sum_{i=1}^{i=n} \alpha_i^2 \right)^{1/2} = 1$. On a alors nécessairement $|\alpha_i| \leq 1$, d'où

$$\|A\underline{x}\| = \left\| \sum_{i=1}^{i=n} \alpha_i A(\underline{e}_i) \right\| \leq \sum_{i=1}^{i=n} |\alpha_i| \|A\underline{e}_i\| \leq \sum_{i=1}^{i=n} \|A\underline{e}_i\| \leq n < +\infty$$

Puisque nous avons pour tout $\underline{x}, \underline{y} \in \mathbb{R}^n$, $\|A\underline{x} - A\underline{y}\| \leq \|A\| \|\underline{x} - \underline{y}\|$, on en déduit l'uniforme continuité.

— De la linéarité de A , nous obtenons facilement

$$\|(A + B)\underline{x}\| \leq \|A(\underline{x}) + B(\underline{x})\| \leq \|A(\underline{x})\| + \|B(\underline{x})\| \leq (\|A\| + \|B\|) \|\underline{x}\|$$

de même

$$\|\alpha A(\underline{x})\| = |\alpha| \|A\underline{x}\|.$$

— Enfin, on a

$$\|(BA)\underline{x}\| \leq \|B(A\underline{x})\| \leq \|B\| \|A\underline{x}\| \leq \|B\| \|A\| \|\underline{x}\|$$

1.9.3 Preuve de la proposition 1.4.4

Soit $\underline{x} \in X$, de coordonnées α_i dans la base $(\underline{x}_1, \dots, \underline{x}_n)$, d'après l'inégalité de Schwarz on a

$$\|A(\underline{x})\|^2 = \sum_{i=1}^m \left(\sum_{j=1}^n a_{ij} \alpha_j \right)^2 \leq \sum_{i=1}^m \left(\sum_{j=1}^n a_{ij}^2 \sum_{k=1}^n \alpha_k^2 \right) \leq \sum_{i,j} a_{ij}^2 \|x\|^2.$$

CHAPITRE 2

Calcul différentiel

On connaît bien depuis le secondaire la notion de fonction réelle définie sur un intervalle. L'étude d'une fonction consiste généralement dans l'étude de sa continuité, de sa dérivabilité ou de son taux de croissance. Dans cette section, on étend ces notions aux *fonctions de plusieurs variables*, c'est-à-dire des applications définies sur des parties (ouvertes) de \mathbb{R}^m à valeurs dans \mathbb{R}^n .

2.1 Éléments de topologie métrique

Le titre de cette section est trompeur, il ne s'agit ici que de rappeler quelques définitions usuelles en topologie métrique, il n'y a aucune propriété énoncée. Il s'agit pour l'essentiel de donner un sens à la notion de convergence pour des objets autre que des scalaires...

Rappelons que dans \mathbb{R}^n , un *ouvert* O est une partie de \mathbb{R}^n telle que pour tout élément x_0 de O , la boule (fermée) de centre x_0 et de rayon $\varepsilon > 0$,

$$B(x_0, \varepsilon) = \{x \in \mathbb{R}^n, \text{ tel que } |x_0 - x| \leq \varepsilon\},$$

soit incluse dans O pour tout ε suffisamment petit.

Par ailleurs un ensemble sera dit *fermé* si son ensemble complémentaire est ouvert. Autrement dit, une partie de \mathbb{R}^n est ouverte si elle ne contient aucun point de sa frontière, tandis qu'une partie de \mathbb{R}^n est fermée si elle contient tous les points de sa frontière. On remarquera, qu'au contraire d'une porte, *un ensemble peut être ni ouvert ni fermé*.

La raison pour laquelle le calcul différentiel se fait sur des ensembles ouverts est, disons, technique. Il s'agit pour l'essentiel d'éliminer des petites subtilités pouvant survenir sur le bord des ensembles considérés (par exemple la notion de continuité à gauche ou à droite dans le cas d'une fonction d'une variable réelle). Il n'y a donc pas lieu d'insister trop sur cet aspect de la théorie dans un premier temps.

Dans toute la suite, la norme d'un vecteur \underline{x} de \mathbb{R}^n est la norme euclidienne usuelle, si

$$\underline{x} = \sum_{i=1}^n x_i \underline{e}_i$$

les $\{\underline{e}_i\}$ désignant la base canonique de \mathbb{R}^n , alors

$$\|\underline{x}\| = \sqrt{\sum_{i=1}^n (x_i)^2}.$$

Il est facile de vérifier que cette norme satisfait aux conditions définissant une norme :

$$\|\underline{x} + \underline{y}\| \leq \|\underline{x}\| + \|\underline{y}\|,$$

$$\|\lambda \underline{x}\| = \lambda \|\underline{x}\|,$$

$$\|\underline{x}\| = 0 \Leftrightarrow \underline{x} = 0.$$

Définition 2.1.1. On dira d'un élément (ou vecteur) \underline{x} de \mathbb{R}^n qu'il converge vers un autre élément \underline{x}_0 de \mathbb{R}^n si la norme de la différence $\underline{x}_0 - \underline{x}$ tend vers 0 :

$$\underline{x} \longrightarrow \underline{x}_0 \Leftrightarrow \|\underline{x}_0 - \underline{x}\| \longrightarrow 0.$$

Définition 2.1.2. Une suite \underline{x}_n de \mathbb{R}^n est une suite de Cauchy si et seulement si

$$\lim_{n,m \rightarrow +\infty} \|\underline{x}_n - \underline{x}_m\| = 0.$$

Par ailleurs, on est amené à utiliser souvent la notion de *reste négligeable*, notés petit o et grand O dans le sens où

$$\lim_{\|\underline{h}\| \rightarrow 0} O(\underline{h}) = 0,$$

$$\lim_{\|\underline{h}\| \rightarrow 0} \frac{o(\underline{h})}{\|\underline{h}\|} = 0.$$

Remarque 3. Les notions de convergences et de suite de Cauchy, présentées ici, se généralisent sans modification à tout espace vectoriel de dimension finie ou non, muni d'une norme.

2.2 Fonctions continues de plusieurs variables

Revenons tout d'abord au cas où les fonctions sont définies sur \mathbb{R}^n (ou sur des ouverts de \mathbb{R}^n) dans le cas familier où $n = 1$ afin d'y jeter un nouveau regard.

Si f est une fonction à valeurs réelles définie sur un intervalle ouvert $]a, b[$ de \mathbb{R} et si $x_0 \in]a, b[$ alors

Définition 2.2.1. On dit que f est continue en x_0 si

$$\lim_{x \rightarrow x_0} f(x) = f(x_0).$$

Remarquons ici que la notion x tends vers x_0 naturelle dans \mathbb{R} ne l'est plus si x désigne des objets abstraits, d'où la nécessité d'une topologie.

On va voir qu'il est possible de reformuler cette définition de façon "légèrement" plus abstraite afin de la généraliser : On peut dire que f est continue en x_0 si $\forall \varepsilon > 0, \exists \eta > 0$ tel que

$$|x_0 - x| \leq \eta \Rightarrow |f(x_0) - f(x)| \leq \varepsilon$$

autrement dit f est continue en x_0 si

$$f(x) = f(x_0) + O(|x - x_0|).$$

Pour plus d'informations sur la notion d'ensemble ouvert, on peut consulter les manuels de topologie élémentaire ; il est conseillé de se limiter à la topologie métrique dans un premier temps. Il faut savoir que même si la topologie est une science trop abstraite pour ce cours, elle est une base fondatrice du calcul différentiel et donc de toutes les mathématiques appliquées faisant intervenir des équations aux dérivées partielles, et en premier lieu la mécanique. En fait, on utilise la topologie sans le savoir, un peu comme M. Jourdain.

2.3 Fonction différentiable – Application dérivée

2.3.1 Cas des fonctions réelles définie sur un intervalle réel

De façon classique on la définition pour une fonction d'une variable réelle à valeur dans \mathbb{R} :

On dit que f est dérivable en x_0 s'il existe un réel $f'(x_0)$ défini par

$$f'(x_0) = \lim_{h \rightarrow 0} \frac{f(x_0 + h) - f(x_0)}{h}.$$

Cependant, comme pour la continuité, on va adopter une autre présentation de la définition afin de la généraliser

Définition 2.3.1. On dit que f est dérivable en x_0 s'il existe un réel $f'(x_0)$ tel que

$$f(x_0 + h) - f(x_0) = f'(x_0)h + o(h)$$

le "reste" $o(h)$ étant *négligeable* dans le sens que

$$\lim_{h \rightarrow 0} \frac{o(h)}{h} = 0.$$

On remarque que la quantité $f(x_0 + h) - f(x_0)$ s'exprime comme la somme d'une application linéaire qui à h associe $f'(x_0)h$ et d'un reste négligeable. On peut donc considérer la dérivée de \underline{f} au point x_0 *non seulement comme un réel mais également comme un opérateur linéaire sur \mathbb{R}* qui à h associe $f'(x_0)h$ (en fait cette considération a été établie à partir de la bijection naturelle entre \mathbb{R} et $\mathcal{L}(\mathbb{R})$).

C'est cette définition qu'il faudra retenir. En général, les étudiants ont beaucoup de mal à oublier la définition qu'ils ont apprise dans le Secondaire valable uniquement pour les fonctions réelles d'une variable réelle, et ils l'utilisent sans scrupule même s'il n'a plus de sens comme c'est le cas pour tout les autres cas de fonctions. Il va sans dire que ce genre d'erreurs coûte quelques points aux examens ;-)

2.3.2 Cas général

Considérons maintenant une fonction \underline{f} définie sur un intervalle $]a, b[$ de \mathbb{R} à valeurs dans \mathbb{R}^m . Rappelons que cela signifie que $\underline{f}(x) = (f_1(x), \dots, f_m(x))$ ou encore, si on représente une base canonique de \mathbb{R}^m par $\{\underline{u}_1, \dots, \underline{u}_m\}$:

$$\underline{f}(x) = f_1 \underline{u}_1 + \dots + f_m \underline{u}_m.$$

En un point $x_0 \in]a, b[$, la dérivée de \underline{f} en x_0 , noté $\underline{f}'(x_0)$, est défini comme *le vecteur de \mathbb{R}^m* (si il existe) tel que

$$\underline{f}'(x_0) = \lim_{h \rightarrow 0} \frac{\underline{f}(x_0 + h) - \underline{f}(x_0)}{h} \quad \text{ou plutôt} \quad \underline{f}(x_0 + h) - \underline{f}(x_0) = \underline{f}'(x_0)h + o(h).$$

Soulignons encore que l'égalité précédente est *vectorielle* :

$$\underline{f}'(x_0) = f_1(x_0)\underline{u}_1 + \dots + f_m(x_0)\underline{u}_m,$$

et donc peut également s'écrire sous forme matricielle dans la base canonique :

$$\begin{pmatrix} f_1(x_0 + h) \\ \vdots \\ f_m(x_0 + h) \end{pmatrix} - \begin{pmatrix} f_1(x_0) \\ \vdots \\ f_m(x_0) \end{pmatrix} = \begin{pmatrix} f'_1(x_0) \\ \vdots \\ f'_m(x_0) \end{pmatrix} + o(h),$$

ou encore sous forme vectorielle :

Définition 2.3.2. \underline{f} est dérivable en x_0 si et seulement si il existe un vecteur $\underline{f}'(x_0)$ tel que :

$$\underline{f}(x_0 + h) - \underline{f}(x_0) = \underline{f}'(x_0)h + \underline{o}(h),$$

La dérivée de \underline{f} en x_0 apparaît non seulement comme un vecteur de \mathbb{R}^m mais également comme une application linéaire de \mathbb{R} dans \mathbb{R}^m qui associe à tout réel h le vecteur $\underline{f}'(x_0)h$ (il s'agit là de la bijection naturelle entre \mathbb{R}^m et $\mathcal{L}(\mathbb{R}^m, \mathbb{R})$).

On va maintenant généraliser ces notions.

Définition 2.3.3. Soit \underline{f} une application (on ne dit plus fonction ?) d'un ouvert Ω de $E = \mathbb{R}^n$ dans $F = \mathbb{R}^m$. On dit que \underline{f} est différentiable en un point $\underline{x} \in \Omega$, $\underline{x} = (x_1, \dots, x_n)$, si il existe une application linéaire de \mathbb{R}^n dans \mathbb{R}^m , notée $\underline{f}'(\underline{x}_0)$ telle que

$$\underline{f}(\underline{x}_0 + \underline{h}) = \underline{f}(\underline{x}_0) + \underline{f}'(\underline{x}_0)(\underline{h}) + o(h) = 0$$

où le "reste" $\underline{o}(h)$ est cette fois-ci négligeable au sens que

$$\lim_{h \rightarrow 0} \frac{\|\underline{o}(h)\|}{\|h\|} = 0.$$

On dit alors que $\underline{f}'(\underline{x}_0)$ est la différentielle (on ne dit plus dérivée ?) de \underline{f} en \underline{x} et on également parfois

$$\underline{f}'(\underline{x}) = \underline{d}\underline{f}_{\underline{x}}.$$

Ainsi, si \underline{f} est différentiable sur un ouvert Ω , alors pour tout $\underline{x} \in \Omega$,

$$\underline{f}'(\underline{x}) \in \mathcal{L}(E, F) = \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m),$$

i.e. la différentielle de \underline{f} en \underline{x} est une application linéaire de \mathbb{R}^n (tout entier) dans \mathbb{R}^m mais on peut remarquer que \underline{f}' est une application de Ω vers $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$.

Remarque 4. Pour des raisons, disons pédagogiques, nous nous sommes limités dans le cadre de ces rappels aux fonctions définies sur des parties de \mathbb{R}^n à valeurs dans \mathbb{R}^m . Mais, tout ce qui suit est également valable si on remplace les \mathbb{R}^n et \mathbb{R}^m par des objets "plus abstraits" que sont les espaces de Banach (i.e. des espaces normés complets) qu'on pourrait noter également E et F . Les énoncés et les démonstrations demeurant strictement identiques.

Remarque 5. On parle également d'application tangente pour désigner la différentielle en un point.

Théorème 2.3.4. Soit \underline{f} une application d'un ouvert Ω de \mathbb{R}^n dans \mathbb{R}^m . Si \underline{f} est différentiable en $\underline{x} \in \Omega$, alors sa différentielle est unique.

Démonstration. Supposons qu'on ait, avec $t \in \mathbb{R}$,

$$f(x + th) = f(x) + df_1 th + o(th) \quad \text{et} \quad f(x + th) = f(x) + df_2 th + o(th)$$

Alors on obtient en effectuant la différence :

$$(df_1 - df_2)th = o(th) = t^2 o(h)$$

d'où pour tout t

$$(df_1 - df_2)h = to(h)$$

c'est à dire

$$(df_1 - df_2)h = 0$$

□

Quelques remarques

1. Si f est différentiable en \underline{x} alors f est nécessairement continue en \underline{x} (exercice).
2. La différentielle notée $f'(\underline{x})$ ou \underline{df}_x est appelée parfois *dérivée totale* de f en \underline{x} , afin de la distinguer des *dérivées partielles*.
3. Si $A \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ et si $x \in \mathbb{R}^n$ alors

$$A'(\underline{x}) = A$$

Autrement dit, une application linéaire est sa propre différentielle. Remarquez que \underline{x} apparaît explicitement dans le premier membre mais pas dans le second : c'est parce que la fonction dérivée d'une application linéaire est constante (indépendance par rapport au "point" où on dérive).

Théorème 2.3.5. *Règle de composition* – Soit \underline{f} une application d'un ouvert $\mathcal{O} \subset \mathbb{R}^n$ vers \mathbb{R}^m différentiable en un point \underline{x}_0 de \mathcal{O} et \underline{g} une application d'un ouvert de \mathbb{R}^m contenant $\underline{f}(\mathcal{O})$ vers \mathbb{R}^k différentiable en $\underline{f}(\underline{x}_0)$. Alors l'application, composé de \underline{f} par \underline{g} , \underline{F} de $\mathcal{O} \subset \mathbb{R}^n$ vers \mathbb{R}^k définie par

$$\underline{F}(\underline{x}) = \underline{g} \circ \underline{f} = \underline{g}(\underline{f}(\underline{x}))$$

est différentiable en \underline{x}_0 et

$$\underline{F}'(\underline{x}_0) = \underline{g}'(\underline{f}(\underline{x}_0))\underline{f}'(\underline{x}_0).$$

Remarquer que le second membre est le produit de deux applications linéaires : $\underline{g}'(\underline{f}(\underline{x}_0)) \in \mathcal{L}(\mathbb{R}^m, \mathbb{R}^k)$ et $\underline{f}'(\underline{x}_0) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$; on a bien que $\underline{F}'(\underline{x}_0) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^k)$.

Démonstration.

$$\begin{aligned} \underline{F}(\underline{x}_0 + \underline{h}) &= \underline{g}(\underline{f}(\underline{x}_0 + \underline{h})) \\ &= \underline{g}(\underline{f}(\underline{x}_0) + \underline{f}'(\underline{x}_0)\underline{h} + o(\underline{h})) \\ &= \underline{F}(\underline{x}_0) + \underline{g}'(\underline{f}(\underline{x}_0))\underline{f}'(\underline{x}_0)\underline{h} + o(\underline{f}'(\underline{x}_0)\underline{h} + \underline{h}) + o(\underline{h}) \\ &= \underline{F}(\underline{x}_0) + \underline{g}'(\underline{f}(\underline{x}_0))\underline{f}'(\underline{x}_0)\underline{h} + o(\underline{h}) \\ &= \underline{F}(\underline{x}_0) + \underline{F}'(\underline{x}_0)\underline{h} + o(\underline{h}) \end{aligned}$$

□

2.3.3 Matrice Jacobienne et dérivées partielles

Puisque la différentielle de \underline{f} en un point est une application linéaire de \mathbb{R}^n vers \mathbb{R}^m , on peut la représenter sous forme matricielle. Plus précisément par une matrice $n \times m$. Soient $(\underline{e}_1, \dots, \underline{e}_n)$ et $(\underline{u}_1, \dots, \underline{u}_m)$ les bases canoniques des espaces \mathbb{R}^n et \mathbb{R}^m respectivement :

$$\begin{aligned} \underline{x} &= x_1\underline{e}_1 + \dots + x_n\underline{e}_n \\ \underline{f}(\underline{x}) &= f_1(\underline{x})\underline{u}_1 + \dots + f_m(\underline{x})\underline{u}_m. \end{aligned}$$

Définition 2.3.6. La représentation matricielle de la différentielle de \underline{f} dans ces bases canoniques est la matrice Jacobienne. Usuellement, elle est notée $J(\underline{f})$ ou encore $J_{\underline{f}}$.

Définition 2.3.7. Soit x_j la j -ème variable de \mathbb{R}^n , la dérivée partielle de \underline{f} par rapport à x_j est la dérivée de \underline{f} par rapport à x_j , les autres variables étant considérées fixes. On la note :

$$\frac{\partial \underline{f}}{\partial x_j} = \frac{\partial f_1}{\partial x_j}\underline{u}_1 + \dots + \frac{\partial f_m}{\partial x_j}\underline{u}_m.$$

On peut remarquer qu'on a encore

$$\frac{\partial \underline{f}}{\partial x_j}(\underline{x}) = \lim_{t \rightarrow 0} \frac{\underline{f}(\underline{x} + t\underline{e}_j) - \underline{f}(\underline{x})}{t}, \quad \text{ou plutôt} \quad \underline{f}(\underline{x} + t\underline{e}_j) - \underline{f}(\underline{x}) = \frac{\partial \underline{f}}{\partial x_j}(\underline{x})t + o(t).$$

Proposition 2.3.8. Si \underline{f} est différentiable en \underline{x} et si $\underline{a} = (a_1, \dots, a_n) \in \mathbb{R}^n$, alors toutes les dérivées partielles existent et

$$\underline{f}'(\underline{x})\underline{a} = a_1 \frac{\partial \underline{f}(\underline{x})}{\partial x_1} + \dots + a_n \frac{\partial \underline{f}(\underline{x})}{\partial x_n}.$$

Attention : la réciproque n'est pas vraie : On peut construire une fonction dont les dérivées partielles existent et sont continues mais qui n'est pas elle-même différentiable.

