

Cours de mathématiques
Partie III – Algèbre
MPSI 4

Alain TROESCH

Version du:

4 juin 2015

Table des matières

16 Fang cheng, ou l'élimination de Gauss-Jordan...	7
I Position du problème et reformulation matricielle	8
I.1 Position du problème	8
I.2 Rappels sur les matrices et transcription matricielle du système	8
I.3 Structure de l'ensemble des solutions	10
II Échelonnement d'une matrice par la méthode du pivot	10
II.1 Opérations sur les lignes d'une matrice	10
II.2 Échelonnement de la matrice	11
III Résolution d'un système échelonné	13
III.1 Inconnues principales	13
III.2 Recherche d'une solution particulière	14
III.3 Recherche de la solution générale de l'équation homogène associée	14
17 Structures algébriques	17
I Lois de composition	17
I.1 Définitions	17
I.2 Propriétés d'une loi de composition	18
I.3 Ensembles munies de plusieurs lois	21
I.4 Stabilité	22
II Structures	22
II.1 Généralités	22
II.2 Morphismes	23
II.3 Catégories (HP)	24
III Groupes	25
III.1 Axiomatique de la structure groupes	25
III.2 Exemples importants	26
III.3 Sous-groupes	26
III.4 Congruences modulo un sous-groupe	28
III.5 Ordre d'un élément	29
IV Anneaux et corps	30
IV.1 Axiomatiques des structures d'anneaux et de corps	30
IV.2 Sous-anneaux, sous-corps	32
IV.3 Calculs dans un anneau	33
IV.4 Éléments inversibles	35
IV.5 Idéaux (HP)	36

18 Arithmétique des entiers	37
I Divisibilité, nombres premiers	38
I.1 Notion de divisibilité	38
I.2 Congruences	40
I.3 Nombres premiers	40
II Décomposition primaire d'un entier	42
II.1 Décomposition primaire	42
II.2 Valuations p -adique	42
III PGCD et PPCM	44
III.1 PGCD et PPCM d'un couple d'entiers	44
III.2 Identité de Bézout	45
III.3 PGCD et PPCM d'une famille finie d'entiers	47
III.4 PGCD et PPCM vus sous l'angle de la décomposition primaire	48
IV Entiers premiers entre eux	49
IV.1 Couple d'entiers premiers entre eux	49
IV.2 Famille finie d'entiers premiers entre eux	50
IV.3 Fonction indicatrice d'Euler	51
V Théorème des restes chinois	51
V.1 Cas de modulo premiers entre eux	51
V.2 Résolution d'un système quelconque	53
19 Polynômes et fractions rationnelles	55
I Polynômes à coefficients dans un anneau commutatif	55
I.1 Polynômes formels	55
I.2 Opérations arithmétiques sur les polynômes	56
I.3 Indéterminée formelle	57
I.4 Dérivation	58
I.5 Degré et valuation	59
II Arithmétique dans $\mathbb{K}[X]$	61
II.1 Division euclidienne	62
II.2 Idéaux de $\mathbb{K}[X]$	62
II.3 Divisibilité	63
II.4 PGCD et PPCM	63
II.5 Polynômes premiers entre eux	65
II.6 Décomposition en facteurs irréductibles	66
III Racines d'un polynôme	67
III.1 Spécialisation, évaluation	67
III.2 Racines et multiplicité	69
III.3 Majoration du nombre de racines	70
III.4 Interpolation de Lagrange	71
III.5 Polynômes scindés	72
IV Polynômes irréductibles dans $\mathbb{C}[X]$ et $\mathbb{R}[X]$	73
IV.1 Factorisations dans $\mathbb{C}[X]$	73
IV.2 Facteurs irréductibles dans $\mathbb{R}[X]$	73
V Fractions rationnelles	74
V.1 Définition des fractions rationnelles formelles	74
V.2 Degré, racines, pôles	75
V.3 Décomposition en éléments simples dans $\mathbb{C}(X)$	77
V.4 Décomposition en éléments simples dans $\mathbb{R}[X]$	78

20	Espaces vectoriels	79
I	Notion d'espace vectoriel	80
I.1	Définition	80
I.2	Combinaisons linéaires	80
I.3	Un exemple important : espace de fonctions	81
I.4	Produits d'espaces vectoriels	81
I.5	Sous-espaces vectoriels	82
I.6	Intersections de sev	83
I.7	Sous-espace vectoriel engendré par un sous-ensemble	83
I.8	Sommes de sev	84
I.9	Sommes directes	85
II	Familles de vecteurs	86
II.1	Familles libres	86
II.2	Familles génératrices	87
II.3	Bases	88
III	Espaces vectoriels de dimension finie	88
III.1	Notion de dimension	88
III.2	Dimension, liberté et rang	90
III.3	Dimension de sous-espaces et de sommes	91
21	Applications linéaires	93
I	Généralités sur les applications linéaires	93
I.1	Définitions et propriétés de stabilité	93
I.2	Image et noyau	95
I.3	Endomorphismes	97
I.4	Automorphisme	99
I.5	Projecteurs et symétries	100
II	Applications linéaires et familles de vecteurs	101
II.1	Détermination d'une application linéaire	101
II.2	Caractérisations de l'injectivité et de la surjectivité par l'image de bases	102
II.3	Recollements	103
III	Applications linéaires en dimension finie	104
III.1	Rang d'une application linéaire	104
III.2	Théorème du rang	105
IV	Formes linéaires	106
IV.1	Formes linéaires, espace dual, hyperplan	106
IV.2	Qu'est-ce que le principe de dualité? (hors-programme)	107
22	Matrices	109
I	Matrice d'une application linéaire et opérations matricielles	109
I.1	L'ensemble des matrices de type (n, p)	110
I.2	Matrice d'une application linéaire	111
I.3	Structure d'espace vectoriel de $\mathcal{M}_{n,p}(\mathbb{K})$	112
I.4	Définition du produit matriciel	113
I.5	Expression matricielle de l'évaluation d'une AL	115
I.6	Transposition	116
I.7	Produit matriciel revisité	116
II	Matrices carrées	119
II.1	L'algèbre $\mathcal{M}_n(\mathbb{K})$	119
II.2	Matrices carrées de type particulier	120
II.3	Matrices inversibles	123
II.4	Expression matricielle du pivot de Gauss	124

	II.5	Calcul pratique de l'inverse d'une matrice	125
III		Rang d'une matrice	127
	III.1	Image et noyau d'une matrice	127
	III.2	Calcul du rang	128
	III.3	Caractérisation du rang par les matrices extraites	129
IV		Changements de base	129
	IV.1	Changements de base pour des applications linéaires	130
	IV.2	Matrices équivalentes	131
	IV.3	Matrice d'un endomorphisme, matrices semblables	132
	IV.4	Matrices semblables	132
	IV.5	Trace d'une matrice, trace d'un endomorphisme	133
	IV.6	Introduction à la réduction des endomorphismes (Spé)	135
V		Produit matriciel par blocs	136
23 Groupe symétrique et déterminants			139
I		Groupe symétrique	139
	I.1	Notations liées à des permutations	139
	I.2	Signature d'une permutation	141
	I.3	Décomposition cyclique d'une permutation	143
	I.4	Cycles et signature	145
II		Déterminants	145
	II.1	Formes multilinéaires	145
	II.2	Formes n -linéaires symétriques, antisymétriques, alternées	147
	II.3	Déterminant d'une famille de vecteurs	148
	II.4	Orientation d'un espace	150
	II.5	Déterminant d'un endomorphisme	152
	II.6	Déterminant d'une matrice carrée	153
III		Calcul des déterminants	154
	III.1	Opérations sur les lignes et colonnes	155
	III.2	Calcul par blocs	155
	III.3	Développements suivant une ligne ou une colonne	156
	III.4	Caractère polynomial du déterminant	157
24 Espaces préhilbertiens réels			159
I		Produits scalaires	159
	I.1	Formes bilinéaires	159
	I.2	Matrice d'une forme bilinéaire	160
	I.3	Produits scalaire	162
	I.4	Formes bilinéaires symétriques, définies, positives	162
	I.5	Produits scalaires	163
	I.6	Normes euclidiennes	165
	I.7	Espaces préhilbertiens réels, espaces euclidiens	166
II		Orthogonalité	166
	II.1	Vecteurs orthogonaux	166
	II.2	Sous-espaces orthogonaux	168
	II.3	Projeté orthogonal	169
	II.4	Orthonormalisation de Schmidt	170
III		Espaces euclidiens	171
	III.1	Bases orthonormales d'un espace euclidien	171
	III.2	Changements de base et matrices orthogonales	172
	III.3	Projecteurs orthogonaux	174
	III.4	Distance d'un point à un sous-espace vectoriel	175

IV	Géométrie affine	175
IV.1	Sous-espaces affines d'un espace vectoriel	175
IV.2	Définition d'un hyperplan par vecteur normal	178
25	Isométries vectorielles	181
I	Isométries d'un espace euclidien	181
II	Isométries vectorielles en dimension 2	183
II.1	Description de $O(2)$	183
II.2	Isométries positives en dimension 2	183
II.3	Isométries négatives en dimension 2	185

Fang cheng, ou l'élimination de Gauss-Jordan, ou l'art de pivoter

On a donc autant d'équations linéaires qu'il n'y a d'inconnues à trouver ; les valeurs de ces inconnues seront obtenues par l'élimination ordinaire.

Voyons maintenant, si cette élimination est toujours possible, ou si la solution peut quelquefois devenir indéterminée ou même impossible

Carl Friedrich Gauss (traduction Edmond Dubois)

Ce sont les tournesols, ce merveilleux cadeau

D'origine céleste et de divine essence.

Et c'est en pivotant tous dans le même sens

Qu'ils adorent leur père sans lui tourner le dos.

(Vette de Fonclare)

Et les shadoks pivotaient, pivotaient, pivotaient...

(Libre adaptation, d'après Jacques Rouxel)

Introduction

Ce chapitre est consacré à la résolution d'équations linéaires. Ces équations interviennent dans de nombreux problèmes d'algèbre linéaire, il est donc important d'avoir une démarche permettant d'arriver au bout des calculs de façon ordonnée et méthodique. L'algorithme du pivot de Gauss fournit une telle méthode, qui de plus, par son aspect algorithmique simple, a l'avantage de pouvoir être très facilement implémentée sur un ordinateur.

Note Historique 16.0.1 (Équations)

La recherche de solutions d'équation n'est pas un problème récent :

- À Babylone et en Égypte (2e millénaire avant J.-C.), on trouve déjà trace de résolutions de problèmes se ramenant à des équations de degré 2. Les méthodes de résolution sont exposées sur des exemples concrets, mais sont celles utilisées actuellement (mise sous forme canonique).
- Vers 300 après J.-C., Diophante formalise la notion d'équations. Il s'intéresse en particulier à la recherche de solutions rationnelles d'équations à coefficients rationnels, ce qu'on appelle actuellement des *équations diophantiennes*.
- Vers 800 après J.-C., le mathématicien arabe Al Khwarizmi écrit le premier traité de résolution systématique des équations de degré 2. Le titre de ce traité est *kitab al-mukhtasar fi hisabi al-jabr wa'l-muqabalah*, soit, à peu près : *Abrégé du calcul par la réduction et la comparaison*

- Le terme arabe « Al-jabr » signifie « par réduction » (ie « en se ramenant à des situations-type par manipulation des termes »). Il a donné naissance au terme « algèbre ».

Nous nous intéressons ici aux équations linéaires à n indéterminées, autrement dit aux systèmes d'équations linéaires. La méthode que nous exposons ici est celle appelée couramment « méthode d'élimination de Gauss-Jordan », ou encore « méthode du pivot de Gauss », mais ses origines remontent à des temps bien plus anciens

Note Historique 16.0.2 (Pivot de Gauss)

- Le nom de la méthode du pivot est un hommage aux deux mathématiciens Gauss et Jordan.
- Gauss utilise cette méthode dans ses ouvrages, en l'appelant *élimination ordinaire*, ou, en latin (langue qu'il emploie pour ses publications scientifique), *eliminatio vulgaris*
- Gauss et Jordan utilisent cette méthode d'élimination ordinaire notamment dans le cadre de la classification des formes quadratiques.
- Ce n'est que vers 1880 que Frobenius publie plusieurs mémoires faisant un état des lieux de la théorie des matrices, et élucide complètement à l'occasion la théorie des systèmes linéaires à coefficients réels ou complexes.
- Mais la méthode est en fait beaucoup plus ancienne : elle est déjà exposée dans un ouvrage chinois du III^e siècle *Jiuzhang suanshu* (Prescriptions de calcul en 9 chapitres) de Liu Hui. Le huitième chapitre est entièrement consacré à la méthode d'élimination par pivot, appelée *fang cheng* (disposition, ou modèle rectangulaire).
- La méthode elle-même est sans doute plus ancienne, puisque Liu Hui en attribue la paternité à Chang Ts'ang, 3 ou 4 siècles plus tôt, auteur d'un ouvrage aujourd'hui disparu.

I Position du problème et reformulation matricielle

I.1 Position du problème

Nous nous intéressons à la résolution d'un système de n équations à p inconnues réelles (ou complexes), les équations étant toutes linéaires, c'est-à-dire s'écrivant sous forme d'une combinaison linéaire des inconnues (combinaison à valeurs dans \mathbb{R} ou \mathbb{C}) :

$$\begin{cases} a_{1,1}x_1 + \cdots + a_{1,p}x_p = b_1 \\ a_{2,1}x_1 + \cdots + a_{2,p}x_p = b_2 \\ \cdots \\ a_{n,1}x_1 + \cdots + a_{n,p}x_p = b_n \end{cases} \quad (16.1)$$

Les x_i sont les *inconnues* du système, les $a_{i,j}$ les *coefficients*, et les b_i constituent le *second membre*.

L'idée principale de l'algorithme du pivot de Gauss est de se ramener, par des combinaisons de lignes, à un système échelonné équivalent, c'est à dire un système où l'inconnue de plus petit indice apparaissant dans une ligne n'apparaît plus dans les lignes suivantes. Par exemple, un système triangulaire est échelonné. Un tel système est facile à résoudre en partant par le bas.

I.2 Rappels sur les matrices et transcription matricielle du système

Le système (16.1) peut se réécrire matriciellement. Pour ce faire, rappelons quelques faits à propos des matrices. Nous notons $\mathbb{K} = \mathbb{R}$ ou \mathbb{C} , les descriptions étant semblables dans \mathbb{R} et dans \mathbb{C} .

Définition 16.1.1 (Matrices)

Une matrice à coefficients dans \mathbb{K} est une famille rectangulaire d'éléments de \mathbb{K} , autrement dit une famille doublement indexée :

$$A = (a_{i,j})_{(i,j) \in [1,n] \times [1,p]}$$

pour certaines valeurs entières strictement positives n et p . On représente cette matrice de la façon suivante :

$$A = \begin{pmatrix} a_{1,1} & \cdots & a_{1,p} \\ \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,p} \end{pmatrix}$$

Dans la description $(a_{i,j})_{(i,j) \in \llbracket 1,n \rrbracket \times \llbracket 1,p \rrbracket}$, l'indice i est appelé indice de ligne et l'indice j est appelé indice de colonne. L'entier n représente le nombre de lignes de la matrice, et p le nombre de colonnes. Enfin, on note $\mathcal{M}_{n,p}(\mathbb{K})$ l'ensemble des matrices à n lignes et p colonnes à coefficients dans \mathbb{K} , et si $n = p$, on note simplement $\mathcal{M}_n(\mathbb{K})$, dont les éléments sont appelés matrices carrées d'ordre n .

La somme matricielle est définie pour des matrices de même taille uniquement, coefficient par coefficient.

Nous rappelons également l'expression du produit de 2 matrices, possible uniquement lorsque le nombre de colonnes de la première est égal au nombre de lignes de la seconde.

Définition 16.1.2 (Produit matriciel)

Soit $A \in \mathcal{M}_{n,p}(\mathbb{K})$ et $B \in \mathcal{M}_{p,q}(\mathbb{K})$. On définit le produit $AB \in \mathcal{M}_{n,q}(\mathbb{K})$ par :

$$AB = (c_{i,k})_{(i,k) \in \llbracket 1,n \rrbracket \times \llbracket 1,q \rrbracket}, \text{ où } c_{i,k} = \sum_{j=1}^p a_{i,j} b_{j,k}.$$

Ainsi, pour obtenir l'élément en position (i,k) du produit AB , on considère la i^{e} ligne de A et la k^{e} colonne de B , et on forme la somme des produits coefficients par coefficients des coordonnées de cette ligne et cette colonne.

En d'autres termes, $c_{i,k}$ est le résultat du produit scalaire canonique dans \mathbb{R}^n de la colonne obtenue en redressant la i^{e} ligne de A et de la k^{e} colonne de B .

Nous admettons à ce stade toutes les règles usuelles sur les opérations matricielles, notamment les propriétés d'associativité et de distributivité, similaires à celles des réels. Attention cependant au fait que le produit matriciel n'est pas commutatif, et qu'une égalité $MX = MY$ n'implique pas toujours $X = Y$, même si $M \neq 0$ (la bonne condition est l'inversibilité de M ; plus généralement, dans une structure algébrique, la possibilité de faire une telle simplification définit la notion d'élément régulier; tout élément inversible est régulier).

Le système (16.1) peut être traduit par une égalité matricielle $AX = B$, où

$$\begin{pmatrix} a_{1,1} & \cdots & a_{1,p} \\ \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,p} \end{pmatrix} \times \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_p \end{pmatrix}.$$

Ainsi, $A = (a_{i,j})_{(i,j) \in \llbracket 1,n \rrbracket \times \llbracket 1,p \rrbracket}$, $X = (x_i)_{i \in \llbracket 1,p \rrbracket}$ et $B = (b_i)_{i \in \llbracket 1,n \rrbracket}$. On a donc, en notant $\mathbb{K} = \mathbb{R}$ ou \mathbb{C} :

$$A \in \mathcal{M}_{n,p}(\mathbb{K}), \quad B \in \mathcal{M}_{n,1}(\mathbb{K}) \simeq \mathbb{K}^n \quad \text{et} \quad X \in \mathcal{M}_{p,1}(\mathbb{K}) \simeq \mathbb{K}^p.$$

L'intérêt de cette présentation matricielle est d'une part la rapidité apportée par le fait qu'on se dispense de réécrire les variables, et d'autre part une présentation plus claire, du fait de l'alignement oblique des coefficients dans la matrice. Ce n'est rien de plus que le principe de la disposition rectangulaire des chinois. Une disposition méthodique de la sorte supprime une source importante d'erreurs d'inattention.

Dans ce qui suit, A désigne une matrice à n lignes et p colonnes, B une matrice colonne à n lignes, et X la matrice colonne des inconnues, à p lignes.

I.3 Structure de l'ensemble des solutions

Nous nous donnons dans la suite de ce cours une matrice $A \in \mathcal{M}_{n,p}(\mathbb{K})$ et B un vecteur colonne de $\mathcal{M}_{n,1}(\mathbb{K})$, et nous nous intéressons à l'équation $AX = B$, de l'inconnue $X \in \mathcal{M}_{p,1}(\mathbb{K})$. Ainsi, l'équation $AX = B$ représente un système linéaire de n équations à p inconnues.

Définition 16.1.3 (Système homogène associé)

Étant donné le système donné matriciellement par $AX = B$, on appelle système homogène associé au système $AX = B$ le système $AX = 0$.

Nous renvoyons au chapitre concernant les équations différentielles pour la notion de sous-espace affine, que nous retrouvons dans la situation présente :

Théorème 16.1.4 (Structure de l'ensemble des solutions)

L'ensemble des solutions \mathcal{S} de l'équation $AX = B$ est un sous-espace affine de \mathbb{K}^p . Ainsi, si X_0 désigne une solution particulière, et si \mathcal{S}_0 désigne l'ensemble des solutions de l'équation homogène associée $AX = 0$, alors :

$$\mathcal{S} = \{X_0 + X \mid X \in \mathcal{S}_0\} = X_0 + \mathcal{S}_0.$$

II Échelonnement d'une matrice par la méthode du pivot

II.1 Opérations sur les lignes d'une matrice

Étant donné une matrice de $\mathcal{M}_{n,p}(\mathbb{K})$ dont les lignes sont désignées par L_1, \dots, L_n , les opérations admissibles pour le pivot sont les trois opérations suivantes décrites ci-dessous

Définition 16.2.1 (Opérations admissibles sur les lignes d'une matrice)

- Opération de permutation : échange des lignes L_i et L_j de la matrice, codée par $L_i \leftrightarrow L_j$
- Opération de dilatation : multiplication d'une ligne L_i par un scalaire (réel ou complexe) *non nul* λ , codée par $L_i \leftarrow \lambda L_i$
- Opération de transvection : ajout à une ligne donnée L_i d'une autre ligne L_j éventuellement multipliée par un scalaire λ (le résultat remplaçant la ligne L_i). Cette opération est codée par $L_i \leftarrow L_i + \lambda L_j$.

Remarque 16.2.2

- L'itération de la première opération (permutation) autorise les permutations quelconques des lignes d'une matrice. En effet, nous justifierons plus tard que toute permutation de \mathcal{S}_n peut s'écrire comme composition de transpositions, c'est-à-dire des permutations particulières n'effectuant que l'échange de deux valeurs.
- La combinaison des deux dernières règles amène la règle suivante souvent bien pratique :

$$L_i \leftarrow \lambda L_i + \mu L_j, \text{ si } \lambda \neq 0.$$

Attention à ce que le coefficient devant la ligne modifiée par cette opération ne soit pas nul !

Avertissement 16.2.3

Les différentes opérations s'effectuent successivement (même si on les note ensemble dans la même étape) : on ne peut pas effectuer des opérations simultanées. Ainsi, si on a dans la même étape deux opérations $L_1 \leftarrow L_1 + L_2$ et $L_2 \leftarrow L_1 + L_2$, cela signifie que la seconde est effectuée avec la ligne L_1 obtenue à l'issue de la première opération, et non avec la ligne L_1 initiale.

Théorème 16.2.4 (Effet des opérations sur les lignes sur un système)

Soit $AX = B$ un système linéaire. Effectuer l'une des opérations ci-dessus (permutation, dilatation ou transvection) à la fois sur A et sur B fournit un système équivalent, donc ne modifie pas l'ensemble des solutions.

En effet, il suffit de savoir faire les opérations inverses, ramenant au système initial :

- L'opération inverse de $L_i \leftrightarrow L_j$ est $L_i \leftrightarrow L_j$
- L'opération inverse de $L_i \leftarrow \lambda L_i$ (pour $\lambda \neq 0$) est $L_i \leftarrow \frac{1}{\lambda} L_i$
- L'opération inverse de $L_i \leftarrow L_i + \lambda L_j$ est $L_i \leftarrow L_i - \lambda L_j$.

Ainsi que suggéré dans le théorème ci-dessus, nous allons donc transformer le système initial petit à petit, en appliquant les mêmes transformations à la matrice A et à la matrice B . Par commodité, nous adopterons la présentation suivante pour représenter les coefficients du système, en tenant compte de la matrice B :

$$(A | B) = \left(\begin{array}{ccc|c} a_{1,1} & \cdots & a_{1,p} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{n,1} & \cdots & a_{n,p} & b_n \end{array} \right)$$

Ainsi, effectuer les mêmes opérations sur les lignes de A et B revient à effectuer ces opérations sur la matrice $(A | B)$.

II.2 Échelonnement de la matrice

Ainsi que nous l'avons dit dans l'introduction, nous allons résoudre le système $AX = B$ en nous ramenant, à l'aide d'opérations élémentaires sur les lignes, à un système échelonné équivalent.

Définition 16.2.5 (Matrice échelonnée)

Soit m et n deux entiers non nuls, et $M = (a_{i,j})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$ dans $\mathcal{M}_{m,n}(\mathbb{K})$. On dit que M est une matrice échelonnée s'il existe un entier $k \in \llbracket 1, m \rrbracket$ et une suite croissante $j_1 < j_2 < \cdots < j_k$ d'éléments de $\llbracket 1, n \rrbracket$ tels que :

- (i) $\forall i \in \llbracket 1, k \rrbracket, a_{i,j_i} \neq 0$;
- (ii) $\forall i \in \llbracket 1, k \rrbracket, \forall j \in \llbracket 1, j_i - 1 \rrbracket, a_{i,j} = 0$;
- (iii) $\forall i \in \llbracket k + 1, m \rrbracket, \forall j \in \llbracket 1, n \rrbracket, a_{i,j} = 0$

Autrement dit, les lignes nulles sont regroupées au bas de la matrice (lignes $k + 1$ à m), les autres lignes sont classées suivant la position de leur premier élément non nul, ces positions étant deux à deux distinctes. Une matrice échelonnée admet donc la représentation suivante :

$$M = \begin{pmatrix} 0 & \cdots & 0 & a_{1,j_1} & \bullet & & \cdots & & \bullet \\ 0 & \cdots & & \cdots & 0 & a_{2,j_2} & \bullet & & \cdots & \bullet \\ \vdots & & & & & & & & & \vdots \\ 0 & & \cdots & \cdots & & & 0 & a_{k,j_k} & \bullet & \cdots & \bullet \\ 0 & & & \cdots & & & & \cdots & & & 0 \\ \vdots & & & & & & & & & & \vdots \\ 0 & & & \cdots & & & & \cdots & & & 0 \end{pmatrix},$$

les coefficients indiqués d'un \bullet étant quelconques.

Définition 16.2.6 (Système échelonné)

Un système $AX = B$ est dit échelonné si A est une matrice échelonnée.

Et voici maintenant l'algorithme lui-même, permettant de se ramener à un système échelonné :

Méthode 16.2.7 (Algorithme du pivot de Gauss, ou élimination de Gauss-Jordan)

1. On cherche la première colonne non nulle de la matrice $(A | B)$.
2. Sur cette colonne, on effectue un choix de pivot : n'importe quel coefficient non nul de la colonne convient, mais on a intérêt à choisir un pivot donnant le moins de calculs possible. Il y a trois critères pour cela :
 - Le pivot lui-même doit être facile à inverser. L'idéal est un pivot égal à 1.
 - Les autres coefficients de la ligne du pivot doivent être « simples », de préférence des entiers.
 - Plus il y a de zéros sur la ligne contenant le pivot, moins il y aura de calculs !
3. On fait un échange de lignes pour ramener le pivot choisi sur la première ligne.
4. On annule tous les coefficients situés sous le pivot à l'aide d'opérations élémentaires $L_i \leftarrow L_i + \lambda L_1$, ou bien pour éviter d'introduire des fractions, $L_i \leftarrow \alpha L_i + \beta L_1$, avec $\alpha \neq 0$
5. On recommence récursivement en considérant la sous-matrice située strictement en-dessous à droite du pivot.

Remarque 16.2.8

Bien entendu, les critères de choix du pivot sont donnés ici en vue d'un calcul à la main. En vue d'une implémentation sur ordinateur, le choix du pivot doit se faire de sorte à diminuer au maximum les erreurs d'arrondi. Les critères sont, dans cette optique, différents de ceux énoncés ci-dessus. Nous ne discuterons pas de ces critères ici.

Nous donnons ci-dessous une description purement algorithmique en pseudo-code un peu lâche.

Algorithme 16.1 : Pivot de Gauss

Entrée : A, B matrices définissant le système $AX = B$

Sortie : A', B' définissant un système équivalent, A' étant échelonné

Les opérations sur les lignes sont effectuées directement sur la matrice $(A | B)$.

Initialiser l'indice j de colonne à 1;

Initialiser l'indice i de ligne à 1;

tant que les indices i et j nous gardent à l'intérieur de la matrice $(A | B)$ **faire**

si le bas de la colonne j (en-dessous de la ligne i au sens large) est nul **alors**
 | Passer à la colonne suivante ($j \leftarrow j + 1$)

sinon

 | Placer un élément non nul en position (i, j) par une opération $L_i \leftrightarrow L_k$;

pour $k \leftarrow i + 1$ à dernière ligne **faire**

 | $L_k \leftarrow L_k + \lambda L_i$, où $\lambda = -\frac{L_k[j]}{L_i[j]}$

fin pour

 | Passer à la ligne suivante ($i \leftarrow i + 1$);

 | Passer à la colonne suivante ($j \leftarrow j + 1$)

fin si

fin tant que

renvoyer A, B , sous la forme $(A | B)$

Théorème 16.2.9 (Correction de l'algorithme du pivot de Gauss)

Soit A' et B' les matrices issues de l'algorithme du pivot de Gauss. Alors le système $A'X = B'$ est équivalent au système $AX = B$, et A' (et même $(A' | B')$) est une matrice échelonnée.

Définition 16.2.10 (réduite de Gauss)

La matrice échelonnée A' obtenue à l'aide de la méthode du pivot appliqué à la matrice A s'appelle *réduite de Gauss de la matrice A* . Il n'y a pas unicité d'une réduite de Gauss.

Exemple 16.2.11

Recherche d'une réduite de Gauss A' , et de la matrice B' associée lorsque :

$$A = \begin{pmatrix} 1 & -4 & -2 & 3 & 2 \\ 2 & 2 & 1 & 0 & 1 \\ -1 & 0 & 0 & 1 & 0 \end{pmatrix} \quad \text{et} \quad B = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}.$$

III Résolution d'un système échelonné

L'algorithme du pivot de Gauss nous ramène donc à la résolution d'un système échelonné $AX = B$, ce qui se fait facilement en partant de la fin. Nous commençons par introduire un peu de vocabulaire.

III.1 Inconnues principales

Un système échelonné n'admet pas toujours une solution unique, ni même une solution tout court. Pour trouver une solution particulière, il y aura donc des choix à faire. Ainsi, si la dernière équation fait par exemple intervenir 3 inconnues, en fixer 2 à sa guise impose la troisième. Ainsi, en remontant un système échelonné, on ajoute à chaque nouvelle ligne considérée de nouvelles inconnues n'intervenant pas dans les lignes suivantes. Si on a déjà obtenu des valeurs pour les inconnues des dernières lignes, donner des valeurs quelconques à toutes les nouvelles inconnues sauf une détermine alors la dernière inconnue.

On peut donc remonter le système en donnant des valeurs quelconques à toutes les nouvelles inconnues de chaque ligne, sauf une. L'inconnue qui jouera ce rôle particulier est l'une quelconque des nouvelles inconnues (en remontant), mais un choix s'impose naturellement : celui de la première inconnue intervenant de façon effective dans la ligne. Autrement dit, l'inconnue se plaçant à la position du pivot utilisé sur cette ligne pour obtenir l'échelonnement.

Définition 16.3.1 (Inconnue principale)

Nous appelons *inconnue principale* du système échelonné $AX = B$ une inconnue x_i se plaçant en tête d'une des lignes du système. Il s'agit donc des inconnues d'indice j_i de la définition 16.2.5

Remarque 16.3.2

- La définition d'inconnue principale donnée ici est fortement dépendante de l'ordre des variables. Une permutation des variables préservant l'échelonnement (c'est en général possible) donnerait un autre système d'inconnues principales.
- Il existe une notion plus générale d'inconnues principales, se définissant bien à l'aide de déterminants, portant sur des systèmes non nécessairement échelonnés. En gros, il s'agit d'un système d'inconnues tel que le choix quelconque des autres inconnues déterminent de façon unique des inconnues principales. Mais pour un système donné, le choix des inconnues principales n'est pas unique.

- Dans ce cadre plus général, notre définition des inconnues principales n'est qu'un des plusieurs choix possibles pour le système échelonné donné.

III.2 Recherche d'une solution particulière

Méthode 16.3.3 (recherche d'une solution particulière d'un système échelonné)

Soit A une matrice échelonnée, et $AX = B$ un système.

1. S'il existe dans ce système une ligne du type $0 = b_i$, avec b_i non nul, alors le système n'admet pas de solution. On dit que $AX = B$ est un système non compatible.
2. Sinon, donner des valeurs quelconques aux variables non principales réduisent le système à un système triangulaire (à coefficients diagonaux non nuls) portant sur les inconnues principales. Un tel système se résout de façon unique en remontant les équations.
3. On obtient de la sorte une description de toutes les solutions de $AX = B$, en donnant toutes les valeurs possibles aux inconnues non principales
4. D'après le théorème de structure, on peut se contenter d'exprimer une solution particulière. Une solution particulière est par exemple obtenue en donnant aux inconnues non principales la valeur nulle.

Remarque 16.3.4

On peut normaliser le système triangulaire de sorte à ce que le premier coefficient non nul sur chaque ligne soit égal à 1. Ainsi, si $AX = B$ est un système échelonné normalisé compatible, une solution particulière est obtenue en annulant les inconnues non principales et en prenant pour le vecteur des inconnues principales le vecteur B .

Exemple 16.3.5

Recherche d'une solution particulière dans le cas de l'exemple 16.2.11.

Nous retenons de la méthode ci-dessus le résultat important suivant :

Théorème 16.3.6

Soit $A \in \mathcal{M}_{n,p}(\mathbb{K})$ et $B \in \mathcal{M}_n(\mathbb{K})$. Le système $AX = B$ admet au moins une solution si et seulement si le système $A'X + B' = 0$ obtenu après échelonnement par la méthode de Gauss Jordan ne possède aucune ligne du type $0 = b_i$, avec $b_i \neq 0$.

III.3 Recherche de la solution générale de l'équation homogène associée

À ce stade, l'essentiel du travail a été fait : on réutilise les résultats des calculs effectués dans la phase précédente. La recherche des solutions de l'équation homogène associée $AX = 0$ équivaut à la recherche des solutions du système homogène échelonné $A'X = 0$, pour la matrice A' obtenue précédemment par la méthode de Gauss-Jordan.

Ainsi que nous l'avons évoqué plus haut :

Proposition 16.3.7

L'ensemble des solutions du matrice échelonné $AX = 0$ est obtenu en résolvant le système triangulaire obtenu sur les inconnues principales en donnant toutes les valeurs possibles quelconques pour les inconnues non principales.

Ainsi, si X_s désigne un vecteur de valeurs pour les inconnues non principales (secondaires), et qu'on note $\varphi(X_s)$ l'unique solution du système obtenu sur les variables principales, et enfin, qu'on note $\psi(X_s)$ le vecteur solution obtenu en remettant en place les inconnues principales et secondaires (c'est un shuffle des vecteurs X_s et $\varphi(X_s)$), l'ensemble des solutions est

$$\mathcal{S} = \{\psi(X_s), X_s \in \mathcal{M}_{p-r,1}(\mathbb{K})\},$$

où r est le nombre d'inconnues principales.

Lemme 16.3.8

La fonction ψ est linéaire et bijective (on dit que c'est un isomorphisme). La linéarité s'exprime par la relation :

$$\psi(\lambda X_s + \mu Y_s) = \lambda \psi(X_s) + \mu \psi(Y_s).$$

On note (e_1, \dots, e_{p-r}) la base canonique de $\mathcal{M}_{p-1,1}(\mathbb{R})$: e_i a toutes ses coordonnées nulles, sauf la i -ième égale à 1.

Théorème 16.3.9 (Base de l'ensemble des solutions de l'équation homogène)

La famille $(\psi(e_1), \dots, \psi(e_{p-r}))$ est une base de \mathcal{S}_0 , ensemble des solutions du système échelonné $AX = 0$. Autrement dit, toute solution s'exprime de façon unique comme une combinaison linéaire des $\psi(e_i)$.

Ainsi, pour obtenir une base de l'ensemble des solutions, il suffit de poser, de toutes les façons possibles, l'une des inconnues non principales égale à 1, et toutes les autres nulles. La description de la base suffit alors à décrire l'ensemble des solutions de l'équation homogène. On écrit :

$$\mathcal{S}_0 = \text{Vect}(\psi(e_1), \dots, \psi(e_{p-r})).$$

Exemple 16.3.10

Recherche de la solution générale de l'équation homogène, puis de l'équation non homogène dans le cas de l'exemple 16.2.11.

Structures algébriques

THÉORÈME. Soit une équation donnée, dont a, b, c, \dots sont les m racines. Il y aura toujours un groupe de permutations des lettres a, b, c, \dots qui jouira de la propriété suivante :

1. *Que toute fonction des racines, invariable par les substitutions de ce groupe, soit rationnellement connue ;*
2. *Réciproquement, que toute fonction des racines, déterminable rationnellement, soit invariable par les substitutions.*

[...] Nous appellerons groupe de l'équation le groupe en question.

(Évariste Galois)

Restez groupir !

(hum...)

Note Historique 17.0.11

Il est fréquent de trouver des propriétés communes dans des situations qui au départ semblent totalement sans rapport. Une des grandes découvertes des mathématiques du 19^e siècle a été de réussir à unifier ces problèmes en apparence distincts, en faisant ressortir de ces différents problèmes des structures ensemblistes et opératoires ayant des propriétés similaires.

C'est Évariste Galois le premier à mettre en avant ces études de structure à l'occasion de ses travaux visant à étudier la résolubilité des équations polynomiales par radicaux. Il y parle de groupes de permutations des solutions d'une équation, et est amené à étudier des propriétés de certains sous-ensembles de ces groupes de permutations. C'est lui qui introduit la terminologie de « groupe », même si la formalisation précise de cette notion est beaucoup plus tardive.

Le groupe des permutations d'un ensemble avait déjà été étudié auparavant par Lagrange (mais sans en faire ressortir cette structure bien particulière de groupe). Il a notamment établi à cette occasion un résultat important, généralisé plus tard pour tout groupe sous le nom de « théorème de Lagrange ».

La notion de structure algébrique repose de façon essentielle sur la notion de loi de composition (c'est-à-dire d'opération définie sur un ensemble, comme l'addition ou la multiplication) et sur les différentes propriétés que ces lois de composition peuvent vérifier. Nous commençons donc notre étude par l'examen de ces propriétés, après avoir défini de façon précise ce qu'est une loi de composition.

I Lois de composition

I.1 Définitions

Dans ce qui suit, E est un ensemble quelconque.

Définition 17.1.1 (Lois de composition)

On distingue deux types de lois de compositions (opérations), suivant que la loi décrit une opération entre deux éléments de l'ensemble E , ou entre un élément de E et un élément d'un ensemble externe Ω , appelé domaine d'opérateur.

- Une *loi de composition interne* est une application de $\varphi : E \times E$ dans E , souvent notée de façon opérationnelle plutôt que fonctionnelle (par exemple $x + y$ au lieu de $\varphi(x, y)$ pour désigner une addition)
- Une *loi de composition externe à gauche* sur E , d'ensemble d'opérateurs Ω , est une application de $\Omega \times E$ dans E , également notée de façon opérationnelle le plus souvent (par exemple $\lambda \cdot x$ au lieu de $\varphi(\lambda, x)$)
- De même, une *loi de composition externe à droite* sur E d'ensemble d'opérateurs Ω est une application $E \times \Omega \rightarrow E$.

Exemples 17.1.2

- Les lois $+$ et \times sont des lois de composition interne sur \mathbb{N} , \mathbb{Z} , \mathbb{R} ou \mathbb{C} .
- La loi $+$ est une loi de composition interne sur \mathbb{R}^n ou \mathbb{C}^n .
- $(\lambda, X) \mapsto \lambda X$ (multiplication d'un vecteur par un scalaire) est une loi de composition externe sur \mathbb{R}^n (ou \mathbb{C}^n), d'ensemble d'opérateurs \mathbb{R} (ou \mathbb{C})
- De même pour la multiplication des polynômes par des scalaires.
- La composition \circ définit une loi de composition interne sur E^E .
- Le produit scalaire sur \mathbb{R}^n n'est pas une loi de composition (interne ou externe), car le résultat de l'opération n'est pas un élément de \mathbb{R}^n .

I.2 Propriétés d'une loi de composition

Soit E un ensemble, muni d'une loi de composition interne que nous noterons \star . Nous étudions ici quelques propriétés pouvant être vérifiées par la loi \star

Définition 17.1.3 (Associativité, commutativité)

- On dit que \star est *associative* ssi : $\forall (x, y, z) \in E^3, (x \star y) \star z = x \star (y \star z)$
- On dit que \star est *commutative* ssi : $\forall (x, y) \in E^2, x \star y = y \star x$.

Ainsi, lorsque E est muni d'une loi associative, on peut effectuer les opérations dans l'ordre que l'on veut, à condition de respecter la position respective des éléments les uns par rapport aux autres. Si la loi est commutative, on peut échanger la position respective des éléments (mais pas nécessairement faire les opérations dans l'ordre qu'on veut si la loi n'est pas associative).

Exemples 17.1.4 (Lois commutatives, associatives)

1. Les lois $+$ et \times définies sur \mathbb{N} , \mathbb{Z} , \mathbb{R} et \mathbb{C} sont associatives et commutatives.
2. Le produit matriciel définit une loi associative sur $\mathcal{M}_n(\mathbb{R})$ (ensemble des matrices carrées d'ordre n), mais pas commutative.
3. La composition définit une loi associative sur E^E mais pas commutative.
4. La soustraction dans \mathbb{Z} est non associative et non commutative.
5. La loi définie sur \mathbb{R} par $(a, b) \mapsto (a + b)^2$ est commutative mais non associative.

Notation 17.1.5 (Suppression des parenthèses)

Lorsque \star est associative, nous nous permettons d'omettre le parenthésage, en notant $x \star y \star z$ au lieu de $(x \star y) \star z$ ou $x \star (y \star z)$, la propriété d'associativité levant toute ambiguïté sur l'interprétation de cette expression.

Avertissement 17.1.6

Attention à toujours bien indiquer le parenthésage lorsque la loi n'est pas associative, ou lorsque plusieurs lois sont en jeu sans qu'il n'ait été défini de façon explicite de relation de priorité sur les opérations.

Convention 17.1.7 (Commutativité d'une loi d'addition, usage)

Nous réserverons la notation additive (signe opératoire $+$) pour des lois de composition commutatives. Cela n'empêche pas en revanche de considérer des lois commutative notées multiplicativement.

Nous voyons maintenant des propriétés liées à l'existence de certains éléments particuliers de E .

Définition 17.1.8 (Élément neutre)

Soit e un élément de E .

1. On dit que e est un *élément neutre à droite* pour la loi \star ssi : $\forall x \in E, x \star e = x$.
2. On dit que e est un *élément neutre à gauche* pour la loi \star ssi : $\forall x \in E, e \star x = x$.
3. On dit que e est un *élément neutre* pour la loi \star s'il est à la fois neutre à gauche et neutre à droite.

Pour une loi commutative, e est neutre ssi e est neutre à droite ssi e est neutre à gauche.

Exemple 17.1.9 (Éléments neutres)

1. 0 est élément neutre pour $+$ dans $\mathbb{N}, \mathbb{Z}, \mathbb{R}, \mathbb{C}$. C'est le seul élément neutre pour $+$.
2. 1 est élément neutre pour \times dans $\mathbb{N}, \mathbb{Z}, \mathbb{R}, \mathbb{C}$. C'est le seul élément neutre pour \times .
3. I_n est élément neutre pour \times sur $\mathcal{M}_n(\mathbb{R})$, 0_n est élément neutre pour $+$ sur $\mathcal{M}_n(\mathbb{R})$.
4. id_E est élément neutre pour \circ sur E^E .
5. Sur un ensemble E de cardinal supérieur ou égal à 2, la loi $(x, y) \mapsto y$ admet plusieurs neutres à gauche (tout $x \in E$ est neutre à gauche). En revanche, il n'y a pas de neutre à droite.

Une loi ne peut pas admettre plusieurs éléments neutres, comme le montre la propriété suivante.

Proposition 17.1.10 (Unicité du neutre)

- Si E admet pour \star un neutre à gauche e_g et un neutre à droite e_d , alors $e_g = e_d$, et cet élément est alors un élément neutre pour \star .
- L'élément neutre, s'il existe, est unique; dans ce cas, il s'agit aussi de l'unique neutre à gauche, et de l'unique neutre à droite.

Notation 17.1.11 ($0_E, 1_E$)

- On note généralement 0_E (ou 0 s'il n'y a pas de risque d'ambiguïté) le neutre (s'il existe) d'une loi notée additivement $+$.

- On note généralement 1_E (ou 1 s'il n'y a pas de risque d'ambiguïté) le neutre (s'il existe) d'une loi notée multiplicativement \times .

Définition 17.1.12 (Élément symétrique)

Supposons que E admet un élément neutre e pour la loi \star . Soit $x \in E$.

- On dit que ${}^s x$ est un symétrique à gauche de x pour la loi \star si ${}^s x \star x = e$.
- On dit que x^s est un symétrique à droite de x pour la loi \star si $x \star x^s = e$.
- On dit que \bar{x} est un symétrique de x pour la loi \star si et seulement si \bar{x} est un symétrique à droite et à gauche de x .
- On dit que x est symétrisable (*resp.* symétrisable à gauche, *resp.* symétrisable à droite) si x admet au moins un symétrique (*resp.* un symétrique à gauche, *resp.* un symétrique à droite).

Terminologie 17.1.13 (Opposé, inverse)

- Dans le cas d'une loi notée additivement, on parle plutôt d'opposé, et en cas d'unicité, on note $-x$ l'opposé de x .
- Dans le cas d'une loi notée multiplicativement (on omet dans ce cas souvent d'écrire le signe), on parle plutôt d'inversibilité (inversibilité à droite, à gauche), et en cas d'unicité, on note x^{-1} l'inverse de x .

Proposition 17.1.14 (Unicité du symétrique)

Si \star est associative, alors, en cas d'existence, le symétrique est unique.

Exemples 17.1.15

1. Dans \mathbb{N} seul 0 admet un opposé pour $+$.
2. Dans $\mathbb{Z}, \mathbb{R}, \mathbb{Q}, \mathbb{C}$, tout élément admet un opposé pour $+$.
3. Dans \mathbb{N} seul 1 admet un inverse, dans \mathbb{Z} , seuls 1 et -1 admettent un inverse. Dans \mathbb{R}, \mathbb{Q} et \mathbb{C} tous les éléments non nuls admettent un inverse.
4. Dans E^E muni de \circ , les éléments symétrisables à gauche sont les injections, les éléments symétrisables à droite sont les surjections, les éléments symétrisables sont les bijections. Une injection non surjective admet plusieurs symétriques à gauche; une surjection non injective admet plusieurs symétriques à droite.

Proposition 17.1.16 (Symétrique de $x \star y$)

Supposons \star associative. Soit $(x, y) \in E^2$. Si x et y sont symétrisables, de symétriques x^s et y^s , alors $x \star y$ est symétrisable de symétrique $y^s \star x^s$. Notez l'inversion !

Traduisons pour une loi multiplicative : si x et y sont inversibles, d'inverses x^{-1} et y^{-1} , alors xy aussi, d'inverse $y^{-1}x^{-1}$.

Définition 17.1.17 (Élément absorbant)

Soit $x \in E$.

- On dit que x est un élément absorbant à gauche pour \star ssi : $\forall y \in E, x \star y = x$.
- On dit que x est absorbant à droite pour \star ssi : $\forall y \in E, y \star x = x$.
- On dit que x est absorbant s'il est à la fois absorbant à gauche et à droite.

Exemples 17.1.18

1. 0 est absorbant pour \times dans $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$.
2. Pour la loi $(x, y) \mapsto y$, tout élément y de E est absorbant à droite. Il n'y a pas d'élément absorbant à gauche si E est de cardinal au moins 2.

Définition 17.1.19 (Élément régulier ou simplifiable)

- Un élément x est dit régulier (ou simplifiable) à gauche ssi :

$$\forall (y, z) \in E^2, x \star y = x \star z \implies y = z.$$

- Un élément x est dit régulier (ou simplifiable) à droite ssi :

$$\forall (y, z) \in E, y \star x = z \star x \implies y = z.$$

- Un élément x est dit régulier (ou simplifiable) s'il est à la fois régulier à gauche et à droite.

Proposition 17.1.20 (Régularité des éléments symétrisables)

Supposons que E soit muni d'une loi \star associative.

- Soit x un élément admettant un symétrique à gauche. Alors x est régulier à gauche.
- Soit x un élément admettant un symétrique à droite. Alors x est régulier à droite.
- Soit x un élément admettant un symétrique. Alors x est régulier.

Ainsi, le fait de pouvoir simplifier une égalité par un réel ou complexe non nul x ne vient pas tant de la non nullité que de l'inversibilité de x . Par exemple, la non nullité n'est pas un critère suffisant de régularité dans $\mathcal{M}_n(\mathbb{R})$: il est nécessaire d'avoir l'inversibilité de la matrice que l'on veut simplifier. Il convient toutefois de noter que la condition d'inversibilité, si elle est suffisante, n'est en général pas nécessaire.

Exemple 17.1.21

Dans l'ensemble $\mathbb{R}[X]$ des polynômes à coefficients réels muni du produit, tout polynôme non nul est régulier, alors que seuls les polynômes constants non nuls sont inversibles.

I.3 Ensembles munies de plusieurs lois

Soit E un ensemble muni de deux lois de composition \star et \diamond .

Définition 17.1.22 (Distributivité)

- On dit que la loi \star est distributive à gauche sur \diamond ssi : $\forall (x, y, z) \in E^3, x \star (y \diamond z) = (x \star y) \diamond (x \star z)$.
- On dit que la loi \star est distributive à droite sur \diamond ssi : $\forall (x, y, z) \in E^3, (y \diamond z) \star x = (y \star x) \diamond (z \star x)$.
- On dit que la loi \star est distributive sur \diamond ssi elle est distributive à droite et à gauche.

Exemples 17.1.23

1. La loi \times est distributive sur $+$ dans $\mathbb{N}, \mathbb{Z}, \mathbb{R}, \mathbb{C}, \mathcal{M}_n(\mathbb{R})$...
2. La loi \cap est distributive sur \cup sur $\mathcal{P}(X)$. Inversement, la loi \cup est distributive sur \cap .

Remarque 17.1.24

La première relation $x \star (y \diamond z) = (x \star y) \diamond (x \star z)$ a également un sens lorsque \star est une loi externe. Ainsi, dans \mathbb{R}^n , muni de l'addition en loi interne, et de la multiplication des scalaires en loi externe, on a $\lambda(X + Y) = \lambda X + \lambda Y$: la loi externe est distributive sur la loi interne.

Définition 17.1.25 (pseudo-associativité)

Soit E un ensemble muni d'une loi de composition externe \diamond sur \mathbb{K} , lui-même muni d'une loi de composition interne \star . On dit que les lois \star et \diamond vérifient une propriété de pseudo-associativité si pour tout $(\lambda, \mu) \in \mathbb{K}^2$ et $x \in E$

$$(\lambda \star \mu) \diamond x = \lambda \diamond (\mu \diamond x).$$

Cette propriété est par exemple satisfaite sur \mathbb{R}^n pour la multiplication par un scalaire : $(\lambda\mu)X = \lambda(\mu X)$. Plus généralement, un espace vectoriel vérifiera cette propriété de pseudo-associativité.

I.4 Stabilité**Définition 17.1.26**

Soit E un ensemble muni d'une loi \star et $F \subset E$ un sous-ensemble de E . On dit que F est stable par \star , si la restriction de la loi de E à $F \times F$ peut se corestreindre à F , autrement dit si :

$$\forall (x, y) \in F^2, x \star y \in F.$$

Dans ce cas, la loi de E se restreint en une loi $\star_F : F \times F \rightarrow F$, appelée *loi induite sur F par \star* .

II Structures**II.1 Généralités****Définition 17.2.1**

- Une structure de « truc » est la donnée d'un certain nombre d'axiomes (définissant ce qu'on appelle un « truc ») portant sur un ensemble fini de lois de composition (internes et/ou externes).
- On dit que E est muni d'une structure de truc ssi E est muni d'un nombre fini de lois de composition vérifiant les axiomes de structure de truc.

Exemples 17.2.2

1. Une structure de magma se définit comme la donnée d'une loi de composition, et un ensemble vide d'axiomes. Ainsi, tout ensemble E muni d'une loi de composition est muni d'une structure de magma.
2. Une structure de monoïde se définit comme la donnée d'une loi de composition, et de deux axiomes : l'associativité de la loi et l'existence d'un élément neutre. Par exemple $(\mathbb{N}, +)$ est muni d'une structure de monoïde (on dit plus simplement que $(\mathbb{N}, +)$ est un monoïde). L'ensemble des mots sur un alphabet \mathcal{A} , muni de l'opération de concaténation est aussi un monoïde (appelé monoïde libre sur l'alphabet \mathcal{A}). Contrairement à \mathbb{N} , le monoïde libre n'est pas commutatif.
3. Ainsi, la structure de monoïde est plus riche que celle de magma : tout monoïde est aussi un magma ; un monoïde peut être défini comme un magma dont la loi est associative et possède un élément neutre.

4. Une structure de groupe est une structure de monoïde à laquelle on rajoute l'axiome d'existence de symétriques. Par exemple $(\mathbb{Z}, +)$ est un groupe, mais pas $(\mathbb{N}, +)$.

Définition 17.2.3 (Structure induite)

Soit E un ensemble muni d'une structure de truc, et F un sous-ensemble de E . Si F est stable pour chacune des lois de E , l'ensemble F muni des lois induites sur F par les lois de E est appelée structure induite sur F par la structure de E .

Avertissement 17.2.4

En général, F ne peut pas être muni d'une structure de truc, mais seulement d'une structure moins riche, certains des axiomes de la structure de truc pouvant ne pas être préservée par restriction.

Exemple 17.2.5

$(\mathbb{N}, +)$ est la structure induite sur \mathbb{N} par la structure de groupe additif de $(\mathbb{Z}, +)$. En revanche, $(\mathbb{N}, +)$ n'est pas un groupe. On a perdu l'existence des opposés par restriction.

Définition 17.2.6 (Sous-truc)

Soit E un ensemble muni d'une structure de truc et F un sous-ensemble de E . On dit que F est un sous-truc de E si F est stable par les lois de E , si F contient les neutres imposés de E , et si les lois induites sur F par les lois de E vérifient les axiomes de la structure de truc.

Nous verrons comment traduire de façon effective cette notion dans le cas de sous-groupes et sous-anneaux.

II.2 Morphismes

Lorsqu'on dispose d'une structure de truc, on est souvent amené à considérer des applications entre ensembles munis de la structure de truc. Cependant seules nous intéressent les applications compatibles dans un certain sens avec la structure de truc : les autres ne sont pas pertinentes dans le contexte (si on a à s'en servir, c'est qu'on sort de la structure de truc, et que la structure de truc n'est plus le contexte adapté).

Définition 17.2.7 (Homomorphisme)

Soit E et F deux ensembles munis d'une structure de truc, E étant muni des lois de composition interne $(\star_1, \dots, \star_n)$ et F des lois $(\diamond_1, \dots, \diamond_n)$, et des lois de composition externes (\top_1, \dots, \top_m) et $(\perp_1, \dots, \perp_m)$ sur K . On dit qu'une application $f : E \rightarrow F$ est un homomorphisme de truc (ou plus simplement un morphisme de truc) ssi :

- L'application f est compatible avec les lois internes :

$$\forall k \in [1, n], \quad \forall (x, y) \in E^2, \quad f(x \star_k y) = f(x) \diamond_k f(y).$$

- L'application f est compatible avec les lois externes :

$$\forall \ell \in [1, m], \quad \forall \lambda \in K, \quad \forall x \in E, \quad f(\lambda \top_\ell x) = \lambda \perp_\ell f(x).$$

- Si l'existence du neutre e_i pour la loi \star_i est imposée dans les axiomes (et donc le neutre f_i pour la loi \diamond_i existe aussi), f doit être compatible avec le neutre : $f(e_i) = f_i$.

On peut avoir à rajouter certaines propriétés liées à la structure étudiée. On peut aussi ajouter l'existence d'un homomorphisme nul (ne vérifiant pas la compatibilité avec les neutres non additifs), afin d'obtenir une structure intéressante sur l'ensemble des morphismes.

Pour chaque structure étudiée, nous redéfinirons de façon précise la notion d'homomorphisme associée, si celle-ci est à connaître. Nous donnons une propriété générale, dont la démonstration dans le cadre général nous dispensera des démonstrations au cas par cas.

Proposition 17.2.8 (Composition d'homomorphismes)

Soit $f : E \rightarrow F$ et $g : F \rightarrow G$ deux morphismes de truc. Alors $g \circ f$ est un morphisme de truc.

Nous définissons alors :

Terminologie 17.2.9

- Un isomorphisme de truc est un homomorphisme de truc bijectif.
- Un endomorphisme de truc est un homomorphisme de truc de E dans lui-même (muni des mêmes lois)
- Un automorphisme de truc est un endomorphisme qui est également un isomorphisme.

Proposition 17.2.10

Si $f : E \rightarrow F$ est un isomorphisme de truc, alors f^{-1} est un isomorphisme de truc.

Ainsi, la réciproque d'un isomorphisme est bijective (ça, ce n'est pas une surprise), et, ce qui est moins évident, c'est aussi un homomorphisme de truc.

II.3 Catégories (HP)

La notion de structure et de morphisme associé permet de définir la notion de catégorie. Grossièrement, une catégorie est la donnée :

- d'une classe d'objets ;
- de flèches entre ces objets ;
- d'une règle de composition entre les flèches.

Par exemple, la catégorie des monoïdes est la catégorie dont les objets sont les monoïdes, les flèches sont les homomorphismes de monoïdes, et la composition des flèches correspond à la composition usuelle des homomorphismes (la composée de deux homomorphismes étant encore un homomorphisme, donc une flèche de la catégorie). On définit de même la catégorie des groupes, ou la catégorie des anneaux, ou encore la catégorie des corps.

Cette notion de catégorie nous permet de travailler dans un certain contexte. Se donner une catégorie permet de se concentrer sur un certain type d'objets, et un certain type d'applications, et de les étudier d'un point de vue formel.

La notion de catégorie dépasse largement le cadre de l'étude des structures algébriques, car si les structures algébriques fournissent des catégories, de nombreuses catégories sont issues d'autres contextes, comme :

- la catégorie des ensembles, les morphismes étant toutes les applications ;
- la catégorie des ensembles ordonnés, les morphismes étant les applications croissantes
- la catégorie des espaces topologiques, les morphismes étant les applications continues
- ou encore, la catégorie des catégories, les morphismes étant les foncteurs de C dans D , associant à chaque objet de C un objet de D , et à chaque flèche de C une flèche de D , tout en respectant un certain nombre de règles de compatibilité.
- ou encore, des catégories de foncteurs entre deux catégories, les objets étant cette fois des foncteurs, et les flèches étant des « transformations naturelles » entre foncteurs...

III Groupes

III.1 Axiomatique de la structure groupes

Définition 17.3.1 (Groupe)

Soit G un ensemble. On dit que G est muni d'une structure de groupe si G est muni d'une loi de composition \star telle que :

- \star est associative ;
- il existe un élément neutre e pour la loi \star ;
- tout élément x admet un symétrique x^s .

En vertu de résultats antérieurs, on peut énoncer :

Proposition 17.3.2 (Unicité du neutre et des symétriques)

Soit (G, \star) un groupe. Alors :

- G admet un unique élément neutre pour \star
- Pour tout $x \in G$, il existe un unique symétrique x^s de x .

Corollaire 17.3.3 (régularité des éléments d'un groupe)

Tous les éléments d'un groupe sont réguliers pour la loi du groupe.

Définition 17.3.4 (Groupe abélien ou commutatif)

On dit qu'un groupe (G, \star) est abélien (ou commutatif) si la loi de G est commutative.

Notation 17.3.5 (Notation additive, notation multiplicative)

La loi d'un groupe est le plus souvent notée additivement (signe $+$) ou multiplicativement (signe \times , parfois remplacé par un point, voire omis, comme dans \mathbb{R}). La notation additive est réservée au cas de groupes abéliens. Nous avons alors les notations suivantes pour désigner des itérées de la loi de composition sur un même élément x :

- loi multiplicative : $x \times \cdots \times x$ (avec n occurrences) est noté x^n ;
le neutre est noté 1 ;
par convention, $x^0 = 1$;
- loi additive : $x + \cdots + x$ (avec n occurrences) est noté $n \cdot x$ ou nx ;
le neutre est noté 0 ;
par convention $0x = 0$.

Une définition plus rigoureuse par récurrence pourrait être donnée pour ces itérées.

Notation 17.3.6 (Simplifications d'écriture pour la notation additive)

Soit $(G, +)$ un groupe commutatif. Comme mentionné plus haut, l'opposé d'un élément x est noté $-x$.

On note alors $x - y$ au lieu de $x + (-y)$. On a alors les règles suivantes :

- $\forall (x, y, z) \in G^3, \quad x - (y + z) = x - y - z$
- $\forall (x, y, z) \in G^3, \quad x - (y - z) = x - y + z.$

En vertu des définitions générales données dans le paragraphe précédent, nous donnons la définition suivante :

Définition 17.3.7 (Homomorphisme de groupe, HP)

Soit (G, \star) et (H, \diamond) deux groupes.

- On dit qu'une application $f : G \rightarrow H$ est un homomorphisme de groupe si pour tout $(x, y) \in G$, $f(x \star y) = f(x) \diamond f(y)$.

On note $\text{Hom}(G, H)$ l'ensemble des homomorphismes de G dans H .

- Si $(G, \star) = (H, \diamond)$, on dit que f est un endomorphisme de (G, \star) .
- Un homomorphisme bijectif est appelé isomorphisme ; en vertu de ce qui précède, la réciproque d'un isomorphisme est un isomorphisme.
- Un endomorphisme bijectif est appelé automorphisme ; en vertu de ce qui précède, la réciproque d'un automorphisme est un automorphisme.

On note $\text{Aut}(G)$ l'ensemble des automorphismes de G .

III.2 Exemples importants**Exemples 17.3.8**

1. $(\mathbb{Z}, +)$, $(\mathbb{Q}, +)$, $(\mathbb{R}, +)$, $(\mathbb{C}, +)$ sont des groupes commutatifs notés additivement.
2. (\mathbb{Q}^*, \times) , (\mathbb{R}^*, \times) , (\mathbb{C}^*, \times) , (\mathbb{Q}_+^*, \times) , (\mathbb{R}_+^*, \times) sont des groupes commutatifs notés multiplicativement.
3. $(\mathbb{N}, +)$, (\mathbb{Q}, \times) , (\mathbb{Z}, \times) , « (\mathbb{R}_-^*, \times) » sont-ils des groupes ?
4. (\mathbb{U}, \times) et (\mathbb{U}_n, \times) sont des groupes.
5. Pour $n \geq 2$, $(\mathbb{Z}/n\mathbb{Z}, +)$ est un groupe.
6. $(\mathbb{Z}/n\mathbb{Z} \setminus \{0\}, \times)$ est-il en général un groupe ?
7. Étant donné X un ensemble, (\mathfrak{S}_X, \circ) , l'ensemble des permutations de X est un groupe pour la loi définie par la composition.
8. $\exp : (\mathbb{R}, +) \rightarrow (\mathbb{R}_+^*, \times)$ est un homomorphisme de groupes. C'est même un isomorphisme.
9. Sa réciproque est donc aussi un isomorphisme de groupes : $\ln : (\mathbb{R}_+^*, \times) \rightarrow (\mathbb{R}, +)$.
10. L'application $x \mapsto e^{ix}$ est un morphisme de groupes de $(\mathbb{R}, +)$ à (\mathbb{U}, \times) .
11. L'application $z \mapsto e^z$ est un morphisme de groupes surjectif (mais non injectif) de $(\mathbb{C}, +)$ sur (\mathbb{C}^*, \times) .
12. Soit $n \geq 2$ et pour tout $k \in \llbracket 0, n-1 \rrbracket$, $\omega_k = e^{i \frac{2k\pi}{n}}$. Alors, étant donné $k \in \llbracket 0, n \rrbracket$:

$$\begin{aligned} f : (\mathbb{Z}, +) &\longrightarrow (\mathbb{U}_n, \times) \\ \ell &\longmapsto \omega_k^\ell \end{aligned}$$

est un homomorphisme de groupe. Il est surjectif si $k = 1$, et plus généralement si k est premier avec n (d'après le théorème de Bézout).