On note aussi parfois :

$$d\underline{f} = \frac{\partial \underline{f}}{\partial x_1} dx_1 + \dots + \frac{\partial \underline{f}}{\partial x_n} dx_n.$$

Théorème 2.3.9. *La composante de la i -ème ligne et de la j -ème colonne de la matrice Jacobienne est la dérivée partielle de la i -ème composante de \underline{f} par rapport à la j -ème variable. On note :*

$$J_{ij} = \frac{\partial f_i}{\partial x_j}.$$

ou encore

$$(J) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \dots & \frac{\partial f_1}{\partial x_j} & \dots & \frac{\partial f_1}{\partial x_n} \\ \vdots & & \vdots & & \vdots \\ \frac{\partial f_i}{\partial x_1} & \dots & \frac{\partial f_i}{\partial x_j} & \dots & \frac{\partial f_i}{\partial x_n} \\ \vdots & & \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1} & \dots & \frac{\partial f_m}{\partial x_j} & \dots & \frac{\partial f_m}{\partial x_n} \end{pmatrix}.$$

Autrement dit : la j -ème colonne de la matrice Jacobienne est la dérivée partielle de \underline{f} par rapport à la j -ème variable x_j .

2.4 Gradient, divergence, rotationnel

2.4.1 Gradient

La matrice Jacobienne étant relative aux bases canoniques, il est important, afin de faire du calcul dans des bases autres que les canoniques (telles que les coordonnées cylindriques...), de se rappeler l'objet intrinsèque dont elle est le représentant. Il s'agit du *gradient* de \underline{f} , qui n'est autre que l'application dérivée :

$$\underline{f}'(\underline{x})\underline{a} = \nabla \underline{f}(\underline{x}) \underline{a},$$

mais qui est aussi parfois défini par

$$d\underline{f} = \nabla \underline{f} d\underline{x}.$$

ou encore si les \underline{e}_j désignent une base fixe et si les \underline{x} se décomposent dans cette base :

$$\frac{\partial \underline{f}}{\partial x_j}(\underline{x}) = \nabla \underline{f}(\underline{x}) (\underline{e}_j).$$

Ainsi la représentation matricielle du gradient de \underline{f} dans les bases canoniques est la matrice Jacobienne.

2.4.2 Divergence

Dans le cas particulier où $n = m$, la matrice Jacobienne est une matrice carré. On peut alors définir de nouveaux opérateurs voir le paragraphe 4.4 sur la divergence dans le chapitre sur le théorème de Stokes :

Proposition 2.4.1. *L'opérateur divergence est la trace du gradient :*

$$\operatorname{div} \underline{f} = \operatorname{tr}(\nabla \underline{f}).$$

2.4.3 Rotationnel

Toujours dans le cas où $n = m$,

Proposition 2.4.2. *Le rotationnel de \underline{f} , noté $\overrightarrow{\operatorname{rot}} \underline{f}$, est le vecteur associé à la partie antisymétrique du gradient de \underline{f} . Dans \mathbb{R}^3 , on a :*

$$[\overrightarrow{\operatorname{rot}} \underline{u}] = \begin{bmatrix} u_{3,2} - u_{2,3} \\ u_{1,3} - u_{3,1} \\ u_{2,1} - u_{1,2} \end{bmatrix}.$$

2.4.4 Quelques remarques et propriétés des opérateurs différentiels

Remarque 6. *Dans la littérature mathématique, le gradient d'une fonction scalaire à plusieurs variables est souvent représentée par un vecteur, ainsi si $f(x) = f(x_1, \dots, x_n) \in \mathbb{R}$, le gradient est défini par*

$$\begin{aligned} f'(x)h &= (\nabla f(x), h) \\ &= \nabla f(x) \cdot h \\ &= [\nabla f(x)]^\perp[h] \end{aligned}$$

si bien qu'il est noté

$$\nabla f(x) = \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{pmatrix}$$

Cependant, cette notation n'a plus de sens lorsque la fonction f est vectorielle.

Remarque 7. *L'opérateur divergence est parfois noté (sur wikipédia, dans de nombreux livres, surtout en physique) comme un produit scalaire par un pseudo vecteur ∇ , mais c'est un abus d'écriture : cette notation n'est pas une définition et ne peut pas s'étendre à d'autres objets que des fonctions vectorielles, notamment la divergence d'un tenseur si importante en Mécanique.*

Terminons ce paragraphe par quelques propriétés sur la divergence et le rotationnel :

Proposition 2.4.3. *Pour tout champ de vecteur défini sur $\Omega \subset \mathbb{R}^3$ à valeur dans \mathbb{R}^3 :*

$$\overrightarrow{\operatorname{rot}} \nabla f = 0 \operatorname{div} \overrightarrow{\operatorname{rot}} \underline{u} = 0 \tag{2.1}$$

$$(\overrightarrow{\operatorname{rot}} \underline{u}) \wedge \underline{u} = \underline{u} \cdot \nabla \underline{u} - \frac{1}{2} \nabla (u^2) \tag{2.2}$$

$$\overrightarrow{\operatorname{rot}} \overrightarrow{\operatorname{rot}} \underline{u} = \nabla \operatorname{div} \underline{u} - \Delta \underline{u}. \tag{2.3}$$

Théorème 2.4.4. Si un champ de vecteur \underline{u} , défini de $\Omega \subset \mathbb{R}^n$ à valeur dans \mathbb{R}^n , est tel que $\overrightarrow{\text{rot}} \underline{u} = 0$ alors il existe une fonction scalaire f définie sur Ω telle que

$$\underline{u} = \nabla f .$$

On dit que \underline{u} dérive d'un potentiel.

De façon analogue, on a

Théorème 2.4.5. Si un champ de vecteur \underline{u} , défini de $\Omega \subset \mathbb{R}^n$ à valeur dans \mathbb{R}^n , est tel que $\text{div} \underline{u} = 0$ alors il existe un champs \underline{v} tel que

$$\underline{u} = \overrightarrow{\text{rot}} \underline{v} .$$

2.5 Dérivation le long d'une courbe – Dérivation partielle dans une direction

Soit \mathcal{C} une courbe de \mathbb{R}^n , c'est à dire l'image dans \mathbb{R}^n d'une application \underline{c} définie sur un intervalle $]a, b[\subset \mathbb{R}$. Pour fixer les idées on peut prendre $n = 2$ ou encore $n = 3$ (courbe dans le plan, l'espace...).

Soit \mathcal{O} un ouvert de \mathbb{R}^n contenant la courbe \mathcal{C} et soit \underline{f} comme précédemment, une application de \mathbb{R}^n vers \mathbb{R}^m .

Définissons l'application composée \underline{g} de $]a, b[$ vers \mathbb{R}^m :

$$\underline{g}(t) = \underline{c} \circ \underline{f}(t) = \underline{f}(\underline{c}(t))$$

D'après le théorème de composition, on a

$$\underline{g}'(t) = \nabla \underline{f}(\underline{c}(t)) \underline{c}'(t),$$

où $\underline{c}'(t) = c'_1(t)\underline{u}_1 + \dots + c'_m(t)\underline{u}_m$.

$\underline{g}'(t)$ représente la dérivée de \underline{f} le long de la courbe \mathcal{C} . Cela nous permet également de définir la dérivée dans une direction \underline{y} :

En prenant $\underline{c} = \underline{x} + t\underline{y}$, on a $\underline{c}'(t) = \underline{y}$. D'où, d'après la proposition précédente :

$$\underline{g}'(0) = \nabla \underline{f}(\underline{x}) \underline{y} .$$

ou encore

$$\nabla \underline{f}(\underline{x}) \underline{y} = \lim_{t \rightarrow 0} \left[\frac{\underline{f}(\underline{x} + t\underline{y}) - \underline{f}(\underline{x})}{t} \right] .$$

Cette dernière limite étant la dérivée partielle de \underline{f} suivant le vecteur unitaire \underline{y} au point \underline{x} .

2.6 Fonction dérivée

Soit \underline{f} une application d'un ouvert Ω de \mathbb{R}^n dans \mathbb{R}^m différentiable en tout $\underline{x} \in \Omega$. Alors l'application dérivée de \underline{f} est l'application qui à tout $\underline{x} \in \Omega$ associe $\underline{f}'(\underline{x})$, la différentielle de \underline{f} en \underline{x} .

2.7 Dérivée seconde

Soit \underline{f} une application d'un ouvert $\mathcal{O} \subset E$ vers F où comme précédemment E et F désignent respectivement \mathbb{R}^n et \mathbb{R}^m (ou encore tout espaces de Banach).

On a vu que la différentielle (ou application dérivée) de \underline{f} en un point \underline{x} de \mathcal{O} est une application linéaire de E vers F :

$$\underline{f}'(\underline{x}) \in \mathcal{L}(E, F) \quad \underline{f}' : \mathcal{O} \subset E \longmapsto \mathcal{L}(E, F).$$

Ainsi, en réapplicando la définition de différentielle à \underline{f}' , on a :

$$\underline{f}''(\underline{x}) \in \mathcal{L}(E, \mathcal{L}(E, F)) \simeq \mathcal{L}(E \times E, F).$$

$\underline{f}''(\underline{x})$ apparaît donc comme une application bilinéaire.

Puisque $E = \mathbb{R}^n$, \underline{f}' se décompose en dérivées partielles :

$$\underline{f}'(\underline{x}).(\underline{h}) = \underline{f}'(\underline{x}).(h_1, \dots, h_n) = \sum_{j=1}^n \frac{\partial \underline{f}}{\partial x_j}.h_j$$

Il en est de même pour l'application dérivée seconde :

$$\underline{f}''(\underline{x}).(\underline{h}) = \sum_{j=1}^n \frac{\partial \underline{f}'}{\partial x_j}.h_j$$

Ainsi,

$$\underline{f}''(\underline{x}).(\underline{h}).(\underline{k}) = \sum_{j=1}^n \frac{\partial \underline{f}'(\underline{x})}{\partial x_j}.h_j.(\underline{k})$$

et finalement

$$\underline{f}''(\underline{x}).(\underline{h}).(\underline{k}) = \sum_{i,j=1}^n \frac{\partial^2 \underline{f}(\underline{x})}{\partial x_i \partial x_j}.h_j.k_i \in F.$$

Théorème 2.7.1. Si \underline{f} une application d'un ouvert $\mathcal{O} \subset E$ vers F est une application deux fois différentiable en \underline{x} alors la dérivée seconde est une application bilinéaire symétrique :

$$\forall \underline{h}, \underline{k} \in E \quad \underline{f}''(\underline{x}).(\underline{h}).(\underline{k}) = \underline{f}''(\underline{x}).(\underline{k}).(\underline{h}).$$

C'est un résultat non-trivial (excellent exercice en calcul différentiel) mais qui donne immédiatement le

Corollaire 2.7.2. *Théorème de Schwarz – Si f une application d'un ouvert $\mathcal{O} \subset E$ vers F est une application deux fois différentiable en \underline{x} alors*

$$\frac{\partial^2 f(\underline{x})}{\partial x_i \partial x_j} = \frac{\partial^2 f(\underline{x})}{\partial x_j \partial x_i}.$$

Ces résultats s'étendent à toutes les dérivées successives.

2.8 Principaux théorèmes sur les fonctions de plusieurs variables

Dans cette section, on énonce divers résultats théoriques fondamentaux en calcul différentiel. Cependant ces résultats ne sont pas nécessaire au niveau de la Licence de Mécanique.

Définition 2.8.1. *On dit qu'une application différentiable f d'un ouvert \mathcal{O} un ouvert de \mathbb{R}^n vers \mathbb{R}^m est continûment différentiable sur \mathcal{O} si f' est une application continue de \mathcal{O} vers $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$. On dit alors que f est de classe \mathcal{C}^1 , ce qu'on écrit $f \in \mathcal{C}^1(\mathcal{O})$.*

Théorème 2.8.2. *Soit f une application d'un ouvert \mathcal{O} de \mathbb{R}^n vers \mathbb{R}^m . Alors $f \in \mathcal{C}^1(\mathcal{O})$ si et seulement si toutes les dérivées partielles existent et sont continues sur \mathcal{O} .*

Théorème 2.8.3. *Théorème de la Moyenne – Soit f une fonction continue et dérivable sur un intervalle $[a, b]$. Si la fonction dérivée f' est continue, alors il existe $c \in [a, b]$ telle que $f(b) - f(a) = (b - a)f'(c)$.*

Lemme 2.8.4. *Théorème du point fixe – Soit φ une application de E dans F , deux espaces de Banach. Si φ est contractante, alors φ possède un unique point fixe.*

Cette partie est à compléter

2.9 Formule de Taylor

Théorème 2.9.1. *Soit f une fonction k -fois différentiable en $x_0 \in \Omega \subset \mathbb{R}^n$ à valeur dans \mathbb{R}^m , alors pour tout $h \in \mathbb{R}^n$:*

$$f(x_0 + h) = f(x_0) + f'(x_0)h + \frac{1}{2}f''(x_0)h.h + \cdots + \frac{1}{k!}f^{(k)}(x_0)(h)^k + o(h^k)$$

où $f^{(k)}(x_0)(h)^k$ signifie $f^{(k)}(x_0) \underbrace{h \dots h}_{k \text{ fois}}$.

Démonstration. Il suffit d'appliquer la formule de Taylor à la fonction $g(t) = f(x_0 + th)$ en $t=0$, et montrer que

$$g^{(k)}(0) = f^{(k)}(x_0)(h)^k.$$

□

2.10 Coordonnées cylindriques et sphériques

Nous terminons ce chapitre sur le calcul différentiel par des expressions particulièrement utiles en Mathématiques appliquées, les expressions des différents opérateurs différentiels en coordonnées cylindriques et en coordonnées sphériques.

2.10.1 Expressions en coordonnées Cylindriques

Plaçons dans $E = \mathbb{R}^3$ et désignons par $(\underline{x}_1, \underline{x}_2, \underline{x}_3)$ vecteur de la base canonique de \mathbb{R}^3 et désignons par O le point $(0,0,0)$. Tout point M de E peut être défini par son vecteur position \underline{OM} , dont on donne les coordonnées dans la base canonique :

$$\underline{OM} = x_1 \underline{x}_1 + x_2 \underline{x}_2 + x_3 \underline{x}_3,$$

notation que l'on préfère à :

$$\underline{OM} = x \underline{x}_1 + y \underline{x}_2 + z \underline{x}_3.$$

On définit les coordonnées cylindriques par

$$r = \sqrt{x_1^2 + x_2^2}, \quad \theta = \arctan \frac{x_2}{x_1}, \quad x_3 = x_3.$$

Alors le vecteur position se ré écrit :

$$\underline{OM} = r \underline{e}_r + x_3 \underline{x}_3,$$

où

$$\underline{e}_r = \cos \theta \underline{x}_1 + \sin \theta \underline{x}_2.$$

En associant de plus le vecteur

$$\underline{e}_\theta = \underline{x}_3 \wedge \underline{e}_r,$$

on aura défini une base $(\underline{e}_r, \underline{e}_\theta, \underline{x}_3)$ orthonormée de E , appelé base base cylindrique. Les coordonnées (r, θ, x_3) sont les coordonnées cylindriques de M .

2.10.1.1 Cas d'une fonction à valeur réelle

Considérons maintenant une fonction f définie sur une partie de \mathbb{R}^3 , adoptons la notation abusive :

$$f(x_1, x_2, x_3) = f(r, \theta, x_3).$$

Nous notons également le produit scalaire de deux vecteurs (colonnes) indifféremment des trois manières suivantes :

$$\begin{aligned} \underline{u} \cdot \underline{v} \\ [\underline{u}]^\perp [\underline{v}] \\ (\underline{u}, \underline{v}) \end{aligned}$$

Rappelons nous que nous avons identifié

$$\frac{\partial f}{\partial x_1}(x_1, x_2, x_3) = \lim_{h \rightarrow 0} \frac{f(x_1 + h, x_2, x_3) - f(x_1, x_2, x_3)}{h} = \nabla f \cdot \underline{x}_1,$$

puisque la variation de $(x_1 + h, x_2, x_3)$ à (x_1, x_2, x_3) est $h\underline{x}_1$.

Le problème avec les coordonnées cylindriques est que si la variation de $(r + \delta r, \theta, x_3)$ à (r, θ, x_3) est bien de $\delta r \underline{e}_r$, ce qui entraîne

$$\frac{\partial f}{\partial r}(r, \theta, x_3) = \lim_{h \rightarrow 0} \frac{f(r + h, \theta, x_3) - f(r, \theta, x_3)}{h} = \nabla f \cdot \underline{e}_r,$$

en revanche la variation de $(r, \theta + \delta \theta, x_3)$ à (r, θ, x_3) n'est malheureusement pas $\delta \theta \underline{e}_\theta$, ce qui signifie que

$$\frac{\partial f}{\partial \theta}(r, \theta, x_3) = \lim_{h \rightarrow 0} \frac{f(r, \theta + h, x_3) - f(r, \theta, x_3)}{h} \neq \nabla f \cdot \underline{e}_\theta.$$

Plus précisément, si M est un point de coordonnées cylindrique (r, θ, x_3) , et M' un point de coordonnées $(r, \theta + \delta \theta, x_3)$ alors la variation \underline{MM}' est en fait :

$$\underline{MM}' = r \delta \theta \underline{e}_\theta.$$

D'où nous avons :

$$\frac{\partial f}{\partial \theta}(r, \theta, x_3) = \nabla f \cdot r \underline{e}_\theta,$$

autrement dit les composantes cylindriques de ∇f sont :

$$\nabla f = \frac{\partial f}{\partial r} \underline{e}_r + \frac{1}{r} \frac{\partial f}{\partial \theta} \underline{e}_\theta + \frac{\partial f}{\partial x_3} \underline{x}_3.$$

ou encore

$$[\nabla f]_{\text{cylindrique}} = \begin{pmatrix} \frac{\partial f}{\partial r} \\ \frac{1}{r} \frac{\partial f}{\partial \theta} \\ \frac{\partial f}{\partial x_3} \end{pmatrix} = \begin{pmatrix} f_{,r} \\ \frac{1}{r} f_{,\theta} \\ f_{,3} \end{pmatrix}.$$

2.10.1.2 Cas d'une fonction vectorielle

Considérons maintenant une fonction \underline{f} définie sur une partie de \mathbb{R}^3 , et à valeur dans \mathbb{R}^3 adoptons également la notation abusive :

$$\underline{f}(x_1, x_2, x_3) = \underline{f}(r, \theta, x_3).$$

On décomposera \underline{f} dans les coordonnées cylindriques par

$$\underline{f}(r, \theta, x_3) = f_r \underline{e}_r + f_\theta \underline{e}_\theta + f_3 \underline{x}_3.$$

D'après leur définition, il est facile de voir que

$$\begin{aligned} \frac{\partial \underline{e}_r}{\partial r} &= 0, & \frac{\partial \underline{e}_\theta}{\partial r} &= 0, & \frac{\partial \underline{x}_3}{\partial r} &= 0, \\ \frac{\partial \underline{e}_r}{\partial x_3} &= 0, & \frac{\partial \underline{e}_\theta}{\partial x_3} &= 0, & \frac{\partial \underline{x}_3}{\partial x_3} &= 0, \\ \frac{\partial \underline{e}_r}{\partial \theta} &= \underline{e}_\theta, & \frac{\partial \underline{e}_\theta}{\partial \theta} &= -\underline{e}_r, & \frac{\partial \underline{x}_3}{\partial \theta} &= 0. \end{aligned}$$

Ainsi,

$$\begin{aligned} \frac{\partial \underline{f}}{\partial r} &= \nabla \underline{f} \underline{e}_r = f_{r,r} \underline{e}_r + f_{\theta,r} \underline{e}_\theta + f_{3,r} \underline{x}_3, \\ \frac{\partial \underline{f}}{\partial \theta} &= \nabla \underline{f} r \underline{e}_\theta = (f_{r,\theta} - f_\theta) \underline{e}_r + (f_{\theta,\theta} + f_r) \underline{e}_\theta + f_{3,\theta} \underline{x}_3, \\ \frac{\partial \underline{f}}{\partial x_3} &= \nabla \underline{f} \underline{x}_3 = f_{r,3} \underline{e}_r + f_{\theta,3} \underline{e}_\theta + f_{3,3} \underline{x}_3. \end{aligned}$$

On en déduit alors l'expression du gradient d'une fonction vectorielle en composantes cylindriques :

$$\nabla \underline{f} = \begin{pmatrix} f_{r,r} & \frac{1}{r}(f_{r,\theta} - f_\theta) & f_{r,3} \\ f_{\theta,r} & \frac{1}{r}(f_{\theta,\theta} + f_r) & f_{\theta,3} \\ f_{3,r} & \frac{1}{r}f_{3,\theta} & f_{3,3} \end{pmatrix}.$$

La divergence étant la trace du gradient, on a :

$$\operatorname{div} \underline{f} = f_{r,r} + \frac{1}{r}(f_{\theta,\theta} + f_r) + f_{3,3}$$

D'où le Laplacien d'une fonction scalaire en coordonnées cylindriques

$$\Delta f = f_{,rr} + \frac{1}{r}(\frac{1}{r}f_{,\theta\theta} + f_{,r}) + f_{,33}.$$

Le rotationnel étant le vecteur associé à la partie antisymétrique du gradient, on a

$$\overrightarrow{\operatorname{rot}} \underline{f} = \begin{pmatrix} \frac{1}{r}f_{3,\theta} - f_{\theta,3} \\ f_{r,3} - f_{3,r} \\ f_{\theta,r} - \frac{1}{r}(f_{r,\theta} - f_\theta) \end{pmatrix}.$$

2.10.2 Expressions en coordonnées sphériques

C'est la même chose qu'en cylindrique, sauf qu'il y a encore plus de calcul à faire : - ((

Si f est une fonction scalaire :

$$\nabla f = \left(f_r, \frac{1}{r} f_\theta, \frac{1}{r \sin \theta} f_\phi \right)$$

si \underline{f} est une fonction vectorielle :

$$\nabla \underline{f} = \begin{pmatrix} f_{r,r} & \frac{1}{r}(f_{r,\theta} - f_\theta) & \frac{1}{r \sin \theta}(f_{r,\phi} + f_\phi) \\ f_{\theta,r} & \frac{1}{r}(f_{\theta,\theta} + f_r) & \frac{1}{r \sin \theta} f_{\theta,\phi} - \frac{\cot \theta}{r} f_\phi \\ f_{\phi,r} & \frac{1}{r} f_{\phi,\theta} & \frac{1}{r \sin \theta} f_{\phi,\phi} + \frac{f_r + \cot \theta f_\theta}{r} \end{pmatrix}.$$

A compléter ... par le lecteur ; -p

2.11 Extrema d'une fonctions différentiable

Le calcul des variations et les conditions d'optimalité mériterait un chapitre entier : à défaut de l'avoir rédigé, signalons le minimum syndical :

Théorème 2.11.1. Soit f une fonction deux fois différentiable dans $\Omega \subset \mathbb{R}^n$ à valeur dans \mathbb{R} , f possède un minimum local en x_0 , i.e. il existe $r > 0$ tel que

$$f(x_0) \leq f(x) \quad \forall x \in B(x_0, r) \subset \Omega,$$

si et seulement si :

1. $f'(x_0) = 0$,
2. $f''(x_0) \geq 0$.

Démonstration. Soit $x \in B(x_0, r)$, alors pour tout $t \in [0, 1]$, $x_0 + t(x - x_0) \in B(x_0, r)$. Si bien que

$$\begin{aligned} f(x_0) &\leq f(x_0 + t(x - x_0)) \\ &\leq f(x_0) + t f'(x_0)(x - x_0) + o(t) \\ \implies & \\ 0 &\leq t f'(x_0)(x - x_0) + o(t) \\ \implies & \\ 0 &\leq f'(x_0)(x - x_0) \end{aligned}$$

De même en prenant $t \in [-1, 0]$, on a de même

$$0 \geq f'(x_0)(x - x_0)$$

d'où la condition d'Euler

$$f'(x_0) = 0.$$

Reprenons, nous avons

$$\begin{aligned} f(x_0) &\leq f(x_0) + tf'(x_0)(x - x_0) + \frac{t^2}{2}f''(x_0)(x - x_0)^2 + o(t^2) \\ \implies \\ 0 &\leq f''(x_0)(x - x_0)^2 \end{aligned}$$

d'où

$$f''(x_0) \geq 0$$

La réciproque est immédiate. □

2.11.1 Dérivation d'une fonctionnelle d'énergie

En mécanique et bien souvent en science de l'ingénieur, nous rencontrons des fonctionnelles d'énergies, c'est à dire des fonctions définies sur un espace de Banach E et à valeur dans \mathbb{R} . Ce sont souvent des fonctionnelles quadratiques, c'est à dire basé sur un produit scalaire.

$$J(\underline{x}) = \frac{1}{2}(A\underline{x}, \underline{x}) - (b, \underline{x}) + c$$

où A peut représenter un opérateur borné ou non, auto-adjoint ou non.

On peut voir que J est dérivable, pour tout $\underline{h} \in E$:

$$\begin{aligned} J(\underline{x} + \underline{h}) &= \frac{1}{2}(A\underline{x} + \underline{h}, \underline{x} + \underline{h}) - (b, \underline{x} + \underline{h}) + c \\ &= \frac{1}{2}(A\underline{x}, \underline{x}) - (b, \underline{x}) + c + \frac{1}{2}[(A\underline{x}, \underline{h}) + (\underline{h}, A\underline{x})] + \frac{1}{2}(A\underline{h}, \underline{h}) \\ &= J(\underline{x}) + \frac{1}{2}([A + A^\top]\underline{x} - b, \underline{h}) + \frac{1}{2}(A\underline{h}, \underline{h}) \end{aligned}$$

d'où on déduit

$$J'(\underline{x}) = \frac{A + A^\top}{2}\underline{x} - b$$

et même

$$J''(\underline{x}) = A.$$

CHAPITRE 3

Intégrale de Riemann

On se propose dans cette section de rappeler la définitions et les principales propriétés de l'intégrale de Riemann. Ce sont des résultats et définitions qui sont connus depuis le Secondaire, mais une mise au point sur cette théorie est utile en vue du théorème de Stokes ou de la théorie de Lebesgue. On notera qu'il n'existe pas à proprement parler d'une définition de l'intégrale, il y en existe plusieurs : intégrales de Riemann, de Stieljes, de Lebesgue. On parle plutôt de *notion d'intégrale*.

3.1 Définition de l'intégrale de Riemann

Soit $[a, b]$ un intervalle de \mathbb{R} et soit $\{x_0, \dots, x_{n+1}\}$ une subdivision ordonnée de $[a, b]$:

$$a = x_0 \leq x_1 \leq \dots \leq x_{n+1} = b$$

On supposera de plus que $\max_k(x_{k+1} - x_k)$ tends vers 0 lorsque n tends vers l'infini.

On pose pour toute fonction f définie sur $[a, b]$:

$$\bar{I}_n(f) = \sum_{k=0}^n \sup_{x \in [x_k, x_{k+1}]} f(x)(x_{k+1} - x_k)$$

$$\underline{I}_n(f) = \sum_{k=0}^n \inf_{x \in [x_k, x_{k+1}]} f(x)(x_{k+1} - x_k)$$

Définition 3.1.1. Soit f une fonction définie sur $[a, b]$, on dit qu'elle est Riemann-intégrable si pour toute subdivision ordonnée,

$$\lim_{n \rightarrow \infty} \bar{I}_n(f) = \lim_{n \rightarrow \infty} \underline{I}_n(f)$$

On note alors l'intégrale de f sur $[a, b]$ par

$$\int_a^b f(x) dx = \lim_{n \rightarrow \infty} \bar{I}_n(f) = \lim_{n \rightarrow \infty} \underline{I}_n(f).$$

Théorème 3.1.2. Une condition suffisante pour qu'une fonction soit Riemann-intégrable est qu'elle soit continue, ou continue par morceaux.

3.2 Principales propriétés

Proposition 3.2.1. Si f et g sont Riemann-intégrables alors $f + g$ l'est également.

Proposition 3.2.2. Si f et g sont Riemann-intégrables alors fg l'est également.

Proposition 3.2.3. Si $f(x) \leq g(x) \forall x \in [a, b]$ alors

$$\int_a^b f(x) dx \leq \int_a^b g(x) dx.$$

Proposition 3.2.4. Relation de Chasles – Si f est Riemann-intégrable sur $I \subset \mathbb{R}$,

$$\int_a^b f(x) dx = \int_a^c f(x) dx + \int_c^b f(x) dx \quad \forall a, b, c \in I$$

Proposition 3.2.5. Si f et g sont Riemann-intégrables sur $[a, b]$ et que c est une constante, alors

$$\begin{aligned} \int_a^b cf(x) dx &= c \int_a^b f(x) dx \\ \int_a^b (f(x) + g(x)) dx &= \int_a^b f(x) dx + \int_a^b g(x) dx \end{aligned}$$

On retiendra en particulier que l'opérateur intégrale est une forme linéaire sur l'ensemble des fonctions Riemann-intégrables.

Proposition 3.2.6. Si $|f|$ est Riemann-intégrable sur $[a, b]$ alors,

$$\left| \int_a^b f(x) dx \right| \leq \int_a^b |f(x)| dx$$

On retiendra tout particulièrement le résultat fondamental :

Théorème 3.2.7. Changement de variable – Soit f une fonction Riemann-intégrable sur $[a, b]$ et φ une fonction dérivable définie sur $[A, B]$ et telle que $\varphi(A) = a$ et $\varphi(B) = b$

$$\int_a^b f(x) dx = \int_A^B f(\varphi(y)) \varphi'(y) dy$$

3.3 Intégration et dérivation

Proposition 3.3.1. *soit f une fonction Riemann-intégrable et soit F la fonction définie par*

$$F(x) = \int_a^x f(t)dt$$

Alors, F est une fonction continue et dérivable : on a de plus $F'(x) = f(x)$.

Théorème 3.3.2. *Théorème fondamental – Soit f une fonction Riemann-intégrable sur $[a, b]$, si $F'(x) = f(x)$, alors*

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

Corollaire 3.3.3. *Intégration par partie – Soient u et v deux fonctions dérivables sur $[a, b]$.*

$$u(b)v(b) - u(a)v(a) = \int_a^b u'(x)v(x)dx + \int_a^b v'(x)u(x)dx$$

3.4 Formule intégrale de Taylor

Théorème 3.4.1.

$$f(x_0+h) = f(x_0) + f'(x_0)h + \frac{1}{2}f''(x_0)h.h + \dots + \frac{1}{n}f^{(n)}(x_0)(h)^n + \int_0^1 \frac{(1-t)^n}{n!} f^{(n+1)}(x_0+th)(h)^{n+1}dt$$

3.5 Intégrale sur un contour

Soit C une courbe de \mathbb{R}^n définie par une carte (\underline{c}, I) où $I = [a, b]$ est un intervalle de \mathbb{R} et \underline{c} est une application de I dans \mathbb{R}^n .

Soit f une fonction définie sur un domaine $\Omega \subset \mathbb{R}^n$ contenant la courbe C .

Définition 3.5.1. *L'intégrale de f le long de la courbe (\underline{c}, I) (la carte définissant l'orientation) est*

$$\int_C f = \int_a^b f(\underline{c}(t)) |\underline{c}'(t)| dt$$

3.6 Intégrale sur un pavé de \mathbb{R}^n .

Soit $\Omega = [a_1, b_1] \times \cdots \times [a_n, b_n]$ un pavé de \mathbb{R}^n et soit f une fonction définie et continue sur ω . L'intégrale de f sur Ω est définie par

$$\int_{\Omega} f(x) dx = \int_{a_1}^{b_1} \int_{\dots} \int_{a_n}^{b_n} f(x) dx_1 \dots dx_n = \int_{a_1}^{b_1} \left(\dots \left(\int_{a_n}^{b_n} f(x) dx_n \right) \dots \right) dx_1$$

Cette intégrale est bien définie puisque f est continue pour chaque variable x_i . De plus, cette intégrale est indépendante de l'ordre d'intégration choisie).

3.7 Intégrale multiple – Formule de Jacobi

On généralise naturellement la définition de l'intégrale sur un pavé à tout domaine se décomposant en une réunion finies et disjointe de pavés. Plus généralement pour domaine $\omega \subset \mathbb{R}^n$ on peut trouver une suite de domaine ω_n qui sont chacun des décompositions finies de pavés disjoints de \mathbb{R}^n . Alors,

$$\int_{\Omega} f(x) dx = \lim_{n \rightarrow \infty} \int_{\Omega_n} f(x) dx$$

Par ailleurs si un domaine Ω est l'image d'un pavé D par une application T :

$$T(D) = \Omega,$$

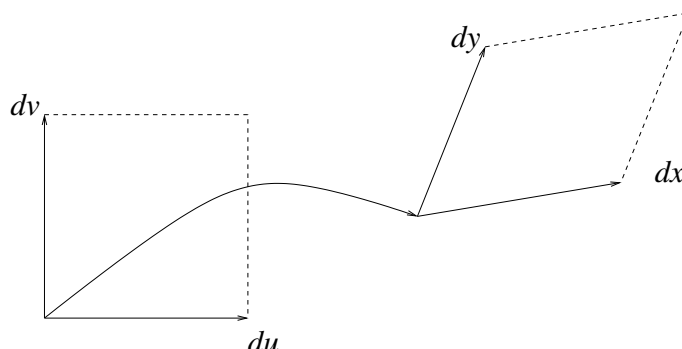
alors on peut montrer que le résultat suivant, généralisation aux intégrales multiples de la formule de changement de variable (3.2.7) :

Théorème 3.7.1. *Formule de Jacobi – Formule de changement de variable.*

$$\int_{\Omega} f(x) dx = \int_D f(T(u)) |J_T(u)| du$$

où $|J_T(u)|$ est la Jacobienne (déterminant du gradient ou Jacobien) de l'application $T(u) = x$.

Démonstration. Pseudo-preuve dans \mathbb{R}^2 : un élément d'aire $dudy$ se transforme en un élément $dx dy$ par le changement de variable T .



l'aire du parallélogramme $dx dy$ peut être calculée par la formule

$$\begin{aligned} dx dy &= |dx \wedge dy| \\ &= \left| \begin{bmatrix} x, u \\ y, u \end{bmatrix} du \wedge \begin{bmatrix} x, v \\ y, v \end{bmatrix} dv \right| \\ &= |x, u y, v - y, u x, v| du dv \\ &= |J_T(u)| du dv. \end{aligned}$$

□

Exemples :

1. Considérons le cylindre $\Omega = \{(x, y, z) \in \mathbb{R}^3 / x^2 + y^2 - 1 \geq 0 \text{ et } 0 \leq z \leq 1\}$, alors Ω est l'image de $D = [0, 1] \times [0, 2\pi] \times [0, 1]$ en posant

$$T(r, \theta, z) = (r \cos \theta, r \sin \theta, z)$$

si bien que

$$J_T = \frac{\partial(x, y, z)}{\partial(r, \theta, z)} = \begin{pmatrix} \cos \theta & -r \sin \theta & 0 \\ \sin \theta & r \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

d'où $|J_T(u)| = r$. On a ainsi :

$$\int_{\Omega} f(x, y, z) dx dy dz = \int_D f(r, \theta, z) r dr d\theta dz.$$

2. Si $\Omega = \{(x, y, z) \in \mathbb{R}^3 / xyz \geq 0, x + y + 1 \leq 1\}$, alors Ω est l'image de $D = [0, 1]^3$ en posant

$$T(u, v, w) = (u(1-v), uv(1-w), uvw)$$

si bien que

$$J_T = \frac{\partial(x, y, z)}{\partial(u, v, w)} = \begin{pmatrix} 1-v & -u & 0 \\ v(1-w) & u(1-w) & -uw \\ vw & uw & uv \end{pmatrix}$$

d'où $|J_T(u)| = u^2 v$.

CHAPITRE 4

Théorème de Stokes et Formule de la divergence

On introduit ici un objet mathématique particulier : les formes différentielles. L'étudiant non-mathématicien peut ignorer les définitions et la forme générale du théorème de Stokes. On pourra s'attarder uniquement sur les différentes formes du théorème de Stokes qui apparaît comme une généralisation des formules d'intégration par partie et ce indépendamment de la définition d'intégrale choisie.

On regardera en particulier les théorèmes de la divergence définissant la notion de divergence et les formules de Green, fondamentales en Mécanique et en mathématiques appliquées, puisque ces notions régissent les principales lois en Mécanique (équations d'équilibre en solides, équation de Navier-Stokes en fluides) mais partout en Mécanique appliquée où les formulations variationnelles (principe des travaux virtuels) interviennent avec toutes les conséquences dans la résolution numérique par la méthode des éléments-finis, voir le chapitre 12 sur l'analyse fonctionnelle.

4.1 Formes différentielles

Définition 4.1.1. Une k -surface Ω dans \mathbb{R}^n , est l'image par un difféomorphisme T d'un domaine $D \in \mathbb{R}^k$ à valeur dans \mathbb{R}^n . On dit que la k -surface est de classe C^m , si l'application T est de classe C^m .

On dit aussi que la k -surface $\Omega \in \mathbb{R}^n$ est une sous-variété de dimension k plongée

dans \mathbb{R}^n .

Définition 4.1.2. *On note*

$$\omega = \sum a_{i_1, \dots, i_k} dx_{i_1} \wedge \dots \wedge dx_{i_k} \quad (4.1)$$

une k -forme différentielle ou forme différentielle d'ordre k , la forme linéaire qui à toute k -surface $\Omega \in \mathbb{R}^n$ de classe C^1 , image par T de $D \in \mathbb{R}^k$, fait correspondre la quantité :

$$\int_{\Omega} \omega = \int_D \sum a_{i_1, \dots, i_k}(T(u)) |J_T(u)| du$$

où

$$|J_T(u)| = \det \frac{\partial(x_{i_1}, \dots, x_{i_k})}{\partial(u_1, \dots, u_k)}$$

désigne le Jacobien de l'application $(u_1, \dots, u_k) \longrightarrow (T_{i_1}(u), \dots, T_{i_k}(u))$.

Remarque 8. *Convention : une forme différentielle d'ordre 0 est une fonction.*

Remarque 9. *La notation (4.1) est unique si les x_{i_j} sont numérotés et pris dans un ordre croissants.*

On remarque que cette définition reprend simplement la formule de Jacobi de changement de variable dans les intégrale multiples, si bien que cette définition est indépendante du choix de D et de T .

Cette définition permet de généraliser la notion d'intégrale sur une courbe définie dans la section précédente. En effet, considérons un contour γ de \mathbb{R}^2 défini par

$$\gamma = \{(x, y) \in \mathbb{R}^2 \mid (x, y) = (x(t), y(t)) = T(t) \quad t \in [a, b]\}.$$

Et posons

$$\omega = f(x, y) ds \quad \text{avec} \quad ds = \underline{t} \cdot \begin{pmatrix} dx \\ dy \end{pmatrix},$$

où \underline{t} est la tangente unitaire à γ , on peut prendre

$$\underline{t} = \frac{1}{\sqrt{(x')^2 + (y')^2}} \begin{pmatrix} x' \\ y' \end{pmatrix}.$$

Alors, ω est la forme linéaire qui fait correspondre à γ l'intégrale de f sur γ :

$$\int_{\gamma} \omega = \int_{\gamma} f(x, y) ds = \int_a^b f(x(t), y(t)) \sqrt{(x')^2 + (y')^2} dt$$

Proposition 4.1.3. *Dans \mathbb{R}^n une $n + 1$ forme différentielle est une forme nulle.*

4.2 Différentielle d'une forme différentielle

La notation

$$\omega = \sum a_{i_1, \dots, i_k} dx_{i_1} \wedge \dots \wedge dx_{i_k}$$

d'une forme différentielle provient de la définition du produit de deux formes différentielles :

Définition 4.2.1. *Le produit extérieur des deux 1-formes différentielles dx et dy est la 2-forme différentielle $dx \wedge dy$. Elle est notée avec le symbole \wedge .*

Proposition 4.2.2. $dx \wedge dy = -dy \wedge dx$.