13. Puisque $\omega_k^n = 1$, le morphisme précédent « passe au quotient » et définit un homomorphisme de groupe :

$$\begin{aligned} f : (\mathbb{Z}/n\mathbb{Z}, +) &\longrightarrow (\mathbb{U}_n, \times) \\ \ell &\longmapsto \omega_k^\ell \end{aligned}$$

Cet homomorphisme est un isomorphisme si $k = 1$, et plus généralement si k est premier avec n .

III.3 Sous-groupes

Toujours en suivant les définitions plus générales, nous donnons la définition suivante :

Définition 17.3.9 (Sous-groupe)

Soit (G, \star) un groupe. Un sous-ensemble H de G est appelé *sous-groupe de G* si H est stable pour la loi de G et si la loi induite définit sur H une structure de groupe.

Remarquez qu'on n'a pas donné l'appartenance du neutre à H dans la définition, celle-ci étant automatique en vertu de :

Proposition 17.3.10 (Appartenance de l'élément neutre à H)

Soit H un sous-groupe de G . Alors l'élément neutre e de G est dans H et est l'élément neutre du groupe H .

Dans la pratique, pour vérifier que H est un sous-groupe de G on utilise le résultat suivant, ou sa version compactée :

Théorème 17.3.11 (Caractérisation des sous-groupes)

Un sous-ensemble H d'un groupe (G, \star) (de neutre e_G) est un sous-groupe de G si et seulement si :

- (i) H est non vide (on vérifie par exemple $e_G \in H$);
- (ii) H est stable pour \star : $\forall (x, y) \in H, x \star y \in H$,
- (iii) H est stable par prise de symétrie : $\forall x \in H, x^s \in H$.

On peut rassembler les deux dernières propriétés en une seule vérification :

Théorème 17.3.12 (Caractérisation des sous-groupes, version compactée)

Un sous-ensemble H d'un groupe (G, \star) (de neutre e_G) est un sous-groupe de G si et seulement si :

- (i) H est non vide (on vérifie par exemple $e_G \in H$);
- (ii) $\forall (x, y) \in H^2, x \star y^s \in H$,

On traduit cette dernière propriété dans les deux cas les plus fréquents :

- pour un sous-groupe d'un groupe additif, la vérification de stabilité à faire est donc :

$$\forall (x, y) \in H^2, x - y \in H;$$

- pour un sous-groupe d'un groupe multiplicatif, la vérification de stabilité à faire est donc :

$$\forall (x, y) \in H^2, xy^{-1} \in H.$$

De façon évidente, étant donné G un groupe, d'élément neutre e , $\{e\}$ et G sont des sous-groupes de G .

Définition 17.3.13 (Sous-groupe propre)

Un sous-groupe propre de G est un sous-groupe de G distinct de $\{e\}$ et de G .

Proposition 17.3.14 (Transitivité de la notion de sous-groupe)

Soit G un groupe. Alors si H est un sous-groupe de G et K un sous-groupe de H , alors K est un sous-groupe de G .

Proposition 17.3.15 (Intersection de sous-groupes)

Soit G un groupe, et $(H_i)_{i \in I}$ une famille de sous-groupes de G . Alors $\bigcap_{i \in I} H_i$ est un sous-groupe de G .

Théorème 17.3.16

Les sous-groupes de \mathbb{Z} sont exactement les $n\mathbb{Z}$, $n \in \mathbb{N}$.

Les sous-groupes de \mathbb{Z} sont donc des groupes monogènes (additifs), dans le sens suivant (donné multiplicativement) :

Proposition/Définition 17.3.17 (Sous-groupe monogène, HP)

Soit G un groupe multiplicatif, et $x \in G$. Soit

$$H = \{x^n, n \in \mathbb{Z}\}$$

(dans cet ensemble, certains éléments peuvent être égaux).

Alors H est un sous-groupe de G . On dit que H est un *sous-groupe monogène* de G .

Nous avons déjà eu l'occasion de prouver en exercice le résultat suivant, pour l'employer dans des situations particulières :

Proposition 17.3.18 (caractérisation des sous-groupes additifs de \mathbb{R} , HP)

Les sous-groupes de $(\mathbb{R}, +)$ sont soit égaux à $a\mathbb{Z}$, $a \in \mathbb{R}_+$, soit denses dans \mathbb{R} .

III.4 Congruences modulo un sous-groupe

Soit G un groupe (multiplicatif) et H un sous-groupe de G .

Définition 17.3.19 (Congruence modulo H , HP)

- On définit la relation de congruence à droite modulo H sur G par :

$$\forall (x, y) \in G, x \equiv_d y \pmod{H} \iff xy^{-1} \in H.$$

- On définit la relation de congruence à gauche modulo H sur G par :

$$\forall (x, y) \in G, x \equiv_g y \pmod{H} \iff x^{-1}y \in H.$$

Dans le cas de groupes abéliens, les deux notions coïncident évidemment.

Exemple 17.3.20

Si $G = \mathbb{Z}$ et $H = n\mathbb{Z}$, cela correspond (en notation additive cette fois) à la relation de congruence modulo n .

Proposition 17.3.21

Les relations de congruence à gauche et à droite modulo H sont des relations d'équivalence.

Définition 17.3.22 (Classes à droite et à gauche modulo H , HP)

- Les classes d'équivalence à droite modulo H sont appelées classes à droite modulo H . Il s'agit des ensembles Ha , $a \in G$, qui ne sont pas des groupes (sauf si $a \in H$).
- Les classes d'équivalence à gauche modulo H sont appelées classes à gauche modulo H . Il s'agit des ensembles aH , $a \in G$, qui ne sont pas des groupes (sauf si $a \in H$).

On en déduit une des propriétés les plus importantes des cardinaux des groupes finis.

Définition 17.3.23 (Ordre d'un groupe fini, HP)

Soit G un groupe fini. Son ordre est son cardinal.

Lemme 17.3.24 (Cardinaux des classes à gauche et à droite, HP)

Soit G un groupe fini. Pour tout $a \in G$, H , aH et Ha ont même cardinal.

Théorème 17.3.25 (Lagrange, HP)

Soit G un groupe fini, et H un sous-groupe de G . Alors l'ordre de H divise l'ordre de G .

Remarquez aussi que pour que les deux relations de congruence coïncident il faut et il suffit que pour tout $a \in G$, $aH = Ha$, ou encore $aHa^{-1} = H$. Cela motive la définition suivante, très importante dans l'étude des tours de composition, à la base de la théorie de Galois. C'est sous cette condition en effet que la loi quotient de la loi G par la relation de congruence munira le quotient d'une structure de groupe.

Proposition/Définition 17.3.26 (Sous-groupe distingué, HP)

On dit que H est un sous-groupe distingué (ou normal) de G si l'une des deux propriétés équivalentes est satisfaite :

- (i) $\forall a \in G, aH = Ha$
- (ii) $\forall a \in G, \forall h \in H, aha^{-1} \in H$.

III.5 Ordre d'un élément**Définition 17.3.27 (Ordre d'un élément d'un groupe, HP)**

L'ordre d'un élément $x \neq 1$ d'un groupe (multiplicatif) G , dont le neutre est noté 1, est

$$\text{ord}(x) = \min\{n \in \mathbb{N}^* \mid x^n = 1\}$$

Cet ordre peut être $+\infty$ par convention si l'ensemble ci-dessus est vide.

Proposition 17.3.28

Si $\text{ord}(x)$ est fini, alors les éléments $x^0, \dots, x^{\text{ord}(x)-1}$ sont deux à deux distincts, et

$$\{x^n, n \in \mathbb{Z}\} = \{x^n, n \in \llbracket 0, \text{ord}(x) - 1 \rrbracket\}.$$

On en déduit :

Théorème 17.3.29 (encore Lagrange, HP)

Soit x un élément d'un groupe fini G . Alors l'ordre de x divise l'ordre de G .

IV Anneaux et corps**IV.1 Axiomatiques des structures d'anneaux et de corps****Définition 17.4.1 (Anneau)**

Soit A un ensemble, muni de deux lois de composition internes (généralement notées $+$ et \times). On dit que $(A, +, \times)$ (ou plus simplement A) est un anneau si :

- (i) $(A, +)$ est un groupe abélien
- (ii) (A, \times) est un monoïde (autrement dit \times est associative et il existe un élément neutre 1 pour \times)
- (iii) \times est distributive sur $+$

Remarque 17.4.2

Certains ouvrages (notamment anciens) n'imposent pas l'existence de l'élément neutre 1 pour le produit et parlent alors d'*anneau unifère* ou *unitaire* pour ce que nous appelons ici simplement un *anneau*. La convention que nous adoptons concernant l'existence d'un élément neutre est celle généralement adoptée actuellement, et nous suivons en cela le programme officiel de la classe de MPSI.

Exemples 17.4.3

1. $\{0\}$ muni des opérations triviales est un anneau ; ici le neutre pour le produit est 0.
2. $\mathbb{Z}, \mathbb{R}, \mathbb{Q}$ et \mathbb{C} munis des opérations usuelles sont des anneaux.
3. Pour tout $n \in \mathbb{N}^*$, $\mathbb{Z}/n\mathbb{Z}$ est un anneau. La structure circulaire de ces anneaux explique la terminologie.
4. L'ensemble $\mathbb{R}[X]$ des polynômes à coefficients réels est un anneau. De même pour $\mathbb{Z}[X], \mathbb{Q}[X]$ ou $\mathbb{C}[X]$.
5. \mathbb{N} n'est pas un anneau.
6. L'ensemble $\mathcal{M}_n(\mathbb{R})$ des matrices carrées est un anneau.
7. L'ensemble $(\mathcal{P}(E), \Delta, \cap)$ est un anneau (anneau de Boole).
8. $(\mathbb{R}^{\mathbb{R}}, +, \circ)$ est-il un anneau ?

Proposition 17.4.4

Si A est un anneau ayant au moins deux éléments, alors $1 \neq 0$.

Définition 17.4.5 (Anneau commutatif)

On dit qu'un anneau $(A, +, \times)$ est commutatif si et seulement si la loi \times est commutative.

Les exemples donnés ci-dessus sont des exemples d'anneaux commutatifs, à l'exception d'un exemple. Lequel ?

Enfin, conformément à la définition générale, nous donnons :

Définition 17.4.6 (Homomorphisme d'anneaux)

Soit A et B deux anneaux. Un homomorphisme d'anneaux de A à B est une application $f : A \rightarrow B$, soit égale à la fonction nulle, soit vérifiant :

$$\forall (x, y) \in A^2, \quad f(x + y) = f(x) + f(y), \quad f(xy) = f(x)f(y) \quad \text{et} \quad f(1_A) = 1_B.$$

Ainsi, un homomorphisme d'anneaux (à part le morphisme nul un peu particulier) est à la fois un homomorphisme du groupe $(A, +)$ et du monoïde (A, \times) .

Un corps est un anneau vérifiant une condition supplémentaire :

Définition 17.4.7 (Corps)

Soit K un ensemble muni de deux lois $+$ et \times . On dit que $(K, +, \times)$ (ou plus simplement K) est un corps si K est un anneau commutatif tel que (K^*, \times) soit un groupe.

Ainsi K est un corps si et seulement si c'est un anneau commutatif non réduit à $\{0\}$, et tel que tout élément non nul soit inversible. En particulier, $\{0\}$ n'est en général pas considéré comme un corps.

Remarque 17.4.8

- Conformément au programme, nous adoptons la convention stipulant que tout corps doit être commutatif. Là encore, les ouvrages anciens n'imposent pas cette condition. Il est d'usage actuellement d'appeler *corps gauche* un ensemble muni d'une structure vérifiant tous les axiomes de la structure de corps, à l'exception de la commutativité de la multiplication. On rencontre aussi parfois la terminologie *anneau à divisions*, traduction de la terminologie anglaise *division ring*.
- Dans le cas des corps finis, les deux notions coïncident, d'après le théorème de Wedderburn, stipulant que « tout corps fini est commutatif », ce qui, avec notre terminologie, se réexprime : « tout corps gauche est un corps. ».

Exemples 17.4.9

1. \mathbb{R} , \mathbb{Q} et \mathbb{C} sont des corps.
2. \mathbb{Z} n'est pas un corps
3. En général $\mathbb{Z}/n\mathbb{Z}$ n'est pas un corps. Par exemple, 2 n'est pas inversible dans $\mathbb{Z}/4\mathbb{Z}$.
4. $\mathbb{Z}/p\mathbb{Z}$ est un corps si et seulement si p est premier. Ce corps est en général noté \mathbb{F}_p (comme « field », terminologie anglaise pour « corps »). Exemple \mathbb{F}_2 , corps à 2 éléments.
5. On peut montrer que tout corps fini a un cardinal égal à p^n , où p est un nombre premier et n un entier. On peut également montrer que pour de telles données, il existe (à isomorphisme près) un unique corps à $q = p^n$ éléments, qu'on note \mathbb{F}_q . Lorsque $n = 1$, on retrouve les corps \mathbb{F}_p du point précédent.

Définition 17.4.10 (Homomorphisme de corps)

Soit K et L deux corps. Un homomorphisme de corps $f : K \rightarrow L$ est un homomorphisme des anneaux sous-jacents.

Proposition 17.4.11 (Injectivité des homomorphismes de corps, HP)

Un homomorphisme de corps est toujours injectif.

Définition 17.4.12 (Caractéristique d'un corps)

Soit K un corps, d'élément neutre $1_K \neq 0_K$. Soit $H = \{n \cdot 1_K, n \in \mathbb{Z}\}$ le sous-groupe monogène de $(K, +)$ engendré par 1_K .

- Si H est infini, on dit que K est de caractéristique nulle.
- Si H est fini, de cardinal $p \in \mathbb{N}$, on dit que K est de caractéristique p .

Théorème 17.4.13 (Primalité de la caractéristique d'un corps)

Soit K un corps de caractéristique non nulle. Alors sa caractéristique p est un nombre premier.

Remarque 17.4.14

- Un corps fini est toujours de caractéristique non nulle, donc première.
- Il existe des corps infinis de caractéristique p (par exemple le corps des fractions rationnelles à coefficients dans \mathbb{F}_p)
- Si K est un corps de caractéristique p , alors pour tout x de K , $px = x + \dots + x = 0$ (avec p facteurs dans la somme).

IV.2 Sous-anneaux, sous-corps

Conformément à la définition générale, nous avons :

Définition 17.4.15 (Sous-anneau)

Soit $(A, +, \times)$ un anneau. Un sous-ensemble $B \subset A$ est un sous-anneau de A si et seulement si B est stable pour les lois $+$ et \times , et si $1_A \in B$

Remarquez qu'encore une fois, on ne dit rien de l'appartenance de 0_A à B , celle-ci étant ici aussi automatique (puisque $(B, +)$ est un sous-groupe de $(A, +)$). En revanche, l'appartenance de 1_A à B n'est pas automatique, comme le montre l'exemple de $B = \left\{ \begin{pmatrix} \lambda & 0 \\ 0 & 0 \end{pmatrix}, \lambda \in \mathbb{R} \right\}$, sous-ensemble de $\mathcal{M}_2(\mathbb{R})$, stable pour les lois $+$ et \times . Ce n'est pas un sous-anneau au sens que nous en avons donné puisque $I_2 \notin B$. En revanche, les restrictions de \times et $+$ définissent tout de même une structure d'anneau sur B , le neutre étant alors $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$.

Proposition 17.4.16 (Caractérisation des sous-anneaux)

Un sous-ensemble B d'un anneau A est un sous-anneau de A si et seulement si :

- (i) $1_A \in B$.
- (ii) pour tout $(x, y) \in B$, $x - y \in B$
- (iii) pour tout $(x, y) \in B$, $xy \in B$

Exemples 17.4.17

1. \mathbb{Z} est un sous-anneau de \mathbb{Q} qui est un sous-anneau de \mathbb{R} qui est un sous-anneau de \mathbb{C} .
2. $\mathbb{Z}/n\mathbb{Z}$ n'a d'autre sous-anneau que lui-même.

Définition 17.4.18 (Sous-corps)

Soit $L \subset K$ un sous-ensemble d'un corps K . On dit que L est un sous-corps de K si L est stable par $+$ et \times , $1_K \in L$, et si les lois induites sur L par celles de K le munissent d'une structure de corps.

Remarque 17.4.19

On pourrait remplacer l'hypothèse $1_K \in L$ par le fait que L contient un élément $x \neq 0_K$.

Proposition 17.4.20 (Caractérisation des sous-corps)

$L \subset K$ est un sous-corps de K si et seulement si :

- $1_K \in L$
- pour tout $(x, y) \in L$, $x - y \in L$
- pour tout $(x, y) \in L$ tel que $y \neq 0$, $xy^{-1} \in L$.

Exemples 17.4.21

\mathbb{Q} est un sous-corps de \mathbb{R} , \mathbb{R} est un sous-corps de \mathbb{C} .

Proposition 17.4.22 (Image par un homomorphisme)

1. Soit $f : A \rightarrow B$ un homomorphisme d'anneaux et A' un sous-anneau de A . Alors $f(A')$ est un sous-anneau de B .
2. Soit $f : K \rightarrow L$ un homomorphisme de corps et K' un sous-corps de A . Alors $f(K')$ est un sous-corps de L .

Définition 17.4.23 (Corps premier)

Un corps premier est un corps ne contenant pas de sous-corps autre que lui-même.

En particulier, tout corps contient un sous-corps isomorphe à un sous-corps premier. Par ailleurs, on a aussi :

Théorème 17.4.24 (Tout corps contient \mathbb{Q} ou $\mathbb{Z}/p\mathbb{Z}$)

Soit K un corps.

- Si K est de caractéristique nulle, il existe un homomorphisme injectif de corps $i : \mathbb{Q} \rightarrow K$.
- Si K est de caractéristique p , il existe un homomorphisme injectif de corps $i : \mathbb{F}_p \rightarrow K$.

Cela résultat permet de donner la liste de tous les corps premiers :

Proposition 17.4.25 (Description des corps premiers)

Les corps premiers sont exactement \mathbb{Q} et les \mathbb{F}_p , pour p premier.

IV.3 Calculs dans un anneau

Du fait de l'existence d'une addition et d'une multiplication dans un anneau et dans un corps, et des règles d'associativité et de commutativité, tous les calculs que l'on a l'habitude de faire dans \mathbb{Z} , \mathbb{R} ou \mathbb{C} peuvent se généraliser à un anneau ou un corps quelconque. Il faut toutefois faire attention que dans un

anneau, contrairement à ce qu'il se passe dans \mathbb{R} ou \mathbb{C} , tous les éléments ne sont pas inversibles, et que par ailleurs, les calculs nécessitant de permuter l'ordre de certains facteurs multiplicatifs ne peuvent pas être effectués en toute généralité dans un anneau non commutatif. Ainsi :

- Pour des calculs dans un anneau : considérer l'analogie avec \mathbb{Z} (plutôt que \mathbb{R}), et se méfier :
 - * des inversions intempestives
 - * des simplifications abusives
 - * des problèmes de commutativité, qui nécessitent parfois l'introduction d'hypothèses supplémentaires, à vérifier scrupuleusement.

L'analogie avec \mathbb{Z} n'est pas toujours suffisante, puisqu'il peut se produire des situations bien particulières, n'ayant pas lieu dans \mathbb{Z} , comme par exemple l'existence de « diviseurs de zéro » (voir un peu plus loin)

- Pour des calculs dans un corps : toute analogie avec \mathbb{R} est permise.

Nous rappelons qu'on peut définir dans un anneau A le produit nx où $n \in \mathbb{Z}$, et $x \in A$ de la manière suivante :

- Si $n = 0$, $nx = 0$
- Si $n > 0$, $nx = x + \dots + x$ (n facteurs)
- Si $n < 0$, $nx = -|n|x$.

Nous pouvons définir même a^n , pour tout $n \in \mathbb{N}$, et même pour tout $n \in \mathbb{Z}$ si a est inversible.

Nous voyons, outre les règles usuelles découlant des règles d'associativité et de distributivité, deux résultats déjà évoqués dans le cas de \mathbb{R} ou \mathbb{C} , et que nous voyons plus généralement dans le cadre d'anneaux, mais qui nécessitent une hypothèse de commutativité.

Théorème 17.4.26 (Factorisation de $a^n - b^n$)

Soit a et b deux éléments d'un anneau A tels que $ab = ba$. Alors pour tout $n \in \mathbb{N}^*$

$$a^n - b^n = (a - b) \sum_{k=0}^{n-1} a^{n-1-k} b^k.$$

Corollaire 17.4.27 (Factorisation de $1 - a^n$)

Pour tout élément a d'un anneau A ,

$$1 - a^n = (1 - a) \sum_{k=0}^{n-1} a^k.$$

Si $1 - a$ est inversible (condition plus forte que $a \neq 1$), on peut alors écrire :

$$(1 - a)^{-1}(1 - a^n) = \sum_{k=0}^{n-1} a^k$$

En revanche, évitez d'écrire cela sous forme de fraction lorsqu'on n'est pas dans une structure commutative, et attention à placer l'inverse du bon côté (même si, pour l'expression considérée, ce ne serait pas gênant car les facteurs considérés commutent, même si globalement l'anneau n'est pas commutatif ; mais autant prendre dès maintenant de bonnes habitudes)

Théorème 17.4.28 (Formule du binôme)

Soit a et b deux éléments d'un anneau tels que $ab = ba$. Alors, pour tout $n \in \mathbb{N}$,

$$(a + b)^n = \sum_{k=0}^n \binom{n}{k} a^k b^{n-k}.$$

Attention en revanche au cas où on n'a pas commutativité de a et b : il convient de bien distinguer les deux facteurs ab et ba apparaissant dans le développement de $(a + b)(a + b)$ (par exemple, pour $n = 2$) :

$$(a + b)(a + b) = a^2 + ab + ba + b^2 \neq a^2 + 2ab + b^2,$$

si $ab \neq ba$. Cette situation peut se produire notamment dans le cadre du produit matriciel. Il faut être toujours bien vigilant à vérifier l'hypothèse de commutativité $ab = ba$.

IV.4 Éléments inversibles

Un anneau n'étant pas nécessairement commutatif, il convient de distinguer la notion d'inversibilité à droite, inversibilité à gauche. Un inverse est alors à la fois un inverse à gauche et à droite, et en cas d'existence, il est unique, comme nous l'avons montré dans une situation générale. De plus, dans le cas d'un anneau, l'ensemble des éléments inversibles possède une structure particulière.

Théorème 17.4.29 (Groupe des inversibles d'un anneau)

Soit A un anneau. Alors l'ensemble des éléments inversibles de A , généralement noté A^ ou $U(A)$, est stable pour la loi \times , et la loi induite munit A^* d'une structure de groupe multiplicatif.*

Remarque 17.4.30

Remarquez la cohérence avec les notations déjà rencontrées \mathbb{R}^* , \mathbb{Q}^* , \mathbb{C}^* ... mais pas avec \mathbb{Z}^* .

Exemples 17.4.31

1. $(\mathcal{M}_n(\mathbb{R}))^* = \text{GL}_n(\mathbb{R})$, ensemble des matrices inversibles, appelé *groupe linéaire*
2. L'ensemble des inversibles de \mathbb{Z} : $\{-1, 1\}$
3. $(\mathbb{Z}/4\mathbb{Z})^* = \{1, 3\}$
4. $(\mathbb{Z}/p\mathbb{Z})^* = (\mathbb{Z}/p\mathbb{Z}) \setminus \{0\}$, si p est premier. On peut montrer que ce groupe multiplicatif est isomorphe au groupe additif $\mathbb{Z}/(p-1)\mathbb{Z}$.
5. Que dire plus généralement de $(\mathbb{Z}/n\mathbb{Z})^*$?

L'étude de $\mathbb{Z}/4\mathbb{Z}$ amène un résultat peu commun pour qui n'a pas l'habitude de travailler dans des structures algébriques abstraites : $2 \times 2 = 0$. Autrement dit, on a deux éléments a et b non nuls, et vérifiant $ab = 0$. La vieille règle, bien pratique pour résoudre des équations, qui nous dit que si $ab = 0$, alors $a = 0$ ou $b = 0$, ne s'applique donc pas dans ce contexte. Comme elle est bien pratique tout de même, nous allons établir un contexte dans lequel elle est vraie, en définissant une propriété adéquate des anneaux nous permettant de l'utiliser.

Définition 17.4.32 (Diviseurs de zéro, HP)

Soit a un élément d'un anneau A . On dit que a est un diviseur de 0 à gauche si $a \neq 0$ et s'il existe $b \in A$, $b \neq 0$, tel que $ab = 0$. On définit de façon symétrique les diviseurs de zéro à droite.

Remarque 17.4.33

La notion de diviseur de 0 caractérise en fait la non régularité : un élément a non nul d'un anneau est régulier à droite si et seulement si il n'est pas diviseur de 0.

Définition 17.4.34 (anneau intègre, HP)

Un anneau A est dit intègre s'il est non nul, et s'il n'admet aucun diviseur de zéro (ni à gauche, ni à droite).

En particulier, dans un anneau intègre, toutes les simplifications par des éléments non nuls sont possibles, puisque le seul élément non régulier est 0.

Exemples 17.4.35

1. \mathbb{Z} est intègre, $\mathbb{R}[X]$ est intègre, tout corps est intègre.
2. $\mathcal{M}_n(\mathbb{R})$ n'est pas intègre.
3. À quelle condition sur n , $\mathbb{Z}/n\mathbb{Z}$ est-il intègre ?

IV.5 Idéaux (HP)

La notion de sous-anneau est souvent trop restrictive, et on est souvent amené à considérer une structure moins riche :

Définition 17.4.36 (Idéal d'un anneau commutatif, HP)

Soit A un anneau commutatif, et I un sous-ensemble de A . On dit que I est un idéal si et seulement si I est un sous-groupe du groupe $(A, +)$ et si pour tout $a \in I$ et tout $\lambda \in A$, $\lambda a \in I$.

Ainsi, I est un sous-groupe de $(A, +)$, stable par multiplication par un élément de A .

Nous n'étudierons pas les idéaux cette année, mais nous illuminerons parfois quelques résultats à l'éclat de cette notion, notamment en arithmétique. Nous donnons tout de même quelques exemples importants :

Exemples 17.4.37

1. Pour tout $n \in \mathbb{N}^*$, $n\mathbb{Z}$ est un idéal de \mathbb{Z} .
2. L'ensemble des polynômes de $\mathbb{R}[X]$ s'annulant en 0 est un idéal de $\mathbb{R}[X]$. Comment généraliseriez-vous ce résultat ?
3. L'ensemble des polynômes $\{XP(X, Y) + YQ(X, Y), (P, Q) \in \mathbb{R}[X, Y]\}$ est un idéal de l'anneau $\mathbb{R}[X, Y]$ des polynômes à deux indéterminées.
4. Que peut-on dire des idéaux d'un corps ?

Dans les deux premiers exemples, on constate que l'idéal considéré est de la forme $\{\lambda a, \lambda \in A\}$, donc engendré par un unique élément a , par multiplication par les éléments λ de A . Un idéal vérifiant cette propriété est appelé *idéal principal*. Tout idéal n'est pas principal, comme le montre le troisième exemple. Un anneau dont tous les idéaux sont principaux est appelé *anneau principal*. C'est le cas par exemple de \mathbb{Z} . Cette définition, qui peut paraître anodine, est à la base d'une généralisation possible de la notion de pgcd et de ppcm à des anneaux autres que \mathbb{Z} , comme vous le verrez l'an prochain.

Arithmétique des entiers

La mathématique est la reine des sciences, mais la théorie des nombres est la reine des sciences mathématiques.

(Carl-Friedrich Gauss)

Les nombres sont le plus haut degré de la connaissance. Le nombre est la connaissance même.

(Platon)

Les nombres sont l'essence des choses.

(Pythagore)

L'arithmétique, c'est être capable de compter jusqu'à vingt sans enlever ses chaussures.

(Walt Disney)

Note Historique 18.0.38

- L'arithmétique désigne dans un premier temps l'étude des opérations élémentaires entre entiers (arithmétique élémentaire), et les algorithmes permettant de faire ces opérations (algorithmes de multiplication, division euclidienne...). C'est une des disciplines fondamentales des mathématiques dans le sens où, avec la géométrie et le calcul numérique (algébrique), elle constitue le point de départ de toutes les mathématiques.
- En découle de façon naturelle (et déjà en Grèce antique) l'étude des propriétés de divisibilité, et donc de primalité.
- Plus généralement, l'arithmétique désigne l'étude de problèmes relatifs à des nombres entiers. Ainsi, les problèmes de Diophante, liés à la recherche de solutions entières d'équations relèvent de l'arithmétique. De telles équations sont d'ailleurs appelées *équations diophantiennes*. L'exemple le plus célèbre en est certainement le fameux théorème de Fermat-Wiles stipulant que pour tout $n \geq 3$, il n'existe pas de triplet (a, b, c) d'entiers naturels non nuls tels que $a^n + b^n = c^n$. Il est instructif d'ailleurs de noter que l'ouvrage dans lequel Pierre de Fermat a écrit en marge sa fameuse note concernant une preuve de ce résultat n'est autre que *Arithmetica* de Diophante.
- L'exemple du théorème de Fermat-Wiles justifie la nécessité de sortir du cadre des entiers pour résoudre des problèmes arithmétiques en apparence simples. Ainsi, l'arithmétique a évolué en diverses branches (théorie algébrique des nombres, théorie analytique des nombres, géométrie algébrique...)
- Les concepts essentiels de l'arithmétique ont également été généralisés dans des contextes différents de celui des entiers. C'est une des motivations de l'introduction de la notion d'anneau et d'idéal. Un exemple que vous aurez l'occasion d'étudier prochainement est l'étude de l'arithmétique des polynômes. Mais cela ne s'arrête pas là !

I Divisibilité, nombres premiers

I.1 Notion de divisibilité

Définition 18.1.1 (Divisibilité, diviseur, multiple)

- Soit a et b deux entiers relatifs, $b \neq 0$. On dit que b *divise* a , et on écrit $b \mid a$, si et seulement s'il existe $q \in \mathbb{Z}$ tel que $a = bq$.
- On dit dans ce cas que b est un *diviseur* de a , et que a est un *multiple* de b .

On note $a \mid b$ pour dire que a divise b .

Ainsi, $2 \mid 4$, $-2 \mid 4$, $2 \mid -4$, et $-2 \mid -4$.

Proposition 18.1.2 (Caractérisation de la divisibilité en termes d'idéaux)

Soit a et b deux entiers positifs, $a \neq 0$. Alors $a \mid b$ si et seulement si $b\mathbb{Z} \subset a\mathbb{Z}$.

Définition 18.1.3 (couple d'entiers associés)

On dit que deux entiers a et b sont associés si et seulement si $a \mid b$ et $b \mid a$, c'est-à-dire $a\mathbb{Z} = b\mathbb{Z}$.

Proposition 18.1.4 (Caractérisation des entiers associés)

Les entiers a et b sont associés si et seulement si il existe $\varepsilon \in \{-1, 1\}$ tel que $a = \varepsilon b$.

Remarque 18.1.5

- Ce résultat peut sembler trivial et sans intérêt. Sa version plus générale, pour un anneau intègre A , est plus intéressante, et affirme que les éléments associés diffèrent d'une constante multiplicative appartenant au groupe A^* des inversibles de A .
- Par exemple, dans $\mathbb{K}[X]$, les éléments associés à un polynôme P sont tous les λP , pour $\lambda \in \mathbb{K}^*$.
- Dans $\mathbb{Z}[i]$ (entiers de Gauss), les éléments associés à un nombre z sont les 4 éléments z , $-z$, iz et $-iz$.
- Les éléments associés de x sont les éléments qui jouissent des mêmes propriétés de divisibilité que x .

Théorème/Définition 18.1.6 (Théorème de la division euclidienne)

Soit $(a, b) \in \mathbb{Z}^2$, $b \neq 0$.

- Il existe un unique couple $(q, r) \in \mathbb{Z} \times \mathbb{N}$ tel que :

$$a = bq + r \quad \text{et} \quad 0 \leq r < |b|.$$

- L'entier q est appelé *quotient* de la division euclidienne de a par b .
- L'entier r est appelé *reste* de la division euclidienne de a par b .

Remarquez que b peut être négatif.

Exemples 18.1.7

1. $27 = 6 \times 4 + 3 = 6 \times 3 + 9 = 6 \times 6 - 3$.

Ainsi, des identités $a = bq + r$, il y en a beaucoup, mais une seule vérifie la condition imposée sur r . Ici, le quotient de la division de 27 par 6 est 4, et son reste est 3.

$$2. 27 = (-6) \times (-4) + 3 = (-6) \times (-5) - 3.$$

Sans la valeur absolue dans la condition sur r , c'est la deuxième égalité qui aurait été la bonne. Mais la valeur absolue impose un reste positif. Ainsi, le quotient de la division de 27 par -6 est -4 , et le reste est 3

$$3. -27 = 6 \times (-4) - 3 = 6 \times (-5) + 3.$$

Ici, on voit que si on change le signe du nombre divisé, le quotient n'est pas simplement l'opposé (attention, cela ne correspond pas à la plupart des implémentations informatiques de la division euclidienne). Ainsi, la première identité ne convient pas. Le quotient de la division euclidienne de -27 par 6 est -5 , le reste est 3.

Remarquez que la situation est la même que pour la partie entière, pour laquelle $\lfloor -x \rfloor \neq -\lfloor x \rfloor$, sauf lorsque x est entier. C'est normale, puisque la partie entière n'est autre que le quotient de la division euclidienne (réelle) par 1.

$$4. -27 = (-6) \times 5 + 3.$$

Sans surprise, le quotient de la division euclidienne de -27 par -6 est 5, le reste est 3.

La plupart des propriétés arithmétiques de \mathbb{Z} (pour ne pas dire toutes) découlent de l'existence de cette division euclidienne. On peut définir de façon similaire dans certains anneaux une division euclidienne, la condition sur le reste étant un peu plus dure à exprimer. On parle dans ce cas d'anneau euclidien.

Définition 18.1.8 (Anneau euclidien, HP)

Soit A un anneau. On dit que A est euclidien s'il est intègre, et muni d'un stathme, c'est-à-dire d'une application $v : A \setminus \{0\} \rightarrow \mathbb{N}$ telle que :

$$\forall a \in A, \forall b \in A \setminus \{0\}, \exists (q, r)^2 \in A, a = bq + r \quad \text{et} \quad (r = 0 \text{ ou } v(r) < v(b))$$

Exemple 18.1.9

- Quel est le stathme pour la division euclidienne dans \mathbb{Z} ?
- Quel est le stathme pour la division euclidienne dans $\mathbb{C}[X]$?
- On peut montrer que $\mathbb{Z}[i]$ est euclidien, de stathme $z \mapsto |z|^2$.

Ainsi, \mathbb{Z} et $\mathbb{R}[X]$ sont des anneaux euclidiens. Cette dernière propriété nous permettra d'établir un certain nombre de propriétés arithmétiques pour les polynômes, très similaires à celles qu'on a pour les entiers.

Remarques 18.1.10

- Certains auteurs appellent préstathme la notion de stathme telle que nous l'avons définie. Ils imposent une condition supplémentaire pour les stathmes. La différence n'est pas trop gênante dans la mesure où on peut montrer qu'avec leur terminologie, tout anneau intègre muni d'un préstathme peut aussi être muni d'un stathme.
- Dans la notion générale de division euclidienne définie par stathme, on n'impose pas de propriété d'unicité. Par exemple, dans $\mathbb{Z}[i]$, on n'a pas de propriété d'unicité.

Note Historique 18.1.11

La division euclidienne est appelée ainsi par référence à Euclide qui décrit dans ses éléments le procédé algorithmique de soustractions répétées permettant d'obtenir le quotient et le reste. Cependant, on trouve trace de cette notion à des époques antérieures, notamment en Égypte.

C'est Gauss le premier, avec l'étude de $\mathbb{Z}[i]$, qui remarque que de nombreuses propriétés arithmétiques ne sont pas spécifiques à \mathbb{Z} et découlent de façon plus générale de l'existence d'une division euclidienne dans un anneau. Cette remarque est évidemment à la base de la notion d'anneau euclidien.

I.2 Congruences

De façon quasi-indissociable de la notion de division euclidienne, nous définissons :

Définition 18.1.12 (Congruences d'entiers)

Soit $n \in \mathbb{N}^*$, et $(a, b) \in \mathbb{Z}^2$. On dit que a et b sont *congrus modulo n* , et on écrit $a \equiv b [n]$, si et seulement si n divise $b - a$, ou encore si les divisions euclidiennes de a et b par n ont même reste.

On trouve aussi assez souvent la notation $a \equiv b \pmod{n}$, ou un mélange des 2 : $a \equiv b [\text{mod } n]$. Nous rappelons les résultats suivants, que nous avons déjà eu l'occasion de démontrer.

Théorème 18.1.13

La relation de congruence modulo n est une relation d'équivalence.

Théorème 18.1.14

La relation de congruence modulo n est compatible avec le produit et la somme : soit $(a, a', b, b') \in \mathbb{Z}^4$ tels que $a \equiv a' [n]$ et $b \equiv b' [n]$. Alors $a + b \equiv a' + b' [n]$ et $ab \equiv a'b' [n]$

En d'autre terme, c'est une congruence sur les monoïdes $(\mathbb{Z}, +)$ et (\mathbb{Z}, \times) , au sens vu dans le chapitre sur les ensembles.

Ces règles sont importantes pour pouvoir mener à bien le calcul modulaire de façon efficace : il permet de faire lors d'une succession d'opérations, des réductions modulo n étape par étape, plutôt que de tout calculer dans \mathbb{N} et de réduire à la fin.

Exemples 18.1.15

- Calculer le reste de la division euclidienne de $12 \times 21 \times 28 \times 18 \times 75 \times 23$ par 11.
- Calculer le reste de la division euclidienne de 1685^{1750} par 12.

Cette possibilité de réduire les opérations à chaque étape est également important pour l'implémentation informatique du calcul modulaire, permettant ainsi de travailler avec des entiers plus petit, diminuant de la sorte la complexité des calculs. On peut ainsi, contrairement au cas du calcul dans \mathbb{Z} , borner explicitement le temps de calcul des opérations modulo n par un réel dépendant de n mais ne dépendant pas des opérandes.

I.3 Nombres premiers

Nous les avons déjà rencontrés, évidemment. Nous rappelons :

Définition 18.1.16 (Nombres premiers)

Soit $p \in \mathbb{N}^*$. On dit que p est un nombre premier si p admet exactement 2 diviseurs positifs distincts (à savoir 1 et p lui-même)

Remarquez que l'existence de deux diviseurs distincts exclut d'office 1 de l'ensemble des nombres premiers, puisqu'il n'a qu'un diviseur.

Définition 18.1.17 (Nombres composés)

Soit $n \in \mathbb{N}^*$. On dit que n est un nombre composé si n possède au moins 3 diviseurs positifs distincts, ou en d'autres termes, si n possède un diviseur positif distinct de 1 et de n .

Proposition 18.1.18

Tout nombre composé admet un diviseur strict premier.

Cette proposition est à la base de l'existence de la décomposition primaire. L'unicité, quant à elle découle de ce vieux lemme, dont nous démontrons la contraposée de façon élémentaire, par descente infinie (cette démonstration est très proche de celle qu'en donne Gauss) :

Lemme 18.1.19 (Euclide)

Soit a et b deux entiers et p un entier premier tel que $p|ab$. Alors $p|a$ ou $p|b$.

Cette propriété se traduit sur les idéaux par $ab \in p\mathbb{Z} \implies a \in p\mathbb{Z}$ ou $b \in p\mathbb{Z}$, ou encore, dans $\mathbb{Z}/p\mathbb{Z}$:

$$\overline{ab} = 0 \implies \overline{a} = 0 \text{ ou } \overline{b} = 0$$

Ainsi, le lemme d'Euclide traduit le fait que $\mathbb{Z}/p\mathbb{Z}$ est intègre.

Remarque 18.1.20

Cette propriété est même une caractérisation des nombres premiers, puisque si n n'est pas premier et différent de 1, et si $n = ab$ est une décomposition non triviale, alors l'égalité $\overline{ab} = 0$ dans $\mathbb{Z}/n\mathbb{Z}$ contredit l'intégrité. Ainsi $n > 1$ est premier si et seulement si $\mathbb{Z}/n\mathbb{Z}$ est intègre.

De façon plus général, étant donné un anneau commutatif A , et un idéal I , I est en particulier un sous-groupe distingué de A , et on peut donc définir une structure de groupe quotient sur A/I . Il n'est pas dur de se rendre compte que la relation de congruence modulo I respecte également le produit et donc que celui-ci passe au quotient. Des vérifications immédiates montrent que A/I est alors muni d'une structure d'anneau. C'est par exemple ainsi qu'on obtient la structure d'anneau de $\mathbb{Z}/n\mathbb{Z}$, par quotient de l'anneau \mathbb{Z} par l'idéal $n\mathbb{Z}$. La remarque précédente incite à donner la définition suivante :

Définition 18.1.21 (Idéal premier, HP)

Soit A un anneau commutatif, et I un idéal de A . On dit que I est un idéal premier de A si et seulement si A/I est intègre.

Ainsi, dans \mathbb{Z} , p est premier si et seulement si $p\mathbb{Z}$ est un idéal premier de \mathbb{Z} différent de $\{0\}$ et \mathbb{Z} (qui sont aussi des idéaux premiers).

Théorème 18.1.22 (Combien de nombres premiers ?)

Il y a une infinité de nombres premiers.

C'est bien joli tout ça, mais comment faire pour déterminer les nombres premiers (pas trop gros) ? Erathostène, mathématicien, astronome, bibliothécaire en chef d'Alexandrie (excusez du peu), astéroïde et cratère lunaire, répondit à cette question il y a déjà très longtemps, par un procédé d'élimination.

Méthode 18.1.23 (Crible d'Ératostène)

Pour trouver tous les nombres premiers inférieurs ou égaux à n :

1. Écrire tous les nombres entiers de 2 à n .
2. Le plus petit d'eux, à savoir 2, est premier (il n'a pas de diviseur strictement plus petit que lui, autre que 1)
3. Les multiples stricts de 2 ne sont pas premiers, on les barre tous.
4. Parmi les nombres restants (en excluant les nombres premiers précédents, à savoir 2 dans la première étape, et en excluant les nombres barrés), le plus petit est premier (il n'est divisible par aucun nombre premier strictement plus petit que lui et différent de 1, sinon il serait barré). On barre tous ses multiples stricts qui ne peuvent pas être premiers, et on recommence cette étape jusqu'à épuisement de tous les entiers de la liste.

Cet algorithme est très facile à implémenter dans un langage informatique. Il n'est évidemment efficace que pour des petites valeurs de n , mais ne peut pas servir à la recherche de très grands nombres premiers. Notamment, il est à peu près inutilisable pour répondre à la question de savoir si un très grand nombre donné est premier ou non (question cruciale dans certaines situations en rapport avec des cryptages, comme la méthode RSA).

II Décomposition primaire d'un entier

II.1 Décomposition primaire

Un théorème incontournable de l'arithmétique est bien sûr :

Théorème 18.2.1 (Décomposition primaire)

Tout entier strictement positif n s'écrit de façon unique sous la forme

$$n = p_1 \times \cdots \times p_k,$$

où $p_1 \leq \cdots \leq p_k$ sont des nombres premiers, ce produit étant éventuellement vide si $n = 1$.

Un anneau dans lequel on a une propriété d'existence et d'unicité (à facteurs multiplicatifs inversibles près, et à l'ordre près des facteurs) d'une décomposition en facteurs irréductibles est appelé *anneau factoriel*. Ainsi, quitte à multiplier par l'élément inversible -1 pour obtenir la décomposition d'un entier relatif, ce résultat se réexprime en disant que \mathbb{Z} est un anneau factoriel. On peut montrer que tout anneau principal est factoriel. Par ailleurs tout anneau euclidien est principal. Donc tout anneau euclidien est factoriel. C'est par exemple le cas de $\mathbb{K}[X]$, lorsque \mathbb{K} est un corps (ainsi tout polynôme se décompose de façon unique, à éléments inversibles près, comme produit de polynômes irréductibles). C'est par exemple aussi le cas de l'anneau $\mathbb{Z}[i]$ des entiers de Gauss. La question se pose alors, notamment dans ce dernier cas, de savoir décrire les éléments irréductibles. C'est une question pas complètement triviale dans $\mathbb{Z}[i]$, en rapport avec le théorème des deux carrés (donnant la description des entiers s'écrivant comme somme de deux carrés).

II.2 Valuations p -adique

Un nombre premier p pouvant apparaître plusieurs fois dans la décomposition de n , nous définissons :

Définition 18.2.2 (Valuation p -adique)

Soit n un entier et p un entier premier. On appelle valuation p -adique de l'entier n le nombre d'occurrences (éventuellement nul) de l'entier p dans la décomposition primaire de n .

Il s'agit donc de l'unique entier v tel que p^v divise n mais pas p^{v+1} .

En notant \mathbb{P} l'ensemble des nombres premiers, il vient donc :

Proposition 18.2.3 (Reexpression de la décomposition primaire)

Pour tout $n \in \mathbb{N}^*$

$$n = \prod_{p \in \mathbb{P}} p^{v_p(n)},$$

ce produit ayant un sens, puisque constitué d'un nombre fini de termes non égaux à 1.

Proposition 18.2.4 (Règles sur les valuations)

Soit a et b deux entiers strictement positifs, et p un nombre premier.

1. On a : $v_p(ab) = v_p(a) + v_p(b)$.
2. Si b divise a , on a : $v_p\left(\frac{a}{b}\right) = v_p(a) - v_p(b)$.

Exemples 18.2.5

1. Déterminer, pour p premier, $v_p((p^2)!)$, et plus généralement $v_p((p^k)!)$, puis plus généralement $v_p(n!)$ (formule de Legendre)
2. Déterminer $v_p\left(\binom{n}{k}\right)$ en fonction de n et k .

On obtient en particulier le :

Lemme 18.2.6

Soit p un nombre premier. Alors pour tout $k \in \llbracket 1, p-1 \rrbracket$, $\binom{p}{k} \equiv 0 \pmod{p}$.

De ce lemme, on tire :

Proposition 18.2.7

Soit a et b deux entiers et p un nombre premier. Alors $(a+b)^p \equiv a^p + b^p \pmod{p}$.

Ce résultat affirme en fait que l'application $x \mapsto x^p$ est un endomorphisme du corps \mathbb{F}_p (c'est en fait l'identité d'après le petit théorème de Fermat), et plus généralement sur tout corps de caractéristique p . Il est appelé *morphisme de Frobenius*.

La proposition précédente, ainsi que nous venons de le suggérer, a un rapport très étroit avec le petit théorème de Fermat. On peut en fait démontrer ce dernier à partir de cette proposition par une récurrence assez immédiate.

Théorème 18.2.8 (Petit théorème de Fermat)

- Soit p un nombre premier, et $x \in \mathbb{Z}$. Alors : $x^p \equiv x \pmod{p}$.
- Si de plus, $x \not\equiv 0 \pmod{p}$, alors $x^{p-1} \equiv 1 \pmod{p}$.

Évidemment, cette preuve explique moins bien la raison profonde du résultat que la preuve voyant ce résultat comme un cas particulier du théorème de Lagrange, appliqué au groupe $(\mathbb{Z}/p\mathbb{Z})^*$.

Remarque 18.2.9

Le petit théorème de Fermat est notamment beaucoup utilisé dans les tests de non primalité (avec un ordinateur !). En effet, pour montrer qu'un entier p n'est pas premier, il suffit de trouver un entier a tel que $a^p \not\equiv a \pmod{p}$. Ainsi, par exemple, à l'aide d'un ordinateur, on peut trouver facilement, pour $n = \frac{1}{9}(10^{31} - 1)$ (nombre constitué de 31 chiffres 1) que $2^n \not\equiv 2 \pmod{n}$. Ainsi, n n'est pas premier. Trouver une décomposition de n est une autre paire de manches...

En revanche, déduire de la validité de tests de Fermat qu'un nombre est premier est beaucoup plus délicat, car le petit théorème de Fermat ne caractérise pas les nombres premiers : il existe des nombres composés vérifiant les identités du théorème de Fermat (la seconde identité étant alors donnée pour tout x premier avec n). Ces nombres sont appelés *nombres de Carmichael*.

III PGCD et PPCM**III.1 PGCD et PPCM d'un couple d'entiers****Lemme 18.3.1 (Somme de deux groupes abéliens)**

Soit H et K deux sous groupes d'un groupe abélien $(G, +)$. Alors $H+K$ est le plus petit groupe contenant $H \cup K$.

Proposition/Définition 18.3.2 (PGCD)

Soit a et b deux entiers positifs tels que l'un au moins des entiers a et b est non nul, et $m \in \mathbb{N}^*$. Les propositions suivantes sont équivalentes :

- (i) l'entier m est le maximum (pour l'ordre usuel) de $\{d \in \mathbb{N}^* \mid d \text{ divise } a \text{ et } d \text{ divise } b\}$
- (ii) l'entier m est le maximum (pour l'ordre de divisibilité) de $\{d \in \mathbb{N}^* \mid d \text{ divise } a \text{ et } d \text{ divise } b\}$.
- (iii) $m = \inf_{(\mathbb{N}^*, |)}(a, b)$
- (iv) $a\mathbb{Z} + b\mathbb{Z} = m\mathbb{Z}$

Si l'une de ces quatre conditions équivalentes est satisfaite, on dit que m est le *plus grand commun diviseur* de a et b (en abrégé : PGCD), et on le note $a \wedge b$.

Ainsi, le PGCD de a et b est entièrement caractérisé par l'égalité des idéaux (en notant (a) l'idéal engendré par a) :

$$(a) + (b) = (a \wedge b).$$

Remarquez que pour établir ce point partant de la description usuelle (premier point), on se sert du fait que tout idéal de \mathbb{Z} s'écrit $\mathbb{Z} \cdot a$, donc que \mathbb{Z} est principal. Le fait que \mathbb{Z} est principal nous assure également que le plus petit idéal contenant a et b est engendré par un élément. C'est là une façon de définir le pgcd, comme élément générateur de l'idéal engendré par a et b .

Cette définition est valide dans tout anneau principal :

Définition 18.3.3 (PGCD dans un anneau principal, HP)

Soit A un anneau principal et a et b deux éléments de A . Un PGCD d de a et b est un élément défini de façon unique à inversibles près par :

$$(d) = (a) + (b).$$

Dans certains cas, un choix de pgcd s'impose (le pgcd positif dans \mathbb{Z} par exemple). Dans ce cas on peut utiliser une notation non ambiguë ($a \wedge b$ par exemple), et parler du PGCD. Dans les autres cas, on parle d'UN PGCD, et on ne peut utiliser une notation qu'à abus près.

Le PGCD se détermine très facilement algorithmiquement en remarquant que si $a = bq + r$, $a \wedge b = b \wedge r$. Ainsi, en prenant des restes divisions euclidiennes successives le dernier reste non nul fournira le PGCD :

Algorithme 18.1 : Algorithme d'Euclide pour le calcul du PGCD

Entrée : a, b : entiers naturels
Sortie : $a \wedge b$
tant que $b > 0$ **faire**
 | $a, b \leftarrow b, a \% b$
fin tant que
renvoyer a

Proposition/Définition 18.3.4 (PPCM)

Soit a et b deux entiers non nuls, et $M \in \mathbb{N}^*$. Les propositions suivantes sont équivalentes :

- (i) l'entier M est le minimum (pour l'ordre usuel) de $\{m \in \mathbb{N}^* \mid a \text{ divise } m \text{ et } b \text{ divise } m\}$
- (ii) l'entier M est le minimum (pour l'ordre de divisibilité) de $\{m \in \mathbb{N}^* \mid a \text{ divise } m \text{ et } b \text{ divise } m\}$
- (iii) $M = \sup_{(\mathbb{N}^*, |)}(a, b)$
- (iv) $M\mathbb{Z} = a\mathbb{Z} \cap b\mathbb{Z}$.

Si l'une de ces quatre propositions équivalentes est satisfaite, on dit que M est le *plus petit commun multiple* de a et b (PPCM en abrégé), et on note $M = a \vee b$.

Encore une fois, ce dernier point peut être pris comme définition dans un anneau principal.

Définition 18.3.5 (PPCM dans un anneau principal, HP)

Soit A un anneau principal et a et b deux éléments non nuls de A . Alors un PPCM de a et b est un élément m tel que

$$(m) = (a) \cap (b).$$

Ce qui est parfois évident sur les idéaux ne l'est pas toujours autant pour les autres descriptions :

Proposition 18.3.6 (Distributivité du produit sur \wedge et \vee)

Soit a et b deux entiers naturels, et c un entier naturel non nul.

1. Si a et b ne sont pas tous les deux nuls, $(a \wedge b) \cdot c = (ac) \wedge (bc)$.
2. Si a et b sont non nuls, $(a \vee b) \cdot c = (ac) \vee (bc)$.

III.2 Identité de Bézout

L'identité de Bézout est elle aussi une conséquence immédiate de la caractérisation par les idéaux :

Théorème 18.3.7 (identité de Bézout, ou théorème de Bachet-Bézout)

1. Soit a et b deux entiers dont l'un au moins est non nul. Alors il existe des entiers relatifs x et y tels que $ax + by = a \wedge b$.
2. Réciproquement, étant donné un entier $d \in \mathbb{N}^*$, s'il existe des entiers relatifs x et y tels que

$$d = ax + by,$$

alors $a \wedge b \mid d$.

Note Historique 18.3.8

- C'est le nom d'Étienne Bézout, mathématicien français du 18^e siècle, qui est le plus souvent associé à ce résultat. C'est pourtant à Claude-Gaspard Bachet de Méziriac que l'on doit la première preuve, parue dans son ouvrage *Problèmes plaisans et délectables qui se font par les nombres*, paru en 1624. Sa preuve est celle que nous présentons ci-dessous (par l'algorithme d'Euclide)
- Qu'a fait Bézout alors pour avoir droit à tous ces honneurs ? Il a généralisé le résultat à d'autres situations, notamment au cas des polynômes.
- Il est intéressant de noter que le fameux ouvrage dans lequel Fermat écrivit dans une marge qu'il savait démontrer ce qu'on appelle aujourd'hui le théorème de Fermat-Wiles est en fait une traduction par Bachet de Méziriac de l'*Arithmétique* de Diophante. Le monde est petit...

La démonstration passant par les idéaux peut se généraliser dans un anneau principal. Elle possède l'inconvénient de ne pas être constructive. Il peut être intéressant de trouver explicitement des entiers x et y assurant l'égalité $ax + by = a \wedge b$. L'algorithme de la division euclidienne itéré permet à la fois de déterminer $a \wedge b$, et d'obtenir une identité de Bézout : il s'agit de l'algorithme d'Euclide. Il est intéressant de noter que cet algorithme est valide à partir du moment où l'on dispose d'une notion de division euclidienne : il peut donc être généralisé à tout anneau euclidien, dans le sens évoqué précédemment. Ainsi, par exemple, il nous permettra d'obtenir des identités de Bézout dans $\mathbb{R}[X]$.

Lemme 18.3.9

Soit a et b deux entiers positifs, $b \neq 0$. Soit r le reste de la division euclidienne de a par b . Alors $a \wedge b = b \wedge r$.

Théorème 18.3.10 (Algorithme d'Euclide)

- Soit a et b deux entiers positifs, $b \neq 0$. En effectuant la division euclidienne de a par b , puis en effectuant la division euclidienne de b par le reste obtenue, et en continuant ainsi en divisant l'ancien reste par le nouveau reste, on finit par obtenir un reste nul.
- Le dernier reste non nul est alors égal au PGCD de a et de b .
- L'identité de la division euclidienne permet alors d'écrire, étape par étape, les restes successifs comme combinaison linéaire de a et b à coefficients dans \mathbb{Z} . La dernière étape fournit une identité de Bézout.

Ainsi, en écrivant $r_0 = a$, $r_1 = b$ puis les divisions euclidiennes successives :

$$\begin{cases} r_0 &= r_1 q_2 + r_2 \\ r_1 &= r_2 q_3 + r_3 \\ &\vdots \\ r_{k-2} &= r_{k-1} q_k + r_k \\ r_{k-1} &= r_k q_{k+1} + r_{k+1}, \end{cases}$$

avec $r_2 \neq 0$, $r_3 \neq 0, \dots, r_k \neq 0$ et $r_{k+1} = 0$, on a $r_k = a \wedge b$. De plus, en posant $x_0 = 1$, $x_1 = 0$, $y_0 = 0$, $y_1 = 1$ et pour tout $i \leq [3, k]$

$$x_i = x_{i-2} - q_i x_{i-1} \quad \text{et} \quad y_i = y_{i-2} - q_i y_{i-1},$$

on obtient pour tout $i \in [1, n]$,

$$r_i = ax_i + by_i,$$

donc en particulier pour $i = k$, on obtient une identité de Bézout :

$$a \wedge b = ax_k + by_k.$$

On peut donc décrire de façon plus algorithmique :

Algorithme 18.2 : Algorithme d'Euclide étendu

Entrée : a, b : entiers naturels non nuls
Sortie : m, u, v tels que $m = a \wedge b = ua + bv$
 $u, v, w, x, r, s \leftarrow 1, 0, 0, 1, a, b$;
tant que $s \neq 0$ **faire**
 $q, s, r \leftarrow r // s, r \% s, s$;
 $w, u \leftarrow u - qw, w$;
 $x, v \leftarrow v - qx, x$
fin tant que
renvoyer (r, u, v)

En pratique, pour ne pas s'embrouiller, mieux vaut écrire les différentes relations obtenues par la division euclidienne, en remplaçant étape par étape les restes obtenus par leur expression obtenue récursivement en fonction de a et b .

Exemple 18.3.11

- Trouver à l'aide de l'algorithme d'Euclide le pgcd de 27 et 33, ainsi qu'une identité de Bézout.
- Comment trouver une autre identité de Bézout ?
- À retenir : on n'a pas unicité de la relation de Bézout !

III.3 PGCD et PPCM d'une famille finie d'entiers

La notion de PGCD et de PPCM de deux entiers peut être généralisée à un plus grand nombre d'entiers :

Proposition/Définition 18.3.12 (PGCD d'un nombre fini d'entiers)

Soit a_1, \dots, a_n des entiers naturels, non tous nuls, et m un entier naturel. Les propriétés suivantes sont équivalentes :

- m est le maximum (au sens de l'ordre usuel) des entiers d qui divisent chacun des $a_i, i \in \llbracket 1, n \rrbracket$.
- m est le maximum (au sens de la divisibilité) des entiers d qui divisent chacun des $a_i, i \in \llbracket 1, n \rrbracket$.
- $m = \inf_{(\mathbb{N}^*, |)}(a_1, \dots, a_n)$
- $m\mathbb{Z} = a_1\mathbb{Z} + a_2\mathbb{Z} + \dots + a_n\mathbb{Z}$.

Si l'une de ces quatre propositions équivalentes est satisfaite, on dit que m est le PGCD de la famille (a_1, \dots, a_n) et on note $m = a_1 \wedge a_2 \wedge \dots \wedge a_n$.

De la même façon :

Proposition/Définition 18.3.13 (PPCM d'un nombre fini d'entiers)

Soit a_1, \dots, a_n des entiers naturels, non nuls, et m un entier naturel. Les propriétés suivantes sont équivalentes :

- m est le minimum (au sens de l'ordre usuel) des entiers m multiples de chacun des $a_i, i \in \llbracket 1, n \rrbracket$.
- m est le minimum (au sens de la divisibilité) des entiers m multiples de chacun des $a_i, i \in \llbracket 1, n \rrbracket$.
- $m = \sup_{(\mathbb{N}^*, |)}(a_1, \dots, a_n)$
- $m\mathbb{Z} = a_1\mathbb{Z} \cap a_2\mathbb{Z} \cap \dots \cap a_n\mathbb{Z}$.

Si l'une de ces quatre propositions équivalentes est satisfaite, on dit que m est le PPCM de la famille (a_1, \dots, a_n) et on note $m = a_1 \vee a_2 \vee \dots \vee a_n$.

La caractérisation par idéaux, ou encore la caractérisation par borne inférieure (et l'associativité des

bornes inférieures) nous assure que ces notions correspondent aux PGCD et PPCM itérés :

$$a_1 \wedge \cdots \wedge a_n = ((a_1 \wedge a_2) \wedge \cdots) \wedge a_n \quad \text{et} \quad a_1 \vee \cdots \vee a_n = ((a_1 \vee a_2) \vee \cdots) \vee a_n.$$

En particulier, on en tire l'associativité de \wedge et de \vee .

On peut étendre le théorème de Bachet-Bézout à cette situation, toujours en utilisant la caractérisation par les idéaux :

Théorème 18.3.14 (Relation de Bézout)

Soit a_1, \dots, a_n des entiers naturels non tous nuls. Alors il existe des entiers relatifs x_1, \dots, x_n tels que

$$a_1 \wedge \cdots \wedge a_n = x_1 a_1 + \cdots + x_n a_n.$$

Réciproquement, s'il existe des entiers x_1, \dots, x_n tels que

$$d = x_1 a_1 + \cdots + x_n a_n,$$

alors d est un multiple de $a_1 \wedge \cdots \wedge a_n$.

Méthode 18.3.15

Les coefficients x_1, \dots, x_n peuvent se trouver explicitement, par itération de l'algorithme d'Euclide : on cherche d'abord une relation de Bézout entre $d_1 = a_1 \wedge a_2$, a_1 et a_2 , puis entre $d_2 = d_1 \wedge a_3$, d_1 et a_2 ; en substituant à d_1 la première relation trouvée, on obtient une relation de Bézout entre $a_1 \wedge a_2 \wedge a_3$, a_1 , a_2 et a_3 . On continue alors de la sorte, de proche en proche.

Enfin, toutes les notions introduites dans ce paragraphe peuvent être généralisées à des entiers relatifs quelconques ; le pgcd et le ppcm ne sont alors définis correctement qu'au signe près (c'est le cas général dans un anneau principal, ou le pgcd ne peut être déterminé qu'à un facteur multiplicatif inversible près). Dans le cas de \mathbb{Z} , on a un choix privilégié qui consiste à prendre la valeur positive. Le pgcd et les relations de Bézout se trouvent de la même façon, en les cherchant d'abord pour les valeurs absolues, puis en modifiant les signes de façon adéquate.

III.4 PGCD et PPCM vus sous l'angle de la décomposition primaire

Nous traduisons d'abord la divisibilité en terme de dévomposition primaire :

Lemme 18.3.16 (Caractérisation de la divisibilité par les valuations)

Soit a et b deux entiers non nuls. Alors $a|b$ si et seulement si pour tout $p \in \mathbb{P}$, $v_p(a) \leq v_p(b)$.

Étant donné deux nombres a et b , le pgcd et le ppcm de a et b s'obtiennent facilement à l'aide de leur décomposition primaire. Par exemple :

$$150 = 2 \times 3 \times 5^2 \quad \text{et} \quad 180 = 2^2 \times 3^2 \times 5.$$

Ainsi, $150 \wedge 180 = 2 \times 3 \times 5 = 30$ et $150 \vee 180 = 2^2 \times 3^2 \times 5^2 = 900$.

Plus généralement, en utilisant le lemme énoncé ci-dessus, on obtient :

Proposition 18.3.17

Soit a et b deux entiers strictement positifs. Alors, pour tout $p \in \mathbb{P}$,

$$v_p(a \wedge b) = \min(v_p(a), v_p(b)) \quad \text{et} \quad v_p(a \vee b) = \max(v_p(a), v_p(b)).$$

Cette propriété peut bien sûr être généralisée au calcul du PGCD et du PPCM d'une famille finie quelconque d'entiers naturels. On en déduit en particulier la relation suivante entre PGCD et PPCM :

Proposition 18.3.18 (relation liant le PGCD et le PPCM)

Soit a et b deux entiers naturels non nuls. Alors

$$(a \wedge b) \times (a \vee b) = ab.$$

Nous retrouverons cette relation plus tard sans nous servir de la décomposition primaire.

IV Entiers premiers entre eux

IV.1 Couple d'entiers premiers entre eux

Définition 18.4.1 (Entiers premiers entre eux)

Soit a et b deux entiers naturels non tous les deux nuls. On dit que a et b sont premiers entre eux si et seulement si $a \wedge b = 1$, donc si a et b n'ont pas d'autre diviseur positif commun que 1.

Cela peut aussi s'exprimer par la relation $a\mathbb{Z} + b\mathbb{Z} = \mathbb{Z}$.

Note Historique 18.4.2

La première apparition de cette notion est dans le Livre VII des *Éléments* d'Euclide.

Proposition 18.4.3 (Simplification des fractions)

- Soit a et b deux entiers naturels, $b \neq 0$. Alors $\frac{a}{a \wedge b}$ et $\frac{b}{a \wedge b}$ sont premiers entre eux.
- En particulier, il est toujours possible d'écrire un rationnel $\frac{a}{b}$ sous forme irréductible $\frac{a'}{b'}$, c'est-à-dire de sorte que $a' \wedge b' = 1$, en simplifiant par $a \wedge b$.

On déduit des résultats de la section précédente :

Théorème 18.4.4 (Bézout, ou Bachet-Bézout)

Deux entiers naturels a et b sont premiers entre eux si et seulement s'il existe des entiers relatifs x et y tels que $ax + by = 1$.

En particulier :

Corollaire 18.4.5 (Éléments inversibles de $\mathbb{Z}/n\mathbb{Z}$, HP)

1. Soit $n \in \mathbb{N}^*$, et $k \in \llbracket 0, n-1 \rrbracket$. La classe de k modulo n est inversible dans $\mathbb{Z}/n\mathbb{Z}$ si et seulement si k et n sont premiers entre eux.
2. En particulier, si p est premier, $\mathbb{Z}/p\mathbb{Z}$ est un corps.

Méthode 18.4.6 (Calcul d'un inverse modulo n)

Soit k premier avec n . Pour calculer l'inverse de k modulo n , c'est-à-dire l'inverse de k dans $\mathbb{Z}/n\mathbb{Z}$, déterminer une relation de Bézout $xk + yn = 1$, par l'algorithme d'Euclide. On obtient alors $xk \equiv 1 [n]$.

Un résultat souvent très utile pour les propriétés de divisibilité, et assez voisin du lemme d'Euclide (le lemme d'Euclide est d'ailleurs un cas particulier du lemme de Gauss) :

Lemme 18.4.7 (Lemme ou théorème de Gauss)

Soit a, b et c trois entiers naturels tels que $a \mid bc$ et $a \wedge b = 1$. Alors $a \mid c$.

Note Historique 18.4.8

Gauss démontre le lemme d'Euclide de façon élémentaire (la démonstration qu'on en a donnée), puis en déduit l'existence et l'unicité de la décomposition primaire, qu'il utilise pour démontrer son propre théorème ci-dessus. La démonstration que nous en donnons est indépendante du lemme d'Euclide, et le lemme d'Euclide peut alors être vu comme conséquence du lemme de Gauss.

En utilisant le lemme de Gauss de façon convenable sur une relation de Bézout, on obtient, indépendamment de la décomposition primaire :

Proposition 18.4.9

Si a et b sont premiers entre eux et $a \mid c$ et $b \mid c$, alors $ab \mid c$.

Corollaire 18.4.10 (PPCM de deux nombres premiers entre eux)

Si a et b sont premiers entre eux, $a \vee b = ab$

Cela permet de retrouver, indépendamment de la décomposition primaire, la relation entre PGCD et PPCM :

Proposition 18.4.11 (relation liant PGCD et PPCM)

Soit a et b deux entiers strictement positifs. Alors $ab = (a \wedge b)(a \vee b)$.

IV.2 Famille finie d'entiers premiers entre eux

Enfin, nous définissons deux notions sur un nombre quelconque d'entiers, à bien distinguer l'une de l'autre :

Définition 18.4.12 (Nombres premiers entre eux deux à deux)

Soit a_1, \dots, a_n des entiers naturels. On dit que a_1, \dots, a_n sont premiers entre eux deux à deux si deux entiers pris au hasard parmi ces n entiers sont toujours premiers entre eux, c'est-à-dire :

$$\forall (i, j) \in \llbracket 1, n \rrbracket^2, (i \neq j) \implies a_i \wedge a_j = 1.$$

La notion suivante est moins forte :

Définition 18.4.13 (Nombres premiers entre eux dans leur ensemble)

Soit a_1, \dots, a_n des entiers naturels. On dit que (a_1, \dots, a_n) sont premiers entre eux dans leur ensemble si $a_1 \wedge \dots \wedge a_n = 1$, ou de façon équivalente, s'il existe des entiers x_1, \dots, x_n tels que

$$x_1 a_1 + \dots + x_n a_n = 1.$$

Par exemple 10, 12 et 15 sont premiers entre eux dans leur ensemble. Vous remarquerez en revanche que deux quelconques d'entre eux ne sont pas premiers entre eux !

La réciproque, en revanche est vraie :

Proposition 18.4.14

Soit $(a_1, \dots, a_n) \in \mathbb{N}^n$. Si a_1, \dots, a_n sont premiers entre eux deux à deux (il suffit même en fait que deux d'entre eux soient premiers entre eux) alors ils sont premiers entre eux dans leur ensemble.

IV.3 Fonction indicatrice d'Euler

Arrivé à ce stade, nous ne pouvons nous empêcher de parler de la fonction indicatrice d'Euler, d'une importance capitale en arithmétique.

Définition 18.4.15 (Fonction indicatrice d'Euler, ou fonction phi d'Euler, HP)

La fonction φ d'Euler est la fonction qui à tout n de \mathbb{N}^* associe $\varphi(n)$ le nombre d'entiers de $[1, n - 1]$ premiers avec n .

En particulier, les résultats précédents amènent :

Proposition 18.4.16 (Cardinal de $(\mathbb{Z}/n\mathbb{Z})^*$, HP)

Soit $n \in \mathbb{N}^$. Alors $(\mathbb{Z}/n\mathbb{Z})^*$ est de cardinal $\varphi(n)$.*

Grâce au théorème de Lagrange, on en déduit notamment la généralisation suivante du petit théorème de Fermat

Corollaire 18.4.17 (Théorème d'Euler, HP)

Soit $n \in \mathbb{N}^$. Alors pour tout $x \in \mathbb{N}^*$ tel que $x \wedge n = 1$, $x^{\varphi(n)} \equiv 1 [n]$.*

On peut montrer (voir exercices ou problèmes) que φ est multiplicative : $\varphi(ab) = \varphi(a)\varphi(b)$. On peut également calculer facilement $\varphi(p^k)$, pour p premier. On peut en déduire une expression de $\varphi(n)$ pour tout n , à condition de connaître la décomposition primaire de n .

V Théorème des restes chinois

Nous nous intéressons dans cette section à la résolution de systèmes de congruences.

Note Historique 18.5.1

Le nom de théorème des restes chinois provient du fait que le mathématicien chinois Sun Zi du III^e siècle répond à la question suivante : « Soit une armée. Si on range les soldats par 3 il en reste 2, si on les range par 5, il en reste 3 et si on les range par 7 il en reste 2. Combien y a-t-il de soldats ? ». La réponse de Sun Zi est : « Multiplie le reste de la division par 3, c'est-à-dire 2, par 70, ajoute-lui le produit du reste de la division par 5, c'est-à-dire 3, avec 21 puis ajoute le produit du reste de la division par 7, c'est-à-dire 2 par 15. Tant que le nombre est plus grand que 105, retire 105. »

V.1 Cas de modulo premiers entre eux

La justification du théorème des restes chinois repose sur le lemme suivant portant sur les groupes (évidemment, c'est un point de vue dont ne disposait pas Sun Zi) :

Théorème 18.5.2 (Produit de groupes cycliques d'ordre premiers entre eux)

1. Soit a et b deux entiers premiers entre eux. Alors $\mathbb{Z}/ab\mathbb{Z}$ est isomorphe à $\mathbb{Z}/a\mathbb{Z} \times \mathbb{Z}/b\mathbb{Z}$, un isomorphisme explicite étant donné par $\bar{x} \mapsto x(1, 1) = (x \bmod a, x \bmod b)$.
2. Plus généralement, si a_1, \dots, a_n sont deux à deux premiers entre eux, alors

$$\mathbb{Z}/(a_1 \cdots a_n)\mathbb{Z} \simeq \mathbb{Z}/a_1\mathbb{Z} \times \cdots \times \mathbb{Z}/a_n\mathbb{Z},$$

l'isomorphisme étant donné explicitement par $\bar{x} = (x \bmod a_1, \dots, x \bmod a_n)$.

Corollaire 18.5.3 (Unicité modulo $a_1 \cdots a_n$ de la solution d'un système)

Soit a_1, \dots, a_n des entiers deux à deux premiers entre eux, et b_1, \dots, b_n des entiers. Alors le système

$$\begin{cases} x \equiv b_1 [a_1] \\ \dots \\ x \equiv b_n [a_n] \end{cases}$$

admet une solution, unique modulo $a_1 \cdots a_n$.

Corollaire 18.5.4 (cas d'un second membre nul)

En particulier, si $b_1 = \cdots = b_n = 0$, la seule solution modulo $a_1 \cdots a_n$ est 0.

En notant $A = a_1 \cdots a_n$, et pour tout $B = (b_1, \dots, b_n)$, $X(B)$ l'unique élément de $\mathbb{Z}/A\mathbb{Z}$ solution du système ci-dessus, on obtient :

Proposition 18.5.5

L'application $B \mapsto X(B)$ est un morphisme de groupes, de $(\mathbb{Z}^n, +)$ dans $\mathbb{Z}/A\mathbb{Z}$.

Ainsi, il suffit de déterminer les valeurs de $X(e_i)$ pour les vecteurs $e_i = (0, \dots, 0, 1, 0, \dots, 0)$, le 0 étant en position i . En effet, pour un vecteur $B = (b_1, \dots, b_n)$, on aura alors

$$X(B) = \sum_{i=1}^n b_i X(e_i).$$

Pour déterminer $X(e_i)$, on part de la constatation que les résultats précédents permettent de réduire le système au système à deux inconnues :

$$\begin{cases} x \equiv 1 [a_i] \\ x \equiv 0 [\hat{a}_i], \end{cases}$$

où nous avons noté $\hat{a}_i = \prod_{j \neq i} a_j$. Ce système peut être résolu en utilisant une relation de Bézout, du fait que a_i et \hat{a}_i sont premiers entre eux. On commence donc par déterminer (par l'algorithme d'Euclide étendu) u_i et v_i tels que

$$u_i a_i + v_i \hat{a}_i = 1.$$

Nous avons alors :

$$v_i \hat{a}_i \equiv 0 [\hat{a}_i] \quad \text{et} \quad v_i \hat{a}_i \equiv 1 [a_i].$$

Ainsi, $X(e_i) = \overline{v_i \hat{a}_i}$ (classe dans $\mathbb{Z}/A\mathbb{Z}$).

On énonce :

Théorème 18.5.6 (Théorème des restes chinois)

Si les a_i sont deux à deux premiers entre eux, il existe modulo $a_1 \dots a_n$ une unique solution au système $x \equiv b_i \pmod{a_i}$, $i \in \llbracket 1, n \rrbracket$ donné par

$$x \equiv \sum_{i=1}^n b_i v_i \hat{a}_i \pmod{a_1 \dots a_n},$$

où $\hat{a}_i = \prod_{j \neq i} a_j$ et v_i est un coefficient d'une relation de Bézout $u_i a_i + v_i \hat{a}_i = 1$.

V.2 Résolution d'un système quelconque

Les a_i n'étant plus supposés premiers entre eux, en écrivant chaque a_i sous la forme $p_1^{\alpha_1} \dots p_k^{\alpha_k}$, on remplace chaque ligne du système par un sous-système équivalent :

$$\begin{cases} x \equiv a_i \pmod{p_1^{\alpha_1}} \\ \dots \\ x \equiv a_i \pmod{p_k^{\alpha_k}} \end{cases}$$

Ainsi, toutes les équations du nouveau système obtenu sont réduites modulo une puissance d'un nombre premier. Un même nombre premier peut intervenir dans plusieurs lignes, avec un exposant éventuellement différent. La compatibilité de ces équations est aisée à vérifier (pour chaque paire d'équations modulo p^α et p^β , avec disons $\alpha \leq \beta$, on doit obtenir les mêmes équations en réduisant la seconde modulo p^α). Si les équations ne sont pas compatibles, il n'y a pas de solution, sinon, on garde la plus contraignante des équations (à savoir celle faisant intervenir le plus grand exposant). On est alors ramené à un système tel qu'étudié plus haut, auquel on peut appliquer le théorème des restes chinois.

Exemple 18.5.7

Résoudre les deux systèmes suivants :

1. $\begin{cases} x \equiv 3 \pmod{42} \\ x \equiv 10 \pmod{49} \end{cases}$
2. $\begin{cases} x \equiv 3 \pmod{42} \\ x \equiv 9 \pmod{49} \end{cases}$

Polynômes et fractions rationnelles

On a ainsi traité le problème comme s'il s'agissait simplement de déterminer la forme des racines, dont l'existence est admise sans démonstration, manière de raisonner qui est ici entièrement illusoire et en fait une véritable petitio principis.

(C. F. Gauss, à propos des « démonstrations » antérieures du théorème de d'Alembert-Gauss)

Le but de ce chapitre est d'étudier les fonctions polynomiales $x \mapsto a_d x^d + \dots + a_1 x + a_0$. On s'intéresse notamment aux propriétés arithmétiques (produit, somme, divisibilité...) et aux propriétés analytiques (racines, dérivation...)

On se placera dans un cadre plus formel dans le but notamment de généraliser des constructions *a priori* uniquement valables pour des polynômes à coefficients réels (comme la dérivation) à des polynômes à coefficients dans des anneaux plus généraux.

Seuls les polynômes à coefficients dans \mathbb{R} ou \mathbb{C} sont théoriquement au programme. Nous donnerons les définitions formelles plus généralement pour les polynômes à coefficients dans un anneau. Ce point de vue a une certaine importance, car c'est lui qui permet d'itérer ensuite la construction pour obtenir les polynômes de plusieurs indéterminés, puisque si \mathbb{A} est un anneau, $\mathbb{A}[X]$ hérite de cette structure d'anneau. En revanche, si \mathbb{K} est un corps, on n'a pas de structure de corps sur $\mathbb{K}[X]$, mais uniquement d'anneau.

Pour l'étude des propriétés de l'anneau des polynômes, nous nous limiterons au cas où les coefficients sont dans un corps. On dispose dans ce cas de propriétés plus fortes que dans le cas général des polynômes sur un anneau, en particulier toutes les propriétés permettant de faire de l'arithmétique. On a même parfois besoin de certaines hypothèses supplémentaires (par exemple sur la caractéristique du corps). Ainsi, pour certaines propriétés, nous reviendrons aux exigences du programme (\mathbb{R} ou \mathbb{C}), en précisant parfois ce qu'il en est dans les autres situations. Il convient de remarquer que dans ce cas, ces propriétés ne sont en général plus satisfaites pour les polynômes à plusieurs indéterminées, puisque $\mathbb{A}[X]$ n'est pas un corps.

I Polynômes à coefficients dans un anneau commutatif

Soit \mathbb{A} un anneau, qu'on supposera commutatif.

I.1 Polynômes formels

Remarque 19.1.1 (Motivation de la définition)

Une fonction polynomiale réelle est entièrement déterminée par la suite de ses coefficients. Les différentes constructions telles la somme, le produit, la dérivation, peuvent s'écrire uniquement sur les coefficients.

La remarque précédente semble justifier de considérer un polynôme comme une suite de coefficients. Seul un nombre fini de ces coefficients doit être non nul.

Définition 19.1.2 (Polynômes formels)

- Un polynôme formel P à coefficients dans \mathbb{A} est une suite $(a_n)_{n \in \mathbb{N}}$ d'éléments de \mathbb{A} , nulle à partir d'un certain rang.
- Le réel a_k est appelé k -ième coefficient de P , ou coefficient du monôme de degré k de P .
- L'ensemble des polynômes (formels) à coefficients dans \mathbb{A} est noté $\mathbb{A}[X]$.

Exemples 19.1.3

1. $\mathbb{R}[X]$ ou $\mathbb{C}[X]$, polynômes formels à coefficients réels ou complexes ;
2. $\mathbb{Q}[X]$, ensemble des polynômes à coefficients rationnels ;
3. $\mathbb{Z}[X]$, ensemble des polynômes à coefficients entiers ;
4. $(\mathbb{Z}/n\mathbb{Z})[X]$, polynômes à coefficients dans $\mathbb{Z}/n\mathbb{Z}$, dont un cas particulier important est $\mathbb{F}_p[X]$.

I.2 Opérations arithmétiques sur les polynômes

Les polynômes considérés dans cette section sont à coefficients dans un anneau commutatif.

Définition 19.1.4 (Somme de polynômes de $\mathbb{A}[X]$)

La somme de deux polynômes $P = (a_n)_{n \in \mathbb{N}}$ et $Q = (b_n)_{n \in \mathbb{N}}$ de $\mathbb{A}[X]$ est la suite $P + Q = (a_n + b_n)_{n \in \mathbb{N}}$, qui est bien nulle à partir d'un certain rang.

Définition 19.1.5 (Produit d'un polynôme de $\mathbb{A}[X]$ par un élément de \mathbb{A})

Le produit de $P = (a_n)_{n \in \mathbb{N}}$ par $\lambda \in \mathbb{A}$ est $\lambda P = (\lambda \cdot a_n)_{n \in \mathbb{N}}$.

Lorsque \mathbb{A} est un corps, on parle de multiplication par un scalaire.

Définition 19.1.6 (Produit de deux polynômes de $\mathbb{A}[X]$)

Soit $P = (a_n)_{n \in \mathbb{N}}$ et $Q = (b_n)_{n \in \mathbb{N}}$ deux polynômes de $\mathbb{A}[X]$. Soit pour tout $n \in \mathbb{N}$, $c_n = \sum_{k=0}^n a_k b_{n-k}$.

Alors $(c_n)_{n \in \mathbb{N}}$ est un polynôme. On définit alors $PQ = (c_n)_{n \in \mathbb{N}}$.

La suite $\mathbf{c} = (c_n)_{su}$ est appelée *produit de convolution* des suites $\mathbf{a} = (a_n)$ et $\mathbf{b} = (b_n)$, et est parfois notée $\mathbf{c} = \mathbf{a} \star \mathbf{b}$.

Théorème 19.1.7 (Structure d'anneau de $\mathbb{A}[X]$)

La somme et le produit définis ci-dessus munissent $\mathbb{A}[X]$ d'une structure d'anneau commutatif.

Avertissement 19.1.8

Ne généralisez pas trop vite « \mathbb{A} anneau $\implies \mathbb{A}[X]$ anneau » en « \mathbb{K} corps $\implies \mathbb{K}[X]$ corps ». Cette dernière affirmation est fautive ! Ainsi, si \mathbb{K} est un corps, tout ce qu'on peut dire, c'est que $\mathbb{K}[X]$ est un anneau.

Cependant, lorsque \mathbb{K} est un corps on peut munir $\mathbb{K}[X]$ d'une structure plus riche. En effet, la multiplication par un scalaire munit $\mathbb{K}[X]$ d'une structure d'espace vectoriel sur le corps \mathbb{K} , compatible d'une

certaine manière avec la structure d'anneau (voir chapitre ultérieur). La structure totale (espace vectoriel + anneau) est appelée structure d'algèbre. Dans le cas où \mathbb{A} n'est pas un corps, on peut adapter la définition des espaces vectoriels, en définissant la notion de *module* sur un anneau $\mathbb{A} : \mathbb{A}[X]$ est alors muni d'une structure de module sur \mathbb{A} .

Note Historique 19.1.9

Les premiers polynômes apparaissant dans l'histoire des mathématiques sont des polynômes de petits degrés associés à des équations traduisant des problèmes concrets : ainsi trouve-t-on dès l'époque babylonienne des résolutions d'équations polynomiales de degré 2. Ces méthodes sont systématisées par Al Khwarizmi à la fin du premier millénaire. Ainsi, les polynômes sont d'abord introduits de façon fonctionnelle. La notation par exposant pour les puissances apparaît plus tard ; elle est introduite par Nicolas Chuquet au 15^e siècle. Auparavant, on répétait l'inconnue autant de fois que le degré, ce qui était une limitation à une étude générale. Cependant, dès le 14^e siècle, le point de vue formel apparaît dans les travaux de Ibn al-Banna, qui présente les polynômes sous la forme de suites de coefficients. C'est exactement l'approche que nous venons d'en faire.

I.3 Indéterminée formelle

Par commodité, on adopte une notation plus proche de la notation fonctionnelle qu'on connaît pour les fonctions polynomiales ; cette notation est plus facile à manipuler que la définition formelle par les suites. De ce fait, à partir du moment où nous aurons défini l'indéterminée formelle X (remplaçant la notion de variable pour les fonctions polynomiales), nous n'utiliserons plus la définition formelle des polynômes par les suites.

On rappelle que par définition, l'anneau \mathbb{A} considéré contient un élément neutre pour le produit, noté 1.

Définition 19.1.10 (Indéterminée formelle)

On définit dans $\mathbb{A}[X]$ l'indéterminée formelle X comme étant le polynôme $X = (0, 1, 0, 0, \dots)$.

Ainsi, X n'est pas une variable (au sens fonctionnel), mais un polynôme bien précis, auquel on donne un nom particulier, et auquel on attribue une notation bien particulière, dont le but est l'analogie avec les fonctions polynomiales.

Avertissement 19.1.11

- En particulier, l'indéterminée formelle X n'étant pas une variable, elle ne doit pas être quantifiée, et ne peut pas être utilisée pour résoudre des équations.
- Un polynôme n'est pas une fonction de l'indéterminée formelle, donc la notation $P(X)$ en lieu et place de P n'est pas de mise. On l'utilise néanmoins dans certaines situations, notamment lorsque plusieurs indéterminées sont en jeu. Cette notation peut être justifiée rigoureusement par la notion de spécialisation qu'on verra un peu plus loin.

Proposition 19.1.12 (Monômes)

Pour tout $n \in \mathbb{N}$, on a $X^n = (\underbrace{0, \dots, 0}_n, 1, 0, \dots)$, le 1 étant donc à l'indice n .

Corollaire 19.1.13 (Expression d'un polynôme à l'aide de l'indéterminée formelle)

Soit $P = (a_n)_{n \in \mathbb{N}}$ un polynôme de $\mathbb{A}[X]$. Alors $P = \sum_{k=0}^{+\infty} a_k X^k$, cette somme ayant un sens puisqu'elle est en fait finie, les a_k étant nuls à partir d'un certain rang.

Encore une fois, il faut bien comprendre ce que signifie cette égalité : il s'agit bien d'une somme de polynômes, et non d'éléments de \mathbb{A} (signification de l'indéterminée).

De la définition même, il vient :

Proposition 19.1.14 (principe d'identification)

Soit P et Q deux polynômes de $\mathbb{A}[X]$. Notons $P = \sum_{k=0}^{+\infty} a_k X^k$ et $Q = \sum_{k=0}^{+\infty} b_k X^k$, en étant bien conscient que ces sommes sont en fait finies. Alors $P = Q$ si et seulement si pour tout $k \in \mathbb{N}$, $a_k = b_k$.

Les règles de calcul sur les polynômes de $\mathbb{A}[X]$ résultent alors des règles usuelles de calcul dans un anneau découlant des associativités, des commutativités, et des distributivités. La règle élémentaire sur laquelle reposent ces calculs est la suivante :

Lemme 19.1.15 (Produits de monômes)

Soit $(i, j) \in \mathbb{N}^2$. Alors $X^i X^j = X^{i+j}$.

Ce résultat est une évidence, et découle simplement de l'associativité du produit dans l'anneau $\mathbb{A}[X]$.

Exemples 19.1.16

1. Calcul de $(2 + 3X + 2X^2)(3X + X^2 + 2X^3)$ dans $\mathbb{F}_5[X]$.
2. Calcul de $(X^2 + X + 2)^7$ dans $\mathbb{F}_7[X]$.

Définition 19.1.17 (Monômes)

Un monôme est un polynôme de la forme aX^n , $a \neq 0$. L'entier n est appelé *degré* du monôme.

Ainsi, tout polynôme est une somme de monômes de degrés deux à deux distincts.

Les polynômes formels peuvent aussi se composer. En effet, étant donné un polynôme P , on peut considérer, pour tout $n \in \mathbb{N}$, le polynôme P^n (en a en particulier $P^0 = 1$ et $P^1 = P$). On définit alors la composée de deux polynômes de la sorte :

Définition 19.1.18 (Composition de polynômes formels)

Soit P et Q deux polynômes de $\mathbb{A}[X]$. On note $Q = \sum_{k=0}^d a_k X^k$. On définit alors le polynôme $Q \circ P$ par :

$$Q \circ P = \sum_{k=0}^d a_k P^k.$$

I.4 Dérivation

On sait facilement dériver (au sens analytique) une fonction polynomiale à coefficients réels, $x \mapsto x^n$ se dérivant en $x \mapsto nx^{n-1}$. Cette règle de dérivation peut être vue de façon purement formelle, permettant de généraliser la dérivation des polynômes à un anneau quelconque (dans lequel on ne dispose pas des techniques d'analyse, spécifiques à \mathbb{R})

Définition 19.1.19 (Dérivée formelle d'un polynôme)

Soit $P = \sum_{k=0}^d a_k X^k$ un polynôme à coefficients dans un anneau commutatif \mathbb{A} . Le *polynôme dérivé* est défini par :

$$P' = \sum_{k=1}^d k a_k X^{k-1}.$$

Proposition 19.1.20 (Linéarité de la dérivation)

Soit P, Q deux polynômes de $\mathbb{A}[X]$, et $a \in \mathbb{A}$.

1. $(P + Q)' = P' + Q'$.
2. $(aP)' = aP'$.

La linéarité s'exprime en terme de structures en affirmant que la dérivation est une application linéaire (c'est-à-dire un homomorphisme d'espaces vectoriels) lorsque \mathbb{A} est un corps, ou un homomorphisme de \mathbb{A} -modules sinon.

Vu que la définition de la dérivation est calquée sur la dérivée analytique des fonctions polynomiales réelles, on a, sans surprise, des règles de dérivation similaires, et notamment :

Proposition 19.1.21 (Dérivée de produits)

1. Soit P et Q deux polynômes à coefficients dans \mathbb{A} . Alors

$$(PQ)' = P'Q + PQ'.$$

2. Soit P_1, \dots, P_n des polynômes à coefficients dans \mathbb{A} . Alors

$$(P_1 \cdots P_n)' = \sum_{i=1}^n P_1 \cdots P_{i-1} P_i' P_{i+1} \cdots P_n.$$

3. (Formule de Leibniz) Soit P et Q deux polynômes à coefficients dans \mathbb{A} . Alors

$$P^{(n)} = \sum_{k=0}^n \binom{n}{k} P^{(k)} Q^{(n-k)}.$$

Corollaire 19.1.22 (Dérivée de P^n)

En particulier, étant donné $n \in \mathbb{N}^*$, $(P^n)' = nP'P^{n-1}$.

Avec un petit abus de notation, on pourrait considérer cette égalité également au rang $n = 0$ (le terme P^{-1} n'est pas bien défini, mais le produit par n annule l'ensemble). On généralise le corollaire précédent de la sorte

Proposition 19.1.23 (Dérivée d'une composition)

Soit P et Q dans $\mathbb{A}[X]$. Alors

$$(Q \circ P)' = P' \times (Q' \circ P).$$

Exemples 19.1.24

1. Que peut-on dire de $(PQ)^{(p)}$ dans $\mathbb{F}_p[X]$, p étant premier ?
2. La formule de Faa di Bruno est-elle valable dans $\mathbb{A}[X]$?

I.5 Degré et valuation**Définition 19.1.25 (Degré et valuation)**

Soit $P = (a_n)_{n \in \mathbb{N}}$ un polynôme à coefficients dans un anneau commutatif \mathbb{A} .

1. Le *degré de P* est $\deg(P) = \max\{n \in \mathbb{N} \mid a_n \neq 0\}$.
Si P est non nul, cet ensemble est non vide, et majoré. Ainsi, $\deg(P) \in \mathbb{N}$.
Si $P = 0$, par convention, $\deg(P) = -\infty$.
2. La *valuation de P* est $\text{val}(P) = \min\{n \in \mathbb{N} \mid a_n \neq 0\}$.
Si P est non nul, cet ensemble est non vide, et minoré. Ainsi, $\text{val}(P) \in \mathbb{N}$.
Si $P = 0$, par convention, $\text{val}(P) = +\infty$.

On trouve aussi parfois la notation $\omega(P)$ pour la valuation.

On utilise souvent la filtration suivante de $\mathbb{A}[X]$ (une filtration de E est une chaîne d'inclusions d'union totale E)

Notation 19.1.26 (Filtration par les degrés)

Soit \mathbb{A} un anneau et $n \in \mathbb{N}$. On note $\mathbb{A}_n[X]$ l'ensemble des polynômes de degré au plus n .

Proposition 19.1.27

On a évidemment $\mathbb{A}_0[X] \subset \mathbb{A}_1[X] \subset \dots \subset \mathbb{A}_n[X] \subset \dots$ et

$$\mathbb{A}[X] = \bigcup_{n=0}^{+\infty} \mathbb{A}_n[X].$$

Définition 19.1.28 (Monôme dominant, coefficient dominant, polynôme unitaire)

Soit $P = \sum_{k=0}^d a_k X^k$ un polynôme de $\mathbb{A}[X]$, de degré d .

1. Le monôme dominant de P est le monôme $a_d X^d$, donc le monôme de plus haut degré de P .
2. Le coefficient dominant de P est l'élément a_d de \mathbb{A} , donc le coefficient du monôme dominant.
3. Le polynôme P est dit unitaire si son coefficient dominant vérifie $a_d = 1_{\mathbb{A}}$.

Proposition 19.1.29 (Degré d'une somme, d'un produit, d'une dérivée)

Soit P et Q deux polynômes de $\mathbb{A}[X]$, et $\lambda \in \mathbb{A}$. Alors :

1. $\deg(P + Q) \leq \max(\deg(P), \deg(Q))$.
Si $\deg(P) \neq \deg(Q)$, alors $\deg(P + Q) = \max(\deg(P), \deg(Q))$.
2. Si \mathbb{A} est intègre, et si $\lambda \neq 0$, $\deg(\lambda P) = \deg(P)$.
3. Si \mathbb{A} est intègre (en particulier si \mathbb{A} est un corps) et si P et Q sont non nuls, $\deg(PQ) = \deg(P) + \deg(Q)$.
4. $\deg(P') \leq \deg(P) - 1$.

Exemples 19.1.30

1. Trouver dans $(\mathbb{Z}/6\mathbb{Z})[X]$ un exemple contredisant le point 2.
2. Trouver dans $(\mathbb{Z}/6\mathbb{Z})[X]$ un exemple contredisant le point 3.

Corollaire 19.1.31 (Théorème de permanence de l'intégrité)

Si \mathbb{A} est intègre, alors $\mathbb{A}[X]$ est intègre.

Corollaire 19.1.32 (Propriétés de stabilité)

1. $\mathbb{A}_n[X]$ est un sous-groupe additif de $\mathbb{A}[X]$.
2. La dérivation $D : \mathbb{A}[X] \rightarrow \mathbb{A}[X]$ induit un homomorphisme de groupes $D_n : \mathbb{A}_n[X] \rightarrow \mathbb{A}_{n-1}[X]$ (et même un homomorphisme de \mathbb{A} -modules dans le sens où D_n respecte aussi le produit externe).
3. Si \mathbb{K} est un corps de caractéristique nulle, $D_n : \mathbb{K}_n[X] \rightarrow \mathbb{K}_{n-1}[X]$ est une surjection. Autrement dit, tout polynôme de $\mathbb{K}_{n-1}[X]$ est primitivable formellement dans $\mathbb{K}_n[X]$.

Remarques 19.1.33

- $\mathbb{A}_n[X]$ est-il un sous-anneau de $\mathbb{A}[X]$?
- Trouver un polynôme de $\mathbb{F}_p[X]$ n'admettant pas de primitive.

Corollaire 19.1.34 (Intégrité des anneaux usuels de polynômes)

Si \mathbb{K} est un corps, $\mathbb{K}[X]$ est intègre. En particulier, les anneaux $\mathbb{R}[X]$, $\mathbb{C}[X]$, $\mathbb{F}_p[X]$, $\mathbb{Q}[X]$ sont intègres.

L'anneau \mathbb{Z} étant également intègre, on obtient aussi l'intégrité de $\mathbb{Z}[X]$.

On précise un peu dans certains cas la relation entre dérivée de P et dérivée de P' : dans les situations que vous connaissez, dériver un polynôme non constant baisse son degré de 1. Cependant, ceci n'est pas vrai en toute généralité. Nous avons besoin pour cela d'hypothèses plus fortes.

Proposition 19.1.35 (Degré d'une dérivée dans $\mathbb{K}[X]$, $\mathbb{K} = \mathbb{R}$ ou \mathbb{C})

Soit \mathbb{K} un corps de caractéristique nulle (par exemple \mathbb{Q} , \mathbb{R} ou \mathbb{C}), et P un polynôme non constant de $\mathbb{K}[X]$. Alors $\deg(P') = \deg(P) - 1$.

Remarque 19.1.36

1. Trouver dans $\mathbb{F}_p[X]$ un polynôme non constant pour lequel cette égalité est fautive.
2. Si \mathbb{K} est un corps de caractéristique p , quelle condition donner au degré de P pour avoir $\deg(P') = \deg(P) - 1$?

Corollaire 19.1.37

Soit \mathbb{K} un corps de caractéristique nulle, et soit P et Q deux polynômes de $\mathbb{K}[X]$. Si $P' = Q'$, alors P et Q diffèrent d'une constante additive.

Exemple 19.1.38

Donner un contre-exemple dans le cas où $\mathbb{K} = \mathbb{F}_p$.

On pourrait établir pour les valuations des règles similaires à celles qu'on a pour les degrés. La notion de valuation n'étant pas explicitement au programme, nous laissons le lecteur intéressé établir ces règles par lui-même.

II Arithmétique dans $\mathbb{K}[X]$

On considère ici des polynômes à coefficients dans un corps \mathbb{K} . Vous pouvez considérer $\mathbb{K} = \mathbb{R}$ ou \mathbb{C} , mais, sauf mention explicite du contraire, les résultats exposés sont valables dans le cadre plus général d'un corps quelconque.

II.1 Division euclidienne

L'anneau $\mathbb{K}[X]$ est euclidien (c'est-à-dire qu'il y existe une notion de division euclidienne) :

Théorème 19.2.1 (Théorème de la division euclidienne dans $\mathbb{K}[X]$)

Soit \mathbb{K} un corps. Pour tout polynôme A et $B \neq 0$ de $\mathbb{K}[X]$ il existe d'uniques polynômes Q et R tels que :

- (i) $A = BQ + R$
- (ii) $\deg(R) < \deg(B)$.

Les polynômes Q et R sont appelés respectivement quotient et reste de la division euclidienne de A par B .

Méthode 19.2.2 (Algorithme de la division euclidienne)

- On pose la division euclidienne comme la division des entiers, en disposant A à gauche et B à droite, les monômes étant écrits dans l'ordre décroissant des degrés (donc en marquant d'abord les monômes de plus haut degré).
- On trouve le monôme aX^k tel que $aX^k B$ ait même monôme dominant que A , puis on effectue la différence $A_1 = A - aX^k B$, qui a donc un degré strictement plus petit que A .
- On recommence sur A_1 , et on en déduit A_2
- On recommence ainsi jusqu'à obtenir A_k de degré strictement plus petit que B . Alors A_k est le reste recherché, et le quotient est la somme des monômes par lesquels on a multiplié B pour obtenir les A_i successifs.

Cet algorithme peut facilement être implémenté dans un langage informatique dans $\mathbb{R}[X]$ ou $\mathbb{C}[X]$; un polynôme est dans ce cas représenté par la liste de ses coefficients (on revient à la définition formelle des polynômes sous forme d'une suite finie).

Exemple 19.2.3

Division euclidienne de $X^6 + 3X^2 + 1$ par $X^2 + X + 1$.

On verra un peu plus loin une méthode basée sur l'étude des racines pour déterminer rapidement le reste d'une division euclidienne par un polynôme de petit degré dont on connaît les racines.

Remarque 19.2.4

L'algorithme de la division euclidienne peut-il être mené sans restriction dans $\mathbb{A}[X]$ lorsque \mathbb{A} est un anneau commutatif quelconque ? Donner une condition sur le polynôme B pour qu'on puisse effectuer dans $\mathbb{A}[X]$ la division euclidienne d'un polynôme A quelconque par B .

II.2 Idéaux de $\mathbb{K}[X]$

Comme pour l'arithmétique de \mathbb{Z} , il est commode de raisonner en terme d'idéaux. Le résultat rendant la situation totalement similaire à celle de \mathbb{Z} est le fait que tous les idéaux de $\mathbb{K}[X]$ sont principaux, donc engendrés par un unique polynôme.

Théorème 19.2.5 (Description des idéaux de $\mathbb{K}[X]$)

Soit \mathbb{K} un corps. Alors $\mathbb{K}[X]$ est un anneau principal. De plus, deux polynômes P et Q engendrent le même idéal si et seulement s'il existe un $\lambda \in \mathbb{K}^*$ tel que $Q = \lambda P$.

On notera (P) l'idéal engendré par un polynôme P .

Remarques 19.2.6

- $\{XP + YQ, P, Q \in \mathbb{R}[X, Y]\}$ est un idéal non principal de $\mathbb{R}[X, Y] = \mathbb{R}[X][Y] = \mathbb{R}[Y][X]$. Ainsi, le résultat précédent rentre en défaut lorsque \mathbb{K} n'est pas un corps (dans l'exemple, on considère les polynômes à coefficients dans l'anneau $\mathbb{R}[X]$).
-
- En fait, on peut montrer que $\mathbb{K}[X]$ est principal si et seulement si \mathbb{K} est un corps.
- Nous avons vu en exercice que tout anneau euclidien est principal. Ce n'est donc ici qu'une conséquence de ce résultat plus général.

II.3 Divisibilité

Soit \mathbb{K} un corps. Ainsi, $\mathbb{K}[X]$ est principal. On note (P) l'idéal engendré par un élément P de $\mathbb{K}[X]$, à savoir

$$(P) = \{PQ, Q \in \mathbb{K}[X]\}.$$

Définition 19.2.7 (Divisibilité dans $\mathbb{K}[X]$)

Soit A et B deux polynômes de $\mathbb{K}[X]$. On dit que B divise A s'il existe un polynôme Q tel que $A = BQ$. Inversement, on dit que A est un multiple de B .

Ainsi, B divise A si et seulement si le reste de la division euclidienne de A par B est nul. Comme dans \mathbb{Z} , on a la caractérisation suivante :

Proposition 19.2.8 (Caractérisation en termes d'idéaux)

Soit A et B deux polynômes de $\mathbb{K}[X]$. Alors A divise B si et seulement si $B \in (A)$, ou encore si et seulement si $(B) \subset (A)$.

Comme dans le cadre général, on dit que le couple (A, B) est un couple de polynômes associés si A divise B et B divise A . Il vient alors de la description des idéaux de $\mathbb{A}[X]$ que :

Proposition 19.2.9 (Polynômes associés)

Soit $(A, B) \in \mathbb{K}[X]^2$. Alors (A, B) est un couple de polynômes associés si et seulement s'il existe $\lambda \in \mathbb{K}^*$ tel que $A = \lambda B$.

II.4 PGCD et PPCM

Proposition/Définition 19.2.10 (PGCD de deux polynômes)

Soit \mathbb{K} un corps. Soit A et B deux polynômes de $\mathbb{K}[X]$, dont l'un au moins est non nul et $P \in \mathbb{K}[X]$. Les propositions suivantes sont équivalents :

- P divise A et B et est de degré maximal pour cette propriété.
- P divise A et B et tout autre diviseur de A et B est aussi un diviseur de P
- $(P) = (A) + (B)$.

Si ces propriétés sont vérifiées on dit que P est un PGCD de A et B .

Il n'y a pas unicité d'un PGCD de A et B . Plus précisément :

Proposition 19.2.11 (Description des PGCD)

Soit A et B deux polynômes de $\mathbb{K}[X]$, dont l'un au moins est non nul, et P un PGCD de A et B . Alors un polynôme Q est un PGCD de A et B si et seulement s'il existe $\lambda \in \mathbb{K}^*$ tel que $Q = \lambda P$.

Notation 19.2.12 ($A \wedge B$)

En particulier, si A et B sont deux polynômes de $\mathbb{K}[X]$ dont l'un au moins est non nul, il existe un unique PGCD unitaire de A et B . Ce PGCD unitaire est noté $A \wedge B$.

Comme dans le cas de \mathbb{Z} , on déduit du troisième point équivalent de la définition l'existence de relations de Bézout.

Proposition 19.2.13 (Relation de Bézout)

Soit A et B deux polynômes de $\mathbb{K}[X]$ dont l'un au moins est non nul.

1. Il existe des polynômes U et V tels que $AU + BV = A \wedge B$
2. Soit $P \in \mathbb{K}[X]$ tel qu'il existe U et V dans $\mathbb{K}[X]$ tels que $AU + BV = P$. Alors P est un multiple de $A \wedge B$.

Comme dans \mathbb{Z} , on peut déterminer un PGCD et une relation de Bézout par l'algorithme d'Euclide étendu, en utilisant le lemme suivant :

Lemme 19.2.14

Soit A et B deux polynômes tels que $B \neq 0$. Soit Q et R le quotient et le reste de la division de A par B . Alors $A \wedge B = B \wedge R$

Méthode 19.2.15 (Calcul d'un PGCD et d'une relation de Bézout)

La méthode est la même que dans \mathbb{Z} , par divisions euclidiennes successives, jusqu'à obtenir un reste nul. Le dernier reste non nul est le PGCD, et en combinant les relations de division obtenues, on trouve de la même façon que dans \mathbb{Z} une relation de Bézout (quitte à diviser par un scalaire, pour obtenir le PGCD unitaire)

Exemple 19.2.16

Trouver les PGCD de $X^8 - 1$ et $X^{12} - 1$, et une relation de Bézout.

Comme pour le cas de \mathbb{Z} , la définition du PGCD s'étend au cas du PGCD de n polynômes. On obtient alors :

Proposition 19.2.17 (Propriétés du PGCD)

L'opération \wedge est commutative et associative. Par ailleurs, si C est unitaire, $(A \wedge B)C = AC \wedge BC$.

Évidemment, on peut aussi définir les PPCM :

Proposition/Définition 19.2.18 (PPCM de deux polynômes)

Soit \mathbb{K} un corps. Soit A et B deux polynômes non nuls de $\mathbb{K}[X]$, et $P \in \mathbb{K}[X]$. Les propositions suivantes sont équivalents :

- (i) A et B divisent P et P est de degré minimal pour cette propriété.
- (ii) A et B divisent P et tout autre multiple de A et B est aussi un multiple de P
- (iii) $(P) = (A) \cap (B)$.

Si ces propriétés sont vérifiées on dit que P est un PPCM de A et B .

Proposition 19.2.19 (Description des PPCM)

Soit A et B deux polynômes non nuls de $\mathbb{K}[X]$, et P un PPCM de A et B . Alors un polynôme Q est un PPCM de A et B si et seulement s'il existe $\lambda \in \mathbb{K}^*$ tel que $Q = \lambda P$.

Notation 19.2.20 ($A \vee B$)

En particulier, si A et B sont deux polynômes non nuls de $\mathbb{K}[X]$, il existe un unique PPCM unitaire de A et B . Ce PPCM unitaire est noté $A \vee B$.

Exemple 19.2.21

$P = (X + 1)^2$ et $Q = (X + 1)(X - 1)$.

II.5 Polynômes premiers entre eux

Définition 19.2.22

Soit A et B deux polynômes de $\mathbb{K}[X]$. On dit que A et B sont premiers entre eux si $A \wedge B = 1$.

Autrement dit, les seuls diviseurs communs à A et B sont les polynômes constants non nuls.

Plus généralement, on définit comme dans \mathbb{Z} la notion de famille finie de polynômes deux à deux premiers entre eux, ou premiers entre eux dans leur ensemble.

Ici encore, les propriétés valables dans \mathbb{Z} se généralisent :

Théorème 19.2.23 (Théorème de Bézout)

Soit A et B deux polynômes de $\mathbb{K}[X]$. Alors A et B sont premiers entre eux si et seulement s'il existe deux polynômes U et V tels que $AU + BV = 1$.

Exemple 19.2.24

Soit $\lambda \neq \mu$ dans \mathbb{K} . Alors les polynômes $X - \lambda$ et $X - \mu$ sont premiers entre eux.

Lemme 19.2.25 (Lemme de Gauss)

Soit A , B et C deux polynômes de $\mathbb{K}[X]$ tels que A divise BC et A et B soient premiers entre eux. Alors A divise C .

Corollaire 19.2.26

Soit A , B et C trois polynômes tels que A et B divisent C et A et B soient premiers entre eux. Alors AB divise C .

Comme dans \mathbb{Z} , on a une relation simple entre PPCM et PGCD, à ceci près que comme ces notions sont définies à constante multiplicative près, il faut faire attention au coefficient dominant :

Proposition 19.2.27 (relation entre PGCD et PPCM)

Soit A et B deux polynômes de coefficients dominants a et b respectivement. Alors

$$ab(A \wedge B)(A \vee B) = AB.$$

II.6 Décomposition en facteurs irréductibles**Définition 19.2.28 (Polynôme irréductible)**

Un polynôme non constant P de $\mathbb{K}[X]$ est irréductible si et seulement s'il n'est, à une constante multiplicative non nulle près, divisible que par lui-même et par 1.

Exemples 19.2.29

1. Les polynômes $X - \lambda$ sont irréductibles ($\lambda \in \mathbb{K}$)
2. Dans $\mathbb{R}[X]$, tout polynôme $aX^2 + bX + c$ tel que $\Delta < 0$ est irréductible.
3. Ces polynômes ne sont pas irréductibles dans $\mathbb{C}[X]$.

Lemme 19.2.30

Soit P un polynôme irréductible de $\mathbb{K}[X]$ et A un polynôme, non multiple de P . Alors A et P sont premiers entre eux.

Le lemme de Gauss fournit facilement la généralisation suivante du lemme d'Euclide :

Lemme 19.2.31 (Euclide)

Soit A et B deux polynômes de $\mathbb{K}[X]$ et P un polynôme irréductible. Alors si P divise AB , P divise A ou P divise B .

De façon équivalente, la contraposée fournit :

Corollaire 19.2.32

Soit A et B deux polynômes de $\mathbb{K}[X]$ et P un polynôme irréductible. Alors, si P ne divise ni A ni B , P ne divise pas AB .

Enfin, voici l'analogie du théorème de la décomposition primaire :

Théorème 19.2.33 (Décomposition en facteurs irréductibles)

Soit P un polynôme non nul de $\mathbb{K}[X]$.

1. Il existe un élément $\lambda \in \mathbb{K}^*$ et des polynômes irréductibles P_1, \dots, P_k tels que

$$P = \lambda P_1 \cdots P_k.$$

2. Cette décomposition est unique, à l'ordre près des facteurs, et à multiplication près de chaque facteur (y compris λ) par un élément non nul de \mathbb{K} .
3. En particulier, si on impose que les P_i soient unitaires, cette décomposition est unique, à l'ordre près des facteurs.

Nous verrons un peu plus loin la description complète des polynômes irréductibles de $\mathbb{R}[X]$ et de $\mathbb{C}[X]$. Pour cela, il nous faut étudier d'un peu plus près les propriétés liées aux racines d'un polynôme.

III Racines d'un polynôme

Pour pouvoir définir la notion de racine d'un polynôme, il faut d'abord pouvoir « appliquer » un polynôme à un élément de \mathbb{A} , donc transformer un polynôme formel en une fonction polynomiale.

III.1 Spécialisation, évaluation

Le lien entre les polynômes formels de \mathbb{R} et les fonctions polynomiales sur \mathbb{R} est assez clair : étant donné un polynôme formel $P = \sum_{k=0}^d a_k X^k$ de $\mathbb{R}[X]$, on peut lui associer la fonction polynomiale

$$\tilde{P} : x \mapsto \sum_{k=0}^d a_k x^k.$$

La seule condition pour pouvoir faire cela de façon plus générale dans $\mathbb{A}[X]$ est de pouvoir faire dans \mathbb{A} des produits (donc calculer des puissances) et des sommes. Comme \mathbb{A} est un anneau commutatif, cela ne pose pas de problème particulier, et on peut donc définir :

Définition 19.3.1 (Fonction polynomiale associée à un polynôme)

1. Soit $P \in \mathbb{A}[X]$, donné par $P = \sum_{k=0}^d a_k X^k$. La fonction polynomiale $\tilde{P} : \mathbb{A} \rightarrow \mathbb{A}$ associée à P est la fonction définie par :

$$\forall b \in \mathbb{A}, \tilde{P}(b) = \sum_{k=0}^d a_k b^k.$$

On rappelle que par convention, $b^0 = 1_{\mathbb{A}}$.

2. L'ensemble des fonctions polynomiales sur \mathbb{A} est l'ensemble :

$$\mathbb{A}[x] = \{\tilde{P} \mid P \in \mathbb{A}[X]\}.$$

Définition 19.3.2 (Évaluation d'un polynôme)

Soit P un polynôme de $\mathbb{A}[X]$. L'évaluation de P en $b \in \mathbb{A}$ est l'élément de \mathbb{A} défini par $\tilde{P}(b)$. Pour simplifier les notations, on désigne souvent cette évaluation plus simplement par $P(b)$.

Proposition 19.3.3 (Respect des structures)

Soit \mathbb{A} un anneau commutatif. L'application $\varphi : \mathbb{A}[X] \rightarrow \mathbb{A}[x]$ définie par $\varphi(P) = \tilde{P}$ est un homomorphisme d'anneaux.

Intuitivement, il apparaît clair que lorsque $\mathbb{K} = \mathbb{R}$ ou \mathbb{C} , on peut identifier les polynômes formels à coefficients dans \mathbb{K} et les fonctions polynomiales sur \mathbb{K} . C'est ce que nous exprimons dans le théorème suivant :

Théorème 19.3.4 ($\mathbb{K}[X] \simeq \mathbb{K}[x]$ pour $\mathbb{K} = \mathbb{R}$ ou \mathbb{C})

Soit $\mathbb{K} = \mathbb{R}$ ou \mathbb{C} . Soit $\varphi : \mathbb{K}[X] \rightarrow \mathbb{K}[x]$ définie par $\varphi(P) = \tilde{P}$. Alors φ est un isomorphisme d'anneaux.

Remarques 19.3.5

1. En considérant le petit théorème de Fermat, montrer que cette propriété n'est pas vraie pour tous les corps \mathbb{K} .
2. On montrera un peu plus loin qu'une condition suffisante pour que cette identification soit vraie est que \mathbb{K} soit un corps infini. C'est le cas en particulier lorsque \mathbb{K} est de caractéristique nulle.

On retiendra donc l'avertissement suivant :

Avertissement 19.3.6

Si \mathbb{K} n'est pas un corps infini (par exemple $\mathbb{K} = \mathbb{F}_p$), deux polynômes distincts P et Q de $\mathbb{K}[X]$ peuvent correspondre à la même application polynomiale. Ainsi, les polynômes sont davantage différenciés dans $\mathbb{K}[X]$ que dans $\mathbb{K}[x]$. On n'a donc pas possibilité en général d'identifier polynômes formels et fonctions polynomiales.

Enfin, dans le cas spécifique de \mathbb{R} (seul cas dans lequel on peut considérer la dérivée au sens analytique), on a également, du fait même des définitions :

Proposition 19.3.7

Pour tout polynôme P de $\mathbb{R}[X]$, $\widetilde{P}' = \widetilde{P}'$.

Ainsi, les opérations définies formellement coïncident avec les opérations sur les fonctions polynomiales, y compris la dérivation dans le cas de \mathbb{R} .

Il est important de constater que le cadre formel qu'on s'est donné pour définir les polynômes permet d'« appliquer » un polynôme à des éléments qui sortent du cadre initialement fixé. Pour prendre un exemple, étant donné un polynôme $P = \sum_{k=0}^d a_k X^k$ de $\mathbb{R}[X]$ et M une matrice carrée à coefficients réels, on peut considérer le polynôme de matrices

$$P(M) = \sum_{k=0}^d a_k M^k,$$

où il faut bien prendre garde au fait que M^0 désigne la matrice identité I_n .

Exemple 19.3.8

Soit $P = 2 + 3X + 3X^2$, et $M = \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}$. Calculer $P(M)$.

On peut formaliser ce type de construction, mais nous resterons assez vague, car nous débordons du programme. Le bon cadre à se fixer est celui donné par la structure de \mathbb{A} -algèbre : un ensemble \mathbb{B} est une \mathbb{A} -algèbre si et seulement si \mathbb{B} est un anneau et est muni d'une loi de composition externe d'ensemble d'opérateurs \mathbb{A} , avec un certain nombre de propriétés (associativité externe, distributivité de la loi externe sur chacune des lois additives de \mathbb{A} et de \mathbb{B} , respect du neutre multiplicatif de \mathbb{A}). En gros, ce qu'il faut en retenir, c'est que dans une \mathbb{A} -algèbre \mathbb{B} , on peut faire, avec des règles de calcul raisonnablement semblables aux situations usuelles, la somme et le produit d'éléments de \mathbb{B} ainsi que le produit d'un élément de \mathbb{A} par un élément de \mathbb{B} . En particulier, étant donné un polynôme $P = \sum_{k=0}^d a_k X^k$ de $\mathbb{A}[X]$ et $b \in \mathbb{B}$, l'expression suivante a un sens :

$$P(b) = \sum_{k=0}^d a_k b^k.$$

Il convient de bien noter que par convention $b^0 = 1_{\mathbb{B}}$.

On parle de *spécialisation* du polynôme P en $b \in \mathbb{B}$.

Exemples 19.3.9

1. L'ensemble des matrices carrées de taille n , à coefficients réels, est une \mathbb{R} -algèbre : la situation décrite plus haut est un cas particulier de cette situation plus générale.

2. On utilisera beaucoup en algèbre linéaire des polynômes d'endomorphismes (applications linéaires d'un espace vectoriel dans lui-même), l'ensemble des endomorphismes d'un espace vectoriel E sur \mathbb{K} étant une \mathbb{K} -algèbre pour la somme usuelle et le produit défini par la composition. Ainsi, f^n désigne dans ce cas la composition itérée de f , et f^0 désigne la fonction identité id_E .

III.2 Racines et multiplicité

Soit \mathbb{A} un anneau commutatif

Définition 19.3.10 (Racine d'un polynôme)

Soit $P \in \mathbb{A}[X]$ et $a \in \mathbb{A}$. On dit que a est une racine de P si $P(a) = 0$.

Théorème 19.3.11 (Caractérisation des racines par la divisibilité)

Soit \mathbb{K} un corps, $P \in \mathbb{K}[X]$ et $r \in \mathbb{K}$. Alors r est racine de P si et seulement si $X - r$ divise P , donc s'il existe $Q \in \mathbb{K}[X]$ tel que $P = (X - r)Q$.

Remarque 19.3.12

Ce théorème reste valable dans $\mathbb{A}[X]$ pour un anneau commutatif quelconque. Pourquoi ?

Si après factorisation $P = (X - r)Q$, r est encore racine de Q , alors r est « plusieurs fois » racine de P . En comptant le nombre de fois qu'on peut mettre $X - r$ en facteur, on obtient la multiplicité de r :

Définition 19.3.13 (Multiplicité d'une racine)

Soit $P \in \mathbb{K}[X]$ et $r \in \mathbb{K}$. On dit que r est racine d'ordre de multiplicité $k \in \mathbb{N}^*$ si et seulement si $(X - r)^k$ divise P et $(X - r)^{k+1}$ ne divise pas P . Autrement dit, il existe $Q \in \mathbb{K}[X]$ tel que $P = (X - r)^k Q$, avec $Q(r) \neq 0$.

Par convention, on dira que r est racine de multiplicité 0 si r n'est pas racine de P . Une racine de multiplicité 1 est aussi appelée racine simple de P .

Cette mise en facteur maximale de $(X - r)^r$ peut être mise en valeur par la formule de Taylor pour les polynômes, nécessitant une hypothèse supplémentaire sur \mathbb{K} .

Théorème 19.3.14 (Formule de Taylor pour les polynômes)

Soit \mathbb{K} un corps de caractéristique nulle, P un polynôme de $\mathbb{K}[X]$, de degré d , et $a \in \mathbb{K}$. Alors,

$$P = \sum_{n=0}^d \frac{P^{(n)}(a)}{n!} (X - a)^n.$$

Ainsi, si v est le plus petit indice pour lequel le terme de la somme est non nul, on obtient

$$P = (X - a)^v \sum_{n=v}^d \frac{P^{(n)}(a)}{n!} (X - a)^{n-v}.$$

Remarques 19.3.15

Pourquoi supposer que \mathbb{K} est de caractéristique nulle ?

Ainsi, l'ordre de multiplicité de r correspond à la valuation de P après changement d'indéterminée $Y = X - r$.

On en déduit de façon immédiate :

Théorème 19.3.16 (Caractérisation de la multiplicité par les dérivées successives)

Soit \mathbb{K} un corps de caractéristique nulle, $P \in \mathbb{K}[X]$ et $a \in \mathbb{K}$. Le réel a est racine d'ordre de multiplicité k de P si et seulement si : $P(a) = P'(a) = \dots = P^{(k-1)}(a) = 0$ et $P^{(k)}(a) \neq 0$.

Ainsi, il faut toujours garder à l'esprit des deux facettes de la multiplicité des racines : la propriété de divisibilité, et la caractérisation par les dérivées.

Corollaire 19.3.17

Soit \mathbb{K} un corps de caractéristique nulle. Soit $P \in \mathbb{K}[X]$ et $r \in \mathbb{R}$. Si r est racine d'ordre k de P , alors r est racine d'ordre $k - 1$ de P' .

III.3 Majoration du nombre de racines

Le corollaire du lemme de Gauss amène :

Théorème 19.3.18

Soit \mathbb{K} un corps. Soit $P \in \mathbb{K}[X]$, et r_1, \dots, r_k des racines deux à deux distinctes de P , de multiplicités respectives $\alpha_1, \dots, \alpha_k$. Alors $(X - r_1)^{\alpha_1} \dots (X - r_k)^{\alpha_k}$ divise P , et r_1, \dots, r_k ne sont pas racines du quotient.

Corollaire 19.3.19 (Majoration du nombre de racines)

Soit $P \in \mathbb{K}[X]$ de degré n . Alors P admet au plus n racines (comptées avec multiplicité).

Exemple 19.3.20

Soit \mathbb{K} un corps dont \mathbb{F}_p est un sous-corps. Montrer que pour tout $x \in \mathbb{K}$, $x^p = x$ si et seulement si $x \in \mathbb{F}_p$.

On en déduit le résultat très important suivant, qu'on décline sous 4 formes équivalentes.

Corollaire 19.3.21 (Rigidité des polynômes)

1. Soit P un polynôme de $\mathbb{K}[X]$ degré au plus n . Alors, si P admet strictement plus de n racines, $P = 0$.
2. Soit P un polynôme de $\mathbb{K}[X]$ s'annulant en une infinité de points de \mathbb{K} . Alors P est le polynôme nul.
3. Si deux polynômes P et Q de $\mathbb{K}_n[X]$ coïncident en strictement plus de n valeurs distinctes. Alors $P = Q$.
4. Soit $n \in \mathbb{N}^*$. Étant donnés x_1, \dots, x_n des éléments deux à deux distincts de \mathbb{K} et y_1, \dots, y_n des éléments de \mathbb{K} non nécessairement distincts, il existe au plus un polynôme $P \in \mathbb{K}_{n-1}[X]$ tel que pour tout $i \in \llbracket 1, n \rrbracket$, $P(x_i) = y_i$. Ainsi, sous réserve d'existence, un polynôme de degré au plus $n - 1$ est entièrement déterminé par sa valeur en n points distincts.

On déduit notamment de cette propriété un résultat annoncé un peu plus haut :

Théorème 19.3.22 (Polynômes formels et fonctions polynomiales)

Soit \mathbb{K} un corps infini. Alors l'application de $\mathbb{K}[X]$ dans $\mathbb{K}[x]$ qui à un polynôme formel associe sa fonction polynomiale est un isomorphisme d'anneaux.

Par ailleurs, le dernier point du corollaire ci-dessus affirme l'unicité sous réserve d'existence d'un polynôme de degré au plus $n - 1$ prenant des valeurs données en n points fixés. Il n'est pas dur de construire explicitement un tel polynôme, fournissant ainsi l'existence. C'est le but du paragraphe suivant.

III.4 Interpolation de Lagrange

On recherche un polynôme de degré au plus n coïncidant en $n + 1$ points distincts avec une fonction f , ou, de façon équivalente, prenant en $n + 1$ points distincts x_0, \dots, x_n , $n + 1$ valeurs (non nécessairement distinctes) imposées y_0, \dots, y_n .

Pour ce faire, on commence par le cas où les valeurs imposées sont toutes nulles, sauf une égale à 1. On trouvera le cas général en formant une combinaison linéaire.