On généralise alors

Définition 4.2.3. *Produit extérieur – On définit le produit extérieur de deux formes différentielles ω_1 et ω_2 par les règles de calcul :*

1. $\omega_1 \wedge \omega_2 = -\omega_2 \wedge \omega_1$
2. $\omega_1 \wedge (\omega_2 + \omega_3) = \omega_1 \wedge \omega_2 + \omega_1 \wedge \omega_3$

Par exemple

$$(f dx \wedge dz) \wedge (g dx + h dy) = -h f dx \wedge dy \wedge dz.$$

Théorème 4.2.4. *Il existe un opérateur nommé différentielle et noté d qui à toute $k-1$ forme différentielle ω fait correspondre une k -forme différentielle $d\omega$ avec les règles suivantes :*

1. $d(d\omega) = 0$
2. $df(x_1, \dots, x_n) = \partial_{x_1} f dx_1 + \dots + \partial_{x_n} f dx_n$
3. $d(f(x_1, \dots, x_n) dx_{i_1} \wedge \dots \wedge dx_{i_k}) = df(x_1, \dots, x_n) \wedge (dx_{i_1} \wedge \dots \wedge dx_{i_k})$

Ainsi, le produit d'une k_1 -forme et d'une k_2 forme différentielle donne une $(k_1 + k_2)$ -forme différentielle.

4.3 Théorème de Stokes – Formule de la divergence

Le théorème de Stokes, qui suit, est fondamental, puisqu'en découle toutes les formules d'intégration par partie (Green, Ostogradski) sur des domaines de \mathbb{R}^n . Nous renvoyons à [Rud95] pour une démonstration

Théorème 4.3.1. *Soit ω une $k-1$ -forme différentielle de classe C^1 et soit K une k -surface de classe C^2 , alors*

$$\int_K d\omega = \int_{\partial K} \omega$$

4.4 Formule de la divergence : définition

Le théorème de Stokes permet de définir la notion de divergence :

Définition 4.4.1. Formule de la divergence – Soit \underline{u} un champs de vecteur défini sur $\Omega \subset \mathbb{R}^n$ à valeur dans \mathbb{R}^n et soit \underline{n} le vecteur normal unitaire extérieur à $\partial\Omega$. On définit l'opérateur divergence par l'unique opérateur différentiel tel que

$$\int_{\Omega} \operatorname{div} \underline{u} = \int_{\partial\Omega} \underline{u} \cdot \underline{n}. \quad (4.2)$$

4.4.1 Cas où Ω est un domaine de \mathbb{R}

Dans ce cas trivial, la divergence est confondue avec la dérivée usuelle d'une fonction.

4.4.2 Cas où Ω est un domaine de \mathbb{R}^2

Soit ω une 1-forme différentielle définie sur \mathbb{R}^2 par deux fonctions scalaire u_1 et u_2 définies sur $\Omega \subset \mathbb{R}^2$:

$$\omega = u_2 dx_1 - u_1 dx_2.$$

Alors, d'après les règles de calculs, on a

$$d\omega = \left(\frac{\partial u_1}{\partial x_1} + \frac{\partial u_2}{\partial x_2} \right) dx_1 \wedge dx_2.$$

D'où on déduit d'après le théorème de Stokes :

$$\int_{\Omega} d\omega = \int_{\partial\Omega} u_2 dx_1 - u_1 dx_2 = \int_{[a,b]} (u_2 t_1 - u_1 t_2) ds$$

où le vecteur $\underline{t} = (t_1, t_2)$ est le vecteur tangent unitaire à $\partial\Omega$. Si bien qu'en posant

$$\underline{u} = (u_1, u_2)$$

et en remarquant que la normale extérieure est donnée par

$$\underline{n} = (-t_2, t_1),$$

on a

$$\begin{aligned} \int_{\Omega} \left(\frac{\partial u_1}{\partial x_1} + \frac{\partial u_2}{\partial x_2} \right) dx_1 dx_2 &= \int_{\Omega} \left(\frac{\partial u_1}{\partial x_1} + \frac{\partial u_2}{\partial x_2} \right) dx_1 \wedge dx_2 \\ &= \int_{[a,b]} (u_2 t_1 - u_1 t_2) ds \\ &= \int_{[a,b]} \underline{u} \cdot \underline{n} ds. \end{aligned}$$

D'où la définition

Théorème 4.4.2. Soit \underline{u} un champ de vecteur différentiable défini sur $\Omega \subset \mathbb{R}^2$ à valeurs dans \mathbb{R}^2 .

$$\operatorname{div} \underline{u} = \frac{\partial u_1}{\partial x_1} + \frac{\partial u_2}{\partial x_2}.$$

4.4.3 Cas où Ω est un domaine de \mathbb{R}^3

Théorème 4.4.3. Soit \underline{u} un champ de vecteur différentiable défini sur $\Omega \subset \mathbb{R}^3$ à valeurs dans \mathbb{R}^3 .

$$\operatorname{div} \underline{u} = \frac{\partial u_1}{\partial x_1} + \frac{\partial u_2}{\partial x_2} + \frac{\partial u_3}{\partial x_3}.$$

Démonstration. On procède de façon analogue au cas dans \mathbb{R}^2 :

Soit ω une 2-forme différentielle définie sur \mathbb{R}^3 par les trois fonctions u_1, u_2 et u_3 :

$$\omega = u_1 dx_2 \wedge dx_3 + u_2 dx_3 \wedge dx_1 + u_3 dx_1 \wedge dx_2.$$

Alors la différentielle vaut :

$$d\omega = \left(\frac{\partial u_1}{\partial x_1} + \frac{\partial u_2}{\partial x_2} + \frac{\partial u_3}{\partial x_3} \right) dx_1 \wedge dx_2 \wedge dx_3.$$

On a donc

$$\begin{aligned} \int_{\Omega} \left(\frac{\partial u_1}{\partial x_1} + \frac{\partial u_2}{\partial x_2} + \frac{\partial u_3}{\partial x_3} \right) dx_1 dx_2 dx_3 &= \int_{\Omega} \left(\frac{\partial u_1}{\partial x_1} + \frac{\partial u_2}{\partial x_2} + \frac{\partial u_3}{\partial x_3} \right) dx_1 \wedge dx_2 \wedge dx_3 \\ &= \int_{\Omega} d\omega. \end{aligned}$$

On écrit alors le théorème de Stokes

$$\begin{aligned} \int_{\Omega} d\omega &= \int_{\partial\Omega} u_1 dx_2 \wedge dx_3 + u_2 dx_3 \wedge dx_1 + u_3 dx_1 \wedge dx_2 \\ &= \int_D \left(u_1(\phi) \frac{\partial(x_2, x_3)}{\partial(t_1, t_2)} + u_2(\phi) \frac{\partial(x_3, x_1)}{\partial(t_1, t_2)} + u_3(\phi) \frac{\partial(x_1, x_2)}{\partial(t_1, t_2)} \right) dt_1 dt_2 \\ &= \int_D \det \left(\underline{f}, \frac{\partial\phi}{\partial t_1}, \frac{\partial\phi}{\partial t_2} \right) dt_1 dt_2 \\ &= \int_D \underline{f} \cdot \left(\frac{\partial\phi}{\partial t_1} \wedge \frac{\partial\phi}{\partial t_2} \right) dt_1 dt_2 \\ &= \int_{\partial\Omega} \underline{u} \cdot \underline{n} ds \end{aligned}$$

d'où la définition de la divergence. □

On admettra de façon plus générale le

Théorème 4.4.4. Soit \underline{u} un champ de vecteur différentiable défini sur $\Omega \subset \mathbb{R}^n$ à valeurs dans \mathbb{R}^n .

$$\begin{aligned} \operatorname{div} \underline{u} &= \sum_{i=1}^n \frac{\partial u_i}{\partial x_i} \\ &= \operatorname{tr}(\nabla \underline{u}). \end{aligned}$$

On étends la notion d'opérateur divergence aux champs de tenseur d'ordre 2 :

Théorème 4.4.5. Soit \mathbb{M} un tenseur d'ordre 2 défini sur $\Omega \subset \mathbb{R}^3$, on définit également :

$$\int_{\Omega} \operatorname{div} \mathbb{M} = \int_{\partial\Omega} \mathbb{M} \underline{n}$$

de sorte que la divergence d'un tenseur d'ordre 2 est un champ de vecteur :

$$\operatorname{div} \mathbb{M} = \sum_{j=1}^n \begin{bmatrix} M_{1j,j} \\ M_{2j,j} \\ \vdots \\ M_{ij,j} \\ \vdots \\ M_{nj,j} \end{bmatrix}$$

Corollaire 4.4.6. Si \mathbb{M} est tenseur d'ordre 2 symétrique, on a :

$$\int_{\Omega} \underline{OM} \wedge \operatorname{div} \mathbb{M} = \int_{\partial\Omega} \underline{OM} \wedge \mathbb{M} \underline{n}$$

Remarque 10. Le théorème 4.4.5 et son corollaire ont une grande importance en Mécanique des milieux continus. Ils permettent notamment de déduire le principe fondamental de la dynamique.

4.4.4 Formules de Stokes

Théorème 4.4.7. Soit \underline{u} un champ de vecteur défini sur $\Omega \subset \mathbb{R}^3$ à valeur dans \mathbb{R}^3 :

$$\int_{\partial\Omega} \underline{n} \wedge \underline{u} = \int_{\Omega} \overrightarrow{\operatorname{rot}} \underline{u}$$

4.4.5 Formules de Green

En appliquant la formule de la divergence à $\operatorname{div} \mathbb{M} \underline{v}$ où \mathbb{M} est un tenseur d'ordre 2 défini sur un domaine Ω de \mathbb{R}^3 et \underline{v} un champ de vecteur de \underline{v} également défini sur Ω .

On a alors

$$\int_{\Omega} \operatorname{div} \mathbb{M} \underline{v} = \int_{\partial\Omega} \mathbb{M} \underline{v} \underline{n} - \int_{\Omega} \mathbb{M} : \nabla \underline{v}$$

et si de plus \mathbb{M} est un tenseur symétrique, on a

$$\int_{\Omega} \operatorname{div} \mathbb{M} \underline{v} = \int_{\partial\Omega} \mathbb{M} \underline{v} \underline{n} - \int_{\Omega} \mathbb{M} : \epsilon(\underline{v})$$

où

$$\epsilon(\underline{v}) = \frac{1}{2} (\nabla \underline{v} + \nabla \underline{v}^T).$$

CHAPITRE 5

Calcul différentiel sur un tenseur d'ordre 2

En mécanique, un grand nombre d'objets apparaissent comme des tenseurs d'ordre 2 définis sur une partie de \mathbb{R}^3 , par exemple, le tenseur des contraintes, le tenseurs des déformations...

Ce chapitre a pour but d'en rappeler les définitions mathématiques et les éléments de calcul différentiel sur ces objets. Pour un exposé complet nous renvoyons à [\[Lic46\]](#).

5.1 Tenseur d'ordre 2

Plaçons nous dans \mathbb{R}^3 et considérons un tenseur \mathbb{M} d'ordre 2, c'est à dire un champ défini sur une partie Ω de \mathbb{R}^3 et à valeur dans $\mathcal{L}(\mathbb{R}^3)$, l'ensemble des opérateurs linéaires sur \mathbb{R}^3 .

Ainsi, si $(\underline{e}_1, \underline{e}_2, \underline{e}_3)$ désigne la base canonique de \mathbb{R}^3 , il est usuel d'exprimer un tenseur sous sa forme matricielle dans la base canonique en chaque point de Ω où il est défini :

$$\mathbb{M} = \begin{pmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{pmatrix}$$

où chaque coefficient m_{ij} est une fonction scalaire défini sur Ω . Il est alors commode d'exprimer le tenseur comme une combinaison linéaire des applications linéaires canoniques,

notées sous la forme d'un produit tensoriel

$$\underline{e}_i \otimes \underline{e}_j$$

dont la représentation matricielle est la matrice nulle partout sauf le coefficient de la i -ème ligne j -ème colonne qui vaut 1. Par exemple,

$$\underline{e}_1 \otimes \underline{e}_2 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Ainsi, on écrit

$$\mathbb{M} = \sum_{i=1}^3 \sum_{j=1}^3 m_{ij} \underline{e}_i \otimes \underline{e}_j.$$

5.2 Gradient d'un tenseur d'ordre 2

Par définition, la dérivée (ou le gradient) d'un tenseur en un point est un élément de $\mathcal{L}(\mathbb{R}^3, \mathcal{L}(\mathbb{R}^3))$. Or il existe une bijection naturelle entre $\mathcal{L}(\mathbb{R}^3, \mathcal{L}(\mathbb{R}^3))$ et $\mathcal{L}(\mathcal{L}(\mathbb{R}^3), \mathbb{R}^3)$.

D'autre part, dans la base canonique (base fixe), on a naturellement

$$\frac{\partial}{\partial x_i} \mathbb{M} = \nabla \mathbb{M} \underline{e}_i.$$

Ce qui signifie que le gradient d'un tenseur peut se décomposer

$$\nabla \mathbb{M} = \frac{\partial}{\partial x_1} (\mathbb{M}) \otimes \underline{e}_1 + \frac{\partial}{\partial x_2} (\mathbb{M}) \otimes \underline{e}_2 + \frac{\partial}{\partial x_3} (\mathbb{M}) \otimes \underline{e}_3.$$

5.3 Divergence d'un tenseur d'ordre 2

La divergence d'un tenseur est donné par la formule de la divergence (voir le chapitre 4) :

$$\int_{\Omega} \operatorname{div} \mathbb{M} = \int_{\partial\Omega} \mathbb{M} \underline{n}.$$

Nous savons alors que c'est un champ de vecteurs :

$$\operatorname{div} \mathbb{M} = \sum_{j=1}^n \begin{bmatrix} M_{1j,j} \\ M_{2j,j} \\ \vdots \\ M_{ij,j} \\ \vdots \\ M_{nj,j} \end{bmatrix}$$

On remarque que

$$\operatorname{div} \mathbb{M} \underline{e}_i = \sum_{j=1}^n (\nabla \mathbb{M} \underline{e}_j) (\underline{e}_i \underline{e}_j).$$

5.4 Tenseur en coordonnées cylindriques

Soit \mathbb{M} un tenseur d'ordre 2 définie sur une partie de \mathbb{R}^3 , alors \mathbb{M} se décompose en coordonnées cartésiennes :

$$\mathbb{M} = \sum_{i=1}^3 \sum_{j=1}^3 m_{ij} \underline{e}_i \otimes \underline{e}_j.$$

On a de même en coordonnées cylindriques

$$\begin{aligned} \mathbb{M} &= m_{rr} \underline{e}_r \otimes \underline{e}_r + m_{r\theta} \underline{e}_r \otimes \underline{e}_\theta + m_{r3} \underline{e}_r \otimes \underline{e}_3 \\ &+ m_{\theta r} \underline{e}_\theta \otimes \underline{e}_r + m_{\theta\theta} \underline{e}_\theta \otimes \underline{e}_\theta + m_{\theta 3} \underline{e}_\theta \otimes \underline{e}_3 \\ &+ m_{3r} \underline{e}_3 \otimes \underline{e}_r + m_{3\theta} \underline{e}_3 \otimes \underline{e}_\theta + m_{33} \underline{e}_3 \otimes \underline{e}_3. \end{aligned}$$

Or toutes les dérivées partielles sont nulles sauf

$$\begin{aligned} \frac{\partial}{\partial \theta} (\underline{e}_r \otimes \underline{e}_r) &= \frac{\partial \underline{e}_r}{\partial \theta} \otimes \underline{e}_r + \underline{e}_r \otimes \frac{\partial \underline{e}_r}{\partial \theta} = \underline{e}_\theta \otimes \underline{e}_r + \underline{e}_r \otimes \underline{e}_\theta, \\ \frac{\partial}{\partial \theta} (\underline{e}_r \otimes \underline{e}_\theta) &= \frac{\partial \underline{e}_r}{\partial \theta} \otimes \underline{e}_\theta + \underline{e}_r \otimes \frac{\partial \underline{e}_\theta}{\partial \theta} = \underline{e}_\theta \otimes \underline{e}_\theta - \underline{e}_r \otimes \underline{e}_r, \\ \frac{\partial}{\partial \theta} (\underline{e}_r \otimes \underline{e}_3) &= \frac{\partial \underline{e}_r}{\partial \theta} \otimes \underline{e}_3 + \underline{e}_r \otimes \frac{\partial \underline{e}_3}{\partial \theta} = \underline{e}_\theta \otimes \underline{e}_3, \\ \frac{\partial}{\partial \theta} (\underline{e}_\theta \otimes \underline{e}_r) &= \frac{\partial \underline{e}_\theta}{\partial \theta} \otimes \underline{e}_r + \underline{e}_\theta \otimes \frac{\partial \underline{e}_r}{\partial \theta} = -\underline{e}_r \otimes \underline{e}_r + \underline{e}_\theta \otimes \underline{e}_\theta, \\ \frac{\partial}{\partial \theta} (\underline{e}_\theta \otimes \underline{e}_\theta) &= \frac{\partial \underline{e}_\theta}{\partial \theta} \otimes \underline{e}_\theta + \underline{e}_\theta \otimes \frac{\partial \underline{e}_\theta}{\partial \theta} = -\underline{e}_r \otimes \underline{e}_\theta - \underline{e}_\theta \otimes \underline{e}_r, \\ \frac{\partial}{\partial \theta} (\underline{e}_\theta \otimes \underline{e}_3) &= \frac{\partial \underline{e}_\theta}{\partial \theta} \otimes \underline{e}_3 + \underline{e}_\theta \otimes \frac{\partial \underline{e}_3}{\partial \theta} = -\underline{e}_r \otimes \underline{e}_3, \\ \frac{\partial}{\partial \theta} (\underline{e}_3 \otimes \underline{e}_r) &= \frac{\partial \underline{e}_3}{\partial \theta} \otimes \underline{e}_r + \underline{e}_3 \otimes \frac{\partial \underline{e}_r}{\partial \theta} = \underline{e}_3 \otimes \underline{e}_\theta, \\ \frac{\partial}{\partial \theta} (\underline{e}_3 \otimes \underline{e}_\theta) &= \frac{\partial \underline{e}_3}{\partial \theta} \otimes \underline{e}_\theta + \underline{e}_3 \otimes \frac{\partial \underline{e}_\theta}{\partial \theta} = -\underline{e}_3 \otimes \underline{e}_r. \end{aligned}$$

Mais le gradient est donné par

$$\nabla \mathbb{M} = \frac{\partial \mathbb{M}}{\partial r} \otimes \underline{e}_r + \frac{1}{r} \frac{\partial \mathbb{M}}{\partial \theta} \otimes \underline{e}_\theta + \frac{\partial \mathbb{M}}{\partial 3} \otimes \underline{e}_3,$$

d'où on tire l'expression du gradient de \mathbb{M} en coordonnées cylindriques :

$$\begin{aligned} \nabla \mathbb{M} &= m_{rr,r} \underline{e}_r \otimes \underline{e}_r \otimes \underline{e}_r + m_{r\theta,r} \underline{e}_r \otimes \underline{e}_\theta \otimes \underline{e}_r + m_{r3,r} \underline{e}_r \otimes \underline{e}_3 \otimes \underline{e}_r \\ &+ m_{\theta r,r} \underline{e}_\theta \otimes \underline{e}_r \otimes \underline{e}_r + m_{\theta\theta,r} \underline{e}_\theta \otimes \underline{e}_\theta \otimes \underline{e}_r + m_{\theta 3,r} \underline{e}_\theta \otimes \underline{e}_3 \otimes \underline{e}_r \\ &+ m_{3r,r} \underline{e}_3 \otimes \underline{e}_r \otimes \underline{e}_r + m_{3\theta,r} \underline{e}_3 \otimes \underline{e}_\theta \otimes \underline{e}_r + m_{33,r} \underline{e}_3 \otimes \underline{e}_3 \otimes \underline{e}_r \\ &+ \frac{1}{r} (m_{rr,\theta} - m_{r\theta} - m_{\theta r}) \underline{e}_r \otimes \underline{e}_r \otimes \underline{e}_\theta + \frac{1}{r} (m_{r\theta,\theta} + m_{rr} - m_{\theta\theta}) \underline{e}_r \otimes \underline{e}_\theta \otimes \underline{e}_\theta \\ &\quad + \frac{1}{r} (m_{r3,\theta} - m_{\theta 3}) \underline{e}_r \otimes \underline{e}_3 \otimes \underline{e}_\theta \\ &+ \frac{1}{r} (m_{\theta r,\theta} + m_{rr} - m_{\theta\theta}) \underline{e}_\theta \otimes \underline{e}_r \otimes \underline{e}_\theta + \frac{1}{r} (m_{\theta\theta,\theta} + m_{r\theta} + m_{\theta r}) \underline{e}_\theta \otimes \underline{e}_\theta \otimes \underline{e}_\theta \\ &\quad + \frac{1}{r} (m_{\theta 3,\theta} + m_{r3}) \underline{e}_\theta \otimes \underline{e}_3 \otimes \underline{e}_\theta \\ &+ \frac{1}{r} (m_{3r,\theta} - m_{3\theta}) \underline{e}_3 \otimes \underline{e}_r \otimes \underline{e}_\theta + \frac{1}{r} (m_{3\theta,\theta} + m_{3r}) \underline{e}_3 \otimes \underline{e}_\theta \otimes \underline{e}_\theta \\ &\quad + \frac{1}{r} m_{33,\theta} \underline{e}_3 \otimes \underline{e}_3 \otimes \underline{e}_\theta \\ &+ m_{rr,3} \underline{e}_r \otimes \underline{e}_r \otimes \underline{e}_3 + m_{r\theta,3} \underline{e}_r \otimes \underline{e}_\theta \otimes \underline{e}_3 + m_{r3,3} \underline{e}_r \otimes \underline{e}_3 \otimes \underline{e}_3 \\ &+ m_{\theta r,3} \underline{e}_\theta \otimes \underline{e}_r \otimes \underline{e}_3 + m_{\theta\theta,3} \underline{e}_\theta \otimes \underline{e}_\theta \otimes \underline{e}_3 + m_{\theta 3,3} \underline{e}_\theta \otimes \underline{e}_3 \otimes \underline{e}_3 \\ &+ m_{3r,3} \underline{e}_3 \otimes \underline{e}_r \otimes \underline{e}_3 + m_{3\theta,3} \underline{e}_3 \otimes \underline{e}_\theta \otimes \underline{e}_3 + m_{33,3} \underline{e}_3 \otimes \underline{e}_3 \otimes \underline{e}_3. \end{aligned}$$