Définition 19.3.23

Polynômes de Lagrange Soit x_0, \dots, x_n des réels 2 à 2 distincts et $i \in \llbracket 1, n \rrbracket$. Le i^{e} polynôme de Lagrange associé à la famille x_0, \dots, x_n est

$$L_i = \frac{\prod_{\substack{j=0 \\ j \neq i}}^n (X - x_j)}{\prod_{\substack{j=0 \\ j \neq i}}^n (x_i - x_j)} = \prod_{\substack{j=0 \\ j \neq i}}^n \frac{X - x_j}{x_i - x_j}.$$

Lemme 19.3.24

Le polynôme L_i est l'unique polynôme de degré au plus $n + 1$ tel que pour tout $j \in \llbracket 0, n \rrbracket \setminus \{i\}$, $L_j(x_i) = 0$, et $L_i(x_i) = 1$.

Théorème 19.3.25 (Polynômes d'interpolation de Lagrange)

Soit \mathbb{K} un corps, $n \in \mathbb{N}^*$, x_1, \dots, x_n des éléments distincts de \mathbb{K} , et y_1, \dots, y_n des éléments de \mathbb{K} . Alors il existe un et un seul polynôme P de $\mathbb{K}_{n-1}[X]$ tel que pour tout $i \in \llbracket 1, n \rrbracket$, $P(x_i) = y_i$, et ce polynôme est donné explicitement par :

$$P = \sum_{i=1}^n y_i L_i.$$

En particulier, si f est une fonction définie sur un intervalle I contenant les x_i , le polynôme d'interpolation de Lagrange de f associé à la famille (x_i) est l'unique polynôme P_f coïncidant avec f sur les x_i , à savoir :

$$P_f = \sum_{i=0}^n f(x_i) L_i.$$

Ces polynômes, appelés polynômes d'interpolation de Lagrange, permettent en particulier d'approcher une fonction réelle f par une fonction polynomiale de degré au plus $n - 1$ coïncidant avec f en n points distincts. Cette approximation est cependant dans les faits souvent assez peu efficace. Il existe d'autres polynômes d'interpolation, comme les polynômes d'interpolation de Hermite, de degré strictement plus petit que $2n$, coïncidant avec f en n points, et de courbe tangente à celle de f en ces points.

Corollaire 19.3.26

Soit P le polynôme d'interpolation de Lagrange associée à la famille (x_i) , et aux valeurs (y_i) . Soit $P_0 = (X - x_1) \dots (X - x_n)$. L'ensemble E des polynômes Q (sans restriction de degré) tels que pour tout $i \in \llbracket 1, n \rrbracket$, $Q(x_i) = y_i$ est alors décrit par :

$$E = P + (P_0) = \{P + (X - x_1) \dots (X - x_n)R, \quad R \in \mathbb{K}[X]\}.$$

Remarque 19.3.27

Quelle est la structure algébrique de l'ensemble E du théorème précédent ?

III.5 Polynômes scindés**Définition 19.3.28 (Polynôme scindé)**

Soit \mathbb{K} un corps. On dit qu'un polynôme non nul P de $\mathbb{K}[X]$ est scindé s'il possède autant de racines (comptées avec multiplicité) que son degré, autrement dit si son nombre de racines est maximal.

Théorème 19.3.29 (Factorisation d'un polynôme scindé)

1. Un polynôme est scindé si et seulement s'il peut se factoriser de la façon suivante :

$$P = \lambda(X - x_1)(X - x_2) \cdots (X - x_n),$$

où λ est un scalaire non nul (égal au coefficient dominant de P), n est le degré de P , et x_1, \dots, x_n sont les racines, non nécessairement distinctes, de P .

2. Si on renomme y_1, \dots, y_k les racines 2 à 2 distinctes de P , de multiplicités respectives $\alpha_1, \dots, \alpha_k$, cette factorisation se réécrit :

$$P = \lambda(X - y_1)^{\alpha_1} \cdots (X - y_k)^{\alpha_k},$$

et on a $\alpha_1 + \cdots + \alpha_k = n$.

Ainsi, un polynôme est scindé si et seulement si sa décomposition en facteurs irréductibles ne fait intervenir que des polynômes irréductibles de degré 1.

Dans $\mathbb{R}[X]$, certaines techniques d'analyse peuvent aider à étudier cette propriété. Ainsi, le théorème de Rolle permet de montrer facilement que :

Proposition 19.3.30

Soit P un polynôme scindé de $\mathbb{R}[X]$, à racines simples. Alors P' est scindé, et ses racines séparent celles de P .

On verra en exercice une propriété plus générale :

Théorème 19.3.31

Soit P un polynôme scindé de $\mathbb{R}[X]$. Alors P' est scindé.

Une propriété importante des polynômes scindés est la possibilité de trouver facilement des relations entre les coefficients et les racines, par développement de la forme factorisée, et par identification des coefficients :

Théorème 19.3.32 (Relations coefficients/racines, ou relations de Viète)

Soit $P = \sum_{k=0}^n a_k X^k$ un polynôme de degré n , scindé, de racines (éventuellement non distinctes, apparaissant dans la liste autant de fois que sa multiplicité) r_1, \dots, r_n . Alors pour tout $k \in \llbracket 1, n \rrbracket$:

$$\sum_{1 \leq i_1 < \dots < i_k \leq n} r_{i_1} r_{i_2} \cdots r_{i_k} = (-1)^k \frac{a_{n-k}}{a_n}.$$

IV Polynômes irréductibles dans $\mathbb{C}[X]$ et $\mathbb{R}[X]$

Cette section étudie spécifiquement les polynômes à coefficients complexes ou réels.

IV.1 Factorisations dans $\mathbb{C}[X]$

Nous avons plus ou moins défini \mathbb{C} comme le corps de rupture du polynôme $X^2 + 1$, donc le plus petit corps contenant \mathbb{R} dans lequel ce polynôme admet une racine i . Un théorème essentiel, parfois appelé *théorème fondamental de l'algèbre* (c'est dire son importance) est le théorème suivant, que d'Alembert croyait avoir démontré, que Gauss a démontré par différentes méthodes :

Théorème 19.4.1 (d'Alembert-Gauss)

Tout polynôme non constant de $\mathbb{C}[X]$ admet au moins une racine.

Corollaire 19.4.2

Tout polynôme de $\mathbb{C}[X]$ est scindé, donc admet exactement autant de racines (comptées avec multiplicité) que son degré.

Corollaire 19.4.3

Dans $\mathbb{C}[X]$, les seuls polynômes irréductibles sont les polynômes de degré 1, c'est-à-dire les polynômes $aX + b$, $a \neq 0$.

Exemple 19.4.4

Quelles sont les racines et leurs multiplicités du polynôme $X^n - 1$? Factoriser ce polynôme en facteurs irréductibles dans $\mathbb{C}[X]$.

Tous les polynômes de $\mathbb{C}[X]$ se factorisant en polynômes non constants de degré minimal, on obtient alors une caractérisation simple de la divisibilité :

Théorème 19.4.5 (Caractérisation de la divisibilité dans $\mathbb{C}[X]$)

Soit P et Q deux polynômes de $\mathbb{C}[X]$. Alors P divise Q si et seulement si toute racine de P est aussi racine de Q , et que sa multiplicité dans Q est supérieure ou égale à sa multiplicité dans P .

IV.2 Facteurs irréductibles dans $\mathbb{R}[X]$

On commence par caractériser les polynômes à coefficients réels parmi les polynômes à coefficients dans \mathbb{C} .

Théorème 19.4.6 (Caractérisation des polynômes à coefficients réels)

Soit $P \in \mathbb{C}[X]$. Les propositions suivantes sont équivalentes :

- (i) P est à coefficients réels ;
- (ii) $P(\mathbb{R}) \subset \mathbb{R}$
- (iii) pour tout $z \in \mathbb{C}$, $P(\bar{z}) = \overline{P(z)}$.

Corollaire 19.4.7 (Racines complexes d'un polynôme réel)

Soit P un polynôme à coefficients réels, et r une racine de P dans \mathbb{C} . Si $r \notin \mathbb{R}$, alors \bar{r} est aussi racine de P , et elles ont même multiplicité.

Ainsi, les racines non réelles d'un polynôme à coefficients réels peuvent être groupées en paires de racines conjuguées de même multiplicité.

Le théorème de d'Alembert-Gauss amène alors :

Théorème 19.4.8 (Polynômes irréductibles de $\mathbb{R}[X]$)

1. Les polynômes irréductibles de $\mathbb{R}[X]$ sont les polynômes de degré 1 et les polynômes de degré 2 de discriminant strictement négatif.
2. Ainsi, tout polynôme $P \in \mathbb{R}[X]$ peut être factorisé en produit de polynômes de $\mathbb{R}[X]$ de degré 1, ou de degré 2, de discriminant strictement négatif.

Exemple 19.4.9

Factorisation dans $\mathbb{R}[X]$ de $X^n - 1$.

V Fractions rationnelles

La construction formelle que nous avons donnée des polynômes nous empêche *a priori* de former des quotients de polynômes (donc des fractions rationnelles), comme nous pouvons le faire pour les fonctions polynomiales. En effet, si \mathbb{K} est un corps, les seuls polynômes inversibles sont les polynômes constants non nuls.

Remarque 19.5.1

1. Quels sont les polynômes inversibles de $\mathbb{A}[X]$ lorsque \mathbb{A} est intègre ?
2. Trouver un polynôme inversible et non constant de $(\mathbb{Z}/4\mathbb{Z})[X]$

Une construction similaire à celle permettant de définir \mathbb{Q} à partir de \mathbb{Z} nous permet cependant de définir formellement des quotients de polynômes.

V.1 Définition des fractions rationnelles formelles

Soit, dans tout ce qui suit, \mathbb{K} un corps. On définit sur $\mathbb{K}[X] \times \mathbb{K}[X]^*$ la relation suivante :

$$(P, Q) \sim (R, S) \iff PS = QR,$$

l'égalité étant donnée dans $\mathbb{K}[X]$.

Proposition 19.5.2

La relation ci-dessus est une relation d'équivalence sur $\mathbb{K}[X] \times \mathbb{K}[X]^*$.

Définition 19.5.3 (Fraction rationnelle)

Une fraction rationnelle est une classe d'équivalence de la relation ci-dessus. La classe d'équivalence de (P, Q) sera notée $\frac{P}{Q}$. L'ensemble des fractions rationnelles sur le corps \mathbb{K} est noté $\mathbb{K}(X)$.

Ainsi, la relation $PS = QR$ amène assez logiquement l'égalité des fractions rationnelles $\frac{P}{Q} = \frac{R}{S}$.

Pour définir les lois de composition sur $\mathbb{K}(X)$, on commence par les définir sur $\mathbb{K}[X] \times \mathbb{K}[X]^*$. On définit, pour tout (P_1, Q_1) et (P_2, Q_2) de $\mathbb{K}[X] \times \mathbb{K}[X]^*$:

$$(P_1, Q_1) \times (P_2, Q_2) = (P_1 P_2, Q_1 Q_2) \quad \text{et} \quad (P_1, Q_1) + (P_2, Q_2) = (P_1 Q_2 + P_2 Q_1, Q_1 Q_2).$$

Lemme 19.5.4

1. Les opérations \times et $+$ sont associatives et commutatives
2. La relation \sim est une congruence sur les monoïdes $(\mathbb{K}[X] \times \mathbb{K}[X]^*, \times)$ et $(\mathbb{K}[X] \times \mathbb{K}[X]^*, +)$.

Théorème 19.5.5 (Structure de $\mathbb{K}(X)$)

Les lois $+$ et \times induisent des lois de composition, également notées $+$ et \times , sur $\mathbb{K}(X)$. L'ensemble $\mathbb{K}(X)$ muni de ces deux lois est un corps.

Les lois de composition ainsi définies se réécrivent sans surprise :

$$\frac{P_1}{Q_1} \times \frac{P_2}{Q_2} = \frac{P_1 P_2}{Q_1 Q_2}, \quad \frac{P_1}{Q_1} + \frac{P_2}{Q_2} = \frac{P_1 Q_2 + P_2 Q_1}{Q_1 Q_2} \quad \text{et} \quad \frac{P_1}{Q} + \frac{P_2}{Q} = \frac{P_1 + P_2}{Q}.$$

Définition 19.5.6 (Inclusion canonique de $\mathbb{K}[X]$ dans $\mathbb{K}(X)$)

L'application $P \mapsto \frac{P}{1}$ de $\mathbb{K}[X]$ dans $\mathbb{K}(X)$ est un homomorphisme injectif d'anneaux. La fraction $\frac{P}{1}$ seront désormais identifiée au polynôme P de $\mathbb{K}[X]$.

En particulier, si $P = AB$, alors $B = \frac{P}{A}$.

Proposition 19.5.7 (Simplification)

Soit P et Q deux polynômes et D un diviseur commun à P et Q . Alors $\frac{P}{Q} = \frac{P/D}{Q/D}$.

Proposition/Définition 19.5.8

Soit $F \in \mathbb{K}(X)$ une fraction rationnelle. Il existe un représentant (P, Q) , unique à multiplication près par un scalaire non nul, tel que $P \wedge Q = 1$ et $F = \frac{P}{Q}$. On dit que $\frac{P}{Q}$ est la forme irréductible de la fraction rationnelle F .

V.2 Degré, racines, pôles

Proposition/Définition 19.5.9 (Degré d'une fraction rationnelle)

Soit $F \in \mathbb{K}(X)$. La quantité $\deg(P) - \deg(Q)$ ne dépend pas de la représentation $\frac{P}{Q}$ choisie de la fraction F . On appelle degré de F et on note $\deg(F)$ la quantité

$$\deg(F) = \deg(P) - \deg(Q) \text{ où } F = \frac{P}{Q}$$

Il s'agit d'un entier relatif, ou de $-\infty$ si $P = 0$.

Les degrés des fractions rationnelles vérifient des propriétés semblables aux degrés des polynômes :

Proposition 19.5.10 (Propriétés des degrés)

Soit F, G des éléments de $\mathbb{K}(X)$.

- $\deg(F + G) \leq \max(\deg(F), \deg(G))$, avec égalité si $\deg(F) \neq \deg(G)$.
- $\deg(FG) = \deg(F) + \deg(G)$
- $\deg(F^{-1}) = -\deg(F)$

Proposition 19.5.11 (Partie entière)

Soit F une fraction rationnelle de $\mathbb{K}(X)$. Il existe un unique polynôme P de $\mathbb{K}[X]$ et une fraction rationnelle G de $\mathbb{K}(X)$ tels que

$$F = P + G \quad \text{et} \quad \deg(G) < 0.$$

Le polynôme P est appelée partie entière de la fraction rationnelle F .

Définition 19.5.12 (Racine, pôle, multiplicité)

Soit F une fraction rationnelle de $\mathbb{K}(X)$, écrit sous forme irréductible $F = \frac{P}{Q}$.

1. Une racine de F est une racine de P , sa multiplicité est sa multiplicité en tant que racine de P .
2. Un pôle de F est une racine de Q , sa multiplicité est sa multiplicité en tant que racine de Q .

Remarque 19.5.13

Puisque $\frac{P}{Q}$ est irréductible, r ne peut pas être à la fois racine de P et racine de Q .

Exemple 19.5.14

Racines, pôles et leurs multiplicités, de $\frac{(X-2)^2}{X^3(X-1)^4}$?

Définition 19.5.15 (Fonction rationnelle associée)

Soit $F = \frac{P}{Q}$ une fraction rationnelle formelle sous forme irréductible, et \mathcal{P} l'ensemble de ses pôles. La fonction rationnelle associée est $\tilde{F} : \mathbb{K} \setminus \mathcal{P} \rightarrow \mathbb{K}$ définie par

$$\forall x \in \mathbb{K} \setminus \mathcal{P}, \quad \tilde{F}(x) = \frac{P(x)}{Q(x)}.$$

V.3 Décomposition en éléments simples dans $\mathbb{C}(X)$

Toutes les fractions rationnelles de $\mathbb{C}(X)$ peuvent se décomposer comme sommes d'un polynôme et de fractions de type simple, plus précisément de fractions $\frac{\lambda}{(X - a)^k}$.

Lemme 19.5.16

Soit F une fraction rationnelle de $\mathbb{C}(X)$, et r_1, \dots, r_k ses pôles, de multiplicités $\alpha_1, \dots, \alpha_k$. Alors, il existe des polynômes P_1, \dots, P_k tels que

$$F = \sum_{i=1}^k \frac{P_i}{(X - r_i)^{\alpha_i}}.$$

Théorème 19.5.17 (Décomposition en éléments simples dans $\mathbb{C}(X)$, DÉS)

Soit F une fraction rationnelle de $\mathbb{C}(X)$, et r_1, \dots, r_k ses pôles, de multiplicités $\alpha_1, \dots, \alpha_k$. Alors il existe un unique polynôme E et d'uniques complexes $\lambda_{i,j}$ ($1 \leq i \leq k, 1 \leq j \leq \alpha_i$) tels que

$$F = E + \sum_{i=1}^k \sum_{j=1}^{\alpha_i} \frac{\lambda_{i,j}}{(X - r_i)^j}.$$

De plus, le polynôme E est la partie entière de la fraction rationnelle F , donc obtenue en effectuant la division euclidienne de P par Q , où $F = \frac{P}{Q}$

Définition 19.5.18 (Partie polaire)

Avec les notations du théorème précédent, la somme $\sum_{j=1}^{\alpha_i} \frac{\lambda_{i,j}}{(X - r_i)^j}$ est appelée partie polaire de F relativement au pôle r_i .

Exemples 19.5.19

1. Forme de la décomposition en éléments simples de $\frac{X^7}{(X - 1)^3(X + 1)^4}$.
2. Forme de la décomposition en éléments simples de $\frac{1}{(1 + X^2)^3}$

Proposition 19.5.20 (cas d'un pôle simple)

Soit r un pôle simple de $F = \frac{P}{Q}$ (sous forme irréductible), et soit \hat{Q} le polynôme $\frac{Q}{X - r}$. Alors le coefficient λ du terme $\frac{1}{X - \lambda}$ de la DÉS de F est :

$$\lambda = \frac{P(r)}{\hat{Q}(r)} = \frac{P(r)}{Q'(r)}.$$

Un cas important de décomposition en éléments simples est le cas de la fraction rationnelle $\frac{P}{P'}$.

Théorème 19.5.21 (Décomposition en éléments simples de $\frac{P'}{P}$)

Soit P un polynôme non nul de $\mathbb{C}[X]$. Soit r_1, \dots, r_k les racines de P , de multiplicités $\alpha_1, \dots, \alpha_k$. Alors les pôles de $\frac{P'}{P}$ sont r_1, \dots, r_k et sont tous des pôles simples. La DÉS de $\frac{P'}{P}$ est :

$$\frac{P'}{P} = \sum_{i=1}^k \frac{\alpha_i}{X - r_i}.$$

Nous avons déjà vu que les racines de la dérivée P' d'un polynôme scindé sont situées entre la racine minimale et la racine maximale de P . De la décomposition de $\frac{P'}{P}$, on déduit une propriété similaire dans $\mathbb{C}[X]$:

Corollaire 19.5.22 (Localisation des racines de P' , HP)

Soit P un polynôme de $\mathbb{C}[X]$. Alors les racines de P' sont dans l'enveloppe convexe des racines de P .

V.4 Décomposition en éléments simples dans $\mathbb{R}[X]$

Théorème 19.5.23 (DÉS dans $\mathbb{R}(X)$)

Soit $F = \frac{P}{Q}$ une fraction rationnelle sous forme irréductible, et $Q = Q_1^{\alpha_1} \cdots Q_k^{\alpha_k}$ la décomposition en facteurs irréductibles de Q dans $\mathbb{R}[X]$. Ainsi, les Q_i sont de degré 1 ou 2. Alors il existe un unique polynôme E , et d'unique polynômes $A_{i,j}$, de degré strictement plus petit que Q_i , tels que

$$F = E + \sum_{i=1}^k \sum_{j=1}^{\alpha_k} \frac{A_{i,j}}{Q_i^j}.$$

Remarques 19.5.24

1. Si Q_i est de degré 1, la partie correspondante dans la DÉS dans $\mathbb{R}(X)$ est la même que dans $\mathbb{C}(X)$.
2. Si Q_i est de degré 2, il admet deux racines complexes conjuguées r et \bar{r} . La partie correspondante dans la DÉS est obtenue en regroupant les parties polaires relatives à r et \bar{r} .
3. On a déjà vu en pratique comment déterminer des DÉS dans des cas simples. On a aussi déjà vu l'intérêt que peuvent avoir ces DÉS, notamment pour le calcul d'intégrales.

Espaces vectoriels

La première impulsion est venue de considération sur la signification des nombres négatifs en géométrie. Habitué à voir AB comme une longueur, j'étais néanmoins convaincu que $AB = AC + CB$, quelle que soit la position de A , B et C sur une droite.

(Herrmann Günther Grassmann)

Il semble que ce soit le destin de Grassmann d'être redécouvert de temps en temps, à chaque fois comme s'il avait été pratiquement oublié.

(A.C. Lewis)

La structure d'espace vectoriel est une structure rigide, généralisant le cadre géométrique usuel.

Il s'agit d'une structure algébrique liée à la donnée d'un corps, qui va constituer l'unité de la dimension : la droite réelle représente l'espace vectoriel typique sur \mathbb{R} de dimension 1, alors que le \mathbb{C} -espace vectoriel typique de dimension 1 ressemblera au plan complexe, donc à un objet géométrique, qui, en tant que \mathbb{R} -espace vectoriel sera en fait de dimension 2.

La rigidité de la structure se traduit par le fait qu'on peut multiplier un élément par un scalaire (un élément du corps de base), ceci de façon injective (sauf pour 0) : ainsi, si un élément x est dans un espace vectoriel E , tous les éléments λx , $\lambda \in \mathbb{K}$ seront aussi dans E , et si $x \neq 0$, l'ensemble des λx , $\lambda \in \mathbb{K}$ « ressemble » à \mathbb{K} (il y a une bijection entre les deux). On parle de la droite engendrée par x . Ainsi un espace vectoriel est une structure droite, qui, dès qu'elle contient un élément non nul, contient tout la droite (au sens du corps \mathbb{K}) contenant x .

La rigidité d'un espace vectoriel est même plus forte que cela : plus que la stabilité par multiplication par un scalaire, on a la stabilité par combinaison linéaire (et encore une fois, l'application qui à (λ, μ) de \mathbb{K}^2 associe $\lambda x + \mu y$ est bijective, sauf si x et y sont colinéaire). Ainsi, si E contient deux points (non colinéaires), il contient tout un \mathbb{K} -plan passant par ces deux points et par l'origine.

Cette structure rigide (plate pourrait-on dire) généralise la situation du plan réel usuel (approximation de la surface localement plate de la Terre sur laquelle nous faisons notre géométrie) ou de l'espace usuel, donc de la géométrie euclidienne classique. Elle ne permet en revanche pas de prendre en compte de façon implicite des phénomènes de courbure intrinsèque (géométrie sphérique définie intrinsèquement sur un objet de dimension 2, sans plongement dans un espace de dimension 3, ou propriétés de courbures de l'espace-temps) : ces structures courbes nécessitent l'introduction d'objets plus complexes (les variétés).

Dans tout le chapitre, on considère un corps \mathbb{K} . Vous pouvez considérer que $\mathbb{K} = \mathbb{R}$ ou \mathbb{C} (seul cas à connaître si on respecte le programme), mais sauf mention explicite du contraire, les définitions et les résultats sont valables pour tout corps (commutatif selon notre définition des corps).

I Notion d'espace vectoriel

I.1 Définition

Définition 20.1.1 (Espace vectoriel)

Soit E un ensemble. On dit que E est un espace vectoriel sur \mathbb{K} (en abrégé \mathbb{K} -ev) si E est muni de deux lois :

- une loi de composition externe $+$: $E \times E \rightarrow E$;
- une loi de composition externe \cdot : $\mathbb{K} \times E \rightarrow E$;

telles que :

- (i) $(E, +)$ soit un groupe abélien
- (ii) pour tout $(\lambda, \mu) \in \mathbb{K}^2$, $(x, y) \in E^2$:
 - $(\lambda\mu)x = \lambda(\mu x)$ (associativité externe de \cdot) ;
 - $1_{\mathbb{K}} \cdot x = x$ (compatibilité du neutre multiplicatif de \mathbb{K}) ;
 - $\lambda(x_1 + x_2) = \lambda x_1 + \lambda x_2$ (distributivité de \cdot sur la loi interne) ;
 - $(\lambda + \mu)x = \lambda x + \mu x$ (distributivité de \cdot sur la somme de \mathbb{K}).

Propriétés 20.1.2

Soit E un \mathbb{K} -ev. Pour tout $x \in E$:

1. $0 \cdot x = 0$, c'est-à-dire $0_{\mathbb{K}} \cdot x = 0_E$;
2. $\lambda \cdot 0_E = 0_E$
3. $(-1) \cdot x = -x$.

Terminologie 20.1.3

- Les éléments de E sont appelés *vecteurs*
- Les éléments de \mathbb{K} sont appelés *scalaires*
- Deux éléments x et y de E sont colinéaires s'il existe $(\lambda, \mu) \in \mathbb{K}^2$ tels que $(\lambda, \mu) \neq (0, 0)$ et $\lambda x + \mu y = 0$.

I.2 Combinaisons linéaires

Une propriété cruciale d'un espace vectoriel E est la stabilité par combinaison linéaire : si $(\lambda, \mu) \in \mathbb{K}^2$ et $(x, y) \in E^2$, alors $\lambda x + \mu y \in E$. La notion de combinaison linéaire étant centrale dans l'étude des espaces vectoriels, nous définissons une notion généralisée de combinaison linéaire.

Définition 20.1.4 (Famille presque nulle)

Soit I un ensemble d'indices, et $(\lambda_i)_{i \in I}$ une famille d'éléments de \mathbb{K} . On dit que la famille $(\lambda_i)_{i \in I}$ est presque nulle, ou qu'elle est à support fini, ou encore que les λ_i sont presque tous nuls, si seul un nombre fini de λ_i est non nul, autrement dit s'il existe $J \subset I$ un sous-ensemble fini de I tel que pour tout $i \in I \setminus J$, $\lambda_i = 0$.

Si I est fini, toute famille est à support fini. Peut-on dire qu'elle est presque nulle ? La terminologie n'est certainement pas très heureuse dans cette situation, puisque toute famille est alors presque nulle même si aucun vecteur n'est nul, mais nous l'utiliserons tout de même.

Définition 20.1.5 (Combinaison linéaire généralisée)

Soit E un \mathbb{K} -ev et $(x_i)_{i \in I}$ une famille de vecteurs de E . Une combinaison linéaire des $(x_i)_{i \in I}$ est un vecteur

$$x = \sum_{i \in I} \lambda_i x_i,$$

où $(\lambda_i)_{i \in I}$ est une famille à support fini de scalaires de \mathbb{K} .

Ainsi, toute combinaison linéaire sur une famille infinie est une combinaison linéaire d'un nombre fini de vecteurs de cette famille.

I.3 Un exemple important : espace de fonctions

Comme nous avons pu nous en rendre compte pour les groupes et les anneaux, on a des critères souvent rapides pour montrer qu'un ensemble est un sous-truc d'un truc plus gros. Nous verrons un peu plus loin que de la même façon, il est beaucoup plus commode de montrer qu'un ensemble est un sous-espace vectoriel d'un espace vectoriel connu plutôt que de montrer de façon directe qu'il s'agit d'un espace vectoriel, ce qui nécessite beaucoup de petites vérifications, élémentaires mais fastidieuses. Pour cette raison, il est important de connaître un certain nombre d'espaces vectoriels de référence, qui seront suffisants pour justifier la structure d'espace vectoriel d'autres ensembles dans la plupart des cas rencontrés.

Proposition 20.1.6 (espace vectoriel de référence)

1. Soit F un ensemble quelconque. Alors l'ensemble de fonctions \mathbb{K}^F est un espace vectoriel sur \mathbb{K} .
2. Plus généralement, soit E un espace vectoriel sur \mathbb{K} et F un ensemble quelconque. Alors l'ensemble de fonctions E^F est un espace vectoriel sur \mathbb{K} .

On déduit de cet exemple les espaces vectoriels usuels, comme cas particulier de la propriété précédente, suivant l'ensemble F choisi :

Exemples 20.1.7 (Espaces vectoriels à bien connaître)

1. $\mathbb{K}^\emptyset = \{0\}$;
2. $\mathbb{K}^{\{1\}} = \mathbb{K}$
3. $\mathbb{K}^{[1,n]} = \mathbb{K}^n$;
4. $\mathbb{K}^{\mathbb{N}}$ l'ensemble des suites à valeurs dans \mathbb{K} ;
5. $\mathbb{K}^{[0,n]} = \mathbb{K}_n[X]$ l'espace des polynômes de degré au plus n ;
6. $\mathbb{C} = \mathbb{R}^2$ est un \mathbb{R} -ev ;
7. $E_2^{E_1}$ les applications entre deux espaces vectoriels ;

I.4 Produits d'espaces vectoriels

Nous voyons maintenant deux façons de construire des espaces vectoriels à partir d'espaces vectoriels de référence : tout d'abord une construction externe (produit cartésien), puis dans la section suivante, une construction interne (sous-espaces vectoriels).

Proposition 20.1.8 (produit cartésien d'ev)

Soit E_1, \dots, E_n des espaces vectoriels sur un corps \mathbb{K} . Alors le produit cartésien $E_1 \times \dots \times E_n$ est un espace vectoriel sur \mathbb{K} , lorsqu'on le munit des lois définies par :

- $\forall \lambda \in \mathbb{K}, \forall (x_1, \dots, x_n) \in E_1 \times \dots \times E_n, \lambda \cdot (x_1, \dots, x_n) = (\lambda x_1, \dots, \lambda x_n)$;
- $\forall (x_1, \dots, x_n), (y_1, \dots, y_n) \in E_1 \times \dots \times E_n, (x_1, \dots, x_n) + (y_1, \dots, y_n) = (x_1 + y_1, \dots, x_n + y_n)$.

On retrouve en particulier la structure d'espace vectoriel de \mathbb{K}^n , déjà obtenue en considérant $\mathbb{K}^{[1,n]}$.

I.5 Sous-espaces vectoriels

Selon les définitions générales sur les structures, on définit :

Définition 20.1.9 (Sous-espace vectoriel, sev)

Soit E un espace vectoriel sur \mathbb{K} . Un sous-ensemble $F \subset E$ de E est un sous-espace vectoriel de E si F est stable par les lois $+$ et \cdot et que les lois induites munissent F d'une structure d'espace vectoriel.

Comme dans le cas des groupes et des anneaux, nous disposons d'un critère simple permettant de court-circuiter un certain nombre de vérifications :

Théorème 20.1.10 (Critère de sous-espace vectoriel)

Soit E un \mathbb{K} -espace vectoriel. Un ensemble F est un sous-espace vectoriel de E si :

- (i) $F \subset E$
- (ii) $0 \in F$
- (iii) F est stable par combinaison linéaire, ce qui équivaut à :

$$\forall x, y \in F, \forall \lambda \in \mathbb{K}, \lambda x + y \in F.$$

Exemples 20.1.11

1. Étant donné un espace vectoriel E , le sous-espace vectoriel nul $\{0_E\}$ et le sous-espace vectoriel total E .
2. Étant donné un vecteur X de \mathbb{R}^2 , la droite $\mathbb{R}X = \{\lambda X, \lambda \in \mathbb{R}\}$
3. Étant donné a, b et c trois réels, le plan de \mathbb{R}^3 d'équation $ax + by + cz = 0$.
4. $\mathbb{R}[X]$ espace des polynômes
5. $\mathcal{C}(\mathbb{R}, \mathbb{R})$ ensemble des fonctions continues sur \mathbb{R} ;
6. plus généralement $\mathcal{C}(I, \mathbb{R})$, ensemble des fonctions continues sur un intervalle I ;
7. $\mathcal{C}^n(\mathbb{R}, \mathbb{R})$ ensemble des fonctions de classe \mathcal{C}^n sur \mathbb{R} ;
8. plus généralement $\mathcal{C}^n(I, \mathbb{R})$, ensemble des fonctions de classe \mathcal{C}^n sur un intervalle I ;
9. Les exemples d'espaces vectoriels de fonctions sont nombreux.

Vous remarquerez dans les premiers exemples les deux points de vue différents pour définir un sous-espace vectoriel : par l'intérieur (sous-espace engendré par un vecteur) ou par l'extérieur (sous-espace défini par une équation sur les coordonnées). On retrouvera souvent ces deux points de vue par la suite.

Définition 20.1.12 (Droite vectorielle, plan vectoriel)

Soit E un \mathbb{K} -espace vectoriel. Une droite vectorielle de E est un sous-ensemble D de E tel qu'il existe $x \in E$ non nul tel que $D = \mathbb{K}x$.

Proposition 20.1.13 (Structure des droites vectorielles)

Les droites vectorielles d'un espace vectoriel E sur \mathbb{K} sont des sous-espaces vectoriels de E .

Proposition 20.1.14

Soit E un \mathbb{K} -espace vectoriel, et D_1 et D_2 deux droites vectorielles. Alors soit $D_1 \cap D_2 = \{0_E\}$, soit $D_1 = D_2$.

Corollaire 20.1.15

Soit D une droite vectorielle d'un espace vectoriel E , et $x \in D$. Si $x \neq 0$, alors $D = \mathbb{K}x$.

Proposition 20.1.16 (Sous-espaces vectoriels de \mathbb{R}^2)

Les sous-espaces vectoriels de \mathbb{R}^2 sont :

- le sous-espace vectoriel nul ;
- les droites vectorielles ;
- le sous-espace vectoriel total \mathbb{R}^2 .

De même, on montrera par la suite que les sous-espaces vectoriels de \mathbb{R}^3 sont exactement l'espace vectoriel nul, les droites vectorielles, les plans vectoriels (plans passant par l'origine) et \mathbb{R}^3 tout entier.

Remarque 20.1.17

L'aspect géométrique d'une droite vectorielle dépend du corps de base \mathbb{K} :

- Si le corps de base est \mathbb{R} , une droite vectorielle a l'aspect d'une droite géométrique usuelle.
- Si $\mathbb{K} = \mathbb{C}$, une droite a l'aspect d'un plan complexe : une droite complexe est donc un objet de dimension géométrique égale à 2.
- Si $\mathbb{K} = \mathbb{F}_p$, alors une droite est constituée d'un nombre fini de points « alignés circulairement » si on peut dire cela ainsi...
- Par exemple, si $\mathbb{K} = \mathbb{F}_2$, une droite est constituée de deux points : il y a dans ce cas autant de droites que de vecteurs non nuls de E , les droites étant les ensembles $\{0, x\}$, $x \neq 0$.

I.6 Intersections de sev**Proposition 20.1.18 (Intersection de sev)**

Soit E un \mathbb{K} -espace vectoriel, et $(E_i)_{i \in I}$ une famille de sous-espaces vectoriels de E . Alors $\bigcap_{i \in I} E_i$ est un sev de E .

En revanche, l'union de deux sev n'est en général pas un sev.

I.7 Sous-espace vectoriel engendré par un sous-ensemble**Définition 20.1.19 (Sous-espace vectoriel engendré par un sous-ensemble)**

Soit E un \mathbb{K} -ev, et X un sous-ensemble de E . Le sous-espace vectoriel engendré par X est le plus petit sous-espace vectoriel de E contenant X . Il est noté $\text{Vect}(X)$.

Remarque 20.1.20

Pourquoi un tel espace existe-t-il ?

Si X est exprimée sous forme d'une famille $(x_i)_{i \in I}$, on note $\text{Vect}(X) = \text{Vect}(x_i)_{i \in I}$. Si X est fini et énuméré par exemple $x = \{x_1, \dots, x_n\}$, on notera $\text{Vect}(X) = \text{Vect}(x_1, \dots, x_n)$.

Proposition 20.1.21 (Minimalité de $\text{Vect}(X)$)

Par définition, tout sous-espace vectoriel de E contenant X contient aussi $\text{Vect}(X)$.

On peut donner une description explicite de $\text{Vect}(X)$ à l'aide de combinaisons linéaires :

Proposition 20.1.22 (Description de $\text{Vect}(X)$)

Soit E un \mathbb{K} -ev et X un sous-ensemble de E . Alors E est l'ensemble des combinaisons linéaires d'éléments de X .

Ainsi, $x \in \text{Vect}(X)$ si et seulement s'il existe $(x_1, \dots, x_n) \in X^n$ et $(\lambda_1, \dots, \lambda_n) \in \mathbb{K}^n$ tels que

$$x = \lambda_1 x_1 + \dots + \lambda_n x_n.$$

Exemples 20.1.23

1. $\text{Vect}(x) = \{\lambda x, \lambda \in \mathbb{K}\} = \mathbb{K}x$
2. $\text{Vect}(x, y) = \{\lambda x + \mu y, (\lambda, \mu) \in \mathbb{K}^2\}$
 - si $x = y = 0$, $\text{Vect}(x, y) = \{0\}$
 - si x et y sont colinéaires, non tous deux nuls (disons $x \neq 0$), $\text{Vect}(x, y) = \text{Vect}(x)$
 - si x et y ne sont pas colinéaires, $\text{Vect}(x, y)$ est un plan vectoriel.

I.8 Sommes de sev

Définition 20.1.24 (Somme de deux sev)

Soit G un \mathbb{K} -ev, et E et F deux sev de G . Alors la somme $E + F$ de E et F est le plus petit sous-espace vectoriel de G , contenant à la fois E et F :

$$E + F = \text{Vect}(E \cup F).$$

Proposition 20.1.25 (Description d'une somme)

Soit G un \mathbb{K} -ev, et E et F deux sev de G . Alors :

$$E + F = \{x + y \mid x \in E, y \in F\}.$$

Définition 20.1.26 (Somme d'un nombre fini de sous-espaces)

Plus généralement, on définit la somme des sous-espaces E_1, \dots, E_n par :

$$E_1 + \dots + E_n = \text{Vect}(E_1 \cup \dots \cup E_n)$$

Proposition 20.1.27 (Description d'une somme d'un nombre fini de sev)

Soit E_1, \dots, E_n et F des sev de E . Les propriétés suivantes sont équivalentes :

- (i) $F = E_1 + \dots + E_n$
- (ii) $F = (((E_1 + E_2) + E_3) + \dots + E_{n-1}) + E_n$
- (iii) $F = \{x_1 + \dots + x_n \mid (x_1, \dots, x_n) \in E_1 \times \dots \times E_n\}$.

Exemples 20.1.28

1. Dans \mathbb{R}^2 : $\text{Vect}((1, 2), (-1, 1)) = \mathbb{R}^2$

2. Soit \mathcal{E} l'ensemble des suites vérifiant une relation de récurrence linéaire de polynôme caractéristique P , ayant n racines deux à deux distinctes x_1, \dots, x_k dans \mathbb{C} . Alors, dans $\mathbb{C}^{\mathbb{N}}$:

$$\mathcal{E} = \text{Vect}((x_1^n)_{n \in \mathbb{N}}, (x_2^n)_{n \in \mathbb{N}}, \dots, (x_k^n)_{n \in \mathbb{N}}).$$

3. Soit f une fonction définie sur \mathbb{R} par une équation différentielle homogène d'ordre 2, dont le polynôme caractéristique admet deux racines réelles distinctes λ et μ . Alors

$$f \in \text{Vect}((x \mapsto e^{\lambda x}), (x \mapsto e^{\mu x})).$$

4. Soit E un ev, et x_1, \dots, x_n des éléments de E . Alors :

$$\text{Vect}(x_1, \dots, x_n) = \mathbb{R}x_1 + \dots + \mathbb{R}x_n.$$

Le dernier exemple peut être vu comme une utilisation de la proposition suivante :

Proposition 20.1.29

Soit $(x_i)_{i \in I}$ et $(x_j)_{j \in J}$ deux familles (pas nécessairement disjointes). Alors :

$$\text{Vect}((x_i)_{i \in I \cup J}) = \text{Vect}((x_i)_{i \in I}) + \text{Vect}((x_j)_{j \in J}).$$

Exemple 20.1.30

$$\text{Vect}(x_1, x_2, x_4, x_5) + \text{Vect}(x_3, x_4, x_6) = \text{Vect}(x_1, x_2, x_3, x_4, x_5, x_6).$$

I.9 Sommes directes

Définition 20.1.31 (Somme directe)

Soit E et F deux sev de G . On dit que la somme $E + F$ est *directe*, et on note $E \oplus F$, si $E \cap F = \{0\}$.

Plus généralement, $E_1 + \dots + E_n$ est directe si $E_1 \oplus E_2$, puis $(E_1 + E_2) \oplus E_3$, etc. On note $\bigoplus_{i=1}^n E_i$.

Proposition 20.1.32 (Caractérisation de \oplus par les intersections)

La somme $E_1 + \dots + E_n$ est directe si et seulement si pour tout $k \in \llbracket 2, n \rrbracket$,

$$F_k \cap \sum_{i < k} F_i = \{0\}.$$

Proposition 20.1.33 (Caractérisation de \oplus par l'unicité)

La somme $\sum E_i$ de sous-espaces vectoriels de E est directe si et seulement si l'application ci-dessous est injective :

$$\begin{aligned} \varphi : E_1 \times \dots \times E_n &\longrightarrow E \\ (x_1, \dots, x_n) &\longmapsto x_1 + \dots + x_n. \end{aligned}$$

En d'autres termes, $E_1 \oplus \dots \oplus E_n$ est directe si et seulement si tout élément x de $E_1 \oplus \dots \oplus E_n$ s'écrit de façon unique sous la forme $x = x_1 + \dots + x_n$, $x_i \in E_i$.

Définition 20.1.34 (Supplémentaire)

Soit E un espace vectoriel, et F et G deux sev de E . On dit que F et G sont supplémentaires dans E si $F \oplus G = E$.

Théorème 20.1.35 (Existence d'un supplémentaire)

Soit E un espace vectoriel quelconque, et F un sev de E . Alors F admet au moins un supplémentaire G .

La démonstration de ce théorème repose sur le lemme de Zorn, donc sur l'axiome du choix. On en verra une démonstration plus élémentaire, indépendante de l'axiome du choix, lorsque E est de dimension finie.

II Familles de vecteurs

Nous rappelons que toute combinaison linéaire d'une famille (finie ou infinie) s'exprime, par définition, comme somme *finie* d'éléments de cette famille, multipliés par des scalaires.

II.1 Familles libres

Proposition/Définition 20.2.1 (famille libre)

Une famille $(x_i)_{i \in I}$ de vecteurs de E est *libre* si une des propriétés équivalentes suivantes est vérifiée :

- (i) Pour toute famille $(\lambda_i)_{i \in I}$ de scalaires presque tous nuls : $\sum_{i \in I} \lambda_i x_i = 0 \implies \forall i \in I, \lambda_i = 0$;
- (ii) Pour tout $x \in \text{Vect}((x_i)_{i \in I})$, il existe une *unique* famille $(\lambda_i)_{i \in I}$ de scalaires presque tous nuls tels que $x = \sum_{i \in I} \lambda_i x_i$

Si de plus, $I = \llbracket 1, n \rrbracket$, les points (i) et (ii) sont équivalents aux points suivants :

- (iii) la somme $\mathbb{R}x_1 \oplus \dots \oplus \mathbb{R}x_n$ est directe ;
- (iv) La fonction $\varphi : \mathbb{K}^n \longrightarrow E$ définie par $\varphi(\lambda_1, \dots, \lambda_n) = \lambda_1 x_1 + \dots + \lambda_n x_n$ est injective.

Définition 20.2.2 (famille liée)

Une famille qui n'est pas libre est dite *liée*.

Remarque 20.2.3

Une famille contenant 0 peut-elle être libre ?

Proposition 20.2.4 (Stabilité de la liberté par restriction)

Toute sous-famille d'une famille libre est libre.

La proposition suivante permet de ramener l'étude de la liberté des familles infinies à l'étude de la liberté de familles finies.

Proposition 20.2.5 (Caractérisation de la liberté pour des familles infinies)

Une famille $(x_i)_{i \in I}$ est libre si et seulement si toutes ses sous-familles finies sont libres. Une famille $(x_i)_{i \in \mathbb{N}}$ est libre si et seulement si pour tout $n \in \mathbb{N}$, la famille (x_1, \dots, x_n) est libre.

Proposition 20.2.6 (Ajout d'un élément à une famille libre)

Soit $(x_i)_{i \in I}$ une famille libre de E et x_j ($j \notin I$) un élément de E . Alors, la famille $(x_i)_{i \in I \cup \{j\}}$, obtenue en ajoutant x_j à la famille libre $(x_i)_{i \in I}$, est libre si et seulement si $x_j \notin \text{Vect}((x_i)_{i \in I})$

Si un tel ajout est impossible, on dira que la famille libre est maximale :

Définition 20.2.7 (Famille libre maximale)

Une famille libre est maximale si et seulement s'il est impossible de lui rajouter un vecteur (quelconque) de E en préservant sa liberté

Proposition 20.2.8 (Caractérisation des sommes directes par la liberté)

Soit E_1, \dots, E_n des sev de E . Alors la somme $E_1 \oplus \dots \oplus E_n$ est directe si et seulement si tout n -uplet (x_1, \dots, x_n) d'éléments tous non nuls de $E_1 \times \dots \times E_n$ est une famille libre dans E .

II.2 Familles génératrices**Définition 20.2.9 (familles génératrices)**

Une famille $(x_i)_{i \in I}$ de vecteurs de E est une famille *génératrice* de E si l'une des propriétés équivalentes suivantes est satisfaite :

- (i) tout $x \in E$ est une combinaison linéaire des x_i , $i \in I$;
- (ii) $\text{Vect}((x_i)_{i \in I}) = E$.

Si de plus $I = \llbracket 1, n \rrbracket$, les points (i) et (ii) sont équivalents à :

- (iii) $E = \sum_{i=1}^n \mathbb{R}x_i$;
- (iv) La fonction $\varphi : \mathbb{K}^n \rightarrow E$ définie par $\varphi(\lambda_1, \dots, \lambda_n) = \lambda_1 x_1 + \dots + \lambda_n x_n$ est surjective.

On dit aussi que la famille $(x_i)_{i \in I}$ engendre E .

On obtient des propriétés symétriques à celles des familles libres :

Proposition 20.2.10 (Stabilité des familles génératrices par ajout)

Toute famille contenant une famille génératrice de E est une famille génératrice de E .

Proposition 20.2.11 (Restriction d'une famille génératrice)

La famille obtenue en retirant un élément x d'une famille génératrice de E est encore génératrice si et seulement si x est une combinaison linéaire des autres vecteurs de la famille.

Lorsque cette situation n'est vérifiée pour aucun élément de la famille, on parle de famille génératrice minimale :

Définition 20.2.12 (famille génératrice minimale)

Une famille génératrice est dite minimale, si et seulement s'il est impossible de lui retirer un élément en préservant son caractère générateur.

II.3 Bases

Définition 20.2.13 (base d'un espace vectoriel)

Soit $(x_i)_{i \in I}$ une famille de vecteurs d'un ev E . On dit que $(x_i)_{i \in I}$ est une *base* de E si elle est une famille à la fois libre et génératrice de E .

Ainsi, $(b_i)_{i \in I}$ est une base de E si et seulement si tout élément x de E s'écrit de façon unique comme combinaison linéaire des éléments b_i : l'existence traduit le caractère générateur, l'unicité traduit la liberté. Les coefficients de cette combinaison linéaire sont appelés *coordonnées de x dans la base $(b_i)_{i \in I}$* .

Le choix d'une base de E permet donc de définir un système de coordonnées : la donnée d'un vecteur x équivaut à la donnée de ses coordonnées dans une base fixée

Exemple 20.2.14

- Les coordonnées cartésiennes dans \mathbb{R}^2 correspondent aux coordonnées dans la base canonique $((1, 0), (0, 1))$.
- Un autre choix de base fournit d'autres coordonnées, par exemple $(2, 3)$ dans la base $((1, 0), (1, 1))$

Proposition 20.2.15 (Caractérisation des bases par minimalité/ maximalité)

Les propriétés suivantes sont équivalentes :

- (i) La famille $(x_i)_{i \in I}$ est une base de E ;
- (ii) La famille $(x_i)_{i \in I}$ est une famille génératrice minimale de E ;
- (iii) La famille $(x_i)_{i \in I}$ est une famille libre maximale de E .

Exemples 20.2.16 (Exemples importants de bases, à connaître)

- Base canonique de \mathbb{K}^n .
- Base canonique $(1, X, X^2, \dots)$ de $\mathbb{K}[X]$; base canonique de $\mathbb{K}_n[X]$.
- $((X - x_0)^k)_{k \in \llbracket 0, n \rrbracket}$ est une base de $\mathbb{K}_n[X]$.

Ce dernier exemple a une généralisation importante, qui mérite d'être citée en tant que proposition :

Proposition 20.2.17 (Famille échelonnée en degrés)

Si (P_0, \dots, P_n) est une famille d'éléments de $\mathbb{K}_n[X]$ telle que pour tout $k \in \llbracket 0, n \rrbracket$, $\deg(P_k) = k$, alors (P_0, \dots, P_n) est une base de $\mathbb{K}_n[X]$.

Avertissement 20.2.18

La réciproque est fausse !

III Espaces vectoriels de dimension finie

Dans tout ce paragraphe \mathbb{K} désigne le corps \mathbb{R} ou \mathbb{C} .

III.1 Notion de dimension

Définition 20.3.1 (Espace vectoriel de dimension finie)

Un espace vectoriel E sur \mathbb{K} est dit *de dimension finie* s'il existe une famille génératrice de cardinal fini $(x_i)_{i \in I}$ de E . Une famille génératrice finie est souvent appelé *système de générateurs*.

Proposition 20.3.2

Soit E un espace vectoriel de dimension finie. Alors de toute famille génératrice de E , on peut extraire une famille génératrice finie.

Le premier théorème important est l'existence d'une base de cardinal fini d'un espace vectoriel de dimension finie. On montre un résultat plus fort, disant que toute famille libre peut être vue comme le début d'une base.

Théorème 20.3.3

Soit E un espace vectoriel de dimension finie. Soit L une famille libre de E , et G une famille génératrice de E . Alors on peut compléter L en une base de E par ajout de vecteurs de G .

En particulier, si $(x_i)_{1 \leq i \leq n}$ est génératrice de E , et $(x_i)_{i \in I}$ est libre, pour $I \subset \llbracket 1, n \rrbracket$, il existe J tel que $I \subset J \subset \llbracket 1, n \rrbracket$, tel que $(x_j)_{j \in J}$ soit une base de E .

Corollaire 20.3.4 (Cardinal des familles libres dans un en de dimension finie)

Soit E un espace vectoriel de dimension finie. Toute famille libre de E est de cardinal fini. En particulier, toute base est de cardinal fini.

Corollaire 20.3.5 (Théorème de la base extraite)

Soit E un espace vectoriel de dimension finie.

- (i) De toute partie génératrice de E on peut extraire une base de E .
- (ii) E admet au moins une base.

Corollaire 20.3.6 (Théorème de la base incomplète)

Toute famille libre d'un espace de dimension fini E peut être complétée en une base de E .

Nous établissons maintenant le théorème de la dimension, qui à la base de la théorie de la dimension. Pour le démontrer, nous utilisons le lemme suivant :

Lemme 20.3.7 (Théorème d'échange)

Soit E un espace vectoriel. Soit F une famille de vecteurs de E , et x et y dans E tels que $x \notin \text{Vect}(F)$ et $x \in \text{Vect}(F \cup \{y\})$. Alors $\text{Vect}(F \cup \{x\}) = \text{Vect}(F \cup \{y\})$.

Théorème 20.3.8 (Théorème de la dimension)

Soit E un espace vectoriel de dimension finie. Alors toutes les bases de E sont finies et de même cardinal.

L'importance de ce théorème provient du fait qu'il permet de définir la notion fondamentale suivante :

Définition 20.3.9 (dimension d'un espace vectoriel)

Le cardinal commun de toutes les bases de E est appelé *dimension de E* , et est noté $\dim E$. Si E n'est pas de dimension finie, on dira que E est de dimension infinie.

La notion de dimension est la formalisation de la notion de nombre de degrés de liberté dont on dispose pour construire un objet. Par exemple, pour définir entièrement une suite par une récurrence linéaire homogène d'ordre 3, on dispose de 3 degrés de liberté, à savoir le choix des trois premiers termes de la suite. L'espace des suites vérifiant une telle relation de récurrence est *de facto* de dimension 3.

Exemples 20.3.10

1. Dimension de \mathbb{K}^n
2. Dimension de $\mathbb{K}_n[X]$
3. Dimension de l'ensemble des suites vérifiant une récurrence linéaire homogène d'ordre k
4. Dimension de l'ensemble des solutions d'une équation différentielle linéaire homogène d'ordre 1, d'ordre 2 à coefficients constants.
5. Dimension de \mathbb{C} en tant que...
6. Dimension de $\mathcal{M}_{n,m}(\mathbb{R})$, de $\mathcal{M}_n(\mathbb{R})$.

Nous terminons cette section par :

Proposition 20.3.11 (Dimension d'un produit cartésien)

Soit E et F deux espaces vectoriels sur \mathbb{K} . Si E et F sont de dimension finie, l'espace vectoriel produit $E \times F$ est de dimension finie, égale à :

$$\dim(E \times F) = \dim(E) + \dim(F).$$

Plus précisément, si (b_1, \dots, b_m) et (c_1, \dots, c_n) sont des bases de E et F , une base de $E \times F$ est $((b_1, 0), \dots, (b_m, 0), (0, c_1), \dots, (0, c_n))$.

III.2 Dimension, liberté et rang**Définition 20.3.12 (Rang d'une famille de vecteurs)**

Soit E un espace vectoriel, $k \in \mathbb{N}^*$, et (x_1, \dots, x_k) une famille de vecteurs de E . Le *rang* de la famille (x_1, \dots, x_k) est la dimension de $\text{Vect}(x_1, \dots, x_k)$ (cet espace est de dimension finie, puisque engendré par une famille finie). On note :

$$\text{rg}(x_1, \dots, x_k) = \dim \text{Vect}(x_1, \dots, x_k).$$

Proposition 20.3.13 (Majoration du rang et cas d'égalité)

$\text{rg}(x_1, \dots, x_k) \leq k$, avec égalité si et seulement si la famille (x_1, \dots, x_k) est libre.

Proposition 20.3.14 (Majoration du cardinal d'une famille libre)

Soit E un espace vectoriel de dimension n . Alors toute famille libre de E est de cardinal au plus n , avec égalité si et seulement si c'est une base.

En particulier, toute famille de strictement plus de n éléments est liée.

Exemple 20.3.15

Montrer que toute matrice $M \in \mathcal{M}_n(\mathbb{R})$ admet un polynôme annulateur, c'est-à-dire un polynôme P tel que $P(M) = 0$. Quelle est la structure de l'ensemble des polynômes annulateurs de M ? Justifier l'existence d'un polynôme annulateur minimal pour la relation de divisibilité.

Proposition 20.3.16 (Minoration du cardinal d'une famille génératrice)

Soit E un espace vectoriel de dimension n . Alors toute famille génératrice de E est de cardinal au moins n , avec égalité si et seulement si c'est une base.

On en déduit en particulier :

Corollaire 20.3.17 (Caractérisation par le cardinal des familles libres maximales)

Soit E un espace vectoriel de dimension n .

- Une famille libre est maximale dans le sens donné plus haut si et seulement si elle est de cardinal n .
- Une famille génératrice est minimale si et seulement si elle est de cardinal n .

III.3 Dimension de sous-espaces et de sommes

Intuitivement, on ne dispose pas de plus de degré de liberté en restreignant l'espace. C'est ce que traduit le théorème suivant :

Théorème 20.3.18 (Dimension d'un sous-espace)

Soit E un espace vectoriel de dimension finie n . Soit $F \subset E$ un sous-espace vectoriel de E . Alors F est de dimension finie, et $\dim F \leq \dim E$. On a égalité si et seulement si $F = E$.

Par ailleurs, étant donné deux sous-espaces vectoriels de E , la dimension de leur somme sera facile à déterminer s'il n'y a pas de redondance (autrement dit, si on a unicité des écritures), ce qui se traduit pas le fait que la somme est directe. Cette absence de redondances correspond aussi intuitivement au cas où la dimension de la somme est maximale par rapport aux sommes des deux sous-espaces, ce que l'on justifiera rigoureusement plus tard.

Théorème 20.3.19 (Dimension d'une somme directe)

Soit E un espace vectoriel, et F et G des sous-espaces vectoriels de dimension finie de E , en somme directe. Alors $F \oplus G$ est de dimension finie, et :

$$\dim F \oplus G = \dim F + \dim G.$$

On en déduit facilement, par récurrence :

Corollaire 20.3.20 (Dimension d'une somme directe de n espaces)

Soit E un espace vectoriel, et (E_1, \dots, E_n) une famille de sous-espaces vectoriels de dimension finie de E , en somme directe. Alors $\bigoplus_{i=1}^n E_i$ est de dimension finie, et :

$$\dim \bigoplus_{i=1}^n E_i = \sum_{i=1}^n \dim E_i.$$

On en déduit en particulier la dimension des supplémentaires, après avoir justifié leur existence de façon plus élémentaire qu'en début de chapitre, dans le cadre de la dimension finie :

Théorème 20.3.21 (Existence et dimension d'un supplémentaire, en dimension finie)

Soit E un espace vectoriel de dimension finie, et F un sous-espace vectoriel de E . Alors il existe un supplémentaire S de F dans E , et :

$$\dim S = \dim E - \dim F.$$

Pour terminer, nous donnons une formule générale exprimant la dimension d'une somme quelconque :

Théorème 20.3.22 (Formule de Grassmann)

Soit E un espace vectoriel, et F et G deux sous-espaces de dimension finie de E . Alors :

$$\dim(F + G) = \dim F + \dim G - \dim F \cap G.$$

Théorème 20.3.23 (Majoration de la dimension d'une somme)

Soit E un espace vectoriel, et E_1, \dots, E_n des sous-espaces vectoriels de dimension finie de E . Alors

$$\dim(E_1 + \dots + E_n) \leq \dim(E_1) + \dots + \dim(E_n),$$

avec égalité si et seulement si la somme est directe.

Comme évoqué plus haut, ces deux derniers résultats affirment que moins il y a de redondances dans l'écriture d'une somme, plus la dimension de la somme est importante ; elle est maximale lorsqu'il n'y a aucune redondance, ce qui s'exprime par le fait que la somme est directe.

Remarque 20.3.24

- Comparez cette formule à $|A \cup B| = |A| + |B| - |A \cap B|$
- Peut-on généraliser, en trouvant pour les dimensions une formule analogue à la formule du crible de Poincaré :

$$\left| \bigcup_{i=1}^n A_i \right| = \sum_{j=1}^n (-1)^{j+1} \sum_{\substack{J \subset \llbracket 1, n \rrbracket \\ |J|=j}} \left| \bigcap_{i \in J} A_i \right| ?$$

Proposition 20.3.25 (Caractérisation des couples de sous-espaces supplémentaires)

Soit E un espace de dimension finie, et F et G deux sous-espaces vectoriels de E . Alors F et G sont supplémentaires l'un de l'autre si et seulement si

$$F \cap G = \{0\} \quad \text{et} \quad \dim F + \dim G = \dim E.$$

Note Historique 20.3.26

Herrmann Günther Grassmann (1809-1877) est le précurseur incompris de toute la théorie de l'algèbre linéaire. Il expose les fondements de cette théorie dans sa thèse *Théorie des flots et des marées*, thèse qui ne sera pas lue par son examinateur, et qui sera publiée uniquement au début du XX-ième siècle. Ce n'est qu'une vingtaine d'années plus tard (vers 1860) que certains mathématiciens se rendent compte de l'importance des travaux de Grassmann. Entre temps, Grassmann s'est contenté d'un poste d'enseignant en lycée. Détourné de la recherche mathématique, c'est dans des études linguistiques (sanskrit, gotique) qu'il finit par trouver la consécration.

Applications linéaires

Comme à chaque fois qu'on définit une nouvelle structure algébrique, il vient une notion de morphisme, adaptée à cette structure. La catégorie des espaces vectoriels sur un corps \mathbb{K} est ainsi formée d'un ensemble d'objets (les \mathbb{K} -ev), et de flèches représentant les morphismes entre espaces vectoriels, c'est-à-dire, selon les définitions générales qu'on en a données, les applications respectant les deux lois définissant un espace vectoriel. La propriété définissant ces morphismes se traduit par une propriété de linéarité $f(\lambda x + \mu y) = \lambda f(x) + \mu f(y)$. Pour cette raison, les morphismes d'espaces vectoriels sont plus fréquemment appelés *applications linéaires*.

Le but de ce chapitre est de donner les propriétés élémentaires des applications linéaires, d'étudier les propriétés de rigidité des applications linéaires (comme pour les polynômes, il suffit en général de connaître l'image d'un petit nombre de vecteurs pour déterminer entièrement une application linéaire), d'étudier certains types d'applications linéaires (endomorphismes, isomorphismes, projecteurs, symétries), de traduire certaines propriétés d'une application linéaire sur son comportement sur une base (caractérisations de l'injectivité, de la surjectivité).

Nous verrons ensuite des résultats spécifiques aux applications linéaires définies sur des espaces de dimension finie, notamment une formule reliant la dimension du noyau et la dimension de l'image (formule du rang).

Pour terminer nous donnons un aperçu de la notion d'espace dual, lié à la notion de forme linéaire et d'hyperplan.

Il est important de remarquer que la notion d'application linéaire en dimension finie est indissociable de la notion de matrice. Nous étudierons ce point de vue plus en détail dans le chapitre suivant.

Dans tout ce qui suit, \mathbb{K} désigne un corps quelconque. Vous pouvez considérer, conformément au programme, que $\mathbb{K} = \mathbb{R}$ ou \mathbb{C} , mais, sauf mention explicite du contraire, les résultats donnés sont valables pour tout corps.

I Généralités sur les applications linéaires

I.1 Définitions et propriétés de stabilité

Définition 21.1.1 (Application linéaire, AL)

Une application $f : E \rightarrow F$ entre deux \mathbb{K} -ev est appelée *application \mathbb{K} -linéaire*, ou plus simplement *application linéaire* (en abrégé : AL), si :

$$\forall \lambda \in \mathbb{K}, \forall x \in E, \quad f(\lambda x) = \lambda f(x) \quad \text{et} \quad \forall (x, y) \in E^2, \quad f(x + y) = f(x) + f(y).$$

Proposition 21.1.2 (Respect du neutre)

Soit $f : E \rightarrow F$ une AL. Alors $f(0_E) = 0_F$.

Proposition 21.1.3 (Caractérisation des AL par respect des CL)

Une application $f : E \rightarrow F$ est une application linéaire si et seulement si :

$$\forall \lambda \in \mathbb{K}, \forall (x, y) \in E^2, \quad f(\lambda x + y) = \lambda f(x) + f(y).$$

Exemples 21.1.4

1. L'application de \mathbb{R}^2 dans \mathbb{R} définie par $f(x, y) = 2x + 3y$?
2. L'application de \mathbb{R}^3 dans \mathbb{R} définie par $f(x, y) = 3x + 4y + 2z + 1$?
3. L'application de \mathbb{R}^2 dans \mathbb{R} définie par $f(x, y) = xy$?
4. La somme \sum , de \mathbb{C}^n dans \mathbb{C} .
5. L'intégrale, de $\mathcal{C}^0([a, b])$ dans \mathbb{R} .
6. L'opérateur de dérivation D de \mathcal{C}^n dans ...
7. L'opérateur de dérivation D de $\mathbb{R}[X]$ dans $\mathbb{R}[X]$.
8. Étant donné une matrice $M \in \mathcal{M}_{n,p}(\mathbb{K})$, et en assimilant \mathbb{K}^n à $\mathcal{M}_{n,1}(\mathbb{K})$, l'application de \mathbb{K}^p dans \mathbb{K}^n qui à une matrice colonne X associe la matrice colonne MX .

Nous verrons dans le chapitre suivant que ce dernier exemple fournit la description générique d'une application linéaire en dimension finie, après choix de bases de E et de F .

Nous obtenons ainsi l'ensemble des flèches entre deux objets de la catégorie des \mathbb{K} -espaces vectoriels

Définition 21.1.5 (Ensemble des applications linéaires)

Soit E et F deux \mathbb{K} -ev. On note $\mathcal{L}(E, F)$ l'ensemble des applications linéaires de E vers F .

Ces ensembles de flèches possèdent eux-même une structure d'espace vectoriel :

Proposition 21.1.6 (Structure de $\mathcal{L}(E, F)$)

$\mathcal{L}(E, F)$ est un espace vectoriel sur \mathbb{K} .

Autrement dit, une combinaison linéaire d'applications linéaires de E vers F est encore une application linéaire.

Étudions maintenant des propriétés liées à la composition.

Proposition 21.1.7 (Composée de deux applications linéaires)

Soit $f \in \mathcal{L}(E, F)$ et $g \in \mathcal{L}(F, G)$. Alors $g \circ f$ est une application linéaire de E vers G .

De manière générale, la composition d'applications à valeurs dans un espace vectoriel est toujours linéaire à gauche, c'est-à-dire $(\lambda f + \mu g) \circ h = \lambda f \circ h + \mu g \circ h$. Si les applications considérés sont linéaires, on obtient aussi la linéarité à droite. On rappelle :

Définition 21.1.8 (Application bilinéaire)

Soit E, F et G trois \mathbb{K} -espaces vectoriels, et $\varphi : E \times F \rightarrow G$. On dit que φ est bilinéaire si elle est linéaire par rapport à chacune de ses deux variables, l'autre étant fixée, c'est-à-dire si pour tout $(x, x', y, y', \lambda) \in E \times E \times F \times F \times \mathbb{K}$,

- $\varphi(\lambda x + x', y) = \lambda\varphi(x, y) + \varphi(x', y)$
- $\varphi(x, \lambda y + y') = \lambda\varphi(x, y) + \varphi(x, y')$.

La propriété énoncée plus haut pour les compositions s'exprime alors de la manière suivante :

Proposition 21.1.9 (Bilinéarité de la composition)

La composition d'applications linéaires est bilinéaire. En termes plus précis, E, F et G étant trois \mathbb{K} -ev, l'application Φ de $\mathcal{L}(E, F) \times \mathcal{L}(F, G)$ dans $\mathcal{L}(E, G)$ définie par $\Phi(u, v) = v \circ u$ est une application bilinéaire.

On pourrait de façon similaire définir une notion d'application n -linéaire (à n variables vectorielles). La composition de n AL est alors n -linéaire. On verra dans un chapitre ultérieur comment cette notion d'application multilinéaire est également liée à la notion de déterminant.

I.2 Image et noyau

Dans ce paragraphe, E et F sont deux espaces vectoriels, et $f \in \mathcal{L}(E, F)$. On étudie ici deux sous-espaces liés à une application linéaire : l'image (qui correspond à la notion usuelle d'image) et le noyau.

Définition 21.1.10 (Image et noyau)

1. L'image de f est $\text{Im}(f) = \{y \in F \mid \exists x \in E, f(x) = y\} = f(E)$;
2. Le noyau de f est $\text{Ker}(f) = \{x \in E \mid f(x) = 0\} = f^{-1}(\{0\})$

La structure algébrique de l'image et du noyau découle d'un résultat plus général de préservation de la structure par image directe et réciproque :

Lemme 21.1.11 (Structure des images directes et réciproques)

1. Soit E' un sev de E . Alors $f(E')$ est un sev de F .
2. Soit F' un sev de F . Alors $f^{-1}(F')$ est un sev de E .

En appliquant ce lemme avec $E' = E$ et $F' = \{0\}$, on obtient :

Proposition 21.1.12 (Structure de l'image et du noyau)

$\text{Im}(f)$ est un sev de F . $\text{Ker}(f)$ est un sev de E .

De façon évidente, l'image mesure le défaut de surjectivité, De façon symétrique, le noyau mesure le défaut d'injectivité : si $\text{Ker}(f)$ n'est pas un singleton, les définitions amènent de façon immédiate la non injectivité de f . La réciproque découle du fait que tout défaut d'injectivité peut être translaté en 0 par linéarité. Nous obtenons :

Théorème 21.1.13 (Caractérisation de la surjectivité et de l'injectivité)

Soit $f \in \mathcal{L}(E, F)$. Alors :

- (i) f est surjective ssi $\text{Im}(f) = F$;
- (ii) f est injective ssi $\text{Ker}(f) = \{0\}$.

Connaissant une famille génératrice de E (par exemple une base), il n'est pas dur de déterminer l'image de f :

Proposition 21.1.14 (Famille génératrice de $\text{Im}(f)$)

Soit $f \in \mathcal{L}(E, F)$ et $(e_i)_{i \in I}$ une famille génératrice de E . Alors $f(e_i)_{i \in I}$ est une famille génératrice de $\text{Im}(f)$:

$$\text{Im}(f) = \text{Vect}(f(e_i), i \in I).$$

En particulier, si f transforme une famille génératrice de E en une famille génératrice de F , alors f est surjective.

Si nous appliquons la propriété précédente à l'application de \mathbb{K}^p dans \mathbb{K}^n définie par $X \mapsto MX$, nous obtenons la description très simple d'une famille génératrice de $\text{Im}(f)$, dans le cas où f est donnée matriciellement :

Corollaire 21.1.15 (Image d'une AL décrite matriciellement)

Soit $M \in \mathcal{M}_{n,p}(\mathbb{K})$ et $f : \mathbb{K}^p \rightarrow \mathbb{K}^n$ (ces ensembles étant vus comme ensemble de vecteurs colonnes), définie par $f(X) = MX$. Alors l'image de f est engendrée par la famille des colonnes de la matrice M .

Exemple 21.1.16

Décrire l'image de $f : X \in \mathbb{R}^3 \mapsto \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 1 & 2 & 3 \end{pmatrix} X$.

On en déduit :

Méthode 21.1.17 (Déterminer l'image et le noyau d'une AL)

- Pour déterminer l'image d'une AL f :
 - * Si f est donnée par une matrice (éventuellement après choix d'une base, voir chapitre suivant), considérer la famille des colonnes de cette matrice, qui engendre $\text{Im}(f)$;
 - * Sinon, trouver une famille génératrice de E (par exemple une base) et considérer son image.
- Pour déterminer le noyau d'une AL f :
 - * Écrire l'équation $f(x) = 0$.
 - * Si nécessaire, décomposer x dans une base \mathcal{B} de E . Grâce à la linéarité de f , l'équation précédente se ramène alors à un système linéaire d'équations portant sur les coordonnées de x dans la base \mathcal{B} .

Conformément à la terminologie générale, nous définissons :

Définition 21.1.18 (Isomorphisme)

1. Une application linéaire bijective de E vers F est appelée un *isomorphisme*.
2. On dit que deux espaces vectoriels E et F sont isomorphes s'il existe un isomorphisme $f : E \rightarrow F$.

Les résultats généraux sur les structures amènent directement :

Théorème 21.1.19 (Réciproque d'un isomorphisme)

Soit f un isomorphisme entre E et F . Alors f^{-1} est une application linéaire, et donc un isomorphisme de F vers E .

I.3 Endomorphismes

Conformément à la terminologie générale, un endomorphisme est une application linéaire d'un espace dans lui-même :

Définition 21.1.20 (Endomorphisme)

Soit E un espace vectoriel sur \mathbb{K} . Une application linéaire de E dans E est appelée *endomorphisme de E* . On note $\mathcal{L}(E)$ l'ensemble $\mathcal{L}(E, E)$ des endomorphismes de E .

Exemples 21.1.21

1. L'identité $\text{Id}_E : x \mapsto x$ de E dans E .
2. L'homothétie (vectorielle) de rapport $\lambda : \lambda \text{Id}_E : x \mapsto \lambda x$, de E dans E ($\lambda \in \mathbb{K}$).
3. La dérivation formelle $D : \mathbb{R}[X] \rightarrow \mathbb{R}[X]$
4. La dérivation analytique $D : \mathcal{C}^2([a, b]) \rightarrow \mathcal{C}^2([a, b])$
5. L'application matricielle $X \mapsto MX$ si M est...

La bilinéarité de la composition des applications linéaires permet de définir sur $\mathcal{L}(E)$ une loi de composition \circ , distributive sur la somme. On obtient alors la structure de l'ensemble $\mathcal{L}(E)$ des endomorphismes de E :

Proposition 21.1.22 (Structure de $\mathcal{L}(E)$)

L'ensemble $(\mathcal{L}(E), +, \circ)$ est muni d'une structure de \mathbb{K} -algèbre.

Nous rappelons ci-dessous la définition d'une \mathbb{K} -algèbre :

Définition 21.1.23 (\mathbb{K} -algèbre)

Étant donné un corps \mathbb{K} , une \mathbb{K} -algèbre est un espace vectoriel A sur \mathbb{K} , muni d'une seconde loi de composition interne \times , compatible avec la loi externe dans le sens suivant :

$$\forall \lambda \in \mathbb{K}, \quad \forall (x, y) \in A^2, \quad \lambda \cdot (x \times y) = (\lambda \cdot x) \times y = x \times (\lambda \cdot y),$$

et telle que $(A, +, \times)$ soit un anneau.

La composition ayant les propriétés usuelles d'un produit, on utilise souvent les conventions de notation usuelles pour les produits. En particulier la notation vu désigne l'endomorphisme $v \circ u$ (omission du signe opératoire), et :

Notation 21.1.24 (Composition itérée)

Étant donné un endomorphisme u de E , et un entier $n \in \mathbb{N}$, on désigne par u^n la n -ième composée itérée de u , définie récursivement par $u^0 = \text{Id}_E$ et $u^n = u \circ u^{n-1}$ pour $n \in \mathbb{N}^*$.

Remarquez que ces conventions de notation ne sont pas gênantes dans la mesure où dans l'espace vectoriel E , on ne dispose pas d'un produit. L'expression $v(x)u(x)$ n'ayant pas de sens, la notation vu ne peut désigner que la composition. De même pour f^n , l'expression $f(x)^n$ n'ayant pas de sens.

Remarque 21.1.25

1. L'anneau $\mathcal{L}(E)$ est-il commutatif? À quelle condition nécessaire et suffisante sur la dimension de E l'est-il?

2. L'anneau $\mathcal{L}(E)$ est-il intègre ?

Cette dernière remarque amène la définition suivante :

Définition 21.1.26 (Endomorphisme nilpotent)

Un endomorphisme nilpotent de E est un endomorphisme u tel qu'il existe $n \in \mathbb{N}$ tel que $u^n = 0_{\mathcal{L}(E)}$. La valeur minimale de n vérifiant cette propriété est appelée *indice de nilpotence de u* .

Exemple 21.1.27

$$1. X \mapsto \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} X$$

2. La dérivation, vue comme endomorphisme de ...

Ainsi, d'une certaine façon, si u est nilpotent d'indice n , l'endomorphisme u « annule » le polynôme X^n . On dira alors que X^n est polynôme annulateur de u . un autre polynôme annulateur est X^{n+1} , ou encore $X^n(X-1)$. La valeur de n étant minimale, X^n est le polynôme unitaire de plus bas degré annihilant u : on dira dans ce cas qu'il s'agit du polynôme minimal de u . Nous généralisons ci-dessous ces notions pour des endomorphismes quelconques.

Définition 21.1.28 (Polynôme d'endomorphisme)

Soit $P = \sum_{k=0}^n a_k X^k$ un polynôme de $\mathbb{K}[X]$ et u un endomorphisme d'un \mathbb{K} -ev E . On définit le polynôme $P(u)$ de l'endomorphisme u par :

$$P(u) = \sum_{k=0}^n a_k u^k : x \mapsto \sum_{k=0}^n a_k u^k(x)$$

La structure d'algèbre de $\mathcal{L}(E)$ nous permet d'affirmer que $P(u) \in \mathcal{L}(E)$.

Remarquez qu'il s'agit d'un cas particulier de spécialisation d'un polynôme à des éléments d'une algèbre, ainsi que nous l'avons vu dans le chapitre sur les polynômes.

Deux polynômes d'un même endomorphisme commutent, ce qui découle de la commutativité dans $\mathbb{K}[X]$, via le lemme suivant :

Lemme 21.1.29 (Spécialisation d'un produit)

Soit u un endomorphisme de E et P et Q deux polynômes de $\mathbb{K}[X]$. Alors $(PQ)(u) = P(u) \circ Q(u)$.

Proposition 21.1.30 (Commutation des polynômes d'endomorphisme)

Soit u un endomorphisme de E et P et Q deux polynômes de $\mathbb{K}[X]$. Alors les deux endomorphismes $P(u)$ et $Q(u)$ commutent :

$$P(u) \circ Q(u) = Q(u) \circ P(u).$$

Cela permet en particulier d'exploiter une factorisation d'un polynôme P pour le calcul de $P(u)$.

Exemple 21.1.31

D désignant l'endomorphisme de dérivation de $\mathcal{C}^\infty(\mathbb{R})$, et P désignant le polynôme $P = (X - 1)(X - 2)(X - 3) = X^3 - 6X^2 + 11X - 6$, on obtient, pour une fonction f de classe \mathcal{C}^∞ :

$$f^{(3)} - 6f^{(2)} + 11f' - 6f = (D - \text{Id}) \circ (D - 2\text{Id}) \circ (D - 3\text{Id})(f),$$

ce qu'on vérifie aisément de façon directe (Id désigne l'identité de $\mathcal{C}^\infty(\mathbb{R})$).

Nous pouvons alors généraliser les notions entrevues lors de l'étude des endomorphismes nilpotents.

Définition 21.1.32 (Polynôme annulateur)

On dit que $P \in \mathbb{K}[X]$ est un polynôme annulateur de $u \in \mathcal{L}(E)$ si $P(u) = 0_{\mathcal{L}(E)}$.

Proposition 21.1.33 (Structure de l'ensemble des polynômes annulateurs)

L'ensemble des polynômes annulateurs de E forme un idéal de $\mathbb{K}[X]$.

L'anneau $\mathbb{K}[X]$ étant principal, on peut alors définir :

Définition 21.1.34 (Polynôme minimal)

Si l'endomorphisme u admet au moins un polynôme annulateur non nul, on définit le polynôme minimal de u comme étant l'unique polynôme unitaire engendrant l'idéal des polynômes annulateurs de u .

Nous verrons plus loin que si E est de dimension finie n , alors $\mathcal{L}(E)$ est également de dimension finie, égale à n^2 . Cela se comprend facilement une fois assimilée la correspondance entre applications linéaires et matrices. Admettant provisoirement ce résultat, nous obtenons :

Proposition 21.1.35 (Existence d'un polynôme annulateur)

Soit u un endomorphisme d'un espace E de dimension finie n . Alors u admet un polynôme annulateur non nul, donc un polynôme minimal. De plus, le degré du polynôme minimal est au plus n^2 .

La théorie de la réduction des endomorphismes (et en particulier le théorème de Cayley-Hamilton que vous verrez l'an prochain) permet d'établir que le polynôme minimal est de degré au plus n . Le théorème de Cayley-Hamilton permet de déterminer de façon explicite un polynôme annulateur d'un endomorphisme dont on connaît une représentation matricielle dans une base, à l'aide des déterminants.

I.4 Automorphisme

Définition 21.1.36 (Automorphisme, $\text{GL}(E)$)

Un automorphisme de E est un endomorphisme bijectif, donc une application linéaire qui est à la fois un endomorphisme et un isomorphisme. On appelle *groupe linéaire de E* , et on note $\text{GL}(E)$, l'ensemble des automorphismes de E .

La terminologie est justifiée par le théorème suivant :

Théorème 21.1.37 (Structure de $\text{GL}(E)$)

L'ensemble $\text{GL}(E)$ muni de la composition des endomorphismes, est un groupe.

Comme on le verra, dans le chapitre suivant, si E est de dimension n , il existe un isomorphisme de groupe entre $\text{GL}(E)$ et le groupe $\text{GL}_n(\mathbb{K})$ des matrices inversibles de taille $n \times n$ à coefficients dans \mathbb{K} .

I.5 Projecteurs et symétries

Définition 21.1.38 (Projecteur, symétrie)

1. Soit $p \in \mathcal{L}(E)$. On dit que p est un projecteur (ou une projection) de E ssi $p \circ p = p$.
2. Soit $s \in \mathcal{L}(E)$. On dit que s est une symétrie ssi $s \circ s = \text{id}$.

Ainsi, un projecteur est par définition un endomorphisme dont $X^2 - X$ est polynôme annulateur, et une symétrie est un endomorphisme dont $X^2 - 1$ est polynôme annulateur.

Proposition 21.1.39 (Caractérisation de l'image d'un projecteur)

Soit p un projecteur de E . Alors $x \in \text{Im}(p)$ si et seulement si $p(x) = x$, ce qui se traduit par l'égalité ensembliste :

$$\text{Im}(p) = \text{Ker}(p - \text{Id}_E).$$

Théorème 21.1.40 (Diagonalisation d'un projecteur)

Soit p un projecteur de E . Alors :

$$E = \text{Ker}(p) \oplus \text{Ker}(p - \text{Id}_E).$$

Cette propriété exprime le fait que p est « diagonalisable ».

De façon générale, dire qu'un endomorphisme f est diagonalisable signifie que l'espace E peut être décomposé en somme directe d'espaces stables par f sur lesquels p induit une homothétie. Ainsi, f est entièrement déterminé par un ensemble de vecteurs engendrant E et sur lesquels f est simplement une dilatation. En termes plus précis :

Définition 21.1.41 (Endomorphisme diagonalisable)

- Un endomorphisme f de E est diagonalisable s'il existe une base $(b_i)_{i \in I}$ de E et une famille $(\lambda_i)_{i \in I}$ de scalaires tels que pour tout $i \in I$, $f(b_i) = \lambda_i b_i$.
- Les λ_i sont appelées valeurs propres de f .
- Étant donné une valeur propre λ , un vecteur x non nul tel que $f(x) = \lambda x$ est appelé vecteur propre associé à λ .
- Si λ est une valeur propre de f , $\text{Ker}(f - \lambda \text{Id})$ est appelé sous-espace propre de f associé à la valeur propre λ .

Ainsi, dire que f est diagonalisable revient à dire que E est somme directe des sous-espaces propres de f . Vous verrez l'année prochaine que cela équivaut en dimension finie à l'existence d'une base (par exemple la base donnée dans l'énoncé de la définition) relativement à laquelle la matrice de f est diagonale (voir chapitre suivant pour la notion de matrice associée à un endomorphisme)

Corollaire 21.1.42

Un projecteur distinct de 0 ou Id , admet exactement deux valeurs propres 0 et 1, et est diagonalisable.

Remarque 21.1.43

Les valeurs propres de u sont les racines du polynôme annulateur $X^2 - X$. Ce n'est pas anodin. L'ensemble des valeurs propres est toujours inclus dans l'ensemble des racines d'un polynôme annulateur, et on a l'égalité s'il s'agit du polynôme minimal.

Le théorème précédent est en fait une caractérisation des projecteurs

Théorème 21.1.44 (Caractérisation géométrique des projecteurs)

Soit $p \in \mathcal{L}(E)$.

- Alors p est un projecteur si et seulement s'il existe deux sev F et G de E tels que $F \oplus G = E$, et :

$$\forall u \in F, \forall v \in G, p(u + v) = u.$$

Cette dernière identité traduit le fait que p est la projection géométrique sur F parallèlement à G .

- Dans ce cas, on a $F = \text{Im}(p)$ et $G = \text{Ker}(p)$.
- Ainsi, un projecteur est une projection géométrique sur $\text{Im}(p)$ parallèlement à $\text{Ker}(p)$.

Le même travail peut être fait pour les symétries.

Théorème 21.1.45 (Diagonalisation d'une symétrie)

Soit s une symétrie de E . Alors :

$$E = \text{Ker}(s + \text{Id}_E) \oplus \text{Ker}(s - \text{Id}_E).$$

Ce dernier résultat traduit le fait que s est diagonalisable, et si s est distinct de Id_E et $-\text{Id}_E$, alors les valeurs propres de s sont exactement 1 et -1 . On peut à nouveau remarquer qu'il s'agit des racines du polynôme annulateur $X^2 - 1$.

Encore une fois, le théorème précédent est une caractérisation des symétries, et donne l'interprétation géométrique des symétries.

Théorème 21.1.46 (Caractérisation géométrique des symétries)

Soit $s \in \mathcal{L}(E)$.

- Alors s est une symétrie si et seulement s'il existe deux sev F et G de E tels que $F \oplus G = E$, et :

$$\forall u \in F, \forall v \in G, s(u + v) = u - v.$$

Cette dernière identité traduit le fait que s est la symétrie géométrique par rapport à F parallèlement à G .

- Dans ce cas, on a $F = \text{Im}(s) = \text{Ker}(s - \text{Id}_E)$ et $G = \text{Ker}(s + \text{Id}_E)$.
- Ainsi, une symétrie au sens algébrique s est une symétrie géométrique par rapport à $\text{Ker}(s - \text{Id}_E)$ (l'ensemble des points fixes), parallèlement à $\text{Ker}(s + \text{Id}_E)$.

II Applications linéaires et familles de vecteurs

II.1 Détermination d'une application linéaire

Nous commençons par un résultat de rigidité, exprimant le fait que l'image d'un nombre limité de vecteurs de E par une application linéaire u permet de déterminer entièrement l'application linéaire u . En effet, par linéarité, la connaissance de u sur une famille de vecteur amène sa connaissance sur tout l'espace engendré par cette famille. On obtient donc le résultat suivant :

Proposition 21.2.1 (Détermination d'une AL par l'image d'une base)

Étant donné $(b_i)_{i \in I}$ une base de E et $(f_i)_{i \in I}$ une famille quelconque de F , il existe une unique application linéaire $u \in \mathcal{L}(E, F)$ telle que pour tout $i \in I$, $f(b_i) = f_i$.

Ainsi, une application linéaire de E dans F est entièrement déterminée par l'image d'une base de E .

Exemples 21.2.2

1. Déterminer l'expression générale de l'application linéaire de \mathbb{R}^2 dans \mathbb{R}^2 telle que $f(1, 0) = (3, 2)$ et $f(0, 1) = (2, 1)$.
2. Montrer que toute application linéaire de \mathbb{R}^p dans \mathbb{R}^n est de la forme $X \mapsto MX$, et décrire M à partir d'une base de \mathbb{R}^p .
3. Soit $(b_i)_{i \in I}$ une base de E et $(c_j)_{j \in J}$ une base de F . Alors pour tout $(i, j) \in I \times J$, il existe une unique application linéaire $u_{i,j}$ telle que $u_{i,j}(b_i) = c_j$ et pour tout $k \neq i$, $u_{i,j}(b_k) = 0$.

Proposition 21.2.3 (Base de $\mathcal{L}(E, F)$)

Si E est de dimension finie, la famille $(u_{i,j})_{(i,j) \in I \times J}$ décrite dans l'exemple ci-dessus est une base de $\mathcal{L}(E, F)$.

Corollaire 21.2.4 (Dimension de $\mathcal{L}(E, F)$)

Si E et F sont de dimension finie, alors :

$$\dim \mathcal{L}(E, F) = \dim(E) \times \dim F.$$

Remarques 21.2.5

1. Pensez au cas des applications linéaires de \mathbb{K}^p dans \mathbb{K}^n , qui sont obtenus, comme on l'a vu plus haut, sous la forme $u_M : X \mapsto MX$, avec $M \in \text{Mat}_{n,p}(\mathbb{K})$. Il n'est pas dur de voir que $M \mapsto u_M$ est un isomorphisme d'espace vectoriel entre $\mathcal{M}_{n,p}(\mathbb{K})$ et $\mathcal{L}(\mathbb{K}^p, \mathbb{K}^n)$. Or, une matrice étant la donnée de $n \times p$ coefficients indépendants, la dimension de $\mathcal{M}_{n,p}(\mathbb{K})$ est np . On retrouve sur cet exemple le résultat précédent.
2. Ce n'est d'ailleurs pas qu'un exemple, car comme on le verra dans le paragraphe suivant, tout \mathbb{K} -espace vectoriel E de dimension finie est isomorphe à \mathbb{K}^n , où $n = \dim(E)$. Via cet isomorphisme, la situation décrite ci-dessus est générique.
3. En partant des bases canoniques de \mathbb{K}^p et de \mathbb{K}^n , la base décrite ci-dessus de $\mathcal{L}(\mathbb{K}^p, \mathbb{K}^n)$ correspond à la base de $\mathcal{M}_{n,p}(\mathbb{K})$ formée des matrices $E_{i,j}$, continuées de coefficients tous nuls, sauf le coefficient en position (i, j) , égal à 1. Il s'agit de la *base canonique de $\mathcal{M}_{n,p}(\mathbb{K})$* .
4. L'espace vectoriel $\mathcal{M}_{n,p}(\mathbb{K})$ est isomorphe à \mathbb{K}^{np} par l'isomorphisme consistant à réordonner les coefficients en les écrivant en une seule ligne (ou en une seule colonne), en juxtaposant les différentes lignes les unes à la suite des autres. Via cet isomorphisme, quitte à réordonner les éléments de la base, la base canonique de $\mathcal{M}_{n,p}(\mathbb{K})$ correspond à la base canonique de \mathbb{K}^{np} .

II.2 Caractérisations de l'injectivité et de la surjectivité par l'image de bases

L'image d'une base par un endomorphisme déterminant entièrement l'application linéaire, toutes les propriétés telles que l'injectivité et la surjectivité peuvent se voir déjà dans la description de l'image d'une base. Ainsi, ces propriétés peuvent être caractérisées par l'image d'une base.

Proposition 21.2.6 (Caractérisation de l'injectivité par l'image d'une base)

Soit $f \in \mathcal{L}(E, F)$. Les propriétés suivantes sont équivalentes :

- (i) f est injective ;

- (ii) l'image de toute famille libre de E par f est une famille libre de F ;
- (iii) l'image de toute base de E par f est une famille libre de F ;
- (iv) il existe une base de E dont l'image par f est une famille libre de F .

Proposition 21.2.7 (Caractérisation de la surjectivité par l'image d'une base)

Soit $f \in \mathcal{L}(E, F)$. Les propriétés suivantes sont équivalentes :

- (i) f est surjective ;
- (ii) l'image de toute famille génératrice de E par f est une famille génératrice de F ;
- (iii) l'image de toute base de E par f est une famille génératrice de F ;
- (iv) il existe une base de E dont l'image par f est une famille génératrice de F .

En combinant ces deux caractérisations, nous obtenons :

Proposition 21.2.8 (Caractérisation de la bijectivité par l'image d'une base)

Soit $f \in \mathcal{L}(E, F)$. Les propriétés suivantes sont équivalentes :

- (i) f est un isomorphisme ;
- (ii) l'image de toute base de E par f est une base de F ;
- (iii) il existe une base de S dont l'image par f est une base de F .

On en déduit en particulier :

Corollaire 21.2.9 (Dimension d'espaces isomorphes)

Soit E et F deux espaces isomorphes. Alors si l'un des deux espaces E ou F est de dimension finie, les deux le sont, et $\dim(E) = \dim(F)$.

Corollaire 21.2.10 (Classification à isomorphisme près des espaces de dimension finie)

- (i) Tout \mathbb{K} -ev E de dimension finie n est isomorphe à \mathbb{K}^n .
- (ii) Si $n \neq m$, \mathbb{K}^n et \mathbb{K}^m ne sont pas isomorphes

Plus précisément, un isomorphisme tel que dans (i) est obtenu par le choix d'une base de E , et correspond alors à la donnée des coordonnées d'un vecteur x dans cette base.

Le dernier résultat peut se réexprimer en remarquant que la relation d'isomorphisme définit une relation d'équivalence, notée \simeq , sur l'ensemble \mathcal{E}_f des \mathbb{K} -ev de dimension finie. L'espace quotient est alors :

$$(\mathcal{E}_f / \simeq) = \mathbb{N}.$$

II.3 Recollements

Une généralisation de la détermination d'une application linéaire par l'image d'une base est la possibilité de définir, et ceci de façon unique, une application linéaire à partir de ses restrictions sur une famille $(E_i)_{i \in I}$ de sous-espaces vectoriels tels que $(E_i)_{i \in I}$. Le cas de la détermination à partir de l'image d'une base en découle en considérant la famille $(\text{Vect}(b_i))_{i \in I}$. On énonce :

Théorème 21.2.11

Soit I un ensemble fini. Soit $(E_i)_{i \in I}$ une famille de sous-espaces vectoriels de E tels que $\bigoplus_{i \in I} E_i = E$, et soit pour tout $i \in I$, $u_i \in \mathcal{L}(E_i, F)$. Alors il existe une et une seule application linéaire $u \in \mathcal{L}(E, F)$ telle que pour tout $i \in I$, $u_i = u|_{E_i}$.

Si, sous les mêmes conditions sur les E_i , les u_i sont des endomorphismes des E_i , on peut faire la même chose en considérant les applications $v_i = j_i \circ u_i$, où j_i est l'inclusion de E_i dans E . L'application u obtenue est alors un endomorphisme de E .

En revanche, contrairement au cas précédent où toute application linéaire de $\mathcal{L}(E, F)$ peut être obtenue comme recollement d'applications linéaires de $\mathcal{L}(E_i, F)$, tout endomorphisme de E ne peut pas être obtenu comme recollement d'endomorphismes des E_i (le problème provenant de la corestriction faite sur l'espace image). Matriciellement, si E est de dimension finie, après choix d'une base de E obtenue par recollement de bases des E_i , un endomorphisme pouvant s'obtenir ainsi est un endomorphisme dont la matrice sera diagonale par blocs (voir chapitre suivant)

Nous donnons une condition pour que ce soit le cas :

Proposition 21.2.12 (Restriction d'un endomorphisme)

Soit E un \mathbb{K} -ev et F un sous-espace vectoriel de E . Soit $u \in \mathcal{L}(E)$. Si F est stable par u (c'est-à-dire si $u(F) \subset F$, ou encore si pour tout $x \in F$, $u(x) \in F$) alors u induit par restriction et corestriction un endomorphisme u_F de F .

Proposition 21.2.13 (Décomposition d'un endomorphisme sur une somme de sev stables)

Soit E un \mathbb{K} -ev et $(E_i)_{i \in I}$ des sev en nombre fini, tels que $E = \bigoplus_{i \in I} E_i$. Soit $u \in \mathcal{L}(E)$. Si les E_i sont tous stables par u , et si u_i désigne l'endomorphisme de $\mathcal{L}(E_i)$ induit par u , alors u est le recollement des endomorphismes u_i dans le sens évoqué ci-dessus.

III Applications linéaires en dimension finie

III.1 Rang d'une application linéaire

Définition 21.3.1 (Rang d'une application linéaire)

Soit $u \in \mathcal{L}(E, F)$ une application linéaire. Si $\text{Im}(u)$ est de dimension finie, on définit le rang de u par :

$$\text{rg}(u) = \dim(\text{Im}(u)).$$

Proposition 21.3.2

Soit $(x_i)_{i \in I}$ une famille génératrice de E . Le rang de $\text{Im}(u)$, s'il existe, est égal au rang de la famille $(u(x_i))_{i \in I}$.

Proposition 21.3.3 (Existence du rang en dimension finie)

- Soit $u \in \mathcal{L}(E, F)$. Si E et/ou F sont de dimension finie alors $\text{Im}(u)$ également, et

$$\text{rg}(u) \leq \dim(E) \quad \text{et/ou} \quad \text{rg}(u) \leq \dim(F).$$

- Sous les conditions idoines d'existence :

- * $\text{rg}(u) = \dim(E)$ si et seulement si u est injective ;
- * $\text{rg}(u) = \dim(F)$ si et seulement si u est surjective.

On en déduit en particulier :

Théorème 21.3.4 (Caractérisation des isomorphismes en dimension finie)

Soit E et F deux espaces vectoriels de même dimension finie n , et soit $f \in \mathcal{L}(E, F)$. Alors les propositions suivantes sont équivalentes :

- (i) f est un isomorphisme ;
- (ii) $\text{rg}(f) = n$;
- (iii) f est injective ;
- (iv) f est surjective.

En particulier, si E est de dimension finie, un endomorphisme $f \in \mathcal{L}(E)$ est un automorphisme si et seulement si f est injective si et seulement si f est surjective.

Théorème 21.3.5 (Effet d'une composition sur le rang)

Soit $u \in \mathcal{L}(E, F)$ et $v \in \mathcal{L}(F, G)$.

1. $\text{rg}(v \circ u) \leq \min(\text{rg}(u), \text{rg}(v))$.
2. Si v est injective, $\text{rg}(v \circ u) = \text{rg}(u)$.
3. Si u est surjective, $\text{rg}(v \circ u) = \text{rg}(v)$.

En particulier :

Corollaire 21.3.6 (Invariance du rang par composition par un isomorphisme)

Tout est dit dans le titre, la composition pouvant se faire à droite ou à gauche.

III.2 Théorème du rang

Cette section courte, mais ô combien importante, a pour objet un résultat reliant la dimension de l'image et la dimension du noyau d'une application linéaire (théorème du rang). En gros, ce théorème dit que, partant du fait qu'il n'y a pas de perte de dimension entre l'espace initial et l'image lorsque u est injective, tout défaut d'injectivité se traduit par une perte sur l'image égale à la dimension du noyau : la dimension du noyau (donc la dimension d'un ensemble d'éléments tous envoyés sur le même point) correspond à la dimension qu'on perd entre l'espace initial et l'image.

Pour établir ce résultat, nous commençons par étudier les propriétés relatives au noyau et à l'image de restrictions.

Lemme 21.3.7 (Noyau et image d'une restriction)

Soit $u \in \mathcal{L}(E, F)$, et E' un sous-espace de E . Soit $v \in \mathcal{L}(E', F)$ la restriction de u à E' . Alors :

- $\text{Ker}(v) = \text{Ker}(u) \cap E'$
- Si $\text{Ker}(u) + E' = E$, $\text{Im}(v) = \text{Im}(u)$.

Corollaire 21.3.8 (Restriction de u à un supplémentaire de $\text{Ker}(u)$)

Soit S un supplémentaire de $\text{Ker}(u)$ dans E . Alors u induit un isomorphisme de S sur $\text{Im}(u)$.

Remarquez que ce dernier résultat ne nécessite pas d'hypothèse de finitude. Cependant, si E n'est pas de dimension finie, il est nécessaire de supposer l'axiome du choix, afin d'assurer l'existence du supplémentaire S .

On déduit du corollaire précédent le très important :

Théorème 21.3.9 (Théorème du rang)

Soit E un espace vectoriel de dimension finie, et F un espace vectoriel quelconque. Soit $f \in \mathcal{L}(E, F)$. Alors :

$$\dim \text{Ker } f + \text{rg } f = \dim E.$$

IV Formes linéaires

Nous terminons ce chapitre par l'étude d'une famille importante d'applications linéaires : les formes linéaires, qui correspondent aux applications linéaires à valeurs dans le corps de base. Les formes linéaires sont intimement liées à la notion d'hyperplan (généralisant la notion de plan en dimension 3). Elles sont aussi à la base des théories de dualité (qui ne sont pas au programme).

IV.1 Formes linéaires, espace dual, hyperplan

Définition 21.4.1 (Forme linéaire)

Une forme linéaire sur un \mathbb{K} -espace vectoriel E est une application linéaire de E vers \mathbb{K} , donc un élément de $\mathcal{L}(E, \mathbb{K})$.

Exemples 21.4.2

1. $f \mapsto \int_a^b f(t) dt$ sur ...
2. la trace sur $\mathcal{M}_n(\mathbb{R})$.
3. L'évaluation des polynômes en a .

Définition 21.4.3 (Dual)

Soit E un espace vectoriel. On appelle *dual* de E , et on note E^* , l'espace vectoriel $\mathcal{L}(E, \mathbb{K})$ constitué des formes linéaires. Le bidual E^{**} est alors le dual de E^* .

On définit alors les hyperplans de la manière suivante :

Proposition 21.4.4 (Hyperplan)

Soit H un sous-espace vectoriel de E . On dit que H est un hyperplan de E s'il existe une forme linéaire non nulle $\varphi \in E^*$ telle que $H = \text{Ker}(\varphi)$. L'équation $\varphi(x) = 0$, caractérisant l'appartenance à H , est appelée équation de H .

Proposition 21.4.5 (Caractérisation des hyperplans en dimension finie)

Si E est de dimension finie n , les hyperplans de E sont exactement les sous-espaces vectoriels de E de dimension $n - 1$.

En appelant *codimension* du sous-espace F de E la quantité $\text{codim}_E(F) = \dim(E) - \dim(F)$, les hyperplans sont donc les sous-espaces de E de codimension 1.

Exemples 21.4.6

1. Les plans vectoriels de \mathbb{R}^3 .
2. Les droites vectorielles de \mathbb{R}^2 .
3. Les polynômes sans terme constant.

Théorème 21.4.7 (Supplémentaire d'un hyperplan)

- Soit H un hyperplan d'un espace (non nécessairement de dimension finie). Alors un sous-espace vectoriel S est un supplémentaire de H si et seulement si S est une droite non contenue dans H .
- De façon symétrique, soit D une droite de E ; un sous-espace vectoriel S de E est un supplémentaire de E si et seulement si S est un hyperplan ne contenant pas D .

Malgré l'utilisation de supplémentaires en dimension non nécessairement finie, ce théorème ne nécessite pas l'axiome du choix.

Proposition 21.4.8 (Comparaison de deux équations de H)

Soit H un hyperplan de E , d'équation $\varphi \in E^*$. Alors pour tout $\psi \in E^*$, $\psi(x) = 0$ est une équation de H si et seulement si $\psi \neq 0$ et $\psi \in \text{Vect}(\varphi)$.

Théorème 21.4.9 (Intersection d'hyperplans)

Soit E un espace de dimension finie n .

1. L'intersection de m hyperplans de E est un sous-espace vectoriel de dimension au moins $n - m$.
2. Réciproquement, tout sous-espace vectoriel F de E de dimension $n - m$ peut s'écrire comme l'intersection de m hyperplans.

Remarquez qu'en dimension finie, après choix d'une base (b_1, \dots, b_n) , une forme linéaire sera décrite par une expression du type

$$\varphi(x) = \sum_{i=1}^n a_i x_i,$$

où les x_i sont les coordonnées de x dans la base (b_1, \dots, b_n) . Ainsi, l'équation d'un hyperplan est de la forme

$$a_1 x_1 + \dots + a_n x_n = 0.$$

On retrouve les équations usuelles d'une droite dans \mathbb{R}^2 ($ax + by = 0$) ou d'un plan dans \mathbb{R}^3 ($ax + by + cz = 0$), les x , y et z correspondant ici aux coordonnées dans la base canonique.

Ce que dit le dernier théorème est alors simplement le fait qu'un sous-espace de dimension $n - m$ (donc de codimension m) peut être décrit par un système de m équations de ce type.

IV.2 Qu'est-ce que le principe de dualité ? (hors-programme)

Nous introduisons ici, sans preuve, quelques notions relatives à la notion de dualité en algèbre linéaire. On suppose que E est de dimension finie. La plupart des notions peuvent se définir en dimension infinie, mais l'identification de E et E^{**} , à la base du principe de dualité, n'est valable qu'en dimension finie.

Soit $\varphi \in E^*$ et $x \in E$. On note $\langle x, \varphi \rangle = \varphi(x)$. Cela définit clairement une application bilinéaire de $E \times E^*$ dans \mathbb{K} . En particulier, en fixant x , on obtient une forme linéaire de E^* dans \mathbb{K} , donc un élément du bidual E^{**} :

Définition 21.4.10 (Application canonique)

L'application $J : E \mapsto E^{**}$ qui à x associe l'élément de E^{**} défini par $\varphi \mapsto \langle x, \varphi \rangle = \varphi(x)$ est appelée application canonique de E dans son bidual.

Théorème 21.4.11 (Identification de E et E^{})**

Si E est de dimension finie, J est un isomorphisme.

On peut alors identifier l'espace E et son bidual E^{**} . Le principe de dualité est alors que toute propriété de dualité démontrée pour les couples (E, E^*) reste vraie pour les couples (E^*, E) .

Il peut être intéressant, de ce fait, de transcrire les objets définis sur E en des objets analogues (leur correspondant par dualité) sur E^* . Nous commençons par les applications linéaires.

Définition 21.4.12 (Transposée d'une AL)

Soit $f \in \mathcal{L}(E, F)$. On appelle transposée de f , et on note ${}^t f$, l'application de $\mathcal{L}(F^*, E^*)$ définie par ${}^t f(\varphi) = \varphi \circ f$.

On peut vérifier que pour tout $x \in E$ et tout $\varphi \in E^*$, $\langle x, {}^t f(\varphi) \rangle = \langle f(x), \varphi \rangle$.

On peut également vérifier que, modulo l'isomorphisme J , on a ${}^t({}^t f) = f$.

La notation ${}^t f$ n'est pas anodine. Par la notion de base duale exposée ci-dessous, et en anticipant les représentations matricielles des applications linéaires (voir chapitre suivant), la matrice de ${}^t f$ dans la base duale est la transposée de la matrice de f dans la base initiale de E .

Cette base duale, qui est donc l'objet associé par dualité à une base de E se définit de la manière suivante. On suppose ici que E est de dimension finie, et on considère une base (e_1, \dots, e_n) de E .

Définition 21.4.13 (Base duale)

La base duale de (e_1, \dots, e_n) est la famille (e_1^*, \dots, e_n^*) d'éléments de E^* , où e_i^* est l'unique forme linéaire prenant la valeur 1 sur e_i et 0 sur les autres e_j . Il s'agit d'une base de E^*

Remarquez qu'il ne s'agit de rien d'autre que la base de $\mathcal{L}(E, F)$ décrite plus haut, dans le cas où $F = \mathbb{K}$, de base (1). Remarquez également que e_i^* est caractérisé par l'équation :

$$\langle e_j, e_i^* \rangle = \delta_{i,j},$$

où $\delta_{i,j}$ est le symbole de Kronecker, égal à 1 si $i = j$, et 0 sinon.

Là encore, il convient de remarquer que la notion définie est bien une notion duale, dans le sens où la base duale de la base duale s'identifie à la base initiale par l'isomorphisme J .

Enfin, nous définissons les version duales des sous-espaces vectoriels de E .

Définition 21.4.14 (Sous-espace orthogonal)

Soit A un sous-espace vectoriel de E , on appelle orthogonal de A dans E^* le sous-espace vectoriel $A^\circ = \{\varphi \in E^* \mid \langle x, \varphi \rangle = 0\}$.

L'orthogonalité inverse le sens des inclusions, et échange intersection et union. On vérifie une fois de plus que $A^{\circ\circ} = A$, via J , ce qui nous assure que la notion est la bonne.

Nous ne poussons pas plus l'étude de la dualité, qui pourrait nous mener fort loin, notre but étant uniquement de donner un premier aperçu d'une notion que vous développerez ultérieurement de façon plus approfondie.

Comme nous l'avons vu dans le chapitre précédent, la notion de matrice est indissociable de celle d'application linéaire. Nous avons déjà constaté au cours d'exemples que toute application linéaire de \mathbb{K}^p dans \mathbb{K}^n peut se représenter matriciellement sous la forme $X \mapsto MX$. Cela reste vrai pour toute application linéaire entre deux espaces vectoriels de dimension finie, à condition d'avoir fixé au préalable une base de chacun de ces espaces : les systèmes de coordonnées définis par ces bases nous ramènent alors au cas de $\mathcal{L}(\mathbb{K}^p, \mathbb{K}^n)$.

Note Historique 22.0.15

La notion de tableau de nombre en soi est vieille (étude de carrés magiques dans l'antiquité). Les règles opératoires, et le rapport avec les applications linéaires est plus récent. C'est Gauss le premier qui met le doigt sur ce rapprochement, à une époque où espaces vectoriels et applications linéaires n'avaient pas encore vu le jour. Le problème qui l'intéressait était de faire des changements de variable dans des formes quadratiques (polynômes de plusieurs variables de degré 2), de sorte à se ramener à des formes simples. Ces changements de variables portant sur plusieurs variables se présentent sous la forme de « substitutions linéaires » dans les termes de Gauss (c'est-à-dire d'applications linéaires). Pour alléger ses notations, il présente ses substitutions linéaires en n'en donnant que les coefficients, rangés dans un tableau... Ce n'est rien d'autre que la représentation matricielle d'une application linéaire. Il fait ensuite remarquer que si on enchaîne deux substitutions, on obtient une règle calculatoire permettant de trouver le tableau associé à la composée : il a découvert le produit matriciel.

Il est important de retenir de cette histoire que l'expression un peu compliquée du produit matriciel n'a pas été donnée au hasard : la définition du produit matriciel a été motivée historiquement par les règles de composition des applications linéaires.

I Matrice d'une application linéaire et opérations matricielles

Nous avons déjà vu que les applications linéaires entre \mathbb{K}^p et \mathbb{K}^n sont exactement les applications de la forme $f : X \mapsto MX$, où $M \in \mathcal{M}_{n,p}(\mathbb{K})$. La matrice M est alors la matrice constituée des colonnes C_i égales aux images des vecteurs de la base canonique de \mathbb{R}^p par f .

Le but de ce paragraphe est de généraliser cette description matricielle au cas d'applications linéaires $f \in \mathcal{L}(E, F)$ dans le cas où E et F sont de dimension finie et munis des bases \mathcal{B} et \mathcal{C} respectivement. Assez logiquement, on considérera alors la matrice dont les colonnes sont les coordonnées dans la base \mathcal{C} des images par f des vecteurs dans la base \mathcal{B} .

Nous verrons que cette description matricielle motive l'expression un peu compliquée du produit matriciel, de sorte que *via* cette correspondance, le produit matriciel correspond à la composition des applications linéaires.

Avant de pouvoir faire cette construction, nous définissons les différents objets qui interviennent.

I.1 L'ensemble des matrices de type (n, p)

Dans tout ce qui suit, \mathbb{K} désigne un corps, et n et p deux entiers naturels.

Définition 22.1.1 (Matrice)

Une matrice de taille $n \times p$ (n lignes et p colonnes) à coefficients dans \mathbb{K} est la donnée d'une famille $A = (a_{i,j})_{(i,j) \in \llbracket 1,n \rrbracket \times \llbracket 1,p \rrbracket}$ d'éléments de \mathbb{K} . On utilise la représentation planaire suivante :

$$A = \begin{pmatrix} a_{1,1} & \cdots & a_{1,p} \\ \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,p} \end{pmatrix}.$$

Définition 22.1.2 (Ensemble des matrices)

L'ensemble des matrices de taille $n \times p$ (on dit aussi *de type (n, p)*) à coefficients dans \mathbb{K} est noté $\mathcal{M}_{n,p}(\mathbb{K})$. Si $n = p$, on dit que la matrice est *carrée*, et on note simplement $\mathcal{M}_n(\mathbb{K})$ l'ensemble des matrices carrées de taille n .

Une matrice est souvent représentée sous forme d'un tableau, explicite dans le cas d'une matrice déterminée de petite taille, ou avec des « ... » dans le cas de matrice explicites ou non, de taille variable. Par exemple :

$$M_1 = \begin{pmatrix} 1 & 4 & 2 \\ 5 & 2 & 2 \\ 7 & 6 & 1 \end{pmatrix} \quad M_2 = (i + j - 1)_{1 \leq i, j \leq n} = \begin{pmatrix} 1 & \cdots & n \\ \vdots & & \vdots \\ n & \cdots & 2n - 1 \end{pmatrix}$$

$$M_3 = (a_{i,j})_{1 \leq i, j \leq n} = \begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,n} \end{pmatrix}.$$

Dans une telle représentation, le premier indice est toujours l'indice de ligne, et le second l'indice de colonne.

Terminologie 22.1.3 (Matrices colonnes, matrices lignes)

- Une matrice $X \in \mathcal{M}_{n,1}(\mathbb{K})$ est appelée matrice colonne, et est représentée par : $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$
- Une matrice $X \in \mathcal{M}_{1,n}(\mathbb{K})$ est appelée matrice ligne, et est représentée par : $X = (x_1 \cdots x_n)$
- On peut identifier \mathbb{K}^n indifféremment à $\mathcal{M}_{n,1}(\mathbb{K})$ ou $\mathcal{M}_{1,n}(\mathbb{K})$, suivant la situation.
- $\mathcal{M}_{1,1}(\mathbb{K})$ peut être identifié à \mathbb{K} (matrice constituée d'un unique coefficient)

Les puristes distinguent \mathbb{K}^n et $\mathcal{M}_{n,1}(\mathbb{K})$ ou $\mathcal{M}_{1,n}(\mathbb{K})$. D'ailleurs, dans la notation sous forme de n -uplet, les virgules permettent de distinguer un tel objet d'une matrice ligne. Sans une telle identification, on ne peut alors pas considérer aussi facilement l'application linéaire $X \mapsto MX$ comme une application de $\mathcal{L}(\mathbb{K}^p, \mathbb{K}^n)$.

Nous ne ferons pas une telle distinction dans la suite du cours.

Même si la plupart d'entre vous connaissez déjà les opérations sur les matrices, on considère au début de ce chapitre qu'on ne sait ni sommer ni multiplier des matrices. On montre dans la suite du chapitre comment la correspondance avec les applications linéaires permet de motiver la définition des opérations sur les matrices.

I.2 Matrice d'une application linéaire

Soit E et F deux espaces vectoriels de dimensions finies respectives p et n .

Définition 22.1.4 (Coordonnées d'un vecteur)

Soit $\mathcal{C} = (c_1, \dots, c_n)$ une base de F , et $X \in F$. Alors il existe d'unique scalaires x_1, \dots, x_n tels que $X = x_1c_1 + \dots + x_nc_n$. On définit la matrice colonne des coordonnées de X dans la base \mathcal{C} par :

$$[X]_{\mathcal{C}} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Définition 22.1.5 (Matrice associée à une famille)

Sous les mêmes hypothèses, soit (X_1, \dots, X_k) une famille de vecteurs de F . Alors la matrice de cette famille dans la base \mathcal{C} est :

$$[X_1, \dots, X_k]_{\mathcal{C}} = \left([X_1]_{\mathcal{C}} \mid \dots \mid [X_k]_{\mathcal{C}} \right).$$

Ainsi, il s'agit de la matrice dont la i -ème colonne comporte les coordonnées dans la base \mathcal{C} du vecteur X_i .

Définition 22.1.6 (Matrice d'une AL relativement à des bases)

Soit $f \in \mathcal{L}(E, F)$, et soit $\mathcal{B} = (b_1, \dots, b_p)$ une base de E et $\mathcal{C} = (c_1, \dots, c_n)$ une base de F . Alors la matrice de f relativement aux bases \mathcal{B} et \mathcal{C} est la matrice $\text{Mat}_{\mathcal{B}, \mathcal{C}}(f)$ de la famille $(f(b_1), \dots, f(b_p))$ dans la base \mathcal{C} :

$$\text{Mat}_{\mathcal{B}, \mathcal{C}}(f) = [f(b_1), \dots, f(b_p)]_{\mathcal{C}} = \left([f(b_1)]_{\mathcal{C}} \mid \dots \mid [f(b_p)]_{\mathcal{C}} \right).$$

Ainsi, il s'agit de la matrice de type (n, p) dont la i -ème colonne donne les coordonnées du vecteur $f(b_i)$ dans la base \mathcal{C} .

Remarque 22.1.7

Si on prend $E = \mathbb{K}^p$, $F = \mathbb{K}^n$, et pour \mathcal{B} et \mathcal{C} les bases canoniques de ces espaces, on retrouve la matrice $\text{Mat}_{b.c.}(f)$. On dit dans ce cas que $\text{Mat}_{b.c.}(f)$ est la *matrice canoniquement associée* à $f \in \mathcal{L}(\mathbb{K}^p, \mathbb{K}^n)$. Réciproquement, étant donné une matrice $M \in \mathcal{M}_{n,p}(\mathbb{K})$, l'unique application linéaire $f \in \mathcal{L}(\mathbb{K}^p, \mathbb{K}^n)$ dont la matrice relativement aux bases canoniques est M est appelée *application linéaire canoniquement associée* à la matrice M .

Exemple 22.1.8

- Soit $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ définie par $f(x) = 2x + 3y$. Déterminer $\text{Mat}_{\mathcal{B}, \mathcal{C}}(f)$ lorsque :
 - $\mathcal{B} = b.c., \mathcal{C} = (1)$
 - $\mathcal{B} = \left(\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \end{pmatrix} \right), \mathcal{C} = (2)$
- Soit $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ définie par $f(x, y) = (2x - y, x + 2y)$.
 - $\text{Mat}_{b.c.}(f)$?
 - $\text{Mat}_{\mathcal{B}, \mathcal{C}}(f)$ lorsque $\mathcal{B} = \mathcal{C} = \left(\begin{pmatrix} 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \end{pmatrix} \right)$
- Matrice de l'opérateur de dérivation $D \in \mathcal{L}(\mathbb{R}_n[X])$ relativement à la base canonique (au départ et à l'arrivée)

4. Trouver une base \mathcal{B} de $\mathbb{R}_n[X]$ telle que

$$\text{Mat}_{\mathcal{B},\mathcal{B}}(D) = J_{n+1} = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ \vdots & & & \ddots & 1 \\ 0 & \cdots & \cdots & \cdots & 0 \end{pmatrix}.$$

La matrice J_{n+1} est appelée matrice de Jordan d'ordre $n + 1$.

5. Plus généralement, soit E un espace de dimension n , et u un endomorphisme nilpotent, d'indice de nilpotence égal à n . Montrer qu'il existe une base relativement à laquelle la matrice de u est J_n .

I.3 Structure d'espace vectoriel de $\mathcal{M}_{n,p}(\mathbb{K})$

Une application linéaire étant entièrement déterminée par les images des vecteurs d'une base, l'application $u \mapsto \text{Mat}_{\mathcal{B},\mathcal{C}}(u)$ est une bijection de $\mathcal{M}_{n,p}(\mathbb{K})$.

Définition 22.1.9 (Transfert de structure)

Étant donnée une bijection $f : E \rightarrow F$, où E est muni d'une certaine structure algébrique, on peut munir F de lois telles que f soit un isomorphisme. Par exemple, si E est muni d'une loi interne $x \times y$, on munit F d'un produit

$$x' \times y' = f(f^{-1}(x') \times f^{-1}(y')).$$

Pour une loi externe, on définit :

$$\lambda \cdot x' = f(\lambda \cdot f^{-1}(x')).$$

La construction du transfert amène de façon immédiate :

Proposition 22.1.10 (Isomorphisme induit par un transfert)

Soit $f : E \rightarrow F$ une bijection, E étant muni d'une structure de truc. Alors, si on munit F de la structure de truc transférée de E par f , f est un isomorphisme de truc.

Ainsi, on peut munir $\mathcal{M}_{n,p}(\mathbb{K})$ de la structure d'espace vectoriel transférée de la structure de $\mathcal{L}(E, F)$ via la bijection $u \mapsto \text{Mat}_{\mathcal{B},\mathcal{C}}(u)$. Cette structure dépend *a priori* de E et F et du choix des bases sur cet espace. La proposition ci-dessous montre que tel n'est pas le cas :

Proposition 22.1.11 (Combinaisons linéaires des matrices)

La structure d'espace vectorielle sur $\mathcal{M}_{n,p}(\mathbb{K})$ définie par transfert est indépendante de E , F et leur base. Plus explicitement, elle est décrite par :

$$(a_{i,j})_{(i,j) \in \llbracket 1,n \rrbracket \times \llbracket 1,m \rrbracket} + \lambda (b_{i,j})_{(i,j) \in \llbracket 1,n \rrbracket \times \llbracket 1,m \rrbracket} = (a_{i,j} + \lambda b_{i,j})_{(i,j) \in \llbracket 1,n \rrbracket \times \llbracket 1,m \rrbracket}.$$

Ainsi, la somme de matrices et la multiplication par un scalaire se fait coefficient par coefficient.

On retrouver la structure usuelle qu'on a déjà eu l'occasion d'utiliser.

La construction même de cette structure nous permet d'affirmer :

Théorème 22.1.12 (Correspondance matrice-AL)

L'application $u \mapsto \mathcal{M}_{\mathcal{B}, \mathcal{C}}(u)$ est un isomorphisme d'espaces vectoriels de $\mathcal{L}(E, F)$ sur $\mathcal{M}_{n,p}(\mathbb{K})$.

Exemples 22.1.13

1. $\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} + \begin{pmatrix} 1 & 1 & 2 \\ 2 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 3 & 5 \\ 6 & 7 & 7 \end{pmatrix}$
2. $3 \cdot \begin{pmatrix} 1 & 2 & 1 \\ 2 & 2 & 3 \end{pmatrix} = \begin{pmatrix} 3 & 6 & 3 \\ 6 & 6 & 9 \end{pmatrix}$
3. $5 \times \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix} \dots ??? \dots$

Avertissement 22.1.14

On ne somme que des matrices de même taille (même nombre de lignes et de colonnes) !

On obtient facilement une base de $\mathcal{M}_{n,p}(\mathbb{K})$, qui n'est d'ailleurs que la base obtenue par l'isomorphisme ci-dessus de la base $(u_{i,j})$ de $\mathcal{L}(E, F)$ décrite dans le chapitre précédent.

Proposition/Définition 22.1.15 (Base canonique de $\mathcal{M}_{n,p}(\mathbb{K})$)

Soit pour tout $(i, j) \in \llbracket 1, n \rrbracket \times \llbracket 1, p \rrbracket$, $E_{i,j}$ la matrice constituée d'un coefficient 1 en position (i, j) et de 0 partout ailleurs. Alors $(E_{i,j})_{(i,j) \in \llbracket 1, n \rrbracket \times \llbracket 1, p \rrbracket}$ est une base de $\mathcal{M}_{n,p}(\mathbb{K})$, appelée base canonique de $\mathcal{M}_{n,p}(\mathbb{K})$.

Corollaire 22.1.16 (Dimension de $\mathcal{M}_{n,p}(\mathbb{K})$)

L'espace $\mathcal{M}_{n,p}(\mathbb{K})$ est de dimension np .

I.4 Définition du produit matriciel

Le cahier des charges pour la construction du produit matriciel est sensiblement similaire au cahier des charges pour la construction de la structure d'espace vectoriel. On souhaite que le produit matriciel corresponde à la composition des applications linéaires. Ainsi, si, relativement à certaines bases, la matrice de u est B , la matrice de v est A , on souhaite que la matrice de $v \circ u$ soit AB .

Comme l'espace d'arrivée de u doit être égal à l'espace de départ de v , cela impose des conditions sur les tailles des matrices :

Avertissement 22.1.17

On ne peut considérer le produit matriciel AB que si le nombre de colonnes de A est égal au nombre de lignes de B . Ainsi, on définit le produit matriciel de $\mathcal{M}_{n,p} \times \mathcal{M}_{p,q}$, à valeurs dans $\mathcal{M}_{n,q}$ (par examen des dimensions des espaces).

On définit alors produit matriciel par transfert, de façon dépendante *a priori* du choix d'espaces vectoriels et de bases.

Définition 22.1.18 (Définition du produit matriciel par transfert)

Soit $A \in \mathcal{M}_{n,p}(\mathbb{K})$ et $B \in \mathcal{M}_{p,q}(\mathbb{K})$. Soit E, F et G trois espaces de dimensions respectives q, p et n , munis de bases \mathcal{B}, \mathcal{C} et \mathcal{D} respectivement. Soit f_1 l'isomorphisme $u \in \mathcal{L}(E, F) \mapsto \text{Mat}_{\mathcal{B}, \mathcal{C}}(u)$, f_2 l'isomorphisme $v \in \mathcal{L}(F, G) \mapsto \text{Mat}_{\mathcal{C}, \mathcal{D}}(v)$ et f l'isomorphisme $w \in \mathcal{L}(E, G) \mapsto \text{Mat}_{\mathcal{B}, \mathcal{D}}(w)$.

On définit le produit AB par :

$$AB = f(f_2^{-1}(A) \circ f_1^{-1}(B)).$$

Proposition 22.1.19 (Description explicite du produit matriciel, par colonnes)

Soit $A \in \mathcal{M}_{n,p}(\mathbb{K})$ et $B = (b_{i,j}) \in \mathcal{M}_{p,q}(\mathbb{K})$. Notons C_1, \dots, C_p les colonnes de la matrice A . Alors la j -ième colonne C'_j du produit AB est :

$$C'_j = \sum_{i=1}^p b_{i,j} C_i.$$

Ainsi, la j -ième colonne du produit AB est obtenu en faisant la combinaison linéaire des colonnes de A par les coefficients de la j -ième colonne de B :

$$\left(\begin{array}{c|c|c} C_1 & \dots & C_p \end{array} \right) \times \left(\begin{array}{c|c|c} \dots & b_{1,j} & \dots \\ \dots & \vdots & \dots \\ \dots & b_{p,j} & \dots \end{array} \right) = \left(\begin{array}{c|c} \dots & b_{1,j}C_1 + \dots + b_{p,j}C_p \\ \dots & \dots \end{array} \right)$$

Corollaire 22.1.20 (Indépendance vis-à-vis des choix effectués)

Le produit matriciel défini par transfert ne dépend pas du choix des espaces vectoriels E, F et G , et des bases de ces espaces.

On en déduit en particulier, vu la construction donnée :

Proposition 22.1.21 (Matrice associée à une composition)

Soit E, F et G trois espaces vectoriels de dimension finie, munis respectivement des bases \mathcal{B}, \mathcal{C} et \mathcal{D} . Soit $f \in \mathcal{L}(E, F)$ et $g \in \mathcal{L}(F, G)$. Alors :

$$\text{Mat}_{\mathcal{B}, \mathcal{D}}(g \circ f) = \text{Mat}_{\mathcal{C}, \mathcal{D}}(g) \cdot \text{Mat}_{\mathcal{B}, \mathcal{C}}(f).$$

Cette formule est la clé de l'interprétation matricielle des applications linéaires. C'est en particulier elle qui fournit les formules de changement de base, en composant par l'identité, dont la matrice sera prise relativement à des bases différentes au départ et à l'arrivée.

Les propriétés de la composition se transfèrent alors immédiatement au produit matriciel :

Proposition 22.1.22 (Propriétés du produit matriciel)

Sous réserve de compatibilité des formats des matrices, le produit matriciel est associatif et bilinéaire (donc en particulier distributif)

Avertissement 22.1.23

Le produit matriciel n'est pas commutatif, même lorsque les tailles sont compatibles pour effectuer les opérations dans les deux sens (par exemple pour des matrices carrées).