C'est à dire, de façon légèrement plus lisible ;) :

$$\begin{aligned} \nabla \mathbb{M} \underline{e}_r &= \begin{pmatrix} m_{rr,r} & m_{r\theta,r} & m_{r3,r} \\ m_{\theta r,r} & m_{\theta\theta,r} & m_{\theta 3,r} \\ m_{3r,r} & m_{3\theta,r} & m_{33,r} \end{pmatrix} \\ \nabla \mathbb{M} \underline{e}_\theta &= \frac{1}{r} \begin{pmatrix} m_{rr,\theta} - m_{r\theta} - m_{\theta r} & m_{r\theta,\theta} + m_{rr} - m_{\theta\theta} & m_{r3,\theta} - m_{\theta 3} \\ m_{\theta r,\theta} + m_{rr} - m_{\theta\theta} & m_{\theta\theta,\theta} + m_{r\theta} + m_{\theta r} & m_{\theta 3,\theta} + m_{r3} \\ m_{3r,\theta} - m_{3\theta} & m_{3\theta,\theta} + m_{3r} & m_{33,\theta} \end{pmatrix} \\ \nabla \mathbb{M} \underline{e}_3 &= \begin{pmatrix} m_{rr,3} & m_{r\theta,3} & m_{r3,3} \\ m_{\theta r,3} & m_{\theta\theta,3} & m_{\theta 3,3} \\ m_{3r,3} & m_{3\theta,3} & m_{33,3} \end{pmatrix}. \end{aligned}$$

La divergence d'un tenseur est définie comme précédemment :

$$\begin{aligned} \operatorname{div} \mathbb{M} \underline{e}_r &= \sum_{j=1}^3 \nabla \mathbb{M} \underline{e}_j (\underline{e}_r \underline{e}_j) \\ &= \nabla \mathbb{M} \underline{e}_r (\underline{e}_r \underline{e}_r) + \nabla \mathbb{M} \underline{e}_\theta (\underline{e}_r \underline{e}_\theta) + \nabla \mathbb{M} \underline{e}_z (\underline{e}_r \underline{e}_z), \end{aligned}$$

de même, on a :

$$\operatorname{div} \mathbb{M} \underline{e}_\theta = \nabla \mathbb{M} \underline{e}_r (\underline{e}_\theta \underline{e}_r) + \nabla \mathbb{M} \underline{e}_\theta (\underline{e}_\theta \underline{e}_\theta) + \nabla \mathbb{M} \underline{e}_z (\underline{e}_\theta \underline{e}_z),$$

et

$$\operatorname{div} \mathbb{M}_{\underline{e}_r} = \nabla \mathbb{M}_{\underline{e}_r}(\underline{e}_z \underline{e}_r) + \nabla \mathbb{M}_{\underline{e}_\theta}(\underline{e}_z \underline{e}_\theta) + \nabla \mathbb{M}_{\underline{e}_z}(\underline{e}_z \underline{e}_z).$$

On obtient donc finalement :

Proposition 5.4.1. *La divergence d'un tenseur d'ordre 2 en coordonnées cylindrique*

$$[\operatorname{div} \mathbb{M}]_{\text{cyl}} = \begin{pmatrix} m_{rr,r} + \frac{1}{r}(m_{r\theta,\theta} + m_{rr} - m_{\theta\theta}) + m_{r3,3} \\ m_{\theta r,r} + \frac{1}{r}(m_{\theta\theta,\theta} + m_{r\theta} + m_{\theta r}) + m_{\theta 3,3} \\ m_{3r,r} + \frac{1}{r}(m_{3\theta,\theta} + m_{3r}) + m_{33,3} \end{pmatrix}.$$

Il peut être parfois avantageux de le récrire sous la forme suivante :

$$[\operatorname{div} \mathbb{M}]_{\text{cyl}} = \begin{pmatrix} \frac{1}{r} \frac{\partial}{\partial r}(r m_{rr}) & + \frac{1}{r} \frac{\partial}{\partial \theta}(m_{r\theta}) + \frac{\partial}{\partial x_3}(m_{r3}) - \frac{1}{r} m_{\theta\theta} \\ \frac{1}{r^2} \frac{\partial}{\partial r}(r^2 m_{\theta r}) & + \frac{1}{r} \frac{\partial}{\partial \theta}(m_{\theta\theta}) + \frac{\partial}{\partial x_3}(m_{\theta 3}) + \frac{1}{r}(m_{r\theta} - m_{\theta r}) \\ \frac{\partial}{\partial r}(m_{3r}) & + \frac{1}{r} \frac{\partial}{\partial \theta}(m_{3\theta}) + \frac{\partial}{\partial x_3}(m_{33}) + \frac{1}{r} m_{3r} \end{pmatrix}$$

5.5 Tenseur en coordonnées sphériques

A compléter... par le lecteur ?^

CHAPITRE 6

Systemes différentiels ordinaires

à écrire entièrement

CHAPITRE 7

Fonction d'une variable complexe

7.1 Nombres Complexes

Définition 7.1.1. On peut définir l'ensemble des nombres complexes par

$$\mathbb{C} = \{(a, b) \in \mathbb{R}^2 \text{ muni des somme et multiplication complexes}\}.$$

On pose $i = (0, 1)$ et on note $z = (a, b) = a + ib$, a désigne alors la partie réelle et b désigne la partie imaginaire de z .

Définition 7.1.2. Les somme et multiplication complexes sont définies, pour $z = a + ib$ et $z' = a' + ib'$, par :

- $z + z' = (a, b) + (a', b') = (a + a', b + b')$.
- $zz' = (aa' - bb', ab' + a'b)$.

Par exemple pour $z = (0, 1) = i$ on a $z^2 = zz = (0, 1)(0, 1) = (-1, 0) = -1$.

Ainsi, à chaque point du plan de \mathbb{R}^2 on fait correspondre un unique nombre complexe ; C'est pourquoi on parle de plan complexe.

Proposition 7.1.3. Pour tout nombre complexe z et z' , on a

1. $\overline{z + z'} = \overline{z} + \overline{z'}$
2. $\overline{zz'} = \overline{z}\overline{z'}$
3. $z + \overline{z} = 2\text{Re}(z)$ et $z - \overline{z} = 2i\text{Im}(z)$

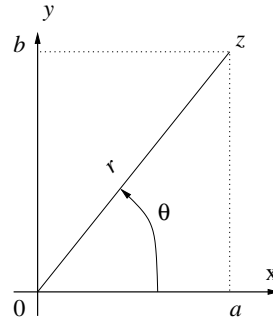


FIGURE 7.1 – Un nombre complexe $z = a + ib$

4. $z\bar{z}$ est réel et si $z \neq 0$, $z\bar{z} > 0$

Définition 7.1.4. On appelle module de z le nombre réel positif $|z| = (z\bar{z})^{1/2}$.

Proposition 7.1.5. Soit z un nombre complexe non nul alors

$$\frac{1}{z} = \frac{\bar{z}}{|z|^2}$$

Proposition 7.1.6. Soit z un nombre complexe

1. $|\bar{z}| = |z|$
2. $|zz'| = |z| |z'|$
3. $|\operatorname{Re}(z)| \leq |z|$ et $|\operatorname{Im}(z)| \leq |z|$
4. $|z + z'| \leq |z| + |z'|$

Proposition 7.1.7. Soient a_1, \dots, a_n et b_1, \dots, b_n des nombres complexes.

$$\left| \sum_{j=1}^n a_j \bar{b}_j \right|^2 \leq \sum_{j=1}^n |a_j|^2 \sum_{j=1}^n |b_j|^2 \quad (\text{inégalité de Schwarz})$$

Nous terminons cette présentation par les propriétés principales de \mathbb{C} :

Théorème 7.1.8. L'ensemble des nombres complexes, avec ses addition et multiplication complexes, possède une structure de corps. On parle de corps des nombres complexes.

Théorème 7.1.9. Le corps des nombres complexes est algébriquement clos.

Autrement dit, les solutions de toute équation algébrique à coefficients complexes sont encore des nombres complexes. Par exemple, le corps des réels \mathbb{R} n'est pas algébriquement clos puisque l'équation $x^2 + 1 = 0$ ne possède pas de solutions réelles mais des solutions complexes.

Théorème 7.1.10. Soit $P(z) = \sum_0^n a_i z^i$ un polynôme de degré n à coefficients complexes. Alors P se factorise

$$P(z) = a_n \prod_1^n (z - r_i)$$

où les r_i sont les n racines de P (comptées avec leurs multiplicités).

7.2 Fonctions holomorphes

7.2.1 Fonctions holomorphes

Soit Ω une partie non vide du plan complexe \mathbb{C} . Considérons une fonction f à valeur complexe définie sur Ω . Soit $z_0 \in \Omega \subset \mathbb{C}$, on sait que f est différentiable en z_0 , si il existe un nombre complexe, noté $f'(z_0)$ tel que :

$$f(z_0 + h) = f(z_0) + f'(z_0)h + o(h).$$

Bien entendu si un tel nombre existe, il est égal à

$$f'(z_0) = \lim_{z \rightarrow z_0} \frac{f(z) - f(z_0)}{z - z_0}$$

C'est la dérivée de f au point z_0 .

Pour $z = (x, y) = x + iy$, introduisons les notations :

$$\frac{\partial f}{\partial z} = \frac{1}{2} \left(\frac{\partial f}{\partial x} - i \frac{\partial f}{\partial y} \right), \quad \frac{\partial f}{\partial \bar{z}} = \frac{1}{2} \left(\frac{\partial f}{\partial x} + i \frac{\partial f}{\partial y} \right).$$

Définition 7.2.1. Si f est différentiable dans tout Ω , alors on dit que f est holomorphe (ou analytique) dans Ω si

$$\frac{\partial f}{\partial \bar{z}} = 0 \quad \text{dans } \Omega.$$

Par exemple, les fonctions analytiques sont évidemment holomorphes dans leurs domaines de convergence, mais aussi la fonction

$$f(z) = \frac{1}{z - z_0}$$

est holomorphe dans $\mathbb{C} - B(z_0, \varepsilon)$, ou même dans $\mathbb{C} - \{z_0\}$.

Par contre la fonction $f(z) = |z|$ n'est pas holomorphe dans aucune partie de \mathbb{C} .

On a les propriétés suivantes :

Proposition 7.2.2. Si f et g sont holomorphes dans Ω , alors

$$- f + g,$$

$$- fg,$$

le sont aussi.

Proposition 7.2.3. *Si f est holomorphe dans Ω et g est holomorphe dans $f(\Omega)$, alors la fonction composée $(f \circ g)$ est holomorphe dans Ω et :*

$$(f \circ g)' = g'(f(z_0))f'(z_0).$$

7.2.2 Conditions de Cauchy-Riemann

Théorème 7.2.4. *Soit f une fonction complexe définie dans $\Omega \subset \mathbb{C}$, on décompose f suivant sa partie réelle P et sa partie imaginaire Q , avec $z = (x, y) = x + iy$:*

$$f(z) = P(x, y) + iQ(x, y).$$

f est une fonction holomorphe si et seulement si P et Q satisfont aux conditions de Cauchy-Riemann :

$$\frac{\partial P}{\partial x} = \frac{\partial Q}{\partial y} \quad \text{et} \quad \frac{\partial P}{\partial y} = -\frac{\partial Q}{\partial x}.$$

Corollaire 7.2.5. *Si $f = P + iQ$ est holomorphe, alors les fonctions réelles P et Q sont nécessairement harmoniques.*

Terminons par cette question que nous laissons au lecteur :

Existe-t-il une fonction définie sur une variable complexe qui soit différentiable mais pas holomorphe ?

7.3 Suite et Séries de nombres complexes

Considérons la série

$$f(z) = \sum_0^{\infty} a_n z^n$$

où $z \in \mathbb{C}$ et $a_n \in \mathbb{C}$. On définit le rayon de convergence R de la série $f(z)$ par :

$$\begin{aligned} \frac{1}{R} &= \lim_{n \rightarrow \infty} \sup |a_n|^{1/n} \quad (\text{règle de Cauchy}) \\ &= \lim_{n \rightarrow \infty} \sup \left| \frac{a_n}{a_{n-1}} \right| \quad (\text{règle de d'Alembert}). \end{aligned}$$

C'est à dire que la série diverge si $|z| > R$ et elle converge si $|z| < R$. Le cas $|z| = R$ étant sujet à discussion.

Par ailleurs, cette convergence est normale et donc uniforme dans tout disque centré de rayon $R' < R$. Alors, d'après les théorèmes d'inversion des limites vus en premier cycle, on a

Proposition 7.3.1. $\forall z, |z| < R$, la fonction d'une variable complexe $f(z)$ est continue, infiniment \mathbb{C} -dérivable, les dérivées s'obtenant en dérivant dans chaque terme de la série :

$$f^{(k)}(z) = \sum_0^{\infty} \frac{(n+k)!}{(n)!} a_{n+k} z^n$$

On dit que $f(z)$ est *analytique* dans $\{z \in \mathbb{C} / |z| < R\}$. Si $f(z)$ est analytique dans \mathbb{C} , on dit que f est *entière*.

7.4 Divers

A compléter : théorèmes des zéros isolés, principe du maximum, théorème d'Abel...

7.4.1 Fonctions exponentielle et trigonométrie

La fonction exponentielle est définie par

$$e^z = \sum_0^{\infty} \frac{z^n}{n!}$$

dont le rayon de convergence est infini.

En dérivant terme à terme, on a la propriété fondamentale :

Proposition 7.4.1.

$$\frac{d}{dz} e^z = e^z$$

d'où on déduit, en écrivant le développement en série de Taylor de la série en z :

$$e^{z+z'} = e^z e^{z'} \quad \forall z, z' \in \mathbb{C}. \quad (7.1)$$

Par ailleurs, on remarque que pour $z = i\theta$, $\theta \in \mathbb{R}$ on a

$$e^{i\theta} = \sum_0^{\infty} \frac{(-1)^n}{(2n+1)!} \theta^{2n+1} + i \sum_0^{\infty} \frac{(-1)^n}{(2n)!} \theta^{2n},$$

c'est à dire la formule d'Euler :

$$e^{i\theta} = \cos \theta + i \sin \theta.$$

On déduit que

$$|e^{i\theta}| = 1 \quad \forall \theta \in \mathbb{R}.$$

Proposition 7.4.2. *Tout nombre complexe $z = a + ib$ peut s'écrire sous la forme*

$$z = re^{i\theta} \quad r \in \mathbb{R}^+, \theta \in [0, 2\pi[$$

où $r^2 = a^2 + b^2$ est le module de z et $\theta = \text{Arctan}(b/a)$ est l'argument de z . Par suite, la fonction exponentielle est périodique de période 2π .

On peut définir les fonctions \sin et \cos d'une variable complexe

$$\sin z = \sum_0^{\infty} \frac{(-1)^n}{(2n)!} z^{2n} \quad \cos z = \sum_0^{\infty} \frac{(-1)^n}{(2n+1)!} z^{2n+1}.$$

Remarquons qu'on a toujours

$$(\sin z)^2 + (\cos z)^2 = 1.$$

7.5 Intégrales complexes

7.5.1 Formule intégrale de Cauchy

Soit Ω un ouvert du plan complexe et soit $\gamma(t)$ une courbe dans Ω , pour $t \in [a, b]$. On décompose

$$\gamma(t) = x(t) + iy(t)$$

Soit $\omega = f(z)dz$ une forme différentielle, où f est une fonction de variable complexe définie sur ω à valeur dans \mathbb{C} , par définition on a

Définition 7.5.1.

$$\int_{\gamma} \omega = \int_{\gamma} f(z)dz = \int_a^b f \circ \gamma(t) \gamma'(t) dt$$

où $\gamma([a, b]) = \gamma$ et $\gamma'(t) = x'(t) + iy'(t)$.

On a la propriété fondamentale :

Théorème 7.5.2. *Soit f une fonction d'une variable complexe z si f est dérivable dans Ω alors on a*

$$f(\zeta) = \frac{1}{2i\pi} \left[\int_{\partial\Omega} \frac{f(z)}{z - \zeta} dz + \int_{\Omega} \frac{\partial f(z) \partial \bar{z}}{z - \zeta} dz \wedge d\bar{z} \right] \quad \forall \zeta \in \Omega.$$

On en déduit immédiatement

Corollaire 7.5.3. *Formule intégrale de Cauchy – Si f est holomorphe dans Ω , un domaine convexe de \mathbb{C} alors*

$$f(\zeta) = \frac{1}{2i\pi} \int_{\partial\Omega} \frac{f(z)}{z - \zeta} dz \quad \forall \zeta \in \Omega.$$

De plus on voit que f est infiniment dérivable avec

$$\frac{d^m f}{dz^m}(\zeta) = \frac{m!}{2i\pi} \int_{\partial\Omega} \frac{f(z)}{(z - \zeta)^{m+1}} dz \quad \forall \zeta \in \Omega.$$

Théorème 7.5.4. *Si f est holomorphe dans Ω alors pour tout chemin fermé γ dans Ω , on a*

$$\int_{\gamma} f(z) dz = 0$$

On dit que $\phi(s) = \int_a^s f \circ \gamma(t) \gamma'(t) dt$ est une primitive de f le long de γ .

Corollaire 7.5.5. *Si f est holomorphe sur Ω , alors f est de classe C^∞ sur Ω .*

Théorème 7.5.6. *Si f est holomorphe dans Ω , alors f se décompose en série entière :*

$$f(\zeta) = \sum_{n=0}^{\infty} a_n (\zeta - z_0)^n \quad \forall \zeta \in \Omega \text{ et } \zeta \in V(z_0).$$

avec

$$a_n = \frac{1}{m!} \frac{d^m f}{dz^m}(z_0) = \frac{1}{2i\pi} \int_{\partial\Omega} \frac{f(z)}{(z - z_0)^{m+1}} dz.$$

Corollaire 7.5.7. *Une fonction à variable complexe est holomorphe dans Ω si et seulement si elle est analytique dans Ω .*

On termine par quelques résultats hors programmes mais néanmoins importants

Théorème 7.5.8. *Si une fonction est holomorphe et non constante alors elle est ouverte*

Corollaire 7.5.9. – *Principe du maximum – Si f est une fonction holomorphe dans un ouvert complexe Ω , et si $|f|$ atteint son maximum à l'intérieur de Ω , alors f est une fonction constante sur Ω .*

Théorème 7.5.10. *Soit f une fonction holomorphe non nulle dans un ouvert complexe Ω , les zéros de f sont isolés.*

En particulier si deux fonctions holomorphes ont des valeurs en un nombre infini de points dans un domaine borné, alors les deux fonctions sont identiques.