Exemple 22.1.24

Comparer $\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ et $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$.

I.5 Expression matricielle de l'évaluation d'une AL**Proposition 22.1.25 (Produit d'une matrice par une matrice colonne)**

Soit $M \in \mathcal{M}_{n,p}(\mathbb{K})$ et $X \in \mathcal{M}_{p,1}(\mathbb{K})$.

On note $X = \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix}$ les coordonnées de X , et $M = (C_1 | \dots | C_m)$ la décomposition en colonnes de M . Le produit $M \cdot X$ est alors la matrice colonne de $\mathcal{M}_{n,1}(\mathbb{K})$ définie par :

$$M \cdot X = x_1 C_1 + \dots + x_m C_m = \sum_{i=1}^m x_i C_i.$$

Corollaire 22.1.26 (Compatibilité du produit matriciel avec l'évaluation)

Soit $f \in \mathcal{L}(E, F)$, et $X \in E$. Soit \mathcal{B} une base de E et \mathcal{C} une base de F . On a alors

$$[f(X)]_{\mathcal{C}} = \text{Mat}_{\mathcal{B},\mathcal{C}}(f) \times [X]_{\mathcal{B}}.$$

On peut aussi voir cette égalité comme un cas particulier de composition d'applications linéaires. En effet, en considérant $\mathcal{D} = (1)$ la base canonique de \mathbb{K} , la matrice colonne X correspond, relativement aux bases \mathcal{D} et \mathcal{B} , à l'application linéaire de \mathbb{K} dans E envoyant 1 sur X , donc l'application $g : \lambda \mapsto \lambda X$. En notant $M = \text{Mat}_{\mathcal{B},\mathcal{C}}(f)$, le produit MX correspond à l'application composée $f \circ g$, donc l'unique application qui à 1 associe $f(X)$. Sa représentation matricielle relativement aux bases \mathcal{D} et \mathcal{C} est donc la matrice colonne $[f(X)]_{\mathcal{C}}$. Le théorème de représentation matricielle d'une composée amène alors le résultat.

Ce résultat admet une réciproque :

Proposition 22.1.27 (Caractérisation de la matrice d'une application linéaire)

Soit $M \in \mathcal{M}_{n,p}(\mathbb{K})$, E et F deux espaces vectoriels de dimension finie, munis de bases \mathcal{B} et \mathcal{C} . Si pour tout $x \in E$, dont les coordonnées dans la base \mathcal{B} sont données par le vecteur colonne X , on a

$$[f(x)]_{\mathcal{C}} = MX,$$

alors $M = \text{Mat}_{\mathcal{B},\mathcal{C}}(f)$.

En particulier, en appliquant cette propriété à l'application nulle, on obtient :

$$(\forall X \in \mathcal{M}_{p,1}, \quad MX = 0) \iff M = 0$$

En revanche, il n'est bien sûr pas suffisant d'avoir $MX = 0$ pour une valeur de X pour pouvoir conclure que $M = 0$ (tout ce qu'on peut en déduire, c'est que X est dans le noyau de l'application linéaire canoniquement associée à M).

I.6 Transposition

On décrit dans ce paragraphe une autre construction, unaire, sur les matrices, définissant une application linéaire sur l'ensemble des matrices.

Définition 22.1.28 (Transposée d'une matrice)

Soit $A = (a_{i,j})_{(i,j) \in \llbracket 1,n \rrbracket \times \llbracket 1,p \rrbracket} \in \mathcal{M}_{n,p}(A)$. Alors la matrice transposée de A , notée tA , est la matrice de $\mathcal{M}_{p,n}\mathbb{K}$ définie par :

$${}^tA = (a_{j,i})_{(i,j) \in \llbracket 1,p \rrbracket \times \llbracket 1,n \rrbracket}.$$

Ainsi, si $A = \begin{pmatrix} a_{1,1} & \cdots & a_{1,p} \\ \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,p} \end{pmatrix}$, alors ${}^tA = \begin{pmatrix} a_{1,1} & \cdots & a_{n,1} \\ \vdots & & \vdots \\ a_{1,p} & \cdots & a_{n,p} \end{pmatrix}$

On trouve aussi assez souvent la notation A^T .

Exemple 22.1.29

Transposée de $\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$?

Proposition 22.1.30 (Linéarité de la transposition)

L'application $A \mapsto {}^tA$ de $\mathcal{M}_{n,p}(\mathbb{K})$ dans $\mathcal{M}_{p,n}(\mathbb{K})$ est une application linéaire.

I.7 Produit matriciel revisité

On explicite un peu plus les règles de produit matriciel. Dans un premier temps, on peut remarquer que l'expression du produit par une colonne donne, dans le cas où la matrice de gauche est une matrice ligne :

$$L \times C = \begin{pmatrix} a_1 & \cdots & a_p \end{pmatrix} \cdot \begin{pmatrix} b_1 \\ \vdots \\ b_p \end{pmatrix} = a_1 b_1 + \cdots + a_p b_p = \sum_{k=1}^p a_k b_k.$$

En notant $\langle X, Y \rangle$ l'application bilinéaire sur \mathbb{K}^p , définie par

$$\left\langle \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix}, \begin{pmatrix} y_1 \\ \vdots \\ y_p \end{pmatrix} \right\rangle = \sum_{k=1}^p x_k y_k,$$

on obtient :

$$L \times C = \left\langle \begin{pmatrix} a_1 \\ \vdots \\ a_p \end{pmatrix}, \begin{pmatrix} b_1 \\ \vdots \\ b_p \end{pmatrix} \right\rangle = \langle {}^tL, C \rangle.$$

Vous pouvez noter que si $\mathbb{K} = \mathbb{R}$, l'application bilinéaire $\langle \bullet, \bullet \rangle$ est le produit scalaire canonique de \mathbb{R}^p . En étudiant ligne par ligne l'expression du produit MX , on obtient alors la description :

Proposition 22.1.31 (Expression coefficient par coefficient du produit MX)

Soit $M = (a_{i,j}) \in \mathcal{M}_{n,p}(\mathbb{K})$ et $X = \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix} \in \mathcal{M}_{p,1}(\mathbb{K})$. En notant L_1, \dots, L_n les lignes de la matrice M , on a alors :

$$MX = \begin{pmatrix} L_1 \\ \vdots \\ L_n \end{pmatrix} X = \begin{pmatrix} L_1 \cdot X \\ \vdots \\ L_n \cdot X \end{pmatrix} = \begin{pmatrix} \langle {}^t L_1, X \rangle \\ \vdots \\ \langle {}^t L_n, X \rangle \end{pmatrix} = \begin{pmatrix} \sum_{k=1}^p a_{1,k} x_k \\ \vdots \\ \sum_{k=1}^p a_{n,k} x_k \end{pmatrix}$$

Les propositions 22.1.19 et 22.1.25 permettent alors d'obtenir une description coefficients par coefficients du produit matriciel.

Proposition 22.1.32 (Description coefficient par coefficient du produit)

Soit $A = (a_{i,j}) \in \mathcal{M}_{n,p}(\mathbb{K})$ et $B = (b_{j,k}) \in \mathcal{M}_{p,q}(\mathbb{K})$. Notons L_i les lignes de A et C_j les colonnes de B . Alors

$$AB = \left(AC_1 \mid \dots \mid AC_q \right) = \begin{pmatrix} L_1 \cdot C_1 & \dots & L_1 \cdot C_q \\ \vdots & L_i \cdot C_k & \vdots \\ L_n \cdot C_1 & \dots & L_n \cdot C_q \end{pmatrix} = \begin{pmatrix} \langle {}^t L_1, C_1 \rangle & \dots & \langle {}^t L_1, C_q \rangle \\ \vdots & \langle {}^t L_i, C_k \rangle & \vdots \\ \langle {}^t L_n, C_1 \rangle & \dots & \langle {}^t L_n, C_q \rangle \end{pmatrix}$$

En particulier, en notant $AB = (c_{i,k}) \in \mathcal{M}_{n,q}(\mathbb{K})$ la matrice produit, on obtient :

$$\forall i \in \llbracket 1, n \rrbracket, \forall k \in \llbracket 1, q \rrbracket, c_{i,k} = \sum_{j=1}^p a_{i,j} b_{j,k} = L_i \cdot C_k = \langle {}^t L_i, C_k \rangle,$$

où L_i est la i -ème ligne de A , et C_k la k -ième colonne de B_k .

Remarque 22.1.33

La description du produit par colonnes est parfois plus efficace que la description coefficient par coefficient, notamment lorsque la matrice de droite possède un grand nombre de 0. On obtient souvent une vision plus immédiate de la matrice, même si formellement le nombre de calcul est exactement le même. Par ailleurs, pour des arguments sur des matrices de taille générique n , la rédaction est souvent simplifiée par la description par les colonnes.

Exemple 22.1.34

1. Calculer $\begin{pmatrix} 2 & 6 & 12 \\ 2 & -7 & 2 \\ 8 & 5 & 11 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 2 \\ -1 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

2. Quel est l'effet de la multiplication à droite par $C_n = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & & & \ddots & 1 \\ 1 & 0 & \dots & \dots & 0 \end{pmatrix}$?

3. Même question avec la matrice « de Jordan » $J_n = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ \vdots & & & \ddots & 1 \\ 0 & \cdots & \cdots & \cdots & 0 \end{pmatrix}$
4. Calculer pour tout $k \in \mathbb{N}$, C_n^k et J_n^k . Que constatez-vous ?

La parfaite symétrie de l'expression obtenue suggère une version duale des descriptions 22.1.25 et 22.1.19 en inversant le rôle des lignes et des colonnes (cela provient de la description complètement symétrique donnée dans la proposition précédente) :

Proposition 22.1.35 (Expression du produit à l'aide des lignes)

1. Soit $A = (x_1 \ \dots \ x_n) \in \mathcal{M}_{1,n}(\mathbb{K})$ une matrice ligne, et $B = \begin{pmatrix} L_1 \\ \vdots \\ L_n \end{pmatrix} \in \mathcal{M}_{n,p}(\mathbb{K})$. Alors

$$AB = x_1 L_1 + \cdots + x_n L_n = \sum_{i=1}^n x_i L_i.$$

Il s'agit d'une matrice ligne.

2. Soit $A = \begin{pmatrix} A_1 \\ \vdots \\ A_n \end{pmatrix} \in \mathcal{M}_{m,n}(\mathbb{K})$, dont les lignes sont A_1, \dots, A_m , et $B \in \mathcal{M}_{n,p}(\mathbb{K})$. Alors

$$AB = \begin{pmatrix} A_1 B \\ \vdots \\ A_n B \end{pmatrix}.$$

La proposition 22.1.32 montre qu'on peut échanger le rôle de A et B , à condition d'invertir lignes et colonnes et d'échanger les indices, autrement dit de transposer les matrices. Ainsi, on a « commutativité du produit à transposition près » dans le sens suivant :

Corollaire 22.1.36 (Transposition d'un produit)

Soit $A \in \mathcal{M}_{n,p}(\mathbb{K})$ et $B \in \mathcal{M}_{p,q}(\mathbb{K})$. Alors

$${}^t(AB) = {}^tB {}^tA.$$

Nous terminons cette section par l'étude du produit des éléments de la base canonique.

Proposition 22.1.37 (Produit des éléments de la base canonique)

Soit $(E_{i,j})$ la base canonique de $\mathcal{M}_{n,p}(\mathbb{K})$, $(E'_{j,k})$ la base canonique de $\mathcal{M}_{p,q}(\mathbb{K})$ et $(E''_{i,k})$ la base canonique de $\mathcal{M}_{n,q}(\mathbb{K})$. Soit $(i,j) \in \llbracket 1, n \rrbracket \times \llbracket 1, p \rrbracket$, et $(k,\ell) \in \llbracket 1, p \rrbracket \times \llbracket 1, q \rrbracket$. Alors :

$$E_{i,j} \times E'_{k,\ell} = \delta_{j,k} E''_{i,\ell} = \begin{cases} E''_{i,\ell} & \text{si } j = k \\ 0 & \text{sinon.} \end{cases}$$

II Matrices carrées

II.1 L'algèbre $\mathcal{M}_n(\mathbb{K})$

On note de façon synthétique $\mathcal{M}_n(\mathbb{K})$ au lieu de $\mathcal{M}_{n,n}(\mathbb{K})$.

Définition 22.2.1 (Matrice d'un endomorphisme)

Si $u \in \mathcal{L}(E)$, où E est de dimension finie, on utilise souvent la même base au départ et à l'arrivée pour exprimer la matrice de u . Ainsi, si \mathcal{B} est une base de E , la matrice de u dans la base \mathcal{B} est la matrice $\text{Mat}_{\mathcal{B},\mathcal{B}}(u)$, et est plus simplement notée $\text{Mat}_{\mathcal{B}}(u)$.

La structure d'algèbre de $\mathcal{L}(\mathbb{K}^n)$ se transfère à l'ensemble des matrices carrées.

Théorème 22.2.2 (Structure de $\mathcal{M}_n(\mathbb{K})$)

L'ensemble $\mathcal{M}_n(\mathbb{K})$, muni de la somme, du produit et de la multiplication par les scalaires, est une \mathbb{K} -algèbre. Cette \mathbb{K} -algèbre est non commutative dès lors que $n \geq 2$.

Le neutre multiplicatif de cette algèbre est l'image par l'isomorphisme $\mathcal{L}(E) \rightarrow \mathcal{M}_n(\mathbb{K})$ de l'identité, à savoir la matrice identité notée I_n :

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

Il s'agit donc de la matrice constituée de 1 sur la diagonale et de 0 partout ailleurs.

Évidemment, on a de façon plus générale $I_n \times M = M$ et $N \times I_n = N$, même si M et N ne sont pas carrées (mais de format compatible).

En particulier le théorème précédent affirme que $\mathcal{M}_n(\mathbb{K})$ est un anneau, et toutes les règles de calcul que nous avons développées de façon générale dans les anneaux sont donc valables dans $\mathcal{M}_n(\mathbb{K})$, en particulier :

Théorème 22.2.3 (Factorisation de $A^n - B^n$)

Soit A et B deux éléments de $\mathcal{M}_n(\mathbb{K})$ tels que $AB = BA$. Alors pour tout $n \in \mathbb{N}^*$

$$A^n - B^n = (A - B) \sum_{k=0}^{n-1} A^{n-1-k} B^k.$$

Corollaire 22.2.4 (Factorisation de $I_n - A^n$)

En particulier, pour toute matrice carrée $A \in \mathcal{M}_n(\mathbb{K})$,

$$I_n - A^n = (I_n - A) \sum_{k=0}^{n-1} A^k.$$

Théorème 22.2.5 (Formule du binôme)

Soit A et B deux éléments de $\mathcal{M}_n(\mathbb{K})$ tels que $AB = BA$. Alors, pour tout $n \in \mathbb{N}$,

$$(A + B)^n = \sum_{k=0}^n \binom{n}{k} A^k B^{n-k}.$$

Exemple 22.2.6

1. Déterminer A^n , pour tout $n \in \mathbb{N}$, où $A = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$
2. Plus généralement, si J_n est définie comme dans l'exemple 22.1.34, calculer $(aI_n + J_n)^k$, pour tout $k \in \mathbb{N}$.

On en déduit une méthode assez efficace, mais pas toujours réalisable, de calcul des puissances.

Remarque 22.2.7

L'anneau $\mathcal{M}_n(\mathbb{K})$ est-il intègre ?

Les notions relatives aux polynômes d'endomorphisme se traduisent en terme de matrices. En particulier, toute matrice carrée d'ordre n admet un polynôme annulateur de degré inférieur ou égal à n^2 (et même inférieur ou égal à n). De là découle l'existence d'un polynôme minimal.

Exemple 22.2.8

1. Déterminer un polynôme annulateur de $\begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$.
2. Déterminer un polynôme annulateur de C_n
3. Déterminer un polynôme annulateur de J_n . Un polynôme minimal.

Le polynôme annulateur peut être efficace pour la recherche des puissances successives d'une matrice A .

Méthode 22.2.9 (Calcul de A^n à l'aide d'un polynôme annulateur)

- Déterminer un polynôme annulateur P de petit degré de A .
- Chercher le reste R_n de la division euclidienne de X^n par P .
- Évaluer l'égalité de division euclidienne en A , il reste $A^n = R_n(A)$. Ainsi A^n s'exprime comme combinaison linéaire des premières puissances de A .

Exemple 22.2.10

Calculer $\begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}^n$.

II.2 Matrices carrées de type particulier

La matrice I_n a une particularité notable : elle est nulle, sauf sur sa diagonale. Les matrices vérifiant cette propriété jouent un rôle central, notamment dans la théorie de la diagonalisation. En effet, les endomorphismes associés laissent stables les axes définis par les vecteurs de la base, et sont donc faciles à étudier. En particulier les produits de matrices diagonales sont simples à exprimer (et donc aussi les puissances). C'est une des motivations de la théorie de la diagonalisation. Nous présentons d'autres formes de matrices (matrices triangulaires)

Définition 22.2.11 (Matrice diagonale)

Soit D une matrice de $\mathcal{M}_n(\mathbb{K})$. On dit que D est une matrice diagonale si tous ses coefficients sont nuls, à l'exception éventuellement de ses coefficients diagonaux.

Ainsi, une matrice diagonale est de la forme :

$$D = \begin{pmatrix} d_1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & d_n \end{pmatrix}$$

Définition 22.2.12 (Matrice triangulaire)

Soit T une matrice de $\mathcal{M}_n(\mathbb{K})$. On dit que T est une matrice triangulaire supérieure (resp. inférieure) si tous ses coefficients situés strictement en-dessous (resp. au-dessus) de sa diagonale sont nuls.

Ainsi, une matrice triangulaire supérieure (resp. inférieure) est de la forme :

$$T = \begin{pmatrix} \bullet & \cdots & \cdots & \bullet \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \bullet \end{pmatrix} \quad (\text{resp. } T = \begin{pmatrix} \bullet & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ \bullet & \cdots & \cdots & \bullet \end{pmatrix})$$

où les \bullet désignent des coefficients quelconques. On définit également les matrices strictement triangulaires supérieures ou inférieures, nulles également sur la diagonale.

Notation 22.2.13

Nous noterons dans ce cours $\mathcal{D}_n(\mathbb{K})$ l'espace des matrices diagonales d'ordre n , $\mathcal{T}_n^+(\mathbb{K})$ l'espace des matrices triangulaires supérieures, $\mathcal{T}_n^-(\mathbb{K})$ l'espace des matrices triangulaires inférieures, $\overline{\mathcal{T}}_n^+(\mathbb{K})$ l'espace des matrices strictement triangulaires supérieures, et $\overline{\mathcal{T}}_n^-(\mathbb{K})$ l'espace des matrices strictement triangulaires inférieures.

Remarque 22.2.14 (Caractérisation des matrices triang. par stabilisation d'un drapeau)

Soit $\mathcal{B} = (b_1, \dots, b_n)$ une base de E et $u \in \mathcal{L}(E)$. On note pour tout $k \in \mathbb{N}$

$$E_k = \text{Vect}(b_1, \dots, b_k),$$

avec la convention $E_0 = \{0\}$. On dit que $(E_k)_{k \in \llbracket 1, n \rrbracket}$ est un drapeau (défini comme une suite strictement croissante de $n + 1$ sous-espaces vectoriels d'un espace vectoriel de dimension n , ce qui impose que $\dim E_k = k$), et que \mathcal{B} est une base adaptée au drapeau (E_k) . On dit que u stabilise le drapeau (E_k) si et seulement si chaque E_k est stable par u .

Alors, $\text{Mat}_{\mathcal{B}}(u)$ est triangulaire supérieure si et seulement si u stabilise le drapeau (E_k) .

Exemple 22.2.15

$(\mathbb{R}_k[X])_{k \in \llbracket 0, n \rrbracket}$, auquel on ajoute $\{0\}$ initialement, est un drapeau, associé à la base canonique de $\mathbb{R}_n[X]$. Les endomorphismes conservant, ou diminuant le degré, stabilisent ce drapeau (par exemple $u(P) = \alpha P + \beta P'$).

Proposition 22.2.16

Nous avons de façon évidente :

$$\mathcal{T}_n^+(\mathbb{K}) = \overline{\mathcal{T}}_n^+(\mathbb{K}) \oplus \mathcal{D}_n(\mathbb{K}) \quad \text{et} \quad \mathcal{T}_n^-(\mathbb{K}) = \overline{\mathcal{T}}_n^-(\mathbb{K}) \oplus \mathcal{D}_n(\mathbb{K}),$$

ainsi que :

$$\mathcal{M}_n(\mathbb{K}) = \overline{\mathcal{T}}_n^+(\mathbb{K}) \oplus \mathcal{D}_n(\mathbb{K}) \oplus \overline{\mathcal{T}}_n^-(\mathbb{K}) = \overline{\mathcal{T}}_n^+(\mathbb{K}) \oplus \mathcal{T}_n^-(\mathbb{K}) = \mathcal{T}_n^+(\mathbb{K}) \oplus \overline{\mathcal{T}}_n^-(\mathbb{K})$$

Nous avons également des propriétés de stabilité :

Proposition 22.2.17 (Produit de matrices diagonales)

Le produit de deux matrices diagonales de même ordre est encore une matrice diagonale. De façon plus explicite :

$$\begin{pmatrix} c_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & c_n \end{pmatrix} \begin{pmatrix} d_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & d_n \end{pmatrix} = \begin{pmatrix} c_1 d_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & c_n d_n \end{pmatrix}.$$

En particulier, pour tout $k \in \mathbb{N}$, on a :

$$\begin{pmatrix} d_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & d_n \end{pmatrix}^k = \begin{pmatrix} d_1^k & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & d_n^k \end{pmatrix}.$$

Proposition 22.2.18 (Produit de deux matrices triangulaires)

- (i) Le produit de deux matrices triangulaires supérieures est une matrice triangulaire supérieure dont les coefficients diagonaux sont les produits deux à deux des coefficients diagonaux des deux matrices.
- (ii) En particulier, si l'une des deux matrices est strictement triangulaire supérieure, le produit l'est également.
- (iii) De même pour les produits de matrices triangulaires inférieures.

Des règles de stabilité précédentes, on déduit :

Théorème 22.2.19 (Structure des sous-ensembles de matrices de type particulier)

Les sous-ensembles $\mathcal{D}_n(\mathbb{K})$, $\mathcal{T}_n^+(\mathbb{K})$, et $\mathcal{T}_n^-(\mathbb{K})$ sont des sous-algèbres de $\mathcal{M}_n(\mathbb{K})$.

Voici d'autres matrices carrées jouant un rôle important, notamment en algèbre bilinéaire :

Définition 22.2.20 (Matrices symétriques, antisymétriques)

Soit $A \in \mathcal{M}_n(\mathbb{K})$.

1. On dit que A est *symétrique* si $A = {}^t A$.
2. On dit que A est *antisymétrique* si $A = -{}^t A$.

On note $\mathcal{S}_n(\mathbb{K})$ l'ensemble des matrices symétriques de $\mathcal{M}_n(\mathbb{K})$, et $\mathcal{A}_n(\mathbb{K})$ l'ensemble des matrices antisymétriques.

En tant de noyau des endomorphismes $\text{id} - t$ et $\text{id} + t$ (où t désigne la transposition), $\mathcal{S}_n(\mathbb{K})$ et $\mathcal{A}_n(\mathbb{K})$ sont des sous-espaces vectoriels de $\mathcal{M}_n(\mathbb{K})$.

Proposition 22.2.21

Les sous-espaces $\mathcal{S}_n(\mathbb{K})$ et $\mathcal{A}_n(\mathbb{K})$ sont supplémentaires l'un de l'autre dans $\mathcal{M}_n(\mathbb{K})$.

Remarque 22.2.22

- À quoi ressemble la diagonale d'une matrice antisymétrique ?
- $\mathcal{A}_n(\mathbb{K})$ et $\mathcal{S}_n(\mathbb{K})$ sont-elles des sous-algèbres de $\mathcal{M}_n(\mathbb{K})$?

II.3 Matrices inversibles

Définition 22.2.23 (Matrice inversible)

Une matrice M est dite inversible s'il existe $n \in \mathbb{N}^*$ tel que $M \in \mathcal{M}_n(\mathbb{K})$, et tel que M soit inversible dans l'anneau $\mathcal{M}_n(\mathbb{K})$.

Avertissement 22.2.24

Par définition, une matrice inversible est toujours une matrice carrée.

Proposition 22.2.25 (caractérisation par l'endomorphisme canoniquement associée)

Une matrice $M \in \mathcal{M}_n(\mathbb{K})$ est inversible si et seulement si l'endomorphisme canoniquement associé $f \in \mathcal{L}(\mathbb{K}^n)$ est un automorphisme.

Cette caractérisation explique qu'on se limite aux matrices carrées : en effet un isomorphisme préserve les dimensions, donc une matrice non carrée ne peut pas être canoniquement associée à un isomorphisme. Nous déduisons alors de la caractérisation des isomorphismes en dimension finie que :

Proposition 22.2.26

Soit $A \in \mathcal{M}_n(\mathbb{K})$. Pour que A soit inversible, il suffit qu'il existe une matrice $B \in \mathcal{M}_n(\mathbb{K})$ telle que $AB = I_n$ OU $BA = I_n$. Dans ce cas $B = A^{-1}$.

Définition 22.2.27 (Groupe linéaire)

L'ensemble des matrices inversibles de $\mathcal{M}_n(\mathbb{K})$ est noté $\text{GL}_n(\mathbb{K})$, et est appelé *n-ième groupe linéaire*.

Proposition 22.2.28 (Structure de $\text{GL}_n(\mathbb{K})$)

$(\text{GL}_n(\mathbb{K}), \times)$ est un groupe (c'est le groupe des inversibles de l'anneau $\mathcal{M}_n(\mathbb{K})$).

Conformément aux propriétés générales concernant l'inverse de produits dans un anneau, on a :

Proposition 22.2.29 (Inverse d'un produit)

Soit A et B deux matrices inversibles de $\mathcal{M}_n(\mathbb{K})$. Alors AB est inversible, et son inverse est $B^{-1}A^{-1}$.

Par ailleurs, l'inversion commute avec la transposition :

Proposition 22.2.30 (Inverse d'une transposée)

Soit A une matrice inversible. Alors tA l'est également, et

$$({}^tA)^{-1} = {}^t(A^{-1}).$$

Voici quelques exemples importants de matrices inversibles.

Proposition 22.2.31 (CNS d'inversibilité pour les matrices triangulaires)

Soit $T \in \mathcal{T}_n^+(\mathbb{K})$ une matrice triangulaire. Alors T est inversible si et seulement si tous ses coefficients diagonaux sont non nuls (donc inversibles), et dans ce cas, T^{-1} est une matrice triangulaire dont les coefficients diagonaux sont les inverses des coefficients diagonaux de T .

En particulier, une matrice diagonale est inversible si et seulement si tous ses coefficients diagonaux sont non nuls, et dans ce cas :

$$D^{-1} = \begin{pmatrix} d_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & d_n \end{pmatrix}^{-1} = \begin{pmatrix} d_1^{-1} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & d_n^{-1} \end{pmatrix}.$$

II.4 Expression matricielle du pivot de Gauss

Les exemples suivants sont importants car ils valident la méthode du pivot de Gauss.

Définition 22.2.32 (Matrices de codage des opérations élémentaires)

Soit $n \in \mathbb{N}^*$. On définit les trois familles suivantes de matrices :

(i) Codage des échanges de lignes (matrice de transposition) :

$$\forall (i, j) \in \llbracket 1, n \rrbracket^2, \quad i \neq j, \quad E(i, j) = \begin{pmatrix} 1 & 0 & \dots & \dots & \dots & \dots & \dots & 0 \\ 0 & 1 & \ddots & & & & & \vdots \\ \vdots & \ddots & 0 & \ddots & & 1 & & \vdots \\ \vdots & & \ddots & 1 & \ddots & & & \vdots \\ \vdots & & & \ddots & 1 & \ddots & & \vdots \\ \vdots & & & & \ddots & 0 & \ddots & \vdots \\ \vdots & & & & & \ddots & 1 & 0 \\ 0 & \dots & \dots & \dots & \dots & \dots & 0 & 1 \end{pmatrix} \begin{matrix} \\ \\ i \\ \\ j \\ \\ \\ \end{matrix}$$

(ii) Codage de la multiplication d'une ligne par un scalaire (matrice de dilatation) :

$$E_i(\lambda) = \begin{pmatrix} 1 & 0 & \dots & \dots & \dots & \dots & 0 \\ 0 & \ddots & \ddots & & & & \vdots \\ \vdots & \ddots & 1 & \ddots & & & \vdots \\ \vdots & & \ddots & \lambda & \ddots & & \vdots \\ \vdots & & & \ddots & 1 & \ddots & \vdots \\ \vdots & & & & \ddots & \ddots & 0 \\ 0 & \dots & \dots & \dots & \dots & 0 & 1 \end{pmatrix} \begin{matrix} \\ \\ \\ i \\ \\ \\ \end{matrix}$$

(iii) Codage d'une combinaison linéaire (matrice de transvection)

$$E_{i,j}(\lambda) = \begin{pmatrix} 1 & 0 & \cdots & \cdots & \cdots & 0 \\ 0 & \ddots & \ddots & \lambda & & \vdots \\ \vdots & \ddots & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & \cdots & 0 & 1 \end{pmatrix} \begin{matrix} \\ i \\ \\ \\ \\ j \end{matrix}$$

Proposition 22.2.33 (Interprétation matricielle des opérations élémentaires sur les lignes)
 Soit $M \in \mathcal{M}_{m,n}(\mathbb{K})$. Soit $i \neq j$ dans $\llbracket 1, m \rrbracket$, et $\lambda \in \mathbb{K}$.

- (i) La matrice N obtenue de M par l'opération $L_i \leftrightarrow L_j$, est $N = E(i, j) \cdot M$;
- (ii) La matrice N obtenue de M par l'opération $L_i \leftarrow L_i + \lambda L_j$ est $N = E_{i,j}(\lambda) \cdot M$;
- (iii) La matrice N obtenue de M par l'opération $L_i \leftarrow \lambda L_i$, $\lambda \neq 0$, est $N = E_i(\lambda) \cdot M$.

Ainsi, partant d'une matrice M à laquelle on applique le pivot de Gauss, on obtient une matrice N , et une matrice P telles que $N = PM$, telle que P soit un produit de matrices de transvection, de dilatation et de permutation.

La validité du pivot de Gauss provient alors du résultat suivant, exprimant que toute opération valide du pivot est réversible :

Proposition 22.2.34 (Inversibilité des matrices de codage des opérations élémentaires)
 Pour $i \neq j$, les matrices $E(i, j)$, $E_{i,j}(\lambda)$ sont inversibles, ainsi que $E_i(\lambda)$ lorsque $\lambda \neq 0$.

Théorème 22.2.35 (Validité de la méthode du pivot)
 Soit $M \in \mathcal{M}_{n,p}(\mathbb{K})$, $B \in \mathcal{M}_{n,1}(\mathbb{K})$, et $P \in \text{GL}_n(\mathbb{K})$. Alors

$$\forall X \in \mathcal{M}_{p,1}(\mathbb{K}), \quad MX = B \iff PMX = PB.$$

Ce théorème exprime la stabilité des solutions d'une équation matricielle par multiplication de l'équation par une matrice inversible.

Au passage, si M lui-même est inversible, ce même théorème affirme qu'on a l'équivalence :

$$MX = Y \iff X = M^{-1}Y.$$

Réciproquement, l'existence et l'unicité d'une solution pour tout second membre Y implique la bijectivité de l'application canoniquement associée à M , donc l'inversibilité de M .

II.5 Calcul pratique de l'inverse d'une matrice

Ainsi, une façon de calculer l'inverse est la résolution d'un système, de second membre Y indéterminé :

Méthode 22.2.36 (Calcul pratique de l'inverse (première méthode))

- Résoudre le système $AX = Y$, où Y est pris en paramètre.

- A est inversible si et seulement si le système admet une et une seule solution, quel que soit le paramètre Y .
- Dans ce cas, écrire le résultat sous forme matricielle $X = BY$.
- B est alors l'inverse de A .

Exemple 22.2.37

Calcul de l'inverse de $A = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$.

Définition 22.2.38 (Système de Cramer)

Un système de Cramer est un système d'équations linéaires admettant une unique solution.

Ainsi, un système est de Cramer si et seulement si la matrice de ses coefficients est inversible. Dans le cas de matrices 2×2 , on a une méthode plus rapide :

Théorème 22.2.39 (Inverse des matrices 2×2 par la comatrice)

Soit $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathcal{M}_2(\mathbb{K})$. Alors M est inversible si et seulement si $ad - bc \neq 0$, et

$$M^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Définition 22.2.40 (Déterminant d'une matrice 2×2)

La quantité $ad - bc$ est appelée *déterminant* de M , et est noté $\det M$ ou $\begin{vmatrix} a & b \\ c & d \end{vmatrix}$.

Remarque 22.2.41

Il existe une notion de déterminant pour des matrices carrées de taille quelconque n . Nous définirons cette notion générale dans le chapitre suivant. La non nullité du déterminant caractérise alors l'inversibilité de la matrice, et comme dans le cas $n = 2$, il existe une formule générale de l'inverse d'une matrice basée sur la notion de comatrice. Dans des situations concrètes, cette formule est cependant assez peu efficace, sauf pour $n = 2$, et éventuellement $n = 3$ (et encore...)

Une deuxième méthode générale de calcul de l'inverse d'une matrice repose sur l'observation suivante : la méthode du pivot de Gauss consiste en des multiplications à gauche par des matrices codant les opérations élémentaires. Ainsi, si à l'aide d'opérations élémentaires sur les lignes, on parvient à transformer la matrice initiale A en la matrice identité, on aura l'existence d'une matrice inversible P (obtenue en multipliant les matrices de codage) telle que $PA = I_n$. Ainsi, cette matrice P est l'inverse de A . Or, $P = PI_n$, et est donc obtenu en appliquant à la matrice I_n les mêmes opérations sur ses lignes que celles qui ont permis de transformer A en I_n . On en déduit la méthode suivante :

Méthode 22.2.42 (Calcul de l'inverse d'une matrice (seconde méthode))

- Juxtaposer la matrice A et la matrice I_n (séparées d'une barre verticale)
- Effectuer un pivot sur A , en faisant les mêmes opérations sur la matrice I_n , pour obtenir une matrice échelonnée à la place de A

- La matrice A est inversible si et seulement si la matrice échelonnée obtenue est inversible (c'est-à-dire s'il s'agit d'une matrice triangulaire supérieure à coefficients diagonaux non nuls)
- Dans ce cas, faire un pivot remontant, pour annuler les coefficients au dessus de chaque pivot, et toujours en effectuant les mêmes opérations sur la matrice de droite.
- En normalisant les coefficients diagonaux, on obtient à gauche la matrice identité, et à droite la matrice A^{-1} .

III Rang d'une matrice

La correspondance entre application linéaire et matrice permet d'étendre aux matrices un certain nombre de notions définies pour les applications linéaires.

III.1 Image et noyau d'une matrice

Définition 22.3.1 (Image et noyau d'une matrice)

Par définition, l'image et le noyau d'une matrice M sont l'image et le noyau de l'application linéaire canoniquement associée. Plus précisément, soit $M \in \mathcal{M}_{n,p}(\mathbb{K})$ et $f : X \mapsto MX$ l'application de $\mathcal{L}(\mathbb{K}^p, \mathbb{K}^n)$ canoniquement associée. Alors

- $\text{Im}(M) = \text{Im}(f) = \{Y \in \mathbb{K}^n \simeq \mathcal{M}_{n,1}(\mathbb{K}) \mid \exists X \in \mathbb{K}^p, MX = Y\}$
- $\text{Ker}(M) = \text{Ker}(f) = \{X \in \mathbb{K}^p \simeq \mathcal{M}_{p,1}(\mathbb{K}) \mid MX = 0\}$.

En particulier, puisque les colonnes de M sont les images par f des éléments de la base canonique, nous obtenons :

Proposition 22.3.2 (Description de l'image d'une matrice)

Soit M une matrice de colonnes C_1, \dots, C_p . Alors

$$\text{Im}(M) = \text{Vect}(C_1, \dots, C_p).$$

Définition 22.3.3 (Rang d'une matrice)

Soit M une matrice. Son rang est égal au rang de l'application linéaire canoniquement associée, c'est-à-dire :

$$\text{rg}(M) = \text{rg}(f) = \dim(\text{Im}(f)) = \dim(\text{Im}(M)) = \dim(\text{Vect}(C_1, \dots, C_n)) = \text{rg}(C_1, \dots, C_n),$$

où les C_i sont les colonnes de M .

Évidemment, le théorème du rang est valable également pour les matrices :

Théorème 22.3.4 (Théorème du rang)

Soit $M \in \mathcal{M}_{n,p}(\mathbb{K})$. Alors

$$\dim(\text{Ker}(M)) + \text{rg}(M) = p.$$

Un exemple important, auquel on se ramène souvent par utilisation du pivot de Gauss est le suivant :

Proposition 22.3.5 (Rang d'une matrice échelonnée)

Soit M une matrice échelonnée (voir chapitre sur les systèmes d'équations linéaires). Alors $\text{rg}(M)$ est le nombre de lignes non nulles de M .

III.2 Calcul du rang

Une propriété importante, permettant le calcul du rang par la méthode du pivot, est la conservation du rang par multiplication par une matrice inversible, donc en particulier par les matrices de codage des opérations. Ainsi, le pivot de Gauss conserve le rang.

Théorème 22.3.6 (Conservation de l'image et du noyau)

Soit $M \in \mathcal{M}_{n,p}(\mathbb{K})$ une matrice quelconque, et $P \in \text{GL}_n(\mathbb{K})$, $Q \in \text{GL}_p(\mathbb{K})$. Alors :

- (i) $\text{Ker}(PM) = \text{Ker}(M)$
- (ii) $\text{Im}(MQ) = \text{Im}(M)$.

Ainsi, la multiplication à gauche par une matrice inversible conserve le noyau et la multiplication à droite conserve l'image.

Corollaire 22.3.7 (Conservation du rang)

Soit $M \in \mathcal{M}_{n,p}(\mathbb{K})$ une matrice quelconque, et $P \in \text{GL}_n(\mathbb{K})$, $Q \in \text{GL}_p(\mathbb{K})$. Alors :

$$\text{rg}(PMQ) = \text{rg}(M)$$

Corollaire 22.3.8 (Conservation de l'image et du noyau par opérations élémentaires)

- (i) Les opérations élémentaires sur les lignes conservent le noyau
- (ii) Les opérations élémentaires sur les colonnes conservent l'image
- (iii) Les opérations élémentaires sur les lignes et les colonnes conservent le rang

Méthode 22.3.9 (calcul du rang d'une matrice)

- Effectuer un pivot pour se ramener à une matrice échelonnée
- Le rang est égal au rang de la matrice échelonnée, donc au nombre de ses lignes non nulles
- Si les opérations sur les colonnes semblent plus simples, c'est possible aussi
- On peut même combiner les deux.

Méthode 22.3.10 (Trouver une base de l'image)

- Cette fois, on veut conservation de l'image, on travaille donc sur les colonnes.
- On effectue un pivot sur les colonnes jusqu'à obtenir une matrice échelonnée sur les colonnes, dont les colonnes engendrent l'image.
- Les colonnes non nulles de cette matrice forment une base de l'image.

Méthode 22.3.11 (Trouver un supplémentaire)

- On suppose donné un sous-espace E de \mathbb{K}^p dont on connaît une famille génératrice de n vecteurs.
- On écrit la matrice de cette famille (chaque vecteur est écrit en colonne) : il s'agit de la matrice de l'application linéaire qui envoie les vecteurs de la base canonique de \mathbb{K}^n sur les vecteurs de la famille génératrice. L'image de cette application est E .

- On effectue un pivot en travaillant sur les colonnes, pour conserver l'image.
- À la matrice échelonnée sur les colonnes obtenue au bout, on ajoute les colonnes correspondant aux vecteurs de la base canonique de sorte à obtenir une matrice triangulaire supérieure.
- Les vecteurs ainsi ajoutés forment une base d'un supplémentaire de E dans \mathbb{K}^p .

On observe que :

Lemme 22.3.12

Une matrice échelonnée a même rang que sa transposée.

En effectuant sur la matrice tA les opérations sur les colonnes correspondant aux opérations sur les lignes effectuées par A pour se ramener à une matrice échelonnée par la méthode du pivot, on en déduit alors :

Théorème 22.3.13 (Rang d'une transposée)

Soit $A \in \mathcal{M}_{n,p}(\mathbb{K})$. On a :

$$\text{rg}(A) = \text{rg}({}^tA).$$

III.3 Caractérisation du rang par les matrices extraites

Définition 22.3.14 (Matrice extraite)

Soit $A = (a_{i,j})$ une matrice de $\mathcal{M}_{n,p}$. Une matrice B est extraite de A si et seulement s'il existe $1 \leq i_1 < \dots < i_q \leq n$ et $1 \leq j_1 < \dots < j_r \leq p$ tels que $B = (a_{i_k, j_\ell})_{(k,\ell) \in \llbracket 1,q \rrbracket \times \llbracket 1,r \rrbracket}$.

Ainsi, la matrice B est obtenue en ne conservant de A que les lignes d'indice i_1, \dots, i_q et les colonnes d'indice j_1, \dots, j_r .

Lemme 22.3.15 (Description vectoriel d'une extraction)

Avec les notations précédentes, en notant (b_1, \dots, b_p) la base canonique de \mathbb{K}^p , (c_1, \dots, c_n) la base canonique de \mathbb{K}^n , f l'application linéaire de $\mathcal{L}(\mathbb{K}^p, \mathbb{K}^n)$ canoniquement associée à M , $I = \{i_1, \dots, i_q\}$, $J = \{j_1, \dots, j_r\}$, en définissant i l'injection de $F = \text{Vect}(b_i, i \in I)$ dans \mathbb{K}^p et p la projection de \mathbb{K}^n sur $G = \text{Vect}(c_j, j \in J)$, parallèlement à $\text{Vect}(c_j, j \notin J)$, et enfin, en notant g l'application linéaire de F dans G associée à la matrice B , relativement aux bases $(b_i)_{i \in I}$ et $(c_j)_{j \in J}$, on a la relation :

$$g = p \circ f \circ i.$$

Proposition 22.3.16 (Rang d'une matrice extraite)

Soit B une matrice extraite de A . Alors $\text{rg}(B) \leq \text{rg}(A)$.

Théorème 22.3.17 (Caractérisation du rang par les matrices carrées extraites)

Le rang de A est l'ordre maximal d'une matrice carrée inversible extraite de A .

IV Changements de base

La correspondance entre applications linéaires et matrice dépend du choix de bases de l'espace de départ E et de l'espace d'arrivée F . Le but de ce paragraphe est d'étudier l'effet d'un changement de base sur E

ou sur F : la matrice relativement aux nouvelles bases se déduit-elle d'une certaine manière de la matrice relativement aux anciennes bases ?

IV.1 Changements de base pour des applications linéaires

Définition 22.4.1 (Matrice de passage)

Soit E un espace vectoriel, et \mathcal{B}_1 et \mathcal{B}_2 deux bases de E . Alors la matrice de passage de la base \mathcal{B}_1 à la base \mathcal{B}_2 est la matrice :

$$P_{\mathcal{B}_1}^{\mathcal{B}_2} = \text{Mat}_{\mathcal{B}_2, \mathcal{B}_1}(\text{Id}_E) = [\mathcal{B}_2]_{\mathcal{B}_1}.$$

Ainsi, la i -ème colonne de cette matrice est constituée des coordonnées du i -ème vecteur de la base \mathcal{B}_2 dans la base \mathcal{B}_1 . Il s'agit donc de la matrice de la famille des vecteurs de la seconde base \mathcal{B}_2 dans la première base \mathcal{B}_1 .

La formule de composition amène facilement :

Proposition 22.4.2 (Inversibilité des matrices de passage)

Toute matrice de passage $P_{\mathcal{B}_1}^{\mathcal{B}_2}$ est inversible. Son inverse est la matrice de passage $P_{\mathcal{B}_2}^{\mathcal{B}_1}$.

Par ailleurs, toute matrice inversible peut être interprétée comme une matrice de passage. Plus précisément :

Proposition 22.4.3

Soit $P \in \text{GL}_n(\mathbb{K})$ une matrice inversible, et E un espace vectoriel de dimension n .

- Pour toute base \mathcal{B} de E , il existe une base \mathcal{C} telle que $P = P_{\mathcal{B}}^{\mathcal{C}}$.
- Pour toute base \mathcal{C} de E , il existe une base \mathcal{B} telle que $P = P_{\mathcal{B}}^{\mathcal{C}}$.

Puisque $[X]_{\mathcal{B}_1} = \text{Mat}_{\mathcal{B}_2, \mathcal{B}_1}(\text{Id})[X]_{\mathcal{B}_2}$, on obtient l'expression de l'effet d'un changement de base sur les coordonnées d'un vecteur :

Proposition 22.4.4 (Effet d'un changement de base sur les coordonnées d'un vecteur)

Soit $X \in E$. Alors $[X]_{\mathcal{B}_1} = P_{\mathcal{B}_1}^{\mathcal{B}_2} \cdot [X]_{\mathcal{B}_2}$.

En composant à gauche et à droite par l'identité, avec bases différentes, on obtient, toujours d'après la formule de composition, la très importante formule de changement de base.

Théorème 22.4.5 (Formule de changement de base)

Soit E un espace vectoriel de dimension finie, muni de deux bases \mathcal{B}_1 et \mathcal{B}_2 , et F un espace vectoriel de dimension finie, muni de deux bases \mathcal{C}_1 et \mathcal{C}_2 . Soit $f \in \mathcal{L}(E, F)$. Alors :

$$\text{Mat}_{\mathcal{B}_2, \mathcal{C}_2}(f) = P_{\mathcal{C}_2}^{\mathcal{C}_1} \cdot \text{Mat}_{\mathcal{B}_1, \mathcal{C}_1}(f) \cdot P_{\mathcal{B}_1}^{\mathcal{B}_2} = (P_{\mathcal{C}_1}^{\mathcal{C}_2})^{-1} \cdot \text{Mat}_{\mathcal{B}_1, \mathcal{C}_1}(f) \cdot P_{\mathcal{B}_1}^{\mathcal{B}_2}$$

Ainsi, cette formule s'écrit $M' = Q^{-1}MP$, où M est la matrice dans les bases initiales, M' la matrice dans les nouvelles bases, P et Q les matrices de passage de la première base vers la seconde, respectivement dans E et dans F .

Exemple 22.4.6

1. Retrouver les matrices des exemples précédents.
2. Donner une relation entre la matrice de D dans la base canonique de $\mathbb{R}_n[X]$ (au départ et à l'arrivée) et la matrice de Jordan J_{n+1} .

IV.2 Matrices équivalentes

On remarque que deux matrices associées à une même application linéaire avec des choix différents de bases, s'obtiennent l'une de l'autre par multiplication à gauche et à droite par des matrices inversibles. Cela motive la définition suivante :

Définition 22.4.7 (Matrices équivalentes)

Soit $(M, N) \in \mathcal{M}_{n,p}(\mathbb{K})$. On dit que M et N sont équivalentes si et seulement si il existe $P \in \text{GL}_n(\mathbb{K})$ et $Q \in \text{GL}_p(\mathbb{K})$ tels que $N = PMQ$.

Proposition 22.4.8

Cela définit une relation d'équivalence.

Ainsi, on obtient :

Théorème 22.4.9

Soit $f \in \mathcal{L}(E, F)$.

- (i) Deux matrices M et N représentant f dans des choix différents de bases sont équivalentes ;
- (ii) Réciproquement, si $M = \text{Mat}_{\mathcal{B}_1, \mathcal{C}_1}(f)$ est la matrice de f relativement à deux bases \mathcal{B}_1 et \mathcal{C}_1 , et si N est équivalente à M , alors il existe deux bases \mathcal{B}_2 et \mathcal{C}_2 de E et F telles que $N = \text{Mat}_{\mathcal{B}_2, \mathcal{C}_2}(f)$.

Soit, pour tout $(n, p) \in (\mathbb{N}^*)^2$, et pour tout $r \in \llbracket 0, \min(n, p) \rrbracket$, soit

$$I_{n,p,r} = \begin{pmatrix} 1 & 0 & \cdots & \cdots & 0 & \cdots & \cdots & 0 \\ 0 & 1 & \ddots & & \vdots & & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots & & & \vdots \\ \vdots & & \ddots & 1 & 0 & \cdots & \cdots & 0 \\ 0 & \cdots & \cdots & 0 & 0 & \cdots & \cdots & 0 \\ \vdots & & & \vdots & \vdots & & & \vdots \\ 0 & \cdots & \cdots & 0 & 0 & \cdots & \cdots & 0 \end{pmatrix} = \left(\begin{array}{c|c} I_r & 0_{r,p-r} \\ \hline 0_{n-r,r} & 0_{n-r,p-r} \end{array} \right)$$

la matrice de type (n, p) , constituée de 1 sur ses r premiers coefficients diagonaux, et de 0 partout ailleurs.

Théorème 22.4.10

Soit E et F des espaces de dimensions respectives p et n . Soit $f \in \mathcal{L}(E, F)$, de rang r . Il existe deux bases \mathcal{B} et \mathcal{C} , respectivement de E et de F , telles que

$$\text{Mat}_{\mathcal{B}, \mathcal{C}}(f) = I_{n,p,r}.$$

Corollaire 22.4.11

Toute matrice de $\mathcal{M}_{n,p}(\mathbb{K})$ de rang r est équivalente à $I_{n,p,r}$.

Corollaire 22.4.12 (Classification des matrices équivalentes par le rang)

Deux matrices de même format sont équivalentes si et seulement si elles ont même rang.

Ainsi, dans $\mathcal{M}_{n,p}(\mathbb{K})$, les classes d'équivalence, pour la relation d'équivalence des matrices, sont les sous-ensembles de matrices de même rang $r \in \llbracket 0, \min(n, p) \rrbracket$. En particulier, l'espace quotient est en bijection avec $\llbracket 0, \min(n, p) \rrbracket$.

IV.3 Matrice d'un endomorphisme, matrices semblables

Dans le cas où f est un endomorphisme, il est fréquent de choisir sur E la même base au départ et à l'arrivée (même si ceci n'est en théorie pas strictement nécessaire). Dans ce cas, on allège un peu les notations, en notant simplement $\mathcal{M}_{\mathcal{B}}(f)$ au lieu de $\mathcal{M}_{\mathcal{B},\mathcal{B}}(f)$. Par ailleurs, un changement de base sur un endomorphisme s'effectue dans ce cas en faisant le même changement de variable au départ et à l'arrivée. Ainsi :

Théorème 22.4.13 (Changement de base pour un endomorphisme)

Soit $f \in \mathcal{L}(E)$, E étant un espace vectoriel de dimension finie, muni de deux bases \mathcal{B}_1 et \mathcal{B}_2 . Alors :

$$\text{Mat}_{\mathcal{B}_2}(f) = (P_{\mathcal{B}_1}^{\mathcal{B}_2})^{-1} \text{Mat}_{\mathcal{B}_1}(f) P_{\mathcal{B}_1}^{\mathcal{B}_2}.$$

Ainsi, cette relation s'écrit : $M' = P^{-1}MP$, où M est la matrice de f dans l'ancienne base, M' la matrice de f dans la nouvelle base, et P la matrice de passage de l'ancienne base vers la nouvelle base.

Exemple 22.4.14

1. Soit p un projecteur de rang r d'un espace de dimension n . Montrer qu'il existe une base \mathcal{B} de E telle que $\text{Mat}_{\mathcal{B}}(p) = I_{n,n,r}$.
2. Décrire de même une matrice simple représentant une symétrie dans un certain choix de base.
3. Déterminer une base relativement à laquelle la matrice de $f : (x, y) \mapsto (3x + 2y, x)$ est diagonale. En déduire une relation entre cette matrice diagonale et la matrice $\begin{pmatrix} 3 & -2 \\ 1 & 0 \end{pmatrix}$.

On dit qu'on a diagonalisé $\begin{pmatrix} 3 & -2 \\ 1 & 0 \end{pmatrix}$, ou de façon équivalente, qu'on a diagonalisé f .

La notion de diagonalisation est liée à l'étude des classes d'équivalence de la relation fournie par les changements de base pour les endomorphismes. Cela motive la définition de la relation d'équivalence idoine, définissant la notion de matrices semblables.

IV.4 Matrices semblables**Définition 22.4.15 (Matrices semblables)**

On dit que deux matrices $A, B \in \mathcal{M}_n(\mathbb{K})$ sont semblables s'il existe une matrice inversible $P \in \text{GL}_n(\mathbb{K})$ telle que $B = P^{-1}AP$.

Proposition 22.4.16 (Relation de similitude)

La relation ainsi définie, appelée relation de similitude, est une relation d'équivalence.

Ainsi :

Corollaire 22.4.17

Les matrices d'un endomorphisme dans différentes bases de E sont semblables.

La classification des matrices semblables est beaucoup plus compliquée que la classification des matrices équivalentes. Cette classification est à l'origine du problème de la réduction des endomorphismes. Le problème de la réduction des endomorphismes consiste à trouver un système simple de représentants des classes de similitude, donc de trouver une matrice simple canonique équivalente à une matrice donnée. La diagonalisation des matrices (ou des endomorphismes) est une des branches de ce problème, mais les matrices diagonales ne suffisent pas à donner un système de représentants des classes de similitude. Par ailleurs, il est important de noter que cette notion dépend beaucoup du corps de base. En effet, toute matrice de $\mathcal{M}_n(\mathbb{C})$ est \mathbb{C} -semblable à une matrice triangulaire supérieure (et même à une matrice de forme assez particulière, avec seulement deux diagonales non nulles), donc toute matrice de $\mathcal{M}_n(\mathbb{R})$ aussi ; en revanche elles ne sont pas toutes \mathbb{R} -semblables à une matrice triangulaire supérieure. Trouver un système de représentant des classes de \mathbb{R} -similitude est un problème plus compliqué que trouver un système de représentant de \mathbb{C} -similitude.

La classification des classes de similitude étant loin d'être triviale, une étape importante est la recherche d'invariants de similitude, c'est-à-dire de propriétés ou quantités préservées par similitude. Ces invariants permettent de distinguer certaines matrices non semblables (si elles n'ont pas même invariant) mais sont en général insuffisants pour prouver qu'elles sont semblable (sauf si l'invariant caractérise la similitude).

Nous avons déjà étudié un invariant de similitude : le rang. En effet, deux matrices semblables sont équivalentes. Cet invariant est assez faible, puisqu'il ne permet de classer les matrices de $\mathcal{M}_n(\mathbb{K})$ qu'en $n + 1$ catégories.

Un deuxième invariant, encore plus facile à calculer, est la trace, que nous étudions dans le paragraphe suivant.

IV.5 Trace d'une matrice, trace d'un endomorphisme**Définition 22.4.18 (Trace d'une matrice)**

Soit $A = (a_{i,j}) \in \mathcal{M}_n(\mathbb{K})$ une matrice carrée d'ordre n . La trace de A est la somme de ses coefficients diagonaux :

$$\operatorname{tr}(A) = \sum_{i=1}^n a_{i,i}.$$

De façon assez immédiate, on obtient :

Proposition 22.4.19 (Linéarité de la trace)

L'application $\operatorname{tr} : \mathcal{M}_n(\mathbb{K}) \rightarrow \mathbb{K}$ est une forme linéaire sur $\mathcal{M}_n(\mathbb{K})$.

Proposition 22.4.20 (Invariance par transposition)

Soit $A \in \mathcal{M}_n(\mathbb{K})$. Alors $\operatorname{tr}(A) = \operatorname{tr}({}^tA)$.

Théorème 22.4.21 (Invariance de la trace par commutation interne)

Soit $A \in \mathcal{M}_{n,p}(\mathbb{K})$ et $B \in \mathcal{M}_{p,n}(\mathbb{K})$. Alors $\text{tr}(AB) = \text{tr}(BA)$.

Remarquez que A et B ne sont pas des matrices carrées, mais de format transposé. Ainsi, AB est une matrice carrée d'ordre n alors que BA est une matrice carrée d'ordre p .

Avertissement 22.4.22

Le théorème affirme qu'on peut inverser l'ordre d'un produit dans la trace, lorsque la matrice dont on cherche la trace s'exprime comme produit de 2 termes. Cela ne signifie pas qu'on peut commuter comme on veut les termes d'un produit de n termes à l'intérieur de la trace. Ainsi, on peut écrire :

$$\text{tr}(ABC) = \text{tr}(CBA) = \text{tr}(BCA),$$

mais en général, **on n'a pas** :

$$\text{tr}(ABC) = \text{tr}(ACB),$$

l'expression ACB ne pouvant pas se déduire de l'interversion (globale) de 2 termes de ABC .

Exemple 22.4.23

Comparer $\text{tr} \left(\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \right)$ et $\text{tr} \left(\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \right)$.

Corollaire 22.4.24 (Invariance de la trace par similitude)

La trace est un invariant de similitude. Autrement dit, si M et N sont semblables, alors $\text{tr}(M) = \text{tr}(N)$.

Cela permet de définir :

Définition 22.4.25 (Trace d'un endomorphisme)

La trace d'un endomorphisme est la valeur commune de la trace des matrices de f relativement à un choix quelconque d'une base.

On en déduit notamment que, comme pour les matrices, la trace est une forme linéaire sur $\mathcal{L}(E)$, et que pour tout $u \in \mathcal{L}(E, F)$ et $v \in \mathcal{L}(F, E)$, on a $\text{tr}(uv) = \text{tr}(vu)$.

En particulier, on a :

Proposition 22.4.26 (Trace d'un projecteur, d'une symétrie)

1. Soit p un projecteur de E , alors $\text{tr}(p) = \text{rg}(p)$.
2. Soit s une symétrie de E . Alors $\text{tr}(s) = n - 2 \text{rg}(s - \text{id})$

Méthode 22.4.27 (Dimension des éléments géométriques d'une projection/symétrie)

- S'assurer que l'endomorphisme considéré est une projection ou une symétrie, en calculant u^2 .
- Déterminer la matrice de u dans une base quelconque
- Déterminer la trace de u grâce à cette matrice.
- la propriété précédente donne la dimension des éléments propres $\text{Ker}(u)$ et $\text{Im}(u)$ de u si u est un projecteur, et de $\text{Ker}(u - \text{id})$ et $\text{Ker}(u + \text{id})$ si u est une symétrie (ces deux espaces étant supplémentaires pour une symétrie).

IV.6 Introduction à la réduction des endomorphismes (Spé)

Ces invariants ne représentent qu'un tout petit pas vers la classification des classes de similitude. Vous définirez l'année prochaine les valeurs propres d'une matrice. L'ensemble des valeurs propres d'une matrice est également un invariant de similitude.

Nous donnons sans preuve la description complète d'un système de représentant des classes de similitude dans $\mathcal{M}_n(\mathbb{C})$. Pour cela, nous définissons, pour tout $\ell \in \llbracket 1, n \rrbracket$ et tout $\lambda \in \mathbb{C}$, la matrice :

$$J_\ell(\lambda) = \lambda I_\ell + J_\ell = \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & & \ddots & \lambda & 1 \\ 0 & \cdots & \cdots & 0 & \lambda \end{pmatrix}$$

Dans le cas où $n = 1$, la matrice $J_1(\lambda)$ est réduite à la matrice à un seul coefficient (λ).

Alors, toute matrice de $\mathcal{M}_n(\mathbb{C})$ est semblable, à permutation près des blocs diagonaux, à une unique matrice

$$\begin{pmatrix} J_{\ell_1}(\lambda_1) & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & J_{\ell_k}(\lambda_k) \end{pmatrix}, \quad \ell_1 + \cdots + \ell_k = n, \quad (\lambda_1, \dots, \lambda_k) \in \mathbb{C}^k,$$

matrice constituée de blocs diagonaux carrés égaux aux matrices $J_{\ell_i}(\lambda_i)$.

Le problème de la réduction d'un endomorphisme u , ou de façon équivalente, d'une matrice M , est de trouver une base relativement à laquelle la matrice de u est de la forme ci-dessus, ou de façon équivalente, de trouver une matrice P inversible telle que $P^{-1}MP$ soit de la forme ci-dessus.

Si les ℓ_k sont tous égaux à 1, la matrice obtenue est diagonale. Ainsi, les matrices diagonales (à permutation près des facteurs) font partie de la famille décrite ci-dessus. Le problème de la diagonalisation est donc un sous-problème du problème plus vaste de la réduction, et si elle n'aboutit pas, il faudra chercher un représentant de la classe de M sous le format plus général donné ci-dessus. Nous définissons :

Définition 22.4.28 (Endomorphisme diagonalisable, matrice diagonalisable)

- (i) Un endomorphisme est diagonalisable s'il existe une base dans laquelle sa matrice est diagonale.
- (ii) Une matrice est diagonalisable si elle est semblable à une matrice diagonale.

Trouver une telle matrice diagonale (et la matrice de passage associée) s'appelle « diagonaliser » l'endomorphisme ou la matrice. Vous verrez l'année prochaine des techniques efficaces pour diagonaliser un endomorphisme ou une matrice.

De façon plus générale, trouver un représentant sous la forme générale exposée ci-dessus (et la matrice de passage correspondante) s'appelle « réduire » ou « jordaniser » l'endomorphisme ou la matrice.

Pour la culture, le résultat qui est à la base du théorème de jordanisation, disant que toute matrice à coefficients complexes est jordanisable, est le théorème suivant :

Théorème 22.4.29 (Théorème des noyaux itérés)

Soit u un endomorphisme de E de dimension finie, on a les inclusions :

$$\text{Ker}(u^0) \subset \text{Ker}(u^1) \subset \text{Ker}(u^2) \subset \cdots \subset \text{Ker}(u^k) \subset \text{Ker}(u^{k+1}) \subset \cdots$$

De plus, cette suite est stationnaire, et « s'essouffle », dans le sens où les sauts de dimension forment une suite décroissante, ultimement nulle.

Nous verrons la démonstration de ce théorème en exercice.

V Produit matriciel par blocs

Nous voyons dans cette dernière section qu'on peut effectuer un produit matriciel en groupant les termes par blocs rectangulaires, en utilisant les règles usuelles, c'est-à-dire en remplaçant dans les formules les coefficients par des blocs matriciels. Il faut cependant prendre garde au fait que cela n'est possible que si le découpage en blocs des deux matrices A et B fournit des formats compatibles pour les produits intervenant de la sorte.

Soit $\mathcal{B} = (b_1, \dots, b_p)$ une base de E , $\mathcal{C} = (c_1, \dots, c_n)$ une base de F . Soit $u \in \mathcal{L}(E, F)$ et soit $A = \text{Mat}_{\mathcal{B}, \mathcal{C}}(u) = (a_{i,j})$. On définit un découpage par blocs de M par les indices

$$0 = i_0 < i_1 < i_2 < \dots < i_{q-1} < i_q = n \quad \text{et} \quad 0 = j_0 < j_1 < j_2 < \dots < j_{r-1} < j_r = p.$$

Les blocs sont alors définis par $A_{k,\ell} = (a_{i,j})_{(i,j) \in [i_{k-1}+1, i_k] \times [j_{\ell-1}+1, j_\ell]}$. La matrice $A_{k,\ell}$ est donc la matrice obtenue par extraction des lignes $i_{k-1}+1$ à i_k et des colonnes $j_{\ell-1}+1$ à j_ℓ de A . On dira que le découpage de A en blocs $(A_{k,\ell})$ est associées aux suites d'indices

$$0 = i_0 < i_1 < i_2 < \dots < i_{q-1} < i_q = n \quad \text{et} \quad 0 = j_0 < j_1 < j_2 < \dots < j_{r-1} < j_r = p.$$

Par ailleurs, étant donné E , muni de la base $\mathcal{B} = (b_1, \dots, b_p)$, nous définissons la décomposition de E associée à la suite

$$0 = j_0 < j_1 < j_2 < \dots < j_{r-1} < j_r = p$$

par

$$E = E_1 \oplus \dots \oplus E_r,$$

où pour tout $k \in \llbracket 1, r \rrbracket$, $E_k = \text{Vect}(\mathcal{B}_k)$, où $\mathcal{B}_k = (b_{j_{k-1}+1}, \dots, b_{j_k})$.

Nous noterons $i_\ell : E_\ell \rightarrow E$, l'injection canonique d'inclusion de E_ℓ dans E .

De même, nous disposons alors d'une décomposition de F associée à la base \mathcal{C} , et à la suite

$$0 = i_0 < i_1 < i_2 < \dots < i_{q-1} < i_q = n,$$

décomposition que nous écrivons $F = F_1 \oplus \dots \oplus F_q$. La suite (j_i) définit également des tranches adaptées \mathcal{C}_ℓ de la base \mathcal{C} , comme ci-dessus.

Enfin, étant donnée la décomposition en sommes directes $F = F_1 \oplus \dots \oplus F_q$, on appelle projection sur le k -ième facteur l'application linéaire de $\mathcal{L}(E, E_k)$ définie par :

$$p_k(x_1 + \dots + x_q) = x_k.$$

Il s'agit d'un projecteur, et on a de façon évidente,

$$\forall x \in F, \quad x = p_1(x) + \dots + p_q(x).$$

Théorème 22.5.1 (Interprétation géométrique d'un découpage par blocs)

Avec les notations introduites ci-dessus, pour tout $(k, \ell) \in \llbracket 1, q \rrbracket \times \llbracket 1, r \rrbracket$, $A_{k,\ell}$ est la matrice de l'application de $\mathcal{L}(E_\ell, F_k)$, obtenue par projection sur F_k de la restriction de u à E_ℓ relativement aux bases \mathcal{B}_ℓ et \mathcal{C}_k :

$$A_{k,\ell} = \text{Mat}_{\mathcal{B}_\ell, \mathcal{C}_k}(p_k \circ u \circ i_\ell).$$

De cette interprétation, on déduit le théorème de produit par blocs :

Théorème 22.5.2 (Produit par blocs)

Soit $A \in \mathcal{M}_{n,p}$ et $B \in \mathcal{M}_{p,m}$ deux matrices, et

$$\begin{aligned} 0 &= i_0 < i_1 < i_2 < \dots < i_{q-1} < i_q = n, \\ 0 &= j_0 < j_1 < j_2 < \dots < j_{r-1} < j_r = p, \\ 0 &= k_0 < k_1 < k_2 < \dots < k_{s-1} < k_s = m. \end{aligned}$$

Les deux premières suites définissent un découpage par blocs $A = (A_{i,j})_{(i,j) \in \llbracket 1, q \rrbracket \times \llbracket 1, r \rrbracket}$ de A et les deux dernières définissent un découpage par blocs $B = (B_{j,k})_{(j,k) \in \llbracket 1, r \rrbracket \times \llbracket 1, s \rrbracket}$.

Le produit AB admet alors une représentation par blocs :

$$AB = (C_{i,k})_{(i,k) \in \llbracket 1, q \rrbracket \times \llbracket 1, s \rrbracket},$$

où pour tout $(i, k) \in \llbracket 1, q \rrbracket \times \llbracket 1, s \rrbracket$, $C_{i,k} = \sum_{j=1}^r A_{i,j} B_{j,k}$.

En particulier, si on a une représentation diagonale par blocs :

$$A = \begin{pmatrix} A_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & A_k \end{pmatrix},$$

où les A_k sont des matrices carrées, alors pour tout $n \in \mathbb{N}$,

$$A^n = \begin{pmatrix} A_1^n & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & A_k^n \end{pmatrix}.$$

Exemple 22.5.3

Calcul des puissances successives de $\begin{pmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 2 \end{pmatrix}$

Groupe symétrique et déterminants

Nous rappelons que le groupe symétrique \mathfrak{S}_n ou S_n est le groupe des permutations de $\llbracket 1, n \rrbracket$, la loi du groupe étant la composition. L'étude algébrique des groupes \mathfrak{S}_n est très intéressante, et aboutit à des résultats aussi importants que la non résolubilité par radicaux des équations de degré au moins 5. Cette étude va bien au-delà des objectifs du programme, qui sont uniquement d'introduire les outils nécessaires pour pouvoir définir correctement les déterminants. Nous nous limiterons donc, avec beaucoup de regrets, à ces objectifs.

I Groupe symétrique

Notre but est d'introduire les outils nécessaires à la définition de la signature d'une permutation, notion utilisée ensuite dans l'étude des déterminants. La signature d'une permutation est défini comme l'unique morphisme de groupes $\varepsilon : \mathfrak{S}_n \mapsto \{-1, 1\}$ non constant. Notre objectif est donc de montrer l'existence et l'unicité de ce morphisme.

I.1 Notations liées à des permutations

Nous rencontrerons deux façons de décrire des permutations. La première, la plus naturelle, est d'associer, dans un tableau, à chaque valeur $i \in \llbracket 1, n \rrbracket$, la valeur de $\sigma(i)$, en rangeant les valeurs de i dans l'ordre croissant. Nous disposerons ces données sous la forme matricielle suivante :

Notation 23.1.1 (Permutation)

Soit $\sigma \in \mathfrak{S}_n$. Nous désignerons explicitement σ par le tableau :

$$\sigma = \begin{pmatrix} 1 & 2 & \cdots & n \\ \sigma(1) & \sigma(2) & \cdots & \sigma(n) \end{pmatrix}$$

Exemple 23.1.2

1. Décrire de la sorte tous les éléments de \mathfrak{S}_2 et \mathfrak{S}_3 .
2. Décrire de la sorte la permutation de \mathfrak{S}_n inversant l'ordre : $\sigma(k) = n + 1 - k$.

Certaines permutations ont un comportement particulier : elles effectuent une rotation sur un ensemble d'éléments (qu'on peut ranger en cercle), et laissent les autres invariants. On appelle *cycle* ces permutations. Pour rendre le caractère cyclique plus visible sur les notations on adopte pour ces permutations une notation plus adaptée. Nous définissons de façon plus précise :

Définition 23.1.3 (Cycle)

Un cycle est la donnée d'un sous ensemble I de $\llbracket 1, n \rrbracket$ et d'une permutation σ de $\llbracket 1, n \rrbracket$ telle qu'il existe i_1, \dots, i_k deux à deux distincts tels que $I = \{i_1, \dots, i_k\}$, et :

- (i) $\forall \ell \in \llbracket 1, k-1 \rrbracket, \sigma(i_\ell) = i_{\ell+1}$ et $\sigma(i_k) = i_1$
- (ii) $\forall i \in \llbracket 1, n \rrbracket \setminus I, \sigma(i) = i$.

Remarque 23.1.4

- Sauf dans un cas bien particulier, la donnée de I est implicitement fournie par la donnée de σ : il s'agit de l'ensemble des points de $\llbracket 1, n \rrbracket$ que σ ne laisse pas fixe.
- Le cas problématique pour lequel I ne découle pas de la seule donnée de σ est le cas d'un cycle de longueur 1 (donc I est un singleton). Dans ce cas σ est l'identité, et tous les points de σ sont invariants. La définition que nous donnons d'un cycle (comme étant un couple formé d'un sous-ensemble et d'une permutation) permet de différencier les cycles de longueur 1. Par exemple les cycles (1) et (2) correspondent à la même permutation, mais il ne s'agit pas du même cycle, n'ayant pas même support.
- Dans la suite, nous identifierons les cycles et les permutations qu'ils définissent, en gardant en tête cette différenciation possible des supports dans le cas de cycles de longueur 1.

Exemple 23.1.5

Sont-ce des cycles : $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 4 & 5 & 3 & 2 & 6 \end{pmatrix}$? $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 4 & 5 & 2 & 6 & 1 \end{pmatrix}$?

Notation 23.1.6 (Cycles)

Avec les conditions de la définition 23.1.3, on note

$$\sigma = (i_1 \ i_2 \ \dots \ i_k).$$

On appelle support du cycle l'ensemble I . La longueur du cycle est l'entier $k = \text{Card}(I)$.

Cette notation signifie que tout élément de la liste est envoyé sur l'élément suivant, le dernier étant envoyé sur le premier.

Exemple 23.1.7

1. Écrire le cycle de l'exemple 23.1.5 avec les notations 23.1.6
2. Écrire le cycle (1 5 8 2 9 7 3) avec les notations 23.1.1

Avertissement 23.1.8

Attention, l'ordre des éléments i_1, \dots, i_n est important.

Remarque 23.1.9

La représentation d'un cycle sous la forme 23.1.6 est-elle unique ?

On rencontre parfois :

Définition 23.1.10 (Grand cycle)

Un grand cycle de \mathfrak{S}_n est un cycle de support $\llbracket 1, n \rrbracket$.

Ainsi, un grand cycle effectue une rotation, dans un certain ordre, entre les n éléments de $\llbracket 1, n \rrbracket$. Un exemple important de grand cycle est :

Définition 23.1.11 (Permutation circulaire)

- (i) On appelle permutation circulaire directe de \mathfrak{S}_n le grand cycle $(1\ 2\ \dots\ n)$.
- (ii) On appelle permutation circulaire indirecte de \mathfrak{S}_n le grand cycle $(n\ \dots\ 2\ 1)$.

Exemple 23.1.12

Écrire les permutations circulaires directes et indirectes avec les notations 23.1.11

Une famille importante de cycles est constituée des cycles de longueur 2. En effet, il s'agit d'une famille génératrice de \mathfrak{S}_n , comme on le verra plus tard.

Définition 23.1.13 (Transpositions)

On appelle transposition de \mathfrak{S}_n un cycle de longueur 2 : $\tau = (i\ j)$.

Ainsi, la transposition $\tau = (i\ j)$ est la permutation consistant en l'échange des valeurs i et j .

Exemple 23.1.14

Écrire la transposition $\tau = (2\ 5)$ de \mathfrak{S}_6 avec les notations 23.1.11.

I.2 Signature d'une permutation**Lemme 23.1.15**

Soit $\sigma \in \mathfrak{S}_n$. Alors σ induit une bijection sur $\mathcal{P}_2(\llbracket 1, n \rrbracket)$, l'ensemble des sous-ensembles de cardinal 2 de $\llbracket 1, n \rrbracket$.

On en déduit :

Lemme 23.1.16

Soit $\sigma \in \mathfrak{S}_n$. On a :

$$\left| \prod_{1 \leq i < j \leq n} (\sigma(j) - \sigma(i)) \right| = \prod_{1 \leq i < j \leq n} (j - i)$$

Puisque $\frac{\sigma(j) - \sigma(i)}{j - i} = \frac{\sigma(i) - \sigma(j)}{i - j}$, cette quantité ne dépend pas de l'ordre respectif entre i et j , mais uniquement de l'ensemble $\{i, j\}$. On peut donc définir une application δ sur $\mathcal{P}_2(\llbracket 1, n \rrbracket)$ par :

$$X = \{i, j\} \mapsto \delta(X) = \frac{\sigma(i) - \sigma(j)}{i - j}.$$

Définition 23.1.17 (Signature)

On définit la signature $\varepsilon : \mathfrak{S}_n \rightarrow \{-1, 1\}$ par :

$$\varepsilon(\sigma) = \frac{\prod_{1 \leq i < j \leq n} (\sigma(j) - \sigma(i))}{\prod_{1 \leq i < j \leq n} (j - i)} = \prod_{X \in \mathcal{P}_2(\llbracket 1, n \rrbracket)} \delta(X).$$

Théorème 23.1.18

La signature ε est un morphisme de groupes de (\mathfrak{S}_n, \circ) dans $(\{-1, 1\}, \times)$.

De l'étude du signe de l'expression définissant $\varepsilon(\sigma)$, il découle que celui-ci dépend du nombre de couples (i, j) avec $i < j$ tel que $\sigma(i) > \sigma(j)$.

Définition 23.1.19 (Inversion)

Soit $\sigma \in \mathfrak{S}_n$. On appelle inversion de σ un couple (i, j) tel que $i < j$ et $\sigma(i) > \sigma(j)$, c'est-à-dire $\delta(\{i, j\}) < 0$.

Ainsi, la signature peut être décrite à l'aide du nombre d'inversions :

Théorème 23.1.20 (Description de la signature par les inversions)

En notant $\text{Inv}(\sigma)$ le nombre d'inversions de σ , on a :

$$\varepsilon(\sigma) = (-1)^{\text{Inv}(\sigma)}.$$

Une transposition $(i \ i + 1)$ n'admettant que le couple $(i, i + 1)$ comme inversion, il vient alors :

Proposition 23.1.21

Notons, pour tout $i \in \llbracket 1, n - 1 \rrbracket$, $\tau_i = (i \ i + 1)$. On a alors $\varepsilon(\tau_i) = -1$.

Or, toute transposition peut s'écrire à l'aide de ces transpositions entre éléments consécutifs :

Proposition 23.1.22 (Décomposition d'une transposition à l'aide des τ_i)

Soit $1 \leq i < j \leq n$, et $\tau = (i \ j)$. Alors

$$\tau = \tau_{j-1} \circ \cdots \circ \tau_{i+1} \circ \tau_i \circ \cdots \circ \tau_{j-2} \circ \tau_{j-1}.$$

En utilisant le fait que ε est un morphisme, on en déduit :

Théorème 23.1.23 (Signature d'une transposition)

Soit τ une transposition. Alors $\varepsilon(\tau) = -1$.

Nous avons donc répondu au problème de l'existence d'un morphisme $\varepsilon : \mathfrak{S}_n \rightarrow \{-1, 1\}$, prenant la valeur -1 sur les transpositions.

Pour étudier son unicité, nous allons établir que toute permutation s'écrit comme composition de transposition. Ainsi, la valeur d'un morphisme ε sur une permutation est entièrement déterminée par la valeur de ε sur les transpositions. Imposer la valeur sur les transpositions fournit dans ce cas l'unicité.

Théorème 23.1.24 (Caractère générateur des transpositions)

Toute permutation $\sigma \in \mathfrak{S}_n$ est un produit de transpositions.

Les transpositions étant elles mêmes engendrées par les τ_i , on se rend compte que toute permutation s'écrit comme produit des τ_i . Cela ne soit pas nous étonner, on l'avait déjà prouvé en étudiant la correction de certains algorithmes de tri (lesquels?)

Remarque 23.1.25

Le rapport obtenu entre signature et décomposition en produit de transpositions permet d'affirmer que la parité du nombre de terme de toute décomposition en produits de transpositions d'une permutation donnée est la même.

Terminologie 23.1.26 (Permutation paire, permutation impaire)

On dira qu'une permutation est paire si $\varepsilon(\sigma) = 1$, et impaire si $\varepsilon(\sigma) = -1$.

Définition 23.1.27 (Groupe alterné)

Le groupe alterné \mathfrak{A}_n est le noyau de la signature, c'est-à-dire l'ensemble des permutations paires. C'est un sous-groupe distingué de \mathfrak{S}_n .

I.3 Décomposition cyclique d'une permutation

Décomposer une permutation en produit de transpositions, ou compter le nombre d'inversions n'est pas quelque chose d'immédiat. Nous donnons dans ce paragraphe une autre façon, plus rapide, de calculer la signature d'une permutation, en s'aidant du type cyclique de cette permutation.

Soit σ une permutation. On définit sur $\llbracket 1, n \rrbracket$ la relation : $i \equiv_{\sigma} j$ si et seulement s'il existe $k \in \mathbb{N}$ tel que $j = \sigma^k(i)$.

On peut interpréter cette relation sous forme d'une action de groupe : le groupe \mathfrak{S}_n agit sur $\llbracket 1, n \rrbracket$ par $\sigma \cdot i = \sigma(i)$. La relation d'équivalence n'est alors autre que la relation définissant les orbites de $\llbracket 1, n \rrbracket$ sous l'action de \mathfrak{S}_n .

Proposition 23.1.28

La relation \equiv_{σ} est une relation d'équivalence.

Soit $x \in \llbracket 1, n \rrbracket$, et S_x l'unique classe d'équivalence contenant x . Par définition de la relation \equiv_{σ} , $S_x = \{\sigma^i(x), i \in \mathbb{N}\}$.

Proposition 23.1.29

Soit k le cardinal de S_x . Alors les $\sigma^i(x)$, $i \in \llbracket 0, k-1 \rrbracket$, sont deux-à-deux distincts, on a $S_x = \{x, \sigma(x), \dots, \sigma^{k-1}(x)\}$, et σ induit sur S_x une permutation de S_x , égale au cycle $(x, \sigma(x), \dots, \sigma^{k-1}(x))$.

Notons C_x le cycle de \mathfrak{S}_n défini par

$$C_x = (x, \sigma(x), \dots, \sigma^{k-1}(x)).$$

Lemme 23.1.30

Soit x et y des éléments d'une même classe d'équivalence modulo \equiv_σ . Alors $C_x = C_y$.

Théorème 23.1.31 (Décomposition en cycles d'une permutation)

Soit σ une permutation de \mathfrak{S}_n . À permutation près des facteurs, il existe une unique décomposition de σ en produits de cycles :

$$\sigma = C_1 \circ \cdots \circ C_k,$$

telle que les supports des cycles forment une partition de $\llbracket 1, n \rrbracket$

Cette dernière hypothèse signifie que les cycles considérés sont à supports disjoints, et que tout élément de $\llbracket 1, n \rrbracket$ est dans un cycle, quitte à considérer dans la décomposition des cycles de longueur 1 (en ayant en mémoire la possibilité de différencier leur support).

Remarque 23.1.32

- Les cycles C_i sont les cycles associés à un système de représentants modulo la relation \equiv_σ .
- L'ordre dans lequel on effectue cette composition importe peu, du fait du lemme qui suit.

Lemme 23.1.33 (Commutation des cycles à supports disjoints)

Soit C_1 et C_2 deux cycles à supports disjoints. Alors $C_1 \circ C_2 = C_2 \circ C_1$.

Méthode 23.1.34 (Comment déterminer la décomposition en produits de cycles)

Soit σ une permutation de \mathfrak{S}_n .

- Partir de 1 et suivre ses images successives jusqu'à retomber sur 1. Cela donne le premier cycle
- Recommencer en partant du plus petit élément de $\llbracket 1, n \rrbracket$ n'appartenant pas au premier cycle trouvé
- Recommencer ainsi jusqu'à épuisement des éléments de $\llbracket 1, n \rrbracket$.

Exemple 23.1.35

Trouver la décomposition en cycles à supports disjoints de :

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 5 & 2 & 8 & 10 & 7 & 9 & 1 & 3 & 4 & 6 \end{pmatrix}$$

Définition 23.1.36 (Support cyclique d'une permutation)

Le support cyclique d'une permutation σ est la partition formée des supports des cycles formant la décomposition en cycles de σ . Ainsi, il s'agit de la partition formée des classes d'équivalence de la relation \equiv_σ .

Définition 23.1.37 (Type cyclique d'une permutation)

Le type cyclique d'une permutation σ est la suite croissante des tailles des parts du support cyclique. Une telle suite croissante de somme n est appelée partition de l'entier n .

I.4 Cycles et signature

Nous voyons enfin comment utiliser la décomposition en cycles pour déterminer la signature. Pour cela nous remarquons qu'il est facile d'écrire un cycle comme produit de transpositions :

Lemme 23.1.38 (Décomposition d'un cycle en transpositions)

Soit $\{i_1, \dots, i_k\}$ des entiers 2 à 2 distincts de $\llbracket 1, n \rrbracket$. Alors :

$$(i_1 i_2 \cdots i_k) = (i_1 i_k) \circ (i_1 i_{k-1}) \circ \cdots \circ (i_1 i_2).$$

Dans le cas où $k = 1$, le terme de droite est réduit à un produit vide de transpositions (égal à l'identité)

Proposition 23.1.39 (Signature d'un cycle)

Soit C un cycle et $\ell(C)$ sa longueur. Alors

$$\varepsilon(C) = (-1)^{\ell(C)-1}.$$

Théorème 23.1.40 (Détermination de ε par le type cyclique)

Soit σ une permutation de \mathfrak{S}_n , et $c(\sigma)$ le nombre de parts dans son support cyclique (ou de façon équivalente dans son type cyclique). Alors :

$$\varepsilon(\sigma) = (-1)^{n-c(\sigma)}.$$

II Déterminants

Dans cette section, nous étudions les déterminants. Nous verrons cette notion tout d'abord vue comme un objet défini sur une famille de vecteurs, puis nous en déduisons une notion de déterminant d'un endomorphisme, puis d'une matrice carrée. Notre but étant en grande partie de caractériser l'inversibilité d'une matrice grâce à cet objet, nous introduisons ensuite des techniques calculatoires efficaces. Nous voyons enfin comment donner une expression de l'inverse d'une matrice à l'aide des déterminants, même si en pratique, la formule obtenue manque d'efficacité.

II.1 Formes multilinéaires

Soit \mathbb{K} un corps.

Définition 23.2.1 (Application multilinéaire)

Soit E_1, \dots, E_n et F des \mathbb{K} -espaces vectoriels. Une application n -linéaire est une application

$$\varphi : E_1 \times \cdots \times E_n \longrightarrow F$$

telle que pour tout $i \in \llbracket 1, n \rrbracket$, φ soit linéaire par rapport à sa i -ième variable, les autres étant fixées quelconques, donc si pour tout x_1, \dots, x_n, x'_i et λ ,

$$\begin{aligned} \varphi(x_1, \dots, x_{i-1}, \lambda x_i + x'_i, x_{i+1}, \dots, x_n) \\ = \lambda \varphi(x_1, \dots, x_{i-1}, x_i, x_{i+1}, \dots, x_n) + \varphi(x_1, \dots, x_{i-1}, x'_i, x_{i+1}, \dots, x_n). \end{aligned}$$

une application est une application multilinéaire si elle est n -linéaire pour un certain $n \geq 2$.

Proposition 23.2.2

Soit φ une application n -linéaire. Alors $\varphi(x_1, \dots, x_n) = 0$ dès lors qu'une des variables x_i est nulle.

Définition 23.2.3 (Forme n -linéaire)

Une forme n -linéaire est une application n -linéaire à valeurs dans \mathbb{K} .

Exemple 23.2.4

1. Voici des exemples de formes bilinéaires ($n = 2$) :
 - $(x, y) \in (\mathbb{R}^n)^2 \mapsto \langle x, y \rangle$, le produit scalaire canonique
 - $(x, y) \in \mathbb{C}^2 \mapsto xy$
 - $(f, g) \in \mathcal{L}(E) \mapsto f \circ g$
 - $(X, Y) \in (\mathbb{R}^n)^2 \mapsto {}^tYMX$
2. $(x_1, \dots, x_n) \mapsto x_1 \cdots x_n$
3. $(f_1, \dots, f_n) \mapsto \int_0^1 f_1(t) \cdots f_n(t) dt$
4. L'aire (signée) du parallélogramme défini par deux vecteurs x et y
5. Le volume du parallélépipède défini par trois vecteurs.

Théorème 23.2.5 (n -linéarité généralisée)

Soit $\varphi : E_1 \times \cdots \times E_n \rightarrow F$ une application n -linéaire. Alors pour tout $(k_1, \dots, k_n) \in \mathbb{N}^*$, pour tout $x_{i,j} \in E_i$ et $\lambda_{i,j} \in \mathbb{K}$, ($i \in \llbracket 1, n \rrbracket$, $j \in \llbracket 1, k_i \rrbracket$), on a :

$$\varphi \left(\sum_{i_1=1}^{k_1} \lambda_{1,i_1} a_{1,i_1}, \dots, \sum_{i_n=1}^{k_n} \lambda_{n,i_n} a_{n,i_n} \right) = \sum_{i_1=1}^{k_1} \cdots \sum_{i_n=1}^{k_n} \lambda_{1,i_1} \cdots \lambda_{n,i_n} \varphi(a_{1,i_1}, \dots, a_{n,i_n}).$$

Avertissement 23.2.6

Les indices des sommes doivent être indépendants !

Exemple 23.2.7

Dans le cas de la bilinéarité, on obtient :

$$\varphi \left(\sum_{i=1}^k \lambda_i a_i, \sum_{j=1}^{\ell} \mu_j b_j \right) = \sum_{i=1}^k \sum_{j=1}^{\ell} \lambda_i \mu_j \varphi(a_i, b_j).$$

Proposition 23.2.8 (L'espace des applications n -linéaires)

L'ensemble, noté $\mathcal{L}(E_1, \dots, E_n; F)$ des applications multilinéaires est un espace vectoriel sur \mathbb{K} .

Proposition 23.2.9

On a un isomorphisme d'espaces vectoriels :

$$\mathcal{L}(E_1, \dots, E_n; F) \simeq \mathcal{L}(E_1, \mathcal{L}(E_2, \dots, E_n; F)).$$

Par itération, on en déduit un isomorphisme d'espaces vectoriels :

$$\mathcal{L}(E_1, \dots, E_n; F) \simeq \mathcal{L}(E_1, \mathcal{L}(E_2, \mathcal{L}(E_3, \dots \mathcal{L}(E_n, F))))).$$

Notation 23.2.10

Lorsque $E_1 = \dots = E_n$, on notera plus simplement $\mathcal{L}_n(E; F)$ l'espace des applications n -linéaires de E dans F , et $\mathcal{L}_n(E)$ l'espace des formes n -linéaires

Comme pour les applications linéaires, pour déterminer entièrement une application n -linéaire (et donc en particulier une forme n -linéaire), il suffit d'en connaître l'image sur les vecteurs d'une base.

Proposition 23.2.11 (Détermination d'une application n -linéaire sur une base)

Soit pour tout $i \in \llbracket 1, n \rrbracket$, $(e_{i,j})_{1 \leq j \leq d_i}$ une base de E_i et pour tout $(j_1, \dots, j_n) \in \llbracket 1, d_1 \rrbracket \times \dots \times \llbracket 1, d_n \rrbracket$, f_{j_1, \dots, j_n} un élément de F . Alors il existe une unique application n -linéaire $\varphi : E_1 \times \dots \times E_n \rightarrow F$ telle que

$$\forall (j_1, \dots, j_n) \in \llbracket 1, d_1 \rrbracket \times \dots \times \llbracket 1, d_n \rrbracket, \quad \varphi(e_{1,j_1}, \dots, e_{n,j_n}) = f_{j_1, \dots, j_n}.$$

Cette application est donnée explicitement par :

$$\varphi \left(\sum_{j_1=1}^{d_1} \lambda_{1,j_1} e_{1,j_1}, \dots, \sum_{j_n=1}^{d_n} \lambda_{n,j_n} e_{n,j_n} \right) = \sum_{j_1=1}^{d_1} \dots \sum_{j_n=1}^{d_n} \lambda_{1,j_1} \dots \lambda_{n,j_n} f_{j_1, \dots, j_n}.$$

En particulier, si $E_1 = \dots = E_n = E$, muni d'une unique base (e_1, \dots, e_d) de E , et si on se donne des éléments f_{i_1, \dots, i_n} de F , pour $1 \leq i_1, \dots, i_n \leq d$, il existe une unique application n -linéaire φ de $\mathcal{L}_n(E; F)$ telle que pour tout $(i_1, \dots, i_n) \in \llbracket 1, d \rrbracket^n$,

$$\varphi(e_{i_1}, \dots, e_{i_n}) = f_{i_1, \dots, i_n}.$$

Exemple 23.2.12

Soit (e_1, \dots, e_n) la base canonique de \mathbb{K}^n . Une forme bilinéaire φ sur \mathbb{K}^n est entièrement déterminée par la donnée de n^2 scalaires $\varphi(e_i, e_j)$. En notant M la matrice carrée définie par ces scalaires ($M = (\varphi(e_i, e_j))_{1 \leq i, j \leq n}$), on peut vérifier que pour tout $X, Y \in \mathbb{K}^n$ (qu'on identifie au vecteur colonne des coordonnées dans la base canonique), on a :

$$\varphi(X, Y) = {}^t X M Y.$$

Plus généralement, toute forme bilinéaire sur un espace vectoriel E de dimension finie s'écrit de la sorte après choix d'une base de E (voir un chapitre ultérieur).

II.2 Formes n -linéaires symétriques, antisymétriques, alternées

Nous définissons maintenant deux types important de formes n -linéaires prenant leurs variables dans un même espace.

Définition 23.2.13 (Formes n -linéaires symétriques, alternées)

Soit $\varphi \in \mathcal{L}_n(E)$ une forme n -linéaire. On dit que :

1. φ est symétrique si elle est invariante par permutation de ses variables : pour tout $\sigma \in \mathfrak{S}_n$,

$$\varphi(x_1, \dots, x_n) = \varphi(x_{\sigma(1)}, \dots, x_{\sigma(n)}).$$

2. φ est antisymétrique si pour tout $\sigma \in \mathfrak{S}_n$,

$$\varphi(x_1, \dots, x_n) = \varepsilon(\sigma)\varphi(x_{\sigma(1)}, \dots, x_{\sigma(n)}).$$

3. φ est alternée si $\varphi(x_1, \dots, x_n) = 0$ dès lors qu'il existe $i \neq j$ tels que $x_i = x_j$.

Exemple 23.2.14

1. Le produit scalaire canonique est une forme bilinéaire symétrique
2. $(A, B) \mapsto \text{tr}(AB)$ est une forme bilinéaire symétrique
3. $(A, B, C) \mapsto \text{tr}(ABC)$ est-elle symétrique ?
4. L'aire du parallélogramme formé par deux vecteurs de \mathbb{R}^2 est une forme bilinéaire alternée.
5. Le volume du parallélépipède formé par trois vecteurs de \mathbb{R}^3 est une forme bilinéaire alternée.

Lemme 23.2.15

Soit φ une forme alternée. Alors pour tout $(x_1, \dots, x_n) \in E^n$, et tout $(i, j) \in \llbracket 1, n \rrbracket^2$ tel que $i \neq j$:

$$\varphi(x_1, \dots, x_i, \dots, x_j, \dots, x_n) = -\varphi(x_1, \dots, x_j, \dots, x_i, \dots, x_n).$$

Réciproquement, si cette condition est satisfaite, alors, si \mathbb{K} n'est pas de caractéristique 2, φ est alternée.

Théorème 23.2.16 (Antisymétrie des formes alternées)

Toute forme n -linéaire alternée est antisymétrique. Si \mathbb{K} n'est pas de caractéristique 2, toute forme antisymétrique est alternée.

Proposition 23.2.17 (Image d'une famille liée par une forme alternée)

Soit (x_1, \dots, x_n) une famille liée, et φ une forme alternée. Alors $\varphi(x_1, \dots, x_n) = 0$.

Nous aurons l'occasion de reparler de formes n -linéaires symétriques (et plus particulièrement pour $n = 2$), lorsque nous étudierons les produits scalaires (la symétrie étant une des 3 propriétés requises pour qu'une forme bilinéaire définisse un produit scalaire). Pour l'heure, nous nous concentrons sur la notion de forme alternée, fortement liée, d'après les exemples, aux problèmes de calculs d'aires et de volumes. Nous allons voir qu'à un scalaire multiplicatif près, en dimension n , la mesure des hypervolumes est la seule forme n -linéaire.

II.3 Déterminant d'une famille de vecteurs

Théorème 23.2.18 (Détermination des formes alternées par l'image d'une base)

Soit E un espace vectoriel de dimension finie, muni d'une base (e_1, \dots, e_n) , et $p \in \mathbb{N}^*$.

- Si $p > n$, la seule forme p -linéaire alternée sur E est la forme nulle.
- Si $p \leq n$, une forme p -linéaire alternée φ est complètement déterminée, de façon unique, par la donnée de $\varphi(e_{i_1}, \dots, e_{i_p})$, pour $1 \leq i_1 < \dots < i_p \leq n$.

Plus précisément, il s'agit de l'unique forme p -linéaire sur E telle que pour tout $(i_1, \dots, i_p) \in \llbracket 1, n \rrbracket^p$,

$$\varphi(e_{i_1}, \dots, e_{i_p}) = \begin{cases} 0 & \text{s'il existe } j \neq k \text{ tel que } i_j = i_k \\ \varepsilon(\tau)\varphi(e_{i_{\tau(1)}}, \dots, e_{i_{\tau(p)}}) & \text{sinon,} \end{cases}$$

où $\tau \in \mathfrak{S}_p$ est l'unique permutation de $\llbracket 1, p \rrbracket$ telle que

$$i_{\tau(1)} < \dots < i_{\tau(p)}.$$

Corollaire 23.2.19 (Formes n -linéaires d'un espace de dimension n)

Soit E un espace vectoriel de dimension n , et (e_1, \dots, e_n) une base de E .

1. Il existe une unique forme n -linéaire alternée φ sur E telle que $\varphi(e_1, \dots, e_n) = 1$.
2. Cette forme n -linéaire est entièrement décrite sur les vecteurs de la base par :

$$\begin{cases} \varphi(e_{i_1}, \dots, e_{i_n}) = 0 & \text{s'il existe } j \neq k \text{ tel que } i_j = i_k \\ \varphi(e_{\sigma(1)}, \dots, e_{\sigma(n)}) = \varepsilon(\sigma) & \text{où } \sigma \in \mathfrak{S}_n \end{cases}$$

3. Toute autre forme n -linéaire alternée sur E est de la forme $\lambda\varphi$, $\lambda \in \mathbb{K}$.

Remarques 23.2.20

- Si la première condition n'est pas satisfaite, les e_{i_k} sont deux à deux distincts, et en nombre n , donc sont constitués des éléments e_i pris chacun une et une seule fois. Aussi la famille $(e_{i,k})$ peut-elle être vue comme une permutation de la famille (e_i) , ce qui permet de l'écrire sous la forme $(e_{\sigma(1)}, \dots, e_{\sigma(n)})$.
- La correspondance entre la permutation σ du corollaire précédent, et la permutation τ qui précède le corollaire est (lorsque $p = n$) : $\sigma = \tau^{-1}$.

Cette unique forme n -linéaire alternée va être notre définition du déterminant (par rapport à une base \mathcal{B}).

Définition 23.2.21 (Déterminant d'une famille de vecteurs)

Soit E un espace de dimension n , $\mathcal{B} = (e_1, \dots, e_n)$ une base de E et (x_1, \dots, x_n) une famille de n vecteurs de E . Soit $\det_{\mathcal{B}}$ l'unique forme n -linéaire alternée telle que $\det_{\mathcal{B}}(e_1, \dots, e_n) = 1$. Le *déterminant de la famille (x_1, \dots, x_n) par rapport à \mathcal{B}* est le scalaire $\det_{\mathcal{B}}(x_1, \dots, x_n)$.

Ainsi, le déterminant par rapport à \mathcal{B} est l'unique forme n -linéaire alternée prenant la valeur 1 sur la famille \mathcal{B} .

Remarque 23.2.22

Si E est un \mathbb{R} -espace vectoriel, cette notion est à relier à la notion d'hypervolume relative à une base : le déterminant de n vecteurs par rapport à \mathcal{B} est le volume (orienté) du parallélépipède défini par les n vecteurs, l'unité de volume étant le parallélépipède défini par les vecteurs de la base \mathcal{B} .

Exemples 23.2.23

1. Voir l'interprétation géométrique dans \mathbb{R}^2 et \mathbb{R}^3 , muni de la base canonique, puis dans \mathbb{R}^2 muni d'une base quelconque.

2. Déterminant par rapport à la base canonique de $\left(\begin{pmatrix} a \\ b \end{pmatrix}, \begin{pmatrix} c \\ d \end{pmatrix} \right)$.

3. Déterminant par rapport à la base canonique de $\left(\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}, \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}, \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} \right)$.

Nous obtenons plus généralement la description suivante :

Théorème 23.2.24 (Description du déterminant par les coordonnées)

Soit E un espace vectoriel de dimension n et $\mathcal{B} = (e_1, \dots, e_n)$ une base de E . Soit (x_1, \dots, x_n) une famille d'éléments de E , dont les coordonnées sont :

$$\forall i \in \llbracket 1, n \rrbracket, [x_i]_{\mathcal{B}} = \begin{pmatrix} a_{1,i} \\ \vdots \\ a_{n,i} \end{pmatrix}, \text{ soit: } x_i = \sum_{j=1}^n a_{j,i} e_j.$$

On a alors :

$$\det_{\mathcal{B}}(x_1, \dots, x_n) = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) a_{\sigma(1),1} \cdots a_{\sigma(n),n} = \sum_{\tau \in \mathfrak{S}_n} \varepsilon(\tau) a_{1,\tau(1)} \cdots a_{n,\tau(n)}.$$

Nous rappelons que nous avons obtenu :

Corollaire 23.2.25 (Formes n -linéaires alternées)

Soit E de dimension n , et \mathcal{B} une base de E . Alors l'ensemble des formes n -linéaires alternées est $\text{Vect}(\det_{\mathcal{B}})$.

Ainsi, changer de base ne nous fait pas sortir de la droite $\text{Vect}(\det_{\mathcal{B}})$. Comme par ailleurs, une autre base \mathcal{B}' , définira également une droite $\text{Vect}(\det_{\mathcal{B}'})$, aussi égale à l'ensemble des formes n -linéaires alternées, on peut donc affirmer que $\det_{\mathcal{B}}$ et $\det_{\mathcal{B}'}$ diffèrent d'une constante multiplicative. En évaluant en \mathcal{B}' , on obtient alors :

Proposition 23.2.26 (Effet d'un changement de base sur le déterminant)

Soit \mathcal{B} et \mathcal{B}' deux bases de E . On a alors :

$$\det_{\mathcal{B}} = \det_{\mathcal{B}}(\mathcal{B}') \times \det_{\mathcal{B}'}$$

Remarque 23.2.27

En échangeant le rôle de \mathcal{B} et \mathcal{B}' , ou en évaluant en \mathcal{B} , on obtient la relation suivante :

$$\det_{\mathcal{B}}(\mathcal{B}') = \frac{1}{\det_{\mathcal{B}'}(\mathcal{B})}.$$

Les deux formes $\det_{\mathcal{B}}$ et $\det_{\mathcal{B}'}$ étant non nulles, et ayant déjà vu l'image par une forme alternée d'une famille liée, on obtient la caractérisation suivante :

Proposition 23.2.28 (Caractérisation des bases)

Soit E de dimension fini, muni d'une base \mathcal{B} . Une famille \mathcal{B}' est une base de E si et seulement si $\det_{\mathcal{B}}(\mathcal{B}') \neq 0$.

Nous verrons un peu plus loin que cette caractérisation équivaut à l'inversibilité de la matrice de la famille \mathcal{B}' dans la base \mathcal{B} , cette matrice étant alors la matrice de passage de la base \mathcal{B} à la base \mathcal{B}' .

II.4 Orientation d'un espace

Dans ce paragraphe, on suppose que $\mathbb{K} = \mathbb{R}$.

Suivant l'ordre dans lequel on donne les vecteurs d'une famille, un hypervolume, défini par le déterminant, peut être positif ou négatif. Ainsi, étant donnée une base $\mathcal{B} = (b_1, b_2)$ de \mathbb{R}^2 , $\det_{\mathcal{B}}(b_2, b_1) = -1$. En définissant $\mathcal{B}' = (b_2, b_1)$, on obtient alors $\det_{\mathcal{B}'} = -\det_{\mathcal{B}}$. Ainsi, toutes les aires vont avoir un signe inversé par rapport à \mathcal{B}' , comparées aux aires par rapport à \mathcal{B} . C'est cette propriété des signes qui va définir la notion d'orientation de l'espace. Grossièrement, deux choix de bases donnant des déterminants différant uniquement d'une constante, soit les déterminants associés à ces deux bases prennent des valeurs qui sont toujours de même signe, soit ils prennent systématiquement des valeurs de signe opposé. On va donc pouvoir classer de la sorte les bases en deux groupes. On dira que ces deux groupes définissent deux orientations différentes de E , et que le choix d'un de ces deux groupes (ou d'un représentant, c'est-à-dire d'une base) définit une orientation de E . Parler d'orientation directe ou indirecte relève alors de la convention.

Nous explicitons un peu l'argument ci-dessus.

Soit E un \mathbb{R} -ev de dimension finie. Soit $\mathcal{B}(E)$ l'ensemble des bases de E , et soit \mathcal{R} la relation définie sur $\mathcal{B}(E)$ par :

$$\forall (\mathcal{B}, \mathcal{B}') \in \mathcal{B}(E)^2, \quad \mathcal{B}\mathcal{R}\mathcal{B}' \iff \det_{\mathcal{B}}(\mathcal{B}') > 0.$$

Proposition 23.2.29

La relation \mathcal{R} est une relation d'équivalence. Pour cette relation d'équivalence, il existe précisément deux classes d'équivalence.

Définition 23.2.30 (Orientation de E)

Une orientation de E est une classe d'équivalence de E modulo \mathcal{R} . Orienter E signifie choisir l'une de ces classes d'équivalence, par exemple par le choix d'un représentant (donc d'une base).

Il y a donc deux orientations de E , donc deux choix possibles d'orientation.

Définition 23.2.31 (Orientation directe, indirecte)

Soit E , muni d'une base de référence \mathcal{B} . On dira qu'une orientation définie par \mathcal{B}' est :

- directe (par rapport à \mathcal{B}) si $\det_{\mathcal{B}}(\mathcal{B}') > 0$ (donc $\mathcal{B}\mathcal{R}\mathcal{B}'$)
- indirecte sinon.

Remarque 23.2.32

La notion d'orientation directe ou indirecte relève de la convention : elle ne peut se définir dans l'absolue, et nécessite la donnée d'une base de référence. Dans certaines situations, il existe une base de référence naturelle. Ainsi, dans le cas de \mathbb{R}^n , la base canonique constituera en général la base directe de référence. Par ailleurs, les espaces vectoriels (y compris \mathbb{R}^n) sont définis indépendamment de toute représentation planaire, spatiale ou autre. Ainsi, définir une orientation directe par certaines caractéristiques d'une représentation planaire ou spatiale relève également de la convention, mais s'avère bien pratique, en physique notamment. Remarquez qu'une telle convention pour \mathbb{R}^2 est inversée si on regarde le plan par l'autre côté ! De même pour les conventions dans \mathbb{R}^3 , si on plonge l'espace \mathbb{R}^3 dans \mathbb{R}^4 , et qu'on le regarde par deux côtés différents.

Proposition 23.2.33 (Effet d'une permutation des vecteurs sur l'orientation)

Soit $\mathcal{B} = (b_1, \dots, b_n)$ une base de E , et $\sigma \in \mathfrak{S}_n$. Soit $\mathcal{B}' = (b_{\sigma(1)}, \dots, b_{\sigma(n)})$ la base obtenue de \mathcal{B} par permutation de ses éléments. Alors \mathcal{B} et \mathcal{B}' définissent la même orientation si et seulement si $\varepsilon(\sigma) = 1$, donc si $\sigma \in \mathfrak{A}_n$.

En particulier, l'échange de deux vecteurs modifie l'orientation, puisqu'une transposition est impaire.

II.5 Déterminant d'un endomorphisme

On aimerait définir la notion de déterminant d'un endomorphisme indépendamment du choix d'une base. L'idée naturelle qui vient à l'esprit pourrait être de se fixer une base $\mathcal{B} = (b_1, \dots, b_n)$ et de définir le déterminant d'un endomorphisme u comme le déterminant de $(u(b_1), \dots, u(b_n))$ par rapport à \mathcal{B} . Cette définition est correcte (elle sera donnée en propriété), mais a l'inconvénient d'introduire une base \mathcal{B} , dont nous aimerions nous affranchir. Prenant une autre base, par exemple $(\lambda b_1, \dots, \lambda b_n)$, nous allons alors multiplier le déterminant de la famille par λ^n , mais également le déterminant de la base (toujours relativement à la base initiale). Pour s'affranchir de la notion de base, on peut constater que cela reste vrai pour toute forme n -linéaire alternée (puisque elles diffèrent du déterminant d'une constante multiplicative seulement). Cela nous motive la définition suivante :

Lemme 23.2.34

Soit $u \in \mathcal{L}(E)$. Soit, pour toute forme n -linéaire alternée φ sur E , φ_u définie par

$$\varphi_u(x_1, \dots, x_n) = \varphi(u(x_1), \dots, u(x_n)).$$

Alors φ_u est une forme n -linéaire alternée.

Proposition/Définition 23.2.35 (Déterminant d'un endomorphisme)

Avec les notations du lemme précédent, il existe un unique scalaire $\det(u)$ tel que pour toute forme n -linéaire alternée φ , on ait :

$$\varphi_u = \det(u) \cdot \varphi.$$

Ce scalaire $\det(u)$ est appelé déterminant de u .

On a alors la propriété attendue :

Proposition 23.2.36 (Caractérisation du déterminant par l'image d'une base)

Soit $\mathcal{B} = (b_1, \dots, b_n)$ une base de E . Alors

$$\det(u) = \det_{\mathcal{B}}(u(b_1), \dots, u(b_n)).$$

Corollaire 23.2.37 (Déterminant de l'identité)

On a $\det(\text{id}) = 1$.

Proposition 23.2.38

Soit n la dimension de E . Soit $u \in \mathcal{L}(E)$ et $\lambda \in \mathbb{K}$. On a $\det(\lambda u) = \lambda^n \det(u)$.

Théorème 23.2.39 (Déterminant d'une composée)

Soit u et v dans $\mathcal{L}(E)$. Alors $\det(v \circ u) = \det(v) \det(u)$.

Enfin, notre but étant de caractériser les matrices inversibles (donc les automorphismes), on obtient :

Théorème 23.2.40 (Caractérisation des automorphismes par le déterminant)

Soit $u \in \mathcal{L}(E)$. Alors u est un automorphisme de E si et seulement si $\det(u) \neq 0$.

II.6 Déterminant d'une matrice carrée

Comme souvent, passer des endomorphismes aux matrices est assez automatique, en utilisant la correspondance usuelle entre des endomorphismes et leur matrice dans un choix de base. La seule difficulté peut résider dans l'indépendance vis-à-vis du choix effectué des bases.

Définition 23.2.41 (Déterminant d'une matrice)

Soit $A \in \mathcal{M}_n(\mathbb{K})$, et $f \in \mathcal{L}(\mathbb{K}^n)$ canonique associé. Le déterminant de M , noté $\det(M)$, est par définition égal au déterminant de f : $\det(M) = \det(f)$.

On obtient alors directement des définitions :

Proposition 23.2.42 (Caractérisation du déterminant de M par les colonnes)

Soit M une matrice de $\mathcal{M}_n(\mathbb{K})$, de colonnes C_1, \dots, C_n . Alors

$$\det(M) = \det_{b.c.}(C_1, \dots, C_n).$$

Les résultats obtenus pour les endomorphismes se transfèrent alors de façon immédiate aux matrices :

Corollaire 23.2.43 (Déterminant de I_n)

On a $\det(I_n) = 1$.

Proposition 23.2.44 (Effet de la multiplication par un scalaire)

Soit $A \in \mathcal{M}_n(\mathbb{K})$ et $\lambda \in \mathbb{K}$. On a $\det(\lambda A) = \lambda^n \det(A)$.

Théorème 23.2.45 (Déterminant d'un produit)

Soit A et B dans $\mathcal{M}_n(\mathbb{K})$. Alors $\det(AB) = \det(A) \det(B)$.

Et pour terminer, la caractérisation annoncée depuis le début :

Théorème 23.2.46 (Caractérisation des matrices inversibles)

Soit $A \in \mathcal{M}_n(\mathbb{K})$. Alors $A \in \text{GL}_n(\mathbb{K})$ si et seulement si $\det(A) \neq 0$. Si cette condition est satisfaite, on a alors :

$$\det(A^{-1}) = \frac{1}{\det(A)}.$$

Ces règles nous permettent d'établir le fait que le déterminant est un invariant de similitude :

Théorème 23.2.47 (Cohérence relative au choix des bases)

Soit $f \in \mathcal{L}(E)$. Alors pour toute base \mathcal{B} de E ,

$$\det(f) = \det(\text{Mat}_{\mathcal{B}}(f)).$$

Pour des matrices décrites par la donnée tabulaire de leurs coefficients, on utilise souvent la notation suivante :

Notation 23.2.48 (Notation tabulaire du déterminant)

Soit $A = (a_{i,j}) \in \mathcal{M}_n(\mathbb{K})$. On note

$$\det(A) = \begin{vmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,n} \end{vmatrix}.$$

Nous transcrivons aussi la description initiale que nous avons donnée pour les familles de vecteurs grâce aux permutations :

Théorème 23.2.49 (Expression du déterminant par les coefficients)

Soit $A = (a_{i,j}) \in \mathcal{M}_n(\mathbb{K})$. Alors

$$\det(A) = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) a_{\sigma(1),1} \cdots a_{\sigma(n),n} = \sum_{\tau \in \mathfrak{S}_n} \varepsilon(\tau) a_{1,\tau(1)} \cdots a_{n,\tau(n)}.$$

Corollaire 23.2.50 (Expression des déterminant 2×2 et 3×3)

1. On a : $\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$
2. (Règle de Sarrus) On a :

$$\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = aei + bfg + cdh - gec - hfa - idb,$$

c'est à dire « diagonales descendantes moins diagonales montantes ».

Avertissement 23.2.51

Prenez garde à ne pas généraliser trop vite la règle de Sarrus à des matrices d'ordre supérieur ! Pour $n \geq 4$, une règle aussi simpliste est fautive ! Il faut considérer toutes les permutations possibles !

La description explicite par les coefficients permet également d'établir de façon quasi-immédiate :

Théorème 23.2.52 (Invariance du déterminant par transposée)

Soit $A \in \mathcal{M}_n(\mathbb{R})$. On a alors $\det({}^t A) = \det(A)$.

III Calcul des déterminants

Nous abordons maintenant les techniques calculatoires. Nous verrons essentiellement quatre techniques, la première étant une adaptation de la méthode du pivot, la seconde étant une façon de se ramener à des déterminants plus petits lorsque la matrice est triangulaire par blocs, l'importante troisième méthode est le développement suivant une ligne, ramenant le calcul d'un déterminant d'ordre n au calcul de n déterminants d'ordre $n - 1$, coefficientés par les coefficients de la ligne (intéressant surtout lorsque la plupart de ces coefficients sont nuls), et la dernière est l'utilisation du caractère polynomial du déterminant en chacune des coordonnées de la matrice.

III.1 Opérations sur les lignes et colonnes

Proposition 23.3.1 (Déterminant d'une matrice triangulaire)

Le déterminant d'une matrice triangulaire est égal au produit de ses coefficients diagonaux :

$$\begin{vmatrix} \lambda_1 & \bullet & \cdots & \bullet \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \bullet \\ 0 & \cdots & 0 & \lambda_n \end{vmatrix} = \lambda_1 \lambda_2 \cdots \lambda_n.$$

Lemme 23.3.2 (Déterminant des matrices de codage des opérations)

On a :

$$\det(E(i, j)) = -1, \quad \det(E_i(\lambda)) = \lambda \quad \det(E_{i,j}(\lambda)) = 1.$$

Corollaire 23.3.3 (Effet des opérations élémentaires sur le déterminant)

- (i) Échanger les lignes i et j change le signe du déterminant ;
- (ii) Multiplier une ligne par un scalaire λ multiplie le déterminant par ce scalaire
- (iii) Faire une combinaison $L_i \leftarrow L_i - \lambda L_j$ ne modifie pas le déterminant (attention à ne pas mettre de coefficient devant L_i).
- (iv) De même pour les opérations sur les colonnes.

Méthode 23.3.4 (Calcul du déterminant par la méthode du pivot)

Ainsi, on peut adapter la méthode du pivot pour le calcul du déterminant :

- Comme usuellement, faire des opérations pour échelonner la matrice, en écrivant des égalités entre les déterminants des différentes matrices, mais :
 - * Changer le signe lorsqu'on fait un échange de lignes
 - * Compenser en divisant à l'extérieur par λ lorsqu'on multiplie une ligne par λ .
- Une fois ramené à une matrice triangulaire, on calcule son déterminant en effectuant le produit des coefficients diagonaux.

Exemple 23.3.5

1. Déterminant de $\begin{pmatrix} 1 & -2 & 1 & 0 \\ 1 & 0 & 1 & 1 \\ 2 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}$

2. Déterminant de $\begin{pmatrix} a & 1 & \cdots & 1 \\ 1 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ 1 & \cdots & 1 & a \end{pmatrix}$

III.2 Calcul par blocs

Dans certaines configurations, on peut se ramener à des déterminants plus petits :

Proposition 23.3.6 (Déterminant d'une matrice triangulaire par blocs)

Soit T une matrice triangulaire par blocs, c'est-à-dire de la forme

$$T = \begin{pmatrix} A_1 & \bullet & \cdots & \bullet \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \bullet \\ 0 & \cdots & 0 & A_k \end{pmatrix}$$

où les A_k sont des matrices carrées. Alors

$$\det(T) = \det(A_1) \cdots \det(A_k).$$

Exemples 23.3.7

1. Déterminant de $\begin{pmatrix} 1 & 2 & 3 & 5 \\ 2 & 2 & 2 & 8 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & -1 & 2 \end{pmatrix}$

2. Déterminant de $\begin{pmatrix} A_1 & \bullet & \cdots & \bullet \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \bullet \\ 0 & \cdots & 0 & A_k \end{pmatrix}$, où $A_i = \begin{pmatrix} \cos(\theta_i) & -\sin(\theta_i) \\ \sin(\theta_i) & \cos(\theta_i) \end{pmatrix}$

3. On peut combiner avec les opérations du pivot, pour éviter la dernière opération (ou les deux dernières), en se ramenant ainsi au calcul d'un déterminant d'ordre 2 ou 3 (par Sarrus).

III.3 Développements suivant une ligne ou une colonne

Sans doute une des techniques les plus importantes, lorsqu'une ligne ou une colonne contient beaucoup de 0. Elle est efficace lorsqu'il y a beaucoup de symétries dans l'expression, pour construire une récurrence, et si ce n'est pas possible, elle peut être combinée à d'autres méthodes pour les calculs des déterminants plus petits obtenus.

Pour exprimer la formule, nous définissons :

Définition 23.3.8 (Mineurs, cofacteurs, comatrice)

Soit $M \in \mathcal{M}_n(\mathbb{K})$.

- Le mineur de position (i, j) de M est le déterminant de la matrice $\tilde{M}_{i,j}$ obtenue en supprimant de M la i -ième ligne et la j -ième colonne. On note $\Delta_{i,j}(M)$ ce mineur
- Le cofacteur de position (i, j) de M est le scalaire $(-1)^{i+j} \Delta_{i,j}(M)$.
- La comatrice de M est la matrice $\text{Com}(M) = ((-1)^{i+j} \Delta_{i,j}(M))_{1 \leq i, j \leq n}$, c'est à dire la matrice des cofacteurs de M .

Nous obtenons alors la formule de développement :

Théorème 23.3.9 (Développement suivant une ligne)

Soit $M = (m_{i,j}) \in \mathcal{M}_n(\mathbb{K})$, et $i \in \llbracket 1, n \rrbracket$. Alors

$$\det(M) = \sum_{j=1}^n (-1)^{i+j} m_{i,j} \Delta_{i,j}.$$

Théorème 23.3.10 (Développement suivant une colonne)

Soit $M = (m_{i,j}) \in \mathcal{M}_n(\mathbb{K})$, et $j \in \llbracket 1, n \rrbracket$. Alors

$$\det(M) = \sum_{i=1}^n (-1)^{i+j} m_{i,j} \Delta_{i,j}.$$

Corollaire 23.3.11 (Caractérisation de l'inversibilité par la comatrice)

Soit $M \in \mathcal{M}_n(\mathbb{K})$. Alors

$$M {}^t\text{Com}(M) = {}^t\text{Com}(M)M = \det(M)I_n.$$

En particulier, M est inversible si et seulement si $\det(M) \neq 0$, et dans ce cas,

$$M^{-1} = \frac{1}{\det(M)} \text{Com}(M).$$

Si la formule elle-même s'avère assez inefficace pour $n > 2$, la caractérisation de l'inversibilité par la non nullité du déterminant est d'une très grande utilité.

Exemples 23.3.12

1. Déterminant de la matrice tridiagonale

$$\begin{pmatrix} a+b & a & 0 & \cdots & 0 \\ b & \ddots & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & a \\ 0 & \cdots & 0 & b & a+b \end{pmatrix}.$$

2. On peut combiner pivot de Gauss et développement suivant les lignes ou colonnes, en annulant d'abord un grand nombre de termes par les opérations élémentaires avant de développer. Nous illustrons ceci sur le calcul du déterminant de Vandermonde :

$$V(x_1, \dots, x_n) = \begin{vmatrix} 1 & 1 & \cdots & 1 & 1 \\ x_1 & x_2 & \cdots & x_{n-1} & x_n \\ x_1^2 & x_2^2 & \cdots & x_{n-1}^2 & x_n^2 \\ \vdots & \vdots & & \vdots & \vdots \\ x_1^{n-1} & x_2^{n-1} & \cdots & x_{n-1}^{n-1} & x_n^{n-1} \end{vmatrix}.$$

Cet exemple est un peu plus qu'un exemple, nous le consignons dans la proposition suivante :

Proposition 23.3.13 (Déterminant de Vandermonde)

Le déterminant de Vandermonde, défini dans l'exemple ci-dessus, est donné par :

$$V(x_1, \dots, x_n) = \prod_{1 \leq i < j \leq n} (a_j - a_i).$$

Nous voyons dans la section suivante une autre façon de calculer ce déterminant de Vandermonde.

III.4 Caractère polynomial du déterminant

La formule explicite du déterminant nous permet d'affirmer :

Proposition 23.3.14 (« Polynomialité » du déterminant)

L'application $M \mapsto \det(M)$ est polynomiale en chacune des coordonnées de M . Il s'agit plus précisément d'une fonction polynomiale en les n^2 variables définies par les coordonnées de M , vérifiant :

- il est homogène de degré total n ,
- les degrés partiels par rapport à chacune des variables $m_{i,j}$ sont 1.
- Les degrés partiels par rapport à l'ensemble des variables constituant une même ligne ou une même colonne sont 1

Exemple 23.3.15

Calcul de $V_n(x_1, \dots, x_n)$.

Espaces préhilbertiens réels

Un espace préhilbertien réel est un espace vectoriel muni d'un produit scalaire. Cet outil permet de parler d'orthogonalité, et donc d'introduire un certain nombre de concepts permettant de généraliser la géométrie euclidienne du plan à des situations plus abstraites.

Pour commencer nous introduisons donc cette notion abstraite de produit scalaire, ce qui nous oblige à faire quelques rappels sur les formes bilinéaires, déjà rencontrées dans le chapitre précédent.

Tous les espaces vectoriels considérés dans ce chapitre sont des espaces vectoriels sur \mathbb{R} .

I Produits scalaires

I.1 Formes bilinéaires

Les résultats de cette sous-section sont valables dans un cadre plus général que les espaces sur \mathbb{R} , mais toute la suite du chapitre nécessitant de travailler avec des espaces vectoriels sur \mathbb{R} , nous nous limitons également à ce cadre dans les rappels qui suivent.

On rappelle la définition d'une forme bilinéaire :

Définition 24.1.1 (Forme bilinéaire)

Soit E un espace vectoriel sur \mathbb{R} . Une forme bilinéaire φ sur E est une application $\varphi : E \times E \rightarrow \mathbb{R}$, linéaire par rapport à chaque facteur, l'autre étant fixé, c'est-à-dire :

- (i) $\forall (x, y, z) \in E^3, \forall \lambda \in \mathbb{K}, \varphi(\lambda x + y, z) = \lambda \varphi(x, z) + \varphi(y, z)$
- (ii) $\forall (x, y, z) \in E^3, \forall \lambda \in \mathbb{K}, \varphi(x, \lambda y + z) = \lambda \varphi(x, y) + \varphi(x, z)$.

On rappelle la propriété de bilinéarité généralisée

Lemme 24.1.2 (Bilinéarité généralisée)

Soit φ une forme bilinéaire sur E . Soit $(k, \ell) \in (\mathbb{N}^*)^2$, et soit $(x_1, \dots, x_k, y_1, \dots, y_\ell) \in E^{k+\ell}$, et $(\lambda_1, \dots, \lambda_k, \mu_1, \dots, \mu_\ell) \in \mathbb{K}^{k+\ell}$. Alors

$$\varphi \left(\sum_{i=1}^k \lambda_i x_i, \sum_{j=1}^{\ell} \mu_j y_j \right) = \sum_{i=1}^k \sum_{j=1}^{\ell} \lambda_i \mu_j \varphi(x_i, y_j).$$

Définition 24.1.3 (Ensemble des formes bilinéaires)

On note $\mathcal{B}(E)$ l'ensemble des formes bilinéaires de E

Nous avons déjà vu le résultat suivant :

Proposition 24.1.4 (Structure de $\mathcal{B}(E)$)

L'ensemble $\mathcal{B}(E)$ est un espace vectoriel sur \mathbb{R} .

Voici des exemples particulièrement importants :

Exemples 24.1.5 (Formes bilinéaires)

1. $\text{cov} : (X, Y) \mapsto \text{cov}(X, Y)$ sur le \mathbb{R} -espace vectoriel \mathcal{V} des variables aléatoires réelles discrètes admettant une variance.
2. $\varphi : (P, Q) \mapsto \int_a^b P(x)Q(x) dx$ sur $\mathbb{R}[X]$, ou sur $\mathcal{C}^0([a, b])$.
3. $\varphi : \left(\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \right) \mapsto \sum_{i=1}^n x_i y_i$ sur \mathbb{R}^n
4. $p_1 : (x, y) \mapsto x$ sur \mathbb{R}
5. $\varphi : (X, Y) \mapsto {}^t X A Y$, sur \mathbb{R}^n , $A \in \mathcal{M}_n(\mathbb{R})$.

Ce dernier exemple est très important, car si E est de dimension finie, toute forme bilinéaire peut être représenté de cette forme après choix d'une base de E . C'est ce que nous étudions maintenant.

Définition 24.1.6 (Forme quadratique)

Une forme quadratique q sur E est une application $q : E \rightarrow \mathbb{R}$ telle qu'il existe φ bilinéaire sur E telle que pour tout $x \in E$, $\varphi(x, x) = q(x)$. On notera q_φ la forme quadratique associée à φ .

La forme quadratique q est en général associée à plusieurs formes linéaires. Par exemple les formes linéaires sur \mathbb{R}^2 définies par

$$\varphi((x, y), (x', y')) = xx' + yy' + 2xy' \quad \text{et} \quad \psi((x, y), (x', y')) = xx' + yy' + xy' + x'y$$

définissent la même forme quadratique q .