7.5.2 Série de Laurent

Soit f une fonction holomorphe sur une couronne $a < |z - z_0| < b$, d'après la formule de Cauchy,

$$f(z) = \frac{1}{2i\pi} \left[- \int_{C_a} \frac{f(u)}{u - z} du + \int_{C_b} \frac{f(u)}{u - z} du \right]$$

or

$$\frac{f(u)}{u - z} = \frac{f(u)}{u} \left[\frac{1}{1 - \frac{z}{u}} \right] = \frac{f(u)}{u} \sum_{n=0}^{\infty} \left(\frac{z}{u} \right)^n = f(u) \sum_{n=0}^{\infty} \frac{z^n}{u^{n+1}}$$

où la série est convergente si $|z| < |u|$ ou encore $|z - z_0| < |u - z_0|$.

D'autre part, on a de même

$$\frac{f(u)}{u - z} = \frac{f(u)}{z} \left[\frac{1}{\frac{u}{z} - 1} \right] = \frac{-f(u)}{z} \sum_{n=0}^{\infty} \left(\frac{u}{z} \right)^n = -f(u) \sum_{n=0}^{\infty} \frac{u^n}{z^{n+1}}$$

où la série est convergente si $|z| > |u|$ ou encore $|z - z_0| > |u - z_0|$.

Par conséquent, en remplaçant $(u - z)$ par $(u - z_0) - (z - z_0)$, on obtient

$$f(z) = \frac{1}{2i\pi} \left[\int_{C_a} f(u) \sum_{n=0}^{\infty} \frac{(u - z_0)^n}{(z - z_0)^{n+1}} du + \int_{C_b} f(u) \sum_{n=0}^{\infty} \frac{(z - z_0)^n}{(u - z_0)^{n+1}} du \right]$$

Ainsi, on a montré le

Théorème 7.5.11. Soit f une fonction holomorphe sur une couronne $a < |z - z_0| < b$, il existe un développement unique de f :

$$f(z) = \sum_{n \in \mathbb{Z}} a_n (z - z_0)^n$$

avec

$$a_n = \frac{1}{2i\pi} \begin{cases} \int_{C_b} \frac{f(u)}{(z - z_0)^{n+1}} du, & n \geq 0 \\ \int_{C_a} \frac{f(u)}{(z - z_0)^{n+1}} du, & n < 0 \end{cases}$$

C'est le développement de f en série de Laurent dans la couronne $a < |z - z_0| < b$.

La notion de série de Laurent permet de définir la notion de Résidu :

Définition 7.5.12. Soit f une fonction holomorphe dans Ω sauf en des points isolés (c'est toujours le cas). Le résidu de f en $z_i \in \Omega$ est la valeur complexe

$$\text{Rés}(f, z_i) = \frac{1}{2i\pi} \int_{\partial B(z_i, \varepsilon)} f(z) dz,$$

pour ε suffisamment petit, de sorte que f est holomorphe dans $B(z_i, \varepsilon) \setminus \{z_i\}$.

C'est aussi le coefficient a_{-1} du développement en série de Laurent de f autour de z_i .

7.5.3 Théorèmes des résidus

Soit f une fonction holomorphe dans Ω sauf en un nombre fini de points isolés $\{z_1, \dots, z_m\}$. Soit ε suffisamment petit pour que f soit holomorphe dans $\Omega \setminus \bigcup_i B(z_i, \varepsilon)$. Alors d'après le théorème 7.5.4, on a

$$\int_{\partial\Omega} f(z)dz - \sum_i \int_{\partial B(z_i, \varepsilon)} f(z)dz = 0.$$

Autrement dit nous avons le

Théorème 7.5.13. – *Théorème des Résidus* – Soit f une fonction holomorphe dans Ω sauf en un nombre fini de points isolés $\{z_1, \dots, z_m\}$, alors

$$\int_{\partial\Omega} f(z)dz = 2i\pi \sum_i \text{Rés}(f, z_i).$$

7.5.4 Classification des Pôles singuliers

Pour faciliter le calcul d'un résidu, une classification est utile et permet de s'affranchir de la détermination du développement en série de Laurent.

Soit a un point de Ω qui apparaît comme un point singulier d'une fonction f , c'est à dire que f n'est pas holomorphe a priori en a . On classe ces pôles ou singularités suivant la possibilité de prolonger $f(z)(z-a)^k$ en une fonction holomorphe au voisinage de a :

1. Singularité ou Pôle illusoire : Si f est prolongeable par continuité en a . Cela signifie que f est holomorphe en a , si bien que

$$\text{Rés}(f, a) = 0.$$

On parle de pôle d'ordre 0.

2. Singularité ou Pôle d'ordre k : Si $(z-a)^k f(z)$ est prolongeable par continuité en a mais pas $(z-a)^{k-1} f(z)$. Le résidu est alors obtenu par la formule intégrale de Cauchy :

$$\text{Rés}(f, a) = \frac{1}{(k-1)!} \frac{d^{k-1}}{dz^{k-1}} [(z-a)^k f(z)]_{z=a}.$$

3. Singularité essentielle : pour tout $k \in \mathbb{N}$ $(z-a)^k f(z)$ n'est pas prolongeable par continuité, on a pas le choix, il faut trouver le développement en série de Laurent.
4. Singularité à l'infini : si a est un point à l'infini on pose :

$$\text{Rés}(f(z), \infty) = \text{Rés}(f(1/z), 0)$$

Théorème 7.5.14. Soit f une fonction holomorphe sauf en des points isolés. Alors la somme des résidus, résidu à l'infini compris est nulle.

Nous terminons par un résultat très utile en calcul des résidus :

Proposition 7.5.15. *Soit a une singularité d'ordre 1 d'une fonction f . Si f s'exprime comme la fraction rationnelle irréductible de deux fonctions holomorphe,*

$$f(z) = \frac{P(z)}{Q(z)},$$

alors

$$\text{Rés}(f, a) = \frac{P(a)}{Q'(a)}.$$

7.5.5 Représentation conforme

Une fonction holomorphe est une application conforme dans le sens où elle laisse les angles inchangés en valeur et en sens.

CHAPITRE 8

Théorie restreinte de la mesure La mesure de Lebesgue

La théorie de la mesure est la base fondatrice de l'intégrale de Lebesgue qui suit dans le chapitre 9.

Nous donnons ici une présentation simplifiée de la théorie de la mesure. En effet, nous nous limitons à la mesure de Lebesgue, c'est à dire la mesure des parties de \mathbb{R}^n . En particulier, nous faisons l'impasse sur les notions fondatrices de la théorie classique de la mesure, comme la notion de tribu, de mesure σ -additive ou d'ensemble Borélien.

Pour un exposé complet nous renvoyons par exemple à "Real and Complex Analysis" de Walter Rudin ou tout cours de Mathématique sur le sujet (niveau Licence de Mathématiques). Nous pensons que cela n'est pas nécessaire pour un étudiant en Mécanique ou pour un élève ingénieur. Ce n'est sans doute pas l'avis de tous... Quant à l'éventuel étudiant en Mathématiques, il pourra voir cette présentation comme une introduction à la théorie abstraite.

Par contre, nous donnons directement une présentation sur un ensemble ouvert $E \subset \mathbb{R}^n$. E peut donc être borné ou non-borné. Puisque l'un des intérêt de la théorie est justement de s'affranchir d'une définition sur un intervalle comme c'est le cas pour l'intégrale de Riemann.

8.1 Ensemble élémentaire

8.1.1 Pavés et mesure d'un pavé

Définition 8.1.1. Un pavé P de E est un ensemble de $E \subset \mathbb{R}^n$ qui peut s'écrire sous la forme de produit (tensoriel) d'intervalles réels contenant ou non leurs extrémités :

$$P = [a_1, b_1] \times \cdots \times [a_n, b_n]$$

ou encore

$$P = [a_1, b_1[\times \cdots \times]a_n, b_n]$$

etc, ... Les intervalles pouvant être ouverts ou non.

Définition 8.1.2. La mesure d'un intervalle $[a, b]$ est sa mesure Euclidienne, c'est à dire sa longueur :

$$m([a, b]) = |b - a|.$$

La mesure d'un pavé de $E \subset \mathbb{R}^n$ est le produit de la mesure des intervalles le constituant :

$$\begin{aligned} m([a_1, b_1] \times \cdots \times [a_n, b_n]) &= m([a_1, b_1[\times \cdots \times]a_n, b_n]) \\ &= \dots \\ &= m(]a_1, b_1[\times \cdots \times]a_n, b_n[) \\ &= \prod_{i=1}^n |b_i - a_i| \end{aligned}$$

Ainsi, dans \mathbb{R}^2 les pavés sont des rectangles et $m()$ mesure l'aire, dans \mathbb{R}^3 $m()$ mesure le volume.

8.1.2 Ensemble élémentaire et mesure d'un ensemble élémentaire

Définition 8.1.3. On dit qu'un ensemble est élémentaire s'il est réunion disjointe et dénombrable de pavés.

Si $A = \bigcup P_k$ où les P_k sont des pavés 2 à 2 disjoints, alors

$$m(A) = \sum m(P_k).$$

Proposition 8.1.4. La mesure d'un ensemble élémentaire est indépendante du choix de sa décomposition.

8.2 Ensemble mesurable

8.2.1 Mesure extérieure

Définition 8.2.1. Soit A une partie de $E \subset \mathbb{R}^n$, on définit sa mesure extérieure par

$$\mu^*(A) = \inf \left\{ \sum_k m(P_k); \quad A \subset \bigcup_k P_k \right\}$$

où $A \subset \bigcup_k P_k$ désigne un recouvrement de A par des pavés de E .

8.2.2 Ensemble mesurable

Définition 8.2.2. Une partie A de E est dit mesurable si pour tout $\varepsilon > 0$, il existe un ensemble élémentaire B de E tel que

$$\mu^*(A \Delta B) \leq \varepsilon$$

où $A \Delta B$ désigne la différence symétrique ¹.

Autrement dit, les ensembles mesurables sont les ensembles limites d'ensembles élémentaires. Ce qui signifie que pour trouver un ensemble non-mesurable, il faut se lever de bonne heure !

Malgré tout nous pouvons préciser un peu la définition d'ensemble mesurable. De façon triviale nous avons

Proposition 8.2.3. *Tout ensemble élémentaire est mesurable.*

Théorème 8.2.4. *Tout ensemble ouvert de $\subset \mathbb{R}^n$ est mesurable*

Preuve du théorème 8.2.4 : Soit A un ouvert de E . Par définition, pour chaque point de A , il existe une boule (et donc un pavé) de mesure non-nulle, centré en ce point et contenu dans A .

Il suffit alors de considérer tous les points de A à composantes rationnelles et les pavés de cotés rationnels. Il est clair que tous les points de A sont contenus dans un de ces pavés "rationnels". Ce qui signifie que A peut être approchée par des ensembles élémentaires aussi près que l'on veut. \square

Théorème 8.2.5. *La réunion dénombrable d'ensembles mesurables est mesurable. L'intersection finie d'ensembles mesurables est mesurable.*

1. $A \Delta B = [A \setminus (A \cap B)] \cup [B \setminus (A \cap B)]$

8.2.3 Mesure d'un ensemble mesurable

Définition 8.2.6. Soit A un ensemble mesurable dans E . Alors la mesure de A est sa mesure extérieure.

$$m(A) = \mu^*(A).$$

Proposition 8.2.7. Soit A_n des ensembles mesurables et soit $A = \cup_n A_n$. Alors

$$m(A) \leq \sum_n m(A_n).$$

Si de plus les A_n sont deux à deux disjoints, alors

$$m(A) = \sum_n m(A_n).$$

8.2.4 Ensemble de mesure nulle

Nous terminons cette section par des exemples d'ensemble de mesure nulle. A compléter..

8.3 Fonctions mesurables

8.3.1 Limites de fonctions continues

Définition 8.3.1. Soit f une fonction à valeur réelle définie sur $E \subset \mathbb{R}^n$. On dit que f est mesurable si pour tout $a \in \mathbb{R}$,

$$\{x \in E \mid f(x) > a\}$$

est un ensemble mesurable.

Proposition 8.3.2. Toute fonction continue sur E est mesurable dans E .

Théorème 8.3.3. Une f est mesurable si et seulement si pour tout $\varepsilon > 0$ il existe une fonction continue φ telle que

$$\mu^* \{x, f(x) \neq \varphi(x)\} \leq \varepsilon.$$

Autrement dit, les fonctions mesurables sont toutes les limites simples de fonctions continues. Cela entraîne immédiatement le résultat corollaire :

Corollaire 8.3.4. Toute limite simple de fonction mesurable est mesurable.

8.3.2 Propriétés des fonctions mesurables

Théorème 8.3.5. *Si f est une fonction mesurable sur E , et si f_n désigne une suite de fonctions mesurables sur E , alors les fonctions*

$$\begin{aligned} |f| \\ g(x) &= \sup_n f_n(x) \\ h(x) &= \lim_{\infty} \sup f_n(x) \end{aligned}$$

le sont également.

8.3.3 Classe des fonctions nulles presque partout

Définition 8.3.6. *On dit qu'une fonction mesurable f est nulle presque partout sur E si il existe un ensemble de mesure nulle A tel que*

$$f(x) = 0 \quad \forall x \in E \setminus A.$$

On note $f = 0$ p.p. sur E .

Comme l'union finie d'ensembles de mesure nulle est également de mesure nulle, on peut définir la relation d'équivalence

$$f \sim g \text{ sur } E \quad \iff \quad f - g = 0 \text{ p.p. sur } E$$

8.4 Quelques remarques

8.4.1 Sur les fonctions non-mesurables

L'ensemble des fonctions mesurables apparaît comme la fermeture (différent de la complétion) de l'ensemble ces fonctions continues au sens des limites simples.

Une question se pose alors : existe-t-il des fonctions non-mesurable ?

La réponse est oui, mais comme pour les ensembles non-mesurables, ils sont bien cachés.

On pourrait énoncer un pseudo-théorème : tous les ensembles usuels sont mesurables. toutes les fonctions usuelles sont mesurables.

8.4.2 Sur la théorie abstraite de la mesure

Les ensembles mesurables que nous avons définis correspondent par rapport à la théorie classique de la mesure, aux éléments de la tribu borélienne sur $E \subset \mathbb{R}^n$, à savoir la plus petite σ -algèbre contenant les ouverts de E . On l'appelle la mesure de Lebesgue.

Cela signifie que dans la théorie classique les théorèmes 8.2.4 et 8.2.5 apparaissent comme des définitions. La notion de mesure et d'ensemble mesurable est bien plus générale que telle que nous le présentons.

CHAPITRE 9

Intégrale de Lebesgue Théorème de convergence dominée

C'est en 1901 que Henri Lebesgue a introduit sa définition de l'intégrale d'une fonction, dans une note parue aux comptes-rendus de l'Académie des Sciences.

Cette nouvelle construction est essentiellement due au fait que l'intégrale de Riemann présentait nombre de défauts. En particulier, elle est peu adaptée aux intégrales multiples ainsi que pour les passages à la limite.

Autre inconvénient majeur, l'espace des fonctions Riemann intégrables n'est pas complet, si bien que l'analyse sur ces fonctions fait intervenir des fonctions a priori non intégrables.

L'intégrale de Lebesgue basé sur la théorie de la mesure de Émile Borel répond à tous ces problèmes. C'est ce que nous allons exposer.

Dans toute la suite E représente une partie de \mathbb{R}^n . E peut être borné ou non.

9.1 Fonctions étagées

9.1.1 Fonctions élémentaires

Définition 9.1.1. Une fonction élémentaire de E est une fonction valant 1 sur une partie mesurable de E et 0 partout ailleurs. On parle aussi de fonctions indicatrices¹. Par

1. On parle également de fonctions caractéristiques mais nous ne recommandons pas cette appellation

exemple

$$\mathbb{1}_A(x) = \begin{cases} 1 & \text{si } x \in A \\ 0 & \text{si } x \notin A. \end{cases}$$

En général, de telles fonctions peuvent également être définies sur des ensembles non-mesurables. Nous choisissons une telle restriction car les ensembles non mesurables ne nous intéressent pas. Pour ne pas être gênés, nous les excluons tout simplement de notre étude. Cela est légitime car ces ensembles ne jouent aucun rôle dans la théorie de Lebesgue.

9.1.2 Fonctions étagées

Définition 9.1.2. On appelle fonction étagée toute fonction définie par une combinaison linéaire de fonctions élémentaires ou indicatrices.

Ainsi, f est étagée si il existe des ensembles mesurables A_i et des scalaires α_i tel que

$$f(x) = \sum_{i=1}^n \alpha_i \mathbb{1}_{A_i}.$$

Théorème 9.1.3. Les fonctions étagées sont denses dans l'ensemble des fonctions mesurables.

Autrement dit, les fonctions mesurables peuvent être considérées comme des limites simples de fonctions étagées.

Preuve du théorème 9.1.3 – Soit f une fonction mesurable sur E , sans perte de généralités, on suppose que $f \geq 0$.

On définit alors les ensembles mesurables, pour $n \in \mathbb{N}^*$ et $i = 1, 2, 3, \dots, n2^n$:

$$\begin{aligned} E_{ni} &= \{x \in E \mid (i-1)2^{-n} \geq f(x) \geq i2^{-n}\} \\ F_n &= \{x \in E \mid f(x) \geq n\}. \end{aligned}$$

Nous construisons alors la suite de fonctions de fonctions étagées

$$f_n = \sum_{i=1}^{n2^n} (i-1)2^{-n} \mathbb{1}_{E_{ni}} + n \mathbb{1}_{F_n}.$$

Il est facile de voir que pour tout n

$$f_n \geq f_{n+1} \geq \dots \geq f$$

et que par construction, pour tout $x \in E$

$$\lim_{n \rightarrow \infty} f_n(x) = f(x) \quad \square$$

Remarquons que dans la démonstration, la convergence est uniforme si f est bornée.

9.1.3 Intégrale d'une fonction étagée

Définition 9.1.4. Soit φ une fonction étagée :

$$\varphi(x) = \sum_{i=1}^n \alpha_i \mathbb{1}_{A_i}(x).$$

Si ω est un ensemble mesurable, alors

$$I_{\Omega}(\varphi) = \sum_{i=1}^n \alpha_i \mathfrak{m}(A_i \cap \Omega)$$

est l'intégrale de φ sur Ω .

9.1.4 Intégrale de Lebesgue d'une fonction mesurable

La définition 9.1.4 induit la définition de l'intégrale de Lebesgue

Définition 9.1.5. Soit f une fonction mesurable sur E positive et soit Ω une partie mesurable de E .

L'intégrale de Lebesgue de f sur Ω est la borne supérieure des intégrales sur Ω des φ , fonctions étagées positives majorées par f :

$$\int_{\Omega} f = \sup_{\varphi} I_{\Omega}(\varphi)$$

Remarque 11. On note indifféremment :

$$\int_{\Omega} f = \int_{\Omega} f(x) dx$$

où x joue le rôle d'une variable muette.

Proposition 9.1.6. Soit f fonction mesurable sur E positive et soit Ω une partie mesurable de E . Si f_n désigne la suite de fonctions positives étagées introduite dans la démonstration du théorème 9.1.3, alors

$$\int_{\Omega} f = \lim_{n \rightarrow \infty} I_{\Omega}(f_n).$$

Remarque 12. Ainsi définie, l'intégrale peut tout à fait prendre une valeur infinie.