I.2 Matrice d'une forme bilinéaire

Nous supposons jusqu'à la fin de ce paragraphe que E est de dimension finie n .

Définition 24.1.7 (Matrice associée à une forme bilinéaire)

Soit φ une forme bilinéaire sur E , et $\mathcal{B} = (e_1, \dots, e_n)$ une base de E . Alors on définit la matrice de φ relativement à la base \mathcal{B} par :

$$\text{Mat}_{\mathcal{B}}(\varphi) = (\varphi(e_i, e_j))_{1 \leq i, j \leq n}.$$

Exemples 24.1.8

1. Matrice des variances-covariances si X_1, \dots, X_n sont linéairement indépendants.

2. Matrice dans la base canonique de $\varphi : (P, Q) \mapsto \int_0^1 P(x)Q(x) dx$ sur $\mathbb{R}_n[X]$.
3. Matrice du produit scalaire canonique de \mathbb{R}^n dans la base canonique ; dans la base $\mathcal{B} = (b_1, \dots, b_n)$, où b_i est constitué de 1 sur les coordonnées 1 à i , et de 0 ailleurs.

Théorème 24.1.9 (Expression matricielle de $\varphi(x, y)$)

Soit φ une forme bilinéaire sur E , et \mathcal{B} une base de E . Alors, pour tout $(x, y) \in E^2$,

$$\varphi(x, y) = {}^t[x]_{\mathcal{B}} \text{Mat}_{\mathcal{B}}(\varphi) [y]_{\mathcal{B}} = {}^tXMY,$$

où X et Y sont les vecteurs colonnes représentant les coordonnées de x et y dans la base \mathcal{B} , et M est la matrice de φ relativement à cette même base \mathcal{B} .

Une base \mathcal{B} étant fixée, cette relation caractérise d'ailleurs la matrice $\mathcal{M}_{\mathcal{B}}(\varphi)$:

Théorème 24.1.10 (Caractérisation de $\text{Mat}_{\mathcal{B}}(\varphi)$ par la relation $\varphi(X, Y) = {}^tXMY$)

Soit $M \in \mathcal{M}_n(\mathbb{K})$. La relation $\varphi(x, y) = {}^tXMY$ caractérise la matrice de φ relativement à la base \mathcal{B} : si

$$\forall (x, y) \in E^2, \quad \varphi(x, y) = {}^t[x]_{\mathcal{B}} M [y]_{\mathcal{B}},$$

alors $M = \text{Mat}_{\mathcal{B}}(\varphi)$.

Exemple 24.1.11

L'égalité $\mathcal{M}_{bc}(\varphi) = I_n$ se traduit par $\varphi(X, Y) = {}^tX I_n Y = {}^tXY$; c'est le produit scalaire usuel.

Remarque 24.1.12

À quelle condition nécessaire et suffisante sur leur représentation matricielle relativement à une base donnée deux formes bilinéaires φ et ψ définissent-elles la même forme quadratique q ?

Corollaire 24.1.13

L'application

$$\begin{aligned} \Phi : \mathcal{B}(E) &\longrightarrow \mathcal{M}_n(\mathbb{R}) \\ \varphi &\longmapsto \text{Mat}_{\mathcal{B}}(\varphi) \end{aligned}$$

est un isomorphisme.

Corollaire 24.1.14 (dimension de $\mathcal{B}(E)$)

Si E est de dimension finie n , alors $\mathcal{B}(E)$ est de dimension finie, et

$$\dim(\mathcal{B}(E)) = n^2.$$

Comme pour les applications linéaires, les changements de base s'expriment facilement par des opérations matricielles.

Théorème 24.1.15 (Formule de changement de base pour les formes bilinéaires)

Soit E un espace vectoriel sur \mathbb{R} de dimension finie. Soit φ une forme bilinéaire sur E , et soit \mathcal{C} et \mathcal{D} deux bases de E . Soit P la matrice de passage de \mathcal{C} à \mathcal{D} . Alors

$$\text{Mat}_{\mathcal{D}}(\varphi) = {}^t P \text{Mat}_{\mathcal{C}}(\varphi) P.$$

Exemple 24.1.16

Explicitation du changement de la base canonique à la base \mathcal{B} dans l'exemple 24.1.8, pour le produit scalaire canonique.

I.3 Produits scalaire

Pour définir la notion de produit scalaire, on introduit trois propriétés pouvant être vérifiées par une forme bilinéaire. La conjonction de ces trois propriétés définiront un produit scalaire.

I.4 Formes bilinéaires symétriques, définies, positives**Définition 24.1.17 (Symétrie, positivité, caractère défini)**

Soit φ une forme bilinéaire sur un espace vectoriel E .

1. On dit que φ est symétrique si : $\forall (x, y) \in E^2, \varphi(x, y) = \varphi(y, x)$.
2. On dit que φ est positive si : $\forall x \in E, \varphi(x, x) \geq 0$, c'est-à-dire $\text{Im}(q_\varphi) \subset \mathbb{R}_+$
3. On dit que φ est définie si : $\forall x \in E, \varphi(x, x) = 0 \iff x = 0$, i.e. $q_\varphi(x) \neq 0$ pour tout $x \neq 0$.

Proposition 24.1.18 (positivité des formes définies)

Soit φ une forme définie. Alors φ est soit positive soit négative.

Pour cette raison, on parle assez rarement de forme définie, cette propriété étant indissociable d'une propriété de positivité ou de négativité. On parlera alors de forme définie positive, ou définie négative.

Proposition 24.1.19 (Caractérisation matricielle de la symétrie)

Soit φ une forme bilinéaire sur un espace vectoriel E de dimension finie. Les propriétés suivantes sont équivalentes :

- (i) φ est symétrique ;
- (ii) Il existe une base de E telle que $\text{Mat}_{\mathcal{B}}(\varphi)$ est une matrice symétrique
- (iii) Pour toute base \mathcal{B} de E , $\text{Mat}_{\mathcal{B}}(\varphi)$ est une matrice symétrique.

Exemples 24.1.20

1. $\varphi : (x, y) \mapsto xy$ sur \mathbb{R} est symétrique, définie, positive.
2. cov est symétrique positive, mais pas définie.
3. Le produit scalaire canonique sur \mathbb{R}^n est symétrique, défini, positif.
4. $\varphi : (f, g) \mapsto \int_a^b f(t)g(t) dt$ sur $\mathcal{C}^0([a, b])$ est symétrique, définie et positive.

Proposition 24.1.21 (Symétrie et forme quadratique ; formule de polarisation)

Soit q une forme quadratique sur E . Il existe une unique forme bilinéaire symétrique sur E telle que $q = q_\varphi$, explicitement donnée par l'expression suivante, dite formule de polarisation :

$$\varphi(x, y) = \frac{1}{2}(q(x + y) - q(x) - q(y)).$$

Voici enfin un résultat d'une importance capitale, que nous avons déjà rencontré en cours d'année dans différents contextes, qui sont tous des cas particuliers du contexte général exposé ici.

Proposition 24.1.22 (Inégalité de Cauchy-Schwarz)

Soit φ une forme bilinéaire symétrique positive sur un espace E . Alors, pour tout $(x, y) \in E^2$, on a

$$\varphi(x, y)^2 \leq \varphi(x, x) \cdot \varphi(y, y),$$

et on a égalité si et seulement si il existe $(\lambda, \mu) \neq (0, 0)$ tel que $q_\varphi(\lambda x + \mu y) = 0$.

Exemple 24.1.23

Pour toutes variables aléatoires X et Y admettant une variance, on a :

$$|\text{cov}(X, Y)| \leq \sigma(X)\sigma(Y).$$

Nous verrons d'autres exemples plus loin, lorsque nous aurons réexprimé plus simplement le cas d'égalité lorsque φ est un produit scalaire.

I.5 Produits scalaires

Dans cette section, E désigne un espace vectoriel sur \mathbb{R} .

Définition 24.1.24 (Produit scalaire)

Un produit scalaire sur E est une forme bilinéaire symétrique, définie et positive.

Remarque 24.1.25

La symétrie et la linéarité par rapport à la première variable entraîne la linéarité par rapport à la seconde variable.

Voici les exemples usuels (les 2 premiers), desquels on peut dériver d'autres exemples. Ces exemples sont à considérer comme des exemples du cours.

Exemples 24.1.26 (produits scalaires)

1. $(f, g) \mapsto \int_a^b f(t)g(t) dt$ est un produit scalaire sur $\mathcal{C}^0(\mathbb{R})$ (ou sur $\mathbb{R}[X]$)
2. Le produit scalaire canonique de \mathbb{R}^n est un produit scalaire au sens de cette définition.
3. Un autre produit scalaire sur \mathbb{R}^n : $(X, Y) \mapsto {}^t XMY$, où $M = (\min(i, j))_{1 \leq i, j \leq n}$.

Ainsi, on n'a pas unicité d'un produit scalaire sur un espace vectoriel E .

Notation 24.1.27 (Notations fréquentes pour le produit scalaire)

Soit φ un produit scalaire sur E on note souvent :

$$\varphi(x, y) = \langle x, y \rangle \quad \text{ou} \quad \varphi(x, y) = (x|y).$$

Par ailleurs, on note $\|x\| = \sqrt{\varphi(x, x)} = \sqrt{\langle x, x \rangle}$ la racine de la forme quadratique associée, cette expression étant bien définie par positivité de φ .

Soit φ un produit scalaire, noté $\varphi(x, y) = \langle x, y \rangle$. Alors en particulier, φ est une forme bilinéaire, et si $\mathcal{B} = (e_1, \dots, e_n)$ est une base de E :

$$\text{Mat}_{\mathcal{B}}(\varphi) = (\langle e_i, e_j \rangle)_{1 \leq i, j \leq n}$$

Proposition 24.1.28

La matrice d'un produit scalaire dans une base quelconque est symétrique. La réciproque est fausse.

Un produit scalaire étant une forme bilinéaire symétrique, on dispose aussi de la formule de polarisation, s'écrivant ici :

$$\langle x, y \rangle = \frac{1}{2} (\|x + y\|^2 - \|x\|^2 - \|y\|^2).$$

Cette formule n'est autre que le développement d'un carré :

$$\|x + y\|^2 = \|x\|^2 + 2\langle x, y \rangle + \|y\|^2.$$

On généralise facilement par la formule de bilinéarité généralisée :

Proposition 24.1.29 (carré de la norme d'une somme)

Plus généralement, étant donné $(x_1, \dots, x_n) \in E^n$,

$$\|x_1 + \dots + x_n\|^2 = \sum_{i=1}^n \|x_i\|^2 + 2 \sum_{1 \leq i < j \leq n} \langle x_i, x_j \rangle.$$

L'inégalité de Cauchy-Schwarz est également vraie, et le caractère défini permet de simplifier la condition d'égalité.

Théorème 24.1.30 (Inégalité de Cauchy-Schwarz pour un produit scalaire)

Soit $\langle \cdot, \cdot \rangle$ un produit scalaire sur E . Alors pour tout $(x, y) \in E^2$,

$$|\langle x, y \rangle| \leq \|x\| \cdot \|y\|,$$

avec égalité si et seulement si x et y sont colinéaires.

Comme promis plus haut, nous donnons deux inégalités de Cauchy-Schwarz à connaître impérativement.

Exemples 24.1.31 (Autres exemples d'inégalités de Cauchy-Schwarz)

1. Inégalité de Cauchy-Schwarz numérique :

$$\forall X = (x_1, \dots, x_n), Y = (y_1, \dots, y_n) \in \mathbb{R}^n, \quad |\langle X, Y \rangle| \leq \|X\| \cdot \|Y\|,$$

à savoir :

$$\left| \sum_{i=1}^n x_i y_i \right| \leq \left(\sum_{i=1}^n x_i^2 \right)^{\frac{1}{2}} \left(\sum_{i=1}^n y_i^2 \right)^{\frac{1}{2}},$$

ou encore :

$$\left(\sum_{i=1}^n x_i y_i \right)^2 \leq \left(\sum_{i=1}^n x_i^2 \right) \left(\sum_{i=1}^n y_i^2 \right),$$

avec égalité si et seulement si X et Y sont colinéaires.

2. Inégalité de Cauchy-Schwarz intégrale :

$$\forall (f, g) \in \mathcal{C}^0([a, b]), \left| \int_a^b f(t)g(t) dt \right| \leq \left(\int_a^b f(t)^2 dt \right)^{\frac{1}{2}} \left(\int_a^b g(t)^2 dt \right)^{\frac{1}{2}},$$

avec égalité si et seulement si $g = 0$, ou s'il existe λ tel que $f = \lambda g$.

I.6 Normes euclidiennes

Le but de cette section est de montrer que $x \mapsto \|x\|^2$ définit une norme vectorielle.

Définition 24.1.32 (Norme)

Une norme sur un espace vectoriel E est une application $N : E \rightarrow \mathbb{R}$ telle que :

- (i) $\forall x \in E, N(x) = 0 \iff x = 0$
- (ii) $\forall \lambda \in \mathbb{R}, \forall x \in E, N(\lambda x) = |\lambda|N(x)$
- (iii) $\forall (x, y) \in E^2, N(x + y) \leq N(x) + N(y)$.

Proposition 24.1.33 (Positivité des normes)

Si N est une norme sur E , alors pour tout $x \in E, N(x) \geq 0$.

La fin de ce paragraphe a pour but de justifier que ceci définit bien une norme. Pour commencer, nous obtenons :

Proposition 24.1.34

L'application $x \mapsto \|x\|$ est une norme sur E .

Définition 24.1.35 (Norme euclidienne associée au produit scalaire)

Soit $\langle \cdot, \cdot \rangle$ un produit scalaire sur E . L'application $\|\cdot\|$ est appelée norme euclidienne associée au produit scalaire $\langle \cdot, \cdot \rangle$.

Définition 24.1.36 (Norme euclidienne)

Soit N une norme sur E . La norme N est euclidienne si et seulement s'il existe un produit scalaire $\langle \cdot, \cdot \rangle$, dont N est la norme euclidienne associée

La formule de polarisation donne l'unique candidat possible pour ce produit scalaire :

Méthode 24.1.37 (Montrer qu'une norme est euclidienne)

On définit l'application φ , unique candidat possible :

$$\varphi(x, y) = \frac{1}{2}(N(x + y)^2 - N(x)^2 - N(y)^2).$$

On vérifie ensuite que φ est une application bilinéaire. En cas de bilinéarité, le caractère symétrique défini positif est immédiat par définition de N . Le point à étudier de plus près est donc la bilinéarité : c'est là que se trouve l'éventuelle obstruction.

Exemple 24.1.38

$N : (x_1, \dots, x_n) \mapsto \max(|x_1|, \dots, |x_n|)$ est une norme sur \mathbb{R}^n , mais n'est pas une norme euclidienne.

Ainsi, toutes les normes ne sont pas euclidiennes !

I.7 Espaces préhilbertiens réels, espaces euclidiens

Définition 24.1.39 (Espace préhilbertien réel)

Un espace préhilbertien réel $(E, \langle \cdot, \cdot \rangle)$ est un espace vectoriel E sur \mathbb{R} , muni d'un produit scalaire $\langle \cdot, \cdot \rangle$.

S'il n'y a pas d'ambiguïté sur le produit scalaire, on parlera plus simplement de l'espace préhilbertien E , au lieu de $(E, \langle \cdot, \cdot \rangle)$.

Définition 24.1.40 (Espace euclidien)

Un espace euclidien est un espace préhilbertien réel de dimension finie.

Par exemple \mathbb{R}^n muni du produit scalaire canonique est un espace euclidien. $\mathbb{R}_n[X]$ muni d'un produit scalaire intégral est un espace euclidien. En revanche, $\mathbb{R}[X]$ muni du même produit scalaire n'est qu'un espace préhilbertien réel, ainsi que $\mathcal{C}^0([a, b])$ muni du produit scalaire intégral.

II Orthogonalité

Dans toute cette section, on considère $(E, \langle \cdot, \cdot \rangle)$ un espace préhilbertien réel. Il sera précisé euclidien dans certains cas.

II.1 Vecteurs orthogonaux

Définition 24.2.1 (Vecteurs orthogonaux)

Soit $(x, y) \in E$ deux vecteurs de E . On dit que x et y sont orthogonaux, et on note $x \perp y$, si $\langle x, y \rangle = 0$.

Exemples 24.2.2

1. $\forall x \in E, x \perp 0$.
2. Soit (e_1, \dots, e_n) la base canonique de \mathbb{R}^n . Alors pour tout $i \neq j, e_i \perp e_j$.

3. $\begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix} \perp \begin{pmatrix} -4 \\ -3 \\ 2 \\ 1 \end{pmatrix}$, pour le produit scalaire usuel.

4. $x \mapsto \sin \pi x$ et $x \mapsto \cos \pi x$ sont orthogonaux pour le produit scalaire $\langle f, g \rangle = \int_{-1}^1 f(t)g(t) dt$ sur $\mathcal{C}^0([-1, 1])$.

Définition 24.2.3 (Famille orthogonale, orthonormale)

1. Soit $\mathcal{F} = (x_1, \dots, x_n)$ une famille (finie) d'éléments de E . On dit que la famille \mathcal{F} est orthogonale si et seulement si les x_i sont deux à deux orthogonaux, c'est-à-dire si et seulement si pour tout $(i, j) \in \llbracket 1, n \rrbracket$ tel que $i \neq j$, $x_i \perp x_j$.
2. On dit que la famille \mathcal{F} est orthonormale (ou orthonormée) si et seulement si elle est orthogonale, et que pour tout $i \in \llbracket 1, n \rrbracket$, $\|x_i\| = 1$.

Théorème 24.2.4 (Liberté des familles orthogonales)

Soit \mathcal{F} une famille orthogonale ne contenant pas le vecteur nul. Alors \mathcal{F} est une famille libre.

En particulier, toute famille orthonormale est libre.

Corollaire 24.2.5 (base orthonormale)

On suppose E de dimension finie n (i.e. E euclidien). Soit (e_1, \dots, e_n) une famille orthonormale de E . Alors (e_1, \dots, e_n) est une base de E . On dit qu'il s'agit d'une base orthonormale, et on abrège en b.o.n..

Un fait qu'il est important de garder en mémoire est la facilité d'expression vectorielle des objets dans une base orthonormale. Nous pouvons donner dès maintenant l'expression d'un vecteur dans une b.o.n., nous verrons un peu plus loin l'expression de la matrice d'un endomorphisme.

Proposition 24.2.6 (Expression des coordonnées d'un vecteur dans une b.o.n.)

Soit E un espace euclidien. Soit $\mathcal{B} = (b_1, \dots, b_n)$ une base orthonormale de E . Alors :

$$\forall X \in E, \quad X = \sum_{i=1}^n \langle X, b_i \rangle b_i.$$

En d'autres termes, le vecteur des coordonnées de X dans la base \mathcal{B} est

$$[X]_{\mathcal{B}} = \begin{pmatrix} \langle X, b_1 \rangle \\ \vdots \\ \langle X, b_n \rangle \end{pmatrix}.$$

Exemple 24.2.7

Dans \mathbb{R}^2 , il s'agit des projections orthogonales sur chaque axe.

Ce théorème n'est vraiment utile que si on sait déterminer des bases orthonormales. On verra un peu plus loin comment construire une b.o.n. à partir de n'importe quelle base (procédé d'orthonormalisation de Gram-Schmidt). En particulier, ce procédé nous assurera de l'existence d'une b.o.n.

On ne suppose plus ici que E est de dimension finie

Théorème 24.2.8 (Théorème de Pythagore)

Soit E un espace préhilbertien réel. Soit $(x, y) \in E^2$. Alors

$$x \perp y \iff \|x + y\|^2 = \|x\|^2 + \|y\|^2.$$

Théorème 24.2.9 (Théorème de Pythagore généralisé)

Soit (x_1, \dots, x_n) une famille orthogonale de E . Alors

$$\left\| \sum_{i=1}^n x_i \right\|^2 = \sum_{i=1}^n \|x_i\|^2.$$

II.2 Sous-espaces orthogonaux**Définition 24.2.10 (Sous-espaces orthogonaux)**

Soit E un espace préhilbertien réel.

1. Soit $x \in E$, et F un sous-espace vectoriel de E . On dit que x est orthogonal à F si pour tout $y \in F$, $x \perp y$. On note $x \perp F$, ou $x \in F^\perp$ comme on le verra plus tard.
2. Soit F et G deux sous-espaces vectoriels de E . On dit que F et G sont orthogonaux si et seulement si :

$$\forall x \in F, \forall y \in G, x \perp y.$$

Exemples 24.2.11

1. Dans \mathbb{R}^3 , le plan (Oxy) et la droite (Oz)
2. Dans \mathbb{R}^3 , le plan d'équation $x + y + z = 0$, et la droite engendrée par $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$.
3. Dans $\mathcal{C}^0([-1, 1])$ muni du ps usuel, le sous espace P des fonctions paires, et le sous-espace I des fonctions impaires.

Proposition 24.2.12 (orthogonalité et somme directe)

Soit F et G des sous-espaces vectoriels de E . Si $F \perp G$, alors la somme $F + G$ est directe.

Du fait de la bilinéarité du produit scalaire, il n'est pas nécessaire d'établir l'orthogonalité de toutes les paires de vecteurs de E et F pour obtenir l'orthogonalité des deux espaces. Il suffit en fait de vérifier cette orthogonalité sur des vecteurs de familles génératrices.

Proposition 24.2.13 (Caractérisation de l'orthogonalité par les familles génératrices)

Soit (x_1, \dots, x_m) et (y_1, \dots, y_n) deux familles de E . Alors

$$\text{Vect}(x_1, \dots, x_m) \perp \text{Vect}(y_1, \dots, y_n) \text{ si et seulement si } \forall i \in \llbracket 1, m \rrbracket, \forall j \in \llbracket 1, n \rrbracket, x_i \perp y_j.$$

En particulier, on vérifie souvent l'orthogonalité en se ramenant à des bases :

Corollaire 24.2.14 (Caractérisation de l'orthogonalité par des bases)

Soit F et G deux sous-espaces vectoriels de E , et soit $\mathcal{B} = (b_1, \dots, b_m)$ et $\mathcal{C} = (c_1, \dots, c_n)$ des bases de F et G respectivement. Alors $F \perp G$ si et seulement si

$$\forall (i, j) \in \llbracket 1, m \rrbracket \times \llbracket 1, n \rrbracket, b_i \perp c_j.$$

Définition 24.2.15 (orthogonal d'une partie de E)

Soit X un sev de E . On note $X^\perp = \{x \in E \mid x \perp X\}$, l'ensemble des vecteurs orthogonaux à tout vecteur de X . L'ensemble X^\perp est appelé l'orthogonal de X .

Proposition 24.2.16 (Structure de l'orthogonal)

1. Soit $x \in E$. Alors x^\perp est un sev de E
2. Soit $X \subset E$. Alors X^\perp est un sev de E .

Proposition 24.2.17 (Stabilité de l'orthogonal par Vect)

Soit X une partie de E . Alors $X^\perp = \text{Vect}(X)^\perp$.

Proposition 24.2.18

Soit F un sev de E . Alors $F \perp F^\perp$, donc en particulier, la somme $F \oplus F^\perp$ est directe.

Avertissement 24.2.19

Attention, contrairement à l'idée intuitive qu'on se fait en considérant l'orthogonalité dans \mathbb{R}^3 , F^\perp n'est pas forcément un supplémentaire de F dans E . On montrera que c'est le cas si E est de dimension finie, mais que cette propriété entre en défaut si E est de dimension infinie.

Exemple 24.2.20

Déterminer l'orthogonal dans \mathbb{R}^4 muni de la base canonique de

$$F = \text{Vect} \left(\begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \end{pmatrix} \right).$$

II.3 Projeté orthogonal**Définition 24.2.21 (Projeté orthogonal sur une droite)**

Soit x et y deux éléments de E . On dit qu'un vecteur $z \in E$ est le projeté orthogonal de y sur la droite $\text{Vect}(x)$ si et seulement si $y - z \perp x$ et $z \in \text{Vect}(x)$.

Proposition 24.2.22 (Existence et expression du projeté orthogonal sur une droite)

Soit x et y deux éléments de E . Alors, le projeté orthogonal de y sur $\text{Vect}(x)$ existe, est unique, et vaut :

$$z = \langle y, x \rangle \cdot \frac{x}{\|x\|^2} = \left\langle y, \frac{x}{\|x\|} \right\rangle \cdot \frac{x}{\|x\|}.$$

Définition 24.2.23 (Projeté orthogonal sur un sev)

Soit F un sous-espace vectoriel de E , et $y \in E$. On dit que $z \in E$ est le projeté orthogonal de y sur F si et seulement si :

- (i) $z \in F$
- (ii) $(y - z) \perp F$.

Théorème 24.2.24 (Existence du projeté orthogonal sur sev de dimension finie)

Soit $y \in E$, et F un sous-espace vectoriel **de dimension finie** de E , tel qu'il existe une base orthonormale (b_1, \dots, b_m) de F . Alors le projeté orthogonal de y sur F existe, est unique, et vaut :

$$z = \sum_{i=1}^m \langle y, b_i \rangle b_i.$$

Remarques 24.2.25

1. On verra dans la suite du cours qu'un espace vectoriel de dimension finie et muni d'un produit scalaire admet toujours au moins une base orthonormale pour ce produit scalaire : l'hypothèses d'existence de cette base est donc superflue dans le théorème ci-dessus.
2. Si (b_1, \dots, b_n) est une base orthogonale, mais non orthonormale, on obtient la formule suivante pour le projeté orthogonale de y sur F :

$$z = \sum_{i=1}^m \left\langle y, \frac{b_i}{\|b_i\|} \right\rangle \cdot \frac{b_i}{\|b_i\|}.$$

3. Si $y \in F$, son projeté est bien entendu lui-même, et on obtient, pour une b.o.n. (b_1, \dots, b_m) de F :

$$y = \sum_{i=1}^m \langle y, b_i \rangle b_i.$$

Ainsi, on retrouve l'expression des coordonnées d'un vecteur y dans une base orthonormale.

II.4 Orthonormalisation de Schmidt

Motivation : Étant donné une famille libre (e_1, \dots, e_n) de E , trouver un moyen concret de construire une famille libre orthonormée (f_1, \dots, f_n) engendrant le même espace que (e_1, \dots, e_n) , c'est-à-dire tel que (f_1, \dots, f_n) est une b.o.n. de $\text{Vect}(e_1, \dots, e_n)$.

En particulier, si (e_1, \dots, e_n) est initialement une base de E , on décrit une façon canonique de construire une b.o.n. de E à partir de cette base.

On fait cette construction étape par étape, de manière à avoir, pour tout $k \in \llbracket 1, n \rrbracket$,

$$\text{Vect}(e_1, \dots, e_k) = \text{Vect}(f_1, \dots, f_k).$$

Théorème 24.2.26 (Procédé d'orthonormalisation de Schmidt)

Soit E un espace préhilbertien réel. Soit (e_1, \dots, e_n) une famille libre de E . Il existe une unique famille orthonormale (f_1, \dots, f_n) telle que, pour tout $k \in \llbracket 1, n \rrbracket$, on ait

$$\text{Vect}(f_1, \dots, f_k) = \text{Vect}(e_1, \dots, e_k), \quad \text{et} \quad \langle e_k, f_k \rangle \geq 0.$$

Cette famille peut être construite explicitement par la description par récurrence suivante :

$$f_1 = \frac{e_1}{\|e_1\|} \quad \text{et} \quad \forall k \in \llbracket 2, n \rrbracket, \quad f_k = \frac{u_k}{\|u_k\|}, \quad \text{où} \quad u_k = e_k - \sum_{i=1}^{k-1} \langle e_k, f_i \rangle f_i.$$

Ainsi, pour tout $k \in \llbracket 1, n \rrbracket$, (f_1, \dots, f_k) est une b.o.n. de $\text{Vect}(e_1, \dots, e_k)$. En particulier, si initialement, (e_1, \dots, e_n) est une base de E , alors (f_1, \dots, f_n) en est une base orthonormale.

Exemples 24.2.27

1. Base orthonormale de $F = \text{Vect} \left(\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \right)$

2. Orthonormalisée de Schmidt de $(1, X)$ dans $\mathbb{R}_n[X]$ muni de :

$$\langle P, Q \rangle = \int_0^1 P(t)Q(t) dt.$$

III Espaces euclidiens

On suppose ici que E est un espace euclidien (donc de dimension finie). On commence par montrer l'existence de bases orthonormales. On en déduit en particulier la possibilité de projeter orthogonalement sur tout sous-espace vectoriel.

III.1 Bases orthonormales d'un espace euclidien

Théorème 24.3.1 (Existence d'une b.o.n. d'un espace euclidien)

Tout espace euclidien E admet au moins une base orthonormale.

Théorème 24.3.2 (théorème de la base orthonormale incomplète)

Soit E un espace euclidien.

1. *Toute famille libre orthogonale de E peut être complétée en une base orthogonale de E*
2. *Toute famille orthonormale de E peut être complétée en une base orthonormale de E .*

Corollaire 24.3.3 (Existence et unicité d'un supplémentaire orthogonal)

Soit E un espace euclidien. Tout sous-espace F de E admet un unique supplémentaire G tel que $F \perp G$. De plus, on a $G = F^\perp$. On dit que F^\perp est le supplémentaire orthogonal de F .

Proposition 24.3.4 (base orthonormale d'une somme orthogonale)

Soit E un espace euclidien, et F et G deux sous-espaces vectoriels de E tels que $F \perp G$. Soit (b_1, \dots, b_p) une base orthonormale de F , et (c_1, \dots, c_q) une base orthonormale de G . Alors $(b_1, \dots, b_p, c_1, \dots, c_q)$ est une base orthonormale de $F \oplus G$; en particulier, c'est une famille orthonormale de E .

Corollaire 24.3.5

Soit E un espace euclidien, et F_1, \dots, F_p des sous-espaces vectoriels de E , deux à deux orthogonaux. Alors la somme $F_1 + \dots + F_p$ est directe, et on obtient une base orthonormale de $F_1 \oplus \dots \oplus F_p$ en juxtaposant des bases orthonormales des espaces F_1, \dots, F_p .

Corollaire 24.3.6

Soit E un espace euclidien, F un sous-espace vectoriel de E , et F^\perp son supplémentaire orthogonal. Soit (b_1, \dots, b_p) une base orthonormale de F et (c_1, \dots, c_q) une base orthonormale de F^\perp . Alors $(b_1, \dots, b_p, c_1, \dots, c_q)$ est une base orthonormale de E .

Comme on l'a évoqué plus haut, trouver les coordonnées d'un vecteur X dans une b.o.n. est simple (si on sait calculer le produit scalaire). Il en est de même de la norme.

Théorème 24.3.7 (Coordonnées d'un vecteur dans une b.o.n. et norme)

Soit E un espace euclidien, et $\mathcal{B} = (b_1, \dots, b_n)$ une b.o.n. de E . Soit $x \in E$. Alors :

$$(i) \quad x = \sum_{i=1}^n \langle x, b_i \rangle b_i, \text{ c'est-à-dire } [x]_{\mathcal{B}} = \begin{pmatrix} \langle x, b_1 \rangle \\ \vdots \\ \langle x, b_n \rangle \end{pmatrix}.$$

$$(ii) \quad \|x\|^2 = \sum_{i=1}^n \langle x, b_i \rangle^2$$

Ainsi, la norme de x est la racine de la somme des carrés des coordonnées de x dans la b.o.n. considérée. Cela correspond bien à la situation connue dans \mathbb{R}^n muni de son produit scalaire canonique. La matrice d'un endomorphisme admet également une description très simple :

Théorème 24.3.8 (Matrice d'un endomorphisme relativement à une b.o.n.)

Soit E un espace euclidien muni d'une b.o.n. $\mathcal{B} = (b_1, \dots, b_n)$. Soit $u \in \mathcal{L}(E)$. Alors :

$$\text{Mat}_{\mathcal{B}}(u) = (\langle b_i, u(b_j) \rangle)_{1 \leq i, j \leq n} = \begin{pmatrix} \langle b_1, u(b_1) \rangle & \cdots & \langle b_1, u(b_n) \rangle \\ \vdots & & \vdots \\ \langle b_n, u(b_1) \rangle & \cdots & \langle b_n, u(b_n) \rangle \end{pmatrix}.$$

Enfin, l'expression de la matrice du produit scalaire dans la base \mathcal{B} permet de caractériser facilement les bases orthonormales.

Théorème 24.3.9 (Matrice du produit scalaire relativement à une b.o.n.)

Soit E un espace euclidien, et \mathcal{B} une base de E . Alors la base \mathcal{B} est orthonormale si et seulement si la matrice du produit scalaire dans \mathcal{B} est I_n .

Dans ce cas, pour tout (x, y) dans E^2 , on a :

$$\langle x, y \rangle = {}^tXY,$$

où $X = [x]_{\mathcal{B}}$ et $Y = [y]_{\mathcal{B}}$ sont les vecteurs des coordonnées de x et y dans la base \mathcal{B} .

III.2 Changements de base et matrices orthogonales

Les matrices de passage d'une b.o.n. à une autre vérifient une propriété très forte :

Théorème 24.3.10 (propriété des matrices de passage d'une b.o.n. à une autre)

Soit \mathcal{B} et \mathcal{C} deux b.o.n. de E . Soit $P = [\mathcal{B} \rightarrow \mathcal{C}]$ la matrice de passage de la base \mathcal{B} à la base \mathcal{C} . Alors :

$${}^tPP = I_n = P {}^tP, \quad \text{donc:} \quad P^{-1} = {}^tP.$$

Cette propriété définit la notion de matrice orthogonale :

Définition 24.3.11 (Matrice orthogonale)

Soit $P \in \mathcal{M}_n(\mathbb{R})$ une matrice carrée d'ordre n . On dit que P est une matrice orthogonale si et seulement si ${}^t P P = I_n$.

De la définition même découle de façon immédiate :

Proposition 24.3.12 (Inverse d'une matrice orthogonale)

Soit P une matrice orthogonale. Alors P est inversible, et $P^{-1} = {}^t P$.

Une matrice orthogonale peut se caractériser également par l'orthonormalité de ses colonnes :

Théorème 24.3.13 (Caractérisation d'une matrice orthogonale par ses colonnes)

Une matrice P est orthogonale si et seulement si ses colonnes forment une base orthonormale de $\mathbb{R}^n = \mathcal{M}_{n,1}(\mathbb{R})$, muni du produit scalaire canonique.

Théorème 24.3.14 (Caractérisation des matrices de passage entre b.o.n. par orthogonalité)

Soit E un espace euclidien.

1. Toute matrice de passage d'une b.o.n. de E à une autre b.o.n. de E est une matrice orthogonale.
2. Réciproquement, soit \mathcal{B} une b.o.n. de E , et P une matrice orthogonale. Alors il existe une unique base \mathcal{C} telle que P soit la matrice de passage de \mathcal{B} à \mathcal{C} , et cette base \mathcal{C} est une b.o.n. de E .

Définition 24.3.15 (Groupe orthogonal)

On note $O_n(\mathbb{R})$, ou $O(n)$ l'ensemble des matrices orthogonales de $\mathcal{M}_n(\mathbb{R})$. Cet ensemble $O_n(\mathbb{R})$ est appelé groupe orthogonal.

Théorème 24.3.16 (Structure de $O_n(\mathbb{R})$)

L'ensemble $O_n(\mathbb{R})$ est un groupe multiplicatif, ce qui est cohérent avec la terminologie introduite dans la définition précédente.

Théorème 24.3.17 (Déterminant d'une matrice orthogonale)

Soit $P \in O_n(\mathbb{R})$. Alors $\det(P) \in \{-1, 1\}$. Plus précisément, \det est un morphisme de groupe de $O_n(\mathbb{R})$ dans $(\{-1, 1\}, \times)$.

Le noyau de ce morphisme est donc un sous-groupe de $O_n(\mathbb{R})$.

Définition 24.3.18 (Groupe spécial orthogonal)

Le noyau du déterminant défini sur $O_n(\mathbb{R})$ est appelé groupe spécial orthogonal, et noté $SO_n(\mathbb{R})$ ou $SO(n)$. Ainsi, les éléments de $SO_n(\mathbb{R})$ sont les matrices orthogonales P telles que $\det(P) = 1$.

Le choix d'une matrice orthogonale P telle que $\det(P) = -1$ définit alors une bijection de $SO(n)$ dans $O^-(n) = O(n) \setminus SO(n)$ par $Q \mapsto PQ$.

III.3 Projecteurs orthogonaux

On rappelle la définition du projeté orthogonal :

Définition 24.3.19

Soit E un espace euclidien, et F un sous-espace vectoriel de E . Soit $x \in E$. On dit qu'un vecteur y de E est un projeté orthogonal de x sur F si et seulement si :

- (i) $y \in F$
- (ii) $x - y \in F^\perp$.

Comme nous l'avons vu plus haut, étant donnée une base orthonormale d'un sous-espace de dimension finie F de E , on dispose d'une formule simple de projection orthogonale, assurant l'existence et l'unicité de ce projeté. Comme nous avons montré entre temps que tout espace euclidien de dimension finie admet une b.o.n., on peut réexprimer ce théorème ainsi, en se limitant au cas euclidien (assurant l'hypothèse de la dimension de F) :

Proposition 24.3.20 (Projeté orthogonal dans un espace euclidien)

Soit E un espace euclidien, et F un sous-espace vectoriel de E . Alors tout vecteur x de E admet un et un seul projeté orthogonal sur F . De plus, étant donné une base **orthonormale** (b_1, \dots, b_p) de F , ce projeté orthogonal $p_F(x)$ s'exprime de la manière suivante :

$$p_F(x) = \sum_{i=1}^p \langle x, b_i \rangle b_i.$$

Définition 24.3.21 (projection orthogonale)

Le projeté orthogonal d'un vecteur sur F définit une application $p_F : E \rightarrow E$, qui à x associe son projeté orthogonal sur F . Cette application est appelée « projection orthogonale sur F »

Proposition 24.3.22

Évidemment, p_F est un projecteur.

Définition 24.3.23

Rappel : soit p un projecteur quelconque de E . Alors $\text{id}_E - p$ est un projecteur de E , appelé projecteur associé de p .

Proposition 24.3.24 (projecteur associé à un projecteur orthogonal)

Soit E un espace euclidien, et F un sous-espace vectoriel de E . Alors le projecteur associé de la projection orthogonale p_F sur F est la projection orthogonale p_{F^\perp} sur l'orthogonal de F .

Méthode 24.3.25 (Comment déterminer la matrice d'un projecteur orthogonal de \mathbb{R}^n)

Soit dans \mathbb{R}^3 le plan F d'équation $x + 2y - z = 0$. Déterminer la matrice de p_F , le projecteur orthogonal sur le plan F .

(Réponse : $\frac{1}{6} \begin{pmatrix} 5 & -2 & 1 \\ -2 & 2 & 2 \\ 1 & 2 & 5 \end{pmatrix}$)

- Première méthode : trouver une b.o.n. de F , et utiliser la formule donnant le projeté à l'aide d'une b.o.n.
 - Deuxième méthode : trouver une b.o.n. de F^\perp et projeter sur F^\perp puis passer au projecteur associé
- Les méthodes sont symétriques. Les calculs seront d'autant plus longs que la dimension de l'espace sur lequel on projette est grande. Ainsi, on adoptera la première méthode plutôt lorsque $\dim(F) \leq \dim(F^\perp)$, et la seconde dans le cas contraire.

III.4 Distance d'un point à un sous-espace vectoriel

Théorème 24.3.26 (Distance d'un point à un sous-espace)

Soit E un espace euclidien, F un sous-espace de E , et $x \in E$. Alors :

$$\forall y \in F, \quad \|x - p_F(x)\| \leq \|x - y\|,$$

l'égalité étant réalisée si et seulement si $y = p_F(x)$.

Autrement dit, $p_F(x)$ est l'unique vecteur de F minimisant la distance de x à un point de F .

Définition 24.3.27

On dit que $p_F(x)$ est la meilleure approximation de x dans F . On appelle distance de x à F le réel suivant :

$$d(x, F) = \|x - p_F(x)\| = \min_{y \in F} \|x - y\|.$$

IV Géométrie affine

Comme vous le savez pour l'avoir déjà utilisé en physique, un plan (non vectoriel) peut être défini par la donnée d'un de ses points et d'un vecteur normal. Cela nécessite pour commencer une définition rigoureuse du cadre de la définition affine.

IV.1 Sous-espaces affines d'un espace vectoriel

Définition 24.4.1 (Structure d'espace affine)

Soit T un espace vectoriel sur \mathbb{K} , et E un ensemble. On dit que E est un espace affine attaché à T s'il est muni d'une loi externe de $T \times E$ dans E , notée $(t, x) \mapsto t + x$ ou $x + t$, telle que

- pour tout $(t, t') \in T^2$, tout $x \in E$, $(t + t') + x = t + (t' + x)$
- pour tout $x \in E$, $0 + x = x$
- pour tout $(x, y) \in E^2$, il existe t tel que $y = t + x$
- $(\forall x \in E, t + x = x) \implies t = 0$

Il faut voir E comme un ensemble de points sur lequel un ensemble T de vecteurs agit par translation.

Proposition 24.4.2

Soit E un espace affine attaché à T . Soit $x \in E$. L'application $t \mapsto t + x$ est une bijection de T sur E .

Proposition/Définition 24.4.3 (Translation)

Soit E un espace affine attaché à T . Soit $t \in T$. L'application $\tau_t : x \mapsto t + x$ est une bijection de E dans lui-même. Cette application est appelée *translation de vecteur* t .

Ainsi, $(T, +)$ est le groupe des translations de E .

Définition 24.4.4 (Structure affine sur un espace vectoriel)

Soit E un espace vectoriel sur \mathbb{K} . On peut alors définir une structure d'espace affine sur E , attaché à lui-même, la loi externe correspondant alors à l'addition de E .

Terminologie 24.4.5 (Points et vecteurs)

Soit E un espace vectoriel, muni de sa structure affine usuelle. On distingue les propriétés affines et les propriétés vectorielles de E en considérant, comme dans le cas général d'un espace affine quelconque, que l'espace affine E est distinct de l'espace vectoriel $T = E$ auquel il est rattaché. Ainsi, parlant des éléments de l'espace affine, on parlera de *points* alors que les éléments de l'espace vectoriel E seront appelés *vecteurs*.

La loi externe de l'espace affine E est donc donnée par une relation du type $B = A + \vec{u}$, où A et B sont deux points de l'espace affine et \vec{u} un élément de l'espace vectoriel.

Notation 24.4.6

Soit A et B deux points de l'espace affine A . On notera \overrightarrow{AB} l'unique vecteur \vec{u} de E tel que $B = A + \vec{u}$.

Ainsi, \overrightarrow{AB} est entièrement déterminé par la relation $B = A + \overrightarrow{AB}$.

Définition 24.4.7 (Translaté d'un sous-ensemble de E)

Soit E un espace vectoriel muni de sa structure affine, et X un sous-ensemble de E , vu comme espace affine. Soit t un vecteur de E . Alors le translaté $\tau_t(X)$ de l'ensemble X est le sous-ensemble de l'espace affine E défini par :

$$\tau_t(X) = \{t + x, x \in X\} = \{y \in A \mid \exists x \in X, y = t + x\}.$$

Définition 24.4.8 (Sous-espace affine de E)

Soit E un espace vectoriel, muni de sa structure affine. Un sous-espace affine de E est un translaté par un vecteur t d'un sous-espace vectoriel de E .

En d'autres termes, et en dissociant d'avantage la structure vectorielle et la structure affine, un sous-espace affine de E est un ensemble non vide F tel qu'il existe $x \in F$ et V un sous-espace vectoriel de E tels que

$$F = \{x + u \mid u \in V\}.$$

Cette définition est indépendante d'une correspondance stricte entre les éléments de l'espace affine et les éléments de l'espace vectoriel, et reste vraie dans un espace affine général. Dans le cas de la structure affine usuelle d'un espace vectoriel, les points et les vecteurs étant assimilés, la description précédente correspond à l'espace obtenu par translation de V par le vecteur x .

Proposition/Définition 24.4.9 (Direction)

Soit F un sous-espace affine de E . Il existe un unique sous-espace vectoriel V de E tel que F soit un translaté de V . On dit que V est la direction de F , ou encore que F est dirigé par V .

Proposition 24.4.10

Soit F un sous-espace affine de E de direction V , et $A \in F$. Alors pour tout point B de E , $B \in F \iff \overrightarrow{AB} \in V$.

Assez logiquement nous définissons

Définition 24.4.11 (Hyperplan affine)

Un hyperplan affine d'un espace affine est un sous-espace affine dirigé par un hyperplan vectoriel de E .

Théorème 24.4.12 (Intersection de sous-espaces affines)

L'intersection de sous-espaces affines est soit vide, soit égale à un sous-espace affine. Si cette intersection est non vide, sa direction est l'intersection des directions de chacun des sous-espaces affines.

Certains auteurs définissent parfois la notion de sous-variété affine : il s'agit d'un sous-espace affine, ou de l'ensemble vide. Ainsi, l'intersection de sous-variétés affines est toujours une sous-variété affine.

Exemple 24.4.13 (Sous-espaces affines de \mathbb{R}^2 et \mathbb{R}^3)

Décrivez-les !

Un théorème important fournissant des sous-espaces affines (et on en a déjà vu des cas particuliers) est le suivant :

Théorème 24.4.14 (Fibres d'une application linéaire)

Soit $u \in \mathcal{L}(E, F)$, et $a \in F$. Alors l'image réciproque $u^{-1}(\{a\})$ (appelée fibre en a de u , ou ligne de niveau), est soit l'ensemble vide, soit un sous-espace affine (donc toujours une sous-variété affine).

Les exemples que nous avons déjà rencontrés sont :

Exemple 24.4.15 (Sous-espaces affines obtenus comme fibres)

1. Ensemble des solutions d'un système linéaire
2. Résolution des équations différentielles linéaires non homogènes de degré 1 ou 2.
3. L'ensemble des polynômes interpolateurs en un certain nombre de points.

Définition 24.4.16 (Repère affine)

Un repère affine est la donnée d'un point O de l'espace affine E , et d'une base (vectorielle) (b_1, \dots, b_n) de l'espace vectoriel E .

Définition 24.4.17 (Coordonnées dans un repère affine)

Les coordonnées d'un point A dans un repère (O, b_1, \dots, b_n) sont les coordonnées vectorielles du vecteur \overrightarrow{OA} dans la base (b_1, \dots, b_n) .

Proposition 24.4.18

Soit E un espace affine de dimension n . L'application qui à A associe le vecteur de ses coordonnées est une bijection de E dans \mathbb{R}^n .

IV.2 Définition d'un hyperplan par vecteur normal

On considère un espace euclidien E , muni de sa structure affine. On commence par préciser les notions d'orthogonalité dans le contexte affine.

Définition 24.4.19 (vecteur orthogonal à un sous-espace affine)

Soit \vec{v} un vecteur de E , et F un sous-espace affine. On dit que \vec{v} est orthogonal à F si l'une des trois propriétés équivalentes suivantes est vérifiée :

- \vec{v} est orthogonal à la direction de F ,
- étant donné A fixé, \vec{v} est orthogonal à tout vecteur \overrightarrow{AB} , pour $B \in F$;
- Pour tout $(A, B) \in F^2$, \vec{v} est orthogonal à \overrightarrow{AB}

Proposition 24.4.20 (Hyperplans définis par le vecteur normal \vec{n})

Soit \vec{n} un vecteur non nul. Alors les hyperplans affines orthogonaux à \vec{n} sont exactement les fibres (ou lignes de niveau) de $\vec{u} \mapsto \langle \vec{u}, \vec{n} \rangle$.

Proposition 24.4.21 (Définition d'un hyperplan affine par vecteur normal)

Soit \vec{n} un vecteur non nul de E , et A un point. Il existe un unique hyperplan affine H de E passant par A et orthogonal à \vec{n} . On dit que \vec{n} est un vecteur normal à H

La donnée d'un vecteur normal et de A permet de trouver facilement une équation de l'hyperplan dans une base orthonormale, et réciproquement :

Proposition 24.4.22 (Équation d'un hyperplan de vecteur normal \vec{n})

Soit \mathcal{B} une base orthonormale de l'espace euclidien E , et soit $[\vec{n}]_{\mathcal{B}} = (a_1, \dots, a_n)$ le vecteur de ses coordonnées de \vec{n} . Alors un hyperplan affine H de E admet \vec{n} comme vecteur normal si et seulement s'il admet une équation du type :

$$B \in H \iff a_1x_1 + \dots + a_nx_n = b,$$

où $[\overrightarrow{OB}]_{\mathcal{B}} = (x_1, \dots, x_n)$.

Plus précisément, si on connaît un vecteur A de H , tel que

$$[\overrightarrow{OA}]_{\mathcal{B}} = (y_1, \dots, y_n)$$

la condition $B \in H$ s'écrit $\langle \overrightarrow{AB}, \vec{n} \rangle = 0$, ou encore :

$$a_1(x_1 - y_1) + \dots + a_n(x_n - y_n) = 0, \quad \text{soit:} \quad a_1x_1 + \dots + a_nx_n = a_1y_1 + \dots + a_ny_n.$$

Réciproquement, lire un vecteur normal sur une équation est immédiat !

Exemples 24.4.23

Donner une description par point et vecteur normal dans les cas suivants :

1. D est la droite de \mathbb{R}^2 d'équation $y = 3x - 1$
2. P est le plan de \mathbb{R}^3 d'équation $x + 2y + z = 2$.

Définition 24.4.24 (Distance à un hyperplan)

Soit H un hyperplan défini par le point A et le vecteur normal **unitaire** \vec{n} . Alors la distance d'un point à H est $|\langle \overrightarrow{AM}, \vec{n} \rangle|$.

Isométries vectorielles

Nous étudions dans ce chapitre des applications linéaires particulières entre espaces euclidiens, conservant la norme. Dans le cas de la dimension 2, on retrouve les transformations vectorielles du plan (symétries et rotations vectorielles). Les isométries sont intimement liées à la notion de matrice orthogonale.

I Isométries d'un espace euclidien

Définition 25.1.1 (Isométrie vectorielle)

Une isométrie vectorielle d'un espace euclidien E est un endomorphisme $u \in \mathcal{L}(E)$ vérifiant :

$$\forall x \in E, \|f(x)\| = \|x\|.$$

On note $O(E)$ l'ensemble des isométries de E .

Ainsi, une isométrie est par définition un endomorphisme conservant la norme. La notation $O(E)$ est très voisine de celle utilisée pour désigner le groupe orthogonal. Cela n'a rien d'anodin, comme on le constatera plus tard.

Définition 25.1.2 (Isométrie affine, HP)

Une isométrie affine d'un espace euclidien E est une application f telle que $x \mapsto f(x) - f(0)$ est une isométrie vectorielle.

Exemples 25.1.3

1. $\text{Id}_E, -\text{Id}_E$
2. Dans \mathbb{R}^2 muni de la structure euclidienne canonique, les endomorphismes canoniquement associés aux matrices $\begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$ et $\begin{pmatrix} \cos(\theta) & \sin(\theta) \\ \sin(\theta) & -\cos(\theta) \end{pmatrix}$.
3. Les symétries orthogonales par rapport à un sous-espace F , explicitement données en fonction de la projection orthogonale par $s_F = 2p_F - \text{id}_E$.
4. En particulier, les réflexion (symétries par rapport à un hyperplan)
5. Les projecteurs orthogonaux ?

De façon assez immédiate, en considérant la norme d'un élément du noyau, on obtient :

Proposition 25.1.4 (Bijektivité des isométries)

Une isométrie vectorielle u d'une espace euclidien est un isomorphisme, et u^{-1} est encore une isométrie.

De façon tout aussi évidente :

Proposition 25.1.5 (Composée d'isométries)

La composée $v \circ u$ de deux isométries vectorielles de E est encore une isométrie de E .

On en déduit alors que

Corollaire 25.1.6 (Structure de $O(E)$)

L'ensemble $O(E)$ des isométries est un sous-groupe de $GL(E)$.

Proposition 25.1.7 (Caractérisation des isométries par conservation du produit scalaire)

Soit E un espace euclidien et $u \in \mathcal{L}(E)$. Alors u est une isométrie si et seulement si :

$$\forall (x, y) \in E^2, \quad \langle f(x), f(y) \rangle = \langle x, y \rangle.$$

Sans avoir défini correctement la notion d'angle entre deux vecteurs, mais par analogie à la situation bien connue de \mathbb{R}^2 , cette propriété est à voir comme une propriété de conservation de l'angle (non orienté), en plus de la norme (obtenue pour x et y). En particulier, on a la conservation de l'orthogonalité : si $x \perp y$ alors $u(x) \perp u(y)$.

On peut obtenir une caractérisation par l'orthogonalité de la sorte, mais en rajoutant une information permettant de récupérer la conservation des normes, dans toutes les directions.

Proposition 25.1.8 (Caractérisation des isométries par conservation des b.o.n.)

Soit $u \in \mathcal{L}(E)$. Les propositions suivantes sont équivalentes :

- (i) u est une isométrie
- (ii) u envoie toute b.o.n. sur une b.o.n.
- (iii) il existe une b.o.n. envoyée par u sur une b.o.n.

On en déduit alors la caractérisation matricielle suivante :

Proposition 25.1.9 (Caractérisation matricielle des isométries)

Soit $u \in \mathcal{L}(E)$. Les propositions suivantes sont équivalentes :

- (i) u est une isométrie
- (ii) La matrice de u dans toute b.o.n. \mathcal{B} est orthogonale
- (iii) il existe une b.o.n. \mathcal{B} telle que $\text{Mat}_{\mathcal{B}}(u) \in O(n)$.

Ainsi, le choix d'une base orthonormale \mathcal{B} de E détermine un isomorphisme (de groupes) $O(E) \xrightarrow{\sim} O(n)$. Via cet isomorphisme, on peut considérer le sous-groupe de $O(E)$, image réciproque du sous-groupe $SO(n)$ des matrices orthogonales positives, c'est-à-dire les isométries u telles que $\det(\text{Mat}_{\mathcal{B}}(u)) = 1$, ou, de façon équivalente, par définition du déterminant d'un endomorphisme, telles que $\det(u) = 1$

Définition 25.1.10 (Isométries vectorielles positives)

Une isométrie vectorielle u est dite positive si et seulement si $\det(u) = 1$. On note $SO(E)$ l'ensemble des isométries positives, sous-groupe distingué de $O(n)$ (en tant que noyau d'un morphisme de groupe).

II Isométries vectorielles en dimension 2

II.1 Description de $O(2)$

Nous donnons dans cette section une description complète des isométries en dimension 2. Pour cela, nous commençons par déterminer les matrices orthogonales de $\mathcal{M}_2(\mathbb{R})$.

Proposition 25.2.1 (Matrices orthogonales de $\mathcal{M}_2(\mathbb{R})$)

(i) Soit $M \in SO(2)$, alors il existe $\theta \in \mathbb{R}$ tel que $M = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$;

(ii) Soit $M \in O(2) \setminus SO(2)$, alors il existe $\theta \in \mathbb{R}$ tel que $M = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ \sin(\theta) & -\cos(\theta) \end{pmatrix}$;

Ces descriptions sont uniques modulo 2π .

On a de plus une règle simple pour le produit de matrices de $SO(2)$. En notant, pour tout $\theta \in \mathbb{R}$,

$$R(\theta) = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix},$$

on obtient :

Proposition 25.2.2 (Inverse et produit dans $SO(2)$)

(i) $R(0) = I_n$

(ii) Pour tout $(\theta, \theta') \in \mathbb{R}$, $R(\theta)R(\theta') = R(\theta + \theta') = R(\theta')R(\theta)$

(iii) Pour tout $\theta \in \mathbb{R}$, $R(\theta)^{-1} = R(-\theta)$.

En particulier, on reconnaît en $SO(2)$ un groupe qu'on a déjà rencontré.

Théorème 25.2.3 ($SO(2)$ est isomorphe à \mathbb{U})

L'application qui à $R(\theta)$ associe $e^{i\theta}$ est un isomorphisme de groupe entre $SO(2)$ et \mathbb{U} .

II.2 Isométries positives en dimension 2

Soit E un espace euclidien de dimension 2. On suppose que E est orienté. On obtient la caractérisation matricielle des isométries positives de E :

Théorème 25.2.4 (Caractérisation matricielle des isométries positives en dimension 2)

Soit $u \in \mathcal{L}(E)$. Les propriétés suivantes sont équivalentes :

(i) $u \in SO(E)$

(ii) Il existe une b.o.n. directe \mathcal{B} et un réel θ tel que $\text{Mat}_{\mathcal{B}}(u) = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$

(iii) Il existe θ tel que pour toute b.o.n. directe \mathcal{B} , $\text{Mat}_{\mathcal{B}}(u) = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$

De plus, dans ce cas, θ est unique modulo 2π .

Définition 25.2.5 (Rotation)

On appelle rotation (vectorielle) d'angle $\theta \in \mathbb{R}$ (défini modulo 2π) l'application de $SO(2)$, usuellement notée ρ_θ , dont la matrice dans toute base orthonormale directe est $\begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$

Dans \mathbb{R}^2 muni de son produit scalaire canonique, cette définition correspond bien à la définition géométrique élémentaire et intuitive d'une rotation de centre 0 et d'angle θ . Son effet est de tourner la base canonique d'un angle θ .

Remarque 25.2.6

Si on décide d'inverser l'orientation de E , cela a pour effet sur la rotation de changer son angle θ en $-\theta$.

Les règles de produit matriciel dans $SO(2)$ amènent directement les règles de composition des rotations, dont l'interprétation géométrique est assez intuitive :

Proposition 25.2.7 (Inverse et composée de deux rotations)

- (i) $\rho_0 = \text{id}_E$
- (ii) Pour tout $(\theta, \theta') \in \mathbb{R}$, $\rho_\theta \circ \rho_{\theta'} = \rho_{\theta+\theta'} = \rho_{\theta'} \circ \rho_\theta$
- (iii) Pour tout $\theta \in \mathbb{R}$, $\rho_\theta^{-1} = \rho_{-\theta}$.

Nous pouvons définir la notion d'angle orienté entre deux vecteurs grâce aux rotations. Pour cela, nous utilisons le lemme suivant :

Lemme 25.2.8

Soit E un espace euclidien orienté, et soit x et y deux vecteurs de norme 1 de E . Il existe une unique rotation ρ telle que $\rho(x) = y$.

Définition 25.2.9 (Angle orienté entre deux vecteurs)

Soit E un espace vectoriel orienté, et x et y deux vecteurs non nuls de E . Alors l'angle orienté $\widehat{(x, y)}$ est l'angle θ , défini modulo 2π , de l'unique rotation ρ telle que

$$\rho\left(\frac{x}{\|x\|}\right) = \frac{y}{\|y\|}.$$

Des règles de composition des rotations découlent immédiatement les trois premières des règles suivantes sur les angles :

Proposition 25.2.10 (Propriétés des angles orientés)

Les égalités ci-dessous sont à lire modulo 2π .

- (i) $\forall x \in E \setminus \{0\}$, $\widehat{(x, x)} = 0$
- (ii) $\forall x, y, z \in E \setminus \{0\}$, $\widehat{(x, z)} = \widehat{(x, y)} + \widehat{(y, z)}$
- (iii) $\forall x, y \in E \setminus \{0\}$, $\widehat{(y, x)} = -\widehat{(x, y)}$.
- (iv) $\forall x, y \in E \setminus \{0\}$, et $\lambda, \mu \in \mathbb{R}_+^*$, $\widehat{(\lambda x, \mu y)} = \widehat{(x, y)}$.
- (v) $\forall x \in E \setminus \{0\}$, $\widehat{(-x, x)} = \pi$
- (vi) $\forall x, y \in E \setminus \{0\}$, $\widehat{(-x, y)} = \pi + \widehat{(x, y)}$.

Nous voyons maintenant comment déduire de toutes les définitions ci-dessus l'expression du produit scalaire vue au lycée à l'aide de l'angle entre les vecteurs, ainsi que l'expression du déterminant. Pour cela, nous introduisons une définition générale, valable en dimension n quelconque.

Une matrice de $SO(n)$ ayant un déterminant égal à 1, la formule de changement de base pour les déterminants permet d'affirmer que le déterminant relativement à une base orthonormée d'une famille de n vecteurs d'un espace euclidien de dimension n ne dépend pas du choix de cette base orthonormale.

Définition 25.2.11 (Produit mixte)

Le produit mixte de n vecteurs d'un espace euclidien orienté de dimension n , noté $[x_1, \dots, x_n]$ ou $\det(x_1, \dots, x_n)$, est la valeur commune des $\det_{\mathcal{B}}(x_1, \dots, x_n)$ dans les b.o.n. directes.

Proposition 25.2.12 (Expression du produit scalaire et du produit mixte par l'angle)

Soit x et y deux vecteurs non nuls de E euclidien orienté de dimension 2. Alors

$$\langle x, y \rangle = \|x\| \cdot \|y\| \cdot \cos(\widehat{x, y}) \quad \text{et} \quad [x, y] = \|x\| \cdot \|y\| \cdot \sin(\widehat{x, y})$$

Ainsi, cette proposition permet une formalisation plus rigoureuse de la conservation de l'angle par les isométries, évoquée plus haut à propos de la conservation du produit scalaire : l'expression ci-dessus affirme en fait la conservation du cosinus de l'angle par une isométrie, donc une conservation de la valeur absolue de l'angle. La conservation ou non du signe du déterminant distingue alors les cas de conservation de l'angle orienté ou de passage à l'opposé ; cela distingue donc les isométries positives et les isométries négatives.

Nous donnons, donc, pour les isométries positives de E de dimension 2 :

Proposition 25.2.13 (Conservation de l'angle par une isométrie positive)

Soit $\rho \in SO(E)$, E étant un espace euclidien orienté de dimension 2. Alors

$$\forall (x, y) \in E^2, \quad (\rho(x), \rho(y)) = (\widehat{x, y}).$$

II.3 Isométries négatives en dimension 2

Nous terminons ce chapitre par la description des isométries négatives du plan. Nous rappelons qu'une réflexion est une symétrie orthogonale par rapport à un hyperplan. En dimension 2, les hyperplans sont des droites. Donc les réflexions sont les symétries orthogonales par rapport à une droite.

Théorème 25.2.14 (Isométries négatives en dimension 2)

Soit E un espace euclidien de dimension 2. Les isométries négatives de E sont les réflexions.

Observez que contrairement au cas des rotations, l'angle θ dépend ici du choix de la base orthonormale choisie.

De façon peu surprenante, les symétries inversent les angles. Pour le démontrer nous utilisons le lemme suivant :

Lemme 25.2.15

Soit σ une réflexion et ρ une rotation. Alors $\sigma \circ \rho \circ \sigma^{-1} = \rho^{-1}$.

On obtient

Proposition 25.2.16 (Inversion des angles par isométries négatives)

Soit σ une isométrie négative de E de dimension 2. Alors :

$$\forall (x, y) \in E^2, \quad (\widehat{\sigma(x), \sigma(y)}) = -\widehat{x, y}.$$