Pour une fonction mesurable f désignons par f^+ et f^- deux fonctions mesurables et positives qui sont respectivement ses parties positive et négatives :

$$\begin{aligned} f^+(x) &= |f(x)| \mathbb{1}_{E^+}(x) \\ f^-(x) &= |f(x)| \mathbb{1}_{E^-}(x) \end{aligned}$$

où

$$\begin{aligned} E^+ &= \{x \in E \mid f(x) \geq 0\} \\ E^- &= \{x \in E \mid f(x) \leq 0\}. \end{aligned}$$

Si bien que pour tout $x \in E$

$$f(x) = f^+(x) - f^-(x).$$

Cette décomposition permet de définir l'intégrale (de Lebesgue) pour toute fonction mesurable :

Définition 9.1.7. Soit f une fonction mesurable sur E , alors pour tout $\Omega \subset E$ mesurable, on définit

$$\int_{\Omega} f = \int_{\Omega} f^+ - \int_{\Omega} f^-.$$

Définition 9.1.8. On dit qu'une fonction mesurable f est Lebesgue intégrable ou sommable sur E si

$$\int_E |f| < +\infty.$$

On note par $\mathcal{L}^1(E)$ l'ensemble des fonctions Lebesgue-intégrable sur E .

Le résultat suivant est fondamental :

Théorème 9.1.9. L'intégrale sur Ω d'une fonction positive f est nulle si et seulement si f est nulle presque partout sur Ω :

$$\int_{\Omega} |f| = 0 \iff f = 0 \text{ p.p. sur } \Omega.$$

9.2 Théorèmes de Lebesgue

9.2.1 Théorème de convergence monotone

Théorème 9.2.1. Soient f_1, \dots, f_n, \dots des fonctions mesurables positives sur E telles que

$$0 \leq f_1 \leq \dots \leq f_n \leq f_{n+1} \leq \dots$$

Posons

$$f(x) = \sup_n f_n(x).$$

Alors

$$\int_E f_n \longrightarrow \int_E f.$$

Preuve du théorème 9.2.1 : Posons

$$\alpha_n = \int_E f_n.$$

Alors, puisque la suite f_n est croissante, les α_n convergent vers une limite

$$\alpha_n \longrightarrow \alpha \in [0, +\infty].$$

Par définition f domine les f_n , ainsi

$$\alpha \leq \int_E f.$$

Soit φ une fonction étagée positive et dominée par f et soit $0 < a < 1$. Si on pose

$$E_n = \{x \in E \mid f_n(x) \geq a\varphi(x)\},$$

il est clair que $E_n \subset E_{n+1}$ et $E = \cup_n E_n$.

D'où pour tout n ,

$$\int_E f_n \geq \int_{E_n} f_n \geq a \int_{E_n} \varphi.$$

Autrement dit, pour tout φ dominé par f et pour tout $0 < a < 1$,

$$\alpha \geq \lim_{n \rightarrow \infty} \int_E f_n \geq a \int_E \varphi \implies \alpha \geq \int_E \varphi \implies \alpha \geq \sup_{\varphi} I_E(\varphi) = \int_E f.$$

Si bien qu'on a montré que

$$\alpha = \int_E f. \quad \square$$

9.2.2 Lemme de Fatou

Une conséquence du théorème de convergence monotone 9.2.1 est le

Lemme 9.2.2. Soient f_n des fonctions mesurables sur E et positives. Si

$$f(x) = \liminf_{n \rightarrow \infty} f_n(x) = \lim_{n \rightarrow \infty} \inf_{i \geq n} \{f_i(x)\},$$

alors

$$\int_E f \leq \lim_{n \rightarrow \infty} \int_E \inf f_n.$$

Preuve du lemme de Fatou : Posons

$$g_n(x) = \inf_{i \geq n} \{f_i(x)\}.$$

Par définition,

$$0 \leq g_1(x) \leq g_2(x) \leq \dots \leq g_n(x).$$

Alors, d'après le théorème 9.2.1 de convergence monotone,

$$\begin{aligned} \int_E f(x) dx &= \lim_{n \rightarrow \infty} \int_E g_n(x) dx \\ &= \lim_{n \rightarrow \infty} \int_E \inf_{i \geq n} \{f_i(x)\} dx \\ &\leq \lim_{n \rightarrow \infty} \inf_{i \geq n} \int_E f_n(x) dx. \quad \square \end{aligned}$$

9.2.3 Théorème de convergence dominée

Théorème 9.2.3. Soient $f_n \in \mathcal{L}^1(E)$, une suite de fonctions sommable sur E telles que

1. $f_n(x) \rightarrow f(x)$ p.p. sur E
2. $|f_n| \leq g \in \mathcal{L}^1(E)$

Alors $f \in \mathcal{L}^1(E)$ et

$$\lim_{n \rightarrow \infty} \int_E f_n(x) dx = \int_E f(x) dx.$$

C'est le résultat central de la théorie de Lebesgue. Il en découle tous les théorèmes de passage à la limite.

Son avantage par rapport à la théorie classique de Riemann, c'est son énoncé valable pour tout type de domaine E pourvu qu'il soit mesurable. Son application est aisée puisqu'il ne demande qu'une domination.

De plus, c'est un résultat optimal

Preuve du théorème 9.2.3 de convergence dominée :

Puisque $|f_n| \leq g \in \mathcal{L}^1(E)$ cela signifie que $f \leq g$ et donc que $f \in \mathcal{L}^1(E)$.

D'autre part, puisque $f_n + g \geq 0$, le lemme de Fatou 9.2.2 indique

$$\int_E f + g \leq \liminf_{n \rightarrow \infty} \int_E f_n + g.$$

Ce qui implique, en retranchant g :

$$\int_E f \leq \liminf_{n \rightarrow \infty} \int_E f_n.$$

Mais on a également que $g - f_n \geq 0$ si bien que

$$\int_E g - f \leq \liminf_{n \rightarrow \infty} \int_E g - f_n,$$

qui entraîne, toujours en retranchant g :

$$\int_E -f \leq \liminf_{n \rightarrow \infty} \int_E -f_n$$

D'où

$$\int_E f \geq \limsup_{n \rightarrow \infty} \int_E f_n.$$

En conclusion, on a montré que

$$\limsup_{n \rightarrow \infty} \int_E f_n \leq \liminf_{n \rightarrow \infty} \int_E f_n$$

Cela prouve que la limite existe et que cette limite vaut

$$\lim_{n \rightarrow \infty} \int_E f_n(x) dx = \int_E f(x) dx. \quad \square$$

9.2.4 Théorème de dérivation sous le signe intégral

Théorème 9.2.4. Soit $f(x, y)$ une fonction mesurable et dérivable par rapport à y pour presque tout $x \in E$, telle que cette dérivée partielle est mesurable dans E . Si

$$\left| \frac{\partial f(x, y)}{\partial y} \right| \leq g \in \mathcal{L}^1(E),$$

alors

$$\frac{\partial}{\partial y} \int_E f(x, y) dx = \int_E \frac{\partial f(x, y)}{\partial y} dx.$$

9.2.5 Théorème de Fubini

Théorème 9.2.5. Soit f une fonction mesurable dans le domaine $\Omega_x \times \Omega_y$. Si $0 \leq f < +\infty$, et si

$$\varphi(x) = \int_{\Omega_y} f(x, y) dy \quad \psi(y) = \int_{\Omega_x} f(x, y) dx$$

alors φ et ψ sont mesurables et

$$\int_{\Omega_x} \varphi(x) dx = \int_{\Omega_y} \psi(y) dy = \iint_{\Omega_x \times \Omega_y} f(x, y) dx dy$$

Théorème 9.2.6. *Formule de Jacobi*

$$\int_{\Omega} f(x) dx = \int_D f(T(u)) |J_T(u)| du$$

où $|J_T(u)|$ est la Jacobienne (déterminant du gradient ou Jacobien) de l'application $T(u) = x$.

9.3 Espaces \mathbb{L}^1 et \mathbb{L}^2

On note \mathbb{L}^1 l'espace \mathcal{L}^1 pour lequel on a identifié toutes fonctions égales presque partout : Deux fonctions f et g sont égales dans \mathbb{L}^1 , si elles sont égales presque partout. On dit que \mathbb{L}^1 est l'espace \mathcal{L}^1 quotientée par la relation d'équivalence "égales presque partout".

Ainsi, dans $\mathbb{L}^1(\Omega)$, l'application

$$f \longrightarrow \int_{\Omega} |f|$$

est une norme.

On notera la norme \mathbb{L}^1 :

$$\|f\|_{\mathbb{L}^1} = \int_{\Omega} |f|$$

Théorème 9.3.1. *Muni de sa norme naturelle, l'espace $\mathbb{L}^1(\Omega)$ est complet. On dit que c'est un espace de Banach.*

En mécanique, un espace fonctionnelle est particulièrement important : l'espace des fonctions de carré intégrable :

$$\mathcal{L}^2(\Omega) = \left\{ f \text{ mesurable sur } \Omega / \int_{\Omega} |f|^2 < \infty \right\}.$$

L'espace quotient est notée de manière analogue par $\mathbb{L}^2(\Omega)$. On parle également d'espace d'énergie finie.

L'espace $\mathbb{L}^2(\Omega)$ est muni d'un produit scalaire

$$(f, g) = \int_{\Omega} f \bar{g}$$

définissant une norme

$$\|f\|_{\mathbb{L}^2} = (f, f)^{1/2} \left(\int_{\Omega} |f|^2 \right)^{1/2}.$$

Théorème 9.3.2. *Muni de son produit scalaire induisant sa norme naturelle, l'espace $\mathbb{L}^2(\Omega)$ est complet. autrement dit, c'est un espace de Hilbert.*

Les espaces de Sobolev dérivent de $\mathbb{L}^2(\Omega)$ et constituent le cadre mathématique des formulations variationnelles ou principe des puissance virtuelles en mécanique.

CHAPITRE 10

Transformation de Fourier des fonctions

10.1 Préliminaires

10.1.1 Quelques intégrales utiles

Commençons par des valeurs remarquables (obtenues facilement en appliquant le théorème de Fubini au carré des intégrales recherchées) :

$$\int_{-\infty}^{+\infty} e^{-x^2} dx = \sqrt{\pi}. \quad (10.1)$$

$$\int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}} dx = \sqrt{2\pi}. \quad (10.2)$$

$$\int_{\mathbb{R}^n} e^{-\frac{|x|^2}{2}} dx = (2\pi)^{\frac{n}{2}}. \quad (10.3)$$

10.1.2 Espace $\mathcal{S}(\mathbb{R}^n)$

Définition 10.1.1. $\mathcal{S}(\mathbb{R}^n)$ est l'ensemble des fonctions $f \in \mathcal{C}^\infty(\mathbb{R}^n)$ telles que $\forall \alpha \in \mathbb{N}^n$, $\forall k \in \mathbb{N}$

$$\lim_{|x| \rightarrow \infty} |x|^k |D_\alpha f(x)| = 0$$

où on note $|x| = \sqrt{\sum_i x_i^2}$, $\alpha = (\alpha_1, \dots, \alpha_n)$ et

$$D_\alpha f(x) = \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \dots \partial x_n^{\alpha_n}}$$

On appelle $\mathcal{S}(\mathbb{R}^n)$ l'ensemble des fonctions à décroissance rapide sur \mathbb{R}^n . On l'appelle également espace de Schwarz. C'est un espace vectoriel.

Exemples.

- $\mathcal{D}(\mathbb{R}^n) \subset \mathcal{S}$, noté également par $\mathcal{C}_0^\infty(\mathbb{R}^n)$.
- $e^{-x^2} \in \mathcal{S}(\mathbb{R})$.
- $e^{-|x|}$ n'est pas dans \mathcal{S} (non différentiable en 0).

Nous avons les définitions équivalentes :

Proposition 10.1.2. \mathcal{S} est l'ensemble des fonctions f définies sur \mathbb{R}^n telles que

1. $\forall \alpha, \forall \beta, \lim_{|x| \rightarrow \infty} |x^\beta \partial^\alpha f(x)| = 0$ (si $\beta = (\beta_1, \dots, \beta_n)$ on pose $x^\beta = x_1^{\beta_1} \dots x_n^{\beta_n}$).
2. $\forall \alpha, \forall k, |x|^k |\partial^\alpha f(x)|$ soit bornée.
3. $\forall \alpha, \forall \beta, |x^\beta \partial^\alpha f(x)|$ soit bornée.
4. $\forall \alpha, \forall k, |x|^k |\partial^\alpha f(x)|$ soit Lebesgue-intégrable.
5. $\forall \alpha, \forall \beta, |x^\beta \partial^\alpha f(x)|$ soit intégrable.

Théorème 10.1.3. \mathcal{S} est dense dans \mathbb{L}^1 et dans \mathbb{L}^2 .

Voir les suites régularisantes.

10.1.3 Produit de convolution

Soit f et g deux fonctions de $\mathbb{L}^1(\mathbb{R}^n)$, alors Le produit de convolution de f et g est la fonction

$$f * g(x) = \int_{\mathbb{R}^n} f(x - y)g(y)dy.$$

Les principales propriétés du produit de convolution sont :

- Le produit est commutatif
- L'effet de régularisation : si $D_\alpha f$ est définie et appartient à $\mathbb{L}^1(\mathbb{R}^n)$, alors

$$D_\alpha(f * g)(x) = D_\alpha f * g(x)$$

10.1.4 Suites régularisantes

Soit la fonction ρ de classe C^∞ à support compact :

$$\rho(x) = \begin{cases} e^{\frac{-1}{1-|x|^2}} & |x| \leq 1 \\ 0 & |x| \geq 1 \end{cases}$$

Alors la suite $\rho_m(x) = \rho(mx)$ est une suite régularisante au sens que si $f \in \mathbb{L}^1(\mathbb{R}^n)$ (resp. \mathbb{L}^2) alors, $f_m = f * \rho_m$ est dans \mathcal{S} et converge vers f dans \mathbb{L}^1 (resp. dans \mathbb{L}^2).

10.2 Transformation de Fourier

Définition 10.2.1. Pour $f \in \mathbb{L}^1(\mathbb{R}^n)$ et $y \in \mathbb{R}^n$ on pose

$$\hat{f}(y) = \int_{\mathbb{R}^n} e^{-ix \cdot y} f(x) dx. \quad (10.4)$$

\hat{f} , notée aussi $\mathcal{F}f$, s'appelle la transformée de Fourier de f . Elle existe car $|e^{-ix \cdot y} f(x)| < |f(x)|$ et $f \in \mathbb{L}^1$.

Théorème 10.2.2. Soit $f \in \mathbb{L}^1$, sa transformée de Fourier \hat{f} est continue, bornée et tend vers 0 à l'infini.

Remarque 13. Pour une fonction f quelconque dans \mathbb{L}^1 , \hat{f} n'est pas nécessairement dans \mathbb{L}^1

Lemme 10.2.3. La transformation de Fourier est "invariante" pour $f(x) = e^{-\frac{|x|^2}{2}}$:

$$\mathcal{F} \left(e^{-\frac{|x|^2}{2}} \right) (y) = (2\pi)^{\frac{n}{2}} e^{-\frac{|y|^2}{2}}.$$

Théorème 10.2.4. Pour toute fonction $f \in \mathcal{S}$ et $\hat{f} \in \mathcal{S}$ et nous avons les relations :

$$\widehat{x^\alpha f} = (iy)^\alpha \hat{f}(y) \quad (10.5)$$

$$\widehat{\partial^\alpha f} = (iy)^\alpha \hat{f} \quad (10.6)$$

$$\int \psi(y) \hat{f}(y) dy = \int \hat{\psi}(y) f(y) dy \quad (10.7)$$

La transformation de Fourier a donc la propriété fondamentale d'échanger les multiplications et les dérivations. Les propriétés précédentes restent valables tant qu'elles ont un sens.

Nous avons le formule d'inversion :

Théorème 10.2.5. Si $f \in \mathbb{L}^1$ et $\hat{f} \in \mathbb{L}^1$, alors

$$f(x) = (2\pi)^{-n} \int_{\mathbb{R}^n} e^{ix \cdot y} \hat{f}(y) dy$$

Remarque 14. si $f \in \mathbb{L}^1$ et $\hat{f} \in \mathbb{L}^1$,

$$\hat{\hat{f}}(x) = (2\pi)^n \check{f}(x) = (2\pi)^n f(-x).$$

Corollaire 10.2.6. Formule de Parseval – Si $\theta \in \mathcal{S}$, $\psi \in \mathbb{L}^1$,

$$\int \psi \hat{\theta} = (2\pi)^{-n} \int \hat{\psi} \bar{\theta}$$

Corollaire 10.2.7. *Formule de Plancherel – si $\psi \in \mathcal{S}$, alors*

$$\int |\psi|^2 = (2\pi)^{-n} \int |\psi|^2$$

Théorème 10.2.8. *Si $f \in \mathbb{L}^1$ et $\hat{f} = 0$, alors $f = 0$.*

Autrement dit La transformation de Fourier est une application injective dans \mathbb{L}^1 .

Remarque 15. *Soit i l'injection de \mathcal{S} dans \mathbb{L}^2 ; d'après la formule de Plancherel, si $f \in \mathcal{S}$, $\|i \circ \mathcal{F} f\|_2 = (2\pi)^{\frac{n}{2}} \|f\|_2$ (norme de \mathbb{L}^2), $i \circ \mathcal{F}$ peut donc s'interpréter comme une application linéaire continue de \mathcal{S} muni de la topologie de \mathbb{L}^2 dans \mathbb{L}^2 . Alors, d'après un théorème de Banach, comme \mathcal{S} est dense dans \mathbb{L}^2 , $i \circ \mathcal{F}$ admet donc un prolongement continu linéaire unique à \mathbb{L}^2 , noté provisoirement $\widetilde{\mathcal{F}} : \mathbb{L}^2 \rightarrow \mathbb{L}^2$, dit transformation de Fourier dans \mathbb{L}^2 . On a donc pour toute suite (f_ν) dans \mathcal{S} convergente vers f dans \mathbb{L}^2*

$$\widetilde{\mathcal{F}} f(y) = \lim_{\nu \rightarrow +\infty} \int e^{-ix \cdot y} f_\nu(x) dx$$

[on pourra par exemple prendre les (f_ν) dans $\mathcal{D}(\mathbb{R}^n)$, ou même la suite régularisée par convolution].

Remarque 16. *Pour $f \in \mathbb{L}^2$, on n'a pas $\widetilde{\mathcal{F}} f(y) = \int e^{-ix \cdot y} f(x) dx$, en général.*

Proposition 10.2.9. *$\widetilde{\mathcal{F}}$ est bijective de \mathbb{L}^2 sur \mathbb{L}^2 .*

Par suite, on note abusivement, pour toute fonctions de \mathbb{L}^2 ,

$$\widetilde{\mathcal{F}} = \hat{\cdot}.$$

Théorème 10.2.10. *La transformation de Fourier est un automorphisme topologique de \mathcal{S} .*

10.3 Rapports avec la convolution

Proposition 10.3.1. *Si $f \in \mathbb{L}^1$ et $g \in \mathbb{L}^2$, $\widehat{f * g} = \hat{f} \hat{g}$*

Corollaire 10.3.2.

1. *f et $g \in \mathcal{S}$, $f * g \in \mathcal{S}$:*

$$\begin{aligned} f, g \in \mathcal{S} &\Rightarrow \hat{f}, \hat{g} \in \mathcal{S} \\ &\Rightarrow \hat{f} \hat{g} \in \mathcal{S} \\ &\Rightarrow \mathcal{F}^{-1}(\hat{f} \hat{g}) = f * g \in \mathcal{S} \end{aligned}$$

\mathcal{S} est une algèbre pour la convolution.

2. $f, g \in \mathcal{S}$, $\widehat{fg} = (2\pi)^{-n} \widehat{f} * \widehat{g}$
il suffit de montrer l'égalité de leurs transformées de Fourier, or

$$\begin{aligned}\widehat{fg} &= (2\pi)^n \check{(fg)} \\ &= (2\pi)^n \check{f} \check{g} \\ &= (2\pi)^{-n} \widehat{\widehat{f} \widehat{g}} \\ &= (2\pi)^{-n} \widehat{\widehat{f} * \widehat{g}}.\end{aligned}$$

la transformation de Fourier dans \mathcal{S} échange multiplication et convolution.

10.4 Exemples

Calculer les transformées de Fourier des fonctions définies sur \mathbb{R} suivantes :

1. $f_1(x) = H(x)e^{-ax}$
2. $f_2(x) = \mathbb{1}_{[-a,a]}$
3. $f_3(x) = H(x)x^k e^{-ax}$
4. $f_4(x) = H(x)$

où $H(x)$ est la fonction de Heaviside :

$$H(x) = \begin{cases} 1 & x \geq 0 \\ 0 & x < 0 \end{cases}$$

CHAPITRE 11

Introduction à la théorie des distributions

11.1 Introduction

C'est à Paris, en 1944, que Laurent Schwartz invente, en une nuit, sa théorie des distributions, théorie qui lui vaudra une médaille Fields en 1950. Comme il le décrit lui-même dans son autobiographie [Sch97], la théorie des distributions était inévitable, et plusieurs théories contemporaines en étaient très proches dont notamment la théorie de Sobolev.

La nécessité d'une théorie allant au delà de la définition classique des fonctions et du calcul différentiel s'imposait par des exemples :

1. **Fonction de Heaviside et sa dérivée**
2. **Ondes sans équations.**

L'équation des ondes

$$\frac{\partial^2 u}{\partial x \partial y} = 0$$

a pour solutions générales $u = f(x) + g(y)$ où f et g sont deux fonctions arbitraires deux fois différentiables. La question naturelle est des savoir si dans le cas où f et g ne sont plus différentiables, on a toujours une onde ou non. L'expérience physique montre que oui, bien que cela n'ait pas de sens mathématique, du moins avec les définitions classiques de fonctions et de dérivées.

11.2 Espace des fonctions infiniment dérivables à support compact

Définition 11.2.1. *Le support d'une fonction continue f est la fermeture (ou l'adhérence) de l'ensemble des x tels que $f(x) \neq 0$. On note :*

$$\text{supp } f = \overline{\{x / f(x) \neq 0\}}$$

Proposition 11.2.2. *Dans \mathbb{R}^n les ensembles compacts sont les ensembles fermés bornés.*

Définition 11.2.3. *L'ensemble des fonctions infiniment dérivables à support compact dans $\Omega \subset \mathbb{R}^n$ est noté $\mathcal{D}(\Omega)$.*

Exemples.

$$\varphi = \begin{cases} e^{\frac{-1}{1-|x|^2}} & |x| < 1 \\ 0 & |x| \geq 1 \end{cases}$$

On munit \mathcal{D} d'une pseudo-topologie :

Définition 11.2.4. *On dit qu'une suite $\varphi_n \in \mathcal{D}$ converge vers 0 si et seulement si $\exists K$ un compact tel que $\text{supp } \varphi_n \in K \forall n$ et :*

$$\begin{aligned} \sup_K |\varphi_n| &\longrightarrow 0 \\ \sup_K |\partial^\alpha \varphi_n| &\longrightarrow 0. \end{aligned}$$

On voit que les conditions de convergence sont "sévères" : on dit que c'est une topologie fine (exigeante).

11.3 Théorie des distributions

Définition 11.3.1. *On note $\mathcal{D}'(\Omega)$ le dual topologique de $\mathcal{D}(\Omega)$. On l'appelle l'espace des distributions sur Ω .*

$$\begin{aligned} \mathcal{D}'(\Omega) &= \{\text{Formes linéaires continues sur } \mathcal{D}(\Omega)\} \\ &= \{T \in \mathcal{L}(\mathcal{D}(\Omega), \mathbb{R}) / \varphi_n \rightarrow \varphi \text{ dans } \mathcal{D}(\Omega) \implies T(\varphi_n) \rightarrow T(\varphi)\} \end{aligned}$$

On note indifféremment

$$T(\varphi) = \langle T, \varphi \rangle.$$

11.3.1 Exemples de distributions

On définit l'ensemble \mathcal{L}_{loc}^1 des fonctions mesurables localement intégrable :

$$\mathcal{L}_{loc}^1 = \left\{ f / \int_K f < +\infty, \forall K \text{ compact} \right\}.$$

Proposition 11.3.2. \mathcal{L}_{loc}^1 contient $\mathcal{L}^1, \mathcal{L}^2, \mathcal{C}^m, \dots$

Proposition 11.3.3. A toute fonction f de \mathcal{L}_{loc}^1 est associée une distribution qu'on notera dans un premier temps $[f]$:

$$\langle [f], \varphi \rangle = \int \mathbb{R}^n f(x) \varphi(x) dx \quad \forall \varphi \in \mathcal{D}.$$

On dit que $[f]$ est une distribution régulière.

Les distributions apparaissent ainsi comme une généralisation des fonctions.

Proposition 11.3.4. On note δ la distribution de Dirac, définie par

$$\langle \delta_0, \varphi \rangle = \varphi(0) \quad \forall \varphi \in \mathcal{D}.$$

$$\langle \delta_a, \varphi \rangle = \varphi(a) \quad \forall \varphi \in \mathcal{D}.$$

11.3.2 Dérivation des distributions

Définition 11.3.5. Soit T une distribution dans $\Omega \subset \mathbb{R}^n$, sa dérivée partielle par rapport à la variable x_i est, par définition, la distribution

$$\left\langle \frac{\partial T}{\partial x_i}, \varphi \right\rangle = - \left\langle T, \frac{\partial \varphi}{\partial x_i} \right\rangle \quad \forall \varphi \in \mathcal{D}.$$

De même, on a pour tout multi-indice α :

$$\langle D^\alpha T, \varphi \rangle = (-1)^{|\alpha|} \langle T, D^\alpha \varphi \rangle \quad \forall \varphi \in \mathcal{D}$$

Ainsi, toute distribution est infiniment dérivable. Par ailleurs, cette définition est une extension de la notion classique de dérivée :

Proposition 11.3.6. Soit $f \in \mathcal{L}_{loc}^1$ et supposons que sa dérivée $D^\alpha f$ existe et soit localement intégrable, alors $[D^\alpha f]$ (distribution associée à la dérivée) coïncide avec $D^\alpha [f]$ (dérivée de la distribution associée).

Proposition 11.3.7. Soit $H(x)$ la fonction de Heaviside, alors $[H(x)]' = \delta_0$. Autrement dit, la dérivée de la fonction de Heaviside est la distribution de Dirac en 0.

11.3.3 Opération sur les distributions

Définition 11.3.8. – *Multiplication par une fonction \mathcal{C}^∞* – Soit $T \in \mathcal{D}'(\Omega)$ et soit $f \in \mathcal{C}^\infty(\Omega)$, alors $fT \in \mathcal{D}'(\Omega)$:

$$\langle fT, \varphi \rangle = \langle T, f\varphi \rangle \quad \forall \varphi \in \mathcal{D}.$$

Remarque 17. Si $T \in \mathcal{D}'(\Omega)$ et $U \in \mathcal{D}'(\Omega)$, le produit n'est pas défini a priori.

Définition 11.3.9. – *Translatée d'une distribution* – Soit $T \in \mathcal{D}'(\mathbb{R}^n)$ et $a \in \mathbb{R}^n$, la translatée de T par a est la distribution

$$\langle \tau_a, \varphi \rangle = \langle T_x, \varphi(x - a) \rangle \quad \forall \varphi \in \mathcal{D}.$$

11.3.4 Convergence au sens des distributions

Définition 11.3.10. On dit qu'une suite de distribution T_n converge vers T au sens des distributions, si et seulement si

$$\lim_{n \rightarrow \infty} \langle T_n, \varphi \rangle = \langle T, \varphi \rangle \quad \forall \varphi \in \mathcal{D}.$$

11.3.5 Support d'une distribution

Définition 11.3.11. On note $\text{supp } T$ le support de la distribution $T \in \mathcal{D}'(\Omega)$ défini par

$$\begin{aligned} \text{supp } T &= \{A \subset \Omega / \text{supp } \varphi \in A \implies \langle T, \varphi \rangle = 0\} \\ &= \complement \{A \subset \Omega / \exists \varphi \in \mathcal{D}(A) \text{ telle que } \langle T, \varphi \rangle \neq 0\} \end{aligned}$$

Exemples. $\text{supp } \delta_0 = \{0\}$.

Proposition 11.3.12. Si une distribution est définie par une fonction f , on a

$$\text{supp } [f] \subset \text{supp } f.$$

Proposition 11.3.13. Si $\text{supp } T = \{0\}$ alors T est une combinaison linéaire de distributions de Dirac en 0 et de ses dérivées.

$$T = \sum_{\alpha} C_{\alpha} \delta_0.$$

11.4 Variantes des distributions

11.4.1 Distribution d'ordre finie

Définition 11.4.1. On note $\mathcal{D}^{k'}(\Omega)$ l'ensemble des distribution d'ordre $k \in \mathbb{N}$ le dual topologique de $\mathcal{C}^k(\Omega)$.

Exemples. δ_0 est d'ordre 0, mais δ'_0 est d'ordre 1.

11.4.2 Distribution à support compact

Définition 11.4.2. On note $\mathcal{E}'(\mathbb{R}^n)$ l'ensemble des distributions à support compact.

Proposition 11.4.3. $\mathcal{E}'(\mathbb{R}^n)$ est le dual topologique de \mathcal{C}^∞ .

Exemples.

11.4.3 Distribution tempérée

On rappelle que \mathcal{S} , l'espace de Schwartz est l'ensemble des fonctions à décroissance rapide voir la section 10.1.2

Définition 11.4.4. On note $\mathcal{S}'(\mathbb{R}^n)$ l'ensemble des distributions tempérées, le dual topologique de $\mathcal{S}(\mathbb{R}^n)$.

Exemples.

11.5 Produit de convolution de deux distributions

Définition 11.5.1. Soit $T \in \mathcal{D}'$ une distribution et $\varphi \in \mathcal{D}$, le produit de convolution est la fonction défini par

$$T * \varphi(x) = \langle T_y, \varphi(x - y) \rangle.$$

Exemples.

1. $\delta_0 * \varphi(x) = \varphi(x)$.
2. si $f \in \mathcal{L}_{loc}^1$, alors $[f] * \varphi(x) = f * \varphi(x)$.

Proposition 11.5.2. Si $T \in \mathcal{D}'$ et si $\varphi \in \mathcal{D}$ alors $T * \varphi(x)$ est une fonction \mathcal{C}^∞ . De plus, on a

$$\partial^\alpha (T * \varphi)(x) = \partial^\alpha T * \varphi(x) = T * \partial^\alpha \varphi(x).$$

Proposition 11.5.3. Si $T \in \mathcal{E}'$ et si $\varphi \in \mathcal{D}$, $T * \varphi(x)$ est une fonction à support compact.

Théorème 11.5.4. \mathcal{C}^∞ est séquentiellement dense dans \mathcal{D}' .

Preuve : par le produit de convolution avec la suite régularisante.

Théorème 11.5.5. Si $T \in \mathcal{D}'$ et $T' = 0$, alors T est une constante.

CHAPITRE 12

Petite introduction à l'analyse fonctionnelle

Même prudent, le titre est trompeur, il ne s'agit pas véritablement d'analyse fonctionnelle mais, pour le moment, une collection de résultats théoriques concernant des espaces de fonctions usuels en mathématiques appliquées et en mécanique en particulier. Les résultats qui suivent sont difficiles et peuvent être survolés sans conséquences en Licence de Mécanique, mais pourront être acquises au cours de la maîtrise.

Il est essentiel, avant de vouloir résoudre un problème, de savoir s'il est bien posé ou non : savoir il existe une solution au problème posé et si oui, savoir si il existe une ou plusieurs solutions. Dans un cadre mécanique cela revient à savoir si la modélisation a un sens ou non. En effet, il est vain de vouloir trouver une solution si elle n'existe pas et si il existe plusieurs solutions, laquelle choisir ?

En bref, comment savoir si les équations, les conditions aux limites d'un problème sont bien écrites, s'il n'en manque pas, s'il ne sont pas contradictoires.

A ces questions, seule l'analyse mathématique permet de répondre, en classant notamment les problèmes, en les identifiant et en les regroupant par les propriétés des solutions et de leur dépendance aux données.

L'intérêt de ce cadre mathématique est une maîtrise relative d'un cadre théorique de la modélisation en mécanique : maîtrise des théorèmes d'existence, d'unicité et de la dépendances des solutions aux données du problème, maîtrise également de l'approximation.

Naturellement, un résultat d'existence ne donne pas a priori d'indication sur la manière de rechercher la solution d'un problème. C'est là une des grandes difficultés de ce

cadre théorique : des énoncés difficiles, abstraits, pour une pratique concrète quasi nulle au métier d'ingénieur.

Enfin, c'était sans compter les ingénieurs qui ont mis au point la méthode des éléments-finis, basés sur ce cadre théorique, permettant d'approximer les solutions d'un grand nombre de problème mécanique.

12.1 Espace de Banach

Définition 12.1.1. Soit E un espace vectoriel normé, on dit que c'est un espace de Banach s'il est complet.

Autrement dit, E est complet si toute suite de Cauchy dans E est convergente.

La notion d'espace complet induit des résultats d'existence. C'est la raison essentielle pour laquelle on fait tant d'effort pour travailler dans des espaces de fonctions ayant cette propriété.

12.1.1 Quelques caractérisations abstraites d'un espace de Banach

Théorème 12.1.2. Soit E un espace vectoriel normé, alors E est complet si et seulement si les séries absolument convergentes (resp. normalement) sont convergentes.

Théorème 12.1.3. Soit E un espace vectoriel normé complet et soit un sous espace vectoriel $F \subset E$. Si F est fermé dans E , alors F est lui même complet pour la norme induite.

Théorème 12.1.4. Soit E un espace de Banach, alors son dual topologique E' , c'est à dire l'ensemble des formes linéaires continues sur E est aussi un Banach pour la norme

$$\|A\|_{E'} = \sup_{\|f\|_E=1} |A(f)|$$

Les espaces de Banach permettent de définir simplement la notion de fonction continue et par suite le notion de différentielle :

Définition 12.1.5. Soit E et F deux espaces de Banach réel, et soit f une fonction définie d'une partie ouverte $U \subset E$ dans F . Alors f est continue en $x \in U$ si et seulement si $f(x+h)$ converge vers $f(x)$ dans F , lorsque h converge vers 0 dans E :

$$f(x+h) = f(x) + O(h)$$

Définition 12.1.6. Soit E et F deux espaces de Banach réel, et soit f une fonction définie d'une partie ouverte $U \subset E$ dans F . Alors f est différentiable en $x \in U$ si et seulement si il existe une application linéaire A de E dans F (on note $A \in \mathcal{L}(E, F)$) telle que

$$f(x+h) = f(x) + Ah + o(h)$$

A est la dérivée de f en x , on la note généralement $f'(x)$.

12.1.2 Quelques exemples d'espaces de Banach

Tout espace vectoriel réel ou complexe de dimension finie est un espace complet pour n'importe quelle norme.

$\mathbb{L}^1(\Omega)$ est complet pour tout Ω mesurable avec la norme

$$\|f\|_{\mathbb{L}^1(\Omega)} = \int_{\Omega} |f(x)| dx.$$

$\mathbb{L}^2(\Omega)$ est complet pour tout Ω mesurable avec le norme

$$\|f\|_{\mathbb{L}^2(\Omega)} = \left(\int_{\Omega} |f(x)|^2 dx \right)^{\frac{1}{2}}.$$

$\mathbb{L}^p(\Omega)$, $p \geq 1$ est complet pour tout Ω mesurable avec le norme

$$\|f\|_{\mathbb{L}^p(\Omega)} = \left(\int_{\Omega} |f(x)|^p dx \right)^{\frac{1}{p}}.$$

$C^1([0, 1])$ est complet pour sa norme usuelle

$$\|f\|_{C^1} = \sup_{[0,1]} |f(x)| + \sup_{[0,1]} |f'(x)|.$$

Cela se généralise pour l'ensemble des fonctions de classe C^1 définies sur un compact de \mathbb{R}^n .

12.1.3 Quelques exemples d'espaces qui ne sont pas des espaces de Banach

$C^1([0, 1])$ n'est pas complet pour la norme

$$\|f\|_{\infty} = \sup_{[0,1]} |f(x)|.$$

$C^{\mathbb{R}}$ n'est même pas normable.

$C^1(]0, 1[)$ n'est même pas normable.

12.2 Espace de Hilbert

Les espaces de Hilbert sont très importants car ils sont à la base des théorèmes d'existence pour les problèmes issues de la Mécanique, notamment en ce qui concerne les formulations variationnelles (ou principe des puissances virtuelles) qui entrent dans le cadre de la théorie de Lax-Milgram.

Définition 12.2.1. Soit H un espace vectoriel. Un produit scalaire sur H est une forme bilinéaire définie positive symétrique $a(u, v)$ de $H \times H$ à valeur dans \mathbb{R} . Une telle forme bilinéaire induit une norme

$$\|u\|_a = \{a(u, u)\}^{\frac{1}{2}}.$$

Définition 12.2.2. Soit H un espace vectoriel muni d'un produit scalaire, on dit que H est un espace de Hilbert s'il est complet pour la norme associée au produit scalaire.

Ainsi, un espace de Hilbert est un cas particulier d'espace de Banach.

12.2.1 Quelques exemples d'espaces de Hilbert

$\mathbb{L}^2(\Omega)$ est un Hilbert pour tout Ω mesurable avec le produit scalaire :

$$(f, g)_{\mathbb{L}^2} = \int_{\Omega} f(x)\overline{g(x)}dx$$

$H^1(\Omega)$ est un Hilbert pour tout Ω mesurable avec le produit scalaire :

$$(f, g)_{\mathbb{H}^1} = \int_{\Omega} f(x)\overline{g(x)}dx + \int_{\Omega} \nabla f(x) \cdot \overline{\nabla g(x)}dx$$

12.2.2 Principaux résultats sur les Hilberts

Théorème 12.2.3. Représentation de Riesz – Soit L une forme linéaire continue définie sur un Hilbert H . Il existe un unique $f \in H$ tel que

$$(f, v)_H = L(v) \forall v \in H.$$

Théorème 12.2.4. Soit H un espace de Hilbert séparable, alors il existe une suite $e_n \in H$ telle que

1. $\|e_n\|_H = 1$
2. $(e_n, e_m)_H = \delta_{n,m}$
3. l'espace vectoriel engendré par les e_n est dense dans H :

$$\forall u \in H \quad u = \sum_{n=0}^{\infty} a_n e_n$$

avec

$$a_n = (u, e_n)_H$$

et

$$\|u\|_H = \left\{ \sum_{n=0}^{\infty} a_n^2 \right\}^{\frac{1}{2}}.$$

Index

- x , 52
- Application linéaire, 16
- Base d'un espace vectoriel, 15
- Changement de variables, 51
- Convention d'Einstein, 8
- Dérivée
 - ,fonction, 42
 - seconde, 42
- Déterminant, 24
- Divergence, 39, 53
- Endomorphisme, 16, 23
- Ensemble
 - fermé, 32
 - ouvert, 32
- Espace
 - de dimension finie, 15
 - vectoriel, 11
- Famille
 - liée, 14
 - libre, 14
- Fonction
 - continue, 34
 - dérivable, 34
 - différentiable, 34, 36
- Formule
 - de Green, 58
 - de Jacobi, 51
 - de Taylor, 43
- Formules
 - d'intégration par parties, 53
- Gradient, 39, 46
- homomorphisme, 16
- Méthode
 - des éléments-finis, 53
- Matrice
 - carré, 23
 - d'une application linéaire, 18
 - de passage, 22
 - des cofacteurs, 26
 - jacobienne, 38
 - symétrique, 29
- Norme, 33
- Norme
 - d'une matrice, 20
- Polynôme caractéristique, 28
- Spectre, 27
- Stokes, 58
- Théorème
 - de Cayley-Hamilton, 29

de la moyenne, 43

de Schwarz, 43

du point fixe, 43

Valeur propre, 27

Vecteur propre, 27

Bibliographie

- [Car85] Henri Cartan. *Cours de calcul différentiel*. Hermann, 1985.
- [Dix76] Jacques Dixmier. *Cours de Mathématiques du premier cycle*. Gauthiers-Villars, 1976.
- [Lic46] André Lichnerowitz. *Éléments de Calcul tensoriel*. Jacques Gabay, 1946.
- [Que64] Michel Queysanne. *Algèbre – 1er cycle scientifique – préparation aux grandes écoles*. Armand Colin – Collection U, 1964.
- [Rud95] Walter Rudin. *Principes d'analyse Mathématiques*. Ediscience international, 1995. ISBN : 2-84074-108-3.
- [Sch67] Laurent Schwartz. *Cours d'analyse*. Hermann, 1967.
- [Sch97] Laurent Schwartz. *Un mathématicien aux prises avec le siècle*. Éditions Odile Jacob, 1997